Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Hamburg University of Applied Sciences
Faculty of Life Sciences

Construction of stable HeLa cell lines for the identification of enzyme-induced bias in proximity-dependent labeling techniques

Bachelor thesis
in the study programme BSc Biotechnology

submitted by
Enno, Große Wichtrup,
[redacted]

Hamburg
on 05 July 2022

Reviewer: Prof. Dr. Julien Béthune (HAW Hamburg)
Reviewer: Prof. Dr. Stephan Noll (HAW Hamburg)

The thesis was supervised and prepared in the laboratory for molecular biology and cell culture techniques at HAW Hamburg

Acknowledgements

I would like to take this opportunity to thank all those who supported and assisted me during the preparation of this Bachelor´s thesis.

First, I would like to thank Prof. Julien Béthune, who supervised and assessed my Bachelor´s thesis. I would like to express my thankfulness for the helpful suggestions and constructive criticism during the preparation of this thesis.

Furthermore, my thanks go to Prof. Noll for assessing my thesis as second reviewer and to Elisabeth Schäfer, who supported me in the lab with important advice.

Finally, I would like to thank my parents and my girlfriend, who made my studies possible through their support and always had an open ear for me.

Enno Große Wichtrup

Hamburg, 05.07.2022

Contents

i) List of abbreviations

A

AP-MS .....Affinity purification mass spectrometry
APS.....Ammonium persulfate

B

BCCP.....Biotin carboxyl carrier protein
BPL.....Biotin Protein Ligase

D

ddH2O .....double-distilled water, double-distilled water, double-distilled water
DMEM.....Dulbecco´s Modified Eagle´s Medium
DTT.....Diethiothreitol

E

EF1α ..... elongation factor 1 alpha

F

FBS..... fetal bovine serum
FL……Firefly luciferase
FRT.....Flp-recognition target

H

HRP ..... horseradish peroxidase

M

MS .....Mass spectroscopy

P

PBS.....Phosphate-buffered saline
POI.....Protein of interest
PPI ..... Protein-protein-interaction

R

RMCE ..... Recombinase mediated cassette exchange
rtTA.....reverse tetracycline-controlled transactivator

S

SDS ..... Sodium dodecyl sulfate

T

TEMED ..... Tetramethylethylenediamine
Tet-System .....Tetracycline-controlled transcriptional activation system

X

XL-MS..... Cross-linking mass spectrometry

Y

Y2H.....yeast two-hybrid

ii) List of figures

iii) List of tables

1 Abstract

While different optimized biotin protein ligases (BPL) for proximity dependent labeling techniques are increasingly used in fundamental research, whether or not these enzymes induce bias through their interaction with host cell components has not yet been adequately explored.

This thesis aims to construct the tools which will allow to investigate whether four different BPL´s, used in proximity dependent labeling techniques, could influence the correct targeting of a protein of interest (POI), because of their interaction with cell components.

For this purpose, four stable HeLa cell lines were constructed containing the genetic information for the BPL´s BioID2, TurboID, UltraID or BASU, fused only to a myc-tag, in their genomes. By using the recombinase mediated cassette exchange (RMCE) as the transfection method the isogenicity and with that the comparability of the results was preserved. By varying the dose of the transcriptional activator doxycycline, the dose-response on the expression of enzymes was investigated. Biotinylated proteins were analyzed by western blot, using horseradish peroxidase (HRP) conjugated streptavidin. This allowed a direct comparison of the biotinylation activities of the different enzymes.

All four stable cell lines were constructed successfully and induction was achieved as assumed. For UltraID and TurboID enhanced biotinylation of certain proteins were detected, thus suggesting that enzyme-induced bias may occur.

2 Introduction

2.1 Protein-Protein-Interactions

Protein-protein-interactions (PPI) are physical interactions between two or more proteins caused by intermolecular forces of attraction such as electrostatical forces, hydrogen bonding or hydrophobic effects[1]. They are involved in most biological processes by forming protein complexes which then perform physiological tasks in living cells[1]. Protein complexes formed by PPIs are very diverse and can differ in their composition, affinity, and whether protein binding is permanent or transient[2]. There are an estimated 400.000 PPIs within the human genome[3]. Over 80 % of all human proteins are estimated to not operate alone but in protein complexes[4].

Fundamental biological processes such as metabolic pathways, signal cascades, transcription, translation and cell cycle are highly dependent on PPIs[5–7]. Associated critically important functions such as hormone receptor binding, enzyme allostery, protease inhibition or correction of misfolded proteins by chaperones rely on PPIs to function properly[5,8,9].

As a result protein complexes often play key roles in human disease[10]. Through PPIs harmful protein complexes can form out of endogenous proteins, proteins from pathogens, or both[11]. Diseases such as cancer, infectious diseases and neurodegenerative disease are getting associated with aberrant PPIs[12]. As a result, PPIs are increasingly being perceived as potential drug targets and research is being conducted into novel therapeutics in this area, with some PPIs modulators already approved for marketing[12,13].

In addition, PPIs can be used to deduce the function of a protein. By analyzing what proteins a POI interacts with, a hypothesis can be made about the protein complexes in which the protein is involved and where the protein is localized in the cell. This information provides insight about what the proteins function within the cell is. [14]

Due to the above-mentioned relevance of PPIs in fundamental scientific research, drug discovery and proteomics, analysis of PPIs is of increasing scientific interest and different methods have been developed for their investigation over the years[4].

2.2 Methods for the analysis of protein-protein interactions

With the rapid development in the field of proteomics, different methods have been developed for the identification and characterization of PPIs over the last decades[4]. They can be broadly classified into binary and co-complex technologies as described by Titeca et al.[6 15]. Binary technologies, such as the yeast two-hybrid (Y2H) assay are based on the identification of PPIs between preselected protein pairs[16]. Co-complex technologies include the gold standard, affinity purification coupled with mass spectrometry (AP-MS), and other techniques such as chemical cross-linking of interacting proteins and proximity dependent labeling techniques[6].

A commonly used binary experimental technique for the study of PPIs is the aforementioned Y2H assay, described first by Fields and Song in 1989[16]. To find out whether two proteins interact with each other, one half of a separated *Saccharomyces cerevisiae* transcription factor GAL4 is genetically fused to each of two preselected proteins and expressed[16]. GAL4 is required for the expression of genes which are coding for enzymes of galactose utilization[16]. If the proteins of interest interact with each other, the two halves of the transcription factors that are then in close proximity to each other can resume their function and the enzymes are expressed[16]. That way a galactose selection can indicate whether protein interaction took place[16]. Over the years this technique was further developed in a variety of different versions[17]. Bacterial, yeast and mammalian systems are available with different host strains, vectors and reporter genes[17]. A disadvantage over the co-complex technologies lies in the fact that the interacting protein partners of the POI have to be known in advance and are limited in number[6].

With co-complex technologies previously unknown interacting partners can be identified in great quantity[6]. For example in AP-MS, the cell lysate to be analyzed, which contains the POI, is loaded onto an affinity column[14]. The POI is then captured on a matrix with immobilized antibodies against the POI[14]. Current interaction partners of the POI that are still bound to the protein are, together with the POI, retarded on the column[14]. Proteins which do not interact with the POI are removed from the column during the wash step[14]. If no suitable antibodies against the POI are available, a tag can be fused to the POI by genetic engineering[14]. In this case affinity purification takes place with the corresponding affinity matrix[14]. Subsequently, the purified proteins can be prepared for MS[18]. However, separation of the purified protein by SDS-PAGE is

often used additionally to reduce the sample complexity before MS analysis[18]. A disadvantage of this method is however, that the interacting partners of the POI have to continue the correct binding during lysis of the cells and during affinity purification in order to capture the interacting proteins successfully[14]. Furthermore, only currently interacting partners can be analyzed, and weak or transient partners are likely to be missed[14].

Another co-complex tool for the analysis of PPIs is chemical cross-linking coupled with mass spectrometry (XL-MS)[6]. Here cross-linkers that can covalently tether nearby amino acid residues are added to a protein sample or cell[6]. Interacting proteins are getting cross linked and after a subsequent proteolytic digestion, the cross-linked peptides are analyzed by MS[6]. The distances between the linked residues can provide information for, among other things, subunit topology, molecular modeling and interaction network analyses[6]. With XL-MS transient and weak PPIs can be investigated[6]. However, disadvantages such as the low efficiency of the cross-linking reagents and a cross-linking reaction time of 30 minutes can often worsen the results[6].

Proximity dependent labelling techniques promise to be a good alternative to other co-complex methods for the analysis of PPIs[6]. They are based on the labelling of interaction partners in the direct vicinity of a POI[6]. For this purpose, an enzyme is genetically fused to the POI and expressed in the cell[6]. The fused enzyme labels all proteins that come into the direct vicinity of the POI and are thus presumably interaction partners of the POI[6]. Various enzymes are used for this purpose, such as horseradish peroxidases, ascorbic acid peroxidases or biotin ligases[6]. After disruption of the cell, the labelled proteins can be affinity-purified on the basis of the new label and analyzed with MS[6]. The method has advantages over some fundamental limitations for example in Y2H or AP-MS[19]. The proteins can be expressed in a relatively normal cellular enviroment[19]. Hence the right machinery for post-translational modifications and the normal complement of associated binding partners can be given[19]. Furthermore detection of transient and weak PPIs are possible, because lysis of the cell is set after the labeling[6]. And finally analysis of PPIs among both soluble and membrane proteins is feasible[6].

Due to the promising possibilities of proximity dependent labelling techniques, many variants of the method have been developed. In the next section, a detailed description

of the BioID technique, which uses a biotin ligase, and four improvements are discussed.

2.3 Proximity dependent biotin labeling techniques

2.3.1 BioID

The biotin identification method is a proximity dependent labeling technique that was developed by Roux et al. in 2012 as a new method in the field of PPI analysis in eukaryotic cells[19] .

As briefly described above, for the BioID method a BPL is fused to a POI genetically and expressed in the cell[19]. For this purpose, eukaryotic cells, such as HeLa- or HEK-cells, are transiently or stably transfected with a plasmid, containing the genetic information for the POI fused to a biotin ligase[20,21]. In the case of BioID a mutant of the *E. coli* derived biotin ligase BirA is being used[19]. After expression of the fused proteins and addition of biotin, the biotin ligase mutant begins to biotinylate proteins that are in close proximity to the enzyme and thus represent potential interaction partners of the fusion protein[19]. This is followed by lysis of the cell and affinity capture with streptavidin, which can bind biotin non-covalently with unusually high affinity[19,22]. Subsequent protein identification can be conducted with immunoblot analysis and MS[19]. Figure 1 shows the principle of BioID schematically.

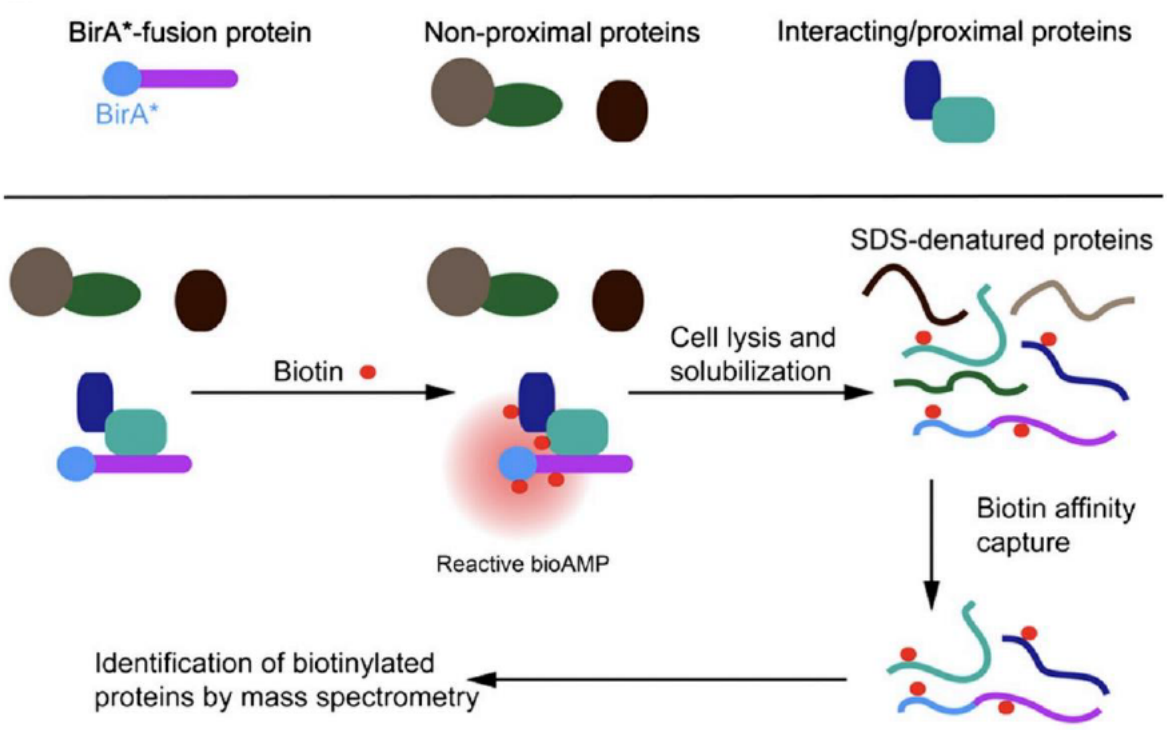

Figure 1: Principle of the BioID method for proximity-dependent biotinylation for the analysis of protein-protein interactions. **Top** Legend of illustrations used. **Bottom** In live cells the biotin ligase BirA* is expressed fused to a POI. After biotin supplementation BirA* biotinylates proteins in the direct vicinity of the fused POI. For the analysis of the biotinylated proteins, which represent potential interaction partners, the cells are lysed and the proteins are prepared with SDS-Page and biotin affinity capture for mass spectroscopy or immunoblot analysis. [19] (Roux et al., 2012)

E. coli`s wild-type protein ligase has a molecular weight of 35 kDa and is a class II protein ligase, which are generally called BirA proteins and function in two ways[19,23,24]. Firstly, it functions as a biotin attachment enzyme and secondly it regulates biotin synthesis by binding to the operator sequences of biotin synthetic operons, subsequently suppressing transcription and biotin synthesis[23,25,26]. The former function includes the major task of catalyzing the biotinylation of a subunit of biotin-dependent carboxylases, which in turn have various functions in the cell[19,23,24]. This biotinylation process catalyzed by BirA is a two-step process[19]. First biotin and ATP are converted to biotinoyl-5'-AMP, which is an activated and highly reactive, yet labile, form of biotin[19,27]. Biotinoyl-5'-AMP is then held in the active site of the enzyme until it reacts with the epsilon amino group of a specific lysine residue of the protein acceptor tag of the carboxylase[19,27,28]. After biotinylation of the biotin carboxyl carrier protein (BCCP) subunit of the carboxylase the apoenzyme becomes a functional holoenzyme (see reaction (1) and (2) and Figure 2)[27].

$$\text{biotin} + \text{ATP} \rightleftharpoons \text{biotinyl} - 5' - \text{AMP} + \text{PPi} \quad (1)$$

$$\text{bio} - 5' - \text{AMP} + \text{apocarboxylase} \rightarrow \text{holocarboxylase} + \text{AMP} \quad (2)$$

Since biotinylation is a rare modification in many organisms including bacteria and mammals, this allows using this modification for labeling of proteins and selective affinity purification with only small background of endogenous biotinylated proteins[19,29]. Therefore BirA was used before in experimental applications such as an acceptor-peptide system[30]. In this technique BirA biotinylates a biotin acceptor tag with high specificity[19 30]. While a binary method for the analysis of PPIs could be based on this technique, for unspecific proximity-dependent biotinylation of proteins certain requirements had to be met[19]. As described in "A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells" from Roux et al. the high specificity of BirA to its acceptor tag had to be overcome and the activated biotin had to be constantly released from the enzyme, regardless of whether the acceptor tag was bound or not[19].

