TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Characterization of a mutation in the transactivator Xyr1 of the *Trichoderma reesei* wild-type strain

Ausgeführt am Institut für Verfahrenstechnik, Umwelttechnik und Technische Biowissenschaften der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Mag. Dr.rer.nat. Robert Mach

Dipl.-Ing. Dr.techn. Astrid Mach-Aigner

durch

Katharina Regnat

Radetzkystraße 23/6

1030 Wien

Wien, 19.11.2015

TECHNISCHE UNIVERSITÄT WIEN

Abstract

by Katharina Regnat

The filamentous and saprophytic fungus *Trichoderma reesei* is a widely used microorganism in both research and industry. Its cellulose and hemicellulose degrading enzymes are applied in the pulp, paper, food, feed, textile industry and perhaps most importantly, in the production of second generation biofuels. Many *T. reesei* strains used in industry are descendants from the strain RUT-C30, in which the carbon catabolite repression is eliminated. Despite this, inducing substances are still needed to ensure a satisfactory protein formation. In one of these strains, a single point mutation in the fungal transcription factor middle homology region (FTFMHR) of the xylanase regulator 1 (Xyr1), leads to an expression of enzymes XynI and XynII in the presence of glucose, a normally repressing carbon source. Investigations show that under normal circumstances the carbon source triggers a conformational change of Xyr1 resulting in its activation or inactivation. Circular dichroism (CD) analyses of the mutated protein Xyr1$_{A824V}$ suggest that an altered secondary structure could hinder conformational change and thus cause the missing response of the protein in the presence of a carbon source.
The aim of this work is to characterize the mutated strain and to identify and characterize the region around the mutation by looking for changes in the secondary structure of different Xyr1 and Xyr1$_{A824V}$ protein fragments. As a main result a deregulation and strongly enhanced expression of the xylanase-encoding genes in the mutated strain, not resulting from a change in chromatin status, was revealed experimentally. Of nine Xyr1 and Xyr1$_{A824V}$ protein fragments, which were produced in this work, non showed a response to carbohydrates and only one showed a change in secondary structure. The latter also suggests that additionally to the mutation in the FTFMHR, the putative activation domain of Xyr1 is responsible for the previously observed changes in secondary structure. Analyses of the tertiary structure support this hypothesis.

TECHNISCHE UNIVERSITÄT WIEN

Kurzfassung

von Katharina Regnat

Der fadenförmige und saprophytische Pilz *Trichoderma reesei* ist ein in Forschung und Industrie weitflächig genutzter Mikroorganismus. Die Cellulose und Hemicellulose abbauenden Enzyme von *T. reesei* werden in der Zellstoff-, Papier-, Nahrungsmittel-, Futtermittel- und Textilindustrie eingesetzt und ist von großer Bedeutung für die Herstellung von Biokraftstoffen der zweiten Generation. Viele *T. reesei* Stämme, die in der Industrie verwendet werden, sind Abkömmlinge des Stammes RUT-C30, bei dem keine Kohlenstoff-Katabolitrepression mehr stattfindet. Trotzdem sind nach wie vor induzierende Substanzen für eine zufriedenstellend starke Enzymproduktion nötig. In einem dieser Stämme führt eine einzelne Punktmutation in der "fungal transcription factor middle homology region" (FTFMHR) des Xylanase-Regulators 1 (Xyr1) zu einer erhöhten Expression der Enzyme XynI und XynII in Gegenwart von Glucose, einer normalerweise reprimierenden Kohlenstoffquelle. Untersuchungen zeigen, dass unter normalen Umständen Xyr1 von der vorliegenden Kohlenstoffquelle ein Signal erhält, das zu einer Konformationsänderung und damit zu einer Aktivierung oder Inaktivierung des Proteins führt. Zirkulardichroismus Analysen des mutierten Proteins $Xyr1_{A824V}$ legen nahe, dass eine veränderte Sekundärstruktur durch die Mutation der Grund für eine fehlende Konformationsänderung und damit einer fehlenden Antwort auf die Gegenwart der Kohlenstoffquelle sein könnte.

Das Ziel dieser Arbeit ist es einersteits den mutierten Pilzstamm zu charakterisieren und andererseits durch etwaige Veränderungen der Sekundärstruktur verschiedener Xyr1 und $Xyr1_{A824V}$ Proteinfragmente die Region der Mutation zu untersuchen. Die Experimente zeigen eine Deregulierung und eine stark erhöhte Expression der Xylanase-kodierenden Gene in dem mutierten Stamm, die nicht auf einen veränderten Chromatin-Status zurückzuführen sind. Die neun für diese Arbeit entwickelten und hergestellten Xyr1 und $Xyr1_{A824V}$ Proteinfragmente zeigen keinerlei Reaktion auf Kohlenhydrate und nur eines der Fragmente zeigt eine Änderung der Sekundärstruktur. Dies lässt vermuten, dass zusätzlich zu der Mutation in der FTFMHR, die mutmaßliche Aktivierungsdomäne von Xyr1 für die zuvor beobachteten Veränderungen in der Sekundärstruktur verantwortlich ist. Die Analyse der Tertiärstruktur eines Proteinfragmentes stützt diese Hypothese.

Acknowledgements

First of all I want to say thanks to Thiago Machado Mello De Sousa for his excellent supervision during the six months of my lab work and the great help during the writing process afterwards. I am glad for learning this much from him.

I am particularly thankful to Prof. Robert Mach and Dr. Astrid Mach-Aigner for letting me be a part of a very exciting project in a friendly, helpful and motivated working group. I always felt welcome and perfectly supervised. For this, I also want to thank the whole gene technology working group with Thiago Machado Mello De Sousa, Dr. Christian Derntl, Alice Rassinger, Petra Till and Daniel Kiesenhofer.

I also like to express my gratitude to my family, especially my parents, without whom my studies would not have been possible.

Finally, I want to say thanks to my boyfriend Max, who always listened and supported me.

Contents

Chapter 1

Introduction

1.1 *Trichoderma*

The very first description of a fungus named *Trichoderma* was in the year 1749. Over 100 years later, in 1865, a link to the sexual state of *Hypocrea* was suggested (Kuhls et al., 1996). Nowadays there are more than 100 phylogenetically described *Trichoderma* species (Schuster and Schmoll, 2010). One of these is the potent enzyme producer *Trichoderma reesei*. This strain was originally isolated from the Solomon islands in World War II, where it destroyed canvas and other cellulose-containing material of the US army (Druzhinina et al., 2010). All strains used nowadays in biotechnology and in basic research are descendants from this one isolate (Seidl et al., 2009). *T. reesei* is a filamentous and saprophytic fungus and easy to cultivate. Because of its good handling and due to its high cellulase and hemicellulase production it is a model organism in research and widely used in the industrial field.

1.1.1 *T. reesei*: history of strain development

The isolate from the Solomon islands was first identified as *Trichoderma viride* and named QM6a as part of the collection at the US Army QuarterMaster Research. QM6a was later identified as a new species and named *T. reesei* in honor of Elwyn T. Reese (Peterson and Nevalainen, 2012) (Seidl et al., 2009). With regard to the need of alternative fuel sources and because *T. reesei* is a good producer of cellulases for the production of cellulose bioethanol, several strain development programs started. Random mutagenesis was used to isolate mutants released from carbon catabolite repression (CCR) with high cellulase and hemicellulase production. CCR is a mechanism wherein a carbon catabolite represses the expression of different genes needed for the

processing of carbohydrates. A promising mutant, QM9414, was found in the Natick laboratories and had a two to four times higher protein and cellulase production than QM6a. But the mutant was still subject to CCR (Peterson and Nevalainen, 2012). This inconvenience could be avoided in a mutant achieved at Rutgers University in New Jersey. The hypercellulolytic and catabolite-derepressed strain RUT-C30 was obtained through a three step process over several other mutants (Montenecourt and Eveleigh, 1977) (Montenecourt and Eveleigh, 1979). First, mutagenesis by UV light and screening for catabolite-derepression led to the mutant M7 (Montenecourt and Eveleigh, 1977). Another mutagenesis by N-nitroguanidine led to the isolation of NG14. Out of this mutant, RUT-C30 resulted from another UV mutagenesis and screening for high cellulase production and catabolite-repression (Montenecourt and Eveleigh, 1979). RUT-C30 produces more protein than QM6a with a cellulase activity nearly three times higher. Also RUT-C30 outperformed a high cellulase-producing mutant MC677 and its parent strain QM9414 (Peterson and Nevalainen, 2012). The history of stain development is shown in figure 1.1.

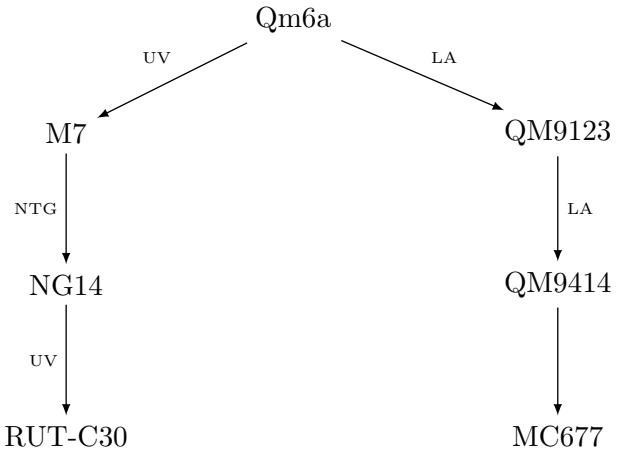

FIGURE 1.1: History of strain development

1.1.2 Applications of *T. reesei* enzymes

With regard to more and more fading resources of fossil fuels, increasing energy costs and the climate change, research is focused on new concepts to gain a sustainable economy. Important applications for substrates from *T. reesei* are in the food and feed, pulp, paper and textile industry. But especially the production of biofuels from waste material is in the focus of today's research.

1.1.2.1 Textile industry

The first success for cellulases produced in *T. reesei* worldwide was surprisingly achieved in the textile industry. The textile industry benefited from the kinetic properties of the enzymes, which allow a controlled treatment of cellulosic fibers (de Lourdes T.M. Polizeli and Rai, 2014). Although the prospect of using cellulases for textile product improvement was introduced only two decades ago, it became the third largest and a very fast growing market for industrial enzymes. Due to cellulases, new ideas for the treatment and the modification of cellulosic fibers arose (Galante et al., 1998a). The best known textile applications are biostoning and biopolishing (de Lourdes T.M. Polizeli and Rai, 2014). In the early 1980s, worn and aged looking jeans were obtained by washing the jeans in industrial laundries with pumice stones. This process was called "stone-washing" and caused a few rough problems. In addition to a huge amount of pumice waste and second quality garments, the washing machines got rapidly worn out. Biostoning via microbial cellulases offered the perfect alternative to the traditional stone-washing with pumice (Galante et al., 1998a). Here the cellulases act on the cotton fabric and break off the small fiber ends in the yarn surface, thereby loosening the indigo dye, which is then easily removed by mechanical abrasion in the wash cycle. Biopolishing is a step in the textile industry for producing high quality garments. The cellulases serve to reduce the tendency for "fuzz" and "pill" formation by removing small fibers from the fabric surface. Microbial cellulases are also added to common laundry detergents to improve the appearance and color brightness. (de Lourdes T.M. Polizeli and Rai, 2014)

1.1.2.2 Food and feed industry

The so called macerating enzymes (cellulases, xylanases and pectinases) are important for the extraction and clarification process. They are especially used for the production of fruit juices and the production and conservation of fruit nectars and purees (de Lourdes T.M. Polizeli and Rai, 2014). In wine production, enzymes like glucanases, hemicellulases and pectinases play an important role by improving color extraction and skin maceration, facilitating wort clarification and filtration rate. They therefore improve the quality and stability of wine (Galante et al., 1998b). In the beer brewing process the microbial enzymes are mainly used for the hydrolyzation of β-glucans resulting from non-starch polysaccharides in poor quality and unmalted barley or other cereals. These β-glucans form gels during the brewing process and lead to poor filtration of the wort, slow run-off times, low extract yields and the formation of haze in the final product (de Lourdes T.M. Polizeli and Rai, 2014). Enzymes used in the feed industry are commonly hydrolases used as additives to achieve either the elimination of anti-nutritional factors or the degradation of several cereal to decrease their viscosity and improve their

nutritional value. The enzymes can also be used as supplements to the animals own digestive enzymes (de Lourdes T.M. Polizeli and Rai, 2014) (Galante et al., 1998b).

1.1.2.3 Pulp and paper industry

The main application for enzymes in this industrial branch is the bleaching of pulp. The enzymes are a good alternative to chlorine-containing chemicals that are toxic and pollutants to the environment. In the Kraft process the wood chips are being cooked for the degradation and solubilization of lignin. But the lignin can not be fully dissolved, leading to brown pulp. To obtain a better brightness the pulp would originally be bleached with chlorine. But if the hemicellulose is hydrolyzed by hemicellulases (xylanases), the lignin is easier to extract and therefore less chlorine is needed.

1.1.2.4 Bioethanol industry

The production of second generation ethanol involves several steps: selection of a proper feedstock and its pretreatment, cellulase production in microorganisms for the following hydrolysis of the feedstock and finally the fermentation of sugar to ethanol (de Lourdes T.M. Polizeli and Rai, 2014). The enzymatic degradation of lignocellulosic material by cellulases and hemicellulases into simple sugars for biofuel production is currently the most investigated step. Because ethanol is a platform chemical with a large market, there is an increased interest in commercializing technologies for ethanol production from inexpensive or waste biomass like wood residues or agricultural residues like straw.

1.2 Cellulose degradation

Cellulose is the most abundant natural polysaccharide and is composed of β-D-glucopyranose units, linked by β-$(1,4)$-glycosidic bonds. Cellobiose, a disaccharide, is the smallest repetitive unit of cellulose and can be converted into two glucose residues (Kumar et al., 2008). The glucose strands form long and unbranched glucose polymers. These polymers are packed together and form highly insoluble crystals. To degrade these complex crystalline structures, cellulolytically potent organisms like *Trichoderma* produce complex mixtures of enzymes for efficient degradation. Over time, the cellulolytic system of *Trichoderma reesei* became the most intensively studied. The whole amount of different enzymes are efficiently secreted into the culture medium and they work together to decompose the highly crystalline native cellulose (Koivula et al., 1998).

1.2.1 The major cellulolytic enzymes in *T.reesei*

The cellulose-hydrolysing enzymes are called cellulases. The class of cellulases is divided into three major groups due to their point of action. The endoglucanases catalyse random cleavage of internal bonds of the cellulose chain. The cellobiohydrolases attack the chain ends and release cellobiose. The β-glucosidases can cleave cellobiose, releasing D-glucose monomer units (Kumar et al., 2008). The cellulolytic enzyme system of *T. reesei* contains two different cellobiohydrolases: CBHI and CBHII, at least four enduglucanases: EG1, EG2, EG3 and EG5, and two β-glucosidases: BGLI and BGLII (Kumar et al., 2008). The cellulolytic system of *T. reesei* is subject to multiple levels of control whereby most of the regulation happens at the level of transcription by CCR mediated by the Carbon catabolite repressor 1 (Cre1). Another regulation occurs via specific induction with easily available carbon sources like D-glucose and by cellulose and other cellulose break-down products as cellobiose or the transglycosylation product sophorose (Seiboth et al., 2011) (Strauss et al., 1995).

1.3 Hemicellulose degradation

Besides cellulose, hemicellulose is the second most prevalent polysaccharide in nature. Hemicelluloses are heterogeneous, non-cellulose polysaccharides composed of very different monosaccharide units such as D-xylose, D-mannose, D-glucose, D-galactose, L-arabinose, D-glucuronic acid and D-galacturonic acid. According to the main sugar units the hemicelluloses are classified in xylans, mannans, glucans, arabinans or galactans. Thus, when a polymer is hydrolysed and yields xylose, it is a xylan. In nature the hemicelluloses are complex structures made of more than just one polymer. But hemicelluloses mainly comprise xylans, which can be degraded by xylanolytic enzymes called xylanases. (Polizeli et al., 2005)

1.3.1 The major xylanolytic enzymes in *T.reesei*

The main xylan degrading enzymes in *T. reesei* are the two major specific endo-$\beta-1,4$-xylanases XYNI and XYNII and one β-xylosidase BXLI (Mach-Aigner et al., 2010) (Toerroenen et al., 1992). All enzymes work as a hemicellulolytic complex together to completely degrade xylan. The transcription of the xylanase-encoding gene *xyn*1 needs to be induced by xylan, cellulose or D-xylose at low concentrations, while *xyn2* is always present at a low basal level. The transcription of this gene can be further induced by xylan, xylobiose, sophorose, the disaccharide cellobiose and low concentrations of D-xylose (Mach-Aigner et al., 2008)(Mach-Aigner et al., 2010)(Zeilinger et al., 1996). The

transcription of *bxl*1 is induced by xylobiose (Margolles-Clark et al., 1997). In addition to this specific induction, the hemicellulase-encoding genes are also subject to CCR mediated by Cre1 (Strauss et al., 1995).

1.4 Transcription factors

The controlled expression of genes is important for organisms to endure changing environmental conditions. For this the initiation of the transcription is an essential step. In the initial phase various proteins, so called transcription factors, bind to specific regulatory sequences on the DNA and activate or repress the transcription. This is also observed for the genes encoding cellulose and hemicellulose degrading enzymes. In general transcription factors consist of at least two characteristic domains: the DNA-binding domain (DBD) for the interaction with DNA and the activation domain for specific interactions with other proteins. Depending on the structure of the DBD, the transcription factors can be classified into different groups. One of these is the zinc finger transcription factor family.

1.4.1 The zinc cluster proteins

The zinc finger motif in transcription factors was first identified 1985 in the *Xenopus* transcription factor TFIIIA (Miller et al., 1985). The zinc finger consists of one alpha helix and a pair of anti-parallel beta strands. In general one or two zinc atoms are bound by cysteins or histidins. The zinc-finger containing transcription factors can be categorized into various families according to their zinc-binding motifs. While some families, for example the Cys2His2, exist in a wide range of eukaryotes, the binuclear cluster is exclusively present in fungi. They are also called zinc cluster proteins and can be separated into three functional domains: the DBD, the regulatory domain, and the acidic region. In addition, the DBD is compartmentalized into subregions: the zinc finger, the linker and the dimerization domain. The family of zinc cluster proteins is best characterized for Gal4p in the budding yeast, *Saccharomyces cerevisiae*. Among the zinc cluster proteins are also the two transcription activators XlnR from *Aspergillus niger* and Xyr1 from *Trichoderma reesei*. These three proteins will be described in the following paragraphs.

1.4.2 Gal4p

The Gal4p transcription activator is one out of 50 zinc cluster proteins in the genome of *Saccharomyces cerevisiae*. Its regulation of genes involved in the catabolism of galactose is well studied. Gal4 binds to the *gal* gene promoters but is inhibited by another GAL protein, the GAL80. In the presence of D-galactose the signal transducer protein Gal3 interacts with GAL80 in the cytoplasm and activates GAL4 in the nucleus (Hasper et al., 2004). Research work on Gal4p gave a first impression on the structural and functional domains in zinc cluster proteins. The DNA-binding domain consists of three regions: the zinc finger, a linker and a dimerization region. The zinc finger has the well-conserved motif $CysX_2CysX_6CysX_{5-12}CysX_2CysX_{6-8}Cys$, where the cysteins bind two zinc atoms and build the DBD. Here the transcription factors can bind to DNA as monomer, homodimer or heterodimer (MacPherson et al., 2006). Downstream of the DNA-binding domain lays the so called fungal transcription factor middle homology region (FTFMHR) with about 80 amino acids. It is highly suggested that this region plays a part in regulating the transcriptional activity in zinc cluster proteins. This hypothesis is supported by the fact that this homology region could only be found in binuclear cluster transcription factors (Schjerling and Holmberg, 1996).

1.4.3 XlnR

Another transcription factor of the zinc cluster protein family is XlnR. This transcription activator is not only a central regulator protein responsible for the activation of more than ten genes involved in the degradation of xylan and cellulose in *Aspergillus niger*, but also contributes to the regulation of D-xylose metabolism. The functional domains of XlnR include again a N-terminal zinc binuclear cluster domain (DBD), a central coiled-coil domain localised in the FTFMHR and a C-terminal activation domain (Hasper et al., 2004). The isolation of a loss-of-function mutation in the putative C-terminal coiled-coil domain of XlnR had indicated that this region is important for functionality. The mutation and deletion studies from Hasper and co-workers show, that maintaining the structure of this region inside the FTFMHR is essential for proper functionality. In general coiled-coil regions are known from other transcriptional activators and are normally located near the N-terminus. These regions are involved in dimerization. In some transcription factors there could be another coiled-coil region in the C-terminus, but there is no known function for this domain in XlnR. When the the C-terminal region, including the coiled-coil region, is deleted, the transcription factor remains in the cytoplasms. So this region is involved in the nuclear import of the protein. This suggests a nuclear-targeting sequence in the C-terminus different from

classical nuclear localization signal (NLS) motifs. (Hasper et al., 2004) Many members of the zinc binuclear cluster family have shown to have a regulatory C-terminal domain. (Gal4, LEU3, PUT3). In the study from Hasper et al. the deletion of the C-terminal region downstream the predicted coiled-coil region leaves a fully active protein: This region responds to repressing and not to inducing signals. This is supported by the hypothesis that the mutation V756F disturbs a D-glucose inhibitory domain. Because this mutation results in xylanase expression under repressing condition. Deletion of a 34 amino acid sequence and the last 78 amino acids leads to an inactive state of XlnR. Here a glucose inhibition region in the C-terminal downstream of Leu_{668} responds to repressing signals via intra- or intermolecular interactions. (Hasper et al., 2004)

1.4.4 Xyr1

The Xylanase regulator 1 (Xyr1) transcription activator in *T. reesei* is a homologue of XlnR from *A. niger* and has a similar structural topology: there is the zinc binuclear cluster domain, the fungal middle homology region and the C-terminal activation domain (Klaubauf et al., 2014) (Derntl et al., 2013) (Fig 1.2). Xyr1 is a wide-domain activator for all important hydrolytic enzyme-encoding genes, except of *bgl2* (Stricker et al., 2006). Analysis of the transcription level of a *xyr1* deletion strain show no transcript formation for the major cellulase-encoding gens *cbh1*, *cbh2* and *egl1* and also no transcript formation for the major xylanase-encoding genes *xyn1* and *xyn2* under repressing or inducing conditions (Mach-Aigner et al., 2008). *Xyr1* itself is subject to CCR mediated by Cre1 in the presence of D-glucose and high concentrations of D-xylose. Cre1 is a wide-domain regulator that binds under repressing conditions to its binding sites in the promoter of *xyr1* what leads to a down-regulation of transcription (Derntl et al., 2013). The inducers of xylanase expression, D-xylose and xylobiose show no inducing effect on the transcription level of *xyr1*, but the transglycosylation product sophorose increases the transcript formation (Derntl et al., 2013). The release from CCR is an evident step for the industrial development and utilisation of *T. reesei* for the production of cellulases and xylanases. Therefore, the Cre1-deficient mutant strain of *T. reesei*, Rut-C30, has been used as a parent strain for many recent industrial strains. Derntl and co-workers from our institute reported on an industrially used *T. reesei* strain that produces high amounts of enzymes, independent of the presence of a certain inducer. The strain was analysed at the transcriptional level in order to characterize his phenotype. It shows a glucose-blind phenotype when it comes to expression of cellulases and xylanases. By DNA-sequencing it was observed that the strain bears a single point mutation in Xyr1. The single point mutation A824V in the C-terminal fungal transcription factor middle homology region of Xyr1, caused an increase and a strong deregulation of xylanase

expression as well as an increase in cellulase expression (Derntl et al., 2013). The single point mutation is located in an predicted α-Helix in the fungal transcription factor middle homology region (FTFMHR). Therefore Derntl and co-workers suggested that a change in the secondary structure of Xyr1 might cause the increased expression of xylanases. This hypothesis is supported by different deletion and mutation studies with the homologue transcription activator XlnR in *A. niger*. Recently, in our lab, the point mutation A824V was introduced in the wild type strain Qm6a. The produced strain is not yet fully characterized and a lot of research focuses on it. Mello-de Sousa and co-workers (unpublished data) demonstrated that the Xyr1 wild type and the Xyr1$_{A824V}$ have indeed different secondary structures and that the complete Xyr1$_{A824V}$ protein shows no response to carbohydrates, inducing or non-inducing, any more. They also suggest that in the wild type there could be a sugar-binding domain which leads to a change in protein conformation when a repressing sugar is present. Then the activation domain in the C-terminus may be hidden. The single point mutation A824V could then lead to a formation where the activation domain is always presented regardless of the presence of an inducer or a repressor (Figure 1.3).

