Doctoral Thesis

Study on gene regulation of protease in filamentous fungus *Trichoderma reesei*

糸状菌 *Trichoderma reesei* におけるプロテアーゼの遺伝子発現制御に関する研究

Submitted June 2020

Nagaoka University of Technology
Department of Bioengineering
Integrated Bioseeince and Technology

Daranagama Arachchige Nayani Dhanushka Daranagama

Under the Supervision of
Professor Wataru Ogasawara

Table of Contents

Chapter 1: Introduction ... *4*

1.1 Background ... **4**

1.2 Utilisation of plant biomass for the bioeconomy ... **8**

1.3 Cellulase producers ... **10**
1.31 *Trichoderma reesei* ... 10
1.32 Cellulases and xylanases ... 11
1.33 Cellulases and Xylanase produced by *T. reesei* ... 13

1.4 Glycoside hydrolase induction mechanisms in *T. reesei.* ... **14**
1.41 Transcription factors involved in cellulase gene expression in *T. reesei* ... 15
1.42 Recently published specific cellulase regulators in *T. reesei* ... 18

1.5 Problem of protein production in *T. reesei* ... **19**
1.51 Global functions of proteases ... 20
1.52 Secreted proteases in *T. reesei.* ... 21
1.53 The main aspartic protease in *T. reesei* –Trichodermapepsin (TrAsP) ... 25
1.54 Purification of the main aspartic protease in *T. reesei* ... 27
1.55 Transcription factors involved in protease production ... 27
1.56 Regulation of extracellular proteases in fungi ... 30

1.6 Aims of the study ... **33**

1.7 References ... **34**

Chapter 2: Role of Trichodermapepsin in T. reesei ... *40*

2.1 Introduction ... **40**

2.2 Materials and Methods ... **41**
2.21 Strain and media ... 41
2.22 Construction of QMΔ*trasp* ... 41
2.23 Total RNA extraction and quantitative reverse transcription PCR (qRT-PCR) ... 43
2.24 Mass spectrometry analysis ... 44
2.25 Purification of TrAsP ... 45
2.26 Measurement of proteolytic activity ... 45
2.27 Biomass saccharification ... 46

2.3 Results ... **47**
2.31 Specificity of proteases secreted by *T. reesei* ... 47
2.32 Identification of TrAsP in the *T. reesei* secretome. ... 48
2.33 Construction of *QMΔtrasp* ... 50
2.34 Specific activity of TrAsP ... 51
2.35 Purification of TrAsP ... 52
2.36 Characterisation of TrAsP ... 52
2.37 Specific activity of purified TrAsP ... 53
2.38 Role of QMΔ*trasp* ... 55
2.3.9 Stability of cellulases in Saccharification conditions ... 57
2.3.10 Saccharification ... 57
2.3.11 *trasp* production in response to environmental sources ... 59

2.4 Discussion ... **63**

2.5 References ... **67**

Chapter 3: Cis acting element analysis of Trichodermapepsin in T. reesei ... *70*

3.1 Introduction ... **70**

3.2 Materials and Methods ... 72
3.21 Strains and media ... 72
3.23 Construction of plasmids for reporter analysis and transformation ... 73
3.24 Preparation of cell-free extracts and β-glucuronidase assay ... 73
3.25 Measurement of proteolytic activity ... 74
3.26 RNA extraction and real-time quantitative reverse transcription PCR (qRT-PCR) ... 77
3.27 Expression and purification of recombinant Xyr1 and Are1 ... 78
3.28 Electrophoretic mobility shift assay (EMSA) ... 78
3.29 Statistical Analysis ... 79

3.3 Results ... 81
3.31 Involvement of Xyr1 and Are1 in *trasp* expression ... 81
3.32 Analysis of promoter deletants ... 83
3.33 Site-directed mutagenesis ... 84
3.34 Binding ability of Xyr1 and Are1 transcription factors to the TrAsP promoter ... 87
3.35 Effect of other cellulase regulators on *trasp* expression ... 89

3.4 Discussion ... 91

3. 5 References ... 98

Chapter 4: General Conclusions ... 103

Chapter 5: Publications ... 106

5.1 Journal Papers ... 106

5.2 Conference Proceedings ... 106

Chapter 6: Acknowledgments ... 107

Chapter 1: Introduction

1.1 Background

As the total human population has increased in the 21st century, industrial activities have also drastically increased. In 2019, the world population is close to 8 billion and is predicated to reach 10 billion by 2050, according to reports from the United Nations. Therefore, the world must be prepared for global issues due to anthropogenic activities. Climate change is a major global issue that can potentially impact food production, result in rising sea levels and have a variety of unforeseeable consequences. Increased levels of greenhouse gases (GHGs) in the atmosphere are a major issue. The emission of GHGs have drastically risen due to 150 years of industrialisation, deforestation and large-scale agriculture, leading to direct effects on global warming of the Earth. The most abundant GHG is carbon dioxide (CO_2), which is primarily produced by the burning of fossil fuels. To solve this issue, the UN devised The Intergovernmental Panel on Climate change (IPCC). In 2018, this group issued a special report that aimed to limit global warming to 1.5 °C and to limit the total global net emission of CO_2 by humans to zero by 2050. This suggests that CO_2 emission must be balanced by the recycling of CO_2. Scientists have been focused on climate change for over 30 years and there are three main United Nations legal facets that have been created to address the climate change problem. These facets will be described in more detail below.

United Nations Framework Convention on Climate Change (UNFCCC)- In 1992, the UNFCCC was established along with the Earth summit. The aim of the UNFCCC is to prevent dangerous anthropogenic interference with the climate system.

Kyoto protocol- In 1997, the Kyoto protocol was created to bind emission targets for developed countries and help establish global carbon markets. In 2008, at the G8 summit, Cool Earth 50 was developed by Japan, which aimed to develop strategies to reduce CO_2 emission by 50% by 2050.

Paris Agreement- In 2015, the Paris agreement was established to combat climate change and to accelerate and intensify the actions and investments needed for a sustainable low carbon future. For the first time, most UN countries agreed on the need to limit global temperature increases well below 2 °C above the pre-industrial levels. They also agreed to strengthen the ability of countries to deal with the impacts of climate change and established a commitment to switch away from dirty fossil fuels and towards clean forms of energy (Unfccc 2016).

In 2019, the climate summit convened to bring world leaders together in order to support and accelerate climate action and ambitions. At this summit, actions were taken to build more concrete plans for a green economy.

On the other hand, population growth results in other serious global environmental issues, such as unsustainable use of natural resources and threats to ecosystems. These problems must be tackled directly and must be resolved in order to secure the availability of natural resources for our society. The sustainable society concept focuses on the maintenance of a sound and rich environment for use by future generations, while also meeting the needs of the present society. This concept has three main pillars, including economic, environmental and social pillars (Purvis et al. 2019). These three pillars are informally referred to as 'people, planet and profits'. According to the report on the cabinet decision in Japan in 2007, a sustainable society is envisioned as a 'low carbon society', a 'sound-material cycle society' and a 'society in harmony with nature' (Cabinet Meeting Decision 2007). Becoming a low carbon society calls for a drastic reduction in greenhouse gas emissions due to fossil fuel consumption. A sound-material cycle society calls for the reduction of waste and the recycling of resources through all stages of our socio-economic activities, including the collection, production, distribution, consumption and disposal of resources. A society in harmony with nature calls for proper conservation of biodiversity and that socio-economic activities, including agricultural, forestry and fishery operations, be carried out in harmony with nature. This also involves the creation of various opportunities and venues for society to experience and enjoy the natural environment. This balance is widely recognised by various international organisations, such as the International Union for Conservation of Nature (IUCN), the United Nations Environment Programme (UNEP) and the World Wildlife Fund (WWF). The revised sustainability society index (SSI) was developed by the sustainable society foundation in the Netherlands in 2006. The SSI is comprised of eight policy categories and three wellbeing dimensions, including human, environmental and economic, and has been calculated for 151 countries, accounting for 99% of the world population (Saisana and Philippas 2012). According to the spider web graph of sustainability, the world is still lacking in sustainability. Over a 10-year period, the environment wellbeing scores have been low, due to low scores for greenhouse gas emissions, energy use and renewable energy (Fig 1.1).

In 2015, the leaders of the United Nations began another agenda to mitigate the world problems of inequality and injustice, end extreme poverty and tackle climate change, and set 17 sustainable development goals (SDGs) (Fig. 1.2). In order to build a sustainable society

and achieve those goals and targets, this task must be undertaken by everyone across the world.

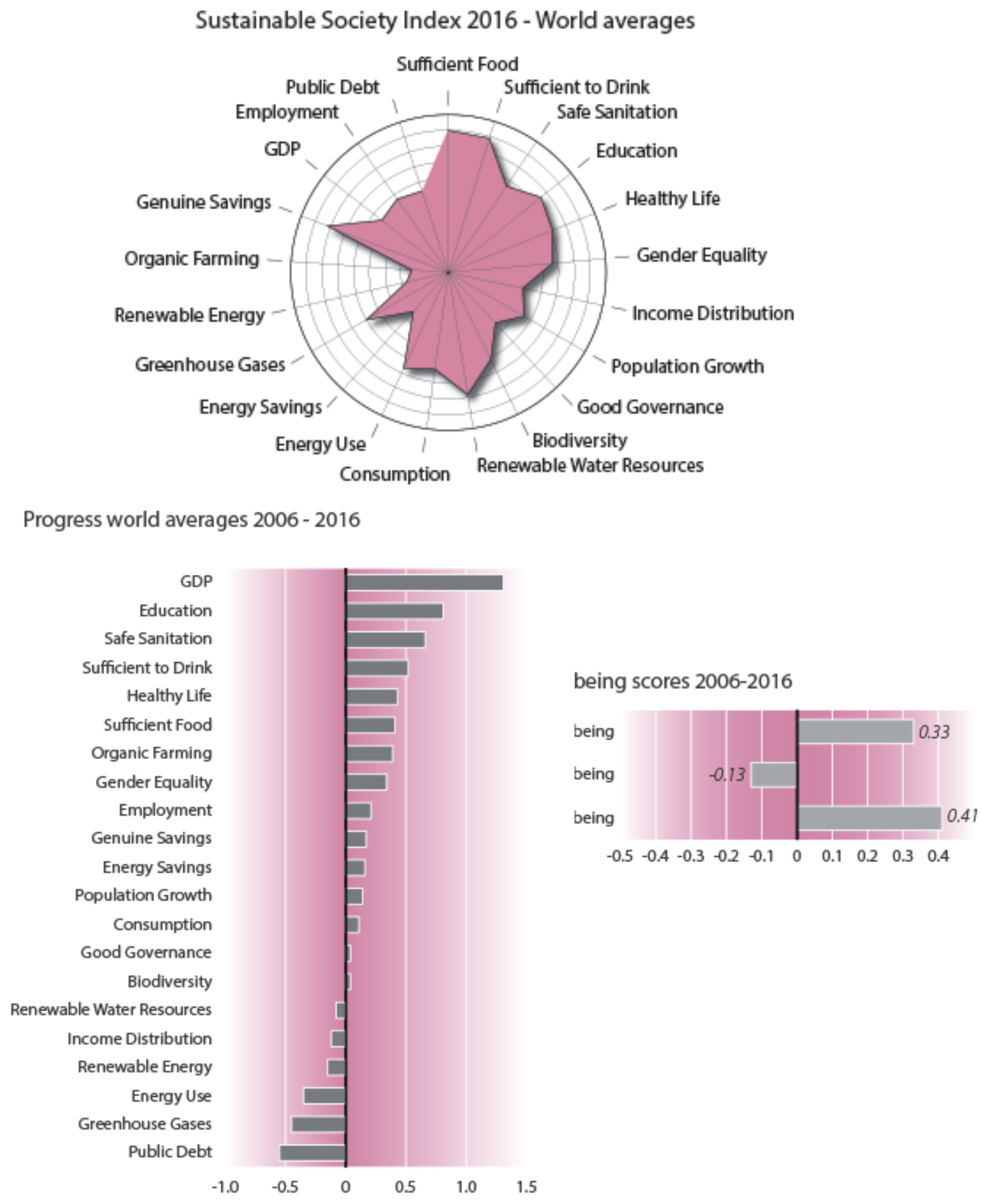


Fig.1.1 Sustainable Society Index (SSI) 2016 world averages. The spider web presents the value of indicator scores on a scale of 1 to 10. All total scores are calculated as the average scores per person, so the number of inhabitants of a country has been taken into account (http://www.ssfindex.com/results/main-results-2016/).

In order to obtain a sustainable society, our energy and transportation should be altered to a low-carbon or zero-carbon society. A low carbon society requires a variety of sustainable energy solutions to replace fossil fuels. Renewable energy sources include solar power, wind power, hydro power, biomass and biofuel, geothermal energy, marine energy and recyclable metal fuels (Bergthorson 2018). Green vehicles or eco-friendly vehicles were introduced for transport. Electric cars were invented in the 1880s and are now popular in many countries. In addition to electric cars, fuel cell powered hydrogen cars and hybrid cars, which only partly use fossil fuels or biofuels and are part electric and hydrogen powered, are also frequently used for transportation.

Fig.1.2 Sustainable development goals.

According to the International Energy Agency, transport biofuel production expanded by 6% in 2019 (96 Mtoe-Millions of tonne of oil equivalent or 161 billion litres) and a 3% annual growth in production is expected over the next five years (Fig 1.3) (https://www.iea.org/reports/tracking-transport-2019/transport-biofuels). On track with sustainable development, transport biofuel consumption must be reduced to 298 Mtoe by 2030. The report mentions that this falls short of the sustained 10% output growth per year needed until 2030 to align with the SDS. In order to increase the production of biofuel, the main biofuel producers, such as the United States, European Union, Brazil, India, China and

ASEAN (Association of Southeast Asian nations) must be engaged with this target. Stronger support of these policies and innovations are required to reduce the production costs. Therefore, novel advanced biofuel production technologies should be developed to achieve sustainable development targets.

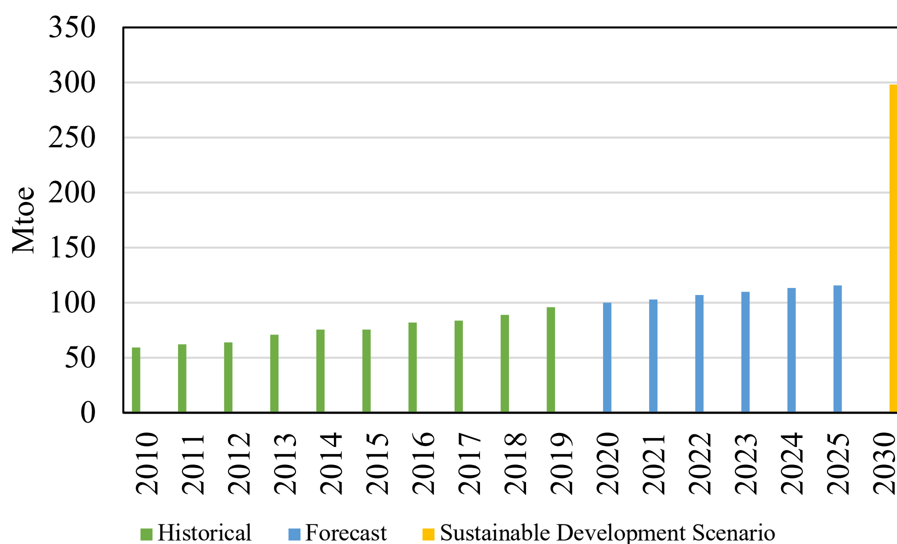


Fig.1.3 Global biofuel production from 2010-2025 as compared to consumption in the sustainable development scenario (Tracking report- May 2019).

1.2 Utilisation of plant biomass for the bioeconomy

Biotechnology uses biological systems or living organisms to solve problems or make useful products. Biotechnological innovations contribute to the bioeconomy, which is comprised of all industries related to biological resources and related services. Biotechnological industries play an essential role in both economic growth and sustainable development. Industrial or white biotechnology uses microorganisms and enzymes to produce materials, such as chemicals, plastics, food, agricultural and pharmaceutical products and energy carriers, for industries by utilising renewable raw materials, such as waste from agriculture and forestry.

Bioenergy or biomass energy is a renewable environmentally friendly alternative energy source used to overcome the issues of the depletion of fossil fuels and increasing greenhouse gases due to combustion of fossil fuels. Bioenergy accounts for three-quarters of all renewable energy use today and makes up half of the most cost-effective options for doubling renewable energy by 2030. The biorefinery concept integrates biomass conversion processes and equipment to produce fuels, power, heat and value-added chemicals from plant

biomass, which is the most abundant renewable resource in nature (Fig. 1.4). This process reduces greenhouse gas emission by 85% as compared to that of reformulated gasoline, as CO_2 is absorbed from the air by plants for photosynthesis. Globally, renewable energy only accounted for 18.1% of the total energy consumption in 2017. According to the 2019 report by REN21, 7.5 % of this is comprised of traditional biomass and 1% is biofuel used for transport (REN21 2019).

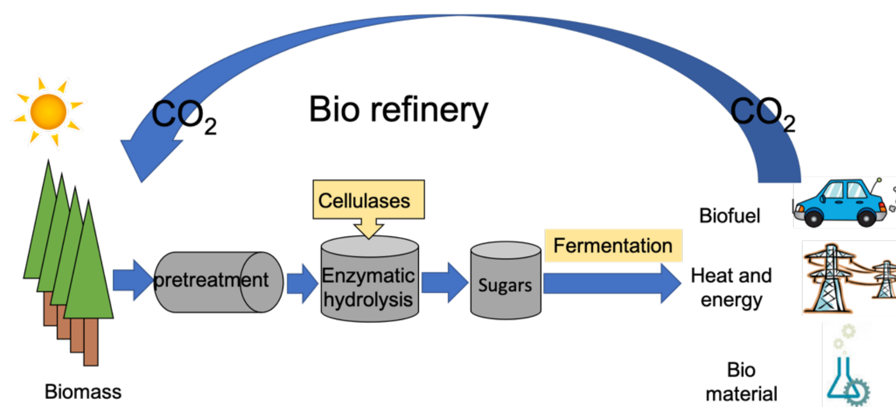


Fig. 1.4 Schematic representation of the biorefinery process.

Plant biomass is made up of lignocellulosic material, such as cellulose, which is the main component (30–50%), hemicellulose, which is the second most abundant component (20–35%), and lignin (5–30%). Plant biomass must be converted to glucose and is then subjected to the fermentation process to produce bioethanol and value-added materials. Bioethanol and lignocellulosic biomass can be converted to many products (Fig. 1.5) (Chandel et al. 2018). The main barrier in this process is efficient degradation of plant biomass. Several technologies have been developed over the past 80 years that allow this conversion process to occur, and the objective is to make this process cost-competitive in today's markets (Himmel et al. 2007).

Acid saccharification methods have been developed over many years. However, this technique has drawbacks, including environmental unsustainability and the production of harmful byproducts. Thus, further development of this technique cannot be expected. Therefore, researchers have been focused on other techniques, such as the use of enzymes for saccharification, as this is an environmentally friendly method. However, a large number of enzymes are needed for efficient degradation. Naturally, some microorganisms secrete large amounts of cellulases and microbial cellulases have a higher demand for the saccharification process. These enzymes can be applied in many industries, including the food, animal, textile, pulp and paper, and bioconversion industries, such as bioenergy.

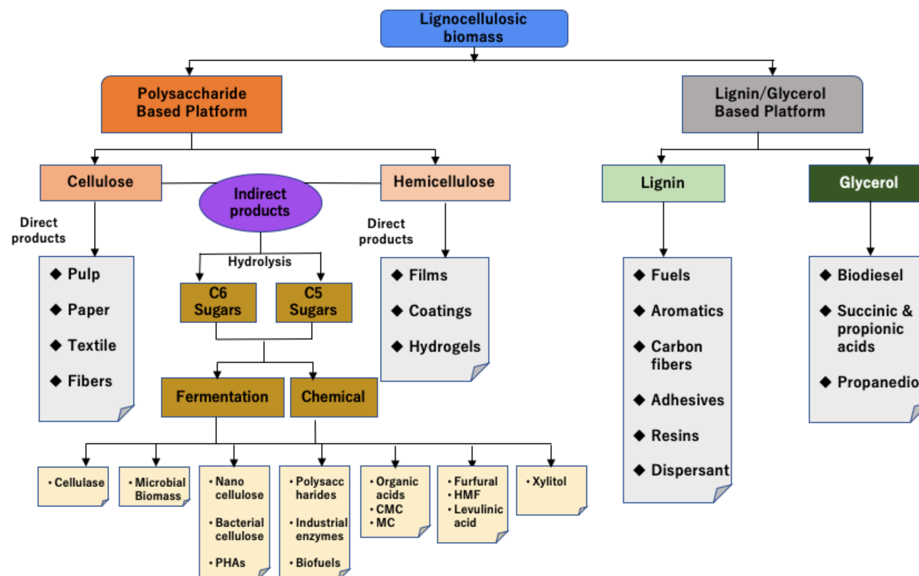


Fig. 1.5 Commercial bioproducts from lignocellulosic biomass in lignocellulose biorefinery platform (Chandel et al. 2018).

1.3 Cellulase producers

Cellulases are enzymes that degrade lignocellulosic materials into simple sugars, such as glucose. There are some natural cellulase producers in nature, which are manly comprised of fungi, bacteria and protozoans. In many herbivorous animals, such as cattle, sheep and horses, cellulases are produced by symbiotic bacteria. Metazoan animals, including termites, snails and earthworms, also produce endoglucanses. Some green microalgae, such as *Chlamydomonas reinhardtii* and *Gonium pectoral,* also produce cellulases.

1.31 *Trichoderma reesei*

Aspergillus, *Trichoderma* and *Penicillium* species are well-known cellulolytic filamentous fungi. Among them, the fungus *T. reesei* (teleomorph: *Hypocrea jecorina*, division Ascomycetes) produces relatively large quantities of cellulases and xylanases and has been used industrially for this reason. This fungus was originally isolated in the Solomon Islands by the US army during World War II, as was used as a degrader of cotton fabrics. This wild type *T. reesei* is called QM6a. The mutant strains were developed using UV irradiation and chemical mutagenesis to improve the production of extracellular proteins. The

mutant strain *T. reesei* QM9414 is now generally used as a standard strain. Industrial strains of *T. reesei* are able to produce more than 100 g/L of extracellular proteins (Cherry and Fidantsef 2003).

1.32 Cellulases and xylanases

Cellulose is made up of a fibrous structure that consists of insoluble bundles of linear D-glucose polymers (60–70 chains) held together by interchain hydrogen bonds and van der Waals interactions. The polymers are comprised of 1,4-linked β-D-glucopyranose units with a degree of polymerisation of up to 15, 000 units. The highly ordered and insoluble nature of this crystalline cellulose makes it more resistant to enzymatic degradation.

Enzymatic degradation of cellulose requires the synergistic action of different types of cellulases, which are classified as endoglucanases (EG: EC 3.2.1.4), exoglucanases/cellobiohydolases (CBH: EC 3.2.1.91) and β-glucosidase (BGL: EC 3.2.1.21) (Fig. 1.6).

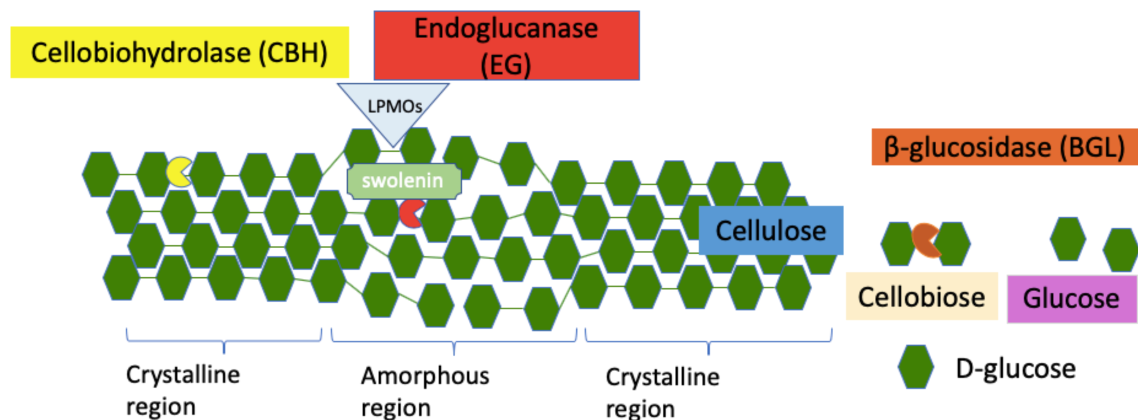


Fig. 1.6 Enzymatic degradation pattern of cellulose.

CBH and EG synergistically act on cellulose to produce cello-oligosaccharides and cellobiose, which is a dimer of glucose and contains a β-1-4 bond. EG generates oligosaccharides of different lengths through random cleavage of the β-1-4 bonds along the amorphous region of the cellulose polymer. This results in more free chain ends in the cellulose that become accessible for CBHs, which are exoglucanases, and act on both the reducing and the non-reducing ends of the cellulose. Cellodextrinase or 1,4-β-D-glucan glucanohydrolases (EC 3.2.1.74), are another type of cellulase that acts specifically on

cellooligosaccharides, whereas BGL hydrolyses cellobiose to glucose (Lynd et al. 2002). A protein that disrupts the crystalline structure of a cellulase, called swolenin (Bai et al., 2012), and lytic polysaccharide monooxygenases (LPMOs) are recently discovered classes of enzymes that have been recognised as effective auxiliary enzymes for cellulose degradation (Harris et al., 2010). LPMOs are currently grouped into (AA) families, and there are three AA9 LPMOs encoded in *T. reesei* genome (Filiatrault-Chastel et al. 2019).

Hemicellulose, the second most abundant component of biomass, is a complex heterogeneous polysaccharide that is mainly composed of xylose and arabinose. The xylan backbone contains D-xylose as its monomeric unit, and its side groups can be replaced with hexose and sugar acids (Fig. 1.7).

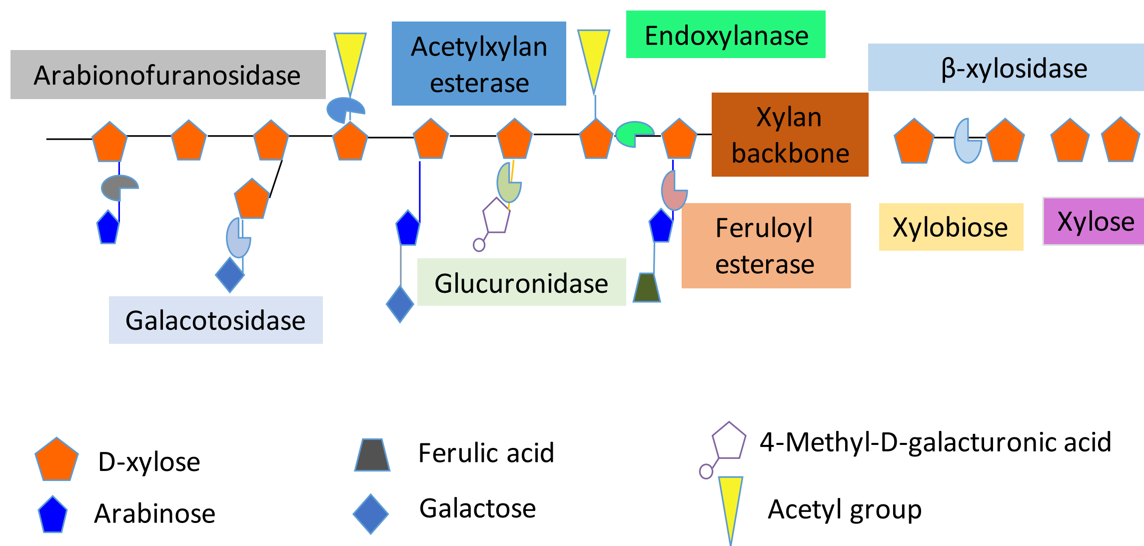


Fig. 1.7 Enzymatic degradation pattern of hemicellulose.

The heterogeneous nature of xylan requires a variety of enzymes for its degradation, including endoxylanases (EC 3.2.18), which cleave the β-1,4 glycosidic linkage in the xylan backbone, arabinofuranosidases (EC 3.2.1.55), which remove arabinose side chains, β-xylosidases (EC 3.2.1.37), which release xylose from xylo-oligosaccharides, acetyl xylan esterases, which remove acetyl groups from the xylan backbone, and feruloyl and ferulic acid esterases (EC 3.11.73), which remove ferulic acid from the xylan side chains (Sun et al. 2012).

1.33 Cellulases and Xylanase produced by *T. reesei*

The glycoside hydrolase enzyme system of *T. reesei* has been subjected to extensive biochemical and genetic analysis, as it has industrial importance. Some enzymes have been characterised and the corresponding genes have been isolated, before sequencing of the whole genome. Previous reports have established that *T. reesei* produced at least two cellobiohydrolases (CBH1 and CBHII), six endoglucanases (EGI to EG VI) and two β-glucosidases (BGLI and BGL II) (S. Shoemaker, V. Schweickart, M. Ladner, D. Gelfand, S. Kwok 1983; Saloheimo et al. 2002). Additionally, for hemicellulases, *T. reesei* was known to produce at least four endo-xylanases (XYNI to XYNIV), one β-xylosidase (BXLI), one α-L-arabinofuranosidase (ABFI) and one acetyl xylan esterase (AXE, EC 3.1.1.72) (Törrönen et al. 1992; Gerber et al. 1997). Several other enzymes have also been identified in the *T.reesei* secretome, such as β-mannmosidase (EC 3.2.1.78) and α-galactosidase (EC 3.21.22) (Margolles-Clark et al. 1996; Margolles-clark et al. 1997). In addition to these enzymes, the transcriptional effect of cellulase inducers for cellulose degradation have been identified by the expressed sequence tag (EST) and microarray analysis (Foreman et al. 2003). In 2008, the whole genome of *T. reesei* QM6a was sequenced (Martinez et al. 2008). After determination of the complete genome sequence of *T. reesei*, 9129 gene models were predicted for the 33 Mb genome sequence. Limitations in genome sequencing were identified (Li et al. 2017) and comparative genome analysis was also performed using different strains of *T. reesei* (Koike et al. 2013). However, after genome sequencing, previously identified genes were altered, specially the endoglucanase, as two endoglucanases, EGIV and EG VII, were classified as LPMO. The cellulases (Druzhinina and Kubicek 2017) and xylanases secreted by *T. reesei* are listed in Table 1.1. The *T. reesei* genome is available at the Joint Genome Institute website (http://genome.jgi psf.org/Trire2/Trire2.home.html).

Table 1.1 *T. reesei* secreted cellulases and xylanases

Name	Protein ID	Glycoside hydrolase Family
Cellobiohydrolase		
CBHI (Cel7A)	123989	GH7
CBHII (Cel 6A)	72567	GH6

Endoglucanase		
EGI (Cel7B)	122081	GH7
EG II (Cel5A)	120312	GH5
EG III (Cel12A)	123232	GH12
EG V (Cel45A	49976	GH45
β-glucosidase		
CEL3A(BGLI)	76672	GH3
CEL3B	121735	GH3
CRL3C	82227	GH3
CEL3E	76227	GH3
CEL3F	104797	GH3
CEL3H	108671	GH3
CEL3J	66832	GH3
CEL1B	22197	GH1
CEL1A	120749	GH1
Xylanase		
XYN I	74223	GH11
XYN II	123818	GH11
XYN III	120229	GH10
XYN IV	111849	GH15
XYNV	112392	GH11
Xylosidase		
BXLI	121127	GH3
XYL3B	58450	GH13
Arabinofuranosidase		
ABF1	123283	GH54
ABFII	76210	GH62
ABFIII	55319	GH54
Mannanase		
MANI	56996	GH5
LPMO		
LPMO 9A (formerly EG IV)	73643	AA9
LPMO 9B (formerly EG VII)	120961	AA9

1.4 Glycoside hydrolase induction mechanisms in *T. reesei.*

The production of cellulolytic and xylanolytic enzymes are regulated at the transcriptional level depending on the carbon source available, in *T. reesei*. Reports suggest that cellulase production is induced by cellulose and its derivatives, such as cellobiose and sorbose (Morikawa 2001). The highest induction level was achieved with cellulose and sophorose, and moderate expression was observed when cellobiose or lactose were used as the carbon source (Ilmèn et al., 1997). Sophorose, whose synthesis from cellobiose involves transglycosylation activity of β-glucosidase, is the most known inducer of cellulase

expression. Carbon catabolite repression (CCR) involves the repression of cellulase induction in glucose medium. Galactose is a part of lactose and hemicellulose derivatives, and does not induce cellulase genes, but rather, only induces the expression of the α-galactosidase encoding genes (*agl1* and *agl2*) and the acetyl xylan esterase encoding gene (*axe1*) (Margolles-Clark et al., 1997).

Recently, the light and pH responsive system in *T. reesei* received attention for its glycoside hydrolase induction mechanisms. Light is thought to be important for extracellular substrate degradation. 16 glycoside hydrolase (GH) genes, were upregulated under light conditions, including the two major cellulase genes, *cel7a/cbh1* and *cel6a/cbh2*, the endoglucanase *cel5a/egl2*, the xylanase *xyn2* and the beta-xylosidase *bxl1*. Transcription of *cel61a/egl4* and an additional 16 genes in the GH-family were decreased under constant light, suggesting that *T. reesei* adjusts its cellulase mixture to different physiological requirements under light and dark conditions (Tisch et al., 2011). The pH-dependent enzyme production of *T. reesei* was assessed and results showed that cellulase activity increased when the pH decreased from 6 to 4, and that xylanases are preferably produced at a higher pH (up to pH7) (Bailey et al., 1993). In other filamentous fungi, the nitrogen responsive pathway is also thought to be involved in cellulase expression. Inorganic nitrogen has been shown to be sufficient for cellulase induction in *T. reesei* and organic nitrogen has been shown to play a role in growth of the fungi (Rodriguez-Gomez and Hobley, 2013). However, in *T. reesei*, there has been little attention given to the nitrogen source responsive gene regulation system. In a recent paper, the GATA-type transcriptional activator, Are1, was identified and characterised in response to available nitrogen sources in the environment (Qian et al. 2019a). According to Qian *et al.* Are1 is involved in the regulation of cellulase expression in the presence of the preferred nitrogen sources, such as ammonium.

1.41 Transcription factors involved in cellulase gene expression in *T. reesei*

Expression of cellulases and hemicellulases are transcriptionally regulated by the signals that not only come from carbon, but also nitrogen, pH and other environmental conditions. Positive and negative transcription factors were identified as being involved in the regulation of the expression levels of cellulases and hemicellulases (Fig. 1.8). Xylanase regulator (Xyr1), activator of cellulase 2 and 3 (Ace2 and Ace3) and the Hap2/3/5 complex have been identified as positive factors, and Ace1 and carbon catobolite repressor, Cre1, have been identified as negative regulators. Many studies have investigated the functions of these

factors under carbon source inductive mediums. Pac1 has been reported as a regulation factor in *T. reesei,* as it is responsive to extracellular pH and it is homolog of PacC in *A. niger.* A possible role in regulation of the production of cellulose-degrading enzymes has been suggested, as up-regulating *pac1* is seen in cellulose (dos Santos Castro et al., 2014).

