Master Thesis

Establishing an assay for studying the molecular mechanisms of drug resistance in cancer

Submitted by Magdalena Teufl
Vienna, February 2017

University of Natural Resources and Life Sciences, Vienna
Department of Chemistry
Division of Biochemistry
Supervised by Dr. Traxlmayr and Prof. Obinger

Kurzfassung

Der epidermale Wachstumsfaktorrezeptor EGFR gehört zur Familie der Rezeptor-Tyrosinkinasen und ist ein Membranprotein auf Zellen von Wirbeltieren. Der EGF-Rezeptor ist ein Mitglied der ERBB Familie und wird in verschiedenen Tumorarten, wie Darmkrebs, Lungenkrebs oder Bauchspeicheldrüsenkrebs, hochreguliert oder mutiert und führt daher zu einem verstärkten Proliferationssignal und Zellwachstum.

EGF, ein Ligand des EGF-Rezeptors, führt bei der Bindung zu einer Konformationsänderung und somit zur Aktivierung des Rezeptors. Daraufhin werden weitere Signalwege, wie der PI3K/AKT/mTOR oder RAS/RAF/MEK/ERK Signalweg, aktiviert. Dies führt zu verschiedensten Reaktionen in der Zelle, wie zum Beispiel Zellwachstum, Migration und Resistenz gegen programmierten Zelltod (Apoptose). Sind diese Signalwege konstitutiv aktiviert, und das ist insbesondere der Fall wenn der Rezeptor entweder überexprimiert oder auch mutiert ist, kann dies zu einem unnatürlichen Zellwachstum und der Entstehung von Krebs führen.

In der Klinik werden zur Behandlung von verschiedenen Krebsarten derzeit gezielt EGFR-spezifische Inhibitoren eingesetzt. Dazu zählen monoklonale anti-EGFR Antikörper wie Cetuximab als auch Tyrosinkinase-Inhibitoren (TKI). Mittlerweile gibt es mehrere Generationen an Tyrosinkinase-Inhibitoren, da die anfangs erfolgreiche Behandlung mit TKIs der ersten Generation nach einiger Zeit, bedingt z.B. durch die somatische Resistenzmutation T790M, wirkungslos wurde. Das Entstehen von Resistenzmechanismen zählt mitunter zu den größten Problemen in der Behandlung und Bekämpfung von Krebs. Daher wird vor allem die Kombinationstherapie mit Medikamenten, welche verschiedene Resistenzmechanismen hervorrufen, als erfolgsversprechend angesehen. Nur wenige Modelle existieren, die das Erforschen und Vorhersagen von Resistenzmechanismen ermöglichen.

In Zuge dieser Arbeit wurde an der Entwicklung eines *in-vitro* Testsystems gearbeitet. Eine Library, bestehend aus EGFR Mutanten, wurde mittels error-prone PCR hergestellt und zur Transfektion von HEK293T Zellen verwendet. Die Transfektion der Zellen wurde optimiert um möglichst nur eine EGFR Mutante pro Zelle zu erhalten. Episomale Replikation des Plasmids wurde durch den SV40 Origin ermöglicht, welcher ein Teil des Plasmids war, und die Anzahl an Plasmiden pro Zelle wurde mittels qPCR bestimmt. Die transfizierte Zellpopulation wurde mittels Durchflusszytometrie auf die Phosphorylierung des EGF-Rezeptors untersucht, nachdem ein gewisser Selektionsdruck angelegt worden war (zum Beispiel ein TKI). Da das Phosphorylierungssignal intrazellulär detektiert werden muss, wurden die Zellen davor fixiert und permeabilisiert. Nur die EGFR Mutanten, welche Resistenzmutationen aufwiesen, zeigten in der Anwesenheit des Inhibitors ein Phosphorylierungssignal und wurden selektiert und zur Isolation der DNA herangezogen.

Im Grunde entspricht dieser Assay einem surface display und ermöglicht die Selektion eines Phänotyps und die Kopplung zu dem entsprechenden Genotyp. Wenn eine Selektion durchgeführt wird, kann die intakte DNA mittels PCR vervielfältigt und zur erneuten Transfektion der Zellen verwendet werden um die resistenten Zellen anzureichern. Schlussendlich kann die DNA isoliert und auf Resistenzmutationen untersucht werden.

Abstract

The epidermal growth factor receptor EGFR is a receptor tyrosine kinase found on the cell surface of vertebrates. EGFR belongs to the ERBB family of transmembrane receptor tyrosine kinases, members of which are known to be associated with cancer. Amplification or mutation of the EGFR gene are known to be involved in the emergence and maintenance of various solid tumors, such as cancers of the colon, head and neck, lung and pancreas.

Upon binding of a ligand, such as EGF, the receptor undergoes a conformational change, which prompts the activation of a variety of signalling pathways, such as the PI3K/AKT/mTOR and the RAS/RAF/MEK/ERK pathway, ultimately leading to cellular growth, migration and survival. When the receptor is overexpressed or mutated, aberrant activation of cellular signalling pathways can lead to cancerous cell growth.

Currently, anti-EGFR monoclonal antibodies like cetuximab as well as tyrosine kinase inhibitors (TKIs) are used in order to halt tumor progression. Though, after several months of treatment the emergence of the resistance mutation T790M towards first generation TKIs was observed in many patients, thus representing a major challenge concerning the use of anti-EGFR kinase inhibitors. By now, next generation TKIs have been developed. However, the incorporation of yet another mutation conferring resistance to those new generation inhibitors has already been detected in cancer patients. Therefore, the development of non-cross-resistant combination therapy is crucial in order to stop or at least delay the onset of resistance. Only few models exist, which allow the prediction of resistance mutations emerging from the treatment with novel (combinations of) anti-EGFR inhibitors.

Therefore, this thesis aimed at the development of an *in-vitro* assay, which would allow the investigation and prognosis of resistance mutations upon EGFR-targeted therapy in cancer. As a model system the HEK293T cell line was used. An EGFR library, which had been randomly mutated using error-prone PCR, was used for the transfection of cells. In order to express only one EGFR mutant per cell, the transfection procedure was optimized to yield only one plasmid per cell. Episomal replication was enabled through the SV40 origin located on the plasmid and the final plasmid number per cell was determined using qPCR. The transfected cell population, expressing mutant EGFR on the cell surface, was subjected to a selection pressure (i.e. a TKI) and the resulting intracellular phosphorylation signal exhibited by EGFR was detected via flow cytometry using phospho-EGFR-specific monoclonal antibodies. Access to the cell interior was granted through prior fixation and permeabilization of cells. Only EGFR mutants harbouring resistance mutations showed a phosphorylation signal in the presence of an EGFR inhibitor and hence were selected and used for the isolation of functional DNA.

This assay resembles a surface display system, thus allows the selection of mutants conferring drug resistance and the coupling to the corresponding genotype. After selection the intact genomic material will be amplified using PCR and further subjected to a subsequent selection cycle in order to expand the population of cells showing a resistance phenotype. Ultimately, the DNA shall be analysed and the mutations conferring resistance can be determined.

Acknowledgements

First of all I want to thank Professor Obinger for the opportunity to do my master thesis in the Protein Biochemistry Group, which made me change my outlook on science once again.

Next, I want to thank the entire BioB working group for the positive and welcoming working atmosphere throughout the months I have spent at the laboratory, for the debates on Nutella and electronic sensors on trash barrels, the shared meal times and numerous coffee breaks. Michael Traxlmayr, who likes to be called a doctor, was a great supervisor, who supported and guided me and was always on hand with help and advice. Thank you! I also want to thank Elisabeth Lobner for her support and guidance in the laboratory and for always lending an open ear.

Further, I want to express my gratitude to my parents and siblings, who enabled me to pursue my studies and always stood by my side.

Last, I want to thank Florian, who is not only my mentor on word and formatting, but also my best friend and always offered me great support.

Abbreviations

EGFR	Epidermal growth factor receptor
EGF	Epidermal growth factor
HER2	Human epidermal growth factor receptor 2
NSCLC	Non-small cell lung cancer
FACS	Fluorescence-activated cell sorting
HEK	Human Embryonic Kidney cells
E. coli	*Escherichia coli*
DMEM	Dulbecco's Modified Eagle Medium
PBS	Phosphate-Buffered Saline
TBS	Tris-Buffered Saline
SV40	Simian Virus 40
RTK	Receptor tyrosine kinase
TKI	Tyrosine kinase inhibitor
ZBF	Zinc-based fixative
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
MFI	Median fluorescence intensity
PEI	Polyethylenimine
CV	Coefficient of Variation
qPCR	Quantitative PCR
Cq	Quantification cycle
EGFR-WT	Wild-type EGFR
EGFR-L858R	EGFR carrying the mutation L858R
EGFR-del19	EGFR carrying a deletion in exon 19
Wt-plasmid	EGFR-plasmid carrying an intact SV40 origin
SV40ko-plasmid	EGFR-plasmid lacking an intact SV40 origin ("SV40-knockout plasmid")

Table of Contents

1 Introduction

1.1 Epidermal growth factor receptor (EGFR)

1.1.1 ERBB receptor family

The epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase present in the cell membrane of vertebrates. EGFR belongs to the ERBB family of transmembrane receptor tyrosine kinases (RTKs), which consists of 4 members including the epidermal growth factor receptor (ERBB1), human epidermal growth factor receptor 2 (HER2 or ERBB2), HER3 (ERBB3) and HER4 (ERBB4). All family members contain an extracellular ligand-binding domain, a single membrane-spanning segment, a juxtamembrane domain, a cytoplasmic tyrosine kinase domain and a C-terminal tail with phosphorylation sites[1]. Binding of a ligand to the extracellular domain of EGFR, HER3 or HER4 leads to receptor activation and the subsequent formation of kinase-active dimers[2]. Both the formation of homo-dimers as well as heterodimeric ERBB combinations are described in the literature [3]. Due to their impaired ability to form active homo-dimers, HER2, which is present in a conformation which resembles a ligand-bound state, and HER3 principally signal through heterodimerization[2,4]. HER3 shows a weak kinase activity compared to the other ERBB receptors but can serve as a phosphotyrosine scaffold[2]. Activation of the ERBB receptors stimulates intracellular signalling pathways, such as RAS/RAF/MEK/ERK, PI3K/AKT/mTOR, Src kinases and STAT transcription factors and finally leads to cell proliferation and resistance to apoptosis[2,3]. These complex interactions between the ERBB family members and their numerous ligands were named the “ERBB network” with its corresponding network theory proposed by Yarden et al. in 2001. The theory proposes that the four receptors and their many ligands would create a “layered network that derives robustness from modularity and from denying autonomy to ERBB2 and ERBB3”[3].

1.1.2 EGFR ligands

Several ligands are known to activate EGFR, such as the epidermal growth factor (EGF), transforming growth factor α (TGF-α), amphiregulin (AR) and epigen (EPG). Another group of ligands is known to bind both EGFR and HER4, including betacellulin (BTC), HB-EGF and epiregulin (EPR)[2].

1.1.3 Structural mechanism of EGFR activation

EGFR consists of an extracellular module accounting for 618 amino acids (AA), a transmembrane helix of 27 AA in length, a Juxtamembrane segment containing 37 AA, a kinase domain with 276 AA length and a C-terminal tail spanning 229 residues[5].

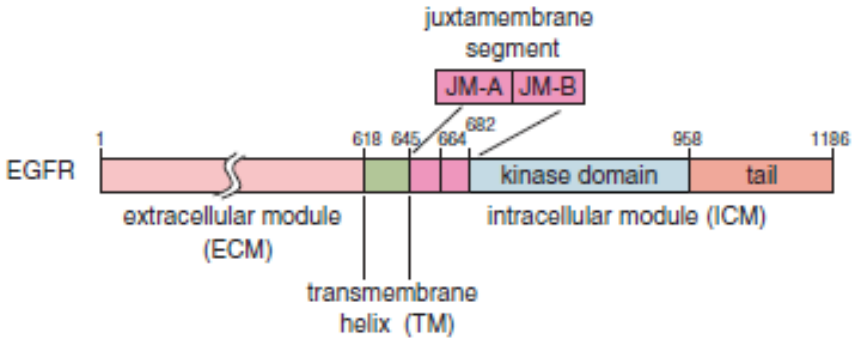


Figure 1: Domain architecture of EGFR consisting of the extracellular module (ECM), the transmembrane helix (TM), the Juxtamembrane segment, the kinase domain and the tail[5]. The numbering system used does not include the signal peptide containing 24 amino acids.

As there are two different residue numbering systems used in the literature (either including or not including the signal peptide), often leading to confusion and misunderstandings, the following table shows the amino acid position according to both numbering systems.

Table 1: EGFR AA position according to both numbering systems used (including and excluding the signal peptide sequence, respectively) of several known mutations and important tyrosine residues.

AA position (including the 24 bp signal sequence)	AA position (without the 24 bp signal sequence)	Description
L8585R	L834R	Activating single-point mutation in EGFR
T790M	T766M	Resistance mutation in the kinase domain towards erlotinib and gefitinib
Y869	Y845	Tyrosine residue in the activation loop of EGFR
Y1092	Y1068	Tyrosine residue on the C-terminal tail of EGFR
Y1016	Y992	Tyrosine residue on the C-terminal tail of EGFR
Y998	Y974	Tyrosine residue on the C-terminal tail of EGFR
Y1197	Y1173	Tyrosine residue on the C-terminal tail of EGFR

The C-terminal tail carries several tyrosine residues which become phosphorylated upon activation of EGFR and serve as docking sites for effector proteins engaged in the downstream signalling pathways. All except for one of the tyrosine residues involved in the signal transduction are located on the C-terminal tail. One tyrosine residue (Tyr869) is located within the activation loop of the kinase domain, which is believed to stabilize the active conformation of EGFR. The C-terminal tail in EGFR is not only known to be required for the recruitment of effector proteins upon phosphorylation of tyrosine residues but is also known to be involved in auto-inhibition of the receptor and critical for the activation loop phosphorylation[5].

Upon binding of a ligand a conformational change occurs including the transition of the un-liganded "tethered" conformation of the extracellular domain to an extended conformation forming "back-to-back" dimers (Figure 2). The ligand does not bridge the two receptors when this conformation is adapted. Instead, the dimer interface is solely formed by the receptors themselves. Apart from the activation by a ligand, enhanced expression, as in certain cancers, can also induce the formation of dimers or higher-order oligomers[4].

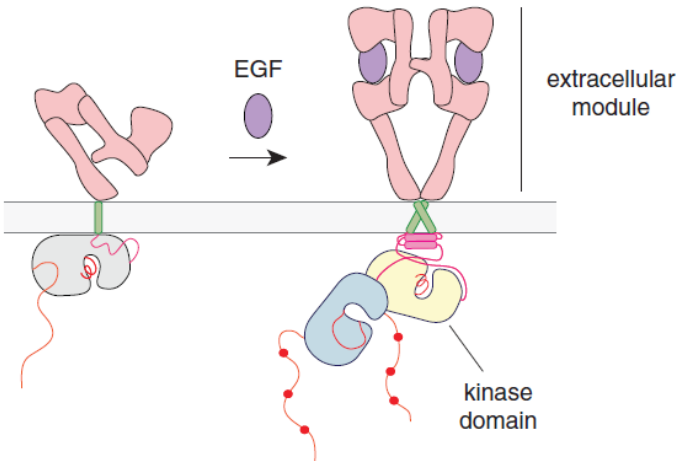


Figure 2: Conformational change from the tethered form adapted by the extracellular domain towards the extended conformation forming "back-to-back" dimers upon the addition of EGF [4].

During the activation, an asymmetric dimer of kinase domains is formed including one kinase termed the "activator" while the other kinase domain behaves as the "receiver". In fact, the activation of EGFR does not require the phosphorylation of the activation-loop according to Kovacs and co-workers but rather the formation of an "intact asymmetric dimerization interface" [4].

1.1.4 EGFR phosphorylation sites and their interaction

Apart from Tyr869, which is located in the activation loop, all other phosphorylation sites on EGFR become phosphorylated upon ligand-induced autophosphorylation[1]. Tyr869 in the

activation loop is believed to either be phosphorylated through a Src-kinase[6] or upon the formation of higher order oligomers, which is proposed to facilitate the otherwise sterically impaired phosphorylation of Tyr869[5].

After phosphorylation, the tyrosine residues on the C-terminal tail serve as docking sites for distinct intracellular signalling proteins[1]. A study conducted in 2006 incorporated the use of microarrays to evaluate the interaction between phosphorylated residues on the four ERBB receptors with virtually every Src homology 2 (SH2) and phospho-tyrosine binding (PTB) domain encoded in the human genome. Their studies revealed that EGFR shows two sites (Y998 and Y1016), which present many "high-affinity" interactions and thus may serve as "multifunctional docking sites". The other tyrosine residues were shown to only engage in significantly fewer interactions and are therefore more specific. The probably most important finding reveals that, when the affinity threshold for the binding is lowered, the ERBB3 network of interactions only changes very little while the binding interactions for both EGFR and ERBB2 become much more promiscuous. These low-affinity interactions become relevant when the levels of activated receptor increase due to overexpression. Therefore, the authors hypothesized that the oncogenic potential of ERBB2 and EGFR might be partially caused through the activation of alternative signalling pathways and suggested to tackle those "secondary pathways" for therapeutic reasons[7].

1.1.5 The role of EGFR in cancer

During the past years it became obvious that ERBB family members are involved in the initiation and maintenance of many solid tumors. During the 1980s numerous reports were published and proposed the impact of EGFR on the pathology of various cancers including cancers of the colon, head and neck, lung and pancreas[1]. Over the course of only 4 years (2003-2006) four EGFR inhibitors including cetuximab (monoclonal antibody), panitumumab (monoclonal antibody), gefitinib (tyrosine kinase inhibitor) and erlotinib (tyrosine kinase inhibitor) were approved by the FDA, representing the importance of targeting EGFR in cancer[1].

1.1.6 Aberrant occurrence of EGFR

Both the amplification and overexpression of the EGFR gene as well as the occurrence of EGFR mutants were reported in cancer. In fact, EGFR is overexpressed or mutated in more than 30% of breast cancers, 60% of non-small cell lung cancers and 40% of glioblastomas[4].

Mutations found in the EGFR gene were stated not to be random but rather located in so-called hotspots[8]. Many (possibly oncogenic) mutations in the EGFR gene were reported and are listed in the Catalogue of Somatic Mutations in Cancer (COSMIC).

In 2004, two particularly well known somatic EGFR mutations leading to aberrant phosphorylation activity were identified in a subset of patients who suffered from non-small cell lung cancer (NSCLC) and showed a remarkable clinical response to small-molecule EGFR tyrosine kinase inhibitors (TKIs)[2]. Lynch et al. investigated the differences between the 10% of patients with NSCLC who showed a dramatic response towards treatment with the TKI gefitinib (1.1.7 Current therapeutic options targeting EGFR) and the remaining 90% of the patients. They analysed the genome of 9 patients with gefitinib-responsive lung cancer and found that eight out of nine patients carried an EGFR mutation within the tyrosine kinase domain of EGFR. In 4 out of 8 patients with mutant EGFR a deletion of four amino acids in exon 19 was found (Exon 19 deletion), while three tumors were shown to harbour an amino acid substitution in exon 21 (either L858R or L861Q). The mutations were found to reside around the ATP cleft of EGFR, where they flank amino acids, which were shown to mediate binding to 4-anilinoquinazoline (of which gefitinib is a derivative) by crystallographic studies. Lynch et al. therefore postulated that those mutations might result in repositioning of these residues, stabilizing not only their interaction with ATP (leading to enhanced phosphorylation and activation) but also with gefitinib, resulting in enhanced inhibition compared to wild-type EGFR (EGFR-WT)[9].

The correlation of the type of EGFR mutation and the actual clinical response towards EGFR-targeted therapy was further investigated in 2009 by Kancha et al., who characterized a panel of EGFR mutations, which were found in NSCLC patients. A previous study[10] had proposed that the mutational status of EGFR did not predict the responsiveness to EGFR-targeted TKIs and hence did not recommend the use of molecular analysis to predict treatment response. Kancha et al. therefore hypothesized that it is necessary to characterize individual mutations in EGFR instead of only investigating whether or not a mutation is present. In fact they found that not all the mutations, which had been detected in tumor samples, were kinase active and would therefore be unresponsive towards treatment with EGFR inhibitors. It was further suggested that these EGFR mutations may result from PCR artefacts due to the use of formalin-embedded tissue in certain cases. Overall, Kancha et al. concluded that not all reported EGFR mutations have pathophysiologic relevance and described the need for the functional characterization of every EGFR mutation detected in the clinic[11].

1.1.7 Current therapeutic options targeting EGFR

Currently, several therapeutics, including tyrosine kinase inhibitors and monoclonal antibodies, are available for targeting EGFR. EGFR mutants carrying activating mutations show great responses towards TKIs, while ligand-activated wild-type EGFR responds well to EGFR-neutralizing antibodies[2].

The first generation of EGFR TKIs blocking the ATP-binding pocket of the intracellular tyrosine kinase domain, gefitinib and erlotinib, were shown to deliver greater results in EGFR mutants, which contain activating mutations such as the L858R mutation or a deletion in exon 19[9]. However, patients suffering from NSCLC often developed a secondary mutation in the gatekeeper residue, T790M, as they became resistant to gefitinib or erlotinib[2]. Second-generation EGFR inhibitors, such as afatinib and dacomitinib are irreversible ATP competitors, which form covalent links with the Cys773 residue in EGFR. These drugs are capable of inhibiting the T790M mutation, but the concentrations required for this purpose showed an unfavourable therapeutic index and dose-limiting toxicities due to the inhibition of wild-type EGFR[2]. Third-generation EGFR inhibitors have recently been developed and were designed to be more potent against the T790M mutation, without inducing toxicity from inhibiting wild-type EGFR. Third-generation drugs AZD9291 and CO-1686 were developed clinically and showed great responses with minimal toxicity[2]. In phase I studies with either AZD9291 or CO-1686 a response rate (RR) of >50% was shown for patients suffering from NSCLC who had developed resistance towards gefitinb or erlotinib through the acquisition of the T790M mutation[12]. Non-surprisingly less skin toxicity was associated with the application of those TKIs, probably due to the mutant specific activity of this drug[12].

Cetuximab (Erbitux®) and panitumumab (Vectibix®) are two antibodies used in EGFR-targeted therapy. Cetuximab was originally developed by Bristol-Myers Squibb as a chimeric (murine-human) molecule and was first used in clinical trials mid 1990s in combination with chemotherapy or radiotherapy[3]. Cetuximab mainly functions by binding to the extracellular domain therefore excluding the biological ligands from binding and subsequently preventing activation of the receptor. This further accelerates and enhances the degradation and internalization of the receptor[1].

Another mechanism of inhibition of EGFR was discussed by Pines et al., 2010. HSP90, a molecular chaperone, facilitates proper folding of proteins in an ATP-dependent manner. Though, it is also known to stabilize viral kinases as well as mutated oncoproteins, such as Bcr-Abl. HSP90 inhibition was shown to reduce mutant EGFR levels and activity. Mutant EGFR,

due to its constitutively active kinase, is probably more dependent on the stabilizing interaction with HSP90 than wild-type EGFR[8].

1.1.8 Mechanisms of resistance towards EGFR-targeted therapies

Both acquired as well as de novo resistance presents a major challenge treating patients suffering from cancer. After an initial relief the rise of resistance towards ERBB-targeted therapies is a common problem in the clinic.

As already mentioned, one particular mutation is known to inhibit the therapeutic effect of the first-generation TKIs gefitinib and erlotinib. The "gatekeeper" mutation T790M shows a "bulky" methionine instead of a threonine, therefore excluding the inhibitors from the ATP binding cleft[1]. A recent study using in-vitro mutagenesis with N-ethyl-N-nitrosourea (ENU) has additionally revealed mutations which might emerge upon treatment with the third-generation TKIs and found that one particular mutant (C797S in the background of exon 19 deletion/T790M) confers resistance to all currently known inhibitors[12].

Apart from point mutations conferring resistance, other strategies are engaged by the tumor to escape the selection pressure exhibited by the TKIs. One important mechanism underlying both acquired as well as de novo resistance (against TKIs and neutralizing antibodies) involves the so-called "oncogenic shift"[1] or "bypass track"[2], a mechanism where other signalling pathways are engaged upon inhibition of one RTK. In fact, many cancers are resistant to single-agent ERBB inhibitors because one of the downstream pathways is constitutively activated by a mutation in one of the involved signalling proteins, resulting in independence from the RTK located upstream in the signalling cascade[2].

1.1.9 Novel therapeutic challenges regarding the treatment of EGFR

Due to the big number of resistance mechanisms and escape routes, there is great need for the development of new agents and the development of combination therapy with innovative dosing and scheduling[2].

One central challenge regarding the development of resistances towards EGFR-directed therapies is the emergence of the T790M gatekeeper mutation. Third-generation EGFR TKIs have shown to harbour tremendous potential in overcoming this resistance. Though, it has been shown that EGFR-mutant lung cancers escape the selection pressure through single agent therapy by incorporating yet another mutation or the activation of a bypass pathway. One

central strategy involves the potent inhibition of mutant EGFR while simultaneously suppressing accessory pathways through the supplementation of additional inhibitors.

An additional challenge was recently described by Hata and co-workers. Investigating de-novo and acquired resistance based on the emergence of the T790M mutation in patients suffering from NSCLC, these authors provided evidence that clinically relevant drug-resistant cancer cells can both pre-exist as well as evolve from what they called "drug-tolerant" cells. Hata et al. speculated that acquired mutations would arise from a pool of drug-tolerant cells escaping the initial therapy. The presence of drug-tolerant cells, which are predicted to acquire one after another mutation in order to alleviate the effect of TKIs, is one of the biggest challenges to face during anti-cancer therapy according to Hata et al.[13].

1.1.10 Laboratory models used to discover mechanisms of resistance

For the reasons discussed above, there is a vital need for the combination of distinct non-cross-resistant therapeutics. Therefore, the prediction of resistance mutations against a given inhibitor is crucial. Few models exist for predicting resistance mechanisms in drug therapy, such as the maintenance of cell lines and xenografts in the presence of a drug or the infection with shRNA libraries thus enabling the generation and identification of drug-resistant clones (Figure 3). Those laboratory models are complemented by biopsy programs, which try to identify the mutations and molecular mechanisms after resistance has already evolved in the clinic[2].

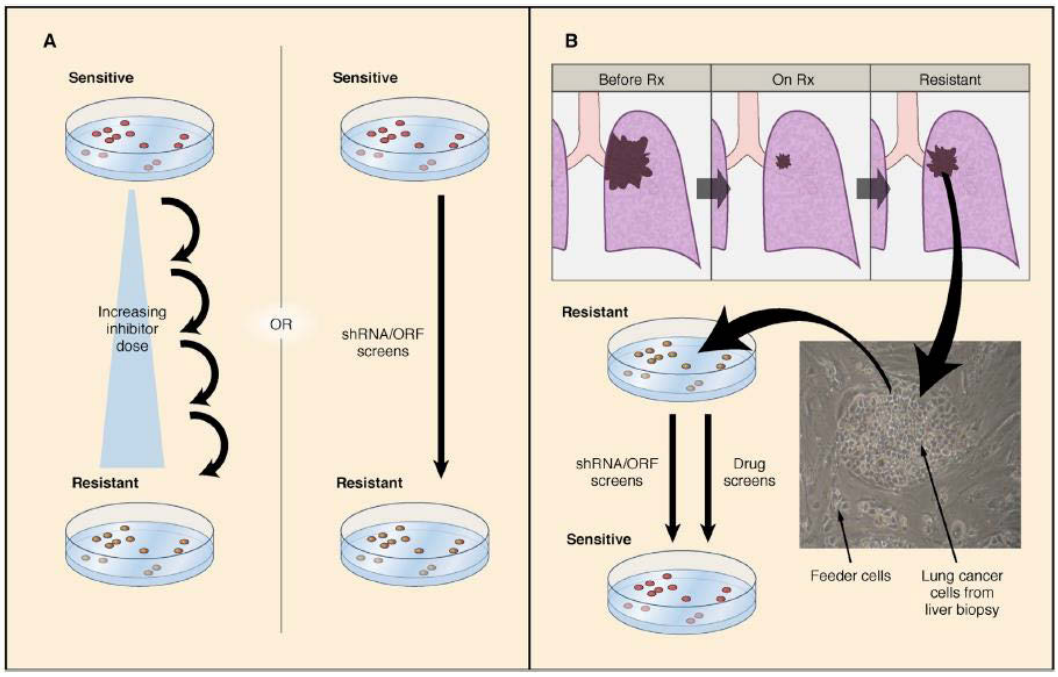


Figure 3: Laboratory models used for the investigation of resistance in cancer

Several drawbacks are associated with those cell line-based assays such as the readout, which is proliferation and resistance to apoptosis in the presence of an inhibitor. As cell proliferation is a result of a vast number of intracellular signalling pathways, it is also dependent on the signalling environment in the host cell, and might therefore not only depict the mutational state of the target gene. Another shortcoming associated with cell line-based assays is that many mutations might be missed, as a recent study had shown that the majority of mutations found in resistant clones after prolonged *in vitro* cultivation were pre-existing[14]. Moreover, none of these models presents a valuable tool for screening combinations of drugs with limited amount of time and money (months pass until the emergence of resistance in a cell line or xenograft).

Therefore, we set out to develop a novel method, which covers virtually all single-nucleotide mutations in the EGFR gene (containing 3630 bp), is fast and inexpensive and delivers a more defined outcome by directly detecting enzymatic function, minimizing influences from other signalling molecules.

1.2 Mammalian cell surface display

In 1985 the first form of surface display was demonstrated by George P. Smith, who used filamentous phage to display peptides. In the following years many more display systems were developed and used. One major drawback of phage display was the missing possibility to identify high affinity interactions, normalization for protein expression and visualization. Those problems could be tackled using yeast display and mammalian cell display[15].

In 2006 Ho et al. have presented a mammalian cell surface-displayed antibody system. They used HEK293T cells for display of functional single chain Fv (scFv) antibodies for affinity maturation. The transmembrane domain of the platelet derived growth factor receptor (PDGFR) was used as an anchor, the scFv and the murine Ig kappa chain signal peptide was expressed under the CMV promoter. The protein additionally contained a myc-tag between the scFv and the transmembrane domain to normalize for the level of protein expression. Using this surface display system, Ho and co-workers were able to show a 240-fold single-pass enrichment of a rare (high-affinity) scFv.

1.3 Transfection of mammalian cells

The transfection of mammalian cells discriminates between transient and stable transfection. Transient transfection refers to a finite change in DNA or protein expression and the

investigation of short-term effects only. Stable transfection requires the integration of the DNA into the host's genome and subsequently delivers long-lasting changes in the cell.

The choice of the transfection method varies with the goal of the transfection (transient vs. stable) and includes a variety of methods based on chemical, physical or biological principles. If stable integration into the genome is required transduction and microinjection are the most commonly used transfection methods. For selection purposes the introduced material normally carries a marker[16].

Transient transfection methods include both chemical as well as physical transfection methods such as the use of cationic lipids, cationic polymer, calcium phosphate, sonoporation, electroporation or laser-irradiation[16] .

One problem associated with the use of transient transfection is the short-term effect of DNA delivery. Apart from stable integration of DNA into the genome, discussed above, the use of episomal vector systems, which behave like additional chromosomes and replicate autonomously in the host cell, is known as an alternative approach to achieve persistence of recombinant DNA in the cell. Several episomal vector systems are known, based on viral plasmid replicons as well as non-viral episomal vectors. Viral replicons, such as the ones present in Simian Virus 40 (SV40) or Ebstein Barr Virus (EBV), consist of two elements: an origin of replication and a transactivating factor that can bind to this origin. Binding of the transactivation factor will initiate DNA replication in the cell. The SV40 plasmid replicon is characterized as a high copy number replicon. The SV40 origin of replication contains one binding site for the SV40 Large T Antigen, the only transactivating factor required. The large T Antigen is a helicase, which allows uncoupling of the plasmid replication from the genome, resulting in a high plasmid copy numbers (of several thousands) in a transfected cell[17].

1.4 Flow cytometry

1.4.1 Principle

Flow cytometry and fluorescence activated cell sorting (FACS) are widely used methods in research as well as biomedical diagnostics. In this technology particles, usually cells, are measured as they pass a laser beam in a fluid stream. The properties measured include particle size, granularity and relative fluorescence.

1.4.2 Detection of intracellular signals

In many cases it is necessary to detect the expression or activation of certain cellular components intracellularly. Intracellular staining usually requires the fixation of the cell or tissue in order to prevent autolysis and degradation of the cellular structure.

Based on their mode of action fixatives can be categorised into crosslinking and denaturing fixatives[18]. The alcohol-based fixatives (such as methanol, ethanol, Carnoy's, Methacarn) are denaturing fixatives. Their mode of action is based on protein denaturation through the removal of water from the free carboxyl, hydroxyl, amino, amido and imino groups of the protein. One major drawback using alcohol as a fixative is the effect of tissue shrinkage and protein coagulation observed with tissue fixation. However, this effect can be abrogated when acetic acid and chloroform are added to the fixative. Crosslinking fixatives on the other hand include glutaraldehyde and formaldehyde, which is by far the most widely used fixative today. Formaldehyde has been found to function through the formation of intra- and intermolecular crosslinks, with its major crosslinks formed between the side chain amino groups of lysine[18].

Although formaldehyde is known to stabilize the overall cell structure well, not to alter cell scatter properties as measured in flow cytometry and further prevent the clumping of cells, a phenomenon which is commonly observed if cells are directly exposed to alcohol[19], its effect on DNA and RNA quality is deleterious through the establishment of crosslinks with the surrounding histones and its detrimental effect on nucleotides, partly degrading nuclear material and altering sequences[18]. Because of this effect the use of "DNA-friendly" fixatives is of crucial importance if the purpose of the assay requires the isolation of intact, correct DNA. DNA-friendly fixatives commonly employ alcohol or acetone as a basis, although a small number of fixatives does not use denaturing liquids but are based on a mixture of Zinc salts, which are non-toxic, inexpensive and seem to preserve the morphology of the tissue comparably well to aldehydes[18,20].

In certain cases the use of formaldehyde is still required, for instance when a higher preservation of cellular structure for histological examinations is required[21]. Several attempts have been made in the past to optimize the size and quality of DNA, which can be extracted from formalin-fixed cellular material[21], though commonly a fragment size of several hundred bp is not exceeded[21,22]. In comparison, DNA read lengths of 2400 bp for Z7-fixed (a commonly used zinc-based fixative) tissue were reported[18].

Most nowadays commercially available fixation kits are based on formaldehyde or are of unknown composition. For the fixation of cells using DNA-friendly fixatives the use of ethanol[23] and methanol[19] have been described in the scientific literature. Zinc-based fixatives on the other hand (originally described by Beckstead 1994) arose interest in 2003 when Wester et al. published a zinc-based fixative (ZBF) which was shown to be superior to formaldehyde "for the analysis of DNA and protein expression"[25]. Much effort was made in order to further develop and optimize the composition of zinc-based fixatives[26], which were found to perform equally well to formaldehyde regarding the preservation of surface epitope labelling and scatter parameters, when flow cytometric analysis were performed[20]. Even the compatibility of a zinc-based fixation with the use of phosphospecific antibodies was proven successful and comparable to that of cells prepared using a "standard formaldehyde fixation procedure"[27].

