

Establishment of the BioID method in *Entamoeba histolytica*  
for the characterization of the putative pathogenicity factor  
EHI\_127670

**Bachelor Thesis**

Degree Course Biotechnology  
Faculty Life Sciences  
Department Biotechnology  
HAW Hamburg

2021

Submitted by

**Arne Hellhund**

First Examiner: Prof. Dr. Jörg Andrä  
Second Examiner: Prof. Dr. Iris Bruchhaus

Submission date: 23.06.2021

# Contents

List of Abbreviations	IV
Abstract	1
1 Introduction	2
1.2 Research Aim	3
2 Theoretical Background	4
2.1 Infection and Life Cycle of <i>E. histolytica</i>	4
2.2 Cell biology of <i>E. histolytica</i>	4
2.3 Amebiasis	5
2.4 State of Research	6
2.4.1 Culture Isolates	6
2.4.2 Pathogenicity Factors	7
2.4.3 BioID	10
3 Materials	11
3.1 Laboratory Devices	11
3.2 Software	12
3.3 Disposables	13
3.4 Chemicals	13
3.5 Buffers and Solutions	16
3.6 Reaction Kits	19
3.7 DNA- and Protein Markers	19
3.8 Enzymes and Reaction Buffers	20
3.9 Antibodies	21
3.10 Vectors	21

3.11	Oligonucleotides	22
3.12	Organisms	23
3.12.1	<i>Escherichia coli</i>	23
3.12.2	<i>Entamoeba histolytica</i>	24
3.13	Culture Media	25
3.13.1	Bacterial Culture Media	25
3.13.2	Amoebic Culture Media	25
4	Methods	27
4.1	Cloning	27
4.1.2	Myc-tag Construct	27
4.1.3	BioID Construct	27
4.1.4	Cloning Process	28
4.1.5	Primer Design	29
4.1.6	Polymerase Chain Reaction	29
4.1.7	TOPO Cloning	30
4.1.8	Ligation	31
4.1.9	Restriction Digest	32
4.1.10	Agarose Gel Electrophoresis and DNA Preparation	33
4.1.11	Measurement of DNA concentration	34
4.1.12	Transformation of <i>E. coli</i> and Selection	34
4.1.13	Cultivation of <i>E. coli</i> and Plasmid Purification	35
4.1.14	Construct Verification	36
4.1.15	Cryopreservation of <i>E. coli</i>	36
4.2	Gene Expression in <i>E. histolytica</i>	37
4.2.1	Cultivation of <i>E. histolytica</i>	37
4.2.2	Harvest of <i>E. histolytica</i>	38

4.2.3	Transfection of <i>E. histolytica</i> and Selection	38
4.2.4	Production of <i>E. histolytica</i> Cell Lysate	39
4.3	SDS-PAGE	40
4.3.1	Glycine-SDS-PAGE after Laemmli	40
4.3.2	Tricine-SDS-PAGE	41
4.4	Western Blot	42
4.5	Biotinylation	43
4.7	Streptavidin Blot	44
4.8	Immunofluorescence Assay (IFA)	45
5	Results	48
5.1	Myc-tag Localization	48
5.2	BioID	55
6	Discussion	68
6.1	Myc-tag Localization	68
6.2	BioID	71
7	Outlook	80
	References	81
	Acknowledgements	87
	Statutory declaration	88

## List of Abbreviations

---

A	Adenine
AB	Antibody
ALA	Amoebic liver abscess
amp	Ampicillin
BioID	Proximity-dependent biotin identification
BNITM	Bernhard-Nocht-Institute for Tropical Medicine
BSA	Bovine serum albumin
C	Cytosine
CP	Cysteine peptidases
ddH <sub>2</sub> O	Double-distilled water
ER	Endoplasmic reticulum
FCS	Fetal calf serum
FD	<i>FastDigest</i>
G	Guanine
G418	Geneticin
Gal	Galactose
GalNAc	N-acetyl-D-galactosamine
GGGGS	Glycine-Glycine-Glycine-Glycine-Serine
HRP	Horseradish peroxidase
IFA	Immunofluorescence assay
lncRNA	Long non-protein coding RNA
OD	Optical density
ODV	Optical density volume
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
rpm	Rounds per minute
SOD	Superoxide dismutase
T	Tyrosine
TBS	Tris-buffered saline
T <sub>m</sub>	Melting temperature
w/v	Weight per volume

---

## Abstract

The enteric protozoan parasite *E. histolytica* is a human pathogen causing amebiasis, a neglected tropical disease that affects nearly 50 million people and causes more than 55,000 deaths annually worldwide. The invasive form of amebiasis is characterized by the invasion of the parasites into the host intestinal tissue, which leads to amoebic colitis or abscesses in other organs, particularly the liver. The exact circumstances that lead to invasive amebiasis are unknown. Proteins that are involved in the virulence of the parasite, so called pathogenicity factors, are expected to have an influence. Recently, the new putative pathogenicity factor EHI\_127670 was discovered. Phenotypic characterizations proved its enhancing influence on pathogenicity. However, structure, function, locus and interaction partners of this molecule remain unknown. It was hypothesized that the role of EHI\_127670 in pathogenicity may be defined by its interaction with other pathogenicity factors. Aim of this work was the characterization of EHI\_127670 by the implementation of BioID in *E. histolytica*, a method that detects interactions partners of a protein by biotinylating them with the protein BirA\*, which is fused to the protein of interest.

First, to determine the gene product of *ehi\_127670*, the gene was expressed under its own promoter in a myc-tag expression vector in *E. histolytica*. Western blotting and immunofluorescence microscopy (IFA) identified EHI\_127670 as a protein, localized in small, spherical vesicles, termed undefined granulae. For BioID, the expression of BirA\* failed, presumably because of its bacterial origin. However, the expression of the EHI\_127670: BirA\* fusion protein succeeded and biotinylation experiments were performed. A strong non-specific biotinylation was observed in the BioID transfectants and the control, presumably by the presence of an unknown biotinylating protein in *E. histolytica*. However, after subtraction of the noise signal, seven potential interaction partners of EHI\_127670 were detected in the pellet fraction. The localization of the fusion protein by IFA failed, assumably because of a massive noise signal caused by fragments of the protein in the cytosol. The BioID method was established successfully, but needs to be optimized by implementation of a suitable control, the reduction of unspecific biotinylation, and the localization of the fusion protein, before the detected potential interaction partners and thereby EHI\_127670 itself may be characterized.

# 1 Introduction

*Entamoeba histolytica* is an enteric protozoan parasite that infects nearly 50 million people worldwide annually, particularly in developing countries. It is the causative pathogen of amebiasis, a neglected tropical disease that is responsible for more than 55,000 deaths per year (Lozano *et al.*, 2012). Whereas 90% of infections remain asymptomatic, 10 % of cases develop an invasive form of amebiasis, with severe symptoms like colitis and amoebic liver abscesses (ALA). This form is characterized by the invasion of parasites into the intestinal tissue and the bloodstream, dispersing the amoeba to other organs, forming abscesses particularly in the liver. The circumstances that lead to the invasive form of amebiasis are not fully understood.

Besides immunological response of the host (Helk *et al.*, 2015) and microbiomic interactions (Burgess *et al.*, 2016), the expression of pathogenicity factors, which are proteins involved in the destruction of host tissue, is expected to have an influence. The identification of new pathogenicity factors is a major research area. By now, the majority of identified pathogenicity factors can be assigned to the protein families amoebapores, cysteine peptidases, and Gal/GalNAc-lectins (Laughlin *et al.*, 2005). Although many mechanisms of those proteins are individually understood, their interplay leading to pathogenicity remains unknown. Moreover, the existence of homologues for most of those proteins in the non-pathogenic *Entamoeba dispar*, indicates the involvement of additional, unidentified factors.

Recently, the new putative pathogenicity factor EHI\_127670 was discovered by comparative transcriptome analysis of pathogenic and non-pathogenic clones of *E. histolytica*. Phenotypic characterization of overexpressing transfectants confirmed its enhancing properties on pathogenicity (Meyer *et al.*, 2016). Besides that, the structure, function, and locus of the putative protein and its role in pathogenicity remain uncharacterized. Since the pathogenicity of *E. histolytica* is presumably a complex mechanism that is composed of the interaction of known and unknown pathogenicity factors, the function of EHI\_127670 might be elucidated by consideration of its interaction partners.

In this work, the BioID method was implemented in *E. histolytica* to detect potential interaction partners of EHI\_127670. BioID is used for proximity-dependent labelling of protein-protein interactions in living cells. The enzyme BirA\* is fused to a protein of interest and biotinylates potential interaction partners. For complex mechanisms that involve a variety of interacting partners, the thereby documented interactome may elucidate the correlation of the interplaying elements.

## 1.2 Research Aim

The role of the recently discovered putative pathogenicity factor EHI\_127670 in the virulence of *E. histolytica* is unknown and might be defined by its interaction with other pathogenicity factors. The aim of this work was the establishment of BioID in *E. histolytica* for the characterization of EHI\_127670. First, to localize the gene product of *ehi\_127670*, it was expressed under its own promoter in a myc-tag expression vector and localized by immunofluorescence microscopy (IFA). Later, the BioID method was established in *E. histolytica* to detect potential interaction partners of EHI\_127670.

## 2 Theoretical Background

### 2.1 Infection and Life Cycle of *E. histolytica*

The enteric protozoan parasite *Entamoeba histolytica* is a human pathogen that infects nearly 50 million people per year worldwide, with a mortality of more than 55,000 deaths annually (Lozano *et al.*, 2012). Although distributed worldwide, the infections have a significantly higher prevalence in tropical and subtropical developing countries, which is linked to contaminated water and food and bad hygienic conditions, promoting faecal-oral transmission (Stanley, 2003).

The life cycle of *E. histolytica* is relatively simple and consists of two phases. Infective cysts of the parasite are ingested by the consumption of contaminated water or food. After ingestion, they resist the acidic environment in the stomach. In the small intestine, basic pH leads to excystation and vegetative trophozoites develop. The developing trophozoites migrate through the small intestine until they reach the colon, where they divide asexually and feed mainly on the phagocytosis of bacteria and cell debris. The parasites encyst in the lower intestine, can then be excreted in the faeces and may infect another host through contaminated water or food. The cysts can remain infectious for several months (Bercu *et al.*, 2007). This life cycle is also referred to as the non-invasive form of amebiasis and remains asymptomatic.

### 2.2 Cell biology of *E. histolytica*

The vegetative trophozoites of *E. histolytica* have an oval to round, constantly shifting shape with a diameter of 10-50  $\mu\text{m}$  (Stanley, 2003). Their relatively fast motility is accomplished by pseudopodia, which are used for active phagocytosis as well. The cytoskeleton has a high actin content (Bailey *et al.*, 1992). The cell is enclosed by a thin, cholesterol-rich cell membrane (Andrä *et al.*, 2004). Approx. 40 % of the cytosol consists of membrane bound vesicles, forming

a complex endomembrane system that has a variety of functions (Scholze *et al.*, 1994, Perdomo *et al.*, 2014). It includes endosomes, lysosomes, secretory vesicles, multivesicular bodies, and mitosomes (Perdomo *et al.*, 2014; Smith *et al.*, 2010). A Golgi-like apparatus was discovered, which is compartmented in multiple singular vesicles (Mazzuco *et al.*, 1996). The endoplasmic reticulum was considered to be compartmented as well, but was characterized to be continuous later (Teixeira *et al.*, 2008). Lysosomes and endosomes are part of the endocytic pathway, enabling phagocytosis, which is crucial for the parasites nutrition when it feeds on bacteria and lysed cells (Andrä *et al.*, 2003). Secretory vesicles contain pathogenicity factors like amoebapores and cysteine peptidases, which are secreted to destroy host tissue by lysing cells or cleaving cell-cell connections (Leippe, 1997; Bruchhaus *et al.*, 2015).

The genomic DNA is located in one or more nuclei. It is 20 Mb big and codes for predicted 8,201 genes, of which 3,788 could be assigned to putative proteins. There are 58 *Entamoeba* specific protein families of five or more members that have no known homologues in other organisms (Lorenzi *et al.*, 2010). The AT content of 75 % is very high. Whole transcriptome sequencing indicated the lack of alternative splicing (Hon *et al.*, 2012). The genome is organized in chromosomes, but the presence of linear and circular DNA molecules makes it difficult to determine the karyotype and the ploidy of *E. histolytica* (Willhoeft *et al.*, 1999).

## 2.3 Amebiasis

It is estimated that 90 % of infections with *E. histolytica* remain asymptomatic (Stanley, 2003). The parasites colonize the intestine without the invasion of the host tissue, causing a non-invasive form of amebiasis. In 10 % of infections, the virulence of the trophozoites causes the invasive form of amebiasis, resulting in the destruction of the colonic mucosa and invasion of other organs (Bercu *et al.*, 2007). Infections that result in invasive amebiasis are accompanied by massive anatomical lesions in the intestine and the liver. The first lesions occur when the parasites invade the colonic mucosa, after degradation of the secreted mucin allowed the migration through the mucous layer (Moncada *et al.*, 2003). The parasites attach to the host epithelial tissue and lyse the cells, followed by phagocytic ingestion (Laughlin *et al.*, 2005). After invasion of the intestine tissue, which may lead to inflammation and amoebic colitis, the

trophozoites can enter the bloodstream, where they are ultimately transported to other organs, particularly the liver. The parasites proceed to destroy the host tissue and may provoke an inflammatory response, which ultimately leads to the formation of ALA (Wells *et al.*, 2004). Without treatment, invasive amebiasis may lead to death. The disease can be treated by medication with metronidazole, which kills the invasive parasites in liver and intestinal tissue. Subsequently, non-invasive parasites in the intestine are removed by diloxanide furoate or paromomycine (Stanley, 2003). The evolutionary background of the invasive amebiasis remains obscure. Since the invasive amoebae do not form cysts and the formation of ALA can lead to the death of the host, the invasive form represents a dead end in the development of the parasites. The circumstances and mechanisms that lead to the invasive form are unknown. Besides immunological response of the host (Helk *et al.*, 2015) and microbiomic interactions (Burgess *et al.*, 2016), the expression of pathogenicity factors, proteins that are associated with the virulence of *E. histolytica*, is expected to have an influence.

## 2.4 State of Research

### 2.4.1 Culture Isolates

For the identification of new pathogenicity factors, comparative analysis of pathogenic and non-pathogenic clones of *E. histolytica* is an important tool in current research. For the comparability of those clones, cells of the same culture isolate need to be used. All clones of *E. histolytica* used in this work derive from the same culture isolate HM-1:IMSS, which was isolated from a patient with amoebic colitis in 1967. Today, there are two sub-types of HM-1:IMSS, the pathogenic cell line B and the non-pathogenic cell line A. Both cell lines are in axenic culture at the Bernhard Nocht Institute since 1991 (B) and 2001 (A). The pathogenicity is defined by the ability of the cells to form ALA in the mouse model, when injected into the liver. Whereas cell line B is still capable of ALA formation, cell line A lost this ability. Since both cell lines derive from the same culture isolate, they share the same genetic background. However, a transcriptomic heterogeneity was observed in both cell lines (Biller *et al.*, 2014).

Therefore, twelve clones, each derived from one cell of the subtype, were generated to achieve transcriptomic homogeneity. The pathogenicity of the clones A1-A12 and B1-B12 was examined in the mouse model. Whereas none of the clones of cell line A induced ALA formation, the pathogenicity of clones B1-12, measured by the size of the generated ALA, showed high variation (Meyer *et al.*, 2016). For the investigation of new pathogenicity factors by comparative transcriptomic analyses of pathogenic and non-pathogenic clones, A1 and B8 were selected as non-pathogenic, and B2 was selected as pathogenic clone.

#### 2.4.2 Pathogenicity Factors

The pathogenicity of *E. histolytica* is characterized by its enormous capability to destroy the host tissue. The lysis process starts with the adhesion of the parasite to the host cell, followed by secretion of various proteins that cleave cell junctions or perforate the cell plasma membrane, which ultimately leads to its death. Subsequently, fragments of the lysed cell are phagocytosed (Laughlin *et al.*, 2005). The activity of the parasites provokes an immune response, which may contribute to abscess formation (Bruchhaus *et al.*, 2015). Proteins, that are involved and contribute to those processes, are called pathogenicity factors. Today, three major protein families could be identified. Gal/GalNAc-lectins, amoebapores and cysteine peptidases mediate the process of cell adhesion, lysis, and tissue disruption.

Gal/GalNAc-inhibitable lectin plays a role in adhesion, signal transduction and escape from the host complement system. It consists of a 205 kDa heavy heterodimer that is attached to a 150 kDa subunit. The protein complex has a specific binding affinity for galactose (Gal) and N-acetyl-D-galactosamine (GalNAc), enabling the trophozoite to attach to the host cell membrane. Inhibition of cell attachment by free Gal and GalNAc or specific antibodies reduced the adherence and cytolytic capacity of the parasite (Laughlin *et al.*, 2005; Mann, 2002). It was hypothesized that the cell attachment via Gal/GalNAc-lectins is followed by a signal cascade, which leads to the regulated secretion of amoebapores (Leippe *et al.*, 1997). Besides this protein family, the proteins EhCPADH112, L220 and SREHP are associated with cell adherence, although their ligands were not discovered yet (Laughlin *et al.*, 2005).

Amoebapores are pore-forming peptides, able to disrupt the membranes of host cells or phagocytosed bacteria (Andrä *et al.*, 2003). There are three similar isoforms. The major isoform, amoebapore A, consists of 77 amino acids and forms an alpha-helical structure. It was suspected that four alpha-helix bundles form the tertiary structure of the amoebapore, which resembles the structure of the antimicrobial peptide NK-Lysin isolated from pig intestine (Leippe *et al.*, 1997). Amoebapores are equipped with signal peptides, which allow the translocation to their storage compartments, the lysosomal granulae. Those are involved in two pathways that are linked to pathogenicity. In the secretory pathway, the vesicles containing amoebapores are secreted after cell attachment through Gal/GalNAc-lectins. The peptides interact with the host cell membrane forming pores, which act as ion channels and permeabilize the membrane, which ultimately leads to the death of the cell (Leippe *et al.*, 1997). Its membrane composition with a high content of cholesterol protects the amoeba from its own weapons (Andrä *et al.*, 2004). In the phagocytic pathway, amoebapore-containing lysosomes are recruited to the phagosomes, where the peptides are discharged to engulfed bacteria, disrupting the membrane and lysing the cells (Andrä *et al.*, 2003).

Cysteine peptidases (CP) are hydrolase enzymes that degrade proteins, forming the biggest group of peptidases in *E. histolytica*. Of the approx. 50 annotated CP encoding genes, 35 can be assigned to a C1 papain-like superfamily (Bruchhaus *et al.*, 2015). CP were discovered in multiple compartments of the cell. They were found to be membrane associated (Jacobs *et al.*, 1998), cytoplasmatic (Que *et al.*, 2002), and located in lysosome-like granulae (Scholze *et al.*, 1994). Like amoebapores, CP can be secreted for host tissue disruption, but instead of lysing cells by membrane permeabilization, CP cleave the junctions between cells, which leads to tissue disassembly (Laughlin *et al.*, 2005). In the phagocytic pathway, CP in the granulae and the cytosol are recruited to phagosomes, where they help digesting bacteria and cell debris (Que *et al.*, 2002). The secretion of CP through lysosome-like granulae allow the parasites to degrade mucin, migrate through the colonic mucosa and ultimately penetrate the colonic lamina propria (Bruchhaus *et al.*, 2015). Furthermore, the hydrolytic properties of CP lead to tissue damages by degradation of the extracellular matrix (Laughlin *et al.*, 2005).

It was discovered that CP modulate the hosts immune response by cleaving the pro-inflammatory cytokine IL-18, thereby suppressing the induction of  $\gamma$ -interferon and the recruitment of macrophages (Que *et al.*, 2003). Moreover, the IL-1 $\beta$  converting activity of CP may initiate the inflammatory response that leads to severe damage to the intestinal tissue (Zhang *et al.*, 2000). Although the exact mechanisms remain unknown, CP could be identified as major factor for the formation of ALA in animal models (Bruchhaus *et al.*, 2015).

Although many mechanisms of the known pathogenicity factors are individually understood, their interplay leading to the disruption of host tissue remains unknown. Moreover, the existence of homologues for most of those proteins in the non-pathogenic *Entamoeba dispar* (Ximénez *et al.*, 2017; Nickel *et al.*, 1999; Pillai *et al.*, 2000), indicates the involvement of additional, unidentified factors. In the past years, technological progress enabled the development of new methods for the discovery of new pathogenicity factors of *E. histolytica*. Next-generation sequencing was used for the comparative transcriptome analysis of the pathogenic clone B2 and the non-pathogenic clones A1 and B8 (Meyer *et al.*, 2016).

In that study, the new putative pathogenicity factor EHI\_127670 was discovered. It demonstrated an enhancing influence on ALA formation in the animal model. Its phenotypic characterization confirmed its pathogenicity in *in vitro* assays (Haferkorn 2018; Nehls, 2020). However, apart from its phenotypic implications, the structure, function, locus, and interaction partners of EHI\_127670 remain unknown. There are no homologues or conserved regions present within the sequence. Since the expression and the localization of the putative protein in *E. coli* repeatedly failed (Haferkorn 2018; unpublished), it was disputed if the sequence coded for a protein or a long non-protein coding RNA (lncRNA).

### 2.4.3 BioID

In this work, the BioID method was implemented in *E. histolytica* for the first time, to detect potential interaction partners of EHI\_127670. BioID is used for proximity-dependent labelling of protein-protein interactions in living cells, enabling the identification of strong and transient interactions over a period of time. Protein-protein interactions are crucial for the function of all living systems and play an important role in infection processes (Taylor *et al.*, 2011). BioID is based on the biotinylating activity of BirA\*, a R118G mutant of the *E. coli* biotin protein ligase BirA. Whereas wild-type BirA specifically biotinylates the bacterial acetyl-CoA carboxylase by adding reactive biotin to a specific lysine residue of the protein, the mutant BirA\* lost its specificity. The promiscuous biotin ligase has a lower affinity for reactive activated biotin and prematurely releases it non-specifically (Roux *et al.*, 2012). Thereby, all proteins in a radius of approx. 10 nm are biotinylated by BirA\* (Kim *et al.*, 2014).

Prerequisite of the biotinylation is the supplementation of biotin to the medium, which can be used as temporal control. A fusion protein of the 35 kDa BirA\* and a protein of interest is expressed in the organism. Incubation of the cells with biotin leads to the biotinylation of potential interaction partners of the protein by BirA\* (Roux *et al.*, 2012). Since the molecular weight of 35 kDa is relatively high and might affect translocation, the proper localization of the fusion protein needs to be examined. There are several ways for the detection of interaction partners. The most valuable method is the detection by mass spectroscopy, which directly identifies interaction partners, but requires titration of the biotin application and purification of the proteins after cell lysis (Sears *et al.*, 2019). Another option is the separation of the proteins by SDS-PAGE and detection of biotinylated bands in a streptavidin blot, which only characterizes the size of potential interaction partners. Since all proteins in spatial proximity of BirA\* are biotinylated, false-positive signals of proteins not interacting with the protein of interest need to be identified. A BirA\*-only control expressed in a second control transfectant is a simple way to detect non-specific biotinylation.

With the BioID method, the interactions of any protein can be recorded in living cells. For complex mechanisms that involve a variety of interacting partners, the thereby documented interactome may elucidate the correlation of the interplaying elements.