1. Xyr1

The primary positive regulator of cellulase and hemicellulose gene expression is represented by Xyr1 (xylanase regulator 1). This is a zinc binuclear cluster protein that binds to GGCWWWW (Furikawa et al., 2009). *xyr1* deletion abolishes cellulases (*cbh1, cbh2, egl1* and *bgl1*) and xylanases (*xyn1, xyn2* and *bxl1*), regardless of the inducer (xylose, xylobiose, α -sophorose, lactose and sophorose). *xyr1* expression was regulated by the Cre1 dependent CCR and Ace1 transcription factors. Recent microarray analysis of *T. reesei* showed that deletion of *xyr1* modulated carbohydrate active enzymes (CAZymes), transcription factors and transporters. Moreover, down regulation of transporters belonging to the MFS and ABC transporter families was also identified in the strain with *xyr1*deletion.

2. Ace2

Activator of cellulase (Ace2) is also a zinc binuclear cluster protein. This protein has been shown to only occur in *T. reesei.* Surprisingly, homologs of this gene have not been found in the genome sequences of *A. nidulans, A. niger* and *N. crassa.* Deletion of *ace2* decreases the expression of major cellulases during growth on cellulose, but not during induction on sophorose, suggesting that two different mechanisms exist for cellulase expression. Ace2 has the ability to bind to the promoter motif of GGC$(T/A)_6$ that is present in *cbh1* and GGGTAA in the *cbh2* and *xyn2* promoters (Shida et al., 2016). In addition, an inverted repeat motif of GGCTGG has been reported to have binding ability to the *xyn2* promoter region (Stricker et al., 2008).

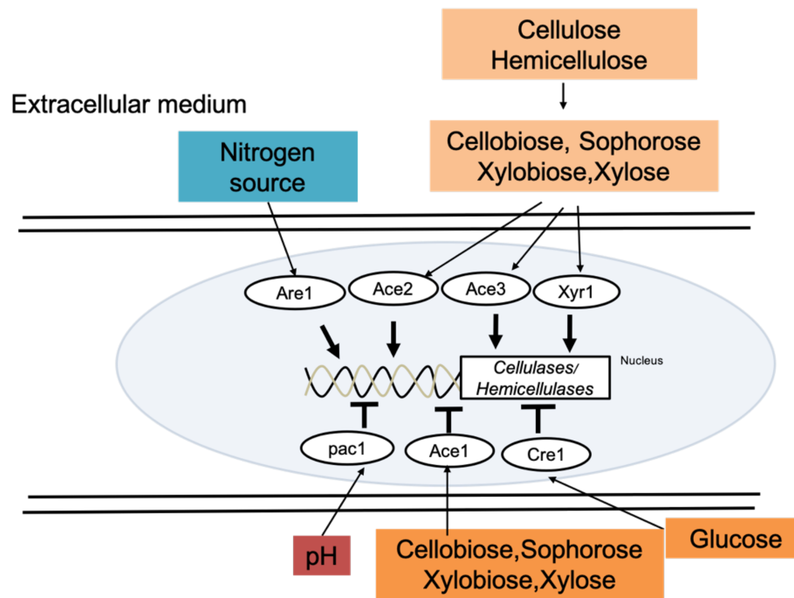


Fig. 1.8 Schematic representation of the regulatory function of transcription factors in *T.reesei.*

3. Ace3

Activator of cellulase expression 3 (Ace3) is a positive regulator of cellulase expression and has a fungal-specific transcription factor domain. Deletion of Ace3 in *T. reesei* led to a significant reduction in both cellulase and hemicellulase genes and also affects *xyr1* expression. In contrast, the *ace3* gene expression was slightly decreased in the deletion strain of *xyr1* in the cellulose medium (Dos Santos Castro *et al.,* 2016).

4. Ace1

Activator of cellulase (Ace1) was first thought to be an activator but has since been identified as a repressor of cellulase expression and has been shown to belong to the three Cys_2His_2-type zinc finger. Ace1 binds to AGGCA motifs within the *cbh1* promoter region and to a GGCTAA motif within the *xyn1* promoter region. Deletion of Ace1 in *T. reesei* led to increased expression of all of the major cellulase and xylanase genes with both cellulose and sophorose induction.

5. Cre1

Cre1 proteins, which are a Cys_2His_2 type transcription factor, mediated CCR and repressed cellulase gene expression in an easily metabolised carbon source, such as D-glucose. In *T. reesei,* the RUT-C30 mutant, *cre1,* has been deleted in its genome, resulting in

greater cellulase production. The functional Cre1 binding site has two closely spaced SYGGRG motifs. Studies suggest that such double binding sites are needed for direct Cre1 repression. In *T. reesei, cbh1* and *xyn1* expression have been demonstrated to be under the direct control of Cre1. Recently, additional factors controlling CCR have also been reported in *T. reesei*, including Cre2 and Cre3, which are orthologs of CreB and CreC of *A. nidulans*. Cre2 is a ubiquitin C-terminal hydrolase that may be involved in the deubiquitination of Cre1. In *A. nidulans,* CreB is stabilised through the formation of a complex with CreC, which had 48% amino acid identity with *T. reesei* (Denton and Kelly, 2011).

6. Pac1

Pac1 is a transcription factor homolog of the PacC transcription factor found in *Aspergillus* spp. and contains a DNA-binding domain with three Cys_2His_2 zinc fingers. Pac1 binds the hexanuclotide sequence, GCCARG, to promoter sites. The genes regulated by PacC, and its homologs, were cellulases, xylanses, proteases, enzymes involved in the synthesis of biological active metabolites, such as penicillin, and an endo polygalacturonease that is involved in pectin degradation. In *T. reesei,* xylanase genes, *xyn2*, *xyn3* and GH30 xylanase genes are abundantly expressed at high pH. At low pH, *xyn1* and *xyn5* are highly expressed. At neutral pH, deletion of *pac1* dramatically increased cellulase production and the transcription levels of major cellulase genes. Therefore, Pac1 acts as an activator for alkaline expressed genes and prevents the expression of acid-expressed proteases.

1.42 Recently published specific cellulase regulators in *T. reesei*

1. Xpp1

Xylanase promoter binding protein 1 (Xpp1) is a transcription factor that regulates gene expression of xylanolytic enzymes. Xpp1 represses *xyn2* due to binding the AGAA region in the promoter. Xpp1 contains a basic-loop-helix protein and its target sequence is a hexameric palindrome, 5′-WCTAGW-3′, with an inverted AGAA-repeat. Xpp1 has no involvement in d-xylose metabolism and on the expression of cellulases. The Xpp1 expression is upregulated under both D-glucose and high D-xylose conditions (Derntl et al. 2015).

2. SxlR

SxlR is a specialised xylanase repressor with a zinc binuclear cluster and may bind

the promoters of GH11 xylanase genes (*xyn1, xyn2 and xyn5*). It does not regulate GH10 (*xyn3*) and GH30 (*xyn4*) xylanases genes. The binding sequence of SxlR is 5′-CATCSGSWCWMSA-3′(Liu et al. 2017).

3. BglR

BglR is a Zn(II)(2)Cys(6)-type fungal-specific transcription factor identified in the analysis of single nucleotide polymorphisms between *T. reesei* PC-3-7 and its parent, KDG-12, which regulates β-glucosidases (Nitta et al. 2012). It was found to share little homology to AmyR *Aspergillus oryzae,* that is an amylase gene activator responsible for maltose. The *bglr* mutant and PC-3-7 mutant had increased cellulase production during the growth of cellobiose. BglR upregulates specific β-glucosidase genes (with the exception of *bgl1*) after the glucose produced in cellobiose triggers CCR. In *N. crassa,* the BglR homolog, Col-26 and VIB1 (vegetative incompatibility blocked) is a p53-like transcription factor are involved in the repression of glucose signalling. Deletion of both *col26* and *cre1* are able to reverse the effects of deletion of vib1(Xiong et al. 2014).

4. Rce1

Rce1 is a transcriptional repressor of cellulase gene expression in *T. reesei* and was identified using a yeast based one hybrid screen. Deletion of *rce1* facilitated the expression of cellulase genes and did not involve xylan-dependent induction of xylanases in *T. reesei*. In addition, Rce1 can directly bind to the *cbh1* gene promoter through competition with Xyr1(Cao et al. 2017).

1.5 Problem of protein production in *T. reesei*

T. reesei is a dominant commercial cellulase producer used in biotechnological industries. The secreted protein in *T. reesei* can be as high as 100 g/L (Cherry and Fidantsef 2003). However, degradation products of cellobiohydolases and endoglucanses were detected in secretomes due to proteolytic activities (Kubicek-Pranz et al. 1991). Therefore, proteases in the secretome influenced the production of cellulases. In order to enhance cellulase production, low-protease level strains or protease deficient strains have to be developed or isolated (Qian et al. 2019b).

1.51 Global functions of proteases

Proteases, also known as peptidases, proteinases and proteolytic enzymes, are enzymes that degrade proteins and peptides (Fig. 1.9). They hydrolyse peptide bonds (CO-NH_2) that link amino acids together in a polypeptide chain.

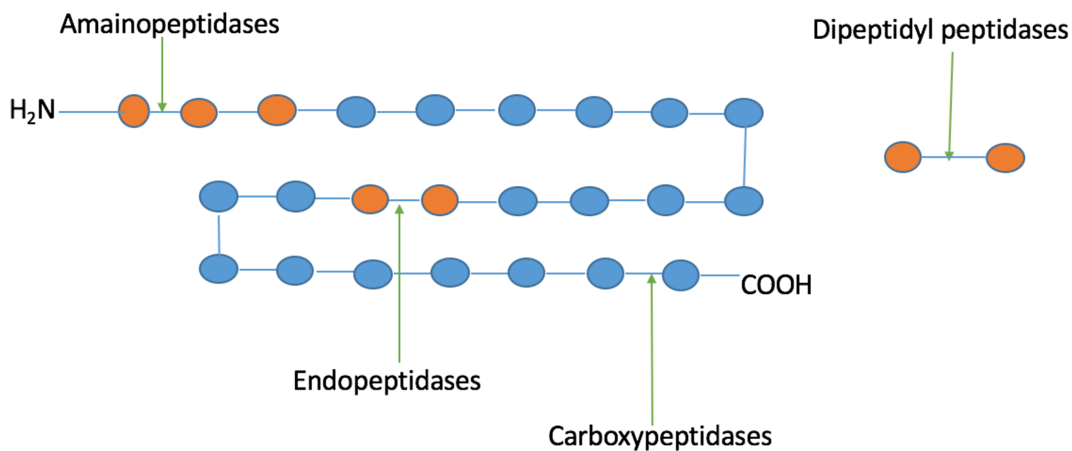


Fig. 1.9 General mechanisms used to break down protein.

Proteases are a highly complex group of enzymes that differ in their substrate specificity, catalytic mechanisms and active site. Proteases are divided in to two groups, depending on the mode of action, either exopeptidases or endopeptidases. Exopeptidases, also known as peptidases, are classified according to whether they release single amino acids from the N-terminus (aminopeptidases) or C-terminus (carboxypeptidases) of peptide chains. Endopeptidases, also known as proteinases, cleave peptide bonds internally within a polypeptide. Based on the functional group present at the active site, endoproteinases are further classified into four prominent groups, either serine, cysteine, aspartic or metalloproteases.

Proteases are essential for all life forms and catalyse proteolysis, which is an irreversible posttranslational modification that affects every protein at all stages, from biosynthesis to degradation. Microorganisms contain a large array of proteases both intracellularly and extracellularly. Generally, intracellular proteases are involved in many physiological pathways, including nutrition, growth, differentiation, signalling and death (Fig. 1.10) and extracellular proteases are important for the hydrolysis of proteins in the cell-free environment. Proteases are also essential in viruses, bacteria, fungi and protozoa for colonisation, invasion, dissemination and evasion of host immune responses, mediating and

sustaining the infectious disease process (Maeda 1996; Ellison et al. 2011)

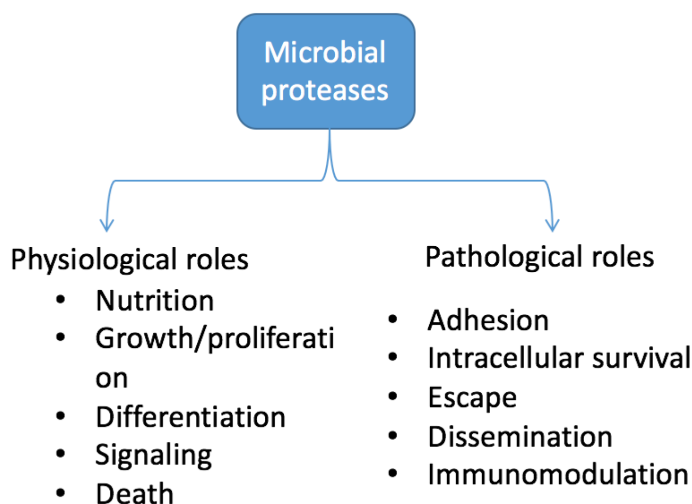


Fig. 1.10 Global function of proteases in microorganisms

1.52 Secreted proteases in *T. reesei.*

Comparing the secreted proteases in cellulolytic filamentous fungi (Table 1.2), *T. reesei* also produces a large number of proteases. Table 1. 3 shows the proteolytic enzymes in the *T. reesei* secretome, according to the *Trichoderma* genome database. The major groups include aspartyl proteases, serine proteases, subtilisin-like proteases and dipeptidyl and tripeptidyl peptidases. However, there are few studies about protease production in *T. reesei*.

Table 1.2: Number of peptidases produced in several cellolytic filamentous fungi. (source: genome database)

Fungi	Number of total peptides
Trichoderma reesei QM6a	279
Trichoderma reesei v2.0	286
Aspergillus niger ATCC 1015	388
Aspergillus nidulans	291
Penicillium oxalicum 114-2	256
Neurospora crassa OR74A v2.0	242

In *T. harizanum,* acid proteases are regulated by pH and nitrogen sources, such as yeast extract, peptone and casein (Delgado-Jarana et al., 2000). Alkaline and neutral proteases seem to be induced only by lactose and chitin, carbon starvation and some organic nitrogen sources, such as casein. Fungal cell wall-induced aspartic protease has also been recorded in *T. harizanum* (Suárez MB et al., 2005). Recently, the cold adaptive proteinase was purified from *T. atroviride* (Kredics L et al, 2008). However, the production and regulation of protease in *T. reesei* is still in its infancy. The trypsin-like alkaline serine protease, Tvp1, role was identified as degrading CBH1 in *T. reesei* QM9414 near pH at 6.0 (Dienes et al., 2007). In a recent study, in response to corn and cellulose medium with a pH value near 4.0, 4 proteases were identified, including the trypsin- like serine protease (tre73897), the aspartic protease (tre79807), the aminopeptidase (tre81070) and the zinc metallopeptidase (tre105808). At a pH near 6.5 two serine proteases, including the subtilase-like protease (tre51365) and subtilisin-related protease (tre123234), were observed. At both pH 4.0 and 6.5, six proteases (tre22459, tre77579, tre81517, tre103039, tre120998 and tre123244) belonging to aspartic protease, serine protease as well as carboxypeptidase groups were shared between the two conditions (Qian, 2019b). The types of proteases that are secreted depends on different cultivation contents. As protease production mechanisms have not been clear, this study also aims to elucidate the protease inductive mechanisms in *T. reesei.*

Table 1.3 Proteolytic enzymes in the *T. reesei* secretome based on the *T. reesei* genome

Annotation/Genome (Merops)		Number of genes in *T.reesei*	Annotation description
AA	Pepsin	15	Water nucleophile; water bound by two Asp from monomer or dimer; all endopeptidases, from eukaryote organisms, viruses or virus-like organisms
AD	Persenilin	3	Water nucleophile; water bound by two Asp; all membrane-bound endopeptidases, known only from eubacteria with active site on cytoplasmic side of cell membrane
CA	Papain Ubiquitin C-terminal hydrolase	4 3	Cysteine nucleophile; catalytic residues in Cys, His, Asn (or

	Ubiquitin-specific protease C39 Aut2 peptidase	12 1 1	Asp) in sequence
CD	Legumin Capase	1 2	Cysteine nucleophile; catalytic residues in His and Cys in sequence
CE	Pyroglutamyl peptidase 1	1	Cysteine nucleophile; catalytic residues in His, Glu (or Asp) and Cys in sequence
CP	C97	1	Cysteine nucleophile; catalytic residues in His and Cys in sequence
GA	G01	6	Water nucleophile; water bound by Gln and Glu; all endopeptidases
I	I78 I87	1 1	Families of peptidase inhibitors not assigned to clans
JE	IC	1	
M	Prenyl protease 2	1	Families of metallopeptidases not assigned to clans
zincin	Asp-zincin Glu-zincin Met-zincin	1 8 2	Water nucleophile; water bound by single zinc ion ligated to two His (within the motif HEXXH) and Glu, His or Asp
MC	Carboxypeptidase A	3	Water nucleophile; water bound by single zinc ion ligated to His, Glu (within the motif HXXE) and His
ME	Pitrilysin	1	Water nucleophile; water bound by single zinc ion ligated to two His (within the motif HXXEH) and Glu
MG	M24	8	Water nucleophile; water bound by two cobalt or manganese ions ligated by Asp, Asp, His, Glu, Glu
MH	M18 M20 Aminipeptidase Y	2 7 9	Water nucleophile; water bound by two zinc ions ligated by His (or Asp), Asp, Glu, Asp (or Glu), His

MJ	Membrane dipeptidase Beta-aspartyl dipeptidase	2 11	Water nucleophile; water bound by two zinc ions ligated by His, His or Asp, Lys or Glu, His, His.
MP	M67	4	Water nucleophile; water bound by single zinc ion ligated to two His and Asp (within the motif HSHP(X)9D)
MT	M81	1	Water nucleophile; water bound by single zinc ion ligated to Asp, His and His.
PA	PA(S)	1	Serine or cysteine nucleophile; catalytic residues in His, Asp, Ser (or Cys) in sequence; all endopeptidases
Ntn-hydrolase	PB(C) PB(T)	3 20	Threonine, serine or cysteine nucleophile at the N-terminus of mature enzyme
PC	PC(C) Gamma-glutamyl hydrolase Pfpl endopeptidase PC(S) S51	 2 1 1	Serine or cysteine nucleophile; catalytic residues in Cys (or Ser), His, Glu in sequence.
PE	DmpA aminopeptidase T05	3 1	Serine nucleophile; catalytic residues in Ser, His, Glu in sequence.
S-	S81	1	Families of serine peptidases not assigned to clans
SB	Subtilisin S53	10 4	Serine nucleophile; catalytic residues in Asp, His, Ser in sequence in family S8, but family S53 differs
Alpha/beta hydrolase	Prolyl oligopeptidase S10 S28 S33	42 4 3 43	Serine nucleophile; catalytic residues in Ser, Asp, His in sequence
SE	D-Ala-D-Ala carboxypeptidase B	6	Serine nucleophile; catalytic residues in Ser, Lys (within the motif SXXK) in sequence
SF	Signal peptidase I	2	Serine nucleophile; catalytic residues in Ser, Lys (or His) in sequence

SK	ClpP endopeptidase	1	Serine nucleophile; catalytic residues in Ser, His, Asp in sequence
ST	Rhomboid	3	Serine nucleophile; transmembrane protein with catalytic residues in the order Ser, His embedded in the membrane
MERC 0:C02 6.952		1	
MERC 0:C10 8.00		15	
MERC 0:U52. UPW		2	
Total		**286**	

1.53 The main aspartic protease in *T. reesei* –Trichodermapepsin (TrAsP)

T. reesei is the major organism used for the production of industrial cellulases and is also moderately useful for the production of heterologous protein. The heterologous production of foreign proteins in *T. reesei* is limited due to its high expression of proteases (Zhang et al., 2014). Previous studies have shown that in some acid proteases contain a Kex2-like dibasic endopeptidase and alkaline protease (spw) (Haab et al,1990, Eneyskaya et al., 1999, Zhang et al., 2014). *T. reesei* produces 15 pepsins, according to the genome database (Table 1.4). According to RNA expression analysis of these proteases, the highest value was recorded in Trichodermapepsin (TrAsP: MEROPS ID: A01.027, peptidase family A1), which is an acidic protease (Fig. 1.11). Therefore, it is thought to be the major aspartic protease secreted by *T. reesei*. The same protein was previously named pep1 in a different report (Landowski et al. 2015). Acid proteases are endopeptidases, with molecular masses that range from 30 to 45 kDa, and depend on aspartic acid residues for their catalytic activity (Leng and Xu, 2011). Its native crystal structure and the structure of the complex it forms with its inhibitor and pepstatin A, were defined by Nascimento et al 2008.

Table 1.4 Pepsin produced by *T.reesei*

Protein ID	Merops Annotation	Merops Name
Trire2 119876	A01.UPA	subfamily A1A unassigned peptidases
Trire2 76887	A01.082	SA76 peptidase (*Trichoderma sp.*)
Trire261536	A01.UPA	subfamily A1A unassigned peptidases
Trire2 81004	A01.044	podosporapepsin
Trire2 108686	A01.UNA	subfamily A1A non-peptidase homologs
Trire2 53961	endothiapepsin	endothiapepsin
Trire2 111818	A01.UNA	subfamily A1A non-peptidase homologs
Trire2 122076	A01.UPA	subfamily A1A unassigned peptidases
Trire277579	saccharopepsin	saccharopepsin
Trire2 121133	A01.080	PepAb peptidase
Trire2 58669	A01.UPA	subfamily A1A unassigned peptidases
Trire2 105788	A01.UNA	subfamily A1A non-peptidase homologs
Trire2 110490	A01.081	PepAc peptidase
Trire2 74156	trichodermapepsin	trichodermapepsin
Trire2 79807	A01.077	CtsD peptidase (*Aspergillus*-type)

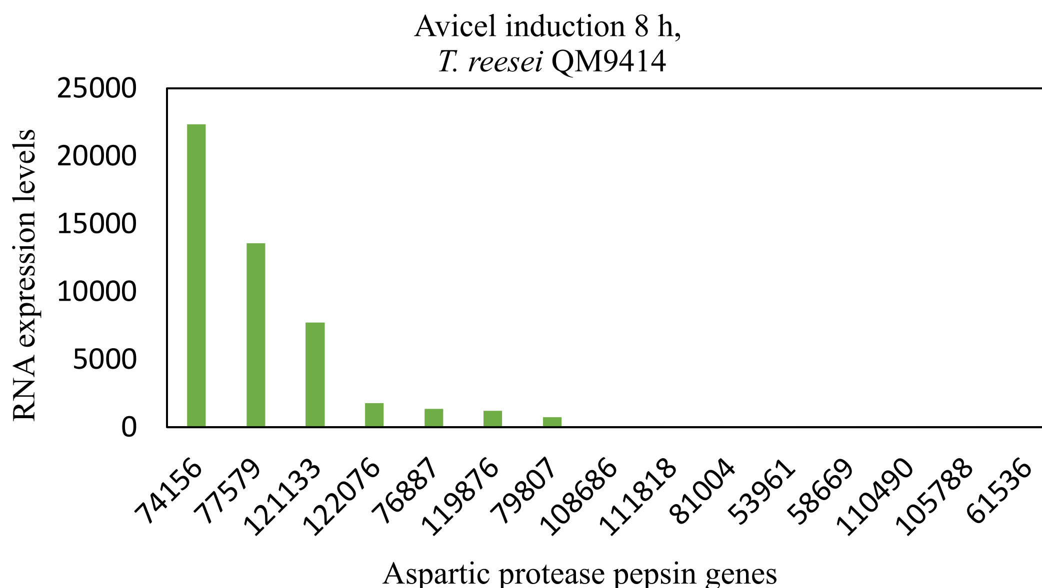


Fig. 1.11 Expression of aspartic acid proteases from microarray data

1.54 Purification of the main aspartic protease in *T. reesei*

Generally, *T. reesei* is cultivated at 28 °C pH 4 for the production of cellulases enzymes. Under these conditions both cellulases and proteases are produced. There are many studies that have investigated the purification and characterisation of the main protease in *T. reesei* (Table 1.5).

Table 1.5 Previous studies of the purification of the main protease of *T. reesei*

References	Factors				
	Molecular weight	Protein ID	Stain of T. reesei	Cultivation media	Specific activity
Landowski et al., 2015	42.7 kDa	74156	Q6a	Lactose	0.04 mg/mL protein[2]
Nascimento et al., 2008	36 kDa (329-residue)	EC 3.4.23.18, (3C9X)	NR	NR[1]	NR
Eneyskaya et al., 1999	32-kDa	EC 3.4.23.18	D38	wheat bran+pepton	NR
Dunne et al.,1992	35.5 kDa	NR	QM9414	Casein	190 unites protease/mg[3]
Haab et al., 1990	42.5 kDa	NR	QM9414	Cellulose+protein	NR

NR: Not recorded 1 dry powder of *T. reesei* used for purification. 2,3 Purification stage and substrate were not mentioned.

1.55 Transcription factors involved in protease production

Even though the mechanisms of cellulase production at the transcription level have been studied in *T. reesei*, the regulation of proteases is not well known. Similarly, studies of

protease production in other filamentous fungi have not been extensively studied. Few transcription factors have been identified in *Aspergillus* spp. (Fig. 1.12).

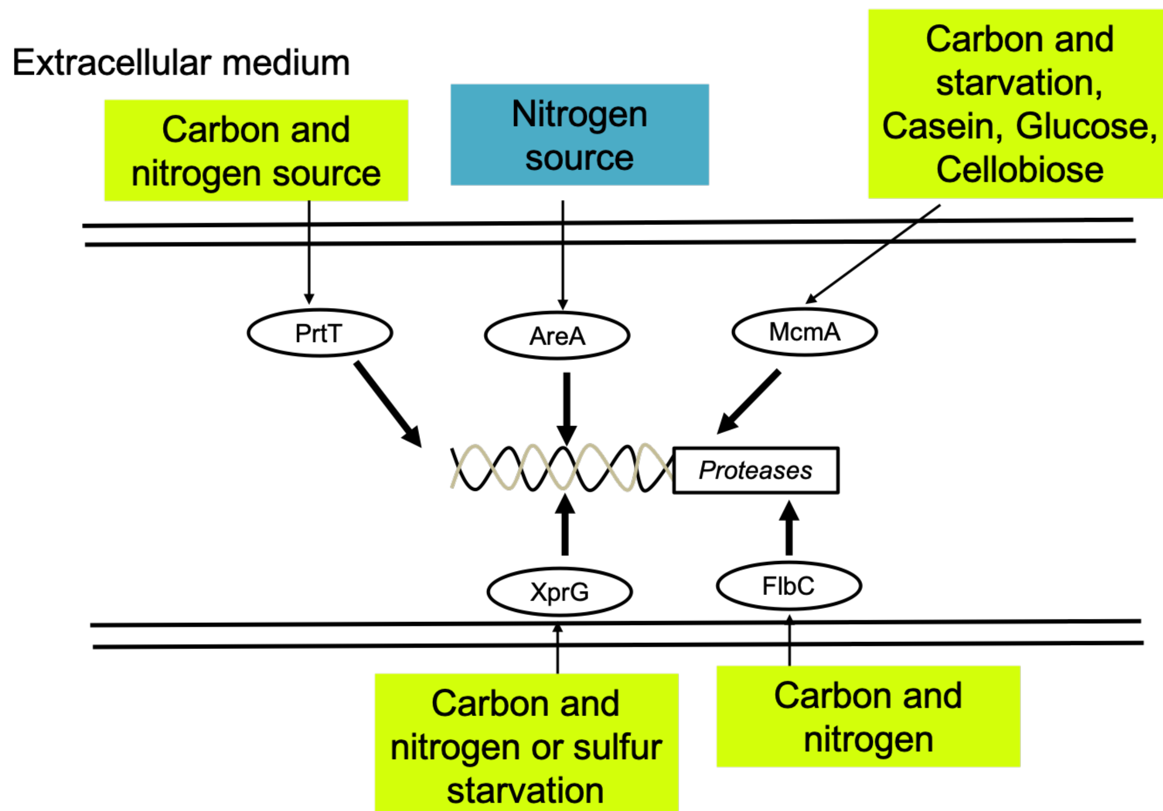


Fig. 1.12 Schematic representation of transcription factors that regulate proteases in *Aspergillus* app.

1. PrtT

Protease regulator (PrtT) is a regulator that was identified in *A. niger* that belongs to the fungal Zn(II)$_2$Cys$_6$ cluster transcription factor family (Punt et al. 2008). The PrtT transcription factor has been identified in *Aspergillus* spp., except *A. nidulans.* No other PrtT orthologs could be found in other non-*Aspergillus* species or any other related species (Punt et al., 2008). Under carbon starvation conditions, PrtT is strongly upregulated in *A. niger*. Disruption of *prtt* resulted in dramatic decreases in secreted proteases, including the major acid proteases in *A. niger,* PepA and PepB. Other extracellular proteases, including neutral metalloprotease I (NpI) in *A. oryzae*, tripeptidylpeptidase A (TppA) in the pathogenic *A. fumigatus* Af293, and alkaline serine protease (Alp), were significantly reduced after disruption of *prtt* (Sharon et al. 2009) Results from a microarray analysis in *A. fumigatus* suggested that PrtT also participates in the regulation of genes related to iron uptake, ergosterol biosynthesis and secondary metabolite biosynthesis. PrtT of *Penicillium oxalicum*

has been shown to regulate the expression of carboxypeptidase, aspergillopepsin tripeptidyl-peptidase, extracellular dipeptidyl-aminopeptidase V and five amylase genes, as well as three major facilitator superfamily transporter genes related to maltose, monosaccharide and peptide transport. Strains with deletion of *prtt* in *P. oxalicum* have reduced expression levels of the protease and increased expression levels of MFS transporters and amylases and in nutrient rich conditions in casein medium, the cellulase genes were upregulated (Chen et al. 2014).

2. AreA

AreA is a transcription factor that is involved in the regulation of nitrogen catabolic enzymes. AreA has a highly conserved DNA binding motif that is comprised of a Cys(4) zinc finger followed by a basic domain that binds to the GATA sequence. High affinity binding of AreA to the HGATAR motif in *A. nidulans* has been reported (Ravagnani et al. 1997). AreA activates the transcription of many structural genes that are involved in nitrogen source catabolism under nitrogen-limited conditions. In *A. nidulans,* under conditions that contain cellulose and ammonium, a mutant with loss of function of the *area* gene had reduced levels of cellulase. However, cellulase activity was elevated in a strain that had a constitutively active *area* allele. AreA mutants have also been shown to regulate extracellular proteases in *A. nidulans* (Hynes 1974; Lamb et al. 1997). Orthologs of AreA in *A. nidulans* were identified in *Neurospora crassa,* NIT2, *Pencillium chrysogenum,* NreA, *Fusarium oxyporum,* Fnr1, *Magnaporthe grisea,* Nut1 and *Saccharomyces cerevisiae*, Gln3 and Gat1. In *Penicillium marneffi,* AreA-dependent protease production was also observed and was found to be dependent on temperature and nitrogen levels (Bugeja et al. 2012). In *T. reesei,* the AreA ortholog (Protein ID: 76817) has 37% amino acid identity with that of *A. nidulans.* In a recent study, this ortholog was characterised as Are1 and was found to be necessary for the activation of expression of extracellular proteases in *T. reesei* in the presence of non-preferred nitrogen sources, such as skim milk (Qian et al. 2019a).

3. XprG

XprG is a transcription factor that plays a key role in the production of extracellular proteases in response to nutrient limitations in *A. nidulans.* With mutations in *xprg*, production of extracellular proteases in response to carbon starvation is abolished and the response to reduced levels of nitrogen and sulphur is also reduced. Genetic data suggests that XprG may modulate AreA-mediated activation of gene expression of extracellular proteases

and that *xprg* may be involved in Cre1-mediated extracellular protease regulation in *A. nidulans* (Katz ME et al., 2008).

4. McmA

In *A. nidulans,* mcmA is an SRF-MADS type transcription factor that has the ability to bind to CeRE and contains the binding consensus sequence $CC(A/T)_6GG$. McmA is involved in the regulation of cellulytic genes, protease genes, as well as genes involved in asexual and sexual development. RNA sequencing analysis in *A. nidulans* revealed that expression levels of predominant proteolytic enzyme (prta), putative tripeptidyl peptidase, putative carboxypeptidase and *pepj* were all significantly decreased in an *mcma* mutant (Li et al. 2016).

5. FlbC

FlbC is a putative nuclear C_2H_2 transcription factor that is involved in the regulation of development in *A. nidulans.* The predicted FlbC-binding sequence has been reported as TGACGAT (Kwon et al. 2010). FlbC is conserved in filamentous Ascomycetes and contains two C_2H_2 zinc fingers at the C-terminus and a putative activation domain at the N-terminus. In *A. oryzae,* production of acid protease in a solid-state culture was markedly decreased by disruption of *flbC*. Northern blot analyses revealed that transcripts of *pepA* were significantly reduced in a strain with disruption of *flbC* (Tanaka et al. 2016).

1.56 Regulation of extracellular proteases in fungi

Fungi can grow under a wide range of environmental conditions and can acquire different compounds as nutrients. Several regulation mechanisms and specific regulatory pathways are involved in nutrient signalling and metabolism. Regulation of proteases also influences levels of carbon, nitrogen, protein and pH in the culture medium. Specially, proteases can lead to variations in the C/N ratio, the presence of easily metabolisable sugars, such as glucose, and metal ions. Extracellular proteases can also impact other physical factors including aeration, inoculum density, pH, temperature and incubation time (de Souza et al. 2015).

- Regulation of acid protease in *Aspergillus* spp.

