University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER THESIS

Curriculum: **Biological Chemistry**	Autumn semester, 2020 Open
Author: **Kaneez Fatima**	 .... (Author signature)
Tutor: **Prof. Dr. Mark van der Giezen**	
Master thesis title: ***Entamoeba invadens*; a study about the role of PLD1 and PLD2 in encystation**	
Keywords: Phospholipase D1 and 2 *Entamoeba histolytica* *Entamoeba invadens* Transfection Cysts formation	Number of pages: 68 .... + appendices/other: 2 .... Stavanger, 4.12.2020 .... date/year

ACKNOWLEDGMENTS

There were a number of people who helped me during the time period of my thesis, none who helped more than my supervisor Professor Dr. Mark van der Giezen. Thank you for giving me the opportunity of being your student. It was a learning experience and one of the most important lesson I learnt during the process was that things do not always go according to the plan and one has to come up with solutions on the go. At the time, not getting proper results was stressful but looking back, it was an opportunity for improvement and finding solutions to very interesting problems. Dr. Mark always had my back and helped throughout the process.

I would like to pay my heartiest regards and compliments to Dr. Dugassa Nemie-Feyissa for his continuous motivation and guidance and for trust in my abilities. Thank you for always helping and offering your guidance. I would also like to thank Dr. Marina Alexeeva, who is such an outstanding and humble person, who was always there to welcome me in the hour of need, who always paid her attention towards my problems regarding research work and guided me by her remarkable knowledge. Thanks to Tia Renee Tidwell (PhD scholar) for helping me with Fluorescence Microscopy.

Last but not the least, this piece of my hard work is dedicated to my family: my parents (Syed Muhammad Sibtain Gillani and Kalsoom Sibtain), and my brothers who always had faith in my abilities and always stood beside me in every moment of my life.

Sincerely,

Kaneez fatima

ABSTRACT

One of the most common problem in the developing world is infectious diseases caused by poor living conditions and one such condition is called amoebiasis. The causative agent of which is *Entamoeba histolytica*; a microaerophilic, parasitic amoebozoan species, which causes intestinal and extraintestinal amoebiasis in humans. Although there are certain drugs available for treating this disease, more research is needed to come up with novel drugs. Furthermore, basic research is needed to understand the different stages of its life cycle and possible therapies during its dormant stages. One such stage is the encystation of its cell, a crucial stage which basically involves the transformation of motile trophozoite into resistant cyst. It is extremely difficult to treat cysts, and a growing body of research has shown that phospholipase D (PLD) activity is significantly increased in *Entamoeba histolytica* during encystation, indicating that it plays a critical role in cyst formation. PLD is a widely distributed enzyme in all branches of life that hydrolyzes phosphatidylcholine. It is believed to play an important role in the regulation of cell physiology by extracellular signals such as hormones, neurotransmitters, growth factors, and cytokines. Due to our inability to effectively induce encystation *in vitro* in *Entamoeba histolytica*, our knowledge about the cyst form remains limited. This also hampers our ability to develop cyst-specific diagnostic tools. This thesis is about role of PLD during cyst formation in *Entamoeba*. In this study we attempted to investigate a potential role of PLD during encystation in the laboratory model organism *Entamoeba invadens* as encystation is not experimentally achievable in human parasite *Entamoeba histolytica*. The research included amplification of PLD1 and PLD2 genes by extraction of genomic DNA and cloning of the genes into standard cloning vector pGEM-T-Easy and then in the *Entamoeba*-specific vector pEiNEO. Electroporation of the cloned *Entamoeba* specific vectors was performed by transfecting the *Entamoeba invadens* trophozoite cells. On other hand, encystation activity was also checked by introducing 0.6% n-butanol and t-butanol in 50% LYG media (encystation medium) but it did not show any significant inhibition. Fluorescent microscopy results showed no effect of n-butanol on cysts formation. Bioinformatics investigation showed that the HKD active site motif of PLD was found conserved throughout *Entamoeba* species as in animal, plants, and fungi.

CONTENTS

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
Bp	Base pairs
BSA	Bovine serum albumin
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
G418	Geneticin 418 (antibiotic)
Hrs	Hours
IPTG	Isopropyl-beta-D-thiogalactopyranoside
Min	Minutes
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PLD1	Phospholipase D1
PLD2	Phospholipase D2
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Introduction

Entamoeba species are unicellular eukaryotes that parasitize all classes of vertebrates and some invertebrates. In 1969, the WHO expert committee adopted the term "Amoebiasis" which is a life threatening infection of the gastrointestinal tract of human beings and defined it as the state of harboring the pathogenic protozoan *Entamoeba histolytica* with or without clinical signs and symptoms (Organization, 1969). In the past, many people infected with *E. histolytica* developed no symptoms and some infections would clear without any treatment. This fact led to the origination of two hypotheses; one was proposed by Emil Brumpt in 1925 according to which, there are two species, one is virulent and invasive while the other one is morphologically identical but harmless. The second hypothesis stated that *E. histolytica* is a commensal organism of the human colon having potential to turn into a pathogenic organism (Aguirre-Beltran et al., 1999).

Only three species of *Entamoeba* have been proven to cause disease and sometimes death in their hosts: *Entamoeba histolytica*, a parasite of humans, *Entamoeba nuttalli*, a parasite of nonhuman primates, and *Entamoeba invadens*, a parasite of reptiles. In humans, six species of *Entamoeba* have been reported out of which only *Entamoeba histolytica* is pathogenic and the remaining five (*E. dispar, E. moshkovskii* (unclear pathogenicity)*, E. coli, E. hartmanni,* and *E. polecki*) are commensal (Ngui et al., 2012; Gomes et al., 2014).

1.2 *E. invadens* as model organism

E. histolytica is the main species related to diarrhea and enteritis in human beings, *E. invadens* on the other hand is the causative agent of amoebiasis in reptiles, including snakes and lizards, and has similar life cycle to *E. histolytica*. *E. histolytica* is endemic in many countries globally and is thought to be a foremost parasitic cause of human morbidity and mortality. Likewise, *E.*

invadens can be responsible for substantial losses of reptiles in zoological collections. Research has focused on usefulness of *E. invadens* as a model organism for studying the process of *Entamoeba* cyst formation (Mi-ichi et al., 2016; Morf & Singh, 2012).

1.3   Biology of *E. invadens*

E. invadens like *E. histolytica* exists in two forms; an invasive and motile trophozoite and an infective cyst (Naiyer et al., 2019).

- ❖ The **trophozoite** varies in size from 10 to 60 µm, although very large forms of up to 90 µm have also been observed occasionally in case of *E. histolytica*. While *E. invadens* trophozoites size varies from 15-25 µm. Actively moving trophozoites produce broad and finger-like pseudopods and contain a spherical nucleus surrounded by nuclear membrane (Bhattacharya et al., 2000; Ali IK et al., 2012).
- ❖ The **cyst** is round or slightly oval bodies having diameter of 10 to 16 µm in *E. histolytica*. The cyst of *E. invadens* has diameter of 11-20 µm. Initially it contains one nucleus but after maturation, it has four nuclei (Bhattacharya et al., 2000).

1.4   Life cycle of *Entamoeba*

The life cycle of *E. invadens* is identical to *E. histolytica* and has four stages; the trophozoite, pre-cyst, cyst and metacyst (Arredondo JL et al., 2014). Infection with *E. histolytica* begins by ingesting fecally contaminated water or food containing *E. histolytica* cysts. The infective cyst of the parasite survives the stomach and intestinal passages. In the bowel lumen, excystation occurs leading to aggregation of motile and potentially invasive trophozoites in the intestinal mucin layer. Trophozoites can form new cysts leading to self-limited and asymptomatic infection. The trophozoite of *E. histolytica* first converts into pre-cyst form with single nucleus. This form

matures into tetra-nucleated cyst as it moves down and out of the colon as shown in Figure 1 (Leiva, 2009; Brunette et al., 2009; Garmie et al., 2016).

Cysts are usually found in formed stool while trophozoites are found in diarrheal stool. In the large intestine, the trophozoites multiply by binary fission and produce cysts, whereas both stages (trophozoite and cyst) are passed in feces. The cyst can survive days to weeks in the external environment due to the protective wall around them and are responsible for transmission (Papadakis et al., 2016). On the other hand, trophozoites passed in the stool are destroyed rapidly once outside the body and would not survive the gastric environment if ingested. In most cases, trophozoites remain confined to the intestinal lumen of non-invasive infection making individuals asymptomatic carriers and pass cysts in their stool. In some patients, the trophozoites invade the intestinal mucosa (intestinal disease), or, through the bloodstream, extra-intestinal sites such as the liver, brain, and lungs (extra-intestinal disease), with resultant pathologic manifestations (Brunette et al., 2009; Hemphill et al., 2019).

In case of *E. invadens*, the infective cysts are passed in the feces and are stable in the environment. Transmission occurs by ingestion of infective cysts present in the feces. The main infective form is cyst. It releases the motile trophozoites after ingestion by the host. The trophozoites invade the intestine and liver and spread hematogenously through the portal vein. Proteolytic enzymes released by the parasite disrupt intestinal mucosa and the epithelial barrier facilitating tissue penetration (Chia et al., 2009).
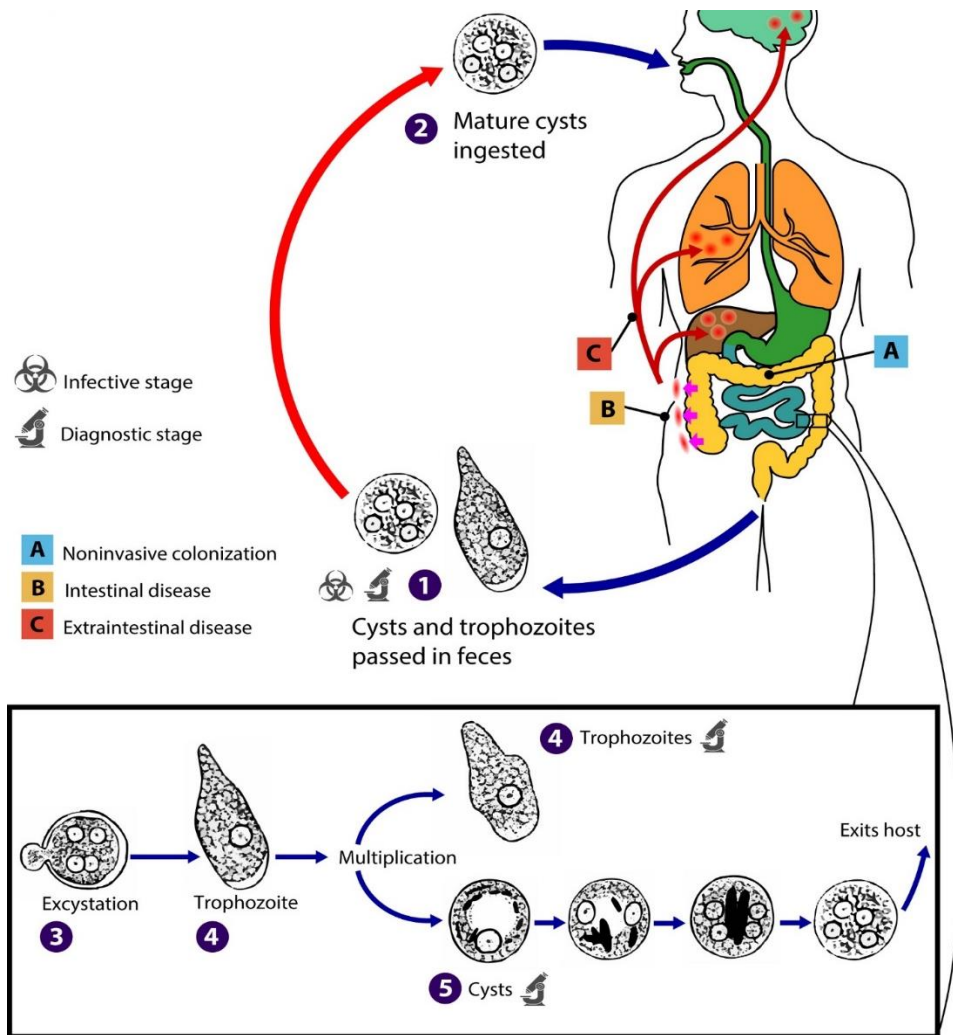
Figure 1. Life cycle of *E. histolytica* (Image taken from CDC, 2010)

Life cycle of *E. histolytica* (CDC, 2010) where cysts and trophozoites are passed in feces. Ingestion of mature cysts from fecally contaminated food, water or hands leads to *Entamoeba histolytica* infection. Trophozoites are released in the small intestine via excystation which migrate to the large intestine. Trophozoites may remain confined to the intestinal lumen (A: noninvasive infection) with individuals continuing to pass cysts in their stool (asymptomatic carriers). Trophozoites can invade the intestinal mucosa (B: intestinal disease), or blood vessels, reaching extra intestinal sites such as the liver, brain, and lungs (C: extra intestinal disease).

Amoebic pathogenesis causes host mucosal barrier depletion, colonic lumen adherence and cytotoxicity and colonic epithelium invasion (Espinosa-Cantellano & Martinez-Palomo, 2000). This parasitic damage leads to colitis and disseminated disease in some cases. The development of disease depends on both host and parasitic genotypes (Garmie, 2016). Parasitic transmission

depends on the host environmental factors. The link between parasite, host, and environmental factors in the virulence of *E. histolytica* has been highlighted by Chelsea Marie and William A Petri. When *E. histolytica* progresses to virulence, the destruction of the colonic environment can lead to degradation of protective mucus, disrupted epithelial barrier function, deregulated ion transport, local and systemic inflammation, impaired nutrient absorption, and disruption of the microbiota (Marie et al., 2014).

Among the known amebic factors involved in pathogenesis are signaling pathways involving heterotrimeric and Ras superfamily G proteins. The current knowledge of roles of heterotrimeric G protein subunits, Ras, Rho and Rab GTPase families in pathogenesis of *E. histolytica* as well as their downstream signaling effectors and nucleotide cycle regulators have been studied (Bosch et al., 2013). Signaling of heterotrimeric G proteins likely regulates amebic motility and attachment to host cell and killing host cell via activation of an RGS-RhoGEFs effector (Bosch et al., 2013). Rho family GTPases and RhoGEFs and Rho Effectors modulate the dynamic actin cytoskeleton of *E. histolytica* and other associated pathogenesis related cellular processes like migration, phagocytosis, invasion and evasion of host responses by surface receptor capping (Labruyère & Guillén, 2006; Meza et al., 2006; Bosch et al., 2013). Multiple roles have been performed by a remarkably large family of 91 Rab GTPases in a complex amoebic vesicular trafficking system which is prerequisite for phagocytosis and pinocytosis and secretion of virulence factors like amoebapores and cysteine proteases (Christy & Petri, 2011; Bosch et al., 2013). A recent study of G proteins signaling in *E. histolytica* has improved the understanding of parasitic pathogenesis although much remains to be discovered. Yet this study highlighted possible targets such as heterotrimeric G protein, Rab GTPase, Rho family GTPases, as well as RhoGEFs and Rho effectors for pharmacological manipulation (Bosch et al., 2013).

1.5 Epidemiology of *E. histolytica*

The prevalence of amoebic infection varies with the level of sanitation and is usually higher in tropics and subtropics as compared to temperate climates (John, 2006). Amoebiasis is the third most common parasitic cause of death after malaria and schistosomiasis (Organization, 1969; Dhanalakshmi & Parija, 2016). The cyst is the transmissive form of *E. histolytica* and one infected person can pass up to 45 million cysts daily (Herman et al., 2017). *E. histolytica* roughly infects 50 million people globally and about 10% of the world population is infected with higher rates occurring in nations where sanitation is poor (Garmie, 2016).

1.6 Disease pathology

Amoebic colitis can result in mucosal thickening to discrete ulcer formation up to necrosis and intestinal wall perforation (Hardin et al., 2007). The process always starts with adherence of the trophozoites to the colonic mucus and epithelium. After adhesion of the amoebae, cytolysis is mediated by the amoebapores, small proteins that are highly effective in lysing lipid bilayers (Serrano-Luna et al., 2013). The attacked host cells undergo apoptosis (Bansal D et al., 2009; Tazreiter, 2010). The third major group of pathogenicity factors are the cysteine proteinases. These enzymes play an important role in tissue invasion, e.g. they were shown to disturb enteric microvilli and thereby mediate a close contact between the amoebae and enteric cells (Ralston & Petri, 2011). After tissue invasion the amoebae enter the host blood stream and there encounter the complement system and host antibodies. These can be cleaved by excreted cysteine proteinases (Nakada-Tsukui & Nozaki, 2016). Interestingly, the strict commensal *Entamoeba dispar* does not express the cysteine proteinase 5 (CP5). The gene encoding CP5 is present, but highly degenerated and therefore not expressed (Wilson IW et al., 2019). Most of the other pathogenicity factors are present in *E. dispar* but are only weakly expressed, e.g. one study showed differences in the

structure and activity of pore forming peptides in *E. dispar* (Espinosa-Cantellano & Martinez-Palomo, 2000; Ximénez et al., 2017). These discrepancies could to some extent explain the differences in pathogenicity of the two closely related species, but more recent studies show that many more factors seem to be involved in *Entamoeba* pathogenicity, for instance protection against oxidative stress by peroxiredoxin (Nakada-Tsukui & Nozaki, 2016; Tazreiter, 2010). Moreover, pathogenicity also depends on the host response during the development of the disease. When trophozoites finally reach the liver, probably via the portal circulation, they cause typical abscesses consisting of areas of dead hepatocytes, liquefied cells, and cellular debris (Prakash & Oliver, 2019; Rigothier et al., 2002). The often surprisingly small number of amoebae relative to the size of the abscess could hint to an ability to elicit apoptosis in hepatocytes without direct contact (Tazreiter, 2010; Ali IK et al., 2012).

1.7 Transmission, diagnosis, and treatment

The foremost factors involved in the transmission of parasitic infections in developing countries are poverty and low educational background (Karan et al., 2012). Defecation near water supplies and indiscriminate squatting are important sources for transmission (Karan et al., 2012; El-Dib, 2017). In addition, unhygienic irrigation practices and fecal contaminated food result in high infection rates mainly in endemic areas (Mohamed et al., 2016; Tharmaratnam et al., 2020). In developed countries, where good sanitation, education and water supplies are available to most members in the community, increased risk of infection is linked to infected asymptomatic food handlers, immigrants and tourists returning from endemic areas (El-Dib, 2017; Shirley et al., 2018). *E. histolytica* infections are diagnosed in the laboratory by detecting parasite in the specimens or by immunologic or molecular technique (Ali IK et al., 2012). Metronidazole is the

first-choise for the treatment of amoebiasis and is recommended by WHO (Ali IK et al., 2012; El-Dib, 2017).