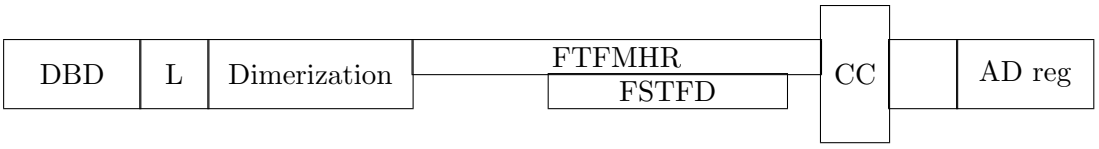

FIGURE 1.2: structural map of Xyr1: DBD= DNA binding domain; L= Linker; Dimerization= Dimerization domain; FTFMHR= fungal transcription factor middle homology region; FSTFD= fungal specific transcription factor domain; CC= putative coiled-coil domain; AD= activation domain

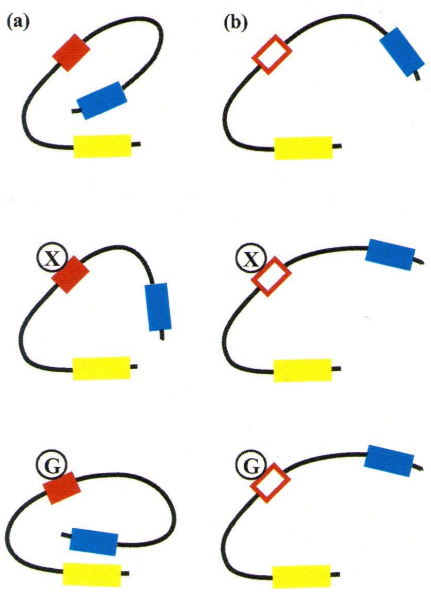

FIGURE 1.3: hypothetical conformation due to the single point mutation A824V in Xyr1. Yellow box: DNA-binding domain, blue box: transactivation domain, red box: hypothetical sugar-binding domain, X: inducing sugar, G: repressing sugar. a) wild type Xyr1, b) $Xyr1_{A824V}$

1.5 Aim of this master thesis

In this thesis, a wild-type strain carrying the mutation A824V within Xyr1 was used. This strain is not released from CCR. The aim of this master thesis is on one hand the initial characterization of this yet unknown mutation strain $Qm6aXyr1_{A824V}$. On the other hand, the hypothesis if the mutation A824V in Xyr1 leads to a change in secondary structure and might be involved in the increased and strong deregulation of xylanase expression as well as an increase in cellulase expression will be tested. For the characterization of the mutation strain the transcription levels of *xyr1* and different other genes encoding for xylanases (*xyn1*, *xyn2*, *cbh1* and *cbh2*) and cellulases will be analysed. Further characterization will involve chromatin analysis of promoters of xylanase- and cellulase-encoding genes. To produce meaningful results all the analysis will be repeated with the wild type Qm6a strain and will then be compared. To investigate the changes caused by the mutation, several recombinant Xyr1 wild type and $Xyr1_{A824V}$ protein fragments will be analysed. Their secondary structure, in the absence and presence of different carbon sources, will be compared via CD-Spectroscopy.

Chapter 2

Material and Methods

2.1 Materials

2.1.1 Kits

Thermo Scientific GeneJET PCR Purification Kit #K0701, #K0702
Thermo Scientific GeneJET Plasmid Miniprep Kit #K0502, #K0503
Thermo Scientific GeneJET Plasmid Midiprep Kit #K0481, #K0482
QIAGEN®Plasmid Midiprep Kit
Thermo Scientific GeneJET Gel Extraction Kit #K0691, #K0692
Thermo Scientific RevertAid H Minus First Strand cDNA Synthesis Kit #K1691
Thermo Scientific CloneJET PCR Cloning Kit #K1231, #K1232
BIO-RAD iQ^{TM} SYBR®Green Supermix #170-8880 #170-8882 #170-8884 #170-8885
His-Bind®protein purification raisin

2.1.2 Vectors

pTS1 (pET28a Xyr1 WT)(8086bp and Kan^r)
pTS2 (pET28a Xyr1 A824V) (8086bp and Kan^r)
pJET2.1 blunt cloning vector (2974bp and Amp^r)
(all vector maps are available in the appendix A)

2.1.3 Bacterial strains

TOP10 competent cells
BL21(DE3) competent cells

2.1.4 Fungal strains

Qm6aΔtmus53Δpyr4
Qm6aXyr1$_{A824V}$ (Qm6aΔtmus53Δpyr4 Xyr1(A824V)+pyr4)

The fungal strains used in this thesis have been provided from a strain bank from the Department of Gene Technology at the Institute of Chemical Engineering at the University of Technology in Vienna.

2.1.5 Technical devices

TABLE 2.1: Technical devices

company	technical device
BINDER	E28 Incubator/Dryer
BIO-RAD	MJ Mini Thermal Cycler
BIO-RAD	T100 Thermal Cycler Thermal Cycler
BIO-RAD	Molecular Image Gel Scanner
Eurotech	-20°C refrigerator
Eppendorf	Thermomixer compact Thermal Cycler
INFORS	MultitonShuttler Inkubator
Inolab	WTW series 720 pHmeter
JASCO	J-815 CD Spectrometer
JASCO	N2 Monitor
JASCO	CDF-426 Temperature control unit
JASCO	V-630 Spectrophotometer
MITSUBISHI	P93D Scan printer
MP	Fast Prep-24 Cell hydrolyser
New Brunswick Scientific	Escella E24 Incubator Shaker Series
New Brunswick Scientific	Ultra Low Temperature Freezer U725 Innova -80°C refrigerator
Peqlab	Perfect Spin Mini Centrifuge/Spinner
QIAGEN	Rotor-Gene Q Thermal Cycler
QIAGEN	QI Agility PCR robot
REICHERT	REICHERT Microscope
Scientific Industries	Vortex-Genie 2
SIGMA	Sigma 1-15K Centrifuge/Spinner
SIGMA	Sigma 3K30 Centrifuge/Spinner
SIGMA	Sigma 3-18K Centrifuge/Spinner
Thermo Scientific	NanoDrop 1000 Spectrophotometer

2.2 Methods

2.2.1 Xyr1 fragment design

To analyse the effects caused by the mutation A824V in Xyr1, six different Xyr1 protein fragments are designed (Figure 2.1). The first fragment, Xyr1-1, consists only of the DNA binding domain and is therefore identical with the same fragment in the wild type Xyr1 protein. Xyr1-2 is its counterpart and misses only the DBD. Therefore there are two versions of this fragment. Xyr1-2.1 is the wild type version, whereas Xyr1-2.2 contains the mutation A824V. Xyr1-3 contains a big part of the FTFMHR and the putative coiled-coil region. The two Xyr1-4 fragments provide the C-terminal of the protein without (Xyr1-4.1) and with (Xyr1-4.2) mutation. The last fragments are Xyr1-5.1 and Xyr1-5.2, a wildtype and a mutated version, and Xyr1-6, which only consists of the putative transactivation domain.

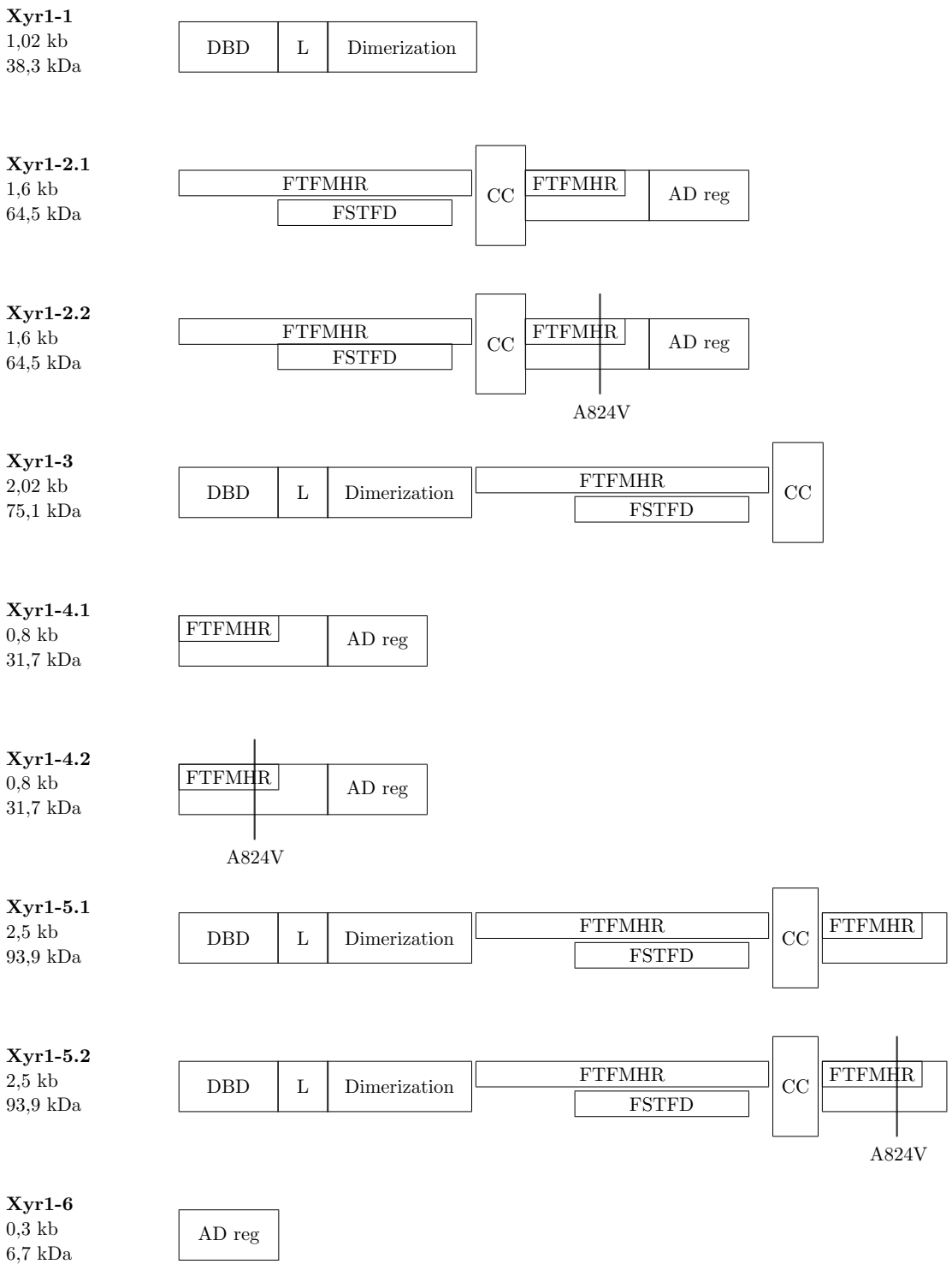

FIGURE 2.1: XYR1 fragment design
DBD= DNA binding domain; L= Linker; Dimerization= Dimerization domain; FTFMHR= fungal transcription factor middle homology region; FSTFD= fungal specific transcription factor domain; CC= putative coiled-coil domain; AD= activation domain

2.2.2 Polymerase chain reaction (PCR)

2.2.2.1 Primer design

All Oligonucleotides were designed according to the Xyr1 fragments to be generated and ordered from Sigmar Aldrich. Stocks were stored as 100μM solutions in H_2O at -20°C

Table 2.2: Primer (P) design with restriction sites

Primer	Sequence 5'-3'	Tm°
xyr1 P1-BglII	GACAGATCTCGATCCCGCGAAATTAATACGA	75.3
xyr1 P2-NotI	GACGCGGCCGCCTGTTTAAAGATCGGCTGACTC	84.1
xyr1 P3-NcoI	CATCCATGGAGAGCGATCTGCGTTATCCGG	80.6
xyr1 P4	GGATATAGTTCCTCCTTTCAGCAAAA	65.1
xyr1 P5-NotI	GACGCGGCCGCAAAATAACCGAAAATGCTACG	82.0
xyr1 P6-NcoI	CATCCATGGTTCTGAGCCTGATGACGATTCTG	78.7
xyr1 P7-NotI	GACGCGGCCGCCGGCATAAATTCCAGGCCAGGATCA	90.6
xyr1 P8-NcoI	CATCCATGGCGTTTTTCTATGGTGTTTACC	73.7
xyr1 P9	CCATGGCAGGTTATAGCGGTATT	69.3

2.2.2.2 Standard PCR

Standard PCR was performed in a Thermal Cycler either with GoTaq®DNA Polymerase or with Phusion®High-Fidelity DNA Polymerase. If not indicated otherwise, the following PCR conditions were used:

Table 2.3: Standard PCR with GoTaq®DNA Polymerase

PCR with GoTaq®DNA Polymerase				
Step	**Temperature**	**Time**	**Cycles**	**PCR mix**
Initial denaturation	95°C	3'	1	1μl template (10ng/μl)
Denaturation	95°C	30"		5μl 5×reaction buffer
Annealing	63°C	30"	29	0.5μl dNTPs (10mM)
Elongation	72°C	1'/kb		0.5μl forward primer (10μM)
Final elongation	72°C	5'	1	0.5μl reverse primer (10μM)
Storage	12°C	∞	1	0.125μl Polymerase (10ng/μl)
				17.375μl ddH_2O

Table 2.4: Standard PCR with Phusion®High-Fidelity DNA Polymerase

PCR with Phusion®High-Fidelity DNA Polymerase				
Step	**Temperature**	**Time**	**Cycles**	**PCR mix**
Initial denaturation	98°C	30”	1	2μl template (10ng/μl)
Denaturation	98°C	10”		10μl 5×reaction buffer
Annealing	primer specific	15”	34	1μl dNTPs (10mM)
Elongation	72°C	15”/kb		2.5μl forward primer (10μM)
Final elongation	72°C	5’	1	2.5μl reverse primer (10μM)
Storage	12°C	∞	1	0.5μl Polymerase (2U/μl)
				31.5μl ddH_2O

Table 2.5: Primer pairs

***xyr*1 Fragment**	**Primer**	**Annealing T** Phusion®High-Fidelity DNA Polymerase	**Annealing T** GoTaq® DNA Polymerase
xyr1-1	for: *xyr1*P1-BglII	72°C	63°C
	rev: *xyr1*P2-NotI	72°C	63°C
xyr1-2	for: *xyr1*P3-NcoI	67°C	63°C
	rev: *xyr1*P4	67°	63°C
xyr1-3	for: *xyr1*P1-BglII	72°C	63°C
	rev: *xyr1*P5-NotI	72°C	63°C
xyr1-4	for: *xyr1*P6-NcoI	67°C	63°C
	rev: *xyr1*P4	67°C	63°C
xyr1-5	for: *xyr1*P1-BglII	72°C	63°C
	rev: *xyr1*P7-NotI	72°C	63°C
xyr1-6	for: *xyr1*P8-NcoI	67°C	63°C
	rev: *xyr1*P4	67°C	63°C

2.2.2.3 Colony PCR

Colony PCR was used to analyse *E.coli* positive transformants, concerning the successful cloning of the Xyr1 constructs into the expression vector pET28a. A few transformants were picked with a sterile toothpick, transferred onto a selective LB masterplate and added to a prepared PCR reaction mixture. On an agarose gel, the PCR product was analysed for the right fragment size.

Table 2.6: reaction mixture for Colony PCR

Component	**Volume**
5 × *Taq* buffer	5μl
dNTPs mix (10mM)	0.5μl
Forward Primer (10μM)(pJET1.2 Sequencing Primer/specific fragment Primer)	0.5μl
Reverse Primer (10μM)(pJET1.2 Sequencing Primer/specific fragment Primer)	0.5μl
Taq DNA Polymerase (5U/μl)	0.125μl
Water	18.3751μl
DNA	one colony
Total volume	25μl

Table 2.7: Colony PCR with GoTaq®DNA Polymerase

Colony PCR with GoTaq®DNA Polymerase			
Step	**Temperature**	**Time**	**Cycles**
Initial denaturation	95°C	3’	1
Denaturation	95°C	30”	29
Annealing	60°C	30”	
Elongation	72°C	1’/kb	
Final elongation	72°C	5’	1
Storage	12°C	∞	1

2.2.2.4 Quantitative PCR

Real-time PCR is the technique to collect data throughout the PCR progress, combining amplification and detection in one single step (Wong and Medrano, 2005). This is achieved by using a fluorescent substance, SYBR®Green, that intercalates with double stranded DNA and therefore the fluorescence intensity correlates with the PCR product concentration. Reactions are characterized by the point in time where the target amplification is first detected, This value is usually referred to as cycle threshold (C_t). This is the time, where the fluorescence intensity is greater then the background fluorescence. This means: The greater the quantity of target DNA in the starting material, the faster the increase in fluorescence intensity appears. This leads to a lower C_t-value.

qPCR for transcript level analysis

Quantitative PCRs for transcript level analysis were performed in a Rotor-Gene Q Thermal Cycler (Qiagen). All reactions were performed in triplicate. The amplification mixture (finale volume: 15μl) contains 7.5 μl 2 × iQSYBR®Green Mastermix (Bio-Rad Laboratories), 6.25 μM forward and reverse primer and 2 μl cDNA (diluted 1:20). Primer sequences are provided is table 2.8. The transcript levels of seven different genes were analyzed: *xyn1, xyn2, cbh1, cbh2, sar1, act1* and *xyr1*. *Sar1* and *act1* serve as reference genes. The transcript levels refer to the wild-type incubation without carbon source. The cycling conditions are described in table 2.9.

During relative quantitation, changes in sample gene expression are measured based on one or more reference genes (*act1* and *sar1*). There are different methods available to calculate the mean normalized gene expression. In this thesis the Comparative C_t ($2^{-\Delta\Delta C_t}$) method is used (Wong and Medrano, 2005). In the comparative method, the C_t-values obtained from two different experimental RNA samples are directly normalized to a housekeeping gene and then compared. First, the difference between the C_t-values (ΔC_t) of the gene of interest and the housekeeping gene is calculated for each experimental sample. Then, the difference in the ΔC_t values between the experimental and control samples $\Delta\Delta C_t$ is calculated. The fold-change in expression of the gene of interest between the two samples is then equal to $2^{-\Delta\Delta C_t}$.

Table 2.8: RT qPCR Primer for transcript level analysis

Primer name	**Sequence 5'-3'**
act1f	TGAGAGCGGTGGTATCCACG
act1r	GGTACCACCAGACATGACAATGTTG
cbh1f	GATGATGACTACGCCAACATGCTG
cbh1r	ACGGCACCGGGTGTGG
cbh2f	CTATGCCGGACAGTTTGTGGTG
cbh2r	GTCAGGCTCAATAACCAGGAGG
taqxyn2f	GGTCCAACTCGGGCAACTTT
taqxyn2r	CCGAGAAGTTGATGACCTTGTTC
taqxyr1f	CCCATTCGGCGGAGGATCAG
taqxyr1r	CGAATTCTATACAATGGGCACATGGG
sar1f	TGGATCGTCAACTGGTTCTACGA
sar1r	GCATGTGTAGCAACGTGGTCTTT
xyn1f	CAGCTATTCGCCTTCCAACAC
xyn1r	CAAAGTTGATGGGAGCAGAAG

TABLE 2.9: RT qPCR cycling conditions for transcript level analysis

qPCR			
Gene	**Temperature**	**Time**	**Cycles**
sar1	95°C	3’	1
	95°C	15”	
	64°C	2’	40
	60-95°C 0,2°C/step	melting analysis	
act1	95°C	3’	1
xyn2	95°C	15”	
	60°C	15”	40
	72°C	15”	
	60-95°C 0.2°C/step	melting analysis	
xyn1	95°C	3’	1
cbh1	95°C	15”	
cbh2	59°C	15”	40
xyr1	72°C	15”	
	60-95°C 0.2°C/step	melting analysis	

qPCR for chromatin accessibility analysis

qPCR analysis of the DNase I-treated samples was performed to measure the relative abundance of target regions in different genes. All PCRs were performed in triplicates in a Rotor-Gene Q Thermal Cycler (Qiagen). The amplification mixture (finale volume: 20μl) contains 10 μl 2 × iQSYBR®Green Mastermix (Bio-Rad Laboratories), 10 μM forward and reverse primer and 2.5 μl DNase I-treated sample (4 ng/μl). Primer sequences are provided in table 2.10. The chromatin accessibility of four different genes and different target regions were analysed: *sar1, act1, xyn1* (extremeURR, URR, CORE) and *xyn2* (URR, CORE). *Sar1* and *act1* serve as reference genes for normalization. The cycling conditions are described in table 2.11.

The amount intact input DNA of each sample was calculated by comparing the threshold values of the PCR amplification with a standard curve generated for each primer set using serial dilutions of genomic, undigested DNA. The chromatin accessibility index (CAI) is defined as: CAI = 1/(Ds/((Dc1+Dc2)/2)), where Ds is the amount of intact DNA detected for each target region and Dc1 and Dc2 are the amounts of intact DNA

detected for the promoter regions of *sar1* and *act* respectively, used as reference genes for normalization.

Table 2.10: CHART qPCR Primer

Primer name	**Sequence 5'-3'**
epiactinTr-f	CTTCCCTCCTTTCCTCCCCCTCCAC
epiactinTr-r	GCGACAGGTGCACGTACCCTCCATT
epixyn1-1Tr-f	GCACTCCAAGGCCTTCTCCTGTACT
epixyn1-1Tr-r	TAGATTGAACGCCACCCGCAATATC
epixyn2-2Tr-f	CTCGAGACGGCTGAGACAGCAGCAT
epixyn2-2Tr-r	TGTCTTTTGGGCTTGGAGGGGTTGT
episar1Tr-f	GTCAGGAAATGCCGCACAAGCAAGA
episar1Tr-r	TGTGTTTTACCGCCTTGGCCTTTGG

Table 2.11: CHART qPCR cycling conditions

CHART qPCR			
Gene	**Temperature**	**Time**	**Cycles**
sar1	95°C	3'	1
act1	95°C	15"	
xyn1	59°C	15"	40
xyn2	72°C	15"	
	60-95°C 0.2°C/step	melting analysis	

2.2.2.5 Purification of DNA from PCR

The PCR products were purified using the (Thermo Scientific) GeneJet PCR Purification Kit. The PCR products were combined and mixed with 1:1 volume Binding buffer. A maximum of 800μL solution were then added to a provided column and centrifuged for one minute at 14800 $\times$ g. If not mentioned otherwise, the centrifugation parameters have been maintained. After centrifugation the flow-through was discarded and 700μl of washing buffer were added. After another round of centrifugation and flow-through discard, 50μl ddH_2O were added and centrifuged into a new sterile eppendorf tube. The purified DNA was quantified with the NanoDrop Spectrometer (Thermo Scientific) and stored at -20°C.

2.2.3 Gel electrophoresis

2.2.3.1 Agarose gel electrophoresis

0.8% agarose gels were used to separate DNA fragments by gel electrophoresis. Gel composition: 0.8% w/v ultra pure agarose in TBE buffer (0.1 M Tris, 0.1 M boric acid, 2mM EDTA). 0.5 μl/ml EtBr were added to the gels to make visualization of the DNA via UV-light possible. Before loading, samples were supplemented with at least 0.5 volumes of gel loading buffer with bromphenol blue. The GeneRuler 1kb DNA Ladder (Thermo Scientific) with a range from 250 bp - 10000 bp, was used as DNA size marker. Voltage (90V) and running times (35 - 50 min) were dependent on the expected fragment sizes. For visualization, the BIO-RAD Molecular Image Gel Scanner was used.