2 Aim of this thesis

This study is conducted in order to establish an assay, which allows the prediction of virtually any drug resistance mutation within the EGFR gene, which will emerge upon treatment with a certain anti-EGFR inhibitor (i.e. TKI).

The setup of this assay includes the establishment of an EGFR library (encoding virtually every mutation within the EGFR gene) which will be prepared through the application of error-prone PCR (Figure 4, step 1). These mutagenized EGFR constructs will be ligated into an appropriate vector carrying the SV40 origin (Figure 4, step 2), thus facilitating episomal replication in cells expressing the large T antigen such as HEK293T cells, which were chosen as an expression system. Since the expression of one type of mutant EGFR on the cell surface is believed to be crucial for the success of this assay, the transfection procedure of HEK293T cells should be optimized in order to obtain only one plasmid per cell (Figure 4, step 3). Conducting this assay, cells will be subjected to a selection pressure (i.e. TKI) and only EGFR mutants harbouring resistance mutations will remain phosphorylated (Figure 4, step 4). Thus, the fixation and permeabilization procedure of HEK293T cells should be optimized in order to allow the detection of EGFR phosphorylation intracellularly without impairing the integrity of the DNA (Figure 4, step 5). Several antibodies specifically recognizing phosphorylated epitopes on EGFR will be tested to obtain a strong signal, which will allow the separation of phospho-EGFR-positive cells from the main population using FACS (Figure 4, step 6). Subsequently, those cells shall be used for the isolation of plasmid DNA (Figure 4, step 7). For this purpose it is crucial, that episomal replication is facilitated through the SV40 origin, as the diversity of EGFR mutants would probably be lost at this stage. The isolated plasmid DNA shall be amplified using PCR without introducing additional mutations (Figure 4, step 8) and the amplified EGFR constructs will either be used for another cycle of selection (ligated into an expression vector and used for the transfection of HEK293T cells) or used for the analysis of mutations via deep sequencing (Figure 4, step 9).

Making use of this assay, the emergence of EGFR mutations upon the addition of an EGFR inhibitor could be evaluated within several weeks only, screening for enzymatic activity instead of cell proliferation. Therefore, this assay harbours tremendous potential regarding the development of combination therapies targeting EGFR in order to avoid or at least delay the emergence of resistance mechanisms upon the treatment with anti-EGFR inhibitors. In addition, the flow cytometric detection system established in this study could possibly even facilitate

major progress in the field of personalized medicine, assessing the optimal combination of therapeutics for each individual patient with only limited amount of time and money.

In addition, the data obtained from this assay would enable a functional characterization of each EGFR mutant, which is currently often missing for reported EGFR mutations (discussed in 1.1.6) and could thus enable the construction of a landscape of single nucleotide mutations in EGFR conferring resistance to a given inhibitor.

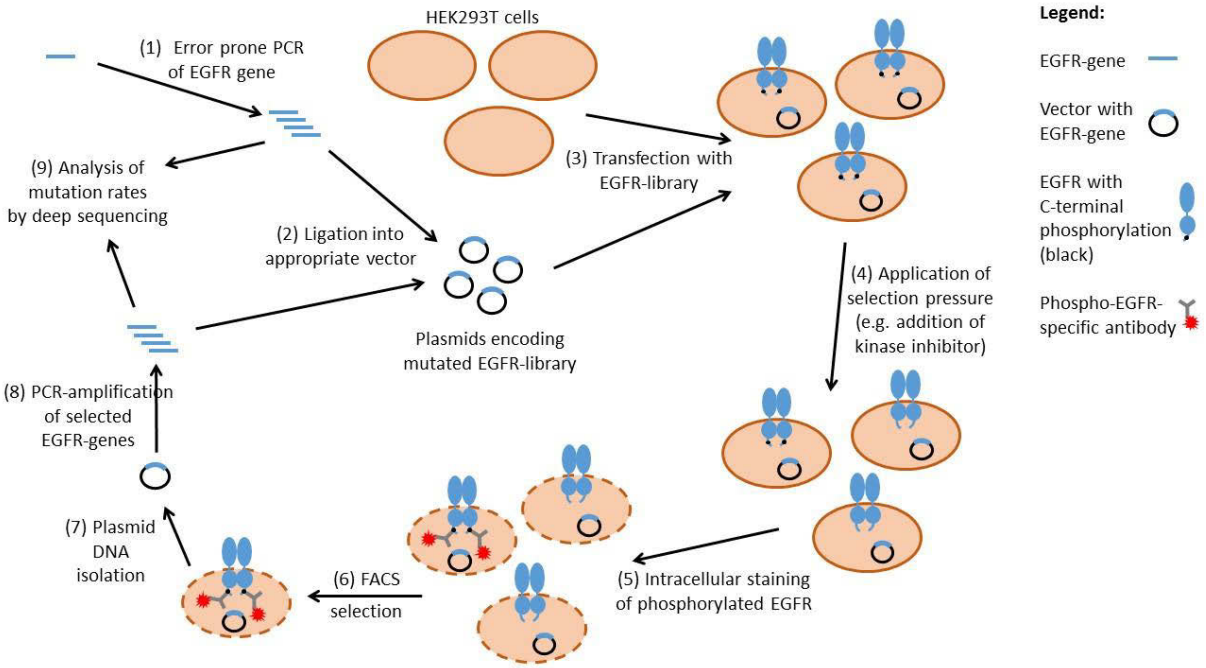


Figure 4: Scheme displaying the necessary steps for the establishment of the assay. © Michael Traxlmayr

3 Materials and Methods

3.1 Mammalian cell culture

3.1.1 Sterile working technique

In order to avoid contamination of the cell line as well as to facilitate protection of the personnel and the environment, all procedures were conducted in a laminar flow hood (Laminar SAFE 2020) using sterile working technique. The UV lamp was activated for 30 min before and after use of the instrument. In addition, the surface inside the hood as well as any of the materials and instruments entering the hood were disinfected with 70% ethanol. All media used for the work with the HEK293T cell line were either prepared in the laminar flood hood to ensure sterility or were prepared in an unsterile manner and autoclaved (121°C, 2 bar absolute, 20 min) before use.

3.1.2 Thawing of cell aliquots

The HEK293T cell line was a generous gift from Dr. Manfred Lehner from the Children's Cancer Research Institute, Vienna. The cell aliquot was thawed at room temperature. As soon as the cell aliquot had thawed 12 ml of complete growth medium (Dulbecco's modified Eagle growth medium + high glucose; Sigma-Aldrich), containing 100 U/ml Penicillin (Sigma-Aldrich), 100 µg/ml Streptomycin (Sigma-Aldrich) and 10% Fetal Bovine Serum (Gibco®, Thermo Fisher Scientific), was added. The cell suspension was centrifuged at 1000 rpm for 5 min using the MegaStar 1.6 R centrifuge (VWR) in order to pellet the cell material. After removing the supernatant, cells were resuspended in 15 ml complete growth medium and incubated in a T75 flask at 37°C and 5% CO_2 in a Heracell 150^{TM} incubator (Heraeus).

3.1.3 Sub-cultivation of HEK293T cells

The cells were passaged every 3 to 4 days in order to avoid outgrowth. The medium was removed from the flask and discarded, subsequently 5 ml PBS were used to wash the cells. In order to avoid the dissociation of the cells, PBS was applied carefully and the culture vessel was rocked back and forth with caution in order to distribute the PBS. After the PBS used for washing had been removed, 5 ml fresh PBS were applied directly to the cell lawn in order to

facilitate the dissociation of the cells from the vessel. Following an incubation step of 3-4 min the culture vessel was hit gently on the sides to ensure the complete dissociation of the cells from the vessel. The cell suspension was pipetted up and down to distribute the cell aggregates evenly. An aliquot (one tenth or one twentieth) of the cell suspension was transferred to a new cell culture flask (T75) holding 15 ml pre-warmed complete growth medium (37°C).

3.1.4 Cryopreservation

In order to prepare a working cell bank and secure cells of a low passage number for future use, HEK293T cells were slowly frozen at -80°C and subsequently transferred into the liquid nitrogen tank.

For this purpose, a T75 flask of 90% confluency was dissociated using PBS buffer and transferred to a CELLSTAR® Falcon tube (greiner bio-one). The cell suspension was spun down at 500 x g for 5 min using the Megastar 1.6R centrifuge (VWR). Afterwards, the supernatant was removed and complete growth medium, containing 5% dimethyl sulfoxide (DMSO, Sigma-Aldrich) in addition, was added to the cell pellet. This cell suspension was aliquoted to a number of cryogenic vials, yielding a number of approximately $3*10^6$ cells per vial. The vials were rapidly transferred to a freezing chamber filled with Isopropanol to enable slow freezing rates (approximately 1°C/minute) and avoid the formation of ice crystals. After 24 hours at -80°C the cryogenic vials were transferred to the nitrogen tank.

3.2 Cloning of the EGFR constructs

3.2.1 Media recipes

The LB-medium, LB-agar and SOC medium used for the routine cultivation of *E. coli* were prepared according to Table 2.

*Table 2: **Media used for the work with E. coli.** Media were prepared with dH_2O and subsequently autoclaved at 121°C at 2 bar absolute for 20 min*

Medium	Composition
LB-medium	1% (w/v) peptone from Casein (PanReac AppliChem) 0.5% (w/v) yeast extract (PanReac AppliChem) 1% (w/v) NaCl (Sigma-Aldrich)
LB-agar	1% (w/v) peptone (PanReac AppliChem) 0.5% (w/v) yeast extract (PanReac AppliChem) 1% (w/v) NaCl (Sigma-Aldrich) 1.5% (w/v) agarose (ROTH)
SOC	2% (w/v) peptone (PanReac AppliChem) 0.5% (w/v) yeast extract (PanReac AppliChem) 0.058% (w/v) NaCl (Sigma-Aldrich) 0.019% (w/v) KCl (Sigma-Aldrich) 0.2% (w/v) $MgCl_2*6H_2O$ (Sigma-Aldrich) 0.25%(w/v) $MgSO_4*7H_2O$ (Sigma-Aldrich) 0.4%(w/v) glucose monohydrate (Sigma-Aldrich)

3.2.2 pSF-CMV-SV40 vector

The pSF-CMV-SV40 vector was obtained from Oxford Genetics. The vector contains a Multiple Cloning Site (MCS) under the control of a cytomegalovirus (CMV) promoter. The bacterial selection marker is a kanamycin resistance gene and the pUC Origin represents the bacterial origin of replication. The SV40 Ori was thought to be of crucial importance in this context as it enables the episomal replication of the plasmid in mammalian cells expressing the SV40 large T antigen. The vector was received from Oxford Genetics at a concentration of 0.5 µg/µl and was used for the transformation of chemically competent *E. coli* (3.2.10) and subsequent cryopreservation (3.2.12).

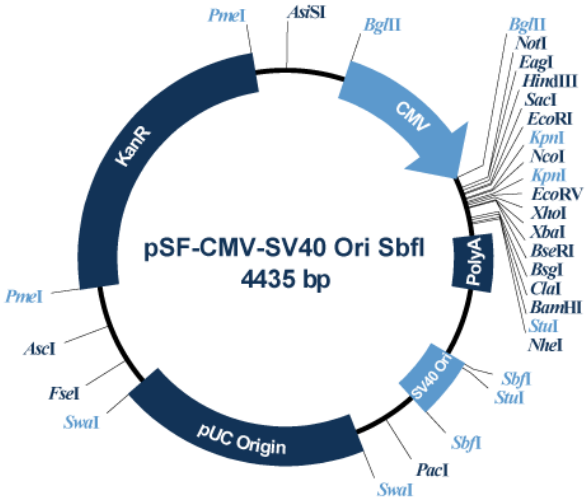


Figure 5: Plasmid map of the pSF-CMV-SV40 plasmid. This figure is made available by Oxford Genetics and shows the kanamycin resistance gene, the pUC origin, the MCS under the control of a CMV promoter and followed by a polyadenylation site. Last, the SV40 origin, flanked by two SbfI restriction sites is shown.

3.2.3 Primer design for cloning of the hEGFR gene

Oxford Genetics suggests the use of the restriction sites NcoI (C/CATGG) and XbaI (T/CTAGA) for cloning as both either contain a start or stop codon, respectively. Though, the restriction site NcoI could not be used for cloning of the hEGFR gene as the NcoI restriction site required a G as the first base of the second codon, which is not the case for the hEGFR gene (the second amino acid is an arginine).

Oxford Genetics further suggests the use of the nearest restriction site possible for cloning of the corresponding gene, but advised to keep the start and stop codon in the same positions (as NcoI and XbaI) in the MCS. The nearest possible restriction site for cloning of the hEGFR gene had been shown to be the KpnI restriction site (GGTAC/C). Therefore, the primer had to contain the same sequence exhibited between the KpnI restriction site and actual start codon (part of the NcoI restriction site).

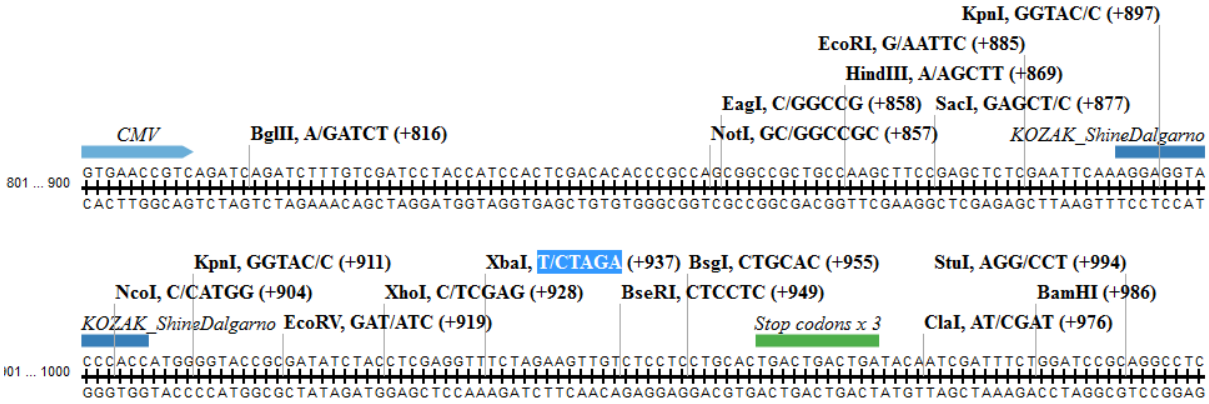


Figure 6: Sequence map of the pSF-CMV-SV40 plasmid showing the restriction sites of the MCS.

The reverse primer was designed to contain the XbaI restriction site T/CTAGA. Two different tags, the human influenza hemagglutinin (HA-) and the myc-tag, were attached to the 3' terminus via Polymerase Chain Reaction (PCR), separated from the hEGFR gene by a GGGGS (G_4S) linker. The myc-tag consists of 10 amino acids (EQKLISEEDL) and ends with a leucine (L), which is one of the amino acids feasible to be used within this restriction site. The amino acid sequence for the HA-tag is YPYDVDPYAL.

The melting temperature (T_m) as well as GC content (GC%) were checked for all primers using the T_m Calculator by New England Biolabs (NEB). The T_m for the annealing region of all primers was designed to exceed 72°C and therefore a 2-step PCR could be carried out avoiding unspecific priming at lower temperatures. Codon optimization with GeneArt® was used to create codons, which are optimized for the expression in human cells though avoiding the creation of the restriction sites KpnI and XbaI within the tag regions.

Table 3: ***Primers used for cloning of the hEGFR gene.*** *The forward primer hEGFR_fwd shows the restriction site KpnI (depicted in bold letters). The part of the primer which anneals to the hEGFR gene is coloured in blue and the nucleotide highlighted in grey depicts a nucleotide which was mutated to an A (G>A) in order to avoid the formation of a hairpin. The part of the primer preceding the KpnI site was added in order to facilitate binding of and successful digestion by the restriction enzyme. For both reverse primers hEGFRHA_rev and hEGFRmyc_rev (here depicted 5'-3') the restriction site (XbaI) is marked in bold letters, the tag (HA- or myc-) is highlighted in green, the G_4S linker is highlighted in purple and the part of the primer annealing to the hEGFR gene is depicted in blue.*

Primer name	Primer sequence
hEGFR_fwd	AATTCAAAGAA**GGTACC**CCACCATGCGACCCTCCGGGACGG
hEGFRHA_rev	AGGAGACAACT**TCTAGA**GGGCGTAATCTGGCACGTCGTAAGGGTAGCTT CCGCCGCCTCCTGCTCCAATAAATTCACTGCTTTGTGGCG
hEGFRmyc_rev	AGGAGACAACT**TCTAGA**GGTCCTCTTCGGAGATCAGCTTCTGCTCAGATC CTCCGCCTCCTGCTCCAATAAATTCACTGCTTTGTGGCG

3.2.4 Preparation of the hEGFR gene

The hEGFR wild-type (wt) gene was a generous gift from Matthew Meyerson (Addgene plasmid #11011) and was received as an *E. coli* colony on solid agar. 20 ml LB-medium (containing 100 µg/ml ampicillin) were inoculated with a pipette tip and grown on a shaker (GFL 3033) at 37°C and 180 rpm for 6 hours. The bacterial suspension was used for cryopreservation (3.2.12.).

1.5 ml bacterial culture were used for plasmid isolation using the illustra plasmidPrep Mini Spin Kit (GE Healthcare). The plasmid concentration was measured using the NanoDrop 1000 spectrophotometer.

3.2.5 Polymerase Chain Reaction (PCR) – amplification of the hEGFR gene

In order to amplify the hEGFR gene from the Addgene plasmid and incorporate the desired restriction sites KpnI and XbaI and protein tags (HA- and myc-tag) to the hEGFR gene, two PCRs (using the Arktik™ thermocycler) were performed in order to create an HA-tagged as well as a myc-tagged EGFR gene carrying the restriction sites KpnI and XbaI. A negative control (no template) was performed for both combinations of primers.

Table 4: PCR using Phusion Polymerase in order to create two EGFR gene constructs carrying a HA- or myc-tag, respectively.

Component	50 µl reaction	Final concentration
5x Phusion buffer HF	10 µl	1x
dH_2O	32.5 µl	
hEGFR_fwd 10 µM	2.5 µl	0.5 µM
hEGFR_rev10 µM	2.5 µl	0.5 µM
10 mM dNTPs	1 µl	200 µM
Template DNA	2.2 ng plasmid #11011	<250 ng
Phusion DNA polymerase	0.5 µl	1.0 unit/50 µl

Table 5: Thermocycling conditions for the PCR used in order to create EGFR gene constructs carrying a HA- or myc-tag, respectively.

Step	Temp	Time
Initial denaturation	98°C	30 seconds
30 cycles	98°C	10 seconds
	72°C	150 seconds
Final extension	72°C	10 min
Hold	4°C	forever

3.2.6 TAE-agarose gel electrophoresis

In order to assess the success of the PCR as well as confirm the identity and size of the gene construct a TAE-agarose gel electrophoresis was performed. A 1% agarose (Biozym) gel was prepared with 50 x TAE, SYBR Safe (Thermo Fisher Scientific) was added in a 1:10000 dilution. 10 µl of sample were mixed with 2 µl 6 x loading dye (Thermo Fisher Scientfic) and loaded onto the gel. A current of 160 V was applied for 45 – 50 min. The result was visualized under UV-light using the Gene Flash imager (Syngene Bioimaging).

3.2.7 Gel purification

Each band representing one EGFR construct was cut from the gel and purified according to the instructions of the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). Capture buffer was added according to the size of the gel fragment (10 µl for every 10 mg of gel) and incubated at 60°C until the agarose had dissolved. The DNA was eluted in elution buffer 6 (buffered dH_2O). The DNA concentration was measured using the Nanodrop 1000 spectrophotometer and the DNA was stored at -20°C in the freezer until further processing.

3.2.8 Restriction digestion

A double digestion was carried out using the KpnI HF (NEB) and XbaI (NEB) enzyme simultaneously. Three reactions, both constructs as well as the pSF-CMV-SV40 plasmid, were set up and incubated at 37°C for 2 hours without shaking.

Both XbaI and KpnI HF are stated to cut 1 µg DNA in 1 hour at 37°C (in a reaction volume of 50 µl) per unit. Both restriction enzymes show a concentration of 20000 U/ml and were therefore available in excess in the reaction mix.

Table 6: Reaction setup for the double restriction digestion of the vector pSF-CMV-SV40 and the two EGFR constructs using XbaI and KpnI HF.

Component	Restriction digestion
dH_2O	10 µl
Cut-smart buffer (10x)	3 µl
Template	15 µl
XbaI	1µl
KpnI HF	1µl

After 2 hours of incubation the enzyme cut vector pSF-CMV-SV40 was digested with calf intestinal alkaline phosphatase (CIP) from New England Biolabs, which is needed to prevent recircularization of the single-digested vector. 3 µl CIP (10 U/µl) were added to the reaction mix, mixed and spun down and further incubated for one more hour at 37°C. The DNA was purified from the enzymatic reaction using the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) and eluted in 20 µl elution buffer 6.

3.2.9 Ligation

The ligation mix was set up in microcentrifuge tubes on ice. The reaction was set up to contain 50 ng of vector and a threefold molar excess of insert and was incubated at room temperature (RT) for 30 min.

Table 7: Reaction setup for the ligation of the pSF-CMV-SV40 vector and the two EGFR constructs.

Component	Ligation
dH_2O	To 20 µl
T4 DNA Ligase buffer 10 x	2 µl
Vector (KpnI HF-, XbaI- and CIP-digested)	50 ng
Insert	125 ng
T4 DNA Ligase	1µl

3.2.10 Transformation of chemically competent *E. coli*

Self-made chemically competent *E. coli* (XL-10) were thawed on ice. 1 µl of each ligation mix (both EGFR constructs as well as the vector only control) was added to 20 µl bacterial suspension, which had been transferred to pre-chilled tubes beforehand, and incubated on ice for 20 min. After the incubation period the cells were heat-shocked at 42°C for 30 seconds and immediately placed on ice for 2 min to cool. 250 µl pre-warmed SOC medium (37°C) were added to each tube and incubated at 37°C for 1 hour.

After addition of 250 µl pre-warmed SOC medium, cells were incubated on LB-agar plates (containing 50 µg/ml kanamycin) in varying dilutions. The remaining suspension was spun down, the supernatant was discarded and the pellet was re-suspended in very little volume in order to seed onto another LB-agar plate. The ratio of colonies obtained from the transformation with an insert and the negative control was used to provide evidence on whether the phosphatase digestion and the ligation were successful. The plates were incubated overnight at 37°C (bottom up).

3.2.11 Confirming the identity of the constructs by sequencing

The identity of the two constructs was confirmed by sequencing. For this purpose three colonies per EGFR construct were used to inoculate LB-medium (containing 50 µg/ml kanamycin) and incubated at 37°C overnight while shaking at 180 rpm.

1.5 ml bacterial culture were used for plasmid isolation using the illustra plasmidPrep Mini Spin Kit (GE Healthcare). The DNA was eluted in elution buffer type 6 and the DNA concentration was subsequently measured using the Nanodrop 1000 spectrophotometer.

The EGFR gene consists of 1210 AA which equals a total of 3630 basepairs (bp). This number and some additional basepairs (tag, linker, restriction site XbaI and KpnI) needed to be analysed by sequencing. 5 sequencing primers (Sigma-Aldrich) were designed to cover the whole hEGFR construct by five overlapping fragments. The first primer was designed to anneal 70 bp upstream of the hEGFR gene, as a proper signal is only obtained 20-30 bps after the annealing site of the sequencing primer.

Table 8: Primers used for sequencing of the hEGFR constructs.

Primer name	Primer sequence
EGFR_seq1	CCGTCAGATCAGATCTTTGTCG
EGFR_seq2	TGGTCTGCCGCAAATTCCG
EGFR_seq3	CTGTTTGGGACCTCCGGTCAG
EGFR_seq4	CTGGATCCCAGAAGGTGAGAAAG
EGFR_seq5	CCTTGTCATTCAGGGGGATG

Microsynth Sequencing requires a volume of 12 µl and a concentration of 60-100 ng/µl per DNA sample. The sequencing primers were used at a concentration of 10 µM. The results were obtained as chromatograms and fasta files and were analysed using MUSCLE multiple sequence alignment to confirm the identity of the two constructs (http://www.ebi.ac.uk/Tools/msa/muscle/).

3.2.12 Cryopreservation of bacterial culture

In order to keep for long-term storage, the bacterial cultures carrying the correct EGFR construct (as confirmed by sequencing) were frozen. For this purpose Glycerol (ROTIPURAN® ROTH) was added to a final ratio of 15% and the bacterial suspension was transferred to cryogenic vials and stored at -80°C.

3.3 Generation of EGFR mutants

3.3.1 Primer design

The primers were designed according to the guidelines specified in the protocol for the QuickChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies). Both a forward and reverse primer were designed to contain the mutation of interest and to anneal to the same sequence on opposite strands of the plasmid. In addition, the mutation of interest should be in the middle of the primer showing 10-15 bases on each side. The primers were designed to show between 25 and 45 bases in length and exhibit a melting temperature larger than 78°C. A distinct formula was used for the calculation of the T_m.

$$T_m = 81.5 + 0.41(\%GC) – (675/N) - \%mismatch$$

The GC content was calculated using GeneRunner, which was used for the design of the primers. “N” in the formula specifies the length of the primer in bases. When the designed primers contains an insertion or deletion, the term %mismatch is redundant und the equation is reduced to the three remaining terms only.

The desired mutations were identified using the Catalogue of somatic mutations in cancer (COSMIC). The primers were designed using GeneRunner and were further checked for the formation of secondary structure (hairpin loops, dimers, internal loops) using this software. Four EGFR mutants were designed including the EGFR mutant carrying the single-substitution L858R (EGFR-L858R), the EGFR mutant carrying a deletion in exon 19 (EGFR-del19) and

the double mutants carrying one of those mutations and the T790M mutation in addition (EGFR-L858R-T790M and EGFR-del19-T790M).

Table 9: Mutagenic primers used for the generation of the EGFR mutants.

EGFR mutant	Mutation	primer	sequence
L858R	2573T>G	forward	GCAGCATGTCAAGATCACAGATTTTGGGC GGGCCAAACTGCTGGGTGCGGAAGAG
		reverse	CTCTTCCGCACCCAGCAGTTTGGCCCGCC CAAAATCTGTGATCTTGACATGCTGC
Exon 19 deletion	E746_A750del	forward	GTTAAAATTCCCGTCGCTATCAAGACATC TCCGAAAGCCAACAAG
		reverse	CTTGTTGGCTTTCGGAGATGTCTTGATAGC GACGGGAATTTTAAC
T790M	2369C>T	forward	GCATCTGCCTCACCTCCACCGTGCAACTC ATCATGCAGCTCATGCCCTTCGGC
		reverse	GCCGAAGGGCATGAGCTGCATGATGAGTT GCACGGTGGAGGTGAGGCAGATGC

Apart from the EGFR mutants created using site-directed mutagenesis, a SV40 origin knockout mutant was established to evaluate the functional efficiency of the SV40 origin and its impact on episomal replication and the subsequent signal amplification. The SV40 origin found on the pSF-CMV-SV40 plasmid shows a length of approximately 200 bp. A rather big stretch of 105 bp was attempted to be removed from the plasmid. For this purpose, the primers were designed especially long to facilitate binding of the primer on at least one side of the deletion. GeneRunner was used for the primer design as well as the evaluation of the formation of secondary structure, which might deteriorate the efficiency of the primer. The melting temperature was calculated according to the equation above. In addition, a sequencing primer was designed to evaluate the success of the SV40 origin knockout.

Table 10: Mutagenic primers designed for the generation of a SV40 origin knockout mutant (SV40ko). Further, the sequencing primer used for the evaluation of the success of the mutagenesis is shown.

Primer name		sequence
SV40ko	forward	CCTAACTCCGCCCATCCCGCCCCTAACTCCCTTTTTTGGA GGCCGAGGCTTTTGCAAAG
	reverse	CTTTGCAAAAGCCTCGGCCTCCAAAAAAGGGAGTTAGGG GCGGGATGGGCGGAGTTAGG
SV40ko_seq	forward	CTATCTACCGGCATTGGCGCA

3.3.2 Site-directed mutagenesis

The hEGFR construct carrying the myc-tag was used as a template for the establishment of the mutants. It is important that the template used is isolated from a dam+ *E. coli* strain.

The sample reactions were prepared, according to the protocol, as follows. The PCR was conducted using the Arkitk™ thermocycler by Thermo Fisher Scientific.

Table 11: Reaction setup for the QuickChange Lightning Site-Directed Mutagenesis Kit.

PCR component	50 µl reaction	25 µl reaction
10 x reaction buffer	5 µl	2.5 µl
dsDNA template	x µl (10-100 ng)	x µl (5-50 ng)
Forward primer	x µl (125 ng)	x µl (67.5 ng)
Reverse primer	x µl (125 ng)	x µl (67.5 ng)
dNTP mix	1 µl	0.5 µl
QuickSolution reagent	1.5 µl	0.75 µl
dH_2O	To 50 µl	To 25 µl

When the reaction mix was complete 1 µl (0.5 µl) QuickChange Lightning Enzyme was added.

The cycling parameters used for the QuickChange Lightning Site-Directed Mutagenesis method required an extension time of 30 seconds/kb and was therefore set to 260 seconds. A 2-step protocol was used (in contrast to the recommended protocol by the manufacturer), which was enabled through the particularly high melting temperatures of the primers.

Table 12: The PCR cycling parameters for the QuickChange Lightning Site-Directed Mutagenesis method used for the generation of the EGFR mutants.

Segment	Cycles	Temperature	Time
1	1	95°C	2 min
2	18	95°C	20 seconds
		68°C	260 seconds
3	1	68°C	5 min

3.3.3 DpnI digestion

Prior to the transformation of the chemically competent cells, a DpnI digestion of the site-directed mutagenesis product was carried out in order to digest the methylated template derived from the dam+ *E. coli* strain. The enzyme specifically recognizes the methylated adenine within the 5'-GATC-3' sequence and leads to cleavage.

2 µl (per 50 µl reaction volume) were added directly to each amplification reaction. The mix was mixed gently by pipetting up and down several times. After a spin down the reaction mix was incubated at 37°C for 5 min to digest the parental (methylated) DNA.

An analytical gel electrophoresis (1% agarose, TAE) was performed in order to determine the success and identity of the PCR product. In some cases the DNA did not show a band, but the transformation with the PCR product was proven to be successful nevertheless.

3.3.4 Transformation, plasmid DNA isolation and sequencing

Self-made XL-10 cells were used for the transformation. The cells were thawed on ice and 45 µl were transferred to a pre-chilled 14 ml BD Falcon polypropylene round-bottom tube. 2 µl of the ß-mercaptoethanol mix (provided with the kit) were added to the cells and the content was mixed gently. Subsequently, the cells were incubated on ice for 2 min. 2 µl of the DpnI-digested reaction mixture were added to the cell aliquot and swirled gently prior to the transformation. Again, the reaction mixture was incubated on ice for 30 min. Subsequently, the heat-shock and further steps were conducted as described in 3.2.10 with the only exception being the addition of 0.5 ml SOC medium (instead of 250 µl).

Finally, three clones per mutant were mini-prepped and sequences were confirmed as described in 3.2.11.

3.4 Transfection of HEK293T cells

3.4.1 Seeding of HEK293T cells

For the various optimization experiments, several distinct transfection vessels, CELLSTAR® 6-well and 24-well plates (greiner bio-one) as well as 12-well plates (TPP), were used depending on the amount of cells needed. A number of 200000 cells was used for each individual FACS sample, thus a multiple of this number was used, dependent on the number of samples which were analysed.

A number of $1.8*10^7$ HEK293T cells is known to cover a surface area of 55 cm^2 at full confluency, according to a protocol for the transfection of HEK293T cells published by Jan Kitajewski on the website of the University of California. This number was used as a reference for the seeding and calculation of the cell number present at a certain confluency.

Table 13: Density of HEK293T cells within different culture vessels at various confluencies.

Culture vessel	Cm^2	Cell number at full confluency	Cell number at 30% confluency	Cell number of 90% confluency
24-well	1.9	$6.2 * 10^5$	$1.9 * 10^5$	$5.6*10^5$
12-well	3.8	$1.2 * 10^6$	$3.7 * 10^5$	$1.1 * 10^6$
6-well	9.6	$3.1 * 10^6$	$9.4 * 10^5$	$2.8 * 10^6$
T25	25	$8.2 * 10^6$	$2.5 * 10^6$	$7.4 * 10^6$
T75	75	$2.5 * 10^7$	$7.4 * 10^6$	$2.2 * 10^7$

From the routine cultivation of HEK293T in T75 flasks cells were used for the experiments within the context of the assay development. The cell confluency was estimated

and the cells were subsequently dissociated in 5 ml PBS and used for the sub-cultivation of cells (3.1.3). The remaining cell suspension was used for seeding various culture vessels at 30% confluency for the subsequent transfection (18 hours afterwards). As the transfection was stated to be inhibited by a high concentration of antibiotics, the culture medium used for the transfection experiments did not contain penicillin/streptomycin.

Table 14: The amount of growth medium and cells seeded as a function of the transfection vessel in use.

	48-well plate	**24-well plate**	**12-well plate**	**6-well plate**	**T25**
Surface area [cm^2]	1	1.9	3.8	9.6	25
Growth medium without pen/strep	263 µl	0.5 ml	1 ml	2.5 ml	6.5 ml
Number of cells seeded	10^5	$1.9 * 10^5$	$3.7 * 10^5$	$9.4 * 10^5$	$2.5 * 10^6$

The recommended cell density at transfection was stated to be >80% confluency by the manufacturer. In fact, as a post-transfection incubation of 48 hours had shown to be most beneficial in this experiment, a lower cell density at transfection was found to be useful in order to avoid outgrowth and cell death.

3.4.2 Plasmid preparation

Several plasmids, including the EGFR constructs carrying two distinct tags (HA- and myc-) as well as the EGFR mutants established using via site-directed mutagenesis, were used for the transfection experiments. For this purpose, the cryopreserved bacterial cultures, which had been shown to carry to correct sequence, were used for the inoculation of liquid LB-medium (containing 50 µg/ml kanamycin). The cultures were grown overnight at 37°C on a shaker (180 rpm). Dependent on the amount of plasmid needed the DNA isolation was performed using the illustra plasmidPrep Mini Spin Kit (GE Healthcare) or the QIAGEN Plasmid Midi Kit. The DNA concentration was determined using the Nanodrop 1000 spectrophotometer.

3.4.3 Ethanol precipitation

In certain cases the transfection was carried out with a particularly high EGFR-plasmid concentration. For this purpose an ethanol precipitation was performed to increase the DNA concentration.

The DNA sample was transferred to a microcentrifuge tube and the pH was adjusted by adding 10% of its volume of 3 M sodium acetate (pH 5.2). Subsequently, at least twice the volume of 100% ethanol was added to the sample. The sample was incubated at room

temperature for 2 min, alternatively an incubation at -20°C overnight was conducted. The sample was centrifuged at 14000rpm for 5 min using the microcentrifuge 1-15PK (Sigma-Aldrich). The supernatant was removed and 500 µl of 70% ethanol were added. After briefly vortexing, the sample was again centrifuged at 14000 rpm for 5 min. The supernatant was removed and 500 µl of 100% ethanol were added. The sample was vortexed briefly and centrifuged at the same conditions. The supernatant was removed and the sample was air-dried for a few hours until all liquid had evaporated from the DNA. At this stage the DNA pellet was dissolved in dH_2O to the desired concentration.