The extracellular proteases are produced under conditions with reduced carbon,

nitrogen or sulphur (Cohen et al., 1973). Exogenous protein has been shown to can act as an inducer of *Aspergillus* extracellular proteases. The extracellular proteases of *A. nidulans* have been shown to be produced in response to limited levels of carbon, nitrogen or sulphur, even in the absence of exogenous protein. *A. niger* produces high levels of aspartic protease. Two genes encoding acid proteases have been isolated from *A. niger.* The *pepa* gene encodes an aspergillopepsin and the *pepb* gene encodes a unique acid protease that is comprised of two polypeptides. The *pepa* and *pepb* genes are regulated by repression of nitrogen and carbon catabolites, environmental pH and the presence of exogenous protein (Jarai and Buxton, 1994). In *A. nidulans,* two extracellular proteases have been identified, PrtA (a serine protease) and PrtB (an aspartic protease aspergillopepsins), which has greater similarity to the *A. fumigatus pepf* gene product (70.7% amino acid identity). Northern blot analysis suggests that at pH 3, the alkaline protease gene, *prta*, is expressed at higher levels than the acid protease gene, *prtb* (VanKuyk et al., 2000). The levels of *prtb* expression under optimal conditions have not yet been determined. Other studies have shown that *A. oryzae* acid protease (PepA) levels are regulated by temperature. This acid protease is produced at a temperature of less than 35 °C. However, when cultured at conditions greater than 38 °C, there is a decrease in its production. These phenomena suggest that regulation of the production of acid proteases is different from that of amylolytic enzymes (Kitano et al. 2002).Addition of phytic acid also increases acid protease activity in mixed cultures of *A. oryzae* and *A. niger* (Leng and Xu 2011).

The GATA transcription factor, AreA, a positive-acting DNA-binding protein, represses nitrogen metabolite extracellular proteases in *A. nidulans* (Cohen 1973). The transcription factors that mediate carbon and sulphur metabolite repression of the extracellular proteases are unknown. However, mutations in the *crea*, *creb*, *crec*, *xprf* and *xprg* genes lead to increased levels of extracellular protease in *A. nidulans* (Katz et al. 2006; Katz et al. 2008) and a mutation in the *xpre* gene results in a protease-deficient phenotype (Katz et al. 2006). The presence of multiple copies of the CreA binding site in the upstream region of the PrtB gene suggests that it may be subject to CreA control. PacC binding sites were not found in the 1027 bp upstream region of the PrtB initiation codon. This suggests that alkaline repression of acid-expressed genes is not mediated by PacC (Vankuyk et al. 2000).

- Regulation of aspartic/acid protease in other fungi

Regulation of aspartic proteases have been reported in *Candidia albicans,* which is a

commensal fungus that appears on the mucosal surfaces of most healthy people as well as immunocompromised patients. *C. albaicans* contains a family of 10 genes that encode aspartic proteases (Saps-**S**ecreted **a**spartic **p**rotease**s**) that are involved in the virulence of the fungus, and of these, Sap2p is the major aspartic protease. Sap2p expression is induced in media that contains proteins as the sole nitrogen source and is repressed when preferred carbon sources, such as ammonium or amino acids, are available (Banerjee et al. 1991). In yeast species, extracellular amino acids are sensed by the SPS sensor, the sensor in the cell membrane of yeast that is responsible for acquiring nitrogen signals and activates the Stp1p and Stp2p transcription factors. The Stp1-transcription factor in yeast regulates aspartic protease. The activated transcription factors, Stp1p and Stp2p, induce aspartic proteases Sap2 oligopeptide transporters, OPT1 and OPT3 (Martínez and Ljungdahl 2005). As in filamentous fungi, in *S. cerevisiae*, the two GATA factors, Gln3p and Gat1p (Nil1p), also activate the expression of catabolic enzymes and uptake systems of secondary nitrogen sources when the preferred nitrogen sources are absent or not at sufficient levels for normal growth. Aspartic protease expression in *C. albaicans* is also regulated by these Gln3p and Gat1p transcription factors under different levels of nitrogen sources. While Stp2p induces the expression of genes involved in amino acid uptake, Stp1p activates genes that are required for the utilisation of proteins as a nitrogen source, including the secreted aspartic protease SAP2 and the oligopeptide transporters OPT1 and OPT3. The link between general transcription factors Gln3p and Gat1p and the specific regulator Stp1 were necessary for Sap2 expression and positive and negative control with the presence/absence of a preferred nitrogen source and protein.

- Regulation of acid protease in *Trichoderma* spp.

Some studies have also been carried out to elucidate the regulation of aspartic/acid proteases in *Trichoderma s*pp. In *T. harzianum,* the gene coding an aspartyl protease, *papa,* was isolated and characterised (Delgado-Jarana et al. 2002). The promoter sequence contained potential AreA and PacC sites, but no potential CreA sites. Expression of the *papa* gene was regulated by pH, was repressed by ammonium, glucose and glycerol, and was induced by organic nitrogen sources (Kredics et al. 2005).

The Basic Local Alignment Search Tool (BLAST) search was carried out to identify factors in *Trichoderma s*pp. that are homologs to other protease regulators found in other fungi. There no direct homologs in *T. reesei* to the PrtT transcription factor of *A. nidulans*. However, there was a similar homolog of *A. nidulans* in *T. reesei* (protein ID 105255) with

an amino acid identity of 34%. Homologs of the protease activators XprG, McmA and FlbC were found in *A. nidilans* (protein IDs 54675, 42249 and 58011, respectively) that had 47 %, 64 % and 53% amino acid similarity, respectively, in *T. reesei.* There was no homolog of Stp1 in *C. albicans* in *T. reesei.*

1.6 Aims of the study

T. reesei produces a large number of cellulases and is utilised as an industrial cellulase producer. Many studies have focused on determining the mechanisms of cellulase regulation in *T. reesei.* Truncation of cellulases due to proteolytic degradation was observed in the secretome, which is a limitation of the cellulase production system. However, so far, few studies have been performed to assess protease regulation in *T. reesei*.

In order to develop a better understanding of protease production mechanisms in *T. reesei*, this study focused on protease gene regulation and the relationship between cellulase and protease production mechanisms. The aims of this study were to **1)** elucidate the role and production mechanisms of TrAsP in QM9414 (Chapter 2) and 2) determine the gene regulation mechanisms of TrAsP as they relate to the regulation of cellulase (Chapter 3).

1.7 References

Banerjee A, Ganesan K, Datta A (1991) Induction of secretory acid proteinase in *Candida albicans*. J Gen Microbiol 137:2455–2461 . doi: 10.1099/00221287-137-10-2455

Bergthorson JM (2018) Recyclable metal fuels for clean and compact zero-carbon power. Prog Energy Combust Sci 68:169–196 . doi: 10.1016/j.pecs.2018.05.001

Bugeja HE, Hynes MJ, Andrianopoulos A (2012) AreA controls nitrogen source utilisation during both growth programs of the dimorphic fungus *Penicillium marneffei*. Fungal Biol 116:145–154 . doi: 10.1016/j.funbio.2011.10.009

Cabinet Meeting Decision T (2007) BECOMING A LEADING ENVIRONMENTAL NATION IN THE 21 ST CENTURY: JAPAN'S STRATEGY FOR A SUSTAINABLE SOCIETY (Tentative Translation)

Cao Y, Zheng F, Wang L, Zhao G, Chen G, Zhang W, Liu W (2017) Rce1, a novel transcriptional repressor, regulates cellulase gene expression by antagonizing the transactivator Xyr1 in *Trichoderma reesei*. Mol Microbiol 105:65–83 . doi: 10.1111/mmi.13685

Chen L, Zou G, Zhang L, de Vries RP, Yan X, Zhang J, Liu R, Wang C, Qu Y, Zhou Z (2014) The distinctive regulatory roles of PrtT in the cell metabolism of *Penicillium oxalicum*. Fungal Genet Biol 63:42–54 . doi: 10.1016/j.fgb.2013.12.001

Chandel AK, Garlapati VK, Singh AK, Antunes FAF, da Silva SS (2018) The path forward for lignocellulose biorefineries: Bottlenecks, solutions, and perspective on commercialization. Bioresour Technol 264:370–381 . doi: 10.1016/J.BIORTECH.2018.06.004

Cherry JR, Fidantsef AL (2003) Directed evolution of industrial enzymes: An update. Curr Opin Biotechnol 14:438–443 . doi: 10.1016/S0958-1669(03)00099-5.

Cohen BL (1973) Regulation of intracellular and extracellular neutral and alkaline proteases in *Aspergillus nidulans.* J Gen Microbiol 79:311–320 . doi: 10.1099/00221287-79-2-311.

de Souza PM, Bittencourt ML de A, Caprara CC, de Freitas M, de Almeida RPC, Silveira D, Fonseca YM, Ferreira Filho EX, Pessoa Junior A, Magalhães PO, Junior, Magalhães PO (2015) A biotechnology perspective of fungal proteases. Braz J Microbiol 46:337–46 . doi: 10.1590/S1517-838246220140359

Delgado-Jarana J, Rincón AM, Benítez T (2002) Aspartyl protease from *Trichoderma harzianum* CECT 2413: Cloning and characterization. Microbiology 148:1305–1315 .

doi: 10.1099/00221287-148-5-1305

Derntl C, Rassinger A, Srebotnik E, Mach RL, Mach-Aigner AR (2015) Xpp1 regulates the expression of xylanases, but not of cellulases in *Trichoderma reesei*. Biotechnol Biofuels 8:112 . doi: 10.1186/s13068-015-0298-8.

Dienes, D., Borjesson, J., Hagglund, P., Tjerneld, F., Liden, G., Reczey, K., et al. (2007). Identification of a trypsin-like serine protease from *Trichoderma reesei* QM9414. Enzyme Microb. Technol. 40, 1087–1094. doi: 10.1016/j.enzmictec. 2006.08.013

Druzhinina IS, Kubicek CP (2017) Genetic engineering of *Trichoderma reesei* cellulases and their production. doi: 10.1111/1751-7915.12726

Ellison CE, Hall C, Kowbel D, Welch J, Brem RB, Glass NL (2011) Population genomics and local adaptation in wild isolates of a model microbial eukaryote. Proc Natl Acad Sci USA 108: . doi: 10.1073/pnas.1014971108

Filiatrault-Chastel C, Navarro D, Haon M, Grisel S, Herpoël-Gimbert I, Chevret D, Fanuel M, Henrissat B, Heiss-Blanquet S, Margeot A, Berrin JG (2019) AA16, a new lytic polysaccharide monooxygenase family identified in fungal secretomes. Biotechnol Biofuels 12:1–15 . doi: 10.1186/s13068-019-1394-y

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem 278:31988–31997 . doi: 10.1074/jbc.M304750200.

Gerber PJ, Heitmann JA, Joyce TW (1997) Purification and characterization of xylanases from *Trichoderma.* Bioresour Technol 61:127–140 . doi: 10.1016/S0960-8524(97)00052-7

Himmel ME, Ding S-Y, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. Science (80-) 315:804–807 . doi: 10.1126/science.1137016

Hynes MJ (1974) Effects of ammonium, L-glutamate, and L-glutamine on nitrogen catabolism in *Aspergillus nidulans*. J Bacteriol 120:1116–1123

Jarai G, Buxton F (1994) Nitrogen, carbon, and pH regulation of extracellular acidic proteases of *Aspergillus niger*. Curr Genet 26:238–244 . doi: 10.1007/BF00309554

Katz ME, Bernardo SM, Cheetham BF (2008) The interaction of induction, repression and starvation in the regulation of extracellular proteases in *Aspergillus nidulans*: Evidence for a role for CreA in the response to carbon starvation. Curr Genet 54:47–55 . doi:

10.1007/s00294-008-0198-6

Katz ME, Gray KA, Cheetham BF (2006) The *Aspergillus nidulans xprG (phoG)* gene encodes a putative transcriptional activator involved in the response to nutrient limitation. Fungal Genet Biol 43:190–199 . doi: 10.1016/j.fgb.2005.12.001.

Kitano H, Kataoka K, Furukawa K, Hara S (2002) Specific Expression and Temperature-Dependent Expression of the Acid Protease-Encoding Gene (PepA) in *Aspergills oryzae* in Solid-State Culture (Rice-Koji). J Biosci Bioeng 93:563–567.

Koike H, Aerts A, LaButti K, Grigoriev I V., Baker SE (2013) Comparative Genomics Analysis of *Trichoderma reesei* Strains. Ind Biotechnol 9:352–367 . doi: 10.1089/ind.2013.0015

Kredics L, Antal Z, Szekeres A, Hatvani L, Manczinger L, Vágvölgyi C, Nagy E (2005) Extracellular Proteases of *Trichoderma Species*. Acta Microbiol Immunol Hung 52:169–184 . doi: 10.1556/AMicr.52.2005.2.3

Kubicek-Pranz E, Gsur A, Hayn M, Kubicek C (1991) Characterization of commercial Trichoderma reesei cellulase preparations by denaturing electrophoresis (SDS-PAGE) and immunostaining using monoclonal antibodies. Biotechnol Appl Biochem 14:317–323 . doi: 10.1111/j.1470-8744.1991.tb00184.x

Kwon NJ, Garzia A, Espeso EA, Ugalde U, Yu JH (2010) FlbC is a putative nuclear C_2H_2 transcription factor regulating development in *Aspergillus nidulans.* Mol Microbiol 77:1203–1219 . doi: 10.1111/j.1365-2958.2010.07282.x.

Landowski CP, Huuskonen A, Wahl R, Westerholm-Parvinen A, Kanerva A, Hänninen AL, Salovuori N, Penttilä M, Natunen J, Ostermeier C, Helk B, Saarinen J, Saloheimo M (2015) Enabling low cost biopharmaceuticals: A systematic approach to delete proteases from a well-known protein production host *trichoderma reesei*. PLoS One 10:1–28 . doi: 10.1371/journal.pone.0134723.

Lamb HK, Dodds AL, Swatman DR, Cairns E, Hawkins AR (1997) Deletion of the N-terminal region of the AREA protein is correlated with a derepressed phenotype with respect to nitrogen metabolite repression. J Bacteriol 179:6649–6656 . doi: 10.1128/jb.179.21.6649-6656.1997.

Leng Y, Xu Y (2011) Improvement of acid protease production by a mixed culture of *Aspergillus niger* and *Aspergillus oryzae* using solid-state fermentation technique. African J Biotechnol 10:6824–6829

Li N, Kunitake E, Endo Y, Aoyama M, Kanamaru K, Kimura M, Kato M, Kobayashi T (2016) Involvement of an SRF-MADS protein McmA in regulation of extracellular

enzyme production and asexual/sexual development in *Aspergillus nidulans*. Biosci Biotechnol Biochem 80:1820–1828 . doi: 10.1080/09168451.2016.1146074.

Li WC, Huang CH, Chen CL, Chuang YC, Tung SY, Wang TF (2017) *Trichoderma reesei* complete genome sequence, repeat-induced point mutation, and partitioning of CAZyme gene clusters. Biotechnol Biofuels 10: . doi: 10.1186/s13068-017-0825-x

Liu R, Chen L, Jiang Y, Zou G, Zhou Z (2017) A novel transcription factor specifically regulates GH11 xylanase genes in *Trichoderma reesei.* Biotechnol Biofuels 10:194 . doi: 10.1186/s13068-017-0878-x

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiol Mol Biol Rev 66:506–577 . doi: 10.1128/mmbr.66.3.506-577.2002

Maeda H (1996) Role of Microbial Proteases in Pathogenesis. Microbiol Immunol 40:685–699 . doi: 10.1111/j.1348-0421.1996.tb01129.x

Margolles-Clark E, Hayes CK, Harman GE, Penttilä M (1996) Improved production of Trichoderma harzianum endochitinase by expression in *Trichoderma reesei*. Appl Env Microbiol 62:2145–51

Margolles-clark E, Ihnen M, Penttilä M (1997) Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. J Biotechnol 57:167–179 . doi: 10.1016/S0168-1656(97)00097-7

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev I V., Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, De Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barbote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26:553–560 . doi: 10.1038/nbt1403.

Martínez P, Ljungdahl PO (2005) Divergence of Stp1 and Stp2 transcription factors in Candida albicans places virulence factors required for proper nutrient acquisition under amino acid control. Mol Cell Biol 25:9435–46 . doi: 10.1128/MCB.25.21.9435-9446.2005

Morikawa Y (2001) induces cellulase gene transcription in the cellulolytic fungus *Trichoderma reesei* L -Sorbose. 329–334

Nitta M, Furukawa T, Shida Y, Mori K, Kuhara S, Morikawa Y, Ogasawara W (2012) A new Zn(II)2Cys$_6$-type transcription factor BglR regulates ß-glucosidase expression in *Trichoderma reesei*. Fungal Genet Biol 49:388–397 . doi: 10.1016/j.fgb.2012.02.009.

Punt PJ, Schuren FHJ, Lehmbeck J, Christensen T, Hjort C, van den Hondel CAMJJ (2008) Characterization of the *Aspergillus niger* prtT, a unique regulator of extracellular protease encoding genes. Fungal Genet Biol 45:1591–1599 . doi: 10.1016/j.fgb.2008.09.007.

Purvis B, Mao Y, Robinson D Three pillars of sustainability: in search of conceptual origins. doi: 10.1017/s0376892900011449.

Qian Y, Sun Y, Zhong L, Sun N, Sheng Y, Qu Y, Zhong Y (2019a) The GATA-Type Transcriptional Factor Are1 Modulates the Expression of Extracellular Proteases and Cellulases in *Trichoderma reesei*. Int J Mol Sci 20:4100 . doi: 10.3390/ijms20174100.

Qian Y, Zhong L, Sun Y, Sun N, Zhang L, Liu W, Qu Y, Zhong Y (2019b) Enhancement of Cellulase Production in Trichoderma reesei via Disruption of Multiple Protease Genes Identified by Comparative Secretomics. Front Microbiol 10: . doi: 10.3389/fmicb.2019.02784.

Ravagnani A, Gorfinkiel L, Langdon T, Diallinas G, Adjadj E, Demais S, Gorton D, Arst HN, Scazzocchio C (1997) Subtle hydrophobic interactions between the seventh residue of the zinc finger loop and the first base of an HGATAR sequence determine promoter-specific recognition by the *Aspergillus nidulans* GATA factor AreA. EMBO J 16:3974–3986 . doi: 10.1093/emboj/16.13.3974.

REN21 (2019) REN21 - 2019 Global Status Report.

S. Shoemaker, V. Schweickart, M. Ladner, D. Gelfand, S. Kwok KM& MI (1983) Molecular cloning of exo-cellobiohydrolase i derived from *trichoderma reesei* strain l27. Bio/Technology 1:691–696 . doi: 10.1038/nbt1083-691

Saisana M, Philippas D (2012) Sustainable Society Index (SSI): Taking societies' pulse along social, environmental and economic issues

Saloheimo M, Kuja-Panula J, Ylösmäki E, Ward M, Penttilä M (2002) Enzymatic Properties and Intracellular Localization of the Novel *Trichoderma reesei*-Glucosidase BGLII (Cel1A). Appl Environ Microbiol 68:4546–4553 . doi: 10.1128/AEM.68.9.4546-4553.2002

Sharon H, Hagag S, Osherov N (2009) Transcription Factor PrtT Controls Expression of Multiple Secreted Proteases in the Human Pathogenic Mold Aspergillus fumigatus. Infect Immun 77:4051–4060 . doi: 10.1128/IAI.00426-09

Sun J, Tian C, Diamond S, Glass NL (2012) Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa.* Eukaryot Cell 11: . doi: 10.1128/EC.05327-11

Tanaka M, Yoshimura M, Ogawa M, Koyama Y, Shintani T, Gomi K (2016) The C_2H_2-type transcription factor, FlbC, is involved in the transcriptional regulation of *Aspergillus oryzae* glucoamylase and protease genes specifically expressed in solid-state culture. Appl Microbiol Biotechnol 100:5859–5868 . doi: 10.1007/s00253-016-7419-6.

Törrönen A, Mach RL, Messner R, Gonzalez R, Kalkkinen N, Harkki A, Rubicek CP (1992) The two major xylanases from *trichoderma reesei:* Characterization of both enzymes and genes. Bio/Technology 10:1461–1465 . doi: 10.1038/nbt1192-1461

Unfccc (2016) Report of the Conference of the Parties on its twenty-first session, held in Paris from 30 November to 11 December 2015. Part one: Proceedings

Vankuyk PA, Cheetham BF, Katz ME (2000) Analysis of two *aspergillus nidulans* genes encoding extracellular proteases. Fungal Genet Biol 29:201–210 . doi: 10.1006/fgbi.2000.1195

Xiong Y, Sun J, Glass NL (2014) VIB1, a Link between Glucose Signaling and Carbon Catabolite Repression, Is Essential for Plant Cell Wall Degradation by *Neurospora crassa*. PLoS Genet 10:1004500 . doi: 10.1371/journal.pgen.1004500

Chapter 2: Role of Trichodermapepsin in *T. reesei*

2.1 Introduction

The filamentous fungus *Trichoderma reesei* can produce many cellulases and hemicellulases. However, the gene regulation in *Trichoderma reesei* is still under investigation. Degradation of cellulases by proteases has been reported in *Trichoderma reesei*. In order to promote the maximum function and stability of cellulases it is necessary to elucidate the mechanisms of protease production and genes regulation. Although many research groups are focused on the investigation of cellulase production mechanisms in filamentous fungi, including *T. reesei,* few researchers are examining protease production mechanisms.

Understanding the substrate specificity of proteases is important for industrial applications. When catalysis occurs, the active site of the protease fits into the substrate, and some substrates are specific to certain enzymes. The catalytic site of the proteases may contain serine, cysteine, aspartyl, threonine, glutamate or metal. Proteases produced by filamentous fungi have diverse specificity. Research has focused on determining the kinetic parameters, catalytic efficiency and substrate efficiency of several proteases produced by *Fusarium* sp., *Aspergillus*, *Taloromyces*, *Rhizomucor* and *Penicillum* sp., among others (Hamin Neto et al. 2017). However, no previous reports have investigated the substrate specificity of proteases produced by *T. reesei.*

Enzymes also respond to other environmental factors, including pH, nitrogen, temperature and light. It is essential to understand the optimal production conditions of each enzyme in the secretome in order to carry out efficient enzyme production or purification. In addition to development of an optimal enzymatic cocktail, it is also important to understand the natural regulation of the enzyme profile as a response to environmental factors. Many studies have been conducted to analyse the production mechanisms of cellulases, but few studies have focused on other enzymes in the secretome of *T.reesei.*

In this study, we aimed to elucidate the substrate specificity of Trichodermapepsin and its role in cellulase-inducing conditions and TrAsP production conditions in regard to carbon, nitrogen and pH values.

2.2 Materials and Methods

2.21 Strain and media

T. reesei QM9414 strain and QMΔ*trasp* were used in this study. Conidia from these strains were obtained from potato dextrose agar plate cultures. Submerged fermentation medium, containing 0.14% $(NH_4)_2SO_4$, 0.2% KH_2PO_4, 0.03% $CaCl_2 \cdot H_2O$, 0.03% $MgSO_4 \cdot 7H_2O$, 0.05% yeast extract, 0.1% bacto polypeptone, 0.1% Tween 80 and 0.1% trace element solution (0.005 mg/L $FeSO_4$, 0.002 mg/L $MnSO_4 \cdot H_2O$, 0.001 mg/L $CuSO_4 \cdot 5H_2O$, 0.002 mg/L $ZnSO_4 \cdot 7H_2O$, 0.001 mg/L $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 0.002 mg/L $CoCl_2 \cdot 6H_2O$ and 0.001 mg/L $NiCl_2 \cdot 6H_2O$) in a 50 mM Na-tartarate buffer (pH 4.0) was used. The medium was supplemented with 1% of galactose, mannose, xylose, arabinose, fructose, maltose, glucose, sucrose, galactomannan, xylo-oligosaccharide, β-1,3-glucan, birch wood xylan, glucomannan, pectin, beech wood xylan, cellobiose, oat-spelt xylan, glycerol, lactose, Avicel, starch, sorbose, CMC, chitin, lignin and xyloglucan as the carbon sources. Spores (10^7 spores per 50 mL medium) were inoculated and incubated in a shaker at 28 °C at 220 rpm for 6 days. To analyse the response to changes in pH, *T. reesei* was grown on 1% Avicel medium adjusted to pH 3.0, 4.0 or 5.0. Spores (10^7 spores per 50 mL medium) were incubated at 28 °C at 220 rpm for 5 days. After cultivation, culture supernatants were obtained by centrifugation at 14,000 rpm. Samples of the culture supernatants were analysed using 12.5% SDS-polyacrylamide gel electrophoresis (PAGE). The intensity of the bands in the gel was measured using ImageJ software (National Institutes of Health) (Yu et al. 2007). Induction experiments were performed as described previously (Mach-Aigner et al. 2008). The carbon source inducers used were 1% Avicel, 1% galactose and 0.125% α-sophorose. The nitrogen source inducers used were 1% bovine serum albumin (BSA), 1% bacto-peptone or 0.5% ammonium sulphate. Samples were taken at different time intervals. After induction, the mycelia were harvested using filtration and were rapidly frozen in liquid N_2 for RNA extraction. For analysis of the saccharification rate, supernatants of 1% Avicel cultivation of QM9414 and QMΔ*trasp* were used after 3 days of cultivation.

2.22 Construction of QMΔ*trasp*

For construction of the gene deletion plasmid, the fragments, including the upstream region (2.0 kbp) and downstream region (2.3 kbp), of *trasp* were amplified using polymerase

chain reaction (PCR). For PCR, the genomic DNA of the standard strain QM9414 was used as a template and the appropriate primers (Table 2.1) were used. The pU*pyr4* plasmid was digested with ECORV for the removal of *pyr4* and pUC118 was digested with HindIII to generate a linear fragment. The deletion cassette-containing plasmid was constructed using a *pyr4* marker and the pUC118 plasmid using a Gibson Assembly Kit (Fig. 2.1). The *trasp* gene deletion cassette was prepared using HindIII enzyme excision of pU*pyr4*Δ*trasp*. The prepared cassette was introduced into *T. reesei* QMΔ*pyr4* as described previously (Margolles-Clark et al. 1996) with the use of 10 mg/mL Yatalase (Takara Bio) for the preparation of the protoplast. Transformants were selected by a uridine autotroph and purified through single conidia using MMRX and PDAX media. For the confirmation of QMΔ*trasp* candidates, colony PCR was performed using the KOD FX Neo enzyme (TOYOBO). Homologous recombination and copy number of strains were confirmed using Southern hybridisation. Independent transformants were analysed after digestion of the genomic DNA with KpnI and XbaI. After electrophoresis, the digested chromosomal DNA was transferred onto a Hybond N + membrane (GE Health- care Bio Science, Waukesha, WI) and hybridised with an alkaline phosphatase-labelled probe of a 0.5 bp NcoI fragment of the *pyr4* gene.

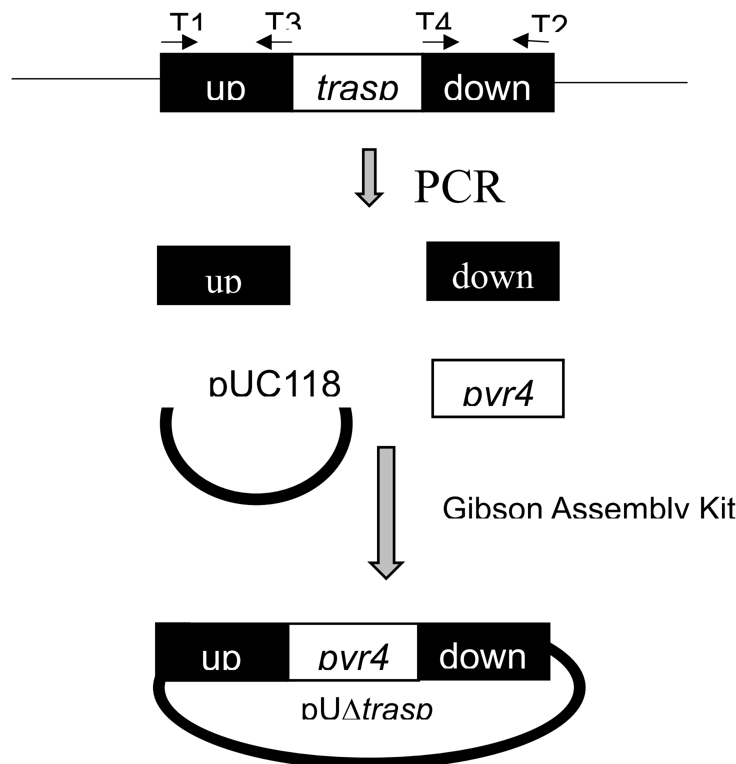


Fig 2.1: Strategy of construction of pUΔ*trasp*

Table 2.1: Primers used for construction of pUΔ*trasp*, red letters indicate the sequence from pUC118 and italic letters indicate the region of enzyme cutting.

Primer Name	Primer sequence (5' → 3')
trasp_clon_fw	GGATCCTCTAGAGTCaagcttACAGAATGACGCCATCATCG
trasp_clon_Rv	GCATGCCTGCAGGTCaagcttCTGGTCCTCCTCGACAAGCT
trasp_inv_fw	GCAGTTGTCGACGATGCGGCGATGGTGGACTTGT
trasp_inv_Rv	GGGAGCAGCGCTGATCTTGAATATCGGAGAAGGTT

2.23 Total RNA extraction and quantitative reverse transcription PCR (qRT-PCR)

Frozen samples were added to 750 µL TES buffer 10 mM Tris–HCl (pH 7.5), 10 mM EDTA, 0.5% SDS and 750 µL acid-phenol preheated to 65 °C. Total RNA extraction was performed using a standard hot acid-phenol RNA extraction protocol. Samples were incubated at 65 °C for 15 minutes, followed by TRIzol LS (Invitrogen, USA) and chloroform extraction. Total RNA was purified using Illustra RNAspin (GE Health Care) as described in the manufacturer's instructions.

qRT-PCR was carried out in a Light Cycler 480 System (Roche Diagnostics). Amplification reactions were per- formed in a final volume of 20 µL using a Light Cycler 480 SYBR Green I Master kit (Roche Diagnostics) with 0.5 mM forward (Fw) primer, 0.5 mM reverse (Rv) primer and 2 µL of 100-fold diluted cDNA. The primers used in the study summarised in Table 2.2.

Thermal cycling was conducted using the following conditions: 5 minutes at 95 °C followed by 45 cycles of 10 seconds at 95 °C, 10 seconds at 60 °C and 10 seconds at 72 °C. All analyses were performed independently in triplicate with a no-amplification control. The specificity of the PCR amplification was documented by a melting curve analysis and *act1* (Protein ID 44504) was used as a reference for quantification.

Table 2.2 Primers used in this study

Gene	
cbh1	Fw (5′-CTTGGCAACGAGTTCTCTT-3′) Rv (5′-TGTTGGTGGGATACTTGCT-3′)
egl1	Fw (5′-CGGCTACAAAAGCTACTACG-3′) Rv (5′-CTGGTACTTGCGGGTGAT-3′)
xyn1	Fw (5′-GGTTGGACGACTGGATCT-3′) Rv (5′-GGTTGT CCTCCATGATGTAG-3′)
xyn2	Fw (5′-CCGTCA ACTGGTCCAACT-3′) Rv (5′-ACACGGAGAGGT AGCTGTT-3′)
trasp	Fw (5′-GCTCAACCGCAA CTCCATC-3′) Rv (5′-GAAGACGACACCCTCCTG CT-3′)
pac1	Fw (5′-ACATCCGTGTTCACGTTC -3′) Rv (5′-GAGTCATCGGCATGAGTC-3′)
act1	Fw (5′-TCCATCATGAAGTGCGAC -3′) Rv (5′-GTAGAAGGAGCAAGAGCAGTG-3′).

2.24 Mass spectrometry analysis

The band corresponding to a 40 kDa protein was digested in the gel using trypsin. Peptides were recovered in the supernatant of the digest, dried by evaporation, reconstituted in 20 µL of H_2O/acetonitrile 97:3 with 0.1% formic acid and analysed by Matrix Assisted Laser Desorption/Ionisation-Time of Flight (MALDI-TOF) mass spectrometry using Spectrometry Axima-CFR plus. From the raw files, the peptide mass fingerprinting (PMF) analysis was assessed against the JGI *T. reesei* QM6a genome database (http://genome.jgi.doe. gov/Trire2/) using Mascot (Matrix Science, London, UK; version 2.3.0) selected for fungi, trypsin digestion and one missed cleavage. Mascot was searched with a fragmented ion mass tolerance of 0.50 Da and a parent ion tolerance of 10.0 ppm.

2.25 Purification of TrAsP

Extracellular TrAsP protease was purified from *T. reesei* QM9414 that was cultivated in 1% galactose media for 72 hours. Culture supernatants were collected using filtration on a miracloth (Calbiochem). Supernatants were precipitated with ammonium sulphate (80% saturation) and were collected by centrifugation at 14,000 rpm for 30 minutes and dissolved in 20 mM sodium acetate buffer (pH 4.0). Solutions were concentrated and buffer-exchanged using Vivaspin 20 (10 kDa) (GE Healthcare) to remove the ammonium sulphate. The solutions were then applied to a HiPrep Q XL 16/10 (GE Healthcare) column, that was previously equilibrated with the same buffer. The fraction containing TrAsP was identified.

2.26 Measurement of proteolytic activity

Total protease activity was measured using a Azocasein assay. A solution of Azocasein was prepared at 5 mg/mL in a 50 mM sodium acetate buffer (pH 5). Azocasein (400 μL) was then mixed with 100 μL of the samples and incubated at 50 °C for 90 minutes. The reaction was halted by the addition of 150 μL of 12% trichloroacetic acid (TCA) and incubated at room temperature for 30 minutes. These samples were then centrifuged at 8000 × *g* for 3 minutes. Subsequently, 150 μL of supernatant was mixed with 150 μL of 1 M NaOH and the absorbance of the solution was measured as 436 nm. One unit of protease activity at OD 436 is equal to 1.0 per mL. Activity of the purified protease was measured by hydrolysis of the fluorogenic peptide MOCAc-Ala-Pro-Ala-Lys-Phe- Phe-Arg-Leu-Lys(Dnp)-NH_2. The optimal pH for protease activity was determined by incubating the purified enzyme at 50 °C for 30 minutes in 0.5 M citrate and 0.5 M NaOH buffer (pH 2–6). To determine the optimal temperature, the enzyme was incubated in 0.5 M citrate and 0.5 M NaOH buffer (pH 5) for 30 minutes from 20–80 °C. Fluorescence was measured in a microplate reader Powerscan HT (DS Pharma Biomedical) at Ex=328 nm and Em=393 nm. Substrate specificity and kinetic parameters of purified TrAsP were measured using the Fluorescence-Quenching Substrate Library FRETS-25 Series, Peptide Institute, Inc (Tanskul et al. 2003). FRESTS-25Xaa substrate contains a highly fluorescent 2-(N-methylamino) benzoyl (Nma) group linked to the side chain of the amino- terminal D-2,3-diamino propionic acid (D-A2Pr) residue, which is efficiently quenched by the 2-4-dinitrophenyl (Dnp) group linked to the amino acid Lys (Fig. 1b). Xaa represents the fixed position for any of the 20 natural amino acids excluding Cys. The Yaa and Zaa positions represent a mixture

of five amino acid residues (P, Y, K, I and D) and (F, A, V, E and R), respectively. FRETS-25Xaa-libraries were prepared with 25 combinations of each Xaa series leading to a total of 475 peptide substrates in 19 separate sections.

