UiT The Arctic University of Norway

The Norwegian College of Fishery Science (Faculty of Biosciences, Fisheries and Economy)

Exploring the Biotechnological Potential of the Secretome from the Marine Fungus *Digitatispora marina*

Proteomics, Bioinformatics and Heterologous Expression of Novel Proteases

Elise Nilsen Heimland

Master's thesis in Marine Biotechnology (BIO-3901) May 2022

Acknowledgements

This master thesis was conducted at ArcticZymes Technologies ASA, Marbio, Norwegian College of Fishery Science and the Department of Chemistry at the Arctic University of Tromsø, UiT.

First and foremost, I would like to express my gratitude to my supervisors Teppo Rämä, Bernd Ketelsen Striberny and Peik Haugen for all the advice and guidance provided throughout the work done in this thesis. I would like to thank my "laboratory supervisor" at ArcticZymes, Sigurd Eidem Gundesø, and also Jenny Johansson Söderberg for helping me, and providing me with the information needed to do the laboratory experiments.

Thank you to my family and friends for all of your encouragement. And to all my study colleagues, I appreciate the motivating conversations and support throughout the master thesis work.

Tromsø, May 2022

Elise Nilsen Heimland

Abstract

Cold-adapted enzymes are of biotechnological value due to their high catalytic powers at low to moderate temperatures and their high thermolability at elevated temperatures. Marine fungi are an understudied group of organisms, which makes them interesting as enzyme producers because of their adaptions to marine habitats. In this thesis study, a bioprospecting pipeline was constructed, and different approaches were assembled to identify the genetic source in the marine fungus *Digitatispora marina* that is responsible for previously observed protease activity in growth medium. Biochemical characterization in form of temperature optimum, salt tolerance and pH optimum were made to acquire information about the properties of the protease(s). The activity was maintained at a range of different NaCl concentrations and pH values of Tris-HCl. Moreover, the protease(s) were also active at lower temperatures with an activity peak at 50°C. Protein extracts from the fungal culture medium were also analyzed by Mass spectrometry for possible identification of proteases contributing to the protease activity. Through bioinformatic analysis, a selection of seven serine proteases (SP1, SP2, SP3, SP4, SP5, SP6 and SP7) and one metalloprotease (MP) was chosen for recombinant protein expression in *E. coli* BL21(DE3) cells. There was a successful expression of SP1, however the activity was not recovered in any of the samples. Further studies are needed to identify the genetic source that is responsible for the protease activity from the secretome of *D. marina.*

Table of contents

Abbreviations

E_a = Activation energy

AZ = ArcticZymes

EQN = Equation

IPTG = Isopropyl-β-D-thiogalactoside

Vmax = Maximal Velocity

OD = Optical density

kDa = Kilo Daltons

PCR = Polymerase chain reaction

SDS-PAGE = Sodium dodecyl sulphate polyacrylamide gel electrophoresis

EDTA = Ethylenediaminetetraacetic acid

PMSF = phenylmethylsulphonyl fluoride

Bp = Base pair

ITS = Internal transcribed spacer

MS = Mass Spectrometry

LB = Lysogeny Broth (medium)

O/N = Overnight

L = Liter

mL = Milliliter

M = Molar

mM = Millimolar

µL = Microliter

µg = Microgram

RPM = Revolutions per minute

x g = Times gravity

TB = Terrific Broth (medium)

1 Introduction

1.1 Marine bioprospecting – different approaches

The North Atlantic Ocean and the Arctic Ocean surrounding the northern parts of Norway generate a cold and seasonally dark environment where the organism's ability to adapt is crucial for survival (1). Microorganisms survive by producing a range of specialized compounds such as secondary metabolites and enzymes that provide advantages at, and protection against cold conditions but also against predators and microbial invaders (1-3). During millions of years of evolution these types of enzymes and chemical compounds have been optimized through natural selection. Today, we can take advantage of their properties by using them as candidates for medical and industrial applications (1, 4).

The continuous demand within modern biotechnology to discover, among others, new novel biocatalysts for use in biocatalytic industrial processes, compels researchers into finding quicker and more accessible solutions for discovering novel enzymes and compounds (5). Marine bioprospecting describes a purposeful and systematic search for new natural chemical compounds, genes or biomolecules derived from the marine environment, with the intention of achieving commercial or societal exploitation of them (6). The bioprospecting workflow can be divided into "top-down" and "bottom-up" approaches that represent different strategies for discovery of novel natural products (5, 7). The top-down approach starts with methods unveiling biological and chemical characterization of compounds (e.g. enzymes) produced by an organism (5). Top-down methods have traditionally been the main strategies for discovering new natural products. This pipeline starts with collecting biological samples for either direct extraction or laboratory cultivation. Retrieved extracts are screened for bioactivity where any "hits" are isolated for structural characterization. The traditional top-down approach has successfully identified several natural products such as for example the Shrimp Alkaline Phosphatase (SAP) produced by ArcticZymes, that was discovered in *P. borealis* shrimp in the late 1980´s (8).

However, many organisms cannot be grown in pure culture at laboratories meaning that they cannot be exploited for biotechnological purposes using the top-down approach (9). In contrast, the bottom-up approach uses genetic information to evaluate the biotechnological potential of the genetic inventory of microorganisms based on bioinformatics and molecular biology techniques (5, 7). This metagenomics-based approach allows cloning of environmental DNA,

screening for desired function and recombinant production of the gene of interest without cultivating the original organism (5).

Moreover, genomics-based bottom-up approaches have revolutionized the potential of discovering and developing new natural products that could not be detected under standard fermentation conditions (7, 10). These strategies contain powerful bioinformatics and genetic manipulating tools that are used to identify genes of interest by expressing them in native or heterologous hosts (7). Bottom-up approaches can roughly be categorized as either function-based screening or sequence-based screening strategies (Figure 1).

1.1.1 Function- and Sequence-based bottom-up strategies

During function-based approaches, cloned metagenomic libraries are constructed and sequences are expressed in a suitable host organism. To identify any desired traits, activity screening is mostly done using activity assays such as spectrophotometric assays (11). Disadvantages of function-based strategies is the low detection frequency of functional proteins. This is connected to difficulties in recombinant expression due to inefficient transcription and translation of the recombinant gene, incorrect folding and/or insufficient secretion of the cloned protein, decreased protein stability and activity due to the lack of necessary cofactors, and presence of inhibitors or other enzyme activities that is endogenous in the expression host affecting proper activity assessment (9). Moreover, the protein product may also pose toxic traits to the host organism and may, when over-expressed, deplete the cells energy storages blocking the growth of the host, or ultimately kill it (12).

Regarding the sequence-based approaches, gene mining, hybridization and direct PCR are used for novel discoveries. Bioinformatic tools are inexpensive and easily accessible, allowing researchers to find sequence alignments between metagenomic DNA libraries and annotated sequences of interest through a variety of databases on any computer. Other advantages are related to the possibility of polishing and modifying protein coding genes prior to their recombinant expression initiation elements in expression vectors (e.g. promotor), addition of protein purification-aiding elements, host-specific codon optimization, and a certain predictability of project size and cost. A downside to sequence-based screenings is the inability to detect novel genes or enzymes as the screenings are based on conserved and already described sequence signatures (9).

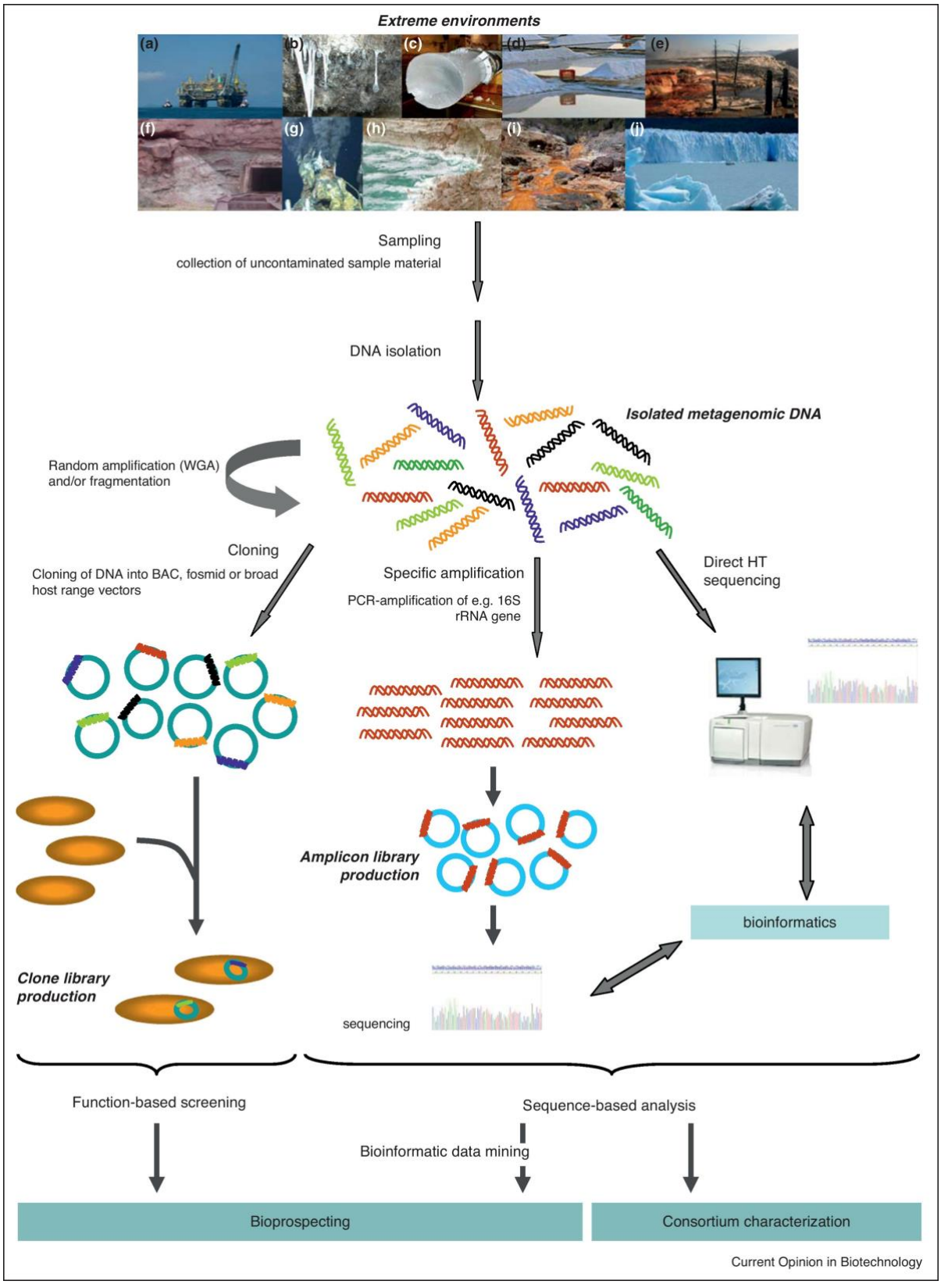

Figure 1: Schematic bioprospecting pipeline presenting the bottom-up approach from sample collection to function- and sequence-based strategies. The figure is retrieved from Lewin et al. (13).

1.2 Marine fungi

Fungi are heterotrophic organisms which obtain their needed energy (for growth and reproduction) from other organisms in the form of nutrients bound to organic matter produced by living organisms or dead organic matter (14). Fungi are known for contributing to the nutrient cycle in ecosystems with an important role as symbionts associated with plants and animals (14).

Fungi constitutes a significantly large Kingdom within the Eukarya domain. Presumedly, fungi first evolved as aquatic organisms between 800-2 400 million years ago because phyla in the basal fungal taxa (*Opisthosporidia* and *Chytridiomycota*) are unicellular flagellated organisms that mostly occur in aquatic environments (15, 16). Later, fungi were the first eukaryotes adapting to life on land (16, 17). These prodigious organisms can cope under different redox conditions with the majority evidently preferring oxic environments although some species inhabit oxygen minimum zones (18, 19). Fungal morphology may be filamentous (e.g., mushrooms), unicellular (e.g., yeasts) or dimorphic (exists in two forms: filamentous hyphae forming or yeasts) (18, 20). Fungi are omnipresent in terrestrial, freshwater and marine environments (21). Readaptation to marine environments have occurred in several linages in the *Dikarya* taxa that consists of the two most derived fungal lineages *Basidiomycota* and *Ascomycota* (16, 22). These two phyla have the ability to form multicellular macroscopic fruiting bodies, and differs in their mycelium septation and the structure of the fruiting body and hymenium where sexual spores are formed (16).

There are various definitions to what marine fungi are (23), among them, the most prominent definition was proposed by Kohlmeyer and Kohlmeyer (1979): *"obligate marine fungi are those that grow and sporulate exclusively in a marine or estuarine habitat, and facultative marine fungi are those from freshwater and terrestrial milieus able to grow and possibly also sporulate in the marine environment"* (24, 25). However, in later decades when molecular techniques and general high throughput sequencing made an entrance, scientists such as Jones (2011) suggested that this definition was too narrow. In 2014 Overy et al., included another term: "marine-derived fungi", which simply indicates that a fungus has been isolated from the marine environment. The current prevailing broad definition of marine fungi is proposed by Pang et al., (2016): *"Any fungus that is recovered repeatedly from marine habitats because: 1) it is able to grow and/or sporulate (on substrata) in marine environments; 2) it forms symbiotic*

relationships with other marine organisms; or 3) it is shown to adapt and evolve at the genetic level or be metabolically active in marine environments" (26).

Marine fungi are an understudied, biotechnologically valuable group of organisms. Due to the vast range of different habitats fungi live in, and the consecutive competition against a diverse array of bacteria, animals, and other fungi, they have developed several mechanisms to compete and survive. Thus, the unique attributes of fungi are propitious for application in biotechnology and industry (27). Marine fungi function as parasites at many trophic levels. In addition, they function as major saprotrophs in oceans by converting detritus into fungal biomass that is consumed by zooplankton (28). The vast majority of fungi identified from marine environments belong to the *Ascomycota* and *Chytridiomycota*, while fairly less *Basidiomycota* and other fungal phyla are reported (29, 30).

The commercial use of fungi have been reported for multiple industrial sectors (31). In general, marine fungi produce a variety of extracellular enzymes, such as proteases, amylases, laccase, cellulases and xylanases that can be applied in a multitude of fields, including medicines, detergents and food and beverages (32). Marine fungi are notable producers of a diverse range of novel enzymes with biological importance (33). They are compelling for researchers as the enzymes obtained from marine fungi are likely to differ from enzymes derived from terrestrial fungi due to their differences in taxonomic diversity and adaptions to the stressful marine environment (32, 34). Marine fungi are unique because of their novel physiological characteristics, such as high salt tolerance, cold adaptivity, barophilicity and their capability of high-level synthesizing of interesting enzymes (33, 35). However, the molecular and physiological processes behind fungal growth are complex and not easily observed. Due to the lack of knowledge of factors that affect fungal growth, successful culturing mimicking natural conditions in the laboratory often fail. The morphology of filamentous organisms in submerged cultivation is of considerable interest due to its impact on process productivity (31).

1.3 Enzymes

Enzymes are proteins with catalytic properties that have an active site containing specific amino acids which allows the enzymes to be highly specific in the recognition, binding and catalyzation of a substrate (36). Enzymes are not consumed by the reactions they catalyze, nor do they alter the equilibrium of these reactions. Therefore, when the product is removed from the active site, the enzyme is ready for a new substrate to bind and repeat the process. Enzymes are central to every biochemical process, and they accelerate chemical reactions immensely. They also function under various conditions of temperature and pH (36).

Activation energy (E_a) is defined as the minimum amount of energy that is needed to activate atoms or molecules to a condition in which they can undergo chemical transformation or physical transport (37). Like other catalysts, enzymes lower the E_a in a reaction. Hence, the accelerating rate of the reaction increases drastically. Factors that affect the activity of an enzyme include substrate concentrations [S], temperature, pH, ionic strength, and nature of salts present. Activity is measured as the initial rate of substrate utilization of an enzyme reaction. Meaning that it is the amount of product made per unit time at the beginning of the reaction that is measured (when the product concentration is increasing linearly). The initial rate of an enzyme reaction (V_0) can be described by the Michaelis-Menten equation (EQN 1):

$$V_0 = \frac{V_{max}[S]}{K_m + [S]} \qquad \text{(EQN 1)}$$

Vmax and Km are kinetic constants, in which Vmax is defined as the maximum reaction rate of the enzyme at saturated [S], and Km is regarded as the substrate concentration in which the enzyme has a reaction rate of half Vmax. Km indicates how efficiently an enzyme binds to the substrate and converts it into product.

Commercially produced enzymes are being extensively used in various sectors to increase quality, speed and yield of different industrial processes, as well as reducing energy consumption and the usage of hazardous chemicals (38). Moreover, enzymes are of fundamental interest in health science because of the high frequency of disease-processes that can be linked to deviant activities of one or more enzymes (39). Enzymes are also increasingly replacing conventional chemical catalysts used in industrial processes because they are more convenient through their ability to function under relatively mild conditions of pH, salinity, temperature, and pressure (40).

1.3.1 Cold-adapted enzymes

More than three-quarters of the Earth consists of cold environments with temperatures that are often below 5°C (41). Cold environments include deep oceans and alpine- and polar-regions (42). Psychrophiles are organisms that are capable of growth in these low temperature environments, and occur as either obligate- or facultative psychrophiles (43). They adapt to low temperatures through a variety of cellular processes, which includes the synthesis of cold shock- and cold acclimation enzymes (42).

Cold-adapted enzymes ensure energy conservation by performing reactions at low temperatures which would otherwise consume a lot of energy if performed at high temperature (44). Heat lability is a general trait of cold adapted enzymes, meaning that they have an unstable active site and are therefore easily inactivated upon the application of heat. Therefore, the inactivation rate constants of heat labile enzymes are immensely faster than their mesophilic or thermophilic homologues (41).

Because of their high enzymatic activity at low temperatures and low thermal stability, cold-adapted enzymes possess high potential for biotechnological applications (45). This includes the considerable economic benefits for example through energy savings in large scale processes that would not require additional heating reactors (41). Cold-adapted enzymes are already applied in various industries where they are used as additives in bioremediation, in food industry, molecular biology and detergent and textile industry (41, 46). Figure 2 gives an overview of the applications of cold-adapted enzymes in various industries.

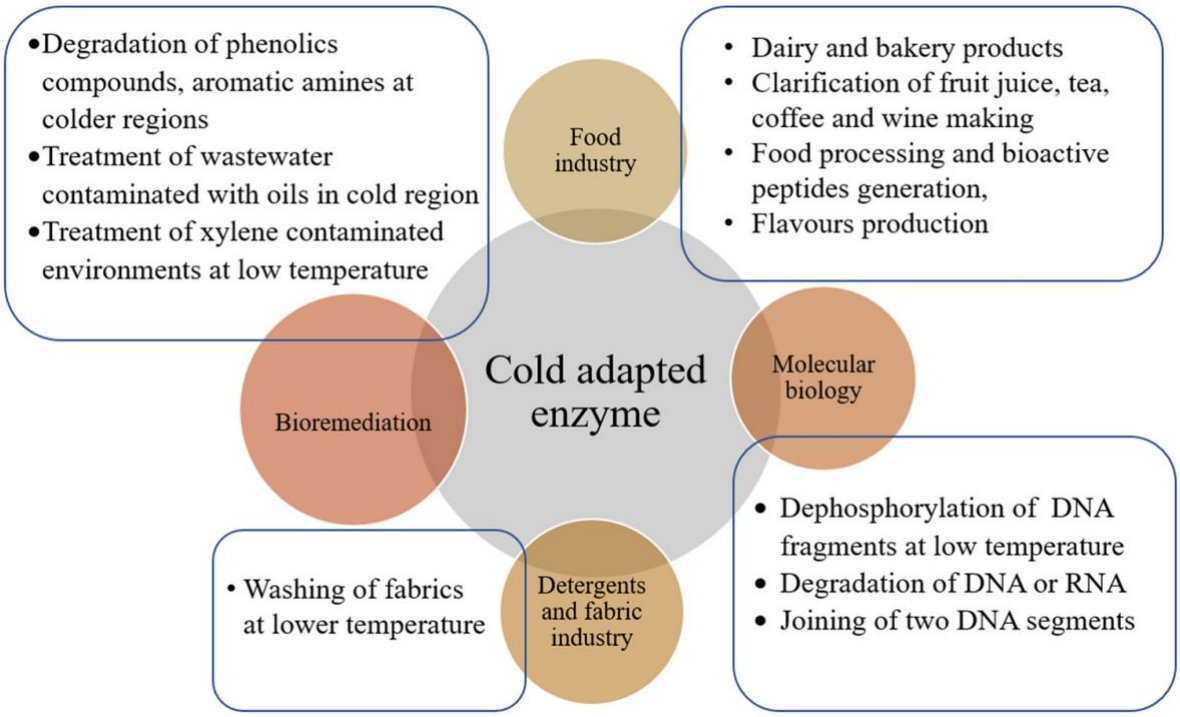

Figure 2: An overview of the applications of cold-adapted enzymes in various industries. The figure is retrieved from Kumar et.al (46).