1.8 *Entamoeba* cyst

It has been discussed earlier that amoebiasis is caused by ingesting food or water contaminated with *Entamoeba* cysts. This form of parasite is responsible for disseminating from one host to another thus spreading the disease. In the small intestine, excystation of cysts occurs releasing trophozoites that cause for the symptoms of amoebiasis (Zulfiqar et al., 2019). Most studies on *E. histolytica* infection have focused on the trophozoite stage which is relatively easy to grow in axenic culture media. The information of the cyst stage is limited due to its incapability to replicate and insufficiency of reproducible methods for endorsing trophozoite encystation and cyst excystation *in vitro* (Naiyer et al., 2019). Currently, amoebiasis is a serious health problem globally due to its epidemiological and clinical status. The encystation process, common within many organisms, is defined as the process of differentiation from a reproductive form to a dormant, resistant cyst form. Much of the knowledge gained regarding *E. histolytica* encystation has been assembled using the model organism *E. invadens*, a reptile pathogen (Siegesmund et al., 2011). Three main methods for inducing encystation in *E. invadens* are commonly used: glucose deprivation, osmotic shock, or a combination of the two (Mi-ichi et al., 2016; Siegesmund et al., 2011). These processes produce cysts which are similar to those of *E. histolytica,* having a chemically similar composition of the cyst wall, containing four nuclei and chromatoidal bodies (aggregations of ribosomes) which are resistant to osmotic lysis (Aguilar-Díaz et al., 2010). However, there are noted differences between *E. invadens* and *E. histolytica.* The commonly used *E. invadens* IP1 strain is "super encysting", withstanding up to 100-fold dilutions of growth medium and complete removal of glucose, both of which would kill *E. histolytica* (Herman et al.,

2017; Turner et al., 2007). *E. invadens* is also more distantly related to *E. histolytica* than other *Entamoeba* species when 18S phylogeny is examined (Wilson IW et al., 2019; Das & Ganguly, 2014). However, little work has been conducted on these species and no model for the induction of encystation has been discovered. Despite these differences a lot has been learnt regarding encystation of *Entamoeba* parasites from the *E. invadens* system (Singh M. S. et al., 2015).

Developing strategies to regulate encystation and excystation will lead to novel measures to block amoebiasis transmission by disrupting the life cycle of *Entamoeba* (Aguilar-Díaz et al., 2010; Mi-ichi et al., 2016). Various chemicals that can target molecules contributing to encystation can be effective. Ideal targets are those molecules that are exclusively present in parasite but not in host (Mi-ichi et al., 2016). Luna-Nacar and coworkers reported the *in vitro* production of cyst-like structures (CLS) having similar features like cysts such as reduction in size, round shape having multiple nucleus and chitin wall formation together with overexpression of glucose 6-phosphate isomerase (Luna-Nacar et al., 2016). They extracted the total proteins from trophozoites, CLS, and partly purified cysts obtained from amoebic human patients' feces. In trophozoites, 1029 proteins were identified out of which 539 proteins were trophozoites specific. Similarly, 550 proteins in CLS and 411 in cyst were identified with 299 and 84 proteins unique to each sample, respectively. Only 74 proteins were shared by all three stages. They observed that almost 70% of proteins of CLS were shared with trophozoites although the relative abundance is different. Proteins linked to metabolically active cell were found abundant in trophozoites while proteins related to proteolysis, stress response and redox homeostasis were highly expressed in CLS (Luna-Nacar et al., 2016). Unlike trophozoites and CLS, proteomic analysis of cysts presented 40% of hypothetical proteins. Luna-Nacar et al concluded that the nature of CLS is more like trophozoites as compared to cyst and they could be produced as a rapid survival response of trophozoites to a stressful condition.

These CLS allow the parasite to endure and live temporary in chitin-like resistant covering having Jacob proteins (cyst wall-specific glycoprotein) (Frisardi et al., 2000; Luna-Nacar et al., 2016). These findings suggested that CLS and cyst formation could be different stress responses. Also, it has been found that cysts showed higher expression of proteins with unknown function (Luna-Nacar et al., 2016).

Various eukaryotic parasites undergo cyst formation that transmits infection. This process is found in several organisms such as *Giardia, Entamoeba, Cryptosporidium, Toxoplasma, Naegleria*, and several nematodes and tapeworms. For enteric protozoa, the transmissible stage is the cyst which allows survival outside the host cell. Understanding the molecular processes controlling stage conversion is crucial for the development of transmission-blocking therapies as well as novel diagnostics (Ali IK et al., 2012; Ehrenkaufer G. M. et al., 2007). Thus, disease transmission can be controlled by blocking the stage conversion from trophozoite to cyst. The knowledge about cyst form of *Entamoeba histolytica* is limited due to inability to produce cyst *in vitro*. This also hindered the development of cyst specific diagnostic tools. To identify proteins in cyst samples and to identify candidate proteins to develop cyst specific diagnostic tools, Ali et al. (2012) purified cysts from stool of infected persons and proteins were identified using LC-MS/MS mass spectrometer. A total of 417 proteins were identified out of which 195 proteins were never detected in trophozoite-derived proteomes or EST datasets, suggesting these proteins are cyst specific. Jacob, Jessie and chitinase, which are cyst wall specific glycoproteins (Chatterjee et al., 2009), were positively identified. Antibodies against Jacob identified cysts in fecal specimens and have potential value as a diagnostic reagent (Ali IK et al., 2012). Some protein kinases such as small GTPase signaling molecules, epigenetic regulators, DNA repair proteins and surface associated proteins were also identified. Proteins identified in this study are probably the most abundant in

excreted cysts and consequently show potential as diagnostic targets (Ali IK et al., 2012). This study provided the first data for naturally occurring *E. histolytica* cysts and provided the vital insights into the infectious cyst form. Moreover, several unique proteins were identified which will assist the development of novel diagnostic markers for *E. histolytica* (Ali IK et al., 2012).

1.9 Change in gene expression during encystation and excystation

During early encystation, various gene families which are involved in signal transduction were found up-regulated (Ehrenkaufer G. M. et al., 2013). These proteins include signaling molecules like protein kinases, GTPase activating proteins and lipid signaling. These proteins may be involved in transducing and affecting the signals that initiate encystation (Ehrenkaufer G. M. et al., 2007; Ehrenkaufer G. M. et al., 2013). A cyst wall is formed from the secretion of lectins and chitin although the underlying processes of membrane trafficking which support encystation is still unclear. Evaluation of expression of membrane trafficking genes during encystation in *E. invadens* identified the up-regulation of genes involved in secretion during encystation (Ehrenkaufer G. M. et al., 2013). Moreover, in mature cyst, endocytic and phagocytic genes showed higher expression suggesting that mRNA and other proteins are preserved in cysts in preparation for excystation. It was found that two distinct dynamin related proteins in *E. histolytica* which are highly expressed during cyst formation (Herman et al., 2017). Phylogenetic analysis of these dynamin proteins revealed that they are paralogous but quite different from human dynamins making it attractive target for drugs to block cyst formation (Herman et al., 2017).

Phospholipase D (PLD) is strongly up-regulated early in encystation and belongs to the class of hydrolases (Dennis, 2015; Ehrenkaufer G. M. et al., 2013). Hydrolases initiate numerous digestive and physiological processes. These enzymes consume a water molecule to cut down all four biological molecules i.e. carbohydrates, proteins, fats or lipids and nucleic acid (Dennis, 2015).

Phospholipases are involved in the catalysis of phospholipids into fatty acids and lipophilic substances. These enzymes cleave ester bonds within phospholipids (Dennis, 2015; Bandana et al., 2018).

Phospholipases are primary digestive enzymes playing important role in various physiological processes like signal transduction, lipid mediator production, metabolite digestion and aggregation of snake venom (Aloulou et al., 2018). Phospholipids are classified into four major classes which are distinguished based on type of reaction they catalyze.

- ❖ **Phospholipase A (PLA)** is further divided into PLA1 and PLA2 according to the specific hydrolysis site within the phospholipid molecule (Bandana et al., 2018). The phospholipase A hydrolyzes the carboxylic ester at the sn-1 site of the glycerol backbone is termed phospholipase A1 while phospholipase A2 hydrolyze the sn-2 position. PLA most notably is found in mammalian cells such as plasma of rat livers and bovine brains, and present in metazoan parasites, protozoan parasites, and snake venom (Selvy et al., 2011).

- ❖ **Phospholipase B (PLB)** is also termed as lysophospholipase, is an enzyme having activities of both PLA1 and PLA2, that is, it can hydrolyze the acyl-chain from both sn-1 and sn-2 positions of a phospholipid. PLB is widely distributed in microorganisms, mammalian tissues and venom (Selvy et al., 2011).

- ❖ **Phospholipase C (PLC)** is a phosphodiesterase responsible for hydrolysis of the glycerophosphate bond. This bond is responsible for linking the polar head group to a glycerol backbone. PLC is found in mammal and bacteria (Kadamur & Ross, 2013; Selvy et al., 2011).

- ❖ **Phospholipase D (PLD)** is responsible for the cleavage of the phosphate releasing phosphatidic acid and alcohol. PLD is also considered phosphodiesterase. PLD widely

present in bacteria, yeast, animals, plants, and viruses (Haas & Stanley, 2007; Selvy et al., 2011); Two main isoforms of phospholipase D have been found in mammalian cells which include PLD1 and PLD2, each encoded by distinct genes. PLD1 is a 120 kDa protein mainly located on the inner membranes of cells. It is mostly present at the Golgi complex, endosomes, lysosomes, and secretory granules. On the other hand, PLD2 is a 106 kDa protein that primarily localizes to the plasma membrane, residing in light membrane rafts (Jenkins & Frohman, 2005)

It has been observed that during encystation and in mature cysts, basic metabolic processes such as glycolysis were significantly decreased, and glucose metabolism is redirected to cell wall synthesis (Ehrenkaufer G. M. et al., 2013; Naiyer et al., 2019). Similarly, numerous virulence factors were down-regulated in mature cyst which is expected as this stage does not cause disease symptoms in the host. Moreover, during excystation, the down-regulation of carbohydrates metabolism observed in the mature cyst is reversed with increased transcription levels of glycoside hydrolases as well as hexokinases (Ehrenkaufer G. M. et al., 2013).

1.10 Structure of PLD

PLD is a widely distributed enzyme involved in the fine regulation of metabolism of cells with different organization (Leiros et al., 2000). This enzyme is necessary for growth and development of animals and plants, especially under conditions of stress and illumination. Apart from phospholipid metabolism, PLD enzymes are involved in vesicle formation, protein transport, signal transduction and mitosis. One of the products of the hydrolytic action of PLD, phosphatidic acid (PA) is a putative second messenger (Li et al., 2020). Under certain circumstances, the formation of phosphatidic acid is also suggested to cause changes in lipid bilayer properties that would facilitate fusion events and vesicle budding in the course of intracellular membrane traffic

(Kooijman et al., 2005; Leiros et al., 2000). Lipids can modify PLD in plant and animal cells. In particular, some phospholipids usually act not as substrates but as cofactors of the enzyme (Kolesnikov et al., 2012). The and three-dimensional structures of PLD from *Streptomyces sp.* strain PMF was determined by the X-ray crystallography and structure was solved by the multiwavelength anomalous dispersion (MAD) phasing method as shown in Figure 2 (Leiros et al., 2000).
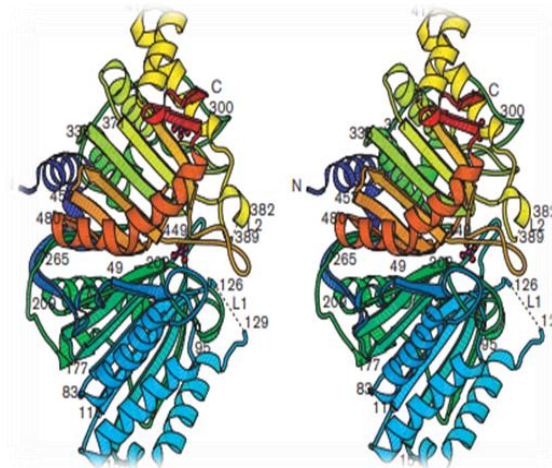
Figure 2. Stereographic illustration of the crystal structure of phospholipase D (PLD) from *Streptomyces* sp. Overall view of 3D-structure.

The structure is colored from dark blue to red according to the amino acid sequence. Two flexible loops (dashed and labeled L1 and L2, accordingly) are the part of the enzyme that docks onto the membrane when hydrolysis takes place, and two phosphate ions (red) are found on the surface of the protein (Leiros et al., 2000).

PLD from *Streptomyces sp.* strain PMF consists of a single polypeptide chain that is folded into two domains. An active site is located at the interface between these domains (Leiros et al., 2000). The structure includes 35 secondary structure elements forming two β sheets, with nine and eight β strands, respectively. There are 18 α helices that flank the two β sheets; two of these have a

somewhat irregular orientation, running diagonally on one side of each of the β sheets. An close-up view from of PLD structure at the outer membrane view is shown in Figure 3.
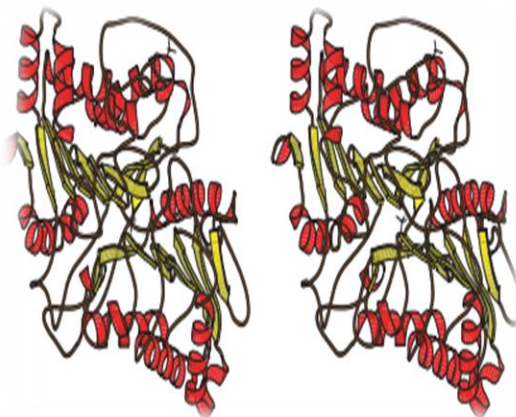
Figure 3. The stereographic illustration of the
PLD in *Streptomyces* sp PMF. The view from the
outer membrane.

Alpha helices are shown in red and β strands in
yellow. (Leiros, 2000).

The two β sheets are sandwiched between surrounding α helices, with the more usual orientation of the helices on the water-exposed side of each β sheet. The α helices α8 and α16–α17 lie between the β sheets and are rotated with respect to the general strand direction. The overall structure of PLD from *Streptomyces sp.* (strain PMF) is a α-β-α-β-α-sandwich (Leiros et al., 2000; Kolesnikov et al., 2012). A 1.8 Å resolution crystal structure of the human PLD1 catalytic domain was presented by Bowling et al (2020). It is characterized by a globular fold with hydrophobic cavity of funnel shaped leading to the active site. Adjacent is a PIP2-binding polybasic pocket at the membrane interface that is crucial for activity. The C-terminus folds into and adds part of the catalytic pocket, which form a phosphohistidine that imitate an intermediate stage of the catalytic cycle. Mapping of PLD1 mutations that interrupt RhoA activation identifies the RhoA-PLD1

binding interface. This structure sheds light on PLD1 regulation by lipid and protein effectors, allowing inhibitor design for this well-studied therapeutic target (Bowling et al., 2020).

1.11 Role of PLD in encystation

Genes encoding PLD were amongst the ones up-regulated during encystation in *E. invadens* (Naiyer et al., 2019; Ehrenkaufer G. M. et al., 2013). Its activities have been observed in simple to complex organisms and has been linked to several important biological processes such as vesicle transport and transduction of signals which are essential for cell shape changes and proliferation. In *E. histolytica*, PLD might be important for efficient cyst formation. Its regulatory role in encystation was examined in *E. invadens* by observing changes in PLD activity during development (Ehrenkaufer G. M. et al., 2013). Its level was increased during early stage of cyst formation and falling back later in encystation supporting its active role in encystation. However, RNA sequencing did not show peak RNA levels of PLD genes indicating that its activity is being regulated at protein level (Ehrenkaufer G. M. et al., 2013). Furthermore, inhibition of PLD activity was tested by suppressing the formation of phosphatidic acid by adding 0.6% n-butanol which serves as a non-productive substrate for PLD. Marked reduction of encystation efficiency was observed (Ehrenkaufer G. M. et al., 2013).

Most of the genes expressed in cysts encode hypothetical proteins i.e. without any known function (Ehrenkaufer G. M. et al., 2013; Ali IK et al., 2012). These hypothetical proteins were also found up-regulated during encystation (Ehrenkaufer G. M. et al., 2013). I screened proteins from these studies that are up-regulated during encystation and which received very little or no work (Table 1). PLD is one of the proteins which is found to be up-regulated during encystation and very little work has been done on it.

Table 1. Proteins up regulated during encystation in *E. invadens* with little (one or two papers) or no published literature

Gene ID	Common Name	Identity percentage with *Homo sapiens*	Identity percentage with *E. histolytica*
EIN_059870	chitotriosidase-1 precursor, putative	30.06%	40.27
EIN_186850	Embryonic protein DC-8, putative	33.04%	83.61%
EIN_313910	DTDP-glucose 4,6-dehydratase, putative	47.58%	84.37%
EIN_147990	vacuolar amino acid transporter, putative	23.94%	71.76%
EIN_186080	transmembrane protein R144.6, putative	26.89%	27.80%
EIN_059470	M-phase inducer phosphatase, putative	no match	40.00%
EIN_136750	glucosamine--fructose-6-phosphate aminotransferase [isomerizing], putative	32.33%	80.00%
EIN_058910	pyruvate-flavodoxin oxidoreductase, putative	no match	35.42%
EIN_053310	chitotriosidase-1 precursor, putative	36.57%	58.56%
EIN_289620	carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase, putative	no match	40.00%
EIN_160630	D-alanine aminotransferase, putative	no match	51.27%
EIN_135440	protein TRANSPARENT TESTA GLABRA, putative	34.13%	62.38%
EIN_107320	A/G-specific adenine DNA glycosylase, putative	33.47%	60.97%
EIN_168100	ER lumen protein retaining receptor, putative	no match	30.81%
EIN_104460	G2/mitotic-specific cyclin-B, putative	30.52%	41.82%
EIN_020180	ankyrin repeat-containing protein, putative	28.32%	42.79%
EIN_024440	actophorin, putative	27.78%	68.12%
EIN_034650	N-acetyltransferase ECO1, putative	21.79%	36.82%
EIN_107750	ubiquitination factor E4, putative	26.01%	51.06%
EIN_037710	protein NUF1, putative	36.54%	42.11%
EIN_243030	inhibitor of apoptosis 1, diap1, putative	47.06%	42.80%
EIN_111270	zinc finger protein DHHC domain containing protein, putative	29.71%	49.51%
EIN_196190	transcription initiation factor IIB, putative	no match	52.84%
EIN_080170	metal resistance protein ycf1, putative	34.93%	36.51%
EIN_161520	proline synthetase associated protein, putative	40.09%	57.96%
EIN_311760	chaperone protein DNAJ, putative	no match	33.99%
EIN_058900	branched-chain-amino-acid aminotransferase, mitochondrial precursor, putative	47.27%	no match
EIN_181450	O-acyltransferase (membrane bound) domain containing protein, putative	27.50%	72.54%
EIN_160910	pyruvate-flavodoxin oxidoreductase, putative	no match	35.45%
EIN_134880	long-chain-fatty-acid--CoA ligase, putative	32.23%	66.25%
EIN_289150	serine palmitoyltransferase, putative	42.39%	69.86%
EIN_111460	protein phosphatase 1E, putative	33.33%	45.50%
EIN_218620	NAD-dependent deacetylase, putative	31.66%	71.64%
EIN_147760	phosphoglycerate mutase, putative	no match	59.95%
EIN_249340	meiotic recombination protein dmc1, putative	60.49%	88.47%
EIN_133760	rab7, putative	47.92%	79.70%
EIN_026170	Phospholipase D, putative	28.95%	59.06%

1.12 Aims of this study

Knowledge about the *Entamoeba histolytica* cyst form is limited due to the inability to produce cysts *in vitro*, this also hinders the development of cyst specific diagnostic tools. Various proteins have been found to be up-regulated during encystation, targeting these proteins can help in developing targeted interventions to disrupt transmission.

