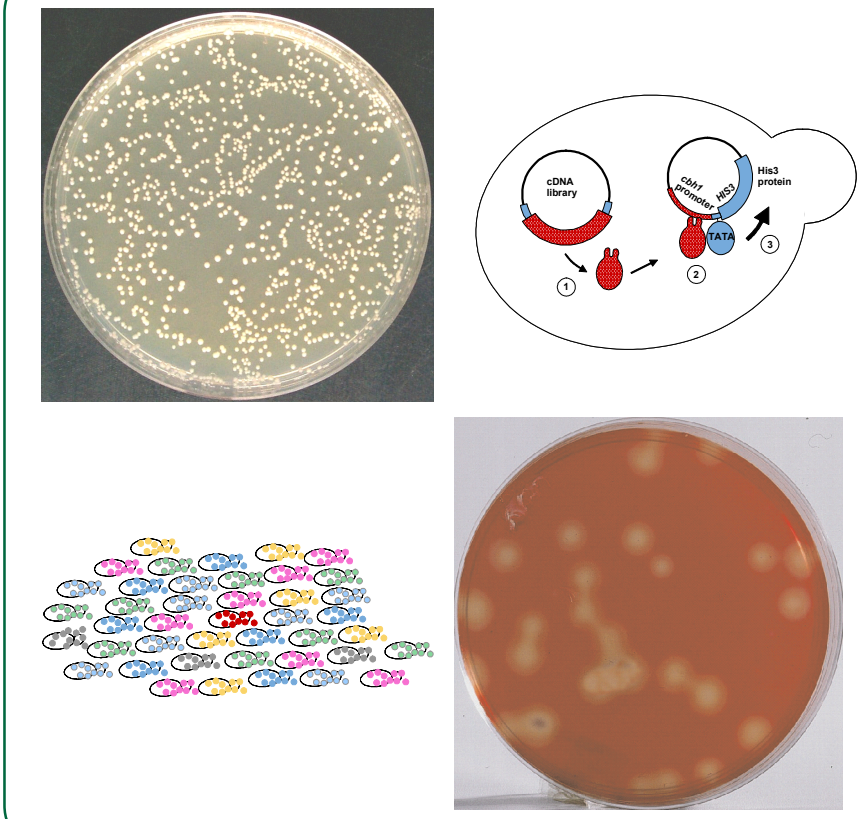
ESPOO 2004 VTT PUBLICATIONS 541



Anu Saloheimo

Yeast *Saccharomyces cerevisiae* as a tool in cloning and analysis of fungal genes

Applications for biomass hydrolysis and utilisation

VTT PUBLICATIONS 541

Yeast *Saccharomyces cerevisiae* as a tool in cloning and analysis of fungal genes

Applications for biomass hydrolysis and utilisation

Anu Saloheimo

VTT Biotechnology

Faculty of Biosciences
Department of Biological and Environmental Sciences
Division of Genetics
University of Helsinki, Helsinki, Finland

To be presented with the permission of the Faculty of Biosciences of the University of Helsinki, for public criticism in the lecture hall of the Natural History Museum, Pohjoinen Rautatiekatu 13, Helsinki, on August 6th, 2004, at 12 o'clock noon.

ISBN 951–38–6400–6 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–6401–4 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 456 4374

VTT, Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel (09) 4561, fax (09) 456 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 5, P.O.Box 2000, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 456 4374

VTT Biotekniikka, Tietotie 2, PL 1500, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 455 2103

VTT Bioteknik, Datavägen 2, PB 1501, 02044 VTT
tel. växel (09) 4561, fax (09) 455 2103

VTT Biotechnology, Tietotie 2, P.O.Box 1501, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 455 2103

Technical editing Maini Manninen

Otamedia Oy, Espoo 2004

Saloheimo, Anu. Yeast *Saccharomyces cerevisiae* as a tool in cloning and analysis of fungal genes. Applications for biomass hydrolysis and utilisation. Espoo 2004. VTT Publications 541. 84 p. + app. 52.

Keywords *Saccharomyces cerevisiae*, *Trichoderma reesei*, *Hypocrea jecorina*, expression-cloning, reporter gene, cellulase, activator, ACEI, ACEII, EGV, permease, Hxt, xylose uptake, xylose utilisation, bio-ethanol

Abstract

The baker´s yeast, *Saccharomyces cerevisiae* has been employed by man for centuries in manufacturing of bread, beer, and wine. In science, it has become a useful tool as well. In this work, methods were developed in order to study the molecular biology of the cellulolytic filamentous fungus *Trichoderma reesei* with the aid of yeast. Cellulose is the most abundant carbon source in nature, and its enzymatic degradation is essential for carbon turnover. In addition, cellulose is used as a raw material in microbial processes. In this work, a previously unknown cellulase-encoding gene was cloned by expression in yeast and detection of hydrolysis halos on substrate plates. This EGV enzyme consists of an exceptionally small core domain, a cellulose-binding domain, and a linker region connecting the two. EGV belongs to family GH45 of glycosyl hydrolases. Additionally, a gene encoding a β-1,3-1,4-glucanase enzyme was cloned and studied. The enzyme was produced in insect cells, and analysis of the degradation products of β-glucan by NMR showed that it was a laminarinase (EC 3.2.1.6).

A yeast-based cloning method for positively acting regulatory proteins was set up, and two regulatory genes of the *T. reesei* cellulases, *ace1* and *ace2,* were isolated. The isolation was based on the ability of the encoded proteins to activate expression of a reporter gene, which was linked to the full-length promoter of the major cellulase gene *cbh1* in yeast. No homologs of the new regulatory proteins were detected outside the Mycota. The DNA-binding properties of the regulatory proteins were studied both *in vitro* and *in vivo* in yeast. Deletion of the *ace1* gene resulted in slower radial growth of the fungus on cellulose-containing plates. However, although isolated as an activator, ACEI was later shown to act as a repressor of hydrolase expression. ACEII, on the

other hand, was shown to be an activator of cellulase expression. However, it is certainly not the only one, since its deletion did not result in a cellulase-negative phenotype.

Additionally, a sugar permease-encoding gene was isolated from *T. reesei* by complementation. The yeast strain used as a host was deleted for the major hexose transporter genes *(hxt1-7, gal2*), and additionally engineered for xylose utilisation. The *T. reesei* permease complemented the growth defect of the mutant strain on xylose-maltose medium. However, adaptive mutation(s) were needed in the host to enable growth on xylose of the permease-expressing strain. The same, engineered yeast strain was used as a host for the native *S. cerevisiae* hexose transporter genes *HXT1, HXT2, HXT4* and *HXT7*, and the kinetics of xylose transport were studied. The affinities of the permeases for xylose varied, K_m values of 190–900 mM were detected. Interesting differences were obtained in the levels of inhibition by the presence of glucose. The single-Hxt strains exhibited a biphasic growth mode on xylose media, where an initially very slow growth was followed by exponential growth after a lag of several days.

Preface

This study was carried out at VTT Biotechnology in the groups Production Microbes and Metabolic Engineering. The financial support from the Foundation for Biotechnical and Industrial Fermentation Research, the former Alko Ltd., Roal Ltd., the Academy of Finland, and the Technology Development Centre (TEKES) is gratefully acknowledged. This work is a part of the research programme "VTT Industrial Biotechnology" (Academy of Finland; Finnish Centre of Excellence programme, 2000-2005, Project no. 64330). I also want to thank the University of Helsinki for a grant for completion of this thesis. I am grateful to Executive Director, Prof. Juha Ahvenainen and Research Director, Prof. Hans Söderlund for excellent working facilities and for their encouraging attitude towards thesis work. Research Managers Sirkka Keränen and Tiina Nakari-Setälä are thanked for their support and interest in my work. I want to express my thanks to my Group Manager Laura Ruohonen for familiarising me with metabolic engineering and for her support in the permease work. Professor Hannu Saarilahti at the Department of Biological and Environmental Sciences, Division of Genetics, I want to thank for his co-operation during the final stages of my studies. Dr. Pirkko Suominen and Prof. Per Saris are thanked for critical reading of my thesis and for useful comments.

My deepest and warmest thanks go to my supervisor, Prof. Merja Penttilä. She has teached me how to make science and write about it. Furthermore, the novel methods developed in this work are due to her creativity. Her enthusiastic attitude towards science has been an important source of inspiration for me. Thank you, Merja, also for your important comments on my thesis. My most sincere thanks are due to my co-authors Marja Ilmén, Nina Aro, Merja Penttilä, Anna-Marja Hoffrén, Bernard Henrissat, Olle Teleman, Jenita Rauta and Laura Ruohonen for their valuable contributions to the generation of the data and writing of the results. Marja, Nina and Merja are especially thanked for the fruitful and enjoyable collaboration we had during the years of studies on cellulase regulation. Thank you, Nina, for taking care of my scientific children, the *ace*s, while I was taking care of my first real baby, and Marja, for writing most of the *ace1* article, while I was having my twins. My “co-authers” Anita Teleman, Maija Tenkanen, Tiina Nakari-Setälä, and Matti Siika-aho are thanked for their contributions to the unpublished results of this thesis. I apologise that I never wrote the article! I am grateful to John Londesborough for his advice with

the uptake studies, and for his endurance in the attempts of making a geneticist understand kinetics. I want to acknowledge Oili Lappalainen for excellent secretarial work with the figures of the thesis and with the manuscript.

I wish to thank Seija Nordberg and Seija Rissanen for many years of skilful technical assistance, as well as for being so helpful. Riitta Nurmi, Aila Siltala and Eila Leino are thanked for help with the experimental work, and Mirjami Pelkonen, Sanna Haasiosalo and Lea Kekki for keeping the lab running. I want to thank all the people of the "hiiva-home" lab for the special atmosphere there is in the lab. My hearty thanks are due to Marja, Tintti, Kari and Markku for everyday companionship over lunch, and Arja and Tarja for sharing your tips on writing of the thesis and the frustration of doing it while being a mother. And Mari, thank you for letting me use your home as a place for writing, and for getting me back in touch with the music of my early youth, that of Ismo Alanko. Marja-Liisa Sumelahti and Markus Färkkilä, my neurologists, are thanked for making my project possible.

I want to thank my parents Eila-Maija and Ilmari Jokinen for having me, an "evening star," long after three older children. Thank you for your support and help, especially in baby-sitting of our children although you live half a thousand kilometres away. I also want to thank my two sisters, Päivi and Raisa, and my brother Heikki, as well as their spouses for taking me seriously already when I was a teenager, and nowadays, for being good friends.

My special thanks I want to express to many friends of mine: Hilu and Leo, Marja and Samuli, Marjut, Selja, Tuomas, Anna and Turo-Kimmo, Mallu, Minna, Varpu and Olli, Nanni, Terhi. Thank you for your friendship and for sharing my life as godparents of my children, as fellow-singers, as parents of little children, or as party-animals ;-)

Above all, I thank you Markku for your friendship and love during our nearly sixteen years together. Thank you for your support, encouragement and belief in me! And thanks for having such wonderful children, Maija, Essi and Juhani, with me. I am grateful to all four of you, my family, for teaching me so much of life in general and emotions in particular.

Anu

List of publications

This thesis is based on the following original articles, which are referred to in the text by their Roman numerals (I–IV). In addition, some unpublished data is presented.

I Saloheimo, Anu; Henrissat, Bernard; Hoffren, Anna-Marja; Teleman, Olle and Penttilä, Merja. 1994. A novel, small endoglucanase gene, *egl5*, from *Trichoderma reesei* isolated by expression in yeast. *Mol. Microbiol.,* 13**,** 219–228.

II Saloheimo, Anu; Aro, Nina; Ilmén, Marja and Penttilä, Merja. 2000. Isolation of the *ace1* gene encoding a Cys2-His2 transcription factor involved in regulation of activity of the cellulase promoter *cbh1* of *Trichoderma reesei*. *J. Biol. Chem.,* 275, 5817–5825.

III Aro, Nina; Saloheimo, Anu; Ilmén, Marja and Penttilä, Merja. 2001. ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. *J. Biol. Chem.,* 276, 24309–24314.

IV Saloheimo, Anu; Rauta, Jenita; Penttilä, Merja and Ruohonen, Laura. Xylose transport studies on homologous and heterologous permeases expressed in the xylose-utilising *Saccharomyces cerevisiae*. Submitted.

Contents

List of abbreviations

aa	amino acid(s)
ATG	translation initiation codon
bp	base pair(s)
BGL	β-glucosidase enzyme
bgl	β-glucosidase-encoding gene
CAE	*cbh2*-activating element
CBD	cellulose-binding domain
CBH	cellobiohydrolase enzyme
cbh	cellobiohydrolase-encoding gene
cDNA	DNA complementary to mRNA
Cel	cellulase enzyme
cel	cellululase-encoding gene
CreA	glucose repressor protein of *Aspergillus*
creA	gene encoding CreA
CREI	glucose repressor protein of *Trichoderma*
cre1	gene encoding GREI
EG	endoglucanase
egl	endoglucanase gene
EST	expressed sequence tag
GH	glycosyl hydrolase
GST	glutathione-S-transferase
GUS	β-glucuronidase enzyme
HCA	hydrophobic cluster analysis
HEC	hydroxyethyl cellulose
HIS3	histidine biosynthetic gene of yeast used as a reporter
Hxt	hexose transporter protein
HXT	hexose transporter-encoding gene of *S. cerevisiae*
kb	kilobase(s)

kD	kilodalton(s)
K_m	Michaelis constant
lacZ	β-galactosidase gene of *Escherichia coli*
MUC	methylumbelliferyl-β-D-cellobioside
MUL	methylumbelliferyl-β-D-lactoside
mRNA	messenger-RNA
NMR	nuclear magnetic resonance
OD, OD_{600}	optical density at 600nm
ORF	open reading frame
PCR	polymerase chain reaction
RBB	remazole brillian blue
TATA box	binding area of the basic transcription machinery
V_{max}	maximal velocity (of transport)
wt	wild type yeast; not a mutant of gene X
XDH	xylitol dehydrogenase enzyme
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
XI	xylose isomerase enzyme
XK	xylulokinase enzyme
XlnR	xylanase activator of *Aspergilli*
xlnR	xylanase activator-encoding gene of *Aspergilli*
XR	xylose reductase enzyme
XYL1	xylose reductase-encoding gene of *P. stipitis*
XYL2	xylitol dehydrogenase-encoding gene of *P. stipitis*
XYL3	xylulokinase-encoding gene of *S. cerevisiae*

1. Introduction

In recent years, the use of biomass and especially plant waste materials for production of fuels, chemicals and materials has increased significantly. It is an important task to find environmentally safe alternatives for replacement of fossil resources, which eventually will be exhausted. Ethanol is already used as a renewable additive in transportation fuels. In the future, gasoline could altogether be replaced by bio-ethanol, ethanol produced by microbial fermentation of plant materials. It is also possible to produce renewable materials by microbial fermentation. These can act as substitutes for plastics or synthetic fibres e.g. in textile industry. In addition, production of most chemicals by microbial metabolism of plant materials is environmentally friendly compared to chemical synthesis. Thus, the use of microbia in modification of biomass supports sustainable development.

Cellulose is the major constituent of plant materials and its degradation is essential for the applications of biomass. It is polymeric, highly crystalline, and mostly associated with hemicellulose and lignin. Thus, its degradation is a challenge. Many micro-organisms, both bacteria and fungi, are capable of efficient cellulose breakdown by secreting cellulose-degrading enzymes. The filamentous fungus *Trichoderma reesei* is one of the most efficient producers of cellulases. Since the isolation of the QM6a strain during the Second World War, it has been subject to extensive mutagenesis programmes leading to hypercellulolytic production strains. *T. reesei* is a multicellular soft rot fungus belonging to Ascomycete, and recently it was identified as an anamorph of *Hypocrea jecorina* (Kuhls *et al.,* 1996). It is a widely used industrial organism in the production of cellulases and other hydrolases, as well as heterologous proteins (for a review, see Nevalainen and Penttilä, 1995) because of its high secretion capacity. Levels of even 40 grams of secreted protein per litre of culture medium have been published for it (Durand *et al.,* 1988). Thus, *T. reesei* has become the eukaryotic model for a cellulolytic organism.

T. reesei lacks the sexual life cycle. Thus, genetics methods based on crossing of two strains cannot be applied for it. Molecular biology methods are largely available, but they are more complicated and time-consuming as compared to those of various unicellular yeasts, e.g. *Saccharomyces cerevisiae.* A major limitation of *T. reesei* is the lack of autonomically replicating plasmids. Instead,

transforming DNA has to be integrated into the genome, and furthermore, targeting of the DNA to a specific location is strain-dependent. In addition, the transformation frequencies of *T. reesei* are moderate. The genome of *T. reesei* is being sequenced but it is not public yet.

The baker´s yeast *S. cerevisiae* has been utilised by man for hundreds of years. The early uses in making of food and beverages have led to sophisticated uses in science. *S. cerevisiae* is a simple unicellular eukaryote, and a vast variety of methods is available for it. It has both haploid and diploid growth modes, which helped initially in setting up genetics methods. In addition, the molecular biology methods were developed early. Yeast was the first eukaryote transformed with replicative plasmids (Beggs, 1978). Knockout strains were also achieved first from yeast among the eukaryotes (Rothstein, 1983). The first eukaryotic genome sequence was from *S. cerevisiae* (Goffeau *et al.,* 1996). Lately, yeast has been at the front line in the development of the modern -omics methods as well. Different methods have been set up for the functional analysis of yeast genes, e.g. global analysis of expression profiles of messenger RNAs (de Risi *et al.*, 1997), analysis of the effects of deletions of one ORF at a time (Giaever *et al.,* 2002), tagging of all ORFs (Ghaemmaghami *et al.,* 2003), analysis of the localisation of the proteome in the cell (Huh *et al.,* 2003), as well as analysis of protein-protein interactions within the whole proteome (Uetz *et al.,* 2000, von Mering *et al.,* 2002). Metabolomics methods have as well been applied for yeast (Allen *et al.,* 2003, Förster *et al.,* 2003). Thus, yeast has become the model eukaryote in cell and molecular biology, as well as systems biology research. Furthermore, it is a useful tool in studies of other eukaryotes as well. In this work, yeast was used as a tool in studying *T. reesei* genes.

1.1 *Trichoderma reesei* cellulases

T. reesei secretes to its environment a complete set of cellulose-degrading enzymes: two exo-acting enzymes, cellobiohydrolases 1 and 2 (CBHI and CBHII), and several endoglucanases (EG). Additionally, β-glucosidases (BGL) degrade cellobiose and other cello-oligosaccharides into the end-product glucose (see Nevalainen and Penttilä, 1995, for a review). Fungal cellulases are modular enzymes consisting of a catalytic core domain, a cellulose-binding domain (CBD), and a flexible linker connecting the former two. However, some

cellulases lack the CBD, e.g. EGIII of *T. reesei* (Ward *et al.*, 1993). The first three-dimensional structure of a fungal cellulose-binding domain, the CBD of CBHI, was solved by NMR (Kraulis *et al.*, 1989), and that of a core domain, CBHII core, by crystallography (Rouvinen *et al.*, 1990), both from *T. reesei*. In addition to classical cellulases, a novel swollenin protein SWOI, related to plant expansins and containing a CBD, has been characterised from *T. reesei*. It is suggested to loosen the structure of crystalline cellulose, thus making it more accessible for other cellulases (Saloheimo *et al.,* 2002b).

1.1.1 Isolation of the *T. reesei* cellulase genes

The major cellulase genes of *T. reesei* were cloned by differential hybridisation between cellulase-producing and non-producing conditions: *cbh1* (Shoemaker *et al.,* 1983, Teeri *et al.,* 1983), *cbh2* (Teeri *et al.,* 1987), *egl1* (Penttilä *et al.,* 1986, van Arsdell *et al.,* 1987), and *egl2* (originally called *egl3*; Saloheimo *et al.,* 1988). Methods based on protein sequences of the purified enzymes were also used. The *cbh2* gene was isolated by another group using as a probe, oligonucleotides designed based on partial amino acid sequences of the protein (Chen *et al.,* 1987). The *bgl1* (Barnett *et al.,* 1991) and *egl3* (Ward *et al.,* 1993) genes, on the other hand, were probed with PCR fragments amplified with degenerate primers based on the protein sequences. Many of the isolated *T. reesei* cellulase genes, or information on them, were later used in the isolation of the corresponding genes from other *Trichoderma* species or other fungi, e.g. *cbh2* from *T. viride*, (Wang *et al.,* 1996), *egl1* from *T. longibrachiatum* (Gonzáles *et al.,* 1992), *cbh1* and *cbh2* from *Phanerochaete chrysosporium* (Sims *et al.*, 1988, Tempelaars *et al.,* 1994), and *cbh1* from *Penicillium janthinellum* (Koch *et al.*, 1993).

A new isolation method for cellulase genes was developed in article I of this work. It is based on functional expression of a cDNA library in yeast, followed by detection of the cellulolytic activities on substrate plates.

Recently, the genomics methods have widened the concept of gene-cloning. Foreman *et al.* (2003) performed a random cDNA-sequencing project resulting in over 5000 partial gene sequences of *T. reesei*. Sequence comparisons of these to available sequences in the databases revealed new cellulase genes: three

predicted endoglucanase (*cel74a, cel61b* and *cel5b*) and five β-glucosidase genes. The *T. reesei* cellulase enzymes and the genes encoding them are summarised in Table 1.

