GENETIC MANIPULATION OF *Saccharomyces cerevisiae* FOR IMPROVED ETHANOL PRODUCTION FROM D-XYLOSE

by

ROSHINI GOVINDEN

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Microbiology in the Faculty of Science at the University of Durban - Westville

Promoter	:	Prof. B. Pillay
Joint Promoter	:	Prof. D. Pillay
Date submitted	:	December, 1999

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy to the University of Durban - Westville, Durban. It has not been submitted before for any degree or examination to any other University.

R. GOVINDEN

December, 1999

TABLE OF CONTENTS

Acknowledgements

The author gratefully acknowledges the following people:

Prof. D. Pillay and Prof. B. Pillay for their encouragement and support during the course of this work,

Prof. W.H. van Zyl for guidance, support and provision of yeast strains and YEp vectors,

Prof. T.W. Jeffries for provision of kit for library construction,

Mrs. D. James for carrying out automated sequencing,

Mr. M. Marimuthu for assistance with the HPLC,

Dr. G. Govender for digital capture of images,

Kugen for his assistance, support and expert proof-reading of the thesis, and

The Foundation for Research Development and the UNESCO/Biotechnology Action Council for financial support and scholarships awarded.

List of Figures

List of Tables

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

Cellulosic biomass such as agricultural residues, paper wastes and wood chips is an ideal, renewable, abundantly available, inexpensive source of sugars for fermentation, especially in processes yielding ethanol which is used as a liquid fuel for transportation. Most of the hydrolysates of cellulosic biomass contain two major fermentable sugars: glucose and xylose. *Saccharomyces* spp. are the most effective microorganisms for fermenting hexose sugars to ethanol. Traditionally, these yeasts have been used in industry to ferment glucose-based agricultural products to ethanol. However, they are unsuitable for the fermentation of sugars derived from cellulosic biomass since they cannot ferment xylose to ethanol or use it for aerobic growth. While yeasts such as *Pichia stipitis* and *Candida shehatae* are naturally occurring xylose-fermenters, they are inefficient fermentative microorganisms since they have a relatively low ethanol tolerance, thus rendering them unsuitable for large scale commercial production of ethanol from cellulosic biomass.

Because of its successful exploitation in the fermentation industry, its proven ability to produce high ethanol concentrations rapidly, its high-level resistance to inhibitors found in lignocellulosic hydrolysates, the wealth of knowledge available about its genetics and physiology and its GRAS (Generally Regarded As Safe) status, *S. cerevisiae* is considered the ideal candidate for genetic engineering to develop a xylose-fermenting strain for ethanol production from xylose-containing hydrolysates.

This chapter provides an overview of microbiological, biochemical and biotechnological findings with regard to the transport, metabolism and fermentation of xylose by natural and genetically engineered yeast and bacterial strains.

1.1 HEMICELLULOSIC MATERIAL

The oil crisis in the mid-1970s generated an interest in transforming the vast reserves of renewable plant biomass to fuel and chemical feedstock (Rosenberg, 1980). Technologies have since become directed toward the utilization of lignocellulosic biomass from crop residues, forest product residues as well as other industrial waste carbohydrate streams. Virtually every nation has sufficient agricultural residues in the form of straws, hulls, stems and stalks, forestry lumber and paper residues and land-filled materials to supply a large portion of its own liquid fuel needs as ethanol. In the United States of America, over 70% of the materials placed in landfills are lignocellulosic materials such as paper, cardboard, yard trash and wood products. It was speculated that approximately 38 billion litres of fuel ethanol could be produced per year from the available material. By 1995, however, only 3.8 billion litres of fuel ethanol were produced from corn starch using traditional yeast-based methods (Ingram and Doran, 1995) and 5 billion litres by 1999 (Bothast *et al.*, 1999). In South Africa, 6 million tons of xylose, the major constituent of hemicellulose, could be recovered from plant refuse like sugarcane bagasse and maize residues (Dekker and Lindner, 1979). The acid hydrolysate from bagasse contains xylose as the main component (du Toit *et al.*, 1984).

Hemicelluloses are major components of lignocellulosic materials. Lignocellulose is a complex polymer comprising fibrous bundles of crystalline cellulose encased in a polymeric matrix of hemicellulose and lignin (Ingram and Doran, 1995). Hemicellulosic sugars average about 26% of the dry weight of hardwoods, 22% of softwoods and 30% of several major agricultural residues (Pettersen, 1984; Schneider, 1989). Other components include pectin, ash and protein.

Although cellulose is twice as abundant as hemicellulose, in general, hemicellulosic sugars can be recovered with milder treatment and in higher yield than glucose from cellulose. A yield of 55% glucose can be recovered from cellulose by dilute acid hydrolysis. In contrast, this

method yields 85% sugars, mainly xylose, from hemicellulose (Jeffries and Kurtzman, 1994). The difference between xylose and glucose yields can be attributed directly to the physical and chemical properties of xylan and cellulose, respectively and is therefore not amenable to process changes. Hemicellulose has a relatively open structure, unlike cellulose which is even impermeable to water. This open molecular architecture facilitates diffusion of acid into the polymer which aids hydrolysis. In addition, hemicellulose facilitates its own hydrolysis: acetyl groups are readily hydrolyzed and the resulting acetic acid catalyzes the partial depolymerization of hemicellulose. One of the main disadvantages of acid hydrolysis is that appreciable amounts of furfural and hydroxymethyl furfural are formed from glucose and xylose even under relatively mild conditions. These products are toxic to yeasts and may therefore inhibit their ability to ferment these hydrolysates.

The bioconversion of celluloses and hemicelluloses to alcohol offers an attractive means of alleviating future fuel shortages, as motor vehicles can be run on a mixture of ethanol and petrol and may in future run on ethanol alone. Current annual world ethanol production is >30 billion litres (Nissen and Nielsen, 1999) of which 80% is produced by anaerobic fermentation of various hexose sugars by *S. cerevisiae*. In Brazil, where fermentation technology is well established, 15 billion litres of ethanol were produced in 1999 from cane sugar (Wyman, 1999). Since the ethanol is primarily used as a renewable source of fuel, this has led to a reduction in petroleum imports by almost 50% (Nissen and Nielsen, 1999). The production of xylitol and furfural from hemicelluloses (xylose) has commercial applications. Xylitol is of the same order of sweetness as sucrose and fructose and therefore has application as a diabetic sweetener in foods. Xylitol is a good anticariogenic sweetener that can be metabolized by diabetic persons, because of its insulin-independent metabolic utilization (Emodi, 1978; Pepper and Olinger, 1988). Furfural is used in the manufacture of furfural-phenol plastics, varnishes, pesticides and as a useful

solvent for cellulosic derivatives (Dekker and Lindner, 1979).

