Study of transposon-mediated cell pool and cell line generation in CHO cells

THÈSE N^{O} 6563 (2015)

PRÉSENTÉE LE 22 MAI 2015

À LA FACULTÉ DES SCIENCES DE LA VIE
LABORATOIRE DE BIOTECHNOLOGIE CELLULAIRE
PROGRAMME DOCTORAL EN BIOTECHNOLOGIE ET GÉNIE BIOLOGIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Sowmya BALASUBRAMANIAN

acceptée sur proposition du jury:

Prof. S. Maerkl, président du jury
Prof. F. M. Wurm, Dr D. Hacker, directeurs de thèse
Prof. H. Hauser, rapporteur
Prof. D. Neri, rapporteur
Dr H. Pick, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

ACKNOWLEDGEMENTS

The journey to the completion of this doctoral studies has been an exciting and wonderful one. Never before did I imagine myself spending 5 years in such a beautiful country and in such wonderful company. I would like to take this opportunity to thank all the people who guided and supported me leading to the culmination of my time here.

First and foremost, I would like to extend my deepest gratitude to Prof. Florian Wurm, for his expert scientific guidance and for sharing his astute vision. With constant motivation and enthusiasm he inspired me to remain inquisitive and encouraged me to improve, both as a scientist and an individual.

Next, I thank Dr. David Hacker for always being available for discussions and helping me resolve the issues faced during my studies, both big and small. I also thank him for his vital contributions in improving my writing skills and for the hours he poured into correcting my written work.

I thank Dr. Lucia Baldi for always being accessible and available to help with anything in her capacity, from corrections of articles to lab safety. I thank Ms. Fabienne Rudin for her prompt help with all administrative and organizational matters. I greatly appreciated her cheerful conduct and delightful discussions.

I thank my colleagues and friends Dr. Dominique Monteil and Dr. Xiao Shen for being wonderful lab mates to bounce scientific ideas off, in the constantly shrinking LBTC. I thoroughly enjoyed our discussions on all topics under the sun. I thank Daniel Blessing and Likuan Zhu for being great office mates and for making lab feel less lonely and more colorful.

I thank my master students Sylvain Genzer, Fabrizio De Angelis, Anne-Catherine Portmann, Adeline Bourgnon and Antonia Šakic for helping me improve my teaching and management skills. A special thanks to Antonia Šakic for being a wonderful student and sharing with me the excitement and thrill of doing science. I thank her in particular for her

contributions to the work done in chapter 5 of this thesis. I would like to thank to Ms. Virginie Bachmann for all her help in the lab and making the lab environment fun and cheerful.

I would also like to thank all the members of LBTC for providing a conducive and stimulating environment for science – Dr. Mattia Matasci, Dr. Divor Kiseljak, Ricardo Valdès-Bango Curell, Guilia Tontodanati, Ana Karina Pitol, Andres Ricardo Lopez, Clara Douet, Valentin Juvet, Dr. Patrik Michel, Dr. Agata Oberbek, Sagar Manoli, Joao Pereira, Guillaume Lüthi, Nicolas-Julian Hilbold, Sarah Thurnheer and Veronika Knapkova.

A huge thank you to my parents Balasubramanian Venkatachellam and Shobha Balasubramanian, for always supporting and believing in me. Their love and encouragement made me what I am today and for that I am eternally grateful to them. I would also like to thank my wonderful brothers Srinath Balasubramanian and Aditya Prakash and my lovely friends Aditi Desai, Ketki Chawla, Joyeeta Sinha, Diego Marcos Gonzalez, Rachana Pradhan, Chhavi Jain, Varun Sharma and Emmanuel Froustey for all the wonderful times we shared.

Last but not the least, I thank Dr. Yashas Rajendra for always being by my side from the day I joined LBTC. His constant encouragement and constructive criticism inspired me to do better and aim to reach my potential, and the long scientific discussions helped me hone my skills. But most of all I thank him for being a wonderful partner and supporting me though all the ups and downs of my time here.

If I have forgotten to mention anyone, please pardon me. I would like to extend my gratitude to everyone who has been a part of my journey in earning my PhD.

Sowmya Balasubramanian

ABSTRACT

The goal of this thesis was to evaluate the use of artificial transposon systems for the generation of recombinant cell pools and cell lines with Chinese hamster ovary (CHO) cells as the production host. Transposons are naturally occurring genetic elements present in the genome of most organisms. Several transposons of metazoan origin have been engineered to facilitate the integration of one or more recombinant genes (transgenes) into the genome of the host cell. In this thesis, three of the most commonly used transposons, piggyBac (PB), Tol2, and Sleeping Beauty (SB), were used for transgene delivery into CHO-DG44 cells. The transposon systems described here involves co-transfection of the host cells with a helper plasmid for the transient expression of the transposase, and a donor plasmid coding for the transgene and a gene for the selection, both positioned between two transposon repeat sequences that are required for DNA transposition. The tumor necrosis factor receptor Fc fusion protein (TNFR:Fc) was used as the model protein for most experiments. Through the optimization of various selection parameters we showed that PB-mediated cell pools can be generated with selection duration of as little as 5 days in the presence of puromycin. All three transposon systems, PB, Tol2, and SB, resulted in cell pools with similar volumetric TNFR:Fc productivities that were about 9 times higher than those generated by conventional plasmid transfection. Transposon-mediated cell pools had 10 – 12 transgene integrations per cell. However, we demonstrated that some the integration events occurred via DNA recombination rather than transposition. We also isolated clonal cell lines from cell pools. As expected, the average volumetric TNFR:Fc productivity of transposon-derived cell lines was higher than that of cell lines generated by conventional transfection. In 14-day fed-batch cultures, protein levels up to 900 mg/L and 1.5 g/L were obtained from transposon-mediated cell pools and cell lines, respectively. The stability over time of the volumetric productivity of cell pools was determined by maintaining the cells in culture for 3 months in the absence of selection. In general, the productivity decreased to 50 % its initial level over the first 7 weeks in culture and then remained constant for the following 5 weeks. In contrast, the volumetric protein yield from transposon-mediated cell lines remained constant for up to 4 months in the absence of selection. We also showed that the three transposon systems could be used for cell pool generation with CHO-K1 and CHO-S with similar volumetric productivities as observed with CHO-DG44 cells. Finally,

we utilized the PB transposon system for generating cell pools co-expressing up to 4 different transgenes, enhanced green fluorescent protein (EGFP), secreted alkaline phosphatase (SEAP), and the light and heavy chains of an IgG1 monoclonal antibody, by simultaneous transfection of all four transgenes. We showed that PB-mediated cell pools had increased volumetric productivity of each of the proteins compared to those generated by conventional co-transfection. The use of the PB transposon system increased the percentage of cells in the pools that were co-expressing all four proteins as compared to the results with cell pools generated by conventional transfection. In conclusion, the transfection of CHO cells with the PB, Tol2 or SB transposon system is a simple, efficient, and reproducible approach to the generation of cell pools and cell lines for the rapid production of recombinant proteins.

Keywords: recombinant protein, CHO cells, piggyBac, Tol2, Sleeping Beauty, transposon, cell pool, cell line, transfection

RESUMÉ

L'objectif de cette thèse était d'évaluer l'application de systèmes basés sur les "transposons" pour obtenir des lignées et des pools de cellules recombinantes utilisant comme hôte de production les cellules CHO (Chinese hamster ovary cells – Cellules ovariennes d'hamster chinois). Les transposons sont des éléments génétiques naturels présents dans le génome de la plupart des organismes vivants. Un grand nombre de transposons d'origine métazoaire ont été modifié génétiquement afin de faciliter l'intégration d'un ou plusieurs gènes recombinants (transgènes) dans le génome de différentes cellules hôtes. Dans cette thèse ont été évalués les trois transposons les plus fréquemment utilisés, c'est à dire piggyBac (PB), Tol2 et Sleeping Beauty (SB), pour leur capacité à insérer les transgènes dans les cellules de la lignée CHO-DG44. Le système de transposons utilisé ici comporte la co-transfection de deux vecteurs plasmidiques, le "helper" pour l'expression transitoire de l'enzyme transposase, et le "donor", codant pour le transgène et le gène de la résistance aux antibiotiques utilisé pour la sélection clonale : les deux gènes du plasmide "donor" sont chacun entourés des deux séquences répétitives nécessaires pour la transposition dans l'ADN. La protéine recombinante utilisée comme modèle dans la plupart des expériences était le TNFR :Fc, produit de la fusion du récepteur au facteur de nécrose tumorale et de la partie constante d'une immunoglobuline (Fc). Grâce à l'optimisation de plusieurs paramètres de sélection nous avons démontré qu'avec PB, il est possible de générer des pools après une période de sélection à la puromycin en seulement 5 jours. Les trois système de transposons, PB, Tol2 et SB ont donné des pools avec des productivités similaires entre eux (TNFR :Fc) et environ 9 fois supérieures aux titres obtenus avec la méthode standard de transfection plasmidique. Les pools générés par transposition ont montré 10 – 12 intégrations individuelles par cellule. Nous avons démontré qu'une partie de ces intégrations était le résultat de recombinaison et non pas de transposition. Nous avons également isolé des lignées cellulaires clonales à partir des pools. La moyenne de production volumétrique de TNFR :Fc dans les lignées dérivés par transposition était supérieure à celle des lignées générées par transfection conventionnelle. Dans des cultures fed batch de 14 jours le niveau de productivité obtenu étaient de 900 mg/L et 1.5 g/L, respectivement pour les pools dérivés par transposition et les lignées générées par transfection conventionnelle. La stabilité dans le temps de cette production a

été testée dans des cultures maintenues durant une période de trois mois en l'absence de sélection. En général, la productivité est descendue à un niveau de 50% par rapport à la valeur initiale pendant les 7 premières semaines de culture puis est restée stable pendant 5 semaines successives. Par contre, la productivité volumétrique des lignées clonales générées par transposons est restée constante jusqu'à 4 mois en absence de sélection. Nous avons aussi validé l'utilisation des trois systèmes de transposons PB, Tol2 et SB pour la production de pools recombinantes avec d'autres lignées cellulaires, CHO-K1 et CHO-S, avec productivités volumétriques similaires à celles observées dans les CHO-DG44. Enfin nous avons utilisé le système PB transposon pour générer des pools recombinants qui co-expriment quatre transgènes différents insérés simultanément par co-transfection: l'EGFP (enhanced green fluorescence protein), la SEAP (secreted alkaline phosphatase), les deux chaines légère et lourde d'une immunoglobuline IgG1. Nous avons pu démontrer que les pools générés par transposition avec PB avaient une productivité supérieure pour toutes les protéines en comparaison avec les pools générés par transfection conventionnelle. En conclusion, la transfection des cellules CHO avec les systèmes de transposons PB, Tol2 ou SB représente une approche simple, reproductible et efficace pour la création de lignées et pools de cellules transgéniques pour la production rapide de protéines recombinantes.

Keywords: protéine recombinante, cellules CHO, piggyBac, Tol2, Sleeping Beauty, transposon, pools de cellules, lignées de cellules, transfection

LIST OF ABBREVIATIONS

BGH-pA	Bovine growth hormone polyadenylation element
bla	Beta-lactamase gene
bsr	Blasticidin S-resistance gene
CHO	Chinese hamster ovary cells
CMV	Cytomegalovirus immediate early promoter
DMA	Dimethyl acetamide
DNA	Deoxyribonucleic acid
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme linked immunosorbent assay
FACS	Fluorescence-associated cell sorting
HEK 293	Human embryonic kidney 293 cells
hph	Hygromycin B phosphotransferase gene
HSV-TK	Herpes simplex virus thymidine kinase
IgG	Immunoglobulin G
IRES	Internal ribosome entry site
ITR	Inverted terminal repeats
PB	PiggyBac
PBS	Phosphate buffer saline
PEI	Polyethylenimine
pDNA	Plasmid deoxyribonucleic acid
RT-qPCR	Quantitative real time polymer chain reaction
SB	Sleeping beauty
SEAP	Secreted alkaline phosphatase
SGE	Stable gene expression

Sh ble	*ble* gene from Streptoalloteichus hindustanus
SV40-pA	Simian virus 40 polyadenylation element
TGE	Transient gene expression
TNFR:Fc	Soluble tumor necrosis factor receptor type II fused to the Fc portion of the human IgG1

TABLE OF CONTENTS

1.

Introduction

1.1. Therapeutic Recombinant Proteins

The biotechnology industry is a $140 billion industry with 246 biopharmaceutical products on the market (Walsh 2014). These products are targeted to a wide range of diseases including cancers and autoimmune disorders. The therapeutic proteins available can be categorized into different molecular types including antibody-based drugs, anticoagulants, blood factors, bone morphogenetic proteins, engineered protein scaffolds, enzymes, Fc fusion proteins, growth factors, hormones, interferons, interleukins, and thrombolytics (Carter 2011).

Recombinant human insulin (Humulin®) created at Genentech, developed by Eli Lilly, and approved by the US Food and Drug Administration (FDA) in 1982 was the first human protein therapeutic derived from recombinant DNA technology in *E. coli* (Walsh 2005). Human tissue plasminogen activator (tPA) produced in Chinese hamster ovary (CHO) cells in 1986 by Genentech was the first therapeutic protein from recombinant mammalian cells (Wurm 2004). Since then, the recombinant therapeutic protein market has been growing at a steady rate. The percentage of therapeutic proteins among the new drugs approved by the USA Food and Drug Administration (FDA) increased from 17% in 2005 to 32% in 2011 (Kling 2011). The total sales of biologics in United States in 2012 reached ~$63.6 billion, an 18.2% increase over 2011 sales (Aggarwal 2014). In the last 15 years (1999-2013), 113 first in class drugs were approved by the FDA of which 30% were biologics (Eder et al. 2014).

Among the different kinds of therapeutic proteins, monoclonal antibodies (mAb) are the highest selling class of biologics with mAb sales reaching $63 billion in 2013 (Walsh 2014). There are currently 40 US Food and Drug Administration (FDA)-approved mAbs, with their sales constituting ~38.5% of the total biologics market (Aggarwal 2014).

In this thesis, most of the studies were conducted using the tumor necrosis factor receptor-Fc fusion protein (TNFR:Fc). Fusion proteins are created by joining the genes encoding 2 or more different proteins. They exhibit properties of their component parts. The fusion of the Fc region of immunoglobulins with other proteins has been the most successful among the fusion protein therapeutics (Walsh 2010). Fusion with Fc has been known to benefit

recombinant proteins by increasing half-life, facilitating expression and secretion, enabling facile purification by protein A affinity chromatography, improving solubility and stability, and enhancing potency by increasing valency (Carter 2011; Huang 2009). The sales in US of Fc fusion proteins had a growth rate of ~35.3% in 2012, which was the highest among biologics (Aggarwal 2014). The most successful drug of this class has been etanercept (Enbrel®), a TNFR2:Fc fusion protein produced by Amgen. It was the top-selling protein therapeutic in 2013 with $8.8 billion in world-wide sales in rheumatoid arthritis and other autoimmune diseases (Walsh 2014).

1.2. Protein production in CHO cells

The variety of hosts for protein production extends from microbial to transgenic animals and plants. However, cultivated mammalian cells remain the principal hosts for therapeutic protein production due to their ability to perform human-like post-translational modifications (PTMs) (Wurm 2004; Zhu 2012). About 60% of the biopharmaceuticals approved between 2010 and 2014 are produced from mammalian cells (Walsh 2014). Some of the cell lines commonly used for mammalian cells include CHO, human embryonic kidney (HEK-293), human retina-derived PerC6, baby hamster kidney (BHK), and mouse myeloma-derived NS0 (Jayapal et al. 2007; Li et al. 2010; Wurm 2004). Among these, CHO cells are most commonly used for therapeutic protein production (Walsh 2014).

CHO cells have been grown in culture for more than 50 years since their discovery in 1957 by Theodore Puck (Puck 1957). CHO cells are well-characterized and have been demonstrated to be a safe host as they are not receptive to human viruses (Kim et al. 2012; Lai et al. 2013). They can also be easily adapted to chemically defined media (Butler and Spearman 2014). The ability of CHO cells to grow to high cell densities in single-cell suspension culture contributed to their attractiveness as a production host (Hacker et al. 2010). Also, CHO cells have the capacity for efficient post-translational modification, and they produce recombinant proteins with glycoforms that are both compatible with and bioactive in humans (Kim et al. 2012). The product yield from recombinant CHO cells has improved more than 20-fold over the past two decades, and this improved product yield has been largely attributed to the development in media and bioprocess design (Hacker et al. 2009).

There are a number of CHO strains that exhibit different phenotypes in cell culture (Wurm 2013). In this thesis, I have worked with three of these: CHO-DG44, CHO-K1, and CHO-S.

1.2.1. CHO-K1

The proline-dependent CHO-K1 cell strain was generated from the original CHO cell line by single-cell cloning in late 1960s (Wurm 2013). The first published CHO genome was that of CHO-K1 (Xu et al. 2011). CHO-K1 cells have been commonly used with the glutamine synthase (GS) selection system (Bebbington et al. 1992; Daramola et al. 2014; Fan et al. 2013; Knox et al. 2013). The GS gene is co-transfected along with the GOI and the GS inhibitor methionine sulphoximine (MSX) is used to inhibit endogenous GS activity such that only transfectants with additional GS activity to overcome this selection can survive. This system based on CHO-K1 is one of the most popular methods for selection of recombinant CHO cell lines. There are several licensed pharmaceutical products that were made with the help of the GS system in combination with CHO-K1 cells (Wurm 2013).

1.2.2. CHO-DG44

The popularity of CHO cells for recombinant protein production increased after the isolation of cell lines harboring mutations in the dihydrofolate reductase (DHFR) gene (Kaufman et al. 1985; McCormick et al. 1984; Scahill et al. 1983). The cell line DXB11 (or DUKX) was generated in 1980 by deletion of one DHFR allele and an inactivating mutation in the second allele from CHO-K1 through chemical mutagenesis (Urlaub and Chasin 1980). The cell line CHO-DG44 was generated by deleting both DHFR alleles from CHO-pro3-strain, another derivative of the original CHO cell line, by two rounds of gamma irradiation and then screening for the absence of DHFR activity. These two DHFR-minus strains require glycine, hypoxanthine, and thymidine (GHT) for growth (Urlaub et al. 1983). They are used as hosts for recombinant protein production since the absence of DHFR activity serves as a method for selection following transfection of the cells with the gene of interest (GOI) and the DHFR gene. Recombinant cells are then selected by growth of the culture in the absence of GHT. However, due to frequent low recombinant protein

productivity by surviving cells, selection for the DHFR gene often involves exposure of the cells to methotrexate (MTX), a chemical inhibitor of DHFR (Wurm 2004). Selection in the presence of MTX usually results in amplification of the exogenous DHFR gene and the GOI. These advantages of selection and amplification made CHO-DG44 cells one of the most widely used host systems for protein production.

1.2.3. CHO-S

The CHO cells were first adapted to grow in suspension cultures in 1960s. Unfortunately, there is lack of clarity and scientific credit for the origin of CHO-S cells (Wurm 2013). CHO-S cells were first mentioned in 1973 (Thompson and Baker 1973). There is now a commercially available cell line adapted for growth in suspension culture using FreeStyle™ CHO medium (Life Technologies), also designated as CHO-S. Even though the commercially available CHO-S could have been derived from the original CHO-S, these two cell lines are expected to differ phenotypic and genetically (Wurm 2013).

1.3. Protein production strategies using mammalian cells

1.3.1. Stable cell line generation

Recombinant protein production from stable CHO cell lines is the most widely used manufacturing approach in the biopharmaceutical industry (Wurm 2004). Stable gene expression involves the expression of the recombinant protein from cell lines in which the transgene(s) has been stably integrated into the genome of the host cell. The process of generation of a recombinant cell line involves several steps. The first step in cell line generation is the delivery of the GOI to the target cells, followed by a genetic selection step to eliminate the non-recombinant cells. Then, by single-cell cloning, cell lines with the phenotypes such as robust cell growth and high protein expression and protein quality are recovered.

To create stably transfected mammalian cell lines an efficient selection system is required to eliminate non-recombinant cells and achieve high protein expression levels (Van Blokland et al. 2011). Raising the stringency of selection often helps to reduce the number of recombinant cells and thereby increases the chance to identify a high protein-producing

cell line (Wurm 2004). This can be achieved by increasing the concentration of the selection agent in the culture medium, but this approach has its limits due to the toxicity of the agent (Van Blokland et al. 2007). Well-established selection strategies rely upon complementation of a host auxotrophy, as described above. In CHO cells, the two most commonly used selection systems are based on the DHFR and GS systems as described in Section 1.2. However, use of MTX and MSX in high concentrations promotes amplification of the integrated recombinant genes, which is known to cause chromosomal rearrangements, leading to clonal variability in regards to growth, productivity, and stability (Chusainow et al. 2009).

A common alternative to the auxotrophic selection method is the use of genes conferring resistance to antibiotics such as geneticin (G418), hygromycin B, zeocin, blasticidin or puromycin. With this strategy, transfected cells are selected in medium containing the appropriate antibiotic. The antibiotics target different cellular pathways from protein synthesis to transcription (Darken 1964; Drocourt et al. 1990; Gonzalez et al. 1978). The pool of cells recovered after selection is highly heterogeneous in terms of specific protein productivity and cell growth. This necessitates the isolation and evaluation of several hundred individual cells to recover a few candidate production cell lines that possess the desired characteristics.

There are several methods for generating single-cell clones. The conventional method, limiting dilution, is time-consuming and low-throughput (Browne and Al-Rubeai 2007). Recently, several automated methods have been developed to select for high-producing cell lines including methods using flow cytometry and cell sorting and methods based on individual protein secretion amounts (e.g. Matrix – based secretion assays) (Akselband et al. 2003; DeMaria et al. 2007; Holmes and Al-Rubeai 1999; Kacmar and Srienc 2005). They have significantly reduced the time frame required to screen large number of clones. However, in some cases these methods are expensive and can be limited by the need for specific antibodies to the therapeutic proteins generated or by the need to optimize the experimental conditions to a specific cell line (Browne and Al-Rubeai 2007). A better strategy to reduce the time frame and screen size for generation of stable cell lines is to

influence the integration of the GOI, which will be described in detail in the following sections.

Protein production from cell lines is a well-established technology with FDA-approved guidelines. It is the preferred method for the production of marketed recombinant proteins as each cell line is derived from a single cell and is expected to be a homologous cell population with little to no variability, a very valuable attribute in terms of product quality. Cell lines are high-yielding with volumetric productivities up to 10 g/L, and they can stably express the protein of interest for extended periods of time. This stability allows them to be scalable up to 20,000 L without any loss in productivity. However, stable cell line development is still a time-consuming process taking several months to complete (Browne and Al-Rubeai 2007; Matasci et al. 2008; Wurm 2004).

1.3.2. Transient gene expression

During preclinical stages of drug development, only milligram to gram quantities of recombinant proteins are required. Due to the long time-frame for the generation of stable cell lines, transient gene expression (TGE) in suspension-adapted mammalian cells, is often employed as a rapid method for the production of moderate quantities of proteins to meet these needs (Geisse and Voedisch 2012; Hacker et al. 2013; Pham et al. 2006). TGE involves expression of transgenes from episomal plasmid DNA in a heterogeneous cell population. The entire process from DNA delivery to protein recovery takes about 1 – 3 weeks (Daramola et al. 2014; Rajendra et al. 2011). Also, TGE can be used for producing proteins that are cytotoxic to cells. Recent improvements in TGE with CHO cells, by employing cell line engineering and process optimization strategies, have resulted in protein yields of 1 – 2 g/L (Cain et al. 2013; Daramola et al. 2014; Rajendra et al. 2011). However, the relatively large amounts of plasmid DNA required for transfection and the difficulty in working at large volumetric scales limit the usefulness of TGE.

1.3.3. Stable cell pools

An alternative approach for the rapid production of a moderate amount of protein is through the generation of cell pools or bulk cultures. These are heterogeneous populations of

recombinant cells obtained by transgene delivery and genetic selection but without the subsequent single cell cloning steps. This methodology is commonly used for the evaluation of expression vectors and cell line engineering (Majors et al. 2008). For cell line development in industry, cell pools are usually generated for the recovery of clonal production cell lines.

Use of cell pools for therapeutic protein production is not traditionally preferred due to the heterogeneity of the cell population and the transgene expression instability. Accordingly, production cell lines are derived from a single cell, in order to minimize heterogeneity. However, we are now aware of the possibility of occurrence of intraclonal variation in the genotype of cells (Pilbrough et al. 2009; Wurm 2013). This contributes to the instability of clones, resulting in a heterogeneous population not unlike a cell pool. Hence, cell pool generation has recently gained interest for rapid production of protein. Even though stably transfected cell lines will continue to be the method of choice for protein manufacturing, cell pools can be used as an alternative to TGE for rapid production of proteins for use during the early phases of drug discovery. The main advantages of cell pools over TGE are the low DNA requirement, the ease of scalability, and the potential for high volumetric yields (Fan et al. 2013; Li et al. 2013; Ye et al. 2010). There is also the possibility of maintaining a frozen cell pool bank to allow multiple consecutive production runs. Cell pools are advantageous to clonal cell lines because of shorter and less costly development times (Ye et al. 2010).

There are few publications on the development of cell pools as a method for recombinant protein production (Fan et al. 2012; Li et al. 2013; Ye et al. 2010). Volumetric productivities of up to 600 mg/L, using fluorescence-activated cell sorting (FACS) to enrich the cell pool in recombinant cells, have been achieved (Ye et al 2010). With the use of specific DNA elements appended to the transgene expression cassette, productivities of up to 1 g/L have also been achieved (Ye et al. 2010). However, the volumetric productivity of these cell pools was reduced by 40 % per week in culture. Cell pools with volumetric productivities up to 900 mg/L were obtained from an optimized fed-batch process by using GS-knockout cells coupled with a weakened promoter for selection of the exogenous GS gene (Fan et al. 2013). Major biopharmaceutical companies are now evaluating use of cell

pools as a method for generating material for preclinical stages of therapeutic protein development such as toxicology tests and first human dose (FHD) analysis They have successfully generated cell pools with high productivities using technologies to express the antibody and selection marker as a single mRNA (personal communication, Dr. Chetan Goudar).

1.4. DNA integration methods for cell line generation

1.4.1. DNA recombination

The majority of the available methods for cell line generation rely on integration through DNA recombination of the exogenous DNA with the cellular genome. The resulting recombinant cells differ in protein expression due to lack of control over gene copy number and the chromosomal context of the integrated genes. This leads to differences in the productivity and stability of transgene expression among the cells. Consequently, a large number of candidate recombinant cell lines (routinely 500-1000) must be screened to find a few which express the protein of interest stably at the desired level (Table 1.1) (Matasci et al. 2008).

For recombinant cell lines, the stability of protein expression over time strongly depends on the integration site of the heterologous DNA in the host cell genome, known as position effect variegation (PEV) (Eissenberg et al. 1992; Wilson et al. 1990). Integration near or within transcriptionally-inactive heterochromatin leads to little or no transgene expression and may result in unstable protein production (Kouzarides 2007; Whitelaw et al. 2001). Even if the gene integrates into transcriptionally-active euchromatin, production instability may occur due to gene loss or transcriptional silencing (Chusainow et al. 2009; Kim et al. 2011; Osterlehner et al. 2011). When the transgene is integrated into the cellular chromosomes, the gene expression levels are affected by the chromatin structures around the integrated genes. Hence several strategies are being employed to improve the efficiency of cell line generation.

1.4.2. DNA regulating elements

Many strategies to obtain position-independent transgene expression rely on the use of DNA regulating elements (Barnes and Dickson 2006; Kwaks and Otte 2006). They are specific DNA sequences appended to the transgene expression cassette within the plasmid. These DNA boundary sequences are able to block the formation of condensed chromatin (heterochromatin), and thus may help to obtain stable transgene expression irrespective of the chromosomal integration site (Matasci et al. 2008). Until now, a number of *cis*-acting elements have been applied to enhance the expression and stability of protein production in CHO cells. These DNA elements include insulators, locus control regions (LCRs), scaffold or matrix attachment regions (S/MARs), antirepressor elements, and ubiquitous chromatin opening elements (UCOEs). LCRs are complex transcriptional enhancers that impart an integration-position-independent and copy-number-dependent expression pattern to linked transgenes (Dean 2006). Insulators are DNA sequence elements that help protect transgenes against chromatin-mediated silencing by increasing the local concentration of factors that promote formation of unfolded structure of chromatin i.e. euchromatin (Gaszner and Felsenfeld 2006). Another class of cis-acting elements is the S/MAR, which maintains the chromatin structure in an "active" configuration through the creation of chromatin loops that curb the silencing of the transgene (Galbete et al. 2009; Gorman et al. 2009). Although this element can increase the protein productivity, the exact mechanism of its action is currently unknown (Kim et al. 2012). Antirepressor or STAR (stabilizing and antirepressor) elements function by affecting the spread of DNA methylation and histone deacetylation patterns from the surrounding genome into the recombinant DNA (Kwaks et al. 2003). Another commonly used element is the UCOE, derived from the promoters of housekeeping genes, which allows position-independent, long-term transgene expression (Antoniou et al. 2003; Williams et al. 2005).

It has been demonstrated that the inclusion of S/MARs or UCOEs in expression vectors increases protein production and the stability of transgene expression in mammalian cells (Benton et al. 2002; Dharshanan et al. 2014; Grandjean et al. 2011). Additionally, the vector engineering required for using DNA regulating elements is simple which makes it easily applicable (Table 1.2). However, the integration requires non-targeted DNA recombination

and the plasmid integration efficiency is only slightly improved by the AT richness of these DNA elements (Allen et al. 2000).

1.4.3. Targeted DNA integration

Alternative approaches to improving volumetric protein production and stable transgene expression rely on targeted integration of a single copy of the gene into highly transcribed sites within the host cell genome. This can be accomplished by means of heterologous site-specific recombination or targeted integration. This allows a stable level of transgene expression over an extended period of time (Barron et al. 2007; Sorrell and Kolb 2005). Cre/loxP (cyclization recombinase/ loxP) and FLP/FRT (flipase/ FLP recombinase target) are the most commonly used site-specific recombination methods for cell line generation (Crawford et al. 2013; Zhou et al. 2010). Both allow the location and timing of gene expression to be closely regulated. It is based on the ability of recombinase genes (cre or FLP) to effect recombination between pairs of DNA sequences (loxP or FRT) that flank a genomic region of interest (Sadowski 1995; Sauer and Henderson 1988; Sternberg 1981). However, they require the flanking recognition sites at the target genomic locus which would enable the homologous recombination. Thus, it requires the construction of a host cell line with these specific recognition sites, which is laborious and time-consuming. Also, these DNA recombinases are known to have low recombination efficiencies (Table 1.1) (Cheng and Alper 2014).

Another strategy for targeted DNA integration is the use of directed DNA nucleases. These are engineered nucleases composed of sequence-specific DNA-binding domains fused to a nonspecific DNA cleavage module. These chimeric nucleases enable efficient and precise genetic modifications by inducing targeted DNA double-strand breaks (DSBs) that stimulate the cellular DNA repair mechanisms, including error-prone non-homologous end joining (NHEJ) and homology-directed repair (HDR) (Gaj et al. 2013). The commonly used nucleases are zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the recently discovered clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated Cas9 nuclease system (Cheng and Alper 2014). Although they can be used for integrating the recombinant transgene, the nucleases are

more commonly used as a genome editing tool for generating engineered cell lines for protein production (Fan et al. 2012).

The challenge for these methodologies, however, is to identify a "hotspot" in the genome that supports sufficiently stable and high production of biologic products with a low transgene copy number, which makes it a difficult technique to adopt for recombinant protein production (Table 1.1) (Crawford et al. 2013).