## 3 Materials

### 3.1 Laboratory Devices

**Table 1:** Devices and manufacturers.

<b>Device</b>	<b>Type</b>	<b>Manufacturer</b>
Heating block	<i>ThermoMixer C</i>	Eppendorf
Benchtop centrifuge	<i>Centrifuge 5424</i>	Eppendorf
Benchtop centrifuge, cooled	<i>5427 R</i>	Eppendorf
Centrifuge	<i>J2-21</i>	Beckmann Coulter
Swing-out rotor centrifuge	<i>GPKR Centrifuge</i>	Beckmann Coulter
Rotor centrifuge	<i>JA-10</i>	Beckmann Coulter
Rotor centrifuge	<i>JA-12</i>	Beckmann Coulter
Shaking incubator	<i>Certomat HK</i>	B.Braun International
Incubator	<i>B 6060</i>	Heraeus
Incubator	<i>Heratherm IGS750</i>	Thermo Scientific
Thermal cycler	<i>Primus</i>	MWG Biotech
Magnet stirrer	<i>RET-GS</i>	IKAMAG
Inverted microscope	<i>Eclipse TS100</i>	Nikon
Confocal microscope	<i>FluoView1000</i>	Olympus
Microscope	<i>Axio Imager M2</i>	Zeiss
Nano-photometer	<i>NanoDrop™ 2000</i>	Thermo Fisher Scientific
pH meter	<i>Lab 850</i>	SI Analytics
Agarose GE system small	<i>Horizon 58</i>	Life Technologies
Agarose GE system big	<i>40-1214-R</i>	Peqlab Biotechnologie
GelDoc	<i>ChemiDoc™ XRS+</i>	Bio-Rad
SDS-PAGE system	<i>45-1010-i</i>	Peqlab Biotechnologie

Electroblotter	<i>PerfectBlue Web S™</i>	Peqlab Biotechnologie
Electrophoresis power Supply	<i>EPS 301</i>	amersham pharmacia
Biosafety cabinet class II	<i>Kendro HS 18</i>	Heraeus
Electroporation system	<i>GenePulser Xcell</i>	Bio-Rad
Pipette controller	<i>S1 Pipet Filler</i>	Thermo Scientific

## 3.2 Software

**Table 2:** Software and manufacturers.

<b>Software</b>	<b>Manufacturer</b>
<i>Mac Vector 12.6.0</i>	International Biotechnologies
<i>NEB Tm Calculator 1.13.0</i>	New England Biolabs
<i>OligoEvaluator™</i>	PREMIER Biosoft
<i>CorelDRAW Graphics Suite 2021</i>	Corel Corporation
<i>NanoDrop™ 2000 Operating Software</i>	Thermo Scientific
<i>Image Lab</i>	Bio-Rad
<i>Imaris</i>	Oxford Instruments

### 3.3 Disposables

**Table 3:** Disposables. Materials of commonplace use are not listed.

<b>Product</b>	<b>Manufacturer</b>
Cell culture flasks (35 mL, 70 mL, 250 mL)	Becton Dickinson
Gene Pulser electroporation cuvettes	Bio-Rad
Semi-micro cuvettes	Sarstedt
Novex gel cassettes 1 mm	Thermo Scientific
Nitrocellulose membrane 0.2 µm	GE Healthcare Life Sciences
Whatmanpaper	GE Healthcare Life Sciences
Cryopreservation tubes	Sarstedt

### 3.4 Chemicals

**Table 4:** Chemicals and manufacturers.

<b>Product</b>	<b>Manufacturer</b>
1-Propanol	Sigma-Aldrich
Acrylamide solution 30%	AppliChem
Adenosine triphosphate	Sigma-Aldrich
Agarose	Invitrogen
Ampicillin	Roth
<i>Ampuwa</i> <sup>®</sup> water	Fresenius Kabi
Ammonium chloride	Sigma-Aldrich
Ammonium iron(III) citrate	Sigma-Aldrich
Ammonium persulfate (APS)	AppliChem

Ascorbic acid	VWR Chemicals
Bovine serum albumin	Sigma-Aldrich
Bromophenol blue	Sigma-Aldrich
CaCl <sub>2</sub> (calcium chloride)	Merck
DAPI (4',6-diamidino-2-phenylindole)	Fluka
Diamond Vitamin <i>Tween</i> <sup>®</sup> 80 solution	JRH Biosciences
DMSO (Dimethyl sulfoxide)	Roth
DTT (Dithiothreitol)	Roth
EDTA (Ethylenediaminetetraacetic acid)	Sigma-Aldrich
Ethanol	Merck
<i>Fetal Calf Serum (FCS)</i>	Capricorn
<i>Genetecin</i> <sup>®</sup> (G418)	Thermo Scientific
Glucose	Roth
Glutathione	Sigma-Aldrich
Glycerol	Roth
Glycine	Roth
HCl (hydrogen chloride)	Roth
HEPES	biomol
Hydrogen peroxide	Sigma-Aldrich
K <sub>2</sub> HPO <sub>4</sub> (Dipotassium hydrogen phosphate)	Merck
KH <sub>2</sub> HPO <sub>4</sub> (Monopotassium phosphate)	Merck
KCl (Potassium chloride)	Merck
KOH (Potassium hydroxide)	Roth
Luminol	Sigma-Aldrich
Methanol	Roth
MgCl <sub>2</sub> (magnesium chloride)	Merck
NaCl (sodium chloride)	Roth
Na <sub>2</sub> HPO <sub>4</sub> (disodium phosphate)	Sigma-Aldrich

NaH <sub>2</sub> PO <sub>4</sub> (sodium dihydrogen phosphate)	Sigma-Aldrich
NaN <sub>3</sub> (sodium azide)	Roth
NaOH (Sodium hydroxide)	Fluka Chemicals
Orange G	Sigma-Aldrich
Paraformaldehyde	Merck
p-coumaric acid	Sigma-Aldrich
Powdered milk	Roth
Saponin	Sigma-Aldrich
SDS (sodium dodecyl sulfate)	Roth
Sodium acetate	Roth
Tetramethyl ethylenediamine (TEMED)	Bio-Rad
Tricine	Roth
Tris-Base	Roth
Trypticase peptone	Becton Dickinson
Tris-HCl	Roth
<i>Tween</i> <sup>®</sup> 20	Merck
Yeast extract	Becton Dickinson
TOPO <sup>®</sup> salt solution	Thermo Scientific

---

## 3.5 Buffers and Solutions

**(50x) TAE Buffer**


---

Sodium acetate	
· 3 H <sub>2</sub> O	250 mM
Tris base	2 M
EDTA	50 mM

---

**(10x) Na PBS**


---

Na <sub>2</sub> HPO <sub>4</sub>	67 mM
NaH <sub>2</sub> PO <sub>4</sub>	33 mM
NaCl	1.4 M
pH	6.8

---

**Stacking gel solution for SDS-PAGE**


---

Tris base	500 mM
SDS (w/v)	0.4%
pH	6.8
sterile filter	

---

**(2x) Laemmli Buffer**


---

Tris HCl	150 mM
Glycerol (v/v)	20%
SDS (w/v)	2%
Bromophenol blue(w/v)	0.25%
pH	6.8
add 10 mM DTT before use	

---

**Electrophoresis buffer SDS PAGE**


---

Tris base	25 mM
Glycin	192 mM
SDS (w/v)	0.1 %
pH	8.8

---

**Separating gel solution for SDS-****PAGE**


---

Tris base	1.5 M
-----------	-------

SDS (w/v)	0.4%
-----------	------

pH	8.8
----	-----

sterile filter

---

**(10x) TBS buffer**


---

NaCl	150 mM
------	--------

Tris Base	10 mM
-----------	-------

---

**TBS-T buffer**


---

TBS Buffer

Tween 20 (v/v)	0.1%
----------------	------

---

**ECL solution A**


---

Tris HCl	100 mM
----------	--------

Luminol (w/v)	2.5%
---------------	------

---

**ECL solution B**


---

DMSO

p-coumaric acid

(w/v)	0.1%
-------	------

---

**Powdered milk blocking solution**


---

TBS buffer

Powdered milk (w/v)	5%
---------------------	----

---

**BSA blocking solution**


---

TBS-T

BSA (w/v)	5%
-----------	----

---

**Transfer buffer Western Blot**


---

Glycine	192 mM
---------	--------

Tris base	25 mM
-----------	-------

SDS	1.3 mM
-----	--------

Methanol (v/v)	20%
----------------	-----

---

**Cytomix**

KCl	120 mM
CaCl <sub>2</sub>	0.15 mM
K <sub>2</sub> HPO <sub>4</sub>	10 mM
KH <sub>2</sub> PO <sub>4</sub>	10 mM
HEPES	25 mM
EDTA	2 mM
MgCl <sub>2</sub>	5 mM
pH (KOH)	7.6

complete with:

ATP	4 mM
Glutathione	10 mM

**(3x) Gel Buffer for Tricine SDS-PAGE**

Tris base	3 M
SDS	10.5 mM
pH	8.45

**Anode Buffer Tricine SDS PAGE**

Tris base	2 M
pH	8.9

**Solution S for Tricine SDS-PAGE**

Tris base	500 mM
SDS	14 mM
NaN <sub>3</sub>	1.5 mM
pH	6.8

**(2x) Sample Buffer Tricine SDS-PAGE**

Solution S (v/v)	2.5 %
SDS (w/v)	2 %
Glycerol (v/v)	30 %
Orange G (w/v)	0.04 %

**Cathode Buffer Tricine SDS-PAGE**

Tris base	1 M
Tricine	1 M
SDS (w/v)	1 %
pH	8.25

All buffers and solutions were prepared with ddH<sub>2</sub>O and sterilized before use, pH was adjusted with 6 M NaOH or 6 M HCl. Sterilization was performed by high-pressure steam sterilization (20 min, 121 °C, 1.2 bar overpressure), or sterile filtration (pore diameter 0.2 µm) for temperature-sensitive reagents.

### 3.6 Reaction Kits

**Table 5:** Reaction kits and manufacturers.

<b>Product</b>	<b>Manufacturer</b>
<i>NucleoSpin® Plasmid mini</i>	Macherey-Nagel
<i>NucleoBond® Xtra Midi</i>	Macherey-Nagel
<i>NucleoSpin® Gel and PCR clean-up</i>	Macherey-Nagel

### 3.7 DNA- and Protein Markers

**Table 6:** Used DNA- and protein markers.

<b>Product</b>	<b>Manufacturer</b>
<i>GeneRuler™ 1 kb Ladder plus</i>	Thermo Scientific
<i>PageRuler™ Prestained Protein Ladder</i>	Thermo Scientific
<i>PageRuler™ Plus Prestained Protein Ladder</i>	Thermo Scientific

### 3.8 Enzymes and Reaction Buffers

**Table 7:** Used enzymes.

<b>Product</b>	<b>Type</b>	<b>Manufacturer</b>
<i>FD BamHI</i>	Restriction Endonuclease	Thermo Scientific
<i>FD EcoRI</i>	Restriction Endonuclease	Thermo Scientific
<i>FD KpnI</i>	Restriction Endonuclease	Thermo Scientific
<i>FD BglII</i>	Restriction Endonuclease	Thermo Scientific
<i>T4 DNA Ligase</i>	Ligase	Invitrogen
<i>Taq-Polymerase</i>	Polymerase	Promega

**Table 8:** Used reaction buffers.

<b>Product</b>	<b>Type</b>	<b>Manufacturer</b>
<i>FD Green Buffer (10x)</i>	digestion buffer	Thermo Scientific
<i>T4 Ligase Buffer (10x)</i>	ligation buffer	Invitrogen
<i>5X Green GoTaq® Flexi</i>	PCR reaction buffer	Promega

### 3.9 Antibodies

**Table 9:** Used Antibodies.

<b>Product</b>	<b>Type</b>	<b>Species</b>	<b>Manufacturer</b>
<i><math>\alpha</math>-c-myc monoclonal</i>	1st Antibody	mouse	Sigma-Aldrich
<i><math>\alpha</math>-HA monoclonal</i>	1st Antibody	rat	Roche
<i><math>\alpha</math>-170 kDa lectin polyclonal</i>	1st Antibody	rabbit	AG Bruchhaus
<i><math>\alpha</math>-SOD polyclonal</i>	1st Antibody	rabbit	AG Bruchhaus
<i><math>\alpha</math>-mouse IgG HRP</i>	2nd Antibody	rabbit	Dako
<i><math>\alpha</math>-rat IgG HRP polyclonal</i>	2nd Antibody	goat	Jackson
<i>Streptavidin-HRP</i>	Streptavidin-HRP	-	Thermo Scientific
<i><math>\alpha</math>-mouse Alexa Fluor 488</i>	2nd Antibody	goat	Thermo Scientific
<i><math>\alpha</math>-rabbit Alexa Fluor 594</i>	2nd Antibody	goat	Thermo Scientific
<i><math>\alpha</math>-rat Alexa Fluor 488</i>	2nd Antibody	goat	Thermo Scientific

### 3.10 Vectors

**Table 10:** Used Vectors.

<b>Plasmid</b>	<b>Type</b>	<b>Resistance</b>	<b>Size</b>	<b>Reference</b>
pCR2.1 TOPO	cloning vector	Ampicillin	3.9 kb	Invitrogen
pNC	expression vector	Ampicillin; G418	6.0 kb	Hamann <i>et al.</i> 1995
pNCcMyc	expression vector	Ampicillin; G418	6.03 kb	Matthiesen, 2009
pNCBID	expression vector	Ampicillin; G418	6.1 kb	This work

## 3.11 Oligonucleotides

**Table 11:** Used Oligonucleotides and target sequences.

<b>Target Sequence; Primer Name</b>	<b>Primer forward (_f) Sequence</b>	<b>Primer reverse (_r) Sequence</b>	<b>purpose</b>	<b>Reference</b>
contig DS571518; P_contigEHI127	CCAAAAGAAAGAAGA AATTGATTAATAAT TCAC	TCTTATACAACACA ATTCATTTTCATGTA AAATATG	sequencing	Eurofins
EHI_127670 +500 bp 5'; P_500EHI127	GAGAGGTACCGAATA AATGGAGTAAATAAG	GAGAGGATCCGTAT AAGCAAGTAACAAT ATGTTC	cloning	Eurofins
pNC Insert; NeoCasSequenzie	GAATTAGTCTCAACTC AACAATGTTTGTTGG	-	sequencing	Eurofins
<i>ehi_127670</i> insert for BirA; P_BioIDEHI127	GAGAGGTACCATGTCC CTTAGTTTGTCAACTCT TCC	GAGAGGATCCGTAT AAGCAAGTAACAAT ATGTTCCCG	cloning	Eurofins
BirA*+linker+HA; P_BID	GAGAGGTACCCGGGA TCCGGAGGTGGAGGT AGTGGAGG	GAGAAGATCTTTAA GCATAATCTGGAAC ATCATATGG	cloning	Eurofins
BirA* 3'- end; P_3'BIDS	GAGAGGCTGAAGTATT AAGAAAATTAGGTGC	-	sequencing	Eurofins

NeoCasSequenzie was provided by AG Bruchhaus. Other Oligonucleotides in Table 11 were designed by me.

## 3.12 Organisms

In this section, all used organisms are described according to their taxonomic classification. The *E. coli* strain *OneShotTop10™* was used for cloning. For all experiments with *E. histolytica*, the clones A1 and B2, derived from the culture isolate HM-1:IMSS, were used.

### 3.12.1 *Escherichia coli*

#### Taxonomic classification

Domain	Bacteria
Phylum	Proteobacteria
Class	Gammaproteobacteria
Order	Enterobacterales
Family	Enterobacteriaceae
Genus	<i>Escherichia</i>
Species	<i>E. coli</i>

**Table 12:** Used strain of *E. coli*.

Product	Strain; Genotype	Type	Manufacturer
<i>OneShotTop10™</i>	<i>DH5α; F-mcrAΔ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 lacX74 deoR rec A1 araD139 Δ(araleu)7697 gal U gal K rpsL (StrR) end A1 nupG</i>	chemically competent	Thermo Scientific

### 3.12.2 *Entamoeba histolytica*

#### Taxonomic classification

Domain	Eukaryota
Phylum	Amoebozoa
Class	Archamoebae
Genus	<i>Entamoeba</i>
Species	<i>Entamoeba histolytica</i>

All clones of *E. histolytica* that were used for experiments derive from the same pathogenic culture isolate HM-1:IMSS, which was isolated from a patient with colitis in 1964 (American Type Culture Collection (ATCC), No. 30459). Today several subtypes are cultivated. Cell line B is a pathogenic subtype of this isolate and capable of forming ALA in rodents. Cell line B derived clone B2 showed high pathogenicity in the mouse model. Cell line A and the derived clone A1 are incapable of forming ALA in rodents, therefore A1 is considered non-pathogenic. Both cell lines are in axenic culture at BNITM since 1991 and 2001, respectively.

**Table 13:** Used clones of *E. histolytica*.

Clone	Cell line	pathogenic	Culture Isolate	Reference, Date
A1	A	no	HM-1:IMSS	ATCC, 2001
B2	B	yes	HM-1:IMSS	ATCC, 1991

### 3.13 Culture Media

#### 3.13.1 Bacterial Culture Media

Bacterial cultures were cultivated in LB medium or on LB agar, which was prepared with ddH<sub>2</sub>O and autoclaved (20 min, 121 °C, 1.2 bar overpressure). LB agar was completed with ampicillin and poured into petri dishes (diameter: 100 mm). LB medium was completed with ampicillin immediately before use. Petri dishes and medium bottles were stored at 4 °C.

<b>LB Medium</b>		<b>LB Agar</b>	
Lennox L Broth Base	20 g	Lennox L Broth Agar	35 g
ddH <sub>2</sub> O	1 l	ddH <sub>2</sub> O	1 l

#### 3.13.2 Amoebic Culture Media

Clones and transfectants of *E. histolytica* were cultivated in TY-I-SS medium (Diamond *et al.*, 1978). For the incomplete medium all components were solved in 4.35 L ddH<sub>2</sub>O in a 5 l Erlenmeyer-flask. pH value was measured by a pH electrode and adjusted to 6.8 with 6 M NaOH. The medium was divided into 10 parts of 400 ml and filled into laboratory glass bottles. High-pressure steam sterilization (20 min, 120 °C, 1.4 bar overpressure) was performed in a pressure cooker. Bovine serum was inactivated two times for 30 minutes at 56 °C. After cooling down, each bottle of medium was completed with 50 mL of inactivated bovine serum, 15 mL *Diamond Vitamin Tween 80 Solution*, 130 µg/ml penicillin and 270 µg/ml streptomycin. The completed medium was stored at 4 °C for no more than two weeks.

**TY-I-SS Medium**


---

Trypticase	100 g
Yeast extract	50 g
Glucose	50 g
NaCl	10 g
K <sub>2</sub> HPO <sub>4</sub>	3,8 g
KH <sub>2</sub> PO <sub>4</sub>	3 g
L-Cysteine	5 g
Ascorbic acid	1 g
Ammonium iron(III) citrate	0,114 g

complete 400 ml with:

Bovine Serum	50 ml
<i>Diamond Vitamin Tween</i>	15 ml
<i>80 Solution</i>	
Penicillin	130 µg/ml
Streptomycin	270 µg/ml

---

All antibiotics used in amoebic and bacterial culture media were sterile-filtered (pore diameter 0.2 µm) and added to the autoclaved media. They are depicted in the following table.

**Table 14:** Used antibiotics for cell culture.

<b>Antibiotic</b>	<b>Stock concentration</b>	<b>Final concentration</b>	<b>Manufacturer</b>
Ampicillin	100 mg/ml	100 µg/ml	Sigma-Aldrich
Penicillin	60 mg/ml	130 µg/ml	Roth
Streptomycin	125 mg/ml	270 µg/ml	Roth

## 4 Methods

### 4.1 Cloning

#### 4.1.2 Myc-tag Construct

For the characterization and localization of the gene product of *ehi\_127670*, the gene was to be expressed under its own promoter in a myc-tag expression vector in *E. histolytica*. For this, a 500 bp long sequence upstream of start ATG, which presumably contains the promoter region, and the entire open reading frame of *ehi\_127670* were amplified using the primers P\_contigEHI127. The annotation of the sequence in the database was checked by sanger sequencing. A *KpnI* restriction sequence was added as 5'-overhang to the forward primer and a *BamHI* restriction sequence was added as 5'-overhang to the reverse primer for PCR amplification of the genomic sequence. Genomic DNA from the pathogenic clone B2 was used as a template. The PCR program is depicted in section 4.1.6. The annealing temperature was set to 53 °C and the extension time was set to 54 s. The amplified insert containing the 5'- *KpnI* and 3'- *BamHI* restriction sites was cloned into the pNCcMyc vector as described in section 4.1.4.

#### 4.1.3 BioID Construct

For BioID, a fusion protein of a genetically modified BirA ligase from *E. coli*, termed BirA\*, and the protein of interest EHI\_127670 was to be expressed in *E. histolytica*. To assure sufficient mobility of BirA\*, a flexible 7xGGGGS linker was placed between BirA\* and EHI\_127670. A 3xHA-tag C-terminal of BirA\* was used to check the expression and the correct localization of the EHI\_127670: BirA\* fusion protein. For amplification of *birA\** together with the 7xGGGGS linker and 3xHA-tag, the forward primer P\_BID\_f with a 5'-overhang containing the cloning site *KpnI-SmaI-BamHI* and the reverse primer P\_BID\_r with a 5'-overhang containing the

restriction site *BglIII* were used for PCR. The construct pSLI-var1-BirA\*-HA, designed for BioID in *Plasmodium falciparum* and provided by Jakob Cronshagen (AG Spielmann, BNITM), was used as template. The PCR program is depicted in section 4.1.6. The annealing temperature was set to 51 °C and the extension time was set to 80 s. The amplified *birA\** sequence containing the 5'-*KpnI-SmaI-BamHI* cloning site and the 3'-*BglIII* restriction site was cloned into the pNC vector as described in section 4.1.4. The *BglIII* restriction site of the insert and the *BamHI* restriction site of the pNC vector pair complementarily and are uncleavable by restriction enzymes after ligation. The thereby constructed pNCBID vector was preserved in a bacterial glycerol stock as described in section 4.1.15. The insert *ehi\_127670* was amplified using the forward primer with a 5'-*KpnI* restriction site overhang and the reverse primer with a 5'-*BamHI* restriction site overhang. DNA from the pathogenic clone B2 was used as a template. The annealing temperature was set to 60 °C and the extension time was set to 26 s. The amplified insert containing the 5'-*KpnI* and 3'-*BamHI* restriction sites was cloned 5'-terminal of *birA\** into the pNCBID vector as described in section 4.1.4.

#### 4.1.4 Cloning Process

Both constructs were assembled by TA-cloning and subcloning. First, primers were designed to amplify a desired DNA sequence by PCR using genomic or plasmid DNA as a template (sections 4.1.5 and 4.1.6). The amplicons were cloned into a *pCR2.1 TOPO* vector by TA cloning (section 4.1.7). After plasmid propagation in *E. coli* (sections 4.1.12 and 4.1.13), the integrity of the plasmids was checked by sanger sequencing (section 4.1.14). The inserts were subcloned from the parental TOPO vector to a destination vector by enzymatic digest, electrophoretic separation, DNA preparation, and ligation (sections 4.1.8 - 4.1.10). Plasmid propagation was performed in *E. coli* and a clone was preserved on a plate (sections 4.1.12 and 4.1.13). Integrity of the plasmids was checked by diagnostic digest and sanger sequencing (sections 4.1.7 and 4.1.12). The finished constructs were propagated in *E. coli* (section 4.1.11) and transfected into *E. histolytica* trophozoites (section 4.2.3).

#### 4.1.5 Primer Design

Oligonucleotides serving as primers were designed to complementarily bind sequences flanking the gene of interest, one forward primer binding 5'-terminal to the 3'-end of the antisense strand and one reverse primer binding 3'-terminal to the 3'-end of the sense strand. A guanine-adenine-guanine-adenine (GAGA) protective sequence at the 5'-end and restriction sites for subsequent subcloning were added. For primer design, a sufficient length > 20 bp, similar melting temperatures  $T_m$ , avoidance of primer-dimer formation, and if possible, a GC-clamp had to be considered. Primer design was performed with *NEB Tm Calculator 1.13.0* and *OligoEvaluator™*. All oligonucleotides were ordered at Eurofins Scientific SE.

#### 4.1.6 Polymerase Chain Reaction

PCR was used for amplification of specific DNA sequences. Depending on the source of the construct to be amplified, either genomic DNA or plasmid DNA was used as a template. One PCR reaction was conducted with following reagents:

##### PCR reagents

<i>5x Green Go Taq Flexi Buffer</i>	5 $\mu$ l
Magnesium chloride (25 mM)	1.5 $\mu$ l
dNTP's (2.5 mM)	2.5 $\mu$ l
Forward primer	2.5 $\mu$ l
Reverse primer	2.5 $\mu$ l
<i>Taq Polymerase</i>	0.5 $\mu$ l
ddH <sub>2</sub> O	9.5 $\mu$ l
DNA template (< 0.02 $\mu$ g/ $\mu$ l)	1 $\mu$ l

Per PCR process, two reaction mixes with DNA template dilutions of 1:100 and 1:1000 were prepared. All reagents were mixed in a 0.2 mL reaction tube and placed into a thermal cycler (*Primus*, MWG Biotech). The PCR program consisted of following steps with specific temperature and duration:

#### PCR Program

Initial Denaturation	95 °C	4 min	
Denaturation	95 °C	30 s	
Annealing	X °C	40 s	30 Cycles
Extension	72 °C	Y s	
Final Extension	72 °C	10 min	

The annealing temperature X and extension time Y were adjusted for every reaction. The annealing temperature was determined by the melting points of the primers and was calculated with *NEB Tm Calculator 1.13.0*. The extension time was defined by the speed of the used *Taq-polymerase* (1 kb/min) and the length of the amplicon. During the final extension step, the non-template-dependent terminal transferase activity of the polymerase adds a single deoxyadenosine to the 3'-end of the amplicates. The thereby created *sticky ends* were required for TOPO cloning (section 4.1.7).

#### 4.1.7 TOPO Cloning

TOPO cloning was the first cloning step in the cloning process, where the PCR product was ligated into the TOPO vector. The enzyme topoisomerase I is covalently bound at both ends of the linear vector and incorporates amplicons with 3'-deoxyadenosine-sticky ends. This allows rapid cloning at room temperature without additional ligase.