The aim of the study was to investigate PLD1 and PLD2 genes which are up-regulated during encystation, with a focus on unique factors that are absent in humans which therefore may provide targeted approaches to interrupt cyst formation and transmission. The role of PLD in cyst

formation was investigated, as it is found to be up-regulated during cyst formation and to analyzed whether perturbation of PLD pathway inhibits stage conversion in *E. histolytica*. To achieve this aim, few objectives were made, which includes amplification of PLD1 and PLD2 from genomic DNA of *E. invadens*, construction of *E. invadens* specific vector with amplified PLD1 and PLD2 genes, then transfection of these cloned vectors into *E. invadens* trophozoite cells. Furthermore, examination of the cyst's formation activity by addition of n-butanol and t-butanol.

2   MATERIAL AND METHODS – BIOINFORMATICS ANALYSIS

In silico approaches were used for analysis of PLD protein in different ways.

2.1   Sequence retrieval of PLD

The FASTA format sequence of PLD protein were retrieved from the National Centre for Biotechnology Information (NCBI) database (NCBI, n.d.). These sequences were selected from different organisms such as plants, animals, fungi, *Entamoeba,* and protists as shown in Table 2 with accession number, organism name and group. These sequences were chosen for further analysis.

Table 2. List of sequence retrieval of PLD

Organisms	Accession
Animals	
Diachasma alloeum	XP_015117785.1
Nematostella vectensis	XP_032241257.1
Anolis carolinensis	XP_008123686.1
Erpetoichthys calabaricus	XP_028656955.1
Plants	
Zea mays	BAA11135.1
Oryza sativa Japonica	BAA19467.1
Arabidopsis thaliana	AAC49656.2
Entamoeba's	
Entamoeba nuttalli P19	EKE42518.1
Entamoeba histolytica HM-1: IMSS	XP_651826.1
Entamoeba dispar SAW760	XP_001735606.1
Entamoeba invadens IP1	XP_004257539.1
Fungi	
Fusarium oxysporum NRRL 32931	EWY91099.1
Moesziomyces antarcticus	XP_014659599.1

2.2 Identification of conserved sequence

Conserved sequences are those regions in a sequence which are the backbone across different species. The FASTA format sequences were subjected one by one to Conserved Domain Database (NCBI, Conserved Domain Database, n.d.) for the identification highly conserved region across the organisms in PLD protein.

2.3 Multiple sequence alignment and phylogenetic tree construction

The sequences were imported into BioEdit software for alignment of PLD protein. The revolutionary relationship among organisms is represented by a diagram which is called phylogenetic tree. It depicts the evolution of a set of taxa from their most recent common ancestor The phylogenetic analysis was carried out by constructing a tree using the WAG Substitution model that includes maximum likelihood approach with phylogeny.fr online tool (Methods of Algorithms in bio-informatique, n.d.).

2.4 Proteomics study of PLD

2.4.1 Homology modelling

Protein modelling or homology modelling is a method to compare structurally unknown protein structure with a structurally known similar protein with above 30% identity. The structurally similar proteins were searched on protein database such as PDB (Protein Database, n.d.). This method is also known as comparative modelling as the protein sequences were compared with known template structures (Krieger et al., 2003).

2.4.2 Superimposition

Super imposition of protein structure is important factor for studying its characteristics. Superimposition is to overlay the image (template) over another image (model) to check whether it stays or differ from the template. In PLD1 case 6ohr.1 chain A (human Phospholipase D1 catalytic domain) was been selected as template and for PLD2 6ohm.1 (human Phospholipase D2 catalytic domain) chain A were taken from protein data bank, furthermore, the constructed model was superimposed over the template model by using Chimera tool for analysis (Chimera molecular modeling system, n.d.). On the other hand, *E. histolytica* PLD sequence was taken from NCBI and protein structure was created by SWISS model (SWISS-model, n.d.). This structure was then superimposed on the *E. invadens* model and structure similarities differences were analysed.

2.4.3   3D model of PLD

Every sequence is determining the structure of protein and the structure determine the function of every protein. The 3D structure of our targeted protein PLD was not present in protein data bank (PDB) that is why we apply the homology modelling approach to predict the 3D structure of PLD. We used online server Swiss model (https://swissmodel.expasy.org/) for homology model. 6ohr.1.A was used as a template with 54% identity and maximum coverage.

2.4.4   Validation of model (Ramachandran plot)

Ramachandran plot was proposed in 1963 by G.N. Ramachandran a biophysicist. To validate the constructed PLD model is legitimate, Ramachandran plot was produced, that determines which torsional angles are acceptable and can achieve perceptiveness into the protein peptide structures.

3   MATERIAL AND METHODS – MOLECULAR WORK

Molecular biology work overview demonstrated as follows: Initially *E. invadens* trophozoite cells were grown in LYI-S media for 14 days at room temperature (23 °C). Then the culture was ready for genomic DNA (gDNA) isolation, which was performed by CTAB protocol specific to *Entamoeba*. After gDNA isolation, PCR amplification was done with PLD 1 and PLD2 specific primers attached with restriction sites and HA-tag. Amplified products were ligated with pGEM-T-Easy (standard vector) and transformed into *E. coli* JM109. After white and blue screening mini cultures were performed on cloned colonies. Cloned plasmids were analyzed by restriction digestion and sanger sequencing. These steps are illustrated in Figure 4.
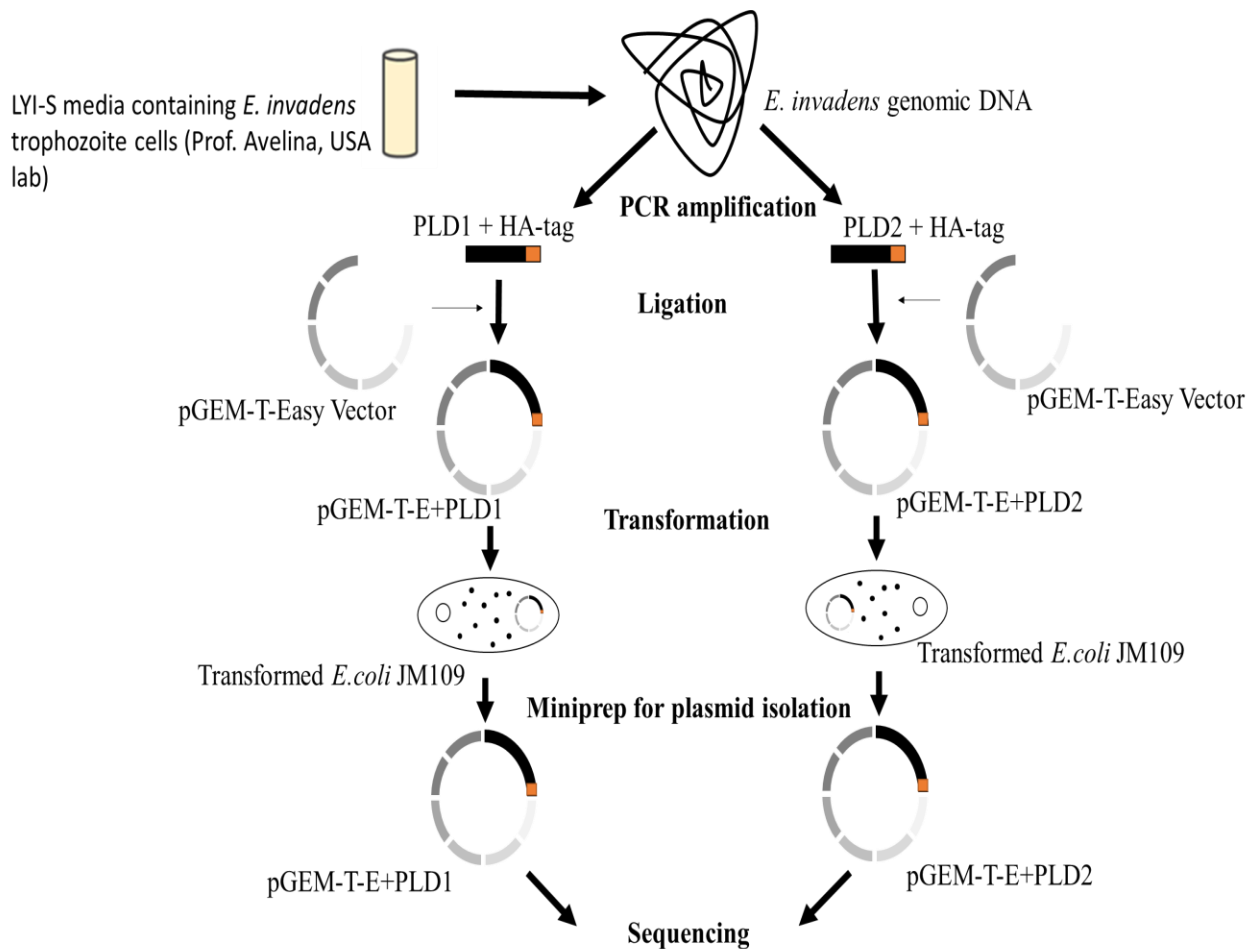

Figure 4. gDNA isolation, PCR amplification & cloning into pGEM-T-Easy standard vector.

After confirmation of PLD1 and PLD2 sequences, these genes were cloned to *E. invadens* specific vector pEiNEO which was firstly digested by *Bam*HI and *Kpn*I restriction enzymes for removal of LUC gene from the vector to ligate PLD1 and PLD2 into it. After ligation and transformation to *E. coli* JM109, different colonies were chosen, and mini cultures were prepared. Mini cultures led to isolation of cloned plasmids which were PLD1+pEiNEO and PLD2+pEiNEO. Before transfection, to ensure the presence of gene of interests (PLD1 and PLD2), these cloned plasmids were confirmed by restriction double digestion. *E. invadens* trophozoite cells were then transfected by successful cloned plasmids by use of electroporation technique. These steps layout can be seen in Figure 5. For screening the transfected *E. invadens* cells, G418 was added, respectively.

LUC (1.6kbp)

pEiNEO-LUC vector (8.1 kbp)
(Kindly provided by Prof. Sudha, India)

pEiNEO-LUC vector digested with *kpn*I and *Bam*HI to remove LUC for insertion of desired genes.

LUC (1.6kbp)

pEiNEO-LUC vector (8.1 kbp)

pEiNEO Vector (6.5 kbp)

pEiNEO Vector (6.5 kbp)

PLD1 + HA-tag

Ligation

PLD2 + HA-tag

Cloned pEiNEO+PLD1

Cloned pEiNEO+PLD2

Transformation

Transformed *E.coli* JM109

Transformed *E.coli* JM109

Miniprep for plasmid isolation

Cloned pEiNEO+PLD2

Cloned pEiNEO+PLD2

Transfection

Transfected *E. invadens* trophosite cell with cloned pEiNEO+PLD1

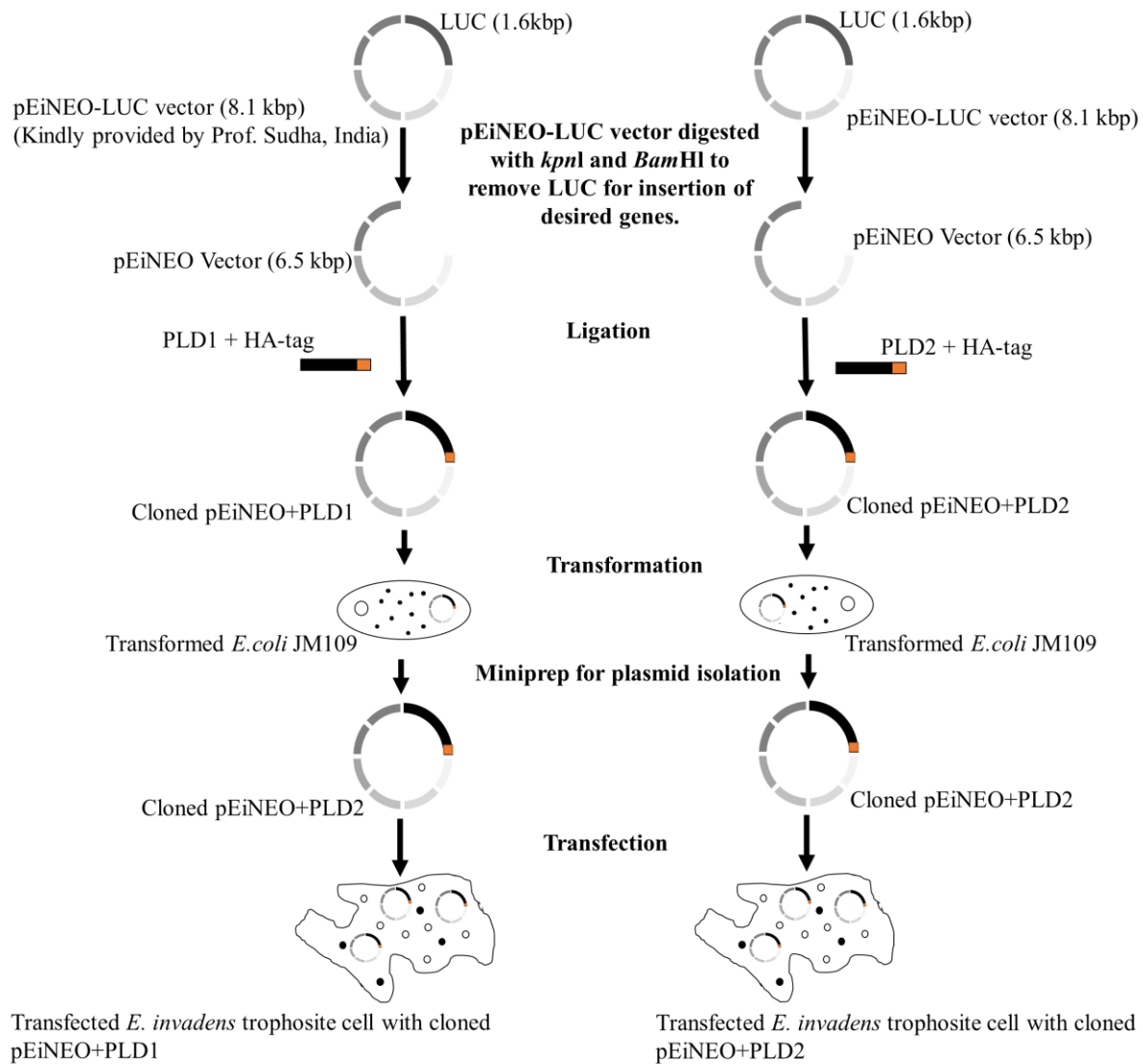Transfected *E. invadens* trophosite cell with cloned pEiNEO+PLD2

Figure 5. Cloning into pEiNEO (*E. invadens* vector)

Second part of the molecular work was to check the activity of n-butanol and t-butanol in cyst formation. *E. invadens* trophozoite cells were transferred to 50% LYG media wild type, 50% LYG media containing 0.6% n-butanol and 50% LYG media containing 0.6% t-butanol. After 24 hrs, 48 hrs, and 72 hrs, cell samples were taken out from the respectively medium and were stained with Calcofluor white stain with 4% paraformaldehyde fixation at the end, just before 72 hrs 0.05% sarkosyl was added to remove remaining trophozoite cells. These stain cells were then observed under 63x objective lens of Leica TCS SP8 FALCON confocal microscope for chitin formation in cyst. These steps are shown in Figure 6.
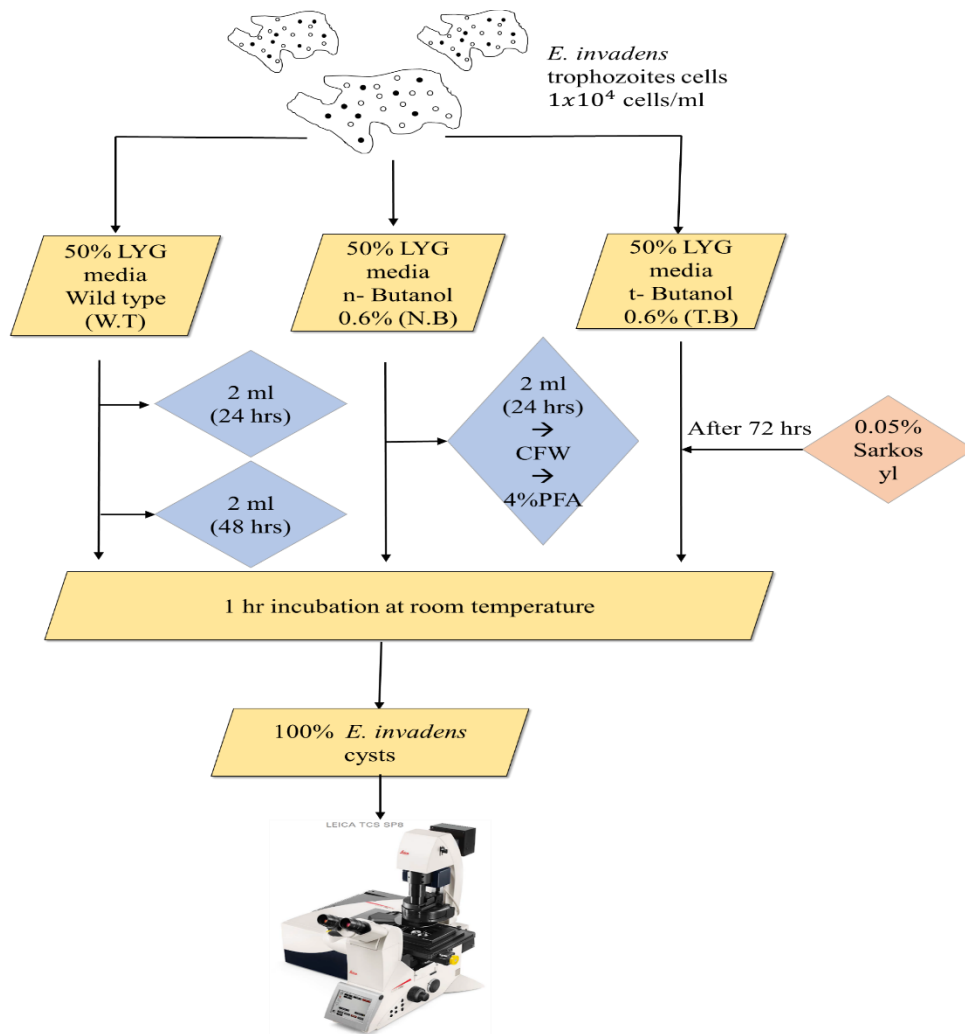

Figure 6. Cysts formation activity in three different 50% LYG media such as wild type, n-butanol & t-butanol

3.1 LYI-S media for culturing *Entamoeba invadens*

Table 3. LYI-S medium composition

Component's	Quantity
K_2HPO_4	1 g
KH_2PO_4	0.6 g
NaCl	2 g
Yeast extract	25 g
Liver digest neutralized	5 g
Glucose	10 g
Cysteine	1 g
Ascorbic acid	0.2 g
Ferric ammonium citrate	0.045 g/ 2 ml
Sodium hydroxide	1 M to adjust pH 6.8
Distilled water	880 ml (final volume)

All the components in Table 3 were dissolved and aliquoted 88 ml in 10 borosilicate glass of 250 ml followed by autoclaving at 121 °C for 15 mins. Afterwards, media was stored at -20 °C.