Table 1.1. Summary of advantages and disadvantages of the different gene integration methods for cell line generation

Method for gene integration/Cell line generation		
Conventional method (Random non-homologous integration)	**Advantages**	Simple to use
	Disadvantages	Low integration efficiency/random integration
		Large screen size required to identify suitable clone
		Labor intensive and time consuming
Site directed method (e.g. Cre/loxP, Flp/FRT)	**Advantages**	Site directed integration
		Low integrated copy number/improved stability
		Potential reduction in screen size to identify suitable clone
	Disadvantages	Complex & requires specific host cell line generation
		Low integration efficiency
Cis-regulating elements method (e.g. UCOE, MAR, STAR)	**Advantages**	Simple to use
		Potential reduction in screen size to identify suitable clone
	Disadvantages	Random integration
		Multiple options/requires optimization
Transposon based method	**EVALUATED DURING THE COURSE OF THIS THESIS**	

1.4.4. Transposon systems

DNA transposon elements have recently been used to catalyze the integration of recombinant genes into mammalian cells for recombinant protein production (Alattia et al. 2013; Li et al. 2013; Matasci et al. 2011). DNA transposons are naturally occurring genetic elements with the ability to move and replicate within the host. They have been used as tools for non-viral gene delivery by insertional mutagenesis and transgenesis with applications in gene therapy (Izsvak and Ivics 2004; Meir et al. 2013). Transposon systems are a particularly attractive option among non-viral vectors due to their ability to integrate

single DNA copies with high frequency at multiple loci (Ivics et al. 2009). The class II transposons are most commonly used for research and gene therapy. They function by the means of a "cut-and-paste" mechanism. The system consists of two components; a transposase and a pair of inverted terminal repeats (ITR) that carry the transposase binding sites necessary for transposition. The dual vector systems that are typically used consist of a donor vector, carrying an artificial transposon with a mammalian expression cassette for the recombinant transgene and selection marker flanked by the ITRs, and a helper vector for transient expression of the transposase (Ding et al. 2005). Co-transfection of the two vectors permits an efficient "cut and paste" transfer of the transposon from the donor plasmid into the genome (Fig. 1.1) (Wilson et al. 2007). Such transposon systems can be effectively used for the generation of recombinant cell lines for protein production; however, only limited progress has been made in this field (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013; Matasci et al. 2011).

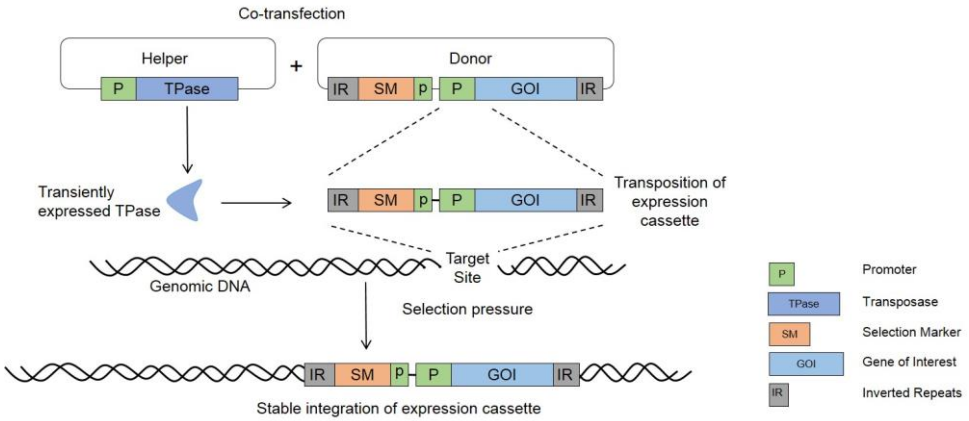

Figure 1.1. Schematic representation of the transposition of an artificial transposon by transiently produced transposase for the generation of stable cell pools.

The most widely used transposon systems for gene transfer in mammalian cells are piggyBac (PB), Tol2, and Sleeping Beauty (SB) (Mates et al. 2007). The PB transposon is derived from the cabbage looper moth *Trichoplusia ni*, and has been shown to transpose efficiently in mice and human cells (Wilson et al. 2007; Wu et al. 2006). The hAT-like (hobo, Activator, Tam3) Tol2 transposon, isolated from the Japanese medaka fish (*Orizyas latipes*), is the only known naturally active vertebrate transposon (Koga et al. 1996). It is a standard tool for manipulating the zebrafish genome and has been shown to transpose

effectively in frog, chicken, mouse, and human cells (Kawakami 2007). The SB transposon was reconstructed from inactive copies of Tc1/mariner like elements found in a fish genome (Ivics et al. 1997). The SB system has been extensively used as a tool for genetic modifications in somatic tissues of a wide range of vertebrate species including humans and the germline of fish, frogs, mice, and rats (Hackett et al. 2005; Huang et al. 2010; Mates et al. 2007). Significant efforts have been made to enhance the transposition efficiency for PB and SB transposases resulting in very effective transposases, namely mPB and SB100X, respectively (Cadinanos and Bradley 2007; Mates et al. 2009).

Apart from having different phylogenetic origins, the three transposon systems also differ in their biochemical properties that affect their activities. Different transposon families display various DNA insertion sequence preferences (Table 1.2). The Tc1/mariner family targets the dinucleotide TA and the piggyBac family targets TTAA tetranucleotides whereas the Tol2 system has a heterogenic target sequence site (Grabundzija et al. 2010). Another main property is overproduction inhibition (OPI), which limits the transposition activity with increasing intracellular transposase concentration. The susceptibility of SB system to OPI is better documented as compared to that of PB and Tol2 (Balciunas et al. 2006; Grabundzija et al. 2010; Lohe and Hartl 1996; Wilson et al. 2007; Wu et al. 2006). Another important factor to consider is the cargo size limitation of the transposons (Table 1.2). Currently, the SB system seems to be the most limited with 50 % reduction in transposition efficiency with cargo size of 6 kB compared to 2 kb (Geurts et al. 2003). Tol2 and PB transposases are known to be able to integrate up to 10 and 14 kb, respectively, without a significant reduction in transposition efficiency (Balciunas et al. 2006; Ding et al. 2005). Comparisons of integration profiles of the three TP systems in human cells showed that PB and Tol2 transposons integrate in intragenic regions and near transcriptional start sites whereas SB transposon integration occurred randomly across the genome, nevertheless in transcriptionally active regions (Grabundzija et al. 2010; Huang et al. 2010; Meir et al. 2011). This reduces the risk of the position effect and transcriptional silencing of the integrated transgene.

Table 1.2. Differences between three transposon systems

	PiggyBac	Tol2	Sleeping Beauty
Origin	Cabbage looper moth (*Trichoplusia ni*)	hAT superfamily-endogenous in medaka fish	Inactive Tc1/mariner-like elements found in fish genome
Commonly used in	Mouse, human cells	Zebrafish (mainly), frog, chicken, mouse cells	Fish, frogs, mice, rats
Transposase gene size	~1.8 kb	~2 kb	~1 kb
Maximum transposon size	Up to 14 kb	>10 kb	~6 kb
Target site	TTAA	-	TA
AT % in ITR	60 – 70%	77%	67%

There have been several studies comparing the activities and integration profiles of PB, Tol2, and SB transposon systems in human cells for applications in gene therapy (Grabundzija et al. 2010; Huang et al. 2010; Meir et al. 2011; Wu et al. 2006). Recently, the PB transposon system has been explored for the generation of stable pools and clonal cell lines for recombinant protein production (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013; Matasci et al. 2011) . It has been shown that recombinant protein production and the efficiency of cell pool and cell line generation can be significantly improved using the PB transposon system in comparison to cell line generation by plasmid integration by DNA recombination (Matasci et al. 2011).

A summary of the advantages and disadvantages of the various DNA integration methods are shown in Table 1.1 and these of the transposon systems will be evaluated during this thesis.

1.5. Thesis Objectives

The goal of this thesis is to study transposon-mediated recombinant cell pool and cell line generation and to determine if this technology results in efficient recombinant protein production with CHO cells. The specific aims can be broadly classified as:

1. Characterization of PB, Tol2, and SB transposon systems for the generation of cell pools and cell lines
2. Generation and characterization of PB-mediated multi-transgene expressing cell pools

The gist of work done in each chapter are as follows:

1.5.1. Optimization of parameters for efficient transposon mediated cell pool generation and comparison with traditional methods

PB transposon has been recently shown to improve the efficiency of recombinant CHO cell line generation (Matasci et al. 2011). However, use of the PB transposon system for cell pool generation was not explored. This chapter focuses on identifying the critical parameters influencing the generation of cell pools with CHO-DG44 cells. Both the duration of the selection period and the time post-transfection for the initiation of selection were investigated. In cell pools generated by conventional transfection methods, the selection agent is known to influence the selection stringency, which may affect volumetric productivity and the stability of transgene expression (Wurm 2004). Hence, four different antibiotics were assessed for the generation of cell pools. After an optimized selection procedure, a simple batch process to assess TNFR-Fc productivities was performed. The stability of protein production with cell pools is an important factor and hence was studied over a period of 90 days in culture. Further experiments were designed and executed to evaluate the integration events in PB-generated cell pools.

1.5.2. Comprehensive study of PB, Tol2 and SB transposon systems for recombinant protein production

This chapter focuses on the comparison of the PB, Tol2 and SB transposon systems for the generation of cell pools and cell lines with CHO-DG44 cells. The overall aim was to study the similarities and differences of the three transposon systems for recombinant protein production. We used the mouse-codon-optimized PB transposase gene (mPBase), a mammalian codon-optimized Tol2 transposase, and the hyperactive SB transposase (SB100X) for this study (Cadinanos and Bradley 2007; Mates et al. 2007). The ratio of helper and donor plasmid DNA necessary for efficient generation of cell pools was evaluated. Real time quantitative PCR (RT-qPCR) was employed for analysis of the integrated transgene copy number and the steady state mRNA levels in comparison with cell pools derived from conventional plasmid transfection. Cell lines were recovered from cell pools generated using all three transposon systems. The stability of transgene expression in cell pools and cell lines was also studied for a period of up to 3 months in the absence of selection. Finally, cell pools and cell lines were subjected to a simple fed-batch production process to assess volumetric productivities.

1.5.3. Study of the three transposon systems for cell pool generation in different CHO cell strains

The heterogeneity of cell lines is a recognized problem which makes the standardization of protein production strategies difficult. CHO-DG44, CHO-K1 and CHO-S are the most commonly used CHO cell strains in the biopharmaceutical industry. Although these cell lines share a common lineage, they are known to exhibit different phenotypes in cell culture. To improve the applicability of transposon systems for cell pool and cell line generation, the PB, Tol2 and SB transposon systems were used to generate cell pools in CHO-DG44, CHO-K1 and CHO-S cells. An analysis of the integrated transgene copy number was performed to compare the activities of the three transposons in the three different CHO strains. The stability of protein production for the cell pools was analyzed over a period of 2 months.

1.5.4. Cell pools expressing multiple transgenes using the PB transposon for gene delivery

In biomedical research and drug development there is often a need for multiple genes to be co-expressed. Conventional approaches include transfection of host cells with multiple expression cassettes either carried in a single plasmid or in multiple plasmids. This often results in inefficient integration and expression of transgene due to the low efficiency of plasmid integration events. Another method involves sequential transfection followed by recombinant cell cloning for each individual transgene (McPhaul and Berg 1986; Paul et al. 1985). In this chapter, we studied the use of the PB transposon system for the simultaneous delivery of multiple genes into CHO-DG44 cells for obtaining cell pools co-expressing multiple proteins, with the goal to simplify multi-transgene cell pool and cell line generation. Two separate strategies were tested in the presence and absence of the PB transposon. First, all the transgenes were selected with a single antibiotic. Then the different transgenes were each selected with a different antibiotic. The average integrated copy number of each transgene in each cell pool was analyzed. Cell lines were recovered from the cell pool expressing 3 proteins. The percentage of cells co-expressing all the transgenes transfected in a pool were also evaluated using three different fluorescent proteins, EGFP, Cerulean, and mCherry, simultaneously.

2.

Material and Methods

2.1. Routine cell culture

Suspension-adapted CHO-DG44, CHO-K1 or CHO-S cells were grown in serum-free ProCHO5 medium (Lonza AG, Verviers, Belgium) supplemented with 13.6 mg/L hypoxanthine, 3.84 mg/L thymidine, and 4 mM glutamine (SAFC Biosciences, St. Louis, MO). The cells were maintained in 5-10 mL of medium in 50-mL Tubespin bioreactors (TS50) (TPP AG, Trasadingen, Switzerland) or in 100 mL of medium in glass bottles agitated by orbital shaking as previously described (Muller et al. 2005). The cells were passed twice per week at an inoculation density of 3×10^5 cells/mL. The cell number and viability were assessed manually by the Trypan Blue exclusion method.

2.2. Plasmids

The donor vector pMP-PB (Fig. 2.1A) and the helper vector pmPBase have been described previously (Matasci et al. 2011). In pMP-PB the transgene and the puromycin N-acetyl-transferase (*pac*) gene were expressed from the mouse cytomegalovirus major immediate early promoter (p-mCMV) and the herpes simplex virus thymidine kinase promoter (p-HSV-tk), respectively. The human elongation factor 1 alpha first intron was present in the 5' untranslated region (5' UTR) of the transgene mRNA. The bovine growth hormone polyadenylation element (BGH-pA) was used for termination and 3' end processing of the transgene mRNA. For each PB donor plasmid, the artificial transposon, enclosed between the inverted terminal repeats (ITRs) specific for the PB transposon system, consisted of the expression cassette of the transgene and that of the selection marker (Fig. 2.1A). The donor plasmids expressing different transgenes were constructed by introducing the respective transgenes into the multiple cloning site (MCS) of pMP-PB. The light and heavy chains for the IgG1 monoclonal antibody were obtained from pMP-PB-LC and pMP-PB-HC, respectively (Matasci et al. 2012). The gene for the fluorescent proteins Cerulean and mCherry were obtained from Dr. Iacovos Michael (EPFL) in the plasmids, PB-CA-Cerulean and PB-CA-mCherry, respectively. The gene for secreted alkaline phosphatase (SEAP) was obtained from Prof. Fussenneger (ETH, Zurich). Donor plasmids coding for resistance against different antibiotics were constructed by replacing the *pac* gene with the appropriate antibiotic resistance gene.

The donor plasmids for the Tol2 and SB transposon systems were constructed such that the only difference between them was the identity of the ITR sequences. The ITRs for Tol2 transposon system were cloned from the plasmid pMiniTol2-CMVpA, obtained from the University of Wuerzburg (Germany) (Balciunas et al. 2006). The ITRs for the SB transposon system were cloned from the vector pT2/HB, purchased from Addgene (Middlesex, UK). The name and organization of the resulting plasmids are provided in Table 2.1. The schematic representation of the resulting plasmids is shown in Fig 2.1A.

The helper plasmid pmPBase carries an expression cassette for the transient production in mammalian cells of a codon-optimized PBase gene. The gene was transcribed from the human cytomegalovirus major immediate early promoter (p-hCMV). The other helper plasmids, pmTol2ase and pSBase were constructed by replacing the mPBase gene in pmPBase with a mammalian codon-optimized Tol2 transposase gene and the SB100X transposase gene to generate pmTol2ase and pSBase, respectively (Fig. 2.1B). The enzyme Tol2ase was cloned from the plasmid pKate-N-Tol2 obtained from the University of Wuerzburg (Germany) (Balciunas et al. 2006). This sequence was codon-optimized for mammalian cells by Life Technologies. The SB100X gene was cloned from the vector pCMV(CAT)T7-SB100, purchased from Addgene (Middlesex, UK).

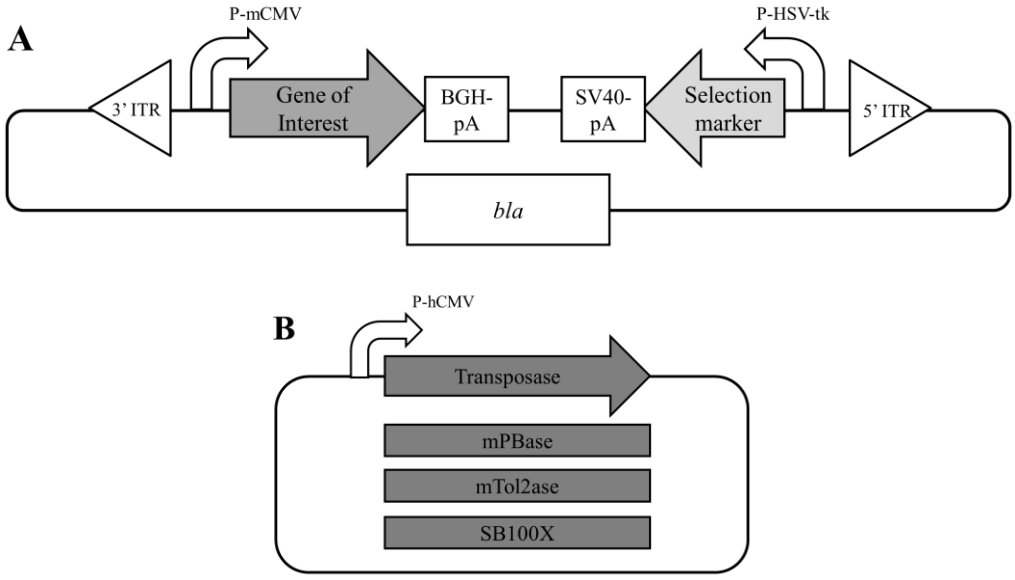

Figure 2.1. Representation of the maps of donor and helper vectors. (A) Donor pDNAs for different transposon systems were generated by substituting the respective 3' and 5' ITRs in the same backbone. Different genes of interest were cloned into each plasmid to create the donor vector for each transposon system expressing the respective protein. (B) All three transposase enzymes were cloned into the same expression cassette driven by human CMV promoter. Abbreviations: p-mCMV – mouse cytomegalovirus major immediate early promoter; p-HSV-TK – herpes simplex virus thymidine kinase promoter; ITR – inverted terminal repeat; BGH-pA – bovine growth hormone polyadenylation element; SV40-pA - simian virus 40 polyadenylation element; *bla* – beta-lactamase gene.

Table 2.1. List of plasmids used

Gene of Interest[a]	Plasmid Name	Selection gene/antibiotic	Transposon ITRs
TNFR:Fc	pMP-PB-TNFR:Fc	*pac*/ Puromycin	PiggyBac
EGFP	pMP-PB-EGFP	*pac*/ Puromycin	PiggyBac
IRES-EGFP	pMPPB-IRES-EGFP	*pac*/ Puromycin	PiggyBac
TNFR:Fc	pSB-BPB-TNFR:Fc	*bsr*/ Blasticidin	PiggyBac
TNFR:Fc	pSB-HPB-TNFR:Fc	*hph*/ Hygromycin	PiggyBac
TNFR:Fc	pSB-ZPB-TNFR:Fc	Sh *ble*/ Zeocin	PiggyBac
TNFR:Fc-IRES-EGFP	pMPIG-TNFR:Fc	*pac*/ Puromycin	PiggyBac
IgG - LC	pMP-PB-LC	*pac*/ Puromycin	PiggyBac
IgG - HC	pMP-PB-HC	*pac*/ Puromycin	PiggyBac
SEAP	pSB-PPB-SEAP	*pac*/ Puromycin	PiggyBac
IRES-EGFP	pSB-BPB-IRES-EGFP	*bsr*/ Blasticidin	PiggyBac
SEAP	pSB-HPB-SEAP	*hph*/ Hygromycin	PiggyBac
IgG - HC	pSB-ZPB-HC	Sh *ble*/ Zeocin	PiggyBac
Cerulean	pSB-PPB-Cer	*pac*/ Puromycin	PiggyBac
mCherry	pSB-PPB-mCh	*pac*/ Puromycin	PiggyBac
EGFP	pSB-PTol2-EGFP	*pac*/ Puromycin	Tol2
IRES-EGFP	pSB-PTol2-IRES-EGFP	*pac*/ Puromycin	Tol2
TNFR:Fc	pSB-PTol2-TNFR:Fc	*pac*/ Puromycin	Tol2
EGFP	pSB-PSB-EGFP	*pac*/ Puromycin	Sleeping Beauty
IRES-EGFP	pSB-PSB-IRES-EGFP	*pac*/ Puromycin	Sleeping Beauty
TNFR:Fc	pSB-PSB-TNFR:Fc	*pac*/ Puromycin	Sleeping Beauty
TNFR:Fc	pMP-TNFR:Fc	*pac*/ Puromycin	-None-

[a]Abbreviations: *bsr*: blasticidin S-resistance gene; *hph*: hygromycin B phosphotransferase gene; Sh *ble*: *ble* gene from Streptoalloteichus hindustanus; *bla*: beta-lactamase gene; IgG – LC: immunoglobulin G light chain; IgG – HC: immunoglobulin G heavy chain; TNFR:Fc: tumor necrosis factor receptor Fc fusion; IRES: internal ribosome entry site (from encephalomyocarditis virus (EMCV)); EGFP: enhanced green fluorescent protein; SEAP: secreted alkaline phosphatase.

2.3. DNA Transfection

Suspension-adapted CHO-DG44 cells were transfected using linear 25 kDa polyethylenimine (PEI) (Polysciences, Eppenheim, Germany) prepared as a solution of 1 mg/mL in deionized water at pH 7.0. The day before transfection, the cells were seeded in ProCHO5 at a density of 2 x 10^6 cells/mL and grown overnight in an orbitally shaken glass bottle as described (Muller et al. 2005). On the day of transfection, cells were centrifuged and resuspended in 5 mL of ProCHO5 in TS50s at a density of 3 x 10^6 cells/mL. To each culture, 22.5 µg of plasmid DNA and 45 µg of PEI were separately added (Rajendra et al. 2011). If necessary, sheared herring sperm DNA (Invitrogen AG, Basel, Switzerland) was used as a non-specific (filler) DNA to keep the total amount of transfected DNA constant (Rajendra et al. 2012). The tubes were then incubated with agitation at 37°C at 180 rpm.

2.4. Generation of recombinant cell pools and cell lines

In the standard protocol for the generation of cell pools, cells were co-transfected with the donor and helper plasmid DNA at a ratio of 9:1 (w/w) unless mentioned otherwise. Control transfections in the absence of a helper plasmid were performed with filler DNA at the same ratio as above. At two days post-transfection, the cells were inoculated at a density of 5 x 10^5 cells/mL in 5 mL of ProCHO5 and 10 µg/mL puromycin in TS50s. Cells were passed every 3-5 days for 10 days in selective medium by inoculation at 5 x 10^5 cells/mL (Fig. 2.2). After the removal of selection pressure, the cell pools were allowed to recover for 2 -3 days. The cells were then either subjected to a 4-day batch process or a 2-week fed-batch process by seeding at 3x10^5 cells/mL (Fig. 2.2). For the 2-week fed-batch cultures of cell pools and clonal cell lines, cells were seeded at a density of 5 x 10^5 cells/mL at 37 °C. After three days, the cultures were fed with 3 g/L glucose and 1.5 g/L essential and non-essential amino acids and treated with 0.25 % dimethyl acetamide (DMA).Incubation was continued at 31°C.

To study the stability of transgene expression over time, the cell pools were maintained in ProCHO5 without selection for up to 3 months.

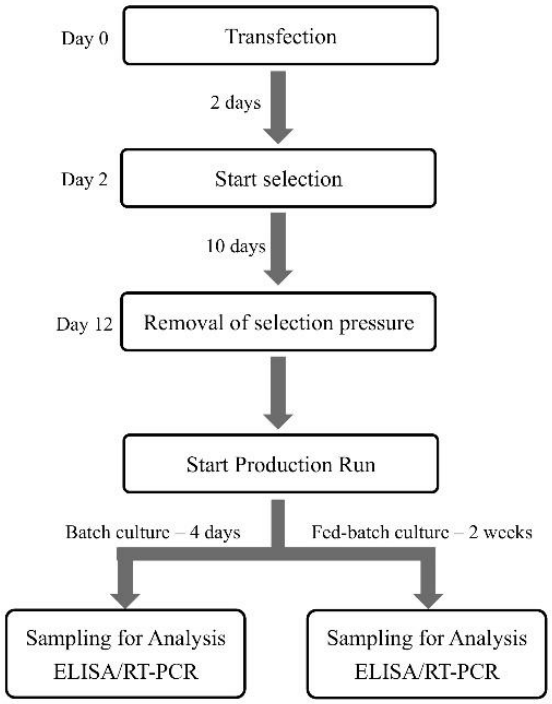

Figure 2.2. Schematic diagram of the standard protocol for the generation of cell pools.

Clonal cell lines expressing TNFR:Fc were recovered by limiting dilution from cell pools after 10 days of selection in puromycin. Briefly, cells were seeded in 96-well plates with 100 µl of ProCHO5 medium (10 % conditioned medium) to a cell density to allow the plating of 0.5 cells per well. After 15-20 days in culture, individual colonies of cells were transferred to a 24-well plate and cultivated in suspension by agitation of the plates on an orbital shaker. After one cell passage in 24-well plates, aliquots of cells were seeded in 24-well plates at a 1:10 dilution in ProCHO5 medium (approximately 0.5 x 10^6 cells/mL). The TNFR:Fc concentration in the medium after 4 days was measured by ELISA as described below.

2.5. ELISA

TNFR:Fc and IgG concentrations in cell culture medium were directly determined by sandwich ELISA using a modified version of a published protocol (Meissner et al. 2001). Briefly, a goat anti-human IgG or an Fc-specific goat anti-human IgG $F(ab')_2$ fragment (Jackson ImmunoResearch Laboratories, Inc. West Grove, PA, USA) was used to coat 96-

well plates for the analysis of IgG and TNFR:Fc, respectively. For both assays, captured protein was detected with alkaline phosphatase-conjugated goat anti-human gamma chain IgG (Life Technologies) using p-nitrophenyl phosphate (p-NPP) (AppliChem GmbH, Darmstadt, Germany) as a substrate. Absorption was measured at 405 nm using a microplate reader (SPECTRAmax TM340; Molecular Devices, Palo Alto, CA, USA). Human TNFR:Fc (Enbrel®) (Amgen, Zug, Switzerland) and human IgG (Jackson ImmunoResearch Laboratories, Inc.) were used to generate standard curves.

2.6. Flow Cytometry

EGFP measurements were performed using a guava easyCyte® microcapillary flow cytometer (Merck-Millipore, Schaffhausen, CH) with excitation and emission wavelengths of 488 and 532 nm, respectively. Cells were prepared for analysis by dilution in PBS to densities of 100 – 400 × 10^5 cells/mL. Each sample was measured in duplicate with a minimum of 5,000 recorded events.

Fluorescence from cell pools expressing Cerulean, EGFP and mCherry was measured using the BD™ LSR-II flow cytometer (BD Biosciences). The excitation and emission wavelengths used for Cerulean, EGFP, and mCherry were 405 – 450/50 nm, 488 – 525/50 nm, and 561 – 610/20 nm, respectively. Cells were prepared for analysis by dilution in PBS to densities of 400 – 1000 × 10^5 cells/mL. Each sample was measured in duplicate with a minimum of 50,000 recorded events.

2.7. Quantification of secreted alkaline phosphatase (SEAP) activity

The SEAP activity in cell culture medium was determined by colorimetry using a modified version of a published protocol (Schlatter et al. 2002). Briefly, an aliquot of cell culture was centrifuged, and 50-µL of the supernatant was transferred into a chamber of a 96- well plate. The plate was then incubated at 37 °C. The substrate, p-nitrophenyl phosphate (AppliChem GmbH, Darmstadt, Germany), was added to 2X SEAP assay buffer (1 mM $MgCl_2$, 21% diethanolamine, pH 9.8) to a final concentration of 12 mM. The solution was heated to 37°C and added to each well having cell culture supernatant. Absorbance readings were taken at 405 nm using a microplate reader (SPECTRAmax TM340; Molecular

Devices, Palo Alto, CA, USA) every minute for 15 min, and the maximum linear reaction rate was determined in order to calculate the SEAP activity as described (Schlatter et al. 2002).

2.8. Quantification of transgene copy number

For each analysis, 1 x 10^6 cells were collected by centrifugation, washed twice with cold PBS, flash frozen in liquid nitrogen, and stored at -80 °C. Total DNA was extracted from cells using the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Real time quantitative PCR (RT-qPCR) was carried out in a LightCycler 480 Real-Time PCR System (Roche Applied Science, Basel, Switzerland) with the ABsolute QPCR SYBR Green ROX mix (Axon Lab AG, Baden-Dättwil, Switzerland) according to the manufacturer's instructions. Oligonucleotide primers for amplification of each gene are listed in Table 2.2. Standard curves to determine the copy number of each gene were created with known amounts of the appropriate plasmid following its quantification by UV absorbance.

Table 2.2. Oligonucleotide primers used in this study for RT-qPCR.

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
TNFR:Fc	GCCAGACCAGGAACTGAAAC	GTGGATGAAGTCGTGTTGGA
bla	ACGATCAAGGCGAGTTACATGA	ACACTGCGGCCAACTTACTTCT
EGFP	GTCTATATCATGGCCGACAAGC	GGTGTTCTGCTGGTAGTGGTC
SEAP	GTATGTGTGGAACCGCACTG	GAGTCTCGGTGGATCTCGTATT
IgG – LC	CTGGAATTCACCATGAGTGTGC	CTACAGATGCAGACAGGGAGGA
IgG – HC	ACAGGTGAAACTGCTCGAGTCT	TGAAGGTGAATCCAGACGCTAC
CHO β – actin	GCTCTTTTCCAGCCTTCCTT	GAGCCAGAGCAGTGATCTCC

2.9. Quantification of transgene mRNA

At the time of sampling for ELISA from a 4-day batch culture, 1 x 10^6 cells were recovered and stored frozen at -80 °C. After thawing the cells, total RNA was extracted from cells using the GenElute mRNA kit (SAFC Biosciences) according to the manufacturer's protocol. The samples were treated with 1 unit of RNase-free DNase I (Invitrogen) for 15 min at room temperature. The DNase I was inactivated by incubation at 75 °C for 15 min. DNA-free RNA (1 mg) was reverse transcribed using M-MLV reverse transcriptase (SAFC

Biosciences) primed with oligonucleotide dT. RT-qPCR was performed as described in section 2.8. The oligonucleotide primers for amplification of the TNFR:Fc and cellular β-actin cDNAs are listed in Table 2.2. The relative amount of TNFR:Fc mRNA was determined using the $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak 2008).

3.

Identifying Parameters for Efficient Transposon-Mediated Cell Pool Generation

3.1. Introduction

Heterogeneous populations of stably transfected cells (cell pools) can be used for rapid production of moderate amounts of recombinant proteins. They have a very low DNA requirement for transfection, are easily scalable, and have the potential for higher volumetric yields compared to TGE and have considerably shorter developmental time frames than SGE. Limited efforts have been made towards the development of efficient cell pool generation methods. Published articles on the generation of cell pools describe selection schemes of relatively long duration, low volumetric protein productivities, and protein production instability as the key issues (Fan et al. 2013; Ye et al. 2010).

The PB transposon system has been shown to have a high integration efficiency in mammalian cells (Ding et al. 2005; Wilson et al. 2007; Wu et al. 2006). This transposon is known to integrate into transcriptionally active regions of the cell's genome, which is expected to improve transgene transcription levels and recombinant protein productivities (Galvan et al. 2009; Grabundzija et al. 2010; Meir et al. 2011; Wilson et al. 2007). The PB transposon system has been recently used for the generation of recombinant CHO and HEK293 cell lines and HEK293 cell pools with considerable success (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013; Matasci et al. 2011). However, the PB system has not been used for the generation of CHO cell pools.

Here, we aim to establish methods for PB-mediated cell pool generation. The PB system consists of one plasmid (helper vector) for the transient expression of the PBase gene and a second plasmid (donor vector) carrying the 5′ and 3′ inverted repeat elements (5′IR and 3′IR) delimiting the ends of the artificial transposon carrying the GOI (Kahlig et al. 2009; Meir et al. 2011). To begin, several parameters including the duration of selection and the ratio of helper:donor plasmid DNA were optimized. The choice of antibiotic and its concentration are known to influence the selection stringency, thereby affecting volumetric productivity and the stability of transgene expression (Wurm 2004). Here, we chose to evaluate 4 different antibiotics for selection, namely puromycin, blasticidin, hygromycin B, and zeocin for the optimization of the selection conditions for cell pool generation. Using these conditions, cell pools were generated for the expression of TNFR:Fc, and the

volumetric productivity was determined in a 14-day batch process. Finally, the stability of protein production in the absence of selection was studied over a period of 3 months.

3.2. Results

3.2.1. Optimization of selection duration and initiation for PB-mediated cell pools

Starting with a previously published method for the selection of CHO cells following transfection with the PB transposon system, we optimized the duration of selection for the generation of CHO cell pools (Matasci et al. 2011). CHO-DG44 cells were co-transfected with pMPIG-TNFR:Fc and pmPBase at a ratio of 9:1 (w/w). At 2 days post-transfection, the cells were grown in the presence of 10 µg/mL puromycin for up to 10 days. The percentage of viable cells in the population was determined daily. Aliquots of cells were also removed from the culture each day and grown in the absence of selection until the cell viability reached at least 90 %. Then, the volumetric productivity of each cell pool was evaluated in a 4-day batch culture that was inoculated at 3 x 10^5 cells/mL in the absence of selection. The volumetric TNFR:Fc productivity of the cell pools increased with the duration of selection up to day 5 and then remained at a similar level for selection durations up to 10 days (Fig. 3.1A). However, the cell viability in the selected cell population reached >90 % only after 8 days of selection (Fig. 3.1A). For this reason, we chose a selection duration of 10 days for all future experiments.