Table 1. Cellulase enzymes of T. reesei and the genes encoding them. Shown are both the original names and the current nomenclature, which is based on classification of glycosyl hydrolases by hydrophobic cluster analysis (HCA) into GH-families having similar protein folds (Henrissat and Bairoch, 1993, Henrissat et al., 1998). The genes with no enzyme information are from the EST sequence information analysed by Foreman et al. (2003). For consistency, the original names of the enzymes and the genes encoding them are used elsewhere in this thesis.

Original names		HCA-based names		Reference
Enzyme	Gene	Enzyme	Gene	
CBHI	*cbh1*	Cel7A	*cel7a*	Shoemaker *et al.*, 1983, Teeri *et al.*, 1983
CBHII	*cbh2*	Cel6A	*cel6a*	Teeri *et al.*, 1987, Chen *et al.*, 1987
EGI	*egl1*	Cel7B	*cel7b*	Penttilä *et al.*, 1986, van Arsdell *et al.*, 1987
EGII	*egl2*	Cel5A	*cel5a*	Saloheimo M. *et al.*, 1988 (originally called *egl3)*
EGIII	*egl3*	Cel12A	*cel12a*	Ward *et al.*, 1993
EGIV	*egl4*	Cel61A	*cel61a*	Saloheimo M. *et al.*, 1997
EGV	*egl5*	Cel45A	*cel45a*	This work, article I
-	-	-	*cel74a*	Foreman *et al.*, 2003
-	-	-	*cel61b*	Foreman *et al.*, 2003
-	-	-	*cel5b*	Foreman *et al.*, 2003
BGLI	*bgl1*	Cel3A	*cel3a*	Barnett *et al.*, 1991
BGLII	*bgl2*	Cel1A	*cel1a*	Saloheimo M. *et al.*, 2002a
-	-	-	*cel1b*	Foreman *et al.*, 2003
-	-	-	*cel3b*	Foreman *et al.*, 2003
-	-	-	*cel3c*	Foreman *et al.*, 2003
-	-	-	*cel3d*	Foreman *et al.*, 2003
-	-	-	*cel3e*	Foreman *et al.*, 2003
SWOI	*swo1*	-	-	Saloheimo M. *et al.,* 2002b

1.1.2 Regulation of the *T. reesei* cellulase genes

Already early reports showed that cellulase production of *T. reesei* is subject to regulation by the carbon sources available (e.g. Mandels and Reese, 1960). Cellulose and its derivatives, lactose, and also the disaccharide sophorose induce cellulase expression in *T. reesei*. On the other hand, glucose prevents cellulase production. Cloning of many of the corresponding genes by differential hybridisation showed the regulation to take place, at least largely, on transcriptional level. Reports on cellulase induction have speculated on how a

highly crystalline and non-soluble compound like cellulose can affect the enzyme production inside the cell. The early hypotheses included contact of cellulose with receptors at the cell surface (Binder and Ghose, 1978), or starvation (Gong and Tsao, 1979) as the cause of the induction. Nowadays, it is generally accepted that sugars released from cellulose, or modified from those, act as inducers. However, it is not clear whether the inducers are formed by cellulases expressed at a low level constitutively (El-Gogary *et al.,* 1989, Carle-Urioste *et al.,* 1997), or whether conidial-bound cellulases form them (Seiboth *et al.,* 1997).

The study by Ilmén *et al.* (1997) confirmed that the cellulase regulation takes place at the transcriptional level, and that different cellulase genes are regulated co-ordinately. The *cbh1* mRNA level is always the highest, and its maximal induction level is several thousand fold. Glycerol and sorbitol were detected as neutral carbon sources, not inducing but not repressing either. After that finding, it was clear that cellulase gene expression is caused by separate induction and repression mechanisms.

Glucose repression is a well-studied phenomenon in *Aspergilli,* and the glucose repressor-encoding gene *cre1* of *T. reesei* was isolated based on the sequence similarity with *creA* (Strauss *et al.,* 1995, Ilmén *et al.,* 1996b, Takashima *et al.,* 1996b). The *cbh1* promoter contains many putative binding sites for the CREI protein. The functional importance in glucose repression of a cluster of three closely spaced sites around –700 bp upstream of the ATG was shown (Ilmén *et al.,* 1996a, Takashima *et al.*, 1996a, Ilmèn, 1997). One of the hyper-cellulolytic *T. reesei* strains, Rut-C30, showing cellulase expression also in the presence of glucose, was shown to contain a partially deleted *cre1* gene, whose product is non-functional (Ilmén *et al.,* 1996b).

A truncated form of the *cbh1* promoter as short as 161 bp, which contains only 30 bp upstream of the TATA box, is still inducible by sophorose (Ilmén *et al.,* 1996a). This was shown by linking the promoter variants to the *lacZ* reporter gene, and detecting the β-galactosidase activity on assay plate and the *lacZ* mRNA by Northern analysis. On the other hand, Henrique-Silva *et al.* (1996) reported the sequence –241 to -72 upstream of the TATA box necessary for the induction by cellulose. The promoter variants were linked to the β-glucuronidase-encoding gene and the GUS activity was measured from the

mycelia grown on cellulose. These results indicate that the induction by sophorose and cellulose work by different molecular mechanisms, and maybe use different transcription factors. The promoter downstream of –72 was enough for a low basal expression on glycerol, which was detected by linking the variants to an antibiotic selection gene and assaying for growth in the presence of selection (Henrique-Silva *et al.,* 1996). However, the integration sites in the genome were random, and such short promoters are easily affected by sequences present at the integration site. Therefore, a large number of individual transformants should have been analysed instead of three. Thus, the existence of the basal expression, which at least is too low to be detected by Northern or slot blot hybridisation methods (Ilmén *et al.,* 1997) remains to be properly shown.

A promoter region responsible for the induction of the *cbh2* gene has been characterised (CAE, *cbh2*-activating element; Zeilinger *et al.,* 1998). This region from –246 to –236 in respect to the ATG contains, partly overlapping and in different strands, a CCAAT box and a GTAATA sequence. CAE was shown to bind a protein complex, which was recognised by antibodies raised against *Aspergillus* HapC protein (Zeilinger *et al.,* 1998). Later, three factors, HAP2, HAP3 and HAP5 were isolated from *T. reesei,* which bind to the CCAAT box as a complex (Zeilinger *et al.,* 2001). This complex in turn binds co-operatively with (an) unknown GTAATA-binding protein(s) (Zeilinger *et al.,* 2003). CAE was detected as a nucleosome-free area in a micrococcal nuclease-based mapping, and the CAE-binding protein complex was found necessary in positioning of nucleosomes around it (Zeilinger *et al.,* 2003). One of the nucleosomes occupies the TATA box in the repressing conditions. In the inducing conditions, the protein complex is needed to remove or reposition the nucleosome, and this makes the TATA box accessible for transcription factors.

Ilmén *et al.* (1997) detected derepression of the cellulase genes from carbon catabolite repression long after the consumption of all glucose from the medium. Messenger-RNA levels of about 10% of the fully induced ones were detected co-ordinately for all major cellulase genes without addition of any cellulosic material. The mechanism behind this phenomenon is not known. One hypothesis is that glucans released from the fungal cell walls during autolysis could act as inducers. Starvation is another option.

From *Aspergilli*, the xylanase activator gene *xlnR* has been isolated (van Peij *et al.,* 1998b, Marui *et al.,* 2002). In addition to xylanase genes, it was shown to activate cellulase expression as well (van Peij *et al.,* 1998a, Gielkens *et al.,* 1999). XlnR was the first positively acting regulatory protein known to regulate cellulase synthesis in any eukaryote.

1.2 Production of bio-ethanol

Cellulosic plant materials are well suited for raw materials in the production of fuel ethanol by microbes. The economically feasible production process, however, relays on the utilisation of all the sugars in the materials. The cheapest, and at the same time most sustainable raw materials would be the wastes from agriculture or the pulp and paper industries. These contain besides cellulose, a large fraction of hemicellulose, which is mainly composed of pentose sugars. Many bacteria, e.g. *Escherichia coli, Erwinia chrysanthemi* and *Klebsiella oxytoga* are capable of consuming both hexose and pentose sugars. The production of ethanol by these organisms is, however; usually low (see Dien *et al.,* 2003, for a review on bacteria in bio-ethanol production). Metabolic engineering has been applied for the pentose-utilising bacteria to enhance the productivity of ethanol (Ingram *et al*., 1987). On the other hand, a good ethanol producer, *Zymomonas mobilis* has been engineered for xylose (Zhang *et al*., 1995) and arabinose (Deanda *et al*., 1996) utilisation. However, bacteria are usually rather sensitive to high ethanol concentrations and even more, to the inhibitory compounds found in the biomass hydrolysates (Zaldivar *et al.*, 2000).

Baker´s yeast *S. cerevisiae*, on the other hand, is one of the most efficient hexose fermenters known, and for this reason, it is a good candidate for an industrial producer of bio-ethanol. *S. cerevisiae* can tolerate ethanol in high concentrations, and the strains are rather resistant to inhibitors (Helle *et al*., 2003). Heterologous expression of a fungal laccase gene has been shown to enhance the inhibitor tolerance even further (Larsson *et al.,* 2001). Furthermore, cellulolytic *S. cerevisiae* strains have been achieved by expression of heterologous cellulase genes, and such strains could be used in simultaneous saccharification and fermentation of cellulose (van Rensburg *et al.*, 1998). A drawback of *S. cerevisiae* is that it cannot metabolise any pentose sugars. Thus the hydrolysates

of rice straw, corn stover or wastes from sulphite-pulping cannot efficiently be used as raw materials for the natural *S. cerevisiae* strains.

1.2.1 Engineering *S. cerevisiae* for xylose metabolism

Xylose-utilising *S. cerevisiae* strains have been achieved by metabolic engineering. The over-expression of the *Pichia stipitis* genes coding for xylose reductase (XR) and xylitol dehydrogenase (XDH) enzymes in *S. cerevisiae* resulted in xylose metabolised to ethanol (Kötter *et al.*, 1990, Walfridsson *et al.*, 1997). This pathway was improved by over-expression of the homologous xylulokinase-encoding gene (Eliasson *et al.,* 2000, Toivari *et al.,* 2001). However, the utilisation of xylose is incomplete and results in formation of side-products acetate, glycerol, and especially xylitol, excreted to the medium along with ethanol. Expression of the three xylose-utilisation genes in multiple copies resulted in simultaneous fermentation of both xylose and glucose (Ho *et al.,* 1998), but glucose was fermented several times faster than xylose. Furthermore, anaerobic growth of the strains on xylose alone was not achieved. The pathway, although it is redox-neutral, results in an imbalance because of the different cofactor preferences of the *P. stipitis* XR and XDH enzymes. Strategies for improving the process further have focused on relieving the imbalance (for a review, see Aristidou and Penttilä, 2000). Recently, random approaches, namely mutagenesis (Wahlbom *et al.,* 2003) and evolutionary engineering (Sonderegger and Sauer, 2003) have been used for improving growth and ethanol productivity from xylose. The latter study produced a strain mixture with at least two different sub-populations, the larger of which was capable of enhanced xylose metabolism and anaerobic growth on xylose as the sole carbon source. Recently, the functional expression of xylose isomerase (XI)-encoding genes of the bacterium *Thermus thermophilus* (Walfridsson *et al.,* 1996, Lönn *et al.,* 2003) or the anaerobic fungus *Piromyces* sp E2 (Kuyper *et al.,* 2003) has also been successful. The affinities of the xylose isomerases for xylose are rather high and, furthermore, no cofactors are needed in the isomerisation reaction of xylose to xylulose.

1.2.2 Uptake of xylose

Efficient uptake of the carbon source is necessary for fermentation to ensure a high substrate flux. Eukaryotic sugar permeases are either passive permeases, which take sugars up by energy-independent facilitated diffusion mechanism, or active, energy-demanding permeases. In *S. cerevisiae*, disaccharide permeases generally work by proton symport; protons and e.g. maltose are transported into the cells together. Hexose transporters, on the other hand, are facilitated diffusers (for a review, see Boles and Hollenberg, 1997). An exception to this is the uptake of the hexose sugar fructose, which in *S. carlsbergensis* is taken up by a specific fructose/proton symporter (Gonçalves et al., 2000). The proton symporters generally have much higher affinities for their substrates. Thus, they can take up sugars from low concentrations, and concentrate them inside the cell against the concentration gradient. The facilitated diffusion permeases are only able to transport the sugars along the gradient.

Although yeast cannot naturally metabolise xylose, it takes it up by the hexose transporter family (Kotyk, 1967, Hamacher *et al.,* 2002, Lee *et al.,* 2002). The hexose transporter system of baker's yeast is well developed. It consists of a family of nearly twenty permease proteins operating by facilitated diffusion. These belong to the major facilitator super family and share a structure consisting of 12 membrane-spanning regions. The genes are similar in amino acid sequence, the identities being between 65–99%. Each protein, Hxt1p-Hxt17p, and Gal2p has its own substrate specificity and affinity for the substrates. However, the affinities of all hexose transporters are relatively low (K_m's for glucose in the range of 1–100 mM; Reifenberger *et al.,* 1997, Wieczorke *et al.,* 1999, Maier *et al.*, 2002). Additionally, the genes are regulated differently to ensure the best possible combination of proteins to work on each condition. The production of the "high-affinity" hexose transporters (Hxt6p and Hxt7p, K_m for glucose around 1 mM) are under glucose repression, while the *HXT1* gene, which encodes a permease with the lowest affinity, is induced by high glucose levels. In addition, the *HXT2* and *HXT4* genes are induced on low-glucose conditions and HXT3 whenever some glucose is present. In addition to the structural proteins, *RGT2* and *SNF3* encode glucose sensors, which are similar in amino acid sequence to the hexose transporters. They are part of the signal transduction pathway, which regulates the expression of the different hexose transporter genes (for a review see Özcan and Johnston, 1999). In many

other yeasts, e.g. *Schizosaccharomyces pombe* or *Kluyveromyces lactis,* glucose is taken up by high-affinity permeases operating by proton symport (Billard *et al.*, 1996, Heiland *et al.*, 2000). This is likely a consequence of the different life styles of baker's yeast and the non-conventional yeasts, which generally live in environments with less hexoses present.

The affinities of the hexose transporters of *S. cerevisiae* for the pentose sugar xylose are very low, one to two orders of magnitude lower compared to that for glucose. The xylose-utilising bacteria take up the substrate by energy-demanding, high-affinity transporters, either xylose-proton symporters or ATP-binding cassette (ABC) transporters (Erbeznik *et al.,* 1998). In many naturally xylose-utilising yeasts, high-affinity xylose-proton symporters are also operating. Xylose uptake has been studied e.g. in *Debaryomyces hansenii* (Nobre *et al.,* 1999), *Kluyveromyces marxianus* (Stambuk *et al.,* 2003), *Pichia stipitis* and *Pichia heedii* (Does and Bisson, 1989, 1990), and several *Candida* species (Lucas and van Uden, 1986, Gárdonyi *et al.*, 2003, Stambuk *et al.,* 2003). However, glucose is the preferred carbon source also in these fungi, and its presence often inhibits xylose uptake. The high- and low-affinity uptake systems are often simultaneously present when the yeasts are growing on xylose. However, in *Candida shehatae*, which is a xylose-fermenting yeast species, starvation induces the expression of the xylose-proton symporter (Lucas and van Uden, 1986). On the other hand, in *K. marxianus*, fully aerobic growth on xylose is needed (Stambuk *et al.*, 2003). Although xylose uptake characteristics of many yeasts have been studied rather extensively, no genes encoding xylose-proton symporters, or specific xylose transporters in general, have been isolated from any eukaryote so far.

In the first xylose-utilising strains achieved, uptake of xylose was probably not the rate-limiting step (Kötter and Ciriacy, 1993). In a strain with a higher XR enzyme level, uptake already seems to restrict fermentation at least on low xylose concentrations (Gárdonyi *et al.,* 2002). After future strain improvements, also the uptake must be addressed. Already now the uptake may be limiting in a strain expressing a xylose isomerase (Kuyper *et al.,* 2003).

1.3 Yeast *S. cerevisiae* as a tool in gene-cloning

First genes isolated by expression in *S. cerevisiae* were native yeast genes cloned by complementation of mutant yeast strains, e.g. the auxotrophic marker genes *HIS3* and *URA3* (McKnight and McConaughy, 1983). Complementation cloning of heterologous genes from cDNA libraries made from other eukaryotes into a yeast expression vector followed; the filamentous fungal *his3* (Goldman *et al.,* 1992), and the human MAP kinase kinase kinase (Takekawa *et al.,* 1997) are examples of genes cloned this way. Genes encoding membrane proteins have as well been cloned by complementation, e.g. amino acid transporters from plants (Frommer *et al*., 1993), and mammals (Lin *et al.,* 1998).

1.3.1 Genetic selection methods

Different genetic selection methods (or reporter gene-based methods) have been developed, which use yeast as a screening host. These methods are based on the presence of two constructs in the same yeast cell, a reporter and a production construct. Chang and Timberlake (1992) identified binding sites of a regulatory protein with a reporter-based method in yeast, already before any reporter systems had been developed for gene-cloning. The developmental regulator BrlA of *Aspergillus* was produced in yeast. In the reporter construct, a reporter gene with the yeast minimal *CYC1* promoter was linked to a developmentally controlled *Aspergillus* promoter. The *in vivo* binding of the BrlA protein to its functional binding sites in the reporter construct activated the reporter gene. Later, the specific promoter regions responsible for the binding were screened for. Shorter fragments of the promoter, as well as short synthetic oligonucleotides were scanned for activation of the reporter by the BrlA protein. A similar strategy was used for the isolation of binding sites from genomic DNA fragments of *Drosophila* for a *Drosophila* homeoprotein (Mastick *et al.,* 1995). Wilson *et al.* (1991) isolated similarly binding sites of a rat hormone receptor protein in yeast, but the regulatory protein was produced as a fusion to a yeast activation domain.

The screening for a transcription factor that is able to bind to known binding sites (Wang and Reed, 1993, Inouye *et al.,* 1994) is commonly known as the yeast one-hybrid strategy. According to this method, multiple copies of the

known binding site are inserted upstream of the coding region of the reporter gene. The TATA box of the reporter gene is included in order for the basic transcription machinery of yeast to bind. Alternatively, a minimal promoter of another yeast gene, e.g. *GAL1* (Wang and Reed, 1993) can be used. A cDNA library is prepared from the same species into a plasmid, which contains a yeast promoter and the Gal4 transcriptional activation domain. Thus, the proteins are expressed as fusion proteins, which activate the transcription of the reporter gene through the Gal4 activation domain, if they bind the binding sites in the promoter region. Both activator and repressor-encoding genes can be cloned. In the early days, the reporter construct was integrated into the genome of the reporter strain. Later, the use of single-copy plasmids in the commercial kits has accelerated the procedure. It is nowadays also possible to include a second, unfused cDNA library in the screen in order to clone regulatory proteins needing auxiliary factors for their function. Recently also a modified one-hybrid method has been developed to detect interactions between a protein and methylated DNA (Feng *et al.*, 2004).

Proteins interacting with the studied protein can be screened from a library by the yeast two-hybrid method (Fields and Song, 1989, Chien *et al.*, 1991). The library is similarly to the one-hybrid method, fused to a yeast activation domain. The studied protein (the "bait") is produced as a fusion to a DNA-binding domain, e.g. LexA, whose binding sites are included in the reporter construct. When the library plasmid contains a protein capable of interacting with the bait, the interaction leads to dragging of the complex to the promoter and activation of transcription of the reporter gene. This method is applicable to protein-protein interactions, but not to those including regulatory proteins.

Additionally, the yeast tribrid method revealing post-translational modifications of interacting proteins has been developed (Osborne *et al.,* 1995, Volpers *et al.*, 2001). In this method, a third construct expressing the gene encoding a modifying protein is included.

A reporter-based method has been developed for cloning of protease-encoding genes as well (Smith and Kohorn, 1991). It uses the Gal4 activator protein, which is needed for the activity of the galactose-induced gene expression and thus, for growth of the yeast on galactose. The cleavage site of the protease in question is inserted in the Gal4 protein sequence between the DNA-binding and

activation domains. These two domains are separated in the yeast cell containing the right protease gene, resulting in a non-functional activator. Thus, this cell is not able to grow on galactose medium. By using toxic 2-deoxygalactose as the substrate, the negative selection can be changed into a positive one. This method could also be used as a screening method for a protease engineered by directed evolution methods, or for screening of cleavage sites of an isolated protease.

The methods discussed above are either based on selection or screening. Most of the selection methods use the expression of the yeast *HIS3* gene, and select for growth of an auxotrophic yeast strain on a medium lacking histidine. Resistance to the antibiotic neomycin has additionally been used. With both selection methods, the expression level of the reporter gene can be studied by using different concentrations of a counter-selection agent, aminotriazol, in the case of histidine biosynthesis, or different concentrations of the antibiotic neomycin. Because only the "right" clones survive, there is hardly any limitation in the amount of colonies plated. However, preparation of large libraries into yeast expression plasmids can be difficult, and in practice, the size of the library can restricts the number of clones to be screened.