The presence of a  $\beta$ -lactamase gene and the location of the multiple cloning site in a  $\beta$ -galactosidase gene allow selection by ampicillin resistance and blue-white screening. The multiple cloning site is flanked by M13 forward and reverse priming sites for construct verification by sequencing. Sequencing was conducted for every construct to ensure the correct length and orientation of the insert.

The amplicates from PCR were pooled and purified with *NucleoSpin® Gel and PCR clean-up* columns by Macherey-Nagel. The ligation into the TOPO vector was performed with following reagents:

#### **TOPO ligation reagents**

TOPO Salt solution	1 $\mu$ l
pCR2.1 TOPO vector	0.5 $\mu$ l
ddH <sub>2</sub> O	3.5 $\mu$ l
DNA insert	0.5 - 1 $\mu$ l

All reagents were mixed in a 1.5 mL reaction tube, spun down, incubated at room temperature for 5 minutes, then put on ice. Later, 2  $\mu$ l of the reaction mixture was used for transformation (section 4.1.12).

#### 4.1.8 Ligation

Ligation of DNA inserts into the destination vectors pNC, pNCcMyc and pNCBID was performed with T4 DNA Ligase by Invitrogen. The ligase ligates DNA along the sticky ends produced by restriction endonucleases during digestion (section 4.1.9). It was used for subcloning DNA inserts, derived from the parental TOPO vector by enzymatic digestion, into the destination vector. Following reagents were used for ligation:

**Ligation reagents**

<i>T4 Ligase buffer (10x)</i>	1.5 $\mu$ l	
<i>T4 DNA ligase</i>	1 $\mu$ l	
ATP (10 mM)	2.5 $\mu$ l	
DNA vector	X $\mu$ l	mass ratio
DNA insert	Y $\mu$ l	insert to vector:
ddH <sub>2</sub> O	Z $\mu$ l	1:2 - 1:5
Total volume	15 $\mu$ l	

All reagents were mixed in a 0.2 ml reaction tube. The volumes of vector, insert and ddH<sub>2</sub>O were adjusted for a mass ratio of insert to vector of 1:2 to 1:5. Per ligation process, three to five reactions with varying mass ratios and volumes were prepared. The tubes were spun down and placed in a thermal cycler (*Primus*, MWG Biotech). Ligation was conducted at 14 °C overnight. Later, 10  $\mu$ l of each reaction mixture was used for transformation (section 4.1.12).

#### 4.1.9 Restriction Digest

Enzymatic restriction digestion of DNA was conducted for diagnostic digests of constructed plasmids or for subcloning DNA fragments. The digestion was performed by endonucleases, which cut DNA at enzyme-specific restriction sites, producing sticky-ends that can be ligated to DNA fragments with complementary sticky-ends. Diagnostic digestion was used to check the correct insertion of a DNA fragment cloned into a vector. Preparative digestion was used to extract inserts from a parental TOPO vector and to cut-open the destination vectors pNC, pNCcMyc and pNCBID for subsequent subcloning. For one digestion reaction, following reagents of the platform *FastDigest (FD)* by *Thermo Scientific* were used:

FD Restriction: diagnostic		FD Restriction: preparative	
<i>FD Green Buffer (10x)</i>	1 $\mu$ l	<i>FD Green Buffer (10x)</i>	4 $\mu$ l
<i>FD enzyme 1</i>	0.5 $\mu$ l	<i>FD enzyme 1</i>	2 $\mu$ l
<i>FD enzyme 2</i>	0.5 $\mu$ l	<i>FD enzyme 2</i>	2 $\mu$ l
ddH <sub>2</sub> O	7 $\mu$ l	ddH <sub>2</sub> O	27 $\mu$ l
DNA (< 1 $\mu$ g)	1 $\mu$ l	DNA (< 5 $\mu$ g)	5 $\mu$ l

All reagents were mixed in a 1.5 mL reaction tube, spun down, and incubated in a heating block at 37 °C for 15 min. Preparative restriction mixes were incubated for 30 min to ensure a complete digestion. The reaction mixture was then separated by agarose gel electrophoresis (section 4.1.10).

#### 4.1.10 Agarose Gel Electrophoresis and DNA Preparation

Agarose gel electrophoresis was used for size-dependent separation of DNA molecules, separating digested DNA fragments for subcloning and for construct verification. In a gel chamber, negative charged DNA molecules migrate along an electric field through an agarose gel to a positively charged anode, with a velocity inversely proportional to their size. DNA fragments of different sizes in samples are separated and can be evaluated with a size marker. A 1% (w/v) agarose-gel was produced by solving agarose in TE-buffer, boiling at 600 W in a microwave oven. The solution was mixed with four drops of a 0.025% ethidium-bromide solution and casted into a horizontal gel chamber. A sample comb for creation of sample pockets was inserted. After polymerization of the gel, the comb was removed and the chamber filled with TE-buffer. DNA samples from restriction digestion containing loading buffer and dye were pipetted into the sample pockets. Sample volume for construct verification was 10  $\mu$ l, for preparative separation 40  $\mu$ l. 10  $\mu$ l of *GeneRuler™ 1 kb Ladder plus* was used as a size marker. The poles of the chamber were connected to the power supply and 90 V was applied for 50 minutes. After electrophoresis, DNA bands with intercalated ethidium

bromide could be detected by illumination with UV light. Gels were photographed under UV in a GelDoc. For DNA preparation, gels were placed on an UV transilluminator desk and bands of DNA, containing cut-out inserts or cut-open vectors, were cut out with a scalpel. Cut out bands were weighted and DNA was purified with a *NucleoSpin® Gel and PCR clean-up* kit.

#### 4.1.11 Measurement of DNA concentration

After each purification step with the kits *NucleoSpin® Gel and PCR clean-up*, *NucleoSpin® Plasmid mini* and *NucleoBond® Xtra Midi*, 1.5 µl sample was used to measure DNA concentration photometrically with the nanophotometer *NanoDrop™ 2000*.

#### 4.1.12 Transformation of *E. coli* and Selection

For plasmid propagation, chemically competent *E. coli* OneShotTop10™ (DH5α) cells were transformed by heat shock transformation and subsequently cultivated on LB-amp agar and LB-amp medium. For heat shock transformation, aliquoted cells stored at -80 °C were thawed on ice. For transformation of TOPO vectors, 2 µl of the ligation mixture, for pNC, pNCcMyc and pNCBID 10 µl of the ligation mixture was added to the cells. The mixture was incubated on ice for 5 minutes. Heat shock transformation was performed in a heating block at 42 °C for 30 seconds, right after that the cells were incubated on ice for 5 minutes. 250 µl SOC-medium was added and the mixture was incubated in an Eppendorf ThermoMixer® C shaking dry block heater at 37 °C and 650 rpm for 60 minutes. For transformations with a TOPO vector, 40 µl X-gal for blue-white screening was added. 200 µl of the cell suspension was plated on a LB-amp agar petri dish and incubated at 37°C overnight. The next day, colonies of ampicillin-resistant transformants were visible. For transformations with a TOPO vector, blue-white screening allowed additional selection of white clones with vectors containing a correctly incorporated insert. Clones were picked using a 10 µl pipet tip and transferred into a 15 mL culture vial or a 1 l Erlenmeyer-flask containing 4 ml or 400 ml LB-amp medium and cultivated (section 4.1.13).

#### 4.1.13 Cultivation of *E. coli* and Plasmid Purification

Transformants of *E. coli* OneShotTop10™ (DH5α) cells were cultivated in a *Certomat HK* shaking incubator at 37 °C overnight. For propagation of small plasmid amounts for cloning and subcloning, 4 ml LB-amp medium in a 15 ml culture vial was inoculated with a picked clone and cultivated. The next day, cells were harvested by transferring 2 ml of cell suspension into a 2 ml reaction tube and subsequent centrifugation at 1200 x *g* for 30 seconds. The supernatant was discarded. This was repeated with fresh 2 ml of cell suspension. The remaining supernatant was removed by carefully knocking down the reaction tube on a piece of tissue paper. Further processing of the pellet and DNA purification was performed with the *NucleoSpin® Plasmid mini* kit. For clone preservation and subsequent inoculation, the empty culture vial was filled with 5 ml of NaPBS. A sterile inoculation loop was used to transfer and spread a minimal amount of cell suspension on a LB-amp agar petri dish, which was incubated at 37 °C overnight and then could be stored at 4°C for several weeks. One petri dish was divided into marked areas and could be used for preservation of several clones. For transfection of *E. histolytica*, higher amounts of plasmid DNA were needed. For that purpose, 400 ml LB-amp medium in a 1 l Erlenmeyer-flask was inoculated with a picked clone and cultivated. The next day, optical density (OD) of the cell suspension was measured photometrically. For optimal cell yield and to avoid overloading of the *NucleoBond® Xtra Midi* column, the mass of cells could not exceed a certain value. To approximate that mass, the product of OD and volume in ml, ODV was set to 400 by the purification protocol. Therefore, only a part of the cell suspension was transferred to a 500 ml centrifuge bottle and centrifuged at 5000 x *g* and 4 °C for 15 minutes. The supernatant was discarded. Further processing of the pellet and DNA purification was performed after the *NucleoBond® Xtra Midi* protocol. Elution was completed with ddH<sub>2</sub>O and a final concentration of at least 1 µg/µl was achieved. After concentration measurement, DNA was inactivated at 80 °C for 20 minutes.

#### 4.1.14 Construct Verification

The integrity of the constructed plasmids was checked by diagnostic digest or sanger sequencing. For sanger sequencing, samples were sent to *microsynth synthlab GmbH*. Forward and reverse M13 primers for TOPO-sequencing are stored there and could be chosen at the ordering process. For the *destination vectors* pNC, pNCcMyc and pNCBID, primers were premixed to the sample. The oligonucleotide NeoCasSequenzie was used as a forward primer to check the integrity of the 5'-part of the insert. The oligonucleotide P\_3'BIDS was used to check the integrity of the 3'-end of the BID construct in the pNCBID vector. Samples were prepared for sequencing as follows:

##### Sanger sequencing

DNA	60-100 ng
ddH <sub>2</sub> O	12 µl
Primer solution (100 pM)	3 µl (optional)

#### 4.1.15 Cryopreservation of *E. coli*

Bacterial glycerol stocks were prepared for long-time plasmid preservation. 4 mL LB-amp medium was inoculated with a picked clone and cultivated overnight. 850 µl of cell suspension was mixed with 150 µl of 99 % glycerol in a 1.5 ml reaction tube and stored at -80 °C.

## 4.2 Gene Expression in *E. histolytica*

For myc-tag localization and BioID experiments, specific genes had to be expressed in *E. histolytica*. During the cloning process, these genes were cloned into vectors and propagated. The constructed plasmids were then used for transfection of *E. histolytica* by electroporation. After successful transfection and selection, the genes of interest were expressed in the transgenic trophozoites. For generation of transfectants, trophozoites of the clones A1 and B2 were cultivated and harvested. After transfection, transfectants were selected and cultivated for subsequent experiments.

### 4.2.1 Cultivation of *E. histolytica*

Trophozoites of the clones A1 and B2, as well as generated transfectants, were cultivated in complete TY-I-SS medium at 37 °C under microaerophilic and axenic conditions. For transfectants, 20 µg/µl *Geneticin* (G418) was supplemented to the medium. Cultivation was conducted in non-ventilated T12.5 cell culture flasks. Every first, third and fifth day of the week, trophozoites were split and fresh medium was added. To determine the right splitting ratio, cell confluency was checked under an inverted microscope and a suitable splitting ratio was estimated. In a class II biosafety cabinet, medium was decanted, then a defined volume of fresh medium was added with a 5 ml pipet and the culture flask was closed. By thorough shaking of the flask, semi-adherent trophozoites were detached and suspended in the medium. A definite volume of the suspension was discarded according to the desired splitting ratio. The flask with the remaining trophozoites was filled with fresh medium, and for transgenic cultures, G418 was added. The flask was shaken to evenly disperse the cells. Every two weeks new culture flasks were used.

#### 4.2.2 Harvest of *E. histolytica*

For experiments with *E. histolytica*, a sufficient amount of trophozoites was needed. Therefore, during the splitting process, excess cell suspension was not discarded, but transferred into a T25 culture flask. The flask was filled up with fresh medium and cultivated for 1-2 days. Cell harvest was performed at approx. 80% confluency during the exponential growth phase to ensure optimal cell condition. Medium was decanted and 10 ml cold NaPBS were poured into the flask. The flask was closed and shaken thoroughly to detach semi-adherent cells. The cell suspension was transferred to a 15 ml falcon tube and centrifuged at 1400 x *g* and 4 °C for 5 minutes. The supernatant was discarded and the pellet washed with 10 ml of cold NaPBS. The pellet was resolved in 1 mL of cold NaPBS, transferred into a 1.5 ml reaction tube, and centrifuged at 1200 *g* for 5 minutes. The supernatant was discarded. The remaining supernatant was removed by carefully knocking down the reaction tube on a piece of tissue paper.

#### 4.2.3 Transfection of *E. histolytica* and Selection

The finished plasmids for myc-tag localization and BioID were transfected into *E. histolytica* by electroporation. The transfection process was performed under sterile conditions in a class II biosafety cabinet. All reagents were put on ice. *E. histolytica* clones were split into a T25 cell culture flask and cultivated until cell confluency reached a monolayer. These cells were transferred into a T75 cell culture flask and cultivated overnight. The medium was decanted. To detach as many cells as possible, the culture flask was put on ice for 5 minutes. The trophozoites were then harvested and washed as described in section 4.2.2, but not transferred into a 1.5 ml reaction tube. Instead, the pellet was resuspended in 10 ml incomplete cytomix and centrifuged at 1400 x *g* and 4 °C for 5 minutes. The supernatant was discarded. DNA was diluted with ddH<sub>2</sub>O to 1 µg/µl, and 100 µl were transferred into a 4 mm electroporation cuvette and put on ice. 8 ml of cytomix was completed. Right after, the cell pellet was resolved in 2 ml completed cytomix, and 800 µl of the cell suspension was pipetted into the cuvette and mixed with the DNA. The contacts of the cuvette were dried with a tissue

paper and the cuvette was inserted into the electroporator. The electroporation parameters were set to 1.2 kV and 25  $\mu$ FD. The sample was pulsed twice. The time constant of the pulses was checked and compared to the expected value of 0.3-0.6 ms. The cell suspension was transferred with a sterile 3 ml Pasteur pipette to a T25 cell culture flask containing prewarmed TY-I-SS medium. The flask was shaken to disperse the cells and cultivated at 37 °C for two days. Every construct was transfected in duplicate into clone A1 and B2, respectively, producing the transfectants A1a, A1b, B2a and B2b. The name of the transfectant was completed by the type of transfected vector and the insert abbreviation superscripted. Example: The first transfectant of B2, transfected with pNCBID containing the insert EHI\_127670, was named B2aBID<sup>127</sup>. Selection was induced after two days by adding 10  $\mu$ g/ $\mu$ l of G418. The medium was changed three times a week and G418 concentration was kept at 10  $\mu$ g/ $\mu$ l until cell confluency reached a monolayer, then G418 was increased to 20  $\mu$ g/ $\mu$ l. When full cell confluency was reached again, a fraction of the cells was split into a T12.5 culture flask and the major part was harvested as described in section 4.2.2 for subsequent experiments.

#### 4.2.4 Production of *E. histolytica* Cell Lysate

Cell lysate was produced for the separation of cellular proteins of transfectants of *E. histolytica* by SDS-PAGE. Transfected trophozoites were lysed by the freeze and thaw method. After cell harvesting (section 4.2.2), 1.5  $\mu$ l of protease inhibitor E64 was added to the pellet to protect the proteins from degradation by peptidases. The reaction tube was frozen in liquid nitrogen and thawed at 37 °C for one minute. During the thawing step, the tube was vortexed repeatedly to create shearing forces that disrupt cell membranes. Completely thawing and heating of the pellet was avoided to prevent protein degradation. The freeze and thaw cycle was repeated five times. After the last thawing step, the sample was centrifuged at 12000 x *g* and 4 °C for 15 minutes. The supernatant containing the cell lysate was removed carefully with a 200  $\mu$ l micropipette.

## 4.3 SDS-PAGE

### 4.3.1 Glycine-SDS-PAGE after Laemmli

To separate the proteins expressed in *E. histolytica* transfectants for subsequent investigation, SDS-PAGE was conducted with cell lysates and pellets. Gel solutions for discontinuous SDS-PAGE were casted into 1 mm Novex Gel Cassettes immediately after addition of APS and TEMED. Three quarters of the cassette was filled with a 10 % separation gel and covered with isopropanol. After polymerization, the isopropanol was decanted and a 4 % stacking gel was casted on top of the separating gel. A sample comb was inserted to produce sample pockets. Polymerization time was approx. 30 minutes. The gels were prepared as follows:

<b>Glycine-SDS-PAGE Separating Gel</b>		<b>Glycine-SDS-PAGE Stacking Gel</b>	
Acrylamide (30%)	2.5 ml	Acrylamide (30%)	0.65 ml
Separating gel solution	1.9 ml	Separating gel solution	1.25 ml
ddH <sub>2</sub> O	3.1 ml	ddH <sub>2</sub> O	3.05 ml
APS (25%)	25 µl	APS (25%)	25 µl
TEMED	5 µl	TEMED	5 µl

Two gel cassettes were placed in a vertical electrophoresis chamber and the chamber was filled with an electrophoresis buffer. For sample preparation, 15 µl of cell lysate was mixed with 15 µl 2x Laemmli buffer and 1.5 µl DTT. Cell pellets were dissolved in 100 µl 2x Laemmli buffer and 5 µl DTT was added. The samples were heated at 90 °C for 10 minutes in a heating block to denature the proteins. The sample comb was removed and 30 µl sample mixture was pipetted into each pocket. As a protein size marker, 10 µl of *PageRuler™ Plus Prestained Protein Ladder* was pipetted into the first pocket. The chamber was closed and connected to the power supply. 40 mA was applied for approx. 20 minutes until the dye front passed the

stacking gel, then the current was increased to 60 mA. The proteins denatured by the anionic tenside SDS were separated by migration through the gel with a velocity inversely proportional to their molecular weight. After approx. 40 minutes, the dye front had passed the separation gel and the separation was stopped. The disposable gel cassettes were cracked open and the gel was removed for subsequent identification of proteins by western blot or streptavidin blot.

#### 4.3.2 Tricine-SDS-PAGE

For small proteins under 10 kDa the separation performance of glycine-SDS-PAGE is limited. To increase the separation resolution of the myc-tagged EHI\_127670 protein, tricine-SDS-PAGE was used. Instead of only one electrophoresis buffer, two buffers, one for the cathode chamber and one for the anode chamber were used. The gels were prepared as follows:

<b>Tricine-SDS-PAGE Separating Gel</b>		<b>Tricine-SDS-PAGE Stacking Gel</b>	
Acrylamide (40%)	2.5 ml	Acrylamide (30%)	0.75 ml
3x Tricine Gel Buffer	2.25 ml	3x Tricine Gel Buffer	1.13 ml
Glycerol (87 %)	1.5 ml	ddH <sub>2</sub> O	3.23 ml
ddH <sub>2</sub> O	1.25 ml	APS (25%)	12 µl
APS (25%)	12.8 µl	TEMED	3.75 µl
TEMED	3.75 µl		

The gel solutions were casted into 1 mm Novex gel cassettes immediately after addition of APS and TEMED.  $\frac{3}{4}$  of the cassette was filled with the 13% separating gel, the stacking gel was casted immediately on top until the solution overflowed. A sample comb was inserted to produce sample pockets. The polymerization time was 2 hours. 2x tricine sample buffer was used for sample preparation. Insertion of the cassettes into the electrophoresis chamber and sample preparation were performed as described in section 4.3.1, but with a different sample buffer. Cathode buffer was poured into the cathode chamber, anode buffer into the anode chamber.

The electrophoresis chamber was closed and connected to the power supply. 20 mA was applied until the dye front had passed the stacking gel, then 40 mA was applied until the dye front had passed the separating gel. After the run was stopped, gel cassettes were cracked open and the gel was removed for subsequent western blotting.

#### 4.4 Western Blot

To check the expression and the coarse localization of the myc-tagged EHI\_127670 and the EHI\_127670:BirA\* fusion protein, western blotting was performed with the gels from glycine- and tricine-SDS-PAGE. In an electroblotter, the protein bands of the gel are transferred to a nitrocellulose membrane by migration through an electric field. Afterwards, the membrane is blocked and incubated with first and secondary antibodies with conjugated HRP in several steps, which bind specifically to one desired protein, resulting in a specific, chemiluminescent detection of that protein.

For electroblotting, two sponges, six pieces of whatman paper, and a nitrocellulose membrane were soaked in transfer buffer. A bracket was opened and one sponge and three pieces of whatman paper were placed inside. Air bubbles were rolled out with a 10 ml pipet. The glycine or tricine gel was placed on top, followed by the nitrocellulose membrane and three pieces of whatman paper. After each layer, air bubbles were rolled out. A sponge was placed on top of the stack, and the bracket was closed and inserted into the electroblotting chamber, with the gel being located between the membrane and the negatively charged cathode. Up to four brackets were placed inside, the blotting chamber was filled with transfer buffer and stirred moderately on a magnet stirrer. The electrodes of the chamber were connected to the power supply and 400 mA was applied. Glycine gels were blotted for 60 minutes, tricine gels for 40 minutes.

The membrane was transferred to a 50 ml falcon tube filled with 20 ml of 5 % milk powder-TBS solution, and incubated at room temperature on a tube roller for 30 minutes. Thereby the membrane was saturated with milk powder protein and blocked against non-specific antibody binding. The blocking solution was discarded and 5 ml of 2.5 % milk powder-TBS pipetted into the tube. 5  $\mu$ l primary antibody solution,  $\alpha$ -c-myc for the myc-tagged EHI\_127670 or  $\alpha$ -HA for

the EHI\_127670:BirA\* fusion protein, was added to create a 1:1000 dilution. The membrane was incubated at 4 °C overnight on a tube roller. After that, the antibody solution was transferred to a 15 ml falcon tube and stored at -20 °C. The membrane was washed three times for 10 min at room temperature on a tube roller, once with a 0.05 % TBS-Tween solution and twice with TBS. Another blocking step was carried out with 20 ml of 5 % milk powder-TBS solution for 30 minutes. The blocking solution was discarded and 5 ml of 2.5 % milk powder-TBS added. 2.5 µl secondary antibody solution,  $\alpha$ -mouse HRP for the myc-tagged EHI\_127670 or  $\alpha$ -rat HRP for the EHI\_127670:BirA\* fusion protein, was added to create a 1:2000 dilution. The secondary antibodies were coupled with the enzyme horseradish peroxidase (HRP), which catalyses the chemiluminescent detection reaction. The membrane was incubated for 120 min at room temperature on a tube roller. The antibody solution was transferred to a 15 ml falcon tube and stored at -20 °C. Three washing steps were conducted as described before. The developer solution was produced by mixing 5 ml of solution A with 0.5 ml of solution B and 1.5 µl of hydrogen peroxide. For developing, the membrane was placed in a petri dish, covered in the developer solution and incubated for 5 min. The antibody-conjugated HRP catalyses the reaction of luminol and hydrogen peroxide in the developer solution, which leads to luminescence of the desired protein bands. For detection the membrane was placed in a *ChemiDoc XRS+* and exposed for 1-300 s in a signal accumulation mode.

## 4.5 Biotinylation

The BioID method uses the biotinylation activity of the enzyme BirA\* to biotinylate endogenous proteins that interact with a protein of interest fused to the enzyme. For that matter, biotin in a sufficient concentration must be present in the cells. Since no paper about the use of this method in *E. histolytica* has been published yet, a suitable biotin application had to be estimated by consideration of the deployment of the BioID method in organisms related to *E. histolytica*. BioID was successfully established in the protozoa *Plasmodium falciparum* (Birnbaum *et al.*, 2020) and *Dictyostelium discoideum* (Batsios *et al.*, 2016), which served as orientation. In both papers, the end concentration of biotin was 50 µM and biotinylation was performed for 24 h before cell harvest.

In this work, 85  $\mu\text{M}$  biotin was applied for 24 h for a first proof-of-concept approach. For every culture of transfectants incubated with biotin, one flask of the same transfectants was cultivated in medium containing 0.91  $\mu\text{M}$  biotin and served as a control. Later, the medium inherent concentration of 0.91  $\mu\text{M}$  was used for biotinylation without addition of extra biotin.

## 4.7 Streptavidin Blot

To investigate the biotinylation of proteins interacting with EHI\_127670:BirA\*, a streptavidin blot was performed with a gel from glycine-SDS-PAGE. Streptavidin blotting resembles western blotting, but instead of a first and secondary antibody system, HRP-conjugated streptavidin with a strong affinity to biotin serves as the binding agent. The biotin content of milk powder makes it necessary to use a special blotting protocol with only one blocking step and intensified washing steps, as well as a BSA solution as a dilution reagent.