The timing of the start of selection with puromycin was investigated next. The cells were co-transfected with pMPIG-TNFR:Fc and pmPBase as described above. At 1, 2 or 3 days post-transfection, selection was started with 10 µg/mL puromycin and was continued for a total of 10 days. The TNFR:Fc productivity of the three resulting cell pools was measured in 4-day batch cultures that were inoculated at 3 x 10^5 cells/mL. All three cell pools produced the same level of TNFR:Fc (Fig. 3.1B), indicating that selection could begin from 1 – 3 days post-transfection for cell pool generation. We arbitrarily chose 2 days post-transfection to initiate selection with puromycin for all future experiments.

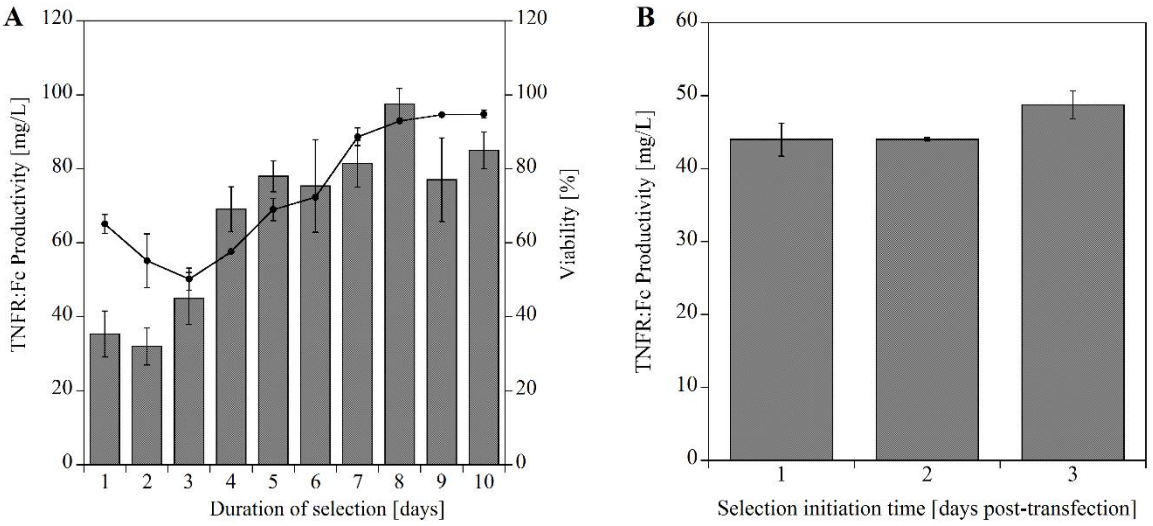

Figure 3.1. Investigation of the duration and start of selection for cell pool generation. Cells were co-transfected in duplicate with pMPIG-TNFR:Fc and pmPBase at a ratio of 9:1 (w/w). (A) The cells were selected in 10 µg/mL puromycin for 1 – 10 days as indicated. The cell viability of the populations under selection was measured daily (black line). Each day, cells were removed from the selected culture and grown in the absence of selection for up to 10 days. The volumetric TNFR:Fc productivity of each cell pool was measured by ELISA at the end of a 4-day batch culture (bars). (B) The cells were grown in 10 µg/mL puromycin starting at day 1, 2 or 3 post-transfection and then selected for 10 days. The volumetric productivity of the pools was analyzed from 4-day batch cultures using ELISA.

3.2.2. Optimization of donor:helper vector ratio for efficient PB-mediated cell pool generation

Different ratios of donor and helper plasmids were tested to find the optimal condition for the efficient generation of PB-mediated clones. Cell pools expressing TNFR:Fc were generated by co-transfection with different amounts of the helper vector pmPBase ranging from 0.01 % to 50 % of the total transfected DNA. The donor vector pMP-PB-TNFR:Fc was used to complete the total transfected DNA to 100%. A control transfection with donor pDNA alone in the absence of any helper plasmid was also performed. The resulting cell pools were subjected to a 4-day batch culture.

A dose-dependent increase in TNFR:Fc productivity was observed as an increasing amount of pmPBase was transfected (Fig. 3.2). The volumetric productivity remained constant for transfection conditions with ≥1 % of pmPBase. The productivity of the PB-mediated cell pools was 4-5 fold higher than for the cell pool generated by conventional transfection (100 % donor plasmid) (Fig. 3.2). This suggested that 1 % of helper plasmid was sufficient to

effectively generate PB transposon – mediated cell pools. However, there was a 2-fold increase in cell pool productivity when 0.1 % pmPBase was used as compared to 1 % (Fig. 3.2). To avoid transfecting under sensitive conditions where the change of external parameters (e.g. media) may affect DNA uptake, we chose to use 10 % of helper plasmid for future experiments.

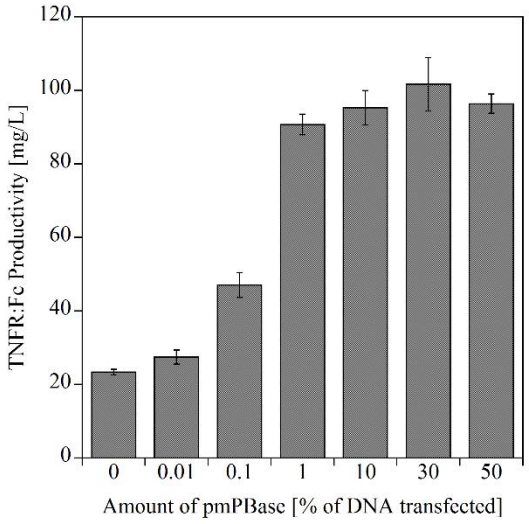

Figure 3.2. Optimization of the amount of helper plasmid for cell pool generation. Cell pools expressing TNFR:Fc were generated by transfecting cells with different amounts of pmPBase, ranging from 0 % to 50 % of the total amount of DNA. The donor vector was added at 100 % to 50% of the total DNA. Each transfection was conducted in duplicate and each ELISA measurement was performed in duplicate. Thus, the error bars represent the standard deviation from 4 different measurements. The TNFR:Fc expression levels of the cell pools were analyzed from 4-day batch cultures.

3.2.3. Generation of PB-mediated cell pools by selection with different antibiotics

To determine if cell pools could be generated by selection with antibiotics other than puromycin, we constructed a family of TNFR:Fc expression vectors for selection in puromycin (pMPPB-TNFR:Fc), blasticidin (pSBBPB-TNFR:Fc), hygromycin B (pSBHPB-TNFR:Fc), and zeocin (pSBZPB-TNFR:Fc) as described in Chapter 2. Cells were co-transfected with one of these donor vectors and pmPBase at a ratio of 9:1 (w/w). As a control, cell pools were also generated by conventional plasmid transfection by replacing pmPBase with filler DNA. For each antibiotic, a range of concentrations was tested to identify the optimal conditions for cell pool generation. Since the selection agents resulted in different levels of cell killing, we maintained the cells under selection until the

cell viability increased to more than 90 %. This occurred after 7 – 14 days of selection for the four antibiotics tested (data not shown).

The volumetric TNFR:Fc productivity of each cell pool was measured at the end of a 4-day batch culture as previously described. For each antibiotic, the maximum TNFR:Fc level for the pools generated with the PB system was 3 – 4 times higher than the maximum level observed in the control cell pools (Fig. 3.3 A-D). In the presence of puromycin, concentrations of 5 – 10 µg/mL were found to be optimal for protein production (Fig. 3.3A). The optimal concentrations of blasticidin (10 µg/mL), hygromycin B (200 µg/mL), and zeocin (50 µg/mL) also resulted in cell pools with volumetric TNFR:Fc productivities of 80 – 90 mg/L (Fig. 3.3 B-D). These findings demonstrated that cell pools could be generated with similar volumetric productivities in the presence of different selection agents. However, we observed a decrease in the volumetric productivities of PB-mediated cell pools at the highest concentrations of puromycin and hygromycin B tested (Fig. 3.3A, C).

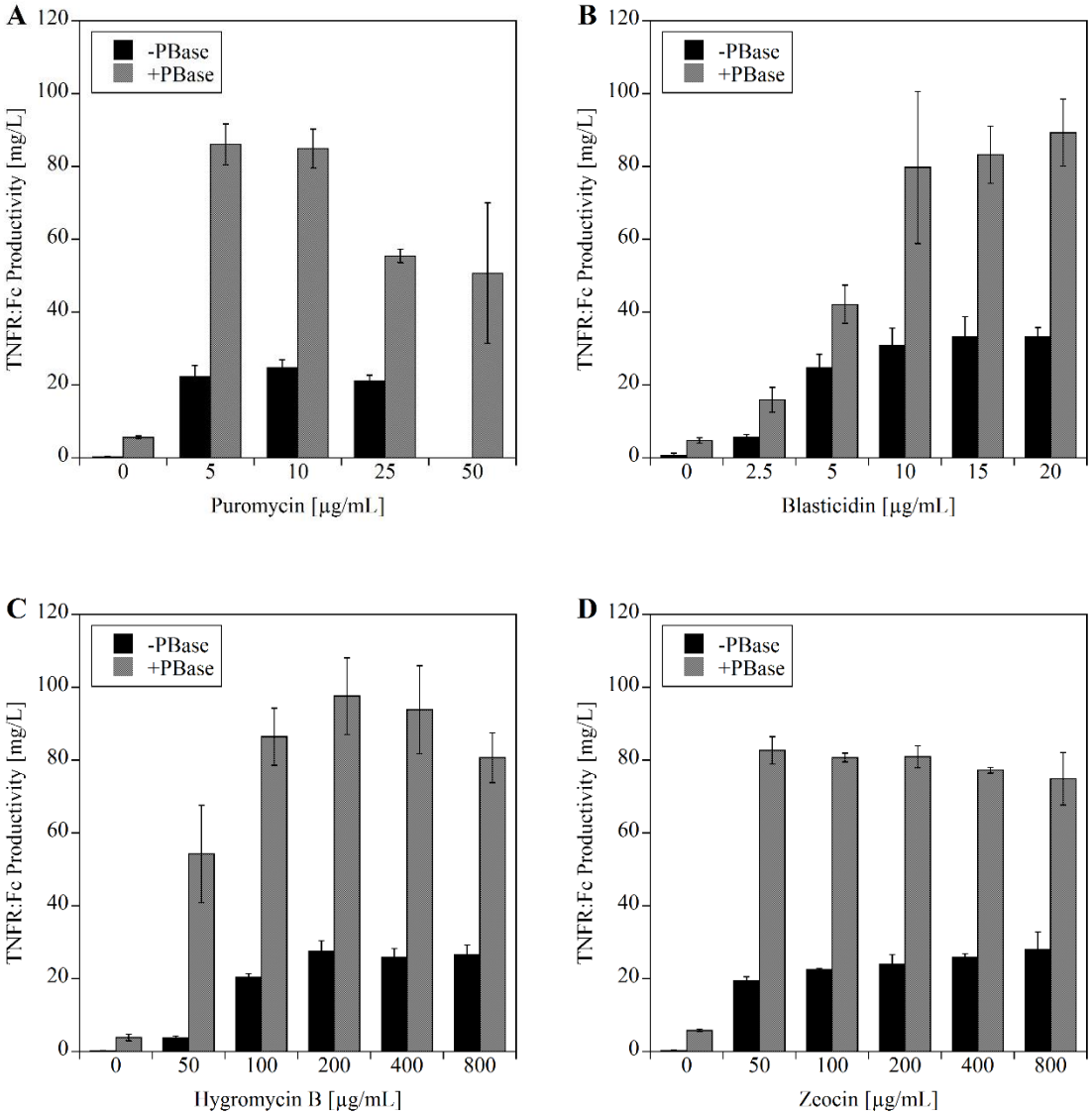

Figure 3.3. Volumetric TNFR:Fc productivity of cell pools generated in the presence of different antibiotics. The cells were co-transfected with either (A) pMP-PB-TNFR:Fc, (B) pSB-BPB-TNFR:Fc, (C) pSB-HPB-TNFR:Fc, or (D) pSB-ZPB-TNFR:Fc and either pmPBase (+PBase) or filler DNA (-PBase) at a 9:1 ratio (w/w). Transfected cells were selected in different concentrations of (A) puromycin, (B) blasticidin, (C) hygromycin B, or (D) zeocin. Cells were maintained under selective pressure until the cell viability reached 90% or more. The cells were then grown in the absence of selection for 10-14 days and used to inoculate 4-day batch cultures for the analysis of TNFR:Fc production by ELISA. The cells (-PBase) selected in the presence of 50 µg/mL puromycin did not survive.

3.2.4. Evaluation of transposition efficiency of PB system

Cells were co-transfected with pMPIG-TNFR:Fc and pmPBase at a ratio of 9:1 (w/w) to generate cell pools labeled PB. The control transfection was performed in parallel by replacing pmPBase with filler DNA to generate cell pools labeled TX. On day 2 post transfection, one set of transfected cells was placed under selective pressure with 10 µg/mL puromycin for 10 days, as per the protocol described in Chapter 2 (Fig. 2.2), resulting in cell pools labeled TX(2) and PB(2). To assess the integration efficiency of the transposon system in the absence of selection pressure, cell pools were generated by allowing the transfected cells to grow in the absence of selection until the episomal DNA was lost before applying the selective pressure. For this purpose, another set of transfected cells labeled TX(0) and PB(0) was maintained by regular subculture in the absence of selection and analyzed daily for EGFP-specific fluorescence using flow cytometry. These unselected cell pools were doubling regularly (data not shown). At 10 days post-transfection, the percentage of EGFP-positive cells had stabilized at ~20 % in the case of PB-mediated transfections and at ~3 % for transfections without PB, indicating the complete loss of transient EGFP expression in these cell pools (Fig. 3.4A). We hypothesized that loss of transient EGFP expression in the cell pools suggested the loss of episomal DNA in the cells. At 10 days post-transfection, the cell pools TX(0) and PB(0) were placed under selective pressure with puromycin (10 µg/mL) for 10 days to generate pools TX(10) and PB(10). The cell pools TX(0) and PB(0) were continued to be cultivated in the absence of selection for 30 days, with regular analyses of EGFP-specific fluorescence. In addition to the ~5 fold higher percentage of EGFP-positive cells in the PB(0) cell pool compared to TX(0), the mean EGFP expression (RFU) for PB(0) was also about 4-5 fold higher than for TX(0), once transient EGFP expression was terminated.

The cell pools generated with (PB) and without (TX) the use of transposon systems showed different viability trends when placed under selection starting from day 2 or day 10 (Fig 3.4B, C). The viability of TX(10) decreased to 10 % while that of TX(2) decreased to ~50 % (Fig. 3.4B). In contrast, irrespective of starting selection before (PB(2)) or after (PB(10)) the loss of transient EGFP expression, the cell pools had very similar patterns of cell viability during selection, reaching a minimum of ~50 % (Fig. 3.4C).

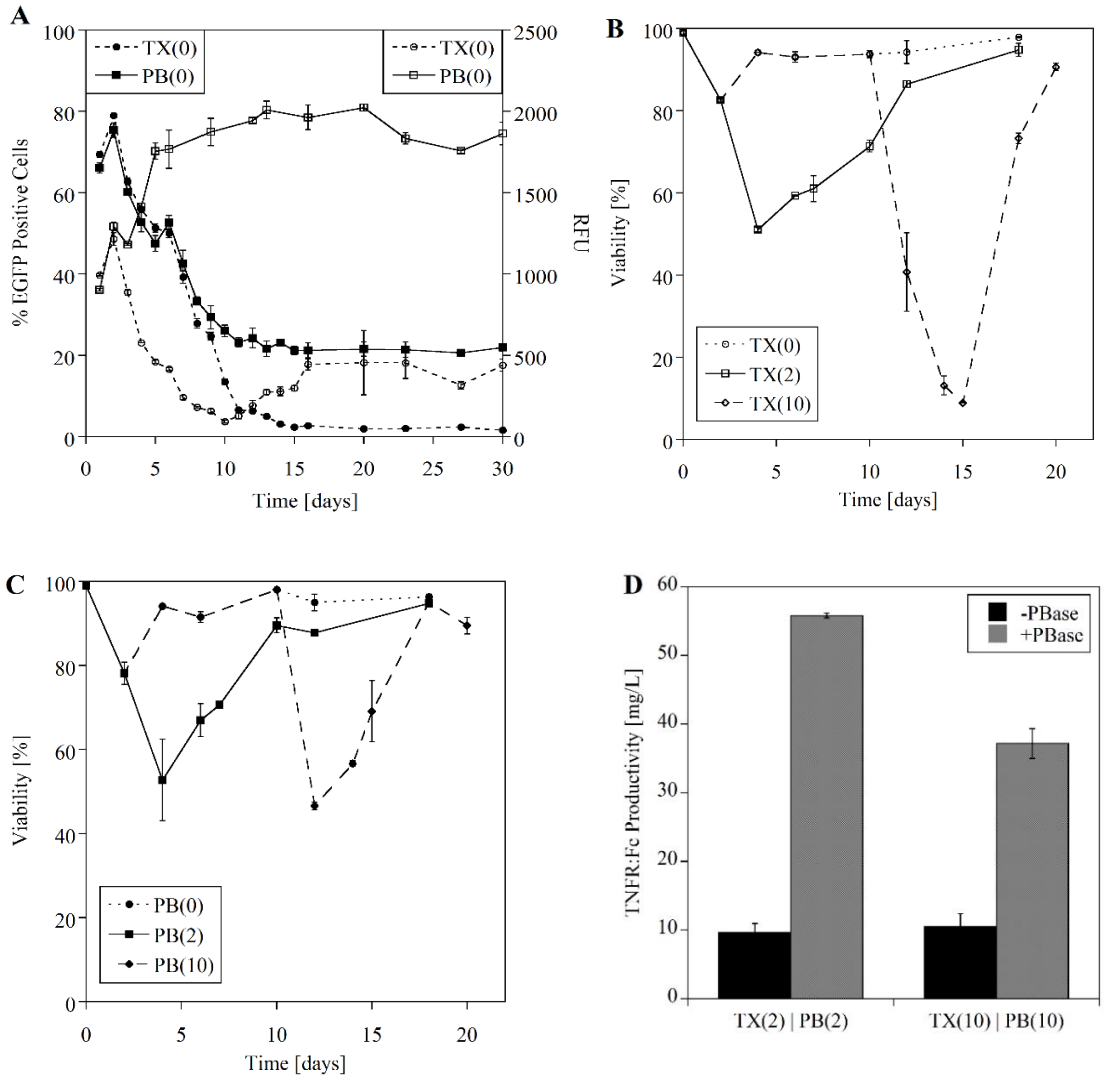

Figure 3.4. Viability and volumetric productivity of cell pools under different selection conditions. Cell pools were generated with pMPIG-TNFR:Fc and pmPBase (PB) or pMPIG-TNFR:Fc and filler DNA (TX) at a ratio 9:1 (w/w). Cell pools were placed under selection pressure on day 2 (TX(2) and PB(2)) or on day 10 (TX(10) and PB(10)) post-transfection. As a control, one set of transfections was cultivated in the absence of selection (TX(0) and PB(0)). (A) The percentage of EGFP-positive cells (filled markers) and the mean fluorescence intensity (empty markers), represented as relative fluorescence units (RFU), in TX(0) and PB(0) cell pools was determined at the times indicated by flow cytometry. The viabilty of the cell pools transfected in the absence (B) and presence (C) of the PB system was assessed during the selection phase. (D) The volumetric productivity of different cell pools were determined by ELISA at the end of 4-day batch cultures.

The expression level of the PB(2) cell pools was 5-fold higher than for the TX(2) cell pools (Fig. 3.4D). Although the TX(2) and TX(10) cell pools had similar volumetric productivities, the PB(10) cell pools produced less TNFR:Fc than did the PB(2) cell pools (Fig. 3.4D). However, it is important to note that the TX(10) cell pool was generated from a very small population of cells (<10 %) that survived the selection pressure (Fig. 3.4B).

3.2.5. Evaluation of protein production and stability of cell pools in 14-day batch cultures

To show the reproducibility of cell pool generation and to study the productivity and viability in a production scenario (14-day culture), cell pools co-expressing TNFR:Fc and EGFP were generated from five independent transfections in which cells were co-transfected with pMPIG-TNFR:Fc and pmPBase at a 9:1 ratio (w:w). Cell pools were recovered by selection with puromycin (10 µg/mL) for 10 days starting on day 2 post-transfection. The volumetric TNFR:Fc productivity of each pool was evaluated in a 14-day biphasic batch culture with a temperature shift to 31 °C from 37 °C on day 4 post-inoculation. The cell density, cell viability, and volumetric TNFR-Fc productivity of the cultures were analyzed daily. All the cell pools had similar cell density and cell viability profiles, with the maximum cell density reaching 7 – 8 x 10^6 cells/mL at day 7 or 8 post-inoculation (Fig. 3.5A). The volumetric TNFR:Fc productivities reached 350 – 550 mg/L by day 14 post-inoculation (Fig. 3.5B).

To assess the stability of TNFR:Fc production over time, the cell pools were maintained in the absence of selection for an additional 2 months. At the end of each month, 14-day batch cultures of 100 mL were inoculated with each cell pool. The cell density and viability trends of the cell pools were similar to those observed at the 1-month time point (data not shown). Over the 3-month cultivation period, the volumetric TNFR:Fc yield declined only in pool 2 (Fig. 3.5C). For the other four cell pools, the productivity was found to be stable for 3 months in the absence of selection (Fig. 3.5C).

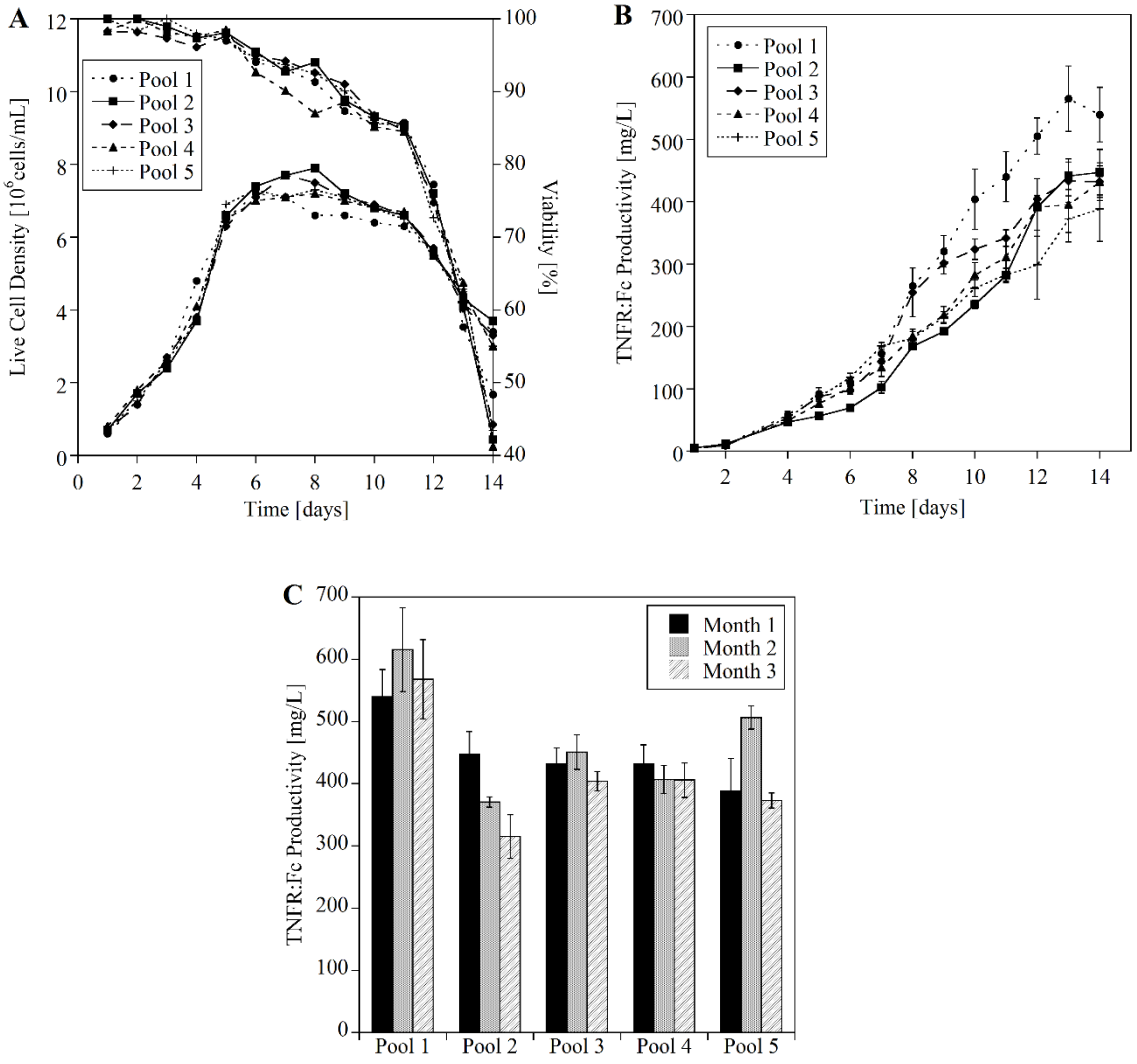

Figure 3.5. Cell growth and TNFR:Fc production by cell pools. Cell pools were generated from five independent co-transfections performed on different days with pMPIG-TNFR:Fc and pmPBase at a ratio of 9:1 (w/w). Selection was performed for 10 days in 10 µg/mL puromycin to generate cell pools numbered 1 – 5. The cell pools were grown in 14-day batch cultures at a volume of 100 mL in 500-mL cylindrical glass bottles. (A) The live cell density and cell viability were measured daily by the Trypan Blue exclusion method using a hemocytometer. (B) Volumetric TNFR:Fc productivities were measured daily by ELISA. (C) The volumetric TNFR:Fc productivities of each pool were compiled from 14-day cultures after 1, 2, and 3 months of growth in the absence of selection.The data was provided by Dr. Mattia Matasci.

3.3. Discussion

Long genetic selection durations, low volumetric protein yields, and loss of protein productivity over time in culture have typically been the key issues limiting the use of cell pools as a protein production method (Fan et al. 2013; Ye et al. 2010). However, with improved expression vectors and the enrichment of the recombinant cell population by fluorescence-activated cell sorting (FACS), it has been possible to generate cell pools expressing up to 1 g/L from a 2-week fed-batch process (Fan et al. 2013; Ye et al. 2010). However, the productivity of the cell pools was not stable over time in the absence of selection as a decrease in volumetric productivity of 40% per week was observed (Ye et al. 2010). Moreover, the entire process from transfection to protein production took a total of about 8 weeks. Hence, there is a need for an improved method to rapidly generate high-yielding cell pools whose productivity is constant over time.

Here we showed that PB-mediated cell pools can be generated with a selection duration of 5-10 days, resulting in cell pools with productivities 3-4 times higher than those generated by conventional plasmid transfection. Furthermore, the addition of a selection agent on 1, 2 or 3 days post-transfection did not affect the growth or productivity of the cell pools, indicating that transposition was a rapid event that occurred within 24 h of transfection. We also compared the use of different antibiotics for the selection of cell pools. Irrespective of the selection agent, the maximum protein yields were similar and were 3-4 fold better than the cell pools resulting from conventional plasmid transfection. However, we observed a decrease in the volumetric productivities of PB-mediated cell pools at the highest concentrations of puromycin and hygromycin B tested (Fig. 3.3A, C). These cell pools had slower growth rates than the other cell pools, and this may have resulted in reduced protein productivities (data not shown). Puromycin and hygromycin B are known to function by inhibiting protein synthesis (Darken 1964; Gonzalez et al. 1978). This mode of action may also have contributed to the decrease in protein productivity at the higher antibiotic concentrations tested.

Our results from the evaluation of the integration efficiency of the PB transposon corroborates our previously published results (Matasci et al. 2011). We observed that (1) ~20 % cells remained EGFP-positive in the absence of selection for the PB system in contrast to only ~3 % for conventional plasmid transfection and (2) the mean EGFP expression level was 4-5 fold higher for the PB system in comparison to conventional transfection (Fig. 3.4A). These results not only demonstrate the superior integration efficiency but also the superior transgene expression of transgenes transposed into the host genome in contrast to the integration of transfected plasmid DNA by recombination. We propose that this advantage could be utilized towards generation of "selection-free" clonal cell lines by coupling with other technologies such as FACS, using fluorescent-labeled antibodies, for enriching the recombinant cell population. Also, many high-yielding cell lines recovered after selection are known lose their expression once the selection pressure is removed (Chusainow et al. 2009; Kim et al. 2011). However, we speculate that 'selection-free' clonal cell lines would be more stable since they never encountered any selection pressure and the integrations have occurred independent of selection.

It is generally agreed that plasmid DNA integration into the genome is independent of selection pressure, and selection is only used to eliminate non-recombinant cells. Thus, loss of integrated transgenes or epigenetic silencing of the transgenes in TX(0) during growth, prior to the start of selection on day 10 post-transfection, are the probable causes of the decrease of the viability in cell pools TX(10) to ~10 % compared to the ~50 % in TX(2). Loss of integrated transgenes has been previously reported in recombinant cells grown in the absence of selection, while this is not the case for recombinant cells grown in the presence of selection (Weidle et al. 1988). In contrast to above observations, PB(2) and PB(10) had very similar viability trends after initiation of selection, with the minimum level reaching ~50 % (Fig. 3.3C). This implied that a similar percentage of cells was resistant to selection pressure, reiterating the evidence that PB-mediated integrations are rapid and stable events.

We have shown PB-mediated cell pool generation to be a rapid and simple technology that can reproducibly yield cell pools with a high level of protein productivity which can be stably maintained in culture for at least 3 months in the absence of selection.

4.

Comprehensive Study of PB, Tol2 and SB Transposon Systems for Recombinant Protein Production

4.1. Introduction

Several studies have compared the activities and integration profiles of PiggyBac (PB), Tol2 and Sleeping Beauty (SB) transposon systems in human cells, mainly focusing on applications in gene therapy (Grabundzija et al. 2010; Huang et al. 2010; Meir et al. 2011; Sharma et al. 2012; Wu et al. 2006). Recently the PB transposon system has been recognized for its potential in the generation of stable pools and clones for recombinant protein production (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013; Matasci et al. 2011). However, to our knowledge, these three transposon systems have not been compared side by side for their application in recombinant protein production in CHO cells.

In this work, the PB, Tol2 and SB transposon systems were used for the generation of stable CHO pools and clones. The mouse-codon-optimized PB transposase gene (mPB), a mammalian codon-optimized Tol2 transposase, and the hyperactive SB transposase (SB100X) were used to establish recombinant cell pools (Cadinanos and Bradley 2007; Mates et al. 2009). The cell pools were generated using the standard protocol described in Chapter 2 (Fig. 2.2).

The amount of donor and helper vectors required for efficient transposition was estimated for the three transposon systems. The transposition efficiencies were analyzed based on the number of transposon-mediated integrations versus the number of integrations resulting from DNA recombination and the percentage of recombinant cells resulting from transfection in the absence of selection. The number of integrated transgenes and the relative level of transgene mRNA transcription in each cell pool were also determined. Finally, cell pools and clones were subjected to simple fed-batch production processes from which the volumetric protein yield was measured. Our data demonstrates the efficacy and robustness of protein expression methods in which the use of transposon-mediated transgenesis is combined with the pool-based rapid generation of high-expressing recombinant cells.

4.2. Results

4.2.1. Generation of cell pools with three transposon systems using various amounts of helper vector

Different ratios of donor and helper vectors were tested to find the optimal conditions for the generation of cell pools expressing TNFR:Fc. In short, the donor plasmids expressing TNFR:Fc (pSBP-PB-TNFR:Fc, pSBP-Tol2-TNFR:Fc or pSBP-SB-TNFR:Fc) were co-transfected with their respective helper plasmids (pmPBase, pmTol2ase, or pSBase, respectively). The amount of helper vector ranged from 0.01 – 50 % of the total DNA transfected. Control transfections with each donor vector alone (100 % of the total DNA) were also performed. The pools were generated as described in Chapter 2 (Fig. 2.2). In short, the cells were placed under puromycin (10 µg/mL) selection for 10 days, from day 2 to day 12 post-transfection. Once generated, each cell pool was grown in the absence of selection in cultures inoculated at 0.3 x 10^6 cells/mL. At day 4 post-inoculation, the volumetric TNFR:Fc production of each cell pool was measured by ELISA.