Frozen adult bovin serum (biovit) was heat inactivated by letting the tube in 56 °C water bath for 30 mins. Before using the media, in one of the 88 ml media flask, 15 ml heat inactivated adult bovine serum was added along with 2 ml vitamin Mixture#18 as described by Clark and Diamond (DIAMOND & CLARK, 1993). Further media was distributed in small glass tubes.

3.2 Culturing of *E. invadens*

Entamoeba invadens IP1 (kindly provided by Prof. Avelina Espinosa, Roger Williams University, USA) trophozoites cells were grown axenically in LYI-S media containing heat inactivated adult bovine serum and vitamin mixture #18 (recipe shown in Appendix A). 0.5 ml of cultured cells were inoculated in 15 ml of media in glass tube. They were incubated at 23 °C at upright position for two weeks.

3.3 Trophozoite cell culture storage

1 ml of *E. invadens* cell culture was poured in 2 ml Eppendorf tube and centrifuged for 1 min at 500 x g. Supernatant was discarded and pellet was resuspended in 1 ml sterile water. Resuspension was performed for 15 mins and later the cells were kept at -20 °C for further use.

3.4 Recombinant DNA technology steps

Recombinant DNA technology is a technique of joining of DNA molecules from two different species that are inserted into a host organism to produce new genetic combinations that are of value to science, medicine, agriculture, and industry. Recombinant DNA technology comprises the following steps.

- Isolation of genetic material
- Amplification using PCR
- Ligation of DNA molecules
- Insertion of recombinant DNA into host
- Enzyme restriction digestion
- Isolation of recombinant cells

3.4.1 Growing *E. coli* JM109

JM109 is a K strain that is recA⁻ and endA⁻ to minimize recombination and improve the quality of plasmid DNA. In addition, the cells carry the F´ episome, which allows blue/white screening. *E. coli* JM109 frozen stock was kindly provided by Prof. Cathrine Lilo's lab and inoculated on fresh LB agar plates. Plates were left at 37 °C incubator for overnight.

3.4.2 Preparation of competent cells

Competency is defined as the ability of the cell to take up DNA. Competency can be induced by chemical method by using divalent or multivalent cations, for example, calcium, magnesium, manganese etc. These cations cause alternation in the permeability of the membranes (outer, inner membrane and cell wall), allowing the DNA to cross the cell envelop. Treatment of *E. coli* cells with $CaCl_2$ at low temperature is the most used method to allow the cells to uptake naked foreign DNA. The selected colony from LB agar plates plated with *E. coli* JM109 was incubated on 5 ml LB broth media at 37 °C overnight on 240 rpm shaker. Later, 500 µl from the overnight culture was added into the 50 ml LB media in Erlenmeyer flask. The culture was incubated at 37 °C shaker

for about 2-3 hours until the OD_{600} reached between 0.4 and 0.5. Afterwards, the sample distributed equally into two centrifuge tubes with labels. Cooled on ice and centrifuged for 10 min at 2509 x g (4 °C). After discarding the supernatant, the pellet was treated with 5 ml ice-cold 0.1 M $MgCl_2$ followed by spinning and the supernatant was discarded. Before mixing the sample in a new tube, 5 ml ice-cold 0.1 M $CaCl_2$ was added into each centrifuge tube and again spined down. The cells, without the supernatant, were resuspended in 500 µl MOPS glycerol and 100 µl competent cells were aliquoted in Eppendorf tubes and stored at -70 °C until use.

3.4.3 Transformation

Bacterial transformation is the direct uptake of exogenous DNA resulting in the acquisition of new genetic traits. Transformation of bacterial cell is mostly carried out by $CaCl_2$ heat shocked method. The $CaCl_2$ neutralizes the charges on DNA and bacterial outer layer and cold-heat shock helps to modify the fluidity of membrane and allow the DNA molecule to enter the cell. 40 ml LB agar containing 40 µl of ampicillin antibiotic was poured in two petri plates. On other hand, 5 µl of pEiNEO-LUC vector (grateful to Prof. Sudha Bhattacharya, Jawaharlal Nehru University, India. for providing *E. invadens* specific vector) was drawn into 100 µl of *E. coli* JM109 competent cell. Incubated on ice for 1.5 hrs followed by heat shock at 42 °C for 1 min and again on ice for 5 mins. Afterwards 750 µl of LB media was poured over same Eppendorf tube. The tunes were incubated for 1 hour at 37 °C on shaker. Media was then spread on the LB agar plates with 100µg/ml ampicillin, which was then incubated for overnight at 37 °C.

3.4.4 Plasmid isolation

Plasmids are very beneficial for bacterial cells because they contain certain genes which allow bacteria to survive in stressful environment. As plasmid is present inside the cell along with other cell components such as genomic DNA, RNA, and proteins, therefore these molecules must be removed in order to isolate pure plasmid. The purification or extraction of plasmid has three basic steps; lysis of the cell followed by neutralization of the lysed cell for the renaturation of plasmid DNA and precipitation of plasmid. Alkaline lysis method (Sambrook & Russell, 2006) is the most widely employed method for plasmid isolation and many commercially available kits are also based on this method. Transformed colony of *E. coli* JM109 containing pEiNEO-LUC plasmid was grown in LB media with 100 µg/ml ampicillin. At their stationary phase i.e after 16 hrs of cultivation, the cells were harvested by centrifugation at 14000 x g for 1 min. Plasmid DNA

purification was followed by manufacturer's instruction in Nucleospin® Plasmid kit (Macherey-Nagel company). Plasmid were stored at -20 °C for further use.

3.4.5 Vectors for cloning

Two separate vectors were used in this research project. One was standard vector pGEM-T-Easy vector and second *Entamoeba* specific vector pEiNEO-LUC.

A: pGEM-T-Easy vector

This is a linearized 3 kbp size, commercially available vector (Figure 7) that contains 3´-T overhangs at the insertion site for compatible overhang towards PCR products. Universal T7 and SP6 RNA polymerase promoters are present within the α-peptide coding region for β-galactosidase, which helps in identification of Blue/White screening on plates by insertional inactivation of these peptides.
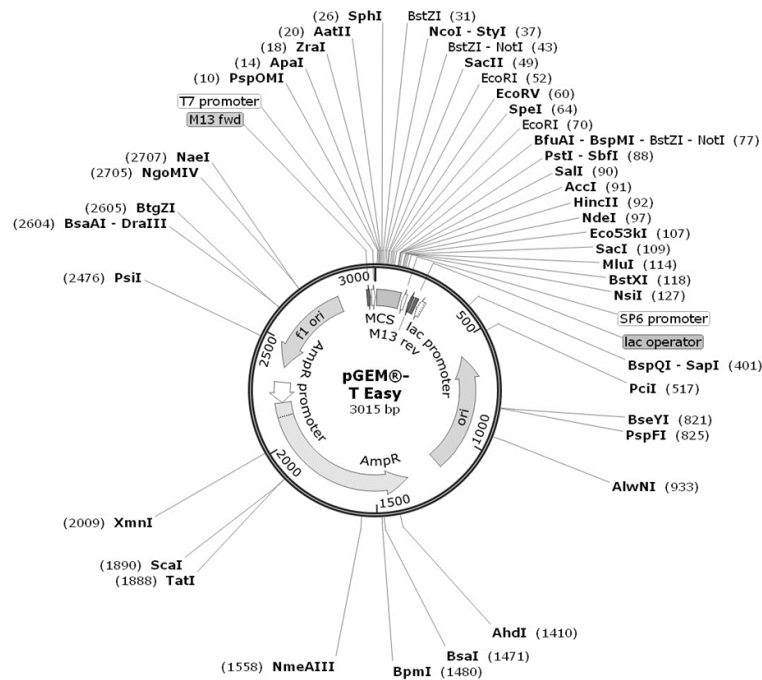

Figure 7. Commercially available Promega pGEM-T-Easy vector map Taken from SnapGene.

Showing different restriction sites, multiple cloning sites for PCR products, having ampicillin resistance gene and universal T7, SP6 promoter sites.

B: pEiNEO-LUC vector

This is a non-commercial *E. invadens* vector 8.1 kbp size, modified from pBlueScript and kindly provided by Prof. Sudha Bhattacharya (Jawaharlal Nehru University, India) (Singh N. O. et al., 2012). This vector contains S10 promotor for *E. invadens* ribosomal protein represented as neomycin phosphotransferase gene (NEO) of approximately 0.8 kbp, L3 promotor which controls luciferase reporter gene (LUC) of 1.6 kbp and having two restriction sites of *Kpn*I and *Bam*HI (Figure 8). The gene of interest (PLD1 and PLD2 HA-tagged) to be inserted were added to the vector at LUC gene site, which was removed by using restriction digestion with *Kpn*I and *Bam*HI.
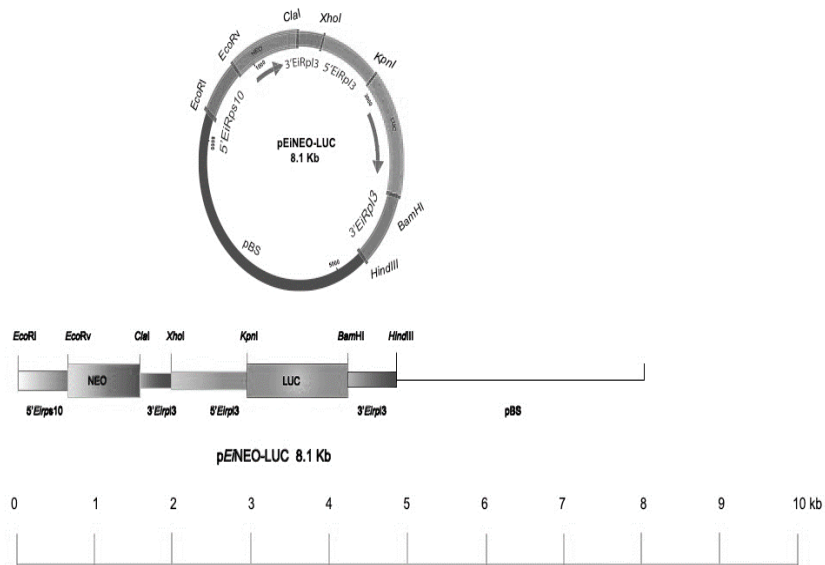

Figure 8. *E. invadens* specific vector pEiNEO-LUC

Figure taken from Dr Maulood Turfah University of Exeter Phd thesis. This illustration based on (Singh N. O. et al., 2012).

3.4.6   pEiNEO-LUC restriction digestion

As mentioned in previous section (3.4.5), the pEiNEO-LUC vector contains luciferase gene which is around 1.6 kbps and was needed to be removed for insertion of desired gene (PLD1 and PLD2). For this purpose, pEiNEO-LUC vector was digested by two different restriction enzymes known as *Bam*HI and *Kpn*I. Reaction mixture for the double digestion was prepared according to the following Table 4, which was then incubated at 37 °C for 2 hrs. Afterwards, the reaction mixture was ran on 1% agarose gel using TAE buffer for 40 mins at 90 V. The gel was observed under UV

radiation and 6.1 kbp band (plasmid pEINEO without LUC gene) cut out the gel. The same Qiagen Gel cleanup Kit and its instructions were used for removal of other impurities.

Table 4. Reaction mixture for the double digestion of pEiNEO-LUC plasmid

Reagents	Volume (µl)
1X fast digest buffer	2
Fast digest *Bam*HI	1
Fast digest *Kpn*I	1
pEiNEO-LUC vector	Up to 1 µg
Sterilized Water	Up to 20 µl

3.5   DNA extraction

Isolation of DNA from *E. invadens* trophozoite cells

The fast CTAB (cetyl trimethylammonium bromide) method was used for genomic DNA isolation from *E. invadens* trophozoite cells as described by Ali (Ali et al., 2005). This method was used for removal of carbohydrates and polysaccharides from cells which inhibits the PCR reaction. Initially cultured cell tube was placed on ice for 10 mins and centrifuged at 2000 x g for 5 mins. The pellet was washed with sterile water and transferred to sterile 1.5 ml Eppendorf tube. Afterwards, 250 µl lysis buffer (0.25% SDS in 0.1M EDTA pH 8.0) and 100 µg/ml Proteinase K (1.4 µl from stock of 21.5 mg/ml from Sigma manufacture) was added to the culture pellet. The tube was vortexed and incubated at 55 °C for 20 mins. After incubation, 75 µl of 3.5 M NaCl was added and mixed. Then, preheated 10% CTAB in 0.7 M NaCl at 55 °C was added to the pellet tube for about 42 µl, mixed and incubated at 65 °C for 10 mins. Later, 400 µl chloroform was added and mixed by inversion. The tube was centrifuged at maximum speed for 5 mins. The supernatant was transferred to new sterile 1.5 ml microcentrifuge tube. 400 µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed by inversion and centrifuged as described above. This results in supernatant and pellet in liquid phases. The supernatant whitish liquid phase was transferred to new sterile centrifuged tube and 2 volumes of 100% ethanol was added to it, mixed properly and stored at room temperature for 5 mins. The tube was again centrifuged for 10 mins as mentioned above and

this time supernatant was discarded carefully from the pellet. Pellet was washed with 200 µl of 70% ethanol and centrifuged for 5 mins at maximum speed. Later the pellet was air dried, resuspended in 50 µl sterile water and left for overnight at 4 °C. By use of GC-Health care MicroSpin S-200 HR Columns, the resuspended DNA was passed over the column and the pellet then used for further PCR and recombinant DNA steps.

3.6   Polymerase chain reaction amplification

Polymerase chain reaction (PCR) is a fast and economical method used to copy small segments of DNA as significant amounts of DNA are essential for molecular and genetic analyses, studies of isolated pieces of DNA are approximately impossible without PCR amplification. Reaction mixture was prepared having volume of 25 µl by adding 14.75 µl nuclease free water, 2.5 µl of dNTP's, 2.5 µl of DreamTaq buffer, 2 µl for each forward and reverse primers (Table 6), 1 µl of DNA template and 0.25 µl of DreamTaq polymerase. Negative control was also used having 15.75 µl of nuclease free water. The PCR cycling steps used in this research are summarized in Table 5 as given below.

Table 5. PCR conditions using DreamTaq Polymerase for DNA template amplification

Cycle Steps	Time	Temperature
1	3 min	95 °C
2	0:30 min	95 °C
3	0:30 min	60 °C
4	2 min	72 °C
5	34x step 2	95 °C
6	10 min	72 °C
7	Forever	4 °C

Primers used for PCR amplification given in Table 6.

Table 6. Specific primers used for gene amplification (uppercase indicates perfect match with target gene; lowercase indicates newly introduced sequence with introduced restriction sites indicates by italics).

Names	Primers
Ei_DLP1-*Kpn*I-F (PLD1 forward Primer)	aga aga *ggt acc* ATG AAG AAG ATT TTG TGG AGA G
Ei_DLP1-*Bam*HI-R-HA (PLD1 Reverse Primer)	tct tct *gga tcc* cta TGC ATA GTC TGG AAC GTC ATA TGG ATA CTC AAC CAT AAT GTT TCC G
Ei_DLP2-*Kpn*I-F (PLD2 forward Primer)	aga aga *ggt acc* ATG CGA AGC CTT TAC TAT CGC C
Ei_DLP2-*Bgl*II-R-HA (PLD2 Reverse Primer)	tct tct *aga tct* tta TGC ATA GTC TGG AAC GTC ATA TGG ATA GGT GAC TAT AGC CGT CAG TCC

3.6.1 Agarose gel electrophoresis

Electrophoresis is the technique which is used to separate macromolecules especially nucleic acids based on size, conformation, and charge. Nucleic acids move towards the anode due to negative charge present on it when placed in electric field. DNA has negative charge due to the presence of phosphate backbone. Due to this negative charge and pH of the buffer (8 or above), the DNA moves towards the anode.

The success of DNA extraction was confirmed by agarose gel electrophoresis. The PCR amplified product was ran on 1% agarose gel using 1% TAE buffer. 1% agarose gel was prepared in 1x TAE/TBE buffer by adding the mixture in an Erlenmeyer flask. The mixture was dissolved by heating it in an oven for 1 min, then letting it cool down to about 50 °C. Gel stand was arranged by placing the comb in the stand and the agarose gel solution was poured in it and left for solidification. The samples were loaded in the wells with appropriate amount of loading buffer and Gel Red (Sigma-Aldrich). The gel was ran with the voltage of about 90 V for 45 mins. Later, the gel was analyzed under UV radiation by ChemiDoc Touch Imaging system.

3.6.2 Purification of PCR products

To ensure that only the main band gets sequenced, purification of PCR products was carried out. PCR amplified product which were PLD1 and PLD2 of appropriate molecular size were cut off from the agarose gel with the help of clean scalp and placed in sterilized 1.5 ml Eppendorf tube. Gel bands in two tubes were weighed and processed according to Qiagen QIAquick Gel Extraction kit. The kit protocol includes dissolving of agarose gel band in buffer QG at 50 °C for 10 mins, vortexed and followed by addition of isopropanol. Later passed through the kit provided column and washed by addition of PE buffer, which was then eluted with sterilized milli-Q water in new 1.5 ml centrifuged tube.

3.7 Cloning of purified PCR products into vectors

PCR cloning is a rapid method for cloning genes, and it allows for the cloning of DNA fragments that are not available in large amounts. A PCR reaction is performed to amplify the sequence of interest, and then it is joined to the vector via a blunt or single-base overhang ligation prior to transformation. Early PCR cloning often used Taq DNA Polymerase to amplify the gene. This results in a PCR product with a single template-independent base addition of an adenine (A) residue to the 3' end of the PCR product, through the normal action of the polymerase. These "A-tailed" products are then ligated to a complementary T-tailed vector using T4 DNA ligase, followed by transformation.

3.7.1 Standard vector (pGEM-T-Eeasy)

Ligation reactions were prepared as described in Table 7 (Promega kit) and incubated at 4 °C for overnight. Later, the plasmid construct was centrifuged and mixed with 50 µl *E. coli* JM109 competent cells (which were thawed on ice for 5 mins), and that was followed by incubation on ice for 20 mins. After that heat shock was given at 42 °C for 50 sec and immediately reactions were returned to ice for 2 more minutes. 950 µl SOC media was added to each ligation reactions, which were then incubated on shaker (150 rpm) at 37 °C for 1 hr. and 30 mins. 100 µl of each transformant cultures were plated on LB Agar containing ampicillin, IPTG and X-Gal for white, blue screening of the cells. Lately, the plates were incubated at 37 °C for overnight.

Table 7. Setup of ligation reaction for pGEM-T-Easy vector and PLD1/PLD2

Reagent	Negative control (µl)	Positive control (µl)	Reaction mixture (µl)
2X ligation buffer	5	5	5
T4 DNA ligase	1	1	1
pGEM-T-Easy vector (50 ng)	1	1	1
PCR product (3:1)	-	-	2
Control insert DNA	-	2	-
Deionized water up to final volume	10	10	10

3.7.2 *E. invadens* vector (pEiNEO)

The same process for cloning *E. invadens* vector with PLD1/PLD2 was performed as described in previous section, and the setup for reaction mixture was followed as given in Table 8.