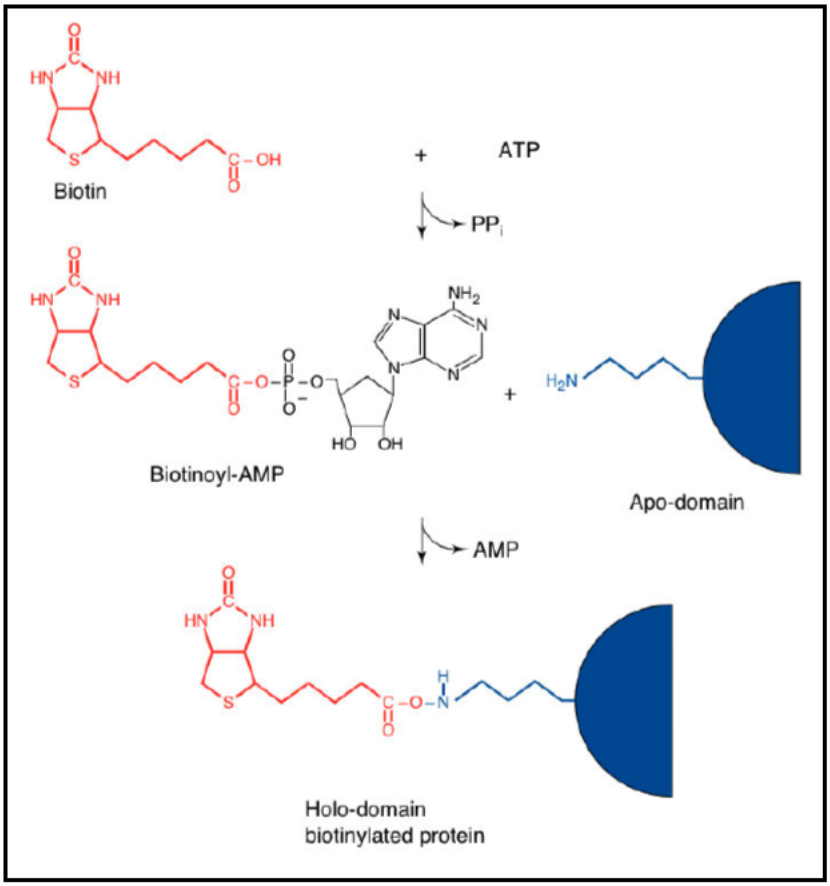

Figure 2: Reaction catalyzed by *E. colis* BirA. Biotin is activated under ATP hydrolysis. Biotinoyl-5'-AMP held in the catalytic center of BirA reacts with the biotin carboxyl carrier protein (BCCP) subunit of the apocarboxylase, yielding the holocarboxylase. [26] (Chapman-Smith et al., 1999)

For this purpose, Roux et al. employed the BirA mutant R118G (hereafter BirA*) for BioID, in which the catalytic domain is destabilized by a single-site mutation[6,19,28]. The mutation is close to the catalytic center of the enzyme, causing the highly reactive biotinoyl-5'-AMP to be released prematurely[19,28,31]. The released molecule then reacts non-specifically with primary amines in its immediate vicinity[19,32]. Due to the lability of

biotinyl-5'-AMP, only proteins in an environment up to approx. 10 nm are biotinylated[19,33].

However, the labeling kinetics of BioID are still rather slow, needing 18-24 hours of incubation with exogenous biotin for sufficient biotinilation[19–21]. In addition, BirA* is comparatively large, which can hinder the correct localization and targeting of the corresponding fusion proteins[21,34]. Finally, as previously mentioned, BirA, as a class II biotin ligase, has a DNA-binding domain and could lead to artefactual non-specific binding to host DNA or chromatin-interacting proteins[21,35,36].

An alternative similar technique allows for labeling times, down to 1 minute with a smaller enzyme called APEX2. However, APEX2 is a peroxidase which is activated with H_2O_2. The toxicity of H_2O_2 limits a broad application of this enzyme. Thus improving BPL-based enzyme has been pursued. [14,20]

Therefore, a number of new biotin ligases have been developed for proximity-dependent labelling in the last 10 years. Aim of new BPLs was to counter the disadvantages of the BioID method described above, trying to improve versatility and speed of the method[21]. In the following section, the four biotin ligases BioID2, BASU, TurboID and UltraID, which were developed as such improvements of the original BioID, are described further.

2.3.2 BioID2

For BioID2 Kim et al. applied the smallest known biotin ligase to that date in 2015, which was the biotin ligase derived from *Aquifex aeolicus*[34]. BioID2 consists of 233 amino acids compared to 321 amino acids in BirA* and has a molecular weight of 26,4 kDa compared to 35 kDa of BirA*[34]. The reduced weight has to do with the missing DNA-binding domain in the class I biotin ligase (see Figure 3)[34].

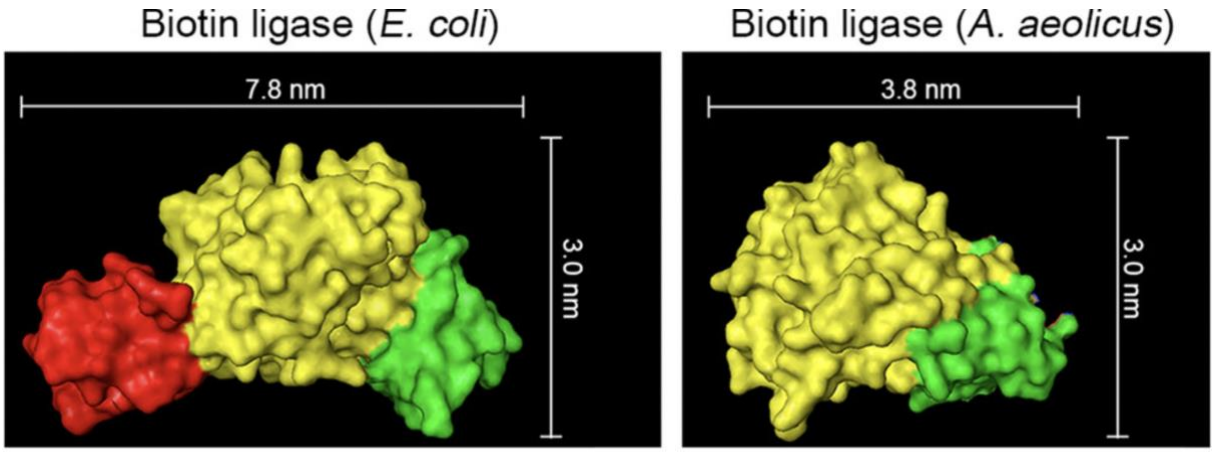

Figure 3: Comparison of two biotin protein ligases used in proximity dependent labelling techniques. **Left** BioID biotin ligase BirA (PDB ID 1BIA) derived from *E. coli*. **Right** BioID2 biotin ligase (PDB ID 2EAY) derived from *A. aeolicus*. The catalytic parts of the enzymes are colored in yellow, ATP- binding site in green and DNA-binding domains in red[34]. (Kim et al., 2016)

Due to the reduced size of the enzyme, the POI is likely to have a better chance of interacting with its interaction partners without being affected by the biotin ligase[34]. In addition to more selective labelling of fusion proteins, the biotin ligases described by Kim et al. also proved to be more efficient at labelling proteins in close proximity[34]. Furthermore less biotin supplementation is required compared to BioID[34].

However the required biotinylation time is still between 12 - 24 hours for both, BioID and BioID2 and they were later shown to have similar enzymatic activity, but not increased[20,21]. Furthermore, whether or not the enzymes used in BioID and BioID2 induce a bias by interacting with cell components has not been thoroughly investigated. Since both originate from prokaryotic bacteria, it cannot be ruled out that the biotin ligases interact with components that are normally not found in bacteria or that they are affected by unknown modifications or different quantities of cell components then found in mammalian cells.

2.3.3 BASU

BASU is a variant of the BirA biotin ligase from *Bacillus subtilis* and was introduced in 2018 by Ramanathan et al.. It was generated by deleting the N terminus (amino acid 1–65) and introducing amino acid substitutions (R142G, E323S, G325R) in the *B. subtilis* biotin ligase sequence (UniProt ID: P0CI75)[37]. Like BioID2 BASU is a class I biotin ligase and lacks a DNA-binding site, making it smaller in size[21,38]. It consist of 259 amino acids compared to 321 amino acids in the E. coli BirA* and has a molecular weight of 28 kDa, which is 7 kDa less than that of BioID[6,37]. As with BioID2, BASU has

first been shown to have enhanced biotinylation activity, although to what extent is again under debate[20,21,37].

2.3.4 TurboID

As described in “Efficient proximity labeling in living cells and organisms with TurboID” from Branon et al. TurboID was engineered by a direct mutation approach of the *E. coli* BirA-R118S biotin ligase[20,21]. Relative to the wild-type BirA, TurboID has 15 mutations[20]. It has the same size as BioID with a molecular weight of 35 kDa[20]. However TurboID catalyzes proximity labeling with much greater efficiency then BioID, BioID2 and BASU[20]. Also the biotinylation time of 10 minutes allows for an extended use, for example enabling application in flies and worms[20].

Nevertheless, TurboID also has some drawbacks. It was shown that caused by the high biotinylation activity, TurboID is drawn to background labeling even before the addition of extra biotin[20,21,39]. Furthermore there are concerns about the inducibility of biotinylation and ligase stability[39]. In addition, when constitutively expressed in mammalian cells or model organisms TurboID can lead to toxicity for the cells[20,21,39].

The extent to which proteins biotinylated by TurboID differ from that of proteins biotinylated by other BPLs has not yet been investigated. Whether it makes a qualitative difference if biotinylation was carried out with TurboID or one of the other BPLs remains unclear.

2.3.5 UltraID

UltraID was introduced by Zhao et al. in 2021 in the paper “ultraID: a compact and efficient enzyme for proximity-dependent biotinylation in living cells”[21]. When Zhao et al. tried to develop a split-BioID2 assay, they split BioID2 between amino acid K^{171}/S^{172}. One of the resulting fragments (BioID2 [2-171]), which contained the catalytic site, was shown to be an enhanced activity enzyme to BioID[21]. One subsequent mutation (L41P) further increased enzymatic activity and yielded UltraID, which is able to outperform current enzymes for proximity dependent biotinylation[21]. UltraID is the smallest biotin ligase to date with a molecular weight of under 20 kDa[21]. Moreover the enzyme kinetics are similar to those of TurboID but with less background activity[21].

The small size of the BPL could lead to the POI being less disturbed by the attached enzyme and thus to an improved targeting. However, whether and to what extent there is a bias due to interactions of the BPL with host cell components has not been investigated in more detail here either.

2.4 Aim of the thesis

The aim of this work was to construct four cell lines that stably express the four single BPLs BioID2, BASU, TurboID and UltraID at identical levels. The cell lines were to form the basis for further investigations to determine enzyme-induced bias in proximity-dependent labelling technique. Additionally in a first characterisation, western blots were used to proof a successful construction of cell lines and to draw initial conclusions on an enzyme-induced bias.

3 Results

The aim of the experiments carried out for this work was, to construct four cell lines, each stably expressing one of the four biotin ligases BioID2, BASU, TurboID or UltraID, each fused to a myc-tag, but without being fused to a POI. This was to enable further investigation on whether the BPLs react with host cell components, thus possibly inducing a bias. Furthermore, a first characterization in the context of that investigation was conducted by screening for the expressed biotin ligases and biotinylated proteins by western blot analysis. The constructed cell lines should in the future allow for further research on bias induction by the respective biotin ligases.

Plasmid stocks for transfection of the master cell line were prepared by replicating existing plasmid stocks in E. coli and subsequent plasmid preparation.

In order to test the prepared plasmids, intended for stable transfection, HeLa-EM2-11ht cells (master cell line) were transfected transiently with each plasmid stock, containing the genetic information for the aforementioned BPLs. After enzyme expression and biotinylation, the four transient cell lines were lysed and expression of the BPLs as well as biotinylation activity was characterized by western blot analysis.

Subsequently HeLa 11ht cells were transfected stably with the prepared plasmid stocks and after biotinylation the preliminary characterization of the BPLs, using

immunoblotting, was conducted. The yielded enzyme expression was analyzed by detection of the myc-tag with anti-myc antibodies and the yielded biotinylation by detection of biotin-tagged proteins with streptavidin-HRP conjugate.

Additionally, dose response to the transcriptional activator doxycycline was investigated by varying the supplemented concentration and subsequent comparison of the yielded BPLs expression levels and biotinylation activity.

3.1 Plasmids

For the transfection of HeLa-EM2-11ht cells four plasmid-stocks were prepared from already existing stocks, since more plasmid was required than was available. The plasmids contained the genetic information for BioID2, BASU, TurboID or UltraID each fused only to a myc-tag. By not fusing the BPL to a fusion protein, it was intended to illustrate whether and to what extent the enzymes in the cell interact with cell components. This should then allow to draw conclusions about biases that are induced by the enzymes. Using the fused myc-tag, the expression level of the enzymes could be monitored in the western blot. Since the myc-tag was fused to all four enzymes, a common detection reagent could be used for all cell lines. Therefore, a direct comparison of the expression level on one blot was enabled.

The stock was prepared by transforming competent DH-5α *E. coli* cells with the already existing plasmids pSF3-BioID2, pSF3-BASU, pSF3-ultraID and pSF3-TurboID. Figure 14 to Figure 17 in the appendix show the four plasmid maps of the plasmids used for the experiments of this thesis. After transfection and cultivation of the *E. coli* cells a plasmid preparation was performed, to purify the plasmids. The plasmids carried an ampicillin resistance gene for the selection of positively transfected cells, which was made use of. Resulting plasmid concentrations are shown in Table 1.

Table 1: Concentrations of the indicated pSF3-plasmid stocks after replication in *E. coli* and plasmid preparation for transfection of HeLa-EM2-11ht cells. Measured with spectrophotometer NanoDrop 1000.

	pSF3-BioID2	**pSF3-BASU**	**pSF3-ultraID**	**pSF3-TurboID**
Concentration in ng/mL	20	62	37	32

Due to the fused myc-tag, the expressed proteins gained a molecular weight of 1,2 kDa each. For one of the plasmids, pSF3-BioID2, a stop codon was added directly after the coding sequence of the enzymes. For the other three plasmids, pSF3-BASU, pSF3-TurboID and pSF3-ultraID, a stop codon further downstream in the MCS of the plasmid was used to stop translation. This adds 31 amino acids to BASU, 31 amino acids to TurboID and 30 amino acids to UltraID. Therefore, the molecular weight of these three expressed proteins differs from the molecular weight of the single BPL. This results in the molecular weights listed in Table 2.

Table 2: Comparison of molecular weights between BPLs used in proximity dependent labelling techniques and same BPLs as encoded on pSF3 plasmids, including myc-tag and amino acid overhang. **Top** Exact molecular weight of the indicated biotin protein ligases used in proximity dependent labeling techniques. **Bottom** Molecular weight of the expressed proteins as coded for on the plasmids pSF3-BioID2, pSF3-BASU, pSF3-ultraID and pSF3-TurboID. Open reading frames on the plasmids pSF3-BASU, pSF3-ultraID and pSF3-TurboID include, next to the genetic information for the BPL and myc-tag the genetic information for an amino acid excess that results from the location of the stop codons.