1.2 XYLOSE METABOLISM

The catabolism of xylose proceeds *via* the pentose phosphate pathway (PPP). After transport of xylose into the cells it is first converted to xylulose (Fig. 1.1). In yeasts this is a two-step reduction and oxidation reaction (Chiang and Knight, 1960). Xylose is reduced to xylitol by xylose reductase (XR) using NADPH and/or NADH as cofactor and xylitol is oxidized to xylulose by xylitol dehydrogenase (XDH) coupling with NAD. Bacteria, on the other hand, achieve this conversion in one step using xylose isomerase (XI). Xylulose is then phosphorylated by xylulokinase (XK) to form xylulose-5-phosphate (Chakravorty *et al.*, 1962).

The conversion of xylose to xylulose is accompanied by an overall net positive free energy change (ΔG^{o} = +1.05 kcal/mol). Xylitol is a symmetrical molecule that does not exist in a ring form and, therefore, should have more rotational degrees of freedom and a lower energy state. Hence, in a two step reduction-oxidation mechanism, the production of xylitol is favoured. The equilibria for XR and XDH from *Pachysolen tannophilus* have been studied at pH 7.0. The equilibrium constant, $K_{eq} = 10^{-10}$ (Ditzelmüller *et al.*, 1984b) for XDH and $K_{eq} = 2.7\text{-}5.8 \times 10^{-8}$ (Ditzelmüller *et al.*, 1984a) for XR. These reactions favour the formation of xylitol. The actual concentration of xylose, xylulose and xylitol *in vivo* would depend on the intracellular NADPH/NADP ratio maintained by the oxidative phase of the PPP and the NADH/NAD ratio obtained from respiration and/or fermentation. The NADH/NAD ratio will be lower during respiration than during fermentation (Ditzelmüller *et al.*, 1984b).

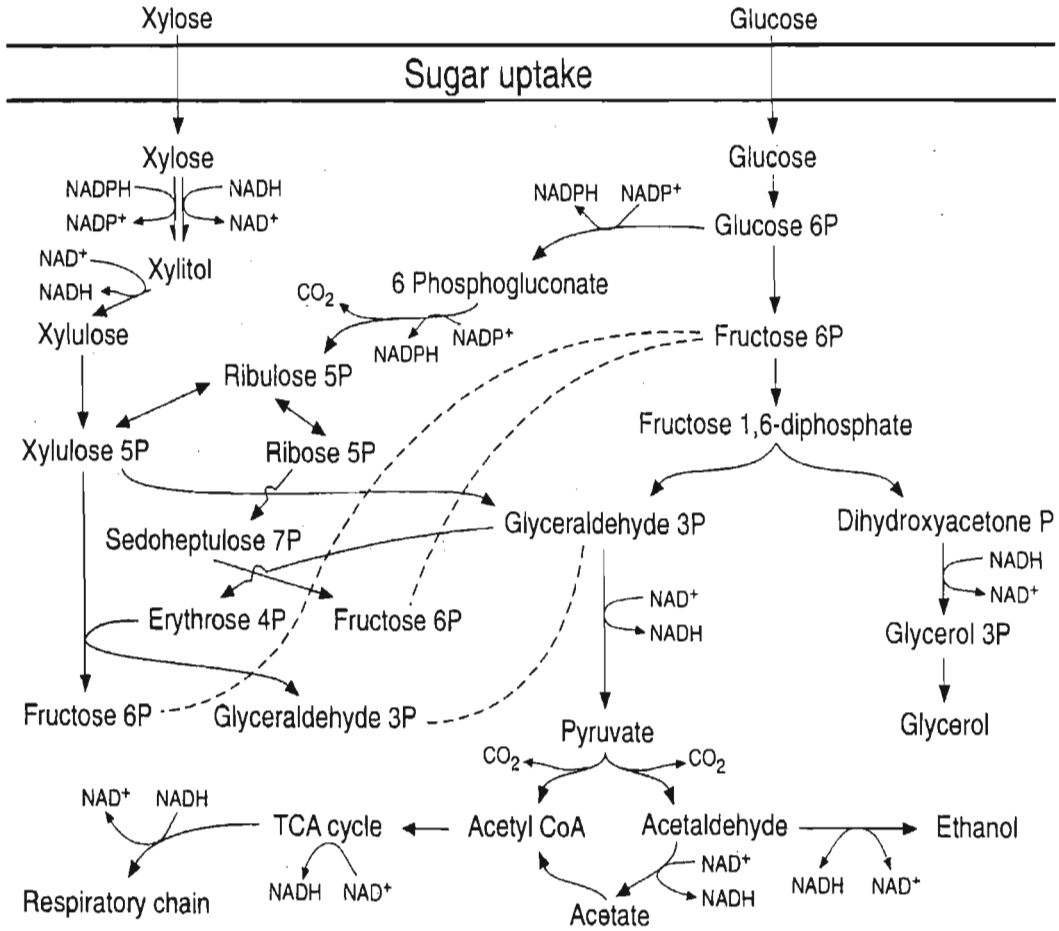


Fig. 1.1 Biochemical pathway of xylose metabolism (Hahn-Hägerdal *et al.*, 1994).

1.2.1 XYLOSE REDUCTASE

XR is the first enzyme in the metabolic pathway of xylose. Its activity and regulation are therefore critical for the fermentation of xylose. Most of the XR enzymes from xylose-utilizing yeasts are inducible. In addition to xylose, other monosaccharides may also act as inducers and the term "aldose reductase" is better suited (Bolen and Detroy, 1985).