Table 8. Setup of ligation reaction mixture for pEiNEO vector and PLD1/PLD2

Reagent	Negative control (µl)	Positive control (µl)	Reaction mixture (µl)
2X ligation buffer	5	5	5
T4 DNA ligase	1	1	1
pEiNEO vector (50 ng)	1	1	1
PCR product (3:1)	-	-	2
Control insert DNA	-	2	-
Deionized water up to final volume	10	10	10

3.8   Mini cultures for mini prep

Transformed cells were inoculated in 5 ml LB broth containing 5 µl 100 mg/ml ampicillin, which were allowed to grow overnight at 37 °C in shaking incubator (220-250 rpm). The next day, isolation of plasmid DNA was performed by using Nucleospin Plasmid kit (Macherey-Nagel company), which includes cell lysis, neutralization of lysed cells and removal of protein and cell debris by centrifugation. The impurities were removed by using the provided column and later plasmid DNA was eluted with distilled water which was stored at -20 °C until further use.

3.9   NanoDrop One

The DNA concentration was determined by using Thermo Scientific Nano Drop One machine. After selecting the appropriate option for the sample (DNA), 2µl of DNA samples were pipetted (blanking it with the correct elution buffer). The machine was cleaned every time with soft tissue before the examination of a new sample.

3.10 Subsequential restriction digestion

For confirmation of plasmid containing gene of interest, restriction digestion was performed on both pGEM-T-Easy vector containing PLD1/PLD2 and pEiNEO with PLD1/PLD2. The double digestion was done by using Fast digest enzymes from Thermo Scientific and their protocol for plasmid digestion as mentioned in Table 9. For PLD1 gene; *Bam*HI and *Kpn*I restriction enzymes were used, while for PLD2; *Bgl*II and *Kpn*I used.

Table 9. Reaction mixture for restriction digestion

Reagents	Volume (µl)
1X fast digest buffer	2
Fast Digest *Bam*HI (PLD1) / *Bgl*II (PLD2)	1
Fast digest *Kpn*I	1
pEiNEO-LUC/pGEM-T-Easy vector containing gene of interests	Up to 1 ug
Sterilized Water	Up to 20

3.11 Long term storage of bacterial clones

Bacterial clones were prepared by adding 500 µl of 50% sterilized glycerol to 500 µl culture broth and they were stored as -80 °C.

3.12 DNA sequencing

To verify the presence of desired inserts in the plasmid, the plasmids were sent for sequencing, which was performed by Microsynth Seqlab GmbH. Universal SP6 forward and T7 reverse primers were used for sequencing the pGEM-T-Easy based construct while designed Rpl3 forward (GCA GAT GTC GCA CAG AAA GTT TG) and Rpl3 reverse (GAT AAC TTA GAG TTT CTT TCC GG) primers were used for pEiNEO based construct containing desired inserts.

3.13 Transfection

Transfection is the process of introducing nucleic acids into eukaryotic cells by non-viral methods. Using various chemical or physical methods, this gene transfer technology enables the study of gene function and protein expression in a cellular environment. The transfection strategies can generally be classified into two types based on whether the introduced nucleic acid exists in the cell for a limited period of time (transient transfection) or whether it persists in the cells long-term and is passed to the progeny of the transfected cell (stable transfection). *E. invadens* trophozoite cells were grown for 7 to 8 days and then the tube containing these cells was placed on ice to harvest these cells at 500 x g for 5 mins at 4 °C. After harvesting, the cells were washed with PBS and, with an incomplete cytomix buffer to have better stability. Subsequently, the washed cells were suspended in 600 µl of cytomix buffer (recipe in Appendix A), which was completed with 2 mM ATP and 5 mM reduced glutathione. This cell suspension was transferred to electroporation cuvette (0.2 mm) which was followed by the addition of 60 µg plasmid DNA. Later, it was placed in Bio-Rad Gene Pulser Xcell Electroporation systems shock pod and 1500 V intensity of two consecutive pulses were applied with the capacitance of 25 µF. These transfected cells were transferred to 13 ml glass tube containing LYI-S medium as well as heat inactivated adult bovine serum (15%), vitamin mix #18 (2%), and penicillin-streptomycin solution (100 µg/ml). For selection of transfected cells, G418 antibiotic (10 µg/ml) was added after 24 hrs to the culture medium and this medium was refreshed in intervals of 72 hrs.

3.14 Encystation of *E. invadens*

50% LYG media (Table 10) was used for encystation of *E. invadens* trophozoite cells.

Table 10. 50% LYG media composition

Component's	Quantity
K_2HPO_4	0.20 g
KH_2PO_4	0.12 g
NaCl	0.4 g
Yeast extract	5 g
Liver digest neutralized	1 g
Cysteine	0.2 g
Ascorbic acid	0.04 g
Ferric ammonium citrate	0.04 g/ ml
Sodium hydroxide	1 M to adjust pH 6.8
Distilled water	380 ml (final volume)

Media was aliquoted in 95 ml glass bottles and were autoclaved at 121 °C for 15 mins for sterilization. Later, before using the media, it was completed by addition of 5 ml adult bovine serum and 2 ml vitamin mix #18.

Trophozoites cells were allowed to grow at room temperature for 7 days, which were then left on ice for 10 mins, followed by centrifugation at 400 x g for 5 mins. The cells pellet were then ones washed with incomplete 50% LYG and then centrifuged same as above. After centrifugation, the cells pellet was suspended in complete 50% LYG media. This media will allow the trophozoite cells to grow and transform to their cysts stage.

3.15 Calcofluor staining and fluorescence microscopy

Calcofluor stain is a fluorescent stain for rapid detection of yeasts, fungi, and parasitic organisms. Calcofluor White is a non-specific fluorochrome that binds to cellulose and chitin in cell walls. Immunofluorescence (IF) microscopy is a form of immunohistochemistry based on the use of fluorophores to visualize the location of bound antibodies. When transferring the *E. invadens* trophozoite cells to the 50% LYG media, cells samples were taken at different time intervals such

as 24 hrs, 48 hrs and 72 hrs. Cells were centrifuged at 500 x g for 5 mins and then pellet was washed with 1X PBS. Afterwards, cells were suspended in 1% calcofluor white (Sigma-Aldrich) stain for 30 mins, washed again with PBS. Furthermore, fixed by 4% paraformaldehyde and lastly 200 ul PBS was added to pellet. The cells were kept in fridge for later fluorescence microscopy.

3.16 n-Butanol and t-Butanol activity on PLD

Two 15 ml 50% LYG media tubes were taken having 100 µg/ml penicillin-streptomycin antibiotics. In addition to the media, one tube had 0.6 % n-butanol, and another contained 0.6 % tert-butanol (as negative control). Afterwards, 1×10^5 trophozoite cells were added in each tube and incubated at room temperature for 72 hrs. Cell samples were taken at different time intervals 24 hrs, 48 hrs and 72 hrs. Before sample collection for staining at 72 hrs, 0.05% sarkosyl was added to each tube and incubated for 1 hr. Afterwards, they were stained with calcofluor white and fixed with paraformaldehyde as described in previous section (3.15).

4 RESULTS

4.1 Bioinformatics analysis

PLD evolutionary relationship with other organisms via phylogenetic tree analysis and its protein structure were analyzed by use of different tools and software's.

4.1.1 Conserved sequences in PLD

Using NCBI database, total 13 sequences in FASTA format were retrieved and then these sequences were applied in CDD database to identify the conserved sequence from all selected organisms from different groups. PLN02866 was identified as the most conserved domain in all species.

4.1.2 Phylogenetic tree

The phylogeny of PLD across different organism were constructed by Phylogeny.fr. The tree showed a clear separation of four clades further dividing into sub-clades, the alignments lead to phylogenetic tree construction and showed the relationship between different species from different groups as shown in Figure 9.
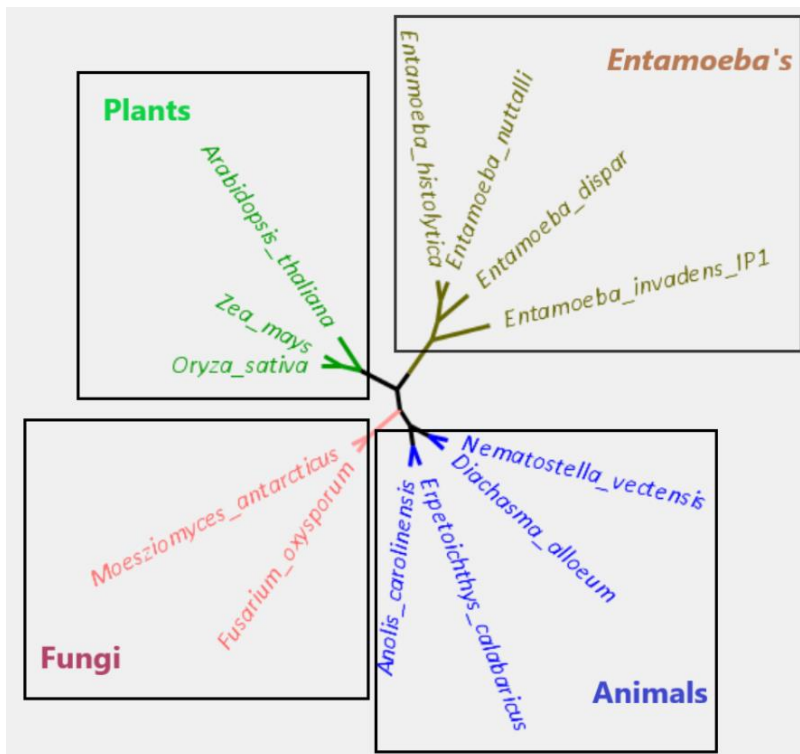
Figure 9. PLD phylogenetic tree generated from Phylogeny.fr and editied in Figtree

4.1.3   Proteomics work

4.1.4   3D model of *E. invadens* PLD

The model of PLN02866 super family domain of PLD was developed by using online source server Swiss model. The model showed 54% of similarities with reference crystal structure and maximum coverage. The theoretical 3D model of PLD PLN02866 super family domain is shown in Figure 10.
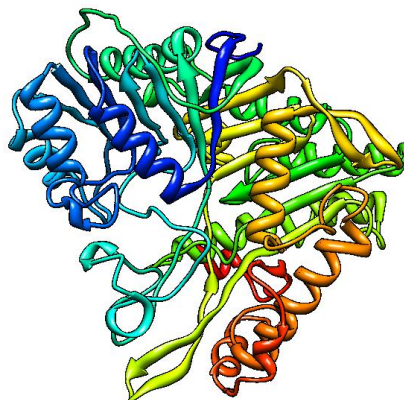
Figure 10. 3D model of *E. invadens* PLD, superfamily domain
PLN02866.

4.1.5   Superimposition of PLD1 protein models

The superimposition of PLD1 template (6ohr.1 human Phospholipase D1 catalytic domain) and E. invadens model was done by chimera, where the green color represented the template and the red colour represented the model of our targeted protein (Figure 11), both the template and model superimposed, in which template had 4566 atoms and 577 residues where as inPLD1 model it was 4343 atoms and 542 residues. The structures are overlapping well in nearly all secondary structurally elements.
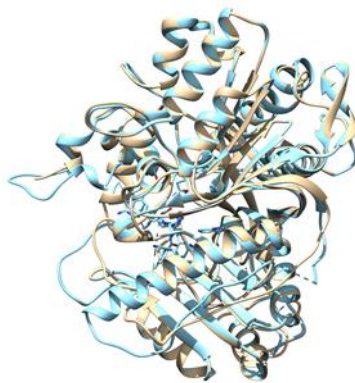
Figure 11. Superimposition of PLD1 template (6ohr.1 chain A) in light blue
and *E. invadens* PLD1 model in golden.

40

The constructed models using Chimera for PLD1 of *E. invadens* and *E. histolytica* were superimposed by using chimera. It showed that the residues in *E. histolytica* model were 544 residues and 4464 atoms. In Figure 12, orange color represents *E. invadens* PLD model while dark blue color represents *E. histolytica* PLD model.
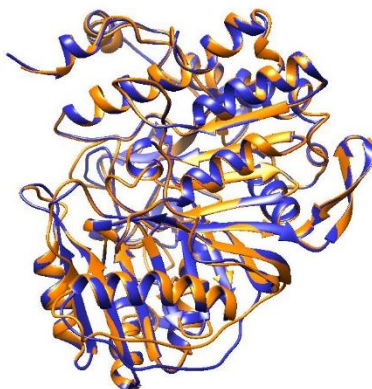
Figure 12. Superimposition of *E. invadens* PLD in orange and *E. histolytica* constructed protein model in dark blue.

4.1.6   Superimposition of PLD2 protein models

For PLD2 the template was a crystal structure of 6ohm.1 human Phospholipase D2 catalytic domain in brown and constructed *E. invadens* PLD2 model in light blue in Figure 13. Both structures were superimposed, where 6ohm.1 structure had 5251 atoms and 1153 residues while PLD2 model had 4326 atoms and 535 residues.
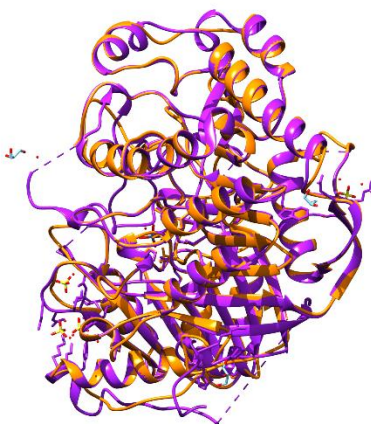
Figure 13. Superimposition of *E. invadens* PLD2 model in orange and 6ohm.1 template in purple.

On other hand the superimposition between *E. histolytica* and *E. invadens* showed in Figure 14. 4359 atoms and 532 residues were found for PLD2 *E. histolytica* model.

Figure 14. Superimposition of *E. invadens* PLD2 model in forest green and *E. histolytica* in yellow.

4.1.7 Validation of model (Ramachandran plot)

In Ramachandran plot, the protein confirmation is displayed graphically where specific region were assigned. Different portion in the plot were the most favorable, generously allowed and disallowed regions which are indicated by different color scheme. For ideal structure it is needed to have 90% of residues in core region (Ho & Brasseur, 2005).

Our result showed that most of PLD residues are located on core region (Figure 15). This orientation is suitable for phi-psi values. About 87% residues were in a favored region and 11% were in allowed region. Approximately 1.02% were present on outlier region.
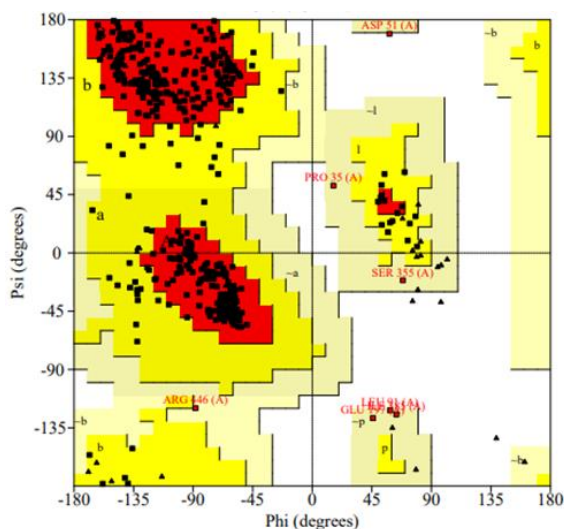
Figure 15. Ramachandran Plot

Ramachandran plot of *E. invaden* PLD structure showed 87.8% residues in favorable region while 0.2% residues in disallowed region. Plot statistics are presented in Table 11

Table 11. Ramachandran plot statistics for constructed PLD model

Plot statistics		
Residues in most favored regions (A, B, L)	430	87.8 %
Residues in additional allowed regions (a, b, l, p)	54	11.0 %
Residues in generously allowed region (~a, ~b, ~l, ~p)	5	1 %
Residues in disallowed regions	1	0.2 %
Number of non-glycine and non-proline residues	490	100 %
Number of end residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	40	
Number of proline residues	14	
Total number of residues	546	

4.2 Molecular Work

E. invadens trophozoites cells were grown at room temp (22 °C) in 15ml LYI-S medium for 2 weeks. Following results were analyzed from these parental cells.

4.2.1 Isolation and PCR amplification of *E. invadens* genomic DNA

The genomic DNA isolated from *E. invadens* trophozoite cells having concentration 238.9 ng/µl, was amplified directly by using specific primers. The PCR product resulted in two genes of interest with their predicted sizes: PLD1 1.7 kbp and PLD2 1.6 kbp. They were analyzed on 1% agarose gel via Gel electrophoresis as shown in Figure 16.
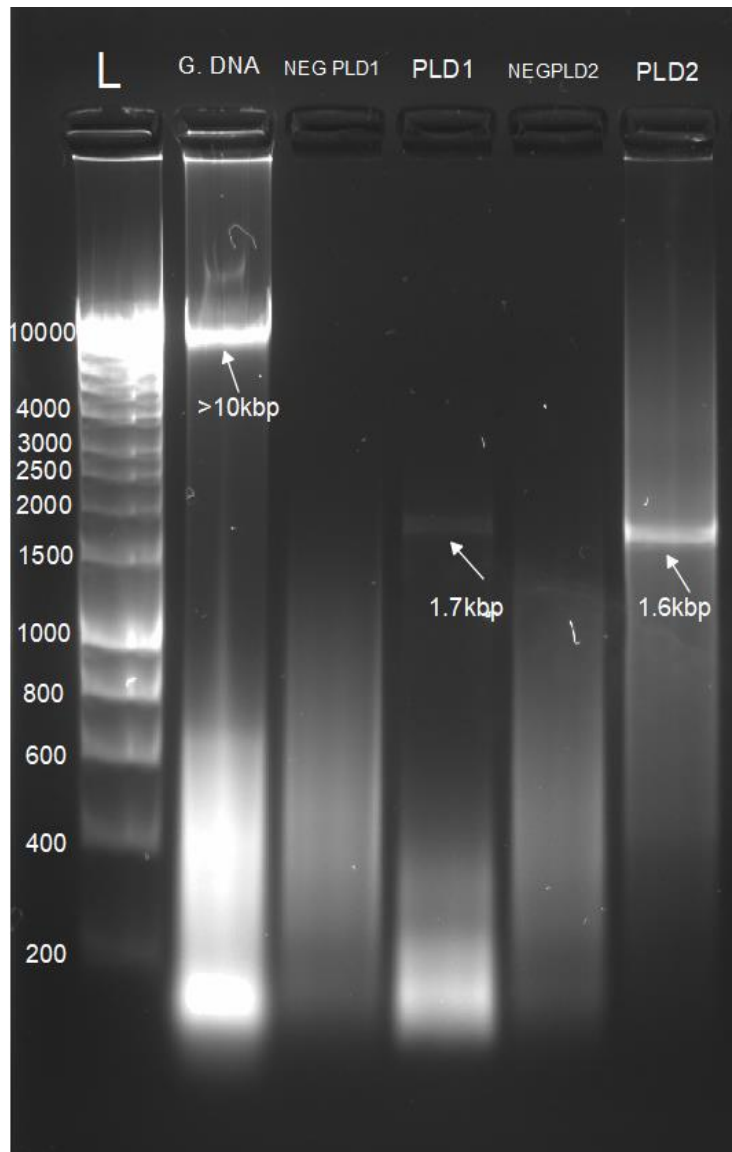
Figure 16. PCR amplification from *E. invadens* genomic DNA

Lane 1: L= Bioline 1 kb Ladder, Lane 2: Isolated genomic DNA from *E. invadens* trophozoite cells, Lane 3: Negative PCR control for PLD1, Lane 4: Amplified PCR product PLD1, Lane 5: Negative PCR control for PLD2, Lane 6: Amplified PCR product PLD2.