	BioID2	**BASU**	**UltraID**	**TurboID**
Molecular weight of the BPL in kDa	26,4	28	20	35
Molecular weight of the expressed protein	27,9	33,7	39,9	39,9

Transcription of the genetic information coding for the protein ligases, each fused to the myc-tag, were under the control of the bidirectional Ptet promoter, as seen in Figure 9. As a part of the tetracycline-controlled transcriptional activation system (Tet-System) the Ptet-controlled transcription unit is only transcribed when doxycycline (a tetracycline derivative) is present. This allowed for the expression of the protein ligases to be controlled. Once doxycycline is present, it binds to the uniformly produced reverse tetracycline-controlled transcriptional activator rtTA2S-M2, which in turn binds to the operator and activates transcription by recruiting the relevant RNA polymerase (see Figure 4). [40]

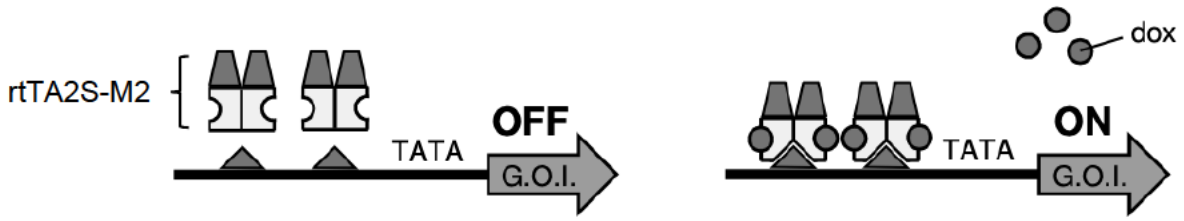

Figure 4: Tetracycline-controlled Tet-On gene expression systems used to regulate the activity of genes in stable HeLa-EM2-11ht cells for the inducible expression of protein ligases for a proximity dependent labeling technique. In the presence of doxycycline, the transcriptional activator rtTA2S-M2 binds doxycycline and thus triggers a conformational change, enabling binding to the operator sequence with subsequent activation of transcription. [40] (figure adapted from Das et al., 2016)

3.2 Transient transfection and characterization of the transient HeLa-EM2-11ht cell lines

HeLa-EM2-11ht cells were transiently transfected to test the prepared plasmid stocks pSF3-BioID2, pSF3-BASU, pSF3-ultraID and pSF3-TurboID. The aim was to test whether this transient transfection of HeLa cells would enable expression of the BPLs and whether the expected biotinylation of cellular proteins through the BPLs would occur. In a transient transfection the transgene is not incorporated in the genome of the cells as it would be the case for stable cell lines, therefore allowing testing the plasmids quickly and easily before starting with the lengthy process of making stable cell lines.

This pre-test was conducted by transfecting HeLa-EM2-11ht cells with the plasmid stocks containing the genetical information for the four BPLs, BioID2, BASU, TurboID and UltraID, which were each fused only to a myc-tag. Expression of the BPLs was controlled by the Tet-system and was induced by the tetracycline derivative doxycycline. In addition to doxycycline, biotin was added, and incubation was carried out overnight. The additional biotin should provide the required amounts of reagents for the biotinylation, catalyzed by the BPLs. After cell lysis characterization of the expression level of the BPLs by detection of the myc-tag and characterization of biotinylation activity, by detection of the biotinylated proteins, was conducted. Therefore, proteins were separated on an SDS-PAGE and after blotting labeled with anti-myc primary antibody and fluorescence secondary antibody. Subsequently streptavidin-HRP conjugate was used for the detection of biotinylated proteins. 10 µg protein per well was loaded, which was ensured by performing a Bradford assay. This should later enable a comparison of biotinylation activity between the enzymes.

For BioID2 no proteins were found in the cell lysate as detected with the Bradford assay. This was probably due to incorrect scrapping of cells conducted for this cell line. Therefore, no lysate was loaded onto the SDS-gel from the cells that were transfected with BioID2.

Figure 5 shows the process flow of the transient cell line construction and subsequent characterization.

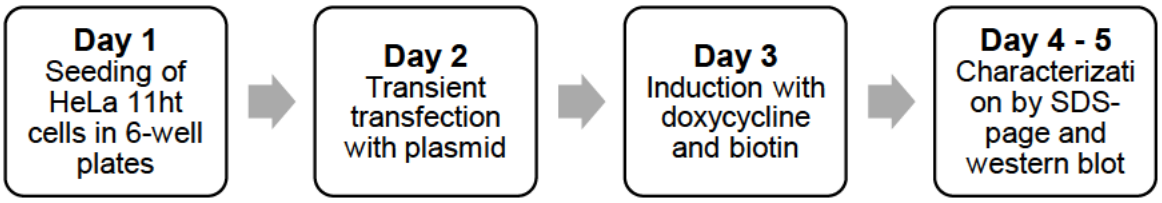

Figure 5: Timeline and process description for the construction of transient HeLa cell lines for proximity dependent labelling and subsequent characterization of BPLs expression level and biotinylation activity.

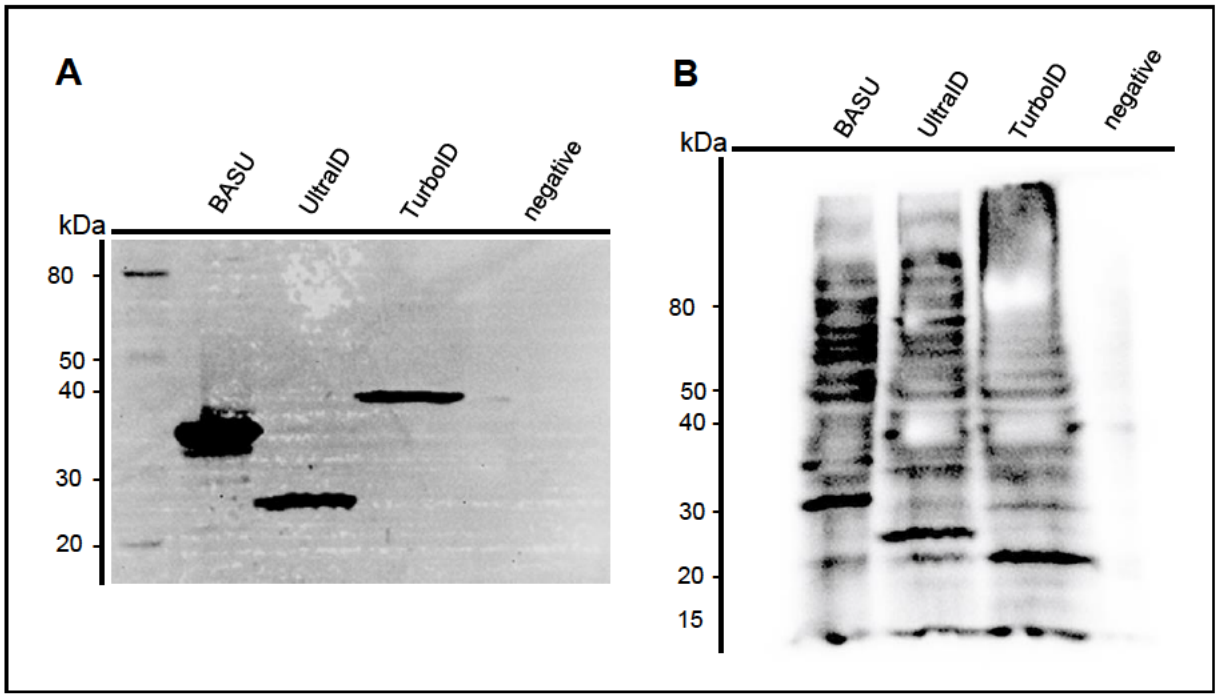

Figure 6: Blots of lysates of HeLa-EM2-11ht cells transiently transfected with the indicated biotin protein ligases for proximity dependent labeling technique. All cell lines were incubated overnight with 50 µM biotin. Negative cells were untransfected HeLa-EM2-11ht cells. **A** Enzyme expression was investigated by IRDye 800CW labeled anti-myc antibody. **B** Biotinylation activity was analyzed by Streptavidin-HRP conjugate.

As can be seen in Figure 6, expression of the BPLs and biotinylation activity was detected for each of the three cell lines. In Figure 6A, a significantly thicker band is visible for BASU than for UltraID and TurboID. While the labelled proteins from the cell lysate of the BASU-transfected cells can be classified at approx. 35 kDa, the potential UltraID is located at approx. 36 kDa and TurboID at approx. 40 kDa. It can be observed

that the determined molecular weights corresponded to the expected weights (see Table 2). In addition, as expected, no bands were detected in the negative control. Control cells were not transfected with a plasmid.

The blot in Figure 5B showed high biotinylation activity of the three BPLs BASU, UltraID, and TurboID. Compared to the negative control cells, the number of biotinylated proteins is greatly increased in cells that expressed one of the BPLs. For TurboID and UltraID, the chemiluminescent light emission in some bands already decreased, probably due to the depletion of the substrate. Nevertheless, the number of bands suggests that many proteins were biotinylated from all BPLs. Many of the labelled proteins were labelled by all three BPLs, as can be seen from the same positions of the bands. However, the intensity of the bands varies considerably among them.

Enzyme expression and biotinylation activity were shown for BASU, UltraID and TurboID. Since the plasmid of BioID2 was prepared in the same way as the others, it was assumed, that protein expression and biotinylation would have been detected for BioID2 as well, if protein lyses had been successful.

3.3 Construction and initial characterization of stable HeLa-EM2-11ht cell lines for enzyme-induced bias investigation

Four HeLa cell lines, each stably expressing a BPL without a fusion protein, were constructed for the purpose of providing a basis for the investigation of enzyme-induced bias in proximity dependent labelling technique.

The construction of the four stable cell lines was performed in two successive rounds. In the first round, the construction of the BioID and BASU cell lines was conducted, followed by the construction of TurboID and UltraID beginning one week later. The shift was intended to distribute the workload more evenly. The general experimental procedure is depicted in Figure 7.

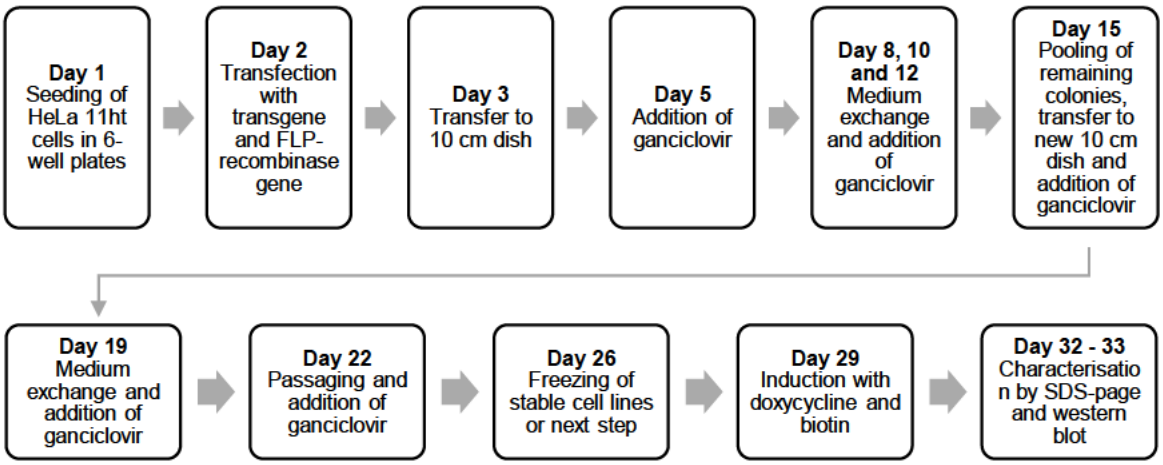

Figure 7: Timeline and process description for the construction of a stable HeLa cell line for proximity dependent labelling.

The general strategy was to integrate the expression cassette on the pSF3 plasmids in the genome of the master cell line HeLa-EM2-11ht. This way a stable integration should be performed. The cassette is depicted in Figure 8.

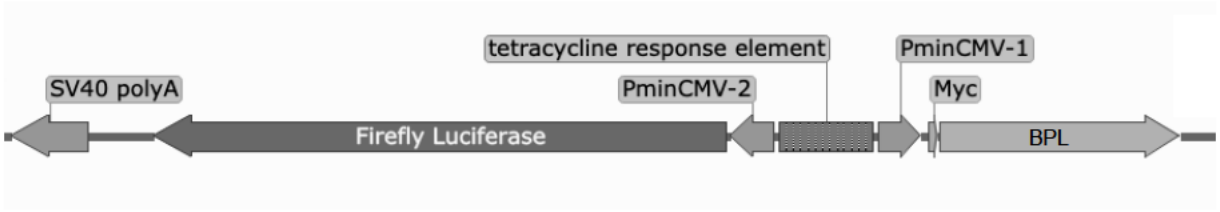

Figure 8: Schematic figure of a cassette, introduced to the genome of the master cell line HeLa-EM2-11ht for the generation of a stable cell lines for proximity dependent labeling technique. Expression of the BPL (could be BioID2, BASU, UltraID or TurboID) and Firefly luciferase (FL) with their respective tag were each under control of the Tet-system for inducible activation of transcription. Firefly luciferase was not of importance for this thesis. The gene was not removed, as no interference was expected from its expression. PminCMV-1 and PminCMV-2 are promoters in the Tet-system. Tetracycline response element is bound by the transcriptional activator rtTA2S-M2 when doxycycline is present, thus activating transcription.

For all experiments conducted for this thesis, the provided HeLa-EM2-11ht master cell line from Mannheim was used, which was a kind gift from Dr. Schönig from the Zentralinstitut für Seelische Gesundheit Mannheim. Cells of this cell line express rtTA2S-M2, which is required for the applied Tet-system. It also contains a locus with recombination sites, which was chosen for two reasons. First, no expression of a transgene, under control of the Tet-system, is carried out here in the absence of doxycycline. Second, robust expression of the same transgene can be achieved when doxycycline is added.

For the construction of that cell line Weidenfeld et al. genetically altered HeLa cells, until the master cell line HeLa-EM2-11ht, was obtained[41]. Weidenfeld et al. first constructed HeLa-EM2 cells, which stably and uniformly expressed the transcription

activator component rtTA2S-M2, as it was under the control of the human elongation factor 1 alpha (EF1α) promoter. EF1α promoter allows for efficient homogeneous expression of rtTA2S-M2, which is a necessary control element for the Tetracycline-controlled transcriptional activation system (Tet System) used in the recombination locus[41]. Weidenfeld et al. then identified a specific genomic locus, in which, when a specific cassette was inserted as a single copy, transcription was highly inducible in a silent but activatable manner by the Tet-system. The clone, containing the cassette at this locus, was called EM2-11 and showed the best regulation properties. Meaning, that expression of the inserted transcription site only took place, when doxycycline (a tetracycline derivative) was added. Weidenfeld et al. inserted a hygtk cassette, in the predefined locus by a special recombination technique. This cassette contained a hygtk positive/negative selection cassette. The resulting cells were called HeLa-EM2-11ht cells and represent the master cell line used in this work.

Since the aim was to compare different BPL in different cell lines, a system was chosen for recombination of the transgene that allowed insertion into the genome of all cells (HeLa-EM2-11ht) in the same way. This was to ensure stable and uniform expression in all cell lines and thus comparability of results. Therefore, the RMCE system as described by Schlake and Bode in “Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci” was used[42]. The system was supposed to ensure at which point in the genome and in which orientation the genetic information for the protein ligases is integrated. Due to the resulting isogenicity, the same expression level for the enzymes could be expected. The accuracy of the system has already been verified in other work[43]. With other systems that work by random integration of the transgene in the genome, it could not be ensured that each cell line expressed the same amount of the BPL at the same time. The transcription rates would probably vary in that case. By using RMCE, however, any differences in the detected BPL levels must result from the expressed enzyme itself, since all the BPL genes should be transcribed at similar levels.