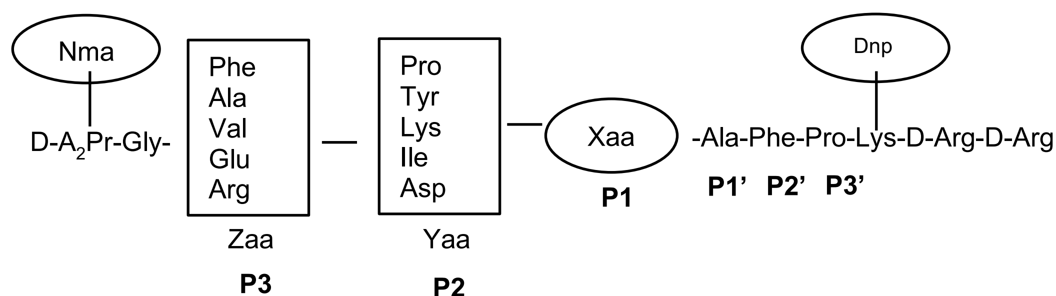


Fig 2.2: FRETS-25Xaa substrate with its substrate of P1, P2, P3, P1', P2', P3'.

Kinetic parameters were determined by fitting the experimental data to the Michaelis Menten equation using Excel Solver (Microsoft) and nonlinear least-square fitting with different concentrations (0.391–25.0 μM) of FRETS-25 as substrates. The enzyme reaction was performed in 50 mM sodium acetate buffer pH 5.0 at 298 K for 5 minutes and the standard deviations were calculated from three independent experiments. The florescence intensity of the released 2-(N-methylamino) benzoyl (Nma) group was measured at an excitation of 340 nm and an emission of 440 using the Infinite 200 PRO microplate reader (Tecan).

2.27 Biomass saccharification

Cellulosic biomass (Erianthus) was subjected to alkaline treatment as previously reported [14]. Saccharification by the cellulases was performed in a 20 ml plastic bottle containing 5% dry biomass (w/v) in 100 mM sodium acetate buffer (pH 5.0), with enzyme loading at 3 mg of protein per gram of dry biomass. The reaction was performed at 50 °C with shaking at 150 rpm for 72 hours. The supernatants were boiled for 5 minutes to inactivate the enzymes, and the sugar produced was measured by the 3, 5-dinitrosalicylic acid (DNS) method. The biomass saccharification ratio (%) was calculated as the ratio of the sugar content to cellulose and hemicellulose in the dry mass.

2.3 Results

2.31 Specificity of proteases secreted by *T. reesei*

In order to determine the protease activity of the *T. reesei* secretome, supernatants of *T. reesei* QM9414 in Avicel and galactose media were used. Azocasein activity, which represents the total proteases in the culture, was measured in after 4 and 6 days of cultivation. The activity was 1.1 ± 0.03 U/mL in 4-day cultivations and 1.4 ± 0.01 U/mL in 6-day cultivations (Fig. 2.3a). Proteinase A activity in the culture was measured using the specific synthetic substrate, MOCAc-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(Dnp)- NH_2. After 4 days of cultivation, the activity was 0. 76 ± 0.21 U/mL, which increased to 1.16 ± 0.07 U/mL after 6 days (Fig. 2.3a). Substrate specificity of protease secretion by *T. reesei* was measured using FRETS-25 series. Primary screening of selection of the favoured Xaa (P1 position) was carried out using *T. reesei* 4-day culture supernatants. The results indicated the proteases in the secretome with preferred residues Phe, Ala, Tyr, Val, Leu, Ile, Met and Glu in the P1 position in descending order (Fig. 2.3b). We next wanted to further characterise main secreted proteases in *T.reesei.*

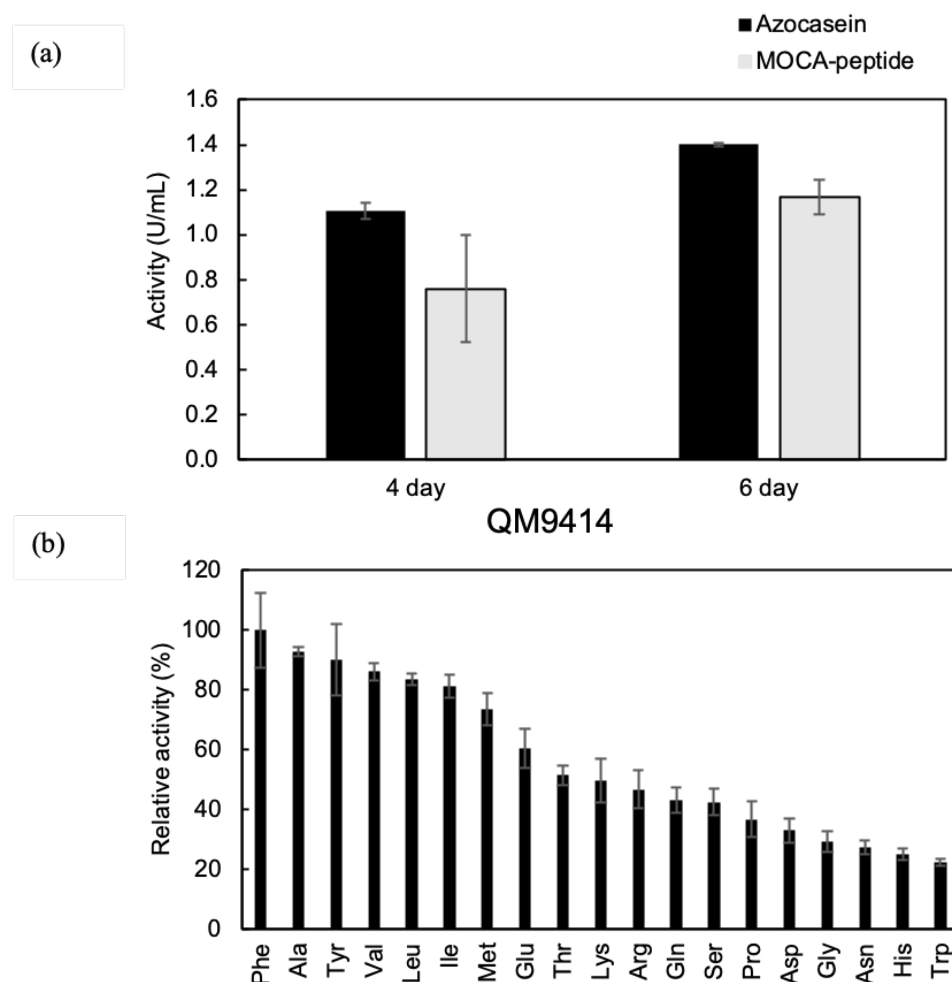


Fig. 2.3 Substrate specificity of the secretome of *T. reesei*

(a) Azocasein activity and proteinase A activity using the fluorogenic peptide MOCAc-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys (Dnp)-NH_2 in 1% Avicel cultivation of *T. reesei* QM9414 after 4 and 6 days. (b) FRETS-25Xaa substrate with its subsites of P1, P2, P3 P1', P2', P3'. (c) Identification of substrate specificity in the *T. reesei* secretome using FRETS-series. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line.

2.32 Identification of TrAsP in the *T. reesei* secretome.

In order to identify new proteins in the *T. reesei* secretome, the extracellular protein expression pattern was assessed using SDS-PAGE from different supernatants with monosaccharides, disaccharides, oligosaccharides, polysaccharides and other carbon sources after 6 days of submerged cultivation (Fig. 2.4a). The different protein expression patterns were compared with Avicel. The same expression pattern as Avicel was obtained for lactose, xylo-oligosaccharide, β-1,3-glucan and xyloglucan. The unknown protein was identified at approximately 40 kDa in all samples. This protein was further analysed by digesting the band with trypsin and analysing the extracted peptides using MALDI-TOF and PMF analysis. The analysis indicated that the peptides were derived from a 40 kDa aspartic protease known as TrAsP (Nascimento et al. 2008). The intensity of this 40 kDa protein band was determined using imageJ analysis (Fig. 2.4b). TrAsP was highly expressed in the *T. reesei* supernatants from cultures grown on monosaccharides, including arabinose, glucose, mannose, galactose, fructose and xylose. Interestingly, its expression was lower when *T. reesei* was grown on well-known inducers of cellulase, including Avicel and L-sorbose.

Total protease activity from supernatants of *T. reesei* grown with the culture conditions mentioned above was measured using non-specific protease substrate, Azocasein. The monosaccharide substrate led to higher protease activity as compared to Avicel (Fig. 2.4c). As TrAsP belongs to the peptidase family AI, its activity was measured by the specific synthetic substrate, MOCAa-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(DNP)-NH_2. Higher proteinase A activity was exhibited in galactose, arabinose, mannose and glucose (Fig. 2.4d). Galactose was the best carbon source for TrAsP as determined via comparison of the band intensity and performing activity analysis. Interestingly Avicel also led to increased proteinase A activity, indicating that there are more proteinase A proteases in the secretome that were not visualised in the SDS-PAGE. The *trasp* expression was also measured using RNA extracted from the mycelia collected from the monosaccharide induction medium. The

highest *trasp* expression was obtained fusing galactose medium, which suggests that galactose is the best inducer for TrAsP.

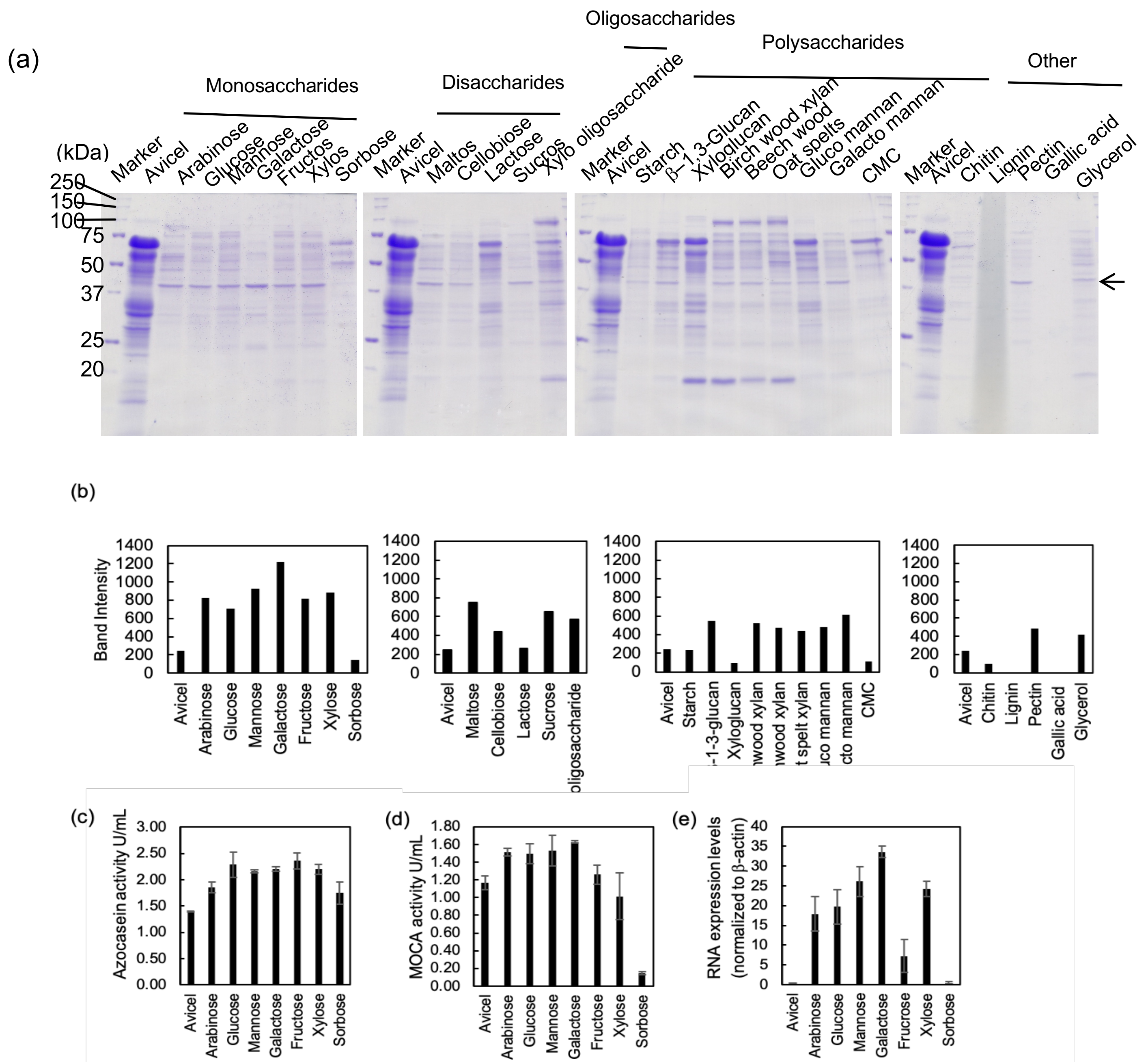


Fig 2.4 *T. reesei* QM9414 was cultivated in medium containing various carbon sources for 6 days. (a) Supernatant (15 µl) was subjected to 12.5% SDS-PAGE. Arrow indicates an approximate 40 kDa band. (b) Bands on SDS-PAGE were quantified using ImageJ software (c) Azocasein activity (d) MOCA activity (e) RNA expression levels

2.33 Construction of *QMΔtrasp*

In order to analyse the role of TrAsP, we constructed a QMΔ*trasp* mutant. Colony PCR was used to confirm the deletion of *trsap* in the constructed strains and southern hybridisation was carried out to confirm homologous recombination (Fig. 2.5).

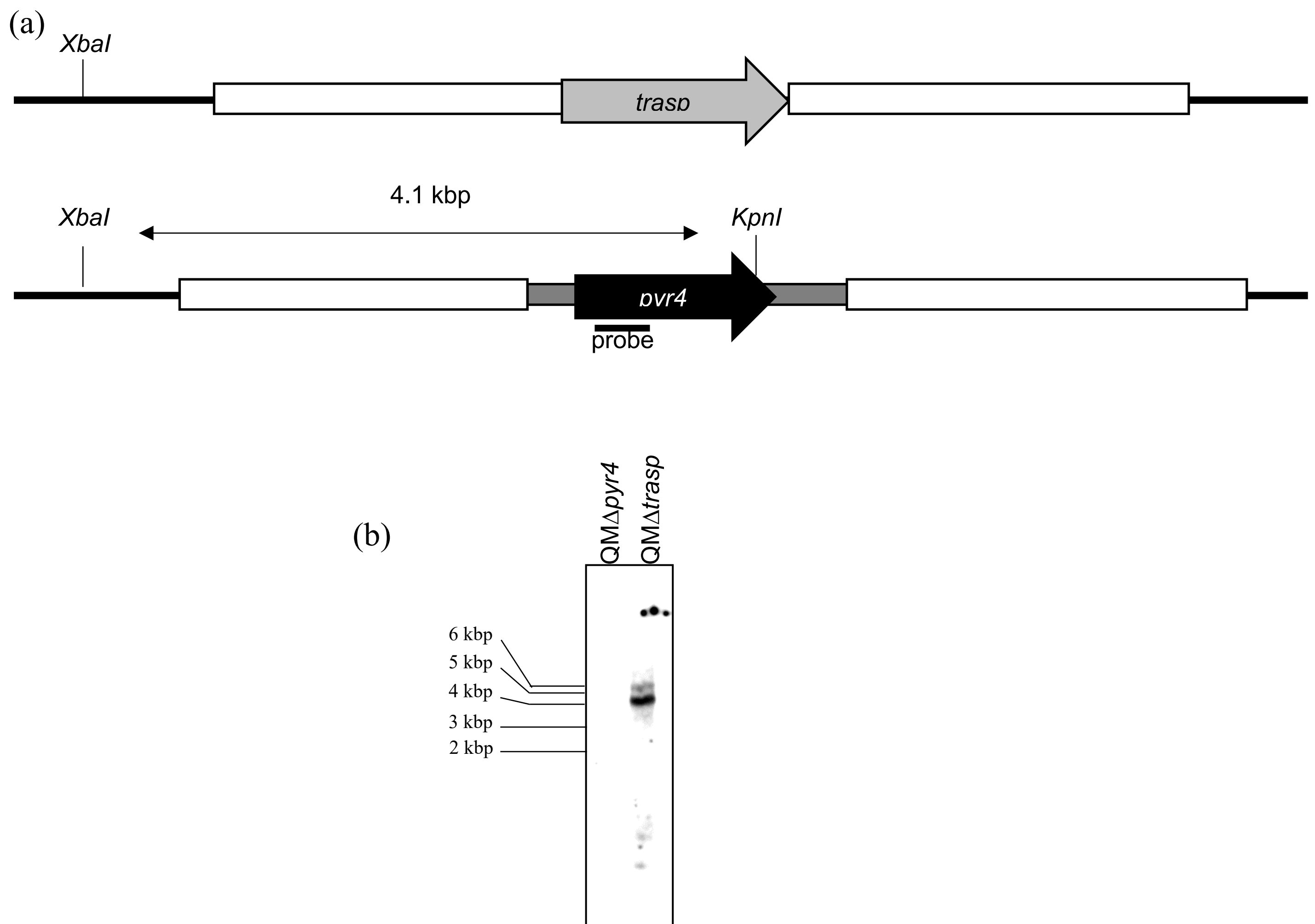


Fig. 2.5 Construction of the QMΔ*trasp* strain (a) Southern hybridisation. Genomic DNA of the strains was digested with KpnI and XbaI and probed with a 0.5 bp *NcoI* fragment of *pyr4* gene. (b) QMΔ*pyr4* showed no band and the insertion of *T. reesei pyr4* gene resulted in the expected hybridisation band in the mutant.

2.34 Specific activity of TrAsP

To determine the effect of *trasp* on the specific activity of the secretome, supernatants of QMΔ*trasp* in Avicel cultivation was used. Azocasein activity was 0.96 ± 0.02 U/mL and MOCA peptidase activity was 0.56 ± 0.23 U/mL after 4 days (Fig. 2.6a). According to these results, the protease activity was reduced as compared to QM9414. The secretome preferred the residues Ala, Leu, Val, Glu, Met, Pro, Lys and Ile in the P1 position (Fig. 2.6b). The preference of the substrate appeared to shift from aromatic (Phe) to aliphatic (Ala) in the deletion strain as compared to QM9414. However, it was difficult to get a clear understanding of TrAsP substrate specificity by increasing or decreasing some of the other proteases in the *trasp* deletant stain. Therefore, we decided to use purified TrAsP analyse the preference of the P1 position by measuring the specific activity.

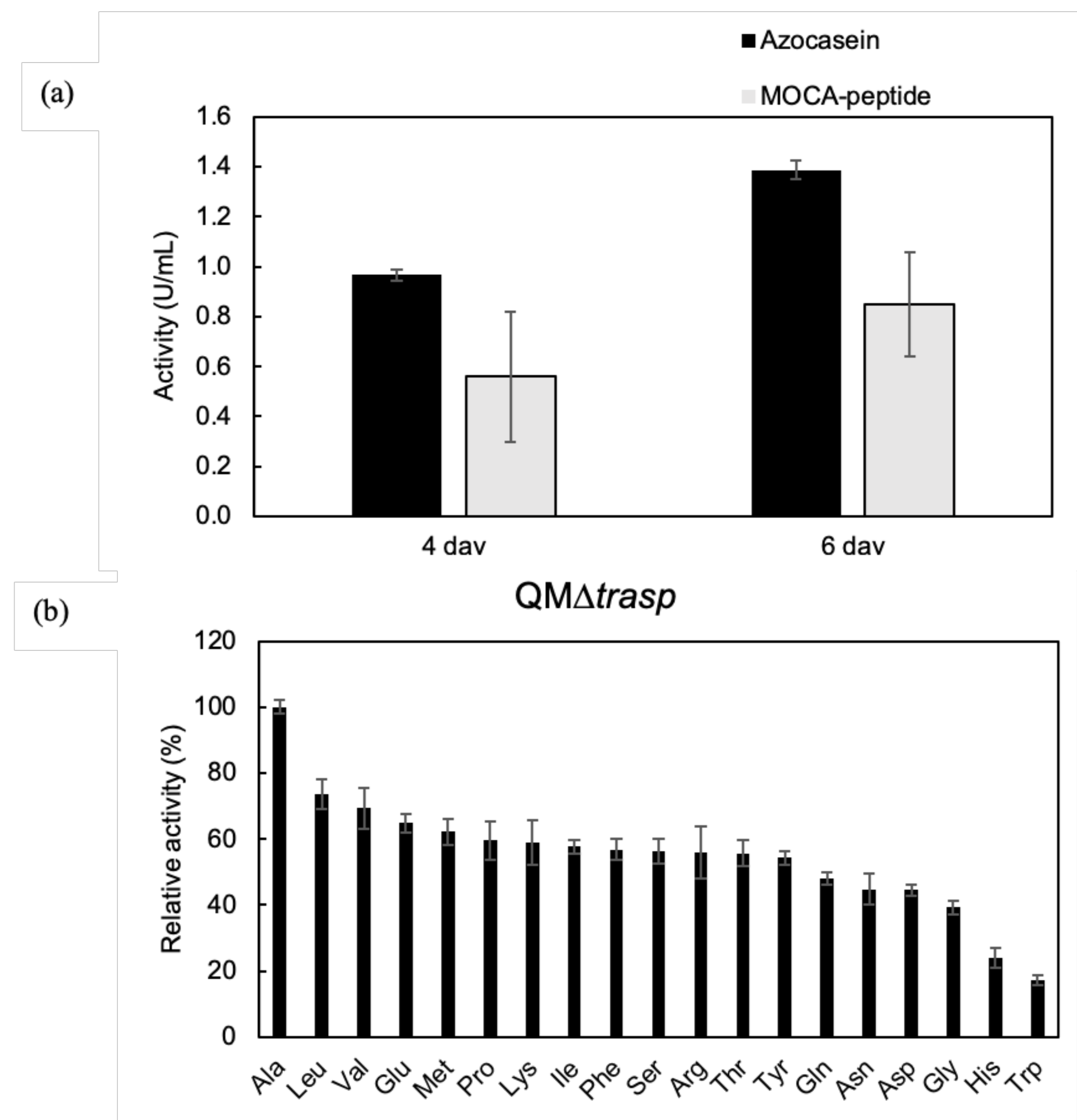


Fig. 2.6 Substrate specificity of the secretome of *T. reesei* QMΔ*trasp*

(a) Azocasein activity and proteinase A activity using fluorogenic peptide MOCAc-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys (Dnp)-NH_2 in 1% Avicel cultivation of *T. reesei* QMΔ*trasp* after 4 and 6 days. (b) Identification of substrate specificity in *T. reesei* secretome by FRETS-Series. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line.

2.35 Purification of TrAsP

TrAsP was purified from the supernatant of *T. reesei* cultivated in 1% galactose medium in order to characterise the protein. The purification steps are summarised in Table 2.3. The total protein amount and total activity were measured in all purification steps. Purified protein had 13% purification fold and 24.1% yield with higher specific activity.

Table 2.3: Summary of purification steps of TrAsP

	Total Protein (mg)	Total activity (U)	Specific activity (U/mg)	Yield (%)	Purification (fold)
Cell free extract	29.7	0.586	0.0197	100	1.00
Ammonium sulphate fractionation	17.0	0.354	0.0208	60.4	1.05
Hiprep Q XL column	0.55	0.141	0.255	24.1	13.0

2.36 Characterisation of TrAsP

The previous studies Haab et al.,1990 found that purified aspartic protease with 42.5 kDa of *T. reesei* had an optimal pH range from 3.0 to 4.0 and Eneysakaya et al.,1999, found that the optimal pH of another aspartic protease (32 kDa) to be 2.8. The *T. harzanium* aspartic protease *papA* gene (85.9% identity with TrAsP) had maximal activity at pH 4.5 (Moreno-Mateos et al., 2007). Experiments were carried out to identify the optimal pH and optimal temperature for the purified protein in this study.

I. Optimal pH

To determine the optimal pH of the purified protein, the enzyme was incubated in 0.5 M citrate and 0.5 M NaOH buffer (pH 2–6) at 50 °C. The protease activity was measured by florescence substrate. The optimal pH was pH 5.0 (Fig. 2.7a).

II. Optimal temperature

To determine optimal temperature, the enzyme was incubated in 0.5 M citrate and 0.5 M NaOH buffer (pH 5.0) for 30 minutes from 20 °C to 80 °C. The optimal temperature for purified TrAsP was 50 °C (Fig. 2.7b).

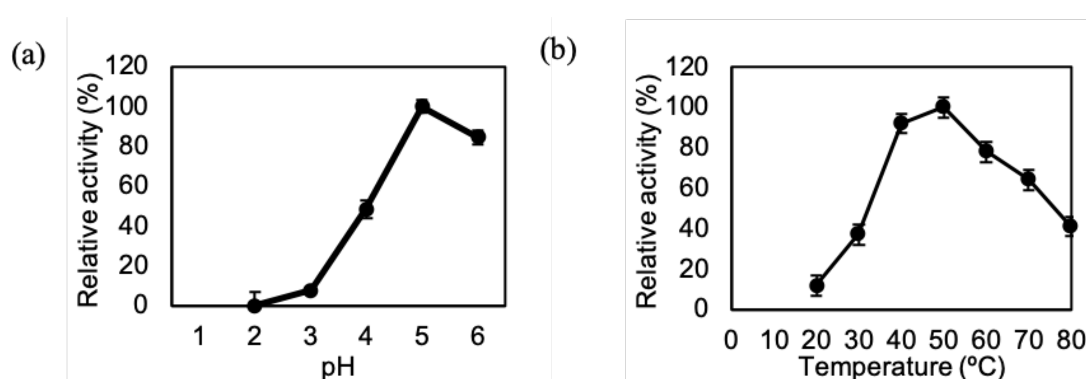


Fig. 2.7 Determination of (a) optimal pH and (b) optimal temperature**.** The data shown is the average of three independent experiments.

2.37 Specific activity of purified TrAsP

In order to determine specific activity of TrAsP, purified TrAsP was analysed using FRETS-25Xaa substrate. Results are graphed with different categories of amino acids (Fig. 2.8). A higher preference was recorded for positively charged (Arg) amino acids and TrAsP had a lower preference for polar uncharged amino acids, such as Asn, Gln, Ser and Thr, and negatively charged Asp. TrAsP showed a significantly low preference for the smallest amino acid, Gly, and the largest amino acid, Trp and a special amino acid Pro with the imino acid group (dihedral angle of − 60°). This suggests that TrAsP is restricted by the size of the amino acid.

In order to further characterise TrAsP, kinetic studies were performed (Table 2.4). The highest k_{cat} value was obtained with Arg (1.22 ± 0.27 s^{-1}). According to k_{cat}/K_m of the protein, higher values were obtained for positively charged amino acids Lys, Arg and His, at 145 ± 12 mM^{-1} s^{-1}, 130 ± 12 mM^{-1} s^{-1} and 108 ± 11 mM^{-1} s^{-1}, respectively. Based on the specific activity and kinetic analysis, we concluded that positively charged amino acids may be more preferable substrates for TrAsP.

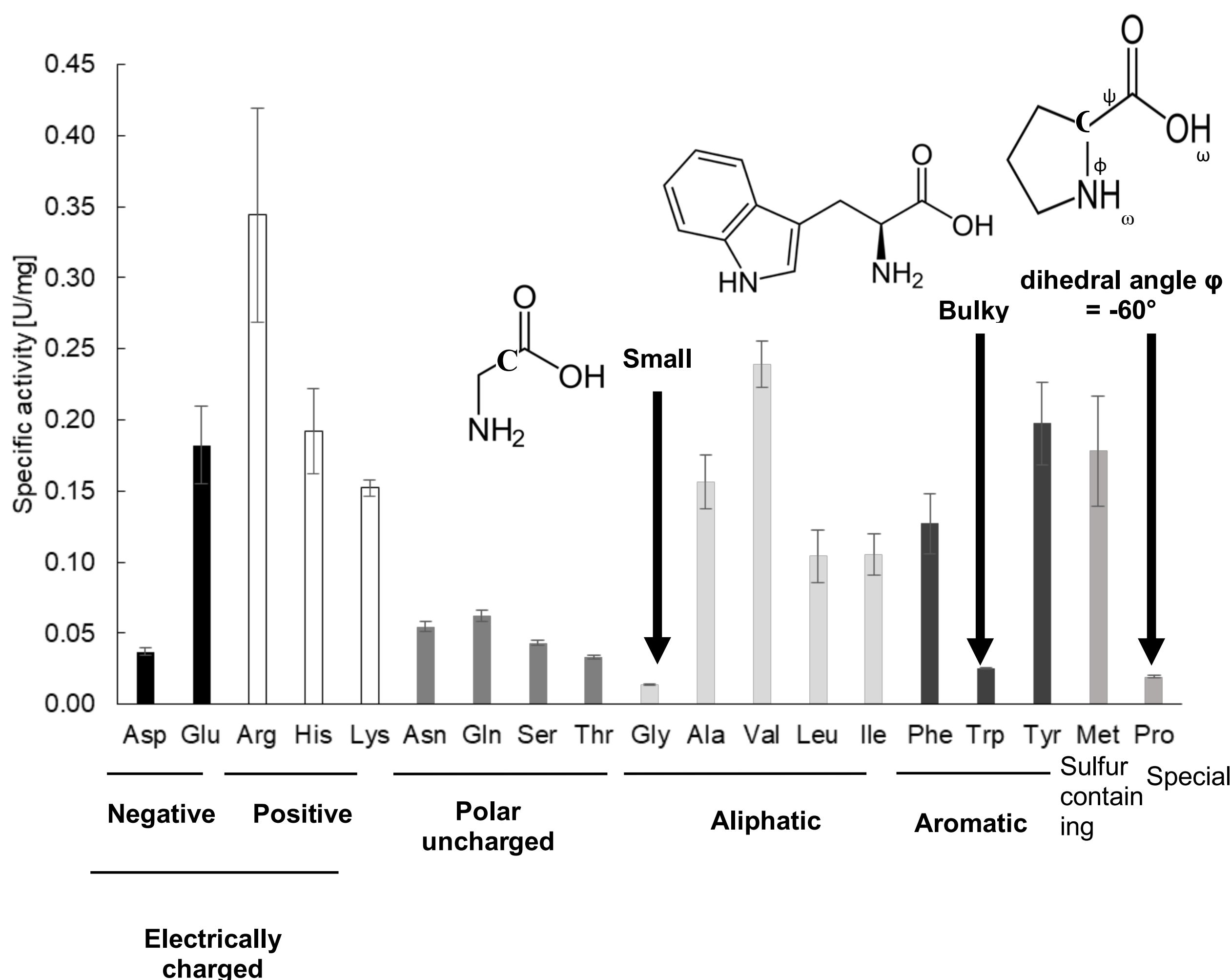


Fig. 2.8 Substrate specificity of purified TrAsP. Specific activity of purified TrAsP was measured using Fluorescence-Quenching Substrate Library FRETS-25 series. Each bar represents the mean of three independent experiments with the SD indicated by a vertical line

Table 2.4 Kinetics values of purified TrAsP

Amino acid	V_{max} [µmol/min•mg]	K_m [µM]	k_{cat} [/sec]	k_{cat}/K_m [/mM•sec]
Asp	0.0368 ±0.0025	19.3±2.26	0.130±0.009	6.85±0.97
Glu	0.182±0.0273	18.3±2.28	0.645±0.096	35.5±5.5
Arg	0.344±0.075	9.60±2.93	1.22±0.27	130±12
His	0.192±0.030	6.43±1.59	0.680±0.105	108±11
Lys	0.152±0.006	3.75±0.48	0.538±0.021	145±12
Asn	0.0545±0.0035	8.53±1.54	0.193±0.012	23.1±2.7
Gln	0.0619±0.0040	5.72±1.30	0.219±0.014	39.6±5.9
Ser	0.0431±0.0015	11.1±0.2	0.152±0.005	13.8±0.4
Thr	0.0332±0.0012	1.17±0.00	0.118±0.004	100±4
Gly	0.0138±0.0005	4.22±0.78	0.049±0.002	11.9±1.8
Ala	0.156±0.019	15.3±2.0	0.553±0.066	36.2±2.9
Val	0.239±0.016	11.7±1.6	0.846±0.057	73.6±9.0
Leu	0.104±0.018	4.21±1.14	0.368±0.065	90.2±10.5
Ile	0.105±0.015	8.46±1.73	0.372±0.052	45.1±7.3
Phe	0.127±0.021	3.87±0.77	0.450±0.074	117±7
Trp	0.0251±0.0003	8.26±0.19	0.089±0.001	10.8±0.3
Tyr	0.198±0.029	10.1±1.3	0.699±0.103	69.1±2.9
Met	0.178±0.039	10.2±2.54	0.631±0.137	62.7±9.0
Pro	0.0194±0.0012	1.63±0.07	0.0685±0.0043	42.3±4.3

2.38 Role of QMΔ*trasp*

In order to elucidate the relationship between *trasp* and cellulases, QM9414 and QMΔ*trasp* were cultivated in 1% Avicel. Supernatants were collected after 2, 4 and 6 days. CMCase activity, Avicelase activity, Xylanase activity and cellobiase activity were measured (Fig. 2.9b-c) and SDS-PAGE was performed (Fig. 2.9a).

CMCase activity of both QM9414 and QMΔ*trasp* increased from 2 to 6 days. A marginal increment was observed with deletion of *trasp* as compared to that of QM9414 (Fig. 2.9b). A

significant increase in Avicelase activity was observed in QMΔ*trasp* as compared to that of QM9414 (Fig. 2.9c).

Higher xylanase activity in 4-day cultures was observed with deletion of *trasp*. However, no significant differences were observed between the strains after 6 days of cultivation (Fig. 2.9d). There was no effect in cellobiase activity with cultivation of QMΔ*trasp* (Fig. 2.9e).