4.2.2 Determination of DNA concentration using NanoDrop One

The DNA samples concentration were determined by using Thermo Scientific NanoDrop One as shown in Table 12. The One's selected for further experiments were pGEM-T-Easy+PLD1 vector (Plasmid purification) from colony 1, whereas pGEM-T-Easy+PLD2 vector (Plasmid purification) from colony 1.

Table 12. Concentration of pGEM-T-Easy vector+PLD1/PLD2 insert using Nanodrop ONE

Sample Name	DNA Concentration	Sample Type
1. colony 1 pGEM-T-easy+PLD1	225.7 ng/µl	DNA
2. colony 2 pGEM-T-Easy+PLD1	200.0 ng/µl	DNA
3. colony 1 pGEM-T-Easy+PLD2	189.2 ng/µl	DNA
3. colony 2 pGEM-T-Easy+PLD2	223.3 ng/µl	DNA

For *E. invadens* vector, pEiNEO+PLD1 vector (Plasmid purification) with concentration 345.2 ng/µl, whereas pEiNEO+PLD2 vector (Plasmid purification) having concentration 241.3 ng/µl were been selected for transfection into *E. invadens* trophozoite cells. Table 13 shows the concentrations of selected pEiNEO cloned plasmid for transfection.

Table 13. Concentration of pEiNEOvector+PLD1/PLD2 insert using Nanodrop ONE

Sample Name	DNA Concentration	Sample Type
1. colony 1 pEiNEO-PLD1	345.2 ng/µl	DNA
2. colony 1 pEiNEO-PLD2	241.3 ng/µl	DNA

4.2.3 DNA fragments on gel after restriction digestion

For confirmation of cloned vectors that either they contain gene of interest or not, subsequential restriction digestion was performed. The restriction digestion products were analyzed on 1% agarose gel as shown in Figure 17. All cloned confirmed the presence of gene of interest. The undigested cloned vectors are shown in lane 1 of Figure 17 (a, b, c, d). Lane 2 and 3, Figure 17(a), after single digestion with *Kpn*I or *Bam*HI, the clear shift in migration of plasmid is shown, which size is 4.7 kbp. Whereas Lane 4, the double digestion by *Kpn*I and *Bam*HI shows the appearance of two band of 1.7 kbp and 3 kbp size, respectively. For Figure 17 (b) pEiNEO+PLD1 vector, Lane 2 and 3 after single digestion with *Kpn*I or *Bam*HI, the migration of cloned plasmid is shown, which size is 8.2 kbp. While Lane 4, the double digestion by *Kpn*I and *Bam*HI shows the presence of two band of 1.7 kbp and 6.5 kbp size, respectively. In Figure 17 (c), single digestions with *Kpn*I and *Bgl*II in Lane 2 and 3 respectively, gave 4.6 kbp size bands of cloned pGEM-T-Easy with PLD2, while the double digestion results in two bands of sizes 1.6 kbp and 3 kbp. Figure 17 (d),

Lane 2 and 3 represent the band of 8.1 kbp, which appeared after single digestion with *Kpn*I and *Bgl*II. Lane 4 shows the double digestion of pEiNEO+PLD2 with band sizes 6.5 kpb (pEiNEO vector) and 1.6 kbp (PLD2 gene).
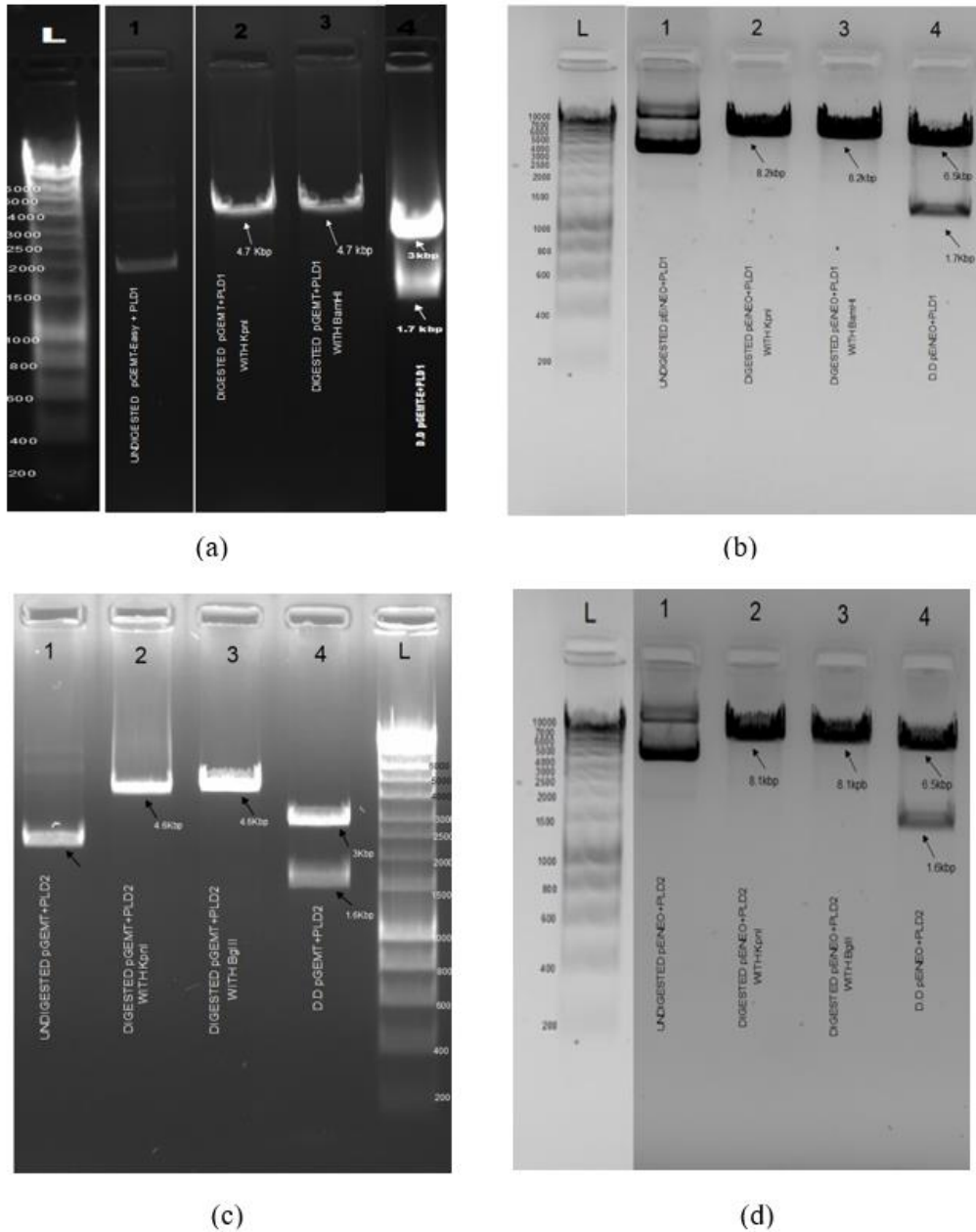

Figure 17. Sub-sequential restriction digestion

Restriction digestion was analyzed on 1% agarose gel where: (a) gel representing the restriction digestion of pGEM-T Easy + PLD1, (b) gel representing the restriction digestion of pEiNEO + PLD1, (c) gel representing the restriction digestion of pGEM-T Easy + PLD2, (d) gel representing

4.3 Sequencing results

The sequences results were compiled, and alignment was performed using Clustal Omega multiple alignment tool online. The DNA sequence was first translated to their protein and then alignment was performed with the reference protein sequence. pGEM-T-Easy vector cloned with PLD1 and PLD2 were confirmed by sequencing result that the vector contains desired gene of interest. (Figure 18 and Figure 19). Alignment can be seen under appendices Appendix B: Sequence alignment of cloned pGEM-T-Easy and PLD1/PLD28.2 (Appendix B).
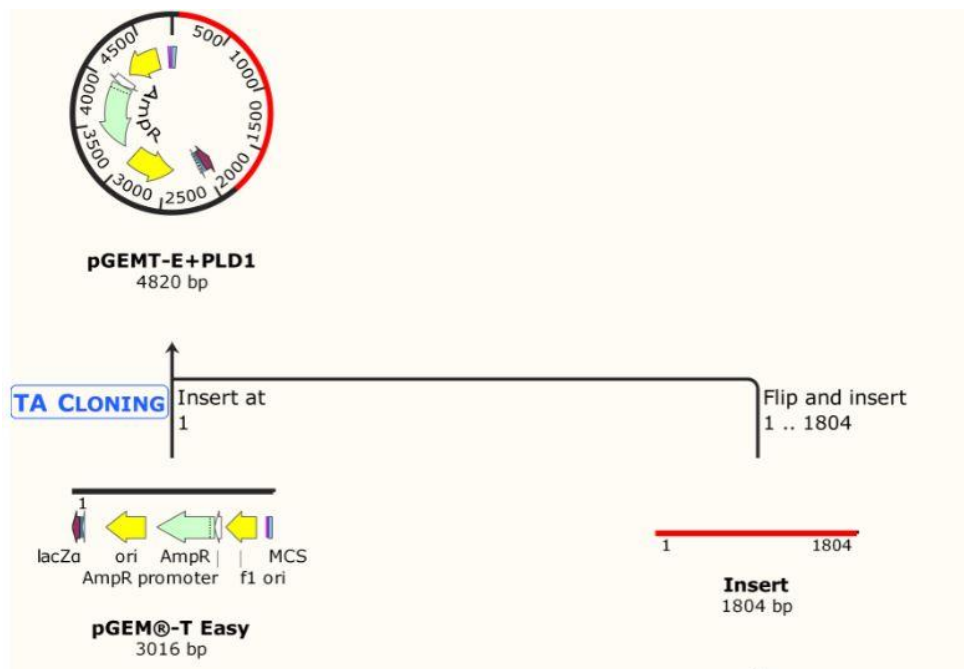
Figure 18. pGEM-T-Easy vector cloned with PLD1 gene constructed by using Snapgene.

Insert in red color: amplified PLD1 with forward and reverse primers.

pGEM-T-Easy vector in black color: standard linearized vector for
TA cloning
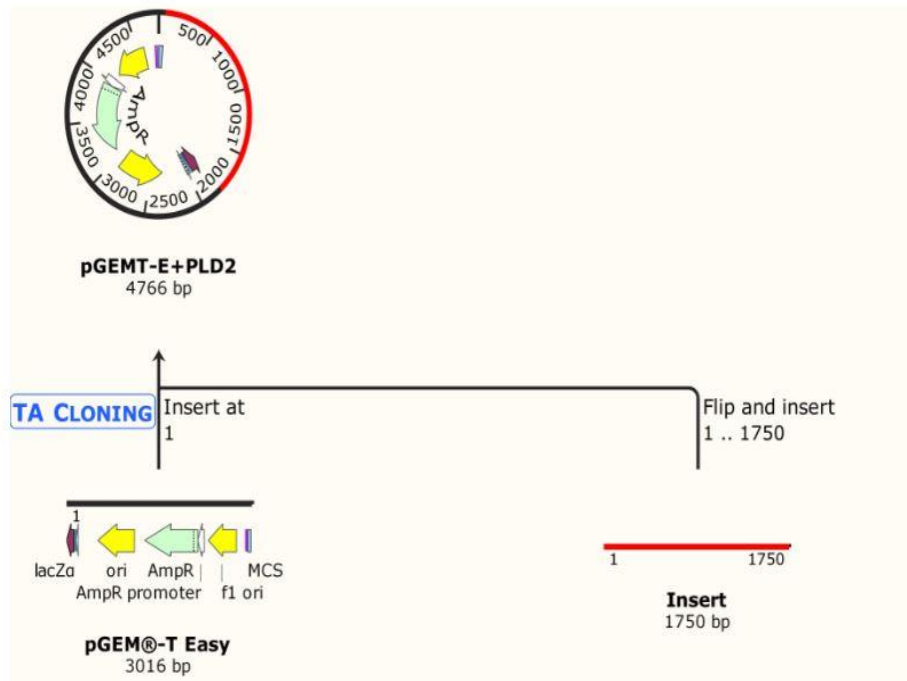
Figure 19. pGEM-T-Easy vector cloned with PLD2 gene constructed by using Snapgene

Insert in red color: amplified PLD2 with forward and reverse primers.

pGEM-T-Easy vector in black color: standard linearized vector for TA cloning

4.3.1 Encystation

The encystation study of *E. invadens* was carried out and the effect of 0.6% n-butanol on PLD activity was also tested as it is considered to repress the activity of PLD. Same amount of t-butanol was also used as control encystation media. The encystation was allowed to proceed for 24, 48 and 72 hrs. Chitin production was observed by fluorescence microscope at mentioned time intervals (Table 14). No effect of n-butanol on encystation was found. Same number of cysts cells were found in wild type, n-butanol cultures and in control (Table 15). In wild type cultures, about 32% cyst was formed after 24 hrs, 66% after 48 hrs and 93% after 72 hrs in stained sample. In n-butanol cultures, about 36% cyst was formed after 24 hrs, 57% after 48 hrs and 100% after 72 hrs in stained sample. In n-butanol cultures, about 28% cyst was formed after 24 hrs, 52% after 48 hrs and 100% after 72 hrs in stained sample as shown in Figure 20. The experiment was performed once.

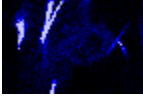Table 14. Cysts chitin observation by using bright field and fluorescence microscope
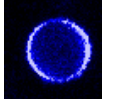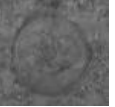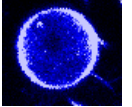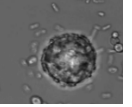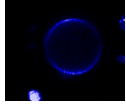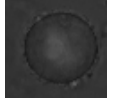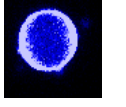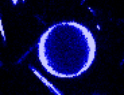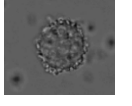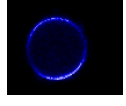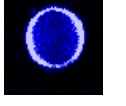
Media type	24 hrs		48 hrs		72 hrs	
	Bright field	CFW stain	Bright field	CFW stain	Bright field	CFW stain
Wild type						
n-Butanol						
t-Butanol						

Table 15. Encystation percentage (%) of *E. invadens* at different time intervals

Time	Wild type	N-Butanol	T-Butanol
24 hrs	32	36	28
48 hrs	66	57	52
72 hrs	93	100	100
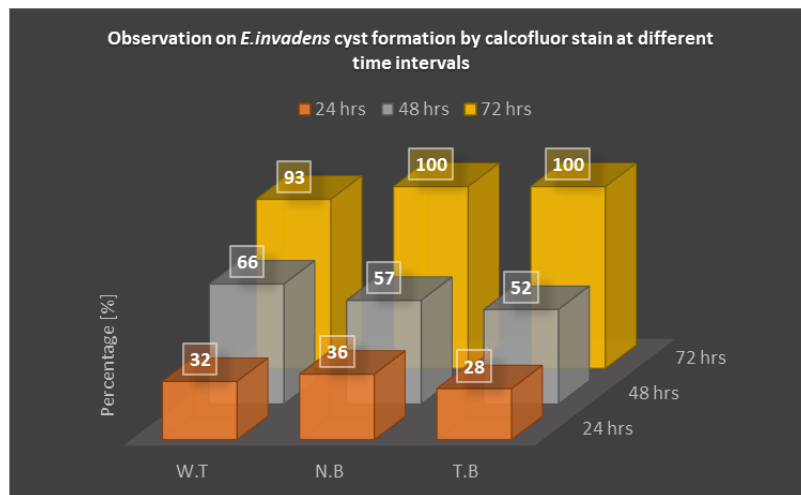
Figure 20. Cysts formation at different time interval.

W.T= wild type 50% LYG media, N.B= 0.6% n-butanol added to 50% LYG media, T.B= 0.6% t-butanol added to 50% LYG media, whereas cells for each sample were analysed at different time intervals such as 24 hrs, 48hrs and 72 hrs.

5   DISCUSSION

Being an important protozoan parasite, knowledge of different stages of its life cycle are of immense importance in interrupting the life cycle of *Entamoeba*. One such conversion stage is encystation, which was the main research topic of this thesis. Encystation in *Entamoeba* has the potential to be targeted in stopping infection and interpreting the life cycle, hence turning down the transmission of this parasite (Aguilar-Díaz et al., 2010). It has been a challenge to study encystation and its role in infection directly in *Entamoeba* (*E. histolytica*) however a related reptile parasite, *E. invadens* has proved to be a competent model organism for such studies. Development of *in vitro* encystation system in *E. histolytica* still remains high priority goal for researchers. To investigate the proteins and genes involved in encystation, frequent studies have been attempted. PLD, an enzyme involved in lipid second messenger signaling has been reported to strongly up regulate during early encystation. Genome sequencing study of *E. histolytica* and *E. invadens* has provided novel information related to amoeba development yet the cellular events, signal transduction, regulation and molecular changes during encystation are poorly understood. Ehrenkaufer and her co-workers (2013) used whole transcriptome sequencing for studying changes in gene expression during encystation and excystation, It was reported that the genomic assembly of *E. invadens* is large (11549 genes) compared to *E. histolytica* (8306 genes) mainly due to intergenic regions expansion but overall number of genes and machinery for regulation of genes were conserved between both species.

5.1   Genomic DNA isolation

Getting the high molecular weight genomic DNA of *E. invadens* was challenging and it was initially difficult to get a good yield following the standard protocols of using CTAB. Primarily 15 ml cultures were used for isolating genomic DNA but that resulted in low yield and improved concentration was attained using 100 ml cultures under the same conditions. Another reason could be because the initial culture was taking time to adopt in the media and there were less visible cells under the microscope. Different temperatures were also checked, and the best growth was observed in room temperature of 23 °C. The genes encoding PLD1 and PLD2 were PCR amplified directly from genomic DNA of *E. invadens* using the samples which were passed through S-200 HR column for removing PCR inhibitors such as salts and restriction enzymes. In the initial experiment of isolating genomic DNA, the S-200 HR column was not used, and it resulted in no amplification.

MicroSpin S-200 HR although reduced the concentration, improved the purity of the genomic DNA and hence it was used before PCR amplification.