Culture conditions have an effect on XR levels. The requirement for oxygen during xylose fermentation is frequently ascribed to the apparent redox imbalance which develops during anaerobic conditions due to the difference in cofactor requirements of the first two enzymes

(Bruinenberg *et al.*, 1983; 1984). Studies on *Candida utilis* provided the results for the redox imbalance theory. *C. utilis* has a XR requiring NADPH and a XDH requiring NAD^+ which results in an overproduction of $NADP^+$ and NADH. The $NADP^+$ can be reduced by recycling *via* fructose-6-phosphate but the NADH cannot be reoxidized in the absence of oxygen. Overproduction of NADH is reflected in an excretion of xylitol which can be prevented if an external electron acceptor is added (Bruinenberg *et al.*, 1983).

P. tannophilus has two XR enzymes, one is NADPH-specific and the other is NADH-specific (Ditzelmüller *et al.*, 1985). In contrast, both *P. stipitis* (Skoog and Hahn-Hägerdal, 1990) and *C. shehatae* (Alexander *et al.*, 1988c) have XR activity with dual cofactor specificity. In *P. stipitis*, kinetic measurements of purified XR indicate that XR prefers NADPH as a cofactor - the activity of XR with NADH is 72% of the activity with NADPH (Verduyn *et al.*, 1985b). The *C. shehatae* XR also has dual cofactor specificity with its activity with NADH being 40% (Bruinenberg *et al.*, 1984; Ho *et al.*, 1990) or 30% (Alexander *et al.*, 1988c) of the activity with NADPH. The ratio of NADH- to NADPH-linked activity shifts under the cultivation conditions employed. Under aerobic conditions, xylose fermentation by *P. stipitis* proceeds *via* NADH-linked xylose reduction *in vivo* although kinetic studies indicate that *in vitro* NADPH is the preferred cofactor. The relative concentration of NADPH and NADH as well as those of $NADP^+$ and NAD^+ are decisive for the choice between NADPH- or NADH-linked xylose reduction *in vivo*. *S. cerevisiae* possesses a XR activity which is very low (Batt *et al.*, 1986a) compared to that of the xylose-fermenting yeasts.

Several groups have studied the effect of oxygen on the redox imbalance by varying the oxygenation when xylose or glucose is used as a carbon source. In *C. shehatae*, Girio and co-workers (1989) showed that NADPH-dependent activity was inversely dependent on oxygen availability in shake-flask cultivation, while NADH dependent activity was independent of oxygen

availability. In most yeasts which do not, or only slowly, ferment xylose anaerobically, XR activity is NADPH-linked and they either lack or exhibit low NADH activity (Bruinenberg *et al.*, 1983). It has been suggested that NADH-linked XR activity is required for the anaerobic fermentation of xylose and that this leads to an imbalance of the NAD^+/NADH redox system (Bruinenberg *et al.*, 1984).

Since XR in *P. stipitis* (Bruinenberg *et al.*, 1984), *P. tannophilus* (Verduyn *et al.*, 1985a) and *C. shehatae* (Verduyn *et al.*, 1985b) can use either NADH or NADPH as cofactor, the redox imbalance caused by these two steps in xylose metabolism could be circumvented. Significant ethanol production reported only for these three yeasts can probably be explained by the fact that they also have NADH-linked XR activity in addition to the normal NADPH-linked activity (Bruinenberg *et al.*, 1984).

Most xylose fermenting yeasts produce some xylitol in addition to ethanol. Xylitol production increases as oxygen decreases (Ligthelm *et al.*, 1988b). Xylitol yields as high as 0.85 g/g consumed xylose have been observed in *P. tannophilus* (Hahn-Hägerdal *et al.*, 1994). Some strains of *P. stipitis* produce negligible amounts of xylitol under oxygen limited conditions and the best ethanol yields are found under these conditions.

Xylitol formation during anaerobic fermentation of xylose indicates that the redox balance has not completely closed, i.e., the cell has a surplus of NADH and a lack of NAD^+. However, during oxygen limited fermentation of xylose by *P. stipitis* CBS 6054, no xylitol was formed (Skoog and Hahn-Hägerdal, 1990). This has been ascribed to the existence of an alternate respiratory system (cyanide-insensitive respiration) which acts as a redox sink providing the XDH reaction with NAD^+ (Jeppsson *et al.*, 1995). Oxygen may thus not only help to alleviate the redox imbalance in different strains but may also serve other purposes. Further support for this interpretation was provided by studies of *P. stipitis* strains under conditions of oxygen limitation

where oxygen was found to be necessary for ethanol production from glucose which does not involve XR or XDH (Ligthelm *et al.*, 1988b; Skoog *et al.*, 1992b). Surprisingly, xylitol was either not, or only poorly, fermented by the xylose fermenting yeasts whereas growth on xylitol was comparable to growth on xylose (du Preez *et al.*, 1986b; Neirinck *et al.*, 1985).

1.2.2 XYLITOL DEHYDROGENASE

XDH catalyzes the conversion of xylitol to xylulose and is inducible by xylose, D-galactose and L-arabinose. Whereas XR catabolizes conversion of L-arabinose to L-arabitol and D-galactose to D-galactitol, these products are not oxidized substrates for XDH which has an NADH-linked activity. Although XDH is present in L-arabinose and D-galactose grown cells, it may serve no metabolic function but occurs as a consequence of simultaneous induction with XR (Bolen and Detroy, 1985). This suggests that the mechanism leading to expression of these two enzymes (either by new enzyme synthesis or activation of existing enzyme) is common for both enzymes.

Studies of the *in vitro* activity of XDH at different oxygen levels in *P. stipitis* and *C. shehatae* strains showed that the level of oxygen supply during fermentation had little effect on XDH activity (Alexander *et al.*, 1988c; Skoog and Hahn-Hägerdal, 1990; Smiley and Bolen, 1982). XDH from *P. tannophilus* consists of four subunits and is reported to have a molecular weight of 172 kDa (Bolen *et al.*, 1986). In *C. shehatae* XDH consists of two subunits and has a molecular weight of 82 kDa (Yang and Jeffries, 1990). In *P. stipitis* this enzyme also comprises two subunits and has a molecular weight of 63 kDa (Rizzi *et al.*, 1989b).

P. tannophilus accumulates more xylitol than *P. stipitis* or *C. shehatae* (du Preez and van der Walt, 1983; Slininger *et al.*, 1985). The Michaelis-Menten constant (K_m) for xylitol in *C. shehatae* is a quarter that observed with the *P. tannophilus* enzyme while the K_m for xylulose

is 1.7 times higher (Ditzelmüller *et al.*, 1984a). This would favour the forward reaction. A lower K_m of XDH for xylitol may explain the higher ethanol yield for *P. stipitis* and *C. shehatae* (du Preez and van der Walt, 1983).