The screening methods are based on different appearance of the cells expressing or not-expressing the reporter. Therefore, the amount of clones screened per plate cannot be as high as in the selection. These methods usually use the blue colour formed from X-gal by the β-galactosidase enzyme coded by the *E. coli lacZ* gene. Both selection and screening methods can also be used together in order to diminish false positives, which occur in all yeast screens.

1.4 Aims of the present study

When this work was initiated, our earlier studies had revealed that the *T. reesei* cellulase genes are induced very strongly and co-ordinately, and that the removal of repression is not sufficient for the expression to begin (results published later as Ilmén *et al.,* 1997). Therefore, it was reasonable to hypothesise that DNA-binding activator(s) are involved. In the promoters of the co-ordinately-expressed genes, no clear similarities could be detected to serve as activator binding sites. No activators of any hydrolytic enzyme genes were known from any eukaryotes then. The initial aim of the work was thus to develop a method

for the isolation of positively acting regulatory genes for the *cbh1* gene. The method had to be designed in such a way that no previous data on the protein(s) or the binding sites would be needed. Yeast was chosen as the host for the novel reporter gene-based method.

During the construction of the cDNA library of *T. reesei* into a yeast expression vector, additional aims arose. Why not use the yeast expression library for cloning of other genes as well? We aimed at setting up a method for the isolation of genes based on detection of hydrolytic activities of the encoded proteins on plate assays. Furthermore, complementation of a yeast hexose transporter-deleted and xylose-metabolising yeast strain for growth on xylose was aimed at by transforming an expression library into it. A further aim was to compare the native yeast hexose transporters in xylose uptake, and to study the kinetics of uptake and growth properties of strains expressing single *HXT* genes separately.

2. Materials and methods

All the materials and methods used, except those described in sections 2.1 and 2.2 are described in detail in the original articles (I–IV).

2.1 Analysis of the β-glucanase-encoding gene and the encoded protein

2.1.1 Isolation of the genomic glucanase gene

The chromosomal β-glucanase gene was isolated from a genomic cosmid library of *T. reesei* (Mäntylä *et al*., 1992) by colony hybridisation using the whole cDNA as a probe, and subcloned into the Bluescript SK (-) vector (Stratagene, CA, USA) as an 8 kb *Eco*RI fragment.

2.1.2 Sequence comparisons

Similarity searches were made by the Blast2 programme at NCBI (www.ncbi.nlm.nih.gov/BLAST) or European Bioinformatics Institute (www.ebi.ac.uk/cgi-nin/blast). The Conserved Domain Database (CDD) was searched with the search engine at NCBI (www.ncbi.nlm.nih.gov/Structure/cdd). Multiple alignments shown in this thesis were made by using the ClustalW multiple sequence alignment programme at www.ch.embnet.org.

2.1.3 *E. coli* expression

The coding region of the β-glucanase gene after the putative signal sequence cleavage site (amino acid 21, Saloheimo and Penttilä, 1993) was amplified from the yeast plasmid by PCR using Vent polymerase (New England Biolabs, MA, USA) and sequence specific primers, which created *Hin*dIII and *Bgl*II sites to the 5´- and 3´-ends of the fragment, respectively. The fragment was digested with the above-mentioned restriction enzymes, followed by cloning into the similarly cut plasmid pFLAG1 (IBI FLAG Biosystems, NY, USA) resulting in plasmid pAS25.

Strains JA221 and RV308 were used as hosts. LB plates used for expression of the protein were supplemented with 100 μg/ml ampicillin and 1 mM IPTG.

2.1.4 Production of antibodies

The RV308 cells transformed with pAS25 were grown at 37°C in ampicillin-containing LB medium (100 μg/ml) to an OD_{600} of about 1.5, followed by an overnight induction at 30°C with 1 mM IPTG. Cells were harvested and fractionated (Sambrook and Russell, 2001). The FLAG-fusion protein was purified from the cytoplasmic fraction using anti-FLAG M2 affinity chromatography according to the IBI FLAG Biosystems manual. The purified, non-denatured protein was used to immunise rabbits (KUO:NZW, Laboratory Animal Centre, Helsinki University).

2.1.5 Baculovirus expression

The coding region with five preceding nucleotides before the ATG and continuing to the amber stop codon (nucleotides 14–882 in Saloheimo and Penttilä, 1993) was amplified from the yeast plasmid as above. The sequence-specific primers created the *Bam*HI site to the 5´ and *Eco*RI site to the 3´ end of the fragment facilitating cloning to the *Bam*HI-*Eco*RI cut expression vector pFastBac1 (Bac-To-Bac expression System, GibcoBRL, NY, USA) under the polyhedrin promoter. The resulting pAS32 plasmid was transformed by electroporation into the *E. coli* strain DH10Bac (GibcoBRL) containing the baculovirus shuttle vector bMON14272 and the helper plasmid pMON7124, resulting in the formation of the recombinant bacmid, which contained the production cassette of the pAS32 vector. Recombinant baculovirus particles were achieved by transfecting the *Spodoptera frugiperda* Sf21 insect cells with the bacmid in the presence of the CELLFECTIN reagent according to the Bac-To-Bac Baculovirus expression system. The insect cell line *Spodoptera frugiperda* Sf9 was used in the viral amplification and *Trichoplusia ni* line High five in the production of the recombinant protein. Insect cells were cultured in Sf-900 II Serum-Free Medium according to the Bac-To-Bac Baculovirus Expression System (GibcoBRL).

2.1.6 Substrates

The substrates used were barley β-glucan (viscosity 20–30 c.s.; Biocon, UK), hydroxyethyl cellulose (HEC; Fluka, Switzerland), laminarin (Fluka, Switzerland), pachyman (Megazyme, Ireland), curdlan (Megazyme, Ireland), pustulan (Calbiochem, CA, USA), Konjak-glucomannan (Megazyme, Ireland), ivory nut mannan (Megazyme, Ireland), and birch xylan (glucuronoxylan, No. 7500; Carl Roth GmbH, Germany).

2.1.7 Activity plate assays

Enzyme activities were visualised on assay plates, which contained 0.1% HEC, 0.1% or 0.4% barley β-glucan, 0.4% pachyman, or 0.2% laminarin in conventional yeast (SC-Ura) or *E. coli* (LB+ampicillin) plates.

2.1.8 SDS-PAGE and Western blotting

For Western analyses and enzyme activity measurements, the yeast supernatant samples were concentrated 70-fold by ultrafiltration (Amicon, Millipore Corp., MA, USA) and the *E. coli* supernatant samples 33-fold by Centricon-10 microconcentrators (Amicon, Millipore Corp., MA, USA). Also the cell fractions of *E. coli,* and both the supernatant and cell fractions of the baculovirus-infected insect cells were analysed with Western. Proteins were separated according to conventional methods in 10–12.5 % SDS-PAGE gels, Western blotted to nitrocellulose filters and treated with the polyclonal antibodies produced in this work. The secondary antibody was alkaline phosphatase-conjugated anti-rabbit IgG (H+L) (Bio-Rad, CA, USA), which was detected by using ProtoBlot Western Blot AP (Promega, WI, USA). For the *E. coli* samples, also the monoclonal anti-FLAG M1 and M2 antibodies of the IBI-FLAG Biosystem were used together with the anti-mouse secondary antibody (Bio-Rad, CA, USA).

2.1.9 Enzyme activity measurements

Enzyme activity assays were performed as follows: The substrates (0.5% barley β-glucan, laminarin or HEC) were incubated with the concentrated yeast or *E. coli* culture supernatants, or with the buffered baculovirus culture supernatant in 50 mM sodium acetate buffer, pH 5.5 at 40 °C for 10 minutes. The reaction was stopped by boiling, and the reducing sugars liberated from the polymers were measured by the DNS reagent using glucose as the standard (Bernfeld, 1955).

The pH optimum of the enzyme was measured similarly using laminarin as a substrate in McIlvain's buffer with various pH values between 3,68 and 7,67.

2.1.10 Hydrolysis experiments

Hydrolysis of different substrates (barley β-glucan, HEC, laminarin, pachyman, curdlan, pustulan, Konjak-glucomannan, ivory nut mannan, and birch glucuronoxylan) by the baculovirus-produced enzyme was studied using two different concentrations of the enzyme, 1000 and 10 000 nkat per gram of the substrate, measured as activity against barley β-glucan. Two mg of the substrates were hydrolysed per ml of 50 mM sodium acetate, pH 5.5 at 40°C. Three different hydrolysis times (1, 5 and 24 hours) were used. The enzyme was buffered prior to use with 25 mM sodium acetate, pH 5, using Bio-Gel P6 columns (Bio-Rad, CA, USA). After hydrolysis, the reaction mixtures were boiled and assayed similarly as the samples in the activity assays.

2.2 Expression of a bacterial xylose-proton symporter in yeast

2.2.1 Strains and vectors

Bacillus megaterium type strain E-93055 was from the strain collection of VTT. The yeast expression vector YEplac195PGKPT (2μ, *URA3, PGK1* promoter) was constructed as follows. The expression vector pMA91 (Mellor et al., 1983) was digested with *Hin*dIII, and the resulting 1.8 kb *PGK1* promoter-terminator

cassette was ligated into the *Hin*dIII site of the YEplac195 vector (Gietz and Sugino, 1988). The orientation of the promoter-terminator fragment in the YEplac195PGKPT vector is *Hin*dIII-*PGK1* promoter-*PGK1* terminator-*Eco*RI, and the cloning site for expression from the *PGK1* promoter is a *Bgl*II site.

2.2.2 Expression of the *B. megaterium xylT* gene in yeast

The xylose-proton symporter-encoding gene *xylT* of the bacterium *Bacillus megaterium* was cloned by PCR using genomic DNA of the *B. megaterium* strain E-93055 as a template and sequence-specific primers based on the published sequence (Schmiedel *et al.*, 1997). *Bgl*II cloning sites were included in both primers to facilitate cloning. The PCR fragment was first cloned using the TOPO TA Cloning Kit (Invitrogen, CA, USA). Sequencing of the gene revealed 28 differences compared to the published sequence, out of which three resulted in changes in the amino acid sequence. These changes were present in all clones, sequenced from three separate PCR reactions, and thus most probably represent variation between the strains. The gene was cloned into the *Bgl*II cloning site of the yeast expression vector YEplac195PGKPT between the *PGK1* promoter and terminator.

3. Results and discussion

The results discussed in this thesis have been described in detail in the original publications, included here as appendices I–IV, except for the analysis of the β-1,3-1,4-glucanase gene, from which only the nucleotide sequence has been published (Saloheimo and Penttilä, 1993). Studies on that gene and the encoded protein are thus described here in more detail. Furthermore, the work done on the bacterial permease is unpublished, and it is described here.

3.1 New isolation method for enzyme genes based on expression in yeast (I, unpublished data)

Expression cloning of cellulase genes using *E. coli* as a host has been applied for many cellulolytic bacteria (Béguin *et al.,* 1987, Silva *et al.*, 1988). This method was also applied for the cellulase genes of the anaerobic fungus *Neocallimastix patriciarum* (Xue *et al.*, 1992). However, the attempts to produce *T. reesei* cellulases in active forms in this bacterium were unsuccessful (Teeri, 1987). Penttilä and co-workers (1984) isolated a β-glucosidase-encoding gene from a chromosomal cosmid library of *A. niger* by expression in yeast, and detected the activity on substrate plates. Yeast was also able to produce CBHI, CBHII, EGI and EGII in active forms, and the activities could be detected on assay plates (Penttilä *et al.*, 1987, 1988). Therefore, a yeast-based cloning method was considered feasible and developed.

Figure 1 shows the principle of the expression-cloning method. *T. reesei* was cultured for the RNA isolation on a medium containing many different polysaccharides and other carbon and nitrogen sources (Solka floc cellulose, distiller's spent grain, locust bean gum galactomannan, lactose, birch acetyl glucuronoxylan, and oat spelt arabinoxylan) in order to induce a maximal number of genes encoding extracellular enzymes (article I, Stålbrand *et al.,* 1995). cDNA was prepared and cloned into the yeast expression vector pAJ401 under the constitutive, strong *PGK1* promoter. The library was transformed into *S. cerevisiae*. In principle, each cDNA of the library is strongly expressed in a separate yeast cell. If a gene codes for a secreted enzyme and contains a signal sequence, the enzyme is secreted into the medium by the yeast cell. Plating of the cells on agar plates separates different yeast clones and the encoded enzyme

activities. The plates are stained to reveal halos around colonies, indicating hydrolysis of the substrate and the presence of the gene encoding the responsible enzyme. The method is not quantitative, because in practice, the efficiency of both expression and secretion of heterologous proteins vary. Thus, the size and intensity of the halo does not necessarily tell about the level of the activity of the native enzyme.

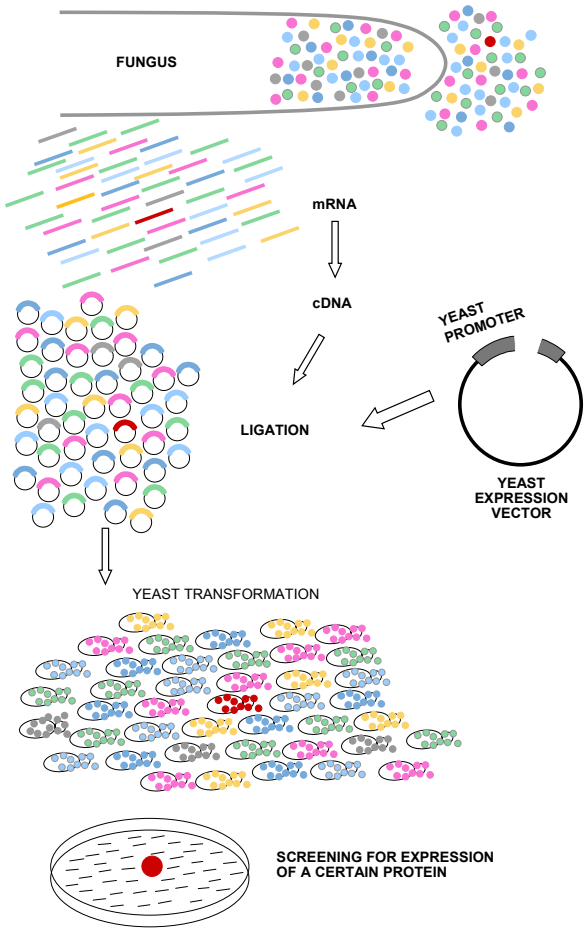


Figure 1. Isolation method for genes encoding hydrolytic enzymes. In this work, the library was screened on plates containing β-glucan from barley. The screened gene and activity are visualised with a red colour.

The library was screened on plates containing barley β-glucan (β-D-1,3-1,4-glucan), and two new genes were isolated. The first one was a novel, small gene of about 1 kb, coding for a previously unknown enzyme (Saloheimo *et al.*, 1993). Further assays on other substrate plates revealed it an endoglucanase, and the gene was named *egl5* and the deduced protein EGV. The second gene (Saloheimo and Penttilä, 1993) coded for a β-(1,3-1,4)-glucanase, which had no activity in the plate assays towards HEC, and thus could not be classified as an endoglucanase. The developed method thus allowed detection of genes encoding totally new, minor enzyme activities normally masked in *T. reesei* by the very abundant major cellulases. The isolation of the *egl5* gene by other methods would probably not have been successful. This method was later applied for other *T. reesei* enzymes: α-L-arabinofuranosidases (Margolles-Clark *et al.,* 1996b), α-galactosidases (Margolles-Clark *et al.,* 1996a), and lipases (Vidgren, V., Saloheimo, A., and Penttilä, M., unpublished).

Similar method was independently developed also at Novo Nordisk A/S, and applied for the isolation of the *H. insolens* cellulase genes (Dalbøge and Heldt-Hansen, 1994), as well as for hemicellulase and pectinase-encoding genes from *A. aculeatus* (Dalbøge, 1997, Pauly *et al.*, 1999).

The progress of genome sequencing has recently revealed new cellulase genes from *T. reesei* (Foreman *et al.,* 2003). From EST sequences, three genes encoding previously unknown endoglucanases and five β-glucosidase genes were predicted based on sequence comparisons to the databases (see Table 1.). Thus, the whole concept of gene-cloning is gradually getting old-fashioned as more and more genomes will be sequenced. However, even though candidates for new cellulases can be predicted from the sequences, it is very hard to be sure of the enzyme activities without functional data. Similar structure does not always predict similar function. E.g. the *T. reesei* cellulases CBHI and EGI are closely related in amino acid sequence and belong to the same GH7 family, even though one is a cellobiohydrolase, degrading cellulose processively from the ends of the glucose chains, and another, an endoglucanase splitting internal linkages. Therefore, yeast expression will remain as a valid technique in the annotation of eukaryotic hydrolytic enzyme genes.

3.1.1 *T. reesei egl5* encoding EGV (I)

The first new glucanase gene isolated with the developed method codes for an unusually small endoglucanase named EGV. It was classified into family 45 of glycosyl hydrolases (Henrissat and Bairoch, 1993) together with the *H. insolens* (Davies *et al.,* 1993) and *Fusarium oxysporum* (Sheppard *et al.,* 1994) enzymes with the same name, and endoglucanase B of *Pseudomonas fluorescens* (Gilbert *et al.,* 1990). Based on the hydrophobic cluster analysis, the catalytic residues of the four enzymes could be predicted, since only two aspartate residues were conserved between the homologous enzymes. The crystal structure of the *Humicola* enzyme has been solved (Davies *et al.,* 1993, 1995). It consists of a six-stranded beta-barrel domain with long inter-connecting loops. The structure and the site-directed mutagenesis performed (Davies *et al.,* 1995) ascertained that those residues, (in *T. reesei* EGV, Asp-27 and Asp-134) indeed are catalytic. Today, the relatively small family GH45 contains endoglucanases also from other filamentous fungi, e.g. *Humicola grisea, T. viride, Magnaporthe grisea* and *Melanocarpus albomyces.* Additionally, some endoglucanases from insect species, as well as mussels can be found. Three-dimensional structures are known from four different enzymes. The structure is distantly related to plant expansins. Only few bacterial enzymes belong to GH45, and there are no structures solved from any of them.

Similarly to most fungal cellulases, EGV contains a cellulose-binding domain (CBD) at its C-terminus, connected to the catalytic core with a flexible, glycosylated linker region. The structure of EGV CBD was molecular modelled in this work based on the NMR structure of the CBHI CBD. Some differences were detected in the structures of the overall very similar domains. The EGV CBD was less wedge-shaped than the CBHI CBD domain, and the hydrophobic surface important in the contact with cellulose was more rounded.

Old literature on *T. reesei* (earlier called *T. viride*) cellulases describes a small endoglucanase with a molecular mass of 20–25 kD (Håkansson *et al.,* 1978, Beldman *et al.,* 1985), or 12.5 kD (Berghem *et al.,* 1976). The former enzyme was not glycosylated, and amino acid sequences obtained from it (EGIV in Ståhlberg, 1991) showed similarity to a CMCase of *A. aculeatus* (Ooi *et al.,* 1990). Thus, this enzyme corresponds to EGIII belonging to family GH12 of glycosyl hydrolases (Ward *et al.,* 1993). The small endoglucanase of Berghem

and co-workers (1976) had an isoelectric point of 4.6 and was heavily glycosylated. Therefore, it could not be EGIII, which has a high pI-value and does not contain a cellulose-binding domain or a linker region, and thus contains less O-glycans. That enzyme was efficient in releasing fibres from filter paper. It is impossible to say, whether that enzyme represented EGV or not.

A study performed recently to *T. reesei* EGV (Karlsson *et al.,* 2002) showed that the enzyme has a limited activity to β-glucan, no activity at all to cello-oligosaccharides, and the main product formed from cellulose by it is cellotetraose. This is in contrast to *H. insolens* EGV, which is able to degrade cello-oligosaccharides, and produces cellobiose and cellotriose as main products from cellulose, the way that other endoglucanases of *T. reesei* also do. *T. reesei* EGV, however, had the highest activity of all isolated *T. reesei* endoglucanases towards the soluble Konjak-glucomannan. Glucomannan thus seems to be the favoured substrate of *T. reesei* EGV although it does not degrade mannan, and thus the enzyme is a glucomannanase rather than a strict endoglucanase. Interestingly, the full-length EGV enzyme and the core enzyme alone behaved very similarly in the hydrolysis experiments on all substrates studied (Karlsson *et al.,* 2002). The core enzyme alone was able to bind to cellulose very tightly. This is in contrast to EGIII, which binds to cellulose rather loosely.

The new cloning method thus enabled us to isolate a gene encoding a previously unknown cellulase enzyme. Although this gene codes for a minor endoglucanase activity, its glucomannanase activity is significant.