5.2 Amplification of genes

In this study, amplification of both genes was challenging, different polymerases were used to get the required bands. Promega mix containing TaqDNA polymerase, dNTPs, MgCl$_2$ and optimal reaction buffer resulted in no amplification of PLD1 and PLD2. However, Dream Taq polymerase, dNTPs mix and dream taq buffer resulted in PCR products which contained PLD1 and PLD2 gene bands when run on the gel. The reason behind DreamTaq yielding the desired genes could be because of its higher sensitivity and high yield ability compared to taq DNA polymerase (Fisher, n.d.). There were a set of two primers used for the amplification of PLD1 and PLD2. One was gene specific primer set and the other one was gene specific primer containing restrictions site and HA tag. The ones containing restriction site and HA tag resulted in successful amplification however the ones without it resulted in no amplification.

5.3 Transformation in *E. coli* JM109

The gene encoding PLD1 and PLD2 was cloned, and recombinantly expressed in *E. coli* JM109 using Promega standard vector pGEM-T-Easy. On PLD2 plate, only two colonies were obtained, one white and one blue, according to the protocol given by Promega, white colored colonies means that it is a transformed colony and blue colonies are the ones which hasn't transformed successfully. To confirm the transformation, miniprep for both colonies was done. Surprisingly, sequencing of inserts of transformants revealed the presence of PLD2 in both blue and white colonies. It was unexpected as it is mentioned in most literatures that blue colonies do not contain gene of interest. However, mentioned on the Promega website was also that sometimes even transformed *E. coli* JM109 can contain blue color cause of a possible mutation (Grooms et al., 2019). They might have a second mutation, *laclq* which upsurges the production of repressor *lacl* thus stopping transcription from the lac operon. The substrate, IPTG, dismisses the repression of the lac operon and lets transcription to occur. These strains will need to be grown on media containing IPTG as well as X-gal. Sometimes, colonies will not appear white but pale blue. So, there is strong chance that pale blue colonies might contain gene of interest (Promega, n.d.).

5.4 Transfection in *Entamoeba invadens*

Previously, DNA-mediated gene transfer has been extensively used as a powerful tool for the genetic alteration of cells or organisms. Transfection and expression of DNA in protozoan parasites was reported to be more difficult as compared to higher eukaryotic cells, as particular growth conditions are required for protozoan parasites and own various peculiarities in gene organization and control of gene expression (Huo et al., 2016).

For the successful transfection in protists, electroporation is a highly effective method that has been extensively used regardless of the high voltages that might upset the viability of cells. Successful transfection of *E. invadens* can be achieved using thousands of volts (1000-4000 V). High voltage is known to kill most of the *E. histolytica* trophozoites at 3000 V/m and 50% of *Leishmania enriettii* at 3000 V/m. *E. invadens* transfection was performed by electroporation through applying two consecutive electrical pulses of 3000 V using a 0.4 mm cuvette or 1500 V using a 0.2 mm cuvette (Turfah, 2020). The transfection efficiency is greatly affected by viability and number of cells, osmolarity, DNA purity and concentrations (Zu et al., 2014). Transfection of *E. invadens* in this study was also challenging because the culture of *E. invadens* trophozoites cells was not dense, instead it was a thin layer. That could be because of bacterial contamination in the LYI-S media. It is common to have more bacterial contamination during cell culturing, although penicillin and streptomycin antibiotics were added into media prior to inoculum addition. Successful transfection was obtained by adding G418 antibiotic in the media. The constructed vector contained neomycin phosphotransferase (NEO) gene as a selectable marker for the transfection of the parasite. We choose NEO gene because studies showed that amoebae are remarkably sensitive to G418 as compared to other eukaryotes (Hamann et al., 1995).

5.5 PLD and its role in Encystation

Role of PLD in *E. invadens* encystation was also assessed. PLD was one of the proteins which is highly expressed during encystation. PLD is involved in lipid second messenger signaling. It is responsible for the conversion of phosphatidyl choline to phosphatidic acid and linked with various crucial biological processes such as vesicle transport and signal transduction. *E. invadens* encode two PLD proteins; PLD1 and PLD2 and both were found up-regulated during encystation. It was also up-regulated in *E. histolytica* cysts (Ehrenkaufer G. M. et al., 2013). To assess the cysts formation during encystation, 0.6% n-butanol was used, which is known to inhibit PLD activity

by serving as non-productive substrate for PLD thus suppressing its activity. T-butanol was used as control as it has no effect on PLD activity. The encystation was allowed to procced for 24 hrs, 48 hrs and 72 hrs. Astonishingly, no decline in encystation efficiency was found in n-butanol treated samples and was approximately similar to that of wild type and t-butanol treated samples. Contrary to our results, marked reduction of encystation efficiency was observed in n-butanol treated samples by Ehrenkaufer and co-workers. There was no major effect on encystation in t-butanol treated samples. So, they concluded that n-butanol is responsible for repression of encystation (Ehrenkaufer G. M. et al., 2013). But in the current study, no change in cyst formation activity was observed in all the three samples.

5.6 Bioinformatics of PLD

The phylogenetic tree of PLD across different organisms was constructed with 100% bootstrap which showed that PLD of *Entamoeba* is distantly related to PLD of human beings. Similarity of sequences or structures usually elicits a discussion on the evolutionary relatedness of the respective proteins and structure/structure alignments are often used in the statistical analysis of protein folds. Structures of different proteins is usually compared by superimposing the coordinates of protein backbones or selected parts of the proteins (Nassoury & Morse, 2005). Superimposition of PLD1 protein (model) with 6ohr.1 (template) which is a human phospholipase D1 catalytic domain was carried out. The superimposition between template and model showed 29.89% similarity. Superimposition between *E. histolytica* PLD1 protein and *E. invadens* PLD1 protein was also carried out which showed similarity of 54.54%. On the other hand, superimposition of PLD2 protein (model) and template (6ohm.1) showed 31.09 % similarity. *E. histolytica* PLD2 protein and *E. invadens* PLD2 protein superimposition showed similarity of 55.45%. The structural dissimilarities between human PLD catalytic domain and *Entamoeba* PLD is desirable as it will help to design targeted approaches to disrupt cyst formation and disease transmission.

PLD protein sequence alignment of *E. nuttali*, *E. histolytica*, *E. dispar, E. invadens, Zea mays*, *Oryza sativa*, *Arabidopsis thaliana*, *Fusarium oxysporum*, *Moesziomyces antarcticus*, *Diachasma alloeum*, *Nematostella vectensis*, *Anolis carolinensis*, *Erpetoichthys calabaricus* and *Homo sapiens* was performed. The sequence alignment showed that HKD is the main active site motif and is present in all the aforementioned organisms. HKD motif appears twice, and both are conserved as shown in Figure 21 and Figure 22.
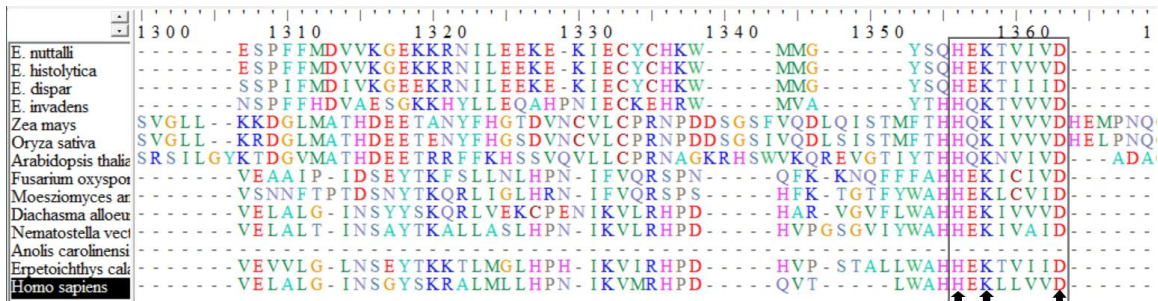
Figure 21. Alignment representation of active domain motif site (HKD) #1

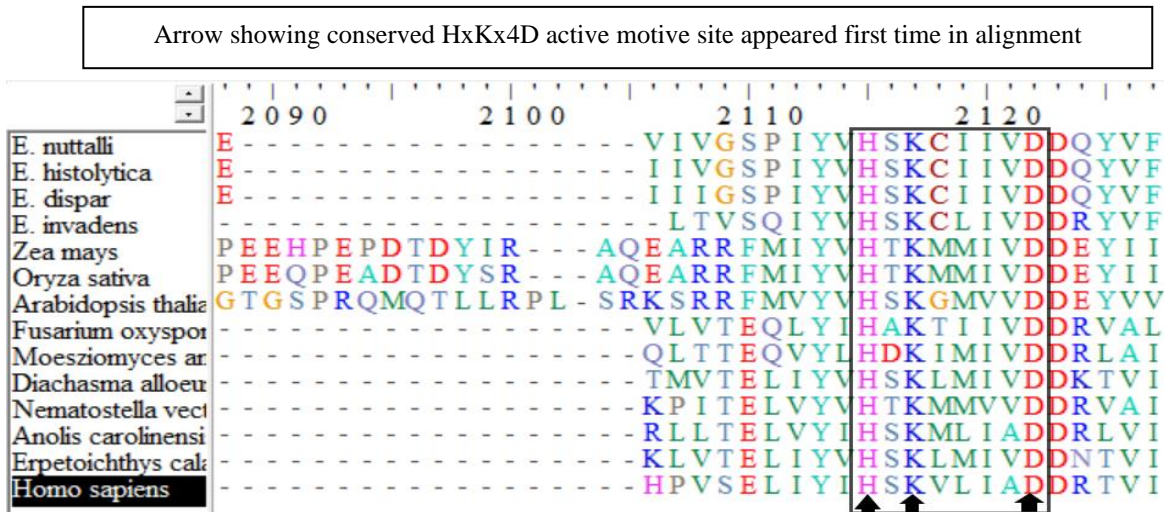Arrow showing conserved HxKx4D active motive site appeared first time in alignment

Figure 22. Alignment representation of Active domain motif site (HKD) #2

Arrow showing conserved HxKx4D active motive site appeared second time in alignment

These two HKD motifs confer hydrolytic activity to PLD and are critical for its enzymatic activity both *in vitro* and *in vivo*. Two HKD motifs from the two domains form a single active site. HKD motif comprises of histidine, lysine, and aspartic acid. The two histidine residues from the two HKD motifs play key roles in the catalysis. Upon substrate binding, a histidine residue from one HKD motif could function as the nucleophile, attacking the phosphodiester bond to create a covalent phosphohistidine intermediate, while the other histidine residue from the second HKD motif could serve as a general acid, stabilizing the leaving group (NCBI, Conserved Protein Domain Family, 2020). Based on structural details, mechanism of human PLD2 was proposed.

Histidine 442 exerts a nucleophilic attack on phosphate moiety of phosphatidyl choline (PC), with the help of aspartate 363. Histidine 756, lysine's 444 and 758 might be involved in stabilizing the tetrahedral intermediate. In the final step, histidine 756 donates hydrogen to the tetrahedral intermediate and activates water, which then exerts a second nucleophilic attack, ultimately leading to the release of choline and phosphatidic acid (PA) respectively (Mahankali, 2015). The conserved catalytic domain might demonstrate the same function of PLD in human and *Entamoeba*. In future, we can introduce site-directed mutagenesis in the active site of PLD protein of *Entamoeba* to study whether PLD performs same function in human and *Entamoeba* or not. Whereas when PLD1 and PLD2 sequence was searched in Pfam database, it showed two same protein family that were PLDc_2 (PLD-like domain) and PLDc (Phospholipase D Active site motif).

6   CONCLUSION

Entamoeba's infection are of immense importance in developing countries and more knowledge about its life cycle will help scientists to identify the novel therapies in treating infections caused by it. In the current study, the aim was to study the role of PLD1 and PLD2 in encystation as it has been suggested in the literature that those two genes are upregulated during cyst formation, and *Entamoeba* cells are the hardest to treat in cyst form. Cloning of these two genes (PLD1 and PLD2) was successful and so was transfection to *Entamoeba invadens* trophozoite cells. Furthermore, imaging of *Entamoeba invadens* cysts under different conditions was also attained. However, more research will be needed to determine the role of PLD1 and PLD2 during encystation and how the whole process works. Important things to understand would be what triggers the upregulation when *Entamoeba* is inside the host and possible therapies to target these changes. N-butanol showed no significant effect on encystation activity, but the experiment needs to be performed multiple times with proper controls to conclude that as a fact.

7 REFERENCES

Aguilar-Díaz, H., Díaz-Gallardo, M., Laclette, J. P., & Carrero, J. C. (2010). In vitro induction of *Entamoeba histolytica* cyst-like structures from trophozoites. *PLoS Negl Trop Dis*, 4(2), e607.

Aguirre-Beltran, A. G. (1999). Standardisation and evaluation of differential diagnostic systems for the detection of *Entamoeba histolytica* and *Entamoeba dispar. (Doctoral dissertation, London School of Hygiene & Tropical Medicine).*

Ali, I.K.M., Haque, R., Siddique, A., Kabir, M., Sherman, N.E., Gray, S.A., Cangelosi, G.A. & Petri Jr, W.A., (2012). Proteomic analysis of the cyst stage of *Entamoeba histolytica*. *PLoS Negl Trop Dis*, 8;6(5):e1643.

Ali, I. K. M., Zaki, M., & Clark, C. G. (2005). Use of PCR amplification of tRNA gene-linked short tandem repeats for genotyping *Entamoeba histolytica*. *Journal of clinical microbiology*, 43(12), 5842-5847.

Aloulou, A., Rahier, R., Arhab, Y., Noiriel, A., & Abousalham, A (2018). Phospholipases: an overview. *In Lipases and Phospholipases*, 69-105.

Ansari, J., Kaur, G., & Gavins, F. N. (2018). Therapeutic potential of annexin A1 in ischemia reperfusion injury. *International Journal of Molecular Sciences*, 19(4), 1211.

Arredondo, J. L., González, M. B., Coria, A. L., Ortega, J. E., Vargas, J., Villarreal, J. L., & Vallarta, M. M. (2014). *Entamoeba histolytica*: trophozoite, precyst and cyst studied by atomic force microscopy.

Bandana, K., Jashandeep, K., & Jagdeep, K. (2018). Phospholipases in bacterial virulence and pathogenesis. *Adv Biotechnol Microbiol*, 10(5), 1-8.

Bansal, D., Ave, P., Kerneis, S., Frileux, P., Boché, O., Baglin, A.C., Dubost, G., Leguern, A.S., Prevost, M.C., Bracha, R.& Mirelman, D. (2009). An ex-vivo human intestinal model to study *Entamoeba histolytica* pathogenesis. *PLoS Negl Trop Dis*, 3(11):e551.

Bhattacharya, A., Campbell, D., Espinosa, C.M., Kris, C., Graham, C.C., Jackson, T.F., Kain, K.C., Keystone, J.S., David, M., Pillai, D.R. & Girija, R. eds. (2000). *Amebiasis.* World Scientific: Vol. 2.

Bosch, D.E., Kimple, A.J., Manning, A.J., Muller, R.E., Willard, F.S., Machius, M., Rogers, S.L. & Siderovski, D.P. (2013). Structural determinants of RGS-RhoGEF signaling critical to *Entamoeba histolytica* pathogenesis. *Structure*, 21(1), 65-75.

Bowling, F. Z., Salazar, C. M., Bell, J. A., Huq, T. S., Frohman, M. A., & Airola, M. V. (2020). Crystal structure of human PLD1 provides insight into activation by PI (4, 5) P 2 and RhoA. *Nature Chemical Biology*, 16(4), 400-407.

Brunette, G. W., Kozarsky, P., Magill, A. J., & Shlim, D. R. (Eds.). (2009). *CDC health information for international travel 2010.* Elsevier Health Sciences.

Chatterjee, A., Ghosh, S. K., Jang, K., Bullitt, E., Moore, L., Robbins, P. W., & Samuelson, J. (2009). Evidence for a "wattle and daub" model of the cyst wall of *Entamoeba*. *PLoS Pathog*, 5(7), e1000498.

Chia, M. Y., Jeng, C. R., Hsiao, S. H., Lee, A. H., Chen, C. Y., & Pang, V. F. (2009). *Entamoeba invadens* myositis in a common water monitor lizard (Varanus salvator). *Veterinary pathology*, 46(4), 673-676.

Chimera molecular modeling system. (n.d.). Retrieved from https://www.cgl.ucsf.edu/chimera/

Christy, N. C., & Petri, W. A. (2011). Mechanisms of adherence, cytotoxicity and phagocytosis modulate the pathogenesis of *Entamoeba histolytica*. *Future Microbiology*, 6(12), 1501-1519.

Das, K., & Ganguly, S. (2014). Evolutionary genomics and population structure of *Entamoeba histolytica*. 12(20-21), 26-33.

Dennis, E. A. (2015). Introduction to thematic review series: phospholipases: central role in lipid signaling and disease. 56(7), 1245-1247.

Dhanalakshmi, S., & Parija, S. C. (2016). Seroprevalence of *Entamoeba histolytica* from a tertiary care hospital, South India. *Tropical parasitology*, 6(1), 78.

DIAMOND, L. S., & CLARK, C. G. (1993). A Redescription of *Entamoeba histolytica* Schaudinn, 1903 (Emended Walker, 1911) Separating It From Entamoeba Dispar Brumpt, 1925. *Journal of Eukaryotic Microbiology*, 40(3), 340-344.

Ehrenkaufer, G. M., Haque, R., Hackney, J. A., Eichinger, D. J., & Singh, U. (2007). Identification of developmentally regulated genes in *Entamoeba histolytica*: insights into mechanisms of stage conversion in a protozoan parasite. *Cellular microbiology*, 9.

Ehrenkaufer, G. M., Weedall, G. D., Williams, D., Lorenzi, H. A., Caler, E., Hall, N., & Singh, U. (2013). The genome and transcriptome of the enteric parasite *Entamoeba invadens*, a model for encystation. 14(7), R77.

El-Dib, N. A. (2017). *Entamoeba histolytica*: an overview. *Current Tropical Medicine Reports*, 4(1), 11-20.

EMBL. (n.d.). *EMBL*. Retrieved from Cloning: https://www.embl.de/pepcore/pepcore_services/cloning/pcr_strategy/primer_design/

Espinosa-Cantellano, M., & Martinez-Palomo, A. (2000). Pathogenesis of intestinal *amebiasis*: from molecules to disease. *Clinical microbiology reviews*, 13(2), 318-331.

Fisher, T. (n.d.). *DreamTaq DNA Polymerases*. Retrieved from Thermofisher.com: https://www.thermofisher.com/no/en/home/brands/thermo-scientific/molecular-biology/thermo-scientific-pcr/thermo-scientific-pcr-enzymes-master-mixes/hot-start-dna-polymerases-master-mixes-thermo-scientific/dreamtaq-dna-polymerase.html

Frisardi, M., Ghosh, S. K., Field, J., Van Dellen, K., Rogers, R., Robbins, P., & Samuelson, J. (2000). The most abundant glycoprotein of amebic cyst walls (Jacob) is a lectin with five Cys-rich, chitin-binding domains. *Infection and immunity*, 68(7), 4217.

Garmie, V. (2016). Prevalence and Intensity of *Entamoeba Histolytica* in Patients Attending Health Centres in Mathare Slums, Nairobi County, Kenya . *Doctoral dissertation, University of Nairobi*.

Gomes, T. D. S., Garcia, M. C., Souza Cunha, F. D., Werneck de Macedo, H., Peralta, J. M., & Peralta, R. H. S. (2014). Differential diagnosis of *Entamoeba* spp. in clinical stool samples using SYBR green real-time polymerase chain reaction. *The Scientific World Journal*.