For all three transposon systems, the TNFR:Fc productivity increased as the amount of helper vector increased until a plateau was achieved (Fig. 4.1A). In the case of the PB system, the highest productivity levels were achieved with 1 - 50 % pmPBase. For the SB system, the highest productivity levels were obtained following transfection with 10 - 50 % pSBase (Fig. 4.1A). For the Tol2 system, the highest productivity levels were observed following transfection with 10 - 30 % pTol2 (Fig. 4.1A). However, the TNFR:Fc productivity of the pools decreased from 90 mg/L to 60 mg/L when the amount of transfected Tol2ase was increased to 50 %. All three transposon systems generated cell pools expressing similar volumetric productivities under optimal conditions for the ratio of donor to helper vector. These values were up to 5-fold higher than the productivities of cell pools resulting from the transfection of 100 % donor vector (Fig. 4.1A).

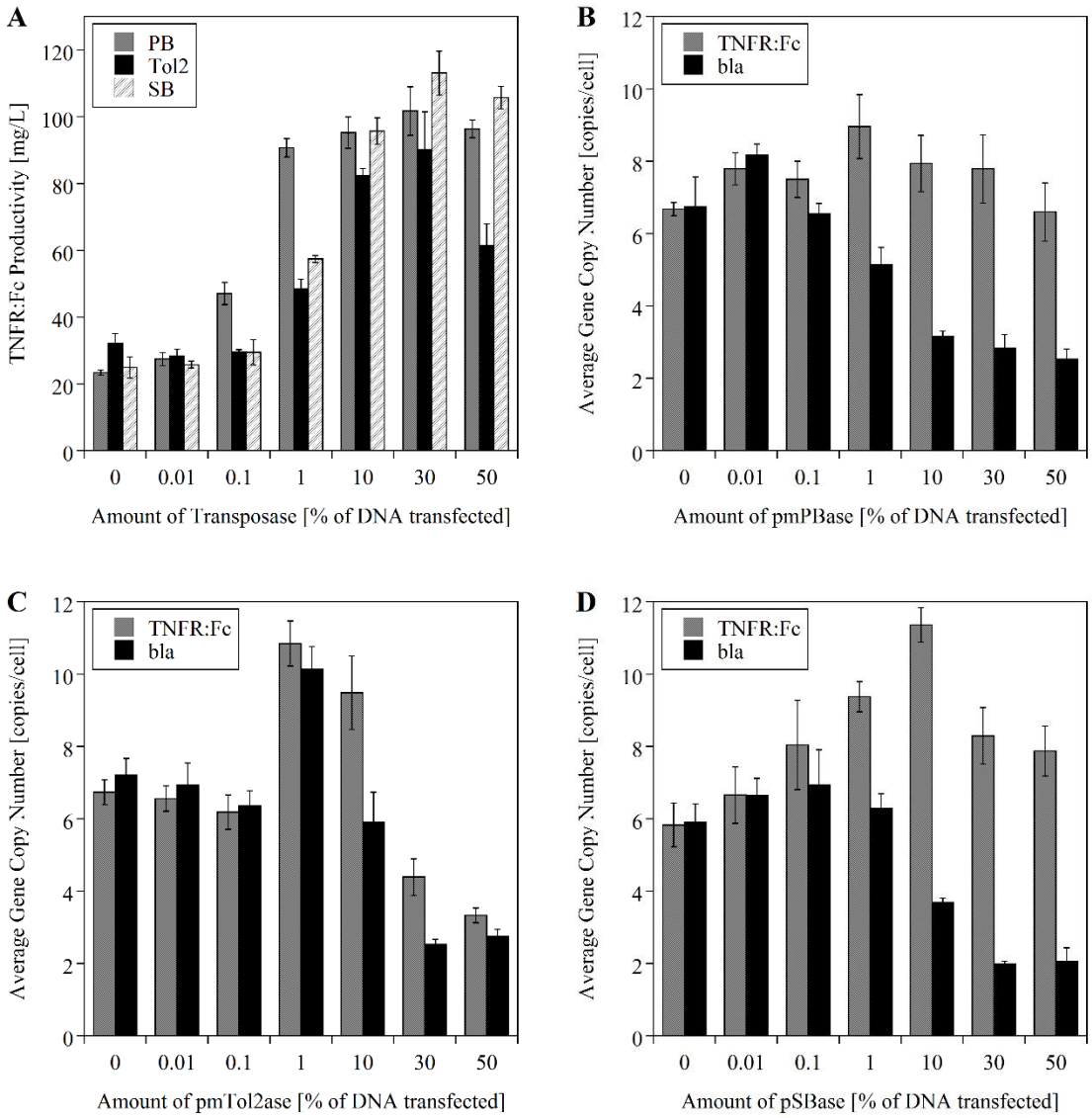

Figure 4.1. Analysis of cell pools generated with various amounts of donor and helper vectors. For each transposon system, cells were co-transfected with different amounts of helper vector, ranging from 0 – 50 % of the total DNA. For each transfection, the donor vector accounted for the remainder of the total DNA. Cell pools were generated by selecting with puromycin (10 µg/mL) for 10 days. For each cell pool, the TNFR:Fc yield was determined by ELISA at the end of 4-day batch culture (A). The TNFR:Fc and *bla* gene copy number per cell was determined by RT-qPCR on the cell pools generated with the (B) PB system, (C) Tol2 system, and (D) SB system. Each transfection was performed in duplicate, and each ELISA and RT-qPCR was performed in duplicate. Thus, each error bars represents the standard deviation from 4 different measurements.

Next, the average transgene copy number per cell was determined for each cell pool by RT-qPCR. On the donor plasmid, the inverted repeats of the transposon allow incorporation of both the TNFR:Fc gene and the *pac* gene but not the *bla* gene in the artificial transposon (Fig. 2.1A). Thus, the latter was not transposed into the CHO genome, but it could have been integrated via DNA recombination. Therefore, the difference between the average

copy numbers of the TNFR:Fc and *bla* genes corresponds to the average number of integration events that occurred by transposition in each cell pool. In PB-mediated cell pools, the average number of copies per cell of the TNFR:Fc gene was similar for all conditions (Fig. 4.1B). However, the number of copies of the *bla* gene decreased as the amount of transfected pmPBase increased (Fig. 4.1B). Thus, the number of transposon-mediated integrations increased as the amount of helper vector increased such that with 10 % or more pmPBase, about two-thirds of the integration events were mediated by transposition. Surprisingly, in the case of Tol2-mediated cell pools, the copy number values of the TNFR:Fc and *bla* genes were similar in all conditions except in cell pools transfected with 10 % pmTol2ase, implying most integrations were occurring by non-homologous recombination in cells transfected with this transposon system (Fig. 4.1C). In the presence of 10 % pmTol2ase, about 40 % of the integration events were transposon-mediated (Fig. 4.1C). In the case of SB-mediated cell pools, the copy number of the *bla* gene decreased as the amount of pSBase transfected increased (Fig. 4.1D). For both PB- and SB-mediated cell pools, at best 60 – 70 % of the integration events were transposon-mediated at best. These results indicated a higher transposition efficiency of these two transposon systems as compared to the Tol2 system. Based on these results, we concluded that the optimal ratio of donor to helper vector is 9:1 (w/w) for all three transposons.

4.2.2. Generation of cell pools with three transposon systems using various amounts of donor vector

Cells were transfected with different amounts of donor vector, ranging from 5 – 90 % of the total DNA, while maintaining the amount of helper pDNA at 10 % of the total DNA. Cells were co-transfected with various amounts of a donor vector expressing EGFP (pMP-PB-EGFP, pSB-PTol2-EGFP, or pSB-PSB-EGFP) and a helper vector (pmPBase, pmTol2ase or pSBase, respectively). The total amount of transfected DNA was kept constant by using filler DNA to replace some of the donor vector. Cell pools were also generated by transfection with each donor vector alone. Pools were generated by selecting with puromycin (10 µg/mL) for 10 days as described in Chapter 2. The % of EGFP-positive cells in each cell pool was measured by flow cytometry.

The co-transfections with the PB system resulted in >99 % EGFP-positive cells even when only 5 % of total DNA was from the donor vector (Fig. 4.2A). In the transfection with the donor vector alone, 90 % of the cells were EGFP-positive (Fig. 4.2A). In contrast, only about 60 % of the cells were EGFP-positive in cell pools resulting from transfection with SB or Tol2 donor vector alone (Fig. 4.2A). For both the SB and Tol2 systems, transfections with donor vector amounts of 25 % and more resulted in cell pools with similar levels of EGFP-positive cells (Fig. 4.2A). However, this was 20 – 30 % lower than the maximum level of EGFP-positive cells in the PB-mediated cell pools. For all three transposon systems, the EGFP expression levels of cell pools resulting from transfections with 25 – 90 % of the donor plasmid were similar (Fig. 4.2B). However, the EGFP levels were 40 % lower for cell pools resulting from transfections with the Tol2 and SB transposon systems as compared to the EGFP levels of the PB-mediated cell pools (Fig. 4.2B).

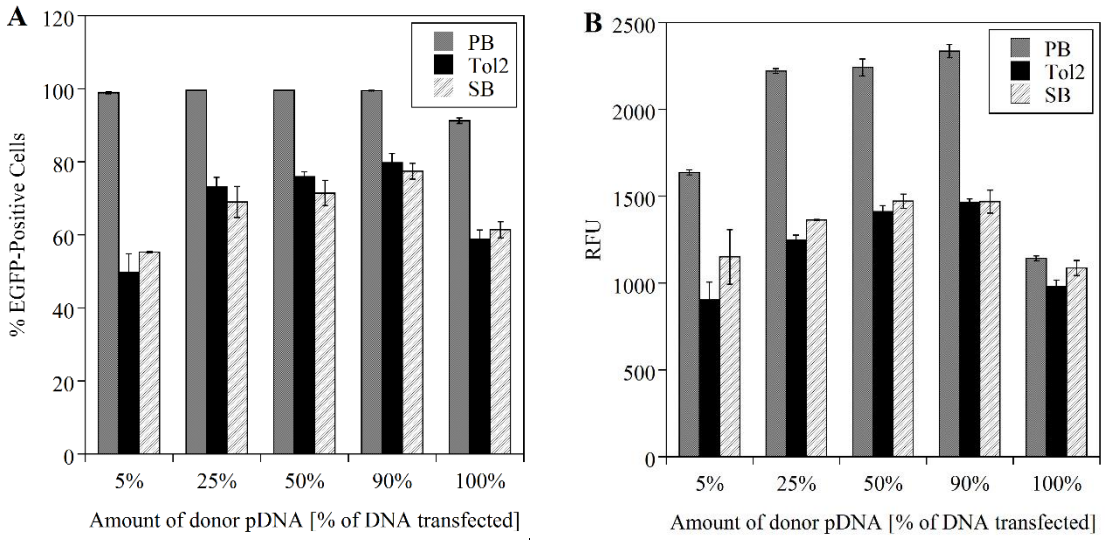

Figure 4.2. Analysis of EGFP expression by transposon-mediated cell pools generated with various amounts of donor plasmid. Cells were co-transfected with different amounts of donor vector (pMP-PB-EGFP, pSB-PTol2-EGFP, or pSB-PSB-EGFP), ranging from 5 % to 90 % of the total DNA, and a helper vector (pmPBase, pmTol2ase or pSBase, respectively) at 10% of the total DNA. The total amount of transfected DNA was kept constant by addition of filler DNA. Cell pools were generated by selecting with puromycin (10 µg/mL) for 10 days. Each transfection was conducted in duplicate and each measurement was performed in duplicate. Thus, the error bars represent the standard deviation from 4 different measurements. The % of EGFP-positive cells (A) and the relative fluorescence units (RFU) (B) in the cell pools were measured by flow cytometry at the end of a 4-day batch culture.

4.2.3. Analysis of transposition efficiency in the absence of selection

Cells were co-transfected with a donor plasmid expressing EGFP (pMP-PB-EGFP, pSBP-Tol2-EGFP or pSBP-SB-EGFP) and helper plasmids (pmPBase, pmTol2ase, or pSBase, respectively) at a ratio of 9:1 (w/w). Cell were also transfected with the donor vector and filler DNA at a 9:1 ratio (w/w). The cells were cultivated in the absence of selective pressure with passage every 3-4 days and analyzed regularly for EGFP-specific fluorescence by flow cytometry during a cultivation period of 30 days. The efficiency of recombinant cell generation was then estimated from the percentage of EGFP-positive cells in the population at 30 days post-transfection, a sufficiently long cultivation time in the absence of selection to eliminate any transiently expressed EGFP.

For all the conditions tested, the percentage of EGFP-positive cells gradually decreased during the first 10 – 12 days post-transfection and then stabilized at defined levels (Fig. 4.3). All three of the control transfections with donor vector alone had about 1 % EGFP-positive cells, while cells transfected with the PB and SB systems stabilized at 22 % and 14 % EGFP-positive cells, respectively (Fig. 4.3). For the cells transfected with the Tol2 system, only about 1 % of cells were EGFP-positive (Fig. 4.3). These results showed that the PB system was the best of the three transposon systems in terms of generating recombinant cells in a transfected cell population.

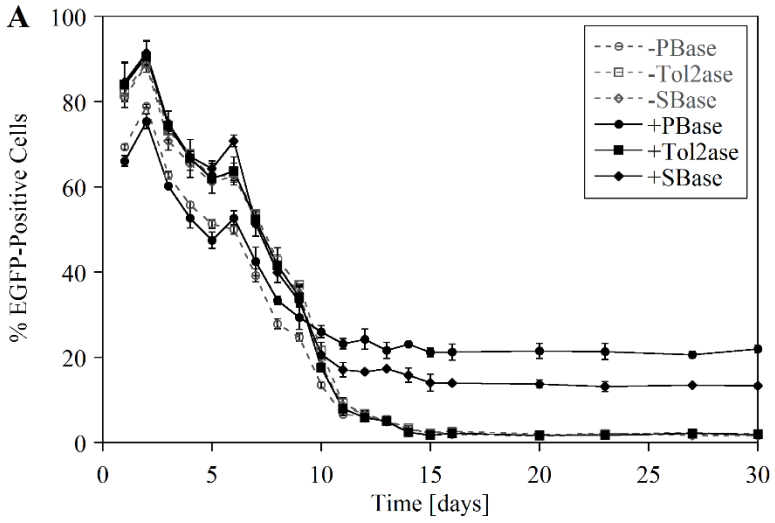

Figure 4.3. Determination of the efficiency of recombinant cell generation with the PB, Tol2 and SB transposon systems. Cells were co-transfected in duplicate with a donor vector expressing EGFP (pMP-PB-EGFP, pSBP-Tol2-EGFP or pSBP-SB-EGFP) and a helper vector (pmPBase, pmTol2ase, or pSBase, respectively) for the respective transposon system (+PBase, +Tol2ase, +SBase) at a 9:1 ratio (w/w). Transfections were also conducted with each donor vector alone (-PBase, -Tol2ase, -SBase). The percentage of EGFP-positive cells in each culture was determined by flow cytometry at the times indicated. The results for each condition are represented as average values of two independent transfections.

4.2.4. Protein productivities of cell pools generated with each transposon system

Cell pools were generated with the donor vectors pSBP-PB-TNFR:Fc, pSBP-Tol2-TNFR:Fc, pSBP-SB-TNFR:Fc in the presence and absence of the respective helper vector pmPBase, pmTol2ase, or pSBase. In addition to the control transfections in the absence of helper vector, we constructed an expression vector without terminal repeats (pMP-TNFR:Fc). The cell pools generated from transfections with pMP-TNFR:Fc alone were termed "Non-TP" cell pools.

For each cell pool, the volumetric protein productivity was analyzed at the end of 4-day batch cultures. Overall, all transposon-mediated cell pools had volumetric TNFR:Fc yields above 80 mg/L (Fig. 4.4) as was described above (Fig. 4.1A). In addition, the protein yields from "Non-TP" cell pools were 2-3 fold lower than the yields from cell pools generated with donor vector alone (Fig 4.4). These results showed that the three transposon systems were able to generate cell pools with similar volumetric productivities, which was ~9 fold higher than that of the "Non-TP" cell pools. It also showed that the presence of the inverted

repeats in the donor plasmids played a role in improving the protein yields. Henceforth, cell pools generated with pMP-TNFR:Fc alone were used as a control.

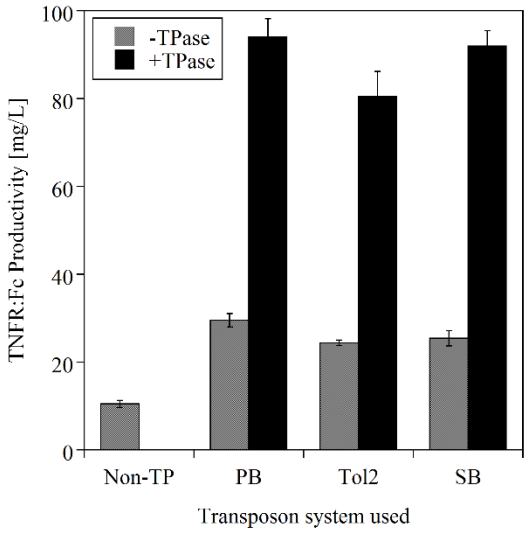

Figure 4.4. Volumetric productivity of cell pools generated with different transposon systems. Cells were transfected as in Fig. 4.3 with different donor vectors, with (+TPase cell pools) and without (-TPase cell pools) the respective helper vectors (PB, Tol2, or SB). Another transfection was performed with pMP-TNFR:Fc alone. This expression vector did not have ITRs for any transposon system (Non-TP cell pools). Cell pools were generated by selecting with puromycin (10 µg/mL) for 10 days. The volumetric productivities were measured by ELISA at the end of 4-day batch cultures. Each transfection was conducted in duplicate and each measurement was performed in duplicate. Thus, the error bars represent the standard deviation from 4 different measurements.

4.2.5. Transgene mRNA analyses of transposon mediate cell pools

For the transposition-mediated cell pools and the Non-TP cell pools described in the previous section, the average integrated TNFR:Fc gene copy number and the relative amount of its mRNA were determined. RT-qPCR was performed on total cellular DNA from the cell pools as described in Chapter 2. The average transgene copy number was similar for each of the cell pools generated with a transposon system but lower than the transgene copy number for the Non-TP cell pools (Fig. 4.5A). RT-qPCR was also used to determine the relative levels of transgene mRNA. The level of TNFR:Fc mRNA was similar for the cell pools generated with a transposon system (Fig. 4.5B). The relative transgene mRNA levels were about 4-fold higher for these cell pools than for the Non-TP cell pools (Fig. 4.5B). Thus, in spite of having more transgene copies integrated, they were

not as well-transcribed in the Non-TP cell pools as they were in the transposon-mediated cell pools.

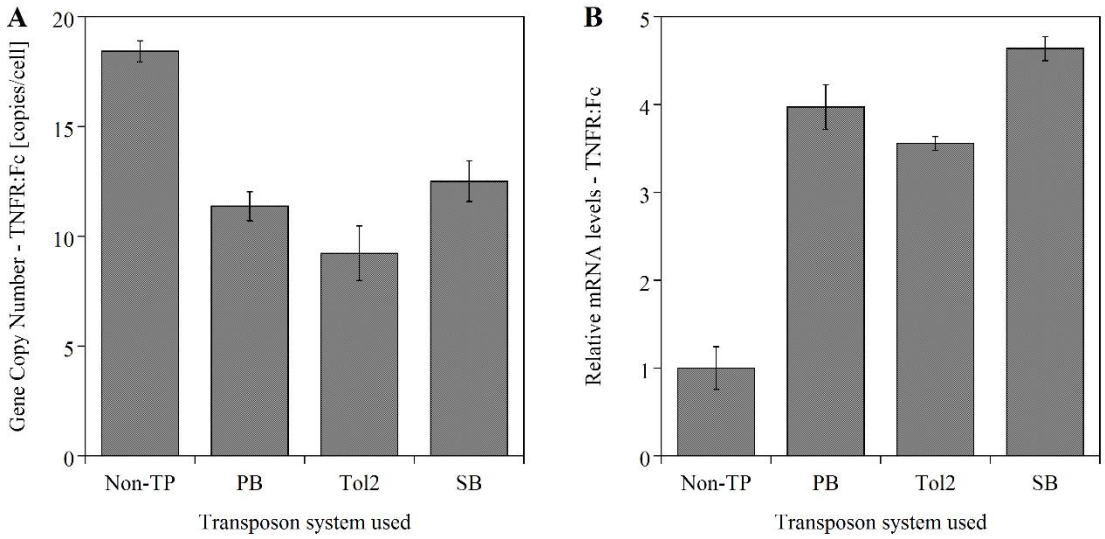

Figure 4.5. Integrated transgene copy number and transgene mRNA levels in cell pools generated with and without transposition. (A) The TNFR:Fc gene copy number for cell pools generated with transposition (PB, Tol2, SB) and without transposition (Non-TP) was determined by RT-qPCR. (B) The TNFR:Fc gene transcription levels were quantified by qRT-PCR and normalized to the level of cellular β-actin mRNA in CHO cells. The error bars represent the standard deviations from 4 measurements as each culture and analytical measurement was performed in duplicate.

4.2.6. Analysis of the stability of protein production in cell pools

It has been shown that the protein production levels of CHO cell pools can decline at a rate of 40 % per week when grown in the absence of selection (Ye et al. 2010). This represents a severe disadvantage for the production of recombinant proteins from cell pools, particularly if repeated productions from the same culture are performed sequentially. Hence, the TNFR:Fc productivity of cell pools generated with different transposons (pools PB, Tol2, and SB) and those generated by conventional transfection (Non-TP), as described in Section 4.2.4, were cultivated in the absence of selection for up to 3 months. The productivity of each cell pool was determined periodically at the end of 4-day batch cultures inoculated at 0.3 x 10^6 cells/mL. The volumetric productivity of all the transposon-based cell pools gradually declined to about 50 % of the initial level over a period of 2 months and then remained constant (Fig. 4.6). Surprisingly, the volumetric productivity of the Non-TP cell pools remained constant during the period of study. Nevertheless, the productivities

of the transposon-mediated cell pools were 3 – 4 fold higher than that of the control cell pools at the end of the stability studies.

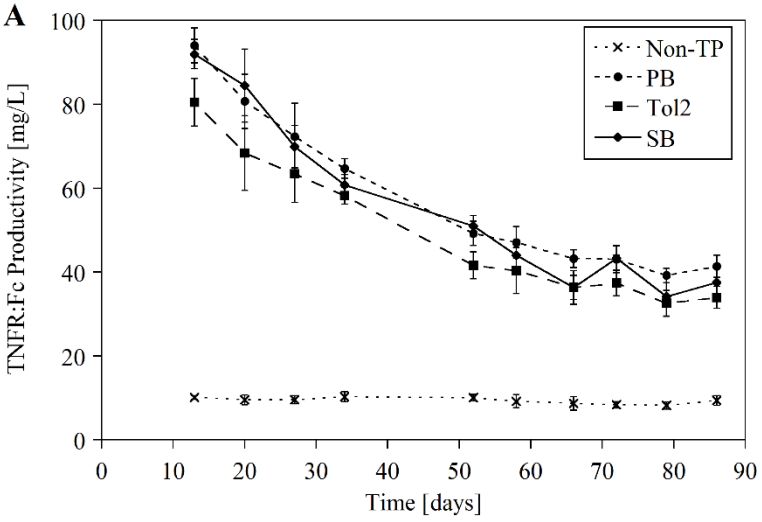

Figure 4.6. Analysis of the volumetric productivity of cell pools over time in the absence of selection. Cell pools were generated with (PB, Tol2, SB) and without (Non-TP) transposition as indicated in Chapter 2. After selection in puromycin, the cell pools were grown in the absence of selection pressure for 3 months. The volumetric productivity was determined by ELISA at the end of 4-day batch cultures at the times indicated.

4.2.7. Analysis of clonal cell lines derived from cell pools

About 100 clonal cell lines were generated by limiting dilution from cell pools generated with one of the three transposons and from the control cell pools. For each clonal cell line, the volumetric TNFR:Fc productivity was measured at the end of a 4-day batch culture by ELISA. The results showed that on average, the cell lines generated with one of the three transposons had higher productivities than did the cell lines generated by conventional plasmid transfection (Fig. 4.7A). The distribution of the volumetric productivities of the cell lines generated with each of the transposons was similar, but the average productivity was the highest for the cell lines generated with the SB system (Fig. 4.7A). The mean TNFR:Fc level of the cell lines resulting from conventional plasmid transfection (Non-TP) was about 1.5 mg/L, whereas it was about 40 mg/L for the cell lines generated with a transposon. About 95 % of the Non-TP cell lines produced less than 1 mg/L TNFR:Fc. For this set of cell lines, the highest volumetric productivity was 30 mg/L. On the other hand,

10 – 15 % of the transposon-mediated cell lines had volumetric productivities over 100 mg/L.

From each set of cell lines, the two with the highest TNFR:Fc productivity were chosen for further studies. The 8 cell lines were cultivated in the absence of selection for over 4 months, and the volumetric productivity was measured monthly from 4-day batch cultures. The volumetric productivity was constant over 4 months for each of the transposon-mediated cell lines (Fig. 4.7B). However, one of the two cell lines generated by conventional plasmid transfection (Non-TP1) exhibited reduced productivity with time in culture (Fig. 4.7B). The specific productivities (Q_p) of all the cell lines were calculated based on data from 4-day batch cultures. The results show that the Q_p of the transposon-derived cell lines ranged from 17 – 23 pg/cell/day while those derived by conventional plasmid transfection each had a Q_p less than 5 pg/cell/day (Fig. 4.7B).

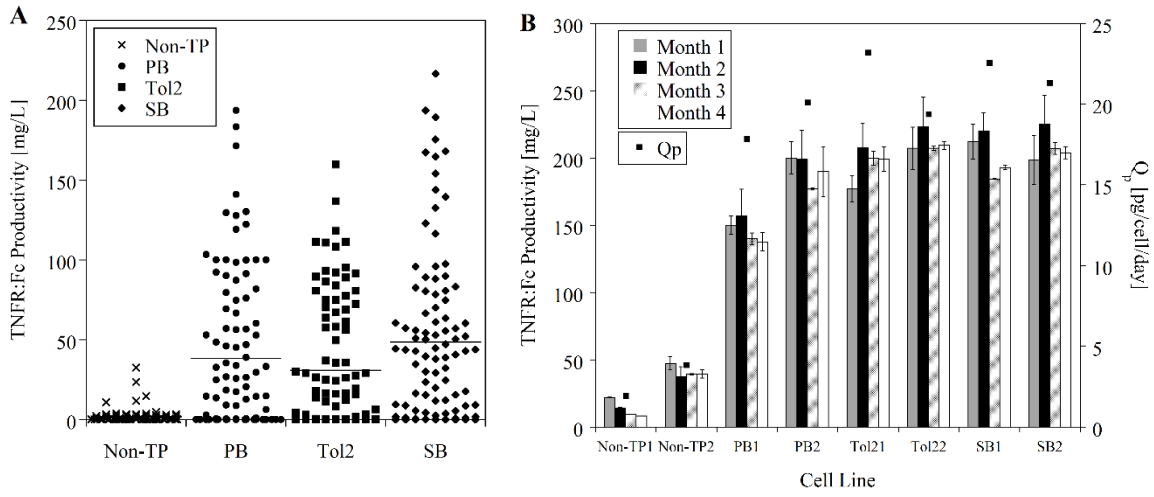

Figure 4.7. Analysis of clonal cell lines generated with different transposons. From each cell pool generated with a transposon (PB, Tol2, and SB) or without (Non-TP), 100 cell lines were generated by limiting dilution. (A) The volumetric TNFR:Fc productivity of each cell line was measured in a 4-day batch culture by ELISA. Each symbol on the plot represents the TNFR:Fc productivity of one cell line. The horizontal bars represent the mean of TNFR:Fc production for each condition. (B) The two highest producers of the 100 cells lines from Non-TP (Non-TP1, Non-TP2), PB (PB1, PB2), Tol2 (Tol21, Tol22) and SB (SB1, SB2) cell pools were chosen for further analysis. The eight cell lines were cultivated for 4 months in the absence of selection, and the volumetric productivities were analyzed monthly at the end of 4-day batch cultures. The cell-specific productivity (Q_p) of each cell line was measured in a 4-day batch culture (diamonds) at the end of the first month of culture.

4.2.8. Expression level analysis from a fed-batch culture for cell lines and pools

The cell pools generated by using PB, Tol2 and SB transposition and by conventional transfection (Non-TP) were subjected to a simple non-optimized fed-batch culture of 2 weeks that included a temperature shift to 31°C at day 3 post-inoculation. The transposon-mediated pools resulted in TNFR:Fc productivities of 800-900 mg/L whereas the cell pool resulting from conventional plasmid transfection produced only about 130 mg/L (Fig. 4.8A). The top two TNFR:Fc-producing cell lines from each condition as described in Section 4.2.7 were also subjected to the same fed-batch cultivation conditions. All transposon-generated cell lines produced from 1.3 to 1.5 g/L, whereas the two Non-TP cell lines produced about 500 mg/L each (Fig. 4.8B).

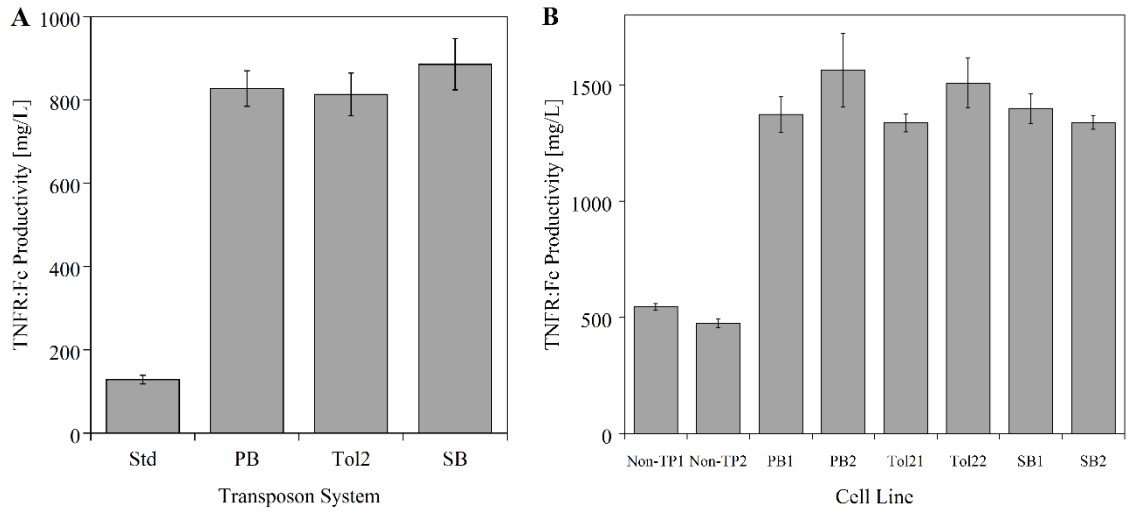

Figure 4.8. Volumetric productivities of fed-batch cultures of cell pools and cell lines. The cell pools (A) and cell lines (B), as described in Figures 4.4 and 4.7B, respectively, were subjected to a 14-day fed batch culture. Each cell pool and cell line was inoculated at a density of 0.5 x 10^6 cells/mL in a 5 mL volume in a TS50. On day 3 post-inoculation, the cultures were fed with 3 g/L glucose and 1.5 g/L essential and non-essential amino acids and treated with 0.25 % dimethyl acetamide (DMA). The cultures were incubated for 3 days at 37 °C and then at 31 °C for 11 days. The volumetric TNFR:Fc productivity was then measured by ELISA. The error bars represent the standard deviation from 4 measurements as each culture and analytical measurement was performed in duplicate.

4.3. Discussion

We investigated the use of PB, Tol2 and SB transposon systems for the rapid generation of cell pools for recombinant protein production using CHO-DG44 cells. Under optimized conditions, the pools generated with each transposon system showed similar volumetric protein productivities, at values 3 – 4 fold higher than the pools generated using the donor plasmid alone, and about 9 fold higher than the pools generated with a donor plasmid without the inverted repeats of the transposon. Using the three transposon systems, we were able to generate cell pools with TNFR:Fc expression levels of 800-900 mg/L and cell lines with expression levels of 1.3-1.5 g/L with a simple, non-optimized fed-batch process. In contrast, PB-derived cell pools expressing EGFP had ~20 % more EGFP-positive cells and ~40 % higher EGFP expression levels than the cell pools generated using the Tol2 or SB transposon systems.