3.1.2 Analysis of the new *T. reesei* β-1,3-1,4-glucanase (unpublished data)

In the screening of the *T. reesei* expression library on barley β-glucan-containing plates (I), four active clones were recovered, which gave an identical restriction pattern with each other, and were clearly different from any earlier-isolated hydrolase genes of *T. reesei*. Retransformation of the plasmids into yeast confirmed that the activities were caused by the cDNA inserts of about 1.4 kb. One of these plasmids, pAS5, was studied further. The 1408 bp long cDNA of this plasmid codes for a deduced protein of 287 aa starting from the first ATG initiation codon at position 436 of the cDNA (position 19 in Saloheimo and

Penttilä, 1993). There are three in-frame stop codons as well as stop codons in the other frames at the beginning of the cDNA, and thus the cDNA is probably full-length. The size of the mRNA, 1.5 kb, is also in accordance with the size of the cDNA (data not shown). There is an A at position -3 (nucleotide 433) as is often the case with fungal genes. The coding region starts with a putative signal sequence of 21 aa, and the proposed cleavage of the signal peptide occurs after Ala-21, leaving 266 aa as the size of the proposed mature enzyme. There is no homology to fungal cellulose-binding domains, which are present also in many non-cellulose-hydrolysing enzymes of *T. reesei*. In the genomic gene, two introns can be found.

Expression of the new β-glucanase gene was studied by Northern hybridisation of *T. reesei* grown on different culture media. Sophorose, the most potent inducer of the cellulase system, induced also the β-glucanase gene when added to a sorbitol culture (Table 2). However, the gene was also expressed on glucose and sorbitol media, which is in contrast to the cellulase genes. Growth on Solka floc cellulose resulted in very low expression, as well as growth on lactose. Thus, the β-glucanase gene is not regulated co-ordinately with cellulases, and it is not under glucose repression. Sophorose induces the expression of many hemicellulose-degrading enzymes along with cellulases (Margolles-Clark *et al.*, 1997). The *T. reesei* RutC30 strain cultured in conditions from which the cDNA library was prepared (Stålbrand *et al.,* 1995), resulted in low expression of the β-glucanase gene (Table 2, inducer mix). Four copies of the gene were found from the library, compared to twenty of the *egl5* gene. The latter gene was expressed in that culture at a rather high level but no comparable hybridisations were made, and thus the comparison of the expression levels is impossible.

*Table 2. Relative expression levels of the new β-glucanase gene on media containing different carbon sources. The intensities of the signals in the Northern hybridisation were compared by eye in the scale: +/-, very low expression; +++++++, very high expression. Relative expression levels of the main cellulase gene cbh1 are shown for comparison (data from Margolles-Clark et al., 1997; the scale: +/-, very low expression; ++++, very high expression). The expression of the two genes can be compared qualitatively, not quantitatively to each other (Tiina Nakari-Setälä, personal comm.). *RutC30 strain was cultured in the same conditions than used in this work for preparation of the library.*

Carbon source	*bgl*	*cbh1*
Glucose	++	-
Glucose derepressed	+++	+++
Glucose + sophorose	+	-
Sorbitol	+++	-
Sorbitol + sophorose	+++++++	++++
Cellobiose	+++	++
Solka floc cellulose	+/-	++++
Lactose	+/-	++
Lenzing xylan	++	+/-
RutC30 strain, inducer mix*	+	++++

The deduced protein sequence of the new gene was analysed with the conserved domain search engine at NCBI. The analysis showed that the enzyme belongs to family GH16 of glycosyl hydrolases. It is a big family containing enzymes that degrade β-1,3- and/or β-1,4-linkages between glucose residues in β-1,3-glucan or mixed glucans, but not in cellulose. Additionally, it contains agarases, xyloglucan endotransferases, β-galactosidases, and bacterial binding proteins. Similarity searches with the Blast2 programme gave highest amino acid similarities (identities 37–55%, similarities 58–65%) to several open reading frames deduced from the *N. crassa* genome. Highest similarities to known enzymes were to endo-1,3-β-glucanases of *Streptomyces sioyaensis* and *Oerskovia xanthineolytica* (identities 33–34%, similarities 48–50%), and to a β-1,3(4)-glucanase of *Bacillus circulans* (identity 31%, similarity 48%) (data not shown).

The soluble, linear β-glucan from barley is composed of glucose units linked together with β-1,3- (29%) and β-1,4- (71%) linkages (Henriksson *et al.,* 1995).

Thus the activity encoded by the new gene could be directed towards either or both of these linkages. Since family GH16 contains both kinds of enzymes, no prediction of the enzyme activity could be made based on the classification of the sequence into this family. The activities of the yeast strain DBY746 carrying the pAS5 plasmid were therefore first studied by plate assays using various carbohydrates as substrates. No activity was detected on plates containing hydroxyethyl cellulose (HEC), the small synthetic substrates methylumbelliferyl cellobioside (MUC) or methylumbelliferyl lactoside (MUL), or on RBB-xylan. This indicates that the new gene codes for neither a classical cellulase nor a xylanase enzyme. The product of the new gene showed clear activity against barley β-glucan but the halos detected were smaller than those of the strains producing the cellulases EGI, EGII or CBHII used as controls (data not shown, see also Penttilä, 1987). However, the expression levels and the secretion efficiencies of foreign proteins in different yeast strains may vary, and thus it is not possible to draw any definite conclusions concerning the level of the enzyme activity based on the plate assays.

Different β-1,3-glucans were also tested in the plate assays. On plates containing laminarin, a soluble β-1,3-glucan polymer, halos on otherwise poorly stained plates could be detected. Somewhat smaller halos were also detected from the control yeast strain. β-1,3-glucans curdlan and pachyman (insoluble β-1,3-glucans) did not work well in the plate assays but some activity could be detected after a long incubation on pachyman (data not shown).

As the plate assays gave no clear answers of the enzyme activity, the yeast strain containing the pAS5 plasmid was next cultured in a 1-liter fermentor, and the culture filtrate was concentrated ~70 times. The activity assays performed showed that the activity against barley β-glucan was rather low (1.63 nkat/ml) but ten times higher compared to the negative control strain containing the vector only (0.17 nkat/ml; Matti Siika-aho and Maija Tenkanen, personal comm.). No endoglucanase activity, measured as activity against HEC, was detected. Both the transformant and the control strain produced similar activities of about 50 nkat/ml against laminarin (Matti Siika-aho and Maija Tenkanen, personal comm.). From an NMR analysis of relations of different oligomers in the end-product pool of β-glucan hydrolysis with the concentrated yeast supernatant, it could be concluded that both β-1,4- and β-1,3-linkage were hydrolysed (Anita Teleman, personal comm.). Traces of hydrolysis products were also seen with the control yeast

containing the vector only. From the yeast expression, however, no definite conclusions of the enzyme activities could be made because of the possible synergistic effects with the β-1,3-glucanases that *S. cerevisiae* is known to secrete (for a review, see Pretorius, 1997).

In order to analyse the activity further and to produce antibodies against the protein, the coding region without the signal sequence was cloned by PCR into the *E. coli* expression vector pFLAG1. From that vector, *E. coli* produces the protein as a fusion protein, which contains a bacterial *ompA* signal peptide of 21 aa and an 8 aa FLAG purification tag. Over 90% of the products after an overnight induction at 30°C were found from the cytoplasmic fraction, the rest being in the periplasmic fraction and the culture medium (data not shown). The M1 antibody, which is designed to detect only an N-terminal FLAG peptide present in the secreted protein, from which the ompA signal peptide has been correctly spliced, did not detect any protein from any fraction. This suggests that the proteins detected from the periplasm and supernatant by the M2 antibody were resulting from cell lysis and not secretion. The protein was purified from the cytoplasmic fraction by using affinity chromatography with the M2 antibody. Only one band could be detected in SDS-PAGE and Western analysis from the purified protein (data not shown). When the purified protein or samples from different cell fractions were analysed, no activity against barley β-glucan, HEC, or laminarin could be detected. On assay plates the results were negative as well (data not shown). Thus the enzyme produced by *E. coli* was inactive and could not be used in the characterisation of the enzyme activity.

The purified protein was used in the immunisation of rabbits. The polyclonal antibodies produced were tested and found out to have only slight cross-reactivity with other *T. reesei* hydrolases, and to detect the *E. coli*-produced protein well. When the concentrated yeast supernatant was used in Western analysis, mainly smaller degradation products were detected, indicating proteolysis of the sample (data not shown).

In order to produce the protein in an active form and in the absence of contaminating enzyme activities, baculovirus expression was applied. Expression of the *T. reesei* CBHI protein was successful in insect cells using a recombinant baculovirus expression vector (von Ossowski *et al.,* 1997) even though this protein could not be produced in active form in *E. coli* (Teeri, 1987). The coding region

and the nucleotides just preceding the ATG start codon of the β-glucanase were cloned by PCR into the pFastBac1 baculovirus expression vector after the polyhedrin promoter. After sequencing, the construct was transformed into the *E. coli* strain containing the bacmid vector, resulting in helper plasmid-mediated transposition of the expression cassette into the bacmid. The recombinant bacmid was used in transfection of insect cells, and the fresh viral stock in infection of a high-producing cell line, leading to high level production of the recombinant protein. In SDS-PAGE, the recombinant protein formed the major band (data not shown). In Western analysis, the size of the product was slightly smaller than the protein produced in *E. coli* and it ran as a duplet (Fig. 2). Major part of the protein was detected from the cell fraction.

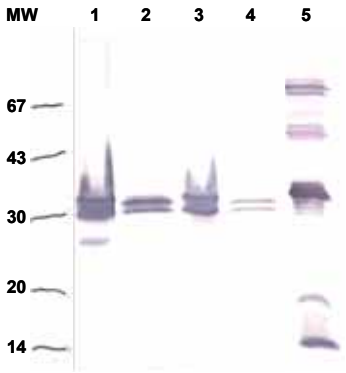


Figure 2. Western analysis of the baculovirus-produced protein and the protein produced by E. coli with the antibodies produced in this work. Lanes 1 and 3, cell fractions of two parallel baculovirus production lines; lanes 2 and 4, supernatant fractions of the former two lines; lane 5, cell fraction of the E. coli production strain. The molecular weight markers are shown on the left.

The baculovirus culture filtrate was buffered and used in activity assays with laminarin as a substrate. The pH optimum of the enzyme was 5.5–6, which is higher than that of most of the *T. reesei* hydrolases (data not shown). Hydrolysis experiment of laminarin and β-glucan with two different dosages of the enzyme and different reaction times is shown in Figure 3. The reaction shows saturation, especially on β-glucan, which could indicate lack of suitable linkages after extending hydrolysis. Table 3 shows the amounts of reducing sugars liberated

from various substrates after 24 h of hydrolysis at 40˚C with the higher enzyme dosage (10 000 nkat/g). The highest amount was achieved with β-glucan, followed by laminarin. The other β-1,3-linked substrates were hydrolysed moderately. No activity towards β-1,6-linkages or mannan could be detected, except for a minor hydrolysis of Konjak-glucomannan. No background activities could be detected from the insect cells infected with the control viruses (data not shown).

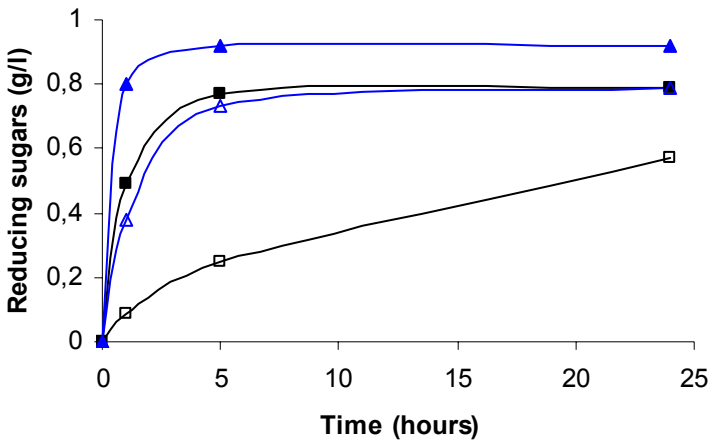


Figure 3. Production of reducing sugars from laminarin (squares) and β-glucan (triangles) by using two different enzyme dosages (1000 and 10 000 nkat/g of substrate) and three reaction times (1, 5 and 24 hours). Liberated reducing sugars, measured as glucose, are shown as the function of the reaction of time. Symbols: □, laminarin hydrolysed by 1000 nkat/g; ■, laminarin hydrolysed by 10 000 nkat/g; ∆, β-glucan hydrolysed by 1000 nkat/g; ▲, β-glucan hydrolysed by 10 000 nkat/g, (Maija Tenkanen and Matti Siika-aho, personal comm.).

Table 3. Hydrolysis of different polysaccharides by the baculovirus culture supernatant, values from hydrolysis of 24 hours with 10 000 nkat/g are shown (Maija Tenkanen and Matti Siika-aho, personal comm.).

Substrate	Composition	Source	Reducing sugars (g/l)
β-glucan	β-1,3-1,4-glucan	Barley	0.92
Laminarin	Soluble β-1,3-glucan	*Laminaria* spp.	0.79
Pachyman	Insoluble β- 1,3-glucan	*Poria cocos*	0.13
Curdlan	Insoluble β-1,3-glucan	*Alcaligenes faecalis*	0.13
Konjak-glucomannan	β-1,4-glucomannan	*Amorphophallus konjak*	0.01
Pustulan	β-1,6-glucan	*Umbilicaria populosa*	0
Mannan	Linear β-1,4-mannan	Ivory nut	0
Roth xylan	Acetyl glucuronoxylan	Birch	0

^{1}H NMR analysis of the hydrolysed β-glucan showed that the enzyme hydrolysed only β-1,4-linkages of β-glucan by retention (Anita Teleman, personal comm.). The number of these linkages decreased from the initial 71% (mol-% of total glucose moieties) to 42%, whereas the number of β-1,3-linkages was unchanged ~29%. However, almost all of the β-1,3-linkages were changed from internal to adjacent to the reducing end. The final hydrolysis products were the trisaccharide 3-*O*-β-cellobiosyl-D-glucopyranose and the tetrasaccharide 3-*O*-β-cellotriosyl-D-glucopyranose. The structure of barley β-glucan consists mainly of cellotriocyl and cellotetracyl moieties, separated by single β-1,3-linkages (Bielecki and Galas, 1991). Thus, the enzyme activity corresponds to that of 1,3(4)-β-D-glucanases, or laminarinases (EC 3.2.1.6), which hydrolyse 1,3- or 1,4-linkages in β-D-glucans when the glucose residue whose reducing group is participating in the cleaved linkage is itself substituted at C-3. The enzyme was thus named LAMI and the gene *lam1*.

Expression in the baculovirus host was necessary for the successful classification of the enzyme activity as that of a laminarinase, EC 3.2.1.6. Baculovirus-insect cell expression system has been used for the expression of similar activities from barley. In the plant, multiple enzyme activities are expressed during the germination of the grain. Doan *et al.* (1993) produced a β-1,3-glucanase and two β-1,3-1,4-glucanase enzymes separately in insect cells, and 10–15 mg of the proteins were secreted per litre of the culture medium. The proteins were purified, and they possessed correct enzyme activities. However, there were differences in

the glycosylation patterns, and in one case, the splicing of the signal peptide was incomplete. The *T. reesei* enzyme produced by the insect cells was present as two equally strong bands, and the majority was detected from the cell fraction (Fig. 2). Both forms were equally present both within the cells and in the medium. They probably represent differentially glycosylated forms of the enzyme. The amount of the protein produced in this work was not studied. However, it was high enough for characterisation of the enzyme activity.

The yeast-based method for cloning of hydrolase genes proved successful in the isolation of glucanase genes from *T. reesei*. The functioning of the method showed that *S. cerevisiae* and filamentous fungi share important components, e.g. signal sequence of T. reesei works in yeast and the enzymes are active after production in this heterologous host. For many substrates, yeast does not produce background activities. The method is especially valuable in cloning of enzymes possessing low activities. At Novo Nordisk, similar method has successfully been used in cloning of various hydrolase genes from many fungal species (for a review, see Dalbøge, 1997).

3.2 New isolation method for regulatory genes, applied for *T. reesei* cellulases (II, III)

The second cloning method developed in this work deals with regulation of the cellulase genes. The expression of the cellulase genes of *T. reesei* is regulated by the carbon sources available. Cellulose and related substances induce the expression. On the other hand, glucose and fructose repress the genes even in the presence of an inducer. Glycerol and sorbitol can be considered as neutral carbon sources, which allow expression in the presence of an inducer. Thus, separate induction and repression mechanisms operate in the cellulase gene regulation of *T. reesei* (Ilmèn *et al.,* 1997). When this study was initiated, only the repression mechanism mediated by the CREI protein had been studied (Ilmèn, 1997). The deletion analysis performed to the *cbh1* promoter had not given clear answers on where the activator binding sites were situated (Ilmèn *et al.*, 1996a). For cloning of gene(s) responsible for the strong induction of the *cbh1* gene, a method was needed that was not based on any previous knowledge of the activator proteins or their binding sites. Yeast was chosen as a host for this screening method as well.

Figure 4 shows the idea of the developed cloning method for activator genes. It relays on the presence of two different plasmids in a single yeast cell: the reporter plasmid and the library plasmid. These two plasmids carry different selection markers. Additionally, the reporter plasmid carries a third selection gene, whose expression in yeast is under the regulation of the studied heterologous promoter, here the *cbh1* promoter. When the library is screened, a heterologous activator protein is expressed from the library plasmid and binds to the heterologous promoter. Mere binding is, however, not enough but in addition, the activation domain present has to play together with the basic transcription machinery of yeast and activate transcription of the third marker gene. The selection for growth on culture plates allows screening of a large library very easily.

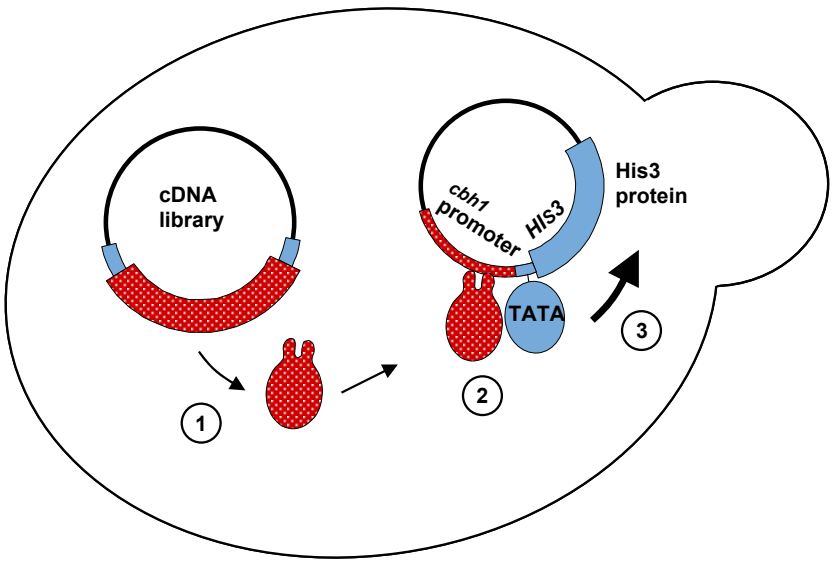


Figure 4. Isolation method for activator genes, here applied for the cbh1 promoter of T. reesei. The yeast cell contains two plasmids, the reporter and the library plasmid. The cell containing the activator cDNA produces the activator from the yeast promoter (1). The protein (red sphere with fingers) binds to its binding sites in the cbh1 promoter (2). As a result of the interplay with the basic transcription machinery of yeast (blue sphere at TATA), the expression of the HIS3 gene is activated (3), which rescues the cell in –his medium. Red colour, T. reesei components; blue colour, yeast components.

In order to detect factors activating cellulase expression, the full-length *cbh1* promoter of 1.15 kb upstream of the TATA region was linked to the promoter-less yeast *HIS3* gene in the reporter plasmid (Fig. 4). Only the regulated TATA region of the *HIS3* promoter was included in order to allow the basic transcription machinery of yeast to bind the promoter. This construct alone did not support growth of the auxotrophic yeast strain on a medium lacking histidine. The same yeast expression library as used in the screening for glucanase genes, was transformed into the reporter yeast strain, and selected for growth on –his medium. The screening resulted in growth of the cells with the frequency of 1/25.000. Along with the *his3* gene of *T. reesei*, two new genes encoding putative regulatory proteins were isolated. The developed method is applicable to any eukaryotes and any regulatory systems. No binding site information of the regulatory proteins is needed. Instead, the full-length promoter can be used, provided that it does not contain elements recognised by the activator proteins of *S. cerevisiae*. Thus, species closely related to baker´s yeast cannot be studied. Conserved activator binding sites can cause background problems with more distantly related species as well.

Yeast had earlier been used as the host when searching for regulatory proteins binding to known targets. According to the yeast one-hybrid method (Wang and Reed, 1993), multiple copies of the binding site are cloned in front of a minimal yeast promoter in the reporter plasmid. The cDNAs are linked to an activation domain in the library plasmid. Thus, DNA binding proteins, irrespective of their nature, are isolated. In another method, yeast was used as a host in promoter analysis for binding sites of a known heterologous DNA-binding activator (Chang and Timberlake, 1992, Wilson *et al.,* 1991). According to the latter method, the promoter under study, or fragments of it, together with a minimal yeast promoter was linked to a reporter gene, and the regulatory protein under study was produced from a separate plasmid. Our method in a way combines the former two. A similar method has been developed, which uses the fission yeast *Schizosaccharomyces pombe* as a host (Remacle *et al.,* 1998). The full-length mammalian promoter was used, but the gene library was expressed as fusions to the VP16 activation domain. Chan *et al.,* (1993) used still another method, where the known binding site of a human activator was used in double in the reporter plasmid. The library screened was a normal yeast expression library without fusion partners. The activator isolated bound DNA with its DNA-binding domain, and the activation domain present activated the expression of

the reporter. Thus, genetic selection methods applying yeasts have widely been used in isolation of regulatory genes from various organisms. The method developed in this work was designed for the isolation of activator proteins instead of any DNA-binding proteins. Furthermore, because long promoters are used, multiple activator genes can be isolated within a single screen. In this work, two regulatory genes were cloned.