The specific cassette that was inserted in the genome of HeLa-EM2-11ht cells by Weidenfeld et al. contain heterospecific Flp-recognition target (FRT) sites (called F3 and F in Figure 9). These sites enabled the Flp-mediated RMCE by double reciprocal crossover between identical pairs of FRT sites. These need to be present on both, the described cassette in the genome of the HeLa-EM2-11ht cells and for example an introduced target plasmid. Thereby the cassette allows the insertion of any transgene

flanked by FRT sites into this specific locus, by Flp-mediated RMCE. The transgene inserted then also needs to be under the control of the inducible Tet-System to allow the inducible and predictable expression, as described by Weidenfeld et al..

The hygtk cassette, present in the preselected locus of HeLa-EM2-11ht cells, allowed for selection, as cells that incorporated this cassette in their genome were hygromycin resistant. The hygtk gene codes for hygromycin-B-phospho-transferase and for the herpes simplex virus thymidine kinase (TK). Hygromycin-B-phospho-transferase causes cells to become resistant to hygromycin, while TK phosphorylates ganciclovir, converting it into a cytotoxic reagent, which in turn causes cells to become sensitive to ganciclovir[44]. Since the aim for the further experiments in this thesis was to work only with cells that contained the cassette, selection with hygromycin was made use of, by adding hygromycin to the HeLa-EM2-11ht cells in culture until before transfection. When this locus is targeted again by another cassette via RMCE, recombination occurs and the hygtk cassette is lost. This makes the recombined cells sensitive to hygromycin and resistant to ganciclovir, thus allowing a negative selection. Figure 9 shows the process of Flp-mediated RMCE as conducted for this work.

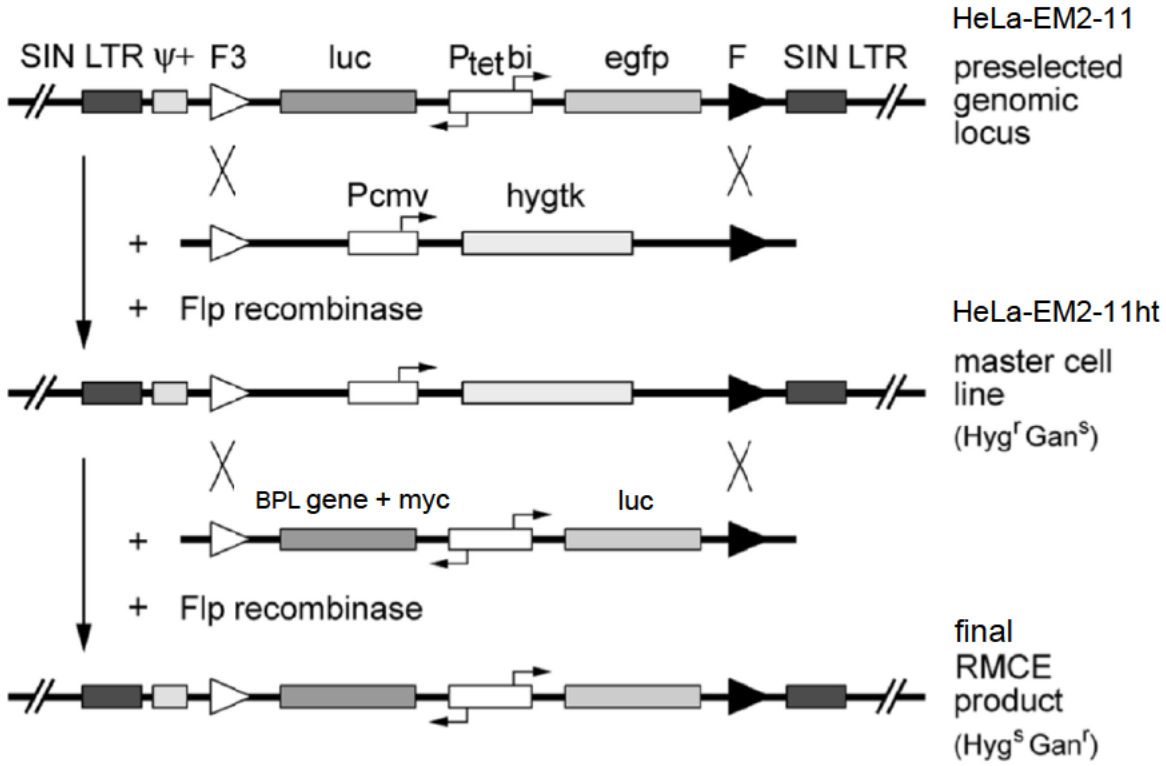

Figure 9: Experimental flow in Flp-mediated RMCE for the construction of stable HeLa cell lines for proximity dependent labeling: HeLa-EM2-11 cells from Weidenfeld et al. contained the illustrated cassette as a single copy in a preselected genomic locus. Luciferase and EGFP expression were under the control of the bidirectional Ptet promoter and was used for prior identification of a highly inducible locus by Weidenfeld et al.. Flp-mediated RMCE cassette exchange, performed by Weidenfeld et al., resulted in the HeLa-EM2-11ht master cell line. The introduced hygtk gene brought resistance to hygromycin and sensitivity to ganciclovir. This enabled cell selection. In a second round of Flp-mediated RMCE, performed for this thesis, four stable HeLa cell lines were constructed, containing the genetic information for the protein ligases BioID2, BASU, TurboID and UltraID, each only fused to a myc-tac. The resulting final RMCE products caused ganciclovir resistance and hygromycin sensitivity in the cells for further selection. [41] (figure adapted from Weidenfeld et al., 2009)

The transcription sites on the plasmids used for transfection of the HeLa-EM2-11ht cells contained, next to the gene for the respective protein ligase and the myc-tag gene, a gene for the Firefly luciferase (FL). This protein also was under control of the bidirectional Ptet promoter (see Figure 8 and Figure 9). It was therefore presumably also expressed after doxycycline addition. However, the luciferase was a by-product and not relevant to this thesis and was therefore not investigated further. It was a remnant on the plasmid and was not specifically removed from the plasmid as no interference was expected from the expression of the protein.

3.3.1 Co-transfection of the master cell line

HeLa-EM2-11ht cells were seeded in 6-well plates and transfected with the plasmids containing the genetic information for a BPL and a plasmid containing the genetic

information for the FLP-recombinase. Flp-recombinase is an enzyme needed for the integration of the BPL-transgene in the genome of the HeLa cells by RMCE. The transfection conditions are stated in Table 3.

Table 3: Conditions for the stable transfection of HeLa-EM2-11ht cells for the investigation of enzyme-induced bias in proximity dependent labelling techniques. Amounts of expression plasmids pSF3-BioID2, pSF3-BASU, pSF3-TurboID and pSF3-UltraID as well as of the expression vector for the FLP-recombinase pJB327-pPGKFLPobpA are stated. Transfection reagent was Lipofectamine 2000. Corresponding volumes are given. See also Table 1.

	pSF3-BioID2	**pSF3-BASU**	**pSF3-TurboID**	**pSF3-UltraID**	**pJB327-pPGKFLPobpA**	**Lipofectamine 2000**
DNA mass in µg	1,5	1,5	1,5	1,5	1,5	-
Volume in µL	75,0	24,2	46,9	40,5	3,0	5,0

3.3.2 Selection of recombined cell lines

The transfected cells were then transferred to a 10 cm dish. Selection of cells which successfully integrated the transgene in the genome was conducted by adding ganciclovir to the cultures. After 15 days clear colonies formed, consisting of cells that must have successfully integrated the transgenes. For each stable cell line, 3 of these single colonies were picked and frozen, for any possible further experiments. The cultivation of the picked colonies required 6 additional days. However, the picked colonies were not further used in experiments for this thesis. Additionally, after picking of the three single colonies all colonies of each cell line were pooled. The pooling was carried out to get a more homogenous population of stable cell lines, by diluting the clonal effects. Since each colony is likely to have slightly different characteristics, this should allow for a more representative result. This approach was therefore assumed to be more appropriate to identify the differences between different BPLs.

3.4 Characterization of the stable cell lines

To show that the construction of stable cell lines was successful, BPL levels and biotinylation of these BPLs were detected. The inducible gene expression using the Tet-system was analyzed by detecting the dose response of the cell lines to doxycycline.

For these purposes, the four constructed cell lines were each incubated overnight with different concentrations of doxycycline and 50 µM biotin. Subsequently, after cell lysis and SDS-PAGE, the present BPL levels, were examined by detection of their myc-tags in two western blots (see Figure 10). Subsequently, biotinylated proteins were detected on the same membrane and visualized by streptavidin-HRP conjugate (see Figure 11).

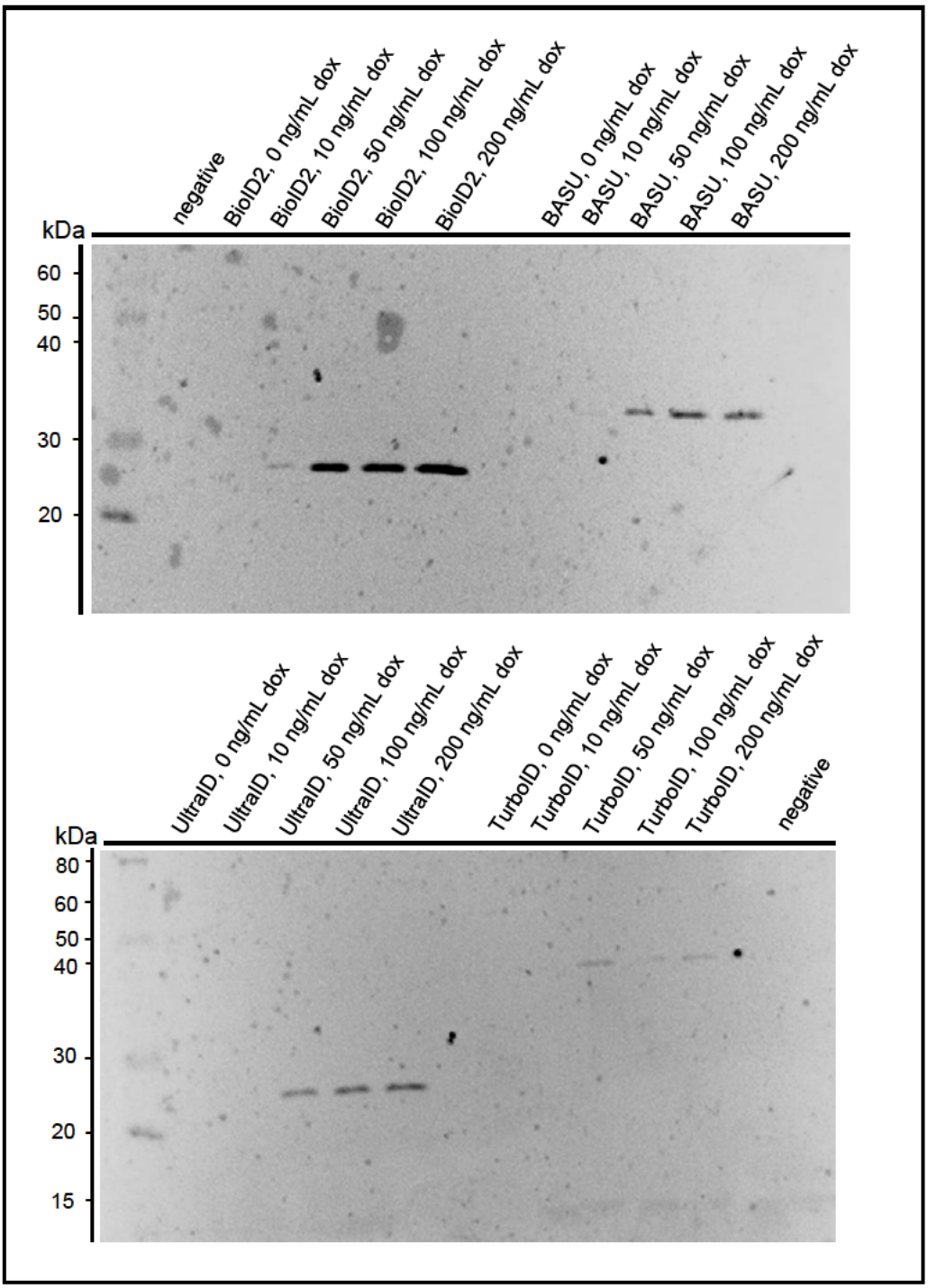

Figure 10: Blots of lysates of HeLa-EM2-11ht cells stably transfected with the indicated biotin protein ligases for proximity dependent labeling technique. All stable cell line were incubated with variable doxycycline concentration as indicated and 50 mM biotin overnight. Enzyme expression was investigated by IRDye 800CW labeled anti-myc antibody, detecting the fused myc-tag. Negative cells were HeLa-EM2-11ht stably expressing BioID2 without doxycycline and biotin addition.

Even though the enzyme expression was supposed to be uniform in all cell lines and the loaded protein mass was regulated to 10 ug per well, location and intensity of bands differed from one cell line to another. Thus, at the time of western blotting, different amounts of BPLs were present in the cell lines at same doxycycline concentrations. Furthermore, the general enzyme concentration over all cell lines can be classified from strongest to least strong: BioID2, BASU, UltraID, TurboID. This is despite the fact that the isogenicity and the Tet-System were supposed to ensure a uniform expression of the enzymes. In the lysate of BioID2 transfected cells, BPL expression was still detected at 10 ng/mL. The same applies to lysate of cells transfected with BASU. UltraID and TurboID, meanwhile, could no longer be detected with the applied method at 10 ng/mL doxycycline induction. As expected, the intensity of the bands differs at different doxycycline concentrations within a cell line. In general, it can be observed that the lower the doxycycline concentration, the lower the intensity of the bands. At 0 ng/mL doxycycline, no band is detectable for any of the four cell lines. Thus, the silent but activatable Tet system was observed to be generally functional.

In addition to the concentration levels of the BPLs, a comparison of the biotinylation activity was of interest. The aim was to investigate the similarities and differences in the biotinylation activity of the four different BPLs, BioID2, BASU, UltraID and TurboID. From the western blots obtained, it was the goal to already draw initial conclusions about the extent to which the various enzymes provoke bias, because they interact with cell components. Therefore, all stable cell lines were incubated with equal concentrations of biotin in addition to the variable concentrations of the doxycycline. The extra biotin was supposed to enable the biotinylation of the interaction partners of the BPLs. Biotinylation was catalyzed by the simultaneously expressed biotin ligases as described in the introduction. Biotinylated proteins were visualized with a streptavidin-HRP conjugate.