The production of xylitol by *P. stipitis* under certain conditions could be expected from the unfavourable thermodynamic equilibrium constants which have been estimated to be 6×10^9 l/mol for the conversion of xylose to xylitol (Rizzi *et al.*, 1988) and 7×10^{-11} mol/l for the conversion of xylitol to xylulose (Rizzi *et al.*, 1989a). Since xylitol production can be overcome by the presence of oxygen or any other electron acceptor, this unfavourable equilibrium would not appear to be the only reason for xylitol production.

Two genes encoding XDH have been cloned from *P. stipitis*. The first is responsible for XDH activity in *P. stipitis* (Kötter *et al.*, 1990). It is not known how the second gene may be involved in xylitol metabolism (Persson *et al.*, 1993). A recent report revealed that the gene YLR070c from *S. cerevisiae* encodes a XDH (Richard *et al.*, 1999). This gene was similar to the *P. stipitis XYL2* gene and its over-expression in *S. cerevisiae* led to NADH-specific XDH activity.

1.2.3 XYLULOKINASE

In addition to the initial steps of xylose metabolism, the results of molecular and genetic studies indicate that the third step in xylose metabolism from xylulose to xylulose-5-phosphate catalyzed by XK may also be limiting (Ho and Chang, 1989; Jeffries, 1984a). The XK activities of *C. shehatae*, *P. tannophilus*, *P. stipitis*, *S. cerevisiae* and *Schizosaccharomyces pombe* were analyzed and compared with their effectiveness in producing alcohol during xylose fermentation. *C. shehatae*, *P. stipitis* and *S. pombe* expressed high levels of XK activity and *P. tannophilus* and *S. cerevisiae* low levels. Since the former are good xylose-fermenters, the low levels of XK activity were considered the reason for the ineffective xylose fermentation to ethanol of the latter

two yeasts (Deng and Ho, 1990). XK from *P. stipitis* NCYC 1541 has been purified and characterized (Flanagan and Waites, 1992) and the native enzyme has a molecular mass of 120-130 kDa and comprises two identical subunits.. The K_m for D-xylulose is 5.2×10^{-4} M which is a very high specificity. XK requires magnesium ions for optimal activity.

It has been suggested that after its formation, xylulose-5-phosphate is converted to D-glyceraldehyde-3-phosphate *via* enzymes of the PPP. Subsequent use of carbon is thought to follow one or more of several different routes, viz., (i) incorporation into biomass; (ii) conversion to ethanol *via* the Embden-Meyerhoff-Parnass (EMP) pathway following the activation of pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH); and (iii) oxidation to carbon dioxide and water *via* respiratory processes. Alternatively, D-xylulose-5-phosphate can be split into glyceraldehyde-3-phosphate and acetyl phosphate by phosphoketolase (Hahn-Hägerdal *et al.*, 1994).

1.2.4 THE PENTOSE PHOSPHATE PATHWAY

D-xylulose-5-phosphate is further metabolized *via* the PPP (also known as the hexose monophosphate shunt or phosphogluconate pathway). The major products of the PPP are fructose-6-phosphate and glyceraldehyde-3-phosphate which can be recirculated *via* the PPP and the glycolysis pathway to yield several intermediates. In the overall metabolism, the main function of the PPP is to produce NADPH, ribose-5-phosphate and erythrose-4-phosphate required for biosynthesis (Fig. 1.1). All xylose has to be shuttled through the PPP, whereas it has been shown that only a minute part (0.9%) of glucose is distributed through this pathway in *S. cerevisiae* (Gancedo and Lagunas, 1973). In contrast to glycolysis, no energy is gained when xylose is metabolized through the PPP.

Mutants of *P. tannophilus* with enhanced specific ethanol productivities also revealed

higher activities of PPP enzymes (Jeffries, 1984a; Lachke and Jeffries, 1986). In *S. cerevisiae*, the ten-fold lower rates of xylulose metabolism compared to that of glucose is explained by a greater accumulation of sedoheptulose-7-phosphate in xylulose-fermenting strains (Senac and Hahn-Hägerdal, 1990). These findings are thought to be the result of competition between PPP and glycolysis for glyceraldehyde-3-phosphate.

Fructose-6-phosphate can be either recycled in the oxidative part of the PPP where carbon dioxide is produced or through shuttling to the glycolysis pathway. There is some uncertainty as to the theoretical yield of ethanol from xylose. Depending on the proportion of fructose-6 phosphate circulated through the PPP (i.e., as opposed to its possible conversion by an active phosphoketolase) the yield may vary between 0.31 and 0.61 g ethanol/g xylose (Evans and Ratledge, 1984; Slininger *et al.*, 1987). If recirculation does not occur, the yield will be 0.51 g/g. However, if fructose-6-phosphate is recycled *via* the PPP the yield may decrease to 0.31 g/g. On the other hand, if an active phosphoketolase is present the yield would increase to 0.61 g/g (Evans and Ratledge, 1984).

Based on the yields of ethanol from xylose in *P. stipitis*, where 45 g per 100 g substrate utilized have been reported (Chiang *et al.*, 1981b) and on a comparison of biomass yields from glucose and xylose (Evans and Ratledge, 1983), Evans and Ratledge (1984) suggested an *à priori* case for pentose metabolism being more efficient than theory would predict. They examined 25 xylose-fermenting yeast species for the presence of phosphoketolase. The enzyme was present at high activity in 19 out of 25 species that are efficient xylose metabolisers. No significant phosphoketolase activity was detected in any yeast grown on glucose and the enzyme was induced by up to 70-fold on xylose as a sole carbon source. These workers also found that biomass yields on xylose were greater than, or equal to, those from glucose in 15 out of 19 species with phosphoketolase activity. While the molar yields of C_2 units from xylose could not account for

the high biomass and ethanol yields from xylose, the presence of phosphoketolase may account for these high yields, since this activity produces two mol C_2 from one mol C_5.