The electroblotting process was performed as described in section 4.4. The membrane was transferred to a 50 ml falcon tube filled with 20 ml of a 0.1 % TBS-tween solution containing 5 % milk powder and incubated at room temperature on a tube roller for 120 minutes. The blocking solution was discarded and the membrane was washed five times at room temperature on a tube roller, three times with a 0.1 % TBS-Tween solution for 5 min and two times with TBS for 5 and 10 min. 5 ml of 5 % BSA solved in 0.1 % TBS-tween solution was pipetted into the tube. 5  $\mu\text{l}$  streptavidin-HRP solution was added to create a 1:1000 dilution. The membrane was incubated at 4 °C overnight on a tube roller. After that, the streptavidin-HRP solution was transferred to a 15 ml falcon tube and stored at -20 °C. The membrane was washed as described above. After that, the membrane was developed as described in section 4.4, but with an exposure time of 1-2400 s to amplify very low signals.

## 4.8 Immunofluorescence Assay (IFA)

IFA was conducted with transfectants of *E. histolytica* to localize the myc-tagged EHI\_127670 and the EHI\_127670:BirA\* fusion protein in cellular compartments, to determine the natural locus of EHI\_127670 and the correct translocation of EHI\_127670:BirA\*. The cytoplasmic superoxide dismutase (SOD) and the cell surface-bound 170 kDa lectin were also localized and served as controls, covering the cytoplasm and the outer membrane for a potential co-localization of EHI\_127670.

During IFA, cellular proteins of interest are stained by a system of a first, protein specific antibody and a second antibody, that carries a conjugated fluorophore and binds to the first antibody. After incubation and staining with both antibodies, cells are placed under a fluorescence microscope and illuminated with a laser of a desired, fluorophore-specific wavelength. The antibody-bound fluorophore is excited by the light and emits light of a different, specific wavelength, that is detected.

For IFA, trophozoites of transfectants were split into a T25 cell culture flask, cultivated and harvested as described in section 4.2.1. and 4.2.2. Instead of a falcon tube, a 15 ml culture tube was used for harvest. The tube was carefully shaken to dissolve the cell pellet in the remaining buffer of the previous washing step. While still shaking the tube, 1 ml of a 4% PFA NaPBS solution was pipetted directly to the pellet to fixate the cells. Fixation occurs by reaction of the aldehyde group of PFA with amine groups of cellular proteins, cross linking them to a static meshwork. The tube was closed tightly and incubated for 30 minutes.

This and all following incubation steps were carried out at room temperature on a tube roller. All centrifugation steps were performed at room temperature and 600 x *g* for 3 min. The cell suspension was divided into two 1.5 ml reaction tubes and centrifuged. The supernatant was discarded and the pellet washed with 500  $\mu$ l NaPBS to remove unbound PFA. For the investigation of proteins inside the cells, 0.05 % saponin was added to NaPBS for this and the following washing steps. Saponin has detergent properties and permeabilizes cell membranes gently by selectively removing cholesterol, thereby antibodies can pass through the membrane. Samples for the investigation of surface proteins were washed without saponin.

All resuspension steps were performed by gently snapping the reaction tube instead of vortexing, to preserve the fragile fixated cells. After centrifugation, supernatant was discarded, the pellet resuspended in 500  $\mu$ l of 50  $\mu$ M ammonium chloride solution with or without saponin, and incubated for 15 min. Ammonium chloride reacts with free aldehyde groups of PFA and blocks them to prevent nonspecific binding to amine groups of antibodies. After a centrifugation step and washing with 500  $\mu$ l NaPBS with or without saponin, the pellet was resuspended in 500  $\mu$ l 2% FCS solution and incubated for 10 min to reduce nonspecific binding of antibodies. After centrifugation, the supernatant was discarded and the pellet washed with 500  $\mu$ l NaPBS with or without saponin. The pellet was resuspended in 200  $\mu$ l NaPBS and 2  $\mu$ l of a primary antibody solution was added, resulting in a 1:100 dilution. Depending on which proteins were to be localized, one or two primary antibody solutions were added to one sample. Besides the myc-tagged EHI\_127670 and the EHI\_127670:BirA\* fusion protein, 170 kDa lectin and SOD were localized as positive controls. As negative control, only the second antibody was applied. The different proteins and the used first and second antibodies are depicted in Table 15. The sample was incubated for 1 h, centrifuged and washed three times with 500  $\mu$ l NaPBS. The pellet was resuspended in 200  $\mu$ l NaPBS and 0.5  $\mu$ l of one or two secondary antibody solutions was added for a 1:400 dilution. The sample tube was placed in a brown coloured, non-transparent 50 ml falcon tube and incubated in the dark for 1h. After centrifugation and discarding of the supernatant, the sample was washed 3 times with 500  $\mu$ l NaPBS. To stain the nucleus, the pellet was resuspended in 200  $\mu$ l NaPBS and 0.5  $\mu$ l of DAPI stain, which is a blue fluorophore and binds to the DNA in the nucleus, was added. The sample was incubated for 10 min in the dark in a non-transparent falcon tube. After centrifugation, the supernatant was discarded.

Before solving the pellet in a final volume of NaPBS, the mass of the over the many washing steps shrunk pellet was examined. The volume was adjusted to weight up an acceptable cell density against the volume sufficient for an expected number of microscopic investigations. Typically, the applied volume ranged from 30 to 200  $\mu$ l. The samples were stored at 4 °C in the dark for no longer than two weeks. For fluorescence microscopy, the sample was mixed gently and 10  $\mu$ l was pipetted to a microscope slide. A coverslip was placed on top and sealed with mineral oil to prevent the sample from drying out. Microscopy was performed under a *Zeiss Axio Imager M2* microscope and an *Olympus FluoView1000* confocal microscope.

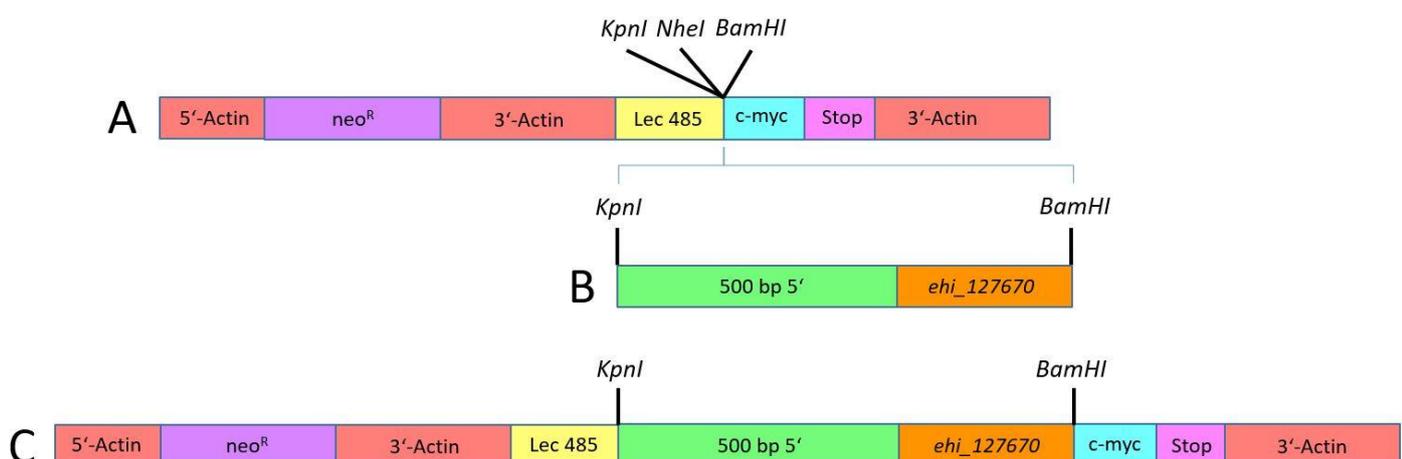
**Table 15:** stained proteins and used antibodies.

<b>Protein</b>	<b>1<sup>st</sup> Antibody</b>	<b>2<sup>nd</sup> Antibody</b>	<b>colour</b>
EHI_127670:myc	<i><math>\alpha</math>-c-myc monoclonal (mouse)</i>	<i><math>\alpha</math>-mouse Alexa Fluor 488</i>	green
EHI_127670:BirA*	<i><math>\alpha</math>-HA monoclonal (rat)</i>	<i><math>\alpha</math>-rat Alexa Fluor 488</i>	green
170 kDa lectin	<i><math>\alpha</math>-170 kDa lectin polyclonal (rabbit)</i>	<i><math>\alpha</math>-rabbit Alexa Fluor 594</i>	red
SOD	<i><math>\alpha</math>-SOD polyclonal (rabbit)</i>	<i><math>\alpha</math>-rabbit Alexa Fluor 594</i>	red

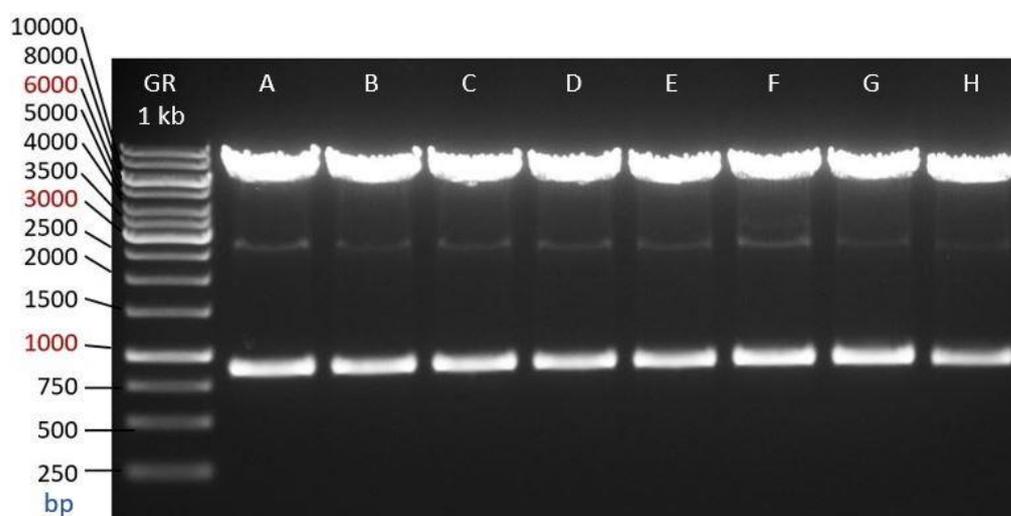
## 5 Results

### 5.1 Myc-tag Localization

The aim of the method was to accomplish the expression of the myc-tagged *ehi\_127670* in a way as close to the natural expression of *ehi\_127670* as possible, first to determine whether it is a protein or lncRNA, and second to localize it in its natural locus. To achieve this, a 500 bp long sequence 5' of *ehi\_127670*, containing the promoter sequence was sequenced and included to the expression vector pNCcMyc. Thereby, the putative protein could be characterized as protein or a long non-protein coding RNA (lncRNA), since a gene for a lncRNA expressed under its own promoter will most likely not be translated into a protein, despite being located on an artificial construct and transfected into a cell. The sequencing of the 500 bp long regions flanking *ehi\_127670* in the contig *DS571518* confirmed the annotation in the database *Amoebadb.org*. The pNCcMyc:500bp+*ehi\_127670* construct was assembled as described in section 4.1.2. The cloning strategy is visualized in figure 1. Eight clones were picked and analysed. The diagnostic digest of the finished constructs, depicted in figure 2, shows the correct fragment sizes for all clones. Clone F was sequenced and propagated for transfection.

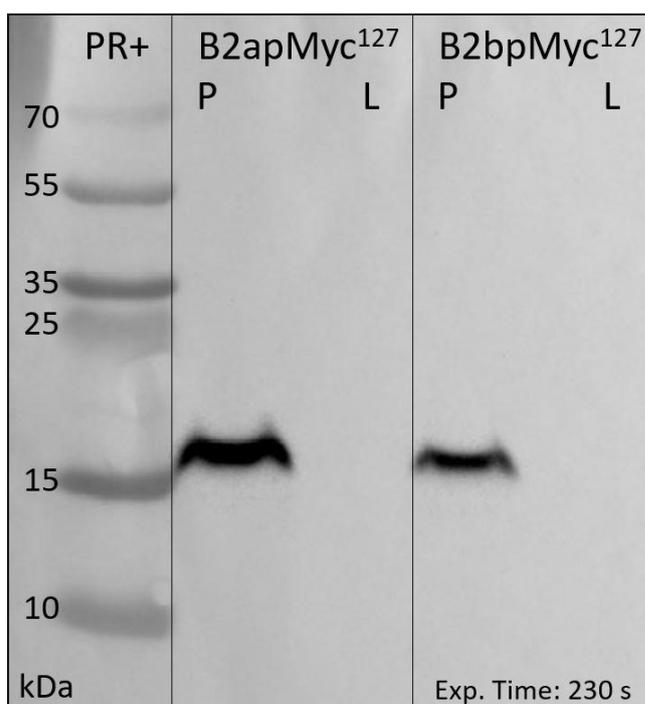


**Figure 1:** Visualization of the cloning strategy for pNCcMyc:500bp+*ehi\_127670*. **A:** pNCcMyc vector with the cloning site *KpnI-NheI-BamHI*. **B:** Insert 500bp+*ehi\_127670* with the restriction sites *KpnI* and *BamHI*. **C:** finished construct pNCcMyc:500bp+*ehi\_127670* with the insert from B inserted in the vector from A.



**Figure 2:** Diagnostic digest of the finished myc-tag construct. Eight different clones A-H were digested with the enzymes *KpnI* and *BamHI* and separated in a 1% agarose gel (80 V, 1h). All samples show the correct fragment sizes of 6 kbp for the cut *pNCcMyc* vector and 833 bp for the 500 bp 5'+*ehi\_127670* insert.

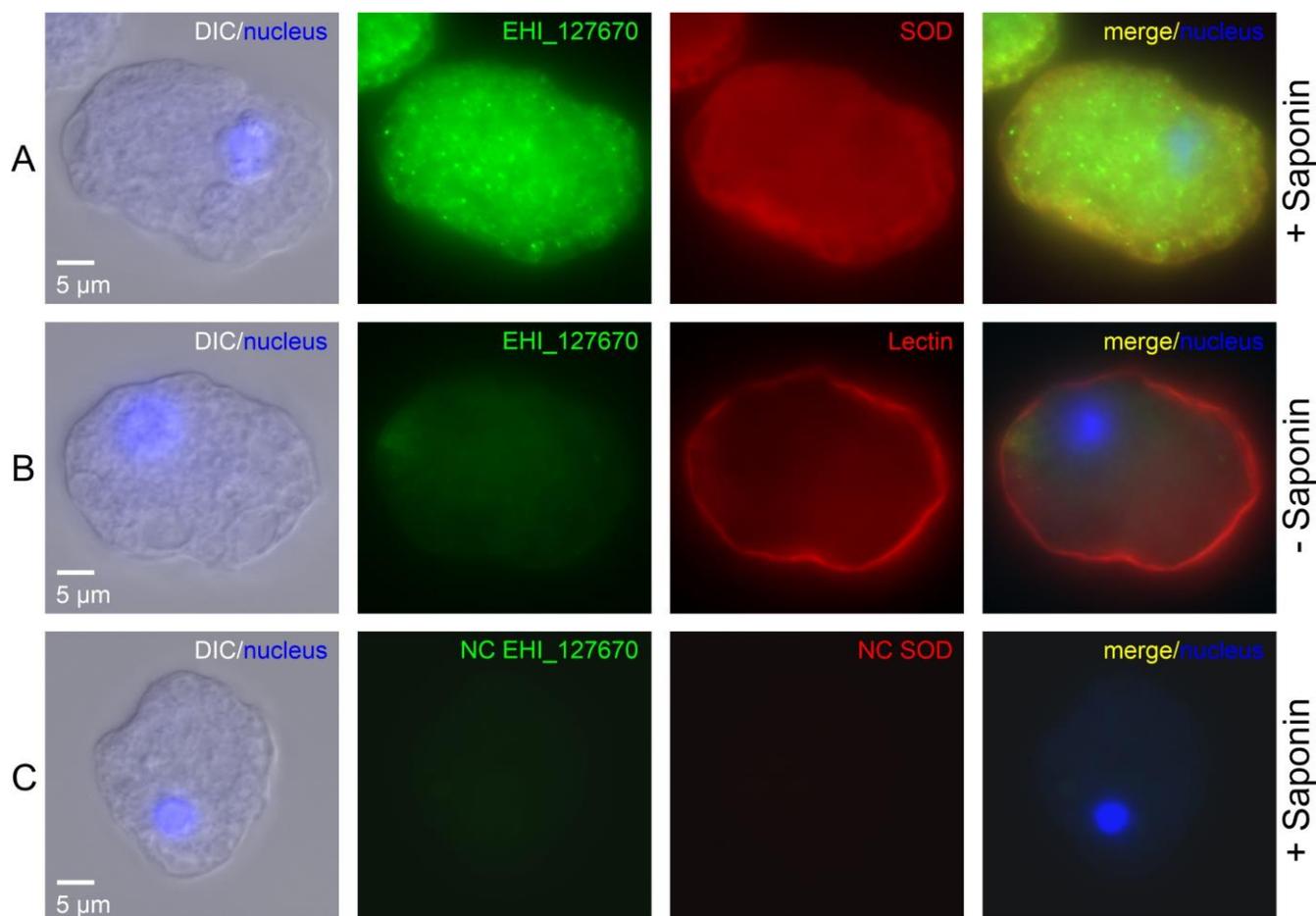
The construct was transfected into trophozoites of clone A1 and clone B2. The transfectants A1apMyc<sup>127</sup> and A1bpMyc<sup>127</sup> died during selection. B2apMyc<sup>127</sup> and B2bpMyc<sup>127</sup> achieved G418 resistance and maintained a stable cell growth. To check the expression of *ehi\_127670* in the myc-tag construct, a tricine-SDS-PAGE with cell lysates and pellets of B2apMyc<sup>127</sup> and B2bpMyc<sup>127</sup>, and subsequently a western blot for the detection of the putative myc-tagged protein was performed. The resulting blot is depicted in figure 3.



**Figure 3:** Western blot of a 13 % tricine-SDS-PAGE gel with 1:1000  $\alpha$ -c-myc primary AB and 1:2000  $\alpha$ -mouse-HRP secondary AB. For each sample B2a and B2b, lysate (L) and pellet (P) was used for SDS-PAGE. B2a and B2b are transfectants of B2 individually transfected with pNCcMyc:500bp+ehi\_127670 to express the gene product of *ehi\_127670*. *Page Ruler Prestained +* was used as a protein ladder. Exposure time was 230 s.

Both samples show a band of approx. 16 kDa in the pellet fraction, which corresponds to the calculated molecular weight of 14.5 kDa of the myc-tagged protein EHI\_127670. This confirms the hypothesis of *ehi\_127670* being a protein-coding gene and contradicts the lncRNA-thesis. The presence of the protein in the pellet fraction allows the characterization of EHI\_127670 as a non-cytosolic protein and therefore a coarse localization.

To specify the location of the protein EHI\_127670 inside the trophozoites, IFA was performed with the B2apMyc<sup>127</sup> and B2bpMyc<sup>127</sup> transfectants. As positive controls, the cytosolic protein superoxide dismutase (SOD) and the membrane-bound 170 kDa lectin were detected together with EHI\_127670. The standard fluorescence microscopy images of the stained trophozoites are depicted in figures 4 and 5.



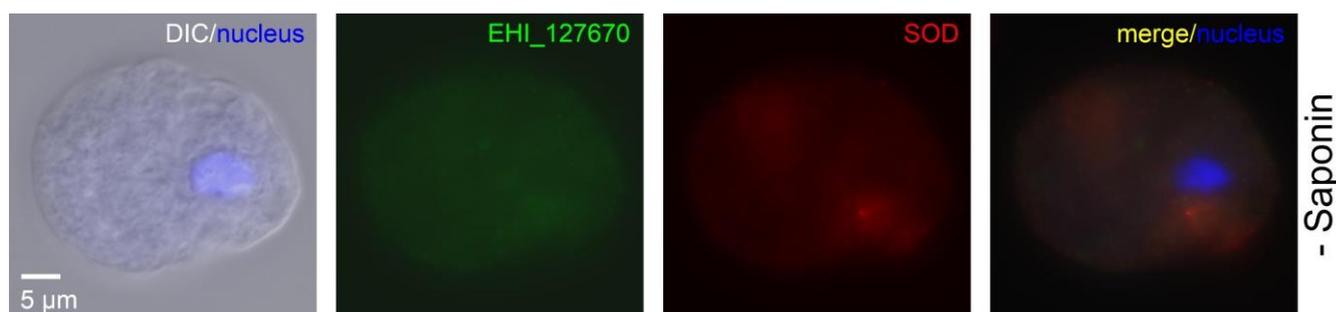
**Figure 4:** Fluorescence images of *E. histolytica* trophozoites fixated with PFA for the localization of EHI\_127670. The modi DIC/nucleus, green fluorescent, red fluorescent and green+red merge/nucleus are depicted in four columns. The samples A, B and C are depicted in three rows. Membrane permeabilization by presence (+) or non-permeabilization by absence (-) of saponin in each sample is indicated on the right edge of every row. Dilution of all primary AB: 1:100; secondary AB: 1:400. DAPI concentration: 25  $\mu$ M. Sample A: B2bpMyc<sup>127</sup> with primary AB  $\alpha$ -c-myc (mouse) and  $\alpha$ -SOD (rabbit), secondary AB  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). Sample B: B2apMyc<sup>127</sup> with primary AB  $\alpha$ -c-myc (mouse) and  $\alpha$ -Lectin (rabbit), secondary AB  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). Sample C: B2apMyc<sup>127</sup> with no primary AB, secondary AB  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red).

In sample A, a strong signal for the control SOD is visible. The fluorescence is homogeneous distributed in the cytosol, which corresponds to descriptions in literature (Dissertation G. Handal, 2010). In sample B, the fluorescence signal for the control 170 kDa lectin is visible as a red ring, which is typical for membrane-bound proteins and coincidences with results found in literature as well (Ravdin, J.I. *et al.* 1986). Compared to those controls, the signals of EHI\_127670 in the samples A and B could be used for localization. In sample A, the fluorescence signal for EHI\_127670 shows a clear signal in the cell lumen. Compared to the SOD signal, EHI\_127670 shows a different pattern of distribution.

The fluorescence is concentrated in small spots, which are evenly dispersed in the cytosol. Between the spots, a green background signal is visible, which could be provoked by unfocussed spots outside of the focal plane, or an inherent fluorescence of the cytosol. The merge of sample A shows a yellowish signal with green spots as expected. An orange ring in membrane proximity indicates an abundance of SOD or a deficiency of the myc-tagged protein in that area.

Sample B displays a very low signal for EHI\_127670, which indicates the absence of the protein on the membrane surface, since the cell was not permeabilized with saponin. A slight green signal on the left side of the cell could be caused by a marginal concentration of EHI\_127670 on the surface, or a slight membrane rupture during the fixation process and thereby entrance of staining antibodies into the cell. The negative controls in sample C indicate no unspecific reactions of the secondary antibodies with cellular proteins.

The weak green signal in Sample B in figure 4 could have been provoked by slight membrane rupture during the fixation process. To assess the quality of the fixation, membrane integrity was checked with a saponin-negative detection of EHI\_127670 and SOD. The results are depicted in figure 5.