3.2.1 ACEI (II)

The first new gene called *ace1* (activator of cellulase expression) was isolated as a 1,9 kb cDNA. The clone was soon noticed to represent a non-full-length mRNA, and the full-length cDNA of 3.2 kb, as well as the chromosomal copy of the gene were cloned. The encoded protein contained two stretches of basic amino acids resembling a bipartite nuclear targeting signal, and partially overlapping, a DNA-binding domain, containing three zinc fingers of the Cys_2His_2 type. Two of the fingers fitted to the consensus of this class of DNA-binding proteins, but the middle one had too long a loop between the second and third zinc-co-ordinating residues. However, a similar atypical finger could be found in the same configuration also in a yeast meiotic inhibitor RME1 (Covitz *et al.,* 1991). Otherwise these two proteins are dissimilar.

Similarity searches performed lately to ACEI showed the highest similarities to sequences encoded by the genomes of *Magnaporthe grisea* (gi38104048; identity 50%, similarity 64%), and *N. crassa* (gi32423175; identity 48%, similarity 60%). Additionally, ACEI of *Talaromyces emersonii* (gi33115142; Murray, P. G., Collins, C. M., and Tuohy, M. G., unpublished) showed 36% identity and 48% similarity to the *T. reesei* ACEI protein. The *Aspergillus nidulans* protein encoded by *stzA* (gi14195703) is also ~35% identical to ACEI. The alignment of the *T. reesei* ACEI protein with the two hypothetical proteins most similar in amino acid sequence with it is shown in Figure 5.

*Figure 5. Alignment of the ACEI protein of T. reesei (TRIRE) with the open reading frames of M. grisea (MAGGR) and N. crassa (NEUCR), which showed the highest amino acid similarities to ACEI. Darker blue, the putative nuclear localisation signal; light blue; the conserved cysteine and histidine residues of the three Zn-fingers. *, identical amino acids; :., similar amino acids.*

The zinc finger region of ACEI (amino acids 382-582) was produced as a GST fusion in *E. coli*, purified and used in *in vitro* binding assays initially with four labelled, partly overlapping fragments of the *cbh1* promoter. Retarded bands were seen at non-denaturing polyacrylamide gel runs with three fragments, indicating formation of DNA-protein complexes. Competition experiments with excess of the same unlabelled fragments showed that the binding was specific. The sequence of one of these fragments, region –941 to –631 relative to ATG, which bound the ACEI-GST fusion protein the best, seemed especially interesting. It contained a repeat of three functional CREI binding sites responsible for the glucose repression of the *cbh1* gene (Ilmén *et al.,* 1996a). Additionally, three CCAAT motifs, and five GGC(T/A)AA motifs of unknown functions could be found at this area. First by using shorter PCR fragments, then 30–40 bp oligonucleotides, and finally mutated oligonucleotides in the *in vitro* binding assays, the consensus sequence of the binding site of ACEI was shown to be AGGCA. Eight such sites out of the nine present in the *cbh1* promoter were able to bind ACEI-GST *in vitro*, and the trombin-cleaved ACEI binding region alone, whenever it was tried.

The alignment of ACEI with the hypothetical *M. grisea* and *N. crassa* proteins showed that the zinc finger regions of the three proteins are almost identical (Fig. 5). Thus, the *M.* grisea and *N. crassa* proteins most probably bind similar sequences as shown for ACEI of *T. reesei*. In addition to the DNA-binding region, the putative nuclear localisation signal is almost completely conserved between the three proteins. At the N-terminal parts of the proteins, short stretches of identical amino acids are alternating with stretches of more variability. The identical parts can be speculated to be important for the shared functions of the proteins. However, no clear regulatory regions can be predicted other than the DNA-binding region. In the *T. emersonii* and *N. nidulans* proteins, there are more differences in the amino acid sequences of the DNA-binding regions compared to the three proteins shown in Figure 5 (data not shown), and the binding sites may differ from those determined for *T. reesei* ACEI.

In parallel with the *in vitro* binding assays, the *in vivo* binding was confirmed by using the yeast one-hybrid method. The construct containing the ACEI DNA-binding domain fused to the Gal4 activation domain was able to activate the expression of the *HIS3* reporter gene linked to a 170 bp *cbh1* promoter region

-843– -676. This region contained one of the AGGCA sequences binding ACEI *in vitro*.

In order to study the role of the *ace1* gene in the cellulase regulation, the coding region of the gene in the genome of a hyper-cellulolytic T. *reesei* strain ALKO2221 was replaced by the acetamidase marker gene. The deletion strains were grown on cellulose-peptone plates, and a reduction of radial growth of the colonies was obtained. This result, together with the isolation method based on activation of expression, was the reason for us to name the gene as an activator of cellulase expression. Later experiments (Aro *et al.,* 2003), however, have shown that this is not true. In a liquid culture on cellulose-containing medium, the *ace1*-deleted strains produced more cellulases, and the expression of the major cellulase as well as xylanase genes was higher in the deleted strains compared to the host ALKO2221. The growth of the deletion strains on cellulose-containing liquid medium was also faster. Thus, it was concluded that *ace1* codes for a repressor of cellulase and xylanase expression.

Regulation of hydrolase gene expression, however, is probably not the only regulatory role that the *ace1* gene has. The growth of the *ace1*-deleted strain on sorbitol was very much retarded compered to the growth of the host strain (Aro *et al.*, 2003). In addition, Aro studied the expression on sorbitol of genes encoding intracellular reductases in the *ace1* deletion strain (Aro, 2003). The expression of the *xdr1* gene, which codes for a xylose reductase, was slightly increased in the deletion strain. However, the *lxr1* gene, which encodes an L-xylulose reductase capable of using xylitol, D-sorbitol and D-mannitol as a substrate, was lowered in the deletion strain on sorbitol. The latter result could indicate that ACEI functions as an activator of certain genes. Regulatory genes capable of both activating and repressing gene-expression are known from eukaryotes. E.g. in *S. cerevisiae*, the transcription factor Rgt1 acts as a repressor in conditions when no glucose is present. However, in the presence of high amounts of glucose, Rgt1 is phosphorylated, which turns it into an activator (Mosley *et al.*, 2003).

Interestingly, ACEI has been shown to function as a repressor of cellulase genes although it was originally isolated by its ability to bind to the *cbh1* promoter and to activate the expression of the linked marker gene. However, the expression vector isolated from the yeast cells contained a truncated cDNA, which lacked

the N-terminus of ACEI. Therefore, it is quite possible that the N-terminal region contains repressor element(s). Alternatively, there might be (weak) activation domain(s) in the truncated protein, the effect of which is masked in the full-length protein by the repressor element(s). These hypotheses are supported by the fact that the activator function of ACEI could not be repeated with a construct containing the full-length *ace1* gene in the reporter yeast (data not shown). On the other hand, the action of ACEI can be promoter-specific, and depend on modification of the protein, or protein-protein interactions with different activator and repressor proteins.

3.2.2 ACEII (III)

The second transcription factor gene, *ace2,* was found from the same relatively small subset of the yeast expression library in four copies. The gene encodes a protein of 341 aa, starting with an N-terminal zinc binuclear cluster DNA-binding domain. These Zn_2Cys_6 domains have been found solely from fungi. Similarity searches performed lately showed the highest similarity of ACEII to a hypothetical protein of *M. grisea* (identity 30%, similarity 41%). The identical amino acids were clustered to the DNA-binding region and some short stretches of amino acids. It is thus unlikely that the two proteins are functional homologs. The transcription factor AflR, which in different *Aspergillus* species regulates aflatoxin synthesis, showed identities of 23–24%. Otherwise, the similarities detected were to the DNA-binding domain or short stretches of amino acids, either to a stretch where seven histidines alternate with other amino acids (HSHEHSHSHSHNH), or to a glutamine- and proline-rich stretch (QQQEQQQGQPQHPPPP). Histidine-rich sequences are present in many regulatory proteins, e.g. AflR and some homeobox-containing regulators of *Drosophila*, but the functions of them are not known. The glutamine-rich area could represent an activation domain. From the EST databases of fungi, the only meaningful similarities were to three EST sequences of *T. reesei*, which together cover the amino acid area 67–300 of ACEII. Other similarities detected were only to the DNA-binding areas of other transcription factors, or to the short stretches mentioned above.

The binding consensus of ACEII was studied with the same strategy as that used for ACEI. Two of the four longer fragments of the *cbh1* promoter were able to

bind the GST-ACEII zinc binuclear cluster fusion protein *in vitro*. From the one having stronger binding properties, a single sub-fragment of 25 bp was responsible for the specific binding. The activator of xylanase genes, XlnR isolated from *A. niger* (van Peij *et al.*, 1998b), contains a DNA-binding domain belonging to the same class. The 25 bp sub-fragment contained a sequence resembling the binding site consensus shown for XlnR (5´ GGCTAAA, van Peij *et al.,* 1998b). Mutagenesis experiments of that sequence restricted the ACEII binding to 5´ GGCTAATAA, where the underlined GGC triplet is essential for binding, and mutations in the subsequent TAA triplets reduce it. This site is at –779 relative to ATG in the *cbh1* promoter. The binding consensus of XlnR has later been shortened to 5´ GGCTAA (van Peij *et al.,* 1998a), and thus the two factors could bind to the same DNA sequence. The amino acid sequences of the DNA binding domains of the two activators are very dissimilar outside the conserved cysteine residues necessary for the folding of the domain, and outside the binding domains, hardly any similarity can be found. Thus, it is highly unlikely that ACEII and XlnR were *Trichoderma* and *Aspergillus* homologs of one factor.

The *ace2* gene was deleted from the genome of the ALKO 2221 strain by replacing it with the hygromycin marker gene. The deletion strains produced much lower cellulase activities and less cellulase mRNAs after growth on a cellulose-containing medium compared to the host. Thus, ACEII was concluded to represent a true activator of cellulase genes. However, multiple activators are responsible for the strong induction of the cellulase genes since the deletion of *ace2* did result in a reduction of cellulase expression, not in a loss of it. Also, the strong and rapid induction seen on a neutral carbon source after the addition of a small amount of sophorose was shown to be independent of the ACEII activator, since the deletion had no effect on that. It is possible that a homolog of the XlnR of *Aspergilli* regulates cellulase and xylanase expression also in *Trichoderma* species. In *Aspergilli*, XlnR regulates the expression of both xylanase and cellulase genes (van Peij *et al.,* 1998a). Additional activators of cellulase expression may also exist.

It was recently shown that ACEII regulates in addition to cellulase genes, also the expression of the *xyn2* gene encoding one of the endoxylanases of *T. reesei* (Würleitner *et al.,* 2003). Expression of the *xyn2* gene is partly constitutive, but the expression level is further induced by the presence of xylobiose, xylan, and

the cellulose inducers, cellulose and sophorose. A mutation in the binding site of ACEII in the *xyn2* promoter completely abolished both the basal and the induced expression (Würleitner *et al.*, 2003). It was concluded that ACEII acts rather as a general activator than a specific xylanase activator. Thus, both ACEI and ACEII seem to be involved in regulation of both cellulolytic and xylanolytic genes, and additionally in more general regulation as well.

The regulatory genes isolated from *T. reese*i seem to have no homologs outside the filamentous fungal species. No probable homologs were found to ACEII even among them. Thus, ACEII may be important for the exceptionally strong induction of the *T. reesei* cellulase genes on cellulose-containing media, although it cannot be the only activator. The life style of the filamentous fungi is dependent on production of hydrolytic enzymes. The enzymes can be directed towards decaying plant material like in e.g. the brown- or white-rot fungi, or towards living plants in the plant-pathogenic fungi. Similar life style is shared with certain bacteria but the long evolutionary distance between prokaryotes and eukaryotes must prevent the occurrence of homologous regulatory systems. Among eukaryotes, filamentous fungal life style is unique. Yeasts usually use mono- and disaccharides as their carbon sources, and often take the disaccharides into the cells. Sucrose is degraded in *S. cerevisiae* outside the cells by a secreted enzyme, invertase. However, the degradation of sugar polymers is not common among yeasts.

3.3 Screening for xylose permeases from *T. reesei* (IV)

The third cloning method applied in this work was based on complementation. A yeast strain deleted for the major hexose transporter-encoding genes, *hxt1-7,* and the galactose transporter gene, *gal2,* was used. This strain is unable to grow on any other fermentable carbon sources except maltose because it lacks the transporter proteins needed for the uptake of hexose sugars. Thus, although it contains the genes *hxt8-17,* no growth on e.g. glucose or galactose can be observed. The xylose-utilisation pathway, consisting of the genes encoding the xylose reductase and xylitol dehydrogenase enzymes was additionally integrated into the genome of this strain, resulting in the strain H1687. The same transporters (Hxt1p-7p, Gal2p) are also responsible for the uptake of the pentose

sugar xylose, although it is not metabolised by the natural *S. cerevisiae* strains, and thus H1687 is unable to grow on xylose medium.

In order to isolate xylose permease-encoding genes from *T. reesei*, the yeast expression cDNA library of a million clones (Margolles-Clark *et al.,* 1996b) was screened for growth of the H1687 strain on xylose-containing plates. After unsuccessful screening trials on plates containing 3–5% xylose with or without another substrate, maltose, the screen on 0.25% xylose-0.001% maltose plates finally gave few growing clones after 2 weeks of culture. After re-streaking of the clones on 0.25% xylose-0.001% maltose plates, only a single clone was able to grow. Plasmid was rescued from this strain and sequenced, and a gene, which showed similarity to sugar permeases, was detected. The highest amino acid identities, 45–48% were detected to high-affinity glucose permeases of *K. lactis, Candida albicans, Trichoderma harzianum,* and *A. nidulans*. The multiple alignment of the Xlt1 protein with the three most similar proteins is shown in Figure 6.

The strain containing the permease expression vector grew rather well on xylose plates but not at all on glucose. The new gene was named *xlt1* (xylose transporter 1). However, after retransformation of the plasmid to the same yeast strain or the strain H2219, integrated with an over-expression construct of the native xylulokinase-encoding gene as well, the growth of the strains could not be repeated. The 5´ non-coding region of the gene was trimmed in order to remove extra nucleotides between the TATA box of the *PGK1* promoter and the initiator ATG of the gene. Such stretches can be harmful for the expression of heterologous genes in yeast. The resulting plasmid was transformed into the H2219 strain. Unfortunately, the trimmed gene did not support growth on xylose plates any better.

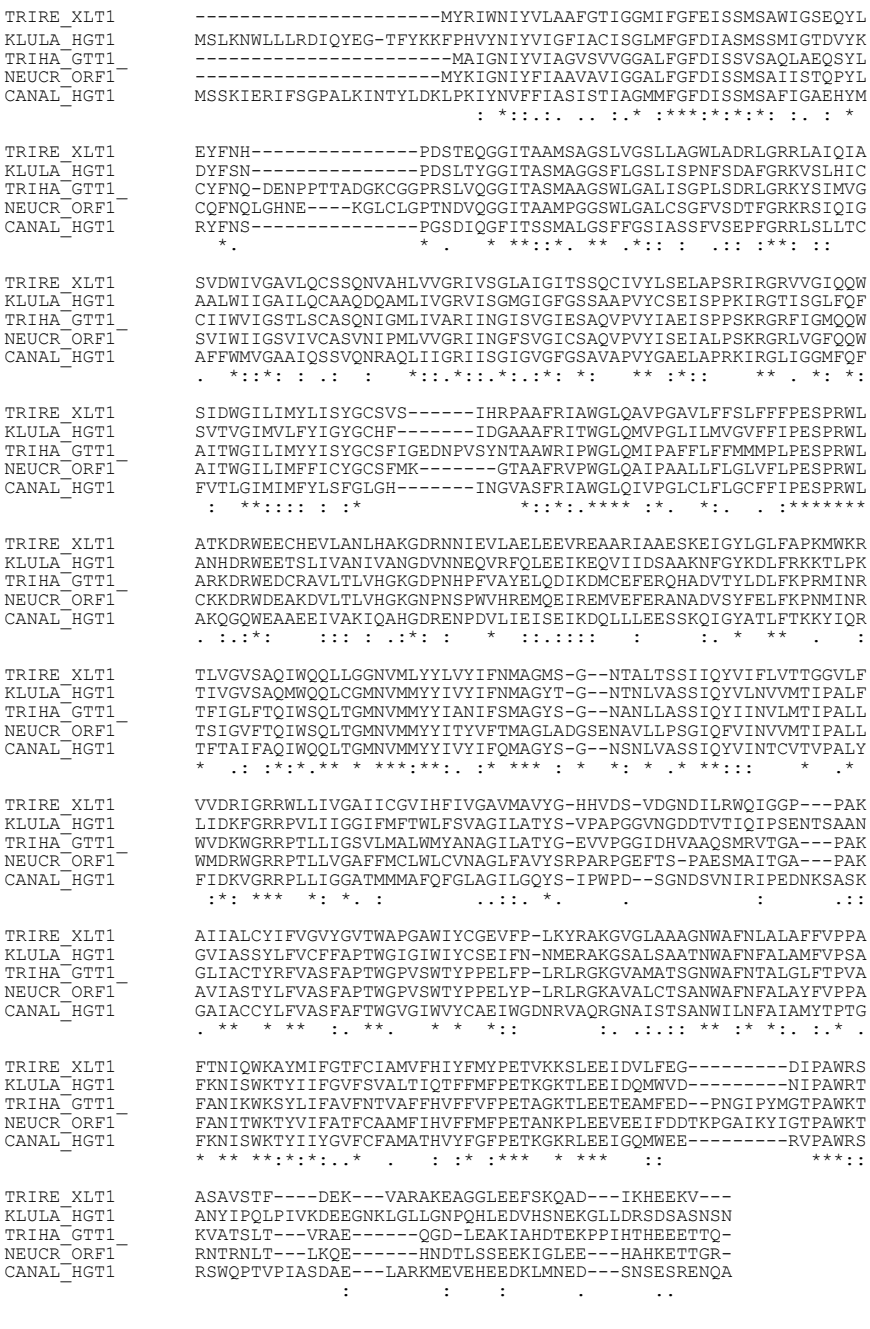


Figure 6. Multiple alignment of the T. reesei (TRIRE), K. lactis (KLULA), T. harzianum (TRIHA) and C. albicans (CANAL) permeases. Identical amino acids in all four sequences are marked with asterisks, similar with dots.

However, a long liquid culture on 2% xylose-0.03% maltose medium resulted in slow growth of the strains with the trimmed *xlt1* gene, as well as some of the strains with the original construct after a lag of two weeks. It appears that the liquid culture of the *xlt1*-expressing strains induced mutations, resulting in adaptive mutation(s) in the host. When the growing strains were cultured again, growth occurred earlier and also on a medium containing 2% xylose as the sole carbon source. The permease sequences of the plasmids from the growing, "adapted" strains were sequenced but no mutations were found. One of the "adapted" strains was cured from the permease plasmid, followed by re-transformation with the *xlt1* expression plasmid and the empty vector. The *xlt1* transformants of the "adapted" strain were able to grow on xylose but the strains transformed with the vector alone could not.

It is likely that the strain has changed permanently. Mere phenotypic adaptation to xylose consumption is not likely since the strains show the altered phenotype also after growth on maltose medium and after the re-transformation procedure. It is probable that multiple mutations had occurred since transformants of the cured strains show variable growth properties on xylose medium. Also the original adapted strains showed variation in their xylose growth. The cured strains were re-transformed also with a hexose transporter gene *hxt1* that had been isolated from *T. reesei* earlier (Ruohonen, L., Walsh, M. C., Margolles-Clark, E., Diderich, J. A., Aristidou, A., Heikkinen, A., van Dam, K., and Penttilä, M., manuscript in preparation). These transformants of the adapted strain grew also faster on xylose media than the same yeast strain without the adaptation (data not shown). Furthermore, after re-transformation with the *xlt1* or *hxt1* constructs, the transformants of one of the two cured strains grew repeatedly faster than the transformants of the other strain. Different mutations could have e.g. different stabilities in the later strain manipulations. It is also probable that the original permease clone picked from the library after two weeks of growth on screening plates and subsequent culturing on xylose plates and liquid medium represents a similarly mutated strain. Plasmid taken from that strain and transformed again to the same original host strain did not support growth on xylose on plates, while the original strain picked from the library did grow.

These mutations could help the heterologous permease e.g. in folding or membrane insertion, and thus make the strain capable of growing on xylose.

Auxiliary factors involved in transport are also known. E.g. the Tpr1 protein of *S. pombe* has been proposed to function as an auxiliary protein interacting with and activating K^+ ion carriers (Lichtenberg *et al.*, 1999). Similarly, the *CSF1* gene of *S. cerevisiae* is needed for nutrient uptake at low temperatures, and was suspected to be a regulator of certain nutrient transporters (Tokai *et al.*, 2000). The mutations occurring during the long liquid culture on xylose-maltose medium could have activated such auxiliary factors.