1.2.5 FINAL FERMENTATION STEPS

Fructose-6-phosphate and glyceraldehyde-3-phosphate, the metabolites from the PPP are further metabolized to pyruvate by glycolysis. Pyruvate constitutes a junction at which the cells may either follow a fermentative pathway and/or an oxidative pathway. In the fermentative pathway, PDC and ADH convert pyruvate to ethanol. The NADH formed in the oxidation of glyceraldehyde-3-phosphate is reoxidized. Studies have been carried out to investigate PDC and ADH in relation to xylose metabolism (Alexander *et al.*, 1988c; Prior *et al.*, 1988; Skoog and Hahn-Hägerdal, 1990; Skoog *et al.*, 1992b).

The *in vitro* PDC levels were similar in *P. stipitis* grown on xylose and glucose at various oxygen concentrations (Skoog and Hahn-Hägerdal, 1990; Skoog *et al.*, 1992b). The *in vitro* PDC level increased with increasing ethanol production. However, during xylitol assimilation the PDC level was ten times lower than the levels for xylose- or glucose-assimilating cells (Hahn-Hägerdal *et al.*, 1994).

In *S. cerevisiae*, ADH occurs as four isozymes, ADHI (Lutsorf and Megnet, 1968), ADHII (Ciriacy, 1975), ADHIII (Pilgrim and Young, 1987) and ADHIV (Paquin and Williamson, 1986). ADHI and ADHIII are responsible for ethanol production, whereas ADHII is responsible for ethanol oxidation. The function of ADHIV remains unknown. Two ADH isozymes (PsADHI and PsADHII) with 71.9 and 74.7% homology to the ADHI from *S. cerevisiae* were isolated from *P. stipitis* (Cho and Jeffries, 1998). PsADH1 plays a major role in xylose fermentation since its disruption resulted in a lower growth rate and high levels of xylitol accumulation. Deletion of PsADH2 did not significantly affect growth or ethanol production. Disruption of both genes,

however, blocked ethanol respiration but not production, implying a pathway for fermentation.

ADH activity in *C. shehatae* increased (four-fold) and the number of isozymes increased from one to three with decreasing oxygen supply during fermentation (Alexander *et al.*, 1988c; Prior *et al.*, 1988). Apart from the fact that the *in vitro* level of ADH is lower when xylose is the carbon source instead of glucose (Prior *et al.*, 1988), there seems to be no limitation on the conversion of xylose to ethanol in the final fermentation steps.

1.2.6 RESPIRATORY PATHWAY

Pyruvate is oxidized *via* the TCA cycle and the respiratory chain (oxidative phosphorylation) in the presence of oxygen. These reactions are carried out by respiratory assemblies within the inner mitochondrial membrane. Slininger *et al.* (1991) showed that *P. stipitis* has an absolute requirement for oxygen in continuous culture on xylose (where specific oxygen uptake was proportional to growth). This was also demonstrated by Neirinck *et al.* (1984) who showed that yeasts are unable to grow on xylose under anaerobic conditions. Studies have shown that anaerobic growth on xylose does not occur for *P. stipitis* (Neirinck *et al.*, 1984), *C. shehatae* (du Preez *et al.*, 1984) or *P. tannophilus* (du Preez *et al.*, 1984; Neirinck *et al.*, 1984). *S. cerevisiae* cannot grow anaerobically on xylulose (Maleszka and Schneider, 1984). However, when supplemented with ergosterol and unsaturated fatty acids, growth occurs on glucose anaerobically. Similar supplementation with xylose-fermenting yeasts was not successful (du Preez *et al.*, 1984; Neirinck *et al.*, 1984).

Since ethanol production is enhanced in the presence of oxygen, it is most likely that ethanol production is either growth-related for both xylose and glucose (Slininger *et al.*, 1990) or dependent on functional mitochondria. The need for an unimpaired mitochondrial function for growth on xylose and optimal xylose fermentation has been proposed based on findings of a study

by Ligthelm *et al.* (1988a) who used respiratory inhibitors (potassium cyanide, antimycin, sodium azide and rotenon) during xylose assimilation by *P. stipitis* and *P. tannophilus* and glucose assimilation by *S. cerevisiae*. The response was markedly different for *S. cerevisiae* in which mutant cells were capable of mitochondrial-independent growth (Alexander and Detroy, 1983) compared to the two xylose fermenters which showed similar effects. The necessity for unimpaired mitochondrial function was further supported by studies with a respiratory-deficient mutant of *P. tannophilus* (Alexander, 1990) and the failure to obtain stable petite mutants of *P. tannophilus* (Neirinck *et al.*, 1984) and *C. shehatae* (Jeffries, 1984b).

The difference between hexose and pentose metabolism is due to the shuttling through the PPP of pentoses. Ethanol production in xylose-fermenting yeasts was approximately twice as high on glucose than on xylose (Grootjen *et al.*, 1990; Ligthelm *et al.*, 1988b). *S. cerevisiae* (Senac and Hahn-Hägerdal, 1990) and *S. pombe* (Lastick *et al.*, 1990) also produce twice as much ethanol from glucose than xylulose. Rizzi and co-workers (1989c) suggested that the absence of anaerobic growth on xylose by *P. stipitis* was a result of low levels of ATP produced under anaerobic conditions since the PPP does not generate energy but rather metabolites for biosynthesis.

1.3 XYLOSE-FERMENTING ORGANISMS

1.3.1 YEASTS

The fermentation of xylose to ethanol was first reported in 1959 by Karczewska. However, the importance of this observation went unnoticed until 1984 (Toivola *et al.*, 1984). Barnett (1976) reported that approximately half of the known yeast species could assimilate xylose but that none could ferment it. The discovery by Wang *et al.* (1980b) that *S. pombe* and

various other yeasts could ferment D-xylulose, which is the keto-isomer of xylose, to ethanol triggered the current efforts to ferment xylose. There are a number of microorganisms capable of making ethanol from xylulose. The best xylulose-fermenters are *C. tropicalis*, *Hansenula anomala*, *Kluveromyces fragilis*, *S. uvarum* and *S. pombe* with yields ranging from 0.41 to 0.47 g ethanol/g xylulose (Gong *et al.*, 1983; Ueng *et al.*, 1981).