2.2.3.2 SDS-Polyacrylamide gel electrophoresis (PAGE)

The separation of protein due to their molecular mass occured with SDS-PAGE. The purified protein solution was mixed with 4 $\times$ Laemmli buffer (240 mM Tris pH 6.8; 400 mMDTT; 8% SDS; 20% Glycerol; 0.004% Bromphenol Blue) and heated for five minutes at 95°C. The proteins were separated in a 12% resolving gel with a 5% stacking gel. Voltage (90V) and running times (90 min) were dependent on the expected fragment sizes. For visualization, the gel was stained with Coomassie Brilliant Blue and scanned with the BIO-RAD Molecular Image Gel Scanner.

2.2.3.3 Blue Native Polyacrylamide gel electrophoresis (BN-PAGE)

For protein interaction studies, a one dimension BN-PAGE was used. The proteins were separated by size and charge. 18 μl of purified protein solution were mixed with 2μl 10$\times$ BN-PAGE loading buffer (750 mM aminocaproic acid, 50 mM Bis-Tris, 5% Coomassie Brilliant Blue G-250). The solution was loaded on a 12% or a 8% resolving gel with a 5% stacking gel on top. The gel was electrophoresed on ice (4°C) with 100V and 35 mA for 2 hours in a BN-PAGE running buffer (250 mM Tris Base, 1.96 M Glycine). Afterwards the gel was stained with Coomassie Brilliant Blue and scanned with the BIO-RAD Molecular Image Gel Scanner.

2.2.4 Preparation of high quality plasmid DNA from recombinant *E. coli* cultures

For the extraction of plasmids from the bacterial cells, the GeneJET Plasmid Miniprep Kit (Thermo Scientific) was used. A single colony from a freshly streaked selective plate was picked and inoculated in 5mL of LB medium, supplemented with the appropriate selection antibiotic. After incubation for 12-16 hours at 37°C while shaking at 200rpm, the bacterial culture was harvested by centrifugation at maximum speed (14000 × g) in a microcentrifuge for two minutes at room temperature. The supernatant was discarded and 250 μl Resuspension Buffer were added. After mixing, the same amount of Lysis Buffer was added and the tubes were carefully inverted for six times. Afterwards 350μl of Neutralization Buffer were added. The cell lysate was centrifuged for five minutes with 12500 × g. The supernatant was transferred to a provided column and centrifuged for one minute under the same conditions. After discarding the flow-through, 500μl Wash Solution were added and centrifuged again for one minute. This step was repeated another time and then the column was transferred into a new sterile eppendorf tube and 50μl ddH_2O were added and the plasmid was eluted via centrifugation for two minutes. The concentrated plasmid sample was quantified with the NanoDrop Spectrometer (Thermo Scientific) and the samples were stored at -20°C.

2.2.5 Cloning of DNA fragments

The cloning into the desired vectors was achieved by a two-step cloning process. First all purified PCR fragments were ligated into the pJET2.1 blunt cloning vector and transformed into competent TOP10 *E.coli* cells. Afterwards the transformants were checked via Colony-PCR. The successfully transformed fragments were cut out of the vector and purified. The digestion was loaded on an agarose gel and the fragments were cut out and extracted from the gel. Then the extracted fragments were ligated into the desired vector pTS1 or pTS2.

2.2.5.1 Restriction digest of plasmid DNA

For the restriction digest, two different restriction systems were needed (Table 2.12). The final reaction volume was 60μl. The reaction mixture consisted of the specific buffer, the restriction enzymes, water and 45μl of plasmid DNA. The amount of plasmid was quantified with the NanoDrop spectrometer. The enzymes were then added in adequate amounts, acording to table 2.12. The digestion was performed over night at

37°C. Afterwards an agarose gel electrophoresis was performed to check if the digestion worked and if the desired DNA fragments with the right size were present.

TABLE 2.12: Restriction digest of plasmid DNA

DNA sequence	**Restriction enzymes**	**amount of enzyme**	**Reaction Buffer**
xyr1-1	BglII/NotI	1U/μg DNA	Buffer O
xyr1-2	NcoI/NotI	2U/μg DNA	Buffer BamH1
xyr1-3	BglII/NotI	1U/μg DNA	Buffer O
xyr1-4	NcoI/NotI	2U/μg DNA	Buffer BamH1
xyr1-5	BglII/NotI	1U/μg DNA	Buffer O
xyr1-6	NcoI/NotI	2U/μg DNA	Buffer BamH1
pTS1	BglII/NotI	1U/μg DNA	Buffer O
PTS2	NcoI/NotI	2U/μg DNA	Buffer BamH1

2.2.5.2 Plasmid vector dephosphorylation

For vector dephosphorylation with FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) the following reaction mixture is prepared:

TABLE 2.13: reaction mixture for vector dephosphorylation

Linear DNA	1μg
10X reaction buffer for AP used in reaction	2μl
FastAP Thermosensitive Alkaline Phosphatase	1μl(1U)
Water, nuclease-free	up to 20μl
Total volume	20μl

The reaction mixture was mixed thoroughly, spun briefly and incubated at 37°C for 10 minutes. To avoid band shifts in the following gel electrophoresis, the reaction was stopped and the enzyme was inhibited by heating at 65°C for 15 minutes.

2.2.5.3 DNA gel extraction

The DNA fragments, resulting from the restriction digest, were cut out of the gel and extracted from it with the GeneJET Gel Extraction Kit (Thermo Scientific). The gel slices were weighted and combined with 1:1 (w/v) volumes binding buffer. The gel mixture was incubated at 55°C for 10 minutes until the gel slices were completely melted.

Then the mixture was loaded on the provided column and centrifuged for one minute at 15000 × g. The flow-through was discarded, 700μl wash buffer were added and centrifuged again. The empty GeneJET purification column was centrifuged for an additional minute to completely remove residual wash buffer. The column was transferred into a new sterile eppendorf tube. 50μl ddH_2O were added and the plasmid was eluted via centrifugation for two minutes. The eluate can be quantified with the NanoDrop Spectrometer (Thermo Scientific) and stored at -20°C.

2.2.5.4 Ethanol DNA precipitation

To purify the extracted DNA fragments, they were precipitated with ethanol. To 50μl of DNA 10μl 0.3M ammonium acetate solution pH 5.2 were added. The mixture was filled up with water to 100μl. For the precipitation, 200μl 100% ethanol were added and stored at -20°C for two hours. After centrifugation (14000 × g, 30 minutes, 4°C) the supernatant was carefully discarded and one ml of 70% ethanol was added to the pellet. After another 15 minutes of centrifugation the supernatant was again discarded and the pellet was dried on air for five minutes. 30μl ddH_2O were added and the purified DNA was quantified with the NanoDrop Spectrometer.

2.2.5.5 DNA ligation into pJET2.1 blunt cloning vector

For the ligation into the pJET1.2 vector, the CloneJET PCR Cloning Kit (Thermo Scientific) was used. The Bunt-end DNA fragments generated by restriction enzyme digestion and purified from gel, were used in a 3:1 molar ratio with the pJET1.2 blunt vector. The ligation mixture was incubated for one hour at room temperature and was then directly used for transformation.

Ligation set up:

Table 2.14: ligation set up

Component	**Volume**
2X Reaction buffer	10μl
Purified PCR product	1,1μl
pJET1.2/blunt Cloning Vector (50ng/μl	1μl
Water, nuclease-free	up to 19μl
T4 DNA Ligase	1μl
Total volume	20μl

2.2.5.6 DNA ligation into pTS1 and pTS2

For the ligation into the pTS1 or pTS2 vector, the DNA fragments generated by restriction enzyme digestion and purified from gel, were used in a 5:1 molar ratio with the linearised vector. All fragments encoding the wild-type versions of Xyr1 were inserted into pTS1, whereas all fragments with the mutation were ligated with pTS2 (Table 2.15). The ligation mixture was incubated for one hour at room temperature and was then directly used for transformation.

Table 2.15: DNA fragments and cloning vector pTS1 or pTS2

DNA fragment	**vector plasmid**
xyr1-1	pTS1
xyr1-2.1	pTS1
xyr1-2.2	pTS2
xyr1-3	pTS1
xyr1-4.1	pTS1
xyr1-4.2	pTS2
xyr1-5.1	pTS1
xyr1-5.2	pTS2
xyr1-6	pTS1

2.2.5.7 Transformation

To 100μl competent and thawed TOP10 or BL21(DE3) *E.coli* cells, 10μl of ligation reaction were added and carefully mixed. After an incubation time of 30 minutes on ice, the cells underwent a heat shock at 42°C for exactly two minutes, to enhance the intake of foreign DNA. With 300μl of pre-warmed LB-media the cells regenerated for 30 minutes at 37°C. Afterwards the cells were plated on LB with the appropriate antibiotics and incubated over night at 37°C. The transformants were checked via Colony PCR (2.2.2.3) the next day.

2.2.6 Sequencing

To check if the genomic sequences, encoding for the desired protein fragments were intact after all procedures happened before, the plasmids were sequenced. The sequencing was performed overnight by Microsynth (Swiss DNA company). For each plasmid the appropriate primers were needed. For a look at the primer pairs go to table 2.5. The

DNA samples and sequencing primers were pre-mixed. The plasmid had to be sequenced with the forward and the reverse primer. For this the same amount of plasmid is transferred into two 1.5 ml tubes, whereas the forward primer is added to one, and the reverse primer is added to the other tube. Each sample had a volume of 12 μl plasmid solution with a concentration of 720-1200 ng, pre-mixed with 3 μl sequencing primer solution (10μM).

2.2.7 Generation of competent *E.coli* BL21(DE3) cells

For generation of competent *E.coli* cells, 400 ml LB were inoculated with a 5 ml overnight culture of the desired strain BL21(DE3) to an initial optical density (OD_{600}) of 0.1. Shaking at 37°C the culture grew until an OD_{600} of approximately 0.6 was reached. The cells were then harvested by centrifugation (15 min, 600 rpm, 4°C), re-suspended in 160 ml MOPS I and incubated on ice for 10 minutes. After another round of centrifugation, the yielding pellet was re-suspended in 160 ml MOPS II an incubated for 30 minutes. The suspension was centrifuged again and the pellet was solved in 8 ml MOPS IIa. Aliquots of 100 μl were pipetted into 1.5 ml eppendorf tubes and stored at -80°C until final use.

2.2.8 Xyr1 production and purification

All protein versions contain a C-terminal His_6-tag, preceded by a thrombin cleavage site.

2.2.8.1 Xyr1 production

A single BL21(DE3) colony with the desired Xyr1 fragment cloned in the respective vector was picked from a plate and inoculated in 5 ml LB-media containing 50 μg/ml kanamycin and 1% glucose. The cells cultivated while shaking with 200rpm at 37°C overnight. The next morning, the total 5 ml overnight culture were added to 200 ml ZYP-5052 containing 100 μg/ml kanamycin. The culture was kept shaking with 250rpm for 4 hours at 37°C and then for 20 hours at 18°C. The cells were harvested by centrifugation at 10000 × g for 10 minutes at 4°C. For each production batch, the cells were divided into two aliquots of 100 ml and the pellets were afterwards stored at -80°C. Prior to purification, the cell pellets are thawed on ice and re-suspended in 10 ml BB5-L (binding buffer with 5 mM imidazole for lysis), incubated on an orbital mixer at room temperature for 30 minutes. The raw cell lysate was then centrifuged (14000 × g, 20 min, 4°C) to pellet the cell debris and filtered through a 0.45μm membrane, then immediately used for purification of Xyr1 expression fragments.

2.2.8.2 Xyr1 purification with His-Bind®

The raw cell lysate contains the recombinant protein among other proteins originating from the bacterial host. The recombinant protein, containing a 6 × Histidine-tag, was purified by Ni-IDA (iminodiacetic acid) agarose chromatography. The method is based on the interaction between a transition Ni^{2+} ion immobilized on a matrix, and the histidine side chains. For this purpose a polypropylene column was filled with 2 ml of uncharged His-Bind®slurry (Novagen) leading to a bed volume of 1 ml. The column was washed with 3 volumes sterile deionized water and charged with 5 volumes of 1 × Charging Buffer (50 mM $NiSO_4$). After the column has been equilibrated with 3 volumes of Binding Buffer BB5 (5 mM imidazole), the filtered cell lysate was added to the column. Afterwards the column was first washed with 10 volumes BB5 and then with 12 volumes of WB40 (40 mM imidazole). After bed settling, the protein was eluted with 6 volumes WB200 (200 mM imidazole) and sampled into 1 ml aliquots. From each wash step and fraction 50 μl were saved for a SDS-PAGE. The samples were stored at 4°C for further preparation.

2.2.8.3 Protein concentration

To achieve a higher protein concentration, all purified protein fractions were combined and concentrated with the Amicon®Ultra-4 Centrifugal Filter Device (50K or 100K). The device contains a membrane that works as a barrier for proteins with the respective size. As shown in table 2.16, two millilitres of protein solution were loaded on the device and centrifuged for a few minutes at room temperature. The flow-through was discarded and the concentrate was combined with the rest of the protein solution. Then again two millilitres were concentrated under the same conditions. This procedure was repeated until only 2.5 ml of concentrated protein solution were left.

Table 2.16: Conditions for protein concentration with different filter devices

Amicon®Ultra-4 Centrifugal Filter Device protein concentration conditions		
	50K	**100K**
centrifugation rotor	fixed angle rotor	fixed angle rotor
loading volume	2 ml	4 ml
rpm	7500 × g	5000 × g
centrifugation time	4 min	5 min

2.2.9 Circular dichroism

Circular dichroism (CD) is a tool which can be used to analyse the secondary structure and folding properties of recombinant and purified proteins. The most widely used applications of protein CD are to determine whether an expressed, purified protein is folded, or if a mutation affects its conformation or stability. In addition, it can be used to study protein interactions (Greenfield, 2006). The interaction between polarised light and matter is the foundation of CD. Polarised light consists of two circularly polarised components of equal magnitude, one rotating counter-clockwise and the other one rotating clockwise. The material, through which the polarised light passes, can absorb the two components differentially. If the absorption of both components is different, the resulting radiation "possesses elliptical polarisation". (Kelly et al., 2005) Proteins have different chromophores that absorb at different wavelengths and give a whole spectra of information about the molecule. The chromophores of interest include the peptide bond (absorption far UV), the aromatic amino acid side chains (absorption near UV) and disulphide bonds (absorption around 260nm) (Kelly et al., 2005). In this thesis, the analysis of the peptide bonds and the aromatic amino acid side chains are the most important. Measuring the absorbance resulting from the peptide bonds gives information about the secondary structure. A spectra in the region 260-320nm arises from the aromatic amino acid side chains and gives information about the tertiary structure. This method is often used to get a hint about changes in tertiary structure between a wild-type and a mutated protein.

2.2.9.1 Buffer exchange

The Elution Buffer (WB200) comprises imidazole which absorbs at far UV. This makes the buffer unsuitable for the following CD measurements of the proteins. Prior to the measurements the buffer needed therefore to be exchanged to the CD working buffer (50 mM Tris, 200 mM NaCl, 50 mM NaH_2PO_4, 10% glycerol pH 7.5). The buffer exchange was performed by employing PD-10 desalting columns (GE Healthcare). After equilibrating the column with 25 ml of CD Buffer, the concentrated 2.5 ml protein solution was added. After the protein solution was totally bedded, the protein was eluted with 3.5 ml CD Buffer.

2.2.9.2 Protein quantification

The protein quantification was performed with the Bradford method, using the Protein Assay Dye Reagent Concentrate (Bio-Rad) and bovine serum albumin (BSA) standards.

2 mg of BSA were dissolved in 1 ml water and filtered through a Millipore filter. The solution was further diluted to 1 mg/ml and to 8 standards of 10, 30, 50, 75, 150, 200, 300 and 500 μg/ml, and kept frozen. One part Dye Reagent Concentrate was diluted with 4 parts water. For the measurement 990 μl Bradford reagent and 10 μl of protein solution or standard were mixed, incubated for 5 minutes, transferred into 1 ml cells and measured at 595nm. With the standards a calibration curve was established and used to calculate the protein concentrations.

2.2.9.3 Far UV spectroscopy

In the far UV, the peptide bonds absorb. This region gives information about the secondary structure of proteins. For the measurements, 400 μl of a 0.5 μM protein solution were used. To study the influence of D-glucose-6-phosphate and D-xylose, the sugars were used at final concentrations of 100 (molar ratio 1:5), 400 (1:1) and 2000 (5:1) nM. Measurements were carried out in 0.2 cm SUPRASIL®quartz cells (HellmaAnalytics) in a J-815 CD Spectrometer (Jasco) at 22°C. CD spectra of the proteins were collected from 260-200 nm as average of 3 scans, smoothed with the Savitsky-Golay method (convolution width of 15) and baseline subtracted in order to exclude buffer influences.

2.2.9.4 Near UV spectroscopy

In the near UV, the aromatic amino acid side chains absorb. This region gives information about the tertiary structure of proteins, particularly of changes in structure between wild-type and mutated protein. For the measurement, 400 μl of a 20 μM protein solution were used. Measurements were carried out in 0,2 cm SUPRASIL®quartz cells (HellmaAnalytics) in a J-815 CD Spectrometer (Jasco) at 22°C. CD spectra of the proteins were collected from 320-260 nm as average of 3 scans, smoothed with the Savitsky-Golay method (convolution width of 15) and baseline subtracted in order to exclude buffer influences.

2.2.9.5 CD data presentation

The data was presented as the mean residue ellipticity $[\theta]$ in deg cm^2 dmol^{-1}, i.e.(millidegrees $\times$ MRW)/(pathlength in mm $\times$ concentration in mg/ml). MRW is the mean residue weight (molecular weight/amount of amino acids - 1) and is different for each protein fragment (Table 2.17).

TABLE 2.17: MRW of all protein fragments

***xyr*1 Fragment**	**M_W**	**AA**	**MRW**
xyr1-1	38.3 kDa	358	107.3 Da
xyr1-2	64.5 kDa	580	111.4 Da
xyr1-3	75.1 kDa	673	111.7 Da
xyr1-4	31.7 kDa	283	112.4 Da
xyr1-5	93.9 kDa	860	109.3 Da
xyr1-6	6.7 kDa	60	113.6 Da

2.2.10 Gene expression and chromatin accessibility analysis

To characterize the wild type and the mutated strain, the transcript levels and the chromatin accessibility of different genes were compared under inducing and repressing conditions.

2.2.10.1 Replacement

The *T. reesei* strains were grown on malt extract agar (MEX). All plates and solutions contain 5 mM uridine, because the wild type strain is missing the *pyr4* gene necessary for pyrimidine biosynthesis. After one week, the spores were harvested and re-suspended in 0.8% NaCl and 0.05% Tween 80 solution. The spores were counted under the microscope and the required amount of 10^6 spores/ml for a final volume of 250 ml was calculated. Adding the spores to a one litre flask containing 250 ml Mandels-Andreotti (MA) media and 10g/l glycerol as carbon source, the mycelium grew for 18 hours with 180rpm at 30°C. For the carbon source replacement, the mycelia was washed with water and re-suspended in MA media containing 1% D-glucose, 0.5 mM D-xylose or 66 mM D-xylose. For control, mycelia was also grown in MA media without any carbon source. After three hours of incubation (180rpm, 30°C), the mycelia was harvested, frozen and stored in liquid nitrogen.

2.2.10.2 RNA extraction

Total RNA from *T. reesei* mycelia was isolated using the PeqGOLD Trifast (Peqlab) reagent. All plastic- and glassware used was RNA free. 1 ml of PeqGOLDD Trifast was transferred to a screw-cap tube containing glass beads. A frozen piece of mycelium was added and homogenised with a Fast Prep-24 Cell hydrolyser (MP Biomedicals) for

30 seconds. Afterwards 200 μl chloroform were added and carefully vortexed. After incubation for 10 minutes at room temperature and centrifugation (5 min, 12000 $\times$ g, RT), 450 μl aqueous phase were transferred into a new eppendorf tube. To precipitate the RNA, the same volume of isopropanol was added and the tubes were carefully mixed. Again the suspension was incubated for 15 minutes at room temperature and centrifuged (10 min, 12000 $\times$ g, 4°C). The pellet was washed with 1 ml ethanol (70%-75%) and again centrifuged under the same conditions. The resulting pellet was air dried and re-suspended in pre-warmed RNA-free water. The total RNA solutions were then stored at -80°C.

2.2.10.3 DNAse digestion

5 μg total RNA were combined with 5 μl 10 $\times$ reaction buffer, 5 μl DNAseI (1U/μg DNA) and 35 μl RNAse free water and incubated for 30 minutes at 37°C. The reaction was stopped by adding 5 μl of stop solution (50 mM EDTA) and the enzyme was inactivated by heating 10 minutes to 65°C.

2.2.10.4 cDNA synthesis

The cDNA is synthesized using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific). The following components and steps shwon in table 2.18 were combined and performed within PCR eppendorf tubes. The final 10 μl were incubated for 5 minutes at 25°C, heated for one hour at 42°C and for 5 minutes at 70°C. Afterwards the cDNA was diluted 1:20 and stored at -20°C. Quantitative PCR was performed according to procedure documentation in section 2.2.2.4.

TABLE 2.18: cDNA syntesis with RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific)

Components	Volume	procedure
RNA DNAse treated	5.0 μl (0.45μg)	
Oligo(dT)$_{18}$ primer (0.5 μg/μl)	0.5 μl	
Random primers (0.2 μg/μl)	0.5 μl	mixing and spinning, 5 min at 65°C, cool on ice, short spin
5 × reaction buffer	2.0 μl	
Ribunuclease inhibitor (20U/μl)	0.5 μl	
dNTPs (10mM)	1.0 μl	
Reverse Transcriptase (200U/μl)	0.5 μl	mixing and spinning, PCR Thermal Cycler
Total volume	10μl	

2.2.10.5 Chromatin digestion and DNA extraction

The frozen mycelia was grinded to powder with mortar and pestle and under liquide nitrogen. The powder was then suspended in ice-cold thawed nuclease digestion buffer (NBD) (250mM sucrose, 60 mM KCl, 15 mM NaCl, 0.05 mM $CaCl_2$, 3 mM $MgCl_2$, 0.5 mM DTT, 15 mM Tris-HCl pH 7.5) to a final concentration of 100 mg mycelium/ml. For the chromatin digestion three eppendorf tubes were filled with 10 μl of NBD. These three tubes served as undigested controls. Three other tubes were filled with 10 μl of DNAse I (Promgea, 1U/μl), the enzyme responsible for the DNA digestion. The digestion started immediately after addition of 100 μl of mycelia suspension into each tube. The mixture was incubated for exactly 2.5 minutes at 37°C. Afterwards the reaction was stopped by adding 100 μl of freshly prepared stop solution (40 mM EDTA, 2% SDS). The resulting mixture was extracted twice with phenol-chloroform (1:1) and once with 1 volume of chloroform. The centrifugations between the extractions were performed with a tabletop centrifuge for 5 minutes at 12000 × g and room temperature. The remaining 125 μl of aqueous phase were treated with 10 μg/ml RNase A at 37°C for 15 minutes, precipitated from sodium acetate 0.3 M pH 5.2 and 2 volumes 100% ethanol. The pellet resulting from centrifugation (15 min, 16000 × g, RT), was washed with 70% ethanol and centrifuged again under the same conditions. The DNA pellet was dried on air and suspended in 50-100 μl 5 mM Tris pH 7.5. The genomic samples were be stored at 4°C. A 0.8 % agarose gel is performed to check the DNA integrity and digestion pattern. For qPCR analysis of the DNAse I treated samples go to section 2.2.2.4

Chapter 3

Results

3.1 characterization of the point mutation in *T.reesei*

For the characterization, the wild type strain and the strain bearing the point mutation $Xyr1_{A824V}$, were cultured in shaking flasks under different inducing and non-inducing conditions. Afterwards the gene expression of *xyn1*, *xyn2xyr1*, *cbh1* and *cbh2*, as well as the chromatin accessibility in *xyn1* and *xyn2* have been analysed.