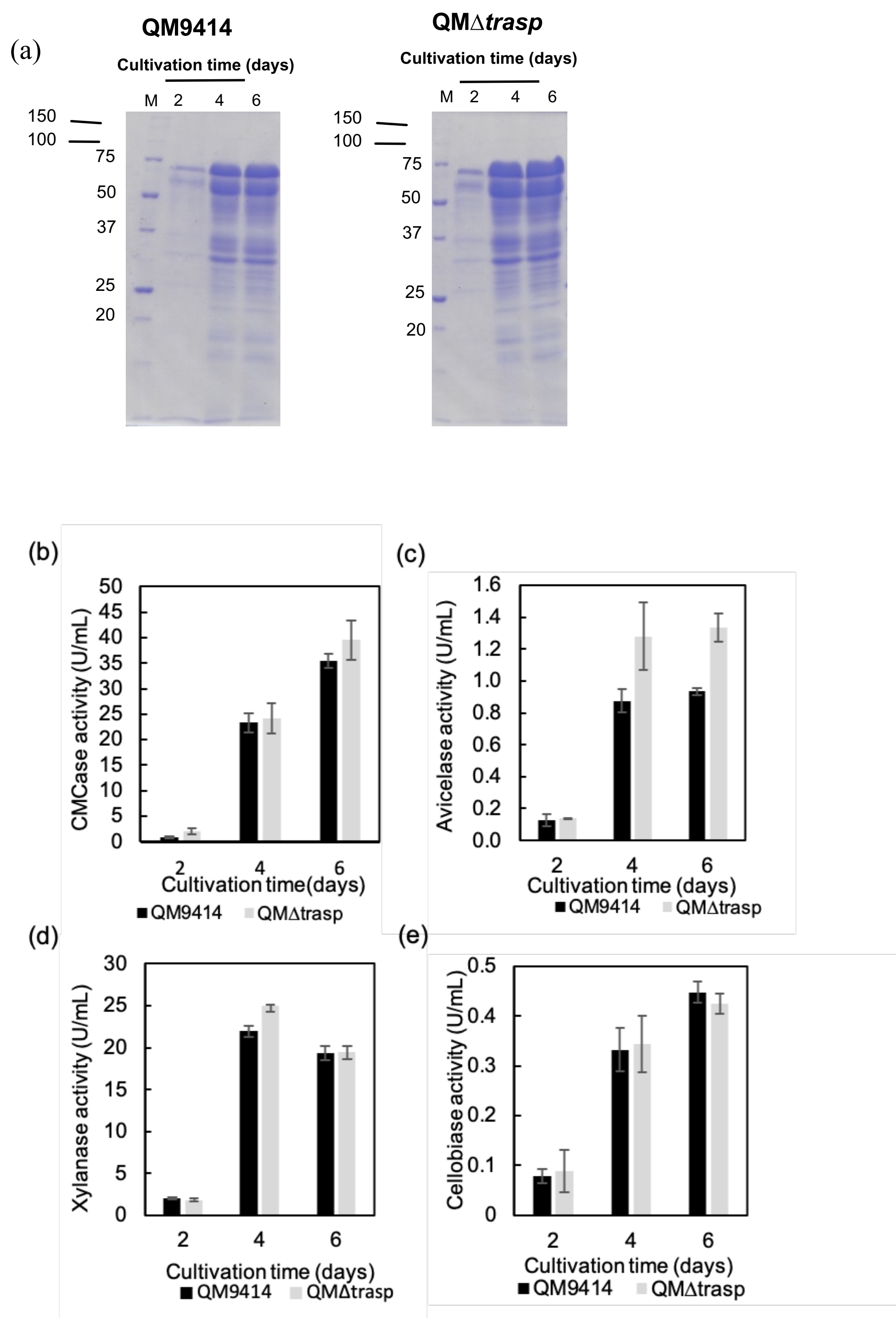


Fig. 2.9 Analysis of cellulase activity in *T. reesei* QM9414 and QMΔ*trasp* after cultivation in 1% Avicel. Samples were collected after 2, 4 and 6 days (a) CMCase activity. (b) Avicelase activity. (e) Xylanase activity. (d) Cellobiase activity. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line.

2.3.9 Stability of cellulases in Saccharification conditions

In order to analyse the effect of *trasp* on cellulase stability, QM9414 and QMΔ*trasp* were cultivated in 1% Avicel and the superannuants were collected after 4 days. Supernatants were incubated at 50 °C and samples were collected after 4 and 6 days. SDS-PAGE and activity analysis were carried out. Higher CMcase, Avicelase and xylanase activity was observed in QMΔt*rasp* (Fig. 2.10).

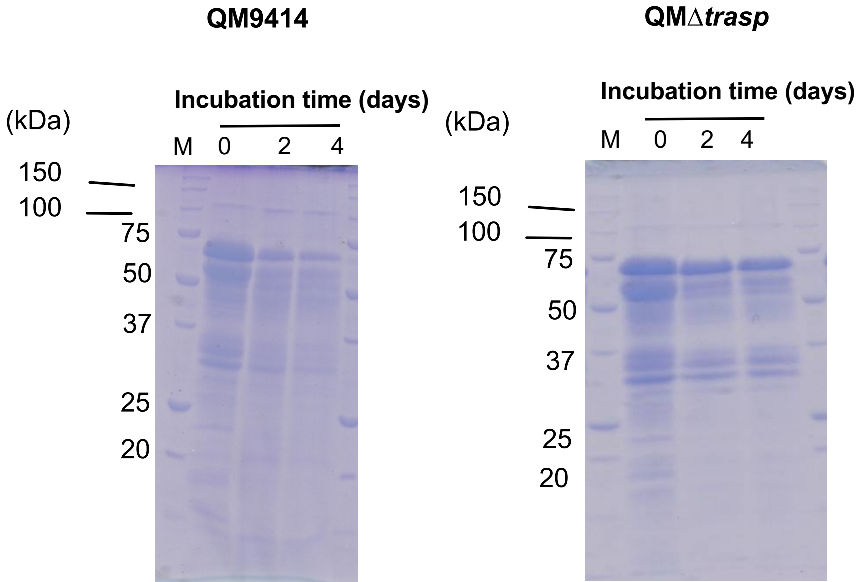
2.3.10 Saccharification

In order to analyse saccharification activity of cellulases with QMΔ*trasp* as compared to QM9414, the saccharification rate was measured. NaOH–pretreated *Erianthus* was used as pretreated biomass, which has previously recorded to contain a higher cellulose and hemicellulose composition. Similar results in cellulose and hemicellulose composition were obtained after pretreatment, which were 50.1% and 25.2%, respectively (Kawai et al. 2013). Glucose and xylose yields were slightly higher for QMΔ*trasp* as compared to that of QM9414 (Table 2.5). A significant change was not observed in the saccharification rate of QMΔtrasp as compared to QM9414. This may be due to deletion of *trasp* leading to no effect on the cellobiase activity. BGL activity is necessary for a higher saccharification rate. However, our results showed that there was a marginal effect of *trasp* on the saccharification rate.

Table 2.5 Saccharification yield (%) by enzymes derived from QM9414 and QMΔ*trasp* using Erianthus biomass pretreated with 3 mg NaOH for 72 hours.

Strain	Saccharification rate (%)	
	Glucose yield	Xylose yield
QM9414	40.7 (±1.37)	36.8 (±1.40)
QMΔ*trasp*	41.8 (±1.5)	37.6 (±1.5)

(a)



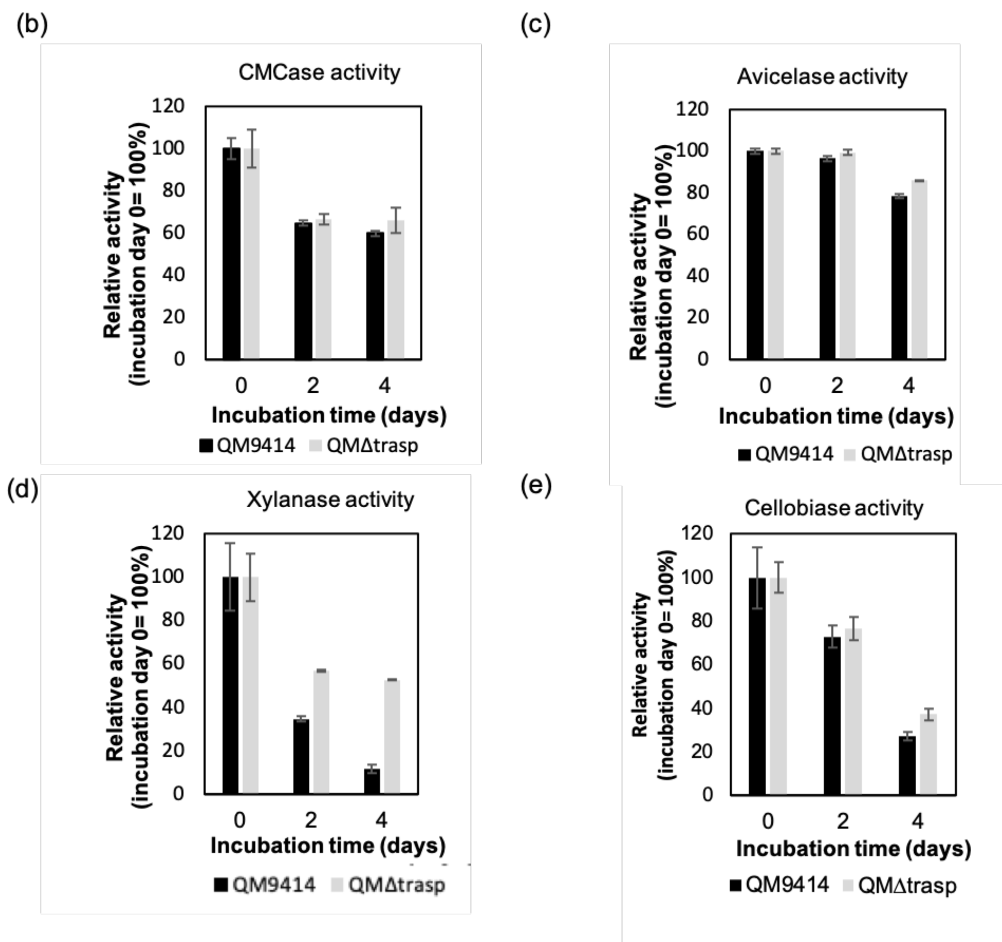


Fig. 2.10 Analysis of cellulase stability in *T. reesei* QM9414 and QMΔ*trasp* after cultivation in 1% Avicel. Samples were collected after 4 and 6 days (a) CMCase activity. (b) Avicelase activity. (e) Xylanase activity. (d) Cellobiase activity. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line.

2.3.11 *trasp* production in response to environmental sources

I. Carbon sources

To analyse *trasp* gene expression as a function of *T. reesei* growth on different carbon sources, *T. reesei* was induced in α-sophorose, Avicel, galactose and mixed media with α-sophorose and galactose. *trasp* expression was compared to that of cellulase and hemicellulase genes for each carbon source. To avoid the effects caused by differential germination on different substrates, the experiments were conducted with glucose- grown mycelia inoculated in the culture. Previous studies have shown that *T. reesei* produces

cellulase when it is grown on cellulose, α-sophorose, L-sorbose and lactose (Mandels and Reese 1957; Mandels et al. 1962; Sternberg and Mandels 1979; Kawamori et al. 1985). In agreement with these previous reports, expression of *cbh1* and *egl1* was significantly higher in *T. reesei* induced on crystalline cellulose Avicel or α-sophorose than on galactose (Fig. 2.11). Minimal expression of *xyn1* was detected with induction on Avicel or α-sophorose. However, compared to α-sophorose, Avicel showed significant *trasp* expression during late induction (Fig. 2.11b). During induction on galactose, *trasp* was highly expressed after 8 and 16 hours, while the expression levels of glycoside hydrolase-encoding genes were quite low. The expression of *trasp* from *T. reesei* cultivated on galactose (at 16 hours) was 106-fold higher than on Avicel and 516-fold higher than on α-sophorose (Fig. 2.11c). In addition, the expression of *trasp* was very low in samples taken at an early induction time during *T. reesei* induction on Avicel or α-sophorose. With the addition of 0.125% α-sophorose to the 1% galactose induction medium, *trasp* expression was significantly reduced and exhibited lower *cbh1* expression (Fig. 2.11d). These findings suggest that there are different regulatory systems for cellulases and proteases when *T. reesei* is induced via a mixed carbon source.

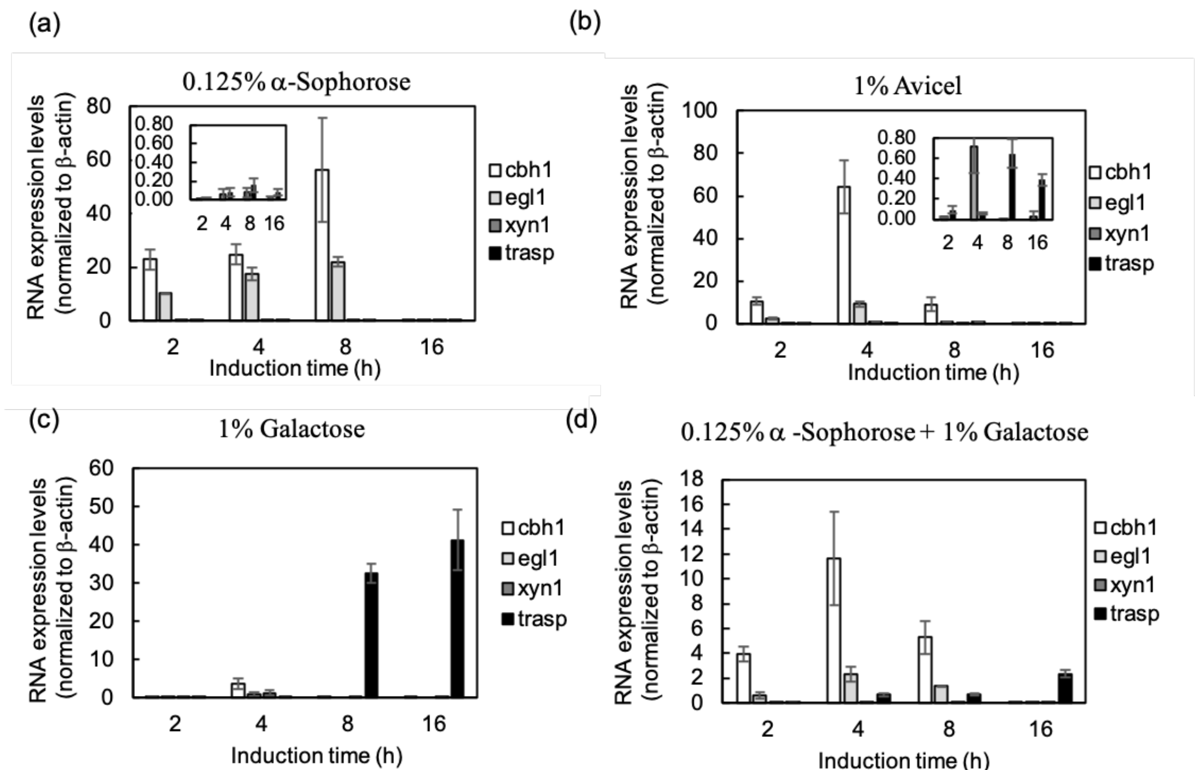


Fig. 2.11 Analysis of the transcription of genes encoding TrAsP and glycoside hydrolases, *cbh1*, *egl1* and *xyn1*after carbon induction. *T. reesei* QM9414 was precultured on glucose and transferred to a medium containing either (a) 0.125% (w/v) α-sophorose, (b) 1% (w/v) Avicel, (c) 1% (w/v) galactose, (d) or a mix of 0.125% (w/v) α-sophorose and 1% (w/v) galactose. All values are normalised to β-actin under the same conditions. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line.

II. Nitrogen sources

In addition to carbon sources, nitrogen sources play an important role in the regulation of proteases. Therefore, it is essential to understand *trasp* expression patterns under different nitrogen sources. To accomplish this, *T. reesei* was induced with 1% BSA, 1% peptone or 0.5% ammonium sulphate. The expression of *cbh1*, *egl1* and *xyn2* was extremely low under these conditions (data not shown). In contrast, *trasp* expression was induced in *T. reesei* by BSA and peptone. While the expression levels of *trasp* with BSA was higher than that with peptone, no induction was observed when *T. reesei* was induced with ammonium sulphate (Fig. 2.12a). The expression of *trasp* was also measured in *T. reesei* induced with a combination of 1% galactose and all of the individual nitrogen sources tested (Fig. 2.12b). We found that the expression of *trasp* was much lower upon addition of 1% galactose, as compared to its expression levels with the individual nitrogen sources alone (Fig. 2.12b).

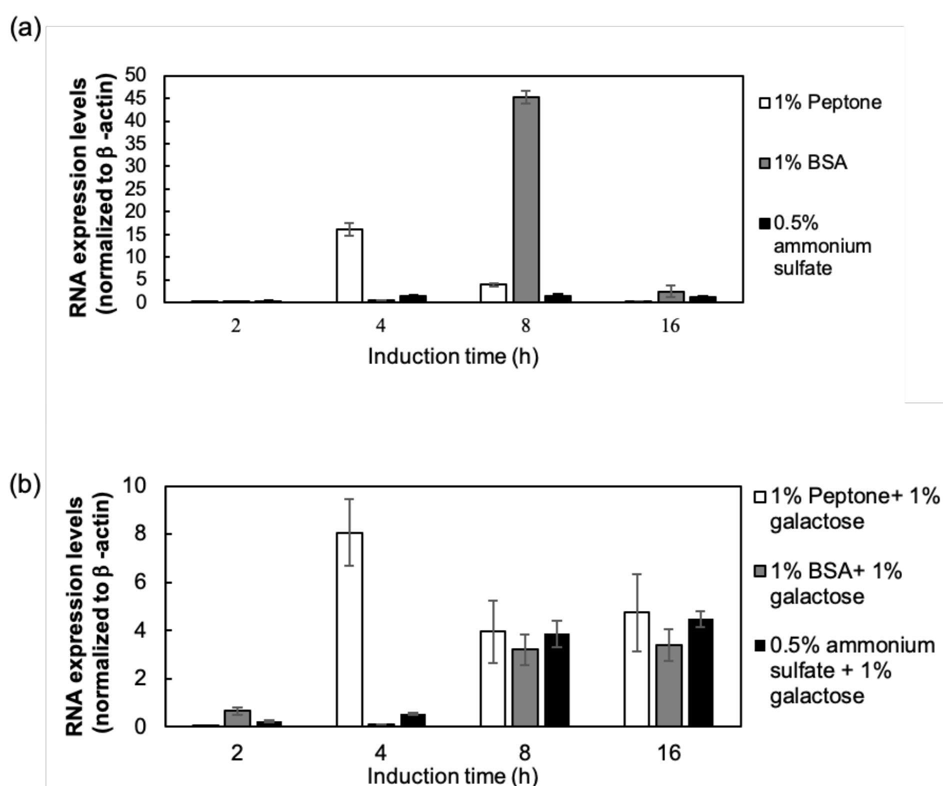


Fig. 2.12 Transcriptional analysis of TrAsP expression under different nitrogen conditions. (a) *T. reesei* QM9414 was precultured on glucose and transferred to a medium containing 1% (w/v) peptone, 1% (w/v) BSA and 0.5% (w/v) ammonium sulphate and was incubated for 2, 4, 8 and 16 hours. (b) T. *reesei* QM9414 was precultured on glucose and transferred to a medium containing 1% (w/v) peptone+1% (w/v) galactose, 1% (w/v) BSA+1% (w/v) galactose or 0.5% (w/v) ammonium sulphate+1% (w/v) galactose and was incubated for 2, 4, 8 and 16 hours. All values were normalised to expression of β-actin under the same conditions. Each bar shows the mean of three independent experiments with the SD indicated by a vertical line

III. pH

In order to analyse the effect of pH in the medium, RNA expression analysis was performed with different pH values. RNA extracted from *T. reesei* grown on galactose medium adjusted to pH 3, 4 or 5 was analysed using qRT-PCR (Fig. 2.13). These pH values were selected as TrAsP is an acidic protease. The relative expression levels of *trasp* were high at pH 3 and 4, with 16.5-fold and 21-fold higher expression, respectively, as compared to low *trasp* expression at pH 5 (after 12 hours). These results indicate that *trasp* expression is favoured under acidic conditions in a galactose medium.

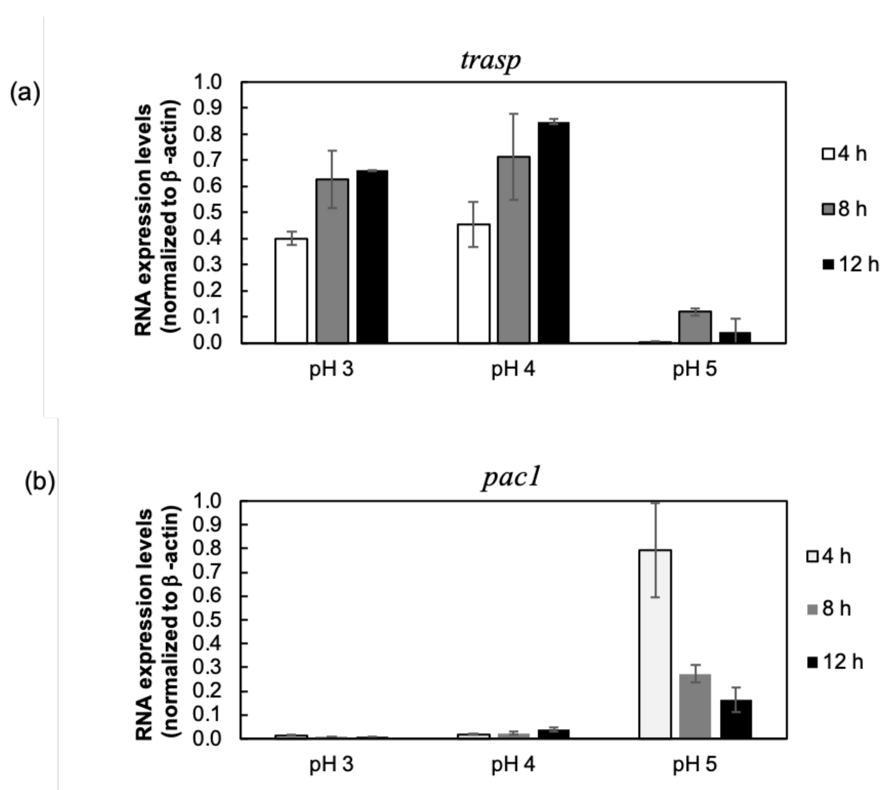


Fig. 2.13 Transcriptional analysis of the pH-dependent regulation of *trasp* and pH regulator *pac1*. *T. reesei* QM9414 was precultured on glucose and transferred to a medium containing 0.1 % galactose with the pH adjusted to 3, 4 or 5 for 5 days. Transcriptional analysis of (a) *trasp* and (b) *pac1* was performed using real-time PCR. All values were normalised to β-actin expression under the same conditions. Each bar shows the mean of three independent experiments, with the SD indicated by a vertical line.

2.4 Discussion

Characterisation of proteases and their function in filamentous fungi is still under investigation. Generally, proteases provide the amino acids necessary to fungi and contribute to fungal survival. In this study, we aimed to identify the role of proteases. In cellulytic filamentous fungi, the stability of cellulases and the levels of secreted cellulases are important for their use as industrial cellulase producers. In this study we elucidated the function of the main protease in *T. reesei*, TrAsP. We investigated the related secreted cellulases and the characteristics of TrAsP.

To this end, we developed a *trasp* deletion strain using homologous recombination and were able to purify TrAsP with a higher specific activity. Cellulase activities were analysed using QMΔ*trasp* and were compared to the wild type QM9414. We confirmed TrAsPs effects on cellulases and xylanases (Fig. 2.9). Using SDS-PAGE, we found that the bands of QMΔtrasp were thicker than those of QM9414. We did not observe effects on BGL with deletion of the *trasp* gene.

We observed an effect of TrAsP on cellulases in Avicel incubated at 50 °C (Fig. 2.10), which is similar to the conditions used for industrial saccharification. Deletion of *trasp* increased the levels of cellulases, which could provide great benefit in commercial applications. However, the saccharification yield of the biomass was not significant (Table 2.5), perhaps because the absence of *trasp* did not affect BGL activity. Alternatively, other proteases with similar functions as *trasp* may compensate for its deletion.

In this study we analysed the substrate specificity of TrAsP as this information is essential to determine the optimal substrates and reaction conditions to improve future biotechnological applications. Various substrates have been used previously for the analysis of substrate specify for several peptidases secreted by filamentous fungi (Hamin Neto et al. 2017). However, little is known about the proteases of *T. reesei.* In this study, we measured the substrate preference of both *T. reesei* QM9414 and QMΔ*trasp,* as the secretome contains a mixture of proteases. Measurement of the substrate preferences for both strains assisted in

the assessment of possible of protein degradation in the medium. It also allowed us to understand how gene deletion affects substrate specificity. The substrate specificity of the P1 position of *T. reesei* secretome was Phe, followed by Ala and Tyr, under cellulase-inducing conditions (Fig. 2.3). We did not obtain the expected results in substrate preference for QMΔ*trasp* (Fig. 2.6) as compared to that of QM9414. This result could be due to the deletion of a gene affecting the composition or by the other proteases in the secretome. Therefore, we used the purified TrAsP protein to measure the specific activity and kinetic parameters. We determined that basic amino acids, lysine and arginine, were more preferable substrates of TrAsP (Fig. 2.8 and Table 2.4). Similar results were reported for an aspartic peptidase from the filamentous fungi *Rhizomucor miehei* (da Silva et al. 2016) and *Phanerochaete chrysosporiumn* (da Silva et al. 2017). The Lys-containing substrates were hydrolysed due to the conserved aspartic acid residue in the flap of the crystal structure (Asp 77 in TrAsP) (Kamitori et al. 2003; Nascimento et al. 2008). This is also explained by the preference of basic amino acids in the P1 position. Additionally, we found that the P1 position was restricted by size. Consistent with our findings, da Silva et al. reported little or no hydrolysis when nonpolar residues, including Gly, Pro and Trp, were at the P1 position but or other subsites (da Silva et al. 2017). As this study is only preliminary, more work is needed to identify other subsites. However, this is the first study focused on elucidating the amino acid substrate specificity of a protease in *T. reesei.* Recent enzyme specificity studies have focused on developing a computational design of remodelling (Li et al. 2013) and on modifying biochemical properties using recombinant DNA techniques (Ahmed 2018). These advances can be extended to the study of the proteases of filamentous fungi, as they have great biotechnological potential and their productivity has room for improvement.

The analysis of protease production mechanisms in *T. reesei* has not been well studied. In this work, we analysed the mechanisms underlying TrAsP production in conditions with various carbon and nitrogen sources and different pH. These results are important for predicting cellulase stability under similar conditions. TrAsP was expressed when *T. reesei* was grown on various carbon sources (Fig. 2.4) and was most highly expressed on monosaccharide-containing medium, except sorbose. TrAsP expression was low when cellulase inducers were present. Under these conditions, we were able to identify galactose as the carbon source that best induced *trasp* expression. In contrast, *T. reesei* induced on Avicel medium led to minimal TrAsP expression. With respect to protease activity, cultivation on Avicel produced proteinase A activity, which may indicate that many other proteinase A isoforms exist in the culture medium. When analysing the *T. reesei* genome, we identified 12

extracellular peptidase A pepsins, besides TrAsP, including JGI protein IDs: 53961, 68662, 81004, 121306, 79807, 105781, 111818, 76887, 77579, 119876, 122076 and 108686. The expression of proteases in response to cellulosic substrates, including Avicel, has been demonstrated in *T. reesei* RUT C-30 (Peciulyte et al. 2014). Peptidase A proteins were found in late cultivations on Avicel, similar to those observed in our results.

This is the first study that highlights the effect of carbon sources on the production of TrAsP, as other reports have mainly focused on the effects of different sources of nitrogen and protein. It is worth noting that *trasp* expression was only induced by galactose after a lengthy period of induction (Fig. 2.11). In *T. reesei*, galactose is reported to induce cellulases gene expression at low growth rates (Karaffa et al. 2006). This fact also supports the delay in TrAsP expression when *T. reesei* was grown on galactose medium, as TrAsP is not well expressed when cellulase inducers are present in the medium. We hypothesised that the true inducer of expression of TrAsP may be a galactose metabolite. A recent study on the optimisation of protease production of some filamentous fungi has recorded that lactose resulted in the highest levels of protease production among the carbon sources tested (Muthukrishnan and Mukilarasi 2016). Several studies have been performed to solve the mechanisms underlying cellulase expression in lactose medium that has been extracellularly hydrolysed to glucose and galactose by ß-galactosidase. These studies found that neither glucose nor galactose led to cellulase expression (Rauscher et al. 2006; Mojzita et al. 2012). Galactose can either be phosphorylated and enter into the Leloir pathway or be reduced to galactitol. Further investigation is necessary to determine the true factor that induces expression of TrAsP following galactose induction in *T.reesei*.

Many fungi produce proteases in the presence of a nitrogen source, and TrAsP expression was also induced by organic nitrogen in *T. reesei*, but was not induced by ammonium sulphate (Fig. 2.12). This result is consistent with previous studies that have shown that peptone strongly stimulates protease formation in *T. reesei* (Haab et al. 1990; Peciulyte et al. 2014). However, expression of TrAsp was lower in response to the combination of galactose and organic nitrogen than it was to galactose or organic nitrogen alone. These results suggest that the mechanisms of induction differ when inducers are mixed, and that these mechanisms of induction may compete given that they act on different regulatory pathways. Interestingly, sequence analysis of the promoter region of *trasp* identified the putative binding sites (GATA) of the nitrogen metabolite regulator AreA (data not shown). These findings indicate that TrAsP expression is influenced by AreA. Under

nitrogen-limited conditions, AreA activates the transcription of many structural gene-encoding enzymes used for nitrogen catabolism. However, there are very few reports on the involvement of AreA in both cellulase and protease production in filamentous fungi.

A number of proteases are induced in different fungal species in response to pH changes (Gente et al. 2001; Delgado-Jarana et al. 2002). In *A. niger*, the pH and the availability of organic nitrogen influence protease production (Margolles-Clark et al. 1996; Braaksma et al. 2009). The induction of *T. reesei* mycelia at pH 3 or 4 resulted in higher expression of TrAsP. *T. reesei* also produces an acidic protease at pH 3 or 4 during the late growth stages (Haab et al. 1990; Margolles-Clark et al. 1996). In contrast, expression of TrAsP was lower at pH 5. These results indicate that TrAsP expression is regulated by the pH of the galactose medium. Interestingly, we found that purified TrAsP has a higher activity at pH 5 (Fig. 2.13). A similar phenomenon was also observed in *T. harzanium* with its aspartic protease, *papA* (accession number: AJ276388), which has 85.9% identity with TrAsP. papA had maximal expression at pH 3.7 and maximal activity at pH 4.5 (Moreno-Mateos et al. 2007). The Pac1 transcription factor responds to extracellular pH, activates genes that are expressed at an alkaline pH and represses genes that are expressed under acidic growth conditions. The putative binding sequence was identified in the promoter region of TrAsP, revealing that Pac1 may regulate *trasp*. Pac1 expression at pH 3 and 4 was extremely low, but was much higher at pH 5 (Fig. 2.13), indicating that Pac1 could repress the expression of *trasp*.

2.5 References

Ahmed ME (2018) Extraction and purification of protease from *Aspergillus niger* isolation. Pharm Pharmacol Int J 6:96–99 . doi: 10.15406/ppij.2018.06.00162

Braaksma M, Smilde AK, van der Werf MJ, Punt PJ (2009) The effect of environmental conditions on extracellular protease activity in controlled fermentations of *Aspergillus niger*. Microbiology 155:3430–3439 . doi: 10.1099/mic.0.031062-0

da Silva RR, de Oliveira LCG, Juliano MA, Juliano L, de Oliveira AHC, Rosa JC, Cabral H (2017) Biochemical and milk-clotting properties and mapping of catalytic subsites of an extracellular aspartic peptidase from basidiomycete fungus *Phanerochaete chrysosporium*. Food Chem 225:45–54 . doi: 10.1016/j.foodchem.2017.01.009

da Silva RR, Souto TB, de Oliveira TB, de Oliveira LCG, Karcher D, Juliano MA, Juliano L, de Oliveira AHC, Rodrigues A, Rosa JC, Cabral H (2016) Evaluation of the catalytic specificity, biochemical properties, and milk clotting abilities of an aspartic peptidase from *Rhizomucor miehei*. J Ind Microbiol Biotechnol 43:1059–1069 . doi: 10.1007/s10295-016-1780-4

Delgado-Jarana J, Rincón AM, Benítez T (2002) Aspartyl protease from *Trichoderma harzianum* CECT 2413: Cloning and characterization. Microbiology 148:1305–1315 . doi: 10.1099/00221287-148-5-1305

Gente S, Billon-Grand G, Poussereau N, Févre M (2001) Ambient alkaline pH prevents maturation but not synthesis of ASPA, the aspartyl protease from *Penicillium roqueforti.* Curr Genet 38:323–8

Haab D, Hagspiel K, Szakmary K, Kubicek CP (1990) Formation of the extracellular proteases from *Trichoderma reesei* QM 9414 involved in cellulase degradation. J Biotechnol 16:187–198 . doi: 10.1016/0168-1656(90)90035-A.

Hamin Neto YAA, da Rosa Garzon NG, Pedezzi R, Cabral H (2017) Specificity of peptidases secreted by filamentous fungi. Bioengineered 5979:1–8 . doi: 10.1080/21655979.2017.1373531

Kamitori S, Ohtaki A, Ino H, Takeuchi M (2003) Crystal structures of *Aspergillus oryzae* aspartic proteinase and its complex with an inhibitor pepstatin at 1.9 Å resolution. J Mol Biol 326:1503–1511 . doi: 10.1016/S0022-2836(03)00078-0

Karaffa L, Fekete E, Gamauf C, Szentirmai A, Kubicek CP, Seiboth B (2006) D-Galactose induces cellulase gene expression in Hypocrea jecorina at low growth rates.

Microbiology 152:1507–1514 . doi: 10.1099/mic.0.28719-0

Kawai T, Nakazawa H, Ida N, Okada H, Ogasawara W, Morikawa Y, Kobayashi Y (2013) A comprehensive analysis of the effects of the main component enzymes of cellulase derived from *Trichoderma reesei* on biomass saccharification. J Ind Microbiol Biotechnol 40:805–810 . doi: 10.1007/s10295-013-1290-6

Kawamori M, Morikawa Y, Shinsha Y, Takayama K, Takasawa S (1985) Preparation of Mutants Resistant to Catabolite Repression of *Trichoderma reesei*. Agric Biol Chem 49:2875–2879 . doi: 10.1080/00021369.1985.10867203

Li Q, Yi L, Marek P, Iverson BL (2013) Commercial proteases: Present and future. FEBS Lett 587:1155–1163 . doi: 10.1016/j.febslet.2012.12.019

Mach-Aigner AR, Pucher ME, Steiger MG, Bauer GE, Preis SJ, Mach RL (2008) Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. Appl Env Microbiol 74:6554–6562 . doi: 10.1128/AEM.01143-08.