3.4.4 Transfection

The transfection was carried out with varying amounts of DNA as well as various transfection reagents. The manufacturer Mirus Bio recommended the use of 1 µg DNA per ml culture volume. This concentration was used as a starting point and reduced by diluting the EGFR-plasmid in sterile water.

The transfection reagents *Trans*IT-X2 (Mirus Bio), *Trans*IT-LT1 (Mirus Bio), *Trans*IT-293 (Mirus Bio) and PEI (Polysciences, Inc.) were thawed and/or warmed to room temperature and vortexed gently. The plasmid dilutions were vortexed at full speed and subsequently transferred into the laminar flow hood. A certain volume (Table 15) of Opti-MEM® Reduced Serum medium (Thermo Fisher Scientific) was transferred to a microcentrifuge tube and a certain volume of the EGFR-plasmid dilution, carrier plasmid pCTCON2-CD20 (optional) and one of the four transfection reagents investigated (Table 15) were added. The carrier plasmid pCTCON2-CD20 is a plasmid commonly used for yeast display and is expected not to interfere with the expression of EGFR or any of the downstream procedures. The mixture was pipetted up and down gently to ensure the distribution of reagents and was subsequently incubated at room temperature for 30 min to allow the formation of transfection complexes. After this incubation period the transfection mixture was transferred to the cell culture vessel dropwise followed by rocking movements of the vessel to ensure a proper distribution of the transfection complexes. No change of medium was needed in order to remove the transfection complexes. The post-transfection incubation time (24-72 hours were recommended by the manufacturer) used in this assay was 48 hours.

Table 15: The volumes of EGFR-plasmid, carrier plasmid and transfection reagents TransIT-293, TransIT-LT1, TransIT-X2 or PEI used for the transfection of HEK293T cells in varying culture vessels.

		48-well plate	24-well plate	12-well plate	6-well plate	T25
Surface area [cm^2]		1	1.9	3.8	9.6	25
Growth medium		263 µl	0.5 ml	1 ml	2.5 ml	6.5 ml
Opti-MEM		26 µl	50 µl	100 µl	250 µl	0.63 ml
EGFR-plasmid		0.26 µl	0.5 µl	1 µl	2.5 µl	5 µl
Carrier plasmid pCTCON2-CD20 (1 µg/µl)		0.26 µl	0.5 µl	1 µl	2.5 µl	5 µl
Transfection reagent	PEI (1 mg/ml)	2.5 µl	5 µl	10 µl	-	-
	*Trans*IT	0.78 µl	1.5 µl	3 µl	7.5 µl	15 µl

3.4.5 Serum starvation

Although no change of culture medium was proposed by the manufacturer, the transfection medium was removed 30 hours after the transfection in order to induce serum starvation. The medium was decanted and the same amount of pre-warmed DMEM (high glucose) without additives was added as indicated in Table 14. The serum starvation was crucial for the assessment of the phosphorylation signal of the EGF-receptor upon the transfection with an EGFR-plasmid, as the FBS contains various growth factors and thus activates the receptor. Therefore, the medium change was carried out approximately 16 hours before the dissociation and subsequent analysis of the samples.

3.5 Flow cytometric analysis

3.5.1 Fixation and permeabilization of HEK293T cells

3.5.1.1 Dissociation of cells

48 hours after the transfection the supernatant in the culture vessel was removed and the cells were dissociated using PBS or DMEM. The cell suspension was pipetted up and down to facilitate an even distribution of cells and segregation of cell aggregates. The volume of the buffer (or DMEM) was chosen in order to obtain a cell number of $2*10^6$/ml. 100 µl of the cell suspension, containing $2*10^5$ cells, were subsequently transferred to a 5 ml round-bottom polystyrene tubes (BD Falcon) and used as one sample in a fixation and permeabilization experiment.

3.5.1.2 Fixation and permeabilization

A number of fixatives was investigated including several DNA-friendly alternatives to PFA already discussed in the introduction. Acetone (J. T. Baker), methanol (Chromasolv® for HPLC, >99.9%, Sigma-Aldrich), 96% ethanol and a zinc-based fixative (ZBF) containing 5 g/l Zinc chloride (MERCK), 0.5 g/l Calcium acetate (ReagentPlus® Sigma-Adrich) and 17.16 mM Zinc trifluoroacetate (Alfa Aesar) in 0.1 M Tris-HCl (this formula was published by Lykidis et al. in 2007 and named "Z7") were investigated for this purpose.

Table 16: Composition of the buffers used for cell permeabilization

buffers	composition
Tris-HCl	100 mM Tris (Trizma® base by Sigma-Aldrich) Adjusted to pH 7.4 with HCl
Tris-buffered saline (TBS)	20 mM Tris (Trizma® base by Sigma-Aldrich) 150 mM NaCl (Sigma-Aldrich) Adjusted to pH 7.5 with HCl
PBS + 1% bovine serum albumin (BSA)	A 1:10 dilution from 10 x PBS (Sigma-Aldrich) was prepared 1% (w/v) of BSA (cold ethanol fraction, Sigma-Aldrich) was added

The use of ZBF required the additional permeabilization utilizing methanol, Triton-X (pro analysi, MERCK) or saponin (for molecular biology, Sigma-Aldrich). Both Triton-X as well as saponin were used at a concentration of 0.2% in PBS containing 10 g/l BSA (PBSA).

- Fixation and permeabilization using methanol / ethanol / acetone

100 µl cell suspension (in 5 ml polystyerne tubes BD Falcon) were fixed with 1 ml of methanol, acetone or ethanol (96%) and vortexed heavily for three seconds. The fixatives, previously stored at -20°C, were transferred to ice shortly before use. After the addition of the fixative, the cell samples were incubated at 4°C for 30 min. At this point the samples fixed with methanol could be transferred to -80°C and stored for several weeks. Normally, after the incubation time, the samples were centrifuged at 500 x g for 5 min and the fixative (ethanol or acetone) was decanted. The samples fixed with methanol did not require the removal of the fixative prior to the washing steps. To ensure the entire removal of the fixatives, the samples were washed twice with 4 ml (3 ml if 1 ml methanol was still present) PBSA. Each washing step usually comprised the addition of a defined volume of solution, subsequent centrifugation at 500 x g for 5 min (at 4°C) and the subsequent removal of the supernatant. After the final washing step the supernatant was removed and the cells were used for antibody staining.

- Fixation with ZBF and subsequent permeabilization

Samples (100 µl cells) were fixed with 1 ml ZBF (pre-chilled at 4°C) and vortexed heavily for 3 seconds. The cells were incubated overnight at 4°C. The following day the cells were washed three times with 4 ml TBS, a total of 30-45 min of washing with TBS was recommended by Jensen et. al. in 2010. Thorough washing with TBS had been shown to be crucial as the phosphate present in PBS will form a precipitate at high concentrations of zinc and calcium ions. The supernatant was removed and the cells were used for the subsequent permeabilization with various reagents. The cells were treated with either 1 ml saponin (0.2% in PBSA), 1 ml Triton-X (0.2% in PBSA) or 1 ml methanol (ice-cold). The incubation time varied for the permeabilization method used and was 30 min when methanol was used and 10 min for the other permeabilizing reagents saponin and Triton-X 100. The supernatant was decanted and the cells were used for staining, but samples permeabilized with Triton-X and methanol were washed twice with PBSA beforehand. Samples, which were permeabilized using saponin, did not employ any washing steps as the staining procedure required the consecutive presence of the detergent saponin. For this reason, all subsequent steps of the staining procedure for saponin-permeabilized samples discussed below, demanded the continuous addition of 0.2% saponin to all staining and washing buffers used.

3.5.2 Staining procedure

3.5.2.1 Intracellular detection of two distinct protein tags HA and myc

After the last washing step the supernatant was decanted and the last drop attached to the tube was removed through agitation of the tube. Commonly about 100 µl were shown to be left and 80 µl of this volume were transferred to a fresh tube for the staining procedure to assure a defined and equal staining volume for all samples.

A mastermix, consisting of PBSA and the desired antibodies, was prepared containing anti-HA (16B12) antibody (BioLegend) and anti-myc-tag antibody (clone 9E10 from Thermo Fisher). 20 µl of mastermix were added to each sample and swirled to mix the content, making up a total staining volume of 100 µl. The anti-HA antibody (labelled with an Alexa Fluor 647 fluorophore) and the anti-myc antibody (labelled with an Alexa Fluor 488 fluorophore) were used at a final concentration of 1.25 µg/ml. The samples were incubated at room temperature

for 30 min (in the dark) and subsequently washed twice with 4 ml PBSA The supernatant was removed and the samples were kept on ice until analysed on the FACS Canto (BD Biosciences).

3.5.2.2 Detection of phosphorylated EGFR by flow-cytometry

- Stimulation with EGF

If stimulation with EGF was required, the cells were dissociated in the required amount of PBS to yield $2*10^5$ cells/90 µl. EGF stimulation was carried out prior to the fixation/permeabilization procedure in order to facilitate the phosphorylation process as the fixation and permeabilization procedure would naturally stop all biological processes including the EGFR activation through EGF. All samples stimulated with EGF received 10 µl EGF (1 µg/ml) making up a total volume of 100 µl and reaching a final EGF concentration of 100 ng/ml. The reference samples were treated with 10 µl PBSA at this point. The incubation with EGF lasted 5 min and was carried out at room temperature. As soon as 5 min had passed the tubes stimulated with EGF were immediately fixed with 1 ml methanol, acetone or ZBF (3.5.1). The incubation time of EGF had to be constant for each sample to ensure reproducibility and facilitate comparison between samples.

- Detection of the phosphorylation signal

After the last washing step (3.5.1) the supernatant was decanted and the last drop attached to the tube was removed through agitation of the tube. Commonly about 100 µl were shown to be left and 80 µl of this volume were transferred to a fresh tube for the staining procedure to assure a defined and equal staining volume for all samples.

Various phosphospecific antibodies were tested including the rabbit mAb anti-pEGFR Tyr1092 (D7A5) XP® (Cell Signalling Technology), the rabbit mAb anti-pEGFR Tyr1092 (D7A5) XP® labelled with phycoerythrin (PE) (Cell Signalling Technology), the rabbit mAb anti-pEGFR Tyr998 (C24A5 by Cell Signalling Technology) and the mouse mAb anti-pEGFR Tyr1197 (9H2), carrying an Alexa Fluor 647 fluorophore (by BD Biosciences). For the expression normalization of the EGF-receptor, staining with the anti-myc-tag antibody 9E10 (Thermo Fisher) was carried out in parallel.

The antibodies pY1092 and pY998 required the detection through a secondary antibody (anti-rabbit IgG (H+L) $F(ab')_2$ fragment Alexa Fluor 647 by Cell Signalling Technology).Thus

an indirect staining protocol was applied. A mastermix containing both the anti-myc as well as the phosphospecific antibody (both antibodies were used at a final dilution of 1:800) in staining buffer was prepared. 20 µl mastermix were added to each sample and samples were subsequently incubated at room temperature for 30 min (in the dark). The samples were washed twice with PBSA to remove unbound antibody. 80 µl of the remaining cell suspension were transferred to a fresh tube for the subsequent staining with the fluorophore-labelled secondary antibody anti-rabbit IgG F(ab')$_2$ fragment. Again, a mastermix was prepared and 20 µl were added to each tube, yielding a final antibody dilution of 1:500. The samples were incubated at room temperature for 30 min (in the dark), washed twice with PBSA and stored on ice until analysed via FACS.

The directly labelled antibody pY1197 (BD Biosciences) was received pre-diluted and was therefore used undiluted (10 µl) simultaneously with the application of the anti-myc-tag antibody (ThermoFisher) yielding a final antibody dilution of 1:800 and a total staining volume of 100 µl. The phosphospecific antibody pY1092-PE (Cell Signalling) was used at a 1:100 dilution while the anti-myc antibody (ThermoFisher) was used at a 1:800 dilution. The final staining volume was 100 µl. The samples were incubated at room temperature for 30 min (in the dark), washed twice with PBSA and stored on ice until analysed via FACS.

3.5.3 Blocking experiments with cetuximab

For the evaluation of the blocking effect of the anti-EGFR antibody cetuximab samples were treated with the antibody prior to the stimulation with EGF. The cells were dissociated in the required amount of PBS in order to hold $2*10^5$ cells/90 µl. Cetuximab, which was a generous gift from Dr. Josef Singer (Medical University Vienna) was adjusted to a final concentration of 100 nM. After an incubation period at 4°C for 20-25 min at 4°C, samples were treated with 100 ng/ml EGF and stained as described in 3.5.2.2.

3.5.4 Inhibition with erlotinib HCl and AZD9291

The inhibitors were received dissolved in DMSO or solid from THP and were subsequently dissolved and diluted to a concentration of 500 µM and stored at 4°C in the dark. Since TKIs operate intracellularly erlotinib HCl and AZD9291 were applied together with the serum-free medium 16 hours prior to the stimulation with EGF und staining procedure. All inhibitors were adjusted to a final concentration of 0.5 µM.

3.5.5 Extracellular staining

For the detection of the extracellular domain of the EGFR or HER2 molecule no preceding fixation and permeabilization was required. 100 µl of previously dissociated cells were diluted with 400 µl PBSA and stained with cetuximab or trastuzumab (a generous gift from Dr. Josef Singer, Medical University Vienna) adjusted to a final concentration of 100 nM. The samples were incubated at 4°C for 30 min in the dark and subsequently washed twice with 4 ml PBSA. 80 µl of the remaining cell suspension were transferred to a new tube and stained with 5 µl of a mouse anti-human IgG (Fc) CH2 domain antibody (BioRad), making up a final antibody dilution of 1:16. After 30 min of incubation at 4°C (in the dark) the samples were washed twice with 4 ml PBSA and kept on ice until the analysed on the FACS Canto (BD Biosciences).

3.6 Generation of a randomly mutated library

A randomly mutated library was established using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies). The polymerase used in this kit is called Mutazyme II DNA polymerase and is stated to be a novel enzyme blend of Mutazyme I DNA polymerase and a novel Taq DNA polymerase exhibiting increased misinsertion and misextension frequencies compared to the wild-type Taq DNA polymerase. The manufacturer claims that all different types of mutations can be obtained with the same set of buffer conditions and dNTPS (said to be balanced) only varying the amount of template or cycles used for the PCR. Naturally, if a low amount of template and a high cycle number is employed, the target gene is replicated more often and errors will accumulate leading to a higher mutation rate.

Different methods are available for the creation of a randomly mutated library including the use of nucleotide analogues such as 2′-deoxy-p-nucleoside-5′-triphosphate (dPTP) and 8-oxo-2′-deoxyguanosine 5′-triphosphate (8-oxo-dGTP)[28] or the use of an increased concentration of $MnCl_2$ and unbalanced dNTP ratios[29]. Importantly, the GeneMorph II Random Mutagenesis kit was shown to exhibit less mutation bias and cover a broader mutational spectrum.

3.6.1 Primer design

The manufacturer of the GeneMorph II Random Mutagenesis kit proposes the primers to exhibit melting temperatures ranging from 55 to 72°C. In our experiment the primers were designed to

exhibit a T_m exceeding 72°C (calculated using GeneRunner). Thus, a 2-step PCR could be carried out avoiding unspecific priming at lower temperatures.

The forward primer was designed to anneal to the pSF-CMV-SV40 vector backbone only in order to allow the randomization of the entire EGFR gene including the signal peptide. The reverse primer was designed to anneal to the myc-tag and G_4S linker as the randomization of those constructs is neither necessary nor desired for the purpose of this assay. The primers were designed using GeneRunner und were checked for the formation of secondary structure elements (hairpin loops, dimers, bulge loops, internal loops).

Table 17: Primers designed for the error-prone PCR using the GeneMorph II Random Mutagenesis Kit. The forward primer carries the KpnI restriction site depicted in bold letters and anneals solely to the vector backbone. The reverse primer shows the XbaI restriction site in bold letters, the myc-tag is highlighted in light pink and the G_4S linker sequence is marked blue.

Primer name	Primer sequence
hEGFR_epPCR_fwd	CGCTGCCAAGCTTCCGAGCTCTCGAATTCAAAGGA**GG TACC**CACC
hEGFR_epPCR_rev	AGGAGACAACT**TCTAGA**GGTCCTCTTCGGAGATCAGC TTCTGCTCAGATCCTCCGCCTCC

3.6.2 Preparation of the template for error-prone PCR

Both EGFR-WT-myc as well as EGFR-L858R-myc were used for the construction of two randomly mutated libraries.

Initially, several amounts of template ranging from 10000 ng to 10 ng were tested. The amount of template mentioned in this context refers to the amount of target DNA only. The amount of EGFR-plasmid DNA added was therefore approximately 2.2 fold greater. In subsequent experiments an amount of 1000 ng target DNA was established as the desired amount of template.

Some experiments employed a preceding plasmid linearization using the XbaI restriction enzyme. The desired amount of EGFR-plasmid (DNA isolate from *E. coli* culture) was supplemented with Cut-smart buffer and an excessive amount of XbaI enzyme before being filled up with dH_2O to the desired volume. The reaction mixture was mixed by pipetting up and down gently, spun down and incubated at 37°C for 2 hours without shaking. Afterwards, the enzyme was heat-inactivated by an incubation step at 65°C for 20 min. To remove the denatured enzyme as well as traces of Cut-smart buffer DNA purification using the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) was conducted.

3.6.3 Error-prone PCR

The error-prone PCR was conducted using the Arktik™ thermocycler by Thermo Fisher Scientific.

Table 18: PCR protocol used for the GeneMorph II Random Mutagenesis Kit

PCR component	50 µl reaction
Mutazyme II reaction buffer	5 µl
EGFR-plasmid template	x µl (1000 ng)
Forward primer	x µl (62.5 ng)
Reverse primer	x µl (62.5 ng)
dNTP mix	1 µl
Mutazyme II DNA polymerase	1 µl
dH_2O	To 50 µl

Initially 125 ng of each primer were used (according to the manufacturer's recommendations). This amount was reduced to 62.5 ng of primer in the final experiments and was shown to provide superior results.

Table 19: PCR cycling conditions for the GeneMorph II Random Mutagenesis Kit

Segment	Cycles	Temperature	Time
1	1	95°C	2 min
2	6	95°C	30 seconds
		72°C	260 seconds
3	1	72°C	10 min

A number of 30 cycles, as proposed by the manufacturer, was initially employed. The incorporation of 6 cycles only was shown to provide superior results.

3.6.4 DpnI digestion

A digestion with the DpnI enzyme was used in order to digest the parental DNA (from a dam+ *E. coli* strain) and dispose of the wild-type EGFR-plasmid.

The total PCR product obtained from the error-prone PCR reaction was used for the DpnI digestion. Cut-smart buffer and an excessive amount of DpnI enzyme were added to the reaction mixture, which was subsequently filled up with dH_2O to the desired volume and mixed gently by pipetting up and down. The reaction mixture was spun down and incubated at 37°C for one hour.

Subsequently, a preparative TAE-agarose gel electrophoresis (1% agarose) was conducted in order to dispose of the DNA fragments. The preparative TAE-agarose gel was prepared as described in 3.2.6 and the entire reaction mixture was loaded onto the gel. The band

representing the mutagenized DNA-fragment was excised from the gel and subsequently purified as described in 3.2.7.

3.6.5 PCR-amplification of randomly mutated genes

In order to increase the amount of DNA available for the subsequent cloning and restriction procedure, the DNA was amplified using Phusion Polymerase (NEB) and the Arktik™ thermocycler. In most cases, several 50 µl PCR reactions per library were prepared in order to obtain the desired amount of mutagenized DNA. The reactions were prepared as described in 3.2.5 using 10-20 ng mutagenized PCR product (DpnI-digested and purified) as template and the error-prone PCR primers (hEGFR_epPCR_fwd and hEGFR_epPCR_rev). The number of cycles was reduced to 20 for this purpose.

The PCR reactions (if several PCR reactions had been performed the products were pooled at this point) were purified using the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) in order to remove traces of the PCR components which might inhibit the subsequent restriction digestion. The DNA was eluted using elution buffer type 6.

3.6.6 Restriction digestion with XbaI and Kpn HF

The entire PCR product, previously purified, was used for the subsequent digestion with XbaI and KpnI HF (as described in 3.2.8). Both enzymes were applied in excessive (at least 10-fold) amounts.

Subsequently, a preparative TAE-agarose gel electrophoresis was conducted in order to dispose of the short DNA fragments obtained from the digestion with the restriction enzymes XbaI and KpnI HF. A TAE-agarose gel was prepared as described in 3.2.6 and the entire reaction mixture was loaded onto the gel. The band representing the mutagenized DNA-fragment EGFR was excised from the gel (around 3.6 kb) and subsequently purified as described in 3.2.7.

3.6.7 Ligation

Several ligases and transformation protocols had been formerly tested but were shown to provide unsatisfying results. Finally, the use of the Electroligase® from NEB, containing a ligation enhancer in combination with the use of the 10-beta electrocompetent *E. coli* (NEB),

especially designed for the transformation of particularly large plasmids, showed to provide superior results.

A molar insert:vector ratio of 1:1.25 was applied and an amount of 100 ng EGFR insert was normally used per ligation. The ElectroLigase and ElectroLigase Buffer were transferred to ice prior to the reaction setup. The tubes were mixed via finger-flicking. 100 ng vector pSF-CMV-SV40 (digested with XbaI, KpnI HF and CIP) was combined with 100 ng EGFR insert and the volume was adjusted to 5 µl with dH_2O. 5 µl ElectroLigase buffer and 1 µl ElectroLigase were added and the mixture was pipetted up and down 7-10 times to ensure thorough mixing of the components. The ligation reaction was incubated at room temperature for 60 min. The ligase was subsequently inactivated by incubating at 65°C for 15 min. After the inactivation the ligation reaction was chilled on ice until used for transformation.

The purification of the ligation reaction using the Monarch® PCR and DNA cleanup Kit by NEB was shown to provide greater results if performed prior to the transformation of the electrocompetent *E. coli* and was therefore employed after the heat inactivation. The DNA was eluted in the smallest elution volume possible (6 µl) in order to increase the DNA concentration and subsequent transformation efficiency.

3.6.8 Transformation of electrocompetent *E. coli*

10-beta electrocompetent *E. coli* were thawed on ice (for about 10 min) and mixed gently by finger-flicking. 45 µl *E. coli* per reaction were added to a pre-chilled Eppendorf tube. The electroporation cuvettes (1 mm) by Bio-rad were transferred on ice. 2 µl ligation mix were transferred to the cell suspension and the *E.Coli*/DNA mixture was transferred into a pre-chilled electroporation cuvette without introducing bubbles. The electroporation was conducted using the Gene Pulser Xcell™ and the conditions were set to 2.0 kV, 200 Ω and 25 µF. The time constant was recorded and should usually be in the range of 4.8 to 5.1 ms. Immediately after the electroporation 960 µl pre-warmed SOC medium (37°C) were added to the cell suspension, mixed up and down twice and transferred to an Eppendorf tube. The tubes were incubated on the thermocycler (Eppendorf Thermomixer comfort) at 37°C until all of the reactions were processed. Then, the cell suspension was transferred to a round-bottom polystyrene tube and shaken vigorously (250 rpm) at 37°C for one hour. In the meantime the LB-agar plates (containing 50 µg/ml kanamycin) were pre-warmed at 37°C for one hour. Finally, the cells were diluted in LB-medium and the dilutions were spread on the pre-warmed selective plates. The plates were incubated at 37°C overnight (bottom up).

3.6.9 Isolation of library plasmids and freezing of *E. coli* libraries

Only a minor volume of the *E. coli* cell suspension was used for the preparation of these dilutions. The major part of the transformed bacterial suspension was used for the inoculation of a volume of 100 ml LB-medium (containing 50 µg/ml kanamycin). This cell suspension (in an Erlenmeyer flask) was shaken (at 180 rpm) overnight at 37°C on the GFL 3033 shaker. The cell suspension in fact represents the randomly mutated EGFR library and was thus used for DNA isolation using the QIAprep Spin Miniprep Kit (Qiagen). In addition, several cryogenic tubes were prepared from this cell suspension (containing 15% glycerol) and frozen at -80°C.

3.6.10 Sequencing of library mutants

Ten colonies per library were mini-prepped and sequences were confirmed in order to assess the mutation rate (as described in 3.2.11).

3.7 DNA isolation from methanol-fixed cells

Usually 200.000 cells per sample were used for plasmid DNA isolation. The cells were transfected (as described in 3.4) and serum starvation was induced 30 hours after the transfection. 48 hours after the transfection cells were dissociated with the desired amount of PBS. To determine the difference in the efficiency of plasmid DNA isolation, both fixed and non-fixed cell samples were used for the extraction of plasmid DNA, making use of a number of different protocols. The fixed cell samples were treated with 1 ml ice-cold methanol and incubated at 4°C for 30 min. One sample was stored at -80°C for the subsequent analysis via FACS evaluating the success of the transfection and resulting EGFR expression. The other samples were washed twice with 4 ml PBSA and the remaining cell pellet was used for the sequent isolation of plasmid DNA.

In order to verify the impact of the episomal replication facilitated through the SV40 origin, cell were transfected with the wild-type SV40 origin plasmid as well as a plasmid lacking the SV40 origin (SV40ko-plasmid). The samples were compared based on their number of plasmids per cell.

3.7.1 DNA isolation using the QIAprep Spin Miniprep Kit

A protocol published by Qiagen for the use of the QIAprep Spin Miniprep Kit for the plasmid DNA isolation from mammalian cells, was used as the basis for several protocols tested. In

most of the cases only one or two alterations were made to the initial protocol. The protocol described in the following paragraph represents the basic protocol, adapted from the one published by Qiagen (https://www.qiagen.com/kr/resources/faq?id=b3a99f9f-150b-44a4-90dc-3f55f5856f54&lang=en&Js=0), and was used for subsequent alterations and variations.

Basic protocol: The samples (fixed as well as non-fixed) were transferred to microcentrifuge tubes (1.5 ml) and spun down at 500 x g for 5 min. The supernatant was discarded with a pipette and as little as possible remaining volume was left. The cells were re-suspended in 250 µl buffer P1 (50 mM Tris-HCl, pH 8.0; 10 mM EDTA; 100 µg/ml RNase A) and subsequently, 250 µl buffer P2 (200 mM NaOH, 1% SDS) were added and the tubes were inverted several times. After an incubation for 5 min at room temperature the genomic DNA and cell debris were precipitated by adding 350 µl buffer N3 (potassium acetate and chaotropic salts). The samples were incubated on ice for 5 min prior to a centrifugation step at 18000 x g for 10 min. The supernatant was loaded onto the column and subsequently centrifuged at 16000 x g for 1 minute. The samples were washed with 500 µl PB buffer and centrifuged. Then, the samples were washed with 750 µl PE buffer and centrifuged twice to ensure complete removal of the ethanol present in the buffer. The column was transferred to a fresh microcentrifuge tube (1.5 ml) and 50 µl elution buffer type 6 (GE Healthcare) were added. The samples were incubated at 37°C for 5 min on the thermocycler (Eppendorf Thermomixer comfort), followed by centrifugation at 16000 x g for 1 minute in order to elute the DNA.

Method 1: The addition of Proteinase K is an integral part of most of the gDNA isolation kits for mammalian cells. A product description of Proteinase K by NEB states that Proteinase K is active in a wide range of temperatures (20 – 60 °C) and pH values (7.5 – 12.0). Further, the activity seems to be stimulated when up to 2% SDS or up to 4 M urea are added to the reaction. Proteinase K even seems to be active in buffers containing chelating agents such as EDTA. For these reasons, the effect of Proteinase K addition on plasmid isolation efficiency was tested. 10 µl Proteinase K (Qiagen) were added to the samples after the addition of buffer P2. The samples were inverted 6 times and incubated at 56°C (Eppendorf Thermomixer comfort) for 10 min. Afterwards, the samples were allowed to cool and 350 µl buffer N3 were added, followed by gently inverting the tubes in order not to shear the genomic DNA. Subsequently, the basic protocol was followed.

Method 2: The samples were prepared according to method 1, with the only exception being a prolonged incubation time of 60 min at 56°C.

Method 3: The samples were prepared according to method 1, with the only exception being the addition of 50 µl Proteinase K instead of 10 µl.

Method 4: The samples were prepared as according to method 1, with the only exception being a prolonged incubation time of 60 min (56°C), shaking at 700 rpm.

Method 5: To test whether Proteinase K would require high pH as in buffer P2, 10 µl Proteinase K were added to buffer P1 in this protocol. The cells were re-suspended in 250 µl buffer P1. 10 µl Proteinase K were added and vortexed for 3 seconds. After an incubation at 56°C for 60 min, the samples were allowed to cool at room temperature prior to the addition of 250 µl buffer P2. Subsequently, the basic protocol was followed.

Method 6: In order to test whether the addition of 1% SDS (detergent) to buffer P1 would benefit Proteinase K, a similar experimental setup to method 5 was chosen, with the only exception being the addition of 1% SDS to buffer P1. After an incubation step at 50°C overnight while shaking at 550 rpm, the samples were allowed to cool at room temperature. 250 µl buffer P2 were added and the basic protocol was followed

Method 7: The samples were prepared according to method 6, with the only exception being an incubation step at 56°C for 60 min only.

Method 8: A step which is commonly employed within gDNA isolation kits is an incubation step at 65°C after the addition of the lysis buffer (instead of the addition of Proteinase K). Cells were prepared as described in method 7 except the addition of 10 µl Proteinase K and were incubated at 65°C (60 min) instead.

Method 9: QIAmp DNA Blood Mini Kit (described in 3.7.2)

Method 10: The formulation of the lysis buffers used in the gDNA isolation kits for the DNA extraction from mammalian cells are unknown, but were stated to contain "anionic detergent with additional DNA stabilizers". To investigate the effect of alternative detergents, cells were re-suspended in buffer P1 containing 5% Tween 20 and 0.5% Triton-X 100. 10 µl Proteinase K were added, following an incubation step at 56°C for 60 minutes. The tubes were allowed to cool before 250 µl buffer P2 were added.

Method 11: To test whether the use of the QIAprep Spin Miniprep kit would benefit from the use of a commercially available lysis buffer used for the cell lysis of mammalian cells, the lysis buffer from the QIamp DNA Blood Mini kit was used within the basic protocol used for the QIAprep Spin Miniprep kit. The cells were re-suspended in 200 µl PBS, before 10 µl Proteinase K and 200 µl AL buffer (QIamp DNA Blood Mini kit) were added. The samples were vortexed for 3 seconds and incubated at 56°C for 60 min. The tubes were allowed to cool, before 250 µl buffer P2 were added. Subsequently, the basic protocol was followed.

3.7.2 DNA isolation using the QIAamp DNA Blood Mini Kit

As a reference the QIAamp DNA Blood Mini Kit was used in parallel. Cells were prepared according to the manufacturer's protocol for cultured cells.

3.7.3 Polymerase Chain Reaction (PCR)

The amplification of the full-length EGFR gene from the DNA isolated from fixed or non-fixed HEK293T cells (described in 3.7.1 and 3.7.2.) was performed using the Phusion DNA polymerase (NEB) in several different PCR setups (using the Arkitk™ thermocycler). Several PCR conditions were tested in order to establish a reproducible PCR protocol for the amplification of the full-length EGFR gene.

The amplification of the full-length EGFR gene using Phusion Polyermase was used as a first estimate for the efficiency of the DNA isolation employing any of the DNA isolation protocols.

Table 20: Phusion PCR cycling conditions for the amplification of the full-length EGFR gene (from the DNA isolated from HEK293T cells using several distinct protocols)

Component	50 µl reaction	20 µl reaction	Final concentration
5x Phusion buffer HF / GC	10 µl	4 µl	1x
10 mM dNTPs	1 µl	0.4 µl	200 µM
10 µM hEGFR_epPCR_fwd	2.5 µl	1 µl	0.5 µM
10 µM hEGFR_epPCR_rev	2.5 µl	1 µl	0.5 µM
Template DNA	variable	variable	<250 ng
Phusion DNA polymerase	0.5 µl	0.2 µl	1.0 units/50 µl
DMSO (optional)	1.2 µl or 5 µl	0.6 µl or 2 µl	3% or 10%
dH_2O	To 50 µl	To 20 µl	

Table 21: Initial PCR cycling conditions used for the amplification of the full-length EGFR gene from DNA isolate from HEK293T cells with Phusion Polymerase.

Step	Temp	Time
Initial denaturation	98°C	30 seconds
30 cycles	98°C 72°C	10 seconds 150 seconds
Final extension	72°C	10 min
Hold	4°C	forever

Table 22: Variation 1 of the PCR cycling conditions used for the amplification of the full-length EGFR gene from DNA isolate from HEK293T cells. The extension time was reduced to 75 seconds only.

Step	Temp	Time
Initial denaturation	98°C	30 seconds
30 cycles	98°C	10 seconds
	72°C	75 seconds
Final extension	72°C	10 min
Hold	4°C	forever

Table 23: Variation 2 of the PCR cycling conditions used for the amplification of the full-length EGFR gene from DNA isolate from HEK293T cells with Phusion Polymerase. The number of cycles was increased to 40.

Step	Temp	Time
Initial denaturation	98°C	30 seconds
40 cycles	98°C	10 seconds
	72°C	150 seconds
Final extension	72°C	10 min
Hold	4°C	forever

Following the PCR reaction, the products were analysed using TAE-agarose gel electrophoresis (1% agarose).

3.7.4 Quantitative PCR (qPCR)

qPCR was conducted in order to quantify the number of plasmids extracted per cell for each distinct DNA isolation method. The real-time PCR Thermocycler MiniOpticon (BioRad) and KAPA Probe FAST Universal Mastermix (Kapa Biosystems) were used for this purpose. The primers and probe were designed by Microsynth to produce an amplicon of 110 bp length. Further, the primers and probe were checked for similarities to the carrier plasmid pCT-CON2 CD20, which was present in the DNA isolate for certain samples, as well as the human genome. In fact, the primers and probe were designed to anneal to the kanamycin resistance gene found on the pSF-CMV-SV40 plasmid backbone. The probe carries a 5'FAM and 3'BHQ modification. 6-carboxyfluorescein (FAM) is a fluorophore and reporter while BHQ (Black Hole Quencher) is a quencher which can quench the activity and fluorescence of the fluorophore FAM as long as those two molecules are in close vicinity. As soon as the probe hydrolyses (this happens when the amplicon is amplified and the DNA polymerase replaces the qPCR probe which is situated in between the forward and reverse primer) and the FAM fluorophore is at distance to its quencher it exhibits a fluorescent signal which is detected by the qPCR thermocycler.

Table 24: Primers and probe used for the quantification of the EGFR-plasmid in DNA isolates from HEK293T cells via qPCR

Primer name	Primer sequence
qPCR_fwd	AGAGGCTATTCGGCTATGA
qPCR_rev	CAGGTCGGTCTTGACAAAA
qPCR_probe	TCTGATGCCGCCGTGTTCCG

Relative as well as absolute quantification is possible when qPCR is performed. Usually, if the whole genome is extracted relative quantification is applied since the number of one particular gene can be referenced to the number of so-called "reference genes". In this present study relative quantification was not possible. Therefore a dilution series of known concentrations of EGFR-plasmids (in the pSF-CMV-SV40 backbone) was prepared and used in order to set up a calibration curve. To create a matrix that is similar to the samples, the DNA isolate from non-transfected cells (should not carry any plasmid) was added to the dilution series.