Alternatively, it is possible that the mutations activated one of the normally silent *HXT* genes present in the hosts. However, the presence of the *Trichoderma* gene was needed for the growth obtained. An explanation for this could be formation of heteromers between the heterologous and homologous permeases active in xylose transport *in vivo*. The formation of heteromers has been suspected between the yeast Hxt proteins (Sherwood *et al.,* 2000). An inactive chimera (Hxt1/4p), where the C-terminus of the Hxt1 protein had been replaced with that of Hxt4, led to a more severe phenotype on high glucose concentration than the deletion of the *HXT1* gene. The strain used in the study lacked the normal *HXT1* gene. The reason for the severe phenotype was concluded to be the formation of inactive heteromultimers with the other Hxt proteins, most likely with the other low-affinity transporter Hxt3p (Sherwood *et al.*, 2000), because these low-affinity transporters are the only hexose transporters that are normally expressed on the high-glucose conditions used (Özcan and Johnston, 1999).

It is also possible that the helper mutation(s) have nothing to do with xylose uptake *per se*. The mutation(s) could also help the subsequent metabolism of xylose in a way that even low concentration of xylose inside the cells supported growth. E.g. lowering of the high K_m of the xylose reductase for xylose could have this effect. However, the *T. reesei* permease gene was needed for the growth to occur. Preliminary xylose uptake studies were performed to the adapted and re-transformed strains. It seems that the velocity of xylose uptake has not been improved in the adaptor mutant (data not shown).

Efficiently xylose-utilising isolates have been recovered in chemostat cultures of a strain containing the XR, XDH and XK over-production constructs integrated, and additionally an empty control vector, on a medium containing a mixture of xylose and glucose (Pitkänen *et al*., submitted). The evolution led to

improvements in growth and uptake of xylose. Adaptive evolution seems to occur readily in xylose-utilising *S. cerevisiae* strains in chemostat cultures (Sonderegger and Sauer, 2003). Maybe the prolonged cultures in this work of the *hxt1-7 gal2* strain containing XR/XDH without or with XK and additionally, the *xlt1* production plasmid on xylose-maltose mixture in shake flasks somehow resembled the chemostat conditions.

3.4 Expression of a xylose-proton symporter of *Bacillus* in yeast

The xylose-proton symporter of *Bacillus megaterium* (Schmiedel *et al.*, 1997) was cloned by PCR into the expression vector YEplac195PGKPT under the *PGK1* promoter. The construct was transformed into the xylose-utilising transporter mutant yeast strain H2219. However, the resulting yeast strains were unable to grow during the 4-week culture in shake flasks on SC-Ura medium supplemented with 2% xylose with or without the addition of 0.03% maltose (data not shown)

Expression of membrane proteins in functional form in heterologous hosts is demanding. Besides mere protein synthesis and folding, the protein has to be directed into the right compartment and assembled correctly into the membrane. However, there are a few successful cases where heterologous sugar permeases have been expressed in functional forms in yeasts. The properties of the MST1 proton-monosaccharide symporter of *Nicotiana tabacum* have been studied after heterologous expression in *S. cerevisiae* (Sauer and Stadler, 1993). Moreover, protein engineering of the *Chlorella* hexose-proton symporters has been achieved in *Schizosaccharomyces pombe* (Will *et al.*, 1998).

Plant (Frommer *et al.,* 1993) and mammalian (Lin *et al.,* 1998) amino acid transporters have been cloned by complementation of yeast strains having defects in the transport of proline. Two putative amino acid permeases were also isolated from a yeast expression library of barley at VTT (Kuusinen, A., Saloheimo, A., Penttilä, M., and Sopanen, T., unpublished). Additionally, an amino acid permease has been cloned from *Arabidopsis* by expression in yeast followed by autoradiographic detection of radio-labelled proline taken up by the yeast cells (Wittstock *et al.,* 2000). Thus, yeast seems to be a relatively good

host for heterologous permeases. The hexose transporter gene *hxt1* has been isolated from the same yeast expression library of *T. reesei* by complementation of the growth defect of the *hxt1-7 gal2* yeast strain on glucose. It complements the growth defect on glucose and transports glucose in the *in vitro* uptake studies (Ruohonen, L., Walsh, M. C., Margolles-Clark, E., Diderich, J. A., Aristidou, A., Heikkinen, A., van Dam, K., and Penttilä, M., manuscript in preparation). Other negative examples are, however, also known in addition to the *Bacillus* permease of this work. Hamacher et al. (2002) tried to complement the xylose growth defect of a mutant strain similar to ours by heterologous xylose transporters from bacteria (a xylose-proton symporter of *E. coli*) and plants (three hexose-proton symporters capable of xylose transport from *Chlorella* and *Arabidopsis*) with no success.

3.5 Studies on the native hexose transporters of yeast (IV)

Yeast was also used as a tool in characterisation of some native hexose transporters. H2219, the hexose transporter-deleted yeast strain (*hxt1-7, gal2*), which expresses the XR, XDH and XK-encoding genes was used as the host. Selected hexose transporter genes of *S. cerevisiae* were individually over-expressed in the H2219 strain. Transporters with different glucose uptake kinetics were chosen: Hxt1p for low, Hxt4p for intermediate, and Hxt7p for high affinity (Reifenberger *et al.,* 1997). Additionally, Hxt2p was chosen because it was more highly expressed during growth of the xylose-utilising yeast strain on xylose compared to glucose (Salusjärvi *et al.,* submitted). The four transporters were expressed separately and constitutively from a multicopy plasmid under the *TPI1* promoter. The xylose-utilising yeast strain H2217, which expresses the native *HXT* genes, was analysed as a reference. The kinetics of xylose uptake, as well as growth properties on xylose-containing media were analysed for the strains.

3.5.1 Xylose uptake kinetics of the single-Hxt strains (IV)

The single-Hxt strains were grown on maltose medium until mid-logarithmic phase and analysed for the uptake of radio-labelled xylose. The H2217 strain

was analysed as a comparison after growth on 2% xylose. Xylose is a relevant carbon source for the xylose-utilising yeast in contrast to maltose used for the single-Hxt over-producing strains. Kinetic values were calculated from analyses made with five or six different xylose concentrations between 25 and 450 or 900 mM, and from data of at least two independent assays (Fig. 3 A-E in IV). In addition, assays with different amounts of competing glucose were performed (Fig. 4A-E in IV).

Generally, the affinities of the permeases for xylose followed the published affinities for glucose (Reifenberger *et al.,* 1997), low affinity for one sugar meaning low affinity for the other, too. However, all affinities were one or two orders of magnitude lower for xylose compered to those for glucose. Velocities of uptake did not differ as dramatically as the affinities, although direct comparison of the values for xylose determined in this work to those for glucose uptake (Reifenberger *et al.*, 1997) is impossible because of different experimental set-ups. In the studies of Reifenberger and co-workers, the single-Hxt strains expressed the transporters from their native chromosomal loci. Only the *HXT1* gene was over-expressed, and the kinetic values for that were calculated by using a strain with two copies of a construct containing the gene under the *ADH1* promoter integrated into the genome. In our study, all the transporters were over-expressed from multi-copy plasmids. Additionally, the velocity of uptake was calculated in respect to the cell dry weight in Reifenberger *et al.*, while correlation to the total protein amount was used in our study. The protein amount of the cells in our study was between 30 and 40% (data not shown). Table 4 shows the comparison of the kinetic values for the two sugars.

Hamacher and co-workers (2002) studied uptake of xylose of single-Hxt strains using one xylose concentration, 50 mM. They used strains, where the *HXT4* and *HXT7* genes were separately expressed under a truncated *HXT7* promoter in a xylose-utilising host strain, where all 22 of the hexose transport-related genes had been deleted. The velocities of xylose uptake of their *HXT4*- and *HXT7*-expressing strains were around two- and five-fold higher compared to ours, respectively (calculatory protein content of 35% of the total CDW used). Our values for the *HXT4*-expressing strain, however, had a large variation between individual experiments (Fig. 3C in IV), and the average value (57 nmol min^{-1} mg total cell $protein^{-1}$) was used in the comparison. Probable cause for the

differences is the very high strength of the truncated *HXT7* promoter on maltose medium.

Recently, Sedlak and Ho (2004) studied intracellular accumulation of xylose and xylitol into cells of single-Hxt strains after growth to the stationary phase and addition of 2% (133 mM) xylose. The efficiency of accumulation by the strains was the following: Hxt7 > Hxt5 > Gal2 > wt > Hxt1 > Hxt4 >>> *hxt*. The results are in contrast to those of ours and partly also to those of Hamacher *et al*. (2002). The reason is probably the different growth phase of the cells (stationary compared to logarithmic) and the significantly longer time period used (120 min compared to 30-60 s).

Table 4. Kinetic values of xylose uptake by the single-Hxt strains after growth to logarithmic phase on maltose medium. The H2217 strain (wt) was analysed after growth on 2% xylose. As a comparison, the values for glucose uptake from Reifenberger et al. (1997) are shown. In the latter study, the strains were first grown on maltose medium until mid-logarithmic phase, followed by culturing on either 0.1% or 2% glucose for several hours. The values for 0.1% glucose are shown. For the Hxt2 and the wt strain (MC996A), the uptake of glucose was biphasic with high and low-affinity components. Both values are thus presented, the values from 2% glucose on the right side of the columns with italics. With the other strains, the differences between the two conditions were insignificant. The xylose uptake was calculated per total protein, while the glucose velocities from Reifenberger et al. (1997) are calculated per total cell dry weight (CDW). The protein content of the cells in our study was 30-40% (data not shown).

	Xylose uptake		**Glucose uptake** from Reifenberger *et al*. (1997)			
	K_m (mM)	**V_{max} (nmol/min/mg prot)**	**K_m (mM)**		**V_{max} (nmol/min/mg CDW)**	
Hxt1	880 ± 8.0	750 ± 94	90 ± 15		690 ± 26	
Hxt2	260 ± 130	340 ± 10	1.5 ± 0.2 60 ± 10	*10 ± 1.0*	97 ± 8.0 450 ± 20	*176 ± 8.0*
Hxt4	170 ± 120	190 ± 23	9.3 ± 0.7		160 ± 8.0	
Hxt7	140 ± 9.3	110 ± 6.8	1.1 ± 0.1		190 ± 7.0	
wt	97 ± 14	280 ± 55	0.8 ± 0.1 21 ± 3.0	*1.7 ± 0.8* *46 ± 7.0*	170 ± 6.0 100 ± 8.0	*15 ± 6.0* *165 ± 6.0*

Hxt7p is clearly a high-affinity, low-capacity transporter in respect to glucose. The same is true for xylose, although the K_m for xylose is over hundred-fold higher compared to that for glucose. Hxt1p is a low-affinity, high-capacity transporter for both the sugars. Hxt2p has for glucose, a high-affinity, low capacity and a low-affinity, high capacity components, when incubated on low glucose. On high glucose concentration, it is an intermediate-affinity transporter similarly to Hxt4p. Thus glucose seems to modulate the affinity of Hxt2p. However, in another paper, also Hxt2p was detected as a medium-affinity permease (K_m 3-5 mM, Maier *et al.,* 2002). For xylose, both Hxt2p and Hxt4p show affinities of intermediate values. However, the K_m values of the different experiments for xylose uptake were highly variable. This is in contrast to the other two transporters giving more constant values.

The inhibition of xylose uptake by the presence of glucose was studied by using a constant xylose concentration (225 or 450 mM) and increasing glucose amounts (8–560 mM). Generally, the uptake of xylose was strongly inhibited by even small amounts of glucose, and almost blocked when the amount of glucose was higher than that of xylose. Hxt1p differed significantly from the others. Small amounts of glucose had little effect on the xylose transport, and 40% of the maximal xylose velocity was achieved even with 560 mM glucose.

As can be seen from Table 4, Hxt1p has about an order of magnitude lower affinity for xylose compared to that for glucose. With the other transporters, as well as with the strains containing intact *HXT* genes (marked "wt" in Table 4), the difference is bigger, between 20 and 100 times. This explains why the xylose uptake by Hxt1p is less inhibited by glucose than that of the other transporters. A xylose transporter less inhibited by glucose would be very useful for the industrial applications of the xylose-utilising *S. cerevisiae*. However, the very low affinity for xylose of the *HXT1*-expressing strain is unfortunate.

3.5.2 Growth of the single-Hxt strains on xylose media (IV)

The single-Hxt strains were cultured on liquid media containing different amounts of xylose either alone or supplemented with minor amounts of maltose or glucose. The growth of the strains follows a similar growth mode on all of the media: The initial, slow growth was followed by a lag of several days, and

finally, by exponential growth after 10–15 days from the start of the cultures. The original slow growth was more profound when a co-substrate for growth was included. All bottles of a certain strain started the exponential growth roughly at the same time, whenever they had been inoculated from the same pre-culture. However, when separate inocula were used, the order of the start of the logarithmic growth was different and seemed to be random. The time frame of the event was the same, always 5–10 days. The onset of the exponential growth on xylose thus seems to depend on the pre-culturing period on maltose medium instead of the kinetic properties of the permeases. Differences in the energy status of the cells may be responsible for the differential growth behaviour. The amount of e.g. storage carbohydrates can vary and effect the growth on xylose. Additionally, the lipid composition of the membranes can vary and affect the function of the cell.

On 2% maltose or 2% glucose media, the growth of the single-Hxt strains was roughly equal. Thus, the low-affinity glucose permease Hxt1p (K_m 90 mM, Reifenberger *et al.,* 1997) supports normal growth on 2% glucose but the "high-affinity" xylose permease Hxt7 (K_m around 130 mM) not on 1–5% xylose without a long lag. The single-Hxt1 strain of Reifenberger and co-workers, which contained the native genomic copy, did not support growth on 0.1% glucose. However, the strain containing two copies of the over-expression construct was able to complement the growth defect. The low expression level was thus the main reason for the lack of growth of that strain on low glucose. Probably, the inefficiency of the xylose-metabolising pathway is the reason for the growth mode obtained for the single-Hxt strains on xylose media. The growth of the strain expressing the *T. reesei hxt1* gene is very similar to the single-Hxt strains both on xylose and on glucose, where they all grow normally (data not shown). Thus, a conclusion could be driven that at least the hexose transporter gene of *T. reesei* is functionally expressed in yeast.

4. Conclusions and future perspectives

In this work, the molecular biology of the filamentous fungus *T. reesei* was studied by using the baker's yeast *S. cerevisiae* as a tool. Five new genes were isolated without any previous knowledge, and the resulting yeast strains studied. Cloning of genes is getting an old-fashioned approach in the post-genomic era. The sequencing of new genomes at an exponential speed results in massive amounts of sequence data. Comparison of the new data to existing, experimentally acquired ones helps in annotation of the new genes. The structures of the encoded proteins can be predicted, and the proteins classified into families of related members. However, the similar structure does not always predict similar function. Therefore, the annotation of the new open reading frames to known functions still needs expression of the putative proteins in heterologous hosts for functional characterisation. *S. cerevisiae* as a well-known unicellular eukaryote is a good choice as a host for eukaryotic proteins.

With the cloning methods developed in this work, it is possible to clone genes encoding specific functions; wanted hydrolytic properties or activators of gene expression in contrast to mere binders. In practise, however, the screening method for cellulases did not result in our hands in cloning of the known cellulase genes *cbh2, egl1* or *egl2,* which all are known to hydrolyse β-glucan. The reason for that must reside in the use of a small sub-library of about 50.000 separate clones. Thus, big genes were probably not represented there as full-length copies. The two new hydrolase genes isolated were 1 kb and 1.4 kb in size. The library used in the cloning of the permease gene contained a million separate clones, which is twenty-fold the size of the previous sub-library and thus, significantly more representative.

A new member of the cellulolytic system of *T. reesei* was found, and the gene cloned and characterised. This *egl5* gene codes for a small endoglucanase, which despite of the small size, contains a cellulose-binding domain. In addition, a laminarinase gene was isolated and the activity analysed. Here, the yeast expression was a successful strategy in the cloning of the gene but not in the analysis of the enzyme activity because of the similar enzyme activities secreted by the yeast. However, the use of other hosts, *E. coli* and baculovirus-insect system, helped in characterisation of the protein. Genes encoding β-1,3-1,4-glucanases different from cellulases have not been isolated from *T. reesei* earlier.

Two regulators of the cellulase genes were as well isolated from *T. reesei*. The binding sites of the proteins in the *cbh1* promoter were determined. The functions were studied by deleting the genes individually from the genome of *T. reesei*. ACEI has been shown to act as a repressor of cellulase and xylanase genes, although the isolation method should reveal only positively acting factors. ACEII, on the other hand, is a true activator. However, other factors are participating in the induction since the deletion strain did not show a cellulase-negative phenotype, and the strong induction seen by the addition of sophorose is not affected by it. It remains to be studied, whether an activator corresponding to *XlnR* of *Aspergilli* is functioning in *T. reesei* as well. The revealing of the regulatory circuits affecting the industrially used production promoter *cbh1* may help in further enhancement of protein production by industrial *Trichoderma* strains.

The replacement of fossil fuels with renewable alternatives is an important task. *S. cerevisiae* yeast is already used in the fermentation of dextrose into ethanol. This approach, however, is not largely accepted because the raw materials like corn could be used as food or feed. Neither is it cost-effective if the product is used as a fuel. The use of sugar-rich wastes, however, could be both economic and acceptable. Therefore the development of the pentose-fermenting yeast is important. The uptake of pentose sugars is not effective by *S. cerevisiae*. After the currently ongoing strain improvements, the uptake is likely to become limiting. The rate of xylose fermentation especially from mixtures of xylose and glucose is low. A permease functioning in xylose uptake in the presence of glucose would be needed in addition to the native yeast hexose transporters. The only clone obtained in the screening for xylose permeases from *T. reesei* is a putative high-affinity xylose permease. It does not support growth of the mutant yeast strain on glucose. However, the strain grows on xylose only after adaptive mutation(s) had occurred to it. The native hexose transporters Hxt1p, Hxt2p, Hxt4p, and Hxt7p, were also studied in respect to their xylose-transporting properties. The affinities of them for xylose were all low, and glucose strongly inhibited xylose transport. The only exception in the latter was Hxt1p, which was much less inhibited by glucose. However, it possessed the lowest affinity of all for xylose. The identification of effective, less glucose-inhibited permeases for xylose is a continuing task.

The effective cellulases and hemicellulases of *T. reesei* can be used in the saccharification of both the cellulose and hemicellulose fractions of the waste materials into monosaccharides. The enzymes could be externally added into the yeast fermentation or *T. reesei* could be first cultured in order to degrade the polymers. The enzymes could also be produced by the yeast itself. Thus, both organisms of this study, *T. reesei* and *S. cerevisiae* can be utilised in the sustainable energy production.

References

Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G. and Kell, D. B. (2003). High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nature Biotechnol. **21,** 692–696.

Aristidou, A. and Penttilä, M. (2000). Metabolic engineering applications to renewable resource utilization. Curr. Opin. Biotechnol. **11,** 187–198.

Aro, N. (2003). Characterisation of novel transcription factors ACEI and ACEII involved in regulation of cellulase and xylanase genes in *Trichoderma reesei*. Espoo, Finland. VTT Publications 488. 83 p. + app. 25 p.

Aro, N., Ilmén, M., Saloheimo, A. and Penttilä, M. (2003). ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. Appl. Environ. Microbiol. **69,** 56–65.

Barnett, C. C., Berka, R. M. and Fowler, T. (1991). Cloning and amplification of the gene encoding an extracellular β-glucosidase from *Trichoderma reesei*: evidence for improved rates of saccharification of cellulosic substrates. Bio/Technol. **9,** 562–567.

Beggs, J. D. (1978). Tranformation of yeast by a replicating hybrid plasmid. Nature **275,** 104–109.

Béguin, P., Gilkes, N. R., Kilburn, D. G., Miller, R. C. Jr, O'Neill, G. and Warren, R. A. J. (1987). Cloning of cellulase genes. CRC Critical Rev. Biotechnol. **6,** 129–162.

Beldman, G., Searle-van Leeuwen, M. F., Rombouts, F. M. and Voragen, F. G. J. (1985). The cellulase of *Trichoderma viride*. Purification, characterisation and comparison of all detectable endoglucanases, exoglucanases and β-glucosidases. Eur. J. Biochem. **146,** 301–308.

Berghem, L. E. R., Pettersson, L. G. and Axiö-Fredriksson, U.-B. (1976). The mechanism of enzymatic cellulose degradation. Purification and some properties of two different 1,4-β-glucan glucanohydrolases from *Trichoderma viride*. Eur. J. Biochem. **61,** 621–630.

Bernfeld, P. (1955). Amylases, α and β. In: Colowick, S. P. and Kaplan, N.O. (Eds.). Methods in Enzymology, Academic Press, New York. Pp. 149–158.

Bielecki, S. and Galas, E. (1991). Microbial β-glucanases different from cellulases. Critical Rev. Biotechnol. **10,** 275–304.

Billard, P., Menart, S., Blaisonneau, J., Bolotin-Fukuhara, M., Fukuhara, H. and Wesolowski-Louvel, M. (1996). Glucose uptake in *Kluyveromyces lactis*: role of the *HGT1* gene in glucose transport. J. Bacteriol. **178**, 5860-5866.