3.1.1 Transcription level analysis

With the quantitative real-time PCR the relative transcript levels of *xyn1*, *xyn2*, *xyr1*, *cbh1* and *cbh2* were measured. These genes were expressed in the presence of 1% D-glucose, low (0.5mM) and high (66mM) concentrations of D-xylose and without any carbon source. The following figures show the relative transcript of each gene in logarithmic form against the carbon source in the culture medium (Figures 3.1 to 3.5).

For *xyr1* there were no big differences observed between the two strains, as shown in figure 3.1. D-glucose led to a repression of transcription as well as high concentrations of D-xylose. Lower concentrations of xylose (0.5mM), which is normally associated with xylanases induction, does not promote *xyr1* induction. The two genes encoding for cellobiohydrolases in figure 3.2 and 3.3 show the same pattern compared to the transcription of *xyr1*. Comparing the two strains showed that the expression of *cbh1* and *cbh2* was a little bit more increased in the mutated strain in the presence of low concentrated D-xylose. The two transcript levels of the xylanase genes in figure 3.4 and 3.5 have shown a different pattern compared to the transcription of *xyr1*. For both genes, the transcript level in the medium without any carbon source was clearly elevated

in the mutated strain whereas the wild type strain revealed less transcript.For *xyn1* the transcript formation with D-glucose could not be detected. Figure 3.4 shows for low concentrations of D-xylose an increased transcript formation in both strains but for the mutated strain the increase was stronger. And again under repression conditions with high concentrations of D-xylose, the transcript decreased for both strains but less pronounced in the mutated strain. For *xyn2*, D-glucose again repressed the expression leading to an decreased gene transcript. Figure 3.5 shows for low concentrations of D-xylose an highly elevated transcript formation in both strains but for the mutated strain the increase was slightly stronger. With high concentrations of D-xylose, the transcript decreased for both strains but there was still more transcript for the mutated strain.

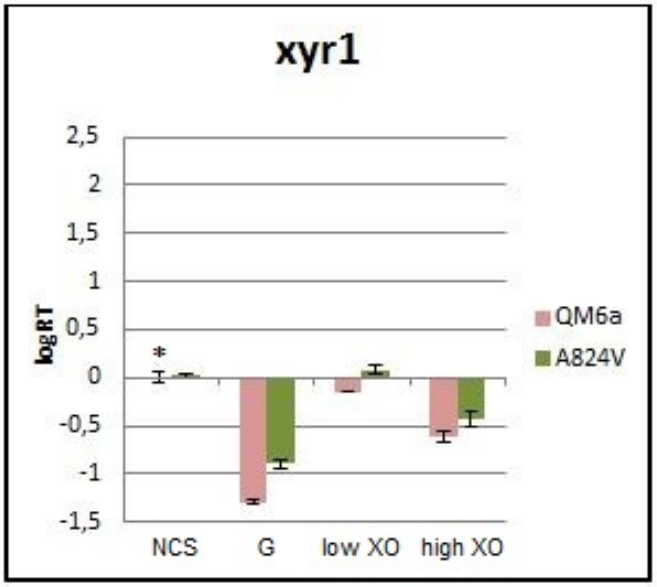

FIGURE 3.1: relative transcript level of *xyr1* under inducing and repressing conditions. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose. $*$ indicates the reference sample.

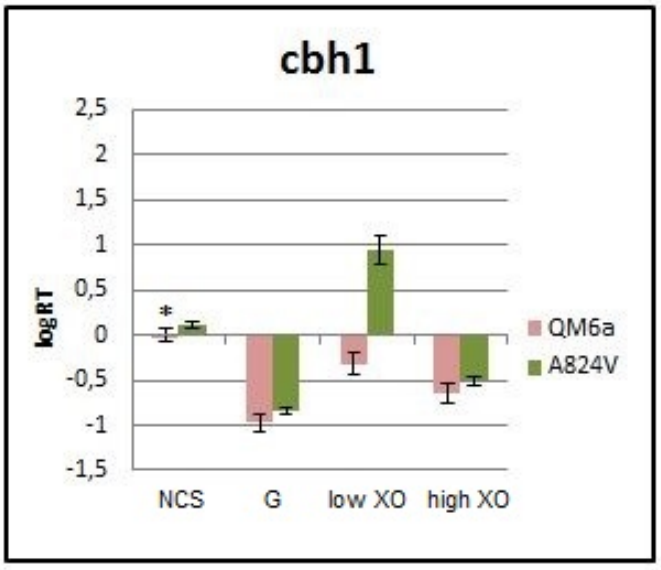

FIGURE 3.2: relative transcript level of *cbh1* under inducing and repressing conditions. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose. $*$ indicates the reference sample.

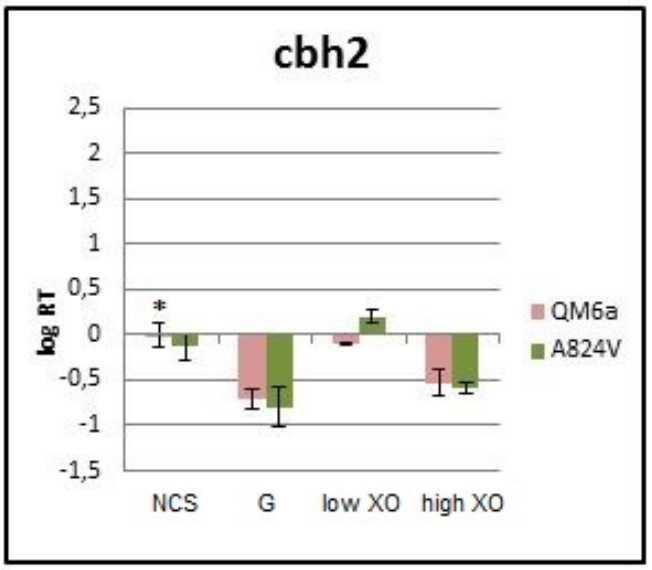

FIGURE 3.3: relative transcript level of *cbh2* under inducing and repressing conditions. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose. $*$ indicates the reference sample.

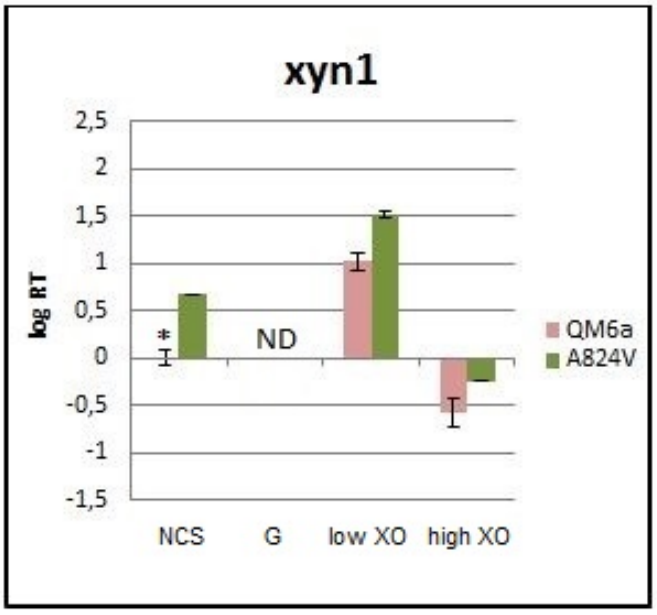

FIGURE 3.4: relative transcript level of *xyn1* under inducing and repressing conditions. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose. ∗ indicates the reference sample, ND means not detected.

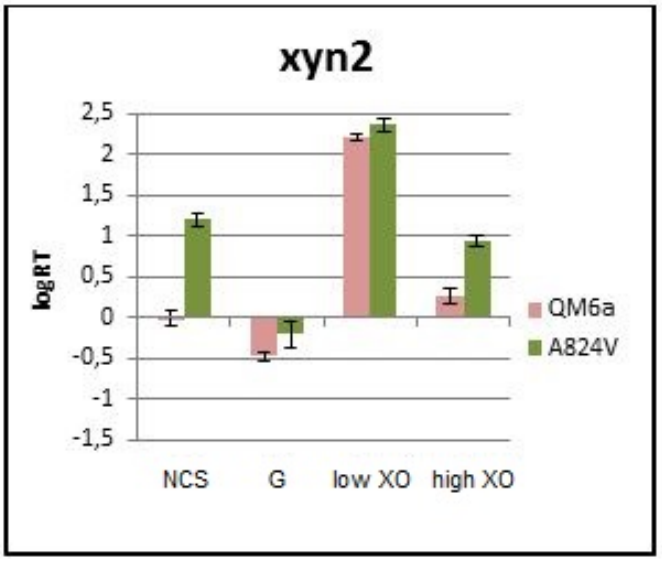

FIGURE 3.5: relative transcript level of *xyn2* under inducing and repressing conditions. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose. ∗ indicates the reference sample.

3.1.2 Chromatin accessibility analysis

As far as the most striking differences between the strains was found for the xylanases-encoding genes and thus our department shorty published a Xyr1 influence in chromatin remodelling of the *xyr1*-regulon (de Sousa, 2015), I decided to check the chromatin accessibility status of *xyn1* and *xyn2*. Because maybe the transcript differences found could be due more accessible chromatin in the strain bearing the point mutation. For the chromatin accessibility real-time PCR (CHART-PCR) the same samples as for the transcript analysis were used. The following figures 3.6 to 3.10 show the relative accessibility index (RAI) of different promoter regions in *xyn1* and *xyn2* against the carbon source in the culture medium. An higher RAI stands for a more condensed chromatin structure in the target region.

All target regions in *xyn1* and *xyn2* revealed for both strains a condensed chromatin structure unter neutral conditions. Under repressing conditions, with D-glucose and with high concentrated D-xylose, the chromatin was even more closed. Only with the inducing sugar, low concentrations of D-xylose, the chromatin was slightly more accessible. As for differences between the two strains, the figures 3.6 to 3.10 show that the chromatin structure in the mutated strain is for all conditions slightly more condensed or equal.

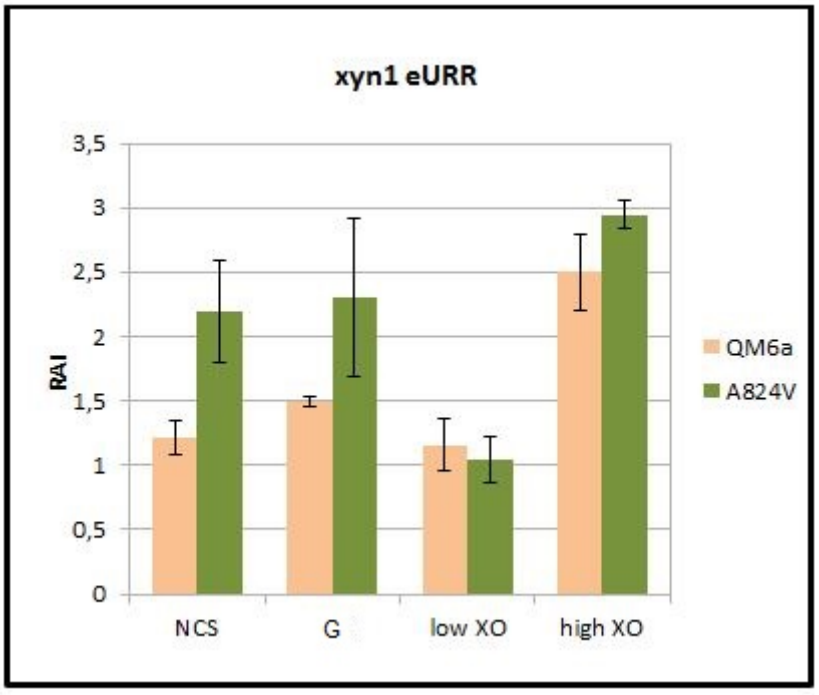

FIGURE 3.6: RAI of the extreme upper regulatory region (eURR) in *xyn1*. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose.

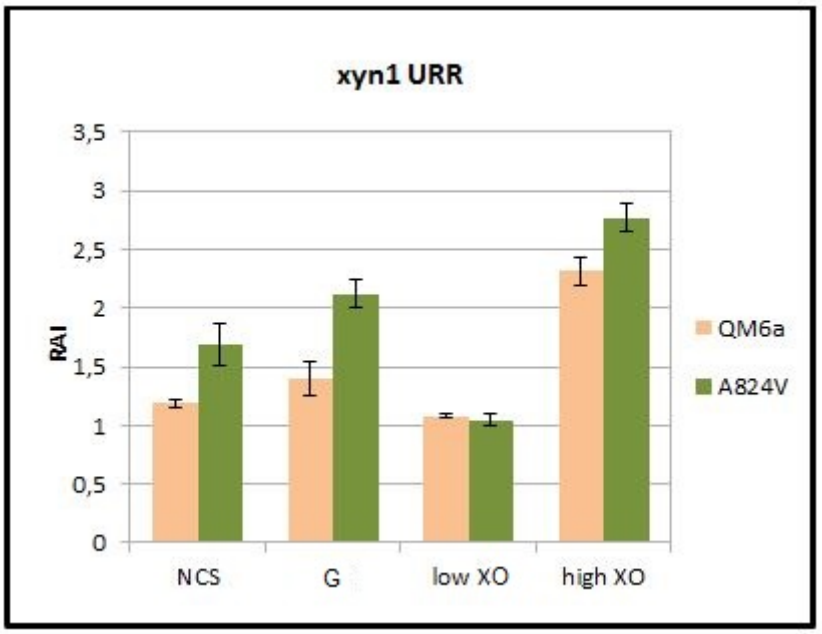

FIGURE 3.7: RAI of the upper regulatory region (URR) in *xyn1*. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose.

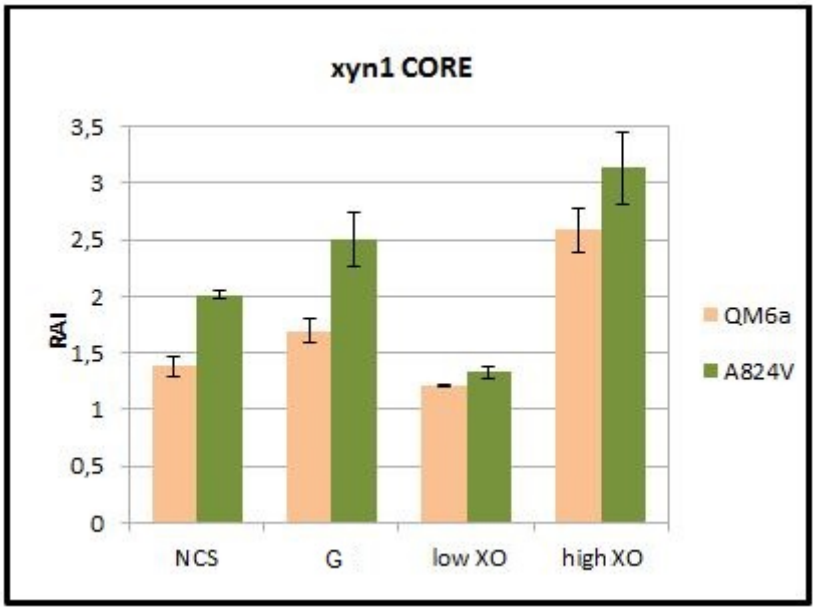

FIGURE 3.8: RAI of the core region in *xyn1*. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose.

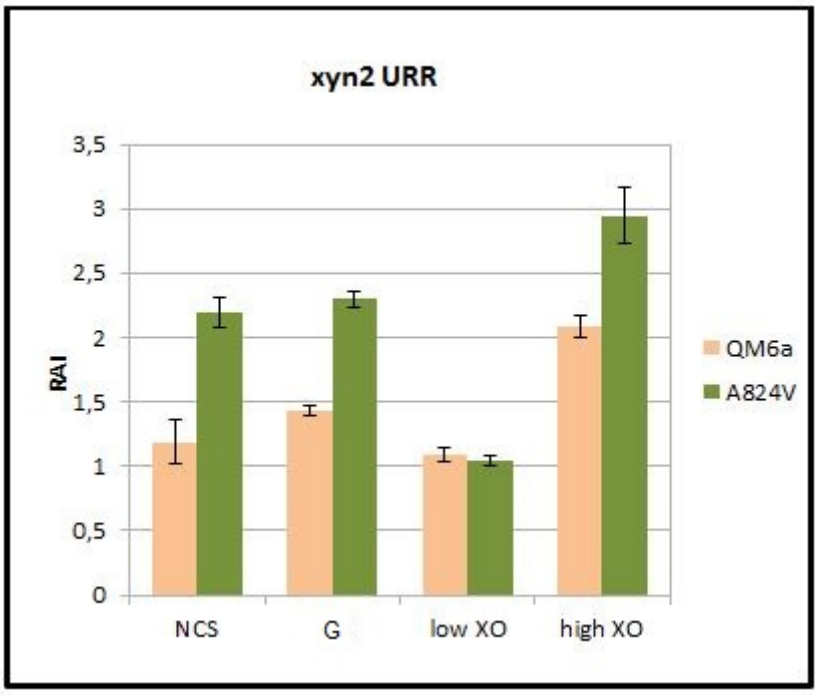

FIGURE 3.9: RAI of the upper regulatory region (URR) in *xyn2*. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose.

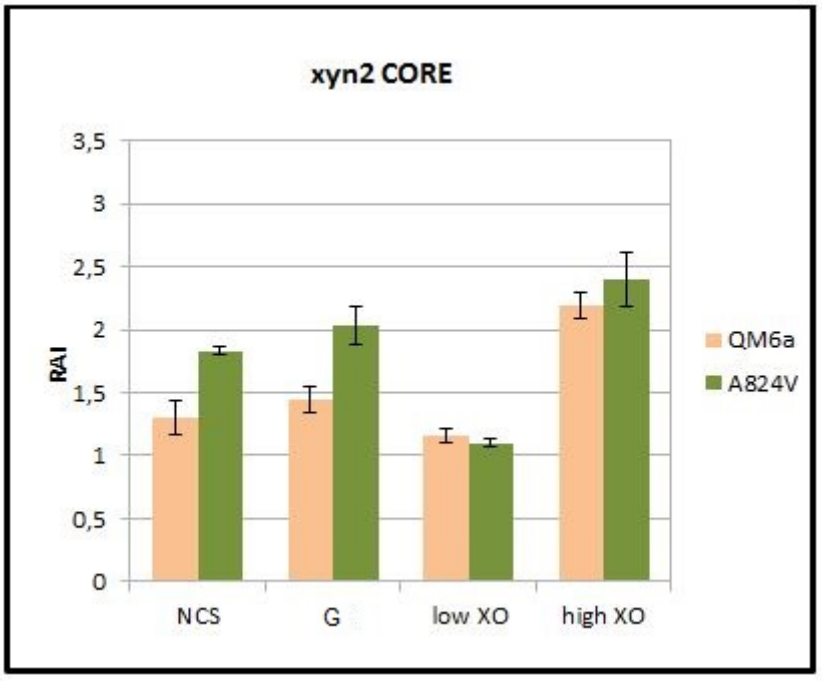

FIGURE 3.10: RAI of the core region in *xyn2*. pink: wild type strain, green: mutated strain. NCS: no carbon source, G: 1% D-glucose, low XO: 0.5 mM D-xylose, high XO: 66 mM D-xylose.

As the figures revealed, the differences in transcript analysis seem not to be caused by chromatin accessibility differences in the strains. Therefore I will try to analyse possible impacts of the point mutation on the Xyr1 protein structure.

3.2 Investigation of changes in secondary and tertiary structure

3.2.1 Primer testing and PCR optimization

First the designed and ordered primers for all nine fragments (2.1), encoding for different versions of Xyr1, needed to be tested. This was achieved with a standard PCR followed by agarose gel electrophoresis. The figure 3.11 shows that there are amplicons for each fragment with the right size. Afterwards the PCR was repeated with the actual working Phusion enzyme to find the optimal PCR conditions.

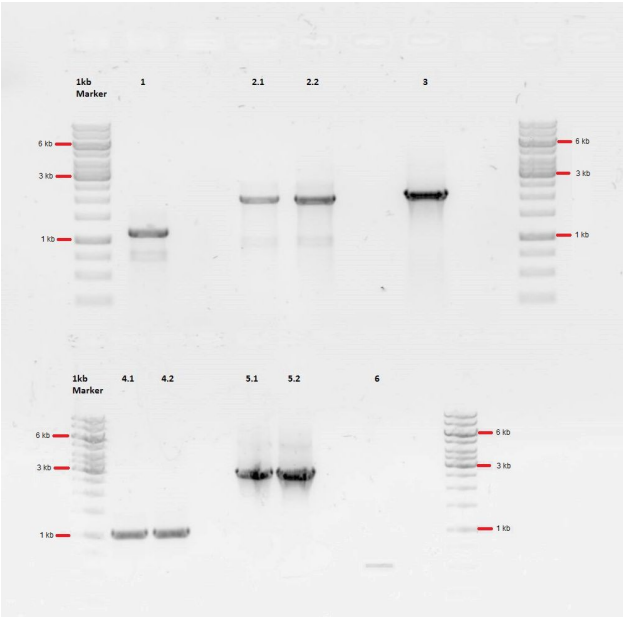

FIGURE 3.11: agarose gelelectrophoresis of all nine fragments from standard PCR with goTaq Polymerase. 1: 1.02kb, 2.1 and 2.2: 1.6kb, 3: 2.02kb, 4.1 and 4.2: 0.8kb, 5.1 and 5.2: 2.5kb, 6: 0.3kb

3.2.2 Fragment ligation into pJET1.2 vector

After the PCR-produced fragments have been purified and digested with the appropriate restriction enzymes (Table 2.5), the fragments were ligated into pJET1.2 and

transformed into competent *E.coli* TOP10 cells. Afterwards a Colony PCR of several clones was performed to screen possible successful transformants (Figure 3.12). Then the plasmids were extracted from a few clones and a mini digestion was performed to confirm successful cloning. (Figure 3.13). The plasmids showing the right digestion band were digested again and the respective fragments were purified for further cloning in the expression vector (pTS1 or pTS2). The following figures show the results for the fragment Xyr1-4.1 as an example. For all other fragments the process was identical and the results were correct.

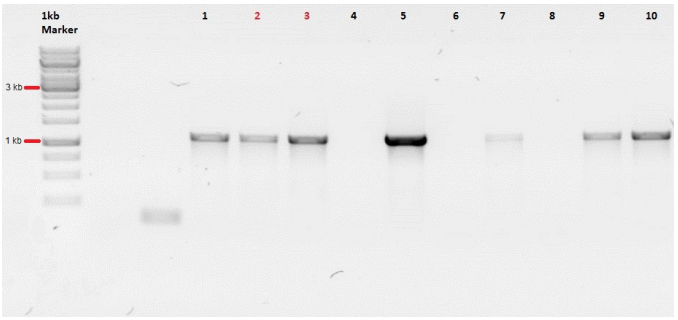

FIGURE 3.12: agarose gel electrophoresis of fragment Xyr1-4.1 (0.8 kb) from colony PCR. Transformants 1, 2, 3, 5, 7, 9 and 10 are positive. Plasmid extraction occurred from transformants 2 and 3

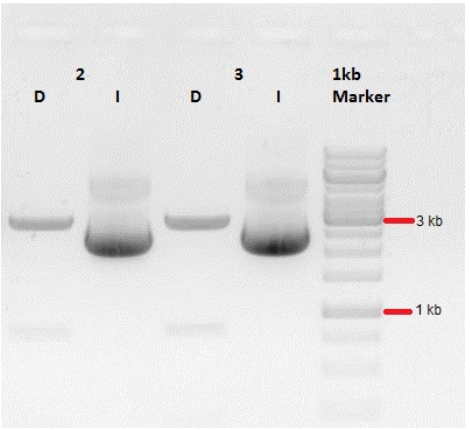

FIGURE 3.13: agarose gel electrophoresis of transformants 2 and 3 from fragment Xyr1-4.1 (0.8 kb) after digestion (D: digested plasmid, I: intact plasmid). All plasmids contained the fragment Xyr1-4.1

3.2.3 Fragment ligation into pTS1 and pTS2

The digested fragments from pJET2.1 were cut out and ligated in the final expressing plasmid (pTS1 or pTS2) and transformed again with competent *E.coli* TOP10 cells.