1.4 Proteases

Proteases are enzymes that catalyze the irreversible cleavage of peptide bonds in proteins and oligopeptides (47). This occurs via a nucleophilic attack on the carbonyl carbon, followed by hydrolysis of the peptide bonds (amide bond) (47). Some alternative names for proteases are proteolytic enzymes, peptidases and proteinases (48). In this thesis, these enzymes will be referred to as proteases.

Proteases cuts peptides into fragments which usually would lead to inactivation but sometimes these fragments fold into functional proteins (49). The protease action is requisite in several of the physiological processes such as digestion of food proteins, mitosis and meiosis, protein turnover (renewal or replacement of protein), cell signaling, processing of polypeptide hormones, apoptosis and the life-cycle of some disease-causing organisms including the replication of retroviruses (48, 50). Because of their vital role in all these processes, they are of great importance for pharmaceutical, medical, and academic purposes (48, 51, 52).

Thanks to their importance in biological systems, proteases have been studied intensively and a number of different proteases have been identified and characterized (53). The MEROPS database is an information resource for proteases and the proteins that inhibit them.

Based on the different mechanism used to hydrolyze peptide bonds and depending on the nature of their active sites, the MEROPS database have classified proteases into 9 groups: **Serine (S),** Aspartic (A), Threonine (T), **Metallo (M),** Cysteine (C), Glutamic (G), Asparagine (N), Mixed (P) and Unknown (U) proteases (54). *In this thesis, the focus will mostly be on serine proteases as well as one metalloprotease.* Further, proteases are also termed as either endo- or exo-proteases, where the endo-proteases hydrolyze peptide bonds internally in peptide chains and the resulting products are oligopeptides, whereas exo-proteases hydrolyze peptide bonds near the terminal ends, and the products of this activity are free amino acids, tri- and di-peptides.

1.4.1 Heat-labile proteases in molecular biology

Several molecular biology approaches depend on heat treatments to remove enzymatic activity (55). For example, some restriction endonucleases are inactivated by heating a digestion reaction to 65°C. During the heat inactivation of the endonucleases, further manipulation of the DNA can occur (55). This poses a problem if the enzyme used in the procedure is not inactivated at temperatures low enough to preserve DNA in a double-stranded form (55).

Heat-labile proteases are highly coveted products as they may be used to digest molecular biology enzymes. Subsequently, a moderate heat treatment would inactivate the protease without damaging the double-stranded DNA (55). Therefore, it is suggested that using heat-labile proteases are beneficial in procedures as further reactions can be accomplished without an intermediate purification step, thereby saving time and avoiding the possibility of product loss (55).

1.4.2 Serine proteases

Serine endo- and exo- proteases are ubiquitous in nature and found in eukaryotes, viruses, archaea and bacteria (56). This group of proteases have a serine residue on their active site and binds covalently to the substrate (53). Their activity optimum is usually in neutral and alkaline conditions (53). They are functional in a range of mechanisms in the body such as blood clotting, inflammation, immunity, and contributes digestive enzymes in both prokaryotes and eukaryotes (57). Serine proteases use the catalytic triad, Ser/His/Asp, to break a peptide bond. The serine is the nucleophile, histidine is the general base and acid, and aspartate helps orient the histidine residue and neutralizes the charge taken by the histidine during the transition state (58, 59).

1.4.3 Metalloproteases

For metalloproteases, the nucleophilic attack on a peptide bond is mediated by a water molecule. The water molecule is activated by a divalent metal cation (usually zinc), and the product of the reaction is a protonated amine and an ionized carboxylate group at neutral pH values (60). In metalloproteases, metal ions are the cofactors at the active site of the enzyme where they function as strong electrophiles that assist the attack initiated by the water molecule (53, 61).

1.5 Recombinant T7 expression of heterologous proteins in *Escherichia coli*

There are several different host organisms available for heterologous protein expression, including bacteria, yeast and filamentous fungi. *Escherichia coli* is frequently used as a host organism because of advantages which include fast and high-density cultivation, well known genetics, ease of genetic manipulation, large number of compatible molecular tools available, and inexpensiveness (62, 63). Moreover, the variety of available plasmids, recombinant fusion partners and mutant strains have advanced the possibilities with *E.coli* as a host organism (63).

In spite of all these advantages, the major drawbacks of using *E. coli* as the recombinant host organism, are its lack of secretion systems for efficient protein-release to growth medium, restricted capability to facilitate extensive disulfide-bond formation and other posttranslational modifications and inefficient cleavage of the amino terminal methionine. These impediments may result in lowered protein stability and increased immunogenicity, as well as deficient folding due to absence of specific molecular chaperones (64-66). Moreover, overexpression of heterologous proteins often results in the production of protein aggregates due to overproduction and improper folding of the expressed proteins, leading to biologically inactive and insoluble proteins (62).

The T7 system is a commonly used approach for expressing recombinant proteins in *E. coli* strains. In this system, the target gene is cloned into an expression vector downstream of the T7 promoter and this construct is introduced into the expression host (67). The plasmid vector should contain a strong and inducible promoter, cloning site, origin of replication, antibiotic resistance marker, ribosome binding site, translational start site, transcription terminator and translation terminator (67). The pET plasmids vectors are examples of T7 based expression systems (68).

The T7 expression system requires a promoter (T7 promoter) to control over expression of the cloned gene (prior to induction) to avoid leaky expression. Leaky expression means that the cloned gene is expressed prior to induction and are caused by basal expression of T7 RNA polymerase from uninduced host cells (68). However, most expression vectors that use the T7 expression system are designed for minimizing leaky expressions (68). For starting the heterologous protein production, the T7 promoter needs to be induced by an inducer such as isopropyl-β-D-thiogalactoside (IPTG). The IPTG will bind to the repressor protein and alter its conformation. The altered repressor is displaced from the operator allowing the T7 RNA polymerase to bind to the promoter which then stimulates the transcription of the gene of interest (67). Figure 3 gives an illustration of the protein regulation in the T7 expression system.

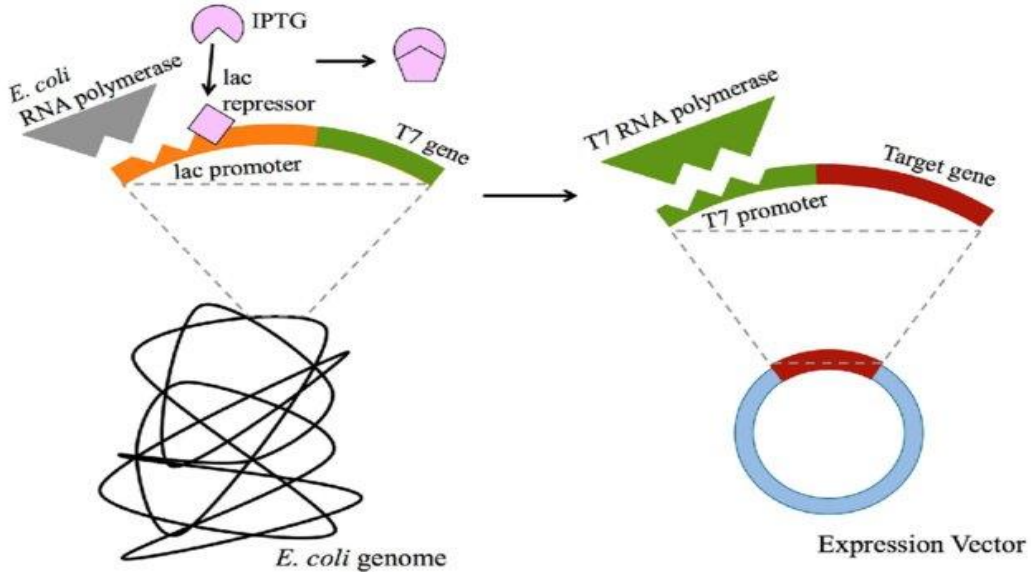

Figure 3: Strategy for regulating protein expression using the T7 expression system. The gene encoding the T7 RNA polymerase is inserted into the chromosome of E. coli and transcribed from the lac promoter and is only expressed if the inducer (IPTG) is present. The T7 RNA polymerase then transcribes the gene of interest (target gene) which is cloned into the pET plasmid vector. In absence of IPTG, the lac operator (positioned between T7 promoter and the cloned gene) reduces transcription of the cloned gene. Figure is adapted by author from Novagen pET System Manual and retrieved from D. Eilerts (69).

2 Project background and workflow

2.1 Project background

The master project is based on a previous project collaboration between the research groups of Marbio and ArcticZymes titled "A new enzymatic Screening Pipeline for exploring a unique marine Fungi collection (FunScreen)" which aimed to discover new proteases and nucleases from marine fungi. Interesting protease activity was found in the growth medium of *Digitatispora marina*, and this activity is the starting point for this thesis project.

2.1.1 Fungal species investigated: *Digitatispora marina*

The genus *Digitatispora* was first described by Gaston Doguet in 1962 (70, 71). The genus consist of two species, *D. marina* and *D. lignicola*, (71, 72). *Digitatispora marina* is one of few species of Agaricomycetes described from marine environments. Agaricomycetes is a class of fungi in the division *Basidiomycota*, which are rare in marine ecosystems (23, 24). From a survey done in 2014, Rämä et al. identified 28 filamentous species of marine fungi, where *D. marina* was the only basidiomycete (71, 73). Other surveys also substantiates that there are few filamentous *Basidiomycota* present in marine habitats: In 2020, Tibell et al. revealed that only 2 (one of which was *D.* marina) out of 77 recorded marine fungi species in the Baltic Sea belonged to *Basidiomycota* (71). *D. marina* forms small, thin and crust-like fruiting bodies on wood material in temperate waters. The fungus has developed a white-rot fungal strategy for decomposing wood (74, 75), meaning that it is able to degrade lignin with high efficiency due to the production of ligninolytic extracellular oxidative enzymes (76). Moreover, the fungus has shown the highest decomposition rate among marine white rot fungi at temperatures below 10°C (74, 75). The fungus has also been suggested to be a relevant enzyme source for biofuel purposes, as well as showing high cellulase, laccase and peroxidase activity in seawater (74, 75).

2.2 Workflow

All laboratory work was performed at Marbio and ArcticZymes Technologies ASA, except for MS-analysis which was done by external engineers at Proteomics and Metabolomics Core Facility (PRiME) at the Faculty of Health Sciences (UiT). In addition, the construction of the sequence library (containing all proteins present in *D. marina*) was done by Prof. Peik Haugen at the Department of Chemistry (UiT). The workflow of this project is illustrated in Figure 4.

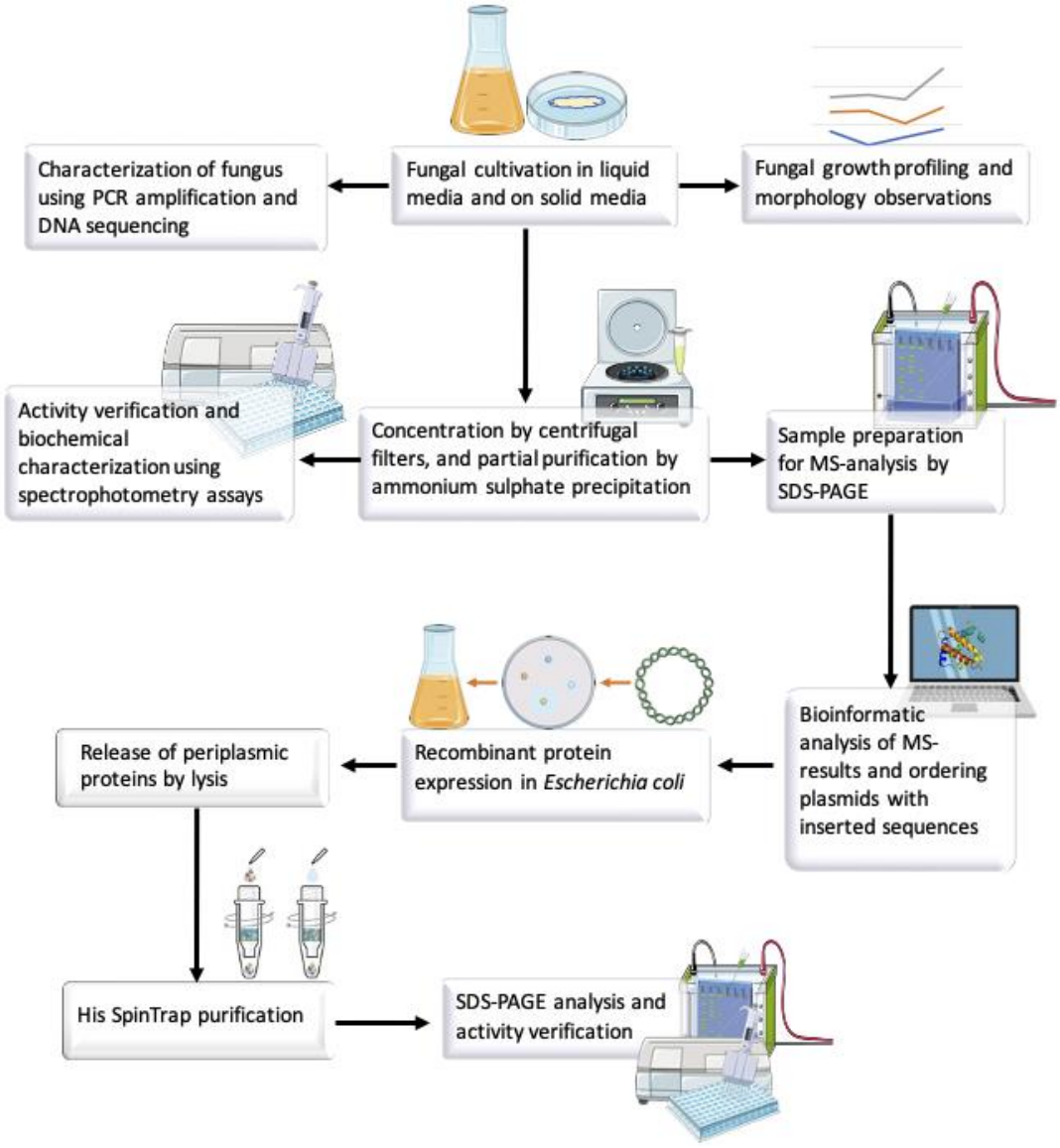

Figure 4: Workflow for the work conducted in this thesis. The fungus, Digitatispora marina, *was cultivated in different solid and liquid media, and growth and morphology studies were performed at Marbio. The remainder of the work was performed at ArcticZymes. The fungal supernatants from the liquid cultures were concentrated and partially purified and tested for activity. Fungal protein samples were analyzed by Mass Spectrometry, and a bioinformatic analysis was conducted for sequence selection and polishing before ordering sequence synthesis and sub-cloning into expression plasmids. Then it was performed recombinant protein expression in E. coli which was lysed and purified with His SpinTrap purification. In addition, SDS-PAGE was run to see if the proteins had been successfully expressed, and the His-tag purified samples were again tested for activity. Biochemical characterization of the protease(s) was also done by performing spectrophotometric assays. Illustration used in this workflow figure is retrieved from Servier Medical Art medical (77) and Thermo Fisher Scientific (78).*

3 Aim of the thesis

The aim of this thesis was to construct and execute a bioprospecting pipeline where different approaches were assembled to identify the genetic source in *Digitatispora marina* that is responsible for previously observed protease activity in growth medium. The key objectives of the thesis were to:

- Succeed in sufficient cultivation of *D. marina* both on solid media (to carry out growth- and morphology- studies) and liquid media (to verify and characterize the observed protease activity)
- Purify and concentrate the protease from fungal growth medium using standard procedures established by ArcticZymes.
- Construct a protein sequence library of *D. marina*
- Analyze protein extracts from culture medium by using Mass spectrometry (MS) for identification of proteases contributing to protease activity in culture medium.
- Using bioinformatic tools for analysis and selection of potential protease sequence candidates for heterologous expression
- Recombinant protein expression in *E. coli* and subsequent protein isolation and verification of activity

4 Materials and methods

The specimen of *Digitatispora marina* used in this thesis was collected and isolated from driftwood found at Vannøya, Norway, by Dr. Teppo Rämä 16th of May 2010.

4.1 Inoculation of fungal cultures (initial cultivation)

Table 1: Products and equipment used in the initial cultivation, solid media cultivation and liquid media fermentation of Digitatispora marina.

Product/Equipment	**Distributor (Country)**
Herasafe biological safety cabinet	Thermo Fisher Scientific (Massachusetts, USA)
Termaks Cooling Incubator	Termaks AS, (Norway)
Petri dishes, 9 cm diameter	VWR International (Pennsylvania, USA)
Filtered seawater	Norwegian College of Fishery Science, UiT (Norway)
Mannitol	Sigma-Aldrich (France)
Agar	Millipore Sigma-Aldrich (Spain)
Peptone	Sigma-Aldrich (Germany)
Homogenized birch wood	Biltema (Norway)
Corn Meal Agar	Sigma-Aldrich (India)
Sea salts	Sigma-Aldrich (Germany)
Moss Malt Extract	Jensen & Co (Norway)
Difco Marine Broth	Becton, Dickinson and Company (New Jersey, USA)
Inoculation needle	Thermo Fisher Scientific (UK)
1-liter Erlenmeyer flasks	Sigma-Aldrich (Germany)
Refrigerator	Electrolux (Sweden)

Cultures of *Digitatispora marina* (isolate id: 008cD1.1) stored in agar lumps in 20% sterile filtered glycerol solution at -80°C, was thawed for 15 minutes and inoculated on three M19 agar plates (composition is described in Table 2) under a sterile hood (Herasafe biological safety cabinet). The inoculated fungi cultures were incubated in a Termaks Cooling Incubator at 10°C until the growth was adequate for further cultivation on different solid media and fermentation in liquid media. Agar pieces of the fungus was preserved using cryopreservation (method described in Appendix A).

4.2 Solid media cultivation of *D. marina* for comparing growth and morphology

Digitatispora marina was cultivated in three different solid media. In addition, all media had two variants, with and without homogenized wood. The media compositions are displayed in Table 2. All media were prepared with Milli-Q water and autoclaved for 25 minutes at 121°C. Petri dishes were filled under a sterile hood with approximately 23 mL of media in each dish.

Table 2: Overview of content in solid media used for cultivating D. marina*. Name of solids, types of materials/solutions and amount of each material/solution is displayed in the table.*

Name of solid media	Materials/solutions	Amount
M19 (W)	Milli-Q water	450,0 mL
	Filtered seawater	450,0 mL
	Mannitol	18,0 g
	Peptone	18,0 g
	Agar	13,50 g
	(Homogenized birch wood)	(0,5 g)
DCAA (W)	Milli-Q water	1,0 L
	Corn Meal Agar	17,0 g
	Sea salt	40,0 g
	(Homogenized birch wood)	(0,5 g)
D2MAA (W)	Milli-Q water	1,0 L
	Moss Malt Extract	4,0 g
	Sea Salts	40,0 g
	Agar	15,0 g
	(Homogenized birch wood)	(0,5 g)

Six petri dishes of each of the 6 different solid media: M19, M19W (with homogenized wood), DCAA, DCAAW (with homogenized wood), D2MAA and D2MAAW (with homogenized wood) were added one inoculum from the established solid cultures under a sterile hood (Herasafe biological safety cabinet) using a sterile inoculation needle. Half of the solid cultures were incubated at 4°C and the rest was incubated in Termaks Cooling Incubator at 10°C.

Monitoring of growth and morphology of the solid *D. marina* cultures was done by measuring the diameter of the fungal mycelia every seven days until one or more of the plates were completely covered with the fungus. In addition, differences in fungal colony shape and structure were inspected.