The whole vector plus EGFR-insert was calculated to be 8079 bp in length. The amount of DNA representing a certain number of EGFR-plasmid copies was calculated according to this equation.

$$\text{DNA [ng]} = \frac{number\ of\ copies * length\ (bp) * 10^9 * 650}{6.022 * 10^{23}}$$

A 1:10 dilution (with dH_2O) was prepared from the DNA isolate from non-transfected HEK293T cells while a 1:20 dilution (with dH_2O) was prepared from the remaining samples. The 2 x KAPA Probe Fast qPCR MasterMix was thawed and stored on ice. The MasterMix contains all the necessary components except for primers, probe and template. The KAPA Taq HotStart DNA Polymerase is an antibody-blocked DNA polymerase and therefore the handling at room temperature for some time is not detrimental. A mastermix (20% excess) was prepared and mixed gently, then aliquoted to the desired number of tubes.

Table 25: qPCR Mastermix prepared from 2 x KAPA Probe FAST qPCR Mastermix, qPCR primers, probe and 2% DMSO

Component	Volume (1 reaction)	Final concentration
PCR-grade water	6.0 µl	
2 x KAPA Probe FAST qPCR MasterMix	10 µl	1x
10 uM qPCR_fwd	0.5 µl	250 nM
10 uM qPCR_rev	0.5 µl	250 nM
10 uM qPCR_probe	0.6 µl	300 nM
DMSO (NEB)	0.4 µl	2%

The dilution series, negative control and DNA isolate samples were prepared to hold 20% excessive volume as a volume of 20 µl was finally transferred to the 48-well PCR plates (Bio-rad) by reverse pipetting. The samples were mixed by pipetting up and down several times before being transferred to the PCR plate, which was closed with a Microseal® PCR Plate sealing film (Bio-rad). On each plate a series of dilutions ranging from 10^2 to 10^7 plasmids per sample was analysed in addition to the samples.

Table 26: Sample preparation of the dilution series ranging from 10^2 to 10^7 plasmid per sample, the negative control and the DNA isolate samples

	Dilution series	**Negative control**	**DNA isolate samples**
Mastermix	21.6 µl	21.6 µl	21.6 µl
Matrix 1:10	1.2 µl	1.2 µl	-
template	1.2 µl Plasmid dilution	1.2 µl dH_2O	2.4 µl template 1:20
Volume [total]	24 µl	24 µl	24 µl

Table 27: qPCR cycling conditions using the MiniOpticon (Bio-rad) and 2 x KAPA Fast Mastermix for the quantification of the plasmid number present in the DNA isolate from HEK293T cells

Segment	**Cycles**	**Temp**	**Time**
1	1	95 °C	3 min
2	40	95 °C 58 °C	5 seconds 30 seconds

The data were analysed using the CFX software and the samples were quantified based on the calibration curve. The quantification cycle (Cq) value was used to determine the number of plasmids present in each sample.

4 Results

4.1 Extracellular detection of EGFR and HER2

4.1.1 Evaluation of HER2 and EGFR expression on HEK293T cells

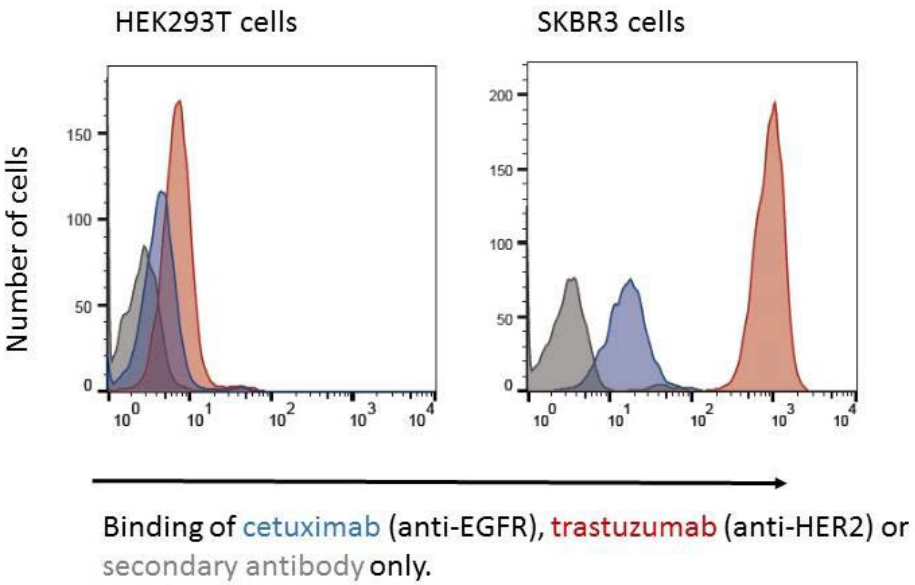


Figure 7: To evaluate HER2 and EGFR expression both HEK293T and SKBR3 cells were incubated with 100 nM of either cetuximab or trastuzumab, respectively. SKBR3 cells were used as a positive control for HER2 expression. Both cetuximab and trastuzumab are human IgG1 antibodies and were detected using an anti-human IgG-Fc antibody (directed against the CH2 domain). The secondary antibody only control was used to show that the signal is specific and due to binding of the primary antibody.

Apart from homodimerization, EGFR is known to heterodimerize with other members of the ERBB family such as HER2. Therefore, an important property of an appropriate cell line for our selection approach is low endogenous background expression of both EGFR and HER2. In order to test whether HEK293T is an appropriate cell line for our selection approach, we analysed the expression level of EGFR and HER2 on these cells. SKBR3 cells, a human breast cancer cell line which overexpresses HER2, served as a positive control for HER2 expression. EGFR and HER2 were analysed using cetuximab and trastuzumab, respectively. The results clearly showed that HEK293T cells only express low levels of EGFR and HER2. In fact, HER2 expression is about two orders of magnitude lower compared to the levels on SKBR3 cells (Figure 7). These results are in line with a previous study, where HEK293 cells were even used as a negative control for EGFR and HER2 expression[30].

The low endogenous expression of EGFR was thought to be beneficial for the purpose of this study, providing the required downstream components of EGFR signalling. Thus, the low expression levels of both EGFR and HER2 detected in this experiment had promoted the use of the HEK293T cell line as a valuable model cell line in this study.

4.1.2 Detection of EGFR on transfected HEK293T cells

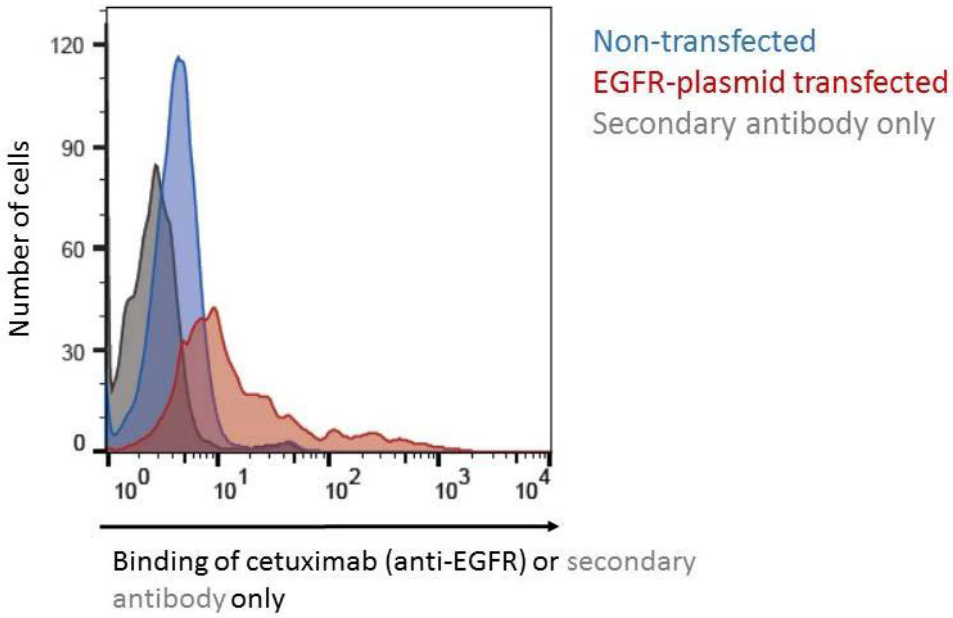


Figure 8: HEK293T cells were transfected with 0.3 ng/ml plasmid carrying the EGFR gene (+ 1 µg/ml carrier plasmid) using the transfection reagent TransIT-X2. Both transfected as well as non-transfected cells were incubated with 100 nM cetuximab. For detection an anti-human IgG-Fc antibody (directed against the CH2 domain) was used.

After successful cloning of the pSF-CMV-SV40 plasmid carrying the EGFR gene, EGFR expression was analysed after transient transfection of HEK293T cells with this plasmid.

The result in Figure 8 clearly shows that both cloning of the EGFR-carrying plasmid as well as the transfection with the EGFR-plasmid was successful and yielded cells which over-expressed EGFR. The expression of EGFR was detected using the anti-EGFR antibody cetuximab. The rather low number of highly expressing cells of about 10% is due to the low concentration of plasmid in order to avoid transfection of one cell with multiple plasmids, as will be discussed below. Transfected cells show EGFR levels that were one or two orders of magnitude above the non-transfected control.

4.2 Optimization of fixation and permeabilization of HEK293T cells

As a first step of assay development the fixation and permeabilization procedure was optimized. For this purpose the transfection reagent *Trans*IT-293, a transfection reagent optimized for the

transfection of HEK293 cells, was used and the manufacturer's recommendations concerning the plasmid concentration were initially accepted in order to provide a starting point for the optimization procedure. The six fixation/permeabilization methods were compared with regard to their effect on the cell morphology, the peak resolution (represented as the Correlation of Variation) and fold change calculated by comparing EGFR expression on transfected (i.e. EGFR-positive) vs. non-transfected (i.e. EGFR-negative) cells.

4.2.1 Change of morphological properties upon fixation and permeabilization

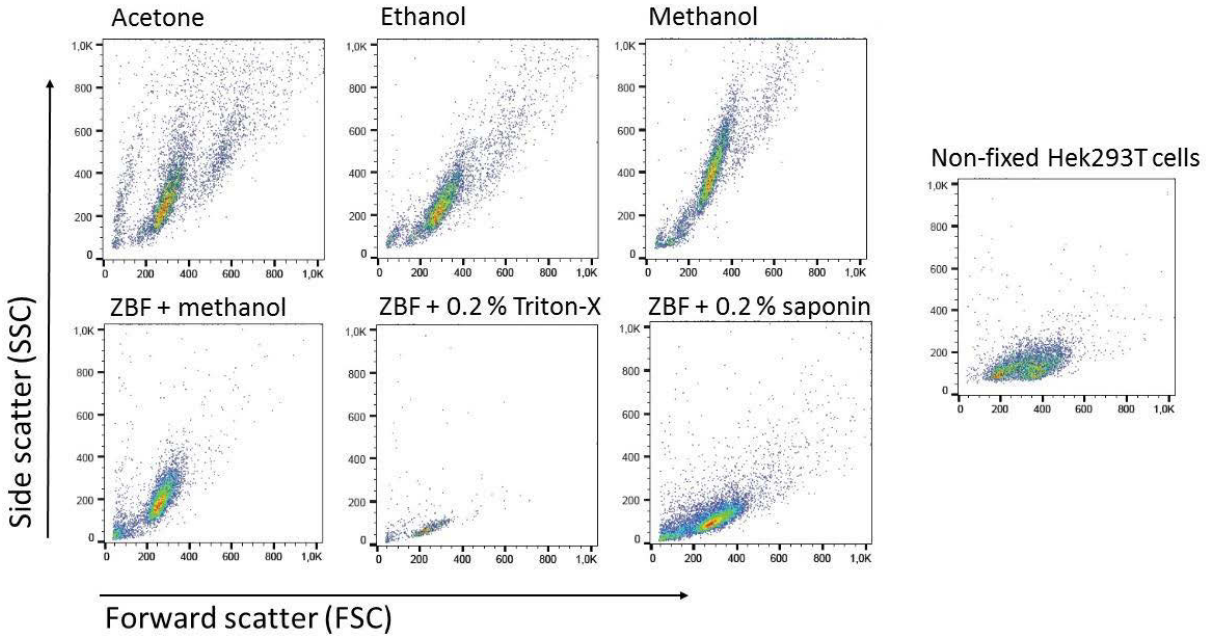


Figure 9: Comparison of the cell morphology of HEK293T cells subjected to a panel of 6 distinct fixation/permeabilization reagents. 48h after transfection with 1 µg/ml of EGFR-plasmid using the transfection reagent TransIT-293, cells were dissociated in PBS/TBS and subjected to a fixation/permeabilization step. To the right, the morphology of non-fixed HEK293T cells is shown as a control.

Figure 9 shows the morphology of HEK293T cells after treatment with several distinct fixation/permeabilization reagents including acetone, methanol, ethanol and zinc-based fixative (ZBF), the latter also included a subsequent permeabilization step. HEK293T cells, not subjected to any fixation/permeabilization reagent nor labeled with any antibody were used as a control in assessing the changing morphology as a function of fixation/permeabilization reagent. Samples fixed/permeabilized using acetone showed an increase in cell fragments and aggregates. The use of methanol as a fixative resulted in a large increase in side scatter, a fact which had already been observed and described by Krutzik et al. (2003), who further described that this effect was largely eliminated if a prior paraformaldehyde (PFA) fixation step was performed. Ethanol also resulted in increased SSC values compared to untreated cells, but the

effect was less pronounced compared to methanol treatment. ZBF was used as a fixative only and required a subsequent permeabilization step. The usage of ZBF in combination with 0.2% saponin (in PBSA) or 0.2% Triton-X (in PBSA) showed to preserve the cell morphology the best of all fixatives investigated (Figure 9). The use of 0.2% Triton-X as a permeabilization reagent though resulted in a major loss of cells which was thought to be due to the harsh chemical nature of the permeabilizing reagent and resulting cell lysis. Apart from its fixative properties methanol is also known to be used in combination with e.g. PFA as a permeabilizing reagent. The use of methanol in combination with ZBF resulted in a slight increase in the side scatter of cells.

4.2.2 Detection of intracellular expression tags

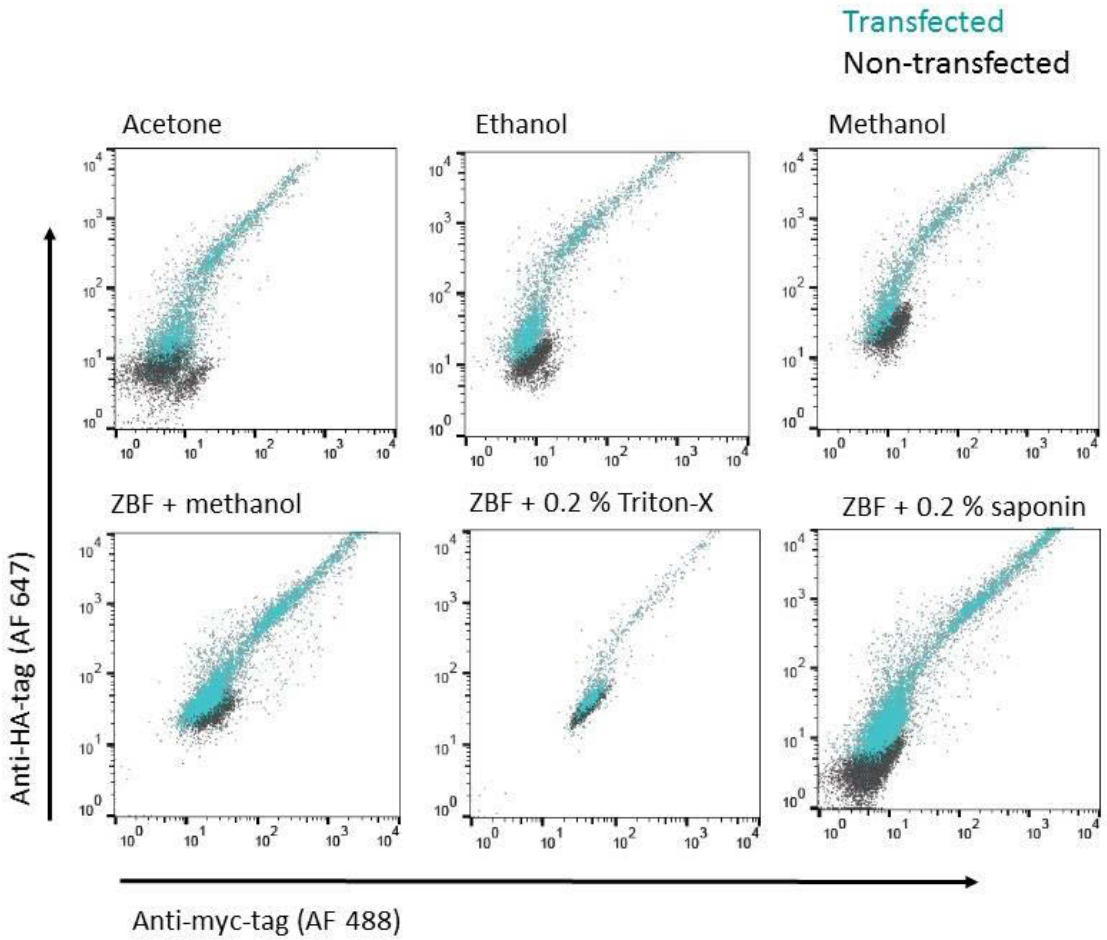


Figure 10: Intracellular staining of expression tags, after fixation/permeabilization with different reagents as indicated. HEK293T cells were transfected with equal amounts of EGFR-plasmid carrying 2 distinct protein tags, an HA-tag and myc-tag, respectively, at the intracellular C-terminus of EGFR. The total plasmid concentration was 1 µg/ml and the transfection reagent TransIT-293 was used. 48 h after transfection cells were harvested and dissociated in DMEM. Both transfected (illustrated in blue) as well as non-transfected (grey) cells were fixed/permeabilized with the reagents indicated on top of the dot plots and stained with anti-myc and anti-HA antibodies carrying distinct fluorophores. Each dot plot illustrates an overlay of transfected and non-transfected cells both treated with the fixative indicated.

Next, we compared the efficiency of intracellular detection after fixation/permeabilization with different reagents. HEK293T cells were transfected with a 1:1 mixture of two different plasmids carrying a myc- and an HA-tag at the C-terminus of EGFR, respectively. Importantly, since the tags were fused to the C-terminus of EGFR, they were expressed intracellularly. After treatment with different fixation/permeabilization reagents, cells were stained with antibodies recognizing the two tags and analysed by flow cytometry. Figure 10 shows dot plot overlays of non-transfected and transfected HEK293T cells.

Remarkably, all six procedures used for fixation/permeabilization of cells enabled detection of myc- and HA-tags, demonstrating that the cells were efficiently permeabilized and that the antibodies were able to reach their intracellular targets. Importantly, only a certain percentage of cells within the transfected cell sample (blue population in Figure 10) actually carries an EGFR-plasmid and thus shows an increase in binding to the myc- as well as HA-tag upon transfection. The remaining cell population shows anti-HA and anti-myc binding, which is comparable to the negative control (grey) and is therefore EGFR-negative. Moreover, the position of the negative population seemed to be dependent on the chemical nature of the fixative and certain characteristics which arose through the use a certain fixatives, such as the formation of aggregates using acetone (Figure 9), could also be observed in this dot plot.

4.2.3 Comparing the use of PBS versus DMEM and its impact on peak resolution

As a next step, we evaluated the impact of the medium (DMEM vs. TBS/PBS) used for cell dissociation on peak resolution (Coefficient of Variation, CV) after intracellular staining of non-transfected (i.e. negative) cells.

CV values were calculated for each fixation/permeabilization method and are presented in Figure 11A. Figure 11 reveals that some changes in peak resolution were observed as a consequence of the medium used. The use of DMEM as a medium showed to have a particularly negative impact on the peak width and resolution when acetone was utilized as a fixation/permeabilization reagent. While the CV was measured to be 29.9 for anti-myc binding in PBS, the peak broadening as a consequence of the use of DMEM is represented in an increased CV value of 64.5 (Figure 11A) and is visualized in a histogram (Figure 11B) as one representative example. This effect was observed for both antibodies (anti-HA as well as anti-myc) used. On the other hand, the peak position and width was shown to be stable if methanol was used as a fixative reagent (Figure 11B). The CV values were comparable for the histograms received for each antibody (32.5 vs. 33.6 for anti-HA binding and 28.1 vs. 28.6 for anti-myc

binding) irrespective of the medium used for cell dissociation. CV values of using Triton-X-permeabilized cells were found to be particularly low for both DMEM and TBS, but here the major loss of cells should not be neglected. In summary, CV values received after the permeabilization with 0.2% saponin or acetone were found to be particularly high, suggesting a broadening of the cell population (Figure 11A).

Peak resolution is of importance as it represents the broadening of population. Importantly, a large peak width will impede the separation of two populations from one another.

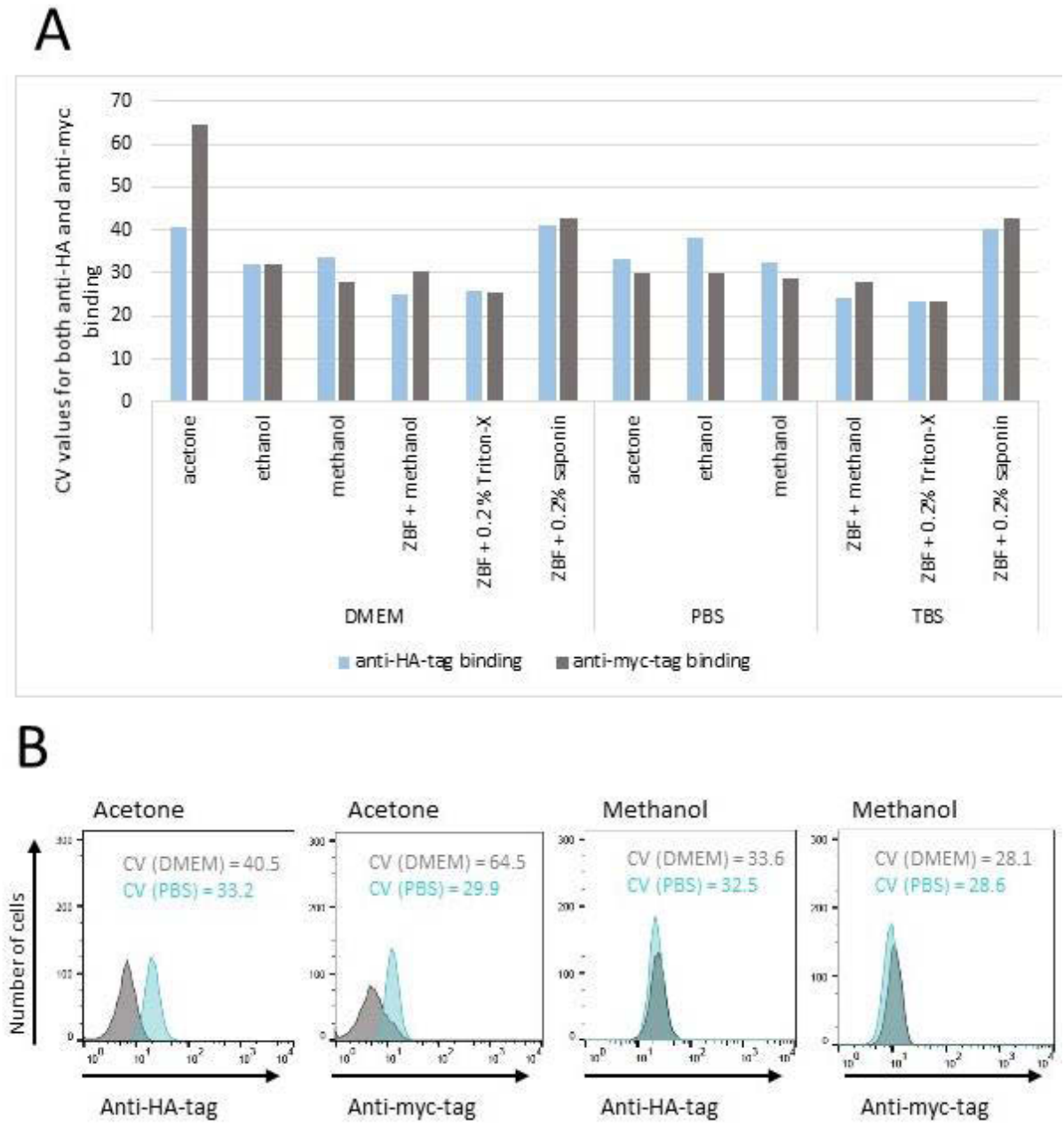


Figure 11: Non-transfected HEK293T cells were dissociated in either PBS or DMEM and the subsequent fixation/permeabilization procedure was performed in this medium. (A) CV values were obtained from the histograms for each antibody in use. (B) The histograms for binding to either myc-tag or HA-tag (AF 488 and AF 647, respectively) are illustrated as an overlay comparing the use of PBS and DMEM for methanol and acetone.

Together, these results indicate that the peak resolution and subsequent separation of populations is strongly dependent on the medium used for the cell detachment, as well as the fixation/permeabilization reagent.

4.2.4 Comparison of the fold change obtained for EGFR expression after fixation/permeabilization with different reagents

A "fold change" for EGFR expression was calculated from the median fluorescence intensity (MFI) for binding to the protein tags of EGFR-positive cells relative to the MFI of EGFR-negative cells within the same sample. The fold change was chosen as a parameter to indicate how well positive cells could be segregated from the EGFR-negative population within the same sample. A gate, containing the same percentage of positive cells (35%), was applied to all samples, providing a way of comparison. The calculated fold change presented in the bar chart in Figure 12B is also reflected in the histograms shown in Figure 12A.

The results clearly show that the fold change depends on the fixative/permeabilizing reagent, as the changes in signal intensity were observed for both antibodies in parallel (Figure 12B). Fixation/permeabilization with acetone for example yielded a rather wide cell population (visible as one wide peak in Figure 12A) which impedes the separation of the EGFR-positive cells from the negative population. This finding is reflected in rather low fold change values of 5-fold for anti-myc binding and 18-fold for anti-HA binding (both in DMEM). In comparison, the use of ZBF in combination with saponin as a permeabilizing reagent resulted in a well-defined population of positive cells that is well separated from the negative population (Figure 12A), which is reflected in a particularly high fold change for both antibodies (44-fold for anti-HA binding and 15-fold for anti-myc binding, both in DMEM). The use of Triton-X on the other hand resulted in a sharp peak, though, the positive cells (which should still account for 35% of the overall cell population) seemed to merge into the negative population and could therefore prevent the separation of populations (Figure 12A). A fold-change of 8-fold (anti-myc) and 2-fold (for anti-HA binding) could be shown for the fixation/permeabilization in DMEM (Figure 12B). Methanol provided equally good results for both DMEM and PBS and showed comparable fold changes to ethanol. The use of ZBF in combination with methanol was shown to provide inferior results (showing fold changes comparable to those received for ZBF + Triton-X) in comparison to the use of methanol only.

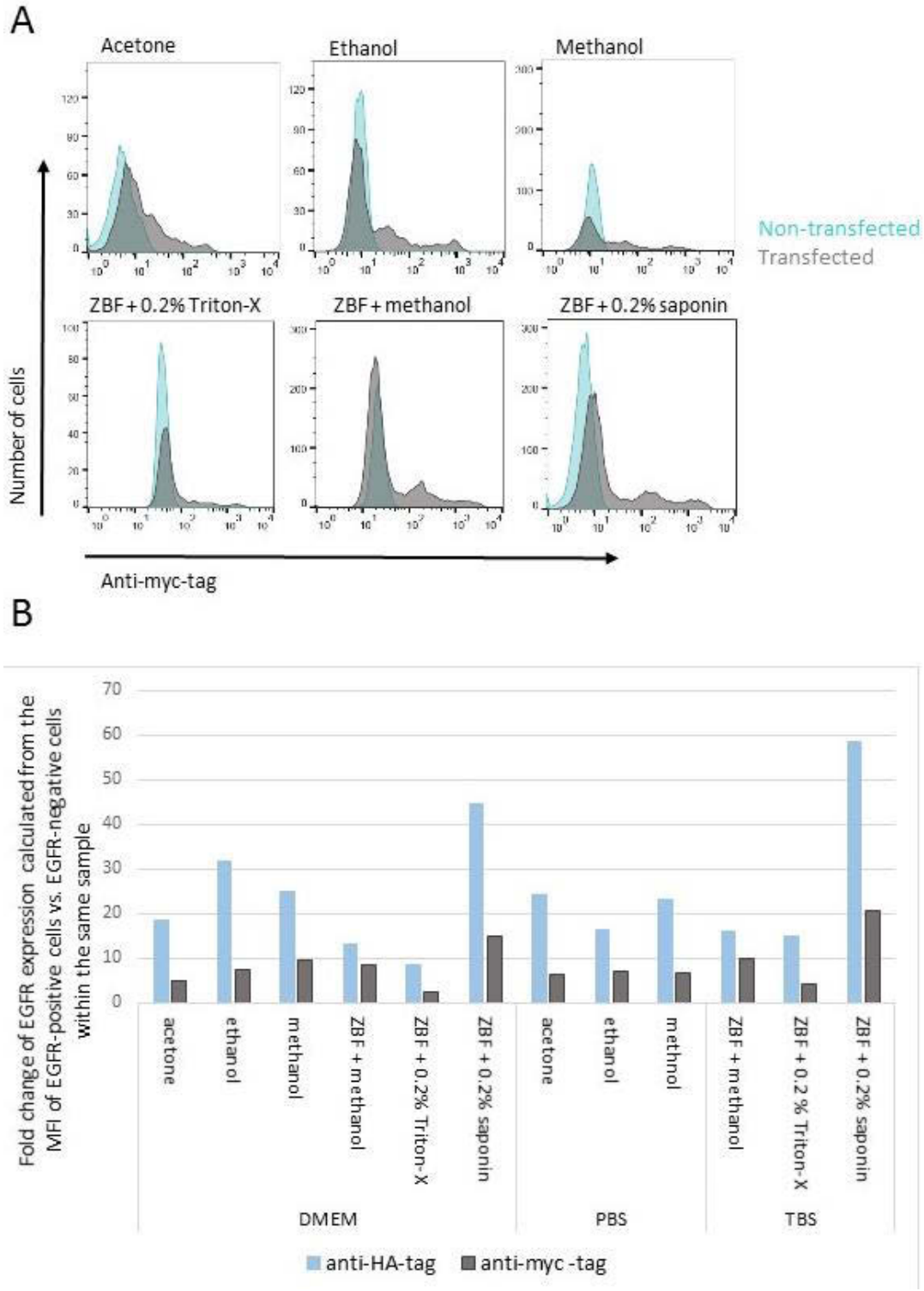


Figure 12: Both transfected (transfection with 1 µg/ml EGFR-plasmid encoding either an HA- or myc-tag) and non-transfected HEK293T cells were subjected to fixation/permeabilization and subsequent antibody staining with an anti-myc antibody (AF 488) and anti-HA (AF 647) antibody. (A) The plots represent overlays of anti-myc binding to transfected and non-transfected cells for each particular fixation/permeabilization procedure and after dissociation in DMEM.(B) Using the FlowJo software a gate was set containing a specific percentage (35%) of EGFR-positive cells. The fold change of EGFR expression was calculated from the MFI of EGFR-positive cells versus the MFI of the EGFR-negative cell population (within the same sample).

4.3 Optimization of transfection yielding single positive cells

The main aim concerning the transfection of HEK293T cells was to maximize the number of "single positive" cells, meaning that those cells would be transfected with one single plasmid and subsequently only express one single EGFR mutant on the cell surface. Still, the percentage of transfected and thus positive cells was required to be as high as possible in order to screen as many mutants as possible.

For this purpose, methanol was chosen as a fixative and permeabilizing reagent as this reagent was shown to facilitate fixation and intracellular staining. After having established the transfection protocol, the desired protocol was tested in combination with every fixation/permeabilization procedure to rule out any bias or interdependencies.

The use of two plasmids containing two different protein tags (HA- and myc-tag) enabled the estimation of the percentage of cells that only took up one plasmid during transfection (single positive cells) and of cells that took up multiple plasmids during transfection (double positive). In this study the term "single positive" cells refers to cells only expressing one of those tags. Although single positive cells may also have been transfected with several plasmids with the same tag, the probability of carrying more than two plasmids is low.

4.3.1 Comparison of the transfection efficiency as a function of transfection reagent

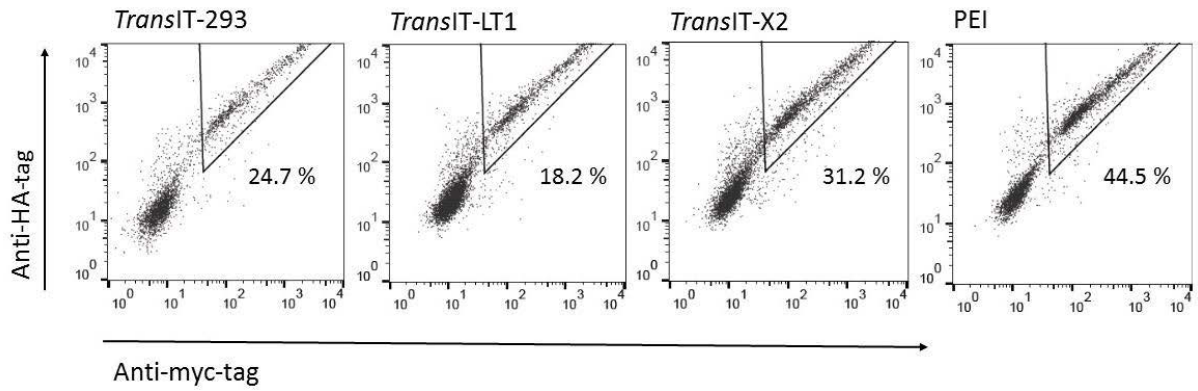


Figure 13: Comparing the number of transfected cells as a function of the transfection reagent used. HEK293T cells were transfected with 1 µg/ml of EGFR-plasmid using a set of 4 distinct transfection reagents. 48 h after transfection cells were dissociated (in PBS), fixed and permeabilized using methanol and stained with an anti-myc (AF 488) as well as an anti-HA antibody (AF 647). The data were analysed using the FlowJo software. One representative example of two independent experiments is shown.

The transfection reagents *Trans*IT-293, *Trans*IT-LT1 and *Trans*IT-X2 were purchased from Mirus Bio, thus their composition is not known.

The results demonstrate that the transfection efficiency greatly varies with the transfection reagent used. The use of PEI as a transfection reagent yielded the highest percentage of EGFR-positive cells accounting for 44.5% of the overall cell population. *Trans*IT-LT1 presented the lowest percentage of transfected cells (18.2%), followed by the transfection reagent *Trans*IT-293, which had been initially used to evaluate the success of the transfection with the EGFR-plasmid as well as establish a fixation/permeabilization protocol. *Trans*IT-293 yielded 24.7% transfected cells. The use of *Trans*IT-X2 delivered 31.2% transfected cells, therefore providing the highest number of transfected cells after PEI. However, independent of the transfection reagent that was used, hardly any "single positive" cells were detected, indicating that, if the transfection is successful, a large number of plasmids enters the cell.