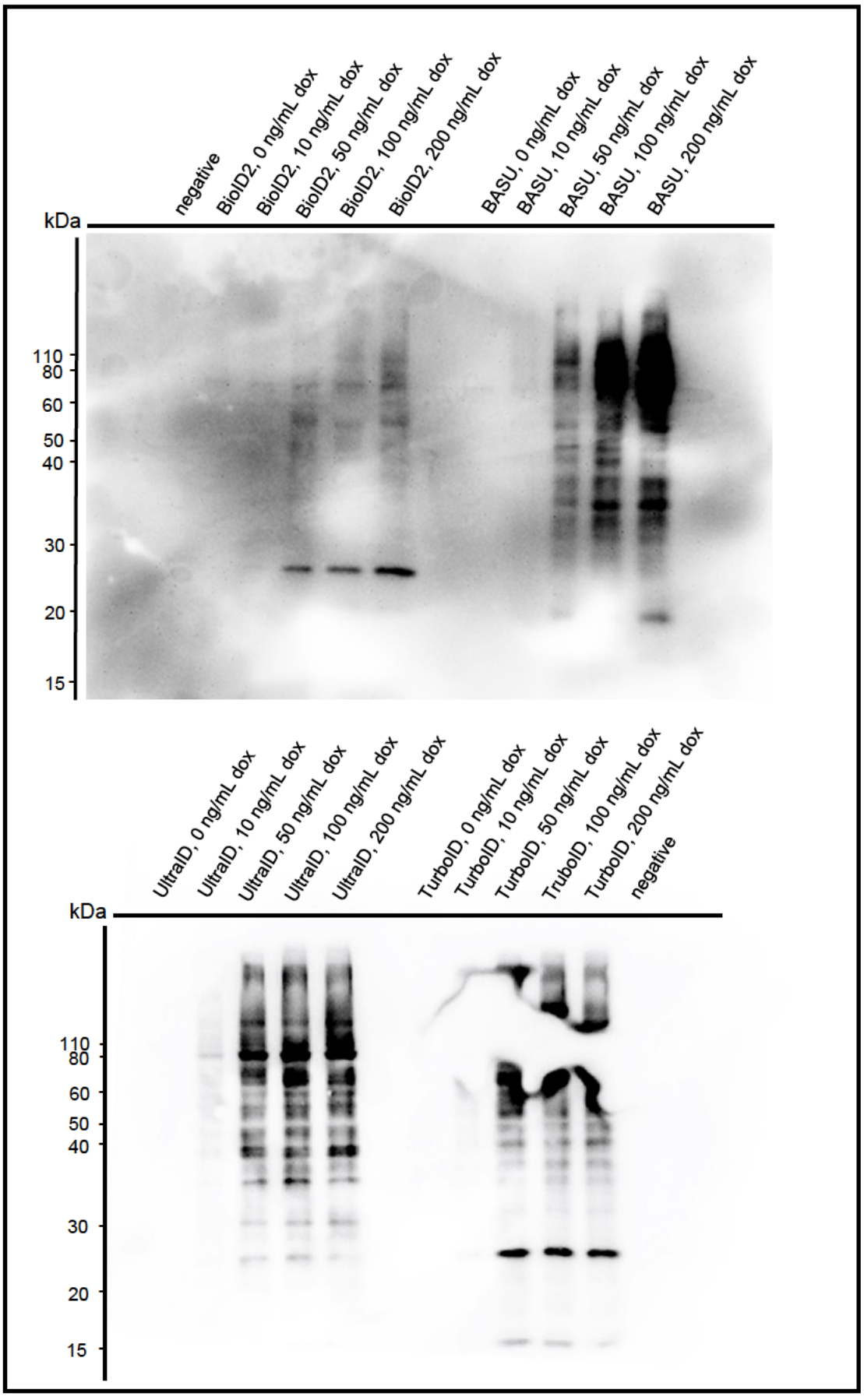

Figure 11: Blots of lysates of HeLa-EM2-11ht cells stably transfected with the indicated biotin protein ligases for proximity dependent labeling technique. All stable cell lines were incubated with variable doxycycline concentrations as indicated and 50 mM biotin overnight. Biotinylation activity of the indicated enzymes was investigated by Streptavidin-HRP conjugate, detecting the biotin tag. Negative cells were HeLa-EM2-11ht stably transfected with BioID2, incubated without doxycycline and biotin addition. Previously, the expression level of BPL was investigated on the same membrane using primary myc antibodies and secondary fluorescent antibodies (see Figure 10).

The quality of the two blots in Figure 11 differ from one another, because slightly different exposure times and cover foils were chosen. In addition, in the bottom blot from Figure 11, a large disturbance from about 50 kDa upwards can be seen in the columns containing TurboID. This could have been caused, for example, by an air bubble blocking the protein transfer during western blotting, or by a lack of substrate for the HRP at this location. Due to this error, bands in the mentioned area are partly or completely invisible.

It was observed that each of the four BPLs, when present, increased the number of biotinylated proteins in the presence of added biotin. Thus, biotinylation activity was successfully demonstrated for each BPL. Within a cell line, higher biotinylation activity was seen where more of the respective enzyme was previously detected. However, when the cell lines were compared with each other, it was noticeable that the biotinylation activity was not highest where the most enzymes were detected before. Rather, although UltraID always had the least enzyme present at all doxycycline concentrations, it exhibited one of the highest biotinylation activities. This could be seen from the high intensity and number of bands in Figure 11. TurboID and BASU also showed a high enzyme activity. BioID2, on the other hand, generally showed a weaker activity compared to the remaining three enzymes, although here the most enzyme was detected before.

At 0 ng/mL doxycycline, a band between 60 and 80 kDa was slightly visible only for BioID2 and BASU. In contrast, no bands were detectable in the lysate of UltraID and TurboID transfected cells, incubated at 0 ng/mL doxycycline.

The positions of the bands within a cell line at different doxycycline concentrations were the same. As the concentration of doxycycline decreased, the intensity of the bands also decreased. When the band positions between the different cell lines was compared, it was noticeable that some bands were in the same position. However, the intensity of that bands often varied between the different cell lines.

3.5 Comparison of expression and activity of the BPLs

In a third blot, the direct comparison of the four stable cell lines was to be carried out in a single western blot. The biotinylation activity of the four BPLs BioID2, BASU, UltraID and TurboID should be preliminarily investigated regarding induced biases. This should allow initial conclusions to be drawn about the behavior of the BPLs in the

cells. Lysates, with prior 50 ng/mL and 200 ng/mL doxycycline induction, were used. Additionally, tubulin was detected, using an additional specific antibody, as an internal loading control.

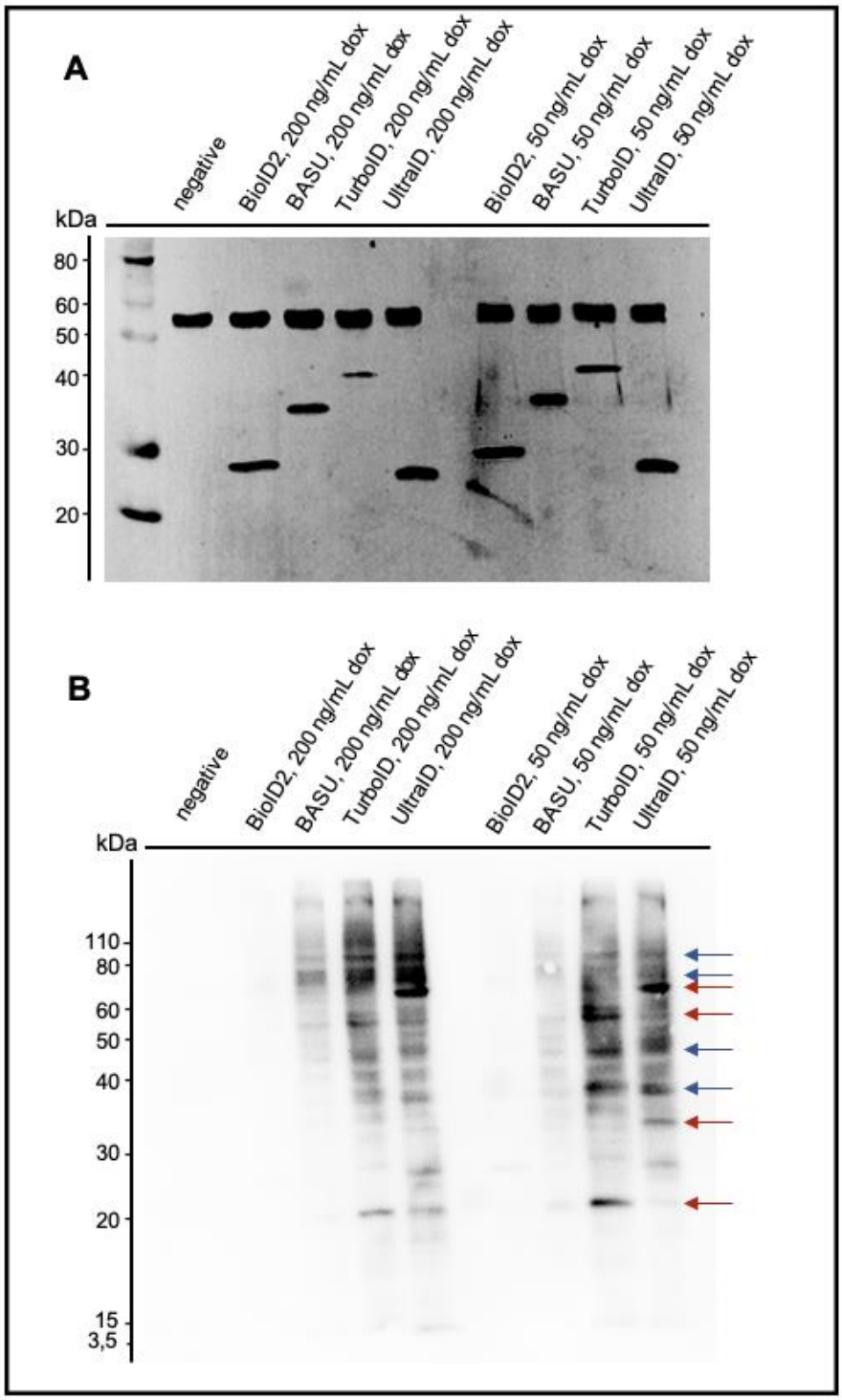

Figure 12: Blots of lysates of HeLa-EM2-11ht cells stably transfected with the indicated biotin protein ligases for proximity dependent labeling technique. All stable cell lines were incubated with variable doxycycline concentration as indicated and 50 µM biotin overnight. In addition to the cell lysates, tubulin fused to a myc-tag was added in each well as an internal loading control. Negative cells were HeLa-EM2-11ht stably transfected with BioID2 without doxycycline and biotin addition. Blue arrows mark proteins that were labelled in similar intensity by all BPLs. Red arrows mark the positions, where one biotinylation signal specifically deviates from the others. **A** Enzyme expression was investigated by IRDye 800CW labeled anti-myc antibody. **B** Biotinylation activity was analyzed by Streptavidin-HRP conjugate.

As can be seen in Figure 12A, expression was successfully detected for all four enzymes. The visualized Proteins were again located at the expected positions and corresponded to the molecular weight of the expressed proteins (cf. Table 2). A comparable amount of present BioID2, BASU and UltraID was observed, as seen from the intensity of the respective bands. The signals of the TurboID levels, on the other hand, turned out with a significantly lower intensity for both doxycycline concentrations. Tubulin, the internal loading control, showed a uniform intensity and localization throughout all columns. The position of the signals at 55 kDa was consistent with the actual molecular weight. Also, when comparing the same cell line at different doxycycline concentrations, the same localizations of bands were visible. Across cell lines, the intensity at 50 ng/mL doxycycline concentration was slightly higher than at 200 ng/mL. This was observed especially for TurboID.

The biotinylation activity of BPLs was demonstrated by the results from Figure 12B. However, biotinylation activity of BioID2 was detected at 50 ng/mL doxycycline just above the detection limit. In order to adjust to the strong signal of TurboID and UltraID in Figure 12B a lower exposure time had to be used. Otherwise TurboID and UltraID would have been overexposed. The difference this makes can be seen when the BioID2 biotinylation activity from Figure 12B and Figure 11 is compared. A significant difference in intensity can be observed, even though both blots contained lysates from the same cell lines.

Since the BioID2 signal, in the enzyme detection blot from Figure 12A, showed a similar intensity to the BASU and UltraID bands, but significantly fewer biotinylated proteins were detected from it, BioID2 was observed to have a lower enzyme activity. BASU showed higher biotinylation activity than BioID2 and lower once than TurboID and UltraID. Despite the low amount of TurboID detected, a very high number of biotinylated proteins was detected here. The biotinylation activity of TurboID is on a comparable high level with UltraID. However, less TurboID was detected before, compared to UltraID.

In some cases, similar bands were recognized between the different cell lines. However, the intensity of these bands was mostly different. In general, bands often merged into one another, or a background formed due to the large number of biotinylated proteins. This was particularly relevant for TurboID and UltraID.

4 Discussion

4.1 Critique of methods

As already explained in the results section, the different cell lines were to be designed in such a way, that the same expression rate could be expected in all cell lines. This should enable a good comparability between the cell lines. For this purpose, cell lines were construct where the transgenes coding for the BPLs were integrated as a single copy at the same position in each cell of a cell line. The BPL transcription sites on the transgenes were under control of the Tet-system. Whether this combination guaranteed similar enzyme expression in all cell lines was not examined again in the context of this thesis. In other scientific studies, however, a uniform and inducible expression between different cell lines has been demonstrated. For example, in "Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells", Béthune et al. were able to show that protein expression with this method takes place at the same rate in different cell lines[43]. For this Béthune et al. constructed two cell lines by inserting a transgene in HeLa-11ht cells, which were used for this thesis too, by RMCE. The transgene coded for Firefly luciferase in both cell lines. Both cell lines were treated in the same way, regarding titer and doxycycline concentration. After doxycycline addition, the light emission was measured over time. The two cell lines showed similar expression levels over time in this Firefly luciferase assay, as seen in Figure 13. [43]

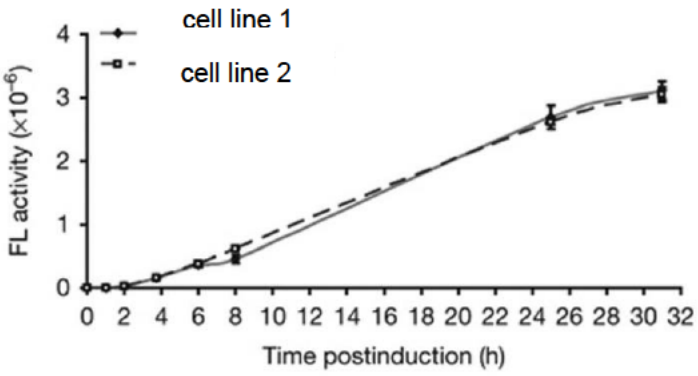

Figure 13: Firefly luciferase assay of two HeLa-11ht cell lines stably expressing FL. The inserted transgene was under the control of the Tet-system. The cell lines were seeded at the same density and induced with the same amounts of doxycycline for the indicated times. [43] (adapted from Béthune et al., 2012)

The experiment from Figure 13 showed a high similarity between the expression of the two cell lines. Therefore, transferred to this thesis, a similar expression was assumed too, since the same master cell line was used. Furthermore, for this thesis too, all cell

lines were treated in the same way and transcription was controlled using the Tet-system. Thus, when contrary to expectations, the amount of BPL detected varied between the different cell lines, it was not assumed that this was due to different expression rates. It was considered that the differences were due to other causes, which are explained in more detail in the substantive discussion below**Error! Reference source not found.**. Thus, the combination of single copy insert in the preselected locus and Tet-system were found to be a suitable method to achieve the goals of this work.

The method used for the characterization of the stable cell lines and comparison of the expression and activities of the BPLs was the western blot. The aim was to carry out a first qualitative and quantitative analysis. An expression of the BPLs was to be detected and the enzyme activity was to be shown. By this, initial conclusions were to be drawn about the possibility of an enzyme-induced bias. Enzyme expression and biotinylation activity of the BPLs were successfully shown in the western blots.

Advantages of this characterization technique are the speed and the low preparative effort. However, the western blot does not provide a comprehensive method for analysis. Based only on the western blot signals, it is barely possible to make assumptions about which biotinylated proteins were represented by which bands. Thus, assured qualitative protein identifications cannot be obtained with this method. However, it was a suitable characterization method and successfully yielded preliminary results. For a more in-depth analysis of the BPLs interaction partners however, further methods such as mass spectroscopy (MS) must be applied.

The western blot was also intended to provide first quantitative results about the expressed enzyme levels and biotinylated proteins. For this purpose, the same loading amount of 10 µg per well and the same expression rate were ensured beforehand. Other parameters, however, would need additional examination, to guarantee a highly accurate method. For example, the extent to which the myc-tags were accessible between all BPLs was not further investigated. Depending on accessibility, the affinity of primary antibody to myc-tag could differ.