Afterwards a colony PCR of several clones was performed (Figure 3.14). Then the plasmids were extracted from a few clones and a digestion was performed from one of them (Figure 3.15). The plasmids showed the right digestion bands and were then sent for sequencing. The sequences are all shown in the appendix (C). The following figures show the results for the fragment Xyr1-4.1 as an example. For all other fragments the process was performed identically and the results were correct.

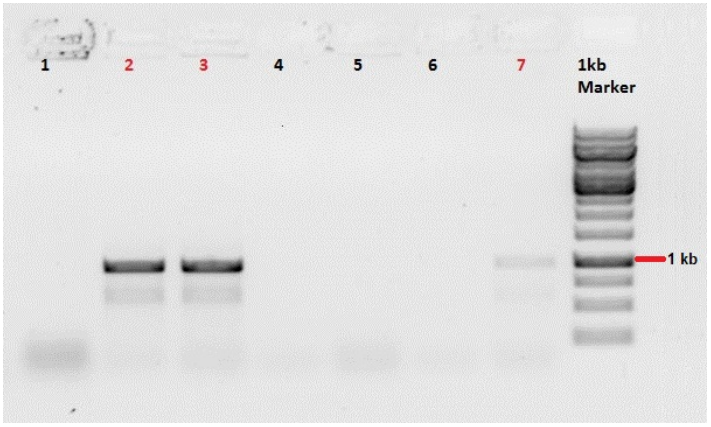

FIGURE 3.14: Agarose gel electrophoresis of fragment Xyr1-4.1 (0.8 kb) from colony PCR. Transformants 2,3 and 7 are positive. Plasmid extraction occured from transformants 2,3 and 7

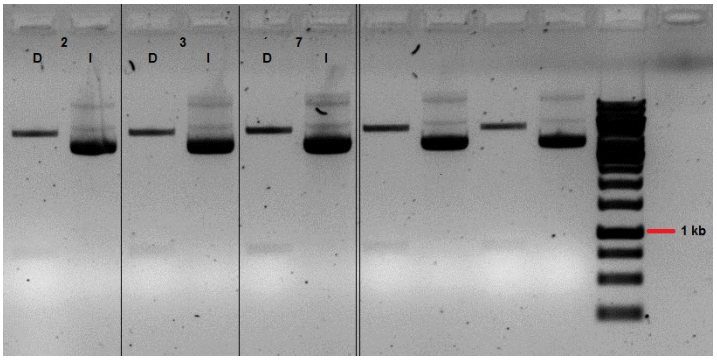

FIGURE 3.15: Agarose gel electrophoresis of fragment Xyr1-4.1 (0.8 kb) after digestion (D: digested plasmid, I: intact plasmid). All plasmids contained the fragment Xyr1-4.1

3.2.4 Transformation into BL21(DE3)

The plasmids for the expression of the Xyr1 fragments were then transformed into the *E.coli* production strain BL21(DE3). From several transformants a colony PCR was performed in order to confirm the transformation and presence of the right plasmid

(Figure 3.16). Afterwards a positive transformant of each versions was picked for protein production.

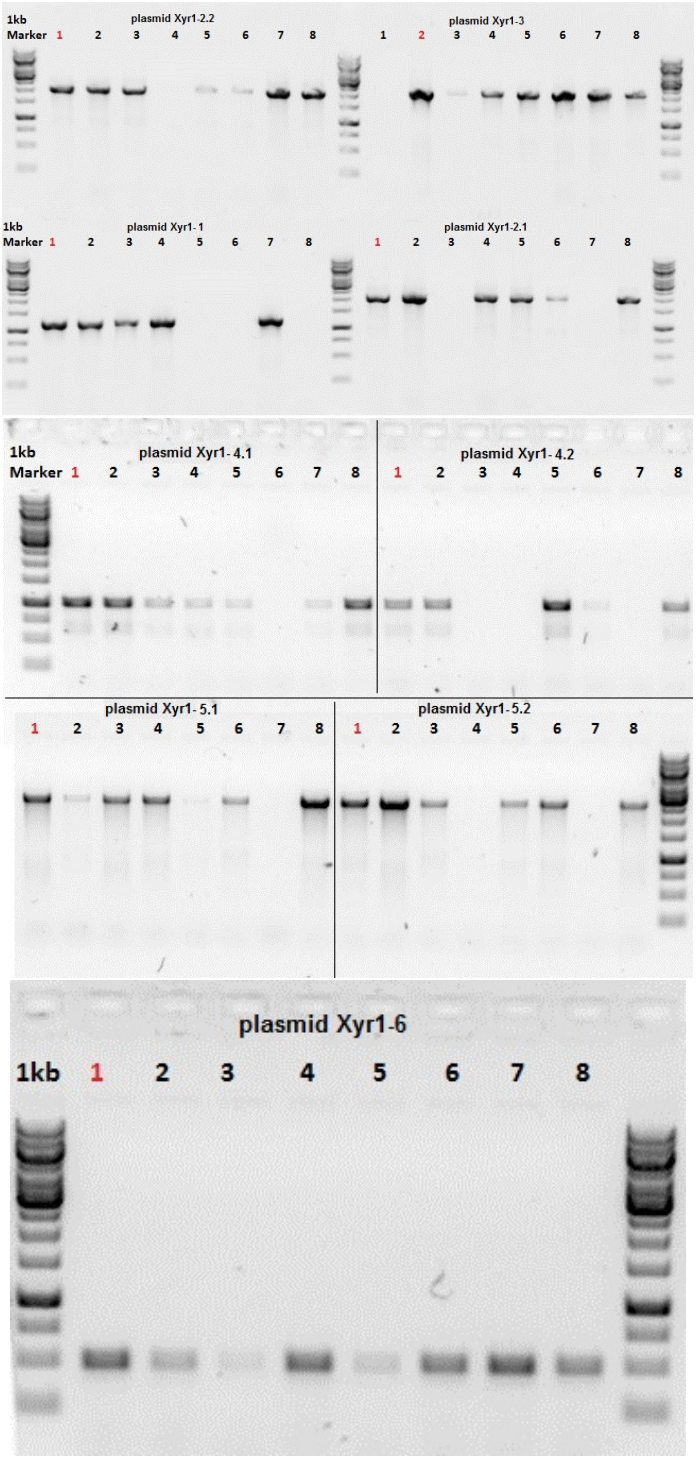

FIGURE 3.16: agarose gel electrophoresis of all plasmids from colony PCR. The red marked numbers are the positive transformants picked for the following protein production

3.2.5 Protein production

The BL21(DE3) transformants carrying the plasmids of interest were cultured and induced to produce the nine different Xyr1 proteins. The proteins were purified by affinity chromatography and afterwards verified via SDS-PAGE (Figure 3.17 to 3.24). All proteins were properly produced.

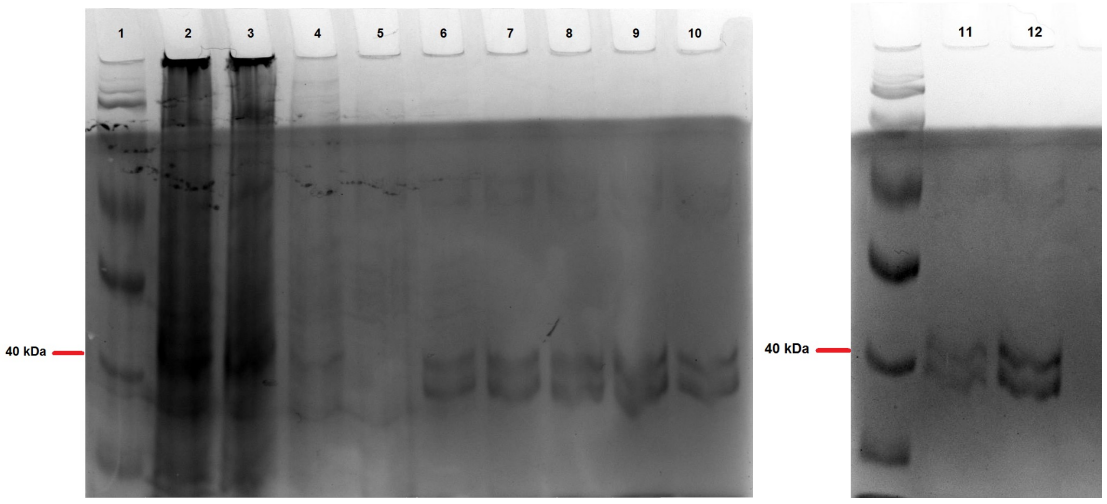

FIGURE 3.17: SDS-PAGE of protein fragment Xyr1-1 (38.3 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6, 12: eluate 7.

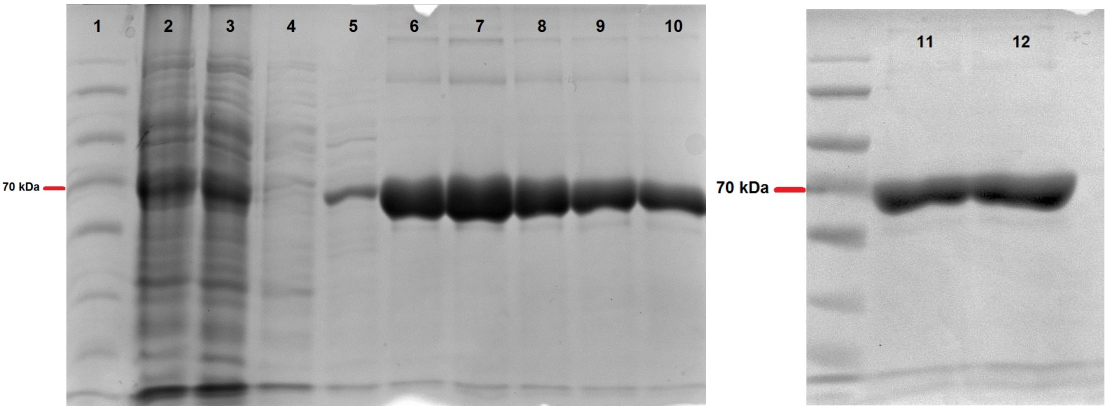

FIGURE 3.18: SDS-PAGE of protein fragment Xyr1-2.1 (64.5 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6, 12: eluate 7.

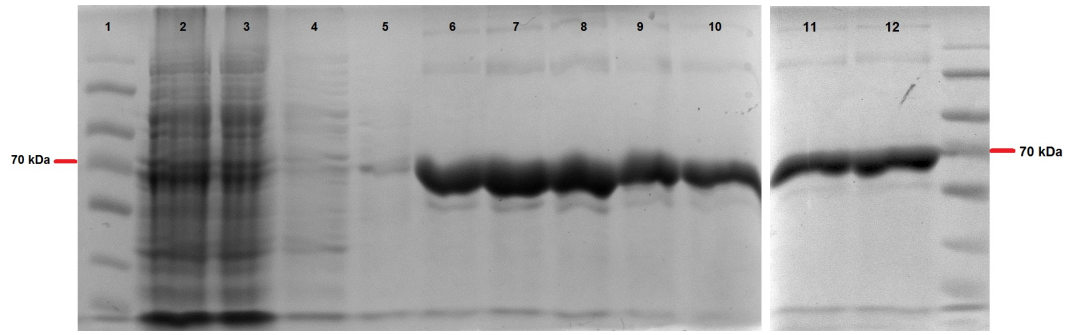

FIGURE 3.19: SDS-PAGE of protein fragment Xyr1-2.2 (64.5 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6, 12: eluate 7.

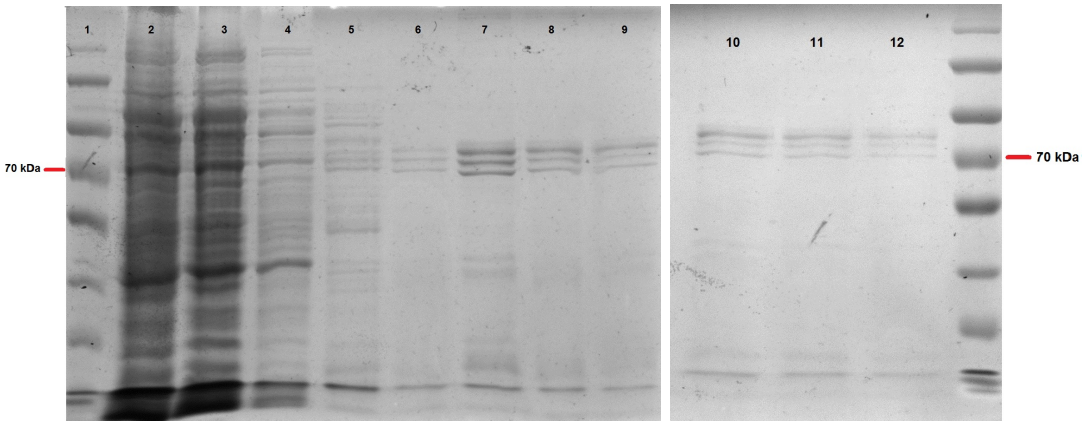

FIGURE 3.20: SDS-PAGE of protein fragment Xyr1-3 (75.1 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6, 12: eluate 7.

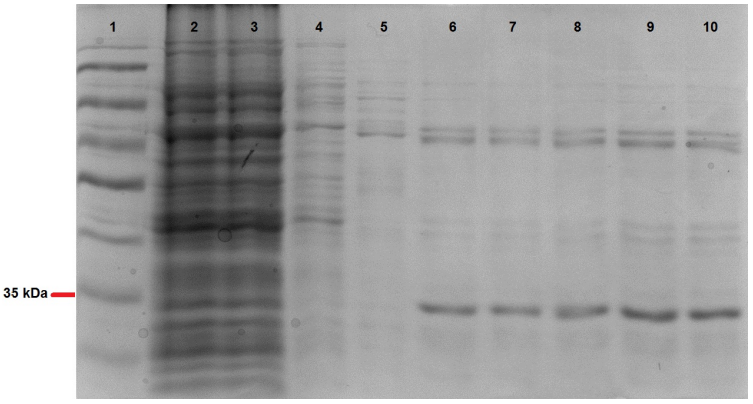

FIGURE 3.21: SDS-PAGE of protein fragment Xyr1-4.1 (31.7 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5.

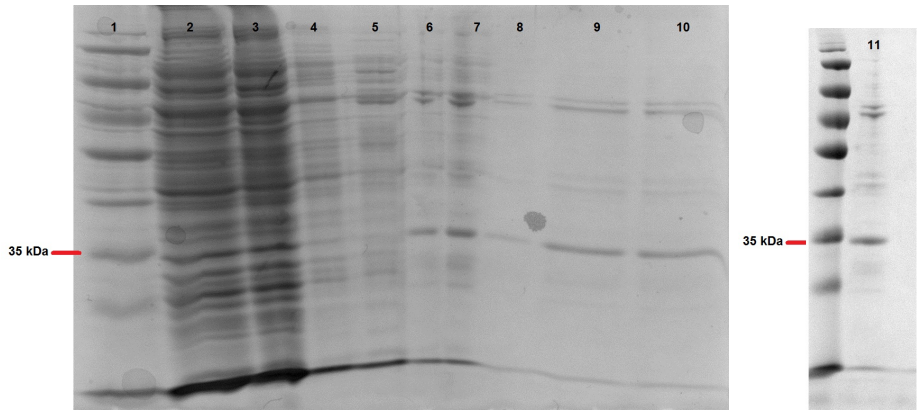

FIGURE 3.22: SDS-PAGE of protein fragment Xyr1-4.2 (31.7 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6.

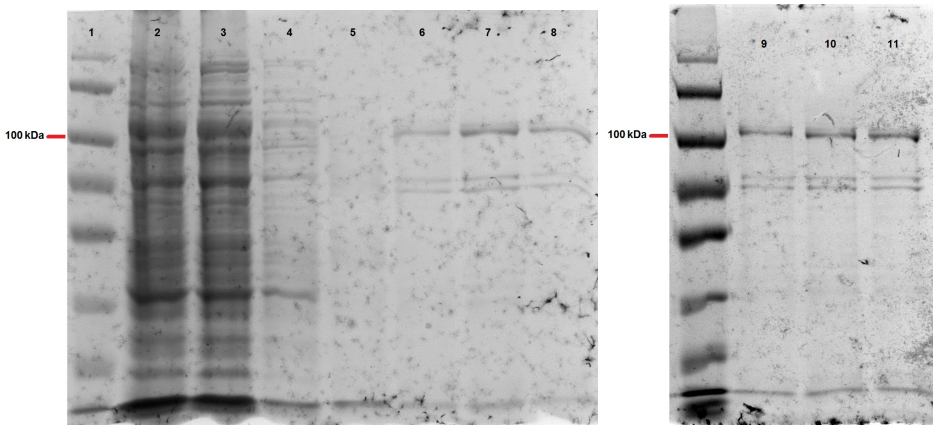

FIGURE 3.23: SDS-PAGE of protein fragment Xyr1-5.1 (93.9 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6.

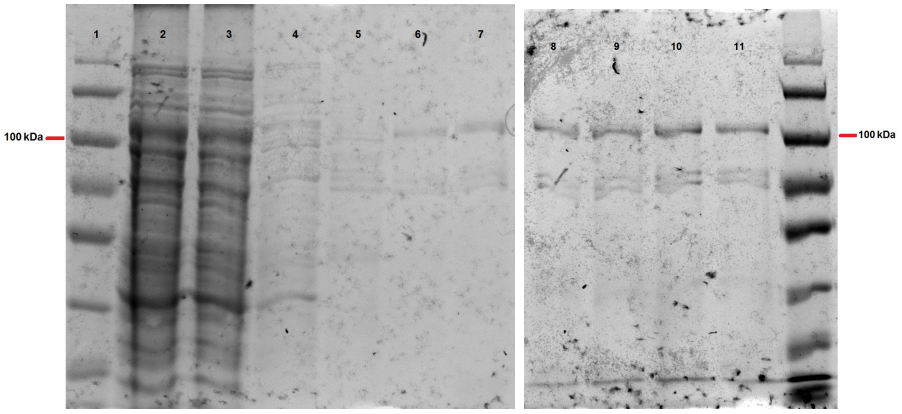

FIGURE 3.24: SDS-PAGE of protein fragment Xyr1-5.2 (93.9 kDa). 1: Marker, 2: cell lysate, 3: flow-through, 4: wash 1, 5: wash 2, 6: eluate 1, 7: eluate 2, 8: eluate 3, 9: eluate 4, 10: eluate 5, 11: eluate 6.

3.2.6 CD spectroscopy

The protein elution buffer (WB200) comprises imidazole which absorbs at far UV. This makes the buffer unsuitable for the following CD measurements of the proteins. Prior to the measurements the buffer needed therefore to be exchanged to the CD working buffer (50 mM Tris, 200 mM NaCl, 50 mM NaH_2PO_4, 10% glycerol pH 7,5). After all proteins were available in the right buffer, the actual analysis for the secondary structure was performed.

The following figures 3.25 to 3.33 show the change of absorbance in the range of 200 to 260 nm for the designed proteins under usually inducing and repressing conditions of xylanases expression. They give information about changes in the secondary structure of the proteins. There are no significant changes under inducing (xylose) and repressing (glucose) conditions (Figure 3.25 to 3.33). The presence of glucose-6-phosphate or D-xylose does not cause changes in secondary structure when the DNA-binding domain plus the putative dimerization domain is isolated from the rest of the protein (Xyr1-1) (Figure 3.25 to 3.27). This same absence of structural changes, when incubating the proteins with different sugars was observed for all other versions of Xyr1 (Xyr1-3, Xyr1-4.1, Xyr1-4.2, Xyr1-5.1, Xyr1-5.2, Xyr1-6) (Figures 3.28 to 3.33). Nevertheless, the point mutation A824V caused local structural changes when the posterior part of the FTFMHR was analysed together with the AD-region (compare figures 3.29 to 3.30). Additionally, by comparing CD spectra of Xyr1 versions 5.1 and 5.2, in which the AD-region is missing, no secondary structural change was observed (compare figures 3.31 to 3.32). Hence, it seems that the structural changes imposed by the point mutation directly influences the folding of the AD-region. On the other hand, at a first glance, the point mutation seems not to interfere in secondary structure contend when only the DNA-binding domain of the protein is missing (Xyr1 versions 2.1 and 2.2) (compare figures 3.26 to 3.27). However an extended CD spectroscopy analysis was performed in order to explore the absorbance changes in the range of 260 - 320 nm, therefore accessing tertiary structure changes. The analysis revealed that these two versions of Xyr1 have indeed differences in their three-dimensional structures (Figure 3.34). Because of their size, the CD spectra for the proteins Xyr1.1 and Xyr1.6 do not really look like common protein curves any more (Figures 3.25 and 3.33). They are to small to reveal a huge amount of secondary structure content.