We observed some differences among the transposon systems tested. The performance of the Tol2 transposon system in terms of the percentage of recombinant cells per transfection and the percentage of transgene integrations by transposition rather than DNA recombination were inferior to those of the PB and SB transposon systems. The transgene copy number analysis showed that transgene integration by DNA recombination in cell pools following transfection with a transposon system was frequent. In PB and SB transposon systems, 60 – 70 % of the transgene integrations resulted from transposition under the optimal transfection conditions. However, only 40 % of the integration events were by transposition in cell pools following transfection with the Tol2 system. In spite of having 60 % of the transgene integrations by DNA recombination, the volumetric productivities of the Tol2-mediated cell pools were the same as those of PB- or SB-mediated cell pools.

With an increase in the amount of pmTol2ase transfected, both the productivity of the pools and the number of transposon-mediated integrations declined. This may indicate the occurrence of the discussed phenomenon of OPI, the phenomenon that limits the transposition activity at increasing transposase levels (Grabundzija et al. 2010). On the other hand, contrary to previously published claims, the PB and SB transposes did not

appear to exhibit OPI (Grabundzija et al. 2010; Wilson et al. 2007; Wu et al. 2006). The decline in the volumetric productivity in Tol2-mediated pools as the amount of pmTol2ase increased may also have been due to integration of the transposase gene which could potentially mediate excision of the integrated transgene.

The PB and SB transposon systems demonstrated a higher percentage of stable recombinant cells in a transfected population compared to the Tol2 system or to conventional plasmid transfection. Efforts have recently been made to enhance the transposition efficiency for PB and SB, resulting in very effective transposases like the hyperactive PBase and SB100X (Cadinanos and Bradley 2007; Mates et al. 2009; Meir et al. 2013; Yusa et al. 2011). However, Tol2 has not been genetically engineered for improved performance. Although all the transposase enzymes in this study were codon optimized for mammalian expression, we believe that further improvements of transposition efficiency can be obtained by modification of the ITRs or by mutagenesis of the respective transposase genes. Nevertheless, the protein productivity of cell pools generated using the three transposon systems tested was very similar.

Interestingly, the ITRs of all three TP systems had a slightly positive impact on recombinant protein production even in the absence of the transposase. The ITRs mainly consist of repeat elements that are AT rich. I propose that, even in the absence of the transposase enzyme, these regions of the plasmids will frequently undergo double-strand – single-strand DNA conversions and thus facilitate integration, possibly even with a preference to widely distributed and frequent AT-rich regions in the genome, homologous in sequence to the incoming DNA. In support of this hypothesis, it has been reported that the presence of sequences of CHO endogenous retroviral origin, when present in transfected plasmids, had a positive effect on the average productivity from recombinant clones (Lewis et al. 2013; Wurm 1994). The probability of such facilitation of integration is high since optimized PEI-mediated transfections, as used here, can deliver 40,000 – 100,000 plasmid molecules per cell (Backliwal et al. 2008; Rajendra et al. 2012).

It was shown that the improved protein productivity from transposon-derived cell pools was mainly due to increased transgene mRNA levels, about 4-fold compared to the Non-

TP cell pool. This was surprising since Non-TP pools had a 2-fold higher integrated transgene copy number. This strongly supports the notion that integrated transgenes in transposon-mediated cell pools were better transcribed than those in Non-TP cell pools. This is in agreement with conclusions drawn from previous publications, where it was suggested that transposases integrate the transgenes into transcriptionally active regions, thereby resulting in higher transcription levels leading to increased productivities (Galvan et al. 2009; Grabundzija et al. 2010; Meir et al. 2011; Wilson et al. 2007).

We observed a gradual decrease in volumetric productivity of the transposon-mediated cell pools to about 50 % of their initial levels within 2 months of cultivation in the absence of selection. Previously, we generated cell pools with the PB system that were stable for 3 months in the absence of selection (Chapter 3). However, the lot of the medium was changed between the previous studies and those reported here. We suspect that this may contribute to the loss in the stability of the productivity of the cell pools. Also, previously published results from cell pools showed a decline in protein production of 40 % per week in spite of the cell pools being selected by FACS for a high level of fluorescent protein expression (Ye et al. 2010). In comparison, the decline observed here was only about 10-15% per week. It is also important to note that a decline of 15-20% in protein productivity over 60 population doublings (about 2 months, assuming a doubling time of 24 hours) is typically considered acceptable for commercial cell lines (Wurm, personal communication). Therefore, our 12-week assay can be considered sufficient to assess the stability of the volumetric productivity in a cell pool. Importantly, even after 2 months, the protein productivity of the transposon-derived cell pools was still 4 times higher than the Non-TP cell pool.

Lastly, the average volumetric productivity of cell lines recovered from transposon-derived cell pools was higher than that from cell lines recovered from the control cell pools. Thus, the use of the transposon systems allowed a much greater efficiency in the generation of highly productive clonal cell lines using CHO-DG44 as the host. This means that a smaller number of cell lines would need to be screened to obtain ones with a high level of protein productivity. In conclusion, all the three transposon systems evaluated in this study can be

used for the generation of cell pools expressing high levels of recombinant protein, with a timeline of less than 4 weeks from transfection to protein recovery.

5.

Study of the Three Transposon Systems for Cell Pool Generation in Different CHO Cell Strains

5.1. Introduction

CHO-DG44, CHO-K1 and CHO-S are three of the most commonly used CHO strains in the biopharmaceutical industry. Although these strains share a common origin, they are known to have different genotypes and phenotypes (Derouazi et al. 2006; Lewis et al. 2013; Wurm 2013). To improve the applicability of transposon systems for cell pool and cell line generation, we decided to evaluate the use of the PB, Tol2 and SB transposon systems for gene delivery in these three CHO strains. The CHO-DG44 cells were isolated in 1983 and have been in culture in the LBTC for more than 15 years (Urlaub et al. 1983). The CHO-K1 cells were suspension-adapted in the LBTC in 2012 from adherent cells obtained from the ATCC. The CHO-S cells are suspension-adapted cells that are commercially available from Life Technologies.

The efficiency of the generation of recombinant cells using the three transposon systems in the different CHO strains was compared. The transposon-mediated cell pools were generated based on the protocol developed for CHO-DG44 cells, and the volumetric productivities of the various cell pools were measured. The stability of the volumetric productivity over 2 months in the absence of selection was also measured. An analysis of average integrated gene copy number was also performed as a way to compare the activity of the three transposons in the three different CHO strains.

5.2. Results

5.2.1. Generation of recombinant cells following transfection of three CHO strains with three different transposon systems

Three different CHO strains were separately co-transfected with three combination of a donor plasmid (pMP-PB-IRES-EGFP, pSBP-Tol2-IRES-EGFP or pSBP-SB-IRES-EGFP) and a helper plasmid (pmPBase, pmTol2ase, or pSBase, respectively) at the ratio of 9:1 (w/w). The donor vectors each express the TNFR:Fc and EGFP genes from a bicistronic mRNA in which the two genes are separated by the EMCV IRES (Table 2.1). For each combination of cell strain and transposon system, a control transfection was performed with the donor plasmid and filler DNA at a 9:1 ratio (w/w). The cells were cultivated in the absence of selective pressure with passage every 3-4 days and analyzed regularly for EGFP-specific fluorescence by flow cytometry for 30 days. The efficiency of recombinant cell generation was then estimated from the percentage of EGFP-positive cells in the population at 30 days post-transfection, a sufficiently long cultivation time to eliminate any transiently expressed EGFP from the population.

For all the conditions tested, the percentage of EGFP-positive cells gradually decreased during the first 10–12 days and then remained constant at defined levels (Fig. 5.1 A-C). All control transfections with donor vector alone had only about 1% EGFP-positive cells at day 30, indicating that this percentage of cells in the transfected population were recombinant (Fig. 5.1 A-C). The % EGFP-positive cells in transfections using the PB transposon system stabilized at 10 – 15 % for all three CHO strains (Fig. 5.1A). For the transfections performed with the Tol2 and SB systems, the percentage of EGFP-positive cells stabilized at 4 – 7 % and 5 – 20 %, respectively (Fig. 5.1B, C). For the transfections performed with the SB transposon system, a higher percentage of EGFP-positive cells was observed using CHO-DG44 cells as the host in comparison to the other two CHO strains (Fig. 5.1C). Transfections using Tol2 and SB transposon systems showed similar EGFP profiles for CHO-K1 and CHO-S cells (Fig. 5.1B, C). For all three transposon systems in all three CHO strains, the % EGFP-positive cells was between 4 – 15 % after 30 days (Fig. 5.1 A-C)

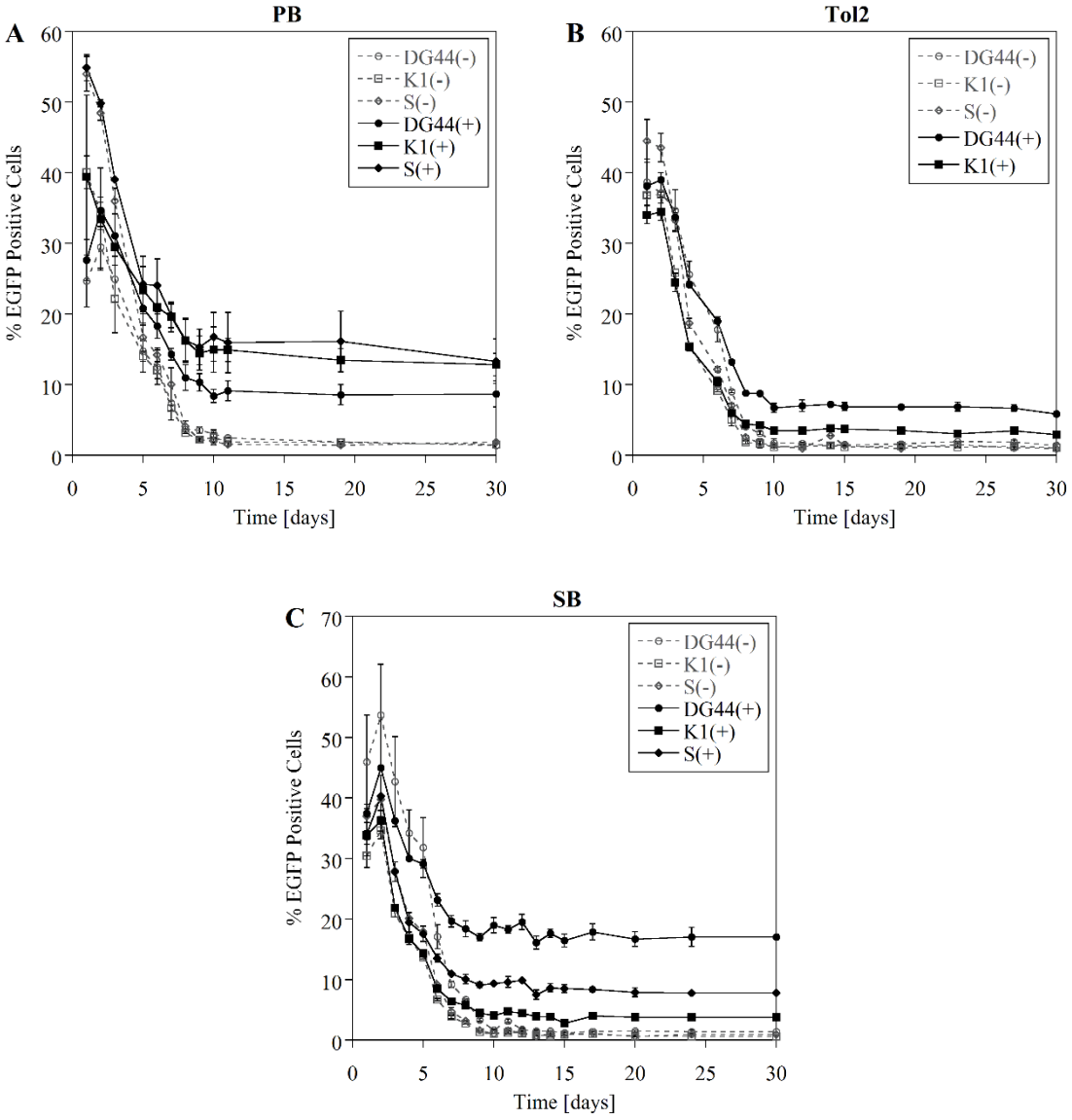

Figure 5.1. Determination of the efficiency of recombinant cell line generation with the three transposon systems. Cells were transfected in duplicate with a donor vector (pMP-PB-IRES-EGFP, pSBP-Tol2-IRES-EGFP or pSBP-SB-IRES-EGFP) and a helper vector (pmPBase, pmTol2ase, or pSBase, respectively) at 9:1 ratio (w/w) to generate the cell populations labeled (A) PB (+), (B) Tol2 (+), and (C) SB (+), respectively. Transfections were also conducted with each donor vector and filler DNA at 9:1 ratio (w/w) to generate cell populations (A) PB (-), (B) Tol2 (-), and (C) SB (-), respectively. For each transposon system, transfections were performed in CHO-DG44, CHO-K1, and CHO-S. The cells were cultivated in the absence of selective pressure with passage every 3-4 days. The percentage of EGFP-positive cells in each culture was determined by flow cytometry at the times indicated. The results for each condition are represented as average values of two independent transfections.

5.2.2. Volumetric productivities of transposon-mediated cell pools derived from different CHO strains

The three CHO strains were co-transfected with the donor vectors pMP-PB-TNFR:Fc, pSBP-Tol2-TNFR:Fc, or pSBP-SB-TNFR:Fc along with their respective helper vectors pmPBase, pmTol2ase, or pSBase, respectively. Cells were also co-transfected with each donor vector and filler DNA at a ratio of 9:1 (w/w). For each transfection, a cell pool was generated by selection in puromycin as described in Chapter 2. Briefly, the cells were cultivated in the presence of 10 µg/mL puromycin for 10 days starting at day 2 post-transfection. The TNFR:Fc productivity of the pools was analyzed from a 4-day batch culture grown in the absence of selection. For all three CHO strains, a 3-4 fold improvement in the volumetric TNFR:Fc productivity of the cell pools was observed in the presence of the three transposases as compared to the productivities of the cell pools generated by conventional plasmid transfection (Fig. 5.2). The highest productivities were achieved in CHO-K-derived cell pools when using the PB and Tol2 systems and in CHO-DG44-derived cell pools when using the SB system (Fig. 5.2 A-C). The lowest productivity for each transposon system resulted from cell pools generated by transfections of CHO-S.

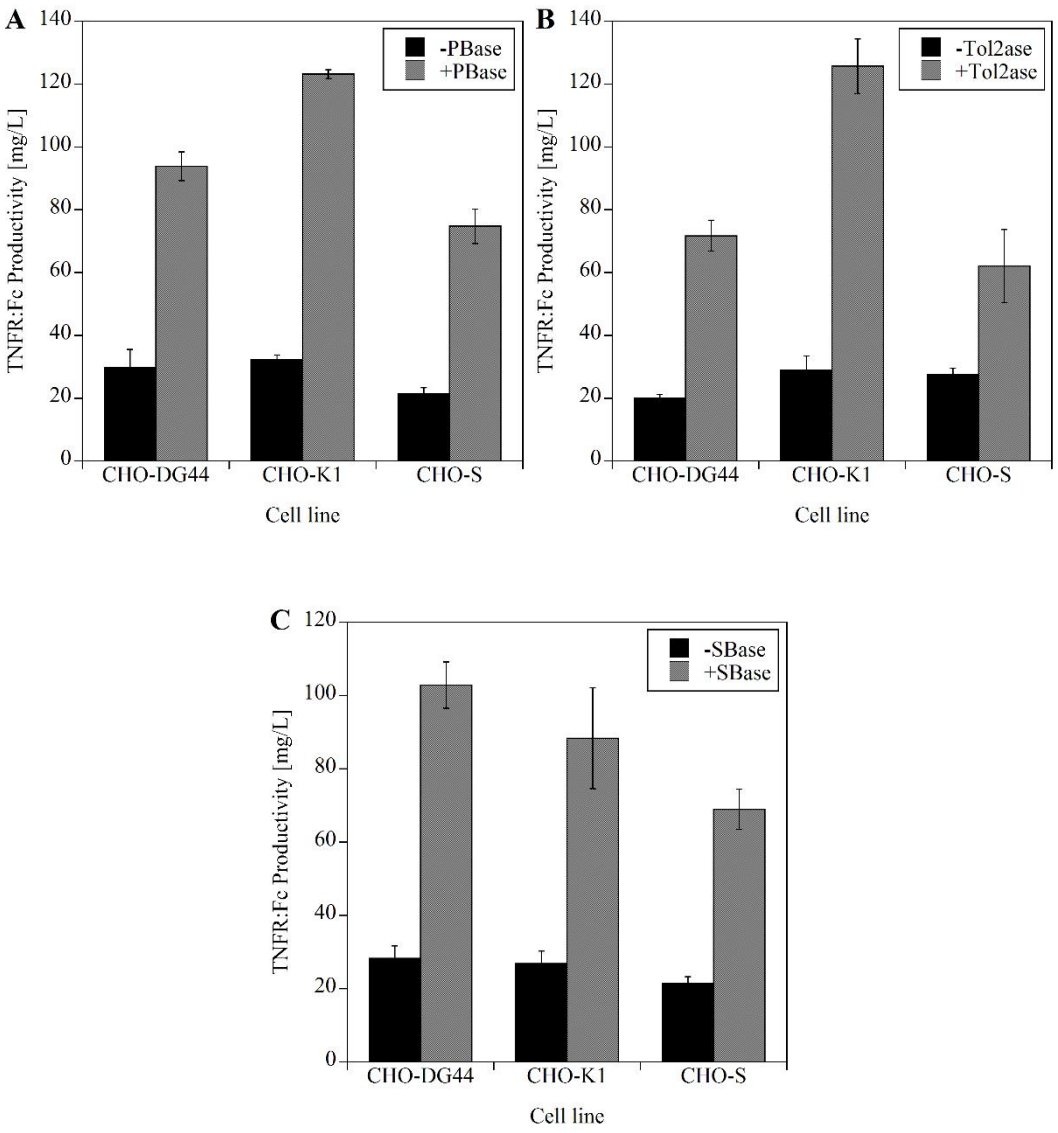

Figure 5.2. V**olumetric productivities of cell pools generated with different transposons and CHO strains.** Different CHO strains as indicated were transfected in duplicate with a donor vector (pMP-PB-IRES-EGFP, pSBP-Tol2-IRES-EGFP, or pSBP-SB-IRES-EGFP) and a helper vector pmPBase, pmTol2ase, or pSBase, respectively, at a ratio of 9:1 (w/w) to generate cell pools labeled (A) +PBase, (B) +Tol2ase, or (C) +SBase, respectively. Cells were also co-transfected with each donor vector alone and filler DNA at a 9:1 ratio (w/w) to generate cell pools labeled (A) -PBase, (B) -Tol2ase, or (C) -SBase. Cell pools were generated by selecting transfected cells with puromycin (10 µg/mL) for 10 days. The volumetric productivities were measured by ELISA at the end of 4-day batch cultures. Each ELISA measurement was performed in duplicate. Thus, the error bars represent the standard deviation from 4 different measurements as each culture was performed in duplicate.

5.2.3. Analysis of stability of volumetric protein production in cell pools

Cell pools generated with different transposons as described in Section 5.2.2 were cultivated in the absence of selection for up to 2 months with passage every 3-4 days. The productivity of each cell pool was determined periodically at the end of 4-day batch cultures starting from a culture seeded at 0.3 x 10^6 cells/mL. The cell pools generated with CHO-DG44 lost 50 % of their productivity within 30 days and then the productivity remained constant (Fig. 5.3 A-C). In contrast, the volumetric productivity of cell pools generated with either CHO-K1 or CHO-S cells declined with time to levels observed in the cell pools generated by conventional plasmid transfection (Fig. 5.3 A-C).

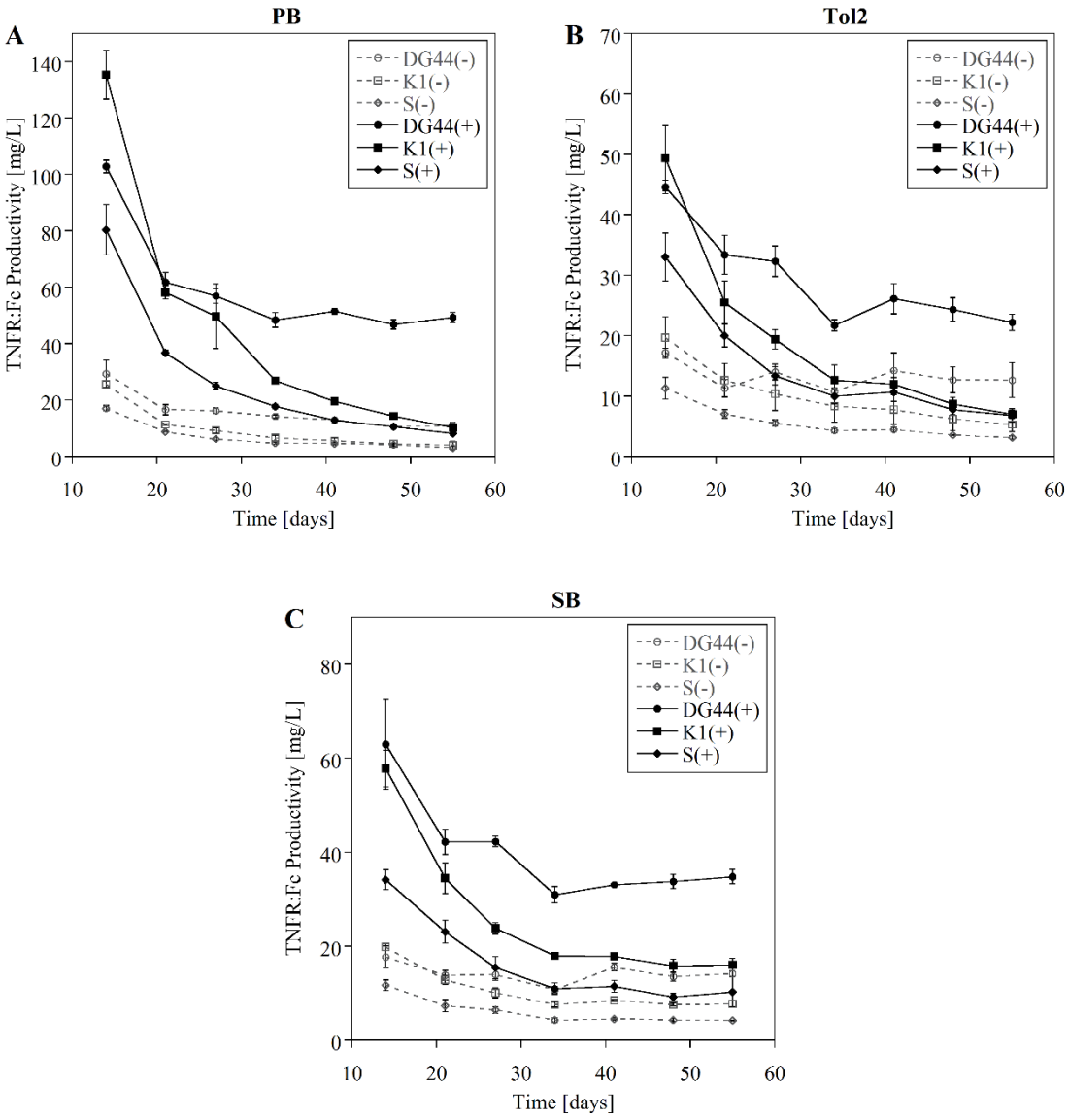

Figure 5.3. Analysis of the volumetric productivity of cell pools generated with PB, Tol2 or SB. Cell pools were generated with the PB (A), Tol2 (B), or SB (C) transposon systems in the presence (+) or absence (-) of the helper vector as indicated in Fig. 5.2. After selection in 10 µg/mL puromycin for 10 days, the cell pools were grown in the absence of selection for 2 months with passage every 3-4 days. The volumetric productivity of each cell pool was determined by ELISA at the end of a 4-day batch culture at the times indicated.

5.2.4. Analysis of the integrated transgene copy number in cell pools.

The average number of integrated TNFR:Fc genes was determined by RT-qPCR at the beginning and end of the cultures described in Fig. 5.3. For all three transposon systems, the average integrated transgene copy number for the cell pools cultivated in the absence of selection for 1 week was similar for those cell pools originating from the CHO-DG44 and CHO-K1 strains (Fig. 5.4 A-C). For the PB and SB transposon systems, however, the average integrated transgene copy number was ~50 % lower for cell pools originating from the CHO-S strain as compared to those cell pools derived from the CHO-DG44 and CHO-K1 strains (Fig. 5.4A, C). The CHO-S-derived cell pools generated with the PB and SB transposon systems also had low volumetric productivities (Fig. 5.2). We observed a loss of ~50% and ~30% of the transgene copies in the cell pools derived from the CHO-K1 and CHO-S strains, respectively, following transfected with the PB transposon system (Fig. 5.4A). These results may explain the decline of the volumetric protein yield over time in these two cell pools (Fig. 5.3). For all other combinations of transposon system and CHO strain, there was no loss in the integrated transgene copy number by cultivating the cells in the absence of selection for 8 weeks (Fig. 5.4 A-C).

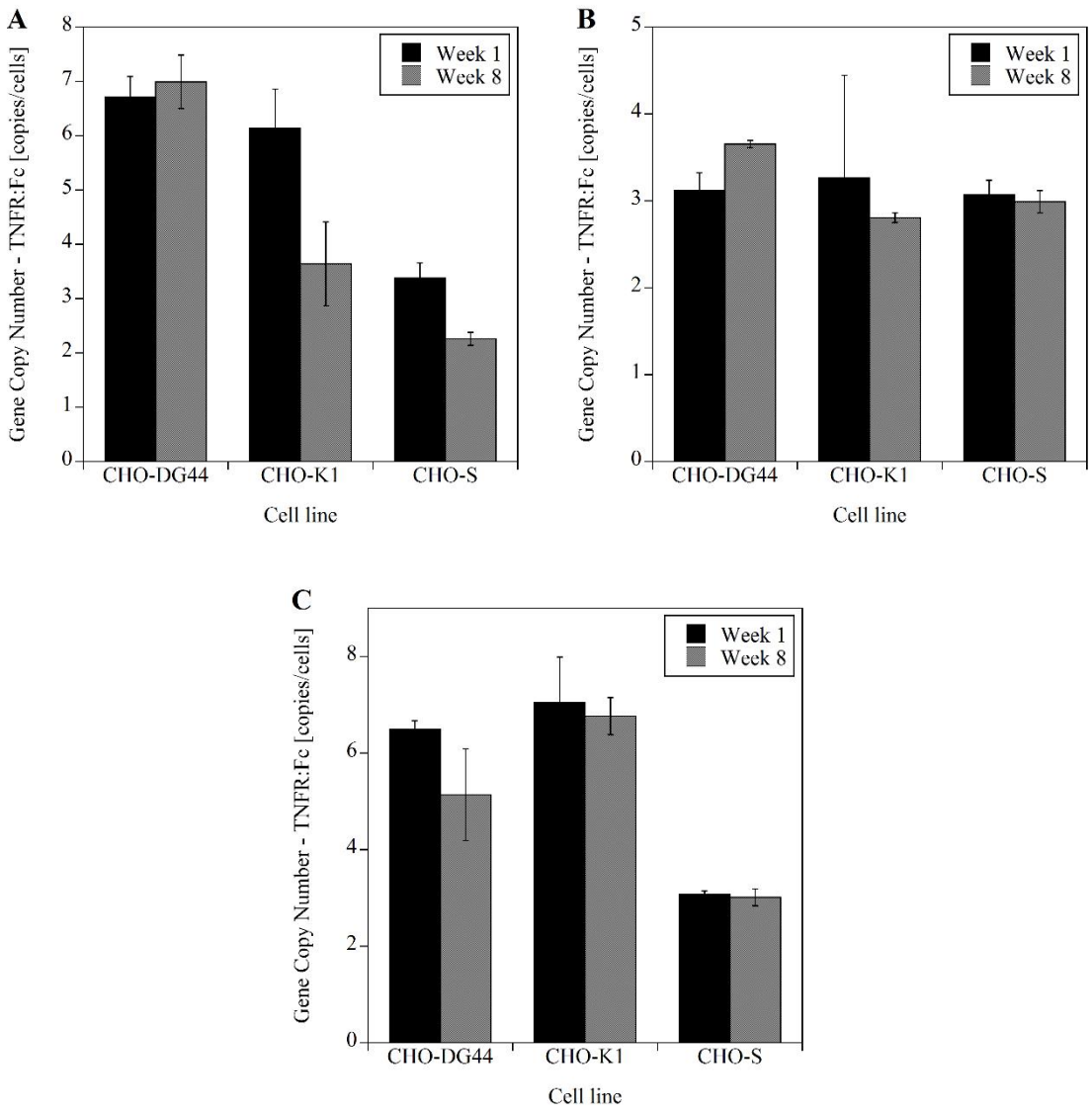

Figure 5.4. Determination of integrated TNFR:Fc gene copy number in cell pools generated with different transposons. Cell pools were generated with the PB (A), Tol2 (B), and SB (C) transposon systems as described in Fig. 5.2. The average number of integrated copies of the TNFR:Fc gene per cell in each pool was determined by RT-qPCR after cultivation of each pool in the absence of puromycin for 1 week and for 8 weeks.

5.3. Discussion

The applicability of the PB, Tol2, and SB transposon systems for the efficient generation of cell pools was extended to two other CHO cell strains, CHO-K1 and CHO-S. The fold improvement in TNFR:Fc productivity of the transposon-mediated cell pools, compared to cell pools generated by conventional transfection, was very similar in all three cell lines. The highest volumetric protein yields were obtained using the PB or Tol2 transposon system in CHO-K1 cells. Of the three strains, the host with the lowest volumetric productivities was CHO-S. However, the cell pools generated with CHO-K1 and CHO-S cells were very unstable and the protein productivity declined to the level of pools generated by conventional transfection in just 4-5 weeks of cultivation in the absence of selection.

Although there was a decline in protein productivity of the cell pools with time in culture for most combinations of transposon system and cell host, this was not due to a loss of copies of the integrated transgene. In cell pools with no loss in gene copy numbers, the decline in protein productivity was assumed to be due to transgene silencing. For cell pools in which there was a decline in the average number of integrated transgenes with time, the loss in protein productivity could be attributed to the physical loss of transgene copies or overgrowth of the culture with cells having lower protein production levels (Chusainow et al. 2009; Yang et al. 2010). Cell pools can be maintained in the presence of selective pressure until used for the protein production phase (Ye et al. 2010). This may help maintain the protein productivity over time.

All the transposon-mediated cell pools had 3-4 fold higher volumetric protein productivity levels compared to cell pools generated by conventional plasmid transfection. The uniform improvement in protein yield using transposon systems was very promising, considering that the protocols (selection parameters, donor: helper DNA amounts, medium etc.) were not optimized for CHO-K1 and CHO-S. Eventually, CHO-DG44 were the most suitable host cells for generating the cell pools with the three transposons, as the cell pools had good volumetric productivities and the production levels were more stable over time compared to those generated with CHO-K1 and CHO-S cells. We speculate that the differences in the final volumetric productivities and stability of protein production of the transposon-mediated cell pools from three cell strains could be a result of the non-optimized conditions

for cell pool generation. Cell strain-specific optimization studies will be required to further improve the systems in terms of the stability of protein production over time. However, these optimization studies for the CHO-K1 and CHO-S cell strains were not in the scope of this thesis. The LBTC has successfully used the PB transposon system for cell pool generation in HEK293 and S2 cells (not published).

Acknowledgements

Part of this work was done by Antonia Šakic, a master student under my supervision.

6.