In order to determine differences in the accessibility of the myc-tags, the following experiment could be carried out in the future. The four myc-tagged BPLs could be expressed and affinity purified. With a subsequent Bradford assay, the concentration of the purified BPLs could be determined. This would allow the same amount of each

BPL to be loaded onto an SDS-page gel. Analysis with anti-myc antibody in a western blot can then provide the relevant information. If the affinity of the myc-tags to the anti-myc antibodies is the same for all BPLs, the intensity of the bands should be the same everywhere. If the intensities differ, it can be assumed that the myc-tag accessibility was influenced by the fused BPLs.

However, no major differences were expected from this. In a limited context, the western blot is thus considered a suitable quantitative method for this initial characterization.

4.2 Substantive discussion

4.2.1 Construction of stable HeLa cell lines

A major goal of this work was to construct stable HeLa cell lines in order to investigate enzyme-induced bias in proximity dependent labeling technique in further experiments. The construction of these cell lines, stably expressing BPLs, was considered successful. This was demonstrated by a detected expression of all BPLs and by the subsequent detection of biotinylation activity of the enzymes.

Regarding enzyme expression, a comparable level of expressed BPLs for all cell lines was expected. The fact that the different BPLs were detected at different levels is presumably not because the cell lines expressed them at different rates (see Figure 10 and Figure 12A). Uniform protein expression with the methods used here has already been demonstrated in other studies, as described above. The observed differences in enzyme levels must come from the expressed enzyme itself, since all the BPL genes should be transcribed at similar levels. In other studies, an increased toxicity for cells has already been demonstrated for TurboID[20,21,39]. In agreement with this, a weaker occurrence of TurboID was detected in the western blot. In addition, it was hypothesized before that TurboID is to some extent unstable[39]. These two factors could provide a plausible explanation for the lower TurboID quantity and may be present in a somewhat similar form in other enzymes. In addition, as already mentioned, differences in the affinity of the myc-tags to the anti-myc antibody as well as varying degradation of the BPL could be reasons for the different enzyme levels.

In future experiments, the amounts of BPLs could be adjusted before biotin addition to allow for a better comparison of biotinylation activities.

4.2.2 Dose-response of stable HeLa-11ht cells to transcription activator doxycycline

It was expected that higher concentrations of doxycycline would generally lead to higher enzyme expression and thus larger amounts of the proteins would be detectable in the lysate. This could be partially confirmed as shown in Figure 10.

At 0 ng/mL doxycycline, no differences were observed compared to the control cell line. Thus, as expected, expression of the transcription site took place below the detection limit in the absence of doxycycline, since the transcriptional activator rtTA2S-M2 was not activated by doxycycline and therefore did not bind to the operator. Furthermore, for BioID2 and UltraID, an increase in BPL levels was observed with increasing doxycycline concentration. For BASU and TurboID, this could only be shown to a limited extent. Here, constant levels of enzymes were detected starting from 50 ng/mL upwards. Since no increase in enzyme levels was detected for TurboID and BASU above 50 ng/mL doxycycline, a toxic effect of the enzymes could possibly have influenced the detected BPL concentration. The same amounts of protein were initially loaded in each lane, which directly correlates to same amounts of live cells. If an enzyme is toxic to the cells, cells might have died from it. But still same amounts of cells were analyzed, meaning that only surviving cells were analyzed. These cells might have only survived, because they managed to reduce the expression of the BPL by one mechanism or another. The toxicity of TurboID has already been described before[20,21,39]. Saturation with doxycycline and an associated maximum expression level was considered unlikely, since BioID2, for example, showed a clear increase in enzyme levels even with an increase from 100 ng/mL to 200 ng/mL doxycycline. The saturation of induction was observed for a concentration of 200 ng/mL in Weidenfeld et al. 2009[41].

When comparing the cell lines, one can see clear differences in the amounts of enzymes expressed at same doxycycline concentrations. BioID2 was generally the most expressed, followed by BASU, UltraID and TurboID. Since the transfection and expression methods used suggested that expression was uniform and precise, other reasons were assumed for the observed differences in BPL levels[41,43]. As already described, some cell lines might be affected by toxicity of individual BPLs. The instability of BPLs could also contribute to the fact that lower amounts of BPLs were detectable despite increased expression induction. Premature degradation of the

enzymes in the cell could be investigated as a possible cause too. The BPLs are of bacterial origin and may be differently modified in the human cells. This could lead to a different stability.

At 0 ng/mL doxycycline, a band between 60 and 80 kDa was visible for BioID2 and BASU. For TurboID and UltraID on the second blot however, no such signals were seen. Bands at 0 ng/mL doxycycline could generally be endogenous biotinylated proteins, since certain carboxylases are known to covalently bind biotin as a coenzyme in mitochondria or cytoplasm in mammals[45]. In the aforementioned paper “ultraID: a compact and efficient enzyme for proximity-dependent biotinylation in living cells” from Zhao et al. blots with control cells are visible. The control cell lanes show signals of endogenous biotinylated proteins. Two bands at approximately 70 kDa and one band at 130 kDa can be seen. The signal at 70 kDa matches the bands yielded in the 0 ng/mL doxycycline lanes in Figure 11 (Top) and could therefore represent the same endogenous protein.

However, the control cells on the BioID2 and BASU blot did not show any bands. If biotinylated proteins were present here, they were below the detection limit. Still, the bands at 0 ng/mL in the BioID2 and BASU blots could be endogenous proteins, as these cell lines were supplemented with additional biotin, whereas the control cell line was not. This could have led to an enrichment of biotinylations also in endogenous proteins that interact with biotin normally. However, the possibility that the bands were caused by inaccuracies during pipetting of the 10 ng/mL lysate into the SDS-PAGE wells cannot be ruled out. The fact that these signals were shown only in the BioID2 and BASU blot, could be caused by different exposure times of the two blots. TurboID and UltraID were described to have a higher activity, thus the blot needed less exposure time then the BioID2 and BASU blot[20,21]. Same amounts of biotinylated proteins are thus visualized to different degrees. It could therefore occur that the same amount of protein is below the detection limit in one blot. In addition, the differences in the quality of the two blot images could have had an impact on the visibility of bands and thus, also could have changed the detection limit of the two blots, complicating the comparability.

It was shown that the expression level is generally tunable with the amount of doxycycline. For the enzymes for which this could not be shown to the full extent

(BASU and TurboID), the reason for this was not assumed to be a different expression rate but instability of the enzymes after expression.

4.2.3 Indications of enzyme-induced bias

One aim of the experiments conducted for this thesis were to allow for first characterizations of the four BPLs BioID2, BASU, UltraID and TurboID with respect to induced biases. This should allow initial conclusions to be drawn about the behavior of the BPL in the cells. It was of interest, whether the four BPLs investigated behave in the same way when expressed alone in HeLa cells, or whether the BPLs might tend to react differently with different cell components. Furthermore, the results should provide initial data on whether the correct targeting of POIs could be influenced by the interactions of the BPLs. For this initial assessment, a western blot was performed containing lysate from all four BPLs discussed so far (see Figure 12).

It was shown that the enzymes were likely to biotinylate many cellular components when not fused to a fusion protein. Generally, a similar biotinylation pattern was found between the BPLs. Enhanced interaction of UltraID with a protein at 35 kDa suggests that the only found bias might be induced by the UltraID.

BioID2 biotinylated proteins were below the detection limit, although large amounts of BioID2 were detected (Figure 12A). This suggests slow enzyme kinetics of BioID2. BASU also showed a significantly weaker biotinylation activity than TurboID and UltraID, but higher than BioID2. Despite the low amount of TurboID expressed, it showed this high enzyme activity. This is consistent with the literature, which reports higher activity of TurboID over BioID2 and BASU[20]. UltraID shows similar enzymatic activity to TurboID, as was also already described in literature before[21]. In general, due to the large number of bands, some of which merge into one another, qualitative protein determinations are problematic with this method.

Ideally, all enzymes would show similar signals in the blot for all BPLs. This would result from BPLs which do not induce a bias. The BPLs would be diffusing freely in the cell after expression and randomly release bioAMP, which would react non-specifically with endogenous proteins. The intensities and localizations of the bands would differ only slightly between the BPLs if none of the BPLs reacted with cellular components in a special way. However, if the cells would react with certain cell components in

different ways, different bands would form or the intensities would differ. This would indicate a bias of one of the enzymes.

The assumption that the BPLs would not react with cell components is supported by the fact that all BPLs used did not originate from eukaryotic cells, but from bacteria. Since the enzymes are foreign to the cells, it was not expected that specific interaction partners would exist for the enzymes. Although this was considered a good precondition, interactions of BPL with cell components in general were to be analyzed by this blot. Due to the free diffusion of the BPLs and continues bioAMP release, cell proteins that are more abundant and accessible for the BPLs were assumed to be biotinylated the most, since chances are high for the BPL to come close to these proteins more often. Based on the streptavidin blot in Figure 12B the BPLs showed comparable biotinylation patterns.

In general, when comparing the biotinylation results (Figure 12B) of the four BPLs, the blots stand out due to same positions of the bands and varying intensities of the signals. Many bands were detected, which were present at the same position for all enzymes. For example, at approx. 95 kDa, 80 kDa, 45 kDa and 38 kDa (cf. blue arrows in Figure 12B). The bands at 70 kDa is probably due to the biotinylation of an endogenous protein. However, intensity of some bands varies between the cell lines. Some bands, for example the one generated by TurboID at approx. 40 kDa, are the autobiotinylated BPLs itself and are not to be confused with possible interaction partners (cf. Table 2). Autobiotinylation of the enzymes can be seen, where the biotinylation signals are at the same position of the expression signals (cf. Figure 12A and Figure 12B). Other noticeable bands are located at approximately 20 kDa and 55 kDa and are clearly more intense when caused by TurboID (cf. red arrows in Figure 12B). Furthermore, UltraID shows a stronger signal at 35 kDa and 70 kDa than the other BPLs (cf. red arrows in Figure 12B). Whether these more prominent bands are due to the BPLs interaction with biotinylated proteins, cannot be determined with certainty. However, if they are interactions, they could cause a bias. Due to the lower enzyme activities and adjusted exposure times, the lanes for BioID2 and BASU cannot be evaluated at all or only partially. BASU's biotinylation pattern shows no noticeable differences to UltraID and TurboID. Proteins biotinylated by BioID2 are below the detection limit.

The results of this initial characterization have created initial insights for enzyme-induced bias investigation in proximity dependent labelling techniques. It was shown that the different BPLs biotinylate a variety of proteins, some of these with different intensity, suggesting a bias could occur. UltraID and TurboID showed one or more differing bands that could indicate an interaction of the BPLs in the cells, thus a bias could be assumed for these enzymes.

4.3 Outlook

The construction of stable HeLa cell lines was found successful, as different expression levels could be justified. The expression level was generally found tunable with the amount of doxycycline and all enzymes were shown to have a biotinylation activity as expected. A first analysis of the biotin detection blot successfully showed that the BPLs interact with cell components and enzyme-related biases due to interactions with host cell components could influence the results of the proximity-dependent labelling technique.

However, the ultimate goal of the overarching research question will be, to carry out further research to obtain more precise and robust results on the same topics, using the constructed stable cell lines. This could include a qualitative analysis to determine which proteins were biotinylated by the BPLs. In addition, a more precise quantitative analysis could allow to determine the intensity of interactions, and thus also the probability that a bias was generated by it. For example, MS could be used as an approach to accurately identify the suspected interaction partners. Furthermore, by using MS, the intensity of the interaction could be determined by quantifying the biotinylated proteins.

An answer to the question, whether the enzymes cause bias when they react with cellular components, could generally be of importance for future research. BPLs that have known specific interaction partners could be selectively used only at locations where these interaction partners do not occur. A more targeted use of the enzymes could thus be made possible, which in turn could lead to more precise results. This could ultimately allow for improved applications within the known fields of application, such as fundamental research, drug discovery and proteomics.

5 Materials and Methods

5.1 Materials

5.1.1 Chemicals and ready to use buffers and solutions

Chemical or buffer	Manufacturer	Product number
Acrylamide/ Rotiphorese Gel 30	Carl Roth	3029.1
Ammonium persulfate (APS)	Carl Roth	9592.3
Ampecillin	Carl Roth	K029.1
Biotin	Sigma-Aldrich	B4639-1G
Bovine serum albumin (BSA)	Sigma-Aldrich	A6003-10G
Bromphenol Blue	Carl Roth	T116.1
Diethiothreitol (DTT)	GE	171318.01
Dimethyl sulfoxid (DMSO)	Sigma	D2438
Doxycycline	Sigma-Aldrich	D9591
Dulbecco´s Modified Eagle´s Medium (DMEM)	Sigma-Aldrich	D6046-500ML
Ethanol	Carl Roth	K928.4
G418	Sigma-Aldrich	1720
Ganciclovir	Sigma-Aldrich	G2536
Glycerol	Carl Roth	4043.1
Glycin	Amersham Biosiences	17-1323-01
Hygromycin	Thermo Fisher	10687010
L-Glutamine	Merck	K0282
LB-Medium	Carl Roth	X964.1
Lipofectamin 2000	Thermo Fisher	11668030
Lipofectamine 2000	Thermo Fisher	11668019
Methanol	VWR	20847.307

Milk powder	Carl Roth	T145.3
NaCl	Carl Roth	3957.1
Novex™ Sharp Pre-Stained Protein Standard	Invitrogen	LC5800
NP40	Thermo Fisher	85124
Opti-MEM	Thermo Fisher	31985062
Optimem	Thermo Fisher	31985062
PBS Dulbecco (cell culture)	Biochrom GmbH	L1825
Penicilin/Streptomycin	PAN-Biotech GmbH	P08-07050
Sodium dodecyl sulfate (SDS)	Serva	20765.03
TET-free fetal bovine serum (FBS)	PAN/Biowest	S181B
Tetramethylethylenediamine (TEMED)	Carl Roth	2367.3
Tris	Serva	37190
Tris-HCl	Carl Roth	9090.
Trypsin / EDTA	PAN-Biotech GmbH	P10-023100
Tween 20	VWR Life Sciences	437082Q

5.1.2 Solutions and buffers

Solution or buffer	**components**	**concentration**
Growth medium for HeLa cells	DMEM	
	TET-free FBS	10 %
	L-Glutamine	1 %
	Penicillin/Streptomycin	1 %
	G418	200 µg/mL
Lysis buffer	ddH_2O	
	Tris-HCl pH (7,4)	50 mM
	NaCl	150 mM
	NP40	0,5 %

	EDTA pH 8	2 mM
	DTT	0,5 mM
Transfer buffer 10x (western blot)	ddH_2O	
	Tris Base	200 mM
	Glycin	1,5 M
Blotting buffer (western blot)	ddH_2O	
	Ethanol	20 %
	Transfer buffer 10x	10 %
Stacking gel	ddH_2O	
	Acrylamide	5 %
	Tris-HCl (pH 6,8)	130 mM
	SDS	0,1 %
	Ammonium persulfate	0,1 %
	TEMED	0,1 %
Seperating gel	ddH_2O	
	Acrylamide	12 %
	Tris-HCl (pH 8,8)	375 mM
	SDS	0,1 %
	ammonium persulfate	0,1 %
	TEMED	0,04 %
Blocking solution (western blot)	Milk powder in PBS	5 %
Electrophoresis buffer	ddH_2O	
	Tris Base	0,25 M
	Glycin	1,92 M
	SDS	1 %
Laemmli buffer	SDS	10 %
	Glycerol	50 %
	Tris-HCl (pH 6,8)	300 mM

	Bromphenol Blue	0,05 %
	DTT	100 mM

5.1.3 Kits

Kit	Manufacturer	Product number
Bradford Assay	Serva	39222.02
GenElute™ HP Plasmid Miniprep Kit	Sigma-Aldrich	NA0150
Super Signal™ West Pico PLUS Chemiluminescent Substrate	Thermo scientific	34577