Yu *et al.* (1995) examined xylulose fermentation by *P. stipitis*, *C. shehatae* and four *S. cerevisiae* strains. Although *P. stipitis* fermented xylose rapidly under oxygen-limitation, it consumed xylulose very slowly. The *C. shehatae* strain studied exhibited the highest consumption rate (0.345 g/g cells/h) under anaerobic conditions but a quarter of the xylulose was back-converted to xylitol, thereby decreasing the ethanol yield to 0.36 g/g. Xylitol was the main by-product in xylulose fermentation with *S. cerevisiae* (Chiang *et al.*, 1981a). However, very high levels of arabinitol (up to 4 g/l) were obtained in this study. Xylose can be readily converted to D-xylulose by XI, an enzyme that is readily available and produced on a large scale. This approach was developed for fermenting xylulose using *S. cerevisiae* (Chiang *et al.*, 1981a; 1981b; Suihko and Poutanen, 1984), *S. pombe* (Ueng *et al.*, 1981; Wang *et al.*, 1980a) and *C. tropicalis* (Jeffries, 1981b). Although the approach is technically feasible, incompatible temperature and pH optima for enzymatic isomerization and fermentation and the expense of the XI have limited its practicality.

This led to the investigations for yeasts capable of direct xylose fermentation and resulted in almost simultaneous reports from four laboratories. Following the screening of yeasts for anaerobic xylose fermentation, Schneider *et al.* (1981) and Slininger *et al.* (1982) reported ethanol production directly from xylose by *P. tannophilus*. Mutation and selection studies on a *Candida* spp. known to assimilate xylose produced a yeast strain capable of fermenting xylose (Gong *et al.*, 1981b). Jeffries (1981a) observed ethanol production from xylose/xylulose mixtures by

C. tropicalis under aerobic conditions.

Following these initial observations, several laboratories screened yeasts for direct conversion of xylose to ethanol. Early comparative studies grouped yeasts capable of xylose metabolism into two distinct classes: those capable of ethanol production under strict anaerobic conditions and those that required some oxygen for ethanol production. *P. tannophilus* was capable of carrying out a strictly anaerobic fermentation (Schneider *et al.*, 1981). Many other xylose fermenting yeasts, however, require oxygen and to date none have been shown to be capable of growth under strictly anaerobic conditions.

du Preez and van der Walt (1983) described *C. shehatae* as a xylose fermenting yeast with higher rates and better yields than *P. tannophilus*. Several other reports on the discovery of xylose fermenting yeasts followed. Toivola *et al.* (1984) investigated approximately 200 type species capable of xylose metabolism. Those exhibiting significant ethanol production included *Bretannomyces naardenensis*, *C. shehatae*, *C. tenuis*, *P. segebiensis*, *P. stipitis* and *P. tannophilus*. du Preez and Prior (1985) identified *P. stipitis* and *C. shehatae* as the best xylose fermenting species. Strains of *P. stipitis* and *C. shehatae* differ considerably in their relative capacities for xylose fermentation. In general, *C. shehatae* strains produce ethanol more rapidly but also produce by-products (e.g., xylitol and glycerol) while *P. stipitis* strains produce higher ethanol concentrations, consume more xylose and produce very little or no by-products (du Preez *et al.*, 1986a; Slininger *et al.*, 1985).

To date, *P. tannophilus*, *P. stipitis* and *C. shehatae* are the most thoroughly investigated and efficient naturally occurring xylose fermenters among yeasts. For optimal performance however, they require carefully controlled oxygenation (Ligthelm *et al.*, 1988b; Skoog and Hahn-Hägerdal, 1990). The two main reasons for yields below the theoretical values are: (i) the production of xylitol, especially by *P. tannophilus*; and (ii) the reassimilation of ethanol

(Lohmeier-Vogel *et al.*, 1989; Skoog *et al.*, 1992a).

1.3.2 BACTERIA

Most fuel ethanol is currently produced from hexose sugars in corn starch or cane sugars by using the bacterium *Zymomonas mobilis* (Murtagh, 1986) or the yeast *S. cerevisiae* (Rosillo-Calle and Hall, 1987). *Z. mobilis* produces ethanol as rapidly as *S. cerevisiae* and with almost no by-product formation from glucose and fructose. It cannot, however, ferment xylose. Most other bacteria have a broad substrate specificity but ethanol is not usually the predominant product from xylose, e.g., *Bacillus macerans* produces ethanol (3.3 g/l), formic acid (1.8 g/l) and acetone (0.7 g/l) from 20 g/l xylose (Schepers *et al.*, 1987).

Bacteria ferment a number of monosaccharides and disaccharides in addition to more complex carbohydrates, such as starch and cellulose (Wiegel and Ljungdahl, 1986). In general, ethanol yield and productivity are lower for bacteria compared to yeasts and fungi. The yields for *Thermoanaerobacter ethanolicus* (0.36 g/g) and *Clostridium thermohydrosulfuricum* (0.39 g/g) are among the higher values. A strain of *Bacteroides xylanolyticus* capable of fermenting xylose and glucose showed neither diauxic growth nor substrate preference. XI was induced immediately in the presence of xylose and XK after 30 minutes (Biesterveld *et al.*, 1994). Klapatch *et al.* (1994) determined the effect of ethanol inhibition on the growth rate at 55°C and 60°C for *C. thermosaccharolyticum* in continuous culture with both xylose and glucose as carbon sources. They showed an almost linear pattern in relation to ethanol concentration. This indicated that the ethanol tolerance of this bacterium was sufficiently high such that this is unlikely to constrain its use in biomass and ethanol production.

1.4 XYLOSE TRANSPORT

Membrane transport has been considered to play an important role in the utilization of sugars present in hemicellulose and cellulose hydrolysates and eventually limit the production of ethanol from lignocellulosic materials by yeasts (Spencer-Martins, 1994). Mono- and disaccharides can be transported into yeasts either by facilitated diffusion or active transport. Xylose is transported by either of these mechanisms. Both of these mechanisms use carrier proteins called permeases. Sugar affinities are 10 to 100 fold higher for active transport systems compared to passive diffusion. This enables efficient transport even in very dilute sugar solutions. The plasma membrane ATPase provides energy for transport against a concentration gradient.

Transport mechanisms are under the control of the cell and the main controlling factors are the (i) concentration of the sugars; (ii) energetic state of the cell; and (iii) availability of oxygen. Specificity of the transport system varies considerably. Unlike active transport, facilitated diffusion is usually not very specific for the sugar.