Moreover, the dot plots reveal that the transfection seemed to deliver two distinct populations of cells, represented as two peaks within the transfected cell population in Figure 12A or as two "clouds" of EGFR-positive cells in Figure 13. This effect though becomes less pronounced when the transfection efficiency increases and is therefore visible the best for the results obtained using the transfection reagents *Trans*IT-293 and *Trans*IT-LT1.

4.3.2 Increasing the ratio of single positive cells by addition of carrier plasmid

The use of carrier plasmid has been described in literature and is meant to facilitate the uptake of especially small amounts of plasmid or keep the plasmid concentration constant.
The titration of EGFR-plasmid started with the plasmid amount recommended by the manufacturer (1 µg/ml) and was decreased in 10-fold dilutions (Figure 14).

The transfection with 1 µg/ml plasmid yielded a high number of double positive transfectants carrying a presumably large number of EGFR-plasmids. With a 10-fold decrease in plasmid concentration (100 ng/ml) the number of double positive cells dropped and the ratio of double positive to single positive cells (cells only showing binding on the x- or y-axis, therefore only expressing one type of tag) slightly improved. At a plasmid concentration of 10 ng/ml (without any carrier plasmid added) the number of transfected cells dropped dramatically (to 0.7%).

In contrast, the addition of a carrier plasmid at a concentration of 1 µg/ml still yielded a high number of positive cells when the EGFR-plasmid concentration was decreased to 10 ng/ml, revealing the impact of the addition of a carrier plasmid on the formation of complexes consisting of plasmid DNA and transfection reagent. The ratio of double positive to single positive cells obtained after using a concentration of 10 ng/ml EGFR-plasmid could be

further improved by reducing the amount of EGFR-plasmid. The use of an EGFR-plasmid concentration of 1 ng/ml yielded about 6% single positive cells (4% myc- and 2% HA-positive), whereas only 0.86% of cells were double positive, indicating that the transfected cells mainly carry one or a few plasmids (Figure 14).

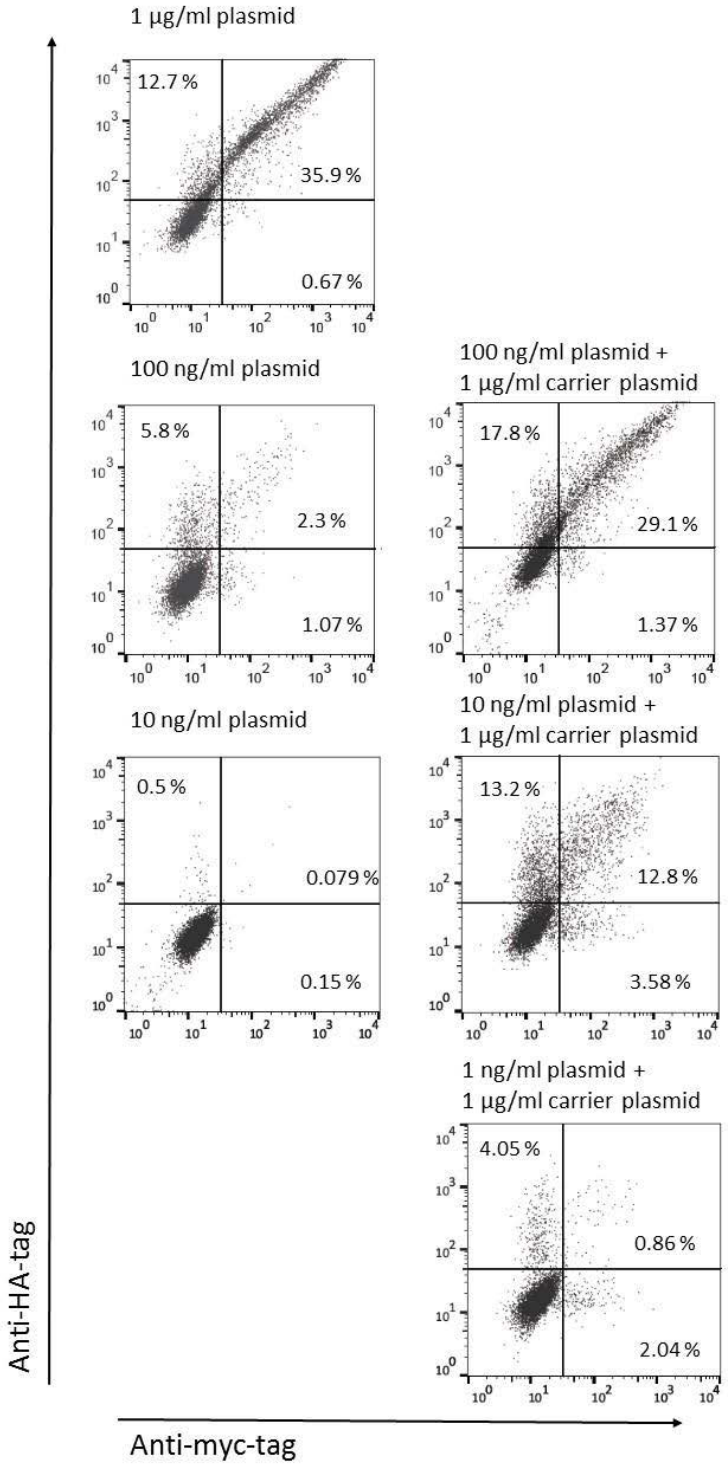


Figure 14: HEK293T cells were transfected with varying amounts of EGFR-plasmid (as indicated) using the transfection reagents TransIT-X2. To some samples 1 µg/ml carrier plasmid was added as indicated. Cells were stained as described in Figure 13. The data were analysed using the FlowJo software and the gates were set on the basis of the main population of the EGFR-negative cells.

4.3.3 Establishing an optimal transfection procedure using *Trans*IT-X2 or PEI

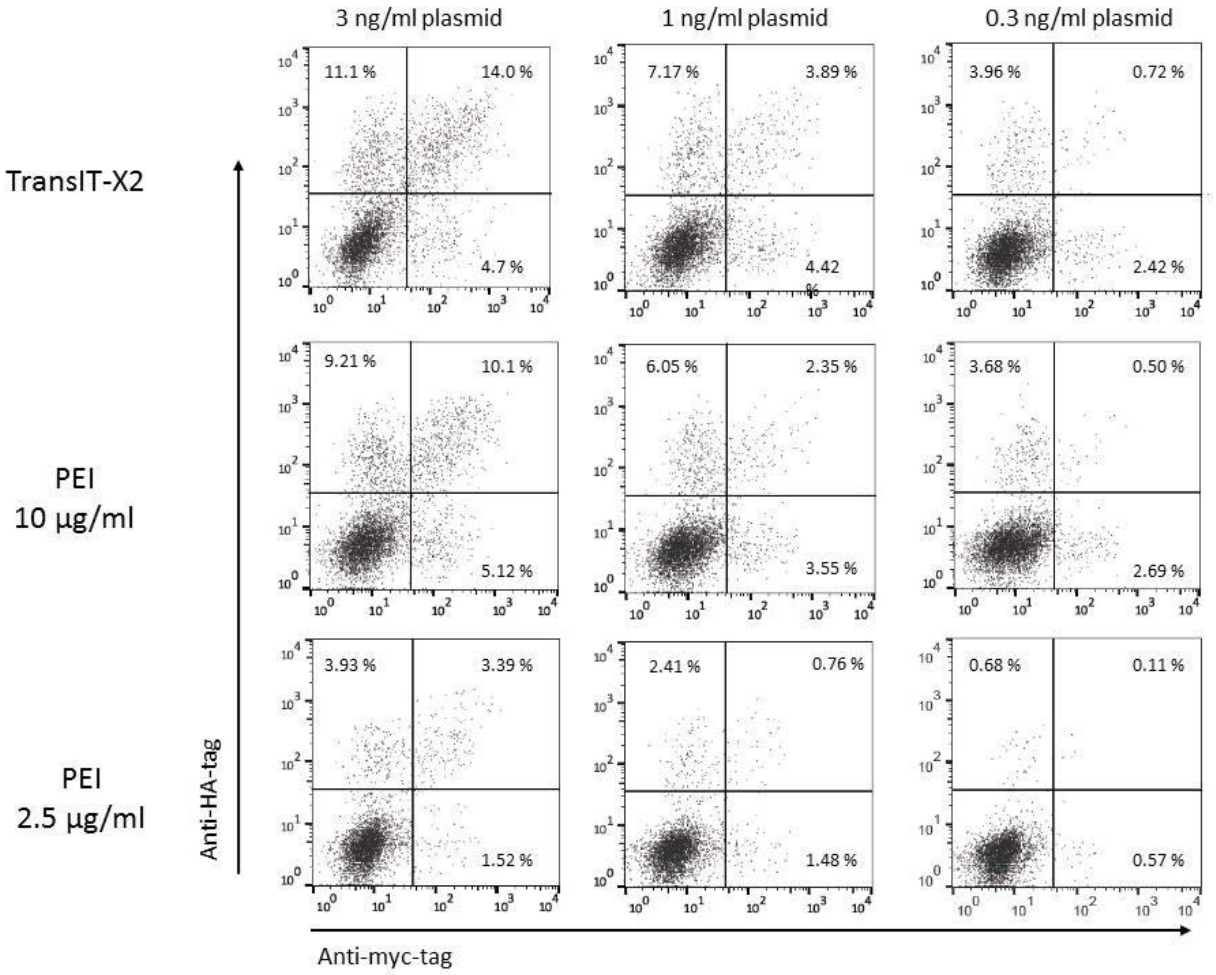


Figure 15: HEK293T cells were transfected with varying amounts of EGFR-plasmid; the actual amount is indicated on top of each column. The addition of 1 µg/ml carrier plasmid is true for all samples. Each row displays a different transfection reagent or dilutions of the same. The cells were stained as described in Figure 13. One representative example of two independent experiments is shown here.

As *Trans*IT-X2 and PEI had been shown to yield superior transfection efficiencies (Figure 13), further optimization experiments were carried out with these two transfection reagents. PEI was originally used at a concentration of 10 µg/ml for transfection and had been shown to have a negative impact on the cell viability. Microscopic inspection revealed that the majority of cells had dissociated from the transfection vessel and was washed off before antibody staining and detection. Therefore dilutions of PEI were used to minimize cell toxicity while maintaining a high transfection efficiency.

Microscopic inspection revealed that cell viability strongly increased when PEI was used at a concentration of 2.5 µg/ml. Though, simultaneously, transfection efficiency was shown to decrease at constant EGFR-plasmid concentrations.

On the basis of the data shown in Figure 15, TransI-X2 used at an EGFR-plasmid concentration of 0.3 ng/ml was established as the optimal transfection protocol, exhibiting little

to no cell toxicity (compared to PEI), yielding a high transfection efficiency and mostly single positive cells.

4.3.4 Testing the fixation and permeabilization with various reagents after transfection with *Trans*IT-X2

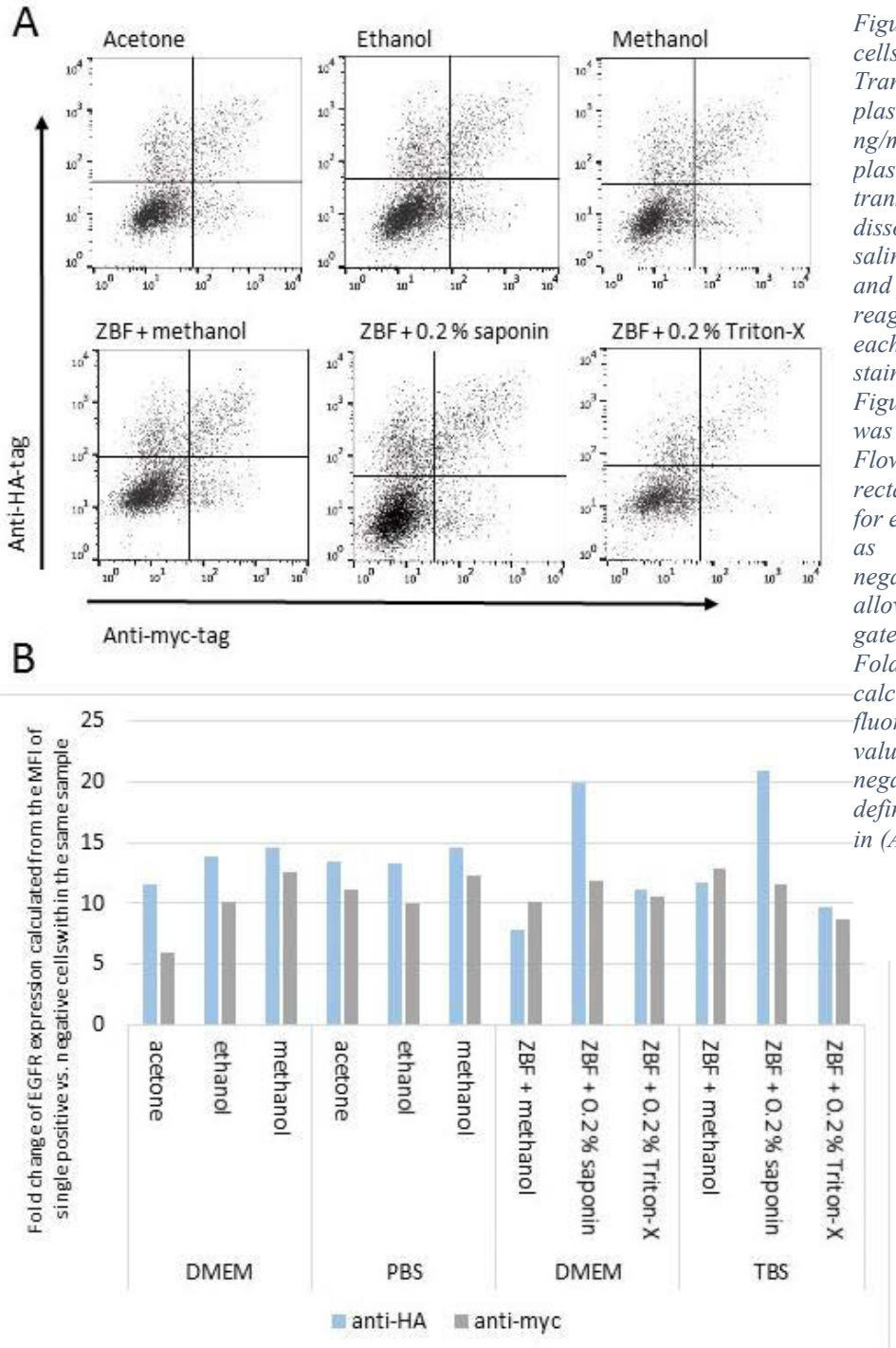


Figure 16 : (A) HEK293T cells were transfected with TransIT-X2 and an EGFR-plasmid concentration of 1 ng/ml (+ 1 µg/ml carrier plasmid). 48 hours after transfection the cells were dissociated in buffered saline (PBS or TBS), fixed and permeabilized using the reagents indicated on top of each dot plot. The cells were stained as described in Figure 13. Data analysis was carried out using the FlowJo software and rectangular gates were set for each sample individually as the shifting of the negative population did not allow the adoption of one gate for each sample. (B) Fold change values calculated from the median fluorescence intensity (MFI) values of single positive vs. negative populations defined by the gates shown in (A).

Having established the transfection protocol yielding mainly single positive cells led to re-evaluation of the six potential fixatives in combination with the desired plasmid dilution.

As expected, all samples yielded a similar percentage of EGFR-positive (i.e. transfected) cells, though some differences in the fold change of EGFR expression calculated from the MFI of single positive cells vs. negative cells was observed. Moreover, a varying distribution of the populations (represented in the CV value) as a function of the fixative could confirm previous findings (4.2.3).

The use of ZBF in combination with 0.2% saponin as a permeabilizing agent had already been shown to yield a high fold change for the signal intensity obtained for binding to both protein tags (Figure 12). The same could be shown for this experiment, utilizing a lower plasmid concentration thus enabling the transfection of single positive cells. A fold change of 20 was received for anti-HA binding (reproducible for both media used), while a fold change of 12 was shown for anti-myc-tag binding (Figure 16B). In general, the fixation and permeabilization using ZBF and saponin could be shown to give constant results irrespective of the amount of plasmid used for transfection. Though, the CV value calculated for the distribution of the negative population (data not shown) is especially large for this fixation/permeabilization method and can be observed as a particularly wide cell population after saponin permeabilization (Figure 16A).

In contrast to saponin-mediated permeabilization, the use of ZBF in combination with 0.2% Triton-X did not yield a fair separation of positive cells from the negative population, reflected by the low fold change changes values (Figure 16B) and also visible in the dot plot in Figure 16A.

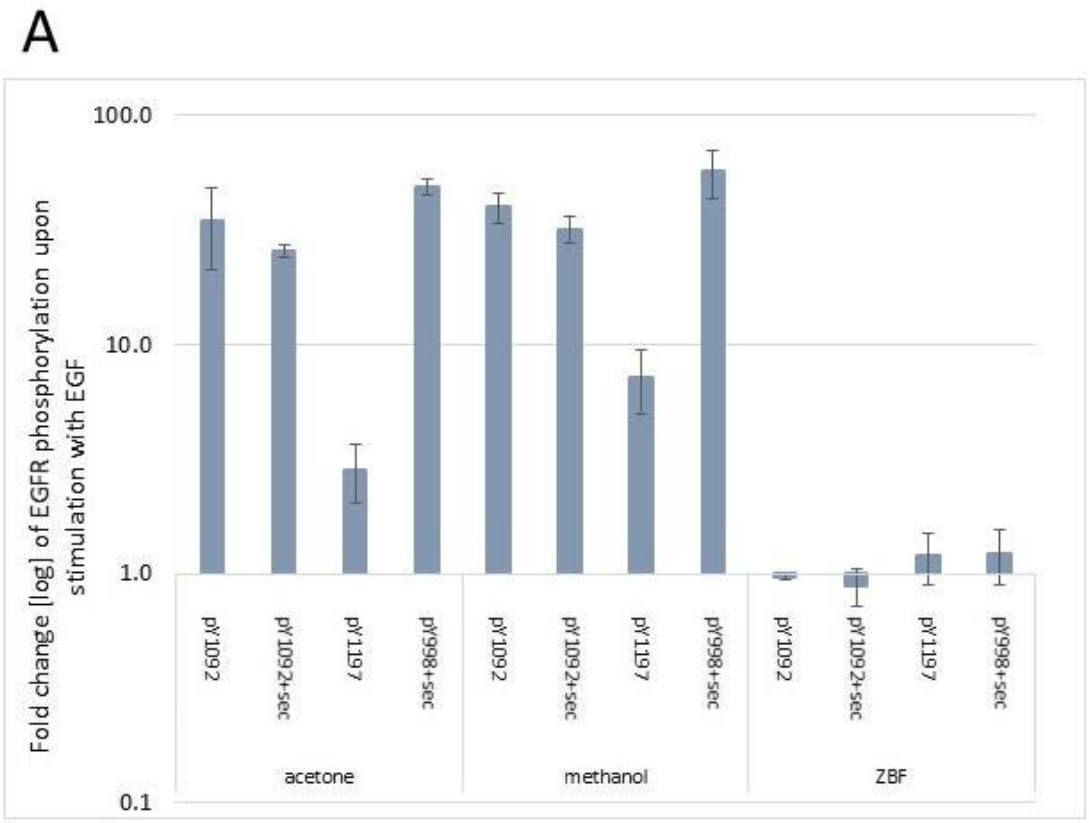
The use of methanol, acetone and ethanol provided a compact negative population (Figure 16A) and a good separation of the populations represented in mediocre to high fold change values for both antibodies (Figure 16B).

Based on this information as well as the morphological features (Figure 9) associated with each of this fixation/permeabilization procedures the use of acetone, methanol as well as ZBF + 0.2% saponin was chosen for further investigation concerning the detection of phosphorylated EGFR. ZBF in combination with saponin was shown to yield a superior separation of single positive cells from the main population, and presented an alternative, reportedly DNA-friendly[18], fixative and was therefore chosen for further investigation. Methanol was chosen for further investigation as it presented reliable results and because it is known for its use for the fixation and permeabilization of cells in the scientific literature[19]. Further, long-term storage in methanol was shown to be possible at -80 °C[19] providing the

scientist with tremendous flexibility. As methanol had been shown to serve the purpose of fixation well, the use of ZBF in combination with methanol was decided to be redundant. Ethanol was excluded because of its similar chemical properties to methanol and slightly inferior results. Due to its harsh mode of action on mammalian cells and the resulting loss of cell integrity (Figure 9), Triton-X was excluded from further analysis. Last, acetone was chosen for further investigation due to its (thought to be) different mode of action compared to methanol and ZBF and because of its satisfying performance in the experiments shown.

4.4 Detection of EGFR phosphorylation

A panel of four distinct antibodies was used for the investigation of EGFR phosphorylation at three phosphorylation positions (Y1197, Y1092 and Y998). Each of these antibodies was tested in combination with methanol, acetone or ZBF + 0.2% saponin in either DMEM or buffered saline (PBS/TBS) as a medium. In order to be able to normalize the phosphorylation signal for EGFR expression, an EGFR-plasmid carrying a myc-tag was used and EGFR expression was detected with an anti-myc-tag antibody.



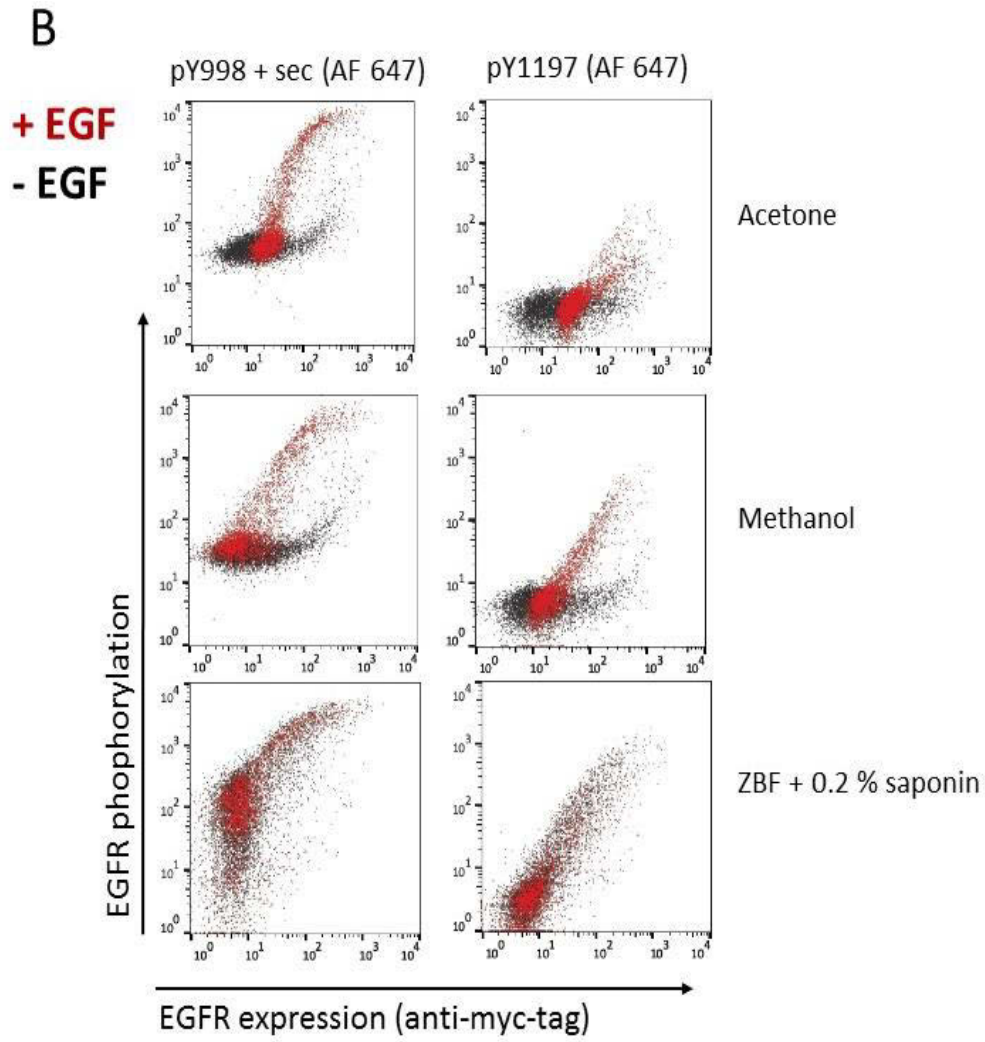


Figure 17: HEK293T cells were transfected with 0.3 ng/ml EGFR-plasmid (+ 1 µg/ml carrier plasmid) using TransIT-X2. 36 hours after the transfection the medium was changed and serum starvation was started. 48 hours after the transfection the cells were dissociated in buffered saline (PBS/TBS) and stimulated with 100 ng/ml EGF for 5 minutes. After the stimulation the cells were immediately fixed (and permeabilized) with methanol, acetone or ZBF as indicated. Cells were stained with an anti-myc-tag antibody (AF 488) as well as one of the antibodies specific for phosphorylated sites on EGFR as indicated in the graph. Unlabelled antibodies were detected with a secondary antibody (goat anti-rabbit). (A) For comparison between the samples, the top 15% of myc-positive cells were gated and the signal intensity for EGFR phosphorylation was compared between the stimulated (+EGF) and unstimulated (-EGF) cells yielding the fold change. Data represent averages ± standard deviations from three independent experiments. (B) Dot plots from one data set are shown representing the signal for EGFR phosphorylation detected with two distinct antibodies. Only the EGFR expressing sub-population (10-15%) should be compared with regard to EGFR phosphorylation.

For the analysis of these data it is very important only to compare the sub-population of cells, which had successfully been transfected with EGFR-plasmid (i.e. myc-positive cells), with regard to EGFR phosphorylation. The main population of HEK293T cells does not express EGFR and thus will not show a change in phosphorylation signal upon the addition of EGF (Figure 17B).

The fold change calculated from the difference in signal intensity for EGFR phosphorylation between EGF-stimulated and unstimulated samples showed to be highly dependent on the fixative as well as the phospho-EGFR-specific antibody used. This

experiment was carried out using both DMEM and buffered saline (PBS or TBS) for the detachment of cells. Since both PBS/TBS and DMEM yielded similar results, buffered saline was chosen to be used in future experiments providing a more defined biochemical system compared to DMEM

Surprisingly, the use of ZBF as a fixative yielded little to no change in phosphorylation signal upon EGF activation (Figure 17A), independent of the antibody (position) used. Figure 17B nicely illustrates the reason for the minor fold change detected when ZBF was applied as a fixative. In fact, the unstimulated samples showed an EGFR level comparable to that of the EGF-stimulated cells, suggesting that this fixation reagent induces crosslinking and thereby activation of EGFR. Therefore, the use of ZBF as a fixative was excluded from further experiments.

The use of methanol and acetone showed comparable outcomes, mostly depending on the antibodies used. The use of the primary antibody recognizing phosphorylated tyrosine 998 (pY998), which is detected through a secondary antibody (Alexa Fluor 647) showed to deliver the best results for both methanol as well as acetone providing a 49-fold or 56-fold change in signal intensity, respectively. The use of the antibody pY1092 PE (a directly labelled antibody) and the same antibody pY1092 unlabelled and detected by a secondary antibody (Alexa Fluor 647) were shown to provide comparable results irrespective of their fluorophore or mode of detection (direct labelling versus indirect staining using a secondary antibody). The use of antibody pY1197 (Alexa Fluor 647) on the other hand was found to be inferior, yielding a fold change of 7.2 using methanol and 2.8 using acetone.

Figure 17B shows the dot plots obtained for the use of the phosphospecific antibodies yielding the largest (pY998 + sec) and lowest fold change (pY1197), according to Figure 17A. The use of acetone as a fixative and permeabilizing reagent was shown to lead to a significant shift in the position of the negative population. This shift was observed for every phosphospecific antibody used (data not shown). Though, using methanol as a fixative, this effect was only visible if the antibody pY1197 (Alexa Fluor 647) was used (Figure 17B). These data clearly demonstrate the large signal difference of distinct antibodies and the importance of screening several antibodies.

4.5 The impact of the SV40 origin

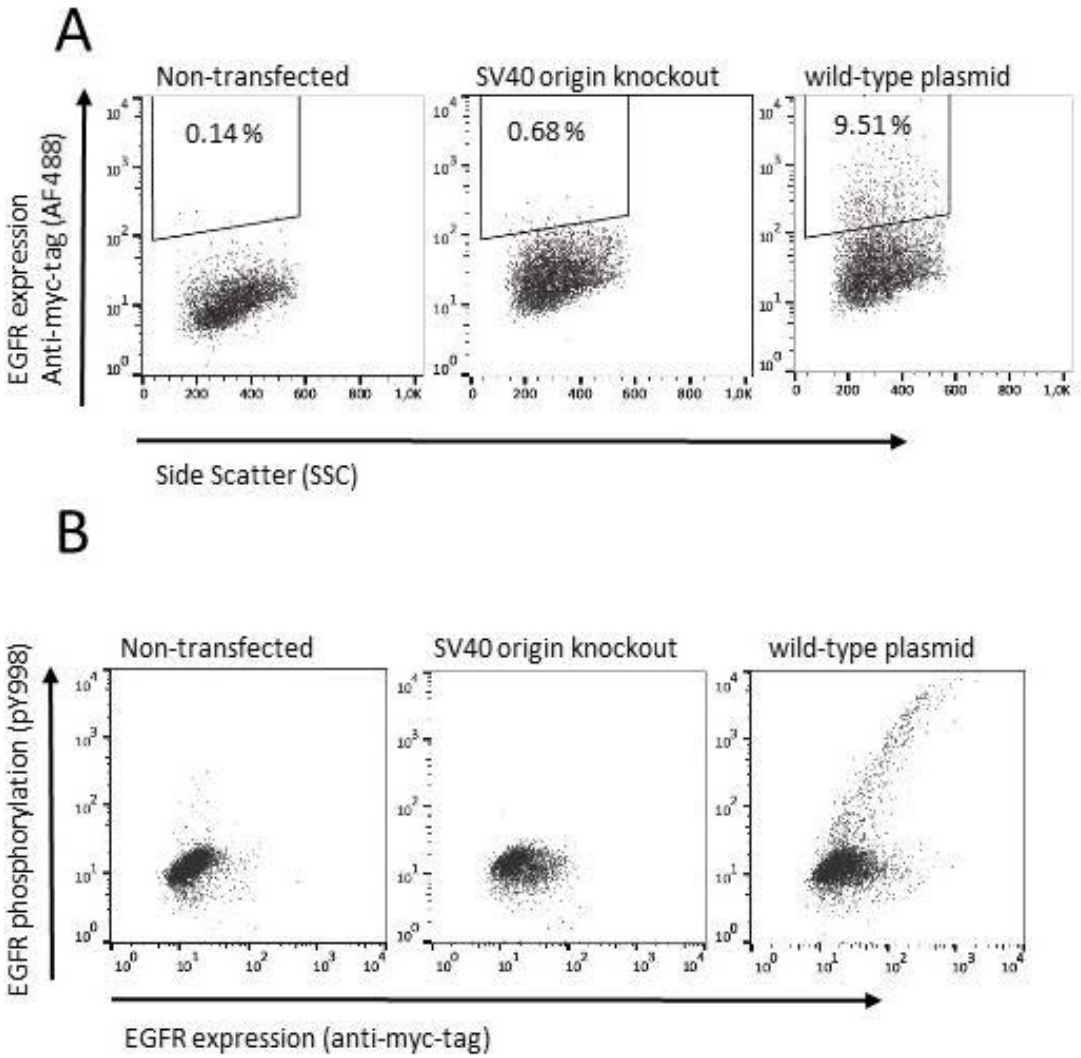


Figure 18: A knockout mutant was created using site-directed mutagenesis (See 3.3). HEK293T cells were transfected with 0.3 ng/ml (+ 1 µg/ml carrier plasmid) of wild-type plasmid still carrying the SV40 origin or the SV40 origin knockout (SV40ko) plasmid using TransIT-X2. 36 h after transfection serum-starvation was started. After 48 h cells were fixed and permeabilized with methanol. (A) Finally, cells were stained with an anti-myc-tag antibody (AF 488). (B) Prior to methanol fixation, cells were activated with 100 ng/ml EGF, and subsequently stained with an anti-myc and an antibody specific for pY998 (AF 647).

The functional efficiency of the SV40 origin was evaluated by establishing a knockout plasmid carrying a deletion in the SV40 origin (SV40ko-plasmid). If the SV40 origin was intact and crucial for episomal replication a difference in EGFR expression would be expected.

In fact, transfection with the SV40ko-plasmid showed an increase of "positive" cells from 0.14% (measured for non-transfected HEK293T cells) to 0.68%. Transfection with the wild-type plasmid, which is the EGFR-plasmid carrying an intact SV40 origin, though showed a steep increase in EGFR expression represented in a percentage of 9.51% myc-positive cells (Figure 18A). Moreover, assessing the phosphorylation intensity as a consequence of the SV40 origin knockout revealed that cells carrying the SV40ko-plasmid show a complete loss of

phosphorylation signal compared to the wild-type plasmid (Figure 18B). This clearly demonstrates that the vast majority of the EGFR phosphorylation signal is dependent on episomal replication. Moreover, these data confirm that episomal replication is faciliated through the SV40 origin and suggest that a certain concentration and closeness of EGFR molecules on the cell surface is crucial for a detectable phosphorylation signal.

4.6 Activating mutations

In order to further validate our detection system, wild-type EGFR (EGFR-WT) was compared with EGFR variants containing activating mutations (EGFR-L858R and EGFR-del19). Figure 19 shows that these activating mutations result in EGFR phosphorylation in the absence of EGF, confirming data known from the literature. Importantly, this experiment clearly shows that cells expressing EGFR carrying activating mutations can be efficiently separated from cells expressing EGFR-WT.