Mandels M, Parrish FW, Reese ET (1962) Sophorose as an inducer of cellulase in *Trichoderma viride*. J Bacteriol 83:400–8

Mandels M, Reese ET (1957) Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. J Bacteriol 73:269–278 . doi: 10.1002/path.1700730133

Margolles-Clark E, Hayes CK, Harman GE, Penttilä M (1996) Improved production of *Trichoderma harzianum* endochitinase by expression in Trichoderma reesei. Appl Env Microbiol 62:2145–51

Mojzita D, Herold S, Metz B, Seiboth B, Richard P (2012) L-xylo-3-hexulose reductase is the missing link in the oxidoreductive pathway for D-galactose catabolism in filamentous fungi. J Biol Chem 287:26010–26018 . doi: 10.1074/jbc.M112.372755

Moreno-Mateos MA, Delgado-Jarana J, Codón AC, Benítez T (2007) pH and Pac1 control development and antifungal activity in *Trichoderma harzianum*. Fungal Genet Biol 44:1355–67 . doi: 10.1016/j.fgb.2007.07.012

Muthukrishnan S, Mukilarasi K (2016) Industrial Important Protease Screening and Optimization from Micro-Fungal Isolates of Ayyanar Falls Forest Samples , Rajapalalyam. World Appl Sci J 34:343–347 . doi: 10.5829/idosi.wasj.2016.34.3.23113

Nascimento AS, Krauchenco S, Golubev AM, Gustchina A, Wlodawer A, Polikarpov I (2008) Statistical Coupling Analysis of Aspartic Proteinases Based on Crystal Structures of the *Trichoderma reesei* Enzyme and Its Complex with Pepstatin A. J Mol Biol 382:763–778 . doi: 10.1016/j.jmb.2008.07.043

Peciulyte A, Anasontzis GE, Karlström K, Tomas P, Olsson L (2014) Morphology and enzyme production of *Trichoderma reesei* Rut C-30 are affected by the physical and structural characteristics of cellulosic substrates. Fungal Genet Biol 72:64–72 . doi: 10.1016/j.fgb.2014.07.011

Rauscher R, Wurleitner E, Wacenovsky C, Aro N, Stricker AR, Zeilinger S, Kubicek CP, Penttila M, Mach RL (2006) Transcriptional Regulation of xyn1, Encoding Xylanase I, in *Hypocrea jecorina*. Eukaryot Cell 5:447–456 . doi: 10.1128/EC.5.3.447-456.2006.

Sternberg D, Mandels GR (1979) Induction of cellulolytic enzymes in *Trichoderma reesei* by sophorose. J Bacteriol 139:761–9

Tanskul S, Oda K, Oyama H, Noparatnaraporn N, Tsunemi M, Takada K (2003) Substrate specificity of alkaline serine proteinase isolated from photosynthetic bacterium, *Rubrivivax gelatinosus* KDDS1. Biochem Biophys Res Commun 309:547–51

Yu L-X, Gray BN, Rutzke CJ, Walker LP, Wilson DB, Hanson MR (2007) Expression of thermostable microbial cellulases in the chloroplasts of nicotine-free tobacco. J Biotechnol 131:362–369 . doi: 10.1016/J.JBIOTEC.2007.07.942

Chapter 3: Cis acting element analysis of Trichodermapepsin in *T. reesei*

3.1 Introduction

T. reesei has a carbon responsive cellulase production mechanism and expresses cellulase when cellulose is available as a carbon source. In addition to cellulose, cellulases are induced by several soluble disaccharides, including cellobiose, α-sophorose and lactose) (Foreman et al. 2003). Generally, cellulase expression is regulated at the transcriptional level. In *T. reesei*, the transcriptional activator xylanase regulator 1 (Xyr1) was shown to act as a general and essential transcription factor to control the expression of major cellulolytic and xylanolytic genes (Stricker et al. 2006). Additional activators ACE2 (Aro et al. 2001) and ACE3 (Häkkinen et al. 2014), the repressor ACE1 (Aro et al. 2003), xylanase repressor Xpp1 (Derntl et al. 2015) and the β-glucosidase activator BglR (Nitta et al. 2012), have also been reported to have a role in the expression of major cellulolytic and xylanolytic genes. Carbon catabolite repressor Cre1 represses cellulase gene expression during glucose-induced growth (Portnoy et al. 2011). Xyr1 normally binds to 5'-GGC(A/T)$_3$-3' in the promoter region (Furukawa et al. 2009). Apart from regulating cellulolytic and xylanolytic genes, Xyr1 has pleiotropic functions, a different regulatory pattern (Klaubauf et al. 2014) and helps regulate D-xylose reductase 1 gene expression in the D-xylose catabolism pathway (Rauscher et al. 2006). Electrophoretic mobility shift assays (EMSA) confirmed that Xyr1 regulates many transporters in the sugar porter family in the major facilitator superfamily (MFS), intracellular β-glucosidase, *bgl2*, putative non-enzymatic cellulose-attacking gene, *cip1*, MFS lactose transporter, *lp*, nmrA-like gene, *endo T*, acid protease, *pepA* and the small heat shock protein, *hsp23,* in *T. reesei* RUT-C30 (Ma et al. 2016).

In addition to carbon, nitrogen is also an important factor for cellulase expression in *T. reesei.* Repression of nitrogen metabolism is a regulatory mechanism that ensures preferential utilization of easily assimilated nitrogen sources, such as ammonium. A large amount of cellulase is secreted into the extracellular medium, but the utilization of nitrogen by the secreted proteins in *T. reesei* is still unknown. Nitrogen starvation has been reported to inhibit cellulase secretion in *T. reesei* Rut-C30 (Callow et al. 2016). This suggests the importance of understanding the utilization of nitrogen by the secreted protein. Little is known about nitrogen regulators and their functions in filamentous fungi. The global regulator AreA has been identified as a cellulase activator in A*spergillus nidulans* (Lockington et al. 2002). Recently a homologue of AreA in *A. nidulans,* has been identified

in *T. reesei*, Are1 (protein ID: tre76817), which is activated by cellulase expression (Qian et al. 2019). In the presence of ammonium as the sole nitrogen source, the transcription levels of the major cellulase genes, *cbh1*, *cbh2*, *egl1* and *egl2,* were dramatically decreased in Δ*are1* mutants. Qian et al. also reported that Are1 is independent of Xyr1 and Cre1 under the same conditions. AreA has a highly conserved DNA binding motif comprised of a Cys(4) zinc finger followed by a basic domain that binds GATA in *A. nidulans* (Wilson and Arst 1998). A binding motif of 5'-HGATAR-3' exhibited high affinity binding to the AreA protein in *A. nidulans* (Ravagnani et al. 1997). However, there are no reports that have assessed the binding affinity of Are1 to any promoter in *T. reesei* promoter.

In contrast to cellulases, the regulation mechanisms of proteases are still unclear and no specific regulators of the transcription of proteases have been identified in *T. reesei.* However, AreA, has been shown to act as a protease regulator in *A. nidulans* (Hynes 1974; Lamb et al. 1997). In addition, an aspartyl protease, PapA from *T. harzianum*, has been shown to potentially regulate AreA (Suárez et al. 2005). The involvement of AreA in protease regulation has also been demonstrated in both *A. oryzae* and *Penicillium marneffei* (Christensen et al. 1998; Bugeja et al. 2012). *In Candida albicans*, AreA orthologues Gat1 and Gln3 are necessary for the protein-induced expression of the secreted aspartyl protease encoding the gene *SAP2 (Dunkel et al. 2014).* In a recent study, Qian et al. reported that the deletion of Are1 abolished the secretion of proteases in skim milk, skim milk with ammonium and skim milk with peptone (Qian et al. 2019). In chapter 2, we demonstrated that expression of TrAsP was increased under organic nitrogen conditions, such as BSA and peptone (Daranagama et al. 2019). The study by Qian et al. demonstrated that Are1 positively regulates the expression of aspartic proteases *apw1* (in this study *trasp*) and *apw2* in peptone conditions. Moreover, the transcription of *apw1* was upregulated in the Δ*are1* strain when ammonium was added, suggesting that the deletion of *are1* led to decreased levels of secreted protease expression by ammonium.

Taken together, these findings and those of chapter 2 demonstrate the importance of Xyr1 and Are1 in protease expression. In this study, we focused on the regulation of transcription of TrAsP in response to the carbon source. In chapter 2, we found that a significant amount of TrAsP was produced when monosaccharides, especially galactose, were used as a carbon source. Therefore, we used known cellulase regulators in response to the galactose and cellulase inducer, Avicel, in our study in order to elucidate the mechanisms of regulation of transcription of TrAsP. Our results highlight the role of the Xyr1 and Are1 transcription factors in the regulation of the protease TrAsP in the presence of a carbon

source. We conducted a promoter deletant analysis and confirmed the binding affinity of Xyr1 and Are1 motifs to the promoter region of TrAsP. Our investigation provides evidence for the involvement of cellulase regulators in the transcriptional regulation mechanisms of TrAsP.

3.2 Materials and Methods

3.21 Strains and media

T. reesei QM9414 (ATCC269221), QMΔ*xyr1*, QMΔ*cre1*, QMΔ*ace1*, QMΔ*xpp1*, QMΔ*pac1* and QMΔ*are1* were obtained from laboratory stocks (Hirasawa et al. 2019). All mutants were constructed as described previously using QM9414Δ*P* (Shida et al. 2015). The gene deletion cassette used contained the 1.5 kbp upstream and downstream region and *pyr4* as a selection marker. All strains were maintained on potato dextrose agar plates (PDA; BD Diagnostics, Franklin Lakes, NJ, USA). *Escherichia coli* DH∝ was used for plasmid amplification. *E. coli* BL21 (DE3) was used as a host for heterologous expression of the DNA-binding domain of Xyr1 and Are1.

For cultivation experiments, 1 × 10^7 spores were grown with 1% Avicel or galactose containing 0.14% $(NH_4)_2SO_4$, 0.2% KH_2PO_4, 0.03% $CaCl_2 \cdot H_2O$, 0.03% $MgSO_4 \cdot 7H_2O$, 0.05% yeast extract, 0.1% bacto polypeptone, 0.1% Tween 80 and 0.1% trace element solution [0.005 mg/L $FeSO_4$, 0.002 mg/L $MnSO_4 \cdot H_2O$, 0.001 mg/L $CuSO_4 \cdot 5H_2O$, 0.002 mg/L $ZnSO_4 \cdot 7H_2O$, 0.001 mg/L $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 0.002 mg/L $CoCl_2 \cdot 6H_2O$ and 0.001 mg/L $NiCl_2 \cdot 6H_2O$] in 50 mM Na-tartrate buffer (pH 4.0). For analysis of β-glucuronidase (GUS) activity, mycelia were harvested by filtration on miracloth and immediately frozen in liquid nitrogen.

For induction experiments, 1 × 10^6 spores were grown in 0.3% glucose medium for 48 hours at 28 °C and centrifuged at 220 rpm, as described previously (Mach-Aigner et al. 2008). Mycelia were collected by filtration on a miracloth, washed twice with a saline solution and transferred to basal medium containing 1% galactose or 1% Avicel. After induction, mycelia were harvested by filtration at 4, 8, 16 and 24 hours and were then used for further experiments.

3.23 Construction of plasmids for reporter analysis and transformation

For analysis of promoter activities, the DNA fragment containing the *trasp* gene and its upstream (2611 bp) and downstream (4428 bp) sequences was obtained by PCR using the *T. reesei* genome as a template. The fragment was ligated with the pUC118 plasmid using restriction enzymes XbaI and HindIII, respectively, and the BKL kit (Takara Bio). The resulting plasmid, pUtrasp was opened by inverse PCR using primers asp1_fw and asp1_rv (Table S1) to remove *trasp*. The 1.8-kbp *gus* gene, encoding β-glucuronidase from *E. coli,* was generated by PCR using pBACgus-1 (Novagen) as a template, as described previously (Ogasawara et al. 2006). *gus* was ligated with pUtrasp using the BKL kit, resulting in the plasmid pUtrasp.gus. In order to insert the selective marker *amdS,* pUtrasp.gus was opened using inverse PCR −277 bp downstream of *gus*. Primers asp2_fw and asp2_rv were used for inverse PCR. The *amdS* gene was obtained from p3SR2 (Shida et al. 2008), which was used as the template for PCR. *amdS* was ligated pUtrasp.gus creating pUtrasp.D0. Using the forward primer at −880 bp (del.880_fw) and the reverse primer at −1080 bp (del.1080_rv), inverse PCR was carried out to delete the region between −1080 bp and −880 bp in pUtrasp.D0, creating the self-ligated plasmid pUtrasp.D1. The following primers were used to construct pUtrasp.D2, pUtrasp.D3 and pUtrasp.D4, respectively: del.670_fw, del.460_fw and del.230_fw. All primers used are listed in Table S1.

For site-directed mutagenesis, the PrimeSTAR MAX mutagenesis kit (Takara Bio) was used with the pUtrasp.D0 plasmid as a template and the forward and reverse primers D0-x1m, D0-x2m, D0-a1m, D0-a2m and D0-a3m. Site-directed mutagenesis was confirmed by sequencing. Primers used are listed in Table S2.

T. reesei protoplasts were prepared as described previously (Shida et al. 2016) and transformed as described previously (Penttila et al. 1987). Transformants were selected on minimum medium containing acetamide as the sole nitrogen source.

3.24 Preparation of cell-free extracts and β-glucuronidase assay

Mycelia cultivated in Avicel were frozen and ground to a fine powder while submerged in liquid nitrogen. The mycelial powder was suspended in an appropriate volume of GUS extraction buffer (Jefferson 1987) and the resulting suspension was centrifuged at 13,000 *g* for 15 minutes at 4°C. The protein concentration of the cell-free extract was determined using the Bradford assay (Bio-Rad Laboratories). GUS activity was measured as

described previously (Jefferson 1987) using 4-methyl-umbelliferyl-β-D-glucuronide (4MUG) as a substrate. Specific GUS activity was expressed as relative fluorescence/l g protein.

3.25 Measurement of proteolytic activity

Activity of protease A was measured by hydrolysis of the fluorogenic peptide MOCAa-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(Dnp)-NH_2, as described in Chapter 2.

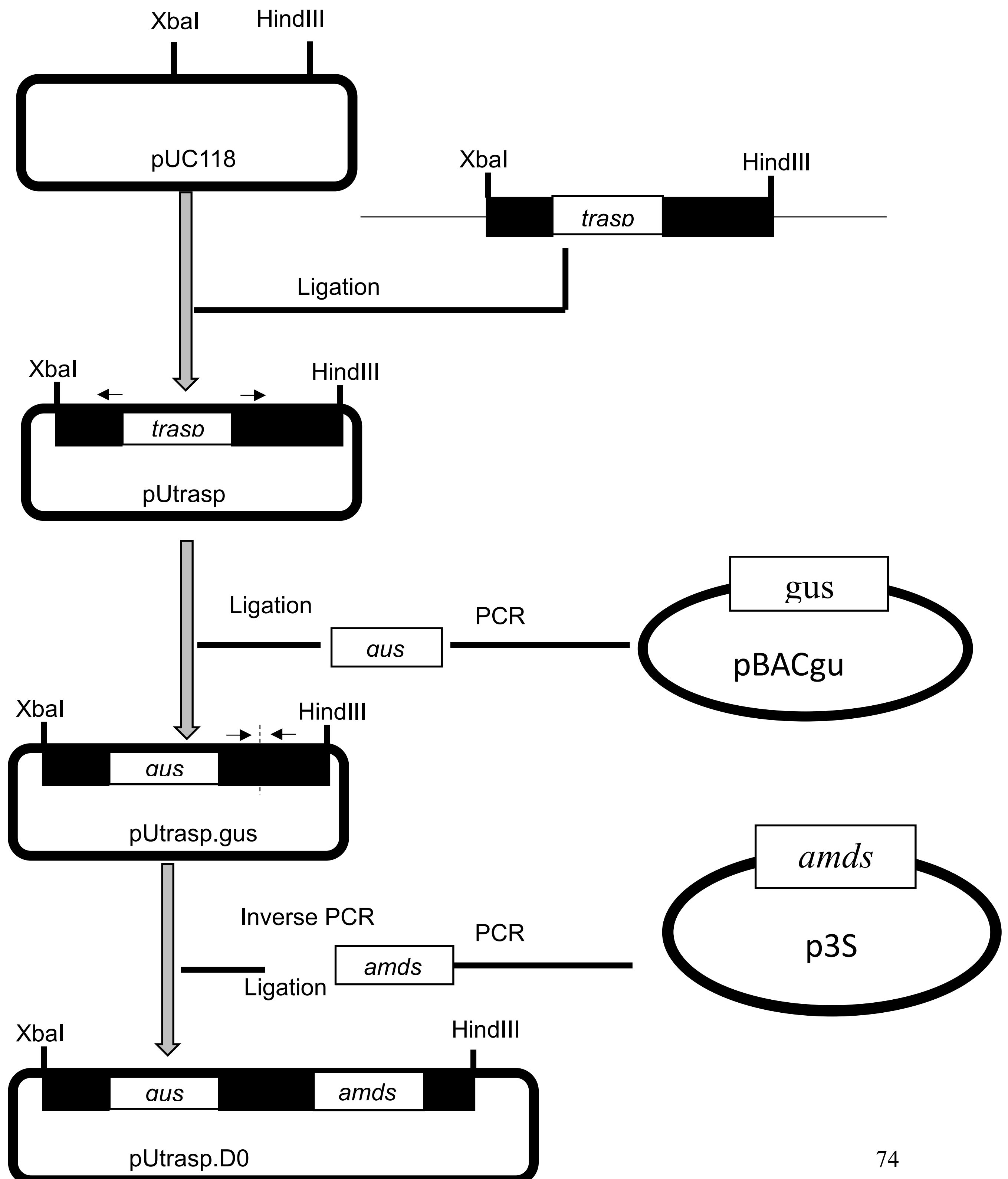


Fig. 3.1 Construction of plasmids for reporter analysis.

Table 3.1 Primers used for constructing promoter deletant mutants (D0 mutants).

Name	Sequence (5’-3’)
asp0_fw	CCGGGGATCCTCTAGAGAAGCAGTCCAGGCTCTCTTGTCCA
asp0_rv	GGCACTGGCCAAGCTTGTGATGCAGAAGCGGATGATTTGG
gus_fw	ATGGTCCGTCCTGTAGAAACCCC
gus_rv	CGACTTATTGTTTGCCTCCCTGCTGC
amds_fw	CTAGTCATCATTGGATAGGCAGATTACTCAGCCTG
amds_rv	CTGGAAACGCAACCCTGAAGGG
asp1_fw	ATATCAACATTAAATCCTGCTTTGCCCCATCCTGACC
asp1_rv	CTTGAATATCGGAGAAGGTTGCTCACGGGCTGGCT
asp2_fw	GGACGATGCCATGAACATGGACTT
asp2_rv	GCCACTACACATCAACACATACCCAGAG
del.1080_rv	ATCCCATCCATGCCTTGCTTGA
del.880_fw	CCTCGGTCATACAAAGCAAAGACG
del.670_fw	ACGGCCAGCCACGACTCGGGATT
del.460_fw	CCTCCATTGGAGCAAGGCAGAAAT
del.230_fw	GGGAGGATGGCGCAAACC
Homo check_ rv	TGAAAGGTATGTATCGTGCCCGG

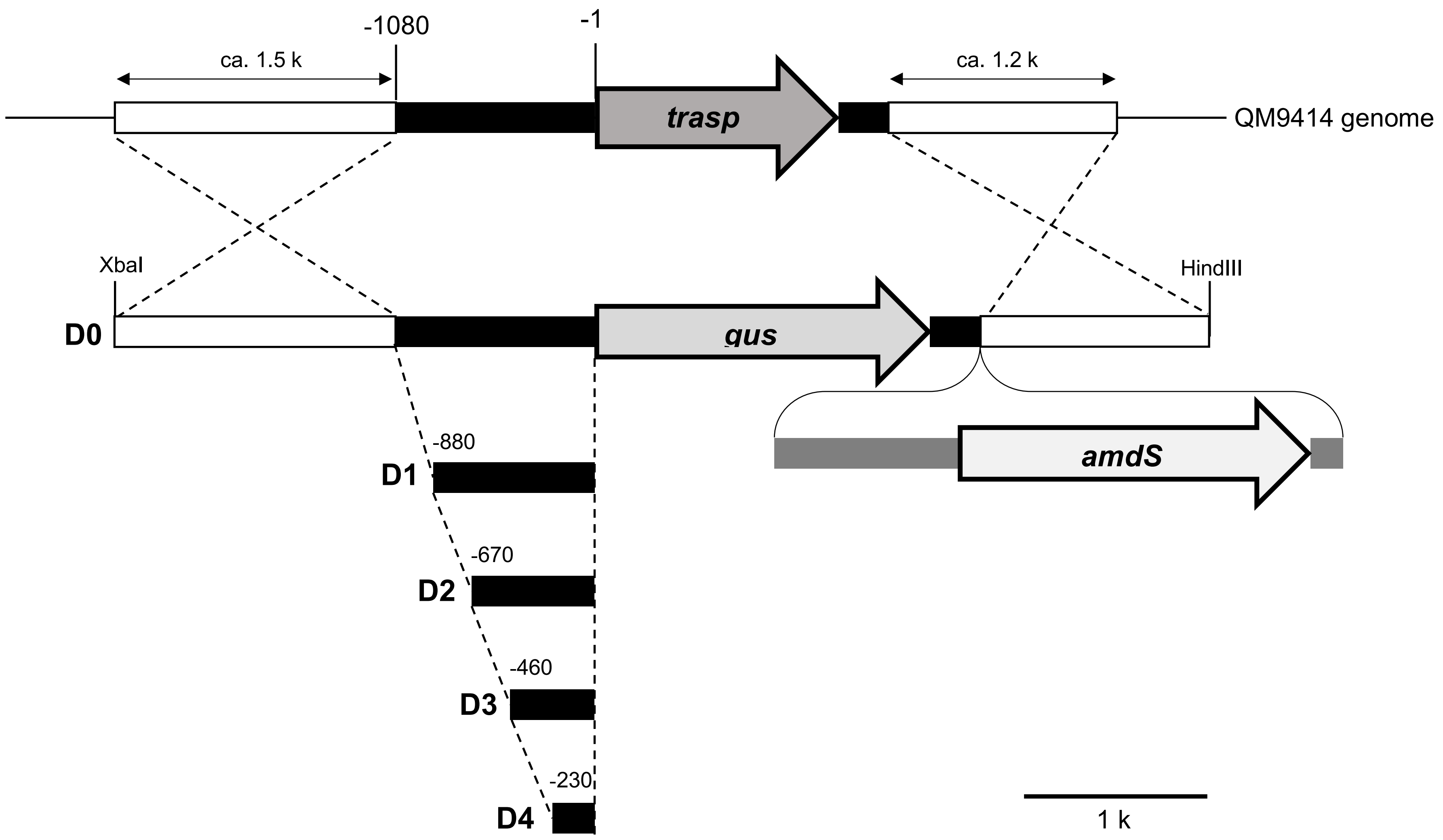


Fig.3.2 *gus* expression cassettes with selection marker *amds* (D0) with deletion of alleles at 880, 670, 460 and 230 base pairs in D1, D2, D3 and D4 transformants, respectively.

Table 3.2 Primers used for EMSA analysis

Binding motif	Name	Sequence (5'-3')
GGCTAA	Prxyn1.1fw	TTGGCAGGCTAAATGCGACATCTTAGCCGGA
	Prxyn1.1rv	TCCGGCTAAGATGTCGCATTTAGCCTGCCAA
AGATAA	pANniaDsite5 Fw	TCTTCCCACCAGAGATAAGAGATTCCGAGG
	pANniaDsite5 Rv	CCTCGGAATCTCTTATCTCTGGTGGGAAGA
GGCAAT	Pt-x1-fw	TAGACAAAGATCGGCAATGGATATCTCTCT
	Pt-x1-rv	TAGACAAAGATCGGCAATGGATATCTCTCT
ATTGCC	Pt-x2-fw	GGGCGCAGCGCCATTGCCCAAGTCATGTCG
	Pt-x2-rv	CGACATGACTTGGGCAATGGCGCTGCGCCC
GATATC	Pt-a1-fw	AGATCGGCAATGGATATCTCTCTCGAGGCC
	Pt-a1-rv	GGCCTCGAGAGAGATATCCATTGCCGATCT
TATC	Pt-a3-fw	CCCCAGCCTGTCATATCGTCACACGGACTG
	Pt-a3--rv	CAGTCCGTGTGACGATATGACAGGCTGGGG
TGATAG	Pt-a2-fw	CAGGAAGACCCTTGATAGACAAAGATCGGC
	Pt-a2-rv	GCCGATCTTTGTCTATCAAGGGTCTTCCTG

3.26 RNA extraction and real-time quantitative reverse transcription PCR (qRT-PCR)

Total RNA was extracted from frozen mycelia using the Illustra RNAspin kit (GE Health Care) and a modified hot-phenol method, as described in chapter 2. Synthesis of cDNA from

total RNA was carried out using the TRanscriptor First-Strand cDNA Synthesis kit (Roche Diagnostics, Germany) according to manufacturer's instructions.

qRT-PCR was carried out using the LightCycler 480 System (Roche Diagnostics, Germany). Amplification reactions were performed in a final volume of 20 μL using a LightCycler 480 SYBR Green I Master kit (Roche Diagnostics) with 0.5 mM forward primer, 0.5 mM reverse primer and 2 μL of 100-fold diluted cDNA. The PCR primers used included *trasp* primers, Fw (5'-GCTCAACCGCAACTCCATC-3') and Rv (5'-GAAGACGACACCCTCCTGCT-3'), as well as β-actin (*act1*) primers, Fw (5'-TCCATCATGAAGTGCGAC-3') and Rv (5'-GTAGAAGGAGCAAGAGCAGTG-3'). All analyses were performed independently in triplicate with a no-amplification control. The specificity of the PCR amplification was documented using melting curve analysis.

3.27 Expression and purification of recombinant Xyr1 and Are1

The DNA-binding domain of Xyr1 (residues 155–230) or the DNA-binding domain of Are1 (residues 664–854) was ligated to the pET22b plasmid. After confirmation by DNA sequencing, the pET plasmid was expressed in *E. coli* BL21(DE3) for protein expression. *E. coli* BL21(DE3) harbouring pET21XyrI or pET21Are1 was grown at 37 °C to an optical density of 600 nm (OD_{600}). The recombinant protein was induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside) to a final concentration of 0.2 mM, and the cultures were incubated for an additional 5 hours at 37 °C.

The cells were harvested, washed twice with wash buffer (50 mM sodium phosphate buffer pH 7.5, 300 mM NaCl, 10 mM imidazole) and resuspended in 10 ml of the same buffer. The cell suspension was sonicated on ice. After centrifugation, the supernatant was recovered and His6-Xyr1, or His6-Are1 was purified from the whole-cell lysate using TALON metal affinity resin (Clontech, Takara Bio), according to the manufacturer's protocol. Purified Xyr1 or Are1 were analysed by SDS-PAGE using Coomassie blue staining and were stored in EMSA binding buffer with protease inhibitor cocktail tablets (Complete mini, Roche Diagnostics, Germany).

3.28 Electrophoretic mobility shift assay (EMSA)

Complimentary oligonucleotides (30 bp) (Table S3) containing Xyr1 and the Are1 binding

site located in the *trasp* promoter were annealed by heating at 95 °C for 5 minutes and were then slowly cooled to room temperature in TE buffer. The annealed probes were purified by non-denaturing PAGE. A DIG Gel Shift kit, 2nd generation (Roche Diagnostics, Germany) was used to detect DNA–protein interactions. Digoxigenin (DIG) was labelled at the 3'-end of the double stranded oligonucleotides. The labelled probes were then purified by ethanol precipitation. The pellet was air-dried and resuspended in ddH_2O to a final concentration of 25 fmol/µL. DIG-labelled (75 fmol) oligonucleotides were mixed with Xyr1 or Are1 protein in binding buffer (25 mM Hepes-KOH (pH 7.9), 50 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 0.1 mM PMSF, 0.1% pepstatin, 10% glycerol and 1 µg of poly [dI–dC]) and kept at 25 °C for 20 minutes. Reaction mixtures were electrophoresed on a 5% polyacrylamide gel in 0.5 M TBE at 150 V, 4 °C for 90 minutes.

The proteins were then electrophoretically transferred on to a positively charged nylon membrane (Roche) by applying a constant current of 12 V/200 mA for 120 minutes. The DNA was cross-linked to the membrane using a UV crosslinker CL-1000 (Funakoshi, Japan). The nylon membrane was blocked in the blocking reagent for 1 hour at room temperature and was subsequently incubated with a 1:20,000 dilution of alkaline phosphatase coupled anti-DIG antibody for 30 minutes. The membrane was washed twice with washing buffer [0.1 M maleic acid (pH 7.5), 0.15 M NaCl, 0.3% (v/v) Tween 20] for 20 minutes each wash. After 5 minutes equilibration in the detection buffer (100 mM Tris/HCl, pH 9.5 and 100 mM NaCl), the membrane was carefully placed on a plastic sheet. 1:100 dilutions of the CSPD substrate was then added dropwise on to the membrane near the edges. The membrane was then tilted to fully cover it with the CSPD substrate. The membrane was incubated at room temperature for 5 minutes and subsequently placed in a plastic folder and incubated at 37 °C for another 15 minutes. Finally, the membrane was exposed to an X-ray film for 3 hours to capture the chemiluminescent signal.

3.29 Statistical Analysis

Statistical analysis was performed using a Student's *t* test. The D0 mutant or QM9414 strain was used as a standard for comparison with each mutant under all conditions analysed. A p-value of $p<0.05$ (*$p<0.05$, ** $P<0.01$, ***$P<0.001$) was considered statistically significant.

Table 3.3 Primers used for EMSA analysis

Binding motif	Name	Sequence (5'-3')
GGCTAA	Prxyn1.1fw	TTGGCAGGCTAAATGCGACATCTTAGCCGGA
	Prxyn1.1rv	TCCGGCTAAGATGTCGCATTTAGCCTGCCAA
AGATAA	pANniaDsite5 Fw	TCTTCCCACCAGAGATAAGAGATTCCGAGG
	pANniaDsite5 Rv	CCTCGGAATCTCTTATCTCTGGTGGGAAGA
GGCAAT	Pt-x1-fw	TAGACAAAGATCGGCAATGGATATCTCTCT
	Pt-x1-rv	TAGACAAAGATCGGCAATGGATATCTCTCT
ATTGCC	Pt-x2-fw	GGGCGCAGCGCCATTGCCCAAGTCATGTCG
	Pt-x2-rv	CGACATGACTTGGGCAATGGCGCTGCGCCC
GATATC	Pt-a1-fw	AGATCGGCAATGGATATCTCTCTCGAGGCC
	Pt-a1-rv	GGCCTCGAGAGAGATATCCATTGCCGATCT
TATC	Pt-a3-fw	CCCCAGCCTGTCATATCGTCACACGGACTG
	Pt-a3--rv	CAGTCCGTGTGACGATATGACAGGCTGGGG
TGATAG	Pt-a2-fw	CAGGAAGACCCTTGATAGACAAAGATCGGC
	Pt-a2-rv	GCCGATCTTTGTCTATCAAGGGTCTTCCTG

3.3 Results

3.31 Involvement of Xyr1 and Are1 in *trasp* expression

In chapter 2, we highlighted the effect of different carbon sources on expression of TrAsP. In this study, in order to elucidate the relationship between carbon sources and Xyr1 and Are1 on TrAsP expression, we carried out secretory protease activity and qRT-PCR analysis in transcription factor deletant mutants QMΔ*xyr1* and QMΔ*are1*. We used Avicel as the carbon source, as it is a prominent cellulase inducer, and galactose due to its ability to induce expression of TrAsP. *T. reesei* QM9414, QMΔ*xyr,* and QMΔ*are1* were cultivated in 1% Avicel and 1% galactose medium and the supernatants were collected after 4 days. We analysed protease activity observed a similar pattern of activity with both Avicel and galactose media (Fig. 3.3a and Fig. 3.3b). QMΔ*xyr1* exhibited higher activity than QM9414 and QMΔ*are1* had significantly lower activity than QM9414. qRT-PCR analysis was performed to determine the effect of these transcription factors on *trasp* (Fig.3. 3c and 3d). QM9414 and the mutants were precultivated for 48 hours in 0.3% glucose medium and were then transferred to 1% Avicel or galactose induction medium. Mycelia were collected at 4, 8, 16 and 24 hours and levels of RNA expression were measured. Similar to the results of the activity analysis, QMΔ*xyr1* exhibited higher expression after 4 hours both in Avicel and galactose. However, expression values were higher in galactose than in Avicel. QMΔ*are1* had significantly lower expression in both media. These results suggest a differential behaviour between Xyr1 and Are1 in response to the carbon source. While Xyr1 had repressive activity, Are1 played an activating role in *trasp* expression.

To elucidate the relationship of Xyr1 and Are1 in the regulation of TrAsP, we investigated the TrAsP promoter region. The upstream 1080 bp in the 5' noncoding region of *trasp* was obtained from the JGI website for *T. reesei*. Analysis of the promoter region revealed one putative TATA box at position –115 and 10 CCAAT boxes as the elements involved in initiation of transcription. Several putative regulatory sequences of Xyr1 (GGCWWW) (Furukawa et al. 2009) and Are1(GATA, HGATAR) (Ravagnani et al. 1997) were identified in the promoter region (Fig. 3.4).