Grooms, K. (2019). *Selecting the Right Colony: The Answer is There in Blue and White*. Retrieved from Promega.com: https://www.promegaconnections.com/selecting-the-right-colony-the-answer-is-there-in-blue-and-white/

Haas, E., & Stanley, D. W. (2007). Phospholipases. Dans xPharm: The Comprehensive Pharmacology Reference. (p. 1-3).

Hamann, L., Nickel, R., & Tannich, E (1995). Transfection and continuous expression of heterologous genes in the protozoan parasite *Entamoeba histolytica*. *Proceedings of the National Academy of Sciences*, 92(19), 8975-8979.

Hardin, R. E., Ferzli, G. S., Zenilman, M. E., Gadangi, P. K., & Bowne, W. B. (2007). Invasive *amebiasis* and ameboma formation presenting as a rectal mass: An uncommon case of malignant masquerade at a western medical center. *World Journal of Gastroenterology: WJG*, 13(42), 5659.

Hemphill, A., Müller, N., & Müller, J. (2019). Comparative pathobiology of the intestinal protozoan parasites *Giardia lamblia*, *Entamoeba histolytica*, and *Cryptosporidium parvum*. *Pathogens*, 8(3), 116.

Herman, E., Siegesmund, M.A., Bottery, M.J., van Aerle, R., Shather, M.M., Caler, E., Dacks, J.B. & Van Der Giezen, M., (2017). Membrane trafficking modulation during *Entamoeba* encystation. *Scientific reports*, 7(1), 1-17.

Ho, B. K., & Brasseur, R. (2005). The Ramachandran plots of glycine and pre-proline. *BMC structural biology*, 5(1), 1-11.

Huo, Z. Y., Xie, X., Yu, T., Lu, Y., Feng, C., & Hu, H. Y. (2016). Nanowire-modified three-dimensional electrode enabling low-voltage electroporation for water disinfection. *Environmental science & technology*, 50(14), 7641-7649.

Jenkins, G. M., & Frohman, M. A. (2005). Phospholipase D: a lipid centric review. *Cellular and Molecular Life Sciences CMLS*, 62(19-20), 2305-2316.

John, D. T. (2006). In *Markell and Voge's medical parasitology. 9th. Ed.* (pp. 178-81). St. Louis, Missouri: Saunders Elsevier.

Kadamur, G., & Ross, E. M. (2013). Mammalian phospholipase C. *Annual review of physiology*, 75, 127-154.

Karan, A., Chapman, G. B., & Galvani, A (2012). The influence of poverty and culture on the transmission of parasitic infections in rural Nicaraguan villages. *Journal of parasitology research*.

Kolesnikov, Y. S., Nokhrina, K. P., Kretynin, S. V., Volotovski, I. D., Martinec, J., Romanov, G. A., & Kravets, V. S. (2012). Molecular structure of phospholipase D and regulatory mechanisms of its activity in plant and animal cells. *Biochemistry*, 77(1),1-14.

Kooijman, E. E., Chupin, V., Fuller, N. L., Kozlov, M. M., de Kruijff, B., Burger, K. N., & Rand, P. R. (2005). Spontaneous curvature of phosphatidic acid and lysophosphatidic acid. *Biochemistry*, 44(6), 2097-2102.

Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology modeling. *Methods of biochemical analysis*, 44, 509-524.

Labruyère, E., & Guillén, N. (2006). Host tissue invasion by *Entamoeba histolytica* is powered by motility and phagocytosis. *Archives of medical research*, 37(2), 252-257.

Leiros, I., Secundo, F., Zambonelli, C., Servi, S., & Hough, E. (2000). The first crystal structure of a phospholipase D. *Structure*, 8(6), 655-667.

Leiva, B. (2009). *Amoebiasis*: Diagnosis and prevalence in León Nicaragua. *Institutionen för mikrobiologi, tumör-och cellbiologi/Department of Microbiology, Tumor and Cell Biology.*

Li, J., Yu, F., Guo, H., Xiong, R., Zhang, W., He, F., Zhang, M. & Zhang, P. (2020). Crystal structure of plant PLDα1 reveals catalytic and regulatory mechanisms of eukaryotic phospholipase D. *Cell Research*, 30(1), 61-69.

Luna-Nacar, M., Navarrete-Perea, J., Moguel, B., Bobes, R. J., Laclette, J. P., & Carrero, J. C. (2016). Proteomic study of *Entamoeba histolytica* trophozoites, cysts, and cyst-like structures. *PLoS One*, 11(5), e0156018.

Mahankali, M., Alter, G., & Gomez-Cambronero, J. (2015). Mechanism of enzymatic reaction and protein–protein interactions of PLD from a 3D structural model. *Cellular signalling*, 27(1), 69-81.

Marie, C., & Petri Jr, W. A. (2014). Regulation of virulence of *Entamoeba histolytica*. Annual review of microbiology. 68, 493-520.

Methods of Algorithms in bio-informatique. (n.d.). Retrieved from https://www.phylogeny.fr/

Meza, I., Talamás-Rohana, P., & Vargas, M. A. (2006). The cytoskeleton of *Entamoeba histolytica*: structure, function, and regulation by signaling pathways. *Archives of medical research*, 37(2), 234-243.

Mi-ichi, F., Yoshida, H., & Hamano, S. (2016). *Entamoeba* Encystation: new targets to prevent the transmission of Amebiasis. *PLoS pathogens*, 12(10), e1005845.

Mohamed, M. A., Siddig, E. E., Elaagip, A. H., Edris, A. M. M., & Nasr, A. A. (2016). Parasitic contamination of fresh vegetables sold at central markets in Khartoum state, Sudan. *Annals of Clinical Microbiology and Antimicrobials*, 15(1), 17.

Morf, L., & Singh, U. (2012). *Entamoeba histolytica*: a snapshot of current research and methods for genetic analysis. *Current opinion in microbiology*, 15(4), 469-475.

Naiyer, S., Bhattacharya, A., & Bhattacharya, S. (2019). Advances in *Entamoeba histolytica* biology through transcriptomic analysis. *Frontiers in Microbiology*, 10, 1921.

Nakada-Tsukui, K., & Nozaki, T. (2016). Immune response of amebiasis and immune evasion by *Entamoeba histolytica*. *Frontiers in Immunology*, 7, 175.

Nassoury, N., & Morse, D. (2005). Protein targeting to the chloroplasts of photosynthetic eukaryotes: getting there is half the fun. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1743(1-2), 5-19.

NCBI. (n.d.). Retrieved from https://www.ncbi.nlm.nih.gov/

NCBI. (2020). *Conserved Protein Domain Family*. Retrieved from Catalytic domain, repeat 2, of vertebrate phospholipases, PLD1 and PLD2, yeast PLDs, and similar proteins: https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=197239

NCBI. (n.d.). *Conserved Domain Database*. Retrieved from https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml

Ngui, R., Angal, L., Fakhrurrazi, S. A., Lian, Y. L. A., Ling, L. Y., Ibrahim, J., & Mahmud, R. (2012). Differentiating *Entamoeba histolytica*, *Entamoeba dispar* and *Entamoeba moshkovskii* using nested polymerase chain reaction (PCR) in rural communities in rural communities in Malaysia. *Parasites & vectors*, 5(1), 187.

Organization, W. H. (1969). *Amoebiasis: report of a WHO expert committee.* HM Stationery Office.

Papadakis, M. A., McPhee, S. J., & Rabow, M. W. (2016). *Medical Diagnosis & Treatment.* New York: McGraw-Hill Education.

Prakash, V., & Oliver, T. I. (2019). *Amebic liver abscess.* In *StatPearls [Internet].* StatPearls Publishing..

Promega. (n.d.). *Transforming Bacteria*. Retrieved from promega.com: https://www.promega.com/-/media/files/resources/product-guides/subcloning-notebook/transforming_bacteria_row.pdf?la=en

Protein Database. (n.d.). Retrieved from https://www.rcsb.org/

Ralston, K. S., & Petri Jr, W. A. (2011). Tissue destruction and invasion by *Entamoeba histolytica*. *Trends in parasitology*, 27(6), 254-263.

Rigothier, M. C., Khun, H., Tavares, P., Cardona, A., Huerre, M., & Guillén, N. (2002). Fate of *Entamoeba histolytica* during establishment of amoebic liver abscess analyzed by quantitative radioimaging and histology. *Infection and immunity*, 70(6), 3208-32.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol: chloroform. *Cold Spring Harbor Protocols*, pdb-prot4455.

Selvy, P. E., Lavieri, R. R., Lindsley, C. W., & Brown, H. A. (2011). Phospholipase D: enzymology, functionality, and chemical modulation. *Chemical reviews*, 111(10), 6064-6119.

Serrano-Luna, J., Piña-Vázquez, C., Reyes-López, M., Ortiz-Estrada, G., & de la Garza, M. (2013). Proteases from *Entamoeba* spp. and pathogenic free-living amoebae as virulence factors. *Journal of Tropical Medicine*, 2013.

Shirley, D. A. T., Farr, L., Watanabe, K., & Moonah, S. (2018). A review of the global burden, new diagnostics, and current therapeutics for amebiasis. *In Open forum infectious diseases, US: Oxford University Press*, Vol. 5, No. 7, p. ofy161.

Siegesmund, M. A., Hehl, A. B., & van der Giezen, M. (2011). Mitosomes in trophozoites and cysts of the reptilian parasite *Entamoeba invadens*. *Eukaryotic cell*, 10(11), 1582-1585.

Singh, M., Sharma, S., Bhattacharya, A., & Tatu, U. (2015). Heat shock protein 90 regulates encystation in *Entamoeba*. *Frontiers in microbiology*, 6, 1125.

Singh, N., Ojha, S., Bhattacharya, A., & Bhattacharya, S. (2012). Stable transfection and continuous expression of heterologous genes in *Entamoeba invadens*. *Molecular and biochemical parasitology*, 184(1), 9-12.

SWISS-model. (n.d.). Retrieved from https://swissmodel.expasy.org/

Tazreiter, M. H. (2010). Reaction of the human parasite *Entamoeba histolytica* to metronidazole. *Doctoral dissertation, uniwien*.

Tharmaratnam, T., Kumanan, T., Iskandar, M.A., D'Urzo, K., Gopee-Ramanan, P., Loganathan, M., Tabobondung, T., Tabobondung, T.A., Sivagurunathan, S., Patel, M. & Tobbia. (2020). *Entamoeba histolytica* and amoebic liver abscess in northern Sri Lanka: a public health problem. *Tropical Medicine and Health*, 48(1), 1.

Turfah, M. (2020). Understanding the role of unusual dynamin-related proteins in the lizard pathogen *Entamoeba invadens* . *Doctoral dissertation, University of Exeter*.

Turner, N. A., & Eichinger, D. (2007). *Entamoeba invadens*: the requirement for galactose ligands during encystment. 116(4), 467-474.

Wilson, I.W., Weedall, G.D., Lorenzi, H., Howcroft, T., Hon, C.C., Deloger, M., Guillén, N., Paterson, S., Clark, C.G. & Hall, N. (2019). Genetic diversity and gene family expansions in members of the genus *Entamoeba*. *Genome biology and evolution*, 11(3), 688-705.

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.

Zu, Y., Huang, S., Liao, W. C., Lu, Y., & Wang, S. (2014). Gold nanoparticles enhanced electroporation for mammalian cell transfection. *Journal of biomedical nanotechnology*, 10(6), 982-992.

Zulfiqar, H., Mathew, G., & Horrall, S (2019). *Amebiasis.* In StatPearls [Internet]: StatPearls Publishing.

8   APPENDICES

8.1   Appendix A: Chemicals and media

Incomplete cytomix medium for electroporation:

Ingredients	Quantity (g)
KCl	8.95
$CaCl_2$	0.02
HEPES	5.96
EGTA	0.76
$MgCl_2$	0.48

1000 ml of buffer (10mM K_2PO_4/KH_2PO_4) with pH 7.6 is used to make solution, where, pH of the solution is adjusted to 7.8-7.9 with the help of KOH. This follows with the filter sterilisation of the buffer. Prior to electroporation 20 mg ATP is used to complete the 8 ml cytomix to achieve the concentration of 4 mM and similarly, 24 mg of reduced gluthathione is used for the 10 mM concentration.

LB agar:

Ingredients	Quantity
LB agar	40 g
Distilled water	1 L

Sterilized by autoclave at 121 °C for 15 minutes.

Luria Broth (LB broth):

Ingredients	Quantity
LB broth	25 g
Distilled water	1 L

Autoclaved at 121 °C for 15 minutes.

Ampicillin:

Ingredients	Quantity
Ampicillin	100 µg/ml
milliQ sterilize water	1

Ampicillin was dissolved in milliQ sterilize water and filter-sterilized via 0.2 µm filter. The final concentration used was 100 ug/ml whereas used concentration was 100 mg/ml.

SOC media composition:

Ingredients	Quantity
Tryptone	20 g
Yeast extract	5 g
5 M NaCl	2 ml
1 M KCl	2.5 ml
1 M $MgCl_2$	10 ml
1 M $MgSO_4$	10 ml
1 M glucose	20 ml
H_2O	1000 ml

Phosphate buffered saline (PBS):

Ingredients	Quantity (g/l)
Na_2HPO_4	0.454
KH_2PO_4	0.068
KCl	0.097
NaCl	7.89
H_2O	1
pH	7.4

Vitamin mixture #18 (Diamond and & CLARK 1993)

Solution 1

Ingredients	Quantity
Niacinamide	45 mg
Pyridoxal hydrochloride	4 mg
Pantothenic acid	23 mg
Thiamine hydrochloride	5 mg
Vitamin B12	1.2 mg
H_2O	25 ml

Solution 2

Ingredients	Quantity
Riboflavin	7 mg
0.1 M NaOH	Minimum amount
H_2O	45 ml

Solution 3

Ingredients	Quantity
Folic acid	5.5 mg
0.1 M NaOH	Minimum amount
H_2O	45 ml

Solution 4

Ingredients	Quantity
D-biotin	2 mg
H_2O	45 ml

Solutions 1-4 were combined in glass tube.

Solution 5

Ingredients	Quantity
DL-6-8-thioctic acid (oxidized form)	1 mg
95 % ethanol	5 ml
Tween-80	500 mg
H_2O	30 ml

Combined solutions of 1-4 was mixed with the solution 5 and final volume adjusted to 200 ml distilled H_2O. Later, autoclaved for 15 minutes at the temperature of 121 ^oC.

DNA Marker

DNA fragments with the size greater than 400 bp was dealt with the Bioline hyperladder I, as Figure 23
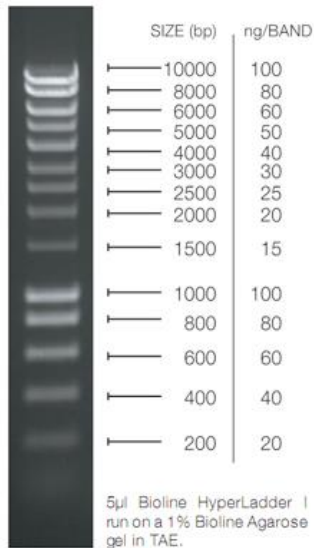
Figure 23. Bioline hyperladder 1

A DNA ladder by the Bioline Hyperladder I as shown in Figure 23, was used throughout this study.

Calcofluor white stain (1%):

Ingredients	Quantity
Calcofluor white stain	1 ml
sterilized water	99 ml

4% paraformaldehyde solution:

Ingredients	Quantity
paraformaldehyde powder	4 g
sterilized water	100 ml
NaOH	1 M to adjust pH 7.4

8.2 Appendix B: Sequence alignment of cloned pGEM-T-Easy and PLD1/PLD2

Figure 24 and Figure 25 confirms the alignment of cloned pGEM-T-Easy vector with PLD1/PLD2 respectively. SP6 as forward primer, T7 as reverse primers used by Microsynth lab for Sanger cycle sequencing, whereas alignment also showed with gene specific forward and reverse primers used for amplification of PLD1/PLD2. EIN-017100 represents reference sequence of PLD1 (Figure 24) and EIN-196230 represents reference sequence for PLD2 (Figure 25).

pGEMT-Easy+PLD1

pGEMT-Easy
F. Primer
PLD1-SP6
EIN_017100 | Entamoeba invade
PLD1-T7 rev
R. Primer
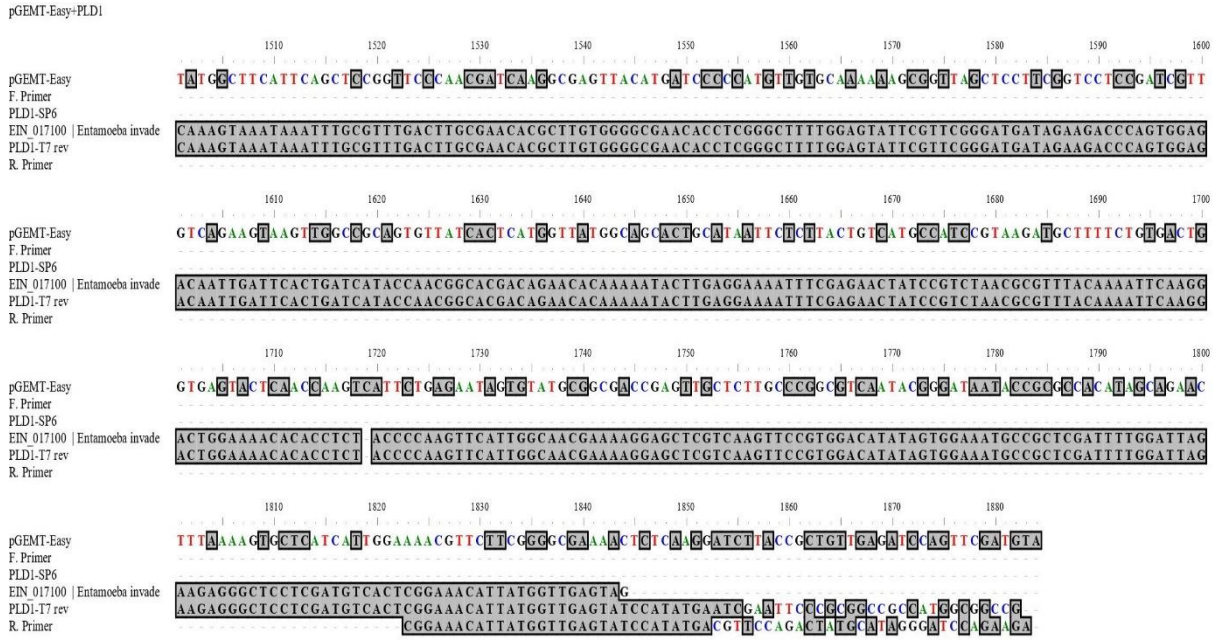
Figure 24. Alignment representation of cloned pGEM-T-Easy+PLD1 along with reference gene and primers

pGEMT-Easy+PLD2

pGEMT-Easy
F.Primer
PLD2-SP6
EIN_196230 | Entamoeba invade
PLD2-T7 rev
R.Primer

Figure 25. Alignment representation of cloned pGEM-T-Easy+PLD2 along with reference gene and primers