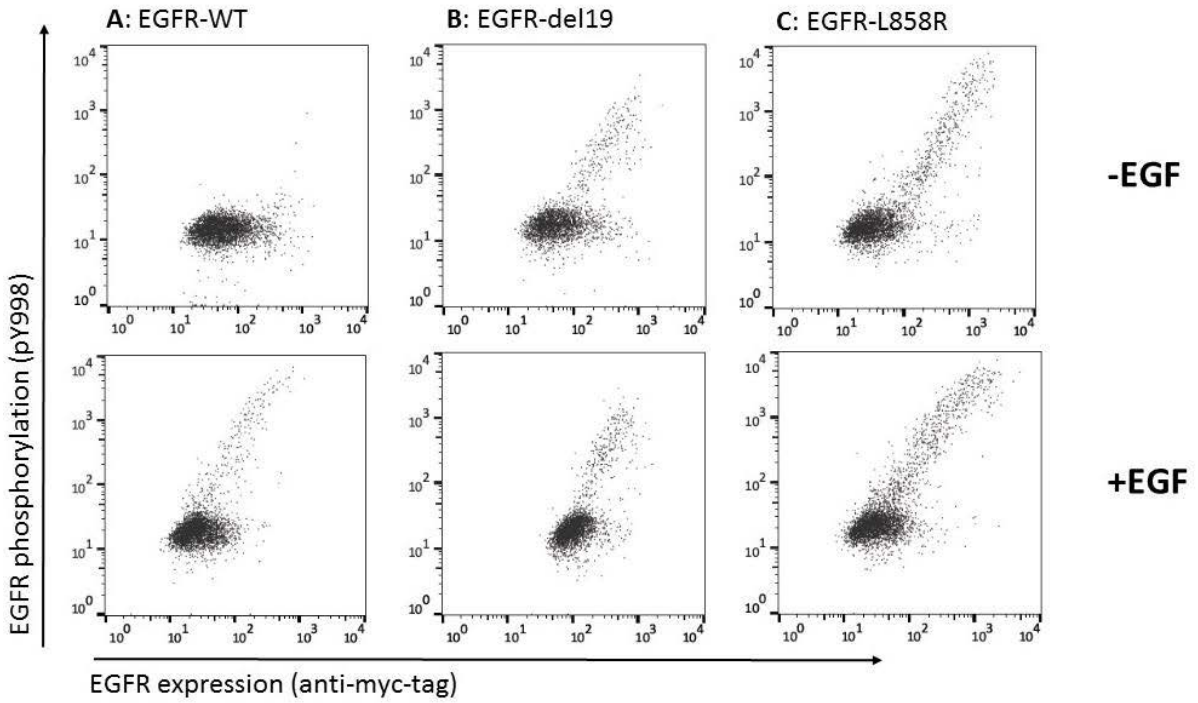


Figure 19: Mutants described and known from the literature, by name the exon 19 deletion and L858R mutation, were established using site-directed mutagenesis. HEK293T cells were transfected with 0.3 ng/ml (+ 1 µg/ml carrier plasmid) of plasmid encoding EGFR-WT, EGFR-L858R or EGFR-del19. Cells were prepared as described in Figure 17 (fixed with methanol). In parallel, samples not receiving EGF as a stimulus were treated with the same amount of PBSA. Finally, the cells were stained with an anti-myc-tag antibody (AF 488) and an antibody specific for pY998, followed by detection with a secondary antibody (AF 647).

4.7 The effect of the addition of cetuximab prior to EGF stimulation on EGFR-WT and EGFR mutants

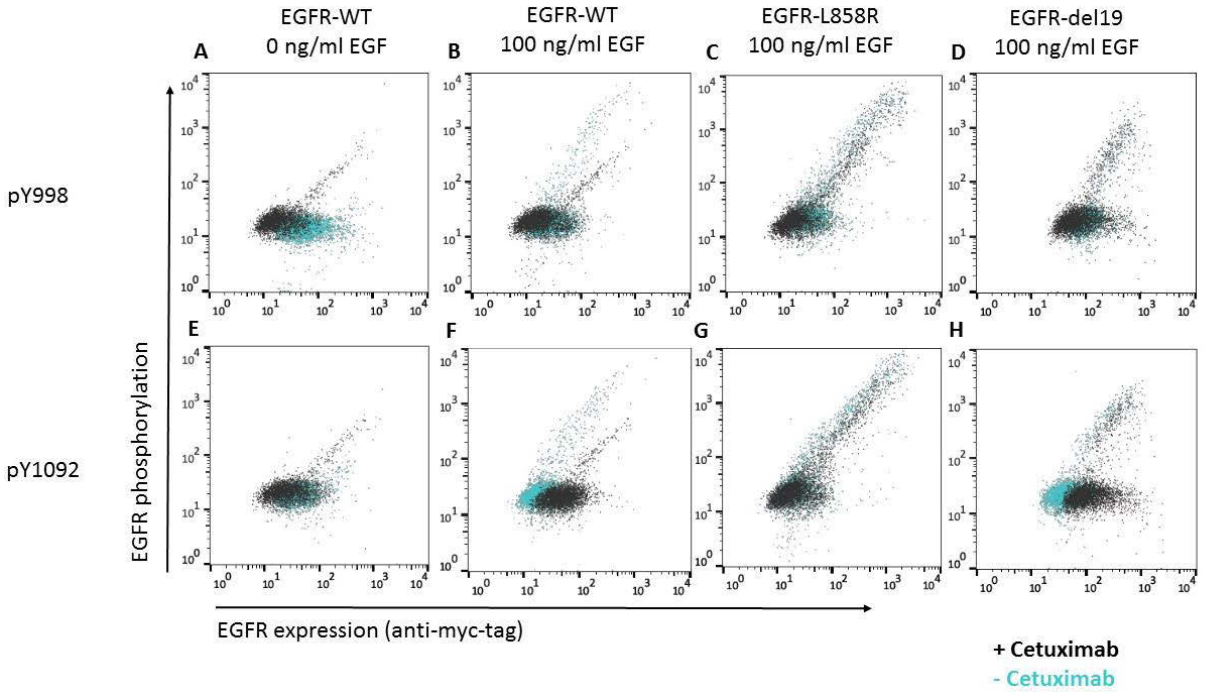
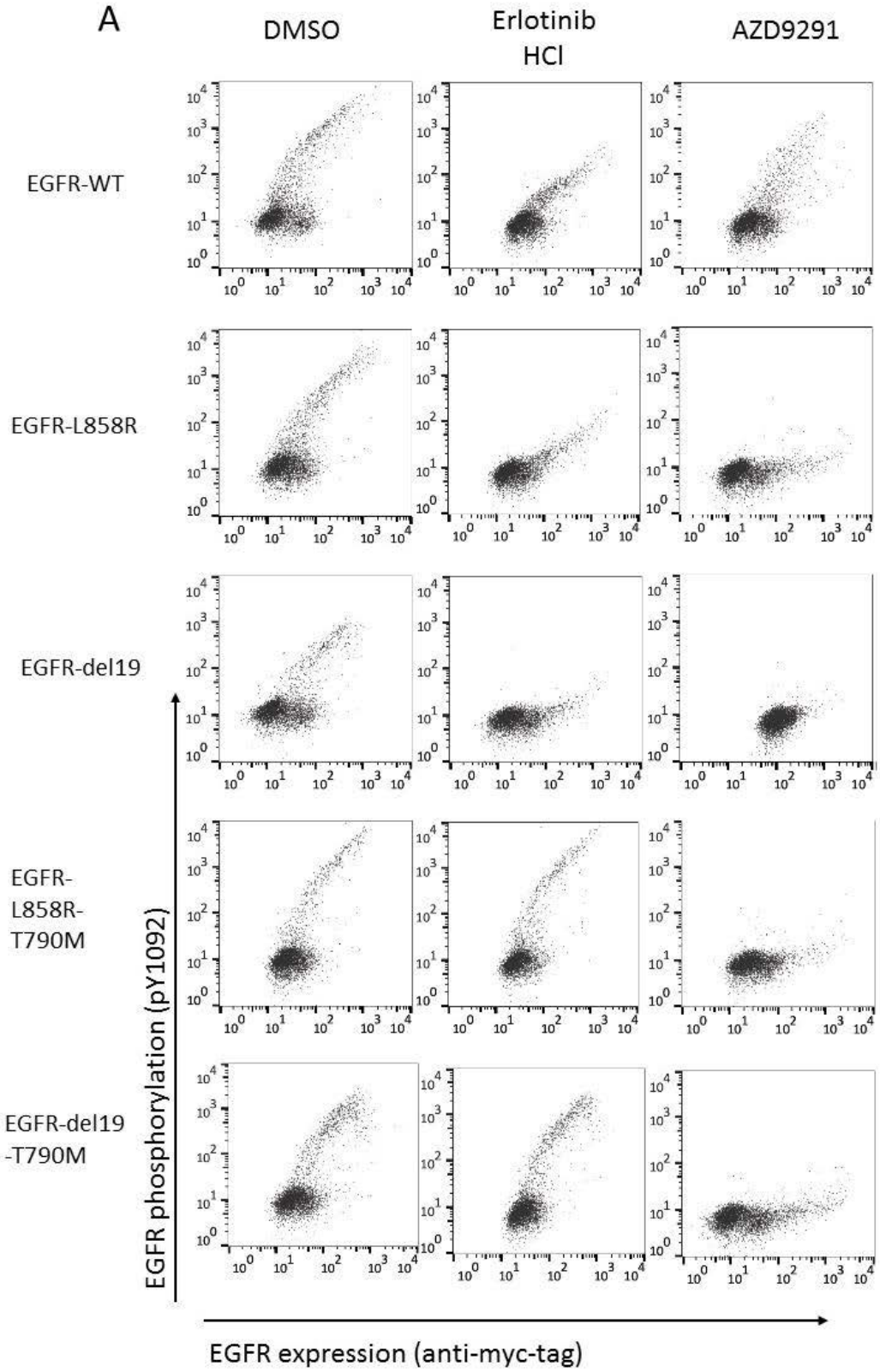


Figure 20: Cells were prepared as described in Figure 19. Some samples were treated with 100 nM cetuximab (gray) prior to EGF stimulation with 100 ng/ml Finally, the cells were stained with an anti-myc-tag antibody (AF 488) and an antibody specific for pY998 (top) or pY1092 (bottom), followed by detection with a secondary antibody (AF 647).

Figure 20 shows the effect of treatment with 100 nM cetuximab prior to EGF stimulation. In this experiment cetuximab was shown to supress the level of EGF activation of the activated EGFR-WT (stimulated with 100 ng/ml EGF) by approximately 3-fold (Figure 20). This result could be reproduced using either of the antibodies specific for pY998 or pY1092, suggesting that the activation was not only limited to one phosphorylation site in EGFR. Surprisingly, the treatment of the unstimulated EGFR-WT with cetuximab seemed to "activate", that is increase, EGFR phosphorylation (Figure 20). In fact, the geometric mean calculated for EGFR phosphorylation of the myc-positive population, could be shown to be nearly equal for both EGFR-WT samples irrespective of being stimulated with EGF or not (gray population in Figure 20 A and E vs. B and F).

Not surprisingly, both mutants, EGFR-L858R and EGFR-del19, showed little to no change in phosphorylation signal upon cetuximab-treatment irrespective of the phosphorylation site which was recognized by the antibody.

4.8 Inhibition with erlotinib and AZD9291



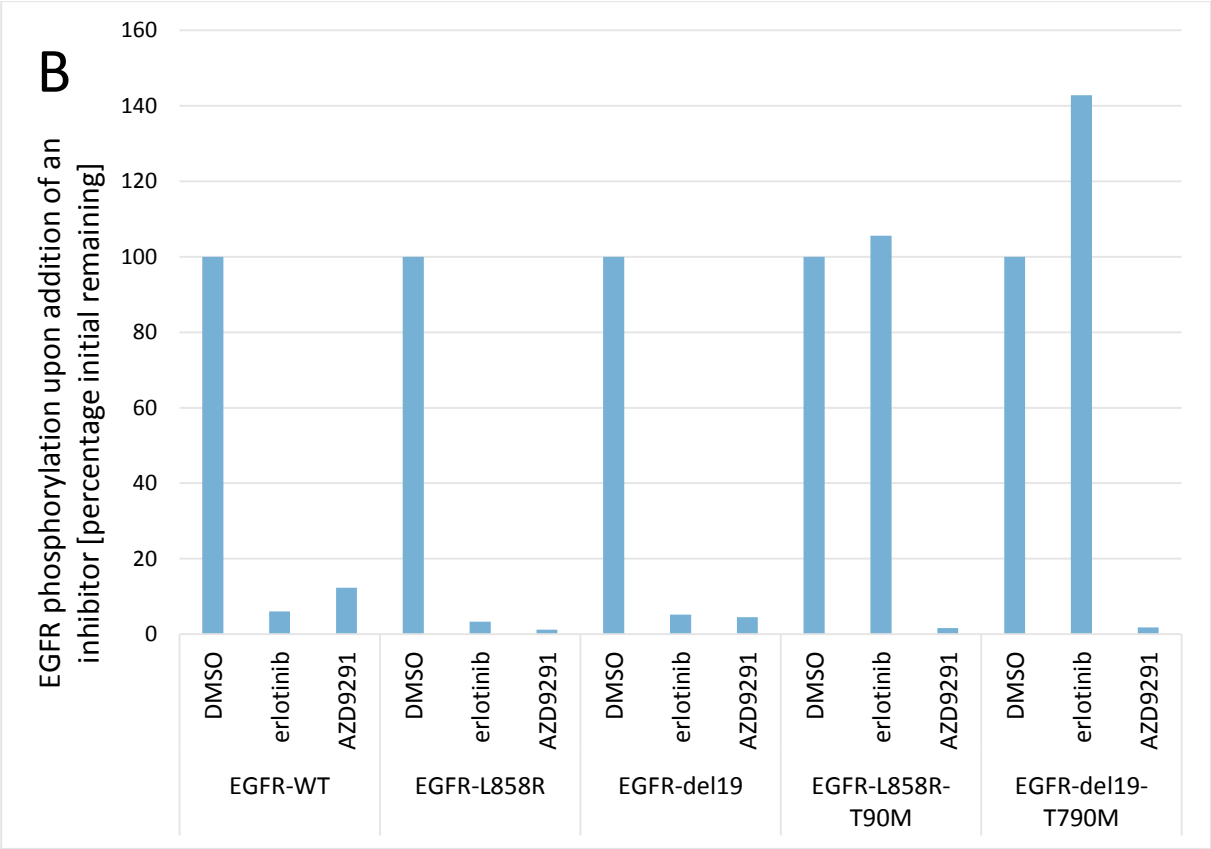


Figure 21: (A) HEK293T cells were transfected with 0.3 ng/ml (+ 1 µg/ml carrier plasmid) plasmids encoding EGFR-WT, EGFR-L858R, EGFR-del19, EGFR-L858R-T790M or EGFR-del19-T790M using the transfection reagent TransIT-X2. 36 hours after the transfection serum starvation was started and the inhibitors erlotinib HCl and AZD9291 (or the same amount of DMSO as a negative control) were adjusted to a final concentration of 500 nM (as indicated on top). All samples shown were stimulated with 100 ng/ml EGF for 5 minutes. After fixation with methanol, cells were stained with an anti-myc-tag antibody (AF 488) and an antibody specific for pY1092, followed by detection with a secondary antibody (AF 647). (B)The geometric mean of EGFR phosphorylation is depicted as the percentage of remaining signal (upon addition of an inhibitor) relative to the negative control (DMSO).

Figure 21A shows the effect of the kinase inhibitors erlotinib HCl and AZD9291 on EGFR phosphorylation on several different EGFR mutants. Cells carrying the EGFR-WT plasmid showed a decrease in EGFR phosphorylation when treated with erlotinib HCl while only showing a minor decrease in phosphorylation signal when subjected to treatment with AZD9291. Cells carrying the EGFR-L858R mutant showed a considerable decrease in phosphorylation signal upon treatment with erlotinib (Figure 21A). An even greater response was observed when the muant EGFR-L858R was treated with AZD9291. Similar results to EGFR-L858R could be observed for cells carrying the EGFR-del19 mutant. The situation for those 2 mutants carrying an additional mutation, T790M, presented itself differently in terms of the response to erlotinib. The T790M mutation has been known to confer resistance towards erlotinib. This result could be reproduced in this assay, therefore verifying the functionality of the assay. Both mutants EGFR-L858R and EGFR-del19 carrying the T790M mutation showed

no response upon addition of erlotinib HCl, but a signficant reduction as a result of treatment with the third-generation TKI AZD9291.

The quantification of data is shown in Figure 21B showing the remaining posphorylation signal upon treatment with a certain inhibitior (erlotinib HCl or AZD9291, as indicated) relative to the same EGFR mutant treated with DMSO as a control. For this purpose 5% of cells showing the highest level of EGFR expression (anti-myc-tag binding) were gated and used for the assessment of the geometric mean of fluorescence intensity.

Cells carrying EGFR-WT showed a reduction in signal to 6% of its original value (DMSO) as a consequence of treatment with erlotinib; a reduction to 12% of its initial signal was observed if AZD9291 was added. A sligthly greater response was observed for the mutant EGFR-L858R, showing a drecrease to 3% and 1% as a result of treatment with erlotinib or AZD9291, respectively. Cells carrying EGFR-del19 showed a remaining signal of 4 and 5% upon treatment with erlotinib and AZD9291, respectively. Cells carrying EGFR, which contained the double mutation L858R and T790M, showed no signal reduction upon addition of erlotinib but a great response (1.6% remaining signal intensity) upon addition of AZD9291. Cells expressing the EGFR mutant carrying the exon 19 deletion and the T790M mutation in addition showed an increase of 43% in signal intensity as a result of treatment with erlotinib HCl, a change which is barely visible on a logarithmic scale. Upon treatment with AZD9291 this mutant showed a decrease in signal intensity to 1.7% of its initial value.

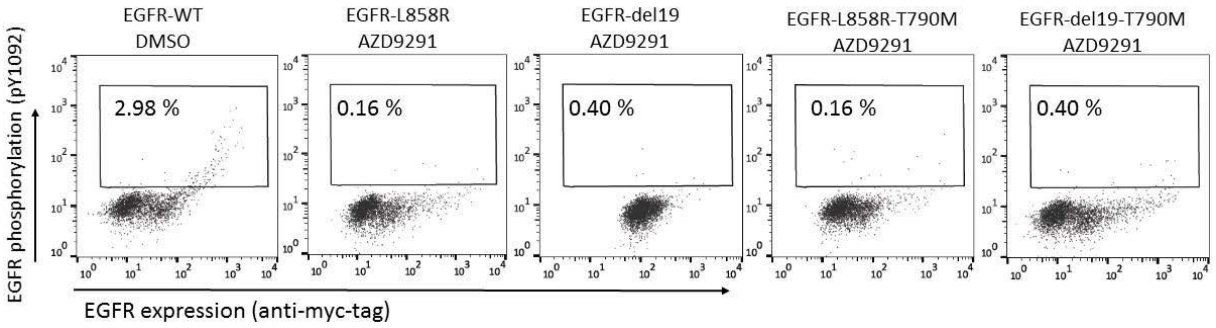


Figure 22: HEK293T cells were prepared as described in Figure 21. Here, the phosphorylation signal of unstimulated EGFR-WT is compared to the signal of EGFR-L858R, EGFR-del19, EGFR-L858R/T790M and EGFR-del19/T790M, which had been treated with AZD9291.

Interestingly, we observed that after treatment with EGF and AZD9291 the phosphorylation signal of EGFR mutants is lower compared to untreated EGFR-WT (i.e. no EGF and no inhibitor added). The percentage of cells excessing a certain phosphorylation signal was reduced upon addition of AZD9291 when compared to the signal exhibited by unstimulated EGFR-WT, suggesting that a "background" phosphorylation exists for the samples which are not treated with EGF (probably induced through traces of FBS or ligand independent

dimerization of EGFR-WT), which is removed upon addition of a TKI. Thus, this background signal observed for unstimulated EGFR-WT is not due to non-specific binding of the phospho-EGFR-specific antibody, but due to real background phosphorylation of EGFR-WT (Figure 22).

4.9 Generation of a randomly mutated library

4.9.1 Randomization of the full-length EGFR gene

In order to construct randomly mutated libraries based on EGFR-WT and EGFR-L858R, respectively, the full-length EGFR gene was amplified by error-prone PCR, followed by further PCR amplification and ligation into the vector. Several experimental set-ups were tested, which failed at generating randomized full-length EGFR-clones. The majority of plasmids isolated from individual *E. coli* colonies after library transformation only contained N- and C-terminal fragments of the EGFR gene, but were missing the majority of the gene in the middle. Obviously, these fragments still ligated into the vector and therefore produced intact plasmids and viable colonies. Only the reduction of the number of cycles used in the PCR cycling protocol as well as the use of XbaI-linearized plasmid, which was thought to relax the supercoiled structure and ease access of the polymerase, was shown to deliver full-length EGFR inserts after ligation.

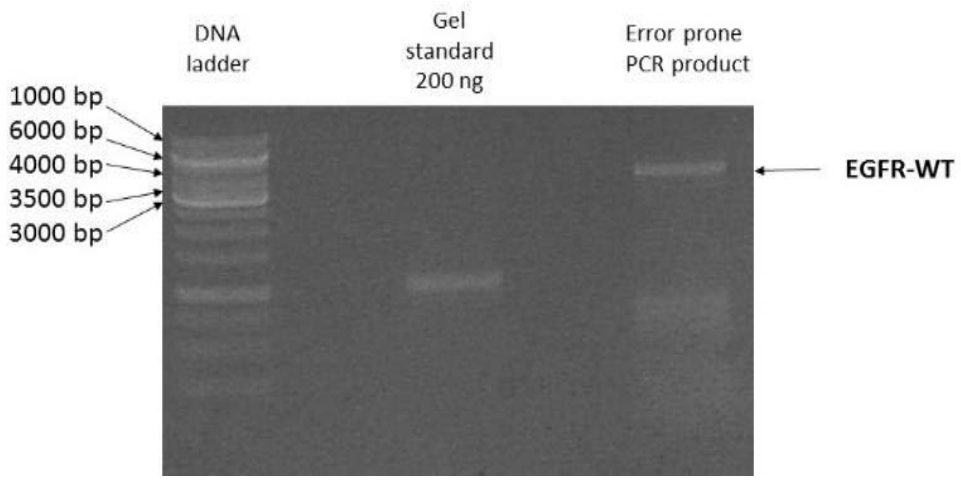


Figure 23: ***Preparative 1% agarose gel after error-prone PCR in order to dispose of short DNA fragments****. XbaI digested DNA template (pSF-CMV-SV40 plasmid containing the EGFR gene) was used for error-prone PCR using the GeneMorph Random Mutagenesis kit using 1000 ng template per reaction (3.6). The PCR reaction was purified and subsequently DpnI digested. A preparative gel was performed to dispose of smaller DNA fragments.*

Figure 23 shows the preparative gel used for the disposal of DNA fragments after error-prone PCR and subsequent digestion with DpnI. The band visible around 4 kb represents the

EGFR fragment after error-prone PCR. In the construction of the libraries based on the EGFR-WT or EGFR-L858R gene gel electrophoresis was performed in separate chambers to avoid cross-contamination and carryover. A gel standard (at 1.1 kb) was used to assess the amount of DNA produced through error-prone PCR. As the amount of product compared to the amount of template enables estimation of the mutation rate, this ratio was initially used to make an estimate on the percentage of mutations. The PCR product obtained from error-prone PCR was extracted and used for subsequent amplification and cloning steps.

4.9.2 Library size and mutation rate

After ligation of the randomly mutated fragments into the vector, *E. coli* was transformed and plasmids from individual colonies were sequenced. The percentage of colonies carrying the closed pSF-CMV-SV40 plasmid without any insert was calculated to be 5.6%. This ratio is already subtracted from the library size thus representing only the true value of randomized, distinct clones.

The library based on the EGFR-plasmid carrying the L858R mutation was calculated to be $1.55*10^6$ whereas the library based on the EGFR-WT plasmid yielded $7.63*10^5$ clones. The EGFR gene holds 3630 basepairs, resulting in a theoretical diversity of $1.01*10^4$ mutants which can be obtained with a single nucleotide change. The EGFR-L858R- and EGFR-WT-library therefore display 142- and 70-fold coverage, respectively.

*Table 28: **Number of clones exhibiting a certain number of mutations**. The cloning and transformation procedure was performed as described in 3.6. 10 single colonies (from each library) were analysed by sequencing. This table shows the distribution of DNA mutations in the 10 analysed genes originating from each of the two libraries.*

Number of DNA-point mutations per EGFR gene	EGFR-WT library	EGFR-L858R library
4 mutations	1	0
3 mutations	1	1
2 mutations	0	0
1 mutation	3	5
0 mutation	4	4
insertion / deletion	1	0

On average, the EGFR-WT-library was shown to carry 1.0 DNA-point mutations/EGFR-gene whereas the EGFR-L858R-library was shown to carry 0.8 DNA-point mutations/EGFR-gene. As the same amount of template was used for error-prone PCR the mutation rate had been expected to be similar between the two libraries; this anticipation could be shown to be true. A considerable number of mutants (5/10 and 4/10 for the EGFR-WT- or EGFR-L858R-library, respectively) was shown not to carry any mutations (or to carry an

insertion or deletion). As the number of mutations rises, the amount of carryover (due to several mutations present in the same gene) as well the probability of the presence of a deleterious mutations on the same gene as the mutation of interest will also increase. Therefore, the number of clones not carrying a mutation was accepted in order to keep the mutation rate low. However, as a consequence, the effective library size (i.e. the number of EGFR mutants) is reduced by approximately 50% for both libraries.

4.10 Testing EGFR phosphorylation in randomly mutated libraries

In order to investigate the phenotype of EGFR mutants in the randomly mutated libraries, cells were transfected with library plasmids (based on EGFR-WT or EGFR-L858R, respectively) and subsequently compared to cells transfected with non-mutagenized EGFR. As the generation of a library was thought to introduce heterogeneity, that is EGFR mutants expressed on the cell surface of cells, a change in the phosphorylation signal was anticipated to be observed upon the introduction of mutant EGFR into HEK293T cells.

4.10.1 Evaluating the effect of the introduction of random mutations into the EGFR-L858R gene on EGFR phosphorylation

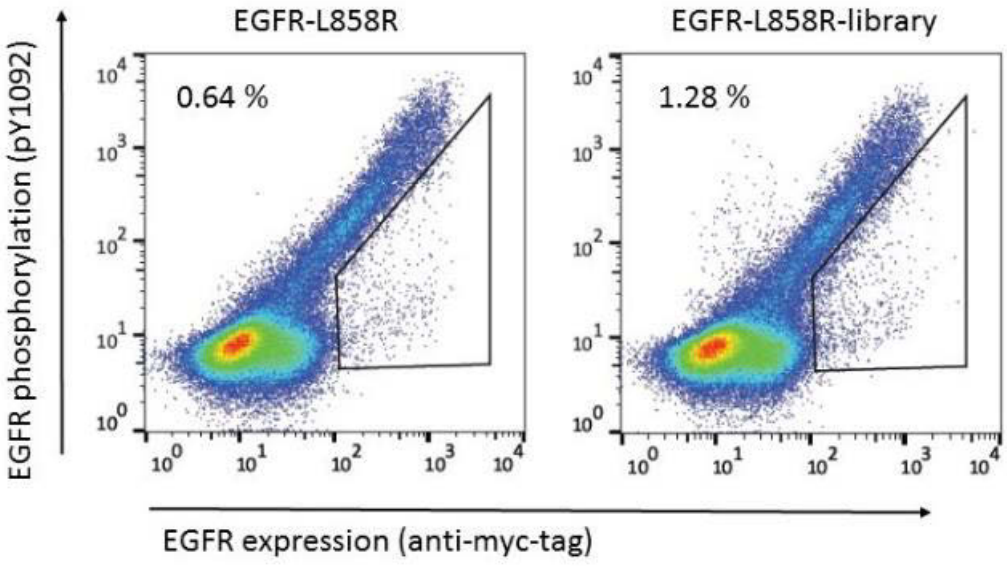


Figure 24: ***Comparison of the EGFR phosphorylation signal of cells expressing the EGFR-L858R mutant versus cells expressing a randomly mutated EGFR-L858R-library****. HEK293T cells were transfected with 0.3 ng/ml of EGFR-L858R plasmids or plasmids derived from the EGFR-L858R-library (+ 1 µg/ml carrier plasmid) using the transfection reagent TransIT-X2. 36 hours after the transfection serum starvation was started. 48 hours after the transfection the cells were dissociated in PBS and fixed/permeabilized using methanol. Finally, cells were stained with an anti-myc-tag antibody (AF 488) and an antibody specific for pY1092, followed by detection with a secondary antibody (AF 647).*

Transfection with a library based on the EGFR-plasmid carrying the L858R mutation was shown to significantly alter the phosphorylation signal of the cells (Figure 24). A significant increase in the myc-positive, but phospho-EGFR-low/negative cell population was observed (1.28% in the library, compared to 0.64% in EGFR-L858R lacking random mutations), indicating that a certain fraction of the EGFR mutants is expressed in full length, but is not functional. The number of transfected cells was shown to amount to the sum of 12.5% of the total population, therefore an increase in 0.64% of knock-down EGFR molecules represents an increase of 5.12% if normalized to the number of transfected cells only. Since each EGFR-gene in the library contains 1 mutation on average (Table 28), this means that about 5% of all EGFR-mutations destroy the enzymatic function of this protein.

Moreover, we also observed a population of cells above the non-transfected population, representing cells that show EGFR phosphorylation but no myc-tag expression. We assume that these cells carry EGFR mutants that had lost their myc-tag due to introduction of a stop codon, but still produce functionally intact EGF-receptors which exhibit a phosphorylation signal.

4.10.2 Evaluating the effect of random mutations in the EGFR-L858R gene on the EGFR phosphorylation after addition of erlotinib

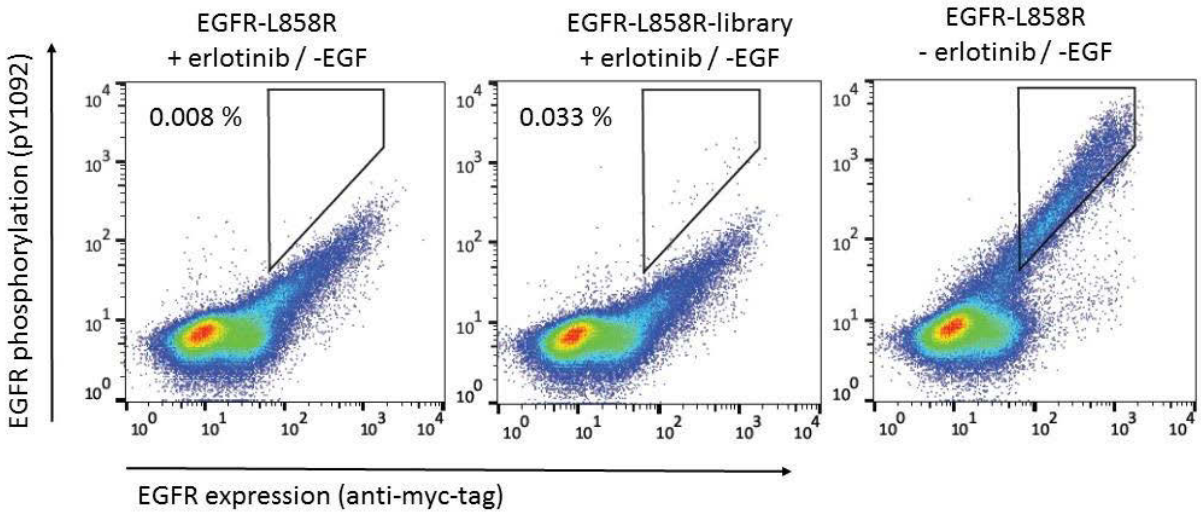


Figure 25: ***Comparison of the EGFR phosphorylation signal of cells expressing the EGFR-L858R mutant versus cells expressing the EGFR-L858R-library after treatment with the inhibitor erlotinib.*** *Cells were prepared as described in Figure 24. When serum starvation was started 500 nM erlotinib was added to the samples as indicated.*

In this experiment cells expressing EGFR-L858R with (EGFR-L858R-library) or without (EGFR-L858R) random mutations were treated with erlotinib and the effect of those random mutations in the library on phosphorylation was evaluated.

As a comparison the L858R mutant exhibiting full activity (no prior addition of erlotinib) is shown (Figure 25, dot plot to the right). The gate set represents the signal intensity usually obtained by the L858R mutant and allows easy comparison if applied to the other dot plots. Upon addition of erlotinib, phosphorylation is strongly reduced (Figure 25, dot plot on the left vs. right). Remarkably, the introduction of random mutations could be shown to lead to an increase in the cell population found in the set gate (0.033% in the library vs. 0.008% for EGFR-L858R). The activity found for those mutants is equivalent to the activity of non-inhibited L858R reinforcing the idea that those mutations conferred resistance to erlotinib.

4.10.3 Evaluating the effect of random mutations in the EGFR-WT gene on EGFR phosphorylation in the presence of EGF

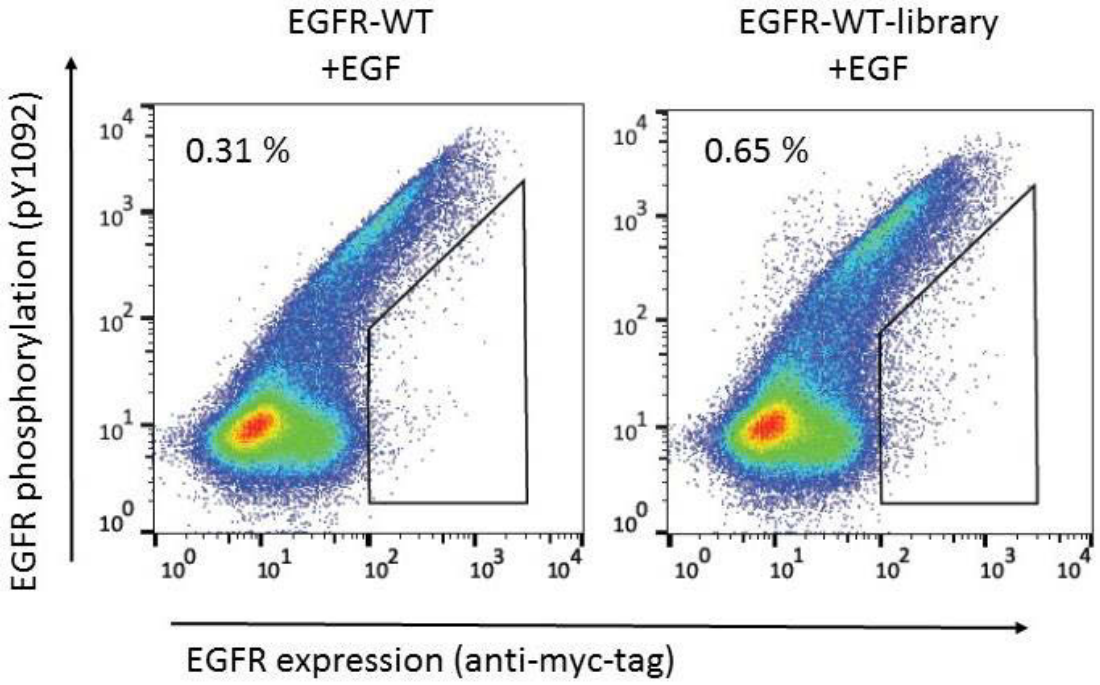


Figure 26: ***Comparison of the EGFR phosphorylation signal of cells expressing EGFR-WT versus cells expressing an EGFR-WT-library after activation through EGF.*** *HEK293T cells were transfected with 0.3 ng/ml of EGFR-WT plasmids or plasmids derived from the EGFR-WT-library (+ 1 µg/ml carrier plasmid) using the transfection reagent TransIT-X2. 36 hours after the transfection serum starvation was started. 48 hours after the transfection the cells were stimulated with 100 ng/ml EGF. After 5 min cells were fixed/permeabilized immediately using methanol. Finally, the cells were stained with an anti-myc-tag antibody (AF 488) and an antibody specific for pY1092, which was subsequently detected with a secondary antibody (AF 647).*

Figure 26 shows the effect of random mutations on the EGFR phosphorylation of EGFR-WT, which had been activated using EGF. As it had already been observed using the L858R-library, the percentage of myc-positive, but phospho-EGFR-low/negative cells, as well as the cell population on top of the non-transfected cells was shown to be increased. Again, the cell

population found on top of the myc-negative population represents cells expressing a functionally intact (therefore phosphorylated) EGFR molecule without any C-terminal myc-tag. The myc-positive, phospho-EGFR-low/negative cell population increased by 2-fold from 0.31 to 0.65% and represents mutants, which are fully expressed (i.e. myc-tag binding) but not functional and therefore not exhibiting EGFR phosphorylation. As already discussed, the number of positive (transfected) cells was determined to be 12.5%, therefore the values shown in Figure 26 (which refer to the total cell population) account for 2.72% if normalized to the number of transfected cells. Thus, through the transfection of the WT-library an increase in cells expressing non-operational EGFR of 2.72% could be shown.

4.10.4 Testing the frequency of activating mutations in a library based on EGFR-WT

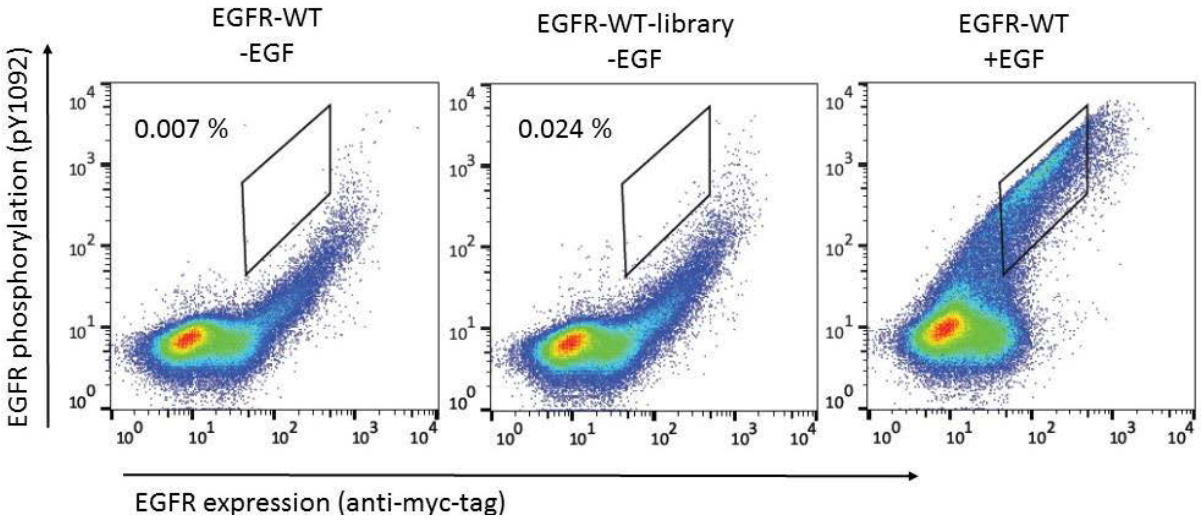


Figure 27: ***Comparison of the EGFR phosphorylation signal of cells expressing EGFR-WT versus cells expressing an EGFR-WT-library without the addition of EGF as a stimulus.*** *Cells were prepared as described in Figure 26 and stimulated with 100 ng/ml EGF as indicated.*

Figure 27 shows potential activating mutations which raise EGFR activity in the absence of the ligand EGF. For comparison, a dot plot showing stimulated EGFR-WT is shown (Figure 27, dot plot to the right). The same gate was applied to all three dot plots presented in Figure 27. Transfection with the EGFR-WT plasmid resulted in a cell count of 0.007% in the set gate, while transfection with the EGFR-WT-library seemed to increase this number by 3-fold (0.024%). The gate was deliberately set in a very restrictive way to avoid the detection of unspecific events. The scattered events observed towards the highly myc-positive end of the transfected cell population were comparable between the samples expressing EGFR-WT or the

corresponding library. The increase in events detected in the set gate indicated that mutations with an activating phenotype are present in the library.

4.10.5 Evaluating the effect of random mutations in the EGFR-WT gene upon blocking the receptor with cetuximab

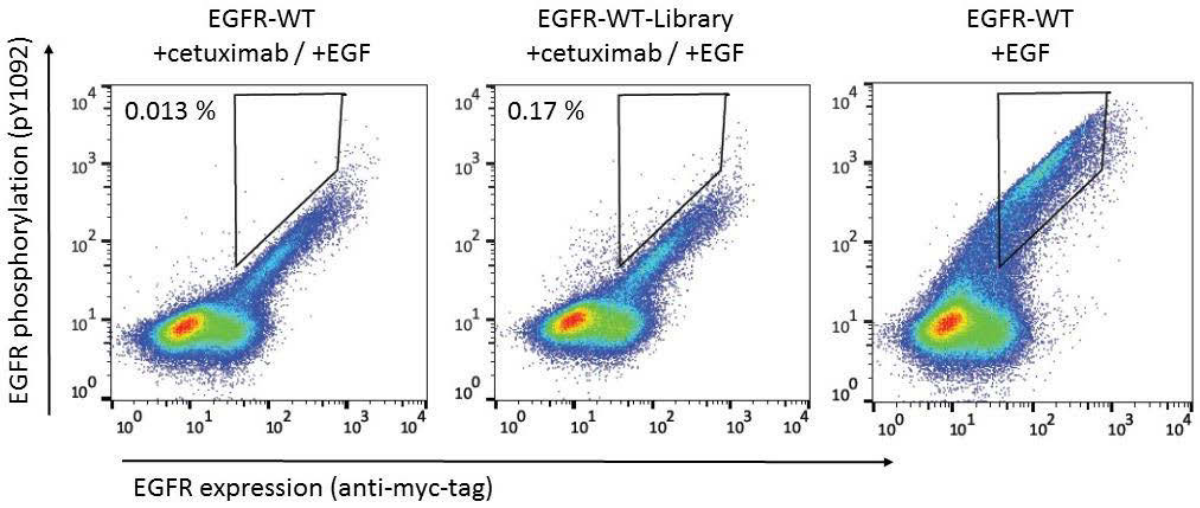


Figure 28: ***Comparison of EGFR phosphorylation of cells expressing EGFR-WT vs. cells expressing a EGFR-WT-library in the presence of both cetuximab and EGF.*** *The cells were prepared as described in Figure 26. Prior to the EGF stimulation cells (as indicated) were incubated with 100 nM cetuximab for 30 minutes.*

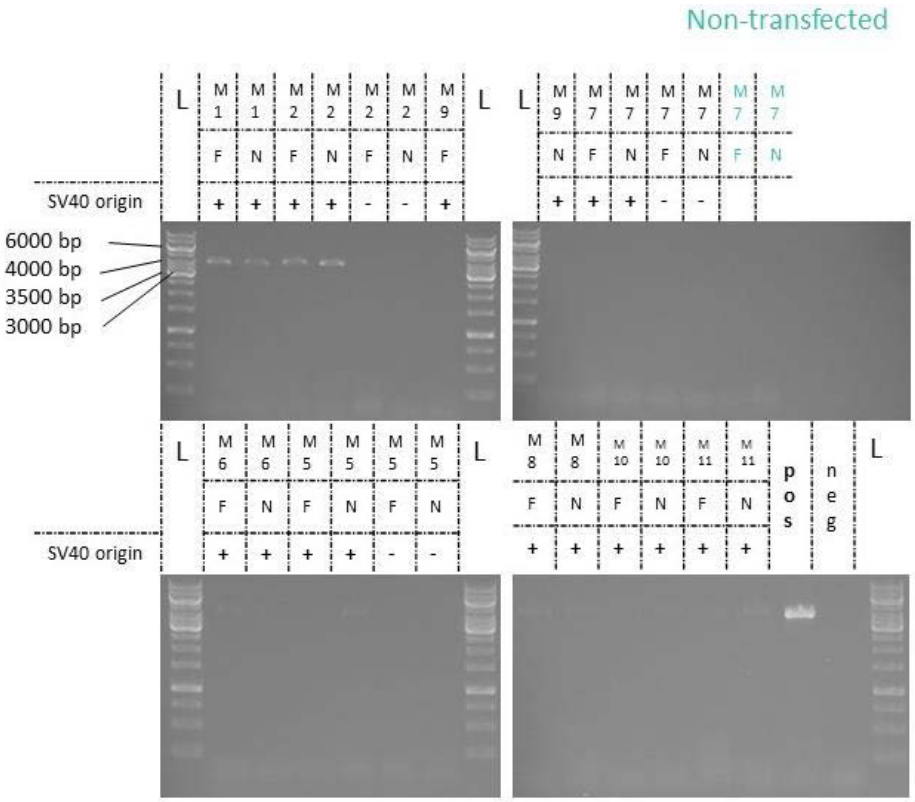
From previous experiments the blocking effect of cetuximab, which results in a decrease in phosphorylation signal, was known. In this experiment the recovery of the phosphorylation signal through the introduction of mutations was evaluated. For comparison, a dot plot showing stimulated EGFR-WT is shown (Figure 28, dot plot to the right).

The addition of cetuximab showed a strong decrease in phosphorylation signal in the presence of the stimulus EGF. Only 0.013% of cells exceeded a certain phosphorylation signal and were found in the set gate. However, transfection with a library based on EGFR-WT showed a more than 10-fold increase in cells present in the gate to 0.17%. Again, as the percentage refers to the total cell population, an effective increase of 1.2% of cells showing unresponsiveness towards cetuximab was found for the transfected cell population.

4.11 Plasmid DNA isolation from methanol-fixed cells

4.11.1 PCR amplification of the full-length EGFR gene

4.11.1.1 Comparison of several distinct DNA extraction methods



M1: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P2 and incubated at 56°C for 10 minutes

M2: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P2 and incubated at 56°C for 60 minutes

M3: QIAprep Spin Miniprep; 50 µl Qiagen Proteinase K was added to buffer P2 and incubated at 56°C for 10 minutes

M4: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P2 and incubated at 56°C for 60 minutes at 700 rpm

M5: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P1 and incubated at 56°C for 60 minutes

M6: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P1 + 1% SDS and incubated at 56°C for 60 minutes; incubation was continued at 50°C overnight

M7: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P1 + 1% SDS and incubated at 56°C for 60 minutes

M8: QIAprep Spin Miniprep; 1% SDS was added to buffer P1 and incubated at 65°C for 60 minutes

M9: QIAamp DNA Blood Mini Kit

M10: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to buffer P1 (+ 5% Tween20 + 0.5 % Triton-X 100 in addition) and incubated at 56°C for 60 minutes

M11: QIAprep Spin Miniprep; 10 µl Qiagen Proteinase K was added to the QIAamp lysis buffer (from QIAamp DNA Blood Mini kit) and incubated at 56°C for 60 minutes. The common protocol established for the use of the QIAprep Spin Miniprep was followed.

Figure 29: Cells were prepared as described in Figure 18. Some samples were fixed/permeabilized using methanol as indicated. After two washing steps with PBSA a total of 9 DNA extraction methods was tested. 10 µl of each DNA extract was used as template in a PCR with Phusion Polymerase (using the protocol and cycling conditions proposed by the manufacturer). The PCR product (10 µl) was analysed on a 1 % agarose gel. Both a positive control (EGFR-plasmid as template) and a negative control (no template) were included.

A number of DNA extraction methods was tested in order to maximize the amount of plasmid DNA isolated. Subsequent to plasmid DNA extraction a PCR was performed in order to assess the efficiency of DNA extraction as well as the quality of extracted DNA.

To evaluate the difference in efficiency both methanol-fixed and non-fixed samples were used for plasmid DNA extraction. In addition, cells transfected with the SV40ko-plasmid were used for plasmid DNA extraction to assess the difference in plasmid number as a function of episomal replication. For the purpose of simplification cells which have been fixed using methanol will be referred to as "F", while cells, which were not treated with methanol as fixative are labelled "N" or "NF". Non-transfected cells were additionally used for DNA extraction in order to assess potential artefacts which might arise through the carryover of genomic material as well as to be used as a "matrix" in qPCR experiments.

Most of the methods used were based on a protocol published by Qiagen, which proposes the use of the QIAprep Spin Minprep kit (usually used for plasmid DNA extraction from bacterial suspensions) for the plasmid DNA extraction from mammalian cells. Most protocols established included only 1 or 2 modifications and additions to the initial protocol. As a control, a commercially available DNA extraction kit (QIAamp DNA Blood Mini kit), especially designated for the use with mammalian cells, was used to evaluate the efficiency and success of the modified protocols.

Figure 29 shows the gel electrophoresis performed with the PCR products obtained for the amplification of the DNA extracts (received from various DNA extraction methods). The positive control was clearly visible, suggesting that the PCR setup as well as cycling parameters were valid and working alright on the EGFR-plasmid. The negative control (no template) did not show a band therefore showing that no carryover had occurred. Most samples did not show any amplification of the EGFR gene, very faint bands are visible for the ampfliciation of the DNA extracts received for the non-fixed cell samples using method 6 or method 9 and for the fixed samples using method 5 or 7 (all of them carrying an intact SV40 origin). Four clearly visible bands around 4 kb were shown in lane 1-4 representing the amplification of fixed and non-fixed samples carrying wild-type plasmid (with an intact SV40 origin) obtained through the use of method 1 and 2. Both methods 1 and 2 required the addition of 10 µl Proteinase K after the addition of the buffer P2. After the addition of Proteinase K the samples were incubated at 56°C for 10 (method 1) or 60 min (method 2), respectively. The fact, that both non-fixed and fixed cells gave a band of comparable size strongly suggests, that the recovery of intact DNA is not majorly impaired by the preceding fixation with methanol. Importantly, the cell samples, which were transfected with the SV40ko-plasmid did not present any band despite being treated

the exact same way (using method 2). On the basis of this experiment, the use of Proteinase K was proposed to be essential for the successful DNA extraction of methanol-fixed cells.

4.11.1.2 Optimization of the PCR amplification

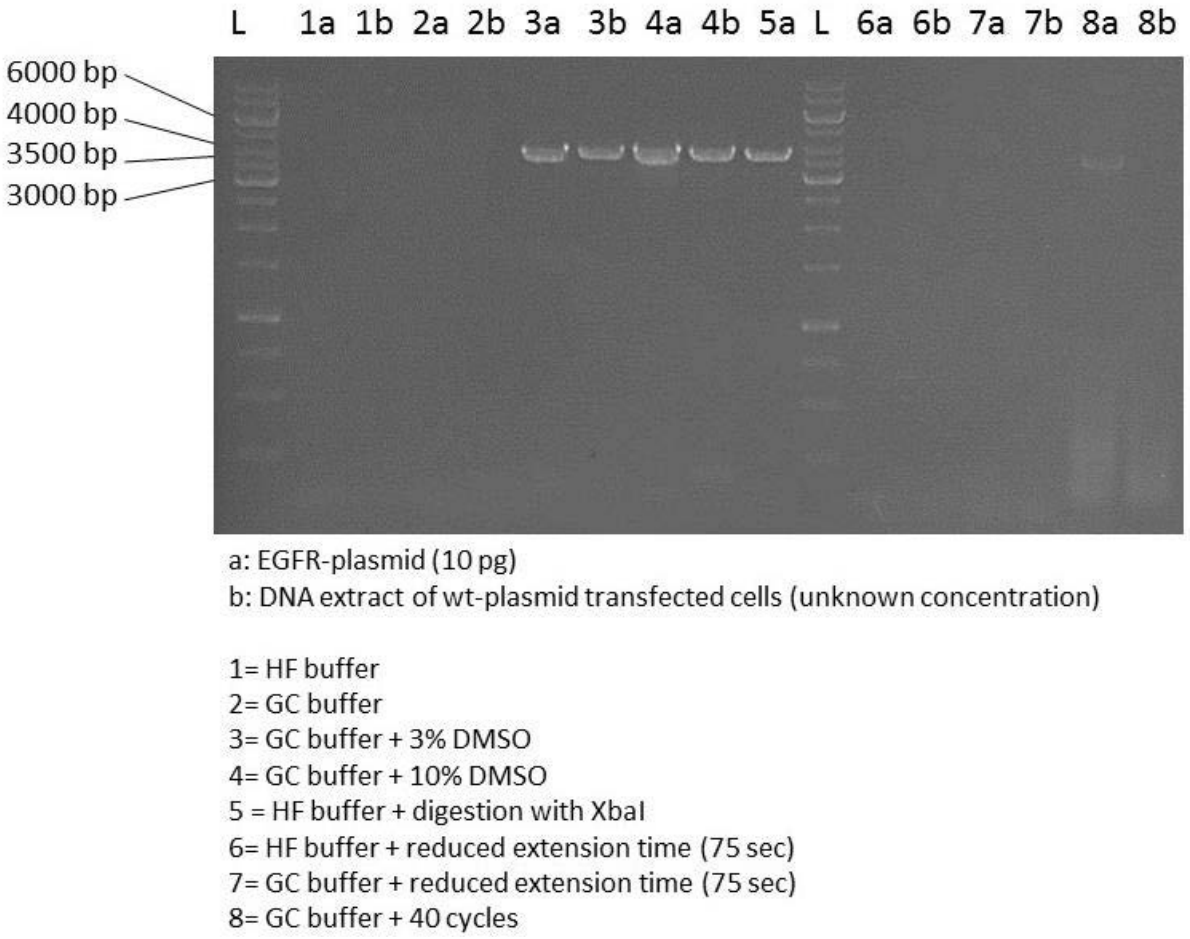


*Figure 30: **DNA gel representing the optimization of PCR by variation of certain parameters**. A defined amount of 10 pg EGFR-plasmid (from an E.coli miniprep), termed "a" in this figure, as well as the DNA extract from wt-plasmid transfected cells (unknown concentration), termed "b", were used in this experiment. In all of these reactions 1 µl template, 200 µM dNTPs, 1x buffer as well as 0.5 µM of epPCR primer (forward and reverse) were used. Reaction setups 1-8 contain one or 2 variations to the initial protocol as indicated. The PCR products (10 µl) were analysed on a 1 % agarose gel.*

The PCR setup using Phusion Polymerase had been shown to provide rather unstable results, probably depending on the amount of plasmids present in the reaction as well as the conformation they adopted.

A set of conditions were tested to eliminiate the unpredictability and create a reproducible PCR protocol. The gel shown in Figure 30 represents one of 3 independent experiments, which were successfully used to establish a PCR protocol that was shown to present reproducible, high-quality results, which allowed predictable DNA amplification.

Figure 30 shows the result obtained for the comparison of a set of 8 distinct modificiations based on the initial PCR protocol proposed by the manufacturer. The intital protocol utilized 1x HF buffer, 200 µM dNTPS, 0.5 µM of each primer and variable template amount (<250 ng). The alterations, as described in the figure legend, have shown to significantly improve the reproducibilty of DNA amplification. A rather low amount of DNA was used in order to establish a protcol which facilitates the amplification of especially small amounts of DNA (as present in the DNA extracts). Reaction setups 4 and 5 could be shown to reproducibly provide clear bands of the correct size in all three experiments. Moreover, in two out of three experiments reaction setup 3 also yielded correct bands. Reaction setup 5 included a previous linearization of the plasmid with the XbaI enzyme. In both reaction setup 3 and 4 the protocol required the addition of DMSO of varying amount (3% for setup 3 and 10% for setup 4). As either linearization of the plasmid by using a restriction enzyme or the addition of DMSO were shown to be crucial for the success of the PCR reaction, the conformation of plasmid was indirectly demonstrated to have a strong impact on the subsequent DNA amplification, because both plasmid-linearization and DMSO have been shown to reduce the fraction of supercoiled plasmid DNA. A fair band appeared in lane 8a, which represents the use of reaction setup 8 (40 cycles, GC buffer) and an amount of 10 pg plasmid DNA. As already mentioned, the gel shown in Figure 30 represents one of three independent experiments and reaction setup 8 only yielded a band in one out of three experiments. Only the use of PCR protocols 4 and 5 delivered well-defined bands of the correct size in a reproducible manner in all three experiments.

As the previous digestion with a restriction enzyme seemed unfeasible for the handling of numerous samples, protocol 4 (utilizing the GC buffer and 10% DMSO) was established as the protocol of choice.

4.11.1.3 Increasing the efficiency of DNA recovery through the addition of a carrier plasmid

If the starting material for DNA extraction is especially low, some manufacturers propose the addition of "carrier DNA" to increase the DNA yield.

The following experiment (represented in Figure 31) assessed the efficiency of DNA amplification of the full-length EGFR gene as a function of the amount of carrier plasmid added. Method 1 was used for DNA extraction and a particular amount of carrier plasmid pCT-Con2 CD20, a plasmid used for yeast display systems, which was thought not to interfere with any downstream procedures, was added after the addition of buffer P2. To evaluate the difference

in efficiency both fixed and non-fixed samples were used for DNA extraction. Additionally, cells transfected with SV40ko-plasmis were used for DNA extraction to assess the difference of plasmid number as a consequence of episomal replication. Quantification of the DNA eluate by measuring A_{260} revealed that the majority of the carrier plasmid was recovered indicating that the major part of free plasmid DNA in the cell lysate was successfully purified. Figure 31 shows a DNA gel presenting DNA amplification using various DNA eluates (from preparation where varying amounts of carrier plasmid had been added) as templates. The addition of 20 µg carrier DNA impaired amplification of the EGFR gene. Most probably the considerable amount of carrier plasmid in the eluate interfered with the amplification reaction in a negative way. This effect could be observed for all three samples (wild-type fixed, wild-type non-fixed, SV40ko fixed) containing 20 µg carrier plasmid.

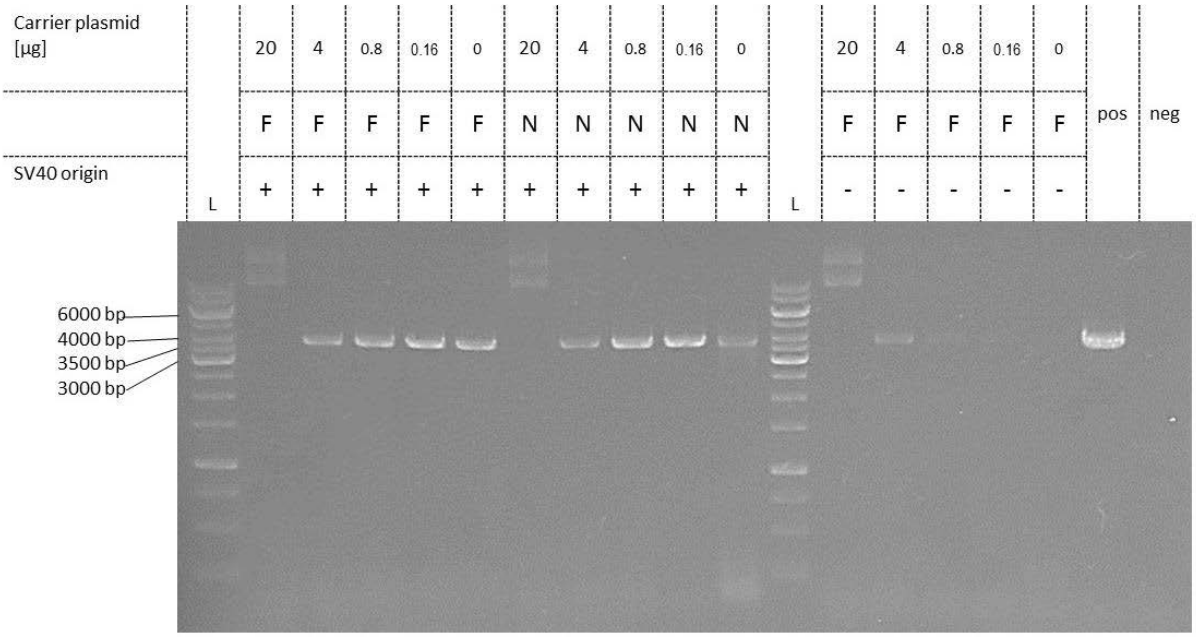


Figure 31: Cells were prepared as described in Figure 18. Plasmid DNA was extracted using method 1. In addition, various amounts of carrier DNA were added after cell lysis as indicated on the top. Finally, full-length EGFR was amplified using the optimized PCR protocol (containing 10%) DMSO and analysed on a 1% agarose gel. Both a positive control (1 pg EGFR-plasmid) and a negative control (1 µl H_2O) were included.

Additionally, Figure 31 reveals that both DNA extracts obtained from fixed as well as non-fixed cell samples carrying wild-type plasmid (with an intact SV40 origin) facilitated the successful amplification of the full-length EGFR gene. Minor differences in DNA band intensities could be observed depending on the amount of carrier plasmid. Though, as little quantitative information can be obtained from a DNA gel, the significance of little variations in signal intensity is minor. For the purpose of quantification quantiative PCR (qPCR) was performed (4.11.2).

In comparison, ampflication of the EGFR gene from the DNA extract obtained from fixed cells carrying the SV40ko-plasmid, only yielded low-intensity bands, suggesting that the amount of DNA recovered from cells lacking the SV40 origin is significantly lower compared to those carrying an intact SV40 origin and thus enabling episomal replication. This result confirms previous findings, which had verified the existence of episomal replication in this assay setup.

4.11.2 Quantification of the DNA eluate using quantitative PCR (qPCR)

4.11.2.1 Comparison of the number of plasmids extracted as a function of the DNA extraction method in use

Three independent experiments were performed comparing 9 distinct DNA extraction methods as described in Figure 29. For 6 conditions tested only 2 valid data points could be obtained. The actual percentage of transfected cells, which had been shown to differ between experiments, was determined by flow cytometric analysis. Construction of a standard curve

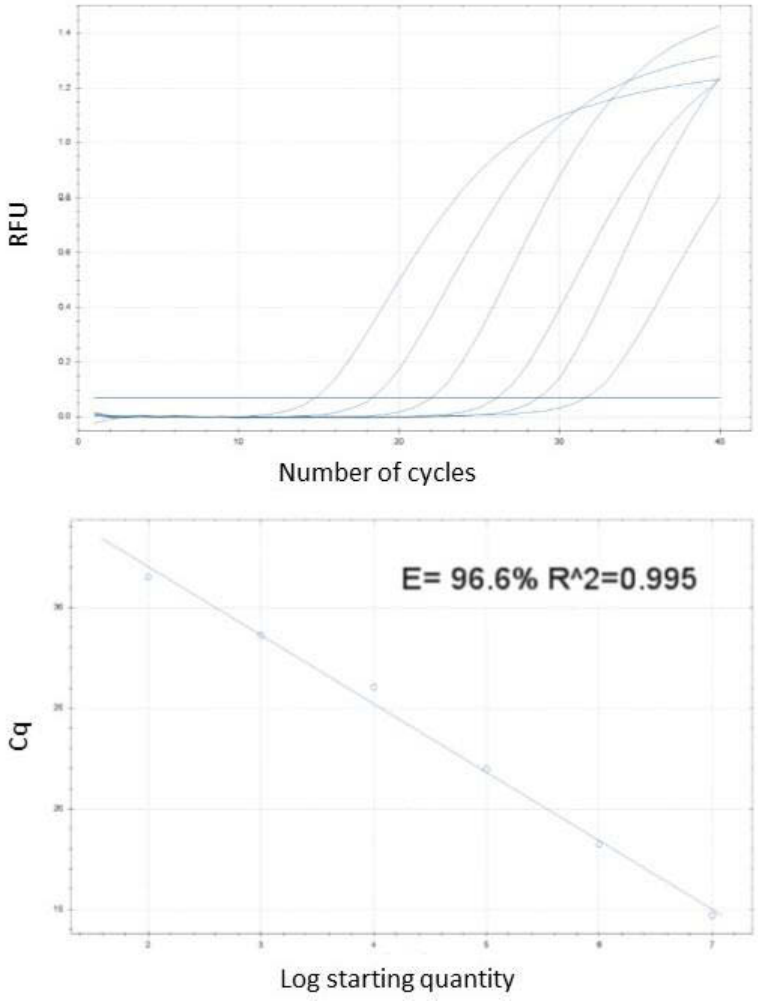


Figure 32: ***qPCR plots and the resulting standard curve received for a dilution series ranging from 10^2 to 10^7 plasmids.*** *A dilution series ranging from 10^2 to 10^7 EGFR-plasmids was prepared and "spiked" with DNA extract derived from non-transfected cells in a 1:10 dilution to imitate the matrix effects exhibited by the samples. One representative example is shown here. In the plot on the top the increase in relative fluorescence units (RFU) is shown as a function of PCR cycles for all six samples of the dilution series. The sample with the highest concentration of EGFR-plasmid showed an increase in RFU first (at the lowest number of cycles). A certain threshold (horizontal line) was set to exclude the detection of experimental fluctuations. The plot on the bottom shows the quantification cycle (Cq) as a function of the logarithmic starting quantity therefore presenting the standard curve used for the plasmid quantification. The efficiency of the qPCR experiment (i.e. the efficiency of duplication in each cycles), calculated on the basis of this dilution series, was 96.6%.*

ranging from 10^2 to 10^7 plasmids enabled the calculation of the numbers of isolated plasmids per sample.

Figure 32 shows the plots received for a typical standard curve in a qPCR experiment. In this case the plot in the top panel shows a series of dilution ranging from 10^2 to 10^7 EGFR-plasmids per cell. Each curve represents a particular dilution and exceeds a certain threshold (horizontal line) at a particular number of cycles. Naturally, the higher the plasmid concentration the lower the number of cycles at which the threshold is exceeded. The "quantification cycle" (Cq) for each dilution is further plotted against the number of plasmids. This standard curve was used to calculate the number of plasmids in each sample based on the corresponding Cq value.

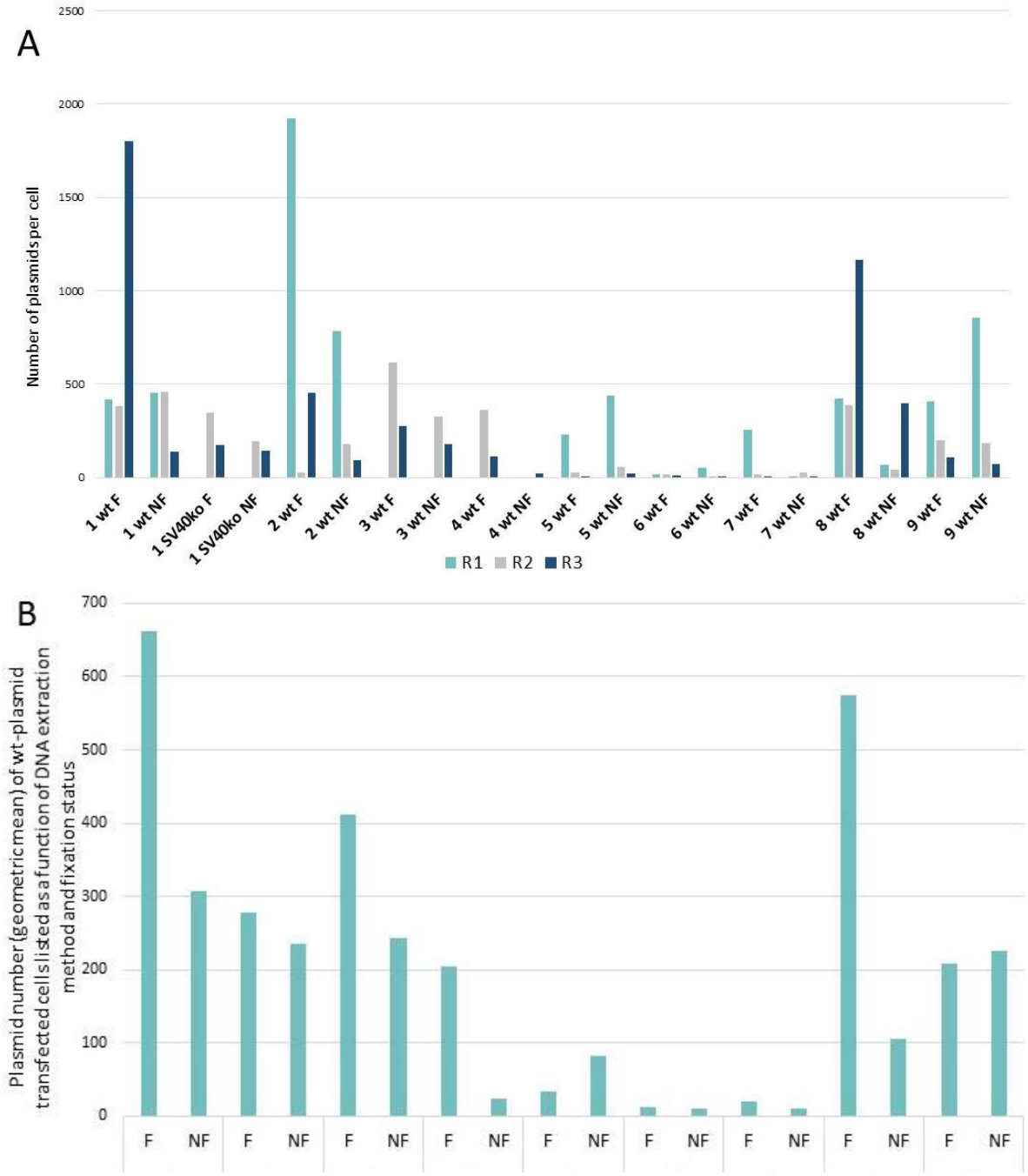


Figure 33: (A) Plasmid DNA was prepared as described in Figure 29. The DNA extracts were used for quantification using qPCR. The template was used in a 1:10 dilution to minimize potential matrix effects of carried over gDNA. A dilution series ranging from 10^2 to 10^7 EGFR-plasmids and „spiked" with DNA extract derived from non-transfected cells in a 1:10 dilution to imitate the matrix effects exhibited by the samples. The starting quantity derived from this experiment was normalized to a defined number of positive cells and the number of plasmids obtained per individual transfected cell was calculated. (B) Presentation of the number of plasmids (geometric mean) for wild-type plasmid (carrying an intact SV40 origin) transfected cells as a function of the DNA extraction method and fixation status.

Data presented in Figure 33A show considerable variations between the individual experiments. Each individual experiment required the repeated transfection of cells, therefore certain experimental varations could not be eliminated. Methods 5, 6 and 7, where Proteinase K was added to buffer P1 (method 6 and 7 contained 1% SDS in addition) were shown to deliver unsatisfactory results for most of the experiments performed. Even an incubation time overnight (as conducted using method 6) did not benefit the plasmid DNA isolation. As the Cq values were comparable to that of the negative control, those DNA extraction methods were basically proven unsuccessful. Interestingly, an increase in temperature to 65°C for 60 minutes after the addition of buffer P1 + 1 % SDS (as conducted within method 8) was shown to provide much better results. In fact, this method did not require the addition of Proteinase K, suggesting that an increase above a certain temperature benefits cell lysis in the presence of a detergent (SDS), irrespective of the addition of Proteinase K. The use of DNA extraction method 4, where the Proteinase K incubation (in buffer P2) was performed under shaking conditions, was shown to be unsuccessful if used on non-fixed cells, while its use on cells, which had previously been fixed with methanol, was shown to be of mediocre success. Consistent with Figure 29, where the full-length EGFR gene was amplified from the plasmid preparations, methods 1 and 2, both include the addition of Proteinase K to buffer P2, were shown to provide good, though rather unstable results ranging from 1921 plasmids/cell to 23 plasmids/cell for the use of method 2 on fixed cell samples, carrying the wild-type plasmid with an intact SV40 origin. Similar, the use of method 1 on fixed cells carrying the wild-type plasmid showed one outlier at 1803 plasmids/cell. Method 9, the use of the QIAamp DNA Blood Mini kit, showed very little variations between the replicate experiments and hardly any differences between fixed and non-fixed samples. These results demonstrate, that the use of the QIAprep Spin Miniprep Kit provided equal to better results compared to the DNA extraction kit for mammalian cells.

Most surprisingly, the DNA extraction of fixed and non-fixed cells carrying the SV40ko-plasmids yielded plasmid numbers that were comparable to those carrying an intact SV40 origin. The DNA extraction of fixed cells carrying SV40ko-plasmids showed a number of 348 and 171 plasmids/cell in two replicate experiments. As the knockout of the SV40 origin is thought to suppress episomal replication, only a few plasmids per cell were expected in these

samples. This finding appears contradictory to previous results like the failed amplification of the full-length EGFR gene from DNA eluates received from cells carrying the SV40ko-plasmids (Figure 31) and shall be discussed in more detail in the discussion.

Overall, the plasmid number received for methanol-fixed samples seemed to be slightly higher than the number received for non-fixed cells, suggesting that methanol-permeabilization supported subsequent cell lysis during plasmid isolation (Figure 33B).

4.11.2.2 Evaluating the effect of the addition of carrier plasmid on DNA recovery

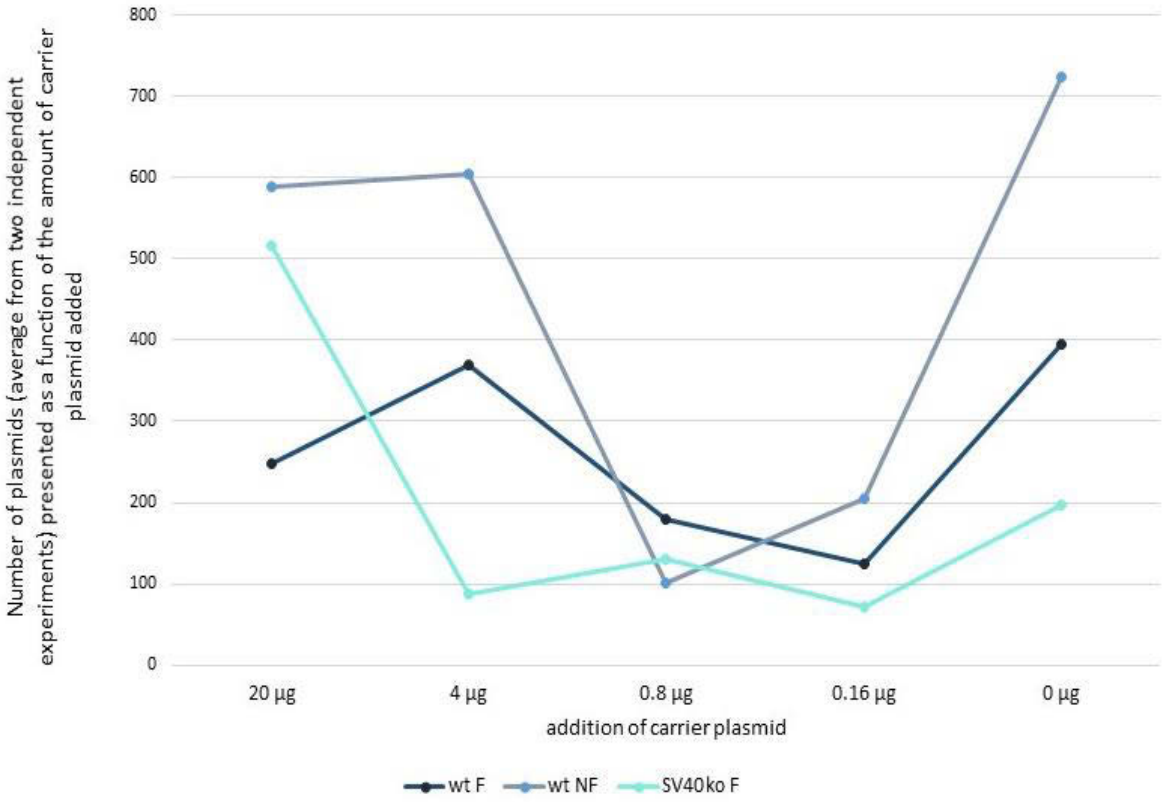


Figure 34: ***Presentation of the relationship between the amount of carrier plasmid added and the DNA yield****. HEK293T cells were prepared as described in Figure 18. After methanol fixation/permeabilization, plasmid DNA was isolated using method 1. In addition, various amounts of carrier DNA were added directly after the addition of Proteinase K.*

After method 1 had been established as the method of choice, the addition of carrier plasmid was tested in order to increase DNA recovery and yield.

Figure 34 shows the average plasmid number per cell (from two independent experiments) as a function of the amount of carrier plasmid added, ranging from 20 µg to 0 µg per sample. Based on these data, no correlation between the plasmid number per cell and the amount of carrier plasmid added could be found. The variations in plasmid number seemed to

be due to experimental fluctuations, rather than a function of the amount of carrier plasmid added.

4.11.2.3 Assessing the impact of the SV40 origin on the number of plasmids per cell

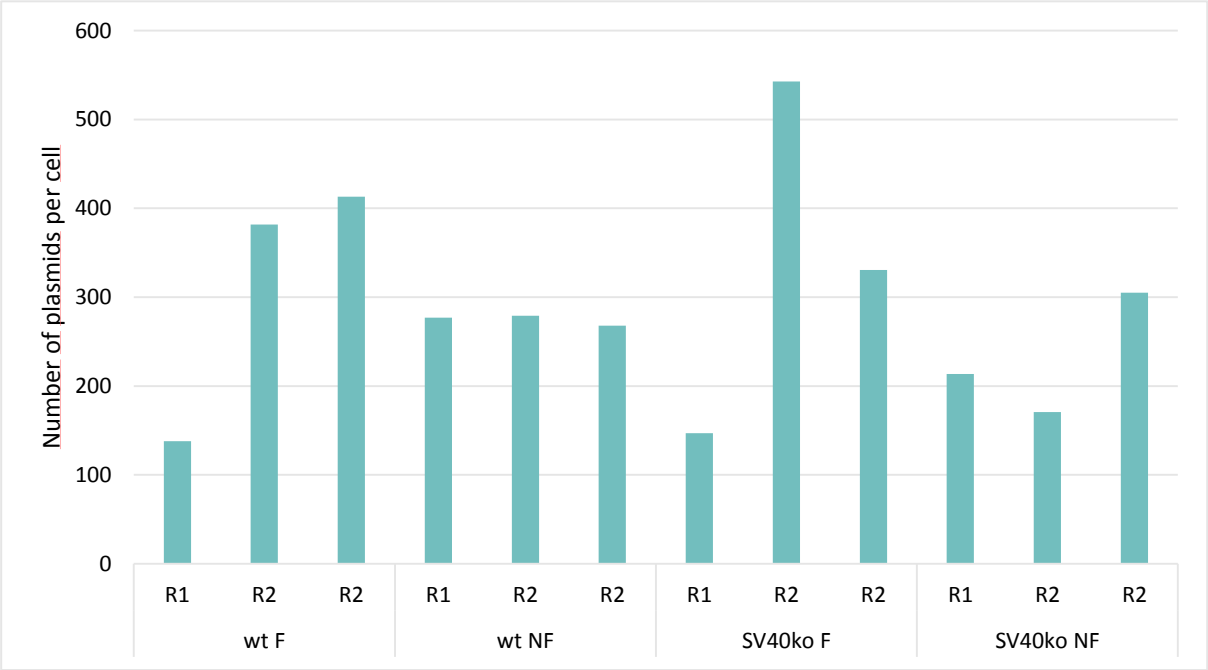


Figure 35: HEK293T cells were prepared as described in Figure 18. After methanol permeabilization (as indicated), plasmid DNA was isolated using method 1. Three replicates were conducted in parallel in order to determine the variations in qPCR data.

To assess the dependence of the plasmid number per cell on episomal replication, a number of replicates (using method 1) was prepared in parallel, indicated R1 to R3. The exact number of transfected cells was determined via FACS after the sample had been separated to two tubes at equal volume. Therefore, the number of plasmids measured using qPCR could be normalized for every individual sample based on the actual cell count.

The results depicted in Figure 35 illustrate, once again, the experimental fluctuations associated with the experimental procedure including DNA extraction and qPCR. Therefore, detailed comparison between samples carrying intact vs. deleted SV40 origin was not possible. However, it is obvious that there are no major differences depending on the presence of an intact SV40 origin, which is contradictory to the amplification of the full-length EGFR gene in Figure 31.

For the purpose of illustrating the difference between samples carrying wt- and SV40ko-plasmids, data pairs, obtained for the use of DNA extraction method 1 (no carrier plasmid) in 6 independent experiments, were correlated and illustrated in the form of a line chart.

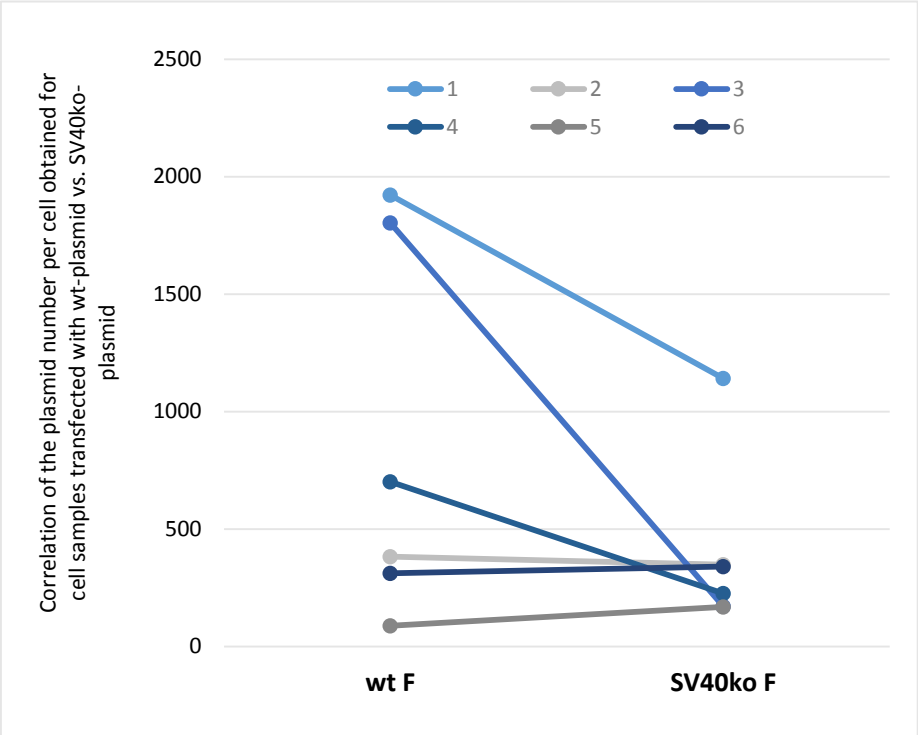


Figure 36: Correlation of the plasmid number/cell obtained for cell samples transfected with wt- vs. SV40ko-plasmid. Data from six experiments were combined and statistically compared using a two-tailed paired t-test, yielding a p-value of 0.14.

In three out of six experiments the lack of the SV40 origin on the EGFR-plasmid was shown to dramatically decrease the number of plasmids per cell (sample 1, 3 and 4), whereas in the other three experiments only minor differences were observed. A two-tailed paired t-test revealed that no significant change in plasmid number is associated with the presence or lack of the SV40 origin, represented in a p-value of 0.14. This result is surprising and contradicts several previous findings, such as the decrease in EGFR-expression as well as phosphorylation upon SV40 origin knockout (Figure 18) or the amplification of the full-length EGFR gene, which had been shown to fail upon SV40 origin knockout (Figure 31).

5 Discussion

5.1 Key experiments for the establishment of the assay

The selection of the HEK293T cell line as a model system was shown to be an adequate choice within this study. The expression levels of both HER2 and EGFR are negligible (Figure 7), therefore the evaluation of recombinant (mutant) EGFR previously introduced through the transfection can be performed with this assay.

The optimization of the fixation and permeabilization procedure presented methanol as the optimal fixative for the purpose of this assay. The use of methanol for the fixation of HEK293T cells could be shown to present a homogenous cell population, evaluated regarding their scatter properties as well as CV value of the fluorescence distribution, and showed a fair separation of the transfected (single positive) cells from the main cell population. Further, its use was shown to be compatible with the application of phospho-EGFR-specific antibodies. In fact, the signal intensity received for the phosphorylation of EGFR upon stimulation with EGF came as a surprise presenting fold changes ranging from 7-fold to 56-fold for methanol, dependent on the antibody in use. A paper published by Kovacs et al. presented the use of flow cytometry in order to evaluate the auto-inhibitory role of the C-terminal tale and revealed a minor fold change of 2- to 4-fold for the comparison of stimulated and unstimulated EGFR-WT[5], showing that compared to previous studies our detection system is highly efficient.

The transfection procedure was optimized in order to obtain predominantly cells carrying one or a few plasmids only. The evaluation of so-called “single positive” cells was enabled by transfection of EGFR constructs carrying two distinct protein tags HA and myc. For the purpose of transfection with only few plasmids, the addition of an inert carrier plasmid was shown to be crucial, suggesting that a certain DNA concentration is necessary for the formation of complexes between DNA and transfection reagent. Even if the number of plasmids (of interest) is lowered, this will simply reduce the number of transfected cells. Naturally, the number of cells carrying multiple plasmids will decrease dramatically when the concentration of plasmid is lowered, still, the number of “single positive” cells will drop simultaneously. For this purpose, the addition of carrier plasmid could be established as a vital tool in order to alter the number of plasmids per cell (Figure 14).

The distinct cell populations, which were observed when the transfection was conducted with high plasmid concentration (1 µg/ml) are likely due to different stages within the cell cycle shown by the HEK293T cells (Figure 13). During mitosis the nuclear membrane disassembles

and preferentially allows the DNA to enter the cell nucleus[31,32]. Cohen discussed this phenomenon when he observed populations of high- and low-expressing cells after transfection (2009).

Comparing the transfection efficiency using the transfection reagents PEI, *Trans*IT-X2, *Trans*IT-293 and *Trans*IT-LT1, revealed that both PEI and *Trans*IT-X2 delivered superior results to the other transfection reagents. *Trans*IT-X2 was described by the manufacturer to provide "novel, non-liposomal, polymeric delivery" and therefore shares certain characteristics with PEI, polyethylenimine. This suggests that the use of polymeric transfection reagents might be superior regarding the transfection efficiency of HEK293 cells (Figure 13).

In fact, the process of transfection faces several cellular as well as extracellular barriers which are known to impede gene delivery such as the uptake of the DNA by endocytosis, the subsequent degradation in endo-lysosomes, the metabolic instability of free DNA in the cytoplasm and last, the nuclear translocation, which requires either the disassembly of the nuclear membrane (such as during mitosis) or the transport via nuclear pore membrane (NPC). Only a small fraction of internalized complexed plasmid will eventually reach the cytoplasm (and subsequently nucleus), while the major part will be degraded in endo-lysosomes[31]. In agreement with our data, polyplexes were discussed to be more likely than lipoplexes to escape from the endosome and reach the nuclear membrane intact by Cohen and co-workers (2009). This finding might thus explain the superior results exhibited through the use of the polymeric transfection reagents PEI and *Trans*IT-X2.

When the investigations were focussed on *Trans*IT-X2 and PEI, the findings revealed that the single positive cells obtained with the use of PEI seemed to exhibit less intense fluorescence signals after intracellular staining compared to the corresponding cell sample transfected with *Trans*IT-X2 (Figure 15). A study performed by Cohen et al. (2009) revealed that the transfection with Lipofectamine or PEI resulted in a similar number of plasmids per cell, but showed huge variation in the signal intensity obtained for protein expression. They proposed that this result could be the consequence of PEI, which was still bound to the plasmid in the nucleus and might therefore not be as transcriptionally active or even partially degraded[32]. *Trans*IT-X2, which is based on a polymeric transfection system according to the manufacturer, might contain additional components which might render the negative effect of bound polymer and facilitate a greater transcription and replication rates compared to PEI.

5.2 Retrieving the genetic information from the phenotypes of interest

In order not to lose the library diversity at the stage of DNA isolation, episomal replication of the one single plasmid reaching the cell nucleus is anticipated through the SV40 origin. The analysis of the plasmid number per cell via qPCR was shown to harbour considerable variability, as observed within the results obtained for triplicate samples analysed in parallel. Though, it is not known whether the preparation and qPCR reaction itself or the previous DNA isolation using the QIAprep Spin Miniprep kit is responsible for the introduction of considerable experimental variation. Still, the comparison of several protocols engaged for the isolation of plasmid DNA from HEK293T cells revealed, that the use of the QIAprep Spin Miniprep kit provided equal or better results compared to the QIAamp DNA Blood Mini kit, which is especially designed for the use with mammalian cells (Figure 33). The addition of Proteinase K after the addition of buffer P2, which consists of 200 mM NaOH and 1% SDS, was shown to provide good results when an incubation time of 10 min was followed. The elongation of the incubation time in the P2 buffer was shown to have a negative effect on the plasmid number obtained using qPCR, though this effect is not profound and might be due to experimental errors. Still, the protocol published by the manufacturer suggests to follow an incubation time of a maximum of 5 min, which might be explained by the fact that the stringent pH conditions could have a deleterious effect on the plasmid DNA. The addition of Proteinase K to buffer P1 (50 mM Tris-HCl, 10 mM EDTA and 100 µg/ml RnaseA) or P1 supplemented with 1% SDS (with varying times of incubation) was shown to provide particularly poor results, suggesting that Proteinase K, which works as an endo- as well as exopeptidase digesting proteins in the cell lysates, requires the previous externalization of cellular material as facilitated through stringent pH conditions. On the other hand, introduction of an incubation step at 65°C for 60 min after the addition of buffer P1 supplemented with 1% SDS could be shown to provide much better results. This suggests that, despite the fact that the cell permeabilization facilitated through 1% SDS might not be sufficient for the Proteinase K to exhibit its full potential, a temperature increase to 65°C might complete the cell lysis. Interestingly, surprisingly poor results were obtained if the addition of Proteinase to buffer P2 was used in combination with the agitation of the cell samples on a thermocycler. If this finding was profound, and further experiments and investigations are needed in order to compare the DNA isolation protocols in more detail, this would again suggest, that the cell lysis facilitated through NaOH in the P2 buffer might lead to deleterious effects on the integrity and stability of DNA. This effect might possibly set in earlier and thus exhibit a more pronounced (negative)

effect on the stability of the DNA if the cell lysis was accelerated through the agitation of the cell samples. This explanation might also apply to the differences in plasmid number found for fixed and non-fixed cell samples, showing a greater plasmid number for cell samples which had been fixed with methanol prior to the DNA isolation. The fixation with methanol leads to the denaturation of cellular proteins and might therefore delay the cell lysis and abbreviate the time the naked plasmid DNA is subjected to stringent lysis conditions (e.g. high pH exhibited through NaOH). Alternatively, prior methanol-mediated permeabilization might simply support the subsequent cell lysis. Though, this effect was not observed for all samples and will therefore need further re-evaluation before definite conclusions can be made. Overall, the optimization of the DNA isolation procedure was proven successful with regard to establishing an efficient DNA isolation procedure in order to obtain a sufficient number of EGFR-plasmids per (methanol fixed) cell. In order to discover the underlying reasons for certain effects observed, such as the difference in plasmid number between fixed and non-fixed samples, more experiments need to be conducted.

5.3 Unravelling the mystery regarding the SV40 origin

The situation regarding the effect of the episomal replication facilitated through the SV40 origin remained baffling until the end of the project. The investigation of EGFR expression or phosphorylation signal upon the knockout of the SV40 origin in comparison to the samples carrying an intact SV40 origin clearly suggested that the SV40 origin is intact and crucial for the replication of plasmid and the subsequent signal (Figure 18). Still, the results obtained for the samples with and without an intact SV40 origin using qPCR, did not provide a definite answer whether and how many plasmids were replicated in the cell nucleus. In fact, through the analysis with a paired two-tailed t-test no significant change in the plasmid number per cell could be found (Figure 36). On the other hand, amplification of the full-length EGFR gene and subsequent analysis using gel electrophoresis revealed that samples carrying a defect SV40 origin could not be successfully amplified (Figure 31). There are several possible reasons, which might, at least partially, explain these contradictory findings.

A study by Cohen and co-workers (2009) describes the evaluation of the number of plasmids per cell nucleus via qPCR, which very much resembles the DNA isolation experiments conducted within this study. They found that a large number of plasmids remains bound to the cell nucleus during cell lysis, which would be analysed and falsify the real result. Cohen therefore proposed several methods, which could be used in order to yield nuclei with less-adhering plasmids. This paper suggested that a considerable background of (possibly

intact) plasmid might be present in the DNA isolate. This finding would in fact explain the considerable amount of DNA detected in the samples carrying the SV40ko-plasmid via qPCR, though, the failed amplification of the full-length EGFR gene cannot be conclusively explained with this hypothesis[32].

As already mentioned, the transfection of cells faces several barriers including the uptake through the cell membrane, the endo-lysosomal degradation, metabolic instability in the cytosol and eventual uptake through the NPC[31]. A considerable number of plasmids is trapped and eventually degraded or fragmented in the endo-lysosomes. This fact presents one possible explanation how SV40ko-plasmids would be detected via qPCR, where an amplicon of only 100 bp length is amplified, whereas the amplification of the full-length EGFR gene is unsuccessful.

A subsequent step regarding the uptake of plasmid DNA presents another reason which might explain the contradictory results. The nuclear membrane presents the ultimate obstacle with only 0.1-0.001% of cytosolically injected plasmid being ultimately transcribed[31]. Lechardeur and co-workers proposed the existence of a nuclease present in the cytosol which exhibits endonucleolytic activity on cytosolically located DNA[33]. Interestingly, this effect was shown to be abrogated when the cytosolic DNA was complexed with PEI or a large amount of bulk-DNA was co-transfected[34]. Although these effects were observed for micro-injected naked cDNA, they might still be true for polyplex/lipoplex released DNA prior to the nuclear uptake.

Taken together, these effects, both concerning the barriers associated with DNA transfection, could explain the contradictory findings regarding the presence or absence of an intact SV40 origin. The same amount of plasmid DNA was originally applied to all the samples, irrespective of their integrity of the SV40 origin, therefore, similar amounts of plasmid would enter the cell cytosol and nucleus. Thus, even if the effect of episomal replication is not evident when performing qPCR, the results obtained for the amplification of the full-length EGFR gene strongly suggest that the DNA isolate for the SV40ko samples contain mainly digested EGFR-plasmids, which cannot serve as templates for the 3600 bp EGFR-gene.

Naturally, the number of EGFR-plasmids in the cell nucleus obtained via epsiomal replication might thus be lower than expected. Luckily, as the plasmids, which are localized in the cellular cytosol, are very likely digested through a variety of mechanisms, little "background" plasmid will be carried over when the full-length EGFR gene from selected cells is amplified for a subsequent cycle of selection. Ultimately, model selections using inhibitors with known resistance mutations will allow the validation of the system. That is, isolation of

most or all of the resistance mutations known from clinical data will demonstrate that library diversity was not lost at any step during the selection process.

Apart from variations in plasmid number due to usage of different plasmid isolation protocols, differences may also be due to qPCR-related errors. Several papers discuss the impact of the plasmid conformation on qPCR-results[35–37]. The presence of the plasmid in a supercoiled or linear conformation was shown to have a considerable impact on the subsequent quantification using qPCR. In our qPCR assays, both the standards as well as the samples were prepared in the same manner in order to ensure adequate quantification, though the impact of the plasmid conformation on the qPCR result cannot be denied and should be kept in mind, when analysing qPCR data. In fact, the introduction of DNA strand breaks was described to change the supercoiled plasmid conformation into a nicked-circular or linear form[37]. Amplification of fragmented cytosolic plasmid vs. nucleus-derived circular supercoiled plasmid might thus be unbalanced, favouring the amplification and detection of cytosolic fragmented plasmid. This in turn suggests, that the EGFR-plasmid replicated and derived from the cell nucleus is underrepresented in the result (or the cytosolic plasmid fraction is overrepresented) and might account for a greater number after all. However, it should be noted that the qPCR reactions contained 2% DMSO, which would be expected to release the supercoiled structure of circular plasmid DNA. Thus, DMSO should reduce the difference in amplification efficiencies between fragmented and circular plasmid.

5.4 Successful application and validation of the developed assay

After key steps of the assay had been developed and optimized, different effects regarding EGFR signalling and resistance mechanisms described in the scientific literature were reproduced with the use of this assay.

The ligand-independent phosphorylation signal exhibited by the activating mutations, initially described by Lynch and co-workers in 2004, could be successfully reproduced using this experimental setup (Figure 19), so could the inhibitory effect on EGFR conveyed through distinct first and third generation TKIs (Figure 21). The resistance towards erlotinib was successfully imitated through the introduction of the T790M mutation into an activating background such as EGFR-L858R or EGFR-del19. Though, these double mutants, could be fully supressed upon the addition of the third generation inhibitor AZD9291, which in fact only had a minor effect on EGFR-WT. This finding is not novel, but is observed as minor drug toxicity exhibited by AZD9291 in the clinics[2,12].

Furthermore, addition of the EGFR-specific monoclonal antibody cetuximab strongly reduced EGFR-phosphorylation, which is in agreement with published data, thereby further validating our assay (Figure 20). Though, both mutants EGFR-L858R and EGFR-del19 could not be supressed with the use of this antibody. A finding, which does not come at a surprise, as those mutants are known to be ligand independent and are therefore not affected by blocking EGF from the extracellular domains. Interestingly though, the phosphorylation signal received for unstimulated EGFR-WT was shown to increase upon the addition of cetuximab. The signal intensity exhibited by those samples was comparable to that of EGFR-WT, which had been treated with cetuximab prior to the stimulation with EGF. This suggests, that through the binding of the bivalent antibody cetuximab, dimerization and therefore activation of two receptors might be induced to a certain extent.

Last, after EGFR mutant libraries had been established, the effect of the introduction of random mutations was evaluated and compared to the signal exhibited by non-mutagenized EGFR. Both non-mutagenized EGFR as well as the EGFR mutant library were subjected to a selection pressure of some sort and the number of EGFR mutants escaping the selection pressure was evaluated, clearly showing that the percentage of resistant or activated EGFR mutants was increased in the library sample compared to the non-mutated counterpart. This finding emphasized the functionality of the assay for the detection of mutations, which might confer resistance to the cell in order to escape the selection pressure, and reinforced the multitude of fields of interest, in which this assay will be a valuable tool.

Overall, it is fair to say that the establishment and future application of this assay will hopefully bring great progress, not only in the field of drug resistance towards EGFR inhibitors, but any other area which might benefit from the rational selection of randomly mutated (mammalian) proteins showing the desired characteristics.

6 References

1. Wheeler DL, Dunn EF, Harari PM. Understanding resistance to EGFR inhibitors-impact on future treatment strategies. *Nat Rev Clin Oncol*. 2010;7(9):493-507. doi:10.1038/nrclinonc.2010.97.
2. Arteaga CL, Engelman JA. ERBB receptors: From oncogene discovery to basic science to mechanism-based cancer therapeutics. *Cancer Cell*. 2014;25(3):282-303. doi:10.1016/j.ccr.2014.02.025.
3. Yarden Y, Pines G. The ERBB network: at last, cancer therapy meets systems biology. *Nat Rev Cancer*. 2012;12(8):553-563. doi:10.1038/nrc3309.
4. Kovacs E, Zorn JA, Huang Y, Barros T, Kuriyan J. A structural perspective on the regulation of the epidermal growth factor receptor. *Annu Rev Biochem*. 2015;84:739-764. doi:10.1146/annurev-biochem-060614-034402.
5. Kovacs E, Das R, Wang Q, et al. Analysis of the role of the C-terminal tail in the regulation of the epidermal growth factor receptor. *Mol Cell Biol*. 2015;1162(June):MCB.00248-15. doi:10.1128/MCB.00248-15.
6. Sato K ichi. Cellular functions regulated by phosphorylation of EGFR on TYR845. *Int J Mol Sci*. 2013;14(6):10761-10790. doi:10.3390/ijms140610761.
7. Jones RB, Gordus A, Krall JA, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*. 2006;439(7073):168-174. doi:10.1038/nature04177.
8. Pines G, Köstler WJ, Yarden Y. Oncogenic mutant forms of EGFR: Lessons in signal transduction and targets for cancer therapy. *FEBS Lett*. 2010;584(12):2699-2706. doi:10.1016/j.febslet.2010.04.019.
9. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med*. 2004;350(21):2129-2139. doi:10.1056/NEJMoa040938.
10. Tsao M-S, Sakurada A, Cutz J-C, et al. Erlotinib in Lung Cancer — Molecular and Clinical Predictors of Outcome. *N Engl J Med*. 2005;353(2):133-144. doi:10.1056/NEJMoa050736.
11. Kancha RK, von Bubnoff N, Peschel C, Duyster J. Functional analysis of epidermal growth factor receptor (EGFR) mutations and potential implications for EGFR targeted therapy. *Clin Cancer Res*. 2009;15(2):460-467. doi:10.1158/1078-0432.CCR-08-1757.
12. Ercan D, Choi HG, Yun C, et al. EGFR Mutations and Resistance to Irreversible

Pyrimidine-Based EGFR Inhibitors. *Clin Cancer Res*. 2015:3913-3924. doi:10.1158/1078-0432.CCR-14-2789.

13. Hata AN, Niederst MJ, Archibald HL, et al. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat Med*. 2016;22(3):262-269. doi:10.1038/nm.4040.
14. Bhang HC, Ruddy DA, Krishnamurthy Radhakrishna V, et al. Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat Med*. 2015;21(5):440-448. doi:10.1038/nm.3841.
15. Ho M, Nagata S, Pastan I. Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. *Proc Natl Acad Sci*. 2006;103(25):9637-9642. doi:10.1073/pnas.0603653103.
16. Kim TK, Eberwine JH. Mammalian cell transfection: The present and the future. *Anal Bioanal Chem*. 2010;397(8):3173-3178. doi:10.1007/s00216-010-3821-6.
17. Lipps G. *Plasmids: Current Research and Future Trends*. Bayreuth: Caister Academic Press; 2008.
18. Howat WJ, Wilson BA. Tissue fixation and the effect of molecular fixatives on downstream staining procedures. *Methods*. 2014;70(1):12-19. doi:10.1016/j.ymeth.2014.01.022.
19. Krutzik PO, Nolan GP. Intracellular phospho-protein staining techniques for flow cytometry: monitoring single cell signaling events. *Cytometry A*. 2003;55(2):61-70. doi:10.1002/cyto.a.10072.
20. Jensen UB, Owens DM, Pedersen S, Christensen R. Zinc fixation preserves flow cytometry scatter and fluorescence parameters and allows simultaneous analysis of DNA content and synthesis, and intracellular and surface epitopes. *Cytom Part A*. 2010;77(8):798-804. doi:10.1002/cyto.a.20914.
21. Dietrich D, Uhl B, Sailer V, et al. Improved PCR Performance Using Template DNA from Formalin-Fixed and Paraffin-Embedded Tissues by Overcoming PCR Inhibition. *PLoS One*. 2013;8(10):1-10. doi:10.1371/journal.pone.0077771.
22. Hansmann L, Schmidl C, Boeld TJ, et al. Isolation of intact genomic DNA from FOXP3-sorted human regulatory T cells for epigenetic analyses. *Eur J Immunol*. 2010;40(5):1510-1512. doi:10.1002/eji.200940154.
23. Smith LJ, Braylan RC, Nutkis JE, Edmundson KB, Downing JR, Wakeland EK. Extraction of cellular DNA from human cells and tissues fixed in ethanol. *Anal Biochem*. 1987;160(1):135-138. http://www.ncbi.nlm.nih.gov/pubmed/3551684. Accessed

December 23, 2016.

24. Beckstead JH. A simple technique for preservation of fixation-sensitive antigens in paraffin-embedded tissues. *J Histochem Cytochem*. 1994;42(8):1127-1134. http://www.ncbi.nlm.nih.gov/pubmed/8027531. Accessed December 23, 2016.
25. Wester K, Asplund A, Bäckvall H, et al. Zinc-based fixative improves preservation of genomic DNA and proteins in histoprocessing of human tissues. *Lab Invest*. 2003;83(6):889-899. doi:10.1097/01.LAB.0000074892.53211.A5.
26. Lykidis D, Van Noorden S, Armstrong A, et al. Novel zinc-based fixative for high quality DNA, RNA and protein analysis. *Nucleic Acids Res*. 2007;35(12). doi:10.1093/nar/gkm433.
27. Zhao H, Li J, Traganos F, et al. Cell fixation in zinc salt solution is compatible with DNA damage response detection by phospho-specific antibodies. *Cytom Part A*. 2011;79A(6):470-476. doi:10.1002/cyto.a.21060.
28. Zaccolo M, Williams DM, Brown DM, Gherardi E. An Approach to Random Mutagenesis of DNA Using Mixtures of Triphosphate Derivatives of Nucleoside Analogues. *J Mol Biol*. 1996;255(4):589-603. doi:10.1006/jmbi.1996.0049.
29. Drummond DA, Iverson BL, Georgiou G, Arnold FH. Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J Mol Biol*. 2005;350(4):806-816. doi:10.1016/j.jmb.2005.05.023.
30. Hosseinzadeh Colagar A, Amjadi O, Valadan R, Rafiei A. Minimal HER1 and HER2 expressions in CHO and HEK-293 cells cause them appropriate negative cells for HERs-related studies. *Res Mol Med*. 2013;1(3):6-12. http://rmm.mazums.ac.ir/browse.php?a_code=A-10-26-19&slc_lang=en&sid=1.
31. Lechardeur D, Lukacs G. Intracellular Barriers to Non-Viral Gene Transfer. *Curr Gene Ther*. 2002;2(2):183-194. doi:10.2174/1566523024605609.
32. Cohen RN, van der Aa MAEM, Macaraeg N, Lee AP, Szoka FC. Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *J Control Release*. 2009;135(2):166-174. doi:10.1016/j.jconrel.2008.12.016.
33. Lechardeur D, Sohn KJ, Haardt M, et al. Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. *Gene Ther*. 1999;6(October):482-497. doi:10.1038/sj.gt.3300867.
34. Pollard H, Toumaniantz G, Amos JL, et al. Ca2+-sensitive cytosolic nucleases prevent efficient delivery to the nucleus of injected plasmids. *J Gene Med*. 2001;3(2):153-164. doi:10.1002/jgm.160.

35. Chen J, Kadlubar FF, Chen JZ. DNA supercoiling suppresses real-time PCR: A new approach to the quantification of mitochondrial DNA damage and repair. *Nucleic Acids Res*. 2007;35(4):1377-1388. doi:10.1093/nar/gkm010.

36. Hou Y, Zhang H, Miranda L, Lin S. Serious overestimation in quantitative pcr by circular (supercoiled) plasmid standard: Microalgal pcnaas the model gene. *PLoS One*. 2010;5(3):e9545. doi:10.1371/journal.pone.0009545.

37. Lin CH, Chen YC, Pan TM. Quantification bias caused by plasmid DNA conformation in quantitative real-time PCR assay. Charbit A, ed. *PLoS One*. 2011;6(12):e29101. doi:10.1371/journal.pone.0029101.