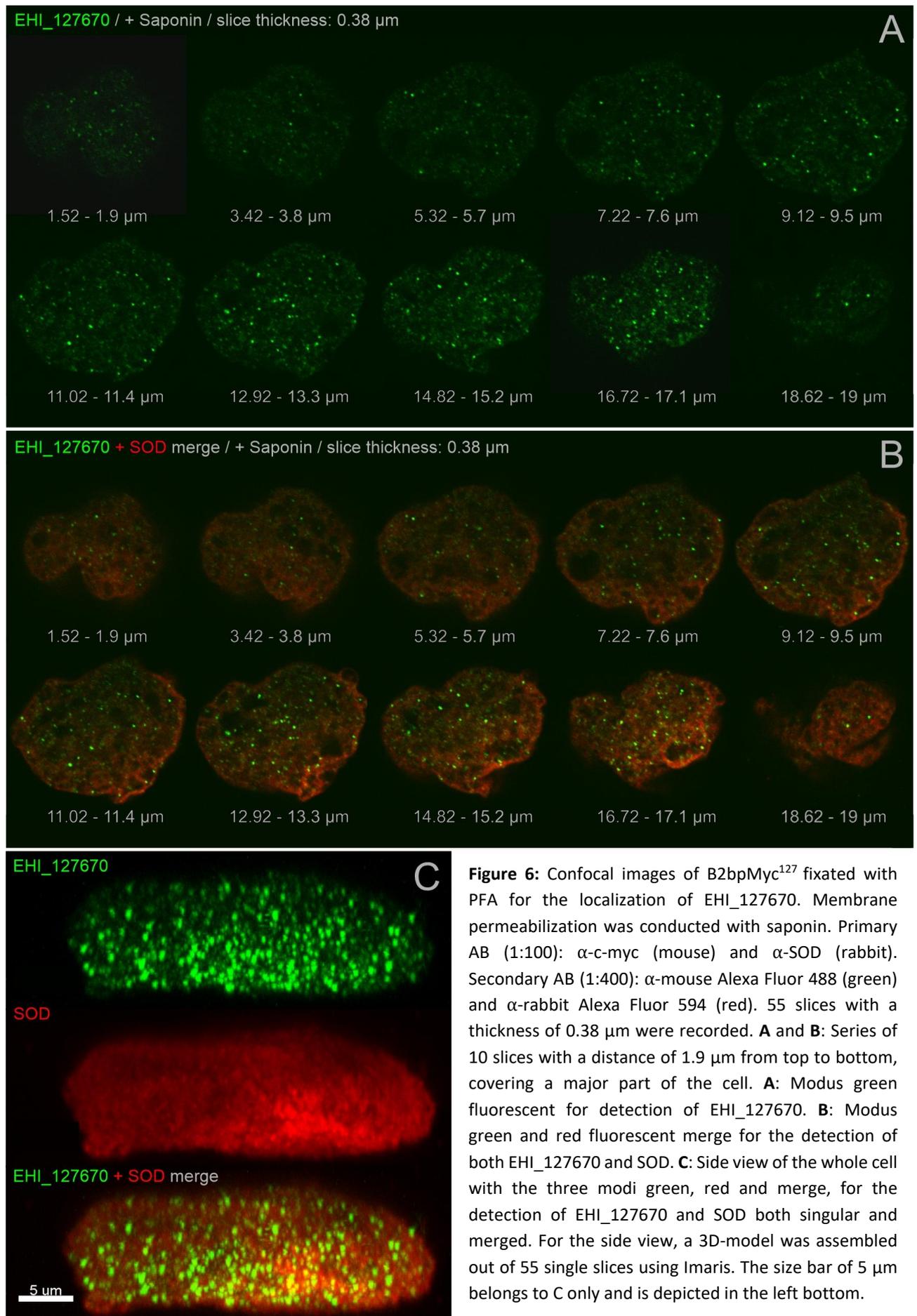


**Figure 5:** Fluorescence images of *E. histolytica* trophozoites fixated with PFA without addition of saponin for membrane integrity check and fixation process assessment. The modi DIC/nucleus, green fluorescent, red fluorescent and green+red merge/nucleus are depicted in four pictures. Non-permeabilization of the membrane by absence (-) of saponin is indicated on the right edge. Dilution of all primary AB: 1:100; secondary AB: 1:400. DAPI concentration: 25  $\mu$ M. Primary AB  $\alpha$ -c-myc (mouse) and  $\alpha$ -SOD (rabbit), secondary AB  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red).

The cytoplasmic SOD in figure 5 shows a weak signal with two slightly fluorescing spots. Because of the absence of saponin during fixation, membrane integrity should be intact and no signal for SOD visible. The weak signal indicates a slight membrane rupture during the fixation process. EHI\_127670 shows a similar signal with barely visible spots.

Because of the similarity of the signals of SOD in figure 5 and EHI\_127670 in sample B in figure 4, and the characterization of the SOD as a cytoplasmic protein, the signal of EHI\_127670 in figure 4 is most likely attributable to a faulty fixation. Since the signal of SOD in figure 5 is significantly lower than in figure 4 sample A, the signal provoked by faulty fixation can be neglected. Thus, a localization of EHI\_127670 on the membrane surface can be excluded.

For the further localization of EHI\_127670, the fluorescent spots from sample A in figure 4 were investigated. To examine the origin of the background signal between the spots, confocal fluorescence microscopy was performed. Thereby, the cell could be divided into several focal planes as slices of a defined thickness, which could be recorded separately. Background noise, potentially provoked by unfocused fluorescent spots outside of the focal plane, could thereby be reduced. Also, a 3D model of the cell could be assembled from the slices. A series of confocal pictures of a cell of B2bpMyc<sup>127</sup>, from the same sample as in figure 4 A, as well as side views of the 3D model, are depicted in figure 6.



**Figure 6:** Confocal images of B2bpMyc<sup>127</sup> fixed with PFA for the localization of EHI\_127670. Membrane permeabilization was conducted with saponin. Primary AB (1:100):  $\alpha$ -c-myc (mouse) and  $\alpha$ -SOD (rabbit). Secondary AB (1:400):  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). 55 slices with a thickness of 0.38 µm were recorded. **A** and **B**: Series of 10 slices with a distance of 1.9 µm from top to bottom, covering a major part of the cell. **A**: Modus green fluorescent for detection of EHI\_127670. **B**: Modus green and red fluorescent merge for the detection of both EHI\_127670 and SOD. **C**: Side view of the whole cell with the three modi green, red and merge, for the detection of EHI\_127670 and SOD both singular and merged. For the side view, a 3D-model was assembled out of 55 single slices using Imaris. The size bar of 5 µm belongs to C only and is depicted in the left bottom.

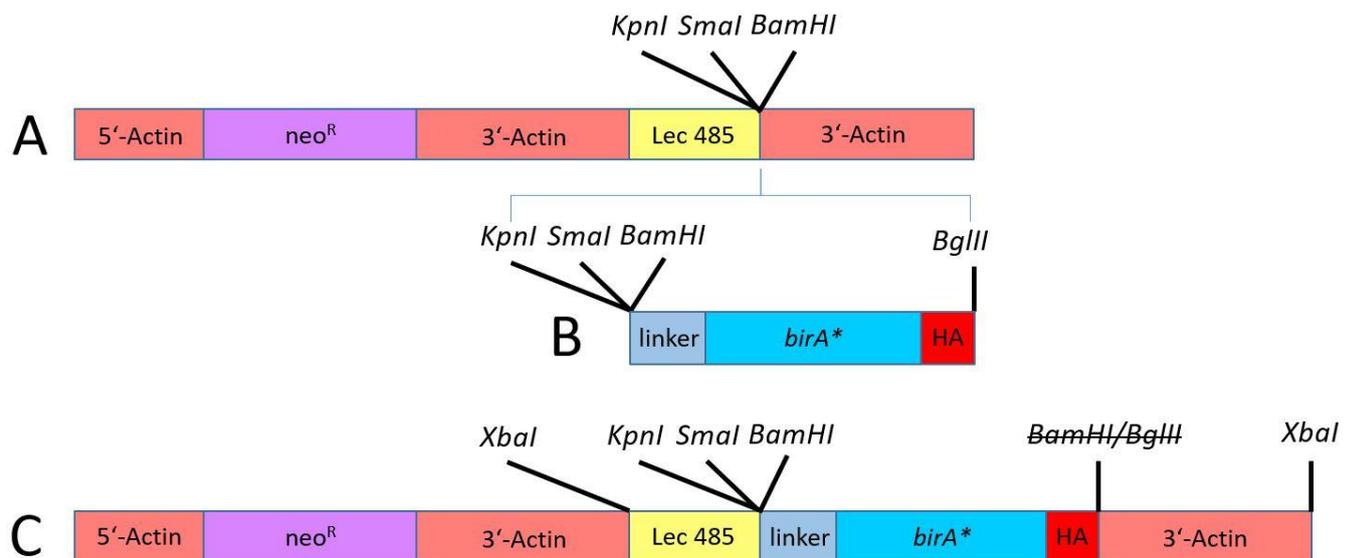
In figure 6 A, many fluorescent round spots of EHI\_127670 are visible in each slice, clearly standing out from a relatively dark background. The distribution in each slice appears to be random. The abundance of the round spots per slice increases with decreasing height of the slice relatively to the ground. In B, the background signal of EHI\_127670 appears to be very low, compared to the cytosolic SOD. Also, the SOD shows a higher abundance in membrane proximity, whereas the spots of EHI\_127670 are evenly distributed. Besides the red signal, circular shadows without a red or green signal are visible.

In C, the side views of the cell reveal a relatively even distribution of green fluorescent spots of EHI\_127670 with an abundance in ground plate proximity. The spots have a spherical shape with a diameter of approx. 500 nm. As already observed in A and B, the background signal between the green fluorescent spots is very low, compared to the cytosolic SOD.

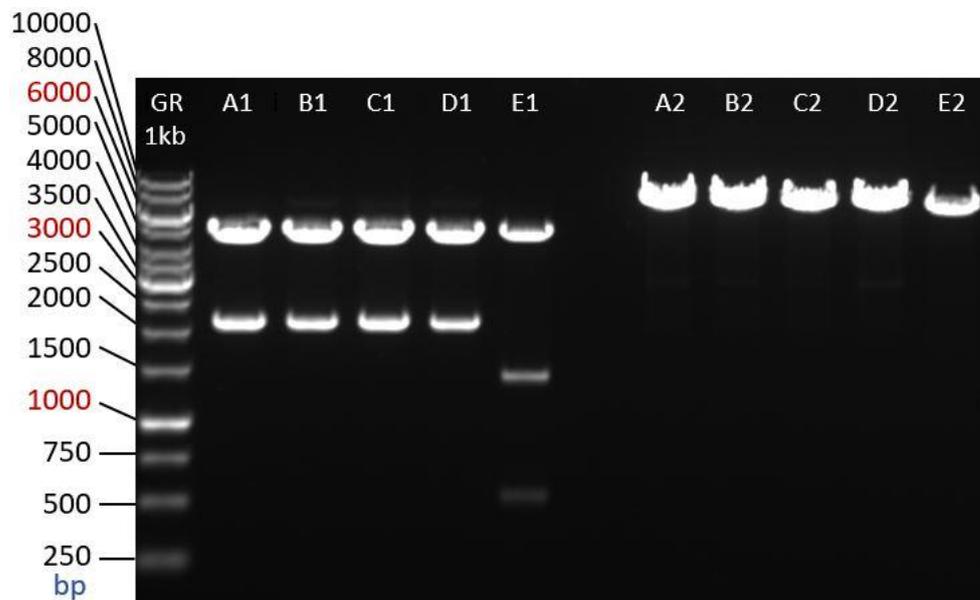
The results in figure 6 indicate an exclusive localization of EHI\_127670 in small, spherical, granule-like cellular compartments, that are evenly distributed in the cytosol, with a higher abundance in ground proximity.

## 5.2 BiOLD

After the localization of EHI\_127670 in undefined granulae, potential reaction partners of the protein were in the focus for further characterization. For the discovery of reaction partners, the BiOLD method was implemented for *E. histolytica*. The pNCBID construct, containing the *birA\** sequence with linker and HA-tag, was assembled as described in section 4.1.3. The cloning strategy is visualized in figure 7. Five clones were picked and analysed. The diagnostic digest of the finished constructs, depicted in figure 8, shows the correct fragment sizes for the clones A-D. Clone B was sequenced and propagated for transfection.

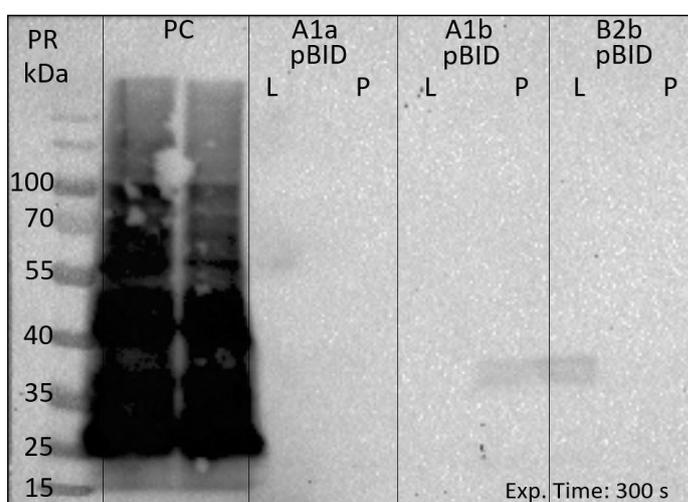


**Figure 7:** Visualization of the cloning strategy for pNCBID. **A:** pNC vector with the cloning site *KpnI-SmaI-BamHI*. **B:** Insert BirA\* with the restriction sites *KpnI-SmaI-BamHI* and *BglIII*. **C:** finished construct pNCBID with the insert BirA\* inserted in pNC. The crossing out of restriction sites indicates their inactivation by complementary pairing. The restriction site *XbaI* is always present but only depicted in C.



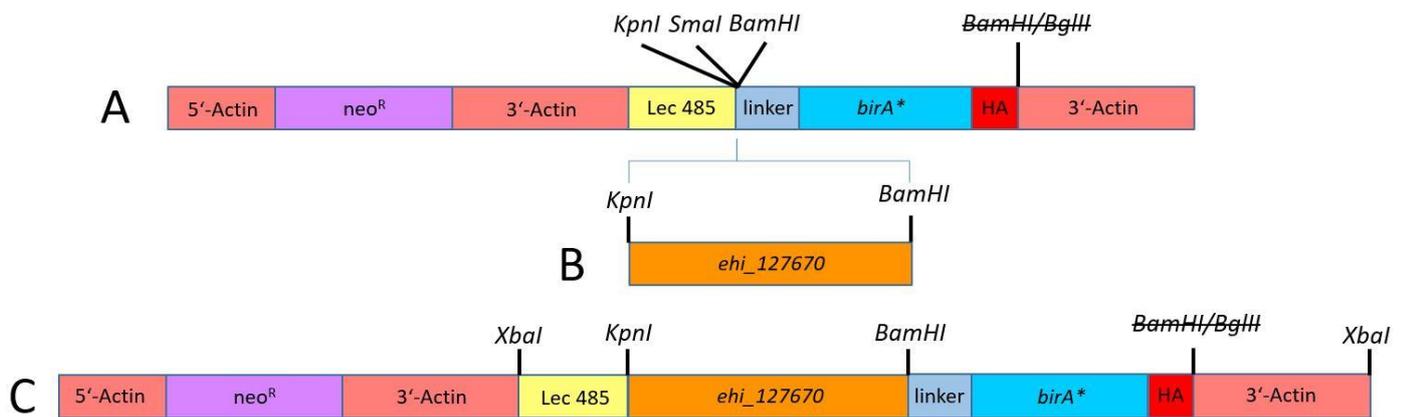
**Figure 8:** diagnostic digest of the BioID construct without fusion with a gene of interest. Five different clones A-E were digested in two reaction mixtures. A1-E1 were digested with the enzyme *XbaI*, A2-E2 were digested with the enzymes *KpnI* and *BglIII*. Incubation time was 1h. The fragments were separated in a 1% agarose gel (90 V, 1h). Samples A1-D1 show the correct fragment sizes of 5 kbp for the *pNC* fragment and 2191 bp for the cut out fragment consisting of 3'Ac (530 bp), *Lec 485* (485 bp) and the BioID sequence including linker and HA-tag (1176 bp). Samples A2-D2 show the correct fragment size of 7191 bp, which is the size of the complete, cut-open construct.

First, to examine the general expression of the bacteria-derived BirA\* in *E. histolytica*, pNCBID with no insert was transfected into trophozoites of clone A1 and clone B2. The transfectant B2apBID died during selection. B2bpBID, A1apBID and A1bpBID achieved G418 resistance and maintained a stable cell growth. To check the expression of BirA\*, a glycine-SDS-PAGE with cell lysates and pellets of the transfectants, and subsequently a western blot for the detection of the HA-tagged BirA\* was performed. A positive control (PC), containing HA-tagged proteins from *P. falciparum* cell lysate, was included. The PC sample was provided by Jan Stacker (AG Spielmann, BNITM). The resulting blot is depicted in figure 9.



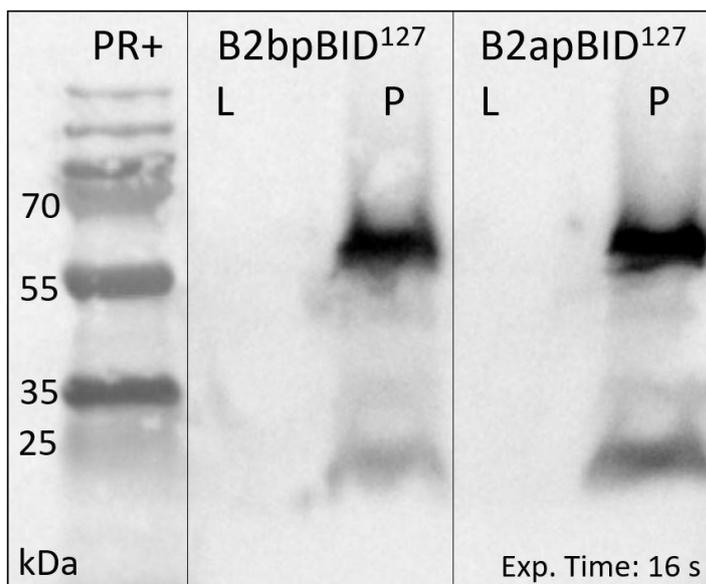
**Figure 9:** Western blot of a 10 % glycine-SDS-PAGE gel with 1:1000  $\alpha$ -HA primary AB and 1:5000  $\alpha$ -rat-HRP secondary AB. For each sample except PC, lysate (L) and pellet (P) was used for SDS-PAGE. A1apBID, A1bpBID and B2bpBID are clones of A1 and B2 transfected with pNCBID for the expression of BirA\*. The positive control PC is cell lysate of *P. falciparum* containing HA-tagged proteins. *Page Ruler Prestained* was used as a protein ladder. Exposure time was 300 s.

The samples B2bpBID, A1apBID and A1bpBID show a negative result while the positive control is strongly positive, thus the protein could not be detected. It was hypothesized that cellular mechanisms prevent the expression of the bacterial derived protein BirA\*. To circumvent those potential mechanisms, the cloning strategy was proceeded and *ehi\_127670* cloned into pNCBID 5'-terminal of the *birA\** sequence. The aim was to trick the cellular recognition system into expressing the fusion protein. The pNCBID:*ehi\_127670* construct was assembled as described in section 4.1.3. The cloning strategy is visualized in figure 10.



**Figure 10:** Visualization of cloning strategy for pNCBID:*ehi\_127670*. **A:** pNCBID vector with the cloning site *KpnI-Smal-BamHI*. **B:** Insert *ehi\_127670* with the restriction sites *KpnI* and *BamHI*. **C:** finished construct pNCBID:*ehi\_127670* with the insert *ehi\_127670* inserted in pNC. The crossing out of restriction sites indicates their inactivation by complementary pairing. The restriction site *XbaI* is always present but only depicted in C.

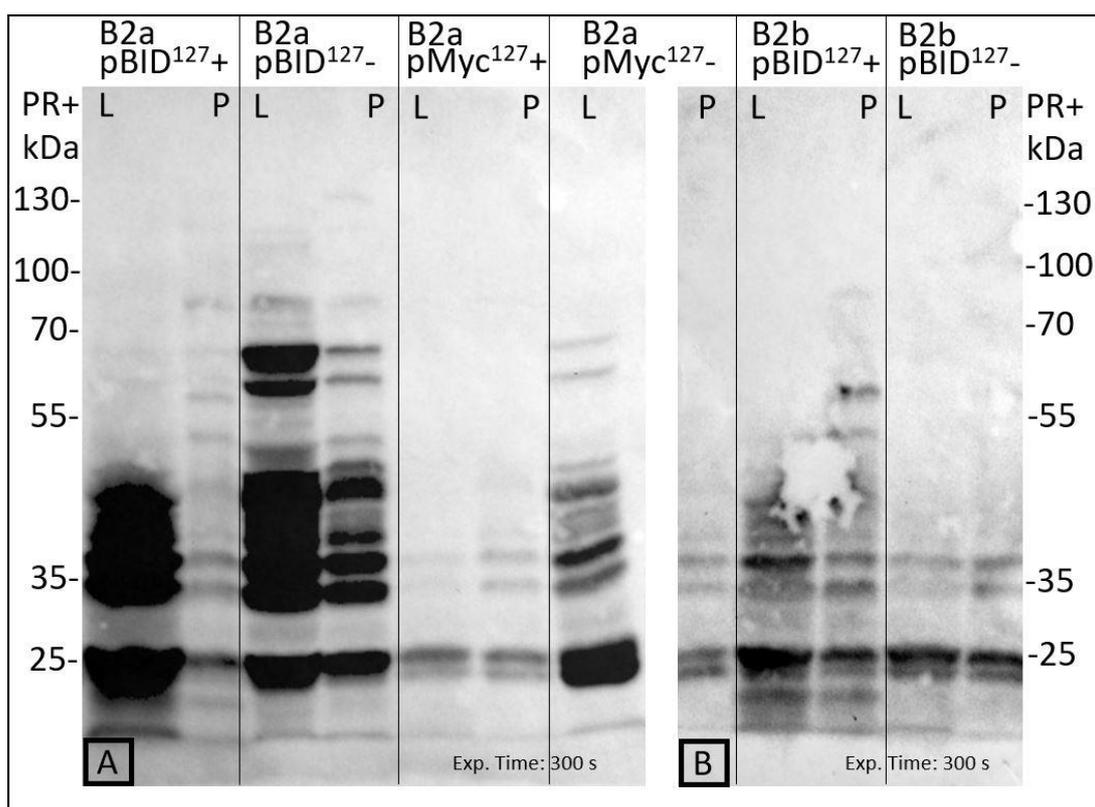
The pNCBID:*ehi\_127670* construct was transfected into trophozoites of clone A1 and clone B2. The transfectants A1apBID<sup>127</sup> and A1bpBID<sup>127</sup> died during selection. B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> achieved G418 resistance and maintained a stable cell growth. To check the expression of EHI\_127670:BirA\*, a glycine-SDS-PAGE with cell lysates and pellets of the transfectants, and subsequently a western blot for the detection of the HA-tagged fusion protein was performed. The resulting blot is depicted in figure 11.



**Figure 11:** Western blot of a 13 % tricine-SDS-PAGE gel with 1:1000  $\alpha$ -HA primary AB and 1:2000  $\alpha$ -rat-HRP secondary AB. For each sample, lysate (L) and pellet (P) was used for SDS-PAGE. B2bpBID<sup>127</sup> and B2apBID<sup>127</sup> are clones of B2 transfected with the pNCBID:*ehi\_127670* construct, resulting in the expression of EHI\_127670:BirA\*. Page Ruler Prestained + was used as a protein ladder. Exposure time was 16 s.

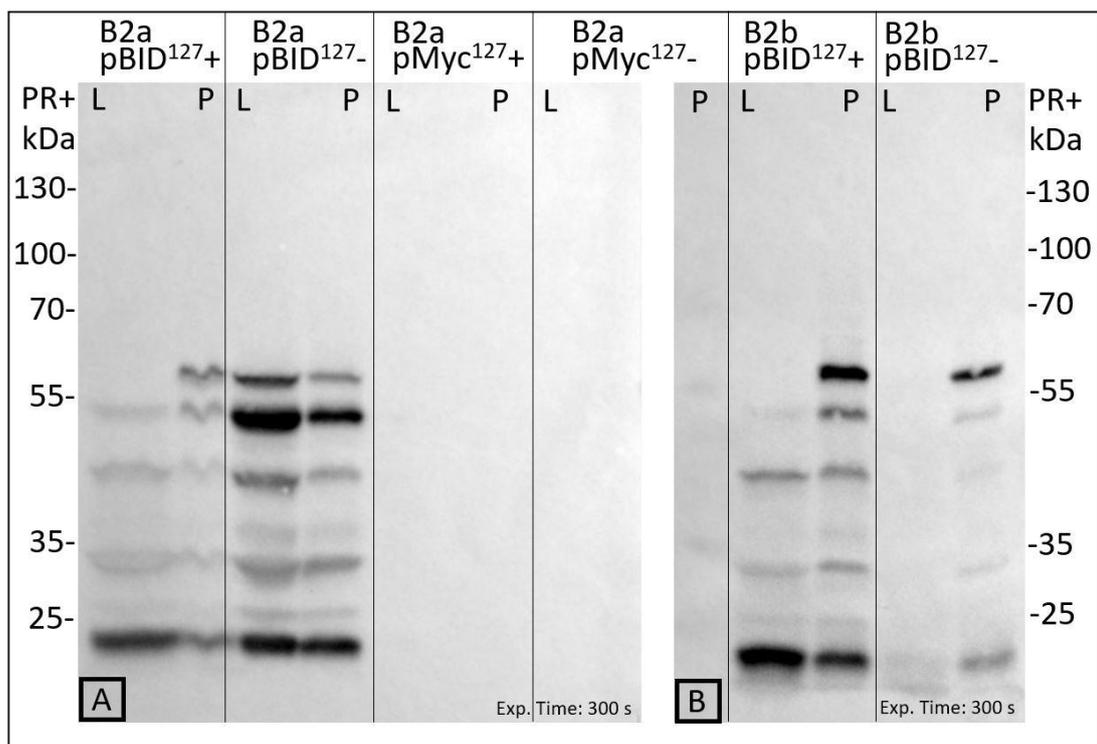
Both samples show a band of approx. 58 kDa in the pellet fraction, which corresponds to the calculated molecular weight of 54.8 kDa of EHI\_127670:BioID\*, meaning the fusion protein was expressed successfully. The presence of the protein in the pellet fraction coarsely indicates the correct localization of the fusion protein in the native locus of EHI\_127670.