Pentose transport in yeasts has been most extensively studied in *Rhodotorula gracilis* (*glutinis*) where it occurs by an active process (Alcorn and Griffin, 1978; Janda *et al.*, 1976). Some transport data is available for *P. stipitis* and *C. shehatae*. Kinetic studies have revealed the presence of more than one uptake system for xylose in *R. glutinis* (Alcorn and Griffin, 1978), *P. stipitis* (Does and Bisson, 1989; Kilian and van Uden, 1988) and *C. shehatae* (Lucas and van Uden, 1986). Xylose transport in *P. stipitis* is mediated by two active transport systems (Does and Bisson, 1989; Kilian and van Uden, 1988) whereas *C. shehatae* has an active transport system as well as a facilitated diffusion system (Lucas and van Uden, 1986).

Several studies have demonstrated that xylose transport is apparently the rate limiting step in xylose metabolism under aerobic and oxygen-limited conditions for *P. stipitis* and *C. shehatae* (Alexander *et al.*, 1988c; Kilian and van Uden, 1988; Ligthelm *et al.*, 1988c). To date the gene(s)

coding for xylose transport in yeasts have not been isolated. However, xylose transport mutants have been isolated from the xylose-fermenting yeast *Pichia heedii* by random mutagenesis (Does and Bisson, 1990). Availability of these mutants should facilitate cloning and identification of the xylose uptake system in yeasts.

In *S. cerevisiae*, which does not metabolize xylose, transport occurs by facilitated diffusion (Kleinzeller and Kotyk, 1967). Xylose transport is apparently related to glucose transport since the influx of xylose is more rapid under both aerobic and anaerobic conditions in the presence of glucose (Ligthelm *et al.*, 1988b). The glucose transport system has a broad specificity for pyranoses. Both the high and low affinity glucose transport system take up xylose (Busturia and Lagunas, 1986). Xylose is transported into the cells *via* hexose carriers and the affinity of the carriers for glucose is high (K_m = 1mM) and that for xylose is low (K_m = 200 mM). Xylose uptake in *S. cerevisiae* and *C. utilis* is less efficient (26%) than glucose transport (Batt *et al.*, 1986a). The affinity is, however, 4 - 25 fold lower for xylose than glucose (Cirillo, 1968; Kotyk, 1967). Kötter and Ciriacy (1993) obtained K_m values of 1.5 mM and 35 mM for glucose and 190 mM and 1.5 M for xylose under the same conditions, for the high and low affinity transport systems, respectively. This implies that the monosaccharide transport system in *S. cerevisiae* has a nearly 200-fold lower affinity for xylose than for glucose.

David and Weismeyer (1970) reported an inducible xylose permease in *E. coli* that is energy dependent and specific for xylose, and that can transport xylose against a 100 fold concentration gradient. Lam and co-workers (1980) suggested that the mechanism of energization for xylose uptake is by a proton motive force rather than by a directly energized mechanism.

1.5 TECHNOLOGY TO IMPROVE XYLOSE FERMENTATION

A xylose fermenting yeast would need to produce 50 - 60 g/l ethanol within 36 h with a yield of at least 0.4 g/g before commercial application could be considered (Jeffries, 1985). To date, *C. shehatae* and *P. stipitis* are superior to all other yeast species in terms of rate and yield of ethanol production from xylose. However, fermentation times of >36 h were required to achieve the desired ethanol concentrations in hydrolysates (Yu *et al.*, 1987) and pure xylose media (Slininger *et al.*, 1985). Several techniques have been employed in an effort to improve xylose fermentation. These include an improved fermentation technology and strain manipulation by molecular and genetic techniques to produce improved or new strains of *C. shehatae*, *P. stipitis* and *S. cerevisiae*. These aspects are discussed in detail below.

1.5.1 FERMENTATION TECHNOLOGY

Several fermentation techniques have been employed to improve xylose fermentation. These include cell immobilization and/or recycling to increase densities, continuous culture, ethanol removal and co-culture. Each technique has been beneficial in some way, but none has provided a total solution to the problem.

The relatively low fermentation rate requires either the use of large reactor volumes or high cell densities. *P. stipitis* immobilized in either agar beads or on a fine nylon mesh attain ethanol concentrations of up to 40 g/l in eight days from 100 g/l xylose (Linko *et al.*, 1986). Dynamic cell immobilization in which cells were continuously recycled by the use of a membrane module also proved successful. With this technique, up to 70 g/l (dry weight) of cells - or about 10 to 20 times the cell density without recycling - was attained. The volumetric fermentation rates increased up to 4.4 g/l/h ethanol, but the specific rates dropped dramatically (Sreenath and Jeffries, 1987).

Continuous fermentation without recycling has not been particularly successful. The principal problem in using continuous cultures was that the biomass increased with aeration whereas ethanol yield decreased. Also, the need to produce new biomass reduced the ethanol yield that could be attained. Although single-stage continuous cultures were not particularly useful for xylose fermentation, a multiple-stage continuous fermentation could enable *C. shehatae* to achieve considerably higher ethanol concentrations (Alexander *et al.*, 1988a). The first stage reactor, operated at an oxygen-limited mode will produce fermentative cells. A low dilution rate must be used because of the oxygen limitation. In the second stage, the cells and fresh sugar would be introduced. Minimal aeration would be employed in this stage and ethanol concentration would be appreciably higher. Although growth stopped, ethanol production continued vigorously. It is important to maintain an influx of viable cells to maintain ethanol production (Alexander *et al.*, 1988a; 1988b).

In an attempt to achieve combined fermentation of the two sugars from lignocellulosic materials by sequential/co-culture processes, Laplace *et al.* (1992) determined the compatibility and typing of associated strains. They reported the occurrence of the killer phenomenon in six *Saccharomyces* spp. and 11 xylose fermenting yeasts, which precludes their utilization in a co-culture process. The killer factor is a proteinaceous toxin whose synthesis is directed by a dsRNA virus-like particle present in many industrial yeasts (Spencer and Spencer, 1983). However, five strains of *C. shehatae* tested did not show any inhibition on the growth of *S. cerevisiae*. Thus, a co-culture process using these strains could be developed.