4.3 Fermentation of the fungus in different liquid media

The following media were prepared for fermentation from inoculated on-plate cultures of *Digitatispora marina:* FMAP, DMA and D2MA. The liquid media compositions are displayed in Table 3. All media were divided equally (250 mL) to four 1-liter Erlenmeyer flasks each. The flasks were autoclaved for 25 minutes at 121°C and further cooled down under laminar flow.

Table 3: Overview of content in liquid media used for fermenting D. marina*. Name of solids, types of materials/solutions and amount of each material/solution is displayed in the table.*

Name of liquid media	Materials/solutions	Amount
FMAP	Distilled water	700,0 mL
	Filtered seawater	300,0 mL
	Difco Marine Broth	15,0 g
	Peptone	5,0 g
DMA	Distilled water	1,0 L
	Moss Malt Extract	20,0 g
	Sea salt	40,0 g
D2MA	Distilled water	1,0 L
	Moss Malt Extract	4,0 g
	Sea Salts	40,0 g

To each flask, ten inoculum pieces were added from the established inoculated cultures under a sterile hood (Herasafe biological safety cabinet) using a sterile inoculation needle. All the liquid cultures were incubated at 10°C.

4.4 Characterization of fungal strains using PCR amplification and DNA sequencing

Table 4: Reagents and equipment used to identify the fungus and check for contaminants by PCR amplification and DNA sequencing.

Reagents/Equipment	Distributor (Country)
DreamTaq Green PCR Master Mix (2X)	Thermo Fisher Scientific (Massachusetts, USA)
Forward primer ITS4	Sigma-Aldrich (Missouri, USA)
Reverse primer ITS5	Sigma-Aldrich (Missouri, USA)
Gel Red (10,000 x)	BioTium, (California, USA)
UltraPure™ Agarose	Life Technologies, (California, USA)
10 x TBE	Life Technologies, (California, USA)
1kb DNA ladder	Life Technologies, (California, USA)
Agarose gel loading dye (6x)	Amresco®, (Ohio, USA)
BigDye 3.1	Sequencing lab at UiT (Norway)
5x sequencing buffer	Sequencing lab at UiT (Norway)
Agarose Gel electrophorese system	Thermo Fisher Scientific (Massachusetts, USA)
GeneFlash® Gel image system	SYNGENE Bio imaging, (UK)
PCR machine	

To verify the identity of the cultivated fungus and exclude the possibility for a contamination, a culture-dependent Polymerase Chain Reaction (PCR) method was performed. This method enables analysis of the internal transcribed spacer (ITS) sequence from a small amount of fungal sample. The ITS region is the spacer DNA located between the small and large ribosomal

subunit rRNA genes, which is used as DNA barcoding of fungal species (79). Reagents and equipment that were used in this experiment are listed in Table 4.
Mycelium was taken from established solid fungal cultures with an inoculation loop and inoculated in 100 µL Milli-Q water in Eppendorf tubes. The established samples were tested at four different dilutions: undiluted, 10x, 100x, and 1000x.

The samples were vortexed and spun down using a table centrifuge. The 10x and 100x solutions were further incubated at -80°C for 15 minutes to lyse the cells and conceivably release DNA into the supernatant. The amplification PCR reaction mix was prepared for each fungal sample by mixing the following: 1 µL fungal template, 12.5 µL 2x Dream Taq, 1 µL forward primer (ITS4), 1 µL reverse primer (ITS5), and 9.5 µL ddH_2O to a volume of 25 µL for each PCR sample. The amplification PCR reaction was run as specified in Table 5.

Table 5: Cycle scheme of amplification PCR

Initial denaturation		96°C	5 min
Cycle x35	Denature	96°C	30 sec
	Annealing	47°C	30 sec
	Elongation	72°C	1 min (1 min <2 kb products)
Final extension		72°C	10 min
Hold		4°C	∞

The PCR amplification results were determined by gel electrophoresis. A 1% solution of agarose was made by dissolving 1 g of agarose in 100 mL 1x TBE buffer in a microwave. After the solution had cooled down, 10 µL of 10,000x GelRed was added and stirred to mix. The gel was cast using a supplied tray and comb and was left for 30 min to set when cooling off.

Further, 3 µL of 1kb ladder was loaded in the first well of the gel followed by 5 µL of each PCR product to the next wells. The gel was run for 30 min at 200 V for the ladder to separate properly. Then the gel was exposed to UV light and photographed.

Lastly, the DNA sequencing PCR was performed. The sequencing PCR master mix was prepared by mixing the following: 1 µL fungal template, 1 µL BigDye 3.1, 2 µL 5x sequencing buffer, 1 µL (1 µM) forward or reverse primer (in separate samples), and 5 µL ddH_2O to a total volume of 10 µL. The sequencing PCR reaction was run as specified in Table 6.

Table 6: Cycle scheme of DNA sequencing PCR

Initial denaturation		96˚C	1 min
Cycle x25	Denature	96˚C	10 sec
	Annealing	47˚C	5 sec
	Elongation	60°C	1 min
Final extension		60°C	4 min
Hold		4˚C	∞

The products were then delivered to the sequencing service at the University Hospital of North Norway (UNN) for sequencing. The sequence results obtained was compared to sequences available in the NCBI Nucleotide Blast Database (80) to verify if there was a sequence match with *Digitatispora marina.*

4.5 Preparation and up-concentration of supernatant

To examine if the protease activity discovered in the supernatant of cultured *Digitatispora marina* during the FunScreen experiments was reproduced, supernatants were prepared from all three culturing media (FMAP, DMA and D2MA) after two and a half months of incubation. An amount of 45 mL of each culturing media with the growing fungus was added to three falcon tubes under a sterile hood (Herasafe biological safety cabinet). The tubes were centrifuged at 4200 rpm for 15 minutes, and the supernatants were transferred to new falcon tubes which were stored at 10°C.

The supernatants were concentrated using Amicon® Ultra-15 10K centrifugal filters (Merck Millipore Ltd, Ireland) with a 10K molecular weight (10 kDa) cutoff. The samples were centrifuged at 4200 rpm in a Multifuge X3R centrifuge (Thermo Scientific, UK) and the flow through was discarded and the up-concentrated supernatant was transferred to 1,5 mL tubes.

4.6 Partial purification and fractionation of fungal supernatant by Ammonium Sulphate Precipitation

The aim of fractionating the supernatant was to further concentrate and partially purify the supernatant and to determine the activity of the active enzyme(s) in the different fractions. The principle of partial purification in this approach is based on the fact that proteins precipitate at different ammonium sulphate concentrations meaning that it is possible to either precipitate contaminants at certain ammonium sulphate concentrations and keep the protein of interest in solution and by that discard the pellet or it is possible to precipitate the protein of interest and discard contaminating protein that remain in solution. In the latter case the protein of interest

containing pellet will need to be reconstituted in an appropriate buffer. After some initial tests we chose the latter approach in parts because this would also allow for simple up-concentration of our protein of interest. A dilution buffer (described in Table 7) was prepared for each of the three supernatants (FMAP, D2MA and DMA).

Table 7: Dilution buffer (10 mL) for ammonium sulphate fractioning of the fungal supernatants. The dilution buffer contained 1M Tris/HCl, 1M NaCl, concentrated supernatant and Milli-Q water

Solutions	**Concentration**
1M Tris/HCl pH 8.0 (ArcticZymes, NO)	75 mM
1 M NaCl (ArcticZymes, NO)	25 mM
Concentrated supernatant	20% v/v
Milli-Q water	Until 10 mL

For each supernatant, six 1,5 mL tubes were added 1 mL of the dilution sample and solid ammonium sulphate ($(NH4)_2SO^-_4$) corresponding to the saturation percentages: 20, 30, 45, 60, 70 and 80 (%). The calculations for how much $(NH4)_2SO^-_4$ that were added to each saturation percentage was done using the Ammonium Sulphate calculator from EnCor Biotechnology Inc (USA). Available online at: https://www.encorbio.com/protocols/AM-SO4.htm. An overview of the corresponding amount of ammonium sulphate to the saturation percentage is presented in Table 8.

Table 8: Saturation percentages and corresponding (NH4)2SO-4 amount (mg) for precipitation at 4 °C after calculations made using the Ammonium Sulphate calculator from EnCor Biotechnology Inc.

% Saturated	**Amount (mg)**
20	110
30	170
45	270
60	370
70	450
80	530

The samples were resuspended until all of the $(NH4)_2SO^-_4$ was dissolved. All tubes were then centrifuged for 30 min at 20.000 g x at 4°C in a Eppendorf centrifuge (Eppendorf AG, Germany). The supernatants were transferred to new 1.5 mL tubes and the pellets were resuspended in 100 µl buffer containing 25 mM Tris/HCl (pH 8.0) and 200 mM NaCl. The samples were tested using the spectrophotometric protease activity assay described in section 4.7.

4.7 Protease activity testing by a spectrophotometry assay

Table 9: Products and equipment used for the protease activity testing

Product/equipment	**Distributor (country)**
Peptide substrate Suc-Ala-Ala-Pro-Phe-pNA	ArcticZymes ASA (Tromsø, NO)
1 M Tris-HCl	ArcticZymes ASA (Tromsø, NO)
0.1 M $CaCl_2$	ArcticZymes ASA (Tromsø, NO)
Enzyme dilution buffer	ArcticZymes ASA (Tromsø, NO)
ArcticZymes Proteinase	ArcticZymes ASA (Tromsø, NO)
SpectraMax® iD3 Spectrophotometer	Molecular Devices, LCC (San Jose, California, USA)
Greiner UV-Star® 96-well plate	Merck KGaA (Darmstadt, Germany and/or its affiliates)
Plate-seal	PerkinElmer International C.V. (Groningen, Netherlands)
SoftMax® Pro Software	Molecular Devices, LCC (San Jose, California, USA)

The spectrophotometric assay is a classic enzyme test that is widely used for testing and characterizing activities of enzymes. During a spectrophotometric assay, the enzyme reaction is measured by changes in the intensity of the light absorbed or scattered by the reaction solution. The protease activity assay performed in this thesis is based on the enzymatic cleavage of a substrate peptide, Suc-Ala-Ala-Pro-Phe-pNA, yielding 4-nitroaniline which expresses a yellow color under alkaline conditions which further leads to a change in light absorbance at 410 nm that is determined using a spectrophotometer. The reaction starts immediately after the addition of an enzyme to the assay buffer at a final concentration of 40-200 ng/mL. Table 9 provides an overview of products and equipment used for the activity testing. Assay buffer was prepared as described in Table 10.

Table 10: Protease activity assay buffer composition.

Solutions	**Amount**
1 M Tris-HCl pH 8.0 at 25 °C	100 mM
0.1 M $CaCl_2$	10 mM
100 mM peptide-substrate	1 mM
Milli-Q water	Until 8,0 mL

The spectrophotometer incubation chamber was heated to 30°C and the photometer program was prepared and set to read mode: ABS (absorbance) at 410 nm and read type: kinetic. Further, the program was set to shake mode (double orbital) to mix the samples prior to measurements,

and intervals was set to six seconds with a total time of 20 minutes (speed read). An amount of 240 µL assay buffer was dispensed in each well of the 96-well plate and inserted in the spectrophotometer incubation chamber for pre-heating. Then, 10 µL duplicates of 100x diluted positive control (ArcticZymes Proteinase), enzyme samples and negative control (culture media) was dispensed to the corresponding wells, and the plate was sealed, and the measurement was started immediately after. Data analysis and reduction was performed using the SoftMax® Pro software.

4.8 SDS-PAGE – preparation for Mass Spectrometry analysis

Mass Spectrometry (MS) analysis of the fungal supernatants was performed three times. The first analysis was performed on concentrated fungal supernatants from FMAP and D2MA using Amicon® Ultra-15 10K centrifugal filters (Merck Millipore, Ireland) (DMA was not included due to concerns with background noise interruptions during MS because of the high concentration of malt extract present). For the second MS analysis, the samples had also been through 80% ammonium sulphate precipitation and this time fungal supernatant grown in DMA medium was also included because the previous MS-analysis showed that there were no indications of background noise. *The last MS-analysis was done on sample bands cut directly from SDS-PAGE gels after recombinant expression, lysis and His-tag purification and did not follow this protocol.*

To submit a complex mixture sample to MS analysis, the samples were run 5-10 mm into the resolving part of a pre-cast NuPAGE™ 4-12% BIS TRIS gel (Invitrogen, USA). The samples proceeded into the resolving area to not remain in any stacking area of the gel. Further, the gel was stained with SimplyBlue™ SafeStain (Thermo Fisher, USA), and the stained area was cut out as 5-10 mm slices (avoiding the loading area). Each slice was placed into high quality 1.5 mL Eppendorf Tubes® (Eppendorf AG, Germany). The MS analysis was performed by senior engineers at Proteomics and Metabolomics Core Facility (PRiME) at the Faculty of Health Sciences (UiT). The method used for MS analysis is described in Appendix B.

In addition, Peik Haugen at Norstruct: Department of Chemistry molecular biosystems and bioinformatics, constructed a sequence library (Fasta-file) that were used as a reference for the MS-analysis. The sequence library consisted of all the proteins that have been detected in, and made available in *Digitatispora marina* from the webpage of the Genome Portal of the Department of Energy Joint Genome Institute that sequenced the *Digitatispora marina* genome (81). Link to webpage: https://mycocosm.jgi.doe.gov/Digmar1/Digmar1.home.html.

4.9 Inhibition testing on protease activity

Prior to the bioinformatic analysis, two known protease inhibitors: phenylmethylsulphonyl fluoride (PMSF) (Sigma-Aldrich, Germany): serine protease inhibitor, and Ethylenediaminetetraacetic acid (EDTA) (ArcticZymes, Norway): metalloprotease inhibitor was tested on the enzyme to possibly find out what type of protease is responsible for the detected activity, and further narrow down interesting candidates for bioinformatic analysis. The testing was performed by adding concentrations of 0.5 mM and 1.0 mM PMSF and EDTA into separate assay buffers (composition described in Table 10). The assaying was done as described in section 4.7.

4.10 Bioinformatic analysis of protein sequences present in MS-results

The MS-results were retrieved as an Excel file. In this file the amino acid sequence-IDs that matched with predicted proteins encoded by the *D. marina* genome were listed. Protease sequences present in the first MS-analysis were first analyzed using the bioinformatic tool for functional annotation, and then structure prediction

Functional annotation was done first by using BLASTP (80). BLASTP identifies the closest match of the query in the NCBI protein databases, and thus the most likely function of the *D. marina* protein. Protein sequences were uploaded to the BLASTP web server in FASTA format and run with default settings. Next, SignalP (82) was used to predict signals peptides, which could indicate that the protein is exported out of the cell, or into specific compartment of the cell. Here, sequences were also used in FASTA format with default settings. The Multiple Sequence Alignment tool, ClustalOmega (83), was then used to identify conserved regions by comparing homologous sequences of the proteases and by this, exclude any sequences that was too similar with each other. The ClustalOmega run was done with default settings and a Proteinase K and ArcticZymes Proteinase was used as reference sequences. For structure prediction and evaluation of potential His-tagging on the C-terminal for downstream purification, SWISS-MODEL (84) was used. SWISS-MODEL were run with default settings and the models generated were chosen based on the highest sequence identity match and highest sequence coverage.

Finally, InterPro platform for HMM (Hidden Markov Models) domain annotation (85) was also used to compare and examine which enzyme families were present after both MS-analysis. By using this approach protein domains conserved in certain enzyme families will be assigned to

an unknown sequence allowing it to be categorized according to its enzyme function or relationship to a certain enzyme family. The InterPro analysis was run with default settings.

4.11 Recombinant protein expression of novel protease genes in

After the bioinformatic analysis, 8 of the sequences (7 serine proteases and 1 metalloprotease) were chosen for recombinant protein expression in *Escherichia coli* cells. Before the 8 sequences were ordered, the predicted cleavage site for the signal peptide was removed from the sequence leaving only the mature peptide sequence. Further, for the sequences that had a predicted protein structure with an exposed C-terminal, SP1, SP2, M6, SP5, SP6 and SP7, a His-Tag (containing ten histidine amino acids) was added for His SpinTrap purification purposes. Table 11 contains an overview of how the 8 protease-sequences are referred to throughout the thesis, as well as information on their amino acid length. The raw amino acid sequence of each of the 8 proteases is available in Appendix E.

Table 11: Protease ID and base pairs for each of the 8 proteases (7 serine proteases (SP) and 1 metalloprotease (MP) that were expressed in E. coli.

Protease ID	**SP1**	**SP2**	**SP3**	**SP4**	**SP5**	**SP6**	**SP7**	**MP**
Amino acids (#)	387	392	391	389	392	594	391	209
Molecular weight (kDa)	41.09	41.09	40.098	41.549	42.107	66.236	41.922	32.226

The sequences, ordered from GenScript Biotech (Netherlands), were pre-sub-cloned into pET-20b (+) expression vectors by GenScript Biotech (Netherlands) B.V. The pET 20b (x) expression vectors contain a signal sequence for periplasmic expression of proteins (PelB) which is why the signal sequences (predicted in SignalP) was removed from the sequences prior to sending the sequences to GenScript.

The plasmids were delivered in vials containing approximately 4 µg of lyophilized plasmid DNA with the sequences inserted. To prepare the plasmid DNA for transformation, the vendors protocol was followed: The vials were centrifuged at 6,000 x g for one minute at 4°C in a Eppendorf centrifuge (Eppendorf AG, Germany). The vials were added 20 µl of Milli-Q water to dissolve the DNA. Each vial was vortexed for one minute in a Genie 2 vortex (Scientific industries, USA).

4.11.1 Transformation of plasmid DNA with inserted sequences in E. coli cells

Plasmid DNA with the inserted sequences were transformed into BL21 (DE3) chemically competent cells (New England Biolabs, Germany). The cells were thawed on ice and gently mixed before transferring 50 µl of cells into transformation tubes on ice. Of the plasmid DNAs, 1 µl was added to each corresponding cell mixture. Each tube was carefully flicked a couple of times to mix cells and DNA. The mixtures were left on ice for 30 minutes, and further heat shocked at 42°C for exactly 10 seconds. After heat shocking the cells, the tubes were again placed on ice for 5 minutes. The mixtures were then added 950 µl of room temperature SOC outgrowth medium (New England Biolabs, Germany) and incubated at 37°C for 60 minutes on a incubation shaker (Ecotron, USA) at 250 rpm. After the incubation, the cells were centrifuged at 6,000 x g for 3 minutes at RT (room temperature). The top layer of the supernatant was discarded leaving approximately 100-150 µl left in the tube. The remaining cells were mixed thoroughly by flicking the tube and inverting and further spreading onto premade LB agar plates (Lysogeny Broth) containing ampicillin (ArcticZymes, Norway). The LB medium composition is available in the Appendix (Table A3). The plates were incubated over night at 37°C.

4.11.2 Making of precultures

For the making of the precultures, 5 mL LB medium with 100 µg/mL ampicillin was added to culture tubes (one for each plasmid DNA). Each culture tube was inoculated with a single colony from the LB Agar plates. This was repeated for all eight of the transformed protease sequences, plates as well as a randomly chosen control (SP1) which was not induced at a later step to see if the induction was successful or not. The cultures were incubated diagonally O/N (overnight) at 37°C on a incubation shaker (Ecotron, USA) at 220 rpm. The next day, OD600 was measured a Ultrospec®10 cell density meter (Amersham Bioscience, USA) for each culture after diluting the cultures 10 times in LB medium.

4.11.3 Main cultures and expression

When making the main cultures, 30 mL of TB medium (composition presented in the Appendix Table A2) with 100 µg/mL of ampicillin was added to autoclaved 200 mL Erlenmeyer flasks. All flasks were inoculated with 1 mL of the pre-cultures. The cultures were then incubated at 37°C on a shaker at 220 rpm while measuring the optical density at 600 nm (OD600) routinely. When OD600 = 0.8-1.1 the cultures were induced with 1 mM IPTG. The cultures were then incubated at 25°C on a incubation shaker (Ecotron, USA) at 220 rpm. The next day, OD600 was measured on all 9 samples (including the control), and 1 mL x 4 of each sample was

transferred to 1,5 mL microcentrifuge tube as well as 13 mL x 2 of each sample was transferred to 15 mL falcon tube and centrifuged at 21,000 rpm at 4°C for 5 minutes. The supernatants were discarded while the pellets were kept in -20°C freezer until lysis was to be performed.