After the successful expression of the EHI\_127670:BioID fusion protein, the biotinylation of potential reaction partners of EHI\_127670 could be investigated. The transfectants B2apBioID<sup>127</sup> and B2bpBioID<sup>127</sup>, as well as B2apBioID<sup>127</sup>, which served as a control, were incubated with and without 85  $\mu$ M biotin for 24h. The trophozoites were harvested and glycine-SDS-PAGE was performed in duplicate containing identical samples of the cell lysates and pellet fractions of the transfectants. Subsequently, a western blot and a streptavidin blot was performed per gel duplicate. The resulting streptavidin blots are depicted in figure 12, the corresponding western blots for the detection of the fusion protein in figure 13.



**Figure 12:** Streptavidin blots of two 10% glycine-SDS-PAGE gels A and B with 1:1000 streptavidin-HRP. For each sample, lysate (L) and pellet (P) was used for SDS-PAGE. A "+" behind the sample's name indicates the incubation with 85  $\mu$ M biotin for 24 h. A "-" behind the sample's name indicates a permanent concentration of 0.97  $\mu$ M biotin in the medium. B2apBioID<sup>127</sup> and B2bpBioID<sup>127</sup> are clones of B2 transfected with the ehi\_127670:BioID construct. B2apMyc<sup>127</sup> is a B2 clone transfected with the ehi\_127670:myc tag construct and serves as control. *Page Ruler Prestained +* was used as a protein ladder. Exposure time for A and B was 300 s.

All samples in figure 12 show a similar biotinylation pattern, varying in intensity, with no general correlation with an incubation with biotin. B2apBID<sup>127</sup> shows a stronger biotinylation for the sample not incubated with biotin than the sample that was incubated, for B2apBID<sup>127</sup> it is the opposite. The control B2apMyc<sup>127-</sup>, that neither contained BirA\*, nor was incubated with biotin, displays a biotinylation pattern similar to B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>, for bands under 70 kDa and with a smaller intensity. Bands over 70 kDa exclusively occur in sample B2apBID<sup>127</sup>, which might be attributed to B2apBID<sup>127</sup> having the overall highest band intensity. B2bpBID<sup>127-</sup> and B2apMyc<sup>127+</sup> show a similar pattern of bands with low intensity, which, considering the signals of B2bpBID<sup>127+</sup> and B2apMyc<sup>127-</sup>, does not correlate with the presence of BirA\* or biotin. In figure 13, the western blots corresponding to the streptavidin blots in figure 12 are depicted.

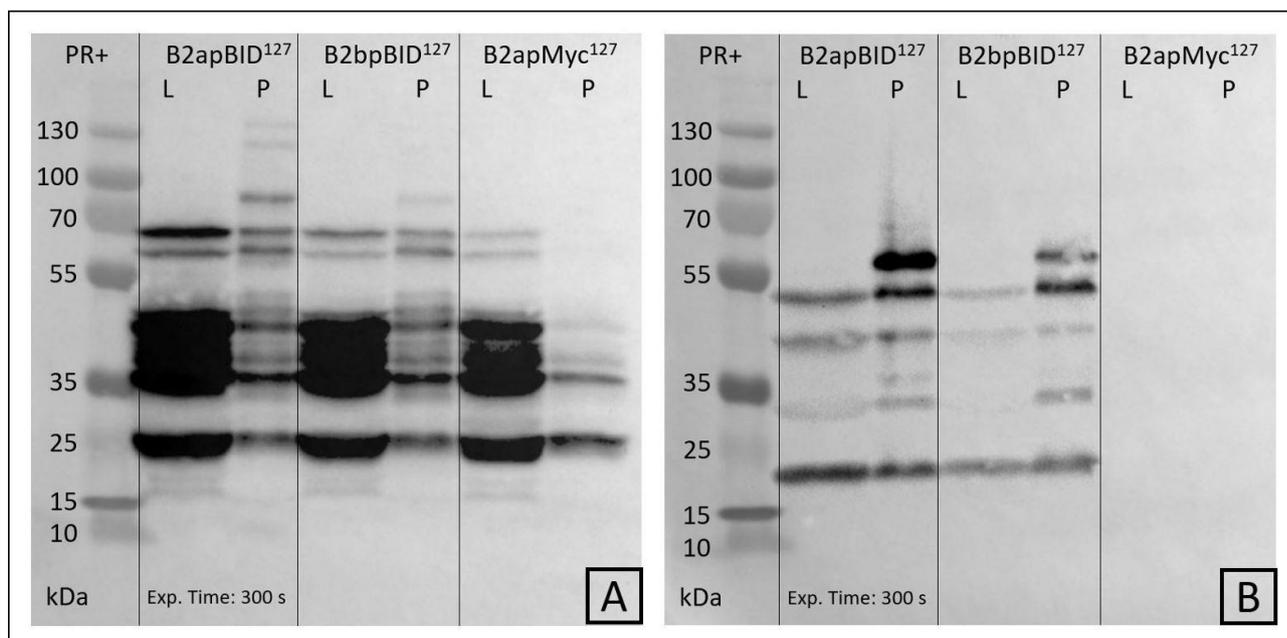


**Figure 13:** Western blots with 1:1000  $\alpha$ -HA primary AB and 1:2000  $\alpha$ -rat-HRP secondary AB of two 10% glycine-SDS-PAGE gels A and B that were produced identical to A and B in figure 12. For each sample, lysate (L) and pellet (P) was used for SDS-PAGE. A "+" behind the sample's name indicates the incubation with 85  $\mu$ M biotin for 24 h. A "-" behind the sample's name indicates a permanent concentration of 0.97  $\mu$ M biotin in the medium. B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> are clones of B2 transfected with the *ehi\_127670*:BioID construct. B2apMyc<sup>127</sup> is a B2 clone transfected with the *ehi\_127670*:myc tag construct and serves as control. *Page Ruler Prestained +* was used as a protein ladder. Exposure time for A and B was 300 s.

As expected, EHI\_127670:BirA\* could not be detected in B2apMyc<sup>127</sup>. In B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> the HA-tagged protein could be detected at approx. 58 kDa in the pellet, in B2apBID<sup>127</sup> - as well in the lysate. Except for B2bpBID<sup>127</sup> - L, all BirA\*-positive samples show a similar pattern of bands of a molecular weight smaller than 58 kDa, which occurs typically when a tagged protein is degraded by proteases. The overall intensity of the bands in figure 13 correlates with the intensity of the bands in figure 12.

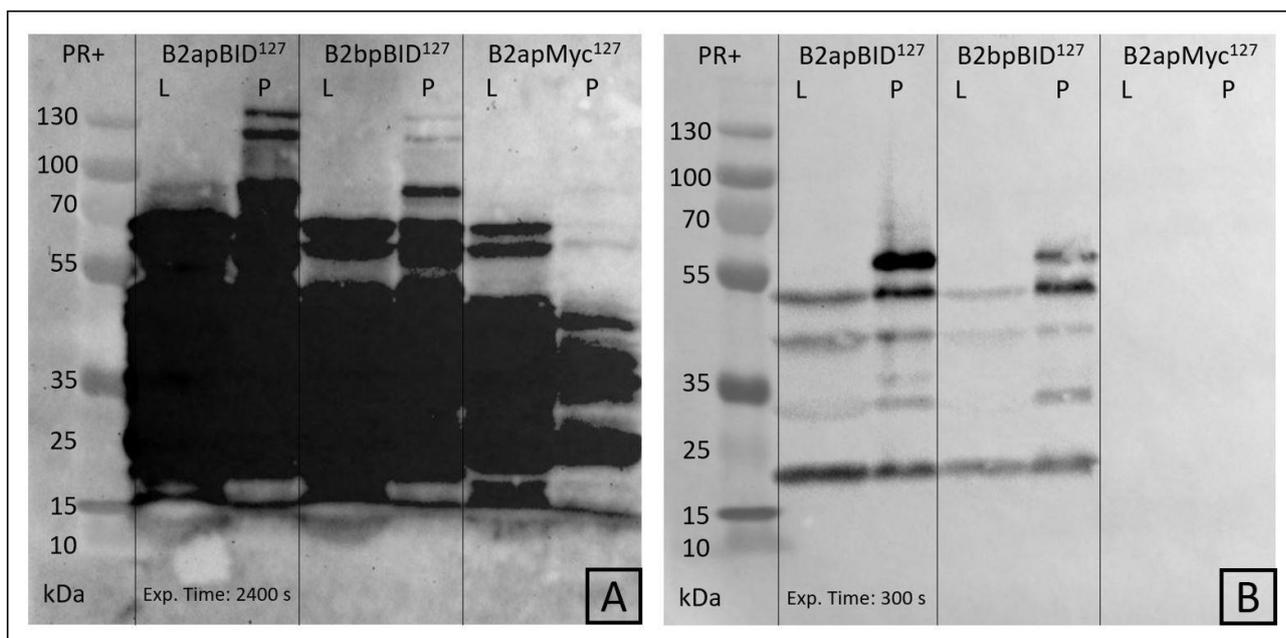
After considerations that are discussed in section 6.2, the unexpected biotinylation pattern in figure 12 was presumably attributed to a medium-inherent biotin concentration sufficient for biotinylation, an unknown catalyst for biotinylation and the varying amount of protein contained in each sample. Because of the strong biotinylation signal of samples that were not incubated with additional biotin in figure 12, the medium-inherent concentration of 0.91  $\mu$ M as a permanent application was considered sufficient for biotinylation. Because of the BirA\*-independent biotinylation pattern in the sample B2apMyc<sup>127</sup>, which indicates a natural biotinylation in *E. histolytica*, this sample was used as a control to subtract natural biotinylation from BirA\*-dependent biotinylation.

To equalize the influence of the total protein amount to the biotinylation signal to some extent, trophozoites of B2apBID<sup>127</sup>, B2bpBID<sup>127</sup> and B2apMyc<sup>127</sup> were count after harvest and split to an equal amount. For Glycine-SDS-PAGE, cell lysates and pellets of the transfectants were loaded into two gels in duplicate. After separation, a western blot and a streptavidin blot was performed with the gel duplicates. The resulting streptavidin blot is depicted in figure 14 A, the corresponding western blot in figure 14 B.



**Figure 14:** Streptavidin blot (A) with 1:1000 streptavidin-HRP and western blot (B) with 1:1000  $\alpha$ -HA primary AB and 1:2000  $\alpha$ -rat-HRP secondary AB of two identical 10% glycine-SDS-PAGE gels. For each sample, lysate (L) and pellet (P) was used for SDS-PAGE. B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> are clones of B2 transfected with the pNCBID:*ehi\_127670* construct. B2apMyc<sup>127</sup> is a B2 clone transfected with the *ehi\_127670*:myc tag construct and serves as control. Biotinylation was performed permanently with 0.97  $\mu$ M biotin. *Page Ruler Prestained +* was used as a protein ladder. Exposure time for A and B was 300 s.

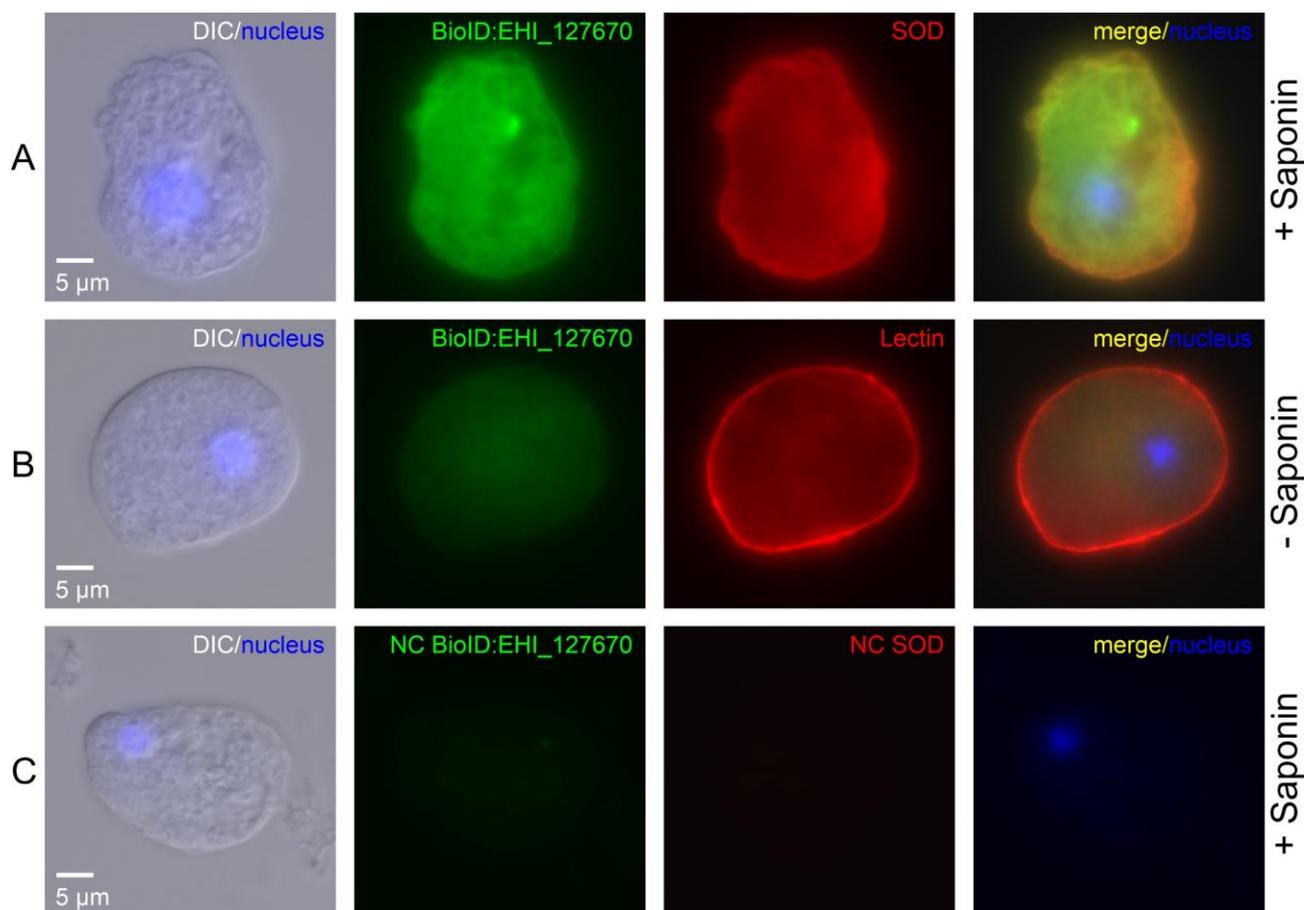
As expected, the  $\alpha$ -HA western blot in figure 14 B shows no signal for B2apMyc<sup>127</sup>. In the samples B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>, a band at approx. 58 kDa in the pellet fraction can be assigned to the EHI\_127670:BirA\* fusion protein. Bands of smaller molecular weight in pellet and lysate fraction may be attributed to enzymatic degradation of the fusion protein. In figure 14 A, all samples show a similar biotinylation pattern for the lysate fraction, with a slight variation in band intensity. Since all bands that occur in B2apMyc<sup>127</sup> are non-BirA\* specific, those signals are subtracted from B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>, leaving no specific biotinylation in the lysate fractions. The pellet fractions of B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> have a similar biotinylation pattern that differs significantly from B2apMyc<sup>127</sup>. Five bands of approx. 48, 50, 58, 60 and 80 kDa, that do not occur in B2apMyc<sup>127</sup>, are visible in both B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>. Two bands of approx. 128 and 132 kDa are visible in B2apBID<sup>127</sup>, but not in B2bpBID<sup>127</sup> or B2apMyc<sup>127</sup>. Those seven bands appear to be BioID-specific in this blot. Since the overall intensity of the bands differs from sample to sample, the signal of some of the seven bands might be too weak to be visible. To amplify those weak signals, exposure time was set to 2400 s for the streptavidin blot. The results are depicted in figure 15.



**Figure 15:** Streptavidin blot (A) with 1:1000 streptavidin-HRP and western blot (B) with 1:1000  $\alpha$ -HA primary AB and 1:2000  $\alpha$ -rat-HRP secondary AB of two identical 10% glycine-SDS-PAGE gels. For each sample, lysate (L) and pellet (P) was used for SDS-PAGE. B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> are clones of B2 transfected with the pNCBID:*ehi\_127670* construct. B2apMyc<sup>127</sup> is a B2 clone transfected with the pNCcMyc:*ehi\_127670* construct and serves as control. Biotinylation was performed permanently with 0.97  $\mu$ M biotin. *Page Ruler Prestained +* was used as a protein ladder. Except for the exposure time of A, both blots A and B are identical to A and B in figure 14. Exposure time for A was 2400 s, for B 300 s.

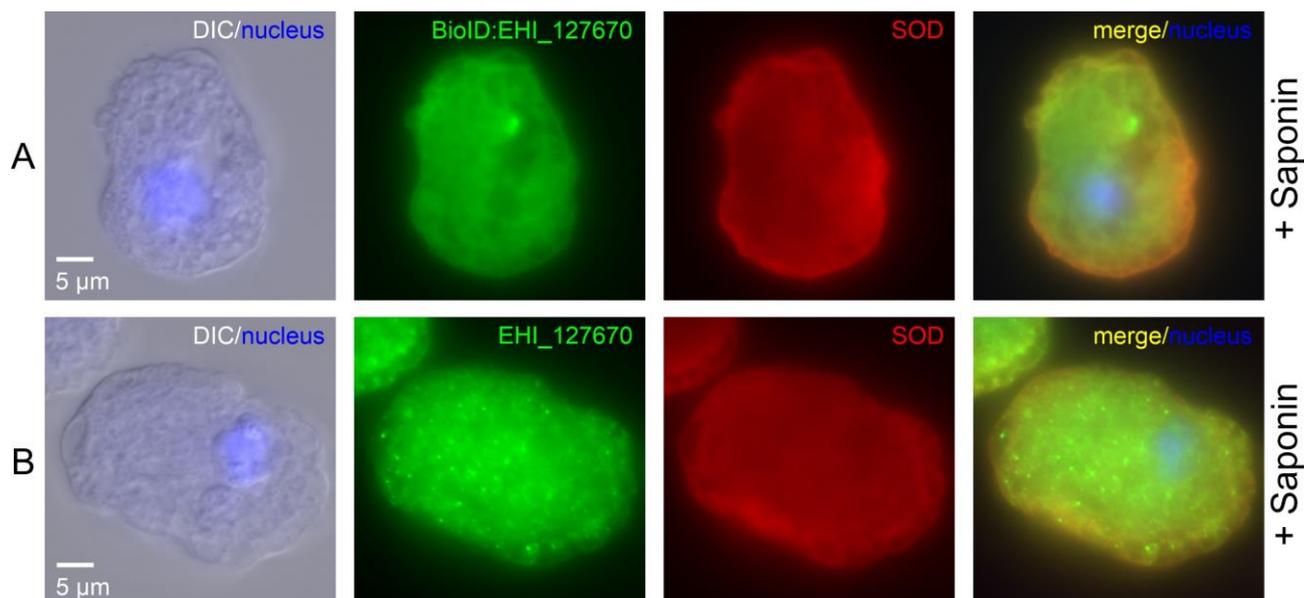
In figure 15 A, two bands at approx. 128 and 132 kDa, that were not detected in figure 14 A, are visible in the pellet fraction of B2bpBID<sup>127</sup>. The biotinylation patterns of B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> are therefore identical, except for B2apBID<sup>127</sup> showing a higher intensity. None of the seven bands from figure 14 A with a presumably specific biotinylation could be definitely detected in the pellet fraction of B2apMyc<sup>127</sup> in figure 15 A. This indicates the presence of at least seven potential candidates for reaction partners of EHI\_127670 with the molecular weights of approx. 48, 50, 58, 60, 80, 128 and 132 kDa.

The correct localization of the EHI\_127670:BirA\* fusion protein in the natural locus of EHI\_127670 is crucial for the correct targeting of the natural interaction partners. IFA was performed with trophozoites of B2apBID<sup>127</sup> and B2bpBID<sup>127</sup> to localize EHI\_127670:BirA\* and compare it to the localization results of EHI\_127670 in section 5.1. As positive controls, the cytosolic protein SOD and the membrane-bound 170 kDa lectin were detected together with EHI\_127670:BirA\*. The images of the stained trophozoites are depicted in figure 16.



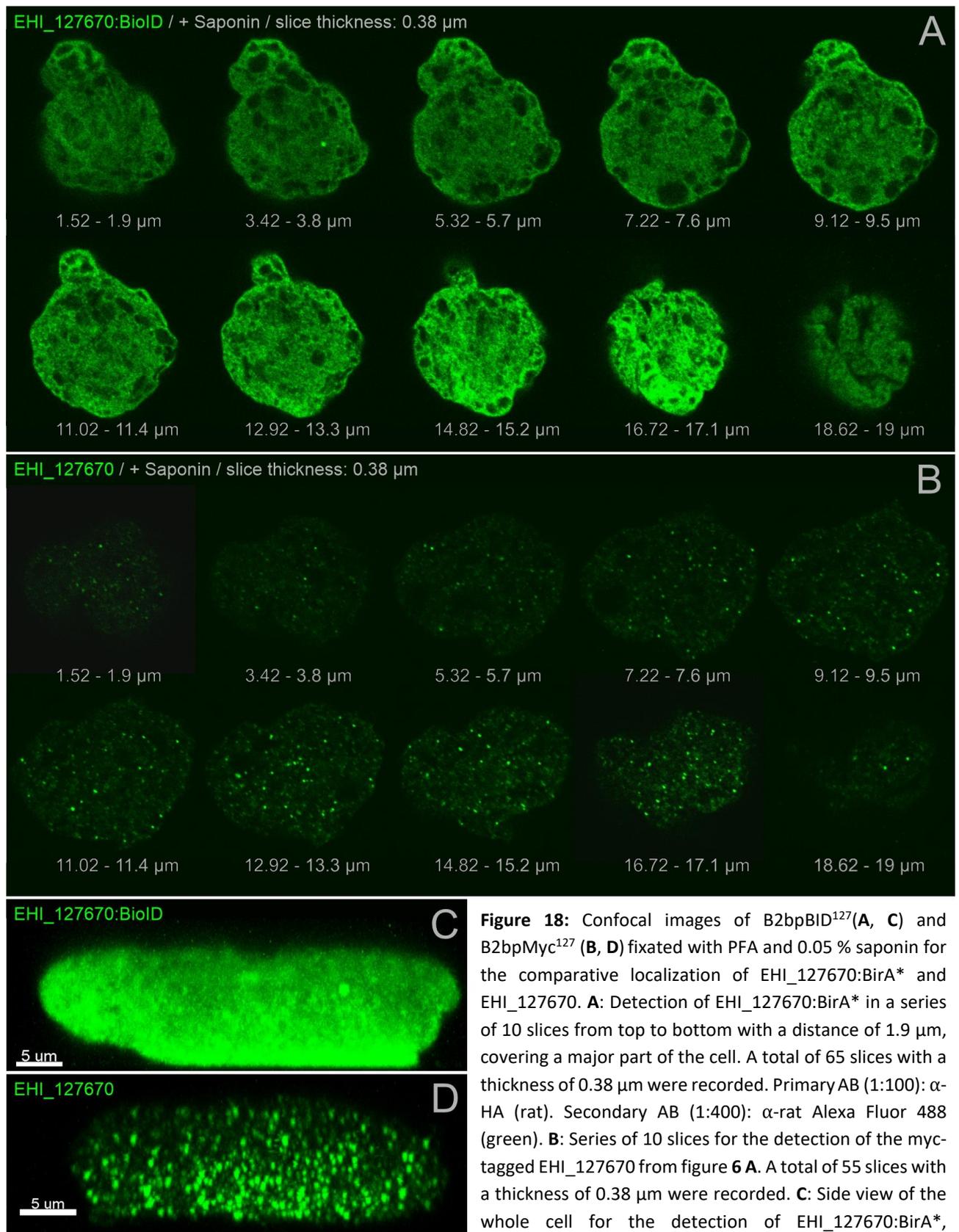
**Figure 16:** Fluorescence images of *E. histolytica* trophozoites fixed with PFA for the localization of the EHI\_127670:BiRA\* fusion protein. The modi DIC/nucleus, green fluorescent, red fluorescent and green+red merge/nucleus are depicted in four columns. The samples A, B and C are depicted in three rows. Membrane permeabilization by presence (+) or non-permeabilization by absence (-) of saponin in each sample is indicated on the right edge of every row. Dilution of all primary AB: 1:100; secondary AB: 1:400. DAPI concentration: 25  $\mu$ M. Sample A: B2bpBID<sup>127</sup> with primary AB  $\alpha$ -HA (rat) and  $\alpha$ -SOD (rabbit), secondary AB  $\alpha$ -rat Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). Sample B: B2apBID<sup>127</sup> with primary AB  $\alpha$ -HA (rat) and  $\alpha$ -170 kDa lectin (rabbit), secondary AB  $\alpha$ -rat Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). Sample C: B2apBID<sup>127</sup> with no primary AB, secondary AB  $\alpha$ -rat Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red)

The signals of 170 kDa lectin and SOD in figure 16 correspond to those detected in figure 4 in section 5.1. In sample B, no signal is detectable for the EHI\_127670:BiRA\* fusion protein, implying the absence of the protein on the cell surface. The negative control C shows no signal for unspecific binding of secondary antibodies to cellular proteins. In sample A, the fluorescence signal for EHI\_127670:BiRA\* resembles the signal of the SOD, and therefore indicates a homogeneous distribution in the cytosol. A single bright green spot is visible. The merge of sample A shows a yellowish cell lumen with an orange ring in membrane proximity. To compare the signals of EHI\_127670:BiRA\* and the myc-tagged EHI\_127670, sample A from figure 4 and sample A from figure 16 are depicted in figure 17.