Laplace *et al.* (1993a) determined the effects of culture conditions on the co-fermentation of glucose and xylose mixtures to ethanol by co-culture of a respiratory deficient mutant of *S. diastaticus* and *P. stipitis*. With a 70% glucose/30% xylose sugar mixture at 50 g/l, xylose was entirely consumed when the dilution rate did not exceed 0.006 h^{-1}. Glucose was entirely

consumed whatever the dilution rate. At a low dilution rate, *P. stipitis* was the dominant species and high dilution rates resulted in washing out of *P. stipitis* due to its low ethanol tolerance. Laplace *et al.* (1993b) also tested separated and co-culture processes for *P. stipitis*/*Z. mobilis* and *P. stipitis*/*S. diastaticus* for ethanol production from mixtures of glucose and xylose. Best results were obtained in separated glucose (180 g/l) and xylose (80 g/l) fermentations using *P. stipitis* and *Z. mobilis*. Maximum ethanol concentrations achieved were 86.2 g/l and 29 g/l, respectively.

In xylose fermentation, both yield and productivity (specific as well as volumetric) are one to two orders of magnitude lower than in hexose fermentation. Therefore, the focus of research in this field has been to understand the limitations of xylose fermentation with a view to finding suitable organisms and process conditions and to use new genetic engineering techniques to improve yield and productivity.

1.5.2 STRAIN IMPROVEMENT

The major limiting factors of pentose fermentations are (i) the rate of bioconversion; (ii) product yield; and (iii) tolerance to solvents produced (Mishra and Singh, 1993; Skoog and Hahn-Hägerdal, 1988). In order to make the fermentation process economically competitive, the microbial strains need to be improved. Current strategies involve the selection and isolation of mutant strains as well as recombinant DNA techniques.

The major objectives for genetic improvement of microorganisms involved in pentose bioconversion are to (i) improve the efficiency of pentose bioconversion; (ii) increase productivity by (a) increase the rate of substrate bioconversion; (b) raise fermentation temperatures; (c) improve tolerance to solvents; (d) increase utilization of pentoses in the presence of hexoses; and (e) minimize by-product formation; (iii) increase resistance to natural inhibitors present in lignocellulosic hydrolysates; and (iv) broaden the pentose utilizing ability of solventogenic

microorganisms (Singh and Mishra, 1995).

Some of the approaches to improve pentose fermentation include screening, mutation, recombination and gene cloning. Two main strategies adopted in screening for xylose fermenting yeasts are to (i) screen for organisms from natural or industrial habitats (Linden *et al.*, 1992; Nigam *et al.*, 1985) or (ii) look at species known by taxonomy to be xylose assimilators (Toivola *et al.*, 1984).

1.5.2.1 Screening and mutagenesis

Screening does not require biochemical or genetic information on the organism and is the most direct and least expensive method. Methods for the selective enrichment of microorganisms for lignocellulosic bioconversion have been employed. Screening protocols involve (i) choice of microorganism; (ii) induction of genetic variability in the cell population; (iii) small scale fermentations of many individuals from the population; and (iv) assays of the fermentation products to identify an improved strain (Singh and Mishra, 1995). By repeated recycling of a *C. shehatae* strain on wood-derived hemicellulose hydrolysates, Parekh *et al.* (1986) obtained increased productivity from 0.31 to 0.44g/l/h and increased yield from 0.29 to 0.4g/g xylose.

The following factors need to be considered for mutation protocols for a particular microorganism, viz., mutagen concentration, time of exposure and conditions of treatment that produce the highest proportion of a particular class of mutants. Both direct and indirect mutagens cause mutations by mispairing mechanisms that involve either template or nucleotide precursors. Indirect mutagens induce a post-replication repair system that is prone to error (Miller, 1983). Cells are allowed to undergo a period of DNA replication and growth after mutagen treatment so that the damaged DNA will be converted into stable altered DNA encoding reproducible and heritable mutants. Several mutants have been isolated for bioconversion of lignocellulosic sugars

into solvents.

Gong and co-workers (1981a) used UV irradiation (indirect mutagen) to isolate a mutant *Candida* spp. (*Candida* XF217) that produced five times more ethanol than the parental strain. The mutant was capable of ethanol production from xylose both aerobically and under oxygen-limiting conditions. The increased ethanol production was attributed to increased XDH and XK activities that enabled a shift from xylitol to ethanol production. Jeffries (1984a) isolated *P. tannophilus* mutants by UV mutagenesis followed by enrichment on nitrate and xylitol as sole sources of nitrogen and carbon, respectively. Mutants exhibited a higher rate of ethanol production and yield from xylose. Ethanol production under aerobic conditions was twice as fast as the parental strain and only 50% faster under anaerobic conditions. Lee *et al.* (1986) selected UV-induced mutants of *P. tannophilus* on the basis of diminished growth on ethanol. The selection strategy was based on enhanced ethanol accumulation by minimizing losses due to its oxidation. Three of the eleven independent mutant loci that conferred the ethanol defective phenotype produced significantly more ethanol than the wild type in aerobic batch cultures.

1.5.2.2 Genetic recombination

1.5.2.2.1 Hybridization

Industrial yeast strains are generally polyploid or aneuploid (Spencer and Spencer, 1983), suggesting that an increase in chromosome number may be advantageous in processes involving such strains. Hybridization or crossbreeding has been widely employed for improvement of brewing yeasts, especially *S. cerevisiae* where genetic recombination has been well studied. The major problems associated with non-conventional yeasts are (i) poor mating ability; (ii) poor sporulation; (iii) spore viability; (iv) homothallism; (v) aneuploidy; (vi) polyploidy; and (vii)

polygenic control (Tubb, 1985).

Classical hybridization techniques have nevertheless been developed in an attempt to increase chromosome number in pentose fermenting yeasts as an approach to improve these strains. Diploids and polyploids of *P. tannophilus* were constructed by James and Zahab (1982; 1983). The technique for polyploid selection involved prototrophic selection and interruption of the normal sequence of events leading from nuclear fusion to meiosis. Maleszka and co-workers (1983) observed a direct correlation between the increased yield of ethanol produced from xylose by the polyploid strain of *P. tannophilus* and the increased number of chromosomes carried by these strains. In addition, the level of by-product formation such as xylitol decreased. These effects have been attributed to complex physiological phenomena rather than significant changes