Cell Pools Expressing Multiple Transgenes Using the PiggyBac Transposon for Gene Delivery

6.1. Introduction

In biomedical research and drug development there is often a need for multiple genes to be co-expressed in the same cell. The applications extend from expression of multi-subunit proteins to co-expression of a chaperone protein or other partner proteins with the protein of interest (Alattia et al. 2013; Chung et al. 2004; Hwang et al. 2003; Kahlig et al. 2009; Mastrangelo et al. 2000; Mohan et al. 2008). Conventional approaches are based on the transfection of host cells with multiple expression cassettes, either carried in a single plasmid or in multiple plasmids. This often results in inefficient integration and expression of one or more of the transgenes, especially when relying on DNA recombination for integration (Mansouri and Berger 2014). Another method is by sequential transfection of the host cell with each transgene to generate cell lines iteratively, which is time-consuming and laborious (Cacquevel et al. 2008; McPhaul and Berg 1986; Paul et al. 1985). Several other strategies are now commonly used for co-expression of multiple proteins including IRES-containing polycistronic genes, polyproteins that contain one or more self-cleaving protease sites, and viruses that contain multiple recombinant genes (Berger et al. 2013; Donnelly et al. 2001; Ghattas et al. 1991; Ho et al. 2013). Recently, the PB transposon system has been successfully used for the generation of cell lines co-expressing multiple proteins in CHO and HEK293 host cells (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013). In Alattia et al. (2013), the co-transposition of two PB transposons, each carrying the genes of two of the four γ – secretase subunits and a selection gene, allowed CHO cell lines expressing fully functional γ – secretase to be generated. Using the PB transposon system, a considerable improvement was shown in active γ – secretase production over other available systems. Li et al. (2013) used the PB transposon system to co-transfect up to 2 artificial transposon to generate doxicyclin-inducible HEK-293 cell pools and cell lines expressing monoclonal antibodies. Kahlig et al. (2009) co-transfected HEK-293 cells with up to 4 artificial PB transposons to generate cell lines expressing a multisubunit neuronal voltage-gated sodium channel (SCN1A).

Despite these successes, a thorough study on the optimization of the conditions for the generation of CHO cell pools expressing multiple transgenes has not been performed. Here, we explored this approach using the PB system. Recombinant cell pools co-expressing

multiple proteins were generated by simultaneous delivery of up to four transgenes into CHO-DG44 cells. Enhanced green fluorescent protein (EGFP), secreted alkaline phosphatase (SEAP), and the light (LC) and heavy chains (HC) of an IgG1 antibody were used as the four model proteins. Studies were conducted to create cell pools expressing different combinations of the four transgenes with one or multiple selection genes. The percentage of cells simultaneously co-expressing all the transgenes in a pool was also determined using three different fluorescent proteins, EGFP, Cerulean, and mCherry.

6.2. Results

6.2.1. Protein productivity of cell pools co-expressing multiple proteins

We generated cell pools expressing up to 4 transgenes with and without the PB transposon system using the genes for EGFP, SEAP, and the LC and HC of an IgG1 antibody. Cell pools expressing 1, 2, 3 or 4 transgenes were generated as indicated in Table 6.1. The amount of each donor plasmid was kept constant at 22.5% of the total transfected DNA, irrespective of the number of different plasmids transfected. The helper vector pmPBase or filler DNA was included to 10% of the total transfected DNA or replaced with filler DNA (Table 6.1). All donor plasmids contained the *pac* gene as the selection marker.

Table 6.1. DNA amounts for multigene transfections with and without the PB system.

Cell Pool	Number of transgenes transfected	Amount of DNA transfected [% of total transfected DNA]					
		pMP-PB-IRES-EGFP	pMP-PB-SEAP	pMP-PB-LC	pMP-PB-HC	pmPBase	Filler DNA
TX	1	22.5%	—	—	—	—	77.5%
	2	22.5%	22.5%	—	—	—	55%
	3	22.5%	—	22.5%	22.5%	—	32.5%
	4	22.5%	22.5%	22.5%	22.5%	—	10%
PB	1	22.5%	—	—	—	10%	67.5%
	2	22.5%	22.5%	—	—	10%	45%
	3	22.5%	—	22.5%	22.5%	10%	22.5%
	4	22.5%	22.5%	22.5%	22.5%	10%	—

The cell pools were generated by selecting with two different puromycin concentrations (10 µg/mL and 50 µg/mL) for a duration of 10 days, starting from day 2 post-transfection. The cell pools transfected with PB transposon, when selected with 10 µg/mL or 50 µg/mL puromycin, resulted in pools PB-10p and PB-50p, respectively. The cell pools generated by conventional plasmid transfection resulted in pool TX-10p (selected with 10 µg/mL puromycin) as the cells did not survive 50 µg/mL puromycin. Cell pools expressing one (EGFP), two (EGFP and SEAP), three (EGFP and IgG), or four (EGFP, SEAP, and IgG) transgenes were recovered. The EGFP-specific fluorescence, SEAP activity, and IgG concentration were measured at the end of 4-day batch cultures performed in the absence of selection.

Use of the PB transposon system for the generation of cell pools resulted in higher percentages of EGFP-positive cells (Fig. 6.1A), levels of EGFP-specific fluorescence (Fig.

6.1B), SEAP activity (Fig. 6.1C), and IgG concentrations (Fig. 6.1D) as compared to the cell pools from conventional plasmid transfection. For the use of the PB transposon system in the presence of a high stringency of selection (PB-50p), protein production was generally higher than in the cell pools resulting from selection at a lower level of stringency (PB-10p) (Fig. 6.1 B-D). However, the expression of individual proteins decreased as the total number of transfected genes increased for all three cell pools (Fig. 6.1 A-D).

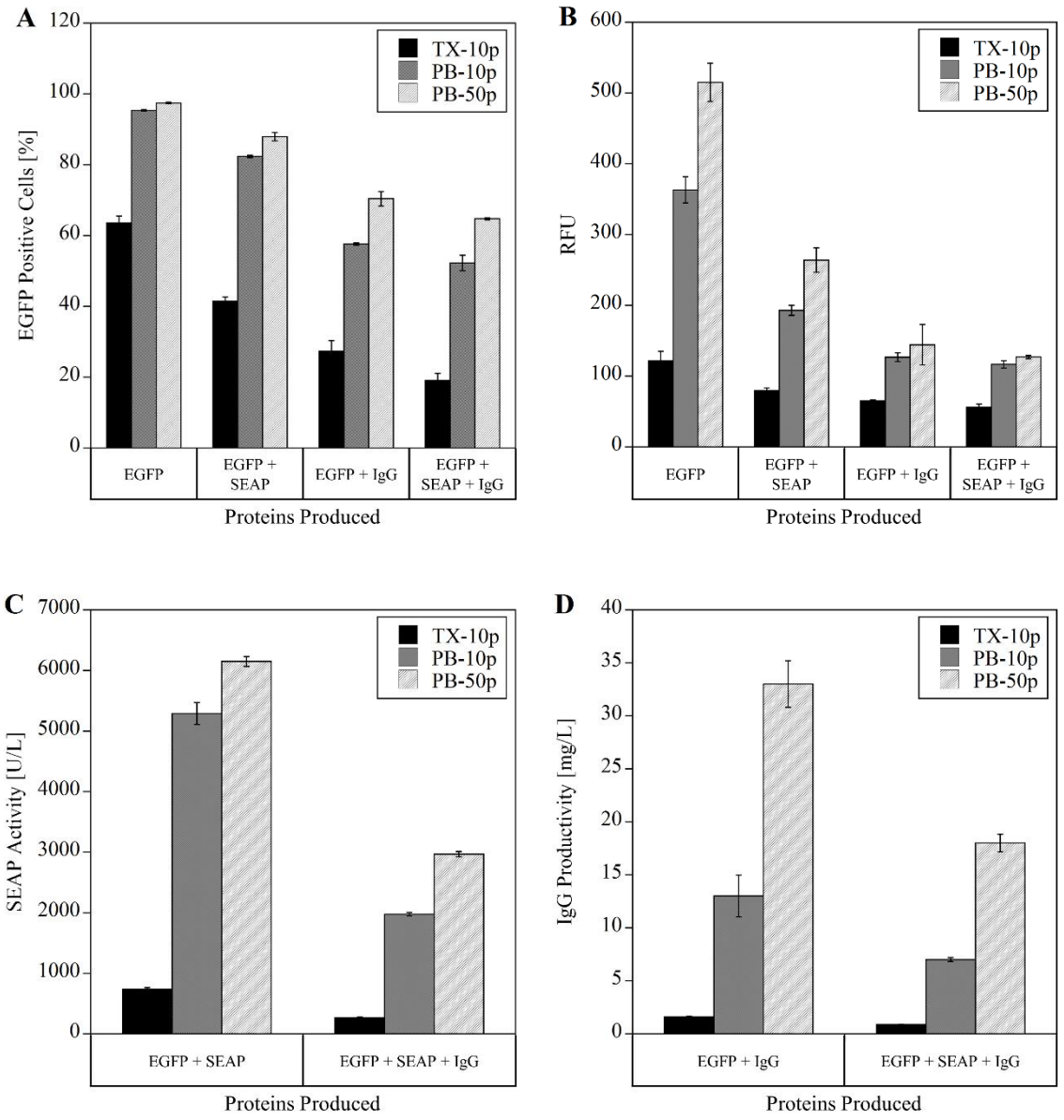

Figure 6.1. Protein production levels for cell pools expressing multiple transgenes. Cell pools expressing 1, 2, 3 or 4 transgenes, as indicated, were analyzed. The cell pools generated in the absence of the PB helper vector were selected with 10 µg/mL puromycin (TX-10p) and the cell pools generated with the PB system were selected with either 10 µg/mL (PB-10p) or 50 µg/mL (PB-50p) puromycin. Protein production was measured at the end of 4-day batch cultures without puromycin. EGFP-specific fluorescence was measured by flow cytometry to determine (A) the % EGFP-positive cells and (B) the relative fluorescence units (RFU). (C) SEAP activity was measured by an enzymatic assay, and (D) the IgG concentration in the medium was measured by ELISA. The error bars represent the standard deviation from 4 different measurements as each culture and each analytical measurement was performed in duplicate.

6.2.2. Analysis of the average integrated transgene copy number in cell pools co-expressing multiple proteins

The number of integrated copies of each transgene in each cell pool was determined by RT-qPCR. In cell pools from all three conditions (TX-10p, PB-10p, PB-50p), the number of copies of each transgene decreased with the number of transgenes co-transfected (Fig. 6.2 A-D). This decrease in transgene copy number correlated to a decrease in the expression level of each protein. The average gene copy number for each gene in the four PB-10p cell pools was lower than that of the corresponding gene in the TX-10p cell pools even though the protein expression levels were higher in the former (Fig. 6.1 A-D). Also, the PB-50p cell pools had a slightly higher number of copies of each gene per cell as compared to the PB-10p cell pools (Fig. 6.2A-D). This difference correlated with higher protein production levels in the PB-50p cell pools (Fig. 6.1 A-D). Interestingly, we observed that the total number of integrated transgene copies for each PB-50p cell pool was similar (Fig. 6.3). Additionally, the number of integrated copies of each transfected transgene for each PB-50p cell pool was similar, with the exception of the IgG-HC gene. A similar trend was also observed for the PB-10p cell pools (Fig. 6.2A-D).

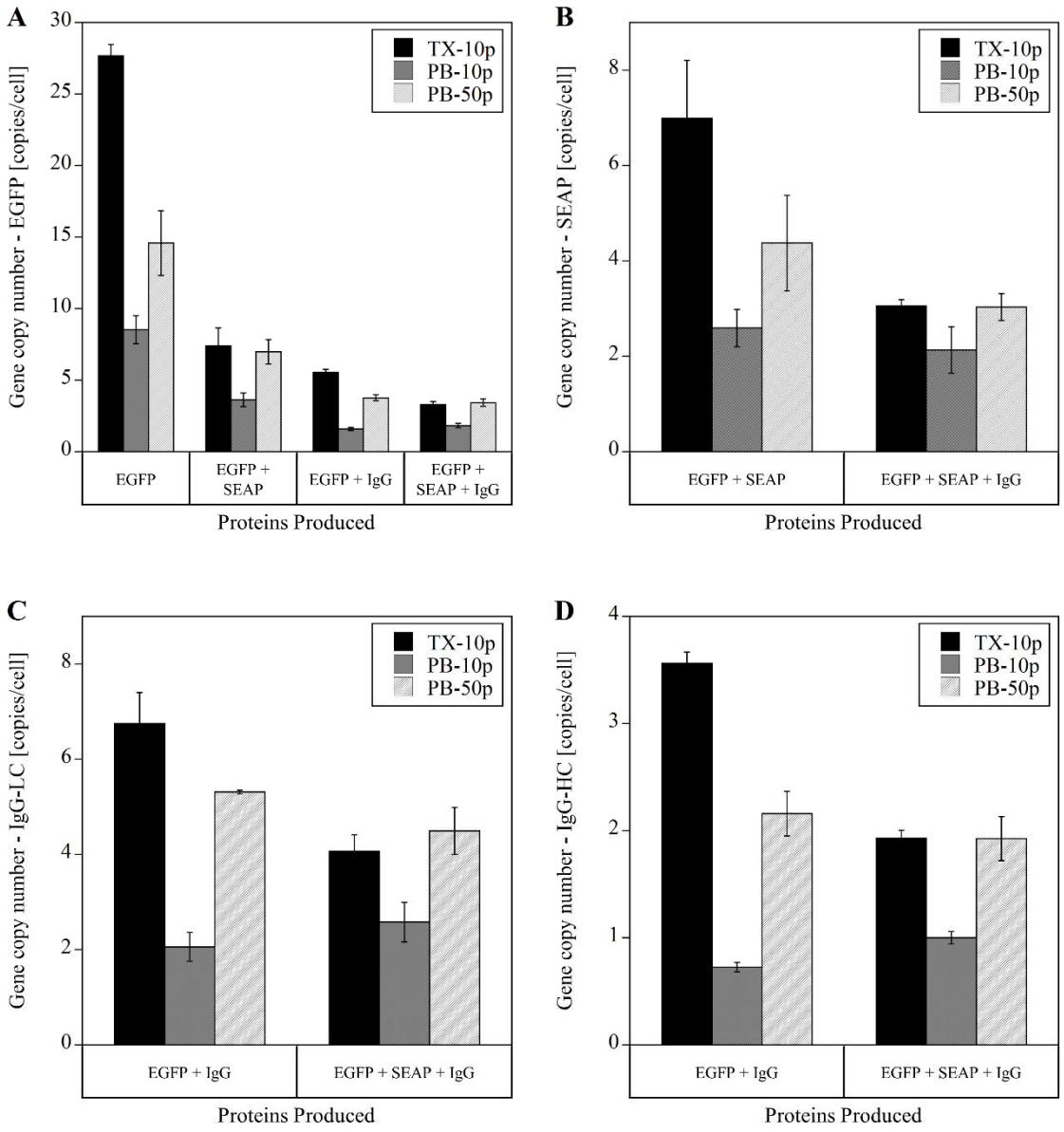

Figure 6.2. Analysis of transgene integration in cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes were generated in the presence (PB-10p, PB-50p) or absence (TX-10p) of PBase. Cell pools expressed the protein(s) as indicated in each bar graph. The cell pools generated in the absence of the PB helper vector were selected with 10 µg/mL puromycin (TX-10p), and the cell pools generated with the PB system were selected with either 10 µg/mL (PB-10p) or 50 µg/mL (PB-50p) puromycin. The average number of integrated (A) EGFP, (B) SEAP, (C) IgG-LC, and (D) IgG-HC genes was estimated by RT-qPCR. The error bars represent the standard deviation from 4 different measurements as each culture and each analytical measurement was performed in duplicate.

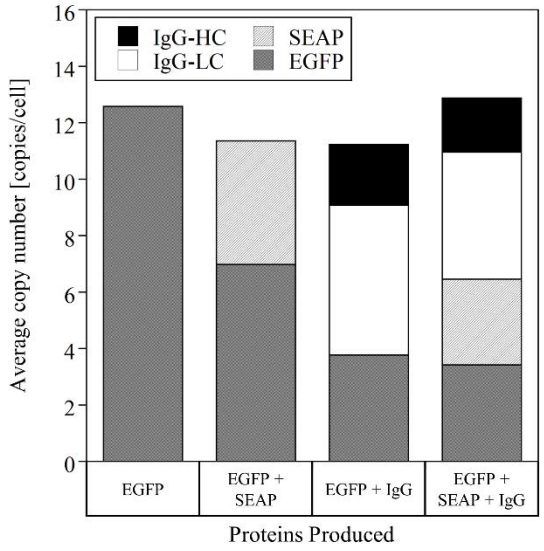

Figure 6.3. Distribution of the integrated gene copies in the PB-50p cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes, as indicated in Table 6.1, were generated in the presence of PBase. The PB-Multiselection cell pools were generated by selecting the cells with 50 μg/mL puromycin (PB-50p). Cell pools were expressing the protein(s) as indicated. The integrated number of each gene was estimated by RT-qPCR. The copy number of each individual transgene is represented in a stack so as to present the total sum integrated gene copy number for each pool.

6.2.3. Protein expression stability of cell pools co-expressing multiple proteins

The stability of protein production of the cell pools described in section 6.2.2 was determined over a 50-day cultivation period in the absence of selection. The production of each protein was determined periodically in 4-day batch cultures. For the four TX-10p cell pools, the % of EGFP-positive cells decreased by 20 – 25 % from the start to the finish of the cultivation period (Fig. 6.4A). But the EGFP expression level, represented by the relative fluorescence units (RFU), remained relatively constant for all four cell pools (Fig. 6.4B). On the other hand, the yields of the secreted proteins SEAP and IgG decreased to about 50 % of the original levels (Fig. 6.4C, D).

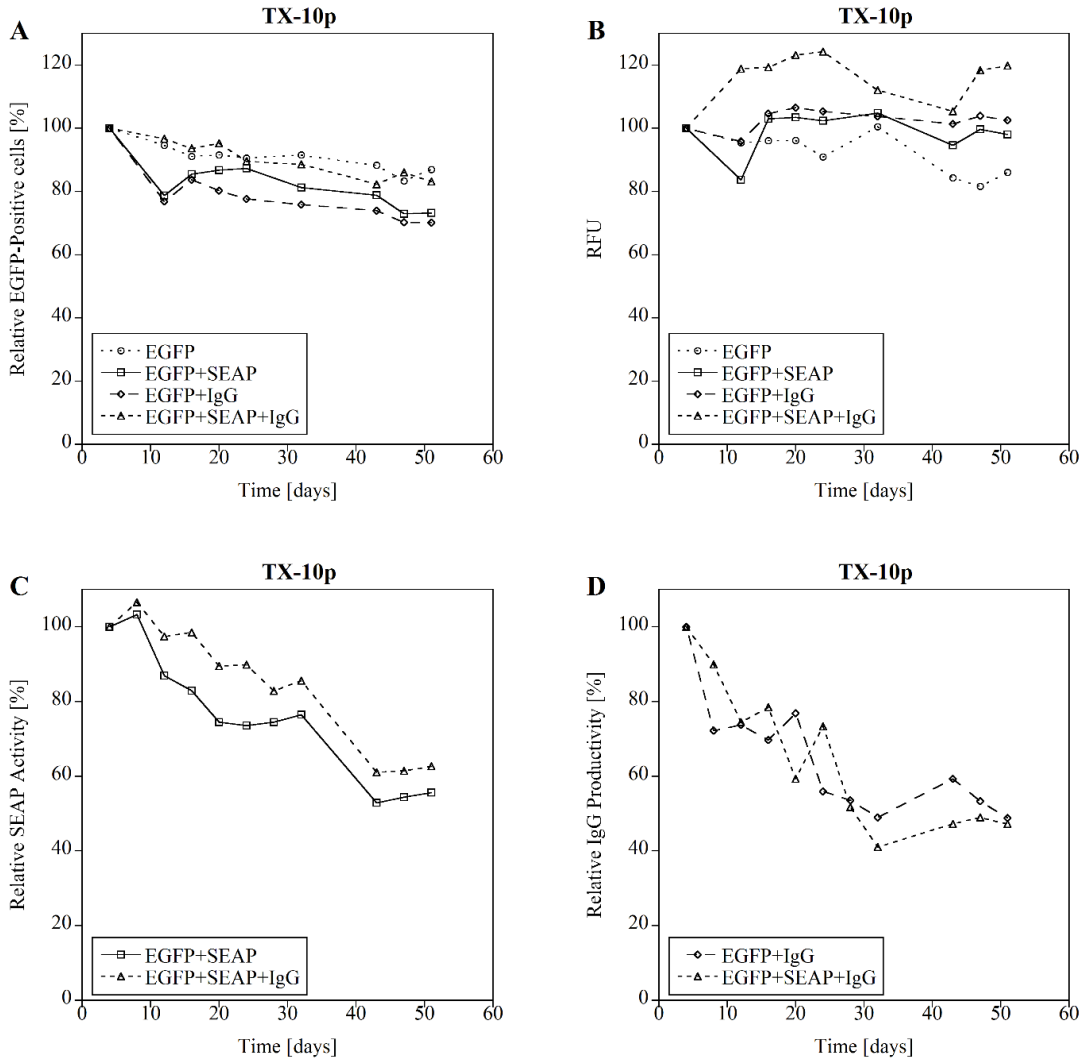

Figure 6.4. Volumetric protein production over time from cell pools resulting from conventional transfection. Cell pools expressing 1, 2, 3 or 4 transgenes were generated in the absence of PBase by selecting with a puromycin concentration of 10 µg/mL (TX-10p). Cell pools expressed the protein(s) as indicated. The stability of protein expression over time was determined by measuring (A, B) EGFP-specific fluorescence, (C) SEAP activity, and (D) IgG concentration. The cell pools were maintained in the absence of selection for 50 days and regularly analyzing at the end of 4-day batch cultures at the times indicated. The data here is represented relative to the expression level at the first point of measurement for each cell pool.

The % of EGFP-positive cells in PB-mediated cell pools was more stable than that observed for the TX-10p cell pools (Fig. 6.5A; Fig. 6.6A). In the PB-10p and PB-50p cell pools at least 90 % of the cells remained EGFP-positive for the duration of the cultivation period in contrast to 70 – 75 % EGFP-positive cells for the TX-10p cell pools (Fig. 6.5A; Fig. 6.6A). As seen for the TX-10p cell pools, the expression levels of EGFP (reported as RFU) appeared to be constant over time for the PB-mediated cell pools (Fig. 6.5B, 6.6B). For PB-10p cell pools, a gradual decline in the SEAP activity and the IgG concentration was

observed over time, as for the TX-10p cell pools (Fig. 6.5C, D). Similar declines in SEAP activity and IgG concentration were observed for the PB-50p cell pools (Fig. 6.6C, D).

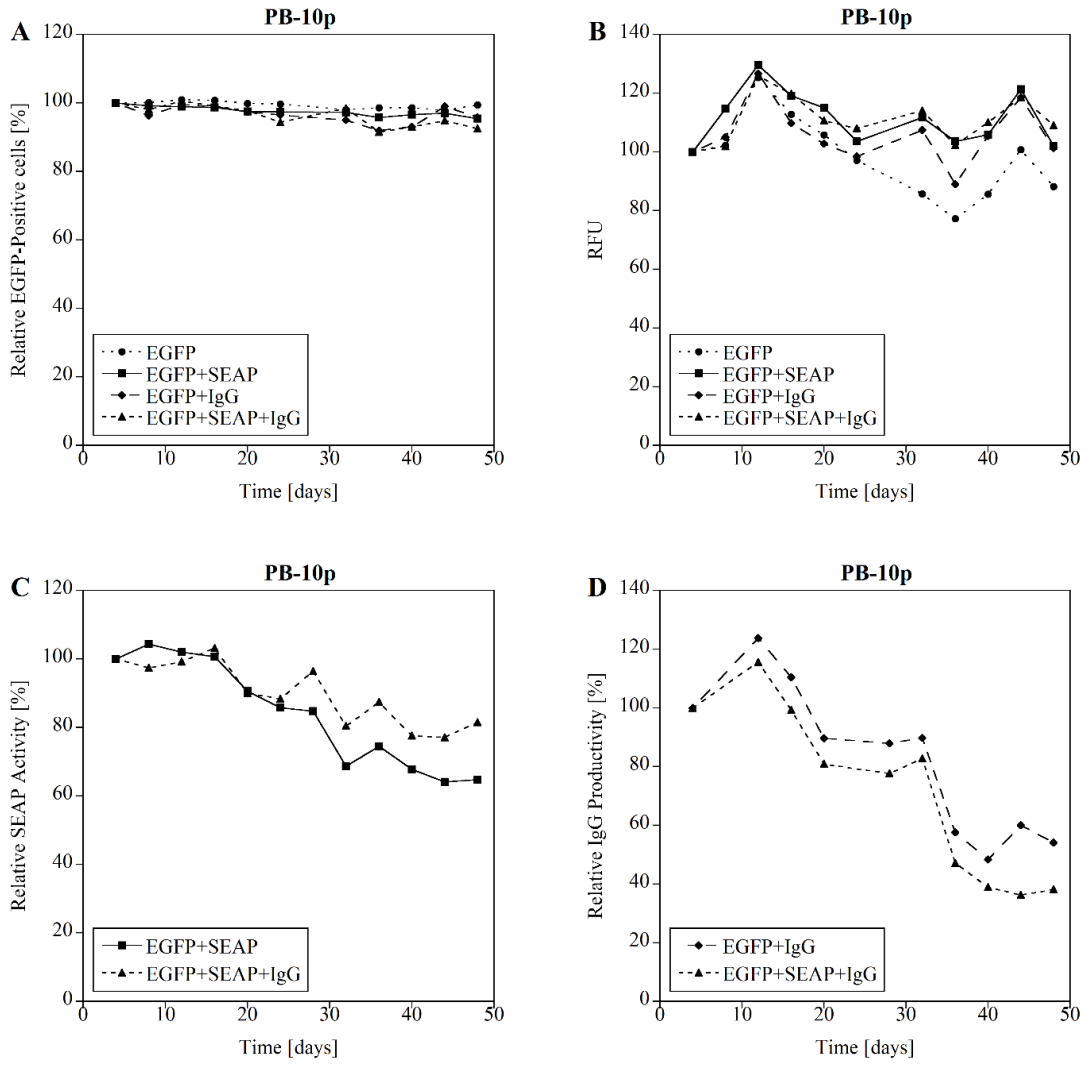

Figure 6.5. Volumetric protein production over time from PB-10p cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes were generated in the presence of PBase by selecting with a puromycin concentration of 10 µg/mL (PB-10p). Cell pools expressed the protein(s) as indicated. The stability of protein expression over time was determined by measuring (A, B) EGFP-specific fluorescence, (C) SEAP activity, and (D) IgG concentration. The cell pools were maintained in the absence of selection for 50 days and regularly analyzing at the end of 4-day batch cultures at the times indicated. The data is represented relative to the expression level at the first point of measurement for each cell pool.

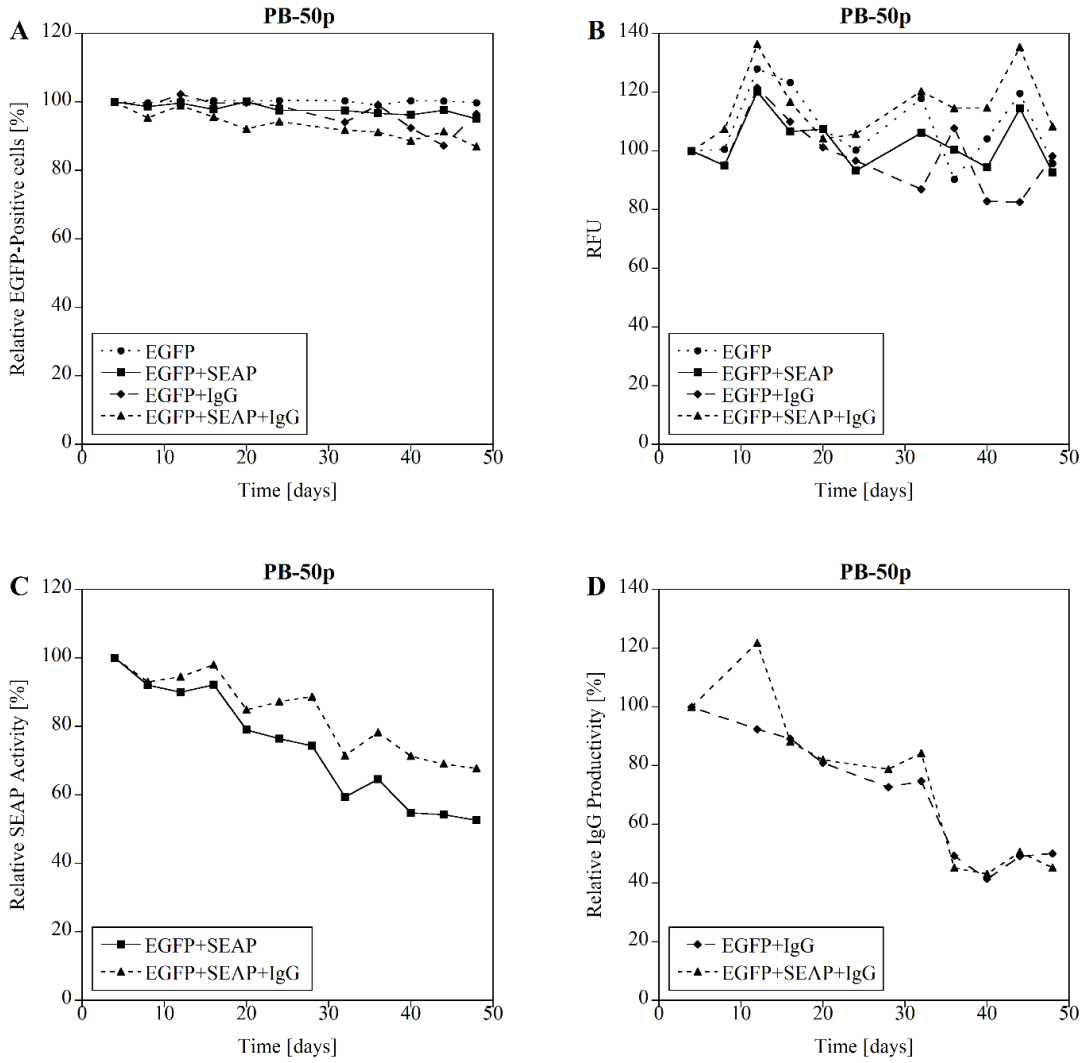

Figure 6.6. Volumetric protein production over time from PB-50p cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes were generated in the presence of PBase by selecting with a puromycin concentration of 50 µg/mL (PB-10p). Cell pools expressed the protein(s) as indicated. The stability of protein expression over time was determined by measuring (A, B) EGFP-specific fluorescence, (C) SEAP activity, and (D) IgG concentration. The cell pools were maintained in the absence of selection for 50 days and regularly analyzing at the end of 4-day batch cultures at the times indicated. The data is represented relative to the expression level at the first point of measurement for each cell pool.

6.2.4. Generation of cell pools co-expressing individually-selected proteins

All of the cell pools described earlier in this chapter were generated by using a single selection gene. However, we observed that in PB-mediated cell pools, the total number of transgene integrations was nearly distributed equally among the different transgenes. With the aim of increasing the number of integrations of each transgene and improving the volumetric productivity and stability of cell pools, cells were transfected with multiple

donor vectors, each having a different selection gene (Table 6.2). The donor plasmids are described in Table 2.1. The selection markers were chosen based on previous results described in Chapter 3 (see Fig. 3.3). PB-mediated cell pools co-expressing 1, 2, 3, or 4 transgenes were generated as indicated in Table 6.2. The cells were then co-selected with the appropriate combination of 10 µg/mL puromycin, 10 µg/mL blasticidin, 200 µg/mL hygromycin B, and 50 µg/mL zeocin to generate PB-Multiselection cell pools for each of the four transfections. The cells were placed under selection pressure until the cell viability increased to >90 %. The EGFP-specific fluorescence, SEAP activity, and IgG concentration were measured at the end of 4-day batch cultures in the absence of selective pressure. The productivities of the cell pools are shown in comparison with PB-mediated cell pools generated by selection with puromycin (50 µg/mL) as described in section 6.2.1 (PB-50p).

Table 6.2. DNA amounts for transfections using multiple selection markers.

Number of transgenes transfected	Amount of DNA transfected [% of total transfected DNA]					
	pSB-BPB-IRES-EGFP[a]	pSB-HPB-SEAP[b]	pMP-PB-LC[c]	pSB-ZPB-HC[d]	pmPBase	Filler DNA
1	22.5%	—	—	—	10%	67.5%
2	22.5%	22.5%	—	—	10%	45%
3	22.5%	—	22.5%	22.5%	10%	22.5%
4	22.5%	22.5%	22.5%	22.5%	10%	—

[a]Contains selection marker *bsr* (blasticidin S-resistance gene); [b]Contains selection marker *hph* (hygromycin B phosphotransferase gene); [b]Contains selection marker *pac* (puromycin N-acetyl-transferase); [d]Contains selection marker Sh *ble* (*ble* gene from *Streptoalloteichus hindustanus)*.

For PB-Multiselection cell pools, EGFP expression (Fig. 6.7A, B) and SEAP activity levels (Fig. 6.7C) were slightly higher than for corresponding PB-50p cell pools. In contrast, IgG production was lower in the PB-Multiselection cell pools as compared to the PB-50p cell pools (Fig. 6.7D).