**Figure 17:** Fluorescence images of *E. histolytica* trophozoites fixated with PFA for the comparative localization of EHI\_127670:BioID\* and EHI\_127670 from figures 4 and 16. The modi DIC/nucleus, green fluorescent, red fluorescent and green+red merge/nucleus are depicted in four columns. Both samples were permeabilized with saponin. Dilution of all primary AB: 1:100; secondary AB: 1:400. DAPI concentration: 25  $\mu$ M. Sample A: B2bpBioID<sup>127</sup> with primary AB  $\alpha$ -HA (rat) and  $\alpha$ -SOD (rabbit), secondary AB  $\alpha$ -rat Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red). Sample B: B2bpMyc<sup>127</sup> with primary AB  $\alpha$ -c-myc (mouse) and  $\alpha$ -SOD (rabbit), secondary AB  $\alpha$ -mouse Alexa Fluor 488 (green) and  $\alpha$ -rabbit Alexa Fluor 594 (red).

The signal of EHI\_127670:BioID\* differs significantly from the signal of EHI\_127670. No fluorescent spots are visible, instead the signal is distributed evenly in the cytosol, with exception of a single spot. This indicates the localization of EHI\_127670:BioID\* in the cytosol, which contradicts its presence in the pellet fraction in figure 11. This makes further investigations necessary. To exceed the limited resolution of standard fluorescence microscopy, confocal fluorescence microscopy was performed. The results are depicted in figure 18. For comparison with EHI\_127670, results from figure 6 are depicted as well.



**Figure 18:** Confocal images of B2bpMyc<sup>127</sup> (**A, C**) and B2bpMyc<sup>127</sup> (**B, D**) fixated with PFA and 0.05 % saponin for the comparative localization of EHI\_127670:BioID\* and EHI\_127670. **A:** Detection of EHI\_127670:BioID\* in a series of 10 slices from top to bottom with a distance of 1.9  $\mu\text{m}$ , covering a major part of the cell. A total of 65 slices with a thickness of 0.38  $\mu\text{m}$  were recorded. Primary AB (1:100):  $\alpha$ -HA (rat). Secondary AB (1:400):  $\alpha$ -rat Alexa Fluor 488 (green). **B:** Series of 10 slices for the detection of the myc-tagged EHI\_127670 from figure 6 **A**. A total of 55 slices with a thickness of 0.38  $\mu\text{m}$  were recorded. **C:** Side view of the whole cell for the detection of EHI\_127670:BioID\*, modelled by the assembly of 65 single slices. **D:** Side view of the whole cell for the detection of EHI\_127670 from figure 6, modelled by assembly of 55 single slices. Imaris was used for modelling. The size bars of 5  $\mu\text{m}$  belong to C and D only.

Figure 18 A shows a strong signal for EHI\_127670:BirA\*, distributed homogeneously in the cytosol and interrupted by dark shadows. The intensity of the signal increases with decreasing height of the slice relatively to the ground. In contrast to B, there is no pattern of spots visible, only one singular bright spot in slice 3.42 – 3.8  $\mu\text{m}$ . The overall signal of the cell area compared to the signal outside of the cell is much higher in A compared to B. The side view of the cell in C reveals a strong signal of EHI\_127670:BirA\* in the cytoplasm, with an increased intensity in ground proximity. A spot-like pattern like in D is not identifiable. Overall, the fluorescence signal of EHI\_127670:BirA\* in A and C does not resemble the signal of EHI\_127670 in B and D, since no characteristic spots over a dark background are identifiable. Therefore, a colocalization of the EHI\_127670:BirA\* fusion protein and EHI\_127670 cannot be confirmed.

## 6 Discussion

### 6.1 Myc-tag Localization

The aim of the method was to accomplish the expression of the myc-tagged *ehi\_127670* in a way as close to the natural expression of *ehi\_127670* as possible. However, since the gene was located on an artificial construct and a 1.2 kDa myc-tag was fused to the 13.2 kDa big EHI\_127670, the expression conditions were not identical to an expression of the genomic *ehi\_127670*. Thus, the results need to be interpreted under this premise. Considering previous studies, the influence of those factors for the significance of the localization results may be estimated.

The pNCcMyc vector was used in previous studies to localize a variety of putative EhAIG proteins of *E. histolytica*, ranging from 23 to 44 kDa (Matthiesen, 2009). Since those results were confirmed by the localization with generated polyclonal antibodies, an influence of the myc-tag to the translocation of the protein is unlikely. Also, the vector pNC (Hamann *et al.*, 1995), which is the basis for pNCcMyc, is a well-established expression vector for *E. histolytica* and was used for many experiments, which results were peer-reviewed and confirmed (Meyer *et al.*, 2016; König *et al.*, 2020). This reinforces the significance of the localization as well. Considering the small size of EHI\_127670, an influence of the myc-tag to the folding and the translocation cannot be excluded completely, although no influence could be observed in prior experiments with slightly bigger proteins. The same applies for the expression with pNC, which is well established, but not identical to the expression of a genomic gene. Overall, the significance of the localization can be estimated to be very high, although the results should be confirmed by IFA with a generated  $\alpha$ -EHI\_127670 antibody.

The fluorescence signal in figures 4 and 6 indicates the localization in numerous vesicle-like compartments in the cytoplasm, with a spherical shape and a diameter of approx. 500 nm. The protein was not released into the lysate during cell lysis. Either EHI\_127670 is solved in the lumen of those compartments and the compartments remain intact during freeze and thaw lysis, or the protein is associated to the compartments membrane.

There is a large number of vesicles present in the cytoplasm of *E. histolytica*, forming membrane-bound organelles like the endoplasmic reticulum, a Golgi-like apparatus (Teixeira *et al.*, 2008; Mazzuco *et al.*, 1996), as well as mitosomes, endosomes, lysosomes and multivesicular bodies (Perdomo *et al.*, 2014; Smith *et al.*, 2010). Although varying in size and shape, the mass of vesicles in *E. histolytica* makes an assignment of the fluorescent spots of EHI\_127670 based on their shape in figures 4 and 6 unreasonable.

Instead, the localization of other, better characterized pathogenicity factors was considered. Of the three major known pathogenicity factors 170 kDa lectin, amoebapores and cysteine peptidases, 170 kDa lectin is bound to the membrane surface, while amoebapores and some cysteine peptidases are stored in lysosome-like secretory granulae in the cytoplasm, and secreted after contact with the host tissue (Leippe, 1997; Mann, 2002; Bruchhaus *et al.*, 2015) or, regarding the cysteine peptidases, also spontaneously (Leippe *et al.*, 1995).

The fluorescence images from the localization of amoebapores and cysteine peptidases (Andrä *et al.*, 2003; Ghosh *et al.*, 1999) show a similar pattern to the pattern observed for EHI\_127670 in figures 4 and 6. Although the pathogenicity mechanism of EHI\_127670 is not investigated yet, considering its characterization as pathogenicity factor in phenotypic experiments (Meyer *et al.*, 2016; Haferkorn, 2018; Nehls, 2020) and its localization in apparently lysosome-like granulae, it is possible that EHI\_127670 is stored in lysosomal granulae too. Those might be part of the phagocytic or secretory pathway or play a double role in both, like it was observed for some cysteine peptidases and amoebapores (Que *et al.*, 2002; Andrä *et al.*, 2003). In this case, the protein might have a direct pathogenic influence, like the permeabilization of the host cell membrane, or work indirectly as an interaction partner of another pathogenicity factor.

Another option is the association of EHI\_127670 to vesicles of the endosomal or secretory pathway, working as a regulatory factor for vesicular transport of pathogenicity factors like it was observed for the protein Rab11b (Mitra *et al.*, 2007). However, those hypotheses are highly speculative and the presence in any other cell organelle cannot be excluded yet.

Assuming the fluorescent spots containing EHI\_127670 are any kind of membrane-bound vesicle, this raises the question of the encoding of the translocation information in *ehi\_127670*. In most eucaryotes, the typical pathway for the translocation of proteins into membrane-bound vesicles starts with the expression of the proteins into the rough endoplasmic reticulum, where they are packed into vesicles and transported into the Golgi apparatus. There they are modified and dispatched in vesicles for further transportation to their intended destination. The information for the synthesis into the ER and subsequent localization of proteins of the secretory pathway is usually encoded in the signal peptide, a short amino acid sequence N-terminal of the protein (Owji *et al.*, 2018).

Since signal peptides are abundant and a continuous endoplasmic reticulum and a Golgi-like apparatus were found in *E. histolytica*, this process most likely defines the secretory pathway in this organism, although the exact mechanisms might vary due to the partial vesiculation of the ER and the Golgi-like apparatus (Teixeira *et al.*, 2008; Mazzuco *et al.*, 1997). However, EHI\_127670 does not contain a known signal peptide. Amino acid sequence analysis in SignalP 3.0 from DTU Denmark was performed to estimate the likelihood of an unknown signal peptide. Whereas the hidden Markov models result predicted none, the neural network model postulated a signal peptide with a cleavage site between positions 14 and 15. The presence of an unknown signal sequence can therefore not be excluded.

There are two more possibilities, first the passive translocation of misfolded EHI\_127670 into endosomes for degradation, and second a signal peptide-independent translocation mechanism. Since the full protein and no degradation fragments were detected in the pellet in figure 3, the degradation hypothesis seems very unlikely.

A signal peptide-independent translocation mechanism was observed in several eucaryotes, including macrophages, *S. cerevisiae* and *D. discoideum*, where proteins lacking a signal sequence were found to be secreted in a ER- and Golgi-independent pathway, a process termed unconventional protein secretion (Nickel *et al.*, 2009; Kinseth *et al.*, 2007). Although this process was not described for *E. histolytica* yet, it could explain a signal peptide-independent localization of EHI\_127670. However, this hypothesis is highly speculative and needs further examination.

For the first subsequent investigations, the possibility of the protein being part of the same secretory pathway like amoebapores or certain cysteine peptidases could be examined by a colocalization with already localized proteins of those groups. Moreover, considering the results from the BioID method in section 5.2 under the presumption of a correct localization of EHI\_127670: BirA\*, the existence of several potential interaction candidates or proteins that were in spatial proximity to EHI\_127670 could be a starting point for further investigations of both locus and function of the protein.

After elaboration and optimization of the BioID method, mass spectroscopy could be used to characterize the interaction partners of which some might be localized already, thereby localizing EHI\_127670 indirectly. For now, the nature of the discovered vesicles containing EHI\_127670 remains unknown. Hence, they will be termed “undefined granulae” from now on.

## 6.2 BioID

The aim of the BioID method was the general implementation for *Entamoeba histolytica* as a proof-of-concept, and using it for first investigations of EHI\_127670. To examine the general expression of BirA\* in *E. histolytica*, it was first attempted to express the protein without fusion to a gene of interest. Considering the strongly positive control and the long exposure time of 300 s in figure 9, the BirA\* protein was presumably not expressed in the transfectants under the described conditions. This might be due to the identification of the foreign DNA or protein by the cell and interruption of the protein expression at transcriptional or translational level.

Little is known about the mechanisms of *E. histolytica* recognizing foreign genes or proteins. In previous studies, the implementation of the CRISPR/Cas9 system in *E. histolytica* failed, because the bacterial Cas9 protein was not expressed (Lender, 2019). At least two general mechanisms are conceivable: First, an active recognition of foreign DNA or mRNA by the cell; this presupposes a putative expression control system that senses foreign DNA or mRNA sequences and intercepts the expression process at transcriptomic or translational level. Second, the recognition of foreign proteins after translation, either actively or passively.

Actively by tagging and degrading the protein, similar to the ubiquitin/proteasome system (Hershko *et al.*, 1991). Passively by using conserved cellular quality control mechanisms of recognizing and degrading proteins that are misfolded due to *E. histolytica*-specific intracellular conditions like pH or ionic strength (Houck *et al.*, 2012).

Those considerations are speculative. However, presuming the recognition of BirA\* occurs at transcriptomic or translational level, the 5'-end of the sequence might play a key role, since it is the starting point for both transcription and translation. Considering this, under the presumption that any gene naturally occurring in *E. histolytica* is not recognized as foreign and therefore will be expressed, the recognition of the foreign BirA\* might be prevented by fusing it C-terminal to an amoebic protein. Since the aim of this work was to implement BioID for the characterization of the putative pathogenicity factor EHI\_127670, it served both as BirA\*-concealing sequence and protein to be investigated. Thus, the cloning strategy was proceeded to assemble the pNCBID:*ehi\_127670* construct. However, since the expression of BirA\* without EHI\_127670 is necessary for a BirA\*-only control, the identification of BirA\*-dependent false-positive signals could not be achieved.

The signal in figure 11 confirms the hypothesis of the concealing function of *ehi\_127670* for the EHI\_127670: BirA\* fusion protein. After successfully expressing the fusion protein, it was proceeded to examine the biotinylation of potential interaction partners of EHI\_127670. The biotinylation pattern in figure 12 also occurs in sample B2apMyc<sup>127-</sup>, which indicates a biotinylation of cellular proteins without the addition of biotin to the medium or the presence of BirA\*. The medium-inherent biotin concentration was 0.97  $\mu$ M, contributed by the components yeast extract and Diamond Vitamin Tween® 80 solution. This does not seem sufficient for the uncatalyzed biotinylation of proteins in B2apMyc<sup>127-</sup>, thus an unknown catalyst for biotinylation in *E. histolytica* is suspected.

An Acetyl-CoA carboxylase-like protein with biotin-carboxylase activity was discovered in *E. histolytica* (Barbosa-Cabrera *et al.* 2012), which implies the occurrence of biotin in the amoebic metabolism. In *E. coli*, the wild type BirA catalyses the biotinylation of the biotin carboxyl carrier protein, which is a subunit of Acetyl-CoA carboxylase. Therefore, the BirA\*-nonspecific biotinylation in sample B2apMyc<sup>127-</sup> in figure 12 might be caused by an unknown BirA-like protein, or a totally different protein with biotin-protein ligase activity, which has evolved in the biotin-dependent metabolism of *E. histolytica*.

The lack of correlation of the intensity of the bands to the absence or presence of additional biotin in the medium in B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>, implies the influence of another hidden parameter. The overall intensity of the bands in figure 13 is an indication for the amount of BirA\* applied for SDS-PAGE, and assuming a constant expression level of BirA\* in the same transfectant, the total amount of protein applied. Since the overall intensity of the bands in figure 13 correlates with the intensity of the bands in figure 12, and the gel duplicates were treated identically, the hidden parameter that influences the varying intensity might be the amount of protein applied for SDS-PAGE. This hypothesis is plausible, since a band with a high amount of a non-specifically and weakly biotinylated protein can carry more biotin than a band with a low amount of a specifically and severely biotinylated protein. Because all samples were prepared identically, the amount of protein was most likely determined by the cell number and therefore the cell confluency of the trophozoites in the T25 flask prior to cell harvest.

Before biotinylation, the trophozoites of a T25 culture flask were split equally into two T25 culture flask, with the original flask being reused. Despite the equal splitting ratio, the flask being reused contained more cells afterwards, because a significant amount of trophozoites does not detach when splitting at room temperature. During the 24 h incubation time before the harvest, exponential growth increased the difference of the cell number between both flasks. Thus, the transfectants B2apBID<sup>127</sup> and B2apMyc<sup>127</sup>, that had the biotin-positive portion of cells split into a new flask, show a higher biotinylation for the biotin-negative portion, which remained in the pre-used T25 flask. For B2bpBID<sup>127</sup>, which had the biotin-negative portion of cells split into a new flask, it is the opposite. The band pattern in figure 13 indicates the presence of the full fusion protein in the pellet, whereas fragments, presumably produced by degradation of EHI\_127670: BirA\*, are present in both lysate and pellet. In sample B2apBID<sup>127</sup>, the full fusion protein is also present in the lysate, which is presumably caused by aspiration of the pellet when the supernatant was isolated.

Following these considerations, the BioID method was adjusted and another run of biotinylation performed. Although the variation of the different sample's band intensity detected in the streptavidin blot in figure 14 A is significantly lower than in figure 12, a small divergence remains. Besides the measurement uncertainty of the counting chamber, other factors are conceivable for this phenomenon.

First, besides the cell number, the cell composition determines the total protein amount of each sample, and might differ between the transfectants. Second, the volume and therefore the protein concentration in the lysate fraction of each sample depends on the residual buffer in the pellet prior to cell lysis. Since the relevant bands were detected in the pellet fraction, this can be neglected, presuming an even suspension of the cell pellets contents before loading the sample into the gel.

Another possibility is the activity of free BirA\*, that was cleaved from the fusion protein, in the cytosol, resulting in a BirA\*-dependent non-specific biotinylation of cytosolic proteins by BirA\* in B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>. The first two bands from top at approx. 50 and 42 kDa in the lysate fractions in figure 14 B presumably do contain the full BirA\* (42 kDa), respectively a bigger fragment of the fusion protein (42 + 8 kDa). Those proteins do most likely have a functional protein-biotin ligase activity, biotinylating cytosolic proteins non-specifically, thereby causing a higher degree of non-specific biotinylation.

Assuming the first explanation of an overall lower protein mass loaded into the gel for B2apMyc<sup>127</sup>, this leaves the possibility of existent, but not detected bands in B2apMyc<sup>127</sup>, which displays the overall lowest band intensity. This was ruled out to some extent by overexposing the blot for 2400 s to amplify the weakest signals. However, to eliminate the risk of not detected bands in the control, the protein mass loaded into the gel for B2apMyc<sup>127</sup> should be exceeded in further experiments, e.g. by increasing the sample volume compared to B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>.

In the case of no undetected bands in sample B2apMyc<sup>127</sup> in figure 15 A, seven potential interaction partners of EHI\_127670 are most likely differential biotinylated in B2apBID<sup>127</sup> and B2bpBID<sup>127</sup>. Self-biotinylation of the fusion protein is an indication for the activity of BirA\* (Choi-Rhee *et al.*, 2004), thus the bands at approx. 58 kDa in the pellet may be assigned to EHI\_127670: BirA\*.

Because of the strong background signal for bands under approx. 50 kDa, some interaction partners of that size may be present, but not distinguishable. Due to the limited specificity of the biotinylation and the streptavidin blot, those results are only a coarse indication for further investigations.

Also, depending on the overall protein concentration in the undefined granulae, the spatial proximity of proteins that do not directly interact to EHI\_127670:BirA\* might provoke false positive signals. A simple control would be the expression of the EHI\_127670:BirA\* fusion protein in a first, and BirA\* in a second control culture, and subtraction of the signals. However, since the expression of BirA\* without an insert failed, this was not possible in this work, but should be achieved in further studies.

A more accurate way to correct false-positive signals would be the deployment of DiQ-BioID, where the target protein, in this case EHI\_127670, and BirA\* are expressed separately (Birnbaum *et al.*, 2020). BirA\* is recruited to EHI\_127670 by the small molecule rapalog, which interconnects the proteins to a fusion protein-like dimer, enabling BirA\* to biotinylate interaction partners of EHI\_127670. As control, a culture of the same transfectants is grown without rapalog, resulting non-specific biotinylation that can be subtracted from the results of the culture grown with rapalog. The correct translocation of BirA\* into the undefined granulae would be crucial for both methods. Since no known signal peptide could be identified in EHI\_127670, and the encoding of the translocation target information remains unknown, further experiments would be needed to achieve a translocation of BirA\* to the locus of EHI\_127670.

Despite potentially false positive results because of spatial proximity of proteins not interacting with EHI\_127670, the identification of those proteins would be a valuable information as well, since the presence in the same granulae indicates the participation in the same pathway. Moreover, by the discovery of already characterized and localized proteins in spatial proximity of EHI\_127670, the nature of the undefined vesicles could be illuminated, assuming the correct translocation of EHI\_127670:BirA\*.

Prospectively, the direct identification of interaction partners of EHI\_127670 by mass spectroscopy is aspired. Since this detection method is very sensitive, the massive noise by BirA\*-non-specific biotinylation in the sample would jeopardize the detection and make an analysis without prior purification unreasonable. Assuming the correct localization of EHI\_127670: BirA\* in the undefined granulae, which is the necessary criterion for significant results, the granulae could be purified by differential centrifugation. Subsequently, the granulae could be lysed and the proteins of interest chromatographically purified from the presumably non-specific biotinylated proteins.

Since the presumably non-specific biotinylated proteins have a size smaller than approx. 48 kDa, whereas all potential interaction partners have a molecular weight higher than or equal to 48 kDa, size exclusion chromatography with columns of a sufficient performance could be used for separation. The thereby purified proteins could then be prepared for mass spectroscopy.

On the long run, the BioID method should be optimized by the deployment of a suitable control for the reduction of BirA\*-dependent false positive signals, as well as a reduction of BirA\*-independent unspecific biotinylation. The former was already discussed above. The latter could be achieved by the cultivation of transfectants in biotin-free medium for a certain time before biotinylation. Depending on the turnover rate for biotinylated proteins, the unspecific biotinylation, cumulated over time, could be reduced. The subsequent application of a definite concentration of biotin might result in a higher degree of BirA\*-dependent biotinylation relative to BirA\*-non-dependent biotinylation, assuming a higher activity of BirA\* compared the unknown catalyst hypothesized before.

The correct translocation of the EHI\_127670: BirA\* fusion protein is crucial for the BioID method, since the natural interaction partners can only be detected when in spatial proximity. In this case, the fusion protein was expected to be localized in the undefined granulae, that were identified as the assumed natural locus of EHI\_127670. However, the relatively high molecular weight of BirA\* compared to EHI\_127670 might impede the translocation of the fusion protein. Whereas the western blots in figures 11 and 14 B show the full fusion protein to be present in the pellet fraction, the signals of EHI\_127670: BirA\* in the figures 16-18 imply a mislocalization in the cytosol.

Those apparently contradicting results may be explained by putting the fluorescence signal in context to the entirety of bands detected in the western blot in figure 14 B. Here, besides the band of the full-sized fusion protein at 55 kDa, bands of smaller fragments are visible in the pellet and the lysate fraction. Those do contain fragments of the degraded fusion protein that still have a HA-tag and are therefore detectable by western blot and IFA. Despite the ratio of degraded to intact fusion protein is unknown, the intensity of the bands indicates a significant mass of delocalized protein fragments in the cytosol.

Since those fragments are detected by IFA, they produce a massive noise signal, which could potentially conceal the signal of the correctly localized EHI\_127670:BirA\*. This hypothesis is substantiated by the presence of the full fusion protein in the pellet fraction of the western blot, which might indicate a correct localization in the undefined granulae.

The overall higher signal in figure 18 A compared to 18 B, relatively to the signal outside of the cell, would underpin that hypothesis as well, under the presumption that those samples could be compared. This is not the case, since the affinity of the two used antibody systems to their antigens is most likely not identical, and the depicted cells do not represent the median fluorescence intensity of the entirety of cells in the sample. However, the overall strong signal in the cytosol in figure 18 A still indicates a concealing of the fluorescence signal of the potentially correctly translocated fusion protein by an overlaying noise signal, provoked by fragments of EHI\_127670:BirA\* in the cytosol.

This leaves the question of the degradation mechanism producing those fragments. The signal in the cytosol in figure 18 A and the high concentration of protease inhibitor E64 during cell lysis indicate a degradation of the fusion protein prior to the cell lysis process. Therefore, the fusion protein must be degraded continuously in the cell, which causes the band pattern in figure 14 B. The fragment size of the bands is identical in all pellet and lysate fractions, thus the protein is cleaved by proteases at specific cleavage sites.