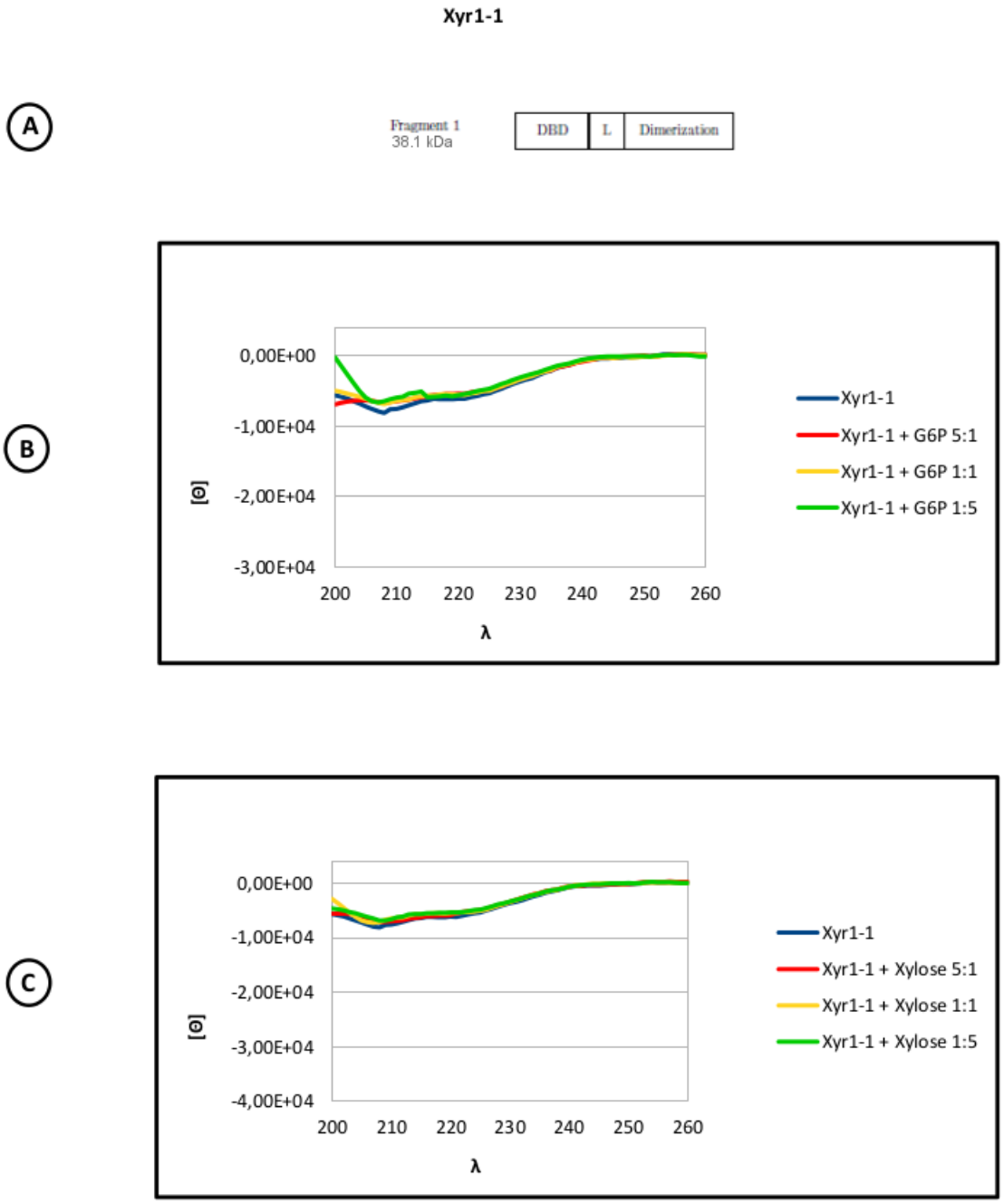

FIGURE 3.25: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-1 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-1 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

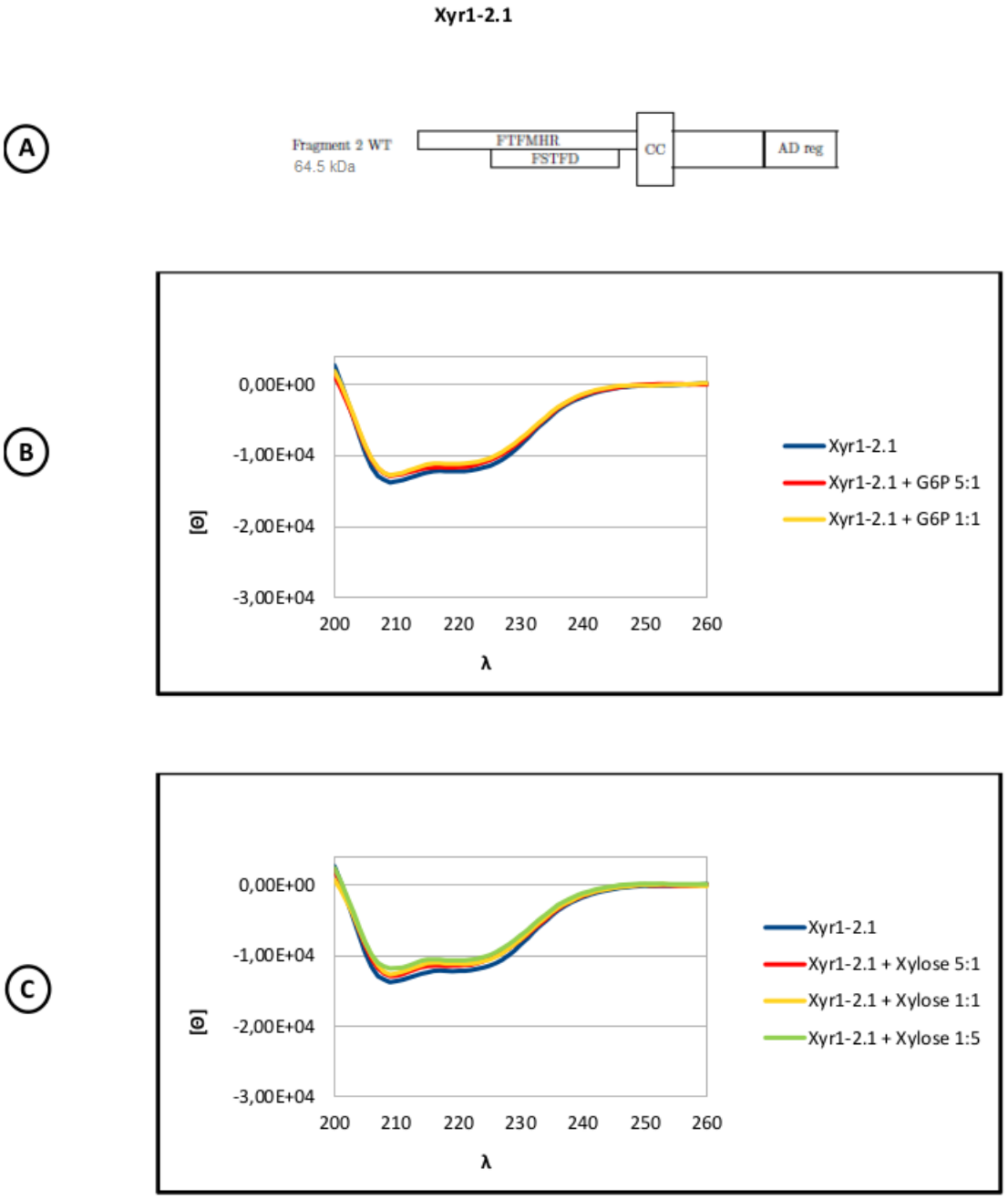

FIGURE 3.26: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-2.1 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-2.1 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

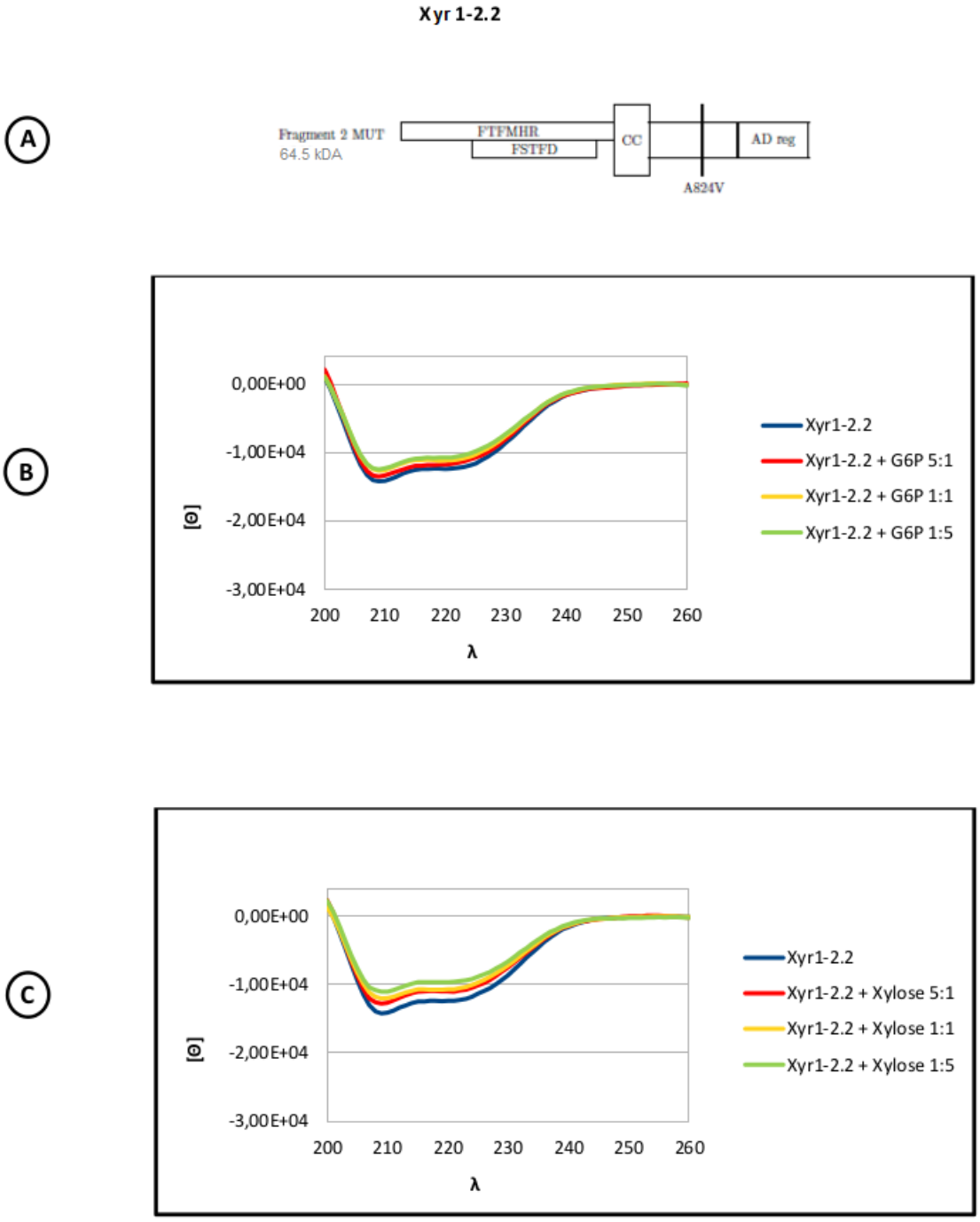

FIGURE 3.27: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-2.2 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-2.2 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

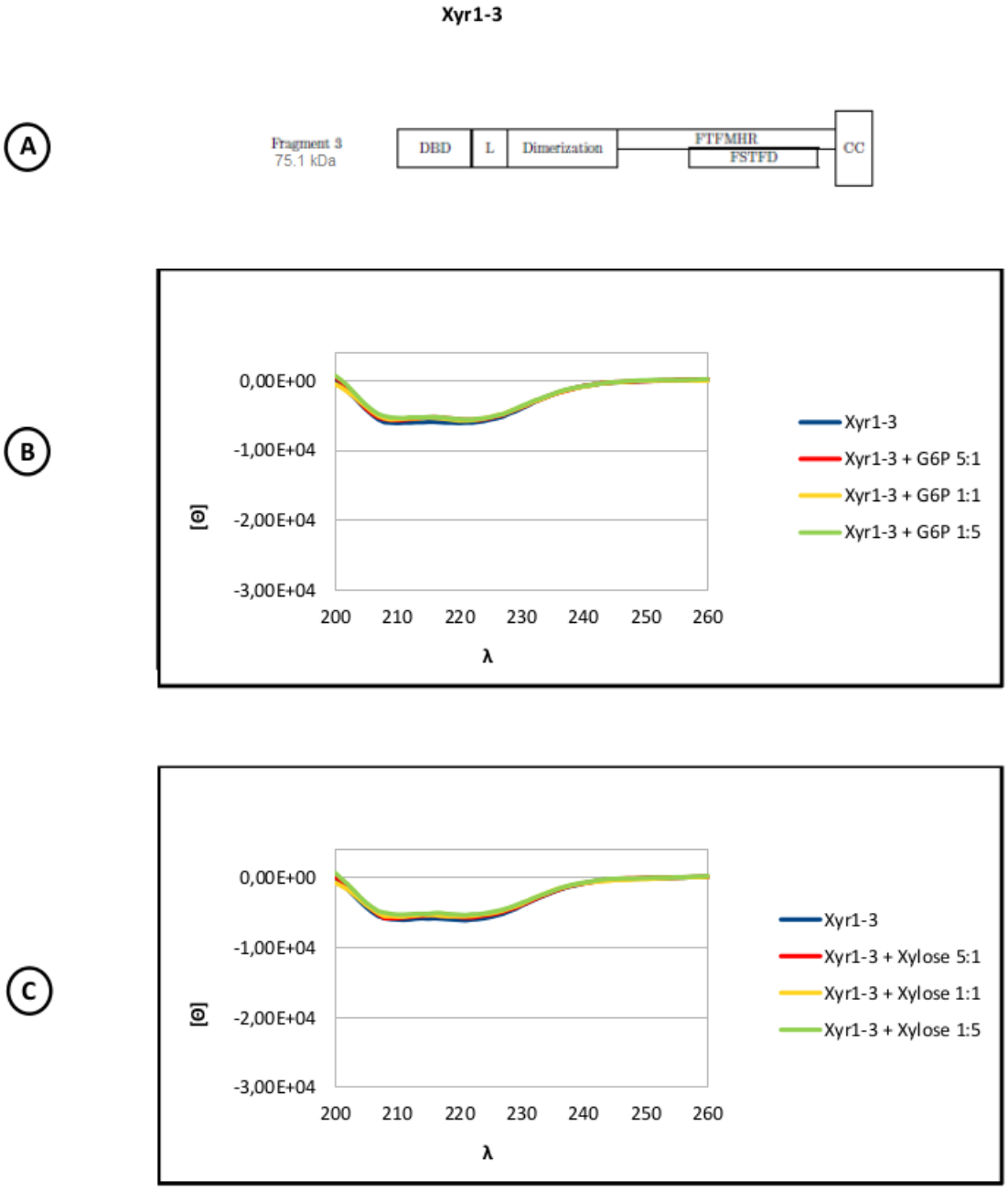

FIGURE 3.28: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-3 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-3 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

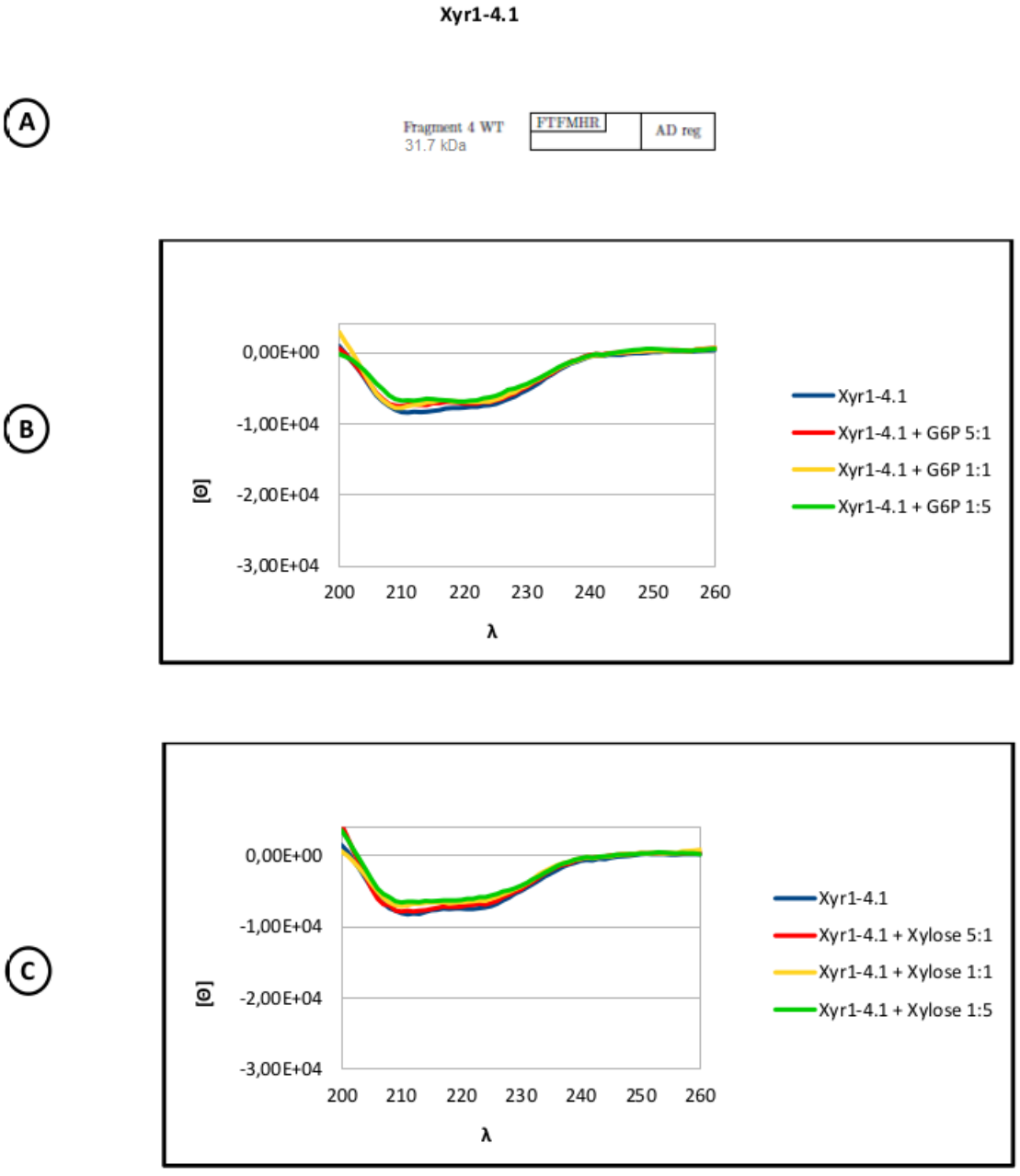

FIGURE 3.29: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-4.1 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-4.1 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

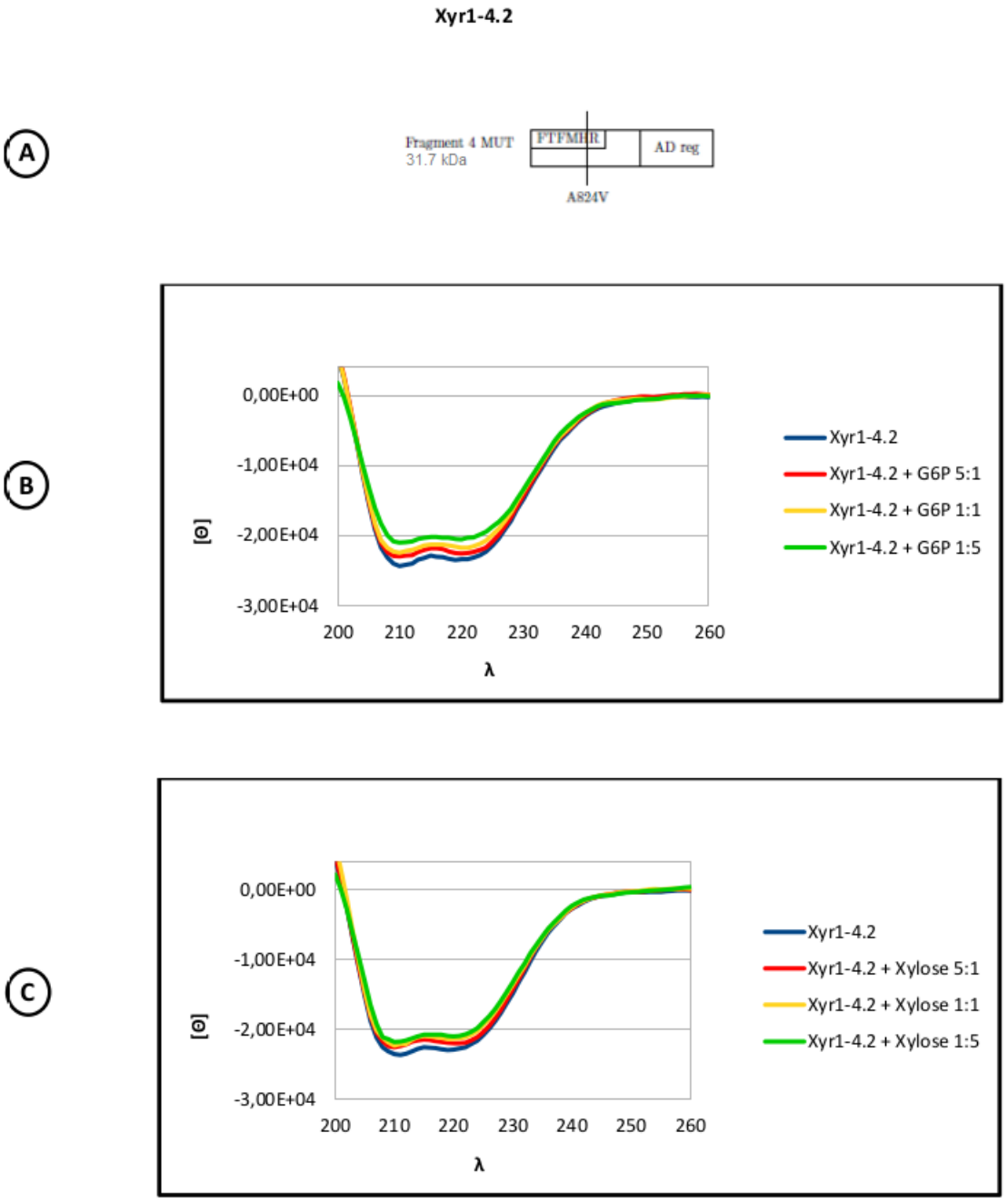

FIGURE 3.30: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-4.2 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-4.2 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

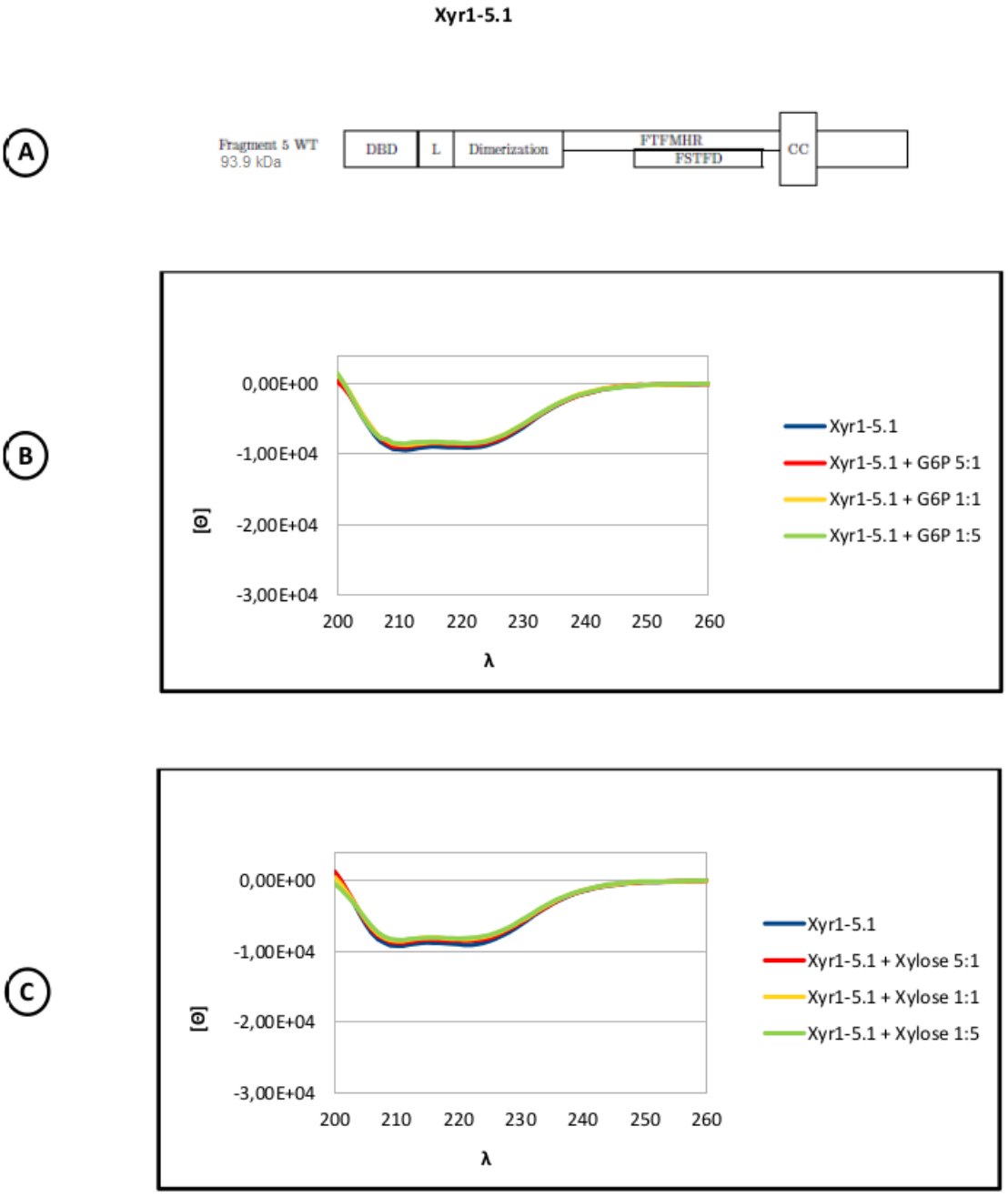

FIGURE 3.31: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-5.1 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-5.1 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