4.11.4 Release of periplasmic protein fraction from E. coli by cold osmotic shock (lysis)

Table 12: Overview of content in Fractioning Buffer 1 and 2 for periplasmatic lysis of E. coli to release protein fractions.

Cell Fractioning Buffer 1	Cell Fractioning Buffer 2
0.2 M Tris-HCl pH 8.0	0.01 M Tris-HCl pH 8.0
200 g L^{-1} sucrose	0.005 M $MgSO_4$
0.1 M EDTA	0.2% SDS
	1% Triton X-100

Lysis by cold osmotic shock was done to break down the *E. coli* cells and release the periplasmic protein fraction. This was done by following a protocol that is a modified version of Neu & Heppel from 1965 (86).

The cells were resuspended in ice-cold Cell fractioning Buffer 1 (Table 12), where the resulting volume was ¼ of the original suspension volume. Samples were incubated on ice for 20 minutes with regular intervals of suspension invertion to counteract sedimentation. After incubation, the cell suspensions were centrifuged for 15 min at 14,000 xg at 4°C. The entire supernatant was discarded. Further, the cells were resuspended in ice-cold Cell Fractioning Buffer 2 (Table 12), where the resulting volume was ¼ of the original suspension volume. Samples were incubated for 15 minutes on ice under regular invertion. The cell suspensions were centrifuged for 15 minutes at 14,000 xg at 4°C. The supernatants, which conceivably contains the periplasmatic proteins and membrane proteins, were transferred to new tubes while the pellets remained in the old tubes, both were retained.

4.11.5 Whole cell lysis of the pellet

For examining the efficiency of the periplasmic lysis, the pellets collected after periplasmic lysis were lysed to see if they contained our protein of interest. This was done by adding a buffer containing a corresponding amount of PB pol lysis buffer (Composition available in Appendix: Table A4), lysozyme, and HL-San to the measured OD600. An overview of the buffer content as well as calculations made for the buffer content is presented in Appendix C.

After the addition of lysis buffer, the samples were incubated on a INFORS HT Labotron orbital shaker (LabX®, Canada) at 15°C O/N. The proteins were collected by centrifugation for 15 minutes at 14 000 xg at 4 °C and the supernatant were collected and stored for further analysis.

4.11.6 SDS-page gel electrophoresis

Pre-casted NuPAGE™ bis-tris 4-12% SDS-page gels (Invitrogen by Thermo Fisher Scientific, USA) (one gel for the lysed supernatants and one for the lysed pellets) were placed in a Novex Mini-Cell (Bio-Rad,USA) chamber and filled with premade MES-buffer (Table A5) until the outer chamber was half-full and the inner chamber was completely filled. An amount of 30 µl of the lysed supernatants and pellets were transferred to new tubes containing 10 µl of loading dye and mixed thoroughly by pipetting. The gels wells were added 3 µl of each sample with 5 µl of a pre-stained protein-ladder on each of the sides. The gel was run at 200 V for 35 minutes.

To remove salts and SDS, the gels were first washed with approximately 100 mL of Milli-Q water and heated in a microwave till it almost started boiling. The gel-wash was left on a horizontal shaker for 1 minute, and the Milli-Q water was discarded. These steps were repeated three times. Further, the gel was stained with approximately 50 mL of Comassieblue/SafeStain staining solution and heated in a microwave for 5 minutes. After the gel staining, the gel was again washed in 100 mL of Milli-Q water for 10 minutes on a shaker to diffuse out the staining dye that did not bind to any proteins. The remaining liquid (Milli-Q water and remaining stain) was discarded and the proteins were detected as blue bands. The gels were photographed in a Gel Doc (Bio-Rad) by using the Image Lab software (Bio-Rad, USA) set to "protein-gel" and "ComassieBlue/SafeStain".

4.11.7 His SpinTrap purification

His SpinTrap purification was performed on the lysed samples containing the protease sequences with an added His-Tag: SP1 (and SP1 control (uninduced), SP2, MP, SP6 and SP7. Sample, SP5, was also included at first but ended up being excluded because it was too viscous to be able to move through the column.

His SpinTrap™ Ni Sepharose™ High Performance columns (Cytiva, UK) was used to purify the samples with added His-Tag. The method was performed as described in the purification protocol from Cytiva. Firstly, the storage solution was removed by shaking the columns repeatedly to resuspend the medium. The bottom closure was removed, and the top cap was loosened before the columns were placed into 2 mL microcentrifuge tubes and centrifuged for

30 sec at 100 x g. For the column equilibration, 600 µl of binding buffer (composition described in Table 13) was added to the columns, and centrifuged for 30 seconds at 100 x g. All columns were transferred to new 2 mL tubes. Then, 600 µl from each sample was added to the corresponding tubes and centrifuged for 30 sec at 100 x g. Further, the columns were added 600 µl binding buffer for washing and centrifuged at the same speed and time as earlier. The last step was elution, where 200 µl of elution buffer (composition described in Table 13) was added and centrifuged for 30 sec at 100 x g. This step was repeated once to elute any remaining protein bound to the column resin. The samples were analyzed by SDS-PAGE (as described in section 4.11.6.

Table 13: Composition of Binding buffer and Elution buffer used respectively in column equilibration and elution during His SpinTrap purification.

Binding buffer		**Elution Buffer**	
Tris-HCl (pH 7.5)	20 mM	Tris-HCl (pH 7.5)	20 mM
NaCl	500 mM	NaCl	500 mM
Imidazole	20 mM	Imidazole	500 mM

After the gel run, any interesting bands were cut out using disposable scalpel knifes, transferred to 1,5 mL microcentrifuge tubes and sent to MS analysis to Proteomics and Metabolomics Core Facility (PRiME) at the Faculty of Health Sciences (UiT) to see if the expressed proteins were the same proteins (sequence) that was inserted in the plasmids. Some results were blasted in BLASTP (80) to look for any indications to which type of protease was present.

4.11.8 Buffer exchange and activity testing of His SpinTrap purified samples

Because imidazole often interferes with enzyme activity, a buffer exchange using Zeba^{TM} spin desalting column (Thermo scientific, USA) was done prior to protease activity testing of the His SpinTrap purified samples.

The Zeba^{TM} spin desalting column (0,5 mL) containing 800 µl of High-Performance Desalting Resin Slurry was used for exchanging buffer after the samples was His SpinTrap purified. The columns were placed into 1.5 mL collection tubes and centrifuged for 1 minute at 1500 x g to remove the storage solution. The flow-through was discarded and washing steps was performed by adding 300 µl of the equilibration buffer on top of the resin. The devices were centrifuged at 1500 x g for 1 minute, and the flow-through was discarded. This was repeated three times. For sample recovery, the columns were transferred to new 1.5 mL collection tubes, and 130 µl

of each sample was added on top of the resin of the corresponding column. To ensure maximum sample recovery, a stacker (15 µl of equilibration buffer) was added as soon as the sample had entered the resin. The devices were centrifuged at 1500 x g for two minutes and the flow-through containing the samples was retained and stored at 4 °C.

To examine if any of the chosen proteases (with added His-Tag) was responsible for the activity observed in the fungal supernatant of *Digitatispora marina*, samples were tested for activity with performing the same protease activity assay on SpectraMax® iD3 spectrophotometer as described in section 4.7.

4.12 Biochemical characterization of the fungal supernatant

Due to time limitations, it was decided to only perform biochemical characterization assays on the 80% ammonium sulphate precipitated supernatant from the FMAP culture. The protease activity was tested at different temperatures, salt concentrations, and different pH values of the Tris-HCl buffer. The enzyme sample were diluted in a salt buffer containing 25 mM Tris/HCl and 200 mM NaCl. This same buffer acted as the negative control. All differently treated samples were tested in triplicates (including the negative control). The spectrophotometer assay for testing the protease activity remained the same as described in section 4.7 with the assaying buffer composition described in Table 10 (with some adjustments for salinity and pH tests). In addition, when testing for temperature optimum, endpoint measurements was done instead of kinetic measurements.

4.12.1 Temperature optimum

The enzyme sample was diluted 20x, whereof 10 µl was added to 1,5 mL microcentrifuge tubes containing 240 ul of the assay buffer and incubated on a Grant QBD2 heating block (Grant Instrument, UK) at seven different temperatures ranging from 20, 30, 40, 50, 60, 70 and 80°C for five minutes. The reactions were stopped by heating the samples at 95°C for 5 minutes on the heating block. Each sample was transferred to a 96-well plate, and the activity was detected with the Spectramax® iD3 spectrophotometer.

4.12.2 Salinity tolerance

The salinity tolerance of the protease activity was determined by adding different concentrations of NaCl: 50 mM, 100 mM, 150mM, 200 mM, 250 mM, 300 mM, 350 mM, 400 mM, 450 mM, 500 mM, 1 M, 1.5 M, 2.0 M, 2.5 M, and 3.0 M, to the assay buffer. The enzyme sample was diluted 10x, and 10 µl of the enzyme was transferred in triplicates to 96-well plates

with 240 µl of each assay buffer, and the activity was tested in a SpectraMax® iD3 spectrophotometer.

4.12.3 Optimum activity within a range of pH values in Tris-HCl

The optimum pH for the protease activity was assayed over a Tris-HCl pH range of 7.5, 8, 8.5 and 9.0. The assay conditions and assay buffers were the same as described under section 5.7 and Table 9, with adjustments of pH in Tris-HCl as stated. The enzyme sample was diluted 10x, and 10 µl of the enzyme was transferred in triplicates to 96-well plates with 240 µl of each assay buffer, and the activity was tested in a SpectraMax® iD3 spectrophotometer.

5 Results

The results from the DNA sequencing done by external engineers at the University Hospital of North Norway was blasted in ncbi nucleotide blast (80) and got a 100% identity match with *Digitatispora marina* "isolate 3027C" meaning that the fungal identity was successfully verified and there were no traces of contamination.

5.1 Growth profiling and morphology observations of the fungus on solid media

The growth of *Digitatispora marina* was measured in colony diameter and observed on six different agar media (D2MAA, D2MAAW, M19, M19W, DCAA, DCAAW) and incubated at two different temperature conditions, 4°C and 10°C. The growth measurements were done approximately once a week and ended when the fungus was completely covering the agar plates of one culturing medium (90 mm) which occurred after 77 days for *D. marina* grown on D2MAA at 10°C. Figure 5 shows a graphic representation of the fungal growth on all agar plates.

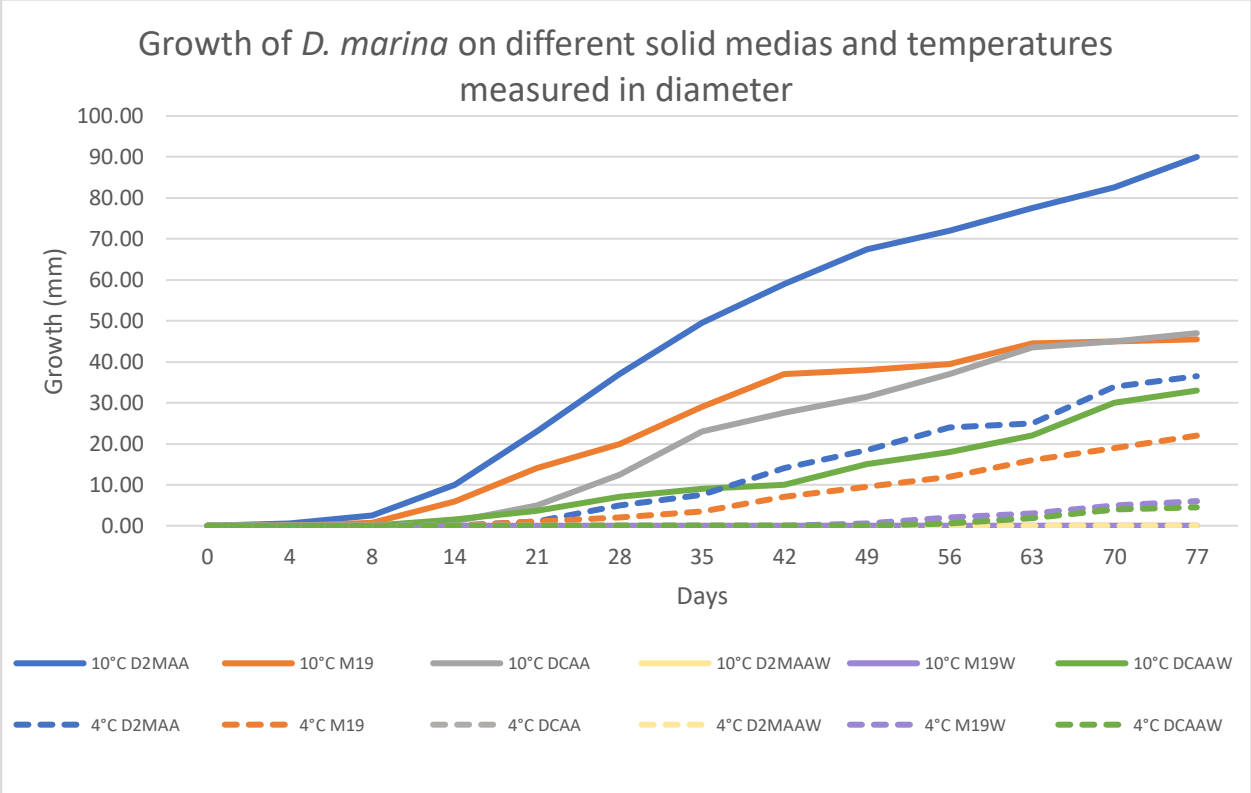

Figure 5: Diagram showing the growth of D. marina *on six different mediums; D2MAA, D2MAAW, M19, M19W, DCAA and DCAAW incubated at 10°C and 4°C. The straight lines represent the fungal growth on plates incubated at 10°C, the stippled lines represent the fungal growth on plates incubated at 4°C*

D. marina grew in all media, except for D2MAAW and M19W (at 10°C) and D2MAAW (at 4°C). The filamentous fungal growth was fastest on D2MAA at 10°C, followed by DCAA (10°C), M19 (10°C), D2MAA (4°C), DCAAW (10°C) and M19 (4°C), with respectively 47 mm, 45,5 mm, 36,5 mm, 33 mm and 22 mm after 77 days of growth. Both DCAA and M19W at 4°C had a growth of 6 mm followed by DCAAW (4°C) with a growth of 4,5 mm.

Differences in mycelium shape and structure were also observed for each plate where the fungus had any growth, which was observed on 9 out of 12 different plate conditions. Figure 6 shows the morphology of *D. marina* on different media. The photos were taken after 80 days of growth.

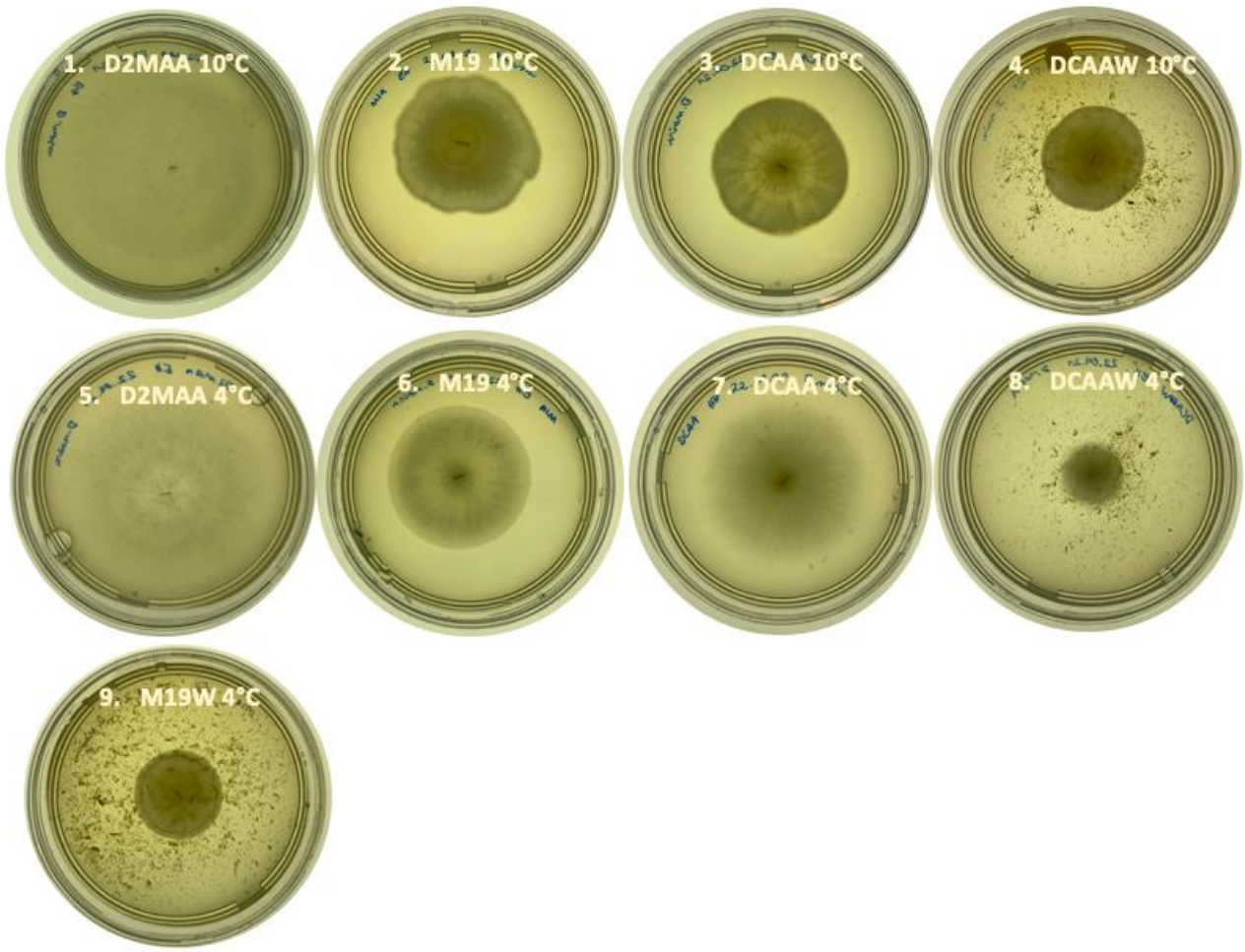

Figure 6: Morphology of Digitatispora marina *grown on different solid media compositions. From upper left to lower right: 1. D2MAA 10°C, 2. D2MAA 4°C, 3. DCAAW 10°C, 4. DCAAW 4°C, 5. M19 10°C, 6. M19 4°C, 7. DCAA 10°C, 8. DCAA 4°C, 9. M19W 4°C.*

D. marina displayed thin filamentous growth (hyphae growth) that were almost see through on D2MAA grown at both temperatures. The hyphae growth on DCAA (4°C) was also thin but more visible. The fungus grew thick hyphae on all media compositions with homogenized

wood. M19 grew thicker hyphae at 10°C than at 4°C. The hyphae growth was also thick on fungus grown at DCAA (10°C).

5.2 Partial purification – ammonium sulphate precipitation effect

Partial purification with ammonium sulphate precipitation of proteases present in the fungal supernatant from *Digitatispora marina* were performed to find at which $(NH_4)_2SO_4$ concentration the protease activity is at its highest level presumably corresponding with precipitation of the protein of interest. Ammonium sulphate precipitation was performed with 20, 30, 40, 50, 60, 70 and 80% saturation of ammonium sulphate, and both the pellets and supernatants were tested for protease activity to find out where most of the active proteases were present. Figure 7 gives a graphic illustration of protease activity obtained at different $(NH_4)_2SO_4$ % saturations.