Binder, A. and Ghose, T. K. (1978). Adsorption of cellulose by *Trichoderma viride*. Biotechnol. Bioeng. **20,** 1187–1199.

Boles, E. and Hollenberg, C. P. (1997). The molecular genetics of hexose transport in yeasts. FEMS Microbiol. Rev. **21,** 85–111.

Carle-Urioste, J. C., Escobar-Vera, J., El-Gogary, S., Henrique-Silva, F., Torigoi, E., Crivellaro, O., Herrera-Estrella, A. and El-Dorry, H. (1997). Cellulase induction in *Trichoderma reesei* by cellulose requires its own basal expression. J. Biol. Chem. **272,** 10169–10174.

Chan, J. Y., Han, X.-L. and Kan, Y. W. (1993). Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. Proc. Natl. Acad. Sci. USA **90,** 11371–11375.

Chang, Y. C. and Timberlake, W. E. (1992). Identification of *Aspergillus brlA* response elements (BREs) by genetic selection in yeast. Genetics **133,** 29–38.

Chen, C.M., Gritzali, M. and Stafford, D.W. (1987). Nucleotide sequence and deduced primary structure of cellobiohydrolase II from *Trichoderma reesei*. Bio/Technol. **5,** 274–278.

Chien, C.-T., Bartel, P. L., Sternglanz, R. and Fields, S. (1991). The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. USA **88,** 9578–9582.

Covitz, P. A., Herkowitz, I. and Mitchell, A. P. (1991). The yeast RME1 gene encodes a putative zinc finger protein that is directly repressed by a1-alpha 2. Genes Dev. **5,** 1982–1989.

Dalbøge, H. (1997). Expression cloning of fungal enzyme genes; a novel approach for efficient isolation of enzyme genes of industrial relevance. FEMS Microbiol. Rev. **21,** 29–42.

Dalbøge, H. and Heldt-Hansen, H. P. (1994). A novel method for efficient expression cloning of fungal enzyme genes. Mol. Gen. Genet. **243,** 253–260.

Davies, G. J., Dodson, G. G., Hubbard, R. E., Tolley, S. P., Dauter, Z., Wilson, K. S., Hjort, C., Mikkelsen, J. M., Rasmussen, G. and Schülein, M. (1993). Structure and function of endocellulase V. Nature **365,** 362–364.

Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C. and Schülein, M. (1995). Structure of oligosaccharide-bound forms of the endoglucanase V from *Humicola insolens* at 1.9 Å resolution. Biochemistry **34,** 16210–16220.

Deanda, K., Zhang, M., Eddy, C. and Picataggio, S. (1996). Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering. Appl. Environ. Microbiol. **62**, 4465–4470.

de Risi, J. L., Iyer, V. R. and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science **278,** 680–686.

Dien, B. S., Cotta, M. A. and Jeffriws, T. W. (2003). Bacteria engineered for fuel ethanol production: current status. Appl. Microbiol. Biotechnol. **63,** 258–266.

Doan, D. N. P., Høj, P. B., Collins, A., Din, N., Hoogenraad, N. J. and Fincher, G. B. (1993). Post-translational processing of barley β-glucan endohydrolases in the baculovirus-insect cell expression system. DNA Cell Biol. **12,** 97–105.

Does, A. L. and Bisson, L. F. (1989) Characterization of xylose uptake in the yeasts *Pichia heedie* and *Pichia stipitis*. Appl. Environ. Microbiol. **55,** 159–164.

Does, A. L. and Bisson, L. F. (1990) Isolation and characterization of *Pichia heedie* mutants defective in xylose uptake. Appl. Environ. Microbiol. **56,** 3321–3328.

Durand, H., Baron, M., Calmels, T. and Tiraby, G. (1988). Classical and molecular genetics applied to Trichoderma reesei for the selection of improved cellulolytic industrial strains. In: FEMS Symposium no. 43. Aubert, J.-P., Beguin,

P. and Millet, J. (Eds.). Biochemistry and genetics of cellulose degradation, Academic Press, London. Pp. 135–152.

El-Gogary, S., Leite, A., Crivellaro, O., Eveleigh, D. E. and El-Dorry, H. (1989). Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. Proc. Natl. Acad. Sci. USA **86,** 6138–6141.

Eliasson, A., Christensson, C., Wahlbom, C. F. and Hahn-Hägerdal, B. (2000). Anaerobic xylose fermentation by recombinant *Saccharomyces cerevisiae* carrying *XYL1, XYL2*, and *XKS1* in mineral medium chemostat cultures. Appl. Environ. Microbiol. **66,** 3381–3386.

Erbeznik, M., Strobel, H.J., Dawson, K.A. and Jones, C.R., (1998). The D-xylose-binding protein, XylF, from *Thermoanaerobacter ethanolicus* 39E: cloning, molecular analysis, and expression of the structural gene. J. Bacteriol. **180,** 3570–7.

Feng, S.-Y., Ota, K., Yamada, Y., Sawabu, N. and Ito, T. (2004). A yeast one-hybrid system to detect methylation-dependent DNA-protein interactions. Biochem. Biophys. Res. Comm. **313,** 922–925.

Fields, S. and Song, O.-K. (1989). A novel genetic system to detect protein-protein interactions. Nature **340,** 245–246.

Foreman, P. K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N. S., Goedegebuur, F., Houfek, T. D., England, G. J., Kelley, A. S., Meerman, H. J., Mitchell, T., Mitchinson, C., Olivares, H. A., Teunissen, P. J. M., Yao, J. and Ward, M. (2003). Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J. Biol. Chem. **278,** 31988–31997.

Förster, J., Famili, I., Fu, P., Palsson, B. Ø. and Nielssen, J. (2003). Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res. **13,** 244–253.

Frommer, W. B., Hummel, S. and Riesmeier, J. W. (1993). Expression cloning in yeast of a cDNA encoding a broad specificity amino acid permease from *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **90,** 5944–5948.

Gárdonyi, M., Jeppsson, M., Lidén, G., Gorwa-Grauslund, M. F. and Hahn-Hägerdal, B. (2002). Control of xylose consumption by xylose transport in recombinant *Saccharomyces cerevisiae*. Biotechnol. Bioeng. **82,** 818–824.

Gárdonyi, M., Österberg, M., Rodrigues, C., Spencer-Martins, I. and Hahn-Hägerdal, B. (2003). High capacity xylose transport in *Candida intermedia* PYCC 4715. FEMS Yeast Research **3**, 45–52.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O´Shea, E. K. and Weissman, J. S. (2003). Global analysis of protein expression in yeast. Nature **425,** 737–741.

Giaever, G., Chu, A. M. *et al.,* (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. Nature **418,** 387–391.

Gielkens, M. M. C., Dekkers, E., Visser, J. and de Graaff, L. H. (1999). Two cellobiohydrolase-encoding genes from *Aspergillus niger* require D-xylose and the xylanolytic transcriptional activator XlnR for their expression. Appl. Environ. Microbiol. **65,** 4340–4345.

Gietz, R. D. and Sugino, A. (1988). New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenised yeast genes lacking six-base pair restriction sites. Gene **74**, 527–534.

Gilbert, H. J., Hall, J., Hazlewood, G. P. and Ferreira, L. M. A. (1990). The *N*-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies *cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. Mol. Microbiol. **4,** 759–767.

Goffeau, A., Barrel, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S. G. (1996). Life with 6000 genes. Science **274,** 546, 563–567.

Goldman, G. H., Demolder, J., Dewaele, S., Herrera-Estrella, A., Geremia, R. A., Van Montagu, M. and Contreras, R. (1992). Molecular cloning of the imidazoleglycerolphosphate dehydratase gene of *Trichoderma harzianum* by

genetic complementation in *Saccharomyces cerevisiae* using direct expression vector. Mol. Gen. Genet. **234,** 481–488.

Gonçalves, P., de Sousa, H. R. and Spencer-Martins, I. (2000). *FSY1*, a novel gene encoding a specific fructose/H^+ symporter in the type strain of *Saccharomyces carlsbergensis.* J. Bacteriol. **182,** 5628–5630.

Gong, C.-S. and Tsao, G. T. (1979). Cellulase and biosynthesis regulation. In: Perlman, D. (Ed.). Annual Reports on Fermentation Processes, vol. 3, Academic Press, New York, pp. 111.

Gonzáles, R., Ramón, D. and Pérez-Gonzáles (1992). Cloning, sequence analysis and yeast expression of the *egl1* gene from *Trichoderma longibrachiatum*. Appl. Microbiol. Biotechnol. **38,** 370–375.

Håkansson, U., Fägerstam, L., Pettersson, G. and Andersson, L. (1978). Purification and characterization of a low molecular weight 1,4-ß-glucan glucanohydrolase from the cellulolytic fungus *Trichoderma viride* QM 9414. Biochim. Biophys. Acta **524,** 385–392.

Hamacher, T., Becker, J., Gárdonyi, M., Hahn-Hägerdal, B. and Boles, E. (2002). Characterisation of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilisation. Microbiol. **148,** 2783–2788.

Heiland, S., Radovanovic, N., Höfer, M., Winderickx, J. and Lichtenberg, H. (2000). Multiple hexose transporters of *Schizosaccharomyces pombe*. J. Bacteriol. **182,** 2153–2162.

Helle, S., Cameron, D., Lam, J., White, B. and Duff, S. (2003). Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. Enzyme Microbial Technol. **33,** 786–792.

Henriksson, K., Teleman, A., Suortti, T., Reinikainen, T., Jaskari, J., Teleman, O. and Poutanen, K. (1995). Hydrolysis of barley (1→3),(1→4)-β-D-glucan by a cellobiohydrolase II preparation from *Trichoderma reesei*. Carbohydr. Polymers **26,** 109–119.

Henrique-Silva, F., El-Gogary, S., Carle-Urioste, J. C., Matheucci, E. Jr., Crivellaro, O. and El-Dorry, H. (1996). Two regulatory regions controlling basal and cellulose-induced expression of the gene encoding cellobiohydrolase I of *Trichoderma reesei* are adjacent to its TATA box. Biochem. Biophys. Res. Comm. **228,** 229–237.

Henrissat, B. and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **293,** 781–788.

Henrissat, B., Teeri, T. T. and Warren, R. A. J. (1998). A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. FEBS Lett. **425,** 352–354.

Ho, N. W. Y., Chen, Z. and Brainard, A. P. (1998). Genetically engineered *Saccharomyces* yeast capable of efficient cofermentation of glucose and xylose. Appl. Environ. Microbiol. **64,** 1852–1859.

Huh, W.-K., Falvo, J. V., Carrol, A. S., Howson, R. W., Weissman, J. S. and O´Shea, E. K. (2003). Global analysis of protein localization in budding yeast. Nature **425,** 686–691.

Ilmén, M. (1997). Molecular mechanisms of glucose repression in the filamentous fungus Trichoderma reesei. Espoo, Finland. VTT Publications 315. 86 p. + app. 48 p.

Ilmén, M., Onnela, M.-L., Klemsdal, S., Keränen, S. and Penttilä, M. (1996a). Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. Mol. Gen. Genet. **253,** 303–314, and corrigendum in Mol. Gen. Genet. **257,** 386.

Ilmén, M., Thrane, C. and Penttilä, M. (1996b). The glucose repressor gene *cre1* of *Trichoderma*: Isolation and expression of a full-length and a truncated mutant form. Mol. Gen. Genet. **251,** 451–460.

Ilmén, M., Saloheimo, A., Onnela, M.-L. and Penttilä, M. E. (1997). Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. Appl. Environ. Microbiol. **63,** 1298–1306.

Ingram, L. O., Conway, T., Clark, D. P., Sewell, G. W. and Preston, J. F. (1987). Genetic engineering of ethanol production in *Escherichia coli*. Appl. Environ. Microbiol. **53**, 2420–2425.

Inouye, C., Remondelli, P., Karin, M. and Elledge, S. (1994). Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: the one hybrid system. DNA Cell Biol. **13,** 731–742.

Karlsson, J., Siika-aho, M., Tenkanen, M. and Tjerneld, F. (2002). Enzymatic properties of the low molecular mass endoglucanases Cel12A (EGIII) and Cel45A (EGV) of *Trichoderma reesei*. J. Biotechnol. **99,** 63–78.

Koch, A., Weigel, C. T. O. and Schulz, G. (1993). Cloning, sequencing, and heterologous expression of a cellulase-encoding cDNA (*cbh1*) from *Penicillium janthinellum*. Gene **124,** 57–65.

Kötter, P. and Ciriacy, M. (1993). Xylose fermentation by *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. **38,** 776–783.

Kötter, P., Amore, R., Hollenberg, C. P. and Ciriacy, M. (1990). Isolation and characterisation of the Pichia stipitis xylitol dehydrogenase gene, *XYL2*, and construction of a xylose-utilising *Saccharomyces cerevisiae* transformant. Curr. Genet. **18,** 493–500.

Kotyk, A. (1967). Mobility of the free and the loaded monosaccharide carrier in *Saccharomyces cerevisiae*. Biochim. Biophys. Acta **135,** 112–119.

Kraulis, P.J., Clore, G.M., Nilges, M., Jones, T.A., Pettersson, G., Knowles, J. and Gronenborn, A.M. (1989). Determination of the three-dimensional structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. Biochemistry **28,** 7241–7257.

Kuhls, K., Lieckfeldt, E., Samuels, G. J., Kovacs, W., Meyer, W., Petrini, O., Gams, W., Börner, T. and Kubicek, C. P. (1996). Molecular evidence that the asexual industrial fungus Trichoderma reesei is a clonal derivative of the ascomycete Hypocrea jecorina. Proc. Natl. Acad. Sci. USA **93,** 7755–7760.

Kuyper, M., Harhangi, H, R., Stave, A. K., Winkler, A. A., Jetten, M. S. M., de Laat, W. T. A. M., den Ridder, J. J. J., Op den Camp, H. J. M., van Dijken, J. P. and Pronk, J. T. (2003). High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? FEMS Yeast Res. **4,** 69–78.

Larsson, S., Cassland, P. and Jönsson, L. J. (2001). Development of a *Saccharomyces cerevisiae* strain with enhanced resistance to phenolic fermentation inhibitors in lignocellulose hydrolysates by heterologous expression of laccase. Appl. Environ. Microbiol. **67,** 1163.1170.

Lee, W.-J., Kim, M.-D., Ryu, Y.-W., Bisson, L. F. and Seo, J.-H. (2002). Kinetic studies on glucose and xylose transport in *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. **60,** 186–191.

Lichtenberg, H., Heyer, M. and Höfer, M. (1999). Tpr1, a *Schizosaccharomyces pombe* protein involved in potassium transport. FEBS Lett. **457,** 363–368.

Lin, G., Cai, X. and Johnstone, R. M. (1998). Expression cloning of a mammalian amino acid transporter or modifier by complementation of a yeast transport mutant. J. Cellular Physiol. **173,** 351–360.

Lönn, A., Träff-Bjerre, K. L., Cordero Otero, R. R., van Zyl, W. H. and Hahn-Hägerdal, B. (2003). Xylose isomerase activity influences xylose fermentation with recombinant *Saccharomyces cerevisiae* strains expressing mutated *xylA* from *Thermus thermophilus*. Enzyme Microbial. Technol. **32,** 567–573.

Lucas, C. and van Uden, N. (1986). Transport of hemicellulose monomeres in the xylose-fermenting yeast *Candida shehatae*. Appl. Microbiol. Biotechnol. **23,** 491–495.

Maier, A., Völker, B., Boles, E. and Fuhrmann, G. F. (2002). Characterisation of glucose transport in *Saccharomyces cerevisiae* with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. FEMS Yeast Res. **2,** 539–550.

Mandels, M. and Reese, E. T. (1960). Induction of cellulase in fungi by cellobiose. J. Bacteriol. **79,** 816–826.

Mäntylä, A. L., Rossi, K. H., Vanhanen, S. A., Penttilä, M. E., Suominen, P. L. and Nevalainen, K. M. H. (1992). Electrophoretic caryotyping of wild-type and mutant *Trichoderma longibrachiatum* (*reesei*) strains. Curr. Genet. **21,** 471–477.

Margolles-Clark, E., Tenkanen, M., Luonteri, E. and Penttilä, M (1996a). Three α-galactosidase genes of *Trichoderma reesei* cloned by expression in yeast. Eur. J. Biochem. **240,** 104–111.

Margolles-Clark, E., Tenkanen, M., Nakari-Setälä, T. and Penttilä, M (1996b). Cloning of genes encoding α-L-arabinofuranosidase and β-glucosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. **62,** 3840–3846.

Margolles-Clark, E., Ilmén, M. and Penttilä, M. (1997). Expression patterns of the hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. J. Biotechnol. **56,** 167–179.

Marui, J., Tanaka, A., Mimura, S., de Graaff, L. H., Visser, J., Kitamoto, N., Kato, M., Kobayashi, T. and Tsukagoshi N. (2002). A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. Fungal Genet. Biol**. 35,** 157-69.

Mastick, G. S., McKay, R., Oligino, T., Donovan, K. and López, A. J. (1995). Identification of target genes regulated by homeotic proteins in *Drosophila melanogaster* through genetic selection of *Ultrabithorax* protein-binding sites in yeast. Genetics **139,** 349–363.

McKnight, G. L. and McConaughy, B. L. (1983). Selection of functional cDNAs by complementation in yeast. Proc. Natl. Acad. Sci. USA **80,** 4412–4416.

Mellor J., Dobson, M. J., Roberts, N. A., Tuite, M. F., Emtage, J. S., White, S., Lowe, P. A., Patel, T., Kingsman, A. J. and Kingsman, S.M. (1983). Efficient synthesis of enzymatically active calf chymosin in *Saccharomyces cerevisiae*. Gene **24**, 1–14.

Mosley, A. L., Lakshmanan, J., Aryal, B. K. and Özcan, S. (2003). Glucose-mediated phosphorylation converts the transcription factor Rgt1 from a repressor to an activator. J. Biol. Chem. **278,** 10322–10327.

Nevalainen, H. and Penttilä, M. (1995). Molecular biology of cellulolytic fungi. In: Kück, U. (Ed.). The Mycota II, Genetics and Biotechnology, Springer-Verlag, Berlin Heidelberg. Pp. 303–319.

Nobre, A., Lucas, C. and Leão, C. (1999). Transport and utilization of hexoses and pentoses in the halotolerant yeast *Debaryomyces hansenii*. Appl. Environ. Microbiol. **65**, 3594–3598.

Ooi, T., Shinmyo, A., Okada, H., Hara, S., Ikenaka, T., Murao, S. and Arai, M. (1990). Cloning and sequence analysis of a cDNA for cellulase (FI-CMCase) from *Aspergillus aculeatus*. Curr. Genet. **18,** 217–222.

Osborne, M. A., Dalton, S. and Kochan, J P. (1995). The yeast tribrid system-Genetic detection of *trans*-phosphorylated ITAM-SH2-interactions. Bio/Technol. **13,** 1474–1478.

Özcan, S. and Johnston, M. (1999). Function and regulation of yeast hexose transporters. Microbiol. Mol. Biol. Rev. **63,** 554–569.

Pauly, M., Andersen, L. N., Kauppinen, S., Kofod, L. V., York, W. S., Albersheim, P. and Darvill, P. (1999). A xyloglucan-specific endo-beta-1,4-glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification and characterisation of the recombinant enzyme. Glycobiology **9,** 93–100.

Penttilä, M. (1987). Construction and characterization of cellulolytic yeasts. Espoo, Finland. VTT Publications 39. 53 p. + app. 63 p.

Penttilä, M. E., Nevalainen, K. M. H., Raynal, A. and Knowles, J. K. C. (1984). Cloning of *Aspergillus niger* genes in yeast. Expression of the gene coding *Aspergillus* β-glucosidase. Mol. Gen. Genet. **194,** 494–499.

Penttilä, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R. and Knowles, J. (1986). Homology between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. Gene **45**: 253–63.

Penttilä, M.E., André, L., Saloheimo, M., Lehtovaara, P. and Knowles, J.K.C. (1987). Expression of two *Trichoderma reesei* endoglucanases in the yeast *Saccharomyces cerevisiae*. Yeast **3,** 175–185.

Penttilä, M.E., André, L., Lehtovaara, P., Bailey, M., Teeri, T.T. and Knowles, J.K.C. (1988). Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. Gene **63,** 103–112.

Pitkänen J.-P., Rintala, E., Aristidou, A., Maaheimo, H., Ruohonen, L. and Penttilä, M. (2004). Selective evolution of *Saccharomyces cerevisiae* for improved xylose metabolism: Isolation in chemostat cultures. Submitted.

Pretorius, I. S. (1997). Utilisation of polysaccharides by *Saccharomyces cerevisiae,* In: Zimmermann, F. K. and Entian, K. D. (Eds.) Yeast sugar metabolism. Technomic Publishing Company, Lancaster. Pp. 435–458.

Reifenberger, E., Boles, E. and Ciriacy, M. (1997). Kinetic characterisation of individual hexose transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanism of glucose repression. Eur. J. Biochem. **245,** 324–333.