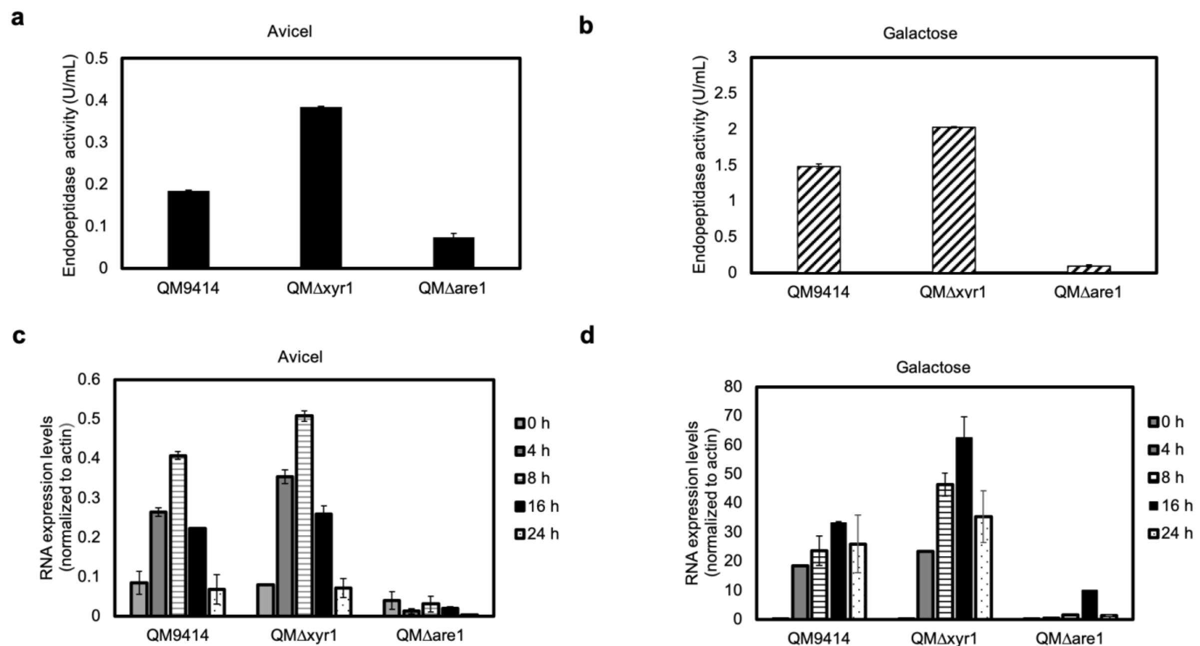
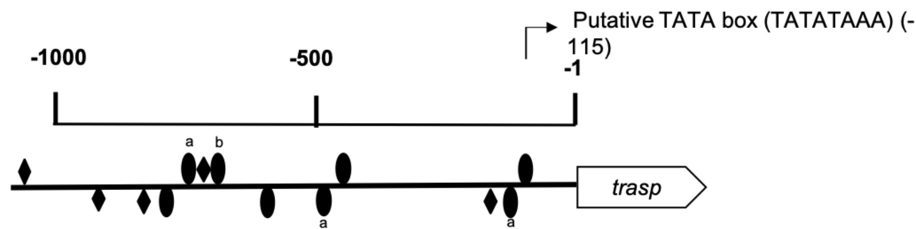


Fig. 3.3 Endopeptidase activity using fluorogenic peptide MOCAa-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(DND)-NH_2. Supernatants of (a) 1% Avicel and (b)1% galactose after 4-day cultivation of *T. reesei* and mutants QMΔ*xyr1* and QMΔ*are1*. Transcriptional analysis of (c) Avicel or (d) galactose induction of *trasp. T. reesei* QM9414 and mutants were precultured on 0.3 % glucose and transferred to a medium containing 1% (w/v) Avicel or galactose. Mycelia were collected after 4, 8, 16 and 24 hours. All values were normalised to levels of β-actin under the same conditions. Each bar represents the mean of three independent experiments with the SD indicated by a vertical line



Transcription factor	Putative binding sequence	Position	Sequence in TrAsP promoter
Xyr1	GGCWWW	-172	AATGCC
		-731	GGCAAT
		-843	ATTGCC
		-926	ATTGCC
		-1066	GGCTTT
AreA	GATA	-152	TTATCG [a]
		-506	CTATCT [a]
		-746	TGATAG [a]
		-117	GATA
		-447	GATA
		-602	TATC
		-724	GATATC [b]
		-803	TATC

a: HGATAR sequence b: GATAs coding and non-coding strand were overlapped.

Fig. 3.4 Schematic representation of the 1080 kb 5'-non coding upstream region of TrAsP in *T. reesei* QM9414 with putative binding sequences, Xyr1: *GGCWWW* (filled diamond), AreA:*GATA* (filled circle).

3.32 Analysis of promoter deletants

To dissect the role of the promoter region of *trasp*, a promoter deletion series (D0-D4) was constructed. We created *gus* expression transformants D0, D1, D2, D3 and D4 containing promoter regions with 1080, 880, 670, 460 and 230 base pairs (bp), respectively (Fig. 3.5). In order to analyse the promoter strength of each mutant, GUS activity of cell-free extracts from mycelia after cultivation of the mutants on 1% Avicel and 1% galactose was measured. Both Avicel and galactose had a similar pattern of expression (Fig. 3.5). An increase in GUS activity was observed in D1, D3 and D4 mutants and significantly lower activity was observed in the D2 mutant as compared to the D0 mutant. Therefore, *trasp* transcription is complex and is difficult to explain using the available results. However, these

results suggest that the –1080 to –880 bp, –670 to –460 bp and –460 to –230 bp regions had a repressive function and the –880 to –670 bp region had an activating function. Therefore, we selected the –880 to –670 bp region for further studies as both the targets of Xyr1 and Are1 binding motifs were located in this region that is pivotal for positive regulation of the promoter.

3.33 Site-directed mutagenesis

In order to find the effects of Xyr1 and Are1 binding sites on *trasp* promoter strength, we conducted site-directed mutagenesis analysis. A site-directed mutation was introduced into the motifs of the D0 reporter construct, which resulted in D0-x1m, D0-x2m, D0-a1m, D0-a2m and D0-a3m constructs (Fig. 3.6). The reporter constructs were introduced into the chromosomal DNA of *T. reesei* QM9414 by homologous recombination. GUS activities of mutants were measured in both Avicel and galactose (Fig. 3.6). With respect to GUS activities of the transformants, mutations of Xyr1 motifs resulted in a 13% increase in D0-x1m (**T**GCAAT) and a 15% increase in D0-x2m (ATT**A**CC) as compared to the D0 mutant in Avicel. In the galactose medium, 16% and 18% increases were observed in GUS activity of the D0-x1m mutant and the D0-x2m mutant, respectively. On the other hand, mutations of AreA motifs at –745 (D0-a2m/5'-T**T**ATAG-3') and motifs at –724 (D0-a1m/5'-GA**C**A-3') resulted in 27% and 16% reductions in relative activity, respectively, as compared to the D0 mutant in Avicel medium. In galactose medium, the D0-a2m mutant and D0-a1m mutant showed 40% and 34% reductions, respectively, as compared to the D0 mutant. These results suggest that both positions are functional. Since mutations in the AreA motif at –803 (D0-a3m/5'-TAT**T**-3') demonstrated no impact on the relative GUS activity in Avicel medium, but a 20% reduction in galactose medium, its functionality seems to depend on the carbon source. Based on these results, we concluded that the Xyr1 and Are1 motifs in the region between 880 bp and 670 bp are functional for *trasp* expression. In addition, protease activity, RNA expression and site mutagenesis analysis confirmed that the Xyr1 transcription factor has repressive activity and that Are1 has an activating function on the *trasp* promoter.

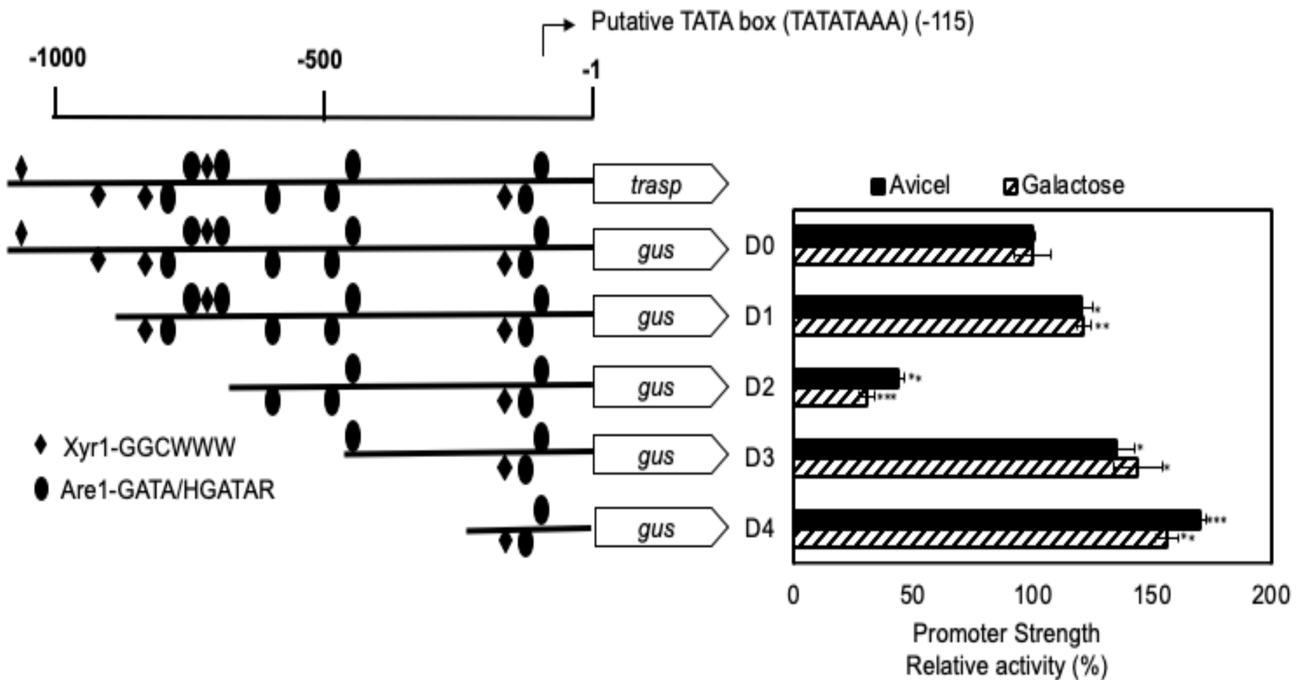


Fig. 3.5 Promoter analysis of *trasp* using the GUS reporter gene. Relative GUS activity measurements from transformants of D0, D1, D2, D3, and D4 cultivated in 1% Avicel and 1% galactose (right panel). Each bar represents the mean of three independent experiments, with SD indicated by a vertical line. Deletions in each transformant and the binding sequences are illustrated in the left panel. Statistically analysis was carried out using a Student's *t*-test, $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

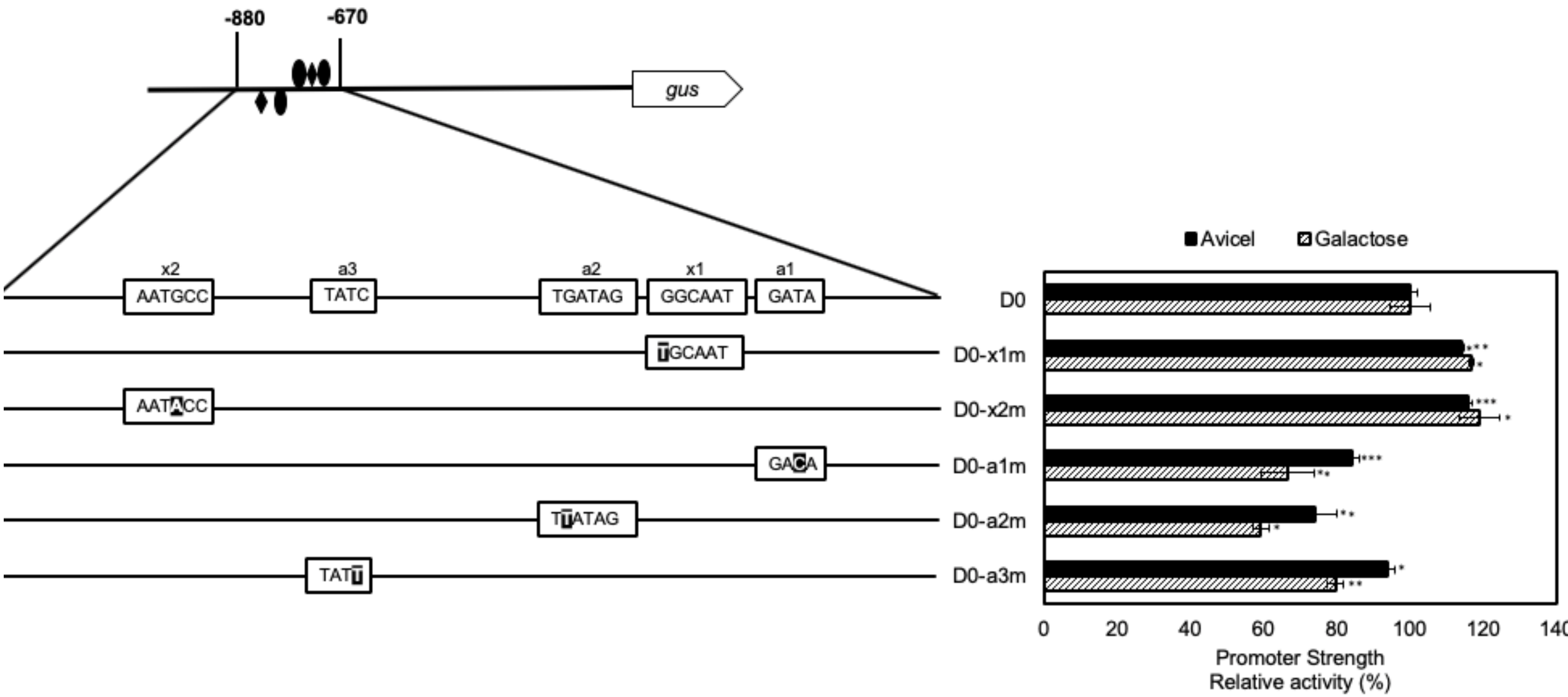


Fig. 3.6 Promoter analysis after site-directed mutagenesis of the Xyr1 and Are1 binding motifs in the *trasp* promoter. Transformants of D0, D0-x1m, D0-x2m, D0-a1m, D0-a2m, and D0-a3m were cultivated in 1% Avicel or 1% galactose for 4 days and mycelia were collected. The relative activity for promoter strength was calculated for each mutant (right panel). Each bar represents the mean of three independent experiments, with SD indicated by a vertical line. Each site mutagenesis location of D0 is illustrated in the left panel. Statistical analysis was carried out using a Student's *t* test, $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

3.34 Binding ability of Xyr1 and Are1 transcription factors to the TrAsP promoter

In order to assess the functionality of the binding sequences, an EMSA analysis was conducted using DIG labelled probes and Xyr1 and Are1 proteins. Oligonucleotide probes containing 5'-GGCAAT-3' (x1), 5'-ATTGCC-3' (x2), 5'-TGCAAT-3' (x1m) and 5'-ATTACC-3' (x2m) were prepared (Table 3.2, Fig. 3.6). 5'-GGCTAA-3' in the *xyn1* promoter was used as the control and a similar binding strength was observed with the purified protein as described previously (Furukawa et al. 2008) (Table 3.2, Fig. 3.8). In the EMSA, Xyr1 exhibited a comparatively weak binding pattern to the selected oligonucleotides and its specificity was confirmed by a competitor assay (Fig. 3.9a). These results demonstrate that Xyr1 can bind to the x1 and x2 motifs found in the upstream region of *trasp*.

Three oligonucleotide probes containing 5'-GATA-3' (a1), 5'-TGATAG-3' (a2) and 5'-TATC-3' (a3) were used for the EMSA analysis of Are1 (Table S3). For the competitor assay, mutant probes 5'-GACATC-3' (a1m), 5'-TTATAG-3' (a2m) and 5'-TATC-3' (a3m) were prepared. As the control, the 5'-AGATAA-3' oligonucleotide, which is site 5 of the NiiA-NiaD bidirectional promoter of *A. nidulans*, was used (Muro-Pastor et al. 2017) and a strong binding pattern was observed (Fig.3.8). The selected sites exhibited weak binding to the *trasp* promoter and the a2 position had comparatively strong binding (Fig. 3.8b). A competitor assay confirmed the specificity of binding. These results suggest that the Are1 protein can bind to the *trasp* promoter.

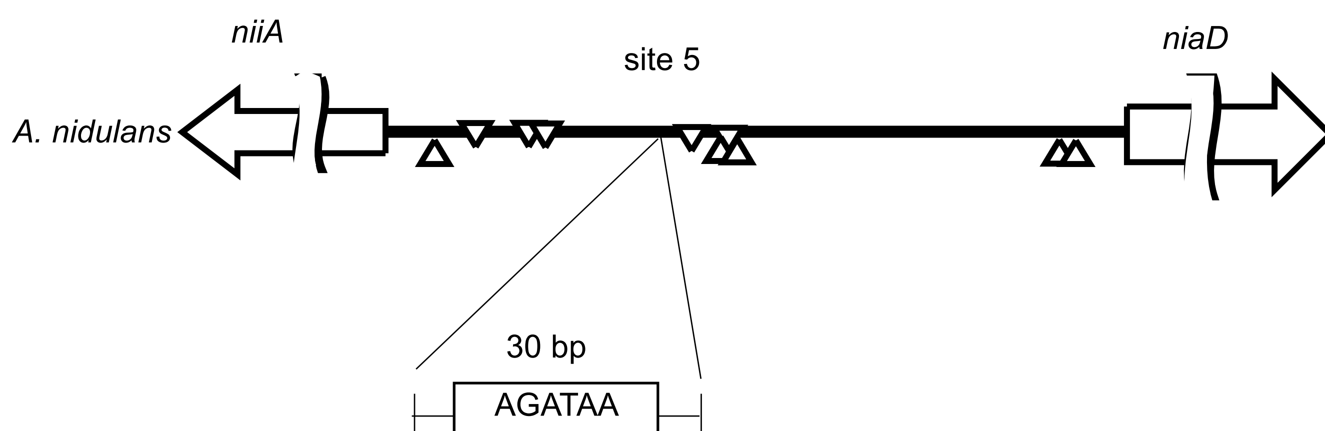
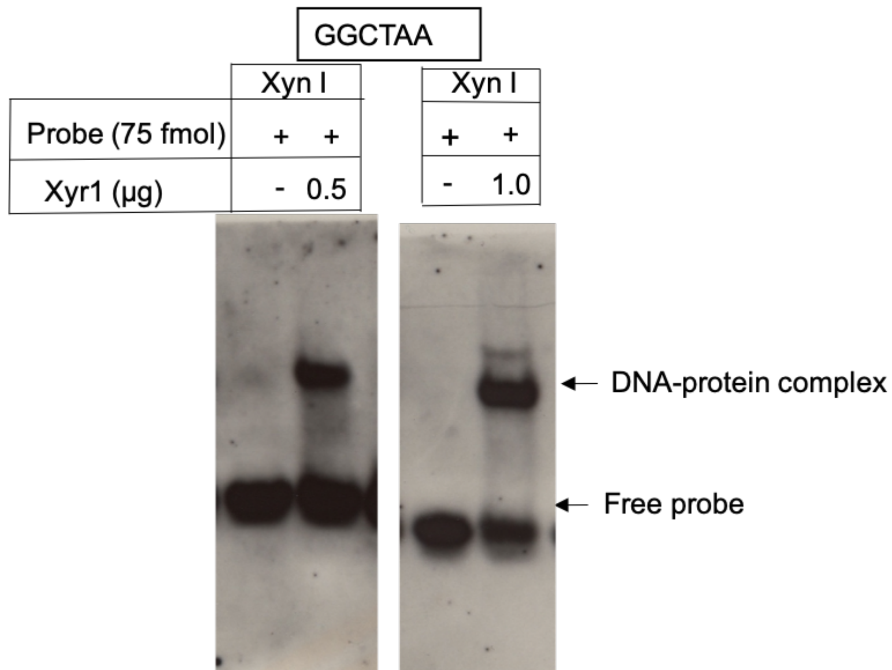


Fig. 3.7 The niiA-niaD bidirectional promoter of *Aspergillus nidulans.*

a



b

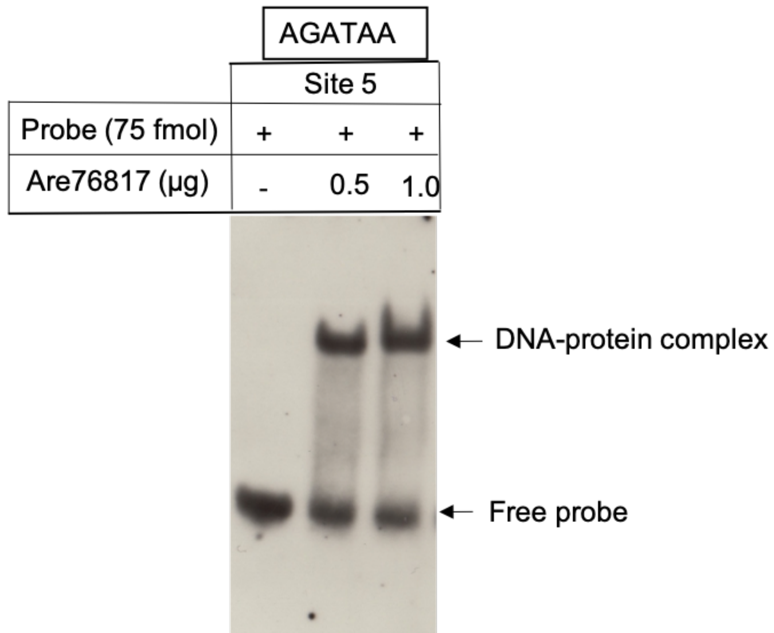


Fig. 3.8 As a positive control, the *xyn1* promoter fragment of *T. reesei* was mixed with Xyr1 (0.5 mg and 1.0 mg). B Site 5 of the NiiA-NiaD bidirectional promoter fragment of *A. nidulans* mixed with Tr76817 at a concentration of 0.5 mg and 1.0 mg. +, addition of oligonucleotide probe to the binding reaction; −, no protein.

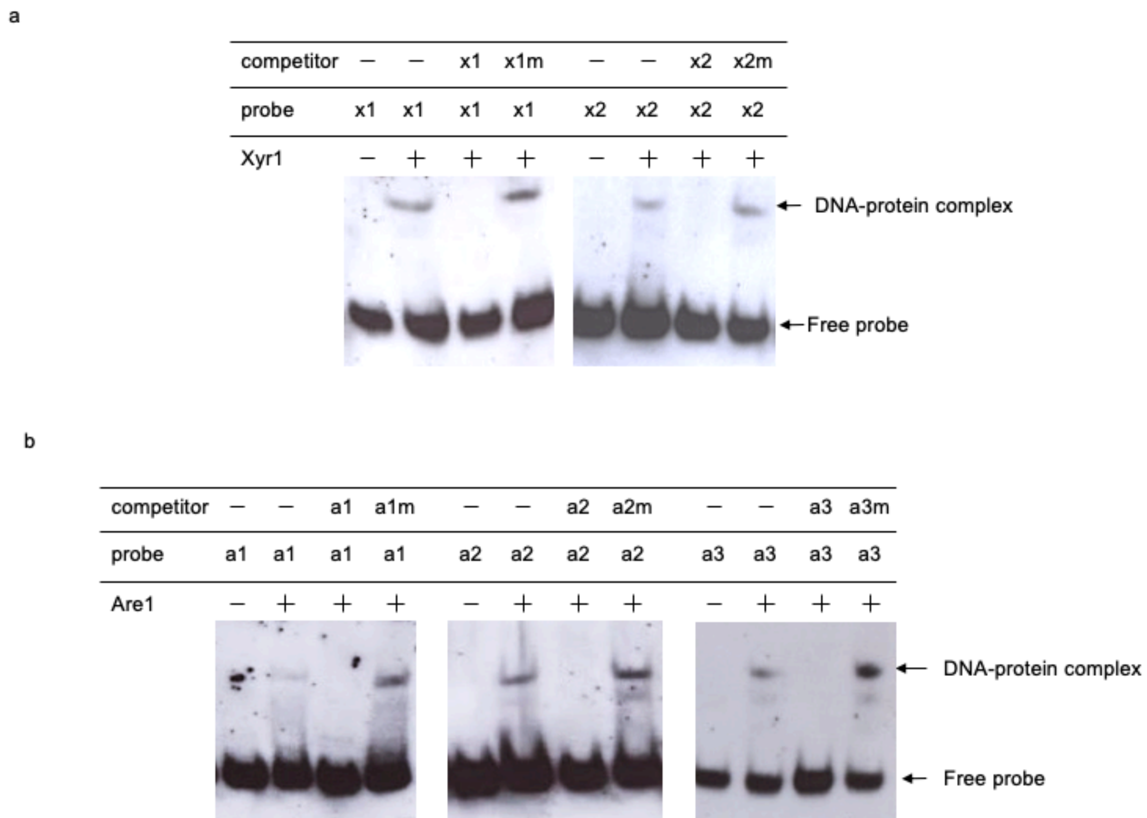
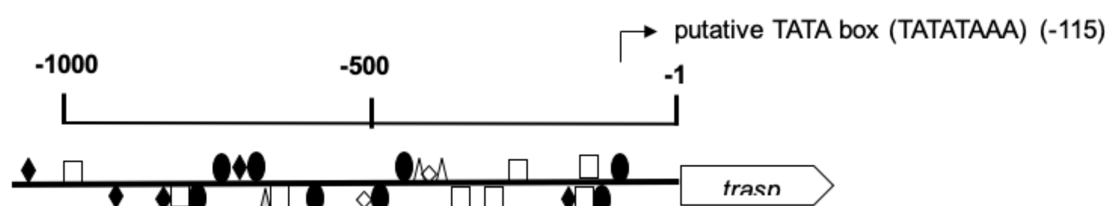


Fig. 3.9 (a) Binding analysis of Xyr1 to the *trasp* promoter fragments at x1 and x2 probes (75 fmol) with Xyr1 at 1.0 µg. (b) Binding analysis of Are1 to the *trasp* promoter fragments at a1, a2 and a3 probes (75 fmol) with an Are1concentration of 1.5 µg. 7.5 pmol of unlabelled probes and unlabelled mutant probes were used as competitors. +, addition of protein to the binding reaction; −, no protein.

3.35 Effect of other cellulase regulators on *trasp* expression

In this study, we demonstrated the involvement of Xyr1 and Are1 transcription factors on expression of *trasp*. The results of the promoter deletant analysis (Fig. 3.5) suggested that other factors were involved in the regulation of expression of *trasp*. We aimed to elucidate the impact of cellulase regulators on *trasp* expression. To this end, we further investigated the

binding motifs of several known cellulase regulators. We found several putative binding sequences of cellulase repressors, including Ace1 (AGGCA), Cre1 (SYGGRG) and Xpp1 (AGAA), as well as the pH regulator Pac1 (GCCARG) (Shida et al. 2016) (Fig. 3.10). In order to elucidate their effect on *trasp* regulation, we carried out a peptidase assay and RNA expression analysis using transcription factor deletants. Results from the enzyme assay suggest that Cre1, Xpp and Pac1 act as repressors for proteases in both Avicel and galactose (Fig. 4.11a and Fig. 3.11b). The relative effect was higher in Avicel than galactose. Relative RNA expression of *trasp* was higher in *cre1*, *xpp1* and *pac1* deletants at 16 hours and 24 hours as compared to QM9414. Ace1 had little effect on *trasp* expression as measured by mRNA levels.



Transcription factor		Function	Putative binding sequence	Position	Sequence in TrAsP promoter
△	Ace1	Repressor for cellulase and xylanase expression	AGGCA	-391	AGGCA
				-446	AGGCA
□	Cre1	Carbon catabolite repressor	SYGGRG	-166	CCGGAG
				-168	CCCCGG
				-253	CCCCAG
				-298	GCGGAG
				-382	CTCCAC
				-690	CTCCGC
				-816	CCCCAG
				-980	CTGGGG
△	Pac1	Global pH regulator	GCCARG	-696	CCTGGC
◇	Xpp1	Repressor for hemicellulose expression	AGAA	-442	AGAA
				-508	TTCT

Fig. 3.10 Schematic representation of 1080 kb of the 5'-non-coding upstream region of *trasp* in *T. reesei* QM9414 with the putative binding sequences represented as follows: Ace1:

AGGCA (open triangle), Cre1: *SYGGRG* (open square), Xpp1:AGAA (open diamond), Pac1:*GCCARG* (filled triangle), Xyr1: *GGCWWW* (filled diamond) and AreA:*GATA* (filled circle).

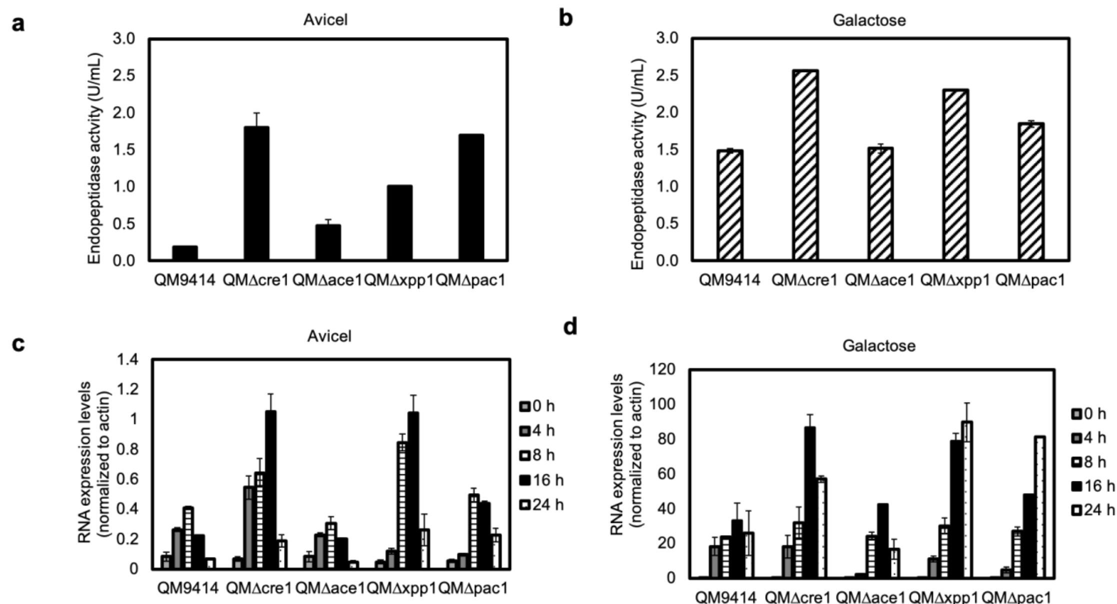


Fig. 3.11 Endopeptidase activity using fluorogenic peptide MOCAa-Ala-Pro-Ala-Lys-Phe-Phe-Arg-Leu-Lys(DND)-NH_2. Supernatants of 1% Avicel (a) and 1% galactose (b) after 4-day cultivation of *T. reesei* and mutants, QMΔ*cre1*, QMΔ*ace1*, QMΔ*xpp1*, QMΔ*pac1*. Transcriptional analysis of (c) Avicel or (d) galactose induction of genes encoding *trasp. T. reesei* QM9414 and mutants were precultured on 0.3 % glucose and were transferred to a medium containing 1% (w/v) Avicel or galactose and mycelia and were collected after 4, 8, 16 and 24 hours. All values were normalised to β-actin expression under the same conditions. Each bar represents the mean of three independent experiments with the SD indicated by a vertical line.

3.4 Discussion

Although fungal proteases have been identified as an impediment to the production of homologous and heterologous proteins, little is known about the protease regulatory mechanisms in fungi. Literature related to the regulatory mechanism of proteases in cellulolytic filamentous fungi is surprisingly scarce (Snyman et al. 2019). Previous studies have shown that protease regulation of the secretome is complex and depends on carbon and

nitrogen in the media in *A. nidulans* (Katz et al. 2006). In this study, we also reported that a carbon- and nitrogen-responsive regulation system exists for TrAsP. General approaches to reduce protease levels in the fungal secretome include UV mutagenesis, protease deletions, deletion of protease regulatory proteins in the genome and regulation of the culture media. Previous studies have developed a system for therapeutic protein production in *T. reesei* through deletion of multiple protease genes, including TrAsP (Landowski et al. 2015). These studies have contributed to the understanding of the regulation of proteases in the secretome.

Based on our previous findings on the carbon and nitrogen responsive mechanisms for TrAsP regulation, we focused on the effects of the main cellulase regulator, Xyr1, and the global nitrogen regulator, Are1, on regulation of *trasp*. We found that Xyr1 and Are1 act as a repressor and activator, respectively, for *trasp* expression in response to Avicel and galactose (Fig. 3.1). Similarly, previous studies have reported that Xyr1 deletion results in the upregulation of five protease genes, including the acid protease *pepA,* in lignocellulosic medium in *T. reesei* RUT-C30 (Ma et al. 2016). In addition, Xyr1-mediated reduction in proteases has been recorded in cellulose, sophorose and glucose media in *T. reesei* (dos Santos Castro et al. 2016), which is consistent with our findings. In a recent study, Are1 was shown to activate proteases in the medium with peptone as the nitrogen source (Qian et al. 2019). However, the role of Are1 with varying carbon sources has not been previously reported in *T. reesei* QM9414. This study showed that Are1 promoted TrAsP expression in response to the cultivation media, suggesting that it has a different regulation mechanism with mixed nutrients. Our SDS-PAGE analysis demonstrated that QMΔ*are1* was unable to produce both cellulases and TrAsP (data not shown). These results suggest that the protease regulation mechanisms of *T. reesei* are connected to the cellulase regulation mechanisms and that *T. reesei* regulates secretion via environmental factor signal transduction. We also demonstrated that Xyr1 and Are1 exhibit a similar pattern in both Avicel and galactose medium. However, galactose resulted in higher *trasp* expression in the deletion strain with the *xyr1* mutant as compared to QM9414 (Fig.3.3d). We also found that expression of TrAsP was higher in the galactose containing medium, among the monosaccharides that were tested with QM9414. Therefore, these results suggest that a different molecular mechanism for *trasp* expression exists for galactose.

We identified binding motifs for Xyr1 and Are1 in the TrAsP promoter region (Fig. 3.4). In order to elucidate the functional region of the TrAsP promoter, we performed deletion analysis on the promoter region of *trasp* in *T. reesei*. However, our results showed a complex expression pattern, indicating that more information is necessary to elucidate the regulation

mechanisms. Nevertheless, we found that the −880 bp to −670 bp region in the *trasp* promoter was crucial for its expression in both Avicel and galactose medium (Fig. 3.5). Both of our targets, Xyr1 and Are1, exist in this region. Therefore, in this study, we selected the −880 bp to −670 bp region for further analysis. Using site-directed mutagenesis analysis, we demonstrated the specific areas that were crucial for TrAsP expression. Our results indicate that both the x1 (−731) and x2 (−843) positions with the sequences 5'-GGCAAT-3' and 5'-ATTGCC-3', respectively, were involved in TrAsP expression (Fig. 3.5). Previous studies have reported the following as functional sequences: the Xyr1 binding sequence, a GGCTAA motif separated by 10 bp spacers within the *xyn1* promoter, the inverted repeat GGCTAA and GGCTGG motif separated by a 12 bp spacer within the *xyn2* promoter (Rauscher et al. 2006; Stricker et al. 2008), and the inverted repeat of GGCTAT and GGCTAA motifs separated by a 16 bp spacer in the *xyn3* promoter (Furukawa et al. 2008). However, the functional Xyr1 motifs were not only inverted repeats, but also possessed a single motif (Shida et al. 2008). In our study, the two binding sites x1 and x2 were more than 100 bp apart. This suggests that Xyr1 may regulate TrAsP expression individually or through interactions with another transcription factors. The function of Xyr1 to reduce expression of TrAsP indicates the presence of a different regulatory mechanism for proteases, as compared to the cellulase regulation mechanisms in *T. reesei*.

With respect to the Are1 motifs, we demonstrated that 5'-TGATAG-3' at the −746 (a2m) position was the most functional motif among those tested, followed by the overlap GATA motif at the −722 and −724 (a1m) positions (Fig. 3.6). According to our results, functionality of the GATA motif at the −803 position is dependent on a carbon source. A central cluster of AreA binding sites (sites 5–8) in *A. nidulans* have been shown to be responsible for 80% of the transcriptional activity of the *niiA* and *niaD* genes (Muro-Pastor et al. 2017). In a similar fashion, it is possible that Are1 motifs of the a1, a2 and a3 cluster have transcriptional activity in the *trasp* gene.