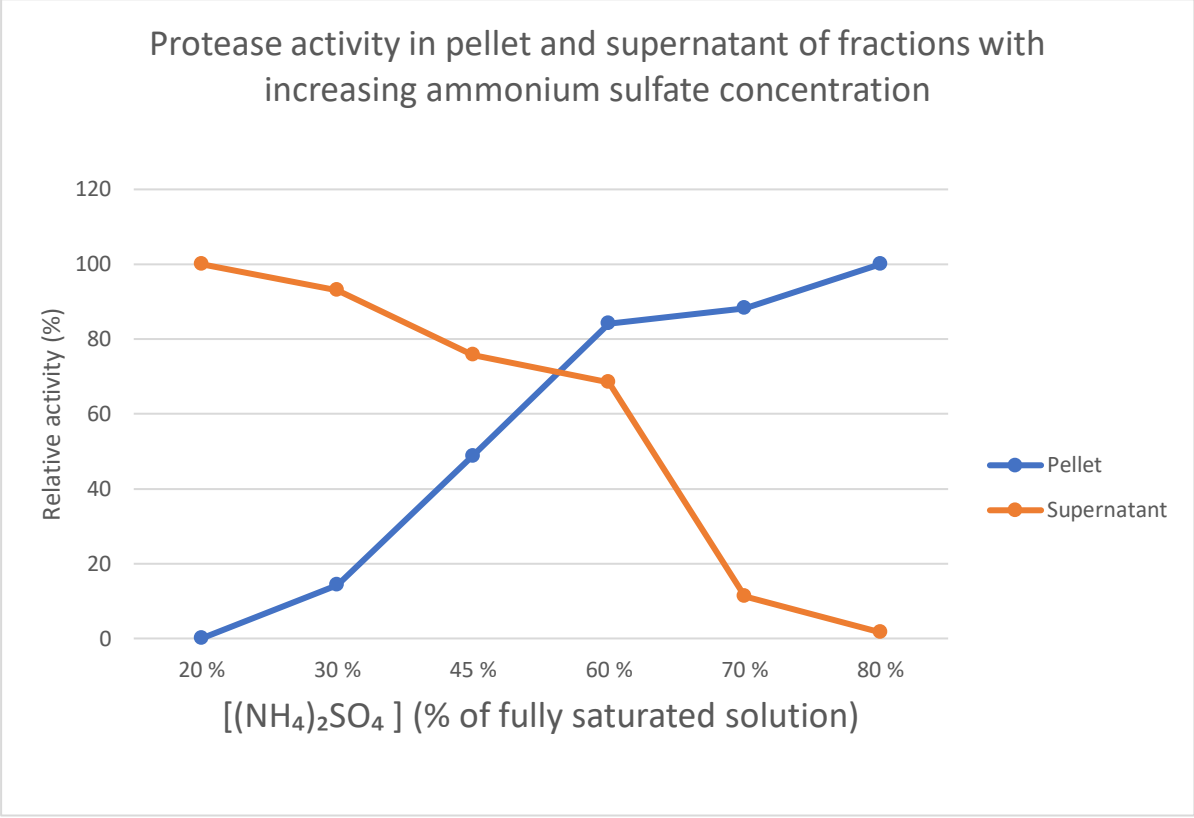

Figure 7: Protease activity in ammonium sulphate precipitated fungal pellet (blue line) and supernatant (orange line) at different saturation percentages: 20, 30, 40, 50, 60, 70 and 80%. Results are depicted as relative to maximum activity (activity measured at 80% saturation of the pellets and 20% saturation for the supernatants).

The activity in the supernatant decreased with increasing $(NH_4)_2SO_4$ saturation. In contrast, the activity in the pellet increased with increasing saturation. The precipitation using 80% saturation of ammonium sulphate gave the highest activity rate (in the pellet) while protease

activity in the supernatant was nearly absent at that high saturation percentage. For illustration purposes the maximum activity for the pellet was set to 100% for relative analysis of the pellet activity measurements. This was also done for the supernatant activity measurements, but with the maximum activity for the supernatant set to 100%.

5.3 Effect of inhibitors on protease activity

Activity inhibition testing with PMSF (a commonly used serine protease inhibitor) and EDTA (a commonly used metalloprotease inhibitor) as protease inhibitors were performed by using a spectrophotometric approach as described earlier. The protease activity was significantly inhibited with the addition of PMSF (0.5 mM and 1 mM). EDTA had a small effect on activity increasing it slightly by x % and y % at 0.5 and 1 mM of EDTA, respectively. The results are displayed in Figure 8.

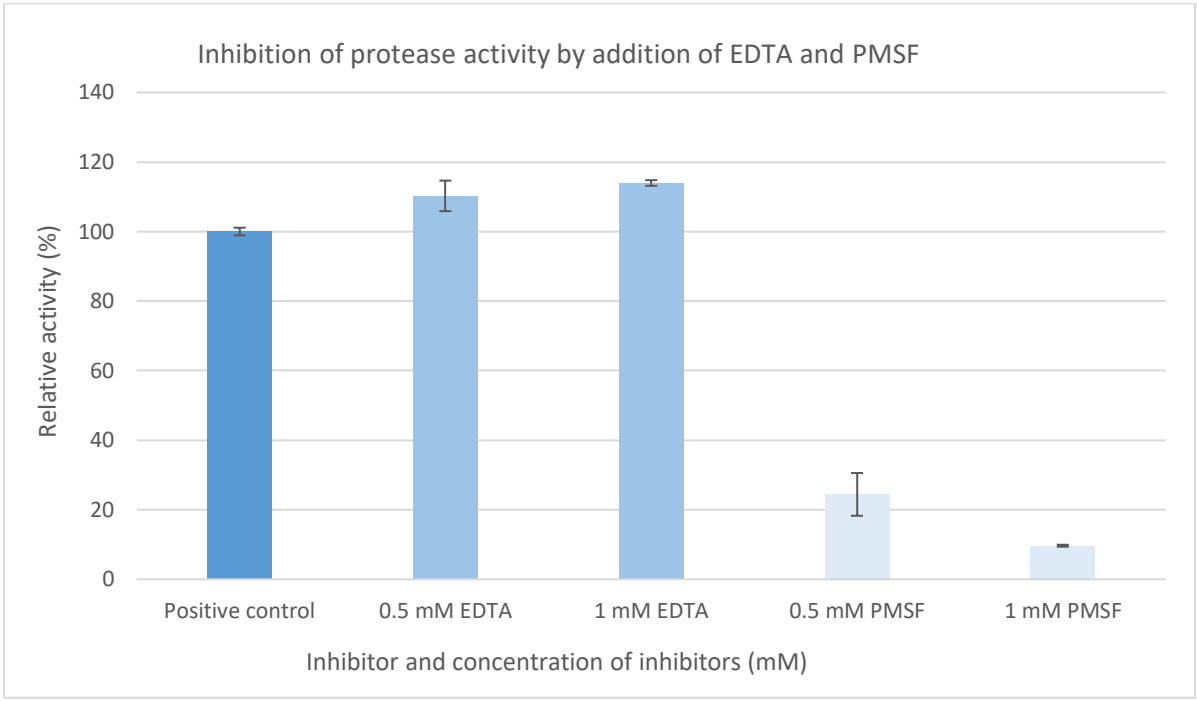

Figure 8: Effect of two inhibitors, EDTA and PMSF, on protease activity found in fungal supernatant after partial purification using spin filtration with 10 kDa cut off, and 80% ammonium sulphate precipitation. Positive control is an enzyme sample at regular buffer conditions.

5.4 MS-results and bioinformatic analysis

Mass Spectrometry (MS) analysis of the fungal supernatants was performed twice. The first analysis was performed on concentrated fungal supernatants from FMAP and D2MA using centrifugal filters with 10 kDa cutoff. For the second MS analysis, all three fungal supernatants (from FMAP, D2MA and DMA) was included and had been partially purified with 80% ammonium sulphate precipitation. Samples were sent as cut out pieces after SDS-PAGE gel electrophoresis run. The MS-results were retrieved as an Excel file.

After the first MS-analysis the results (hits with sequences in constructed library) were blasted in NCBIs protein blast (BlastP). All hits corresponding to serine proteases and one metalloprotease (which was chosen because it was the only protease present in both the FMAP and D2MA growth medium) were chosen for further bioinformatic analysis in SignalP, SWISS-MODEL and ClustalOmega. SignalP was used to predict the presence and cleavage site position of a signal peptide (to find the mature protein sequence) If the match probability were under 60% the proteases were excluded from further investigations. The SWISS-MODEL platform was used to predict protein structures with the aim to evaluate the C-terminal position for addition of a His-Tag. ClustalOmega was used to look for sequence homology between the sequences and to assure that none of them had too many disruptions possibly due to sequencing errors (e.g., too many repeated amino acids). After bioinformatic analysis, 8 protease sequences (7 serine proteases and 1 metalloprotease) were chosen for recombinant protein expression. An overview of the predicted protein structures from SWISS-MODEL is presented in Figure 9.

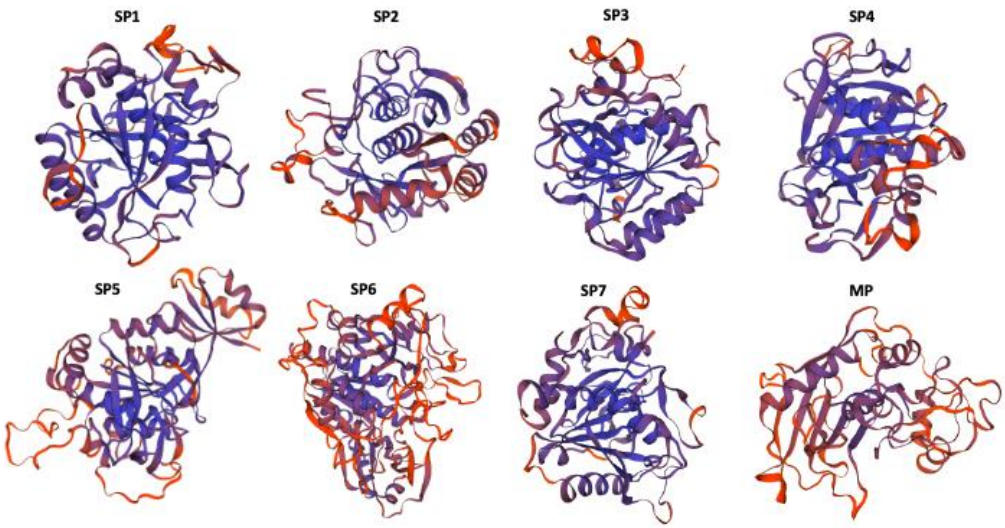

Figure 9: Overview of the predicted protein structures from SWISS-MODEL for SP1, SP2, SP3, SP4, SP5, SP6, SP7 and MP

After the constructions of the predicted protein structures, 6 out of the 8 proteases had an exposed C-terminal where a His-tag was added. An overview of the sequence identity of the predicted protein structures, and the signal peptide position probability is presented in Table 14.

Table 14: Overview of proteases with added His-tag, sequence identity (%) of predicted protein structures acquired from SWISS-MODEL and the probability (%) of the cleavage site position of the signal peptides from SignalP analysis.

Protease	**Seq. Identity (%) of predicted protein structures (SWISS-MODEL)**	**Signal Peptide cleavage site position probability (%) (SignalP)**
SP1 (His-tag)	54,04%	84,16%
SP2 (His-tag)	42,96%	94,75%
SP3	50,91%	79,90%
SP4	45,82%	90,03%
SP5 (His-tag)	30,98%	92,25%
SP6 (His-tag)	35,97%	90,71%
SP7 (His-tag)	47,58%	81,51%
MP (His-tag)	30,74%	85,96%

The sequence homology between the 8 proteases is presented as a sequence similarity matrix from ClustalOmega in Figure 10. Proteinase X (produced at ArcticZymes) and Proteinase K is included for reference purposes.

	SP6	MP	AZ Proteinase	Proteinase K	SP7	SP4	SP1	SP3	SP2	SP5
SP6	100,0	14,3	17,3	21,2	19,8	20,1	18,2	19,5	23,5	21,2
MP	14,3	100,0	13,7	18,0	14,7	16,3	18,6	20,5	17,1	19,4
AZ Proteinase	17,3	13,7	100,0	35,8	36,6	36,4	40,6	39,6	39,3	38,0
Proteinase K	21,2	18,0	35,8	100,0	40,4	44,0	42,4	42,7	40,3	39,7
SP7	19,8	14,7	36,6	40,4	100,0	52,7	52,5	50,5	47,4	45,3
SP4	20,1	16,3	36,4	44,0	52,7	100,0	50,5	49,4	50,0	46,3
SP1	18,2	18,6	40,6	42,4	52,5	50,5	100,0	83,2	57,7	55,3
SP3	19,5	20,5	39,6	42,7	50,5	49,4	83,2	100,0	55,3	55,0
SP2	23,5	17,1	39,3	40,3	47,4	50,0	57,7	55,3	100,0	63,5
SP5	21,2	19,4	38,0	39,7	45,3	46,3	55,3	55,0	63,5	100,0

Figure 10: Sequence similarity matrix between SP1, SP2, SP3, SP4, SP5, SP6, SP7, MP and reference sequences Proteinase X and Proteinase K.

The Sequence similarity matrix shows that the highest sequence similarity is between SP3 and SP1 (83,2%). There was also high similarity between SP2 and SP5 (63,5%), SP2 and SP1

(57,7%), SP2 and SP3 (55,3%). The SP6 and MP sequence are the least related to any of the other sequences. None of the sequences showed high sequence identity with the reference sequences Proteinase X and Proteinase K.

5.4.1 Enzyme families detected in the secretome of *D. marina* by mass spectrometry

A total of 114 sequences from the constructed sequence library, which signatures were detected in the mass spectrometry analysis, were analyzed using the InterPro platform for HMM (Hidden Markov Models) domain annotation.

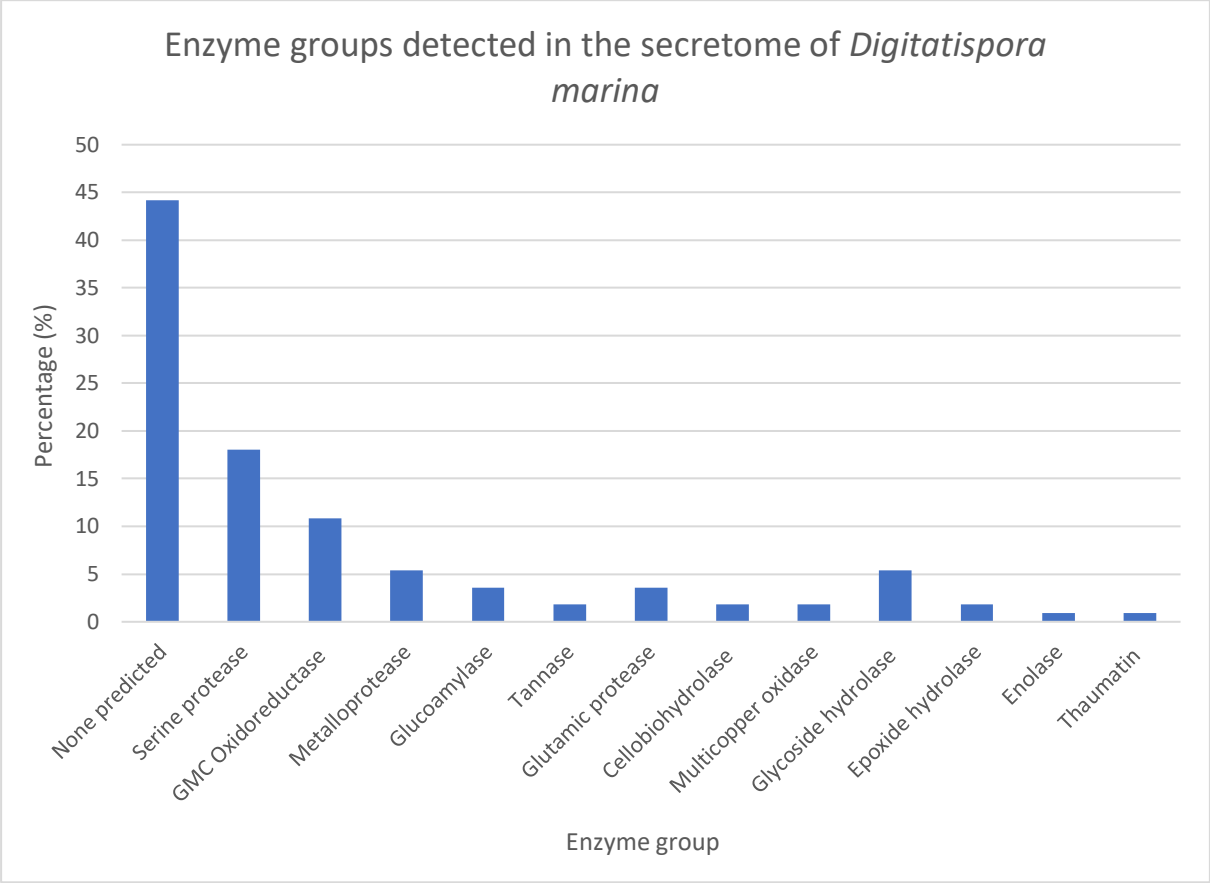

Figure 11: Diagram displaying the enzyme groups present in MS-analysis of up-concentrated FMAP and D2MA growth culture, and 80% ammonium sulphate precipitated FMAP, D2MA and DMA. The enzyme groups are represented as a percentage of the total amount of enzyme groups detected.

Figure 11 shows how many different enzyme families that were present, and the relative contribution of each family that is represented in the secretome of *D. marina*. The largest amount of the sequence hits was not predicted by the Interpro software. Serine protease was the biggest group of the predicted ones, followed by oxidoreductase, metalloprotease and glycoside hydrolase.

5.5 Recombinant protein expression of selected proteases

pET-20b (+) expression vectors with the gene sequences of SP1, SP2, SP3, SP4, SP5, SP6, SP7 and MP inserts assembled by GenScript Biotech (Netherlands) were transformed into competent T7 express E. coli cells and induced with IPTG. Supernatants were lysed by cold osmotic shock to release periplasmic proteins. Pellets left after periplasmic lysis were additionally lysed by whole cell lysis for comparison. Both supernatants and pellets were analyzed by SDS-PAGE. In addition, samples with added His-tag were purified by His SpinTrap purification and further analyzed by SDS-PAGE. Expected bands at molecular weights for each of the proteases: SP1: 41.09 kDa, SP2: 41.09 kDa, MP: 32.226 kDa, SP3: 40.098 kDa, SP4: 41.549 kDa, SP5: 42.107 kDa, SP6: 66.236 kDa, Sp7: 41.922 kDa.

5.5.1 SDS-PAGE analysis of lysed cells

SDS-PAGE analysis of the lysed supernatants (Figure 12) shows a clear indication of over expression of MP at approximately 42 kDa. SP5 sample is present as a smear.

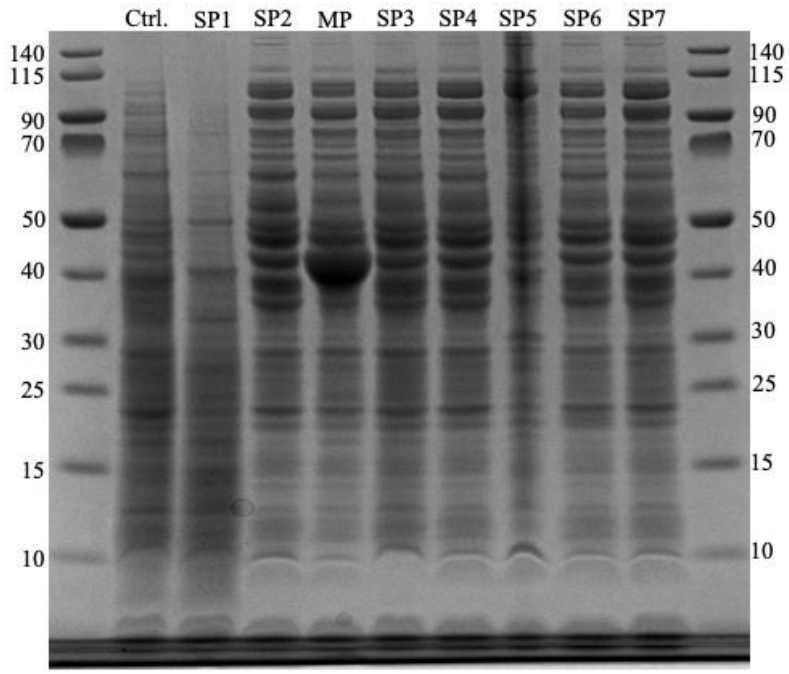

Figure 12: SDA-PAGE analysis of supernatants after lysis of protein samples by cold osmotic shock for periplasmic protein release. Lanes 1 and 11 contains the pre-stained protein ladder 10-140 kDa (Themo Scientific, PageRuler™), lane 2: control sample of SP1 (uninduced), lane 3: SP1 (induced), lane 4: SP2 (induced), lane 5: MP (induced), lane 6: SP3 (induced), lane 7: SP4 (induced), lane 8: SP5 (induced), lane 9: SP6 (induced), lane 10: SP7 (induced).

SDS-PAGE analysis of the lysed pellets (Figure 13) still shows a clear indication of over expression of MP at approximately 42 kDa. The clear bands at 15 kDa is the lysozyme that was used for whole cell lysis of the pellets.