Since the full fusion protein is present exclusively in the pellet fraction, whereas the fragments are present in both lysate and pellet fraction, the degradation process most likely starts in the undefined granulae, assuming the visible HA-tagged fragments in the pellet fraction are cleaved from EHI\_127670 and therefore could not be translocated into the granulae after the degradation. Therefore, the fragments present in the lysate must be ejected from the undefined granulae into the cytosol, where the degradation possibly continued.

An ER-dependent secretory pathway would substantiate that theory, since the protein would be synthesized directly into the ER, followed by the budding of fusion protein-packed vesicles, that would pass the secretory pathway ending with the undefined granulae as destination. In this process, the full fusion protein would never be present in the cytosol, only fragments of it after ejection from the granulae.

Since there might be no known signal peptide present in EHI\_127670, it was hypothesized that the pathway was ER-independent, thus the fusion protein might have been exposed to the cytoplasm during the translocation process. However, assuming the permanent isolation in membrane-bound vesicles during the ER-independent pathway as well, the degradation of the fusion protein could be explained by the presence of cysteine peptidases, under the presumption of a colocalization in the same secretory vesicle. Considering the molecular weight of 44 kDa of BirA\* with linker and tag, which is significantly higher than the 14.5 kDa of EHI\_127670, the degradation could also be provoked during a disturbed translocation process, where BirA\* is cleaved from EHI\_127670.

Another possibility is that the fusion protein is misfolded because of the foreign, *E. coli*-derived BirA\* part and stored in inclusion bodies, and thereby detected in the pellet. Since there were no publications found confirming the presence of inclusion bodies in *E. histolytica*, and the streptavidin blots in figure 14 and 15 indicate a biotinylation by active BirA\*, this hypothesis is very unlikely.

A production of fragments by faulty protein synthesis can be excluded as explanation for the noise signal, since the HA tag is located C-terminal in the protein.

Prospectively, the correct localization of EHI\_127670: BirA\* in the undefined granulae needs to be examined. Therefore, a membrane permeant protease inhibitor like E64d (Murray *et al.*, 1997) could be applied to the cells for a defined time period before harvest, to inhibit the degradation of EHI\_127670: BirA\* and diminish the resulting noise signal. However, by the moment the inhibitor is applied, there will be fragments of the fusion protein present in the cytosol, which still need to be metabolized. Since the turnover rate of the already degraded protein fragments, as well as the general cell viability, will most likely decline by inhibition of proteases, the concentration of the inhibitor and the incubation time needs to be titrated.

Those considerations are based on the presumption, that there is an optimal inhibitor concentration that allows a faster degradation of the fusion protein fragments than degradation of the fusion protein, which would result in a fragments-free cytosol after a certain time. This period would need to be in the time window of cell viability, since the inhibition of proteases will ultimately lead to the cell's death after some time.

If those conditions prove to be wrong or mutually exclusive, another option might be feasible. To eliminate the problem of fusion protein fragments already present by the time of inhibitor application, the *ehi\_127670: birA\** gene could be induced after inhibitor application. Thereby, the protein would be expressed under non-proteolytic conditions, and no pre-existing fusion protein fragments would need to be degraded, thus the degree of inhibition could be set to a maximum instead of an optimum.

A direct gene induction in *E. histolytica* was established by tetracycline-controlled gene expression on an episomal transfected plasmid (Hamann *et al.*, 1997). This method is based on the transient transfection with a vector derived from the vector pNC, which was used in this work as well, hence it might be easily deployable for an induced expression of BirA\* too. Moreover, advances in synthetic biology have led to a variety of novel approaches for induced gene expression in eukaryotes, e.g. in biotechnological applications of yeast (Redden *et al.*, 2015; Peng *et al.*, 2015) and in *P. falciparum* (Goldfless *et al.*, 2014), which might be adaptable to *E. histolytica* as well. Considering the complexity of those sophisticated systems, the expenditure for the implementation in *E. histolytica* is estimated to be very high, but might be worthwhile for a variety of applications in the long run.

## 7 Outlook

After the characterization of EHI\_127670 as protein and the localization in apparently membrane bound vesicles, termed “undefined granulae”, the nature of these vesicles needs to be investigated. Hypothesized to be part of the phagocytic or secretory pathway, EHI\_127670 might possibly be colocalized with known pathogenic factors like certain cysteine peptidases or amoebapores. A colocalization with marker molecules of other defined organelles, e.g. mitochondria, endosomes, or multivesicular bodies should also be attempted. Because of the presumed spatial proximity and possible interaction of EHI\_127670 with proteins that might be localized already, the BioID method might help to colocalize the protein indirectly. However, the localization in other membrane bound vesicles should be examined as well. Moreover, the myc-localization should be verified by the localization with generated antibodies.

The BioID method identified seven potential interaction candidates for EHI\_127670. Those might be identified by mass spectroscopy after purification of the undefined granulae and subsequent size exclusion chromatography. First, to validate the biotinylation of the seven potential interaction partners as BirA\*-specific, the sample volume for the control in the streptavidin blot should be increased to rule out the presence of not detected bands. Furthermore, for a verification of any BioID results, a suitable control needs to be implemented and a colocalization of EHI\_127670: BirA\* with EHI\_127670 needs to be performed, which requires the clarification of the translocation mechanism of EHI\_127670 and the elimination of fusion protein fragments in the cytosol. Moreover, the BirA\*-independent non-specific biotinylation needs to be reduced, potentially by cultivation in biotin-free medium.

After optimization, the BioID method might be a powerful tool for the elucidation of the unknown pathogenicity mechanism of EHI\_127670. By applying the method to other pathogenicity factors, their interactome could be deciphered, which could provide new insights to the pathogenicity of *E. histolytica* and contribute to the fight against amebiasis.

## References

Andrä, J., Berninghausen, O., & Leippe, M. (2004). Membrane lipid composition protects *Entamoeba histolytica* from self-destruction by its pore-forming toxins. *FEBS letters*, *564* (1-2), 109–115. [https://doi.org/10.1016/S0014-5793\(04\)00324-2](https://doi.org/10.1016/S0014-5793(04)00324-2)

Andrä, J., Herbst, R., & Leippe, M. (2003). Amoebapores, archaic effector peptides of protozoan origin, are discharged into phagosomes and kill bacteria by permeabilizing their membranes. *Developmental and comparative immunology*, *27*(4), 291–304. [https://doi.org/10.1016/s0145-305x\(02\)00106-4](https://doi.org/10.1016/s0145-305x(02)00106-4)

Bailey, G. B., Shen, P. S., Beanan, M. J., & McCoomer, N. E. (1992). Actin associated proteins of *Entamoeba histolytica*. *Archives of medical research*, *23*(2), 129–132.

Barbosa-Cabrera, E., Salas-Casas, A., Rojas-Hernández, S., Jarillo-Luna, A., Abarca-Rojano, E., Rodríguez, M. A., & Campos-Rodríguez, R. (2012). Purification and cellular localization of the *Entamoeba histolytica* transcarboxylase. *Parasitology research*, *111*(3), 1401–1405. <https://doi.org/10.1007/s00436-012-2898-6>

Batsios, P., Meyer, I., & Gräf, R. (2016). Proximity-Dependent Biotin Identification (BioID) in *Dictyostelium* Amoebae. *Methods in enzymology*, *569*, 23–42. <https://doi.org/10.1016/bs.mie.2015.09.007>

Bercu, T. E., Petri, W. A., & Behm, J. W. (2007). Amebic colitis: new insights into pathogenesis and treatment. *Current gastroenterology reports*, *9*(5), 429–433. <https://doi.org/10.1007/s11894-007-0054-8>

Birnbaum, J., Scharf, S., Schmidt, S., Jonscher, E., Hoeijmakers, W., Flemming, S., Toenhake, C. G., Schmitt, M., Sabitzki, R., Bergmann, B., Fröhlke, U., Mesén-Ramírez, P., Blancke Soares, A., Herrmann, H., Bártfai, R., & Spielmann, T. (2020). A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. *Science (New York, N.Y.)*, *367*(6473), 51–59. <https://doi.org/10.1126/science.aax4735>

Bruchhhaus, I. (2015). Cysteine Peptidases in Pathogenesis. In T. Nozaki & A. Bhattcharya (Eds.), *Amoebiasis—biology and pathogenesis of Entamoeba* (447–458). Springer. ISBN 978-4-431-55199-7.

Choi-Rhee, E., Schulman, H., & Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein science : a publication of the Protein Society*, *13*(11), 3043–3050. <https://doi.org/10.1110/ps.04911804>

Diamond, L. S., Harlow, D. R., & Cunnick, C. C. (1978). A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *72*(4), 431–432. [https://doi.org/10.1016/0035-9203\(78\)90144-x](https://doi.org/10.1016/0035-9203(78)90144-x)

- Ghosh, S. K., Field, J., Frisardi, M., Rosenthal, B., Mai, Z., Rogers, R., & Samuelson, J. (1999). Chitinase secretion by encysting *Entamoeba invadens* and transfected *Entamoeba histolytica* trophozoites: localization of secretory vesicles, endoplasmic reticulum, and Golgi apparatus. *Infection and immunity*, *67*(6), 3073–3081. <https://doi.org/10.1128/IAI.67.6.3073-3081.1999>
- Goldfless, S. J., Wagner, J. C., & Niles, J. C. (2014). Versatile control of *Plasmodium falciparum* gene expression with an inducible protein-RNA interaction. *Nature communications*, *5*, 5329. <https://doi.org/10.1038/ncomms6329>
- Haferkorn, A. (2018). Charakterisierung des hypothetischen Proteins (127670) mittels Überexpression bzw. Silencing in pathogenen und apathogenen *Entamoeba histolytica* Trophozoiten. Master thesis, University Hamburg
- Hamann, L., Buss, H., & Tannich, E. (1997). Tetracycline-controlled gene expression in *Entamoeba histolytica*. *Molecular and biochemical parasitology*, *84*(1), 83–91. [https://doi.org/10.1016/s0166-6851\(96\)02771-5](https://doi.org/10.1016/s0166-6851(96)02771-5)
- Handal, G. (2010). Regulation of antioxidant enzymes in *Entamoeba histolytica* (Schaudinn, 1903). Dissertation, University Hamburg
- Hershko A. (1991). The ubiquitin pathway of protein degradation and proteolysis of ubiquitin-protein conjugates. *Biochemical Society transactions*, *19*(3), 726–729. <https://doi.org/10.1042/bst0190726>
- Barbosa-Cabrera, E., Salas-Casas, A., Rojas-Hernández, S., Jarillo-Luna, A., Abarca-Rojano, E., Rodríguez, M. A., & Campos-Rodríguez, R. (2012). Hon, C. C., Weber, C., Sismeiro, O., Proux, C., Koutero, M., Deloger, M., Das, S., Agrahari, M., Dillies, M. A., Jagla, B., Coppee, J. Y., Bhattacharya, A., & Guillen, N. (2013). Quantification of stochastic noise of splicing and polyadenylation in *Entamoeba histolytica*. *Nucleic acids research*, *41*(3), 1936–1952. <https://doi.org/10.1093/nar/gks1271>
- Houck, S. A., Singh, S., & Cyr, D. M. (2012). Cellular responses to misfolded proteins and protein aggregates. *Methods in molecular biology (Clifton, N.J.)*, *832*, 455–461. [https://doi.org/10.1007/978-1-61779-474-2\\_32](https://doi.org/10.1007/978-1-61779-474-2_32)
- Jacobs, T., Bruchhaus, I., Dandekar, T., Tannich, E., & Leippe, M. (1998). Isolation and molecular characterization of a surface-bound proteinase of *Entamoeba histolytica*. *Molecular microbiology*, *27*(2), 269–276. <https://doi.org/10.1046/j.1365-2958.1998.00662.x>
- Kim, D. I., Birendra, K. C., Zhu, W., Motamedchaboki, K., Doye, V., & Roux, K. J. (2014). Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), E2453–E2461. <https://doi.org/10.1073/pnas.1406459111>
- Kinseth, M. A., Anjard, C., Fuller, D., Guizzunti, G., Loomis, W. F., & Malhotra, V. (2007). The Golgi-associated protein GRASP is required for unconventional protein secretion during development. *Cell*, *130*(3), 524–534. <https://doi.org/10.1016/j.cell.2007.06.029>

- König, C., Meyer, M., Lender, C., Nehls, S., Wallaschkowski, T., Holm, T., Matthies, T., Lercher, D., Matthiesen, J., Fehling, H., Roeder, T., Reindl, S., Rosenthal, M., Metwally, N. G., Lotter, H., & Bruchhaus, I. (2020). An Alcohol Dehydrogenase 3 (ADH3) from *Entamoeba histolytica* Is Involved in the Detoxification of Toxic Aldehydes. *Microorganisms*, 8(10), 1608. <https://doi.org/10.3390/microorganisms8101608>
- Laughlin, R. C., & Temesvari, L. A. (2005). Cellular and molecular mechanisms that underlie *Entamoeba histolytica* pathogenesis: prospects for intervention. *Expert reviews in molecular medicine*, 7(13), 1–19. <https://doi.org/10.1017/S1462399405009622>
- Leippe M. (1997). Amoebapores. *Parasitology today (Personal ed.)*, 13(5), 178–183. [https://doi.org/10.1016/s0169-4758\(97\)01038-7](https://doi.org/10.1016/s0169-4758(97)01038-7)
- Leippe, M., & Müller-Eberhard, H. J. (1994). The pore-forming peptide of *Entamoeba histolytica*, the protozoan parasite causing human amoebiasis. *Toxicology*, 87(1-3), 5–18. [https://doi.org/10.1016/0300-483x\(94\)90151-1](https://doi.org/10.1016/0300-483x(94)90151-1)
- Leippe, M., Sievertsen, H. J., Tannich, E., & Horstmann, R. D. (1995). Spontaneous release of cysteine proteinases but not of pore-forming peptides by viable *Entamoeba histolytica*. *Parasitology*, 111 ( Pt 5), 569–574. <https://doi.org/10.1017/s0031182000077040>
- Lender, S. (2019). Genetische Manipulation des protozoischen Parasiten *Entamoeba histolytica* (SCHAUDINN, 1903) zur Charakterisierung putativer Pathogenitätsfaktoren. Dissertation, University Hamburg
- Lorenzi, H. A., Puiu, D., Miller, J. R., Brinkac, L. M., Amedeo, P., Hall, N., & Caler, E. V. (2010). New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS neglected tropical diseases*, 4(6), e716. <https://doi.org/10.1371/journal.pntd.0000716>
- Mann B. J. (2002). Structure and function of the *Entamoeba histolytica* Gal/GalNAc lectin. *Int. review of cytology*, 216, 59–80. [https://doi.org/10.1016/s0074-7696\(02\)16003-7](https://doi.org/10.1016/s0074-7696(02)16003-7)
- Matthiesen, J. (2009). Charakterisierung putativer EhAIG-Proteine und Aufbau eines Protein-Markierungssystems bei *Entamoeba histolytica* (Schaudinn, 1903). Diploma thesis, University Hamburg
- Mazzuco, A., Benchimol, M., & De Souza, W. (1997). Endoplasmic reticulum and Golgi-like elements in *Entamoeba*. *Micron (Oxford, England : 1993)*, 28(3), 241–247. [https://doi.org/10.1016/s0968-4328\(97\)00024-3](https://doi.org/10.1016/s0968-4328(97)00024-3)
- Meyer, M., Fehling, H., Matthiesen, J., Lorenzen, S., Schuldt, K., Bernin, H., Zaruba, M., Lender, C., Ernst, T., Ittrich, H., Roeder, T., Tannich, E., Lotter, H., & Bruchhaus, I. (2016). Overexpression of Differentially Expressed Genes Identified in Non-pathogenic and Pathogenic *Entamoeba histolytica* Clones Allow Identification of New Pathogenicity Factors Involved in Amoebic Liver Abscess Formation. *PLoS pathogens*, 12(8), e1005853. <https://doi.org/10.1371/journal.ppat.1005853>

- Mitra, B. N., Saito-Nakano, Y., Nakada-Tsukui, K., Sato, D., & Nozaki, T. (2007). Rab11B small GTPase regulates secretion of cysteine proteases in the enteric protozoan parasite *Entamoeba histolytica*. *Cellular microbiology*, 9(9), 2112–2125. <https://doi.org/10.1111/j.1462-5822.2007.00941.x>
- Moncada, D., Keller, K., & Chadee, K. (2003). *Entamoeba histolytica* cysteine proteinases disrupt the polymeric structure of colonic mucin and alter its protective function. *Infection and immunity*, 71(2), 838–844. <https://doi.org/10.1128/IAI.71.2.838-844.2003>
- Murray, E. J., Grisanti, M. S., Bentley, G. V., & Murray, S. S. (1997). E64d, a membrane-permeable cysteine protease inhibitor, attenuates the effects of parathyroid hormone on osteoblasts in vitro. *Metabolism: clinical and experimental*, 46(9), 1090–1094. [https://doi.org/10.1016/s0026-0495\(97\)90284-5](https://doi.org/10.1016/s0026-0495(97)90284-5)
- Nehls, S. (2020). Einfluss einer Alkoholdehydrogenase und des hypothetischen Proteins EHI\_127670 auf die Virulenz von *Entamoeba histolytica* (Schaudinn, 1903). Master thesis, University Hamburg
- Nickel W, Rabouille C (2009). Mechanisms of regulated unconventional protein secretion. *Nat Rev Mol Cell Biol*. Feb;10(2):148-55. doi: 10.1038/nrm2617. Epub 2008 Dec 24. Erratum in: *Nat Rev Mol Cell Biol*. 2009 Mar;10(3):234. PMID: 19122676
- Nickel, R., Ott, C., Dandekar, T., & Leippe, M. (1999). Pore-forming peptides of *Entamoeba dispar*. Similarity and divergence to amoebapores in structure, expression and activity. *European journal of biochemistry*, 265(3), 1002–1007. <https://doi.org/10.1046/j.1432-1327.1999.00807.x>
- Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., & Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. *European journal of cell biology*, 97(6), 422–441. <https://doi.org/10.1016/j.ejcb.2018.06.003>
- Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., & Vickers, C. E. (2015). Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microbial cell factories*, 14, 91. <https://doi.org/10.1186/s12934-015-0278-5>
- Perdomo, D., Ait-Ammar, N., Syan, S., Sachse, M., Jhingan, G. D., & Guillén, N. (2015). Cellular and proteomics analysis of the endomembrane system from the unicellular *Entamoeba histolytica*. *Journal of proteomics*, 112, 125–140. <https://doi.org/10.1016/j.jprot.2014.07.034>
- Pillai, D. R., & Kain, K. C. (2000). Recent developments in amoebiasis: the Gal/GalNAc lectins of *Entamoeba histolytica* and *Entamoeba dispar*. *Microbes and infection*, 2(14), 1775–1783. [https://doi.org/10.1016/s1286-4579\(00\)01330-7](https://doi.org/10.1016/s1286-4579(00)01330-7)

- Que, X., Brinen, L. S., Perkins, P., Herdman, S., Hirata, K., Torian, B. E., Rubin, H., McKerrow, J. H., & Reed, S. L. (2002). Cysteine proteinases from distinct cellular compartments are recruited to phagocytic vesicles by *Entamoeba histolytica*. *Molecular and biochemical parasitology*, *119*(1), 23–32. [https://doi.org/10.1016/s0166-6851\(01\)00387-5](https://doi.org/10.1016/s0166-6851(01)00387-5)
- Que, X., Kim, S. H., Sajid, M., Eckmann, L., Dinarello, C. A., McKerrow, J. H., & Reed, S. L. (2003). A surface amebic cysteine proteinase inactivates interleukin-18. *Infection and immunity*, *71*(3), 1274–1280. <https://doi.org/10.1128/IAI.71.3.1274-1280.2003>
- Ravdin, J. I., Petri, W. A., Murphy, C. F., & Smith, R. D. (1986). Production of mouse monoclonal antibodies which inhibit in vitro adherence of *Entamoeba histolytica* trophozoites. *Infection and immunity*, *53*(1), 1–5. <https://doi.org/10.1128/iai.53.1.1-5.1986>
- Redden, H., Morse, N., & Alper, H. S. (2015). The synthetic biology toolbox for tuning gene expression in yeast. *FEMS yeast research*, *15*(1), 1–10. <https://doi.org/10.1111/1567-1364.12188>
- Roux, K. J., Kim, D. I., Raida, M., & Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of cell biology*, *196*(6), 801–810. <https://doi.org/10.1083/jcb.201112098>
- Scholze, H., & Tannich, E. (1994). Cysteine endopeptidases of *Entamoeba histolytica*. *Methods in enzymology*, *244*, 512–523. [https://doi.org/10.1016/0076-6879\(94\)44037-9](https://doi.org/10.1016/0076-6879(94)44037-9)
- Sears, R. M., May, D. G., & Roux, K. J. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. *Methods in molecular biology (Clifton, N.J.)*, *2012*, 299–313. [https://doi.org/10.1007/978-1-4939-9546-2\\_15](https://doi.org/10.1007/978-1-4939-9546-2_15)
- Smith, Sherri & Guillen, Nancy. (2010). Organelles and Trafficking in *Entamoeba histolytica*. Structures and Organelles in Pathogenic Protists, Microbiology Monographs 17, 150-164. [https://doi.org/10.1007/978-3-642-12863-9\\_7](https://doi.org/10.1007/978-3-642-12863-9_7)
- Stanley S. L., Jr (2003). Amoebiasis. *Lancet (London, England)*, *361*(9362), 1025–1034. [https://doi.org/10.1016/S0140-6736\(03\)12830-9](https://doi.org/10.1016/S0140-6736(03)12830-9)
- Teixeira, J. E., & Huston, C. D. (2008). Evidence of a continuous endoplasmic reticulum in the protozoan parasite *Entamoeba histolytica*. *Eukaryotic cell*, *7*(7), 1222–1226. <https://doi.org/10.1128/EC.00007-08>
- Wells, C. D., & Arguedas, M. (2004). Amebic liver abscess. *Southern medical journal*, *97*(7), 673–682. <https://doi.org/10.1097/00007611-200407000-00013>
- Willhoeft U, Tannich E. The electrophoretic karyotype of *Entamoeba histolytica*. *Mol Biochem Parasitol*. 1999 Mar 15;99(1):41-53. doi: 10.1016/s0166-6851(98)00178-9. PMID: 10215023.

Ximénez, C., González, E., Nieves, M., Magaña, U., Morán, P., Gudiño-Zayas, M., Partida, O., Hernández, E., Rojas-Velázquez, L., García de León, M. C., & Maldonado, H. (2017). Differential expression of pathogenic genes of *Entamoeba histolytica* vs *E. dispar* in a model of infection using human liver tissue explants. *PloS one*, *12*(8), e0181962. <https://doi.org/10.1371/journal.pone.0181962>

Zhang, Z., Wang, L., Seydel, K. B., Li, E., Ankri, S., Mirelman, D., & Stanley, S. L., Jr (2000). *Entamoeba histolytica* cysteine proteinases with interleukin-1 beta converting enzyme (ICE) activity cause intestinal inflammation and tissue damage in amoebiasis. *Molecular microbiology*, *37*(3), 542–548. <https://doi.org/10.1046/j.1365-2958.2000.02037.x>

# Acknowledgements

Ohne die Unterstützung vieler Menschen hätte ich diese Arbeit nicht schreiben können.

Dafür möchte ich mich hier bedanken. Bei

Prof. Dr. Iris Bruchhaus dafür, dass ich Teil dieser großartigen Arbeitsgruppe sein durfte, für die fachliche Expertise bei der Betreuung dieser Arbeit und den sehr freundlichen Umgang.

Prof. Dr. Jörg Andrä für die Betreuung dieser Arbeit und den Tipp zum BNITM.

Constantin König für die entspannte Arbeitsatmosphäre, viele gute Hinweise und alles, was du mir während meiner Zeit in Labor 5 beigebracht hast, vor allem während meiner Einarbeitung. Ich habe nie so viel in so kurzer Zeit gelernt.

Suzann Ofori und Jana Brehmer für viele Tipps und die entspannte Arbeit in Labor 5.

Jakob Cronshagen für die große Unterstützung mit allen Hinweisen und beantworteten Fragen zu BioID, und für den corel crashkurs.

Der AG Spielmann für die Hilfe wenn Jakob gerade nicht da war und die Unmengen an Antikörpern.

Prof. Dr. Tobias Spielmann für die ausführlichen Einweisungen für das Mikroskop und Imaris.

Der gesamten Arbeitsgruppe Bruchhaus für die entspannte Atmosphäre, die riesige Hilfsbereitschaft und den freundschaftlichen Umgang. Ich habe die Zeit bei euch sehr genossen.

Außerdem danke ich meiner Familie für die bedingungslose Unterstützung, auch was das Studium angeht.

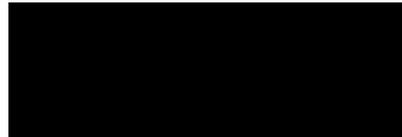
Doreen und meinen Freunden danke ich für das Abhalten von der Arbeit für die wirklich wichtigen Dinge.

## Statutory declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Hamburg, 21.06.2021

Place, Date



Signature