Our study also assessed the binding affinity of Xyr1 to acid proteases in *T. reesei* (Fig. 3.8), which is consistent with the results of the EMSA analysis performed by Ma et al (2016). The binding strength of the motifs is consistent with the findings of our previous study (Furukawa et al. 2009). The important nucleotides within the binding site exhibit specific interactions with Xyr1, such as strong binding to GGCTAA, moderate binding to GGCAAA and GGCTAT, weak binding to GGCTTT, GGCAAT, GGCATT and GGCATA, and no binding to GGCTTA (Furukawa et al. 2009). The 5'-GGCAAT-3' sequence in the *trasp* promoter also elicited a weak binding pattern. On the other hand, binding motifs of Xyr1 that

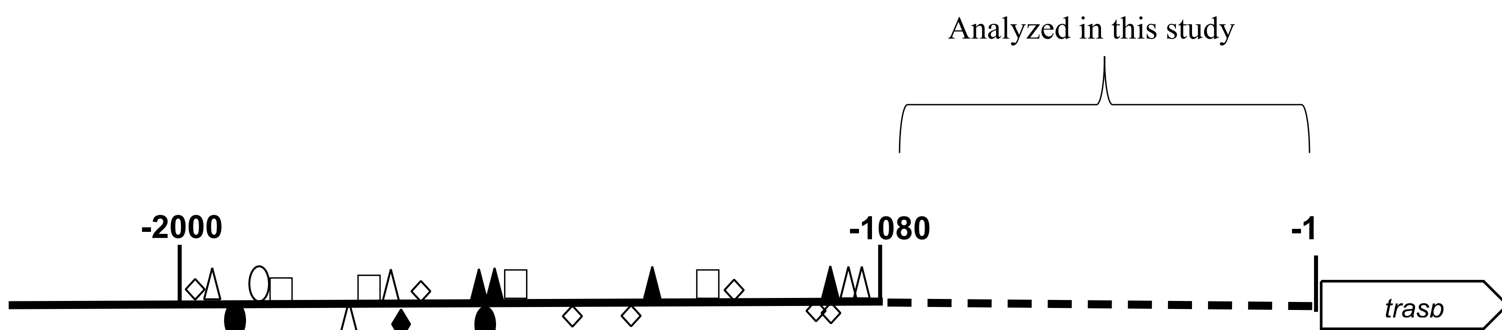
are not strong, such as the 5'-GGCTAA-3', 5'-GGCAAA-3' and 5'-GGCTAT-3' elements, which play important roles as functional Xyr1-binding sites in Xyr1-mediated cellulase and xylanase gene expression, also exist in the *trasp* promoter of *T. reesei*. Therefore, we suggest that even though 5'-GGCAAT-3' motifs showed weak binding, they are functional in TrAsP. In our previous study, we reported that 5'-GGC(A/T)$_3$-3' motifs were present in higher numbers in genes upregulated by Xyr1. However, we could not identify more Xyr1 motifs in the *trasp* promoter. Ma et al (2016) have also reported fewer motifs in genes downregulated by Xyr1, such as *pepA* and *hsp23* in *T. reesei* RUTC30 (Ma et al. 2016). We observed a weaker binding affinity of the Are1 protein to the GATA/HGATAR binding motif in the *trasp* promoter (Fig. 4.8). As no study has investigated the Are1 protein and its binding affinity to *T. reesei* genes, we selected the binding motif of AreA. In a previous study, AreA has been shown to have a strong binding affinity in *A. nidulans*. Therefore, we selected the site 5 position of the NiiA-NiaD bidirectional promoter of *A. nidulans* (5'-AGATAA-3') (Muro-Pastor et al. 2017) (Fig.4.7). 5'-HGATAR-3' had promoter-specific recognition for AreA-dependent promoters in *A. nidulans*. Similarly, we observed strong binding affinity with 5'-HGATAR-3' as compared to other GATA motifs, that naturally lack H (A, T or C) at the first base and last base 'R' (A or G) in the *trasp* promoter.

Our results suggest that other regulators, in addition to Xyr1 and Are1, may also be responsible for *trasp*. Several studies have demonstrated that some transcription factors, such as PrtT, XprG, CreA and Are1, are linked with regulation of proteases in *Aspergillus* (Wilson and Arst 1998; Katz et al. 2006; Katz et al. 2008; Punt et al. 2008). A few studies have also investigated *Trichoderma* (Delgado-Jarana et al. 2002; Kredics et al. 2005). Delgado-Jarana et al. reported that aspartyl protease (PapA) from *T. harzianum* has the potential to regulate AreA and Pac1 transcription factors (Delgado-Jarana et al. 2002). Further analysis of known binding motifs revealed the presence of Ace1, Cre1 and Xpp1, as well as the global regulator Pac1, a pH regulator, motifs in the promoter region of *trasp* (Fig. 3.10). From protease activity of transcription factor deletants, we found that Cre1, Xpp1 and Pac1 transcription factors behave as repressors of *trasp* in both Avicel and galactose (Fig. 3.11), and that the activity of Ace1 was not altered as compared to QM9414. Similarly, with regard to the cellulase regulation mechanisms, Cre1, Xpp1 and Pac1 (at an acidic pH) acted as repressors in *T. reesei.* Poussereau et al. (2001) suggested that the acid protease *acp1* can be regulated by CreA and PacC in the fungus *Sclerotinia sclerotiorum* (Poussereau et al. 2001). However, we noted a different pattern of expression of endopeptidase under regulation of Xpp1, as compared with other regulators, in Avicel and galactose. These findings suggest that Xpp1

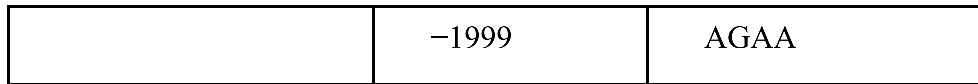
differentially controls other endopeptidases. A time series analysis of *trasp* expression in Avicel and galactose demonstrated that expression of TrAsP depends on the time of induction and the influence of the transcription factor (Fig. 3.11c and Fig. 3.11d). In Avicel, significant upregulation was observed in Δ*cre1* and Δ*xpp1* mutations with cultivation during latter time periods. Upregulation of Cre1, Xpp1 and Pac1 mutants demonstrates the strong repression properties of *trasp* at 16 and 24 hours in galactose. We hypothesise that these results strongly suggest that TrAsP expression depends on the composition of the culture media as well as on the induction time.

Protease regulators may regulate *trasp* expression, although specific protease regulators have not been identified in *T. reesei*. We could not identify binding motifs for the known protease regulators, PrtT (Punt et al. 2008) and McmA (Li et al. 2016), upstream of the *trasp* promoter region. However, we did find the binding motif of FlbC (−803) of the *trasp* promoter (Kwon et al. 2010). Tanaka et al. (2016) reported that the FlbC transcription factor activates the acid protease *pepA* in *A. oryzae* (Tanaka et al. 2016). Kwon et al. (2010) reported that the FlbC transcription factor is necessary for proper activation of conidiation and growth and development of *A. nidulans* and identified its binding motif as TGACGA[T/A] (Kwon et al. 2010). Therefore, FlbC may also affect TrAsP expression.

These findings suggest that TrAsP regulation mechanisms in response to carbon sources is complex. However, in this study, we identified that the region between –880 and – 670 was crucial for TrAsP expression and that Are1 may play a major role in its expression. The region between – 670 and – 460 also elicited an interesting pattern in promoter strength of *trasp*, suggesting that novel protease repressors may exit in the region. Similarly, we hypothesised that Xpp1 in this region may have string repression activity for proteases. Furthermore, we suggest that the promoter region beyond the position at −1080 may contribute to promoter strength. However, we could not identify more Are1 binding motifs between the −1080 and −2000 regions (Fig. 3.12). Therefore, a novel protease activator likely exists beyond −1080 that may contribute to the increased promoter strength.



Transcription factor (Binding motif)	Position in the promoter	Binding motif
♦ Xyr1 (GGCWWW)	−1674	TTAGCC
□ Cre1 (SYGGRG)	−1306	CCCCAG
	−1580	CTGGAG
	−1761	CTCCGG
	−1839	CTCCAG
△ Ace1(AGGCA	−1095	AGGCA
	−1115	AGGCA
	−1740	AGGCA
	−1772	TGCCT
	−1969	AGGCA
▲ Pac1 (GCCARG)	−1127	GCCAGG
	−1385	GCCAAG
	−1599	GCCAAG
	−1631	GCCAAG
● AreA(HGATAR)	−1617	CTATCG
	−1956	TTATCG
○ AreA (GATA)	−1904	GATA
◇ Xpp1 (AGAA)	−1154	TTCTTCT
	−1302	AGAA
	−1419	TTCT
	−1509	TTCT
	−1707	AGAA



	−1999	AGAA

Fig. 3.12 Schematic representation of the distribution of the putative binding sequences of known cellulase regulators in the TrAsP promoter region (between 1080-2000bp).

3. 5 References

Aro N, Ilmén M, Saloheimo A, Penttilä M (2003) ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. Appl Environ Microbiol 69:56–65 . doi: 10.1128/AEM.69.1.56-65.2003.

Aro N, Saloheimo A, Ilmé M, Penttilä M (2001) ACEII, a Novel Transcriptional Activator Involved in Regulation of Cellulase and Xylanase Genes of *Trichoderma reesei*. 276:24309–24314 . doi: 10.1074/jbc.M003624200.

Bugeja HE, Hynes MJ, Andrianopoulos A (2012) AreA controls nitrogen source utilisation during both growth programs of the dimorphic fungus *Penicillium marneffei*. Fungal Biol 116:145–154 . doi: 10.1016/j.funbio.2011.10.009

Callow N V., Ray CS, Kelbly MA, Ju L-K (2016) Nutrient control for stationary phase cellulase production in *Trichoderma reesei* Rut C-30. Enzyme Microb Technol 82:8–14 . doi: 10.1016/J.ENZMICTEC.2015.08.012.

Christensen T, Hynes MJ, Davis MA (1998) Role of the regulatory gene *areA* of *Aspergillus oryzae* in nitrogen metabolism. Appl Environ Microbiol 64:3232-3237.

Daranagama DN, Koki S, Masahiro Y, Haruna S, Yuki O (2019) Proteolytic analysis of *Trichoderma reesei* in celluase - inducing condition reveals a role for trichodermapepsin (TrAsP) in cellulase production. J Ind Microbiol Biotechnol 46:831–842 . doi: 10.1007/s10295-019-02155-9.

Delgado-Jarana J, Rincón AM, Benítez T (2002) Aspartyl protease from *Trichoderma harzianum* CECT 2413: Cloning and characterization. Microbiology 148:1305–1315 . doi: 10.1099/00221287-148-5-1305.

Derntl C, Rassinger A, Srebotnik E, Mach RL, Mach-Aigner AR (2015) Xpp1 regulates the expression of xylanases, but not of cellulases in *Trichoderma reesei*. Biotechnol Biofuels 8:112 . doi: 10.1186/s13068-015-0298-8.

dos Santos Castro L, de Paula RG, Antoniêto ACC, Persinoti GF, Silva-Rocha R, Silva RN (2016) Understanding the Role of the Master Regulator XYR1 in *Trichoderma reesei* by Global Transcriptional Analysis. Front Microbiol 7:1–16 . doi: 10.3389/fmicb.2016.00175.

Dunkel N, Biswas K, Hiller E, Fellenberg K, Satheesh S V., Rupp S, Morschhäuser J (2014) Control of morphogenesis, protease secretion and gene expression in *candida albicans* by the preferred nitrogen source ammonium. Microbiology 160:1599–1608 . doi:

10.1099/mic.0.078238-0

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem 278:31988–31997 . doi: 10.1074/jbc.M304750200.

Furukawa T, Shida Y, Kitagami N, Mori K, Kato M, Kobayashi T, Okada H, Ogasawara W, Morikawa Y (2009) Identification of specific binding sites for XYR1, a transcriptional activator of cellulolytic and xylanolytic genes in *Trichoderma reesei*. Fungal Genet Biol 46:564–574 . doi: 10.1016/j.fgb.2009.04.001.

Furukawa T, Shida Y, Kitagami N, Ota Y, Adachi M, Nakagawa S, Shimada R, Kato M, Kobayashi T, Okada H, Ogasawara W, Morikawa Y (2008) Identification of the cis-acting elements involved in regulation of xylanase III gene expression in *Trichoderma reesei* PC-3-7. Fungal Genet Biol 45:1094–1102 . doi: 10.1016/j.fgb.2008.03.006

Häkkinen M, Valkonen MJ, Westerholm-Parvinen A, Aro N, Arvas M, Vitikainen M, Penttilä M, Saloheimo M, Pakula TM (2014) Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. Biotechnol Biofuels 7:14 . doi: 10.1186/1754-6834-7-14.

Hirasawa H, Shioya K, Mori K, Tashiro K, Aburatani S, Shida Y, Kuhara S, Ogasawara W (2019) Cellulase productivity of *Trichoderma reesei* mutants developed in Japan varies with varying pH conditions. J Biosci Bioeng 128:264–273 . doi: 10.1016/j.jbiosc.2019.03.005.

Hynes MJ (1974) Effects of ammonium, L-glutamate, and L-glutamine on nitrogen catabolism in *Aspergillus nidulans*. J Bacteriol 120:1116–1123

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Report 5:387–405 . doi: 10.1007/BF02667740.

Katz ME, Bernardo SM, Cheetham BF (2008) The interaction of induction, repression and starvation in the regulation of extracellular proteases in *Aspergillus nidulans*: Evidence for a role for CreA in the response to carbon starvation. Curr Genet 54:47–55 . doi: 10.1007/s00294-008-0198-6

Katz ME, Gray KA, Cheetham BF (2006) The *Aspergillus nidulans xprG (phoG)* gene encodes a putative transcriptional activator involved in the response to nutrient limitation. Fungal Genet Biol 43:190–199 . doi: 10.1016/j.fgb.2005.12.001

Klaubauf S, Narang HM, Post H, Zhou M, Brunner K, Mach-Aigner AR, Mach RL, Heck AJR, Altelaar AFM, de Vries RP (2014) Similar is not the same: Differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi. Fungal Genet Biol 72:73–81 . doi: 10.1016/j.fgb.2014.07.007.

Kredics L, Antal Z, Szekeres A, Hatvani L, Manczinger L, Vágvölgyi C, Nagy E (2005) Extracellular Proteases of *Trichoderma Species*. Acta Microbiol Immunol Hung 52:169–184 . doi: 10.1556/AMicr.52.2005.2.3

Kwon NJ, Garzia A, Espeso EA, Ugalde U, Yu JH (2010) FlbC is a putative nuclear C_2H_2 transcription factor regulating development in *Aspergillus nidulans*. Mol Microbiol 77:1203–1219 . doi: 10.1111/j.1365-2958.2010.07282.x.

Lamb HK, Dodds AL, Swatman DR, Cairns E, Hawkins AR (1997) Deletion of the N-terminal region of the AREA protein is correlated with a derepressed phenotype with respect to nitrogen metabolite repression. J Bacteriol 179:6649–6656 . doi: 10.1128/jb.179.21.6649-6656.1997

Landowski CP, Huuskonen A, Wahl R, Westerholm-Parvinen A, Kanerva A, Hänninen AL, Salovuori N, Penttilä M, Natunen J, Ostermeier C, Helk B, Saarinen J, Saloheimo M (2015) Enabling low cost biopharmaceuticals: A systematic approach to delete proteases from a well-known protein production host *trichoderma reesei*. PLoS One 10:1–28 . doi: 10.1371/journal.pone.0134723.

Li N, Kunitake E, Endo Y, Aoyama M, Kanamaru K, Kimura M, Kato M, Kobayashi T (2016) Involvement of an SRF-MADS protein McmA in regulation of extracellular enzyme production and asexual/sexual development in *Aspergillus nidulans*. Biosci Biotechnol Biochem 80:1820–1828 . doi: 10.1080/09168451.2016.1146074.

Lockington RA, Rodbourn L, Barnett S, Carter CJ, Kelly JA (2002) Regulation by carbon and nitrogen sources of a family of cellulases in *Aspergillus nidulans*. Fungal Genet Biol 37:190–196 . doi: 10.1016/S1087-1845(02)00504-2.

Ma L, Chen L, Zhang L, Zou G, Liu R, Jiang Y, Zhou Z (2016) RNA Sequencing Reveals Xyr1 as a Transcription Factor Regulating Gene Expression beyond Carbohydrate Metabolism. Biomed Res Int 2016:4841756 . doi: 10.1155/2016/4841756.

Mach-Aigner AR, Pucher ME, Steiger MG, Bauer GE, Preis SJ, Mach RL (2008) Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. Appl Env Microbiol 74:6554–6562 . doi: 10.1128/AEM.01143-08.

Muro-Pastor M, Gonzalez R, J S, Narendja F, Scazzocchio C (2017) The GATA factor AreA

is essential for chromatin remodelling in a eukaryotic bidirectional promoter. EMBO J 18:1–14 . doi: 10.1093/emboj/18.6.1584.

Nitta M, Furukawa T, Shida Y, Mori K, Kuhara S, Morikawa Y, Ogasawara W (2012) A new Zn(II)2Cys$_6$-type transcription factor BglR regulates ß-glucosidase expression in Trichoderma reesei. Fungal Genet Biol 49:388–397 . doi: 10.1016/j.fgb.2012.02.009.

Ogasawara W, Shida Y, Furukawa T, Shimada R, Nakagawa S, Kawamura M, Yagyu T, Kosuge A, Xu J, Nogawa M, Okada H, Morikawa Y (2006) Cloning, functional expression and promoter analysis of xylanase III gene from *Trichoderma reesei*. Appl Microbiol Biotechnol 72:995–1003 . doi: 10.1007/s00253-006-0365-y

Penttila M, Nevalainen H, Ratto M, Salminen E, Knowles J (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61:155–164 . doi: 10.1016/0378-1119(87)90110-7.

Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, Hartl L, Karaffa L, Druzhinina IS, Seiboth B, Le Crom S, Kubicek CP, Sándor E, Hartl L, Karaffa L, Druzhinina IS, Seiboth B, Le Crom S, Kubicek CP (2011) The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. BMC Genomics 12:269 . doi: 10.1186/1471-2164-12-269.

Poussereau N, Creton S, Billon-Grand G, Rascle C, Fevre M (2001) Regulation of *acp1*, encoding a non-aspartyl acid protease expressed during pathogenesis of *Sclerotinia sclerotiorum*. Microbiology 147:717–726 . doi: 10.1099/00221287-147-3-717

Punt PJ, Schuren FHJ, Lehmbeck J, Christensen T, Hjort C, van den Hondel CAMJJ (2008) Characterization of the *Aspergillus niger* prtT, a unique regulator of extracellular protease encoding genes. Fungal Genet Biol 45:1591–1599 . doi: 10.1016/j.fgb.2008.09.007

Qian Y, Sun Y, Zhong L, Sun N, Sheng Y, Qu Y, Zhong Y (2019) The GATA-Type Transcriptional Factor Are1 Modulates the Expression of Extracellular Proteases and Cellulases in *Trichoderma reesei*. Int J Mol Sci 20:4100 . doi: 10.3390/ijms20174100.

Rauscher R, Wurleitner E, Wacenovsky C, Aro N, Stricker AR, Zeilinger S, Kubicek CP, Penttila M, Mach RL (2006) Transcriptional Regulation of xyn1, Encoding Xylanase I, in *Hypocrea jecorina*. Eukaryot Cell 5:447–456 . doi: 10.1128/EC.5.3.447-456.2006.

Ravagnani A, Gorfinkiel L, Langdon T, Diallinas G, Adjadj E, Demais S, Gorton D, Arst HN, Scazzocchio C (1997) Subtle hydrophobic interactions between the seventh residue of the zinc finger loop and the first base of an HGATAR sequence determine promoter-specific recognition by the Asp*ergillus nidulans* GATA factor AreA. EMBO J 16:3974–

3986 . doi: 10.1093/emboj/16.13.3974.

Shida Y, Furukawa T, Ogasawara W (2016) Deciphering the molecular mechanisms behind cellulase production in *Trichoderma reesei* , the hyper-cellulolytic filamentous fungus. Biosci Biotechnol Biochem 8451:1347–6947 . doi: 10.1080/09168451.2016.1171701.

Shida Y, Furukawa T, Ogasawara W, Kato M, Kobayashi T, Okada H, Morikawa Y (2008) Functional analysis of the egl3 upstream region in filamentous fungus *Trichoderma reesei*. Appl Microbiol Biotechnol 78:515–524 . doi: 10.1007/s00253-007-1338-5.

Shida Y, Yamaguchi K, Nitta M, Nakamura A, Takahashi M, Kidokoro S, Mori K, Tashiro K, Kuhara S, Matsuzawa T, Yaoi K, Sakamoto Y, Tanaka N, Morikawa Y, Ogasawara W (2015) The impact of a single-nucleotide mutation of bgl2 on cellulase induction in a *Trichoderma reesei* mutant. Biotechnol Biofuels 8:230 . doi: 10.1186/s13068-015-0420-y.

Snyman C, Theron L, Divol B (2019) Understanding the regulation of extracellular protease gene expression in fungi: a key step towards their biotechnological applications. 103:5517–5532 . doi: doi.org/10.1007/s00253-019-09902-z.

Stricker AR, Grosstessner-Hain K, Würleitner E, Mach RL (2006) Xyr1 (Xylanase Regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina* . Eukaryot Cell 5:2128–2137 . doi: 10.1128/EC.00211-06.

Stricker AR, Trefflinger P, Aro N, Penttilä M, Mach RL (2008) Role of Ace2 (Activator of Cellulases 2) within the xyn2 transcriptosome of *Hypocrea jecorina*. Fungal Genet Biol 45:436–445 . doi: 10.1016/j.fgb.2007.08.005.

Suárez MB, Sanz L, Chamorro MI, Rey M, González FJ, Llobell A, Monte E (2005) Proteomic analysis of secreted proteins from *Trichoderma harzianum*: Identification of a fungal cell wall-induced aspartic protease. Fungal Genet Biol 42:924–934 . doi: 10.1016/J.FGB.2005.08.002

Tanaka M, Yoshimura M, Ogawa M, Koyama Y, Shintani T, Gomi K (2016) The C_2H_2-type transcription factor, FlbC, is involved in the transcriptional regulation of Aspergillus oryzae glucoamylase and protease genes specifically expressed in solid-state culture. Appl Microbiol Biotechnol 100:5859–5868 . doi: 10.1007/s00253-016-7419-6.

Wilson RA, Arst HN (1998) Mutational analysis of AREA, a transcriptional activator mediating nitrogen metabolite repression in *Aspergillus nidulans* and a member of the “streetwise” GATA family of transcription factors. Microbiol Mol Biol Rev 62:586-596.

Chapter 4: General Conclusions

In order to reach the production targets of the biofuel industry, novel technologies should be developed. *T. reesei* is an industrial cellulase producer and has been the focus of various genetic and biochemical studies for low cost and efficient enzyme production. This study enhances the knowledge about protease regulation in *T. reesei,* which may directly affect cellulase production. Therefore, this study aimed to analyse gene regulation of Trichodermapepsin (TrAsP) in *T. reesei.*

The proteolytic system in *T. reesei* is primarily composed of aspartic, serine, cysteine and metallopoteases and plays a key role in protein degradation. Among these, aspartic acid proteases are also involved in cellulase degradation in the secretome. These are pepsin-type proteases and there are nearly 15 aspartic acid protease genes in the genome. However, none of these genes are well characterised. In this study we characterised TrAsP and elucidated the mechanisms of its production and regulation. According to our findings, pepsin proteases are induced by monosaccharides, especially, galactose, and organic nitrogen, but not by the typical inducers of cellulase, such as Avicel. Pepsins are regulated by pH and organic nitrogen and are involved in cellulase degradation. Degraded or modified cellulase is not useful for plant biomass degradation. Therefore, an understanding of these processes is essential for efficient use of cellulase.

Considering the physiological role of TrAsP, we propose an enzyme secretion regulatory model for *T. reesei* QM9414 (Fig. 4.1).

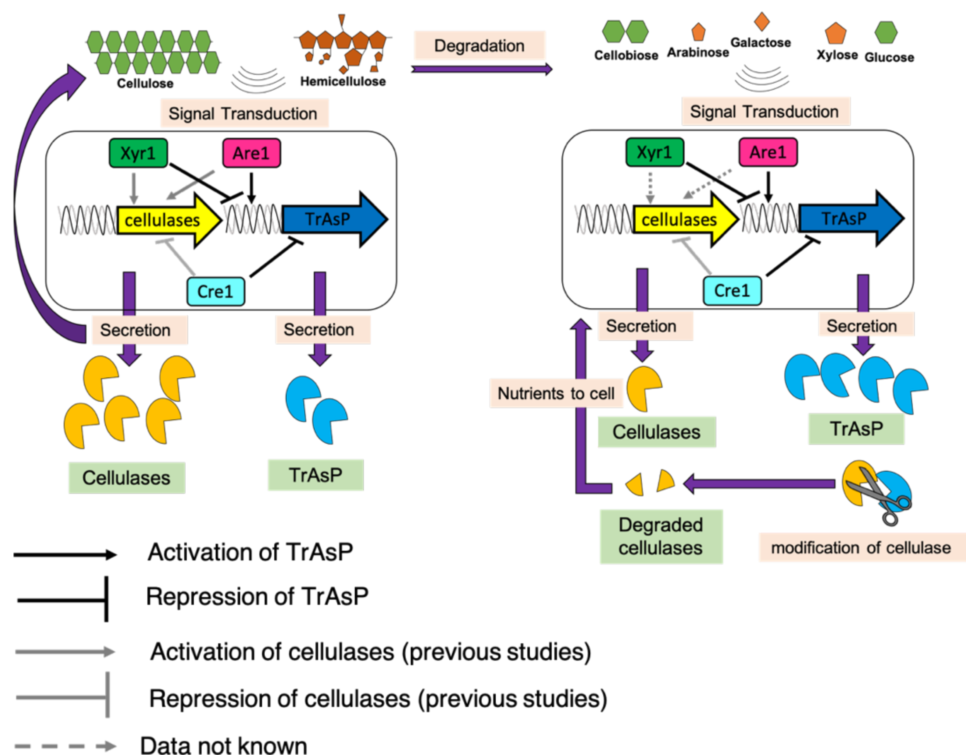


Fig. 4.1 A hypothetical model for an enzyme secretion regulatory mechanism of *T. reesei* QM9414.

Under cellulase-inducing conditions, *T. reesei* produces a large amount of cellulases and a very small amount of pepsins. Xyr1 and Are1 activate cellulase expression, while Cre1 represses cellulase expression. Are1 activates TrAsP expression, while Xyr1 and Cre1 repress its expression. A large amount of cellulases are involved in the degradation of cellulose and hemicellulose. Degraded products include disaccharides and monosaccharides, such as cellobiose, arabinose, galactose, xylose and glucose. These monosaccharides induce a large amount of TrAsP production in the late stage. Under carbon catabolite expression, cellulase secretion is reduced, and a nutrient depleted condition is created. In this situation, TrAsP degrades cellulases in the media allowing *T. reesei* to take up nutrients that are essential for survival in later stages of the cultivation process, and thus exerts a physiological role. However, other proteases, in addition to TrAsP, also play a similar role when nutrients are depleted in the medium and serve to enhance the chance of survival.

Other proteases may also work under this mechanism. Microarray analysis assessing all secreted proteases under galactose induction conditions revelled that some carboxypeptidase proteases (protein ID 22210, 120998) have a similar expression pattern to TrAsP (Fig.4.2). We also noted that some serine proteases were highly expressed in cellulase-inducing conditions (protein ID 103039). In addition, some serine proteases were highly expressed in both cellulase-inducing and non-inducing conditions (Protein ID 123244). Therefore, more extensive studies are necessary to develop a better understanding the proteolytic mechanisms of *T. reesei.*

The findings of our study emphasise that the deletion of *trasp* increases expression of cellulases and some other proteases. Therefore, deletion of these genes in industrial strains may increase the amount of cellulase production. Similarly, altering the composition of the medium can contribute to optimal cellulase condition with less protease secretion.

Future perspectives involve investigation of the proteolytic system in *T. reesei*, as it remains elusive. In this study we characterised TrAsP and developed a model of its regulation. This model may be applicable to other proteases. Furthermore, many aspects of TrAsP regulation with different carbon sources and nitrogen sources remain unknown. With respect to *trasp* expression, future studies should involve a detailed analysis of transcription factors possessing binding sites in the promoter region. Future studies should also aim to identify transporters that are involved in the nitrogen assimilation pathway and sugar uptake in order to develop a better understanding of protease regulation as compared to cellulase regulation. More research is necessary to identify specific protease regulators for *T. reesei* and to unravel the mechanisms of protease regulation. The present analysis of the regulatory mechanisms of

the main protease in *T.reesei* improves our knowledge of this field and may be useful in industrial applications.

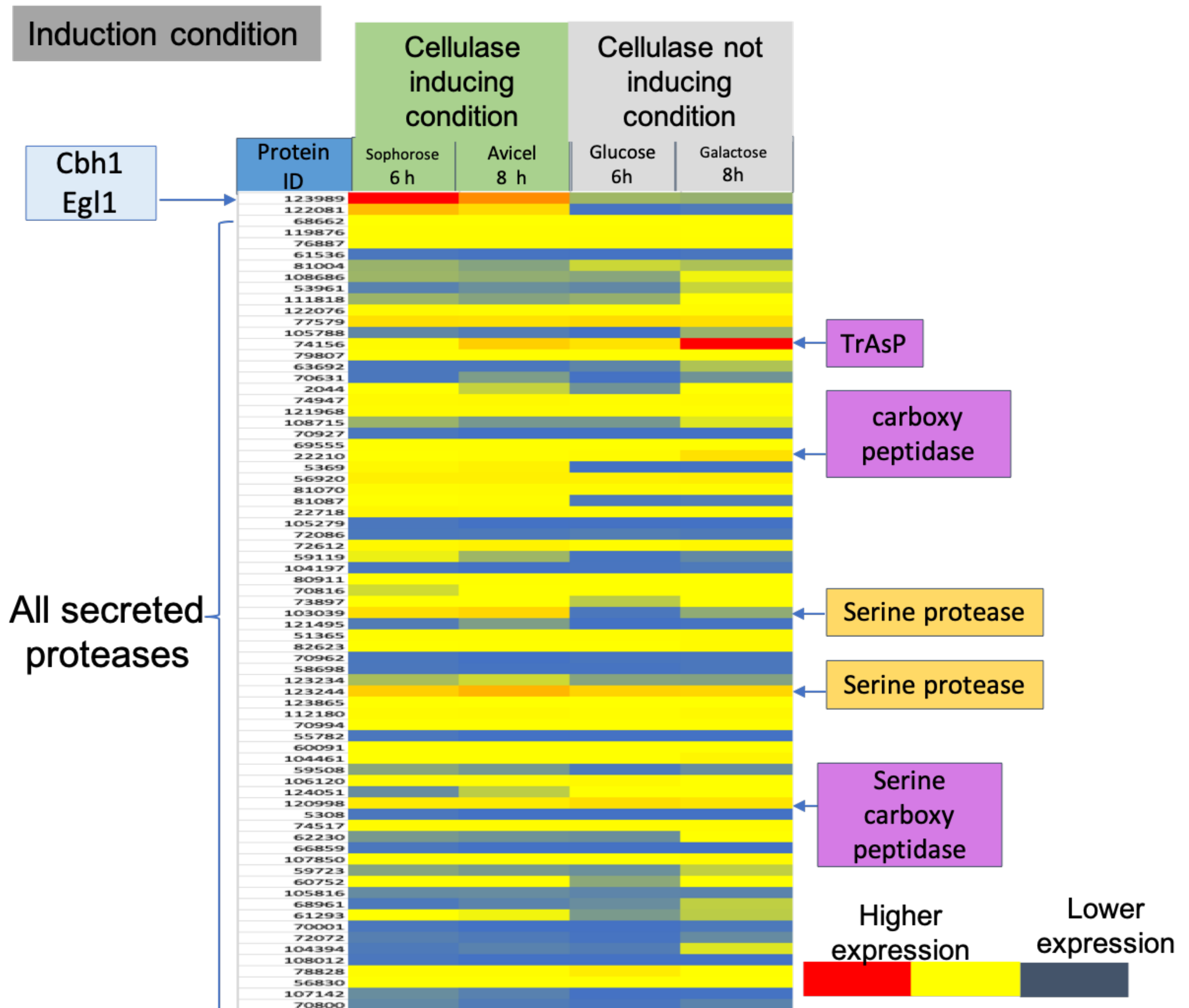


Fig. 4.2 Micro array analysis using all secreted proteases.

Chapter 5: Publications

5.1 Journal Papers

1. **Daranagama N.D., Shioya K., Yuki M, Satao H, Ohataki Y, Suzuki Y, Shida Y. Ogasawara W** (2019) Proteolytic analysis of Trichoderma reesei in cellulase-inducing condition reveals a role for trichodermapepsin (TrAsP) in cellulase production. Journal of industrial microbiology & biotechnology (46) 6.831-842
2. **Daranagama N.D., Suzuki Y, Shida Y. Ogasawara W.** (2020) Involvement of Xyr1 and Are1 for Trichodermapepsin gene expression in response to cellulose and galactose in *Trichoderama reesei.* Current microbiology (in press)

5.2 Conference Proceedings

Daranagama N.D., Aita H., Hirasawa H., Shioya K., Satao H., Suzuki Y, Shida Y. Ogasawara W. (2016) A novel avenue of protease producing mechanism in Trichoderma reesei Proceedings of the conference on Biomass Sceince.11 (0) 99-100

Chapter 6: Acknowledgments

I wish to express my sincere gratitude to my supervisor Professor Wataru Ogasawara and Assistant Professor Yosuke Shida for their guidance throughout my time in this doctoral programme. My special thanks to the Japanese government MEXT scholarship programme for providing me with financial support.

I would like to thank Dr. Koki Shioya, Ms. Haruna Sato and Mr. Yuki Ohtaki who started this research and provided me with the background knowledge to organise my research. A special thank you to Assistant Professor Yoshiyuki Suzuki in National Institute of Technology for the support throughout the study, as well as for assistance with the experimental designs, data analysis and paper publication process. I would also like to thank Dr. Hiroki Hirasawa for sharing his experience and knowledge with me. Furthermore, I would like to thank my senior laboratory members (especially Mr. Hiroki Aita) for teaching me lab techniques and providing me useful advice for success in my experiments. I would also like to thank all of the other laboratory members and research assistants for their support and encouragement.

Finally, I would like to acknowledge the heart-felt love from my parents and brother. Thank you to my husband for his continual support, patience, faith, kindness and love and my baby for the laughter provided during this difficult time.