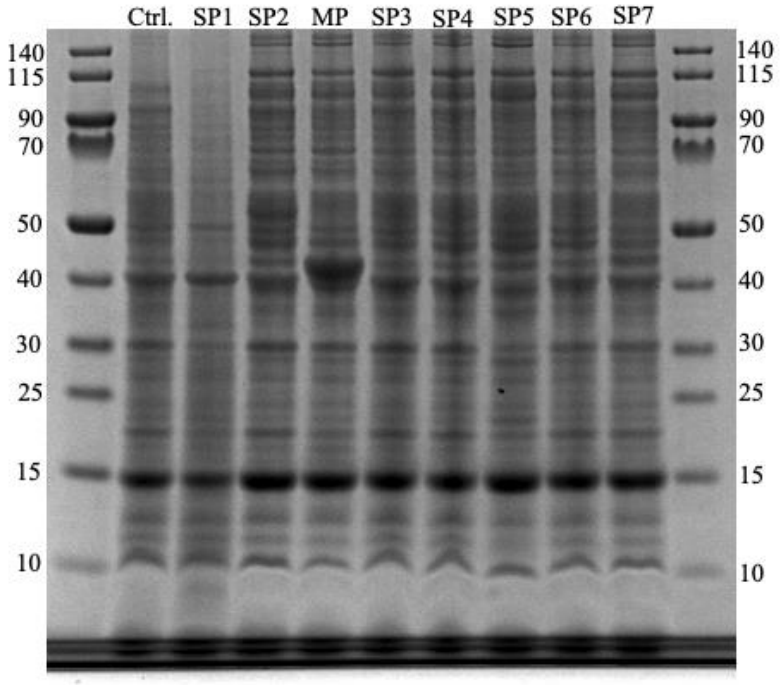

Figure 13: SDS-PAGE analysis of pellets after additional lysis. Lanes 1 and 11 contains the pre-stained protein ladder 10-140 kDa (Themo Scientific, PageRuler™), lane 2: control sample of SP1 (uninduced), lane 3: SP1 (induced), lane 4: SP2 (induced), lane 5: MP (induced), lane 6: SP3 (induced), lane 7: SP4 (induced), lane 8: SP5 (induced), lane 9: SP6 (induced), lane 10: SP7 (induced).

5.5.2 SDS-PAGE gel: After His-tag purification

SDS-PAGE analysis of His-Tag purified supernatant (Figure 14) shows a clear band on MP at approximately 44 kDa. For the His-Tag purified pellet (Figure 15) there is also a clear band on MP at approximately 44 kDa, and another band at approximately 34 kDa in lane SP1.

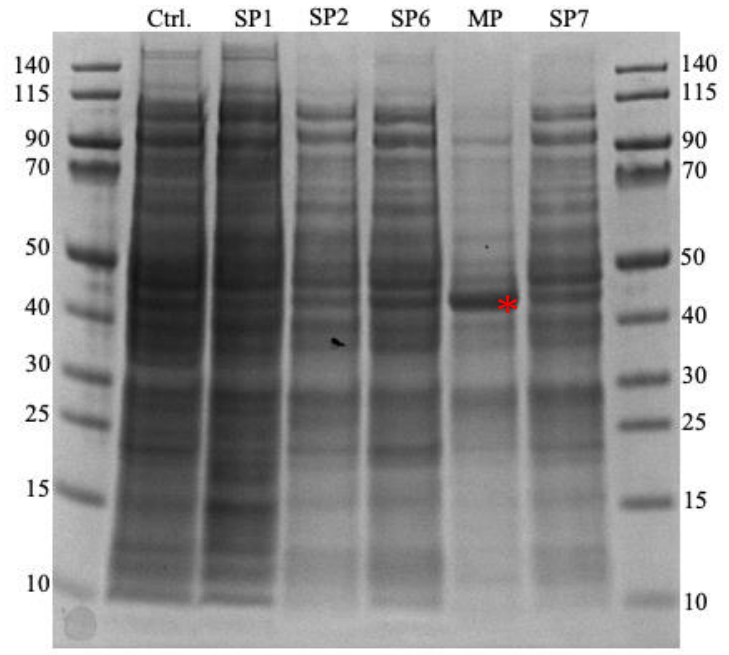

Figure 14: SDS-PAGE analysis His-Tag purified supernatant samples. Lanes 1 and 8 contains the pre-stained protein ladder 10-140 kDa (Themo Scientific, PageRuler^{TM}), lane 2: control sample of SP1, lane 3: SP1, lane 4: SP2, lane 5: SP6, lane 6: MP, lane 7: SP7.

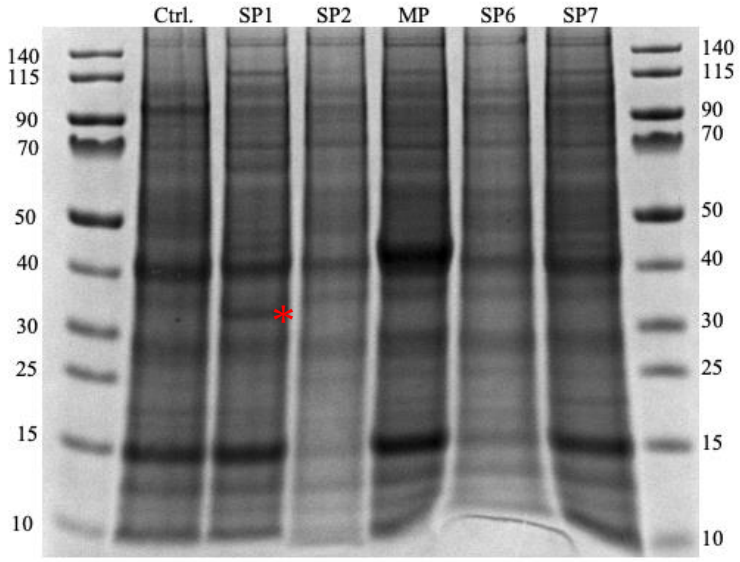

Figure 15: SDS-PAGE analysis His-Tag purified pellet samples. Lanes 1 and 8 contains the pre-stained protein ladder 10-140 kDa (Themo Scientific, PageRuler^{TM}), lane 2: control sample of SP1, lane 3: SP1, lane 4: SP2, lane 5: SP6, lane 6: MP, lane 7: SP7.

5.5.2.1 MS-results of bands cut from His-Tag purified SDS PAGE gels

The bands that were cut from the gels and sent to MS-analysis are marked with a red asterisk (*) and was present in the MP lane of the supernatant sample (**Error! Reference source not found.**), and the SP1 lane of the pellet sample (**Error! Reference source not found.**). The results were retrieved as an Excel file which had been tested against the constructed protein sequence library. The results showed that the band expressed in SP1 did match with the exact sequence that was chosen for protein expression. As for the band present at MP, there were four hits with the protein sequence library, none of them was the MP (metalloprotease) sequence that were chosen for expression. One of the hits was the sequence coding for SP1 (serine protease), and the three others were run in Ncbi´s BLASTP. The BLASTP results showed a high identity hit with a 14-3-3 protein, a translation elongation factor and one histone-fold-containing protein.

5.6 Biochemical characterization of protease activity

Biochemical characterization experiments aiming to determine temperature optimum, salt tolerance and pH optimum for the protease activity was performed on 80% ammonium sulphate precipitated supernatant sample of *Digitatispora marina.*

5.6.1 Temperature optimum for the protease activity

To find indications of a temperature optimum for the protease activity, 10 µl of the 80% precipitated sample was incubated in an assay buffer for five minutes at different temperatures: 20°C, 30°C, 40°C, 50°C, 60°C, 70°C and 80°C. The samples were tested for activity using a spectrophotometer. The activity assay showed that the unidentified protease(s) had an optimum activity at 50°C (set to 100% for relative activity measurements). The standard deviations between the triplicates were low, and the protease(s) showed activity at a range of different temperatures (Figure 16).

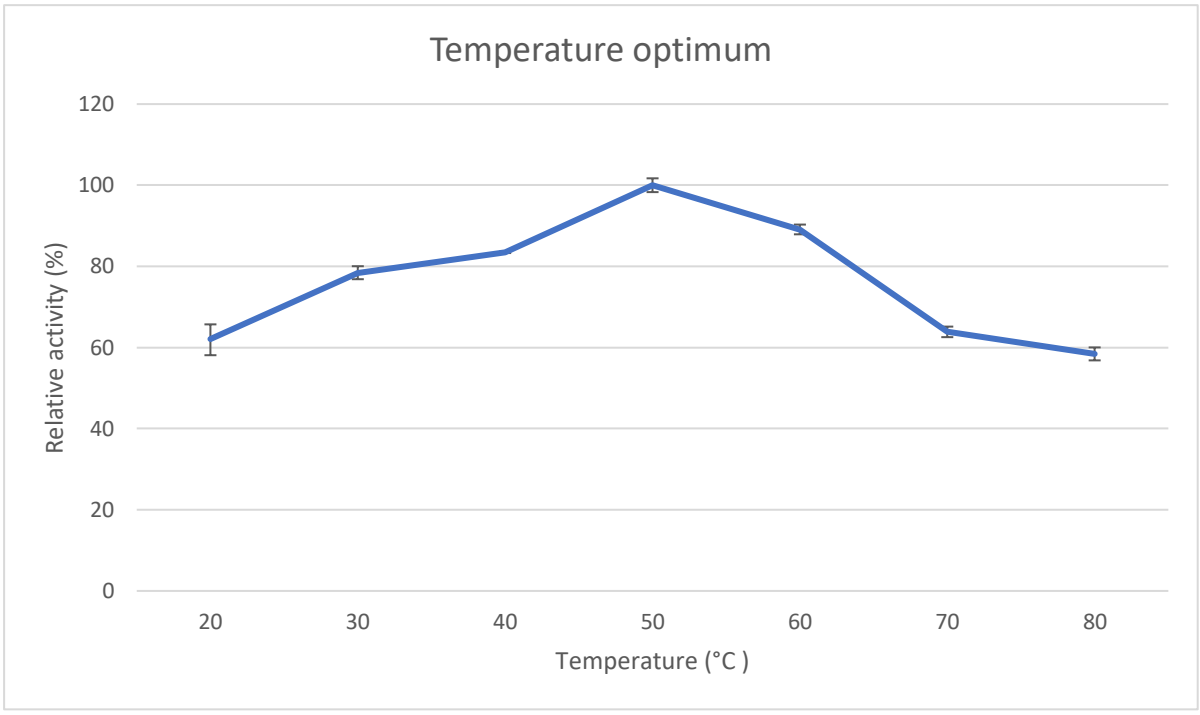

Figure 16: Temperature optimum for the protease activity. The enzyme(s) was assayed at the indicated temperatures in 0.1 M Tris-HCl pH 8.0, 10 mM $CaCl_2$ and peptide substrate. Samples were incubated for 5 minutes and assayed in a spectrophotometer. Results are depicted as relative to maximum measured activity (activity measured at 50°C).

5.6.2 pH optimum for the protease activity

The pH optimum for the protease activity was determined by assaying the sample with different pH values of the Tris-HCl buffer (pH: 7.5, 8, 8.5 and 9). The assay buffer with Tris-HCl at pH 8 (regular assay buffer conditions) gave a slightly higher activity rate and was set to 100% for relative activity measurements. The protease(s) showed activity at all pH values of the Tris-HCl buffer (Figure 17). The standard deviations between the triplicates tested for each sample where lowest for pH 8 and highest for pH 8,5 and 9.

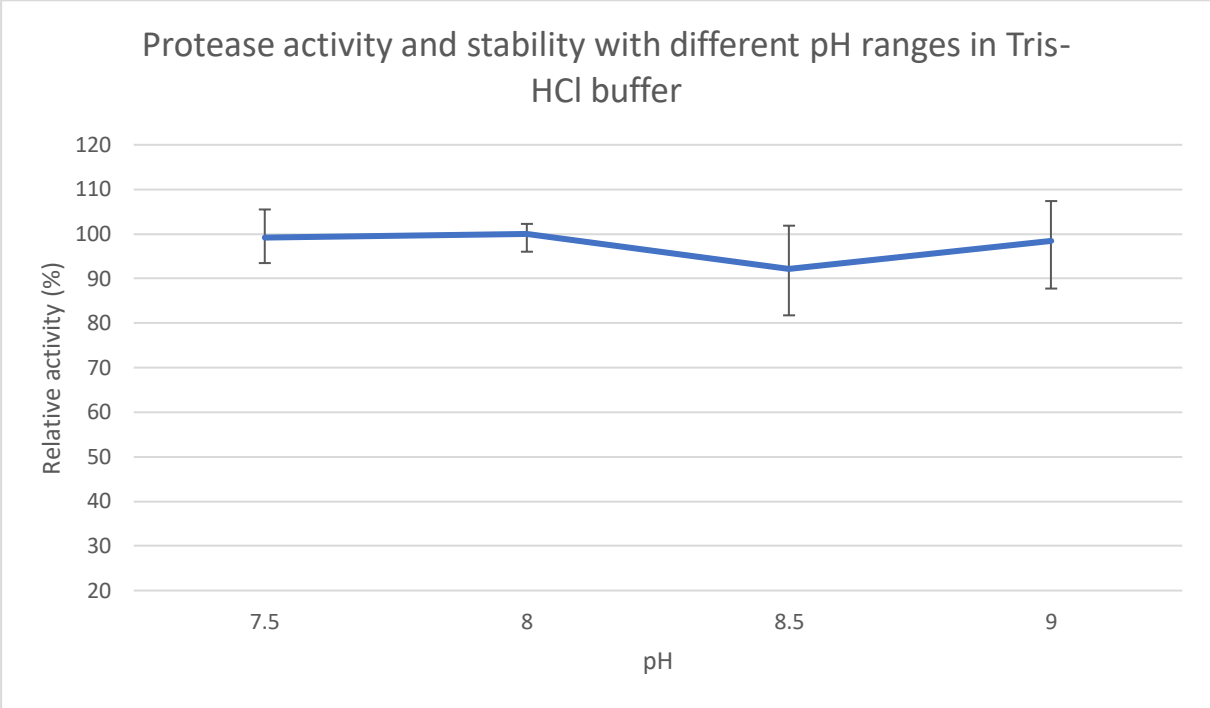

Figure 17: The protease activity was measured in 0.1M Tris-HCl ranging from pH 7,5-9,0, 10 mM $CaCl_2$ and peptide substrate. Results are depicted as relative to maximum activity (activity measured with Tris-HCl pH 8.0 (normal buffer condition)).

5.6.3 Sodium chloride tolerance of the protease activity

The sodium chloride optimum/tolerance was examined by measuring the protease activity at different NaCl concentrations from 0-3 M in a spectrophotometry assay. The enzyme exhibited a broad NaCl tolerance (Figure 18), with maximal activity at 0,15 M (set to 100% for relative activity). The standard deviations between the triplicates tested for each sample where highest with 0,15 M NaCl, and lowest at 0,1M, 0,3M and 1M. All NaCl concentrations gave higher measured protease activity than the assay buffer without added NaCl (0 M).

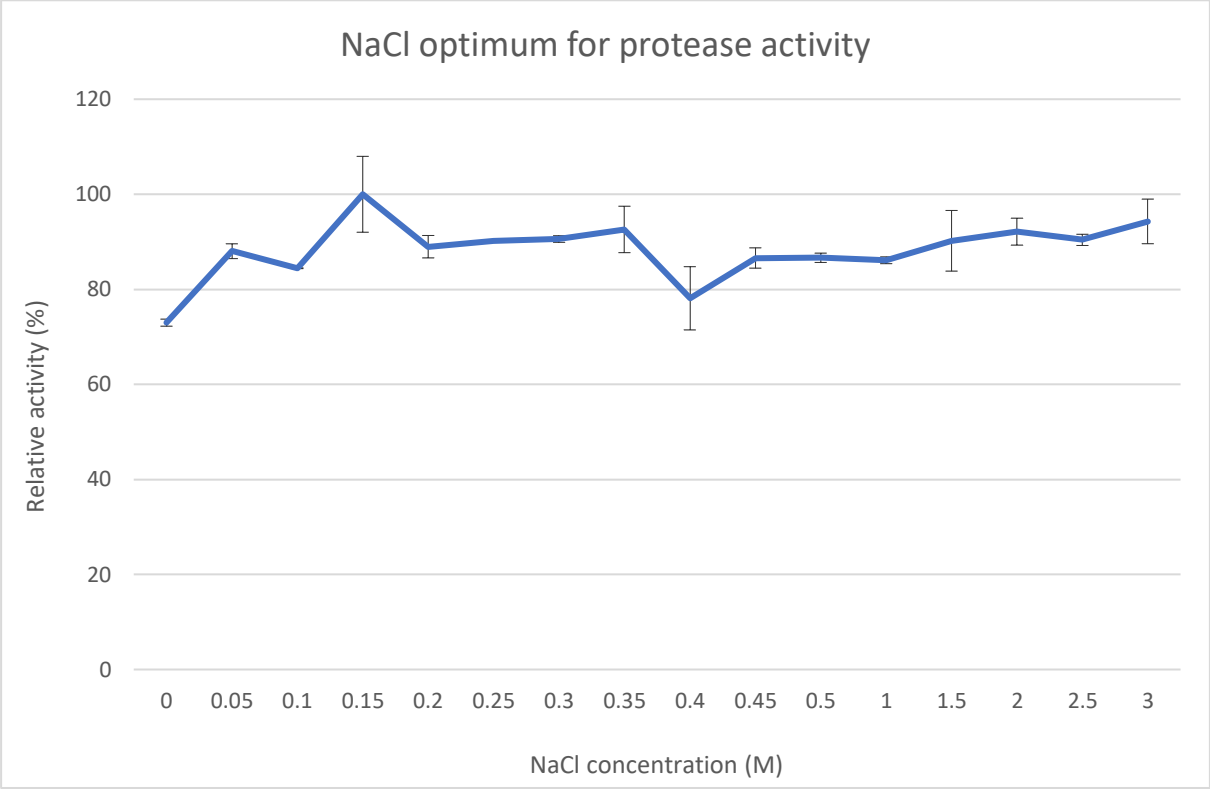

Figure 18: Optimum activity in solutions with varying salinity. The protease(s) has a moderate optimum activity at 0.15M NaCl but operates at a broad range of NaCl. The activity was tested in a 0.1M Tris-HCl buffer, pH 8.0, 10 mM $CaCl_2$ with varying concentrations of NaCl (0-3M). Sample without added NaCl (normal activity assay conditions) was set to 100%.

6 Discussion

6.1 Fungal growth and morphology studies on solid substrate

On average, the fungus thrived better when incubated at 10°C rather than 4°C, which may be explained by its temperature preferences. The fungus occurs in cool temperate waters, but has not been recorded in truly Arctic waters with surface temperatures of the warmest calendar month below 10°C (87, 88). Moreover, there were also significant differences in the growth pattern of the fungus on different media. The fungus grew best (counting in diameters) in D2MAA medium in both temperatures, but the mycelium was composed of very thin hyphae suggesting that the fungus had to grow long hyphae to quickly cover the plate in the prospection of more rich nutrition source. D2MAA medium is poor in nutrients, and this may have been the reason for the different growth pattern on this medium.

Another interesting observation was the differences in hyphae density between M19 at 10°C and 4°C, and between DCAA at 10°C and 4°C. It seems as if the lowering in temperature made the hyphae growth thinner, indicating that temperature also plays a significant role for the morphology of the fungus even though the same nutrients are available. However, the use of growth profiling/growth rates to define optimum growth temperature can be argued to be unfitting because microorganisms behave as thermodynamic units where the reaction- and growth rates increases with increasing culture temperature (41). In addition, they also behave as a biological unit where, at a given temperature, key metabolic steps that are heat labile weaken the function of some pathways (41). Therefore, it is proposed that this might explain why the overall diameter growth of *D. marina* was higher at 10°C than 4°C.

Since *Digitatispora marina* grows on driftwood substrate in its natural marine habitat, it was suggested that the addition of homogenized wood might be beneficial for the filamentous growth. However, the colony diameter was significantly smaller when wood was added to the solid media compositions for both incubations at 4°C and 10°C. Previous studies made on the principal fungus exploited to produce commercial enzymes, *Trichoderma reesei,* have suggested that the presence of lignocellulose results in formation of thick cell walls, and that the fungus is able to anchor more enzymes in the outer cell wall (89, 90) and, thus, increase the cellulolytic capacity of the fungus. This might explain why the *D. marina* hyphae grow denser and less spread within the medium when the homogenized birch wood which contains lignocellulose is present. To conclude, it seems that the fungus obtained its nutrients in a small area instead of having to spread out on the whole agar plate.