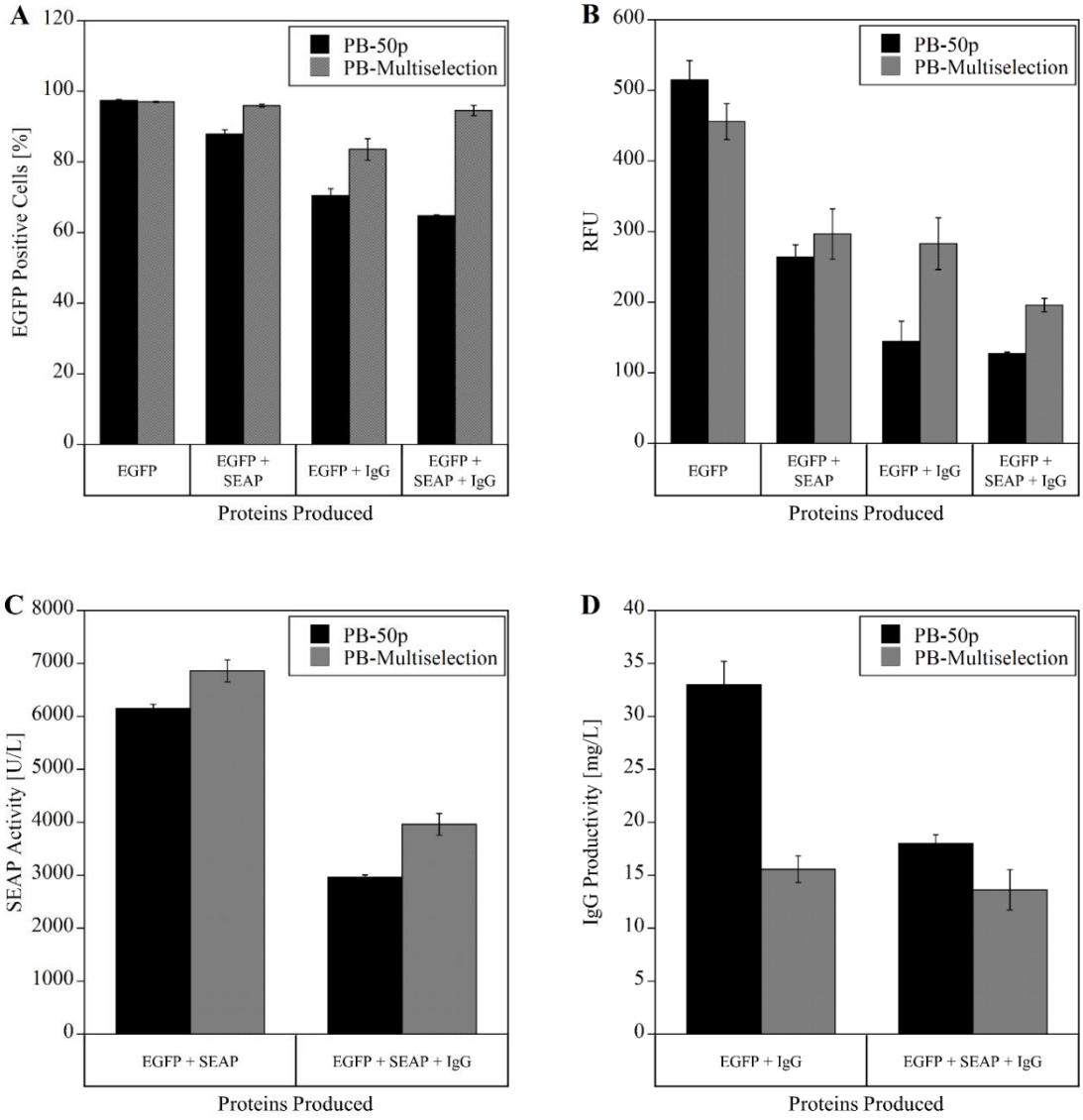

Figure 6.7. Protein expression levels of PB-Multiselection cells pools. Cell pools expressing 1, 2, 3 or 4 transgenes, as indicated in Table 6.2, were generated in the presence of PBase. Cell pools expressed the protein(s) as indicated. The PB-50p cell pools were generated by selecting with a puromycin concentration of 50 µg/mL as described in Section 6.2.1. The PB-Multiselection cell pools were generated by selecting with the appropriate combination of antibiotics as indicated in Table 6.2. (A, B) EGFP-specific fluorescence was measured by flow cytometry, (C) the SEAP activity was measured by an enzymatic assay, and (D) the IgG concentration was measured by ELISA at the end of 4-day batch cultures in the absence of selection. The error bars represent the standard deviation from 4 different measurements as each culture and each analytical measurement was performed in duplicate.

The number of integrated transgene copies in the PB-50p and PB-Multiselection cell pools was determined by RT-qPCR. In contrast to the minor differences in expression levels of the different proteins in the two sets of cell pools, there were significant differences in the integrated copy numbers of each transgene. In case of the PB-50p cell pools, we observed

a decrease in the integrated copy number of each transgene with an increase in the number of transfected genes. For the PB-Multiselection cell pools, however, the integrated copy number of each transgene was similar irrespective of the number of transgenes transfected (Fig. 6.8 A-D).

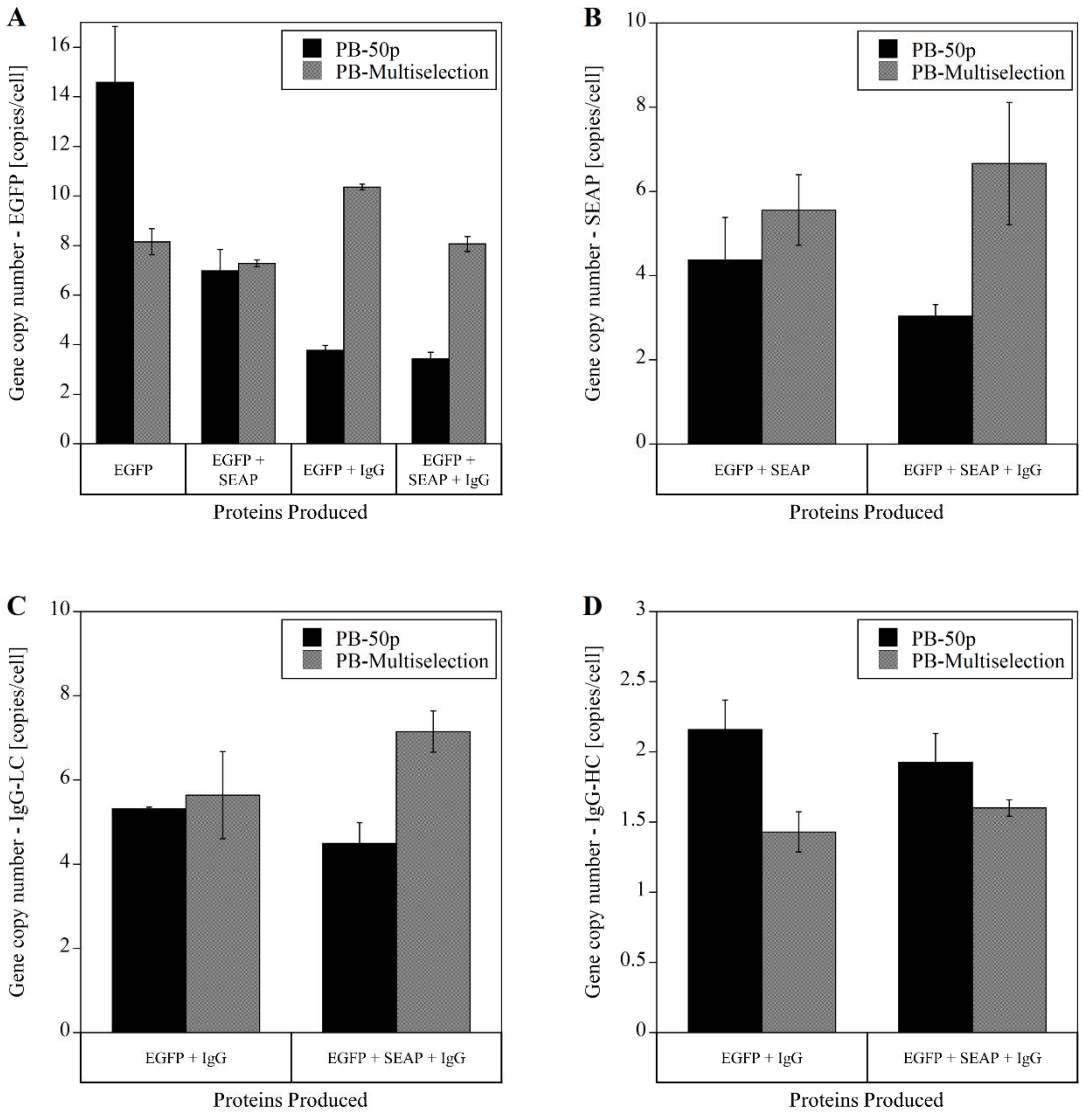

Figure 6.8. Average integrated gene copy number in PB-Multiselection and PB-50p cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes were generated as described in Figure 6.7. The integrated copy number of the (A) EGFP, (B) SEAP, (C) IgG-LC, and (D) IgG-HC genes was estimated by RT-qPCR. The error bars represent the standard deviation from 4 different measurements as each culture and each analytical measurement was performed in duplicate.

In pools PB-10p and PB-50p, we observed a total of 10-12 integrated transgenes per cell with the IgG-HC gene (2 copies) being less represented than the three other genes (3-4

copies) (Fig. 6.3). For the PB-Multiselection cell pools, there were about 6-8 integrated copies of each gene with the exception of the IgG-HC gene (1.5 copies) (Fig. 6.9). Thus, the total number of integrated transgenes per cell increased by selecting each individual GOI with the total number of integrated genes reaching 25 for the cell pool expressing all 4 transgenes (Fig. 6.9). However, this increase in the number of integrated transgenes did not translate into increased protein productivity (Fig. 6.7).

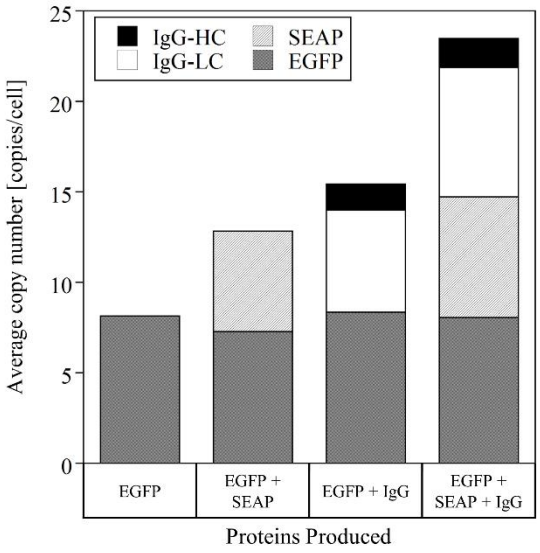

Figure 6.9. Distribution of the integrated gene copies in the PB-Multiselection cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes, as indicated in Table 6.2, were generated in the presence of PBase. The PB-Multiselection cell pools were generated as described in Figure 6.7. The integrated copy number of each gene was estimated by RT-qPCR. The copy number of each individual transgene is represented in a stack so as to present the total number of transgene copies for each pool.

To determine the stability of protein production by the PB-Multiselection cell pools, each one was cultivated in the absence of selection for 4 weeks. The level of each protein was determined periodically from 4-day batch cultures. For the four cell pools, the percentage of EGFP-positive cells remained constant with time (Fig. 6.10A). In contrast, for three of the four cell pools, the level of EGFP-specific fluorescence per cell declined with time (Fig. 6.10B). The levels of SEAP activity and IgG concentration tended to decrease with time as observed for the PB10p and PB-50p cell pools (Fig. 6.10C, D).

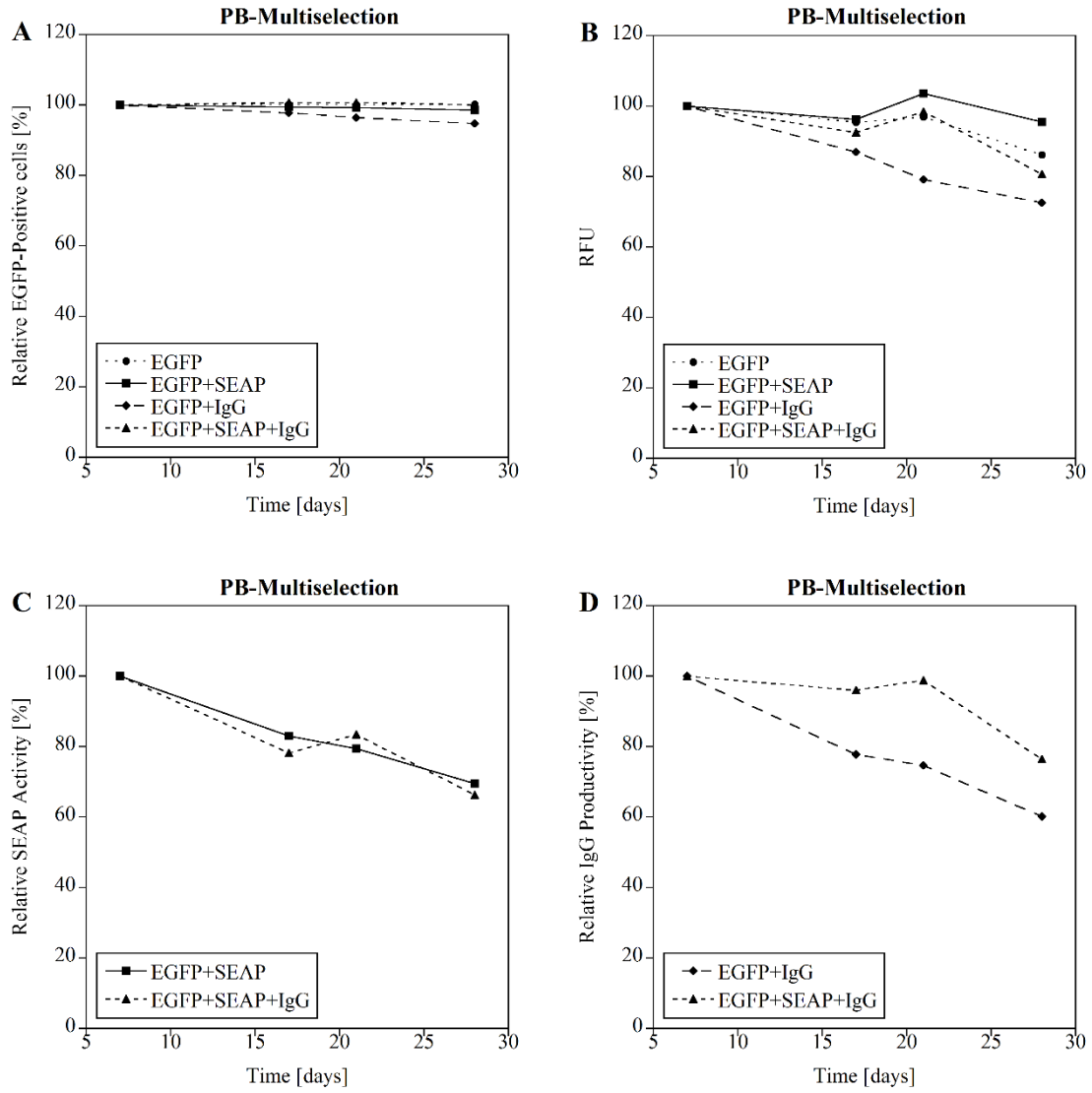

Figure 6.10. Protein production with time in culture for PB-Multiselection cell pools. Cell pools expressing 1, 2, 3 or 4 transgenes, as indicated in Table 6.2, were generated in the presence of PBase by selecting the cells with the appropriate combination of antibiotics as described in Figure 6.7. Cell pools expressed the protein(s) as indicated. The levels of (A, B) EGFP-specific fluorescence, (C) SEAP activity, and (D) IgG concentration were assessed by maintaining the cells in absence of selection for 30 days, and the protein levels were analyzed at the times indicated at the end of 4-day batch cultures. The data is represented relative to the expression level at the first point of measurement for each cell pool.

6.2.5. Analyses of cell pools co-expressing multiple fluorescent proteins

PB-mediated cell pools expressing 3 proteins (PB-10p, PB-50p and PB-Multiselection) had up to 35-fold higher productivity levels compared to those generated by conventional plasmid transfection (TX-10p) (see Fig. 6.1). We hypothesized that these cell pools had a higher percentage of cells expressing all three proteins than the pools generated by conventional transfection. To test this hypothesis, cell pools co-expressing Cerulean (blue),

EGFP (green) and mCherry (red) fluorescent proteins were generated with and without the PB transposon system as described in Table 6.3. All donor plasmids contained the *pac* gene (resistance against puromycin) as the selection marker. The cells pools were generated by selecting with puromycin (10 µg/mL) for 10 days starting from day 2 post-transfection.

Table 6.3. DNA amounts for cell pool generation with 3 different fluorescent proteins.

Cell pool	Percentage of DNA transfected [% of total transfected DNA]				
	pSB-PPB-Cer	**pMP-PB-EGFP**	**pSB-PPB-mCh**	**pmPBase**	**Filler DNA**
TX-fluorescent	33.3%	33.3%	33.3%	—	10%
PB-fluorescent	33.3%	33.3%	33.3%	10%	-

The percentage of cells in each cell pool (TX-fluorescent and PB-florescent) that expressed one or more fluorescent proteins was determined using a flow cytometer. In the TX-fluorescent cell pool, 30% of the cells failed to express any of the fluorescent proteins, while 7% of the cells expressed all three proteins (Fig. 6.11). For the PB-fluorescent cell pool, in contrast, only 4% of the cells failed to express any of the three fluorescent proteins, while 23% of the cells expressed all three proteins (Fig. 6.11). These results, in combination with the protein production improvements seen for the PB-10p, PB-50p and PB-Multiselection cell pools compared to the TX-10p cell pools, indicated that the probability of finding cells co-expressing multiple transgenes would be higher from a transposon-mediated cell pool than from a cell pool generated by conventional transfection. This would in turn increase the efficiency of generating multigene-expressing clonal cell lines from the transposon-mediated cell pools.

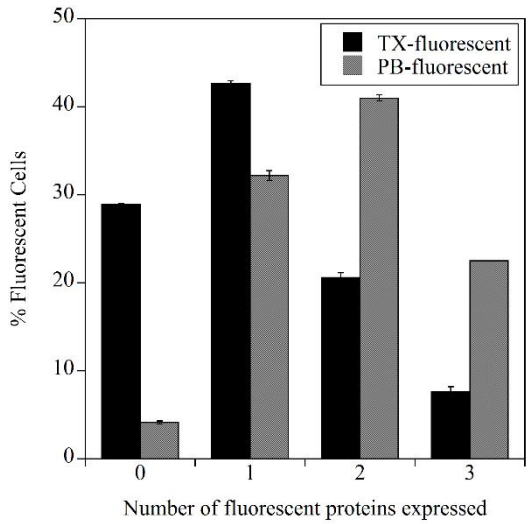

Figure 6.11. Fluorescence distribution profile in PB-fluorescent and TX-fluorescent cell pools. Cell pool expressing the fluorescent proteins Cerulean, EGFP and mCherry were generated by transfecting with pSB-PPB-Cer, pMP-PB-EGFP and pSB-PPB-mCh in a 1:1:1 ratio in the absence (TX-fluorescent) and presence (PB-fluorescent) of PBase. The X-axis indicates the number of fluorescent proteins expressed by a cell. This distribution was generated from 50,000 cells. The error bars represent the standard deviation between experimental duplicates.

6.2.6. Generation of clonal cell lines co-expressing multiple proteins.

We recovered 48 clonal cell lines from the PB-50p cell pool expressing EGFP, SEAP, and IgG as described in Section 6.2.1. The EGFP-specific fluorescence, SEAP activity, and IgG concentration for each cell line was measured at the end of 4-day batch cultures in 24-well plates. About 10 % of the cell lines failed to express one of the proteins, 30 – 35 % expressed one or two proteins, and 25 % expressed all three proteins, albeit at different levels (Fig. 6.12). This correlated with the fluorescence distribution profile of the PB-fluorescent cell pool (Fig. 6.11).

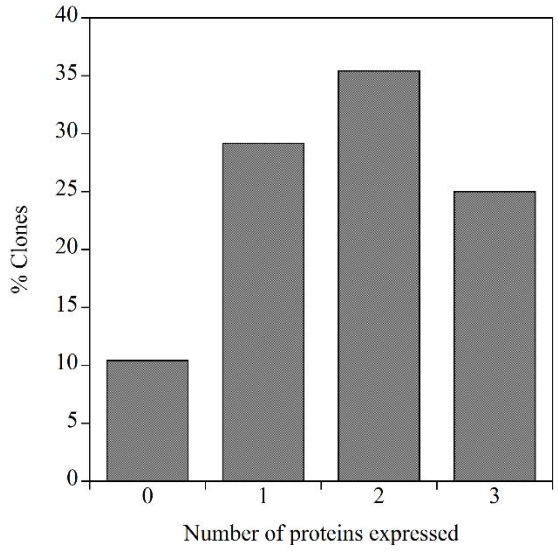

Figure 6.12. Distribution profile of cell lines. Clonal cell lines (n=48) were generated from the PB-50p cell pool expressing EGFP, SEAP, and IgG. The X-axis indicates the number of proteins expressed by the cell lines. The Y-axis indicates the % of cell lines expressing said number of proteins. EGFP-specific fluorescence was measured by FACS, SEAP activity was measured by an enzymatic assay, and the IgG concentration was measured by ELISA from 4-day batch cultures.

6.3. Discussion

The PB transposon system has been successfully used for the generation of cell lines co-expressing multiple proteins in CHO and HEK293 host cells (Alattia et al. 2013; Kahlig et al. 2009; Li et al. 2013). In this chapter, we studied the generation of cell pools and cell lines expressing up to 4 transgenes with the PB transposon system and by conventional transfection using EGFP, SEAP, and the light and heavy chains of an IgG1 antibody as the four model proteins.

The productivities of the PB-10p and PB-50p cell pools were higher than those of the conventionally generated TX-10p cell pools. The expression levels of SEAP and IgG were up to 10 fold and 35 fold higher, respectively, in PB-mediated cell pools compared to the cell pools generated by conventional transfection. Increasing the selection stringency for the generation of the PB-50p cell pools resulted in only a slight improvement in protein production except for IgG, as compared to the PB-10p cell pools. By conventional transfection, we were able to generate cell pools co-expressing all four proteins, albeit with low expression levels. The expression level of each protein for the TX-10, PB-10p and PB-50p cell pools was found to decrease as the number of transfected transgenes increased. For the PB-10p and PB-50p cell pools, the number of integrated copies of each transgene also decreased as the number of transfected transgenes increased.

For PB-10p and PB-50p cell pools, the total number of integrated transgenes was similar (about 10-12 copies/cell) irrespective of the number of different donor vectors transfected. In these cell pools, each gene except the IgG-HC gene, was equally represented among the integrated transgenes. This indicated that with the PB transposon system there was little bias in transgene integration. In cell pools in which IgG was produced, there was usually only 1-2 copies/cell of the IgG-HC gene, in spite of the vector backbone being the same as that of the other donor plasmids. The size of the artificial transposon could not have been a factor since the SEAP and IgG-HC genes have about the same number of base pairs. It has also been shown that cell lines expressing high levels of anti-rhesusD IgG had fewer integrated copies of the IgG-HC gene compared to that of the IgG-LC gene (Chusainow et al. 2009; Ho et al. 2013; Li et al. 2007; O'Callaghan et al. 2010). It may be that similar number of integrated copies of the HC and LC genes are detrimental to IgG production.

Although there was an increase in transgene expression for cell pools derived with the PB transposon system than those resulting from conventional transfection, the PB-10p and PB-50p cell pools had similar losses in protein production over time as did the TX-10p cell pools. Our studies of the stability of protein production demonstrated a difference in the behavior of secreted proteins versus intracellular proteins. The EGFP production level tended to be constant over time whereas the production of SEAP and IgG decreased with time. Further investigation is required to identify if the reason for stable expression of EGFP was protein-specific.

We used genes conferring resistance to different selection markers on each transgene with the aim to increase the number of integrated transgenes and consequently improve the level and stability of protein productivity. The PB-Multiselection cell pools were generated by simultaneously selecting with up to four antibiotics at their optimal concentrations. The resulting cell pools had only slightly higher protein production levels compared to the PB-50p cell pools even though the integrated copy number of each transgene was higher in PB-Multiselection cell pools than in the comparable PB-50p cell pools. For the PB-Multiselection cell pools, 6-8 copies of each transgene, except the IgG-HC gene (1.5 copies/cell), were present. We observed the same phenomenon in the PB-10p and PB-50p cell pools. Thus, the total sum of transgene integration events increased with the increase in the number of different transgenes transfected. The expression levels of the individual proteins in the PB-Multiselection cell pools decreased over a cultivation period of 30 days in the absence of selection. This trend was also observed for the PB-10p and PB-50p cell pools. Further studies are required to understand the principal cause of the decline in protein productivity over time.

The fact that we saw an increase in the average integrated transgene copy number in the PB-Multiselection cell pools compared to the PB-50p cell pools, without a proportional increase in in protein yield, lead us to speculate that there may be other limitations to protein yield besides transgene copy number. There could be transcriptional, translational, or post-translational limitations to protein production. It is also possible that the selection regime necessary for the generation of the PB-Multiselection cell pools increased the frequency of transgene integration by DNA recombination rather than transposition. Transgenes

integrated by DNA recombination would be expected to have had a minimal contribution to transgene expression levels.

We also generated cell pools expressing three different fluorescent proteins. With these cell pools we were able to analyze the percentage of cells expressing 1, 2 or 3 proteins. The results showed that the percentage of cells co-expressing all 3 fluorescent proteins increased by 3 fold in the PB-fluorescent cell pool as compared to the TX-fluorescent cell pool generated by conventional transfection. This indicates that the use of PB-mediated cell pools may significantly reduce the screen size required for generating cell lines co-expressing multiple proteins. This was supported by our study where we were able to obtain 12 out of 48 clonal cell lines from PB-50p cell pools co-expressing the three proteins EGFP, SEAP and IgG.

In our experience, use PB transposon system improved both the volumetric productivities and the percentage of cells co-expressing multiple proteins in the cell pool as compared to those generated by conventional transfection, making it a viable strategy for multi-protein production both from cell pools and cell lines.

7.

Conclusions and Perspectives

7.1. Conclusions

The overall goal of this thesis was to evaluate the efficiency of generation of cell pools and cell lines in CHO cells by using transposon systems for catalyzing transgene integration. Three of the commonly used transposon systems, piggyBac (PB), Tol2, and Sleeping Beauty (SB), were studied.

Through optimizations of several parameters we showed that PB-mediated cell pools could be generated in CHO-DG44 cells with a selection duration of 5-10 days in the presence of puromycin, resulting in cell pools with productivities 3-4 times higher than those generated by conventional plasmid transfection. We were able to generate cell pools with similar volumetric productivities by addition of the selection agent anytime between 1-3 days post-transfection. Similar volumetric productivities were obtained from cell pools by selecting with different antibiotic resistance genes. In 14-day batch cultures, TNFR:Fc productivities of 350-550 mg/L were reproducibly obtained from five independent PB-mediated cell pools. Thereby, we established a reproducible protocol for the generation of transposon-mediated cell pools.

The PB, Tol2 and SB transposon systems behaved similarly in terms of cell pool and cell line generation in CHO-DG44 cells. All three transposon systems generated cell pools with similar volumetric TNFR:Fc productivities (80 – 100 mg/L) and integrated transgene copies (10 – 12 copies/cell). The volumetric productivities of the transposon-mediated cell pools were ~9 fold higher than those generated by conventional transfection using donor vectors without inverted transposon repeats. Some of the integration events in the cell pools occurred via DNA recombination rather than transposition. About 75 % of the integration events in PB- and SB- mediated cell pools occurred via transposition, whereas only 40 % of integrations were via transposition in Tol2-mediated cell pools. The higher levels of protein production observed in the transposon-mediated cell pools correlated with higher levels of transgene mRNA (Table 7.1). As expected, the average volumetric TNFR:Fc productivity of transposon-derived cell lines was higher than that of cell lines generated by conventional transfection (Table 7.1). From 14-day fed-batch cultures, protein levels up to 900 mg/L and 1.5 g/L were obtained from transposon-mediated cell pools and cell lines, respectively. The productivity of the cell pools, however, decreased to 50 % the initial value

when cultivated in the absence of selection for 3 months. However, the volumetric protein yield from transposon-mediated cell lines remained constant for up to 4 months in the absence of selection.

The method of generating transposon-mediated cell pools was also applied to two other CHO cell strains, CHO-K1 and CHO-S. We showed that the three transposon systems could be used for cell pool generation with CHO-K1 and CHO-S with similar volumetric productivities as observed with CHO-DG44 cells (70 – 100 mg/L). However, the expression levels of the transposon-mediated cell pools in CHO-K1 and CHO-S were unstable and demonstrated a decrease in productivity to the level of cell pools generated by conventional transfection. Thus, CHO-DG44 was the best host cell choice for the generation of cell pools using transposon systems for gene delivery.

We also utilized the PB transposon system for generating cell pools co-expressing up to 4 different transgenes for EGFP, SEAP, and the LC and HC of an IgG1 monoclonal antibody, by simultaneous gene transfer. We showed that PB-mediated cell pools had increased volumetric productivity of each of the proteins compared to those generated by conventional co-transfection. We compared two strategies for generation of cell pools, using either the same or different selection markers for each transgene. Although co-selecting the cells with multiple antibiotics showed differences in the viability profiles and the integrated gene copy numbers, it did not have a significant effect on protein productivity and stability compared to the cell pools selected with single selection marker. By analyzing the composition of the cell pools and generation of cell lines we demonstrated that the use of the PB transposon system increased the percentage of cell population that co-expressed all four proteins as compared to the results with cell pools generated by conventional transfection.

Based on our experience with the use of transposon systems for recombinant protein production, we deduced that this method provides a good balance between gene delivery in the presence of DNA regulatory elements and targeted gene integration in terms of the ease of setup and the output generated. The vector engineering required for the construction of artificial transposons is very simple, making it easily applicable (Table 7.1). Transposon

systems are a semi-targeted system which do not require host cell engineering or pre-characterization of integration hotspots in the host cell genome, thereby eliminating the requirement of cell line- or protein-specific optimization for each protein production project. In conclusion, the transfection of CHO-DG44 cells with the PB, Tol2 or SB transposon system is a simple, efficient, and reproducible approach to the generation of cell pools and cell lines for the rapid production of recombinant proteins.

Table 7.1. Summary of advantages of transposon system for recombinant protein production

Method for gene integration/Cell line generation		
Transposon systems	**Advantages**	Simple construction and delivery
		Semi-targeted integration
		Better utilization of integrated transgene copies for gene expression
		Reduction in screen size to identify suitable clonal cell lines
		Easily applicable to different GOIs tested in this study
		Can be used in combination with several other integration methods

7.2. Future Perspectives

Although we have studied the use of transposon systems for cell pool and cell line generation in detail, there are various aspects that still need to be better defined.

7.2.1. Product quality analysis for cell pools

Cell pools are typically not the method of choice for therapeutic protein production due to the heterogeneity of the cell population. However, there is currently some interest in using cell pools as a method for generating material for preclinical stages of therapeutic protein development. Here, we have improved the volumetric productivities of cell pools using transposon systems. To eventually be used for early drug development stages, it is important that the products from cell pools are similar to those produced from cell lines. Hence there is a need to compare product attributes, such as the glycan profile of glycosylated proteins, protein degradation, protein aggregation, and posttranslational modifications other than glycosylation of recombinant proteins produced from cell pools and cell lines.

Additionally, the awareness of intraclonal genetic variation of cell lines in culture, resulting in a heterogeneous population not unlike a cell pool, increases the likelihood of use of cell pools for commercial protein production (Pilbrough et al. 2009; Wurm 2013). Thus, a similar quality of protein produced from cell pools and cell lines will increase the possibility of using cell pools for the production of material for clinical trials, thereby reducing the time-frame to get the product to market.

7.2.2. Co-transfection of donor vector with mRNA of transposase

In this study, a helper vector coding for the transposase was used for gene delivery and the enzyme was transiently expressed. The main drawback of using a helper vector for expression of the transposase is the possibility of integration of the transposase gene into the genome, even though the vector does not carry a selection marker. Another disadvantage of this approach is the presence of the helper plasmid for several days post-transfection. This could result in continued transient expression of the transposase and remobilization of the integrated transposons, potentially leading to genotoxicity. Thus,

generation of transposon-mediated cell lines by using a helper vector for transposase expression would require the cell lines to be tested for transposase activity and integrated copies of the transposase, adding an extra step to the characterization of production cell lines.