Remacle, J. E., Brys, R., Pype, S., Nelles, L. and Huylebroeck, D. (1998). A novel expression cloning method to isolate mammalian transcription factors in *Schizosaccharomyces pombe*. Nucleic Acids Res. **26,** 5223–5224.

Rothstein, R. (1983). One step gene disruption in yeast. Methods Enzymol. **101,** 202–211.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C. and Jones, T.A. (1990). Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. Science **249,** 380–386.

Saloheimo, A. and Penttilä, M. (1993). A new hydrolase gene from *Trichoderma reesei.* In: Suominen, P., and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research, Helsinki. Vol 8. Pp. 291.

Saloheimo, M., Lehtovaara, P., Penttilä, M., Teeri, T.T., Ståhlberg, J., Johansson, G., Pettersson, G., Claeyssens, M., Tomme, P. and Knowles, J.K.C. (1988). EGIII, a new endoglucanase from *Trichoderma reesei*: the characterization of both gene and enzyme. Gene **63,** 11–21.

Saloheimo, A., Henrissat, B. and Penttilä, M. (1993). Small endoglucanase from *Trichoderma reesei*, cloned by expression in yeast. In: Suominen, P. and

Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research, Helsinki. Vol. 8. Pp. 139–146.

Saloheimo, M., Nakari-Setälä, T., Tenkanen, M. and Penttilä, M. (1997). A new *Trichoderma reesei* cellulase: cDNA cloning and demonstration of endoglucanase activity by yeast expression. Eur. J. Biochem. **249,** 584–591.

Saloheimo, M., Kuja-Panula, J., Ylösmäki, E., Ward, M. and Penttilä, M. (2002a). Enzymatic properties and intracellular localization of the novel *Trichoderma reesei* β-glucosidase BGLII (Cel1A). Appl. Environ. Microbiol. **68,** 4546–4553.

Saloheimo, M., Paloheimo, M., Hakola, S., Pere, J., Swanson, B., Nyyssönen, E., Bhatia, A., Ward, M. and Penttilä, M. (2002b). Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expancins, exhibits disruption activity on cellulosic materials. Eur. J. Biochem. **269,** 4202–4211.

Salusjärvi, L., Pitkänen, J.-P., Aristidou, A., Ruohonen, L. and Penttilä, M. (2004). Gene expression analysis of recombinant xylose-fermenting *Saccharomyces cerevisiae* reveals novel responses to xylose as a carbon source. Submitted.

Sambrook, J. and Russell, D. W. (2001). Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 3rd edn.

Sauer, N. and Stadler, R. (1993). A sink-specific H+ /monosaccharide co-transporter from *Nicotiana tabacum*: cloning and heterologous expression in baker´s yeast. Plant J. **4,** 601–610.

Schmiedel, D., Kintrup, M., Küster, E. and Hillen, W. (1997) Regulation of expression, genetic organization and substrate specificity of xylose uptake in *Bacillus megaterium*. Mol. Microbiol. **23,** 1053–1062.

Sedlak, M. and Ho, N. W. Y. (2004). Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinant *Saccharomyces* yeast. Yeast **21**, 671–684.

Seiboth, B., Hakola, S., Mach, R. L., Suominen, P. L. and Kubicek, C. P. (1997). Role of four major cellulases in triggering cellulase gene expression by cellulose in *Trichoderma reesei*. J. Bacteriol. **179**, 5318–5320.

Sheppard, P. O., Grant, F. J., Oort, P. J., Sprecher, C. A., Foster, D. C., Hagen, F. S., Upshall, A., McKnight, G. L. and O'Hara, P. J. (1994) The use of conserved cellulase family-specific sequences to clone cellulase homologue cDNAs from *Fusarium oxysporum*. Gene **150,** 163–167.

Sherwood, P. W., Katic, I., Sanz, P. and Carlson, M. (2000). A glucose transporter chimera confers a dominant negative glucose starvation phenotype in *Saccharomyces cerevisiae*. Genetics **155,** 989–992.

Shoemaker, S., Schweickart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K. and Innis, M. (1983). Molecular cloning of exo-cellobiohydrolase I derived from *Trichoderma reesei* strain L27. Bio/Technol. **1,** 691–695.

Silva, J., Saloheimo, M., Montanez, C., Knowles, J. K. C. and Teeri, T. T. (1988). Molecular cloning of cellulolytic enzyme genes from *Cellulomonas flavigena* in *E. coli*. J. Biotechnol. **8,** 249–256.

Sims, P., James, C. and Broda, P. (1988). The identification, molecular cloning and characterisation of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exo-cellobiohydrolase I gene from *Trichoderma reesei*. Gene **74,** 411–422.

Smith, T. A. and Kohorn, B. D. (1991). Direct selection for sequences encoding proteases of known specificity. Proc. Natl. Acad. Sci. USA **88,** 5159–5162.

Sonderegger, M. and Sauer, U. (2003). Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. Appl. Environ. Microbiol. **69,** 1990–1998.

Ståhlberg, J. (1991). Functional organization of cellulases from Trichoderma reesei. PhD Thesis, University of Uppsala, Sweden. 44 pp.

Stålbrand, H., Saloheimo, A., Vehmaanperä, J., Henrissat, B. and Penttilä, M. (1995). Cloning and expression in yeast of *Trichoderma reesei* β-mannanase

containing a cellulose binding domain. Appl. Environ. Microbiol. **61,** 1090–1097.

Stambuk, B. U., Franden, M. A., Singh, A. and Zhang, M. (2003). D-Xylose transport by *Candida succiphila* and *Kluyveromyces marxianus*. Appl. Biochem. Biotechnol. **105-108**, 255–263.

Strauss, J., Mach, R. L., Zeilinger, S., Hartler, G., Stöffler, G., Wolschek, M. and Kubicek, C. P. (1995). Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*. FEBS Lett. **376,** 103–107.

Takashima, S., Iikura, H., Nakamura, A., Masaki, H. and Uozumi, T. (1996a). Analysis of Cre1 binding sites in the *Trichoderma reesei cbh1* upstream region. FEMS Microbiol. Lett. **145,** 361–366.

Takashima, S., Nakamura, H., Iikura, H., Masaki, H. and Uozumi, T. (1996b). Cloning of a gene encoding a putative carbon catabolite repressor from *Trichoderma reesei.* Biosci. Biotech. Biochem. **60,** 173–176.

Takekawa, M., Posas, F. and Saito, H. (1997). A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinase, MTK1, mediates stress-induced activation of the p38 and JNK pathways. EMBO J. **16,** 4973–4982.

Teeri, T. T. (1987). The cellulolytic enzyme system of *Trichoderma reesei*. Molecular cloning, characterization and expression of the cellobiohydrolase genes. Espoo, Finland. VTT Publications 38. 52 p. + app. 41 p.

Teeri, T., Salovuori, I. and Knowles, J. (1983). The molecular cloning of the major cellulase gene from *Trichoderma reesei*. Bio/Technol. **1,** 696–699.

Teeri, T. T., Lehtovaara, P., Kauppinen, S., Salovuori, I., and Knowles, J. (1987). Homologous domains in *Trichoderma reesei* cellulolytic enzymes: gene sequence and expression of cellobiohydrolase II. Gene **51,** 43–52.

Tempelaars, C. A. M., Birch, P. R. J., Sims, P. F. G. and Broda, P. (1994). Isolation, characterization, and analysis of the expression of the *cbhII* gene of *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. **60,** 4387–4393.

Toivari, M. H., Aristidou, A., Ruohonen, L. and Penttilä, M. (2001). Conversion of xylose to ethanol by recombinant *Saccharomyces cerevisiae*: Importance of xylulokinase (*XKS1*) and oxygen availability. Metabol. Eng. **3,** 236–249.

Tokai, M., Kawasaki, H., Kikuchi, Y. and Ouchi, K. (2000). Cloning and characterisation of the *CSF1* gene of *Saccharomyces cerevisiae*, which is required for nutrient uptake at low temperature. J. Bacteriol. **182,** 2865–2868.

Uetz, P., Giot, L., *et al.,* (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature **403,** 623–627.

van Arsdell, J.N., Kwok, S., Schweickart, V.L., Ladner, M.B., Gelfand, D.H. and Innis, M.A. (1987). Cloning, characterization, and expression in *Saccharomyces cerevisiae* of endoglucanase I from *Trichoderma reesei*. Bio/Technol. **5,** 60–64.

van Peij, N. N. M. E., Gielkens, M. M. C., de Vries, R. P., Visser, J. and de Graaff, L. H. (1998a). The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl. Environ, Microbiol. **64,** 3615–3619.

van Peij, N. N. M. E., Visser, J. and de Graaff, L. H. (1998b). Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol. Microbiol. **27,** 131–142.

van Rensburg, P., van Zyl, W. H. and Pretorius, I. S. (1998). Engineering yeast for efficient cellulose degradation. Yeast **14,** 67–76.

Volpers, C., Lubinus, M., Osborne, M. A. and Kochan, J. P. (2001). cDNA expression cloning and characterization of phosphorylation dependent protein interactors using the yeast tribrid system. Methods Mol. Biol. **124,** 271–293.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S. and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature **417,** 399–403.

von Ossowski, I., Teeri, T., Kalkkinen, N. and Oker-Blom, C. (1997). Expression of a fungal cellobiohydrolase in insect cells. Biochem. Biophys. Res. Comm. **233,** 25–29.

Wahlbom, C. F., van Zyl, W. H., Jönsson, L. J., Hahn-Hägerdal, B. and Cordero Otero, R. R. (2003). Generation of the improved recombinant xylose-utilising *Saccharomyces cerevisiae* TMB 3400 by random mutagenesis and physiological comparison with *Pichia stipitis* CBS 6054. FEMS Yeast Res. **3,** 319–326.

Walfridsson, M., Bao, X., Anderlund, M., Lilius, G., Bülow, L. and Hahn-Hägerdal, B. (1996). Ethanolic fermentation of xylose with *Saccharomyces cerevisiae* harboring the *Thermus thermophilus xylA* gene which expresses an active xylose (glucose) isomerase. Appl. Environ. Microbiol. **62,** 4648–4651.

Walfridsson, M., Anderlund, M., Bao, X. and Hahn-Hägerdal, B. (1997). Expression of different levels of enzymes from the *Pichia stipitis XYL1* and *XYL2* genes in *Saccharomyces cerevisiae* and its effect on product formation during xylose utilisation. Appl. Microbiol. Biotechnol. **48,** 218–224.

Wang., M. M. and Reed, R. R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature **364,** 121–126.

Wang, J. R., Zhang, P., Huang, T. and Zhang, M. F. (1996). Molecular Characterisation of the *cbh2* gene of *Trichoderma viride*. Mycol. **88,** 191–197.

Ward, M., Wu, S., Dauberman, J., Weiss, G., Larenas, E., Bower, B., Rey, M., Clarkson, K. and Bott, R. (1993). Cloning, sequence and preliminary structural analysis of a small, high pI endoglucanase EGIII from *Trichoderma reesei*. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research, Helsinki. Vol 8. pp. 153–158.

Wieczorke, R., Krampe, S., Weierstall, T., Freidel, K., Hollenberg, C. P. and Boles, E. (1999). Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. FEBS Lett. **464,** 123–128.

Will, A., Grassl, R., Erdmenger, J., Caspari, T. and Tanner, W. (1998). Alteration of substrate affinities and specificities of the *Chlorella* hexose/H^+ symporters by mutations and construction of chimeras. J. Biol. Chem. **273,** 11456–11462.

Wilson, T. E., Fahrner, T. J., Johnston, M. and Milbrandt, J. (1991). Identification of the DNA binding site for NGFI-B by genetic selection in yeast. Science **252,** 1296–1300.

Wittstock, U., Chen, S. and Halkier, B. A. (2000). A method for expression cloning of transporter genes by screening yeast for uptake of radiolabelled substrate. J. Exp. Botany **51,** 955–960.

Würleitner, E., Pera, L., Wacenovsky, C., Cziferszky, A., Zeilinger, S., Kubicek, C. P. and Mach, R. L. (2003). Transcriptional regulation of *xyn2* in *Hypocrea jecorina*. Eucaryot. Cell **2,** 150–158.

Xue, G.-P., Orpin, C. G., Gobius, K. S., Aylward, J. H. and Simpson, G. D. (1992). Cloning and expression of multiple cellulase cDNAs from the anaerobic rumen fungus *Neocallimastix patriciarum* in *Escherichia coli*. J. Gen. Microbiol. **138,** 1413–1420.

Zaldivar, J., Martinez, A. and Ingram, L. O. (2000). Effect of alcohol compounds found in hemicellulose hydrolysate on the growth and fermentation of ethanologenic *Escherichia coli*. Biotechnol. Bioeng. **68,** 524–530.

Zeilinger, S., Mach, R. L. and Kubicek, C.P. (1998). Two adjacent protein binding motifs in the *cbh2* (cellobiohydrolase II-encoding) promoter of the *fungus Hypocrea jecorina* (*Trichoderma reesei*) cooperate in the induction by cellulose. J. Biol. Chem. **273,** 34463–34471.

Zeilinger, S., Ebner, A., Marosits, T., Mach, R. and Kubicek, C. P. (2001). The *Hypocrea jecorina* HAP 2/3/5 protein complex binds to the inverted CCAAT-box (ATTGG) within the *cbh2* (cellobiohydrolase II-gene) activating element. Mol. Gen. Genomics **266,** 56–63.

Zeilinger, S., Schmoll, M., Pail, M., Mach, R. L. and Kubicek, C. P. (2003). Nucleosome transactions on the *Hypocrea jecorina* (*Trichoderma reesei*) cellulase promoter *cbh2* associated with cellulase induction. Mol. Gen. Genomics **270,** 46–55.

Zhang, M., Eddy, C., Deanda, K., Finkestein, M. and Picataggio, S. (1995). Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. Science **267**, 240–243.

Appendices of this publication are not included in the PDF version.
Please order the printed version to get the complete publication
(http://www.vtt.fi/inf/pdf/)

Pubshed by

Series title, number and report code of publication

VTT Publications 541
VTT–PUBS–541

Author(s)
Saloheimo, Anu

Title

Yeast Saccharomyces cerevisiae as a tool in cloning and analysis of fungal genes Applications for biomass hydrolysis and utilisation

Abstract

The baker´s yeast, *Saccharomyces cerevisiae* has been employed by man for centuries in manufacturing of bread, beer, and wine. In science, it has become a useful tool as well. In this work, methods were developed in order to study the molecular biology of the cellulolytic filamentous fungus *Trichoderma reesei* with the aid of yeast. Cellulose is the most abundant carbon source in nature, and its enzymatic degradation is essential for carbon turnover. In addition, cellulose is used as a raw material in microbial processes. In this work, a previously unknown cellulase-encoding gene was cloned by expression in yeast and detection of hydrolysis halos on substrate plates. This EGV enzyme consists of an exceptionally small core domain, a cellulose-binding domain, and a linker region connecting the two. EGV belongs to family GH45 of glycosyl hydrolases. Additionally, a gene encoding a β-1,3-1,4-glucanase enzyme was cloned and studied. The enzyme was produced in insect cells, and analysis of the degradation products of β-glucan by NMR showed that it was a laminarinase (EC 3.2.1.6).

A yeast-based cloning method for positively acting regulatory proteins was set up, and two regulatory genes of the *T. reesei* cellulases, *ace1* and *ace2,* were isolated. The isolation was based on the ability of the encoded proteins to activate expression of a reporter gene, which was linked to the full-length promoter of the major cellulase gene *cbh1* in yeast. No homologs of the new regulatory proteins were detected outside the Mycota. The DNA-binding properties of the regulatory proteins were studied both *in vitro* and *in vivo* in yeast. Deletion of the *ace1* gene resulted in slower radial growth of the fungus on cellulose-containing plates. However, although isolated as an activator, ACEI was later shown to act as a repressor of hydrolase expression. ACEII, on the other hand, was shown to be an activator of cellulase expression. However, it is certainly not the only one, since its deletion did not result in a cellulase-negative phenotype.

Additionally, a sugar permease-encoding gene was isolated from *T. reesei* by complementation. The yeast strain used as a host was deleted for the major hexose transporter genes *(hxt1-7, gal2*), and additionally engineered for xylose utilisation. The *T. reesei* permease complemented the growth defect of the mutant strain on xylose-maltose medium. However, adaptive mutation(s) were needed in the host to enable growth on xylose of the permease-expressing strain. The same, engineered yeast strain was used as a host for the native *S. cerevisiae* hexose transporter genes *HXT1, HXT2, HXT4* and *HXT7*, and the kinetics of xylose transport were studied. The affinities of the permeases for xylose varied, K_m values of 190–900 mM were detected. Interesting differences were obtained in the levels of inhibition by the presence of glucose. The single-Hxt strains exhibited a biphasic growth mode on xylose media, where an initially very slow growth was followed by exponential growth after a lag of several days..

Keywords

Saccharomyces cerevisiae, *Trichoderma reesei*, *Hypocrea jecorina*, expression-cloning, reporter gene, cellulase, activator, ACEI, ACEII, EGV, permease, Hxt, xylose, uptake, xylose utilisation, bio-ethanol

Activity unit

VTT Biotechnology, Tietotie 2, P.O.Box 1501, FIN–02044 VTT, Finland

ISBN
951–38–6400–6 (soft back ed.)
951–38–6401–4 (URL:http://www.vtt.fi/inf/pdf/)

Project number

Date	Language	Pages	Price
June 2004	English	84 p. + app. 52 p.	C

Name of project

Commissioned by

Series title and ISSN

VTT Publications
1235–0621 (soft back ed.)
1455–0849 (URL: http://www.vtt.fi/inf/pdf/)

Sold by

VTT Information Service
P.O.Box 2000, FIN–02044 VTT, Finland
Phone internat. +358 9 456 4404
Fax +358 9 456 4374

VTT PUBLICATIONS

523 Soininen, Juha-Pekka. Architecture design methods for application domain-specific integrated computer systems. 2004. 118 p. + app. 51 p.

524 Tolvanen, Merja. Mass balance determination for trace elements at coal-, peat- and bark-fired power plants. 2004. 139 p. + app. 90 p.

525 Mäntyniemi, Annukka, Pikkarainen, Minna & Taulavuori, Anne. A Framework for Off-The-Shelf Software Component Development and Maintenance Processes. 2004. 127 p.

526 Jäälinoja, Juho. Requirements implementation in embedded software development. 2004. 82 p. + app. 7 p.

527 Reiman, Teemu & Oedewald, Pia. Kunnossapidon organisaatiokulttuuri. Tapaustutkimus Olkiluodon ydinvoimalaitoksessa. 2004. 62 s. + liitt. 8 s.

528 Heikkinen, Veli. Tunable laser module for fibre optic communications. 2004. 172 p. + app. 11 p.

529 Aikio, Janne K. Extremely short external cavity (ESEC) laser devices. Wavelength tuning and related optical characteristics. 2004. 162 p.

530 FUSION Yearbook. Association Euratom-Tekes. Annual Report 2003. Ed. by Seppo Karttunen & Karin Rantamäki. 2004. 127 p. + app. 10 p.

531 Toivonen, Aki. Stress corrosion crack growth rate measurement in high temperature water using small precracked bend specimens. 2004. 206 p. + app. 9 p.

532 Moilanen, Pekka. Pneumatic servo-controlled material testing device capable of operating at high temperature water and irradiation conditions. 2004. 154 p.

534 Kallio, Päivi. Emergence of Wireless Services. Business Actors and their Roles in Networked Component-based Development. 2004. 118 p. + app. 71 p.

535 Komi-Sirviö, Seija. Development and Evaluation of Software Process Improvement Methods. 2004. 175 p. + app. 78 p.

537 Tillander, Kati. Utilisation of statistics to assess fire risks in buildings. 2004. 224 p. + app. 37 p.

538 Wallin, Arto. Secure auction for mobile agents. 2004. 102 p.

539 Kolari, Juha, Laakko, Timo, Hiltunen, Tapio, Ikonen, Veikko, Kulju, Minna, Suihkonen, Raisa, Toivonen, Santtu & Virtanen, Tytti. Context-Aware Services for Mobile Users. Technology and User Experiences. 2004. 167 p. + app. 3 p.

540 Villberg, Kirsi, Saarela, Kristina, Tirkkonen, Tiina, Pasanen, Anna-Liisa, Kasanen, Jukka-Pekka, Mussalo-Rauhamaa, Helena, Malmberg, Marjatta & Haahtela, Tari. Sisäilman laadun hallinta. 2004. 172 s. + liitt. 20 s.

541 Saloheimo, Anu. Yeast *Saccharomyces cerevisiae* as a tool in cloning and analysis of fungal genes. Applications for biomass hydrolysis and utilisation. 2004. 84 p. + app. 52 p.

Tätä julkaisua myy	Denna publikation säljs av	This publication is available from
VTT TIETOPALVELU PL 2000 02044 VTT Puh. (09) 456 4404 Faksi (09) 456 4374	VTT INFORMATIONSTJÄNST PB 2000 02044 VTT Tel. (09) 456 4404 Fax (09) 456 4374	VTT INFORMATION SERVICE P.O.Box 2000 FIN–02044 VTT, Finland Phone internat. +358 9 456 4404 Fax +358 9 456 4374

ISBN 951–38–6400–6 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–6401–4 (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0849 (URL: http://www.vtt.fi/inf/pdf/)