6.2 Biochemical characterizations of proteases extracted from growth medium

Enzymes characterized from cold-adapted species have been found to be more temperature and pH-labile. The proposed reason for this is their flexible structure for maintaining enzymatic activity at low temperatures (91). The activity of the proteases extracted from the growth medium of *D. marina* was assayed at different temperatures, salt concentrations and pH values of Tris-Hcl buffer using a spectrophotometer. There was a problem regarding enzyme specificity due to possibility of extracting a mix of proteases, and it is therefore not possible to state if there are one or more proteases that are responsible for the observed activity. Moreover, the peptide substrate used in the assays is also excluding some proteases with different substrate specificities.

The temperature optimum for the observed protease activity was at 50°C for this thesis study which is not unusual for proteases from cold-adapted organisms (56). However, the protease(s) maintained activity at lower ranges of temperature (20°C to 40°C) meaning that the protease(s) are active at quite low temperatures. More in-depth studies need to be done on finding the temperature optimum, and to be able to characterize the proteases(s) as heat labile at moderate temperatures, temperature inactivation studies need to be executed.

As coastal water has an average pH of around 8.0-8.3 (with variations according to special local conditions) (92, 93) it is expected that the proteases would have an optimal activity in this range. However, the pH testing done in this thesis is not representable enough as the activity was only tested at a range of pH values buffered by Tris-Hcl buffer. Therefore, it cannot be drawn any conclusions on optimal pH for the protease activity. Other buffer agents such as HEPES and Bis-tris should be used in assaying the pH-optimum of the protease activity to be able to cover pH ranges outside the Tris-HCl spectrum. Despite this, the protease activity was not significantly affected by the pH adjustments, but there was somewhat higher activity when tested with Tris-HCl at pH 8.

Considering that the protease(s) is collected from a marine derived fungus, it was of interest to investigate how the protein is adapted to salt. Coastal seawater has a salinity of about 3.5% and contains about 470 mM Na^+ and 540 mM Cl^-, giving a salinity of about 0.5 M NaCl (94). Hence, the protease activity was expected to be active at this salt concentration. As represented in Figure 9, the activity peak was measured in presence of 0.15 M NaCl. Overall, NaCl seems to be a positive addition for protease activity at all the tested concentrations since it gave higher

activity values than the protease activity measured without added salt. Because of the high salt tolerance, the protease(s) could be considered a halophilic enzyme. Moreover, since most halophilic enzymes are deactivated by NaCl or KCl concentrations of less than 2 M (95, 96), at very least this enzyme can be classified as a very salt tolerant enzyme. A salt-tolerant proteases could have biotechnological potential when applicable in sample preparation in presence of high salt such as in extraction of viral particles or blood.

6.2.1 Inhibition tests of protease activity

The protease activity was significantly inhibited with the addition of PMSF, indicating that one or more serine proteases is responsible for the observed activity. Another interesting feature was the increased protease activity towards EDTA. Although this effect was rather moderate, this could suggest that EDTA has a possible influence on the protease activity. Further, it might indicate that the enzymes(s) are less dependent on calcium or that the calcium ions are bound tight to the structure. This could be an advantage if the protease(s) are to be utilized in detergents because enzymes used for this purpose needs to withstand the presence of complex calcium and magnesium (97).

6.3 Recombinant protease expression, purification and activity testing

Eight protease genes were expressed in *E. coli* BL21 cells: Seven serine proteases (SP1, SP2, SP3, SP4, SP5, SP6, SP7) and one metalloprotease (MP). Based on the sequence similarity matrix obtained by running the sequences in ClustalOmega, there was high sequence homology (>50%) between SP7, SP4, SP1, SP3, SP2 and SP5, and it was especially high between SP1 and SP3 indicating that the two proteases are highly related to each other which also means that they probably share similar functions. None of the sequences had high homology with the reference sequences Proteinase K and ArcticZymes Proteinase, indicating that they are not highly related to them and does not necessarily share a lot of enzyme functions.

Expression in *E. coli* cells was chosen because of the fast growth kinetics of *E. coli* which was essential due to time limitations at the final stage of the thesis work. Although the *E. coli* expression system has several advantages including rapid multiplication, simple nutritional requirements and fast and easy transformation processes, there are some limitations. These include formation of aggregates of insoluble proteins (inclusion bodies) and misfolding of heterologous proteins, production of lipopolysaccharide, lack of posttranslational modification, and protein degradation due to proteases (67, 98, 99). These limitations may become more

evident when trying to express eukaryotic proteins which may be dependent on e.g. post-translational modifications or correct folding machineries only present in eukaryotic hosts.

The formation of aggregates may have been evident in SP5 which was completely insoluble and was not able to be purified by a His SpinTrap column and was therefore removed from further analysis. An expression band for SP1 was only present in the insoluble fraction (the pellet which was further whole cell lysed) suggesting that the periplasmic protein release was not successful.

Perhaps a more fitting host organism for the recombinant protein expression would be the yeast, *Pichia pastoris.* There are a lot of benefits to the *P. pastoris* expression system, including appropriate folding and secretion of recombinant proteins to the external environment of the cell (98). Moreover, the expression system has limited production of endogenous secretory proteins making the purification of recombinant proteins easier (98). The downsides to using *P. pastoris* expression system is the slow growth kinetics.

The His SpinTrap purification done on SP1, SP2, MP, SP6 and SP7 did not seem to have worked optimally as there is still a clear background in all lanes which probably is background protein from the *E. coli* host. MP lane had less background and seemed to have had the best effect of the purification. There was a clear expressed band present at around 32-35 kDa in the SP1 lane, which is a lower molecular weight then what the expressed protein should have (41 kDa). As the difference is not significant, it might just be due to regular errors in protein migration in the gel. To verify that it was in fact SP1 that had been expressed, the band was cut out and sent to MS-analysis. The results did match with SP1 and no other proteins present from the constructed library. Therefore, it might be said that the expression for SP1 was successful, however it was not overexpressed, but this can be because of the low amount (3 ul) that was loaded on to the gel.

There was a clear overexpression in the MP lane at approximately 44 kDa. However, MP have a molecular weight of approximately 32 kDa, which is usually too big of a difference to just be errors in protein migration through the gel. This band was also cut out and examined by MS-analysis. The results from the MS-analysis did not have a sequence hit with MP. However, it had a high sequence hit with SP1 (which also have similar molecular weight as the overexpression band). Moreover, there were also hit with three other sequences from the constructed protein sequence library of *D. marina*: a 14-3-3 protein, a translation elongation

factor and a histone-fold-containing protein, with molecular weights of respectively 28,9 kDa, 50,1 kDa and 11,4 kDa, none of which matched the molecular weight of the overexpressed bands. Thus, the overexpressed band is most likely stemming from a contamination of the SP1 sample. The sequence hits with the other three proteins, can be explained by the fact that the MS analysis predicted a match with a sequence if there was a presence of 2 or more peptides. Thus, since the MS-analysis also was run against proteins produced by *E. coli,* the remnants of *E. coli* cells present can have had similar sequence and identical peptides as the 14-3-3 protein, the translation elongation factor and the histone-fold-containing protein which might explain the sequence hit.

None of the His-purified proteases had activity when tested by spectrophotometric assay. Which might mean that none of them are responsible for the activity found in the initial extracts from the growth medium, or that an additional factor is needed to activate the protein e.g. moderate heat treatment or addition of other commercial proteases. As for the other two protease candidates who were not added to his-tag, and thus were not cleaned up, it is not possible to know, at this point, if they are responsible for the activity. They must be purified by other means and then tested to be evaluated. It would also be interesting to test with other substrates to see if any of the proteases would show activity against them.

There might be several reasons to why the expression was not optimal for all 8 proteases. One of them being that *E. coli* was used as the expression host and not yeast cell. Because the sequence identity from the SWISS MODEL protein structure prediction was 54% and 30% the His-tag addition might not have been precise, and the C-terminal might not be exposed after all. The proteases have a natural ability to cleave itself (56), which in turn means that the His-tag purification might not have worked because the His-tag could have been cleaved off prior to affinity purification. SignalP had significantly higher probability of the cleavage position of the signal sequence present in the proteases. Although the probability was high it is still not certain that it is the correct position which might mean that parts of the active protein have been cleaved off before insertion.

6.4 Analysis of other enzyme families present in the culture medium of *D. marina*

The InterPro scan determining which enzyme families that were present in the MS-analysis (Figure 11) estimated that the biggest portion of the enzyme sequences resulted in "None predicted" enzymes. Some of the none predicted enzymes is probably due to sequence deficiencies related to the original genome assembly. Moreover, for the sequences that did not have any indications of sequence deficiencies, might be novel or at least less studied enzymes because there are no reference sequence data available in the database.

The protease families (serine-, metallo- and glutamic proteases) and the glutamic glucomylase family detected are all found to be relevant in biofilm degradation (100) which might be interesting to look further into. The other families that were detected all seems to be involved in the degradation of lignin and sugars to fuel the fungus with energy for nutrient up-take.

6.5 Further work

The proposed further work of this project would be to purify the two proteases that did not have an added His-tag, and to further test to see if the activity is recovered here. Other possible future work is to apply either moderate heat treatment or another protease to the His-tag purified samples to see if they need to be activated by external factors. Regarding the heterologous expression, it is also suggested to test other host organisms, for examples *Pichia pastoris*. If the protease activity cannot be recovered by the suggestions above, other possibilities would be to order other sequence candidates suitable for expression, or to run another MS-analysis on even more up-concentrated and purified supernatant sample to find other hits that might be possible candidates for heterologous expression. A repeated and extended biochemical characterization of the protease activity should also be done to be able to make any concrete conclusions as the experiments were preliminary with only testing in triplicates. Another suggestion for future work, would be to test and characterize the protease activity of *D. marina* grown on the solid substrates as this might give other enzyme candidates that might better mimic the enzyme production of the fungus in its natural habitat which is on the surface of driftwood in coastal sea water and not in the free water column.

7 Conclusion

The thesis presented has used the Arctic marine fungus *Digitatispora marina* to demonstrate the potential of the bioprospecting pipeline for the discovery of protease activity. The fungus was sufficiently cultivated and the fungal secretome contained active proteases that functioned at low temperatures, varying pH values and showed high salt tolerance. However, since the experiments were not performed several times, more data is required to assure and verify these results. The expression was carried out with a successful expression on SP1, but the activity was not recovered in any of the tested proteases. Further investigations to find the gene(s) that are responsible for the activity needs to be made to be able to pursue the commercial potential of the proteases in *Digitatispora marina*.

8 References

1. Svenson J. MabCent: Arctic marine bioprospecting in Norway. Phytochemistry reviews. 2013;12(3):567-78.
2. Kubanek J, Jensen PR, Keifer PA, Sullards MC, Collins DO, Fenical W. Seaweed resistance to microbial attack: a targeted chemical defense against marine fungi. Proceedings of the National Academy of Sciences. 2003;100(12):6916-21.
3. Matz C, Webb JS, Schupp PJ, Phang SY, Penesyan A, Egan S, et al. Marine biofilm bacteria evade eukaryotic predation by targeted chemical defense. PloS one. 2008;3(7):e2744.
4. Molinski TF, Dalisay DS, Lievens SL, Saludes JP. Drug development from marine natural products. Nature reviews Drug discovery. 2009;8(1):69-85.
5. Schneider YK. Bacterial Natural Product Drug Discovery for New Antibiotics: Strategies for Tackling the Problem of Antibiotic Resistance by Efficient Bioprospecting. Antibiotics. 2021;10(7):842.
6. Capon Robert J. Marine Bioprospecting − Trawling for Treasure and Pleasure. European Journal of Organic Chemistry. 2001;2001(4):633-45.
7. Luo Y, Cobb RE, Zhao H. Recent advances in natural product discovery. Curr Opin Biotechnol. 2014;30:230-7.
8. Nilsen IW, Øverbø K, Olsen RL. Thermolabile alkaline phosphatase from Northern shrimp (Pandalus borealis): protein and cDNA sequence analyses. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 2001;129(4):853-61.
9. Streit WR, Daniel R, Jaeger K-E. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. Current Opinion in Biotechnology. 2004;15(4):285-90.
10. Winter JM, Behnken S, Hertweck C. Genomics-inspired discovery of natural products. Current opinion in chemical biology. 2011;15(1):22-31.
11. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews. 2005;69(1):195-.
12. Tabor S. Expression using the T7 RNA polymerase/promoter system. Current protocols in molecular biology. 1990;11(1):16.2. 1-.2. 1.
13. Lewin A, Wentzel A, Valla S. Metagenomics of microbial life in extreme temperature environments. Current Opinion in Biotechnology. 2013;24(3):516-25.
14. Volk TJ. Fungi. In: Levin SA, editor. Encyclopedia of Biodiversity (Second Edition). Waltham: Academic Press; 2013. p. 624-40.
15. James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, et al. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature. 2006;443(7113):818-22.
16. Hagestad OC. Bioprospecting of marine fungi from the High Arctic: A study of high latitude marine fungi from understudied taxa; bioactivity potential, taxonomy and genomics. 2021.
17. Naranjo-Ortiz MA, Gabaldón T. Fungal evolution: major ecological adaptations and evolutionary transitions. Biological Reviews. 2019;94(4):1443-76.
18. Zeghal E, Vaksmaa A, Vielfaure H, Boekhout T, Niemann H. The Potential Role of Marine Fungi in Plastic Degradation – A Review. Frontiers in Marine Science. 2021;8.
19. Stief P, Fuchs-Ocklenburg S, Kamp A, Manohar C-S, Houbraken J, Boekhout T, et al. Dissimilatory nitrate reduction by Aspergillus terreus isolated from the seasonal oxygen minimum zone in the Arabian Sea. BMC microbiology. 2014;14(1):1-10.
20. Boyce KJ, Andrianopoulos A. Fungal dimorphism: the switch from hyphae to yeast is a specialized morphogenetic adaptation allowing colonization of a host. FEMS Microbiology Reviews. 2015;39(6):797-811.

21. Raghukumar S. The marine environment and the role of fungi. Fungi in coastal and oceanic marine ecosystems: Springer; 2017. p. 17-38.
22. Spatafora JW, Volkmann-Kohlmeyer B, Kohlmeyer J. Independent terrestrial origins of the Halosphaeriales (marine Ascomycota). American Journal of Botany. 1998;85(11):1569-80.
23. Jones EBG, Pang K-L, Abdel-Wahab MA, Scholz B, Hyde KD, Boekhout T, et al. An online resource for marine fungi. Fungal Diversity. 2019;96(1):347-433.
24. Kohlmeyer J, Kohlmeyer E. Marine mycology: the higher fungi: Elsevier; 2013.
25. Kohlmeyer J, Kohlmeyer E. 1 - Introduction. In: Kohlmeyer J, Kohlmeyer E, editors. Marine Mycology: Academic Press; 1979. p. 1-6.
26. Pang K-L, Overy DP, Jones EBG, Calado MdL, Burgaud G, Walker AK, et al. 'Marine fungi' and 'marine-derived fungi' in natural product chemistry research: Toward a new consensual definition. Fungal Biology Reviews. 2016;30(4):163-75.
27. Hyde KD, Xu J, Rapior S, Jeewon R, Lumyong S, Niego AGT, et al. The amazing potential of fungi: 50 ways we can exploit fungi industrially. Fungal Diversity. 2019;97(1):1-136.
28. Gladfelter AS, James TY, Amend AS. Marine fungi. Current Biology. 2019;29(6):R191-R5.
29. Grossart H-P, Rojas-Jimenez K. Aquatic fungi: targeting the forgotten in microbial ecology. Current opinion in microbiology. 2016;31:140-5.
30. Shearer CA, Descals E, Kohlmeyer B, Kohlmeyer J, Marvanová L, Padgett D, et al. Fungal biodiversity in aquatic habitats. Biodiversity and Conservation. 2007;16(1):49-67.
31. Krull R, Wucherpfennig T, Esfandabadi ME, Walisko R, Melzer G, Hempel DC, et al. Characterization and control of fungal morphology for improved production performance in biotechnology. Journal of Biotechnology. 2013;163(2):112-23.
32. Damare S, Singh P, Raghukumar S. Biotechnology of Marine Fungi. In: Raghukumar C, editor. Biology of Marine Fungi. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 277-97.
33. Jones EG, Pang K-L. Marine fungi: and fungal-like organisms: Walter de Gruyter; 2012.
34. Abe F, Minegishi H, Miura T, Nagahama T, Usami R, Horikoshi K. Characterization of cold-and high-pressure-active polygalacturonases from a deep-sea yeast, Cryptococcus liquefaciens strain N6. Bioscience, biotechnology, and biochemistry. 2006;70(1):296-9.
35. Raghukumar C. Marine fungal biotechnology: an ecological perspective. 2008.
36. Nelson DL, Lehninger AL, Cox MM. Lehninger principles of biochemistry: Macmillan; 2008.
37. Britannica TEoE. "activation energy". Encyclopedia Britannica,. 2021.
38. Nielsen PH, Oxenbøll KM, Wenzel H. Cradle-to-gate environmental assessment of enzyme products produced industrially in Denmark by Novozymes A/S. The international journal of life cycle assessment. 2007;12(6):432-8.
39. Godfrey T, Reichelt J. Industrial enzymology: the application of enzymes in industry. 1982.
40. Chapman J, Ismail AE, Dinu CZ. Industrial applications of enzymes: Recent advances, techniques, and outlooks. Catalysts. 2018;8(6):238.
41. Feller G, Gerday C. Psychrophilic enzymes: hot topics in cold adaptation. Nature Reviews Microbiology. 2003;1(3):200-8.
42. Morgan-Kiss RM, Priscu JC, Pocock T, Gudynaite-Savitch L, Huner NP. Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. Microbiology and molecular biology reviews. 2006;70(1):222-52.
43. Feller G. Psychrophilic enzymes: from folding to function and biotechnology. Scientifica (Cairo) 2013: 512840. 2013.

44. Feller G, Le Bussy O, Gerday C. Expression of psychrophilic genes in mesophilic hosts: assessment of the folding state of a recombinant α-amylase. Applied and environmental microbiology. 1998;64(3):1163-5.
45. Cavicchioli R, Charlton T, Ertan H, Omar SM, Siddiqui K, Williams T. Biotechnological uses of enzymes from psychrophiles. Microbial biotechnology. 2011;4(4):449-60.
46. Kumar A, Mukhia S, Kumar R. Industrial applications of cold-adapted enzymes: challenges, innovations and future perspective. 3 Biotech. 2021;11(10):426.
47. Berg JM, Tymoczko JL, Stryer L. Biochemistry: Enzymes: Basic Consepts and Kinetics. New York: WH Freeman; 2002.
48. Mótyán JA, Tóth F, Tőzsér J. Research Applications of Proteolytic Enzymes in Molecular Biology. Biomolecules. 2013;3(4):923-42.
49. Seife C. Blunting nature's Swiss army knife. Science. 1997;277(5332):1602-3.
50. Neurath H, Walsh KA. Role of proteolytic enzymes in biological regulation (a review). Proceedings of the National Academy of Sciences. 1976;73(11):3825-32.
51. Li Q, Yi L, Marek P, Iverson BL. Commercial proteases: present and future. FEBS letters. 2013;587(8):1155-63.
52. Craik CS, Page MJ, Madison EL. Proteases as therapeutics. Biochemical Journal. 2011;435(1):1-16.
53. Dunn B. Determination of protease mechanism In: Proteolytic Enzymes: A Practical Approach (Beynon, B. and Bond, JS, Eds.). Oxford University Press, New York; 2001.
54. Jensen K, Østergaard PR, Wilting R, Lassen SF. Identification and characterization of a bacterial glutamic peptidase. BMC Biochemistry. 2010;11(1):47.
55. Moran AJ, Hills M, Gunton J, Nano FE. Heat-labile proteases in molecular biology applications. FEMS microbiology letters. 2001;197(1):59-63.
56. Larsen AN, Moe E, Helland R, Gjellesvik DR, Willassen NP. Characterization of a recombinantly expressed proteinase K-like enzyme from a psychrotrophic Serratia sp. The FEBS journal. 2006;273(1):47-60.
57. Solgaard G. Calanus finmarchicus finmarchicus as: Protease Activity and Encapsulation Strategies. 2008.
58. Ashraf NM, Krishnagopal A, Hussain A, Kastner D, Sayed AMM, Mok Y-K, et al. Engineering of serine protease for improved thermostability and catalytic activity using rational design. International Journal of Biological Macromolecules. 2019;126:229-37.
59. Gupta R, Beg Q, Khan S, Chauhan B. An overview on fermentation, downstream processing and properties of microbial alkaline proteases. Applied microbiology and biotechnology. 2002;60(4):381-95.
60. Barrett AJ, Woessner JF, Rawlings ND. Handbook of Proteolytic Enzymes, Volume 1: Elsevier; 2012.
61. Creighton TE. Proteins: structures and molecular properties: Macmillan; 1993.
62. Fakruddin M, Mohammad Mazumdar R, Bin Mannan KS, Chowdhury A, Hossain M. Critical factors affecting the success of cloning, expression, and mass production of enzymes by recombinant E. coli. International Scholarly Research Notices. 2013;2013.
63. Sørensen HP, Mortensen KK. Advanced genetic strategies for recombinant protein expression in Escherichia coli. Journal of Biotechnology. 2005;115(2):113-28.
64. Aune TEV. High level recombinant protein production in Escherichia coli by engineering broad-host-range plasmid vectors containing the Pm/xylS expression cassette. 2008.
65. Makrides SC. Strategies for achieving high-level expression of genes in Escherichia coli. Microbiological reviews. 1996;60(3):512-38.