This could be avoided by delivering the transposase gene as an mRNA. This method has been shown to eliminate the risk of long-lasting side effects due to the rapid degradation of exogenous mRNA by cellular RNases and lack of integration into the genome of transfected HeLa cells (Bire et al. 2013a). Delivery of the transposase mRNA reduced the persistence of the transposase to a narrow window peaking at 18 h post-transfection with only traces of the mRNA remaining at 48 h post-transfection. On the other hand, the intracellular amount of transposase in transfection using helper vector peaked at 36 h post-transfection and then its level plateaued (Bire et al. 2013b). The delivery of the transposase gene as an mRNA needs to be explored in CHO cells for the generation of cell pools and cell lines, thereby improving their biosafety. This would eliminate any concerns of integration of the transposase gene, thereby making it a more reliable system to use for the generation of commercial cell lines.

7.2.3. Combining transposons with other technologies

One of the advantages of transposon systems is that they can be used in combination with other technologies targeting transcriptional, translational, or post-translational improvements in protein production.

The construction of the artificial transposon in this study uncoupled the expression of the selection marker and the GOI by placing them under different promoters. Methods such as IRES-containing polycistronic genes or polyproteins could also be combined with transposon systems for co-expression of the selection marker downstream of the transgene in a polycistronic format. This strategy may increase the stringency of selection and also improve the homogeneity of protein production in cell pools, potentially improving the volumetric productivity and production stability. These methods could also be used to generate artificial transposons expressing different subunits of a multi-subunit protein complex. Combining the ideas of co-expressing multiple transgenes from a single artificial

transposon and co-transfecting cells with multiple artificial transposons, can potentially simplify the expression of large multi-subunit protein complexes.

Cis-regulating elements such as S/MAR and UCOE are known to increase protein yields and the stability of transgene expression in mammalian cells by promoting position-independent transgene expression (Benton et al. 2002; Dharshanan et al. 2014; Grandjean et al. 2011; Kwaks and Otte 2006). Transposon systems can be used along with these *cis*-regulating elements to attempt to improve the stability of protein production over time in cell pools. Use of S/MAR-containing artificial PB transposons has been shown to improve the volumetric productivity of an antibody in mesoangioblasts and CHO-M cells (Ley et al. 2013; Ley et al. 2014).

7.2.4. Investigation of the transposon integration locus in CHO cells

There have been several studies comparing the activities and integration profiles of PB, Tol2, and SB transposon systems in human cells for applications in gene therapy (Grabundzija et al. 2010; Huang et al. 2010; Meir et al. 2011; Wu et al. 2006). Comparisons of integration profiles of the three TP systems in human cells showed that the PB and Tol2 transposons integrate in intragenic regions and near transcriptional start sites whereas the SB transposon integrated with apparent randomness across the genome within transcriptionally active regions (Grabundzija et al. 2010; Huang et al. 2010; Meir et al. 2011). Studies exploring the integration locus of these transposon systems have not been conducted in recombinant CHO cell lines. Recent studies have determined the genome sequence of the CHO-K1, -DG44, and –S strains (Lewis et al. 2013; Xu et al. 2011). This major milestone can be utilized to investigate the preferential integration sites for the three transposon systems. This information would give further insight into the functioning of the transposon systems and would help improve their validity for commercial recombinant protein production in the biotechnology industry.

8.

References

Aggarwal RS. 2014. What's fueling the biotech engine-2012 to 2013. Nat Biotechnol 32(1):32-9.

Akselband Y, Moen PT, Jr., McGrath P. 2003. Isolation of rare isotype switch variants in hybridoma cell lines using an agarose gel microdrop-based protein secretion assay. Assay Drug Dev Technol 1(5):619-26.

Alattia JR, Matasci M, Dimitrov M, Aeschbach L, Balasubramanian S, Hacker DL, Wurm FM, Fraering PC. 2013. Highly efficient production of the Alzheimer's gamma-secretase integral membrane protease complex by a multi-gene stable integration approach. Biotechnol Bioeng 110(7):1995-2005.

Allen G, Spiker S, Thompson W. 2000. Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Molecular Biology 43(2-3):361-376.

Antoniou M, Harland L, Mustoe T, Williams S, Holdstock J, Yague E, Mulcahy T, Griffiths M, Edwards S, Ioannou PA, 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(3):269-79.

Backliwal G, Hildinger M, Hasija V, Wurm FM. 2008. High-density transfection with HEK-293 cells allows doubling of transient titers and removes need for a priori DNA complex formation with PEI. Biotechnol Bioeng 99(3):721-7.

Balciunas D, Wangensteen KJ, Wilber A, Bell J, Geurts A, Sivasubbu S, Wang X, Hackett PB, Largaespada DA, McIvor RS, Ekker SC. 2006. Harnessing a high cargo-capacity transposon for genetic applications in vertebrates. PLoS Genet 2(11):e169.

Barnes LM, Dickson AJ. 2006. Mammalian cell factories for efficient and stable protein expression. Curr Opin Biotechnol 17(4):381-6.

Barron N, Piskareva O, Muniyappa M. 2007. Targeted genetic modification of cell lines for recombinant protein production. Cytotechnology 53(1-3):65-73.

Bebbington CR, Renner G, Thomson S, King D, Abrams D, Yarranton GT. 1992. High level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker. Bio/Technology 10(2):169–175.

Benton T, Chen T, McEntee M, Fox B, King D, Crombie R, Thomas T, 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(1-3):43-46.

Berger I, Garzoni F, Chaillet M, Haffke M, Gupta K, Aubert A. 2013. The multiBac protein complex production platform at the EMBL. J Vis Exp(77):e50159.

Bire S, Gosset D, Jegot G, Midoux P, Pichon C, Rouleux-Bonnin F. 2013a. Exogenous mRNA delivery and bioavailability in gene transfer mediated by piggyBac transposition. BMC Biotechnol 13:75.

Bire S, Ley D, Casteret S, Mermod N, Bigot Y, Rouleux-Bonnin F. 2013b. Optimization of the piggyBac transposon using mRNA and insulators: toward a more reliable gene delivery system. PLoS One 8(12):e82559.

Browne SM, Al-Rubeai M. 2007. Selection methods for high-producing mammalian cell lines. Trends Biotechnol 25(9):425-32.

Butler M, Spearman M. 2014. The choice of mammalian cell host and possibilities for glycosylation engineering. Curr Opin Biotechnol 30C:107-112.

Cacquevel M, Aeschbach L, Osenkowski P, Li D, Ye W, Wolfe MS, Li H, Selkoe DJ, Fraering PC. 2008. Rapid purification of active gamma-secretase, an intramembrane protease implicated in Alzheimer's disease. J Neurochem 104(1):210-20.

Cadinanos J, Bradley A. 2007. Generation of an inducible and optimized piggyBac transposon system. Nucleic Acids Res 35(12):e87.

Cain K, Peters S, Hailu H, Sweeney B, Stephens P, Heads J, Sarkar K, Ventom A, Page C, Dickson A. 2013. A CHO cell line engineered to express XBP1 and ERO1-Lalpha has increased levels of transient protein expression. Biotechnol Prog 29(3):697-706.

Carter PJ. 2011. Introduction to current and future protein therapeutics: a protein engineering perspective. Exp Cell Res 317(9):1261-9.

Cheng JK, Alper HS. 2014. The genome editing toolbox: a spectrum of approaches for targeted modification. Curr Opin Biotechnol 30C:87-94.

Chung JY, Lim SW, Hong YJ, Hwang SO, Lee GM. 2004. Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary cells. Biotechnol Bioeng 85(5):539-46.

Chusainow J, Yang YS, Yeo JH, Toh PC, Asvadi P, Wong NS, Yap MG. 2009. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? Biotechnol Bioeng 102(4):1182-96.

Crawford Y, Zhou M, Hu Z, Joly J, Snedecor B, Shen A, Gao A. 2013. Fast identification of reliable hosts for targeted cell line development from a limited-genome screening using combined phiC31 integrase and CRE-Lox technologies. Biotechnol Prog 29(5):1307-15.

Daramola O, Stevenson J, Dean G, Hatton D, Pettman G, Holmes W, Field R. 2014. A high-yielding CHO transient system: coexpression of genes encoding EBNA-1 and GS enhances transient protein expression. Biotechnol Prog 30(1):132-41.

Darken MA. 1964. Puromycin inhibition of protein synthesis. Pharmacol Rev 16:223-43.

Dean A. 2006. On a chromosome far, far away: LCRs and gene expression. Trends Genet 22(1):38-45.

DeMaria CT, Cairns V, Schwarz C, Zhang J, Guerin M, Zuena E, Estes S, Karey KP. 2007. Accelerated clone selection for recombinant CHO CELLS using a FACS-based high-throughput screen. Biotechnol Prog 23(2):465-72.

Derouazi M, Martinet D, Besuchet Schmutz N, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM. 2006. Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. Biochem Biophys Res Commun 340(4):1069-77.

Dharshanan S, Chong H, Cheah SH, Zamrod Z. 2014. Stable expression of H1C2 monoclonal antibody in NS0 and CHO cells using pFUSE and UCOE expression system. Cytotechnology 66(4):625-33.

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(3):473-83.

Donnelly ML, Luke G, Mehrotra A, Li X, Hughes LE, Gani D, Ryan MD. 2001. Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J Gen Virol 82(Pt 5):1013-25.

Drocourt D, Calmels T, Reynes JP, Baron M, Tiraby G. 1990. Cassettes of the Streptoalloteichus hindustanus ble gene for transformation of lower and higher eukaryotes to phleomycin resistance. Nucleic Acids Res 18(13):4009.

Eder J, Sedrani R, Wiesmann C. 2014. The discovery of first-in-class drugs: origins and evolution. Nat Rev Drug Discov 13(8):577-87.

Eissenberg JC, Morris GD, Reuter G, Hartnett T. 1992. The heterochromatin-associated protein Hp-1 is an essential protein in drosophila with dosage-dependent effects on position-effect variegation. Genetics 131(2):345-352.

Fan L, Kadura I, Krebs LE, Hatfield CC, Shaw MM, Frye CC. 2012. Improving the efficiency of CHO cell line generation using glutamine synthetase gene knockout cells. Biotechnol Bioeng 109(4):1007-15.

Fan L, Kadura I, Krebs LE, Larson JL, Bowden DM, Frye CC. 2013. Development of a highly-efficient CHO cell line generation system with engineered SV40E promoter. J Biotechnol 168(4):652-8.

Gaj T, Gersbach CA, Barbas CF, 3rd. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31(7):397-405.

Galbete JL, Buceta M, Mermod N. 2009. MAR elements regulate the probability of epigenetic switching between active and inactive gene expression. Mol Biosyst 5(2):143-50.

Galvan DL, Nakazawa Y, Kaja A, Kettlun C, Cooper LJ, Rooney CM, Wilson MH. 2009. Genome-wide mapping of PiggyBac transposon integrations in primary human T cells. J Immunother 32(8):837-44.

Gaszner M, Felsenfeld G. 2006. Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 7(9):703-13.

Geisse S, Voedisch B. 2012. Transient expression technologies: past, present, and future. Methods Mol Biol 899:203-19.

Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA, Hackett PB. 2003. Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol Ther 8(1):108-117.

Ghattas IR, Sanes JR, Majors JE. 1991. The encephalomyocarditis virus internal ribosome entry site allows efficient coexpression of two genes from a recombinant provirus in cultured cells and in embryos. Mol Cell Biol 11(12):5848-59.

Gonzalez A, Jimenez A, Vazquez D, Davies JE, Schindler D. 1978. Studies on the mode of action of hygromycin B, an inhibitor of translocation in eukaryotes. Biochim Biophys Acta 521(2):459-69.

Gorman C, Arope S, Grandjean M, Girod PA, Mermod N. 2009. Use of MAR elements to increase the production of recombinant proteins. Cell Line Development 6:1-32.

Grabundzija I, Irgang M, Mates L, Belay E, Matrai J, Gogol-Doring A, Kawakami K, Chen W, Ruiz P, Chuah MK, VandenDriessche T, Izsvak Z, Ivics Z. 2010. Comparative analysis of transposable element vector systems in human cells. Mol Ther 18(6):1200-9.

Grandjean M, Girod PA, Calabrese D, Kostyrko K, Wicht M, Yerly F, Mazza C, Beckmann JS, Martinet D, Mermod N. 2011. High-level transgene expression by homologous recombination-mediated gene transfer. Nucleic Acids Res 39(15):e104.

Hacker DL, Chenuet S, Wurm FM. 2010. Chinese Hamster Ovary cells, recombinant protein production. Encyclopedia of Industrial Biotechnology: John Wiley & Sons, Inc. p 1-8.

Hacker DL, De Jesus M, Wurm FM. 2009. 25 years of recombinant proteins from reactor-grown cells - where do we go from here? Biotechnol Adv 27(6):1023-7.

Hacker DL, Kiseljak D, Rajendra Y, Thurnheer S, Baldi L, Wurm FM. 2013. Polyethyleneimine-based transient gene expression processes for suspension-adapted HEK-293E and CHO-DG44 cells. Protein Expr Purif 92(1):67-76.

Hackett PB, Ekker SC, Largaespada DA, McIvor RS. 2005. Sleeping beauty transposon-mediated gene therapy for prolonged expression. Adv Genet 54:189-232.

Ho SC, Koh EY, van Beers M, Mueller M, Wan C, Teo G, Song Z, Tong YW, Bardor M, Yang Y. 2013. Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability. J Biotechnol 165(3-4):157-66.

Holmes P, Al-Rubeai M. 1999. Improved cell line development by a high throughput affinity capture surface display technique to select for high secretors. J Immunol Methods 230(1-2):141-7.

Huang C. 2009. Receptor-Fc fusion therapeutics, traps, and MIMETIBODY technology. Curr Opin Biotechnol 20(6):692-9.

Huang X, Guo H, Tammana S, Jung YC, Mellgren E, Bassi P, Cao Q, Tu ZJ, Kim YC, Ekker SC, Wu X, Wang SM, Zhou X. 2010. Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. Mol Ther 18(10):1803-13.

Hwang SO, Chung JY, Lee GM. 2003. Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. Biotechnol Prog 19(1):179-84.

Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91(4):501-10.

Ivics Z, Li M, Mátés L, Boeke J, Nagy A, Bradley A, Izsvák Z. 2009. Transposon-mediated genome manipulation in vertebrates. Nature Methods 6(6):415-22.

Izsvak Z, Ivics Z. 2004. Sleeping beauty transposition: biology and applications for molecular therapy. Mol Ther 9(2):147-56.

Jayapal KP, Wlaschin KF, Yap MGS, Hu W-S. 2007. Recombinant protein therapeutics from CHO cells - 20 years and counting. Chem. Eng. Prog. 103(10):40-47.

Kacmar J, Srienc F. 2005. Dynamics of single cell property distributions in Chinese hamster ovary cell cultures monitored and controlled with automated flow cytometry. J Biotechnol 120(4):410-20.

Kahlig KM, Saridey SK, Kaja A, Daniels MA, George AL, Jr., Wilson MH. 2009. Multiplexed transposon-mediated stable gene transfer in human cells. Proc Natl Acad Sci U S A 107(4):1343-8.

Kaufman RJ, Wasley LC, Spiliotes AJ, Gossels SD, Latt SA, Larsen GR, Kay RM. 1985. Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in Chinese hamster ovary cells. Mol Cell Biol 5(7):1750-9.

Kawakami K. 2007. Tol2: a versatile gene transfer vector in vertebrates. Genome Biol 8 Suppl 1:S7.

Kim JY, Kim YG, Lee GM. 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 93(3):917-30.

Kim M, O'Callaghan PM, Droms KA, James DC. 2011. A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. Biotechnol Bioeng 108:2434–2446.

Kling J. 2011. Fresh from the biologic pipeline--2010. Nat Biotechnol 29(3):197-200.

Knox R, Nettleship JE, Chang VT, Hui ZK, Santos AM, Rahman N, Ho LP, Owens RJ, Davis SJ. 2013. A streamlined implementation of the glutamine synthetase-based protein expression system. BMC Biotechnol 13:74.

Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H. 1996. Transposable element in fish. Nature 383(6595):30.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128(4):693-705.

Kwaks TH, Otte AP. 2006. Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. Trends Biotechnol 24(3):137-42.

Kwaks THJ, Barnett P, Hemrika W, Siersma T, Sewalt RGAB, Satijn DPE, Brons JF, van Blokland R, Kwakman P, Kruckeberg AL, Kelder A, Otte AP. 2003. Identification of anti-repressor elements that confer high and stable protein production in mammalian cells. Nat Biotech 21(5):553-558.

Lai T, Yang Y, Ng SK. 2013. Advances in mammalian cell line development technologies for recombinant protein production. Pharmaceuticals (Basel) 6(5):579-603.

Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO. 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. Nat Biotechnol 31(8):759-65.

Ley D, Harraghy N, Le Fourn V, Bire S, Girod PA, Regamey A, Rouleux-Bonnin F, Bigot Y, Mermod N. 2013. MAR elements and transposons for improved transgene integration and expression. PLoS One 8(4):e62784.

Ley D, Van Zwieten R, Puttini S, Iyer P, Cochard A, Mermod N. 2014. A PiggyBac-mediated approach for muscle gene transfer or cell therapy. Stem Cell Res 13(3 Pt A):390-403.

Li F, Vijayasankaran N, Shen AY, Kiss R, Amanullah A. 2010. Cell culture processes for monoclonal antibody production. MAbs 2(5):466-79.

Li J, Zhang C, Jostock T, Dubel S. 2007. Analysis of IgG heavy chain to light chain ratio with mutant Encephalomyocarditis virus internal ribosome entry site. Protein Eng Des Sel 20(10):491-6.

Li Z, Michael IP, Zhou D, Nagy A, Rini JM. 2013. Simple piggyBac transposon-based mammalian cell expression system for inducible protein production. Proc Natl Acad Sci U S A 110(13):5004-9.

Lohe AR, Hartl DL. 1996. Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Mol Biol Evol 13(4):549-55.

Majors BS, Betenbaugh MJ, Pederson NE, Chiang GG. 2008. Enhancement of transient gene expression and culture viability using Chinese hamster ovary cells overexpressing Bcl-x(L). Biotechnol Bioeng 101(3):567-78.

Mansouri M, Berger P. 2014. Strategies for multigene expression in eukaryotic cells. Plasmid 75:12-7.

Mastrangelo AJ, Hardwick JM, Zou S, Betenbaugh MJ. 2000. Part II. Overexpression of bcl-2 family members enhances survival of mammalian cells in response to various culture insults. Biotechnol Bioeng 67(5):555-64.

Matasci M, Bachmann V, Delegrange F, Chenuet S, Hacker D, Wurm F. 2012. Generation of High-Producing CHO Cell Lines by Piggybac Transposition. In: Jenkins N, Barron N, Alves P, editors. Proceedings of the 21st Annual Meeting of the European Society for Animal Cell Technology (ESACT), Dublin, Ireland, June 7-10, 2009: Springer Netherlands. p 129-133.

Matasci M, Baldi L, Hacker DL, Wurm FM. 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(9):2141-50.

Matasci M, Hacker DL, Baldi L, Wurm FM. 2008. Recombinant therapeutic protein production in cultivated mammalian cells: current status and future prospects. Drug Discov Today: Technol 5(2,3):e37-e42.

Mates L, Chuah MK, Belay E, Jerchow B, Manoj N, Acosta-Sanchez A, Grzela DP, Schmitt A, Becker K, Matrai J and others. 2009. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet 41(6):753-61.

Mates L, Izsvak Z, Ivics Z. 2007. Technology transfer from worms and flies to vertebrates: transposition-based genome manipulations and their future perspectives. Genome Biol 8 Suppl 1:S1.

McCormick F, Trahey M, Innis M, Dieckmann B, Ringold G. 1984. Inducible expression of amplified human beta interferon genes in CHO cells. Mol Cell Biol 4(1):166-72.

McPhaul M, Berg P. 1986. Formation of functional asialoglycoprotein receptor after transfection with cDNAs encoding the receptor proteins. Proc Natl Acad Sci U S A 83(23):8863-7.

Meir YJ, Lin A, Huang MF, Lin JR, Weirauch MT, Chou HC, Lin SJ, Wu SC. 2013. A versatile, highly efficient, and potentially safer piggyBac transposon system for mammalian genome manipulations. FASEB J 27(11):4429-43.

Meir YJ, Weirauch MT, Yang HS, Chung PC, Yu RK, Wu SC. 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.

Meissner P, Pick H, Kulangara A, Chatellard P, Friedrich K, Wurm FM. 2001. Transient gene expression: recombinant protein production with suspension-adapted HEK293-EBNA cells. Biotechnol Bioeng 75(2):197-203.

Mohan C, Kim YG, Koo J, Lee GM. 2008. Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells. Biotechnol J 3(5):624-30.

O'Callaghan PM, McLeod J, Pybus LP, Lovelady CS, Wilkinson SJ, Racher AJ, Porter A, James DC. 2010. Cell line-specific control of recombinant monoclonal antibody production by CHO cells. Biotechnol Bioeng 106(6):938-51.

Osterlehner A, Simmeth S, Gopfert U. 2011. Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. Biotechnol Bioeng 108(11):2670-81.

Paul P, Lepage V, Sayagh B, Metzger JJ, Pla M, Boumsell L, Douay C, Cohen D, Colombani J, Dausset J and others. 1985. Serological expression after sequential double transfection with purified HLA-A11 gene of mouse fibroblasts carrying human beta-2 microglobulin. Immunogenetics 22(1):1-8.

Pham PL, Kamen A, Durocher Y. 2006. Large-scale transfection of mammalian cells for the fast production of recombinant protein. Mol Biotechnol 34(2):225-37.

Pilbrough W, Munro TP, Gray P. 2009. Intraclonal protein expression heterogeneity in recombinant CHO cells. PLoS One 4(12):e8432.

Puck TT. 1957. The genetics of somatic mammalian cells. Adv Biol Med Phys 5:75-101.

Rajendra Y, Kiseljak D, Baldi L, Hacker DL, Wurm FM. 2011. A simple high-yielding process for transient gene expression in CHO cells. J Biotechnol 153(1-2):22-6.

Rajendra Y, Kiseljak D, Manoli S, Baldi L, Hacker DL, Wurm FM. 2012. Role of non-specific DNA in reducing coding DNA requirement for transient gene expression with CHO and HEK-293E cells. Biotechnol Bioeng 109(9):2271-8.

Sadowski PD. 1995. The Flp recombinase of the 2-microns plasmid of Saccharomyces cerevisiae. Prog Nucleic Acid Res Mol Biol 51:53-91.

Sauer B, Henderson N. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc Natl Acad Sci U S A 85(14):5166-70.

Scahill SJ, Devos R, Van der Heyden J, Fiers W. 1983. Expression and characterization of the product of a human immune interferon cDNA gene in Chinese hamster ovary cells. Proc Natl Acad Sci U S A 80(15):4654-8.

Schlatter S, Rimann M, Kelm J, Fussenegger M. 2002. SAMY, a novel mammalian reporter gene derived from Bacillus stearothermophilus alpha-amylase. Gene 282(1-2):19-31.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3(6):1101-8.

Sharma N, Hollensen AK, Bak RO, Staunstrup NH, Schroder LD, Mikkelsen JG. 2012. The impact of cHS4 insulators on DNA transposon vector mobilization and silencing in retinal pigment epithelium cells. PLoS One 7(10):e48421.

Sorrell DA, Kolb AF. 2005. Targeted modification of mammalian genomes. Biotechnol Adv 23(7-8):431-69.

Sternberg N. 1981. Bacteriophage P1 site-specific recombination. III. Strand exchange during recombination at lox sites. J Mol Biol 150(4):603-8.

Thompson LH, Baker RM. 1973. Isolation of mutants of cultured mammalian cells. Methods Cell Biol 6:209-81.

Urlaub G, Chasin LA. 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. Proc Natl Acad Sci U S A 77(7):4216-20.

Urlaub G, Kas E, Carothers AM, Chasin LA. 1983. Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. Cell 33(2):405-12.

Van Blokland HJ, Hoeksema F, Siep M, Otte AP, Verhees JA. 2011. Methods to create a stringent selection system for mammalian cell lines. Cytotechnology 63(4):371-84.

Van Blokland HJ, Kwaks TH, Sewalt RG, Verhees JA, Klaren VN, Siersma TK, Korse JW, Teunissen NC, Botschuijver S, van Mer C and others. 2007. A novel, high stringency selection system allows screening of few clones for high protein expression. J Biotechnol 128(2):237-45.

Walsh G. 2005. Therapeutic insulins and their large-scale manufacture. Appl Microbiol Biotechnol 67(2):151-9.

Walsh G. 2010. Biopharmaceutical benchmarks 2010. Nat Biotechnol 28(9):917-24.

Walsh G. 2014. Biopharmaceutical benchmarks 2014. Nat Biotechnol 32(10):992-1000.

Weidle UH, Buckel P, Wienberg J. 1988. Amplified expression constructs for human tissue-type plasminogen activator in Chinese hamster ovary cells: instability in the absence of selective pressure. Gene 66(2):193-203.

Whitelaw E, Sutherland H, Kearns M, Morgan H, Weaving L, Garrick D. 2001. Epigenetic effects on transgene expression. Methods Mol Biol 158:351-68.

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/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol 5:17.

Wilson C, Bellen HJ, Gehring WJ. 1990. Position effects on eukaryotic gene expression. Annu Rev Cell Biol 6:679-714.

Wilson MH, Coates CJ, George AL, Jr. 2007. PiggyBac transposon-mediated gene transfer in human cells. Mol Ther 15(1):139-45.

Wu SC, Meir YJ, Coates CJ, Handler AM, Pelczar P, Moisyadi S, Kaminski JM. 2006. piggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. Proc Natl Acad Sci U S A 103(41):15008-13.

Wurm FM. 1994. Enhancement of expression by gene targeting in endogenous retrovirus-like sequences. Google Patents.

Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. Nat Biotechnol 22(11):1393-8.

Wurm FM. 2013. CHO-Quasispecies-Implications for manufacturing processes. Processes 1(3):296-311.

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat Biotechnol 29(8):735-41.

Yang Y, Mariati, Chusainow J, Yap MG. 2010. DNA methylation contributes to loss in productivity of monoclonal antibody-producing CHO cell lines. J Biotechnol 147(3-4):180-5.

Ye J, Alvin K, Latif H, Hsu A, Parikh V, Whitmer T, Tellers M, de la Cruz Edmonds MC, Ly J, Salmon P, Markusen JF. 2010. Rapid protein production using CHO stable transfection pools. Biotechnol Prog 26(5):1431-7.

Yusa K, Zhou L, Li MA, Bradley A, Craig NL. 2011. A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A 108(4):1531-6.

Zhou H, Liu ZG, Sun ZW, Huang Y, Yu WY. 2010. Generation of stable cell lines by site-specific integration of transgenes into engineered Chinese hamster ovary strains using an FLP-FRT system. J Biotechnol 147(2):122-9.

Zhu J. 2012. Mammalian cell protein expression for biopharmaceutical production. Biotechnol Adv 30(5):1158-70.

Sowmya Balasubramanian
balasubramanian.sowmya@epfl.ch +41786644820

Summary

- Motivated and independent researcher with about 4.5 years of experience in suspension adapted mammalian cell culture
- Scientific expertise in transposition mediated stable CHO cell pool and clone generation
- Experienced with working in multi-cultural teams both as a leader and a team member
- Successfully planned and supervised multiple master thesis projects and semester projects

Education and Research Experience

École Polytechnique Fédérale de Lausanne, Switzerland (EPFL)
Doctoral Studies, Laboratory of Cellular Biotechnology (LBTC) Oct 2011 – Present

- Development of rapid and high titer process for generation of stable CHO pools and clones by piggyBac transposition for recombinant protein production
- Comprehensive analysis of piggyBac, sleeping beauty and tol2 transposases in suspension adapted CHO DG44, CHO-K1 and CHO-S for the generation of stable pools
- Investigation of single step transposition system for generation of pools and clones expressing multiple proteins simultaneously

Master Thesis, Laboratory of Cellular Biotechnology (LBTC) July 2010 – May 2011

- Optimization of experimental conditions for generation of recombinant pools in CHO cells using piggyBac transposon system

University of Cologne, Germany
Research Intern, Institute for Genetics May 2009 – July 2009

- Successfully cloned and purified both Irga8 and Irgb10 of the IRG family which have been documented as participating in resistance to T. Gondii

Birla Institute of Technology and Science, India (BITS)
Integrated course Jan 2006 – May 2011
MSc (Hons). Biological Sciences and BE (Hons) in Computer Science CGPA: 8.2/10

- Ranked first in the department of Biological Sciences (30 students)

Scientific Project – Microbial Fuel Cell (MFC) Jan 2008 – May 2009

- Designed a low cost fuel cell based on sewage water and water plants with 10% improved efficiency.

Scientific Project – Computational genomics of skin cancer Aug 2008 – Dec 2008

- Proposed a mathematical 3-parameter model to characterize the coding sequence of the *myc* gene and its mutants into a numerical values to develop a computational prediction for gene therapy

Sowmya Balasubramanian
balasubramanian.sowmya@epfl.ch +41786644820

BiologicalE, India
Research Intern May 2008 – July 2008

- Analyzed the various quality control techniques used for the production of several vaccines including Hepatitis B vaccine.

Awards and Scholarships

- Secured scholarship to present scientific posters at the Cell culture engineering XIV conference, 2014, PEACE conference, 2013 and ESACT conference, 2013
- Awarded the prestigious DAAD-WISE scholarship for undergraduate students, 2009
- National runner up in The Al Gore Sustainable Technology Venture Competition at Indian Institute of Technology
- Presented a working prototype of Microbial Fuel Cell to Dr. A P J Abdul Kalam (ex-President of India) as one of the top 10 projects of BITS-Pilani

Manuscripts Published/In preparation

- Alattia Jean-René, Mattia Matasci, Mitko Dimitrov, Lorène Aeschbach, Sowmya Balasubramanian, David Hacker, Florian Wurm, and Patrick Fraering. 2013. **Highly efficient production of the Alzheimer's γ-Secretase integral membrane protease complex by a multi-gene stable integration approach.** Biotechnol Bioeng. 110,7:1995-2005.
- Sowmya Balasubramanian, Mattia Matasci, Zuzana Kadlecova, Lucia Baldi, David Hacker, Florian Wurm. **Rapid recombinant protein production from piggyBac transposon-mediated stable CHO cell pools**.
- Sowmya Balasubramanian, Yashas Rajendra, Lucia Baldi, David Hacker, Florian Wurm. **Comparison of three transposons for the generation of highly productive recombinant CHO cell pools and cell lines.** (In preparation)
- Sowmya Balasubramanian, Lucia Baldi, David Hacker, Florian Wurm. **Cell pools expressing multiple transgenes using the piggyBac transposon for gene delivery.** (In preparation)
- Sowmya Balasubramanian, Antonija Sakic, Lucia Baldi, David Hacker, Florian Wurm. **Study of PiggyBac, Tol2 and Sleeping Beauty transposons for cell pool generation in different CHO cell strains.** (In preparation)

Conference Posters

- POSTER - Comparison of PiggyBac mediated cell pool generation with different CHO host systems – CCE XIV 2014, Quebec City, Canada
- POSTER - Synthetic transposons for the generation of stable CHO cell pools expressing recombinant proteins – PEACE 2013, Kananaskis, Canada
- POSTER - PiggyBac transposition for the generation of stable CHO-DG44 cell pools expressing multiple transgenes – ESACT 2013, Lille, France
- POSTER - PiggyBac transposon-mediated integration of transgenes in CHO-DG44 cells – ESACT 2011, Vienna, Austria

Sowmya Balasubramanian
balasubramanian.sowmya@epfl.ch +41786644820

Laboratory Skills

- Mammalian cell culture, Transfection, Molecular Cloning, Quantitative RT-PCR, ELISA, DNA and protein electrophoresis, GUAVA and BDLSR-II Flow cytometry, Western Blotting

Supplementary Activities

- Teaching Assistant for the EPFL team at the International Genetically Engineered Machine competition (iGEM), 2012
- Successfully trained 5 master students for semester projects or master thesis at EPFL
- Member of Abhigyan: an education for all initiative of BITS (India) – Taught English, Hindi and Mathematics to 3 illiterate people for over 3 semesters
- Member of core organizing committee for Quark, 2010 – Technical festival of BITS, Goa
- Trained professionally for 4 years in Indian classical music
- Fluent in English, Hindi, Tamil, Telugu and elementary proficiency in French, Malayalam, Marathi