5.1.4 Cells

Cells	Manufacturer	Product number
Competent DH-5α *E. coli*	New England BioLabs	C2987
HeLa-EM2-11ht	Zentralinstitut für Seelische Gesundheit Mannheim (Dr. Schönig)	-

5.1.5 Antibodys

Antibodys	Manufacturer	Product number	Appplied dilution
Alpha-Tubulin-anti-mouse	Sigma	T6199	1:2.000 in PBS-Tween 20 0,1 % and 1 % milk powder
IRDye-800CW-anti-mouse	abcam	ab216772	1:10.000 in PBS-Tween 20 0,1 % and 3 % milk powder
Myc-Tag-anti-mouse	Cellsignals	2276S	1:2.000 in PBS-Tween 20 0,1 %

			and 1 % milk powder
Streptavidine-HRP Conjugate	Millipore	18-152	1:5.000 in PBS-Tween 20 0,1 % and 2 % BSA

5.1.6 Equipment and consumables

Equipment	**Manufacturer**	**Product specification**
Blotting system	Bio-Rad	Trans-Blot® SD Semi-Dry Transfer Cell
Cell Counter	Invitrogen	Countess™ automated cell counter
Cell Counting Chamber Slides	Thermo Fisher Scientific Inc.	Countess™ Cell Counting Chamber Slides
CO_2 Incubator	Binder	CB150
Incubater shaker	GFL	GFL 3033
Inverted Microscope	Nikon	Eclipse TS 100
Microcentrifuge	Eppendorf	miniSpin
NanoDrop spectrophotometer	Thermo Fisher Scientific Inc.	NanoDrop 1000
Safety Cabinet	Thermo Fisher Scientific Inc.	HERAsafe KS18
spectrophotometer	Bio-Rad	SmartSpecTM Plus
Supercentrifuge	Thermo Scientific	LYNX 6000
Western blot and chemilumineszenz imager	Vilber Lourmat	FUSION SL
Western Blot PVDF Membrane	Amersham Biosciences	10600021

5.2 Methods

5.2.1 Plasmids

For the transfection of the Competent DH-5α *E. coli* cells the cells were incubated with 1 µl plasmid DNA on ice for 30 minutes. A heat shock was performed at 42 °C for 30 seconds, followed by another incubation on ice for 2 minutes. Subsequently 250 µl pre-warmed (37 °C) LB medium was added. The cells were then plated on pre-warmed (37 °C) LB medium agar plates containing 100 µg/mL ampicillin and incubated at 37 °C overnight. On the following day two colonies per plasmid were picked and transferred to LB medium containing ampicillin (100 µg/mL). The cells were incubated at 37 °C overnight with agitation. For the plasmid preparation on the following day the spin method of the GenElute™ HP Plasmid Miniprep Kit was performed. Measurement of the plasmid concentration was conducted with a NanoDrop 1000. The plasmid maps can be found in the appendix (Figure 14, Figure 15, Figure 16, Figure 17).

5.2.2 Cell culture

The Hela-EM2-11ht master cell line used in this thesis was a kind gift from Dr. Schönig from the Zentralinstitut für Seelische Gesundheit Mannheim. Construction of the master cell line was described in “Inducible expression of coding and inhibitory RNAs from retargetable genomic loci” from Weidenfeld et al[41].

For the master cell line, Hela-EM2-11ht cells stored in growth medium with 30 % fetal calf serum and 10 % dimethyl sulfoxid were thawed from the storage in liquid nitrogen and washed with Dulbecco´s Modified Eagle´s Medium (DMEM) by centrifugation (3 minutes, 300 xg) and resuspension of the pellet in 6 mL fresh DMEM containing 10% TET-free FBS, 1 % L-Glutamine, 1 % penicillin/streptomycin and 200 µg/mL G418 (Hereafter called growth medium). The cells were then incubated at 37 °C and 5 % CO_2 until the cell passage was required.

At a cell density of 80 – 90 %, determined with an inverted Microscope, the cells were passaged. For the cell passage, medium was removed and the cells were washed with 5 mL PBS. 1 mL Trypsin-EDTA solution was added and the cells were incubated at 37 °C and 5 % CO_2 for 5 minutes. 4 mL growth medium was added, and the cells resuspended by pipetting up and down. For a 1:20 passage 250 µL of the resuspended cells were added to a new flask, containing 5 mL growth medium.

Until before transfections, hygromycin was added to all master cell line cultures for selection (300 µg/ml).

5.2.3 Transient transfection of master cell line HeLa-EM2-11ht

300.000 cells per well were seeded in a 6 well plate with the growth medium. The next day, volumes corresponding to 1,0 µg of each BPL-encoding plasmid were made up to 250 µL with opti-MEM. 5 µL Invitrogen™ Lipofectamine™ 2000 transfection reagent and 245 µL opti-MEM were added dropwise to each of the plasmid/opti-MEM mixtures after 5 minutes of incubation time. The resulting transfection mix was slowly mixed and incubated for 20 minutes at room temperature. During incubation the growth medium of the cells, that were to be transfected, was changed for growth medium without antibiotics. Subsequently each transfection mix was added dropwise to the respective cells in the 6 well plate and incubated for 4 hours at 37 °C and 5 % CO_2. After incubation, the transfection mixture was replaced with 2 mL of growth medium and incubated overnight at 37 °C and 5 % CO_2.

5.2.4 Construction of stable HeLa cell lines

For the stable transfection of the HeLa-EM2-11ht master cell line, 150.000 to 300.000 cells per well were seeded in a 6 well plate with growth medium. On the next day volumes corresponding to 1,5 µg of all BPL-encoding plasmids and volumes corresponding to 1,5 µg of the FLP-recombinase plasmid were used to prepare the four plasmid mixtures for the transfection. This plasmid mixtures were each made up to 250 µl with opti-MEM. After incubation at room temperature for 5 minutes, a simultaneously prepared mixture of 5 µL Invitrogen™ Lipofectamine™ 2000 transfection reagent and 245 µL opti-MEM was added dropwise to the plasmid/opti-MEM mixtures. The growth medium of the HeLa-11ht cells was changed to growth medium without antibiotics. The transfection mixture was then incubated for 20 minutes and added dropwise to the cells. After 4 hours of incubation at 37 °C and 5 % CO_2 the medium was changed to growth medium and incubated overnight at 37 °C and 5 % CO_2. On the next day the transfected cells were transferred to 10 cm plates containing 8 mL growth medium and incubated for two days. From there every third day the medium was exchanged with fresh growth medium and 50 µM ganciclovir for selection.

After 15 days when colonies were clearly visible without a microscope three single colonies per cell line were picked and cultivated in a 24-well plate with 50 µM ganciclovir until confluent. Picking was conducted by incubating the colonies with trypsin/EDTA-soaked disks and transfer of the disks to a 24-well plate containing growth medium and 50 µM ganciclovir. Subsequently all colonies of each cell line were trypsinated in 2 mL trypsin/EDTA and after 10 minutes of incubation at 37 °C and 5 % CO_2 resuspended with 8 mL growth medium and added to a new 10 cm dish. Picked colonies were transferred when confluent to 6-well plates and then 10-well plates. 14 days after picking the cell lines were finally frozen away in growth medium containing 30 % fetal calf serum and 10 % dimethyl sulfoxid. The pooled cells were held in culture with 50 µM ganciclovir at 37 °C and 5 % CO_2 for 11 further days after pooling. Pooled cell lines were then either directly used for induction of biotinylation or frozen away until characterization was conducted. Freezing was conducted as for the picked cell lines. Cells were generally frozen gradually. The cells were therefore placed in a container surrounded by isopropyl alcohol and stored first at -80 °C for four hours and then at -160 °C.

In preparation for the induction of BPL expression and biotinylation, the pooled stable cell lines were seeded in 6-well plates at 450.000 cells per well in growth medium. Incubation at 37 °C and 5 % CO_2 was carried out for approximately 4 hours to allow the cells to attach.

5.2.5 Induction of BPL expression and biotinylation

Medium in the wells of the 6-well plates was replaced by growth medium containing variable concentrations doxycycline (stable cell lines: 10 ng/mL, 50 ng/mL, 100 ng/mL, 200 ng/mL; transient cell lines: 200 ng/mL) and 50 µM biotin. Incubation at 37 °C and 5 % CO_2 was carried out overnight.

5.2.6 Western blot

Transfected HeLa cells were washed with 1 mL cold PBS per well. 100 µL lysis buffer was added per well and the cells were transferred to 1,5 mL tubes by scraping and centrifuged at 14.000 g, 10 min and 4 °C. The supernatants were transferred to 1,5 mL tubes and kept on ice. A Bradford assay was conducted to assess the protein

concentration in the lysates. Together with 5x Laemmli buffer, lysate containing 10 µg of protein per well were heated to 95 °C for 5 minutes and loaded onto the 12 % SDS-page gel.

The gel was run at 120 V until the samples were stacked and then switched to 200 V until the bromophenol blue front reached the end of the gel. The proteins were blotted on a PVDF membrane for 60 min at 0,8 mA/cm^2 using the semi-dry method. After blotting the membrane was blocked with 5 % milk in PBS for 30 minutes at room temperature (or overnight at 4 °C) and then incubated with the primary antibody myc-tag-anti-mouse 1:2.000 for 60 minutes at room temperature (or overnight at 4 °C). If applicable, the membrane was incubated simultaneously with an alpha-tubulin-anti-mouse antibody 1:2.000. Subsequently the membrane was washed three times with PBS containing 0,1 % tween 20 for 10 minutes each at room temperature. The secondary antibody IRDye-800CW-anti-mouse 1:10.000 was incubated with the membrane for 30 minutes at room temperature. Again, three washings were conducted using PBS-Tween 20. Last wash was conducted with PBS only. Fluorescence signals were measured with a western blot and chemiluminescent imager. Same membranes were then incubated with streptavidin-HRP conjugate 1:5.000 for 30 minutes at room temperature. Following three PBS-Tween 20 washes at room temperature chemiluminescent substrate was applied by pipetting 2 mL on each membrane and incubation for 5 minutes at room temperature. Chemiluminescent signal was measured using a western blot and chemiluminescent imager.

6 References

1. Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences* **93**, 13–20 (1996).

2. Diversity of protein–protein interactions. *The EMBO Journal* **22**, 3486–3492 (2003).

3. Turnbull, A., Boyd, S. & Walse, B. Fragment-based drug discovery and protein–protein interactions. *RRBC* 13 (2014) doi:10.2147/RRBC.S28428.

4. Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein–protein interactions. *PROTEOMICS* **7**, 2833–2842 (2007).

5. Acuner Ozbabacan, S. E., Engin, H. B., Gursoy, A. & Keskin, O. Transient protein–protein interactions. *Protein Engineering, Design and Selection* **24**, 635–648 (2011).

6. Low, T. Y. *et al.* Recent progress in mass spectrometry-based strategies for elucidating protein–protein interactions. *Cell. Mol. Life Sci.* **78**, 5325–5339 (2021).

7. Spirin, V. & Mirny, L. A. Protein complexes and functional modules in molecular networks. *Proceedings of the National Academy of Sciences* **100**, 12123–12128 (2003).

8. Valdar, W. S. & Thornton, J. M. Protein-protein interfaces: analysis of amino acid conservation in homodimers. *Proteins* **42**, 108–124 (2001).

9. Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem Rev* **109**, 839–860 (2009).

10. Chène, P. Drugs Targeting Protein–Protein Interactions. *ChemMedChem* **1**, 400–411 (2006).

11. Ryan, D. P. & Matthews, J. M. Protein–protein interactions in human disease. *Current Opinion in Structural Biology* **15**, 441–446 (2005).

12. Lu, H. *et al.* Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Sig Transduct Target Ther* **5**, 1–23 (2020).

13. Scott, D. E., Bayly, A. R., Abell, C. & Skidmore, J. Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nat Rev Drug Discov* **15**, 533–550 (2016).

14. Egetemaier, S. & Béthune, J. Proteomik-Analyse von dynamischen Proteinkomplexen. *Biospektrum* **25**, 45–48 (2019).

15. Titeca, K., Lemmens, I., Tavernier, J. & Eyckerman, S. Discovering cellular protein-protein interactions: Technological strategies and opportunities. *Mass Spectrometry Reviews* **38**, 79–111 (2019).

16. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).

17. Caufield, J. H., Sakhawalkar, N. & Uetz, P. A comparison and optimization of yeast two-hybrid systems. *Methods* **58**, 317–324 (2012).

18. Dunham, W. H., Mullin, M. & Gingras, A.-C. Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *PROTEOMICS* **12**, 1576–1590 (2012).

19. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology* **196**, 801–810 (2012).

20. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880–887 (2018).

21. Zhao, X. *et al.* ultraID: a compact and efficient enzyme for proximity-dependent biotinylation in living cells. 2021.06.16.448656 (2021) doi:10.1101/2021.06.16.448656.

22. Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. & Salemme, F. R. Structural Origins of High-Affinity Biotin Binding to Streptavidin. *Science* **243**, 85–88 (1989).

23. Feng, Y. *et al.* The Atypical Occurrence of Two Biotin Protein Ligases in Francisella novicida Is Due to Distinct Roles in Virulence and Biotin Metabolism. *mBio* **6**, e00591-15 (2015).

24. Chapman-Smith, A. & Cronan, J. E. Molecular biology of biotin attachment to proteins. *J Nutr* **129**, 477S-484S (1999).

25. Henke, S. K. & Cronan, J. E. Successful conversion of the Bacillus subtilis BirA Group II biotin protein ligase into a Group I ligase. *PLoS One* **9**, e96757 (2014).

26. Chapman-Smith, A. & Cronan, J. E. The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* **24**, 359–363 (1999).

27. Lane, M. D., Rominger, K. L., Young, D. L. & Lynen, F. THE ENZYMATIC

SYNTHESIS OF HOLOTRANSCARBOXYLASE FROM APOTRANSCARBOXYLASE AND (+)-BIOTIN. II. INVESTIGATION OF THE REACTION MECHANISM. *J Biol Chem* **239**, 2865–2871 (1964).

28. Kwon, K. & Beckett, D. Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Science* **9**, 1530–1539 (2000).

29. Li, P., Li, J., Wang, L. & Di, L.-J. Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool. *PROTEOMICS* **17**, 1700002 (2017).

30. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci* **8**, 921–929 (1999).

31. Nonenzymatic biotinylation of a biotin carboxyl carrier protein: Unusual reactivity of the physiological target lysine - Streaker - 2006 - Protein Science - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1110/ps.062187306.

32. Choi-Rhee, E., Schulman, H. & Cronan, J. E. Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Science* **13**, 3043–3050 (2004).

33. Kim, D. I. *et al.* Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences* **111**, E2453–E2461 (2014).

34. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *MBoC* **27**, 1188–1196 (2016).

35. Xu, Y. & Beckett, D. Evidence for interdomain interaction in the Escherichia coli repressor of biotin biosynthesis from studies of an N-terminal domain deletion mutant. *Biochemistry* **35**, 1783–1792 (1996).

36. Groft, C. M., Uljon, S. N., Wang, R. & Werner, M. H. Structural homology between the Rap30 DNA-binding domain and linker histone H5: Implications for preinitiation complex assembly. *Proceedings of the National Academy of Sciences* **95**, 9117–9122 (1998).

37. M, R. *et al.* RNA-protein interaction detection in living cells. *Nature methods* **15**, (2018).

38. P, S.-T., R, S. & Ac, G. Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches. *Molecular & cellular proteomics : MCP* **19**, (2020).

39. May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* **9**, 1070 (2020).

40. Das, A. T., Tenenbaum, L. & Berkhout, B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther* **16**, 156–167 (2016).