66. Yin J, Li G, Ren X, Herrler G. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. Journal of biotechnology. 2007;127(3):335-47.
67. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Frontiers in microbiology. 2014;5:172.
68. Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, et al. Tuning Escherichia coli for membrane protein overexpression. Proceedings of the National Academy of Sciences. 2008;105(38):14371-6.
69. Eilerts D. Recombinant Expression and Potential Autocatalysis of Aedes aegypti Trypsin-like Serine Proteases (AaSPII and AaSPIV): San Jose State University; 2017.
70. Doguet G. MYCOLOGIE-DIGITATISPORA MARINA, NG, N. SP., BASIDIOMYCETE MARIN. COMPTES RENDUS HEBDOMADAIRES DES SEANCES DE L ACADEMIE DES SCIENCES. 1962;254(25):4336-&.
71. Jenssen M, Kristoffersen V, Motiram-Corral K, Isaksson J, Rämä T, Andersen JH, et al. Chlovalicin B, a Chlorinated Sesquiterpene Isolated from the Marine Mushroom Digitatispora marina. Molecules. 2021;26(24):7560.
72. JONES EG. Digitatispora lignicola, a new marine lignicolous Basidiomycotina. Mycotaxon. 1986;27:155-9.
73. Rämä T, Mathiassen GH, Kauserud H. Marine fungi new to Norway, with an outlook to the overall diversity. Agarica. 2014;35:35-47.
74. Rohrmann S, Molitoris H-P. Screening for wood-degrading enzymes in marine fungi. Canadian Journal of Botany. 1992;70(10):2116-23.
75. Molitoris H-P, Schaumann K. Physiology of marine fungi: A screening programme for growth and enzyme production. The Biology of Marine Fungi (ST Moss, ed) Cambridge Univ Pr, Cambridge. 1986:35-47.
76. Abdel-Hamid AM, Solbiati JO, Cann IKO. Chapter One - Insights into Lignin Degradation and its Potential Industrial Applications. In: Sariaslani S, Gadd GM, editors. Advances in Applied Microbiology. 82: Academic Press; 2013. p. 1-28.
77. Servier. Smart servier medical art [cited 2022 17.04.22]. Available from: https://smart.servier.com.
78. Scientific TF. Cytiva His SpinTrap™ Spin Columns and Kit. Thermo Fisher Scientific.
79. Bugni TS, Ireland CM. Marine-derived fungi: a chemically and biologically diverse group of microorganisms. Natural Product Reports. 2004;21(1):143-63.
80. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403-10.
81. Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 2014;42(Database issue):D26-31.
82. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nature Biotechnology. 2019;37(4):420-3.
83. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology. 2011;7(1):539.
84. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research. 2018;46(W1):W296-W303.
85. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res. 2021;49(D1):D344-d54.

86. Neu HC, Heppel LA. The Release of Enzymes from Kwherichia coli by Osmotic Shock and during the Formation of Spheroplasts. J biol Chem. 1965;240(9).
87. Booth T, Kenkel N. Ecological studies of lignicolous marine fungi: A distribution model based on ordination and classification. 1986.
88. Rämä T, Hassett BT, Bubnova E. Arctic marine fungi: from filaments and flagella to operational taxonomic units and beyond. Botanica Marina. 2017;60(4):433-52.
89. Novy V, Nielsen F, Cullen D, Sabat G, Houtman CJ, Hunt CG. The characteristics of insoluble softwood substrates affect fungal morphology, secretome composition, and hydrolytic efficiency of enzymes produced by Trichoderma reesei. Biotechnology for Biofuels. 2021;14(1):105.
90. Messner R, Hagspiel K, Kubicek CP. Isolation of a β-glucosidase binding and activating polysaccharide from cell walls of Trichoderma reesei. Archives of microbiology. 1990;154(2):150-5.
91. Outzen H, Berglund GI, Smalås AO, Willassen NP. Temperature and pH sensitivity of trypsins from Atlantic salmon (Salmo salar) in comparison with bovine and porcine trypsin. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 1996;115(1):33-45.
92. Doney SC, Fabry VJ, Feely RA, Kleypas JA. Ocean acidification: the other CO2 problem. Annual review of marine science. 2009;1:169-92.
93. Jones E, Chierici M, Skjelvan I, Norli M, Børsheim K, Lødemel H, et al. Monitoring ocean acidification in Norwegian seas in 2018, Miljødirektoratet, Report. M-1417-2019.
94. Belkin IM, Levitus S, Antonov J, Malmberg S-A. "Great Salinity Anomalies" in the North Atlantic. Progress in Oceanography. 1998;41(1):1-68.
95. Siddiqui KS, Cavicchioli R. Cold-adapted enzymes. Annu Rev Biochem. 2006;75:403-33.
96. Madern D, Ebel C, Zaccai G. Halophilic adaptation of enzymes. Extremophiles. 2000;4(2):91-8.
97. Bakhtiar S, Andersson MM, Gessesse A, Mattiasson B, Hatti-Kaul R. Stability characteristics of a calcium-independent alkaline protease from Nesterenkonia sp. Enzyme and Microbial Technology. 2003;32(5):525-31.
98. Karbalaei M, Rezaee SA, Farsiani H. Pichia pastoris: A highly successful expression system for optimal synthesis of heterologous proteins. J Cell Physiol. 2020;235(9):5867-81.
99. Sørensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microbial cell factories. 2005;4(1):1-8.
100. Molobela IP, Cloete TE, Beukes M. Protease and amylase enzymes for biofilm removal and degradation of extracellular polymeric substances (EPS) produced by Pseudomonas fluorescens bacteria. 2010.
101. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Analytical chemistry. 1996;68(5):850-8.

Appendix

Appendix A

Cryopreservation of fungal culture in glycerol solution

For long-term preservation of *Digitatispora marina,* a 20% glycerol solution was made in sterile hood by mixing 20 mL 100% glycerol and 80 mL dH_2O between an autoclaved beaker and measuring cylinder. The solution was filtrated through sterile filters into an autoclaved sDuran bottle for storing. Two cryogenic vials were labelled and added 1,0 mL of 20% glycerol solution using filtered pipette tips. Further, 4 small pieces (approximately 0,5 x 0,5 cm) was cut from the fungi culture grown on solid agar using sterile disposable knives and then placed in the vials which were then placed in a -80°C freezer for long-term storage.

Appendix B

Mass spectrometric sequencing of proteins protocol

The MS sequencing of the protein samples (delivered as polyacrylamide gel pieces) was performed by external senior engineers at Proteomics and Metabolomics Core Facility (PRiME) at the Faculty of Health Sciences (UiT). The following protocol was acquired from Jack-Ansgar Bruun (senior engineer at UiT) and is not written or edited by me.

Gel pieces were subjected to in-gel reduction, alkylation, and tryptic digestion using 6 ng/µl trypsin (V511A; Promega) (REF). OMIX C18 tips (Varian) were used for sample cleanup and concentration. OMIX C18 tips (Varian) were used for sample cleanup and concentration. Peptide mixtures containing 0.1% formic acid were loaded onto a Thermo Fisher Scientific EASY-nLC1200 system and EASY-Spray column (C18, 2µm, 100 Å, 50µm, 50 cm). Peptides were fractionated using a 5-80% acetonitrile gradient in 0.1 % formic acid over 60 min at a flow rate of 300 nl/min. The separated peptides were analysed using a Thermo Scientific Orbitrap Exploris 480 or an orbitrap Fusion Lumos mass spectrometer. Data was collected in data dependent mode using a Top20 method. The raw data were processed using the Proteome Discoverer 2.5 software. The fragmentation spectra were searched against (put in database). Peptide mass tolerances used in the search were 10 ppm, and fragment mass tolerance was 0.02 Da. Peptide ions was filtered using a false discovery rate (FDR) set to 5 % for peptide identifications (101).

Appendix C

Whole cell lysis of pellet

Table A1: Amount of PB pol lysis buffer, lysozyme and HL-San added for whole cell lysing of pellets. Calculations were made based on the OD600 measurements for each pellet sample (Control (SP1), SP1, SP2, SP3, SP4, SP5, SP6, SP7 and MP).

Protease	OD600	Volume (mL)	Desired OD600	PB pol lysis buffer (mL)	g Lysozyme (1mg/mL)	µl HL-San (400 U/mL) (Stock 1015 U/µl)
Control (SP1)	11,8	13	50	3,07	0,003068	1,209064039
SP1	5,8	13	50	1,51	0,001508	0,594285714
SP2	11,4	13	50	2,96	0,002964	1,168078818
SP3	11,8	13	50	3,07	0,003068	1,209064039
SP4	10,8	13	50	2,81	0,002808	1,106600985
SP5	9	13	50	2,34	0,00234	0,922167488
SP6	13,4	13	50	3,48	0,003484	1,373004926
SP7	11	13	50	2,86	0,00286	1,127093596
MP	13,6	13	50	3,54	0,003536	1,393497537

Equation for calculating the amount of PB pol lysis buffer added for lysing each sample:

$$\frac{Sample\ Volume\ (mL) * \ OD600}{Desired\ OD600} = PB\ pol\ lysis\ buffer\ (mL)$$

Equation for calculating the amount of Lysozyme added for lysing each sample:

$$\frac{1}{1000} * PB\ pol\ lysis\ buffer\ (mL) = Lysozyme\ (grams)$$

Equation for calculating the amount of HL-San enzymes added for lysing each sample:

$$\frac{400 * PB\ pol\ lysis\ buffer\ (mL)}{1015} = HL - San\ (\mu l)$$

Appendix D

Table A2: Terrific broth growth medium (TB medium) composition.

Solutions	**Amount**
Bacto tryptone	120 g
Bacto yeast extract	240 g
Glycerol (85%)	40 mL
Milli-Q water	9000 mL
Autoclaved potassium phosphate (pH 7,2)	1000 mL

Table A3: LB (Lysogeny Broth) medium composition.

Solutions	**Amount**
Bacto tryptone	10 g
Bacto yeast extract	5 g
Sodium chloride (NaCl)	5 g
Milli-Q water	At to 1000 mL
Sodium hydroxide (pellets)	1 piece

Table A4: PB Pol lysis buffer composition:

Solutions	**Amount**
1 M Tris-HCl pH 8,5 at 25°C	50 mL
2 M Imidazole	5 mL
Sodium chloride (NaCl)	29,2 g
1 M $MgCl_2$	5 mL
Glycerol (85%)	72,4 g
Milli-Q water	990 mL
10% sterile filtrated Triton X-100	10 mL

Table A5: MES SDS running buffer composition:

Solutions	**Amount**
MES (2-(N-morpholino) ethanesulfonic acid) 99,5%	97,6 g
SDS (sodiumdodecylsulfate)	10 g
Trizma base 99,9%	60,6 g
EDTA	3 g
Milli-Q water	10 L

Appendix E

The raw amino acid sequence of seven serine proteases (SP1, SP2, SP3, SP4, SP5, SP6, SP7) and one metalloprotease (MP) that was investigated in this thesis:

>jgi|Digmar1|SP1|CE127244_9529
MLFTATFTTLALFIGSAIAAPSTPLRQLKRADGPTSGRVIVTLKDGASKDAVLSSIGMVSIYADWDIMNA
FAVEANDDAIATLLANPDVAEVEEDGTVTTFVTQTNAPWGLSRLSSSSKLSNQNPGSLSFSYTYNSAGGN
GVDIFVVDTGVRYTHNEFGGRASWGATFGGYASSDGNGHGTHCAGTIAASRYGVAKRANIIAVKVLSDSG
SGYISDVISGLNYARSRAGSSGRPSIVSMSLGGGANTALDNAVVSLTNSGVHVVVAAGNSNTNAANTSPA
RVASAVTVGAMSIADARASFSNYGSVVDIFAPGVSVISTWATSNTATNSISGTSMATPHVAGLMAYLISV
NGNSSPASLQNTVKSMSLKNKISGIPSGTVNYLAHNS*

>jgi|Digmar1|SP2|estExt_Genewise1Plus.C_70211
MFSATFLLTALASLAQLAFAAPTDGRSTRLIVTLKDGIVPAQVVRQYARTAILPIEELTVINGFIVDSTD
ASLDDLKDDPNVESVEEDSVVSVQQDGDGSTQTDAPWGLARLYFRSKLSNQNAGKYNFEYGYHPTAGEGV
DVYVLDTGVRITHTEFEGRASWGATFGGYADADGNGHGTHCAGTIAGHRYGVAKKANIIAVKVLSDGGKG
NVGDVIEGMNYVYNRVQSTGRPSVVSMSLGGGRSGGMDAAVQSLTAAGIHVVVAAGNENEDASNHSPARA
PSAITVGATNIADKRASFSNFGPGVDIFAPGVKIKSAWKKSDTDTKRISGTSMAAPHVAGLVAHTIAWKG
NRSPEKMSTTLKSWSEKNRLKRIPDDTHNRLAHNRPSLLRFD*

>jgi|Digmar1|SP3|fgenesh1_pm.7_#_14
MHFTAAFTTLALFIGSAIAAPSAPLRQVKRADGPTSGRVIVTVKDGASKDAVLGSIGMTSVYADWDILNA
FVVEANDDAIATLLANPDVAEVEEDGTVTTFQLPRQTDAPWGLSRISSVGRLSNQNSAALTFSYTYSSAA
GSGIDVYVVDTGVRYTHTQFGGRARWGATFGGYANSDGNGHGTHCAGTIAASRYGMAKSANIIAVKVLSD
SGSGYISDVVAGLVWVRSQANSSGRPSVVSMSLGGGVNSALDNTVNLLTNAGVHVVVAAGNSNTNAANTS
PARVASAITVGAMSIADARASFSNYGSVVDIFAPGVNVISTWATSNTATNSISGTSMATPHVAGLVATML
SVNGNVSPSAMQSTIKSMAGKNLLSGIPSGTVNYLARNFYD*

>jgi|Digmar1|SP4|fgenesh1_kg.60_#_64_#_TRINITY_DN5310_c0_g2_i1
MNFACLVALLSLAPSVLGSPASRDSPLRELTRANGPSRDKFIVKLKDGANKADVMGVMSSSIAYEWSIIN
GFAVDANEETLSTLLSHPDVERVEEDGMVYADKLVTQPDAYWGMGRLSSTTQLQGKLSDPYNYTYDDSAG
EGVDIYVIDGGIYIQHPDFGGRASWGATFGDYYDYDGSGHGTHCAGTAASNTWGVAKKANLYAVRVLSDI
GEGYWSDLISGMDWVAKQAALTGRPSVATMSLGGDIMEIGDDAAEALVASGVHTTVAAGNEDMDAETHSP
ARAPSVITVGASDITDNRTYYSNYGAGVDIFAPGHFIRSTWKGSANTTRILSGTSMSTPHVGGLIAYYIA
LEGNVSPAEMATKLKSRALKGVLADIPSGTDNYLAHNAI*

>jgi|Digmar1|SP5|estExt_Genemark1.C_1_t30128
MFSAAILLTALASFSQLAFAAPVDAQASSRLIVMLKDGSGPEMITSKYARSAFAIDELQVINAVIVDSND
ISADELRALPEVSSVVEDTVVSIDNLEARARVVQPDAPWGIRRLSPIHKLNRQDDTLLDFKYKYAPSAGK
GADVYVVDTGVRITHEDFEDRATWGATFGGYADADGNGHGTHCAGTVAGARWGVAKKAHIIAVKVLSDSG
FGTVADVVAGMNYVYNTAPSTGRPSIVSMSLGGGGNTALDAAVEGLLSAGIHVVVAAGNNNGDASTRSPA
RVPGAITVGATDISDTKASFSNYGSIVDIFAPGVHITSAWATSDTATNSISGTSMATPHVAGVVAMITAK
KGNTTPPLMADIVKSWGVDNLISGLPDGTPNVLAHNRGSPSD*

>jgi|Digmar1|SP6|fgenesh1_kg.3_#_1583_#_TRINITY_DN2540_c0_g2_i1
MRFSVFVSLVCAFASVSVDAKHKLKESVAHPTAWKKVGVPPASHTITLAIALPQYRFEELDQHLQEVSDP
DHERYGQYLSHDEVNELTAPHPDSLEAVTQWLESHGITEADIVRSPANDWAHITISLPKAESMLNTTYHI
WSHESGSSMVRTTHYSLPEDIFDHIELIQPTTMFARWKGMKSPLFRGDENEAVVEPSTQKIASTGDVPPV
DASCNATITPLCLRQIYNAVGYTPQVPKKGSIAITSYLEQYINDADLQQFYAKHVPEALDSNYTLISVNG
GENLQDPAQAGVEAALDAQYAFGLTYPIPETAWTTGGRPPFNPTLHTPQNTNEPYNDWLSAVLAMRHVPK
VISTSYGEEEQSVPESYGRRACASFAQLGAKGVSVFFSSGDGGVGDGSSDPDNHLCHFDVNGETVTQFMP
SFPASCPYVTSVGGTNYVPETAVFFSGGGFSTWFPTPWYQKRTIRKYLKTLPEGLYEGLYNPKGRAIPDV
AALGRRFEIIYQGRSGLIGGTSASAPAFAAFTALLNDARAAKGLPSLGFLNPLLYSRGYKALNDIADGGT
NPGCGTVGFSADEGWDPVTGLGTPNFGELLKLVT*

>jgi|Digmar1|SP7|fgenesh1_pm.2_#_486
MFSTKLATILLGLTTFSLATHAIQVPRAASEQKKGDWNGPGRYIVTLKQGVNKLDSMDTLSSKVKDQWAI
IPGFAGKFHFFGLHLEADEETFKTLISDPDVEKVEEDGEVYAHDAGWGLARLGADGALNSTDTSANYTYT
YDDSSGRGVDIYVIDTGIFVQHSQFGGRAVWGASFGGNSNADGNGHGTHSAGLIGGQRFGVAKLARLIAV
KVLSDQGSGSISGVISGMNYVMQQAAASGRPSIANIAIGGAANSALDALTASGIHVVVAAGASNTNAGNT
SPARAPSAITVGATDITDARALFSNYGHVVDIWAPGVSIQSAWIGSATDTRVLSGTSMSTAYISGLIAYL
IGLYGNQNPAQMTATMQSFAQYGVLSKIPSGTANIMAHNGV*

>jgi|Digmar1|MP|fgenesh1_kg.3_#_431_#_TRINITY_DN3239_c0_g2_i1
MLFTSLALTLFGATAVFAAPGEVTRICGTEISDAEVAVQESSFPGILSEAKTNSEARIAPATVTLSVYFH
VIYSSTSLSGGYVPDYQIASQIDVLNGAYSNMGIRWRLANTTRTQNSGWFNSAGPDTSAQTSMKNALRQG
GAADLNVYSVGFNSGSGAGLLGYATFPWSYSGSPKDDGVVILYSSVPGGTTGNYNEGHTLTHEAGHWVGL
YHTFQGGCSGSGDSVSDTPAESSPAFGCPTGRDTCSGGGVDPIHNFMDYTYDSCMNQFTSGQGGRTADYM
LAYRGVDTGL*