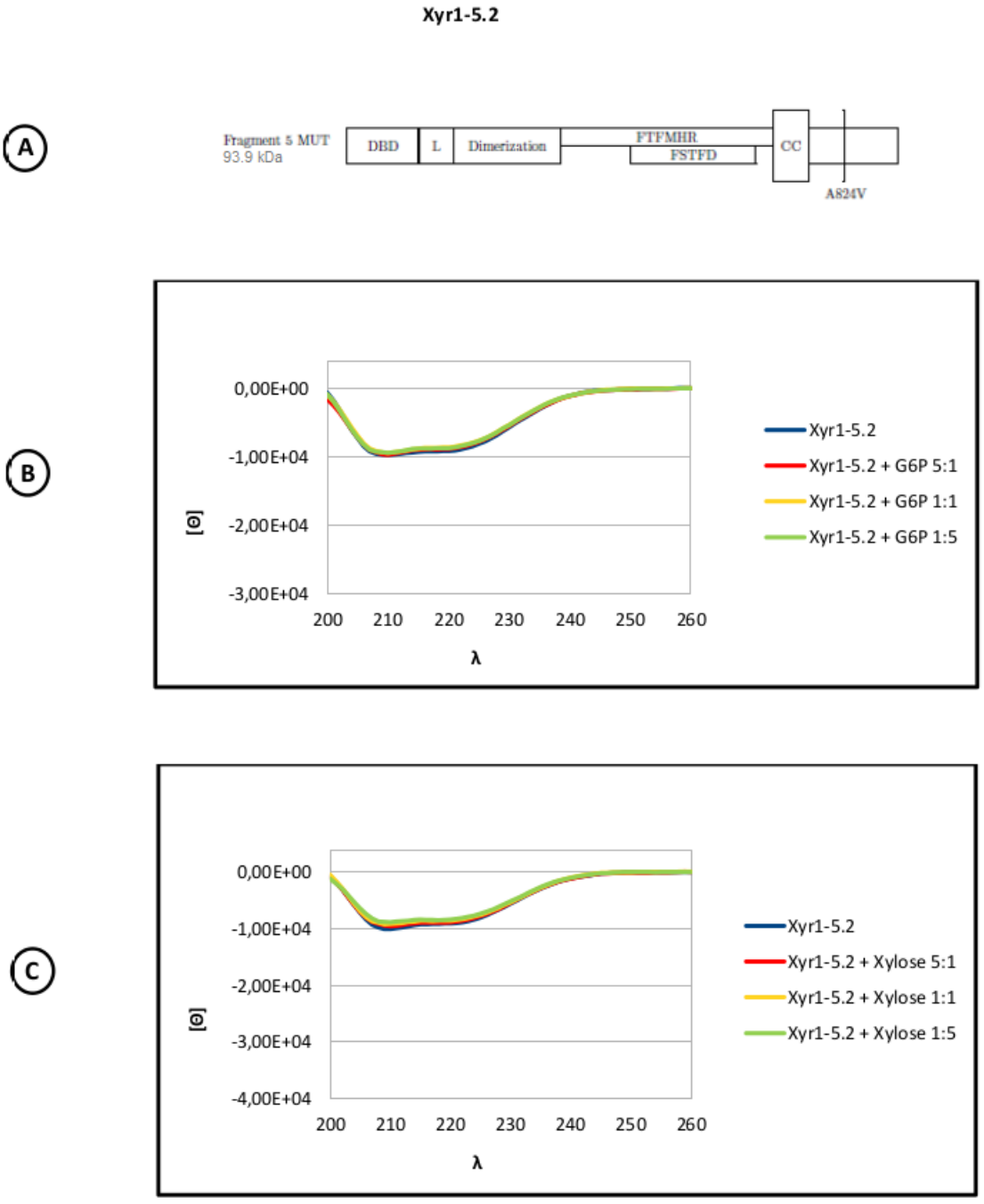

FIGURE 3.32: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-5.2 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-5.2 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

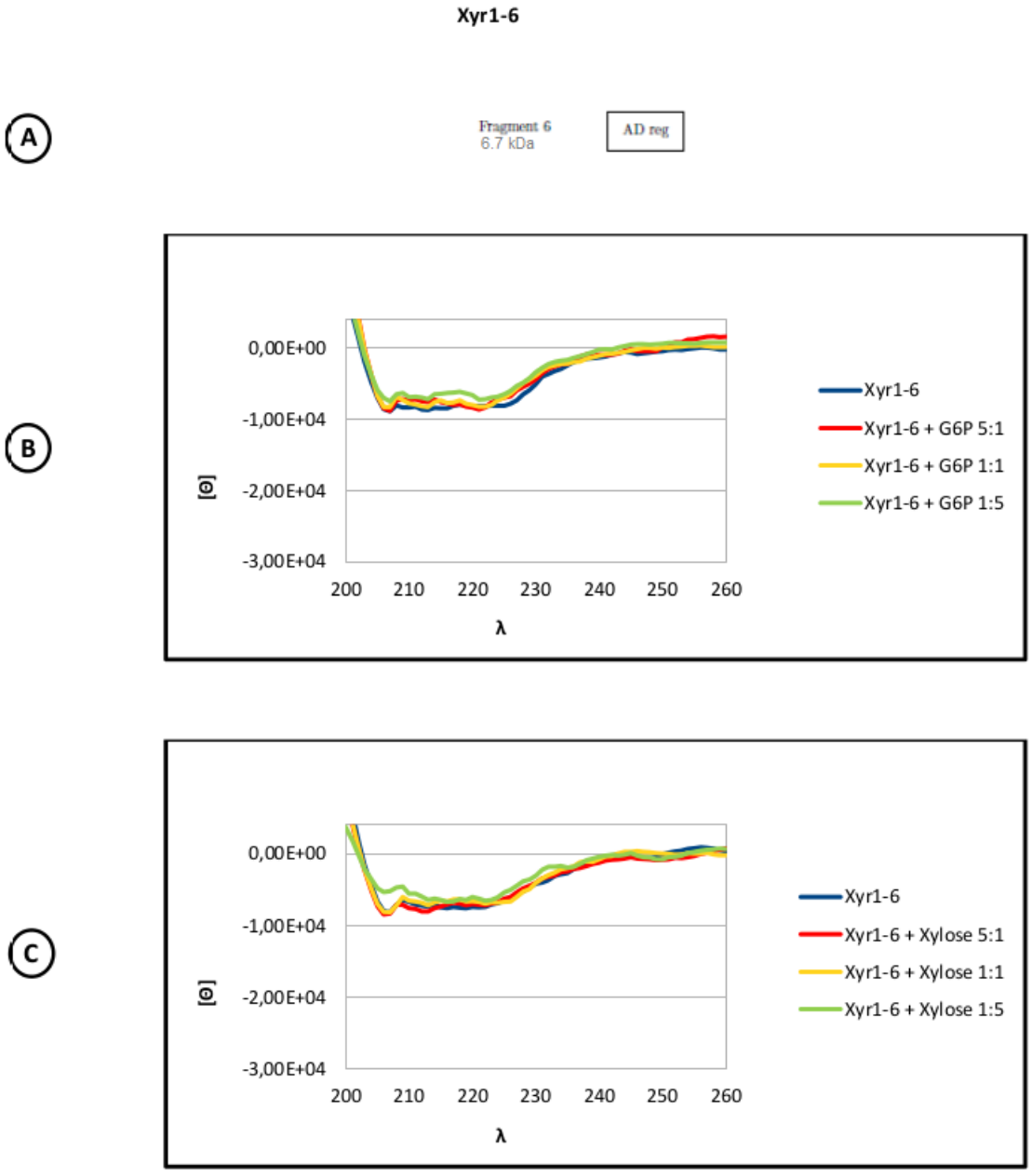

FIGURE 3.33: A: protein map, B: CD spectrum in the range of 200nm-260nm from the protein Xyr1-6 under repressing (G6P: glucose-6-phosphate) conditions, C: CD spectrum in the range of 200nm-260nm from the protein Xyr1-6 under inducing (xylose) conditions. Blue: protein without sugar, red: protein with sugar 5:1, yellow: protein with sugar 1:1, green: protein with sugar 1:5

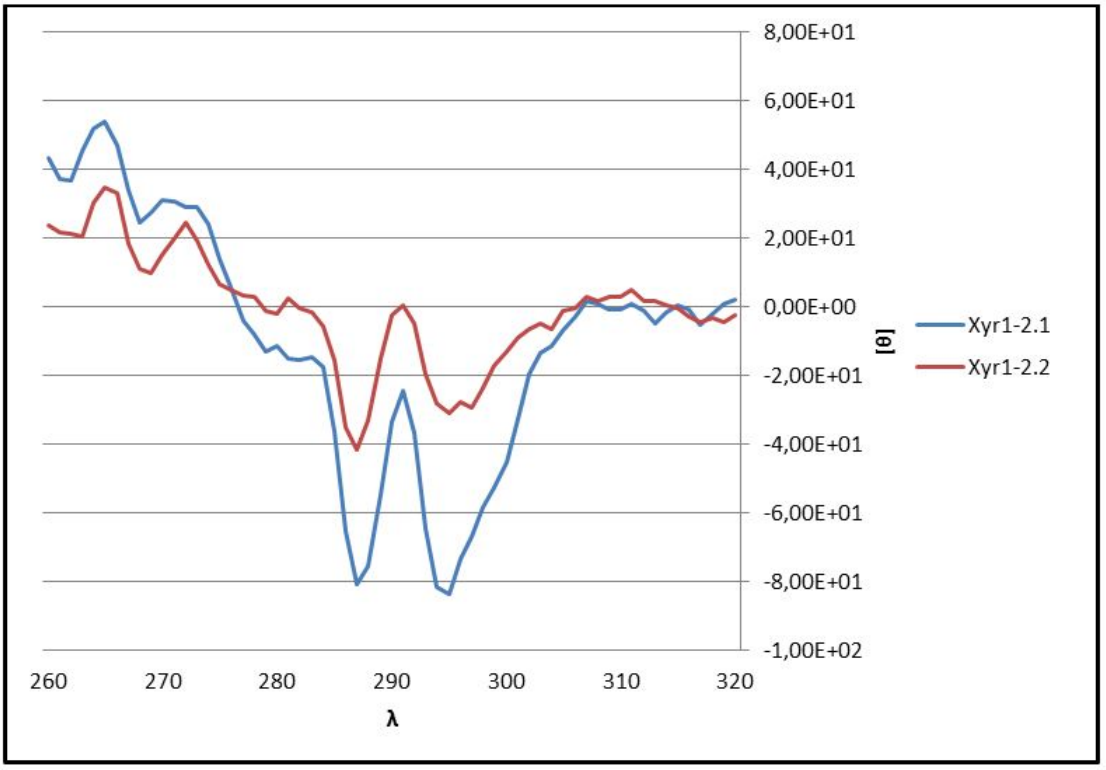

FIGURE 3.34: near UV CD spectrum (260-320 nm) of wild type and mutated version of Xyr1-2. Blue: Xyr1-2.1, red: Xyr1-2.2

Chapter 4

Discussion

The transcription analysis showed that there are differences between the two strains concerning the transcript level with low concentrations of D-xylose. The mutated strain forms a lot more of the genes *xyr1*, *cbh1* and *cbh2* compared to the wild type strain under the same condition. Nevertheless, the transcript formation of the cellulase-encoding follows strictly the expression of *xyr1*. In contrast to this, the expression of the two xylanase-encoding genes *xyn1* and *xyn2* in the mutated strain, does not follow the *xyr1* transcript level. Under non-inducing (no carbon source), inducing (low concentration of D-xylose) and repressing conditions (high concentration of D-xylose) the transcript levels of both genes are clearly elevated in the mutated strain in contrast to the transcription activator *xyr1*. These findings correspond perfectly to the findings from Derntl and co-workers (Derntl et al., 2013). In the strain, bearing the point mutation, the xylanase gene expression is apparently partially deregulated and the transcript formation is higher for all major cellulases and xylanases. But as far as the transcript levels of *xyr1* do not differ between the two strains, the xylanases deregulation does not seem to be related to changes in *xyr1* expression. It is known, that the expression of xylanases requires Xyr1 as activator (Stricker et al., 2006), but does not strictly follow the *xyr1* transcription pattern (Derntl et al., 2013). It is assumed, that the regulation of xylanase gene expression is not directly dependent on the amount of Xyr1 and seems to rely on additional mechanisms (Derntl et al., 2013). Although the expression profile of the xylanases is clearly higher in the mutated strain, the expression is still down-regulated with D-glucose and high concentrations of D-xylose. These findings suggest that the proposed deregulation takes effect during inducing (low D-xylose) and non-inducing (no carbon source) conditions. Taking in account that all experiments were performed in a wild type background with functioning CCR, repressing mechanisms still overrule. Therefore the expression of xylanases with glucose and high concentrations of xylose is still elevated in the mutated strain, but down-regulated due to CCR mediated by Cre1.

Which mechanism could be behind the deregulation? There is data suggesting that Xyr1 is involved in the opening of chromatin together with increasing gene expression in the Qm6a wild-type strain (de Sousa et al., 2015). Checking this hypothesis, we investigated the chromatin status for the xylanases genes in both strains via CHART-PCR. But the real-time qPCR data shows no indication for a more open chromatin status in the upper regulatory or core regions of the xylanse genes *xyn1* and *xyn2* that would justify higher expression (Figure 3.6 - 3.10). On the contrary, the chromatin is slightly more condensed in the A824V strain. Because the chromatin status is not the reason for an enhanced xylanase expression, the investigation shifted to another possible explanation. It is known that in the presence of certain metabolic signals, eukaryotes can change the conformation of some regulatory proteins. There are two important Zn_2Cys_6 binuclear cluster proteins for which metabolic signals have direct impact on the status of the transactivation domain. The GAL4 activator from *S. cerevisiae* has a central region, where D-glucose or its metabolites interact and possibly trigger a change of the activators conformation (Stone and Sadowski, 1993). XlnR from *A. niger* is closely related to Xyr1. Hasper and co-workers suggested that the repression of the gene by D-glucose occurs trough an intra- or intermolecular interaction with the C-terminal of the protein. This would lead to a conformational change and therefore an inactivation of the protein (Hasper et al., 2004). Because Xyr1 is closely related to XlnR, the mechanism could be similar. The mutation A824V in Xyr1 is located in the C-terminal of the protein and it correlates with the V756 residue of XlnR. The mutation V756P in XlnR leads to a xylanase expression under repressing conditions. This means: the mutation in this region disturbs a D-glucose inhibitory domain (Hasper et al., 2004). Recently, it was demonstrated that the complete $Xyr1_{A824V}$ protein shows no response to carbohydrates any more ((de Sousa et al., 2015), in preparation). From these results it was suggested that Xyr1 receives a carbohydrate signal causing a structural change and that the mutation disturbs this signalling of carbohydrates. The findings in this matter allow neither a confirmation nor a disapproval of this hypothesis. It is true, that for all Xyr1 protein fragments there were no changes in secondary structure for the wild type fragments or the mutated ones in the presence of inducing or repressing carbohydrates. But it is simply possible that for a sugar response the whole protein must be available. As seen in the hypothetical figure 1.3, a carbohydrate response can only occur when both, the N- and the C-terminal of the protein are present. This is not the case for any of the protein fragments. Another possible reason could be that the tertiary structure has influence on the interaction between protein and carbohydrates. To examine this approach, we tried to confirm that there are changes in tertiary structure between the full length wild type and the mutated Xyr1 protein. Producing the large, for the measurements necessary amounts of wild type Xyr1 was easily achieved, but the mutated Xyr1 could not be concentrated to a sufficiently large amount and therefore no comparison was allowed. Derntl

and co-workers reported that the mutation is located in the C-terminal of the FTFMHR in an α-helices-rich region. Mello-de-Sousa and co-workers investigated the possibility of changes in secondary structure due to a mutation in this region. Their yet unpublished data showed, that indeed the mutation leads to different CD spectra in the region where peptide bonds absorb and where information about the secondary structure can be concluded from. They further showed, that the mutation changes the secondary structure in $Xyr1_{A824V}$ and suggested that the region in which the mutation is located is important for the folding and could be responsible for the missing carbohydrate response. They suggested that due to the change of secondary structure the sugars can still bind to a hypothetical carbohydrate binding domain but they do not lead to a structural change of the protein. Therefore the transactivation domain would always be presented and the transcription of xylanase genes would be increased (1.3). To investigate this hypothesis we heterologously expressed nine different Xyr1 and $Xyr1_{A824V}$ protein fragments in *E.coli*, aiming to map where the point mutation may interfere, and investigated their secondary structure with CD spectroscopy. Three different fragments have a wild type and a mutated version, but only one pair showed different CD spectra and therefore differences in secondary structure. These findings suggest, that not only the mutation in the α-helices-rich region is responsible for the changes in secondary structure but maybe also another domain or region which interacts with the region around the mutation. The region of interest could be the putative activation domain which is present in the Xyr1-4 fragment, where the difference in secondary structure has been observed. This is supported by the CD spectra from the Xyr1-5 fragment, where no changes have been observed. These fragments miss only the putative activation domain. Surprisingly the Xyr1-2 fragments, which has the putative activation domain but no DNA binding domain, also show no changes in secondary structure. To investigate further on these findings I tried to get information about possible changes in tertiary structure for the fragments Xyr1-2, Xyr1-4 and Xyr1-5. For the CD measurments in the near UV (320-260 nm spectra) the protein had to be very high concentrated (around 20μM). These concentrations could only be gained for the Xyr1-2 protein fragments. But for these a hint of change in tertiary structure could be proofed. The mutated version is suggested to be folded in a different way than the wild type version (Figure 3.34). The results lead to the assumption that the mutation might change the secondary structure of the transactivator Xyr1. But the mutation alone is not sufficient for a change, the putative activation domain may play a role in this.

Chapter 5

Conclusion

In order to investigate the mutated Xyr1$_{A824V}$ strain, I found that the expression of the main cellulase-encoding genes *cbh1* and *cbh2* follows the transcript formation of *xyr1*. On the other hand, the expression of the xylanases is deregulated and enhanced. The increase in transcript formation of these genes can not be explained by a more open chromatin status. Measurements to investigate the accessibility of the promoter region in these genes, showed that the chromatin status is equal or even slightly more condensed in the mutated strain. If the chromatin status is equal, maybe the mutation itself lead to a change in secondary structure and interferes with the transcript regulation by certain metabolic signals. The nine, for this purpose designed and produced Xyr1 protein fragments showed no response to carbohydrates and only one showed a change in secondary structure. But it is possible that ether the whole protein is needed for a conformational change or that the tertiary structure has influence on the interaction between protein and carbohydrates. Experiments to investigate these approaches failed, because the necessary amount of protein could not be produced. Only one fragment pair of Xyr1, wild type and mutated one, showed different CD spectra and therefore differences in secondary structure. This finding suggest, that not only the mutation in the α-helices-rich region is responsible for the changes in secondary structure but maybe also another domain or region which interacts with the region around the mutation. The region of interest could be the putative activation domain which is present in the Xyr1-4 fragments where the differences in secondary structure have been observed. In the other fragment pair (Xyr1-2), containing the putative activation domain, no change in secondary structure could be observed. But for these proteins, measurements showed different tertiary structures. The mutated protein fragment is probably folded in a different way then the wild type one.

Appendix A

vector maps

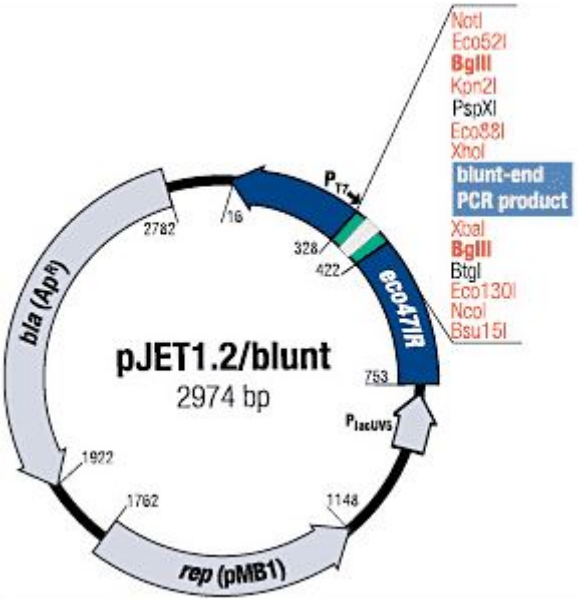

FIGURE A.1: Vector map of pJET2.1 blunt cloning vector

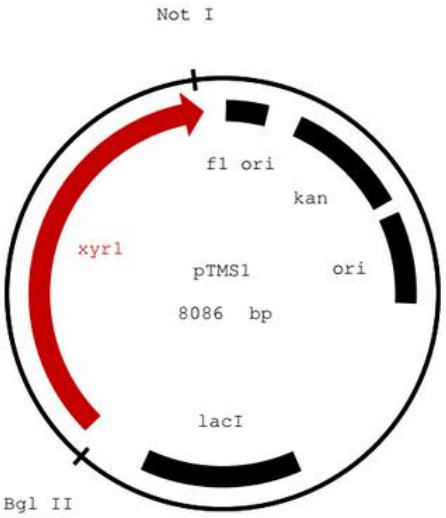

FIGURE A.2: Vector map of pTS1

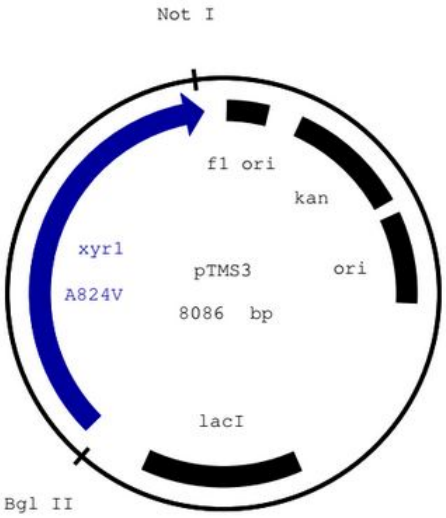

FIGURE A.3: Vector maps of pTS2

Appendix B

sequences of fragments

Xyr1-1:

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCAT**ATG**CTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAG

Xyr1-2:

CAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCATCTGGGTAATATTCTGCCGGTGAGCCTGG
CCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACAG**ATG**CATCCGATGAGCCCGTATGTTCTG
GGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCGTCGTTGTCAGCCTGCCCTGCTGGCAAGT
ATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACCAGCCTGCCGAGCGCACGTAGCAAAGTTT
GTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGATTCATACCGGCACCAATAGCCCGAGCCC
GAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGGTGTTGCAATGCCTGGTAGTCTGAATATG
GATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGCCTGGATGATGTTATTACCTATGTTCATCT
GGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCGTTGGTGGGGTGCAGCATGGTCACTGGC
ACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCTGCAAATCAAGAAGATGGCGAAGGTCT
GAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTTTTGTTACCGAAGAAGAACGCGAAGAA
CGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGCCCTGTGTTATAATCGTCCTCTGTTTCTG
CTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAAATGGCAGGCAGGCAAATTTCGTAGCC
ATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGATGAATTTGGCGATAGTCCGCGTGCAGC
CCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTTTCTGAGCCTGATGACGATTCTGGGTG
AAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTGGTTTTCGTAGCGCACGTGATTGGGAT
GAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAGCCTGAAACGTTTTGTGGCAAAACATC
TGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGCGGTGCAGTTACCGATATGCAGAGTC
CGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAAATTCAGGCAAGCATTGTTGTTGCATA
TAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATGGGATCCGATTAACCTGCTGGATGACG
ATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCATGCAGTTAGCGCAGCAGAAGCAATTAG
CCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTATGGTGTTTACCTGCTGCAGGGTAGTTT
TCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCGAGCGTTATTAAAGCATGTGAAACCATT
GTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCAGCGTAATTTTAGCAAAGTGATGCGTAG
TGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAACAGCAACAGCGTCGCCGTGAACTGCTG

GCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTACCCTGGTGCCGCGCGGCAGCGCGGCC
GCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataa
ctagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggat

Xyr1-3

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCAT**ATG**CTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCAT
CTGGGTAATATTCTGCCGGTGAGCCTGGCCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACA
GATGCATCCGATGAGCCCGTATGTTCTGGGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCG
TCGTTGTCAGCCTGCCCTGCTGGCAAGTATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACC
AGCCTGCCGAGCGCACGTAGCAAAGTTTGTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGA
TTCATACCGGCACCAATAGCCCGAGCCCGAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGG
TGTTGCAATGCCTGGTAGTCTGAATATGGATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGC
CTGGATGATGTTATTACCTATGTTCATCTGGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCG
TTGGTGGGGTGCAGCATGGTCACTGGCACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCT
GCAAATCAAGAAGATGGCGAAGGTCTGAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTT
TTGTTACCGAAGAAGAACGCGAAGAACGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGC
CCTGTGTTATAATCGTCCTCTGTTTCTGCTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAA
ATGGCAGGCAGGCAAATTTCGTAGCCATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGAT
GAATTTGGCGATAGTCCGCGTGCAGCCCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTT
T

Xyr1-4:

CTGAGCCTG**ATG**ACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG
GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG
CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA
ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG

GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGCAGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTA
TGGTGTTTACCTGCTGCAGGGTAGTTTTCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCG
AGCGTTATTAAAGCATGTGAAACCATTGTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCA
GCGTAATTTTAGCAAAGTGATGCGTAGTGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAA
CAGCAACAGCGTCGCCGTGAACTGCTGGCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTA
CCCTGGTGCCGCGCGGCAGCGCGGCCGCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaagg
aagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggagga
actatatccggat

Xyr1-5:

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCAT**ATG**CTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCAT
CTGGGTAATATTCTGCCGGTGAGCCTGGCCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACA
GATGCATCCGATGAGCCCGTATGTTCTGGGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCG
TCGTTGTCAGCCTGCCCTGCTGGCAAGTATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACC
AGCCTGCCGAGCGCACGTAGCAAAGTTTGTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGA
TTCATACCGGCACCAATAGCCCGAGCCCGAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGG
TGTTGCAATGCCTGGTAGTCTGAATATGGATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGC
CTGGATGATGTTATTACCTATGTTCATCTGGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCG
TTGGTGGGGTGCAGCATGGTCACTGGCACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCT
GCAAATCAAGAAGATGGCGAAGGTCTGAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTT
TTGTTACCGAAGAAGAACGCGAAGAACGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGC
CCTGTGTTATAATCGTCCTCTGTTTCTGCTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAA
ATGGCAGGCAGGCAAATTTCGTAGCCATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGAT
GAATTTGGCGATAGTCCGCGTGCAGCCCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTT
TCTGAGCCTGATGACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG
GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG
CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA

ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG
GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGCAGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCG

Xyr1-6:

TTTTTCT**ATG**GTGTTTACCTGCTGCAGGGTAGTTTTCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGC
GAGCCCGAGCGTTATTAAAGCATGTGAAACCATTGTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACC
GAATATCAGCGTAATTTTAGCAAAGTGATGCGTAGTGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATT
TAGCAGAACAGCAACAGCGTCGCCGTGAACTGCTGGCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGC
ATTAGGTACCCTGGTGCCGCGCGGCAGCGCGGCCGCactcgagcaccaccaccaccaccactgagatccggctgctaacaaag
cccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctga
aaggaggaactatatccggat

Appendix C

sequencing data

Xyr1-1

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCAT**ATG**CTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAG

Xyr1-2.1

CAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCATCTGGGTAATATTCTGCCGGTGAGCCTGG
CCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACAGATGCATCCGATGAGCCCGTATGTTCTG
GGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCGTCGTTGTCAGCCTGCCCTGCTGGCAAGT
ATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACCAGCCTGCCGAGCGCACGTAGCAAAGTTT
GTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGATTCATACCGGCACCAATAGCCCGAGCCC
GAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGGTGTTGCAATGCCTGGTAGTCTGAATATG
GATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGCCTGGATGATGTTATTACCTATGTTCATCT
GGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCGTTGGTGGGGTGCAGCATGGTCACTGGC
ACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCTGCAAATCAAGAAGATGGCGAAGGTCT
GAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTTTTGTTACCGAAGAAGAACGCGAAGAA
CGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGCCCTGTGTTATAATCGTCCTCTGTTTCTG
CTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAAATGGCAGGCAGGCAAATTTCGTAGCC
ATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGATGAATTTGGCGATAGTCCGCGTGCAGC
CCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTTTCTGAGCCTGATGACGATTCTGGGTG
AAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTGGTTTTCGTAGCGCACGTGATTGGGAT
GAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAGCCTGAAACGTTTTGTGGCAAAACATC
TGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGCGGTGCAGTTACCGATATGCAGAGTC
CGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAAATTCAGGCAAGCATTGTTGTTGCATA
TAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATGGGATCCGATTAACCTGCTGGATGACG
ATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCATGCAGTTAGCGCAGCAGAAGCAATTAG
CCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTATGGTGTTTACCTGCTGCAGGGTAGTTT
TCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCGAGCGTTATTAAAGCATGTGAAACCATT
GTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCAGCGTAATTTTAGCAAAGTGATGCGTAG
TGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAACAGCAACAGCGTCGCCGTGAACTGCTG

GCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTACCCTGGTGCCGCGCGGCAGCGCGGCC
GCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataa
ctagcataaccccttggggcctctaaacgggtcttgaggggtttttgctgaaaggaggaactatatccggat

Xyr1-2.2

CAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCATCTGGGTAATATTCTGCCGGTGAGCCTGG
CCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACAGATGCATCCGATGAGCCCGTATGTTCTG
GGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCGTCGTTGTCAGCCTGCCCTGCTGGCAAGT
ATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACCAGCCTGCCGAGCGCACGTAGCAAAGTTT
GTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGATTCATACCGGCACCAATAGCCCGAGCCC
GAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGGTGTTGCAATGCCTGGTAGTCTGAATATG
GATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGCCTGGATGATGTTATTACCTATGTTCATCT
GGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCGTTGGTGGGGTGCAGCATGGTCACTGGC
ACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCTGCAAATCAAGAAGATGGCGAAGGTCT
GAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTTTTGTTACCGAAGAAGAACGCGAAGAA
CGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGCCCTGTGTTATAATCGTCCTCTGTTTCTG
CTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAAATGGCAGGCAGGCAAATTTCGTAGCC
ATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGATGAATTTGGCGATAGTCCGCGTGCAGC
CCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTTTCTGAGCCTGATGACGATTCTGGGTG
AAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTGGTTTTCGTAGCGCACGTGATTGGGAT
GAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAGCCTGAAACGTTTTGTGGCAAAACATC
TGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGCGGTGCAGTTACCGATATGCAGAGTC
CGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAAATTCAGGCAAGCATTGTTGTTGCATA
TAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATGGGATCCGATTAACCTGCTGGATGACG
ATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCATGCAGTTAGCGCAGTTGAAGCAATTAG
CCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTATGGTGTTTACCTGCTGCAGGGTAGTTT
TCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCGAGCGTTATTAAAGCATGTGAAACCATT
GTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCAGCGTAATTTTAGCAAAGTGATGCGTAG
TGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAACAGCAACAGCGTCGCCGTGAACTGCTG
GCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTACCCTGGTGCCGCGCGGCAGCGCGGCC
GCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataa
ctagcataaccccttggggcctctaaacgggtcttgaggggtttttgctgaaaggaggaactatatccggat

Xyr1-3

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCATATGCTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG

GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCAT
CTGGGTAATATTCTGCCGGTGAGCCTGGCCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACA
GATGCATCCGATGAGCCCGTATGTTCTGGGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCG
TCGTTGTCAGCCTGCCCTGCTGGCAAGTATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACC
AGCCTGCCGAGCGCACGTAGCAAAGTTTGTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGA
TTCATACCGGCACCAATAGCCCGAGCCCGAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGG
TGTTGCAATGCCTGGTAGTCTGAATATGGATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGC
CTGGATGATGTTATTACCTATGTTCATCTGGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCG
TTGGTGGGGTGCAGCATGGTCACTGGCACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCT
GCAAATCAAGAAGATGGCGAAGGTCTGAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTT
TTGTTACCGAAGAAGAACGCGAAGAACGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGC
CCTGTGTTATAATCGTCCTCTGTTTCTGCTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAA
ATGGCAGGCAGGCAAATTTCGTAGCCATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGAT
GAATTTGGCGATAGTCCGCGTGCAGCCCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTT
T

Xyr1-4.1

CTGAGCCTGATGACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG
GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG
CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA
ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG
GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGCAGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTA
TGGTGTTTACCTGCTGCAGGGTAGTTTTCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCG
AGCGTTATTAAAGCATGTGAAACCATTGTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCA
GCGTAATTTTAGCAAAGTGATGCGTAGTGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAA
CAGCAACAGCGTCGCCGTGAACTGCTGGCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTA
CCCTGGTGCCGCGCGGCAGCGCGGCCGCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaagg
aagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggagga
actatatccggat

Xyr1-4.2

CTGAGCCTGATGACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG
GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG

CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA
ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG
GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGTTGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCGTTTTTCTA
TGGTGTTTACCTGCTGCAGGGTAGTTTTCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGCGAGCCCG
AGCGTTATTAAAGCATGTGAAACCATTGTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACCGAATATCA
GCGTAATTTTAGCAAAGTGATGCGTAGTGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATTTAGCAGAA
CAGCAACAGCGTCGCCGTGAACTGCTGGCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGCATTAGGTA
CCCTGGTGCCGCGCGGCAGCGCGGCCGCactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaagg
aagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggagga
actatatccggat

Xyr1-5.1

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCATATGCTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCAT
CTGGGTAATATTCTGCCGGTGAGCCTGGCCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACA
GATGCATCCGATGAGCCCGTATGTTCTGGGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCG
TCGTTGTCAGCCTGCCCTGCTGGCAAGTATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACC
AGCCTGCCGAGCGCACGTAGCAAAGTTTGTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGA
TTCATACCGGCACCAATAGCCCGAGCCCGAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGG
TGTTGCAATGCCTGGTAGTCTGAATATGGATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGC
CTGGATGATGTTATTACCTATGTTCATCTGGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCG
TTGGTGGGGTGCAGCATGGTCACTGGCACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCT
GCAAATCAAGAAGATGGCGAAGGTCTGAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTT
TTGTTACCGAAGAAGAACGCGAAGAACGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGC
CCTGTGTTATAATCGTCCTCTGTTTCTGCTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAA
ATGGCAGGCAGGCAAATTTCGTAGCCATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGAT
GAATTTGGCGATAGTCCGCGTGCAGCCCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTT
TCTGAGCCTGATGACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG

GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG
CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA
ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG
GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGCAGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCG

Xyr1-5.2

AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAA
ATAATTTTGTTTAACTTTAAGAAGGAGATATCATATGCTGAGCAATCCGCTGCGTCGTTATAGCGCATATCCTG
ATATTAGCAGCGCCAGCTTTGATCCGAATTATCATGGTAGCCAGAGCCATCTGCATAGCATTAATGTTAATACC
TTTGGCAACAGCCATCCGTATCCGATGCAGCATCTGGCACAGCATGCAGAACTGAGCAGCAGCCGTATGATTC
GTGCAAGTCCGGTTCAGCCGAAACAGCGTCAGGGTAGCCTGATTGCAGCACGTAAAAATAGCACCGGCACCG
CAGGTCCGATTCGTCGTCGTATTAGCCGTGCATGTGATCAGTGTAATCAGCTGCGTACCAAATGTGATGGTCTG
CATCCGTGTGCACATTGTATTGAATTTGGTCTGGGTTGCGAATATGTGCGTGAACGTAAAAAACGTGGTAAAG
CAAGCCGTAAAGATATTGCCGCACAGCAGGCAGCCGCAGCGGCAGCACAGCATAGCGGTCAGGTTCAGGATG
GTCCGGAAGATCAGCATCGTAAACTGAGCCGTCAGCAGAGCGAAAGCAGCCGTGGTAGCGCAGAACTGGCCC
AGCCTGCACATGATCCGCCTCATGGTCATATTGAAGGTAGCGTTAGCAGCTTTAGCGATAATGGTCTGAGCCA
GCATGCCGCAATGGGTGGTATGGATGGTTTAGAAGATCATCATGGTCACGTTGGTGTTGATCCGGCACTGGGT
CGTACCCAGCTGGAAGCAAGCAGTGCAATGGGTCTGGGTGCATATGGTGAAGTTCATCCGGGTTATGAAAGT
CCGGGTATGAATGGTCATGTTATGGTTCCGCCTAGCTATGGTGCACAGACCACCATGGCAGGTTATAGCGGTA
TTAGCTATGCAGCACAGGCACCGAGTCCGGCAACCTATAGCAGTGATGGTAATTTTCGTCTGACGGGTCATATT
CATGATTATCCGCTGGCAAATGGTAGCAGCCCGAGCTGGGGTGTTAGCCTGGCAAGCCCGAGCAATCAGTTTC
AACTGCAGCTGAGTCAGCCGATCTTTAAACAGAGCGATCTGCGTTATCCGGTTCTGGAACCGCTGCTGCCGCAT
CTGGGTAATATTCTGCCGGTGAGCCTGGCCTGTGATCTGATTGATCTGTATTTTAGCAGCAGCAGCTCAGCACA
GATGCATCCGATGAGCCCGTATGTTCTGGGTTTTGTTTTTCGTAAACGCAGCTTTCTGCATCCTACCAATCCGCG
TCGTTGTCAGCCTGCCCTGCTGGCAAGTATGCTGTGGGTTGCAGCCCAGACCAGCGAAGCGAGCTTTCTGACC
AGCCTGCCGAGCGCACGTAGCAAAGTTTGTCAGAAACTGCTGGAACTGACCGTTGGTCTGCTGCAGCCGCTGA
TTCATACCGGCACCAATAGCCCGAGCCCGAAAACCAGTCCGGTTGTTGGTGCAGCAGCACTGGGTGTGCTGGG
TGTTGCAATGCCTGGTAGTCTGAATATGGATAGCCTGGCTGGTGAAACCGGTGCATTTGGTGCAATTGGTAGC
CTGGATGATGTTATTACCTATGTTCATCTGGCAACCGTTGTTAGCGCAAGCGAATATAAAGGTGCAAGCCTGCG
TTGGTGGGGTGCAGCATGGTCACTGGCACGTGAACTGAAACTGGGTCGCGAACTGCCTCCGGGTAATCCGCCT
GCAAATCAAGAAGATGGCGAAGGTCTGAGCGAAGATGTTGATGAACATGATCTGAATCGCAATAATACCCGTT
TTGTTACCGAAGAAGAACGCGAAGAACGTCGTCGTGCATGGTGGCTGGTTTATATTGTTGATCGTCATCTGGC
CCTGTGTTATAATCGTCCTCTGTTTCTGCTGGATAGCGAATGTAGCGATCTGTATCATCCGATGGATGACATTAA
ATGGCAGGCAGGCAAATTTCGTAGCCATGATGCAGGTAATAGCAGCATTAACATTGATAGCTCCATGACCGAT
GAATTTGGCGATAGTCCGCGTGCAGCCCGTGGTGCCCATTATGAATGTCGTGGTCGTAGCATTTTCGGTTATTT
TCTGAGCCTGATGACGATTCTGGGTGAAATTGTTGATGTGCATCATGCAAAAAGCCATCCGCGTTTTGGTGTTG
GTTTTCGTAGCGCACGTGATTGGGATGAACAGGTTGCCGAAATTACACGCCATCTGGATATGTATGAAGAAAG
CCTGAAACGTTTTGTGGCAAAACATCTGCCGCTGAGCAGTAAAGATAAAGAACAGCATGAAATGCATGATAGC
GGTGCAGTTACCGATATGCAGAGTCCGCTGAGCGTTCGTACCAATGCAAGCAGTCGTATGACCGAAAGCGAA
ATTCAGGCAAGCATTGTTGTTGCATATAGCACCCATGTTATGCATGTTCTGCACATTCTGCTGGCAGATAAATG
GGATCCGATTAACCTGCTGGATGACGATGATCTGTGGATTAGCAGTGAAGGTTTTGTGACCGCAACCAGCCAT
GCAGTTAGCGCAGCAGAAGCAATTAGCCAGATTCTGGAATTTGATCCTGGCCTGGAATTTATGCCG

Xyr-6

TTTTTCTATGGTGTTTACCTGCTGCAGGGTAGTTTTCTGCTGCTGTTAATTGCAGATAAACTGCAGGCAGAAGC
GAGCCCGAGCGTTATTAAAGCATGTGAAACCATTGTTCGTGCACATGAAGCCTGTGTTGTTACCCTGAGCACC
GAATATCAGCGTAATTTTAGCAAAGTGATGCGTAGTGCACTGGCACTGATTCGTGGTCGCGTGCCGGAAGATT
TAGCAGAACAGCAACAGCGTCGCCGTGAACTGCTGGCTCTGTATCGTTGGACCGGTAATGGCACCGGTCTGGC
ATTAGGTACCCTGGTGCCGCGCGGCAGCGCGGCCGCactcgagcaccaccaccaccaccactgagatccggctgctaacaaag
cccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctga
aaggaggaactatatccggat

FIGURE C.1: sequencing data from Microsynth AG. red marked: A824, blue marked: A824V mutation

Appendix D

list of figures

List of Figures

Appendix E

list of tables

List of Tables

Bibliography

de Lourdes T.M. Polizeli, M. and Rai, M. (2014). *Fungal Enzymes*. CRC Press – Taylor and Francis Group.

de Sousa, M. M. (2015). The transactivator xyr1 of trichoderma reesei ffunction in a nuclear receptor-like way (work still in progress).

de Sousa, T. M. M., Rassinger, A., Pucher, M. E., dos Santos Castro, L., Persinoti, G. F., Silva-Rocha, R., Pocas-Fonseca, M. J., Mach, R. L., Silva, R. N., and Mach-Aigner, A. R. (2015). The impact of chromatin remodelling on cellulase expression in trichoderma reesei. *BMC Genomics 16:588*.

Derntl, C., Gudynaite-Savitch, L., Calixte, S., White, T., Mach, R. L., and Mach-Aigner, A. R. (2013). Mutation of xylanase regulator 1 cause a glucose blind hydrolase expression phenotype in industrially used trichoderma strains. *Biotechnology for Biofuels*, 6:62.

Druzhinina, I. S., Komon-Zelazowska, M., Atanasova, L., Seidl, V., and Kubicek, C. P. (2010). Evolution and ecophysiology of the industrial producer hypocrea jecorina (anamorph trichoderma reesei) and a new sympatric agamospecies related to it. *PLoS ONE*, 5 (2):e9191.

Galante, Conti, and Monteverdi (1998a). *Application of Trichoderma enzymes in the textily industry*, chapter 14, pages 272–286. Gary GE and Kubicek CP.

Galante, Y., Conti, A., and Monteverdi, R. (1998b). *Application of Trichoderma enzymes in the food and feed industry*, chapter 15, pages 286–302. Gary GE and Kubicek CP.

Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols 1(6)*, 1(6):2876?2890.

Hasper, A. A., Trindade, L. M., van der Veen, D., van Ooyen, A. J. J., and de Graaff, L. H. (2004). Functional analysis of the transcriptional activator xlnr from aspergillus niger. *Microbiology*, 150:1367 – 1375.

Kelly, S. M., Jess, T. J., and Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et Biophysica Acta*, 1752:119–139.

Klaubauf, HM, N., H, P., M, Z., K, B., AR, M.-A., RL, M., AJ, H., AF, A., and 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.*, 14:1087–1845.

Koivula, A., Linder, M., and Teeri, T. (1998). *Structure-function relationships in Trichoderma cellulolytic enzymes*, chapter 1, pages 3–25. Gary GE and Kubicek CP.

Kuhls, K., Lieckfeldt, E., Samuels, G. J., Kovacs, W., Meyer, W., Petrini, O., Gams, W., B?rner, T., and Kubicek, C. P. (1996). Molecular evidence that the asexual industrial fungus trichoderma reesei is a clonal derivative of the ascomycete hypocrea jecorina. *Proceedings of the National Academy of Science USA*, 93:7755–7760.

Kumar, R., Singh, S., and Singh, O. V. (2008). Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of Industrial Microbiology and Biotechnology*, 35:377 – 391.

Mach-Aigner, A. R., Pucher, M. E., , and Mach, R. L. (2010). D-xylose as a repressor or inducer of xylanase expression in hypocrea jecorina (trichoderma reesei). *Applied and Environmental Microbiology*, 76(6):1770 – 1776.

Mach-Aigner, A. R., Pucher, M. E., Steiger, M. G., Bauer, G. E., Preis, S. J., and Mach, R. L. (2008). Transcriptional regulation of xyr1, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in hypocrea jecorina. *Applied and Environmental Microbiology*, 74 (21):6554 – 6562.

MacPherson, S., Larochelle, M., and Turcotte, B. (2006). A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiology and Molecular Biology Reviews*, 70(3):583.

Margolles-Clark, E., Ilm?n, M., and Penttil?, M. (1997). Expression patterns of ten hemicellulase genes of the filamentous fungus trichoderma reesei on various carbon sources. *Journal of Biotechnology*, 57:167–179.

Miller, J., Miller, J., McLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor iiia from xenopus oocytes. *EMBO J*, 4:1609?1614.

Montenecourt, B. and Eveleigh, D. (1977). Preparation of mutants of trichoderma reesei with enhanced cellulase production. *Applied and Environmental Microbiology*, 34 (6):777–782.

Montenecourt, B. S. and Eveleigh, D. E. (1979). *Selective Screening Methods for the Isolation of High Yielding Cellulase Mutants of Trichoderma reesei*, chapter 14, pages 289–301. American Chemical Society.

Peterson, R. and Nevalainen, H. (2012). Trichoderma reesei rut–c30 – thirty years of strain improvement. *Microbiology*, 158:58–68.

Polizeli, M. L. T. M., Rizzatti, A. C. S., Monti, R., Terenzi, H. F., Jorge, J. A., and Amorim, D. S. (2005). Xylanases from fungi: properties and industrial applications. *Applied Microbiology and Biotechnology*, 67:577 – 591.

Schjerling, P. and Holmberg, S. (1996). Comparative amino acid sequence analysis of c6 zinc cluster family of transcriptional regulators. *Nucleic Acids Research*, 24:4599 – 4607.

Schuster, A. and Schmoll, M. (2010). Biology and iotechnology of trichoderma. *Applied Microbiology and Biotechnology*, 87:787 – 799.

Seiboth, B., Ivanova, C., and Seidl-Seiboth, V. (2011). *Trichoderma reesei: a fungal enzym producer for cellulosic biofuels Biofuel production – recent Developments and Prospects*. InTech.

Seidl, V., Seibel, C., Kubicek, C. P., and Schmoll, M. (2009). Sexual development in the industrial workhorse trichoderma reesei. *PNAS*, 106(33):13909 – 13914.

Stone, G. and Sadowski, I. (1993). Gal4 is regulated by a glucose–responsive functional domain. *The EMBO Journal*, 12(4):1375 – 1385.

Strauss, J., Mach, R. L., Zeilinger, S., Hartler, G., St?ffler, G., Wolschek, M., and Kubicek, C. (1995). Crel, the carbon catabolite repressor protein from trichoderma reesei. *Federation of European Biochemical Societies Letters*, 376:103–107.

Stricker, A. R., Grosstessner-Hain, K., Wuerleitner, E., and Mach, R. L. (2006). Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and d-xylose metabolism in hypocrea jecorina. *Eukaryotic Cell*, 5(12):2125–2137.

Toerroenen, A., Mach, R. L., Messner, R., Gonzales, R., Kalkkinen, N., Harkki, A., and Kubicek, C. P. (1992). The two major xylanases from trichoderma reesei: Characterization of both enzymes and genes. *Biotechnology*, 10:1461–1465.

Wong, M. L. and Medrano, J. F. (2005). Real time pcr for mrna quantitation. *BioTechniques*, 39:75–85.

Zeilinger, S., Mach, R. L., Schindler, M., Herzog, P., , and Kubicek, C. P. (1996). Different inducibility of expression of the two xylanase genes xyn1 and xyn2 in trichoderma reesei. *The journal of biological chemistry*, 271 (41):25624–25629.