41. Weidenfeld, I. *et al.* Inducible expression of coding and inhibitory RNAs from retargetable genomic loci. *Nucleic Acids Res* **37**, e50 (2009).

42. Schlake, T. & Bode, J. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* **33**, 12746–12751 (1994).

43. Béthune, J., Artus-Revel, C. G. & Filipowicz, W. Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep* **13**, 716–723 (2012).

44. Oumard, A., Qiao, J., Jostock, T., Li, J. & Bode, J. Recommended Method for Chromosome Exploitation: RMCE-based Cassette-exchange Systems in Animal Cell Biotechnology. *Cytotechnology* **50**, 93–108 (2006).

45. Zempleni, J. & Kuroishi, T. Biotin. *Advances in Nutrition* **3**, 213–214 (2012).

7 Affidavit

I herewith declare that

1. I wrote this bachelor thesis independently, under supervision, and that I used no other sources and aids than those indicated throughout the thesis,

2. citations from the literature and the use of ideas of others are declared as such,

3. I did not hand in this thesis for another examination. I am aware that a false declaration has legal consequences.

Enno Große Wichtrup

[redacted]

Hamburg, 05.07.2022

8 Appendix

8.1 Plasmid map pSF3-BioID2

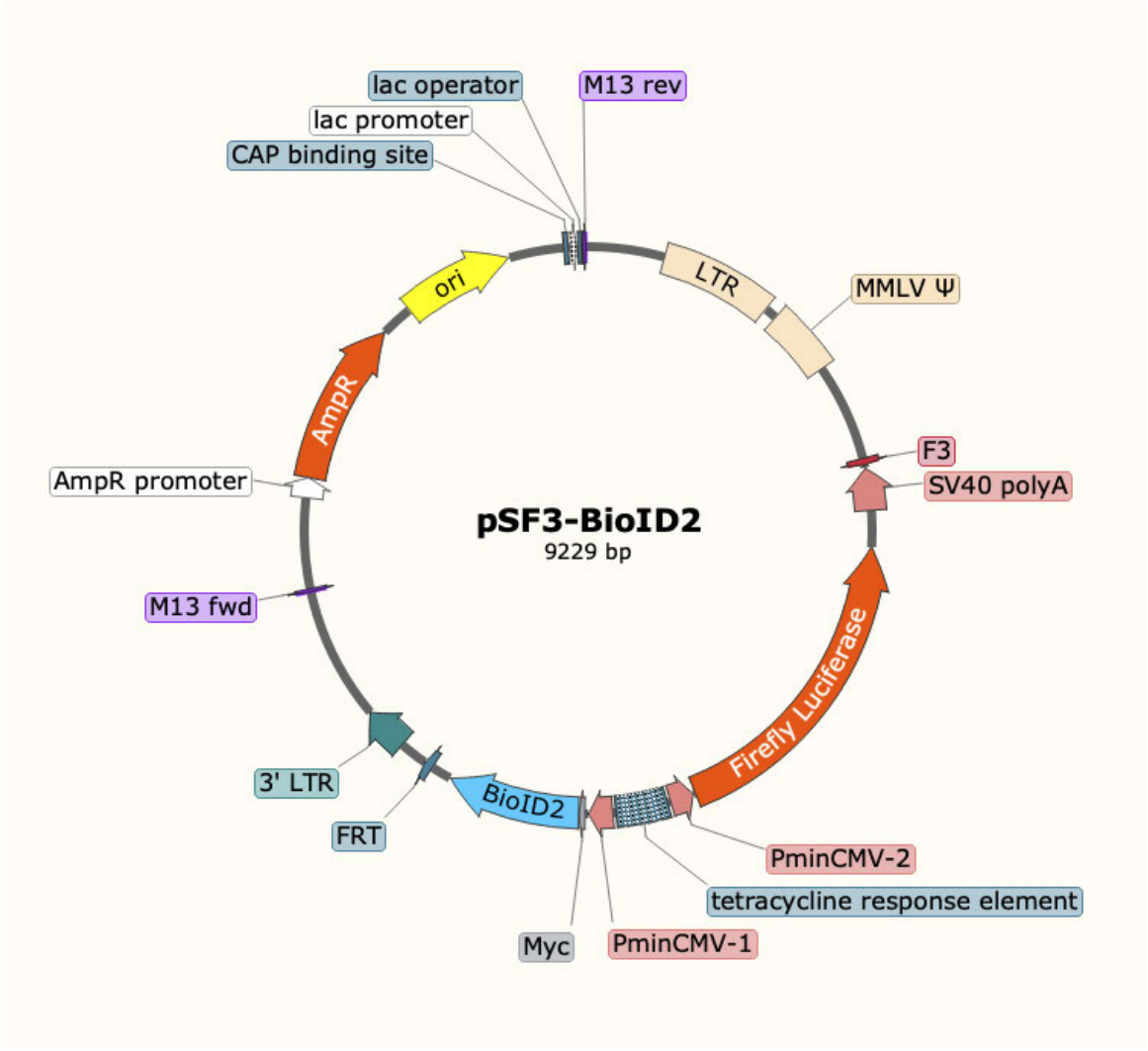

Figure 14: Plasmid map pSF3-BioID2 used for the transfection of transient and stable HeLa-EM2-11ht cells for the use in a proximity dependent labeling technique.

8.2 Plasmid map pSF3-BASU

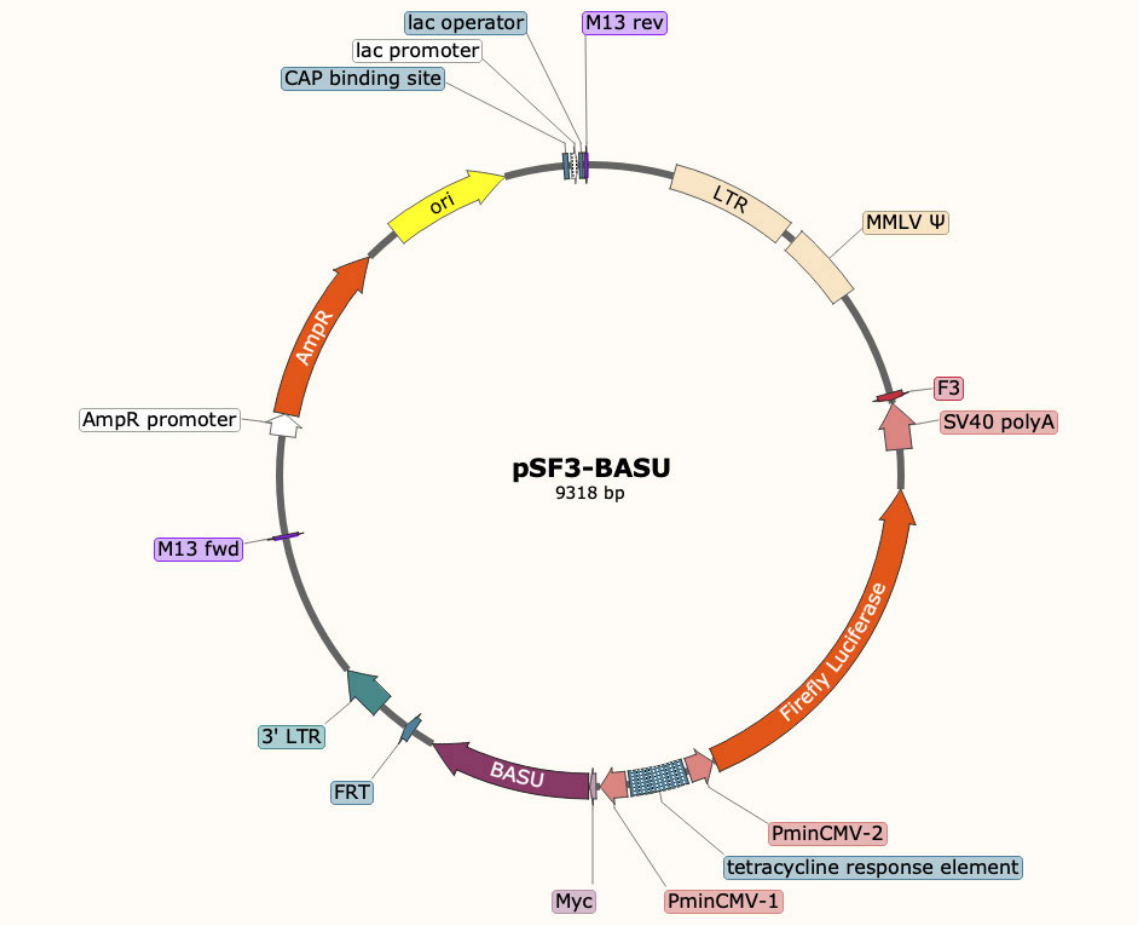

Figure 15: Plasmid map pSF3-BASU used for the transfection of transient and stable HeLa-EM2-11ht cells for the use in a proximity dependent labeling technique.

8.3 Plasmid map pSF3-TurboID

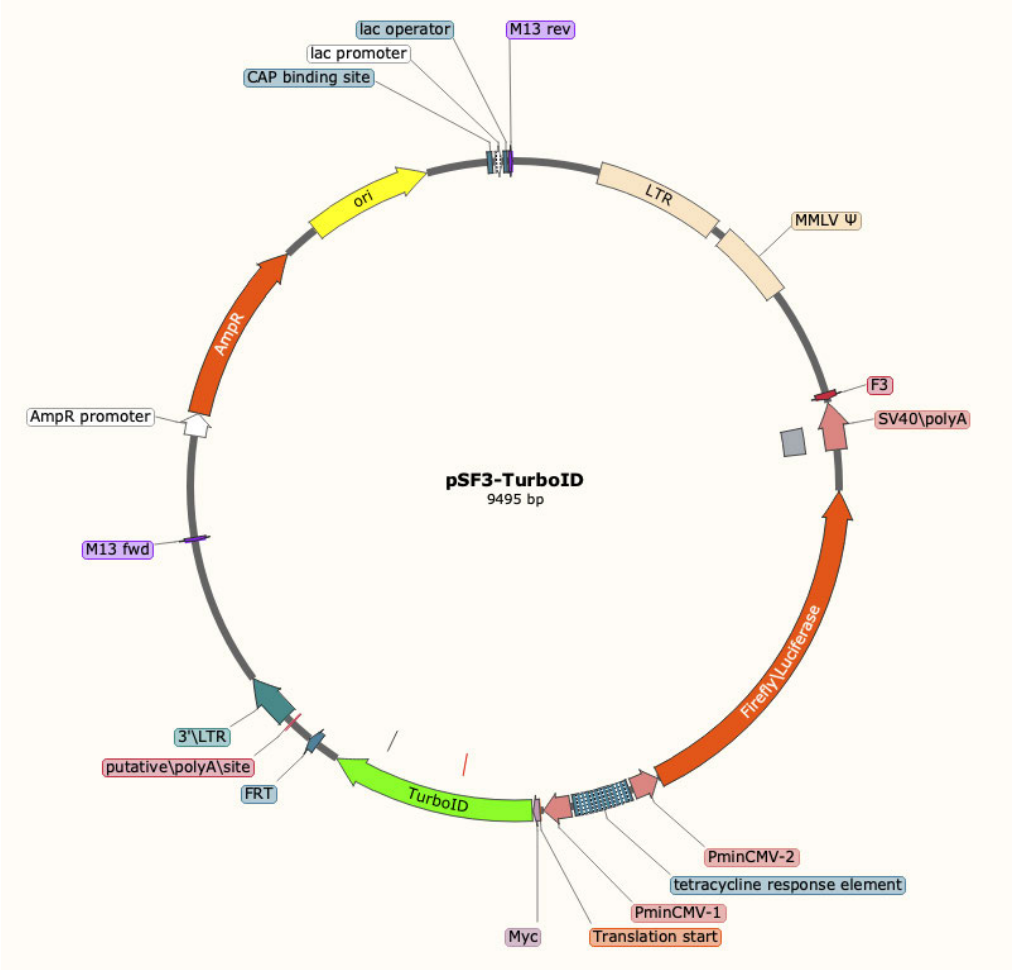

Figure 16: Plasmid map pSF3-TurboID used for the transfection of transient and stable HeLa-EM2-11ht cells for the use in a proximity dependent labeling technique.

8.4 Plasmid map pSF3-ultraID

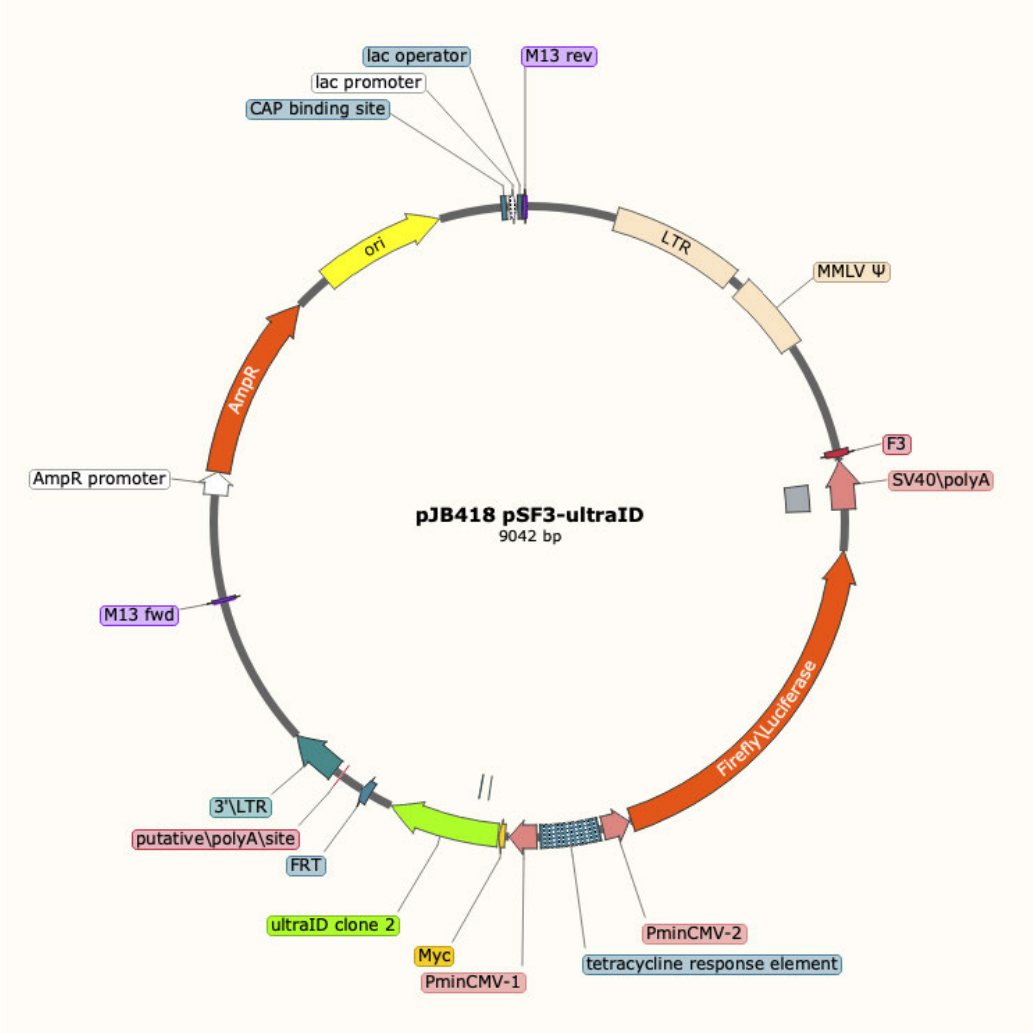

Figure 17: Plasmid map pJB418 pSF3-ultraID used for the transfection of transient and stable HeLa-EM2-11ht cells for the use in a proximity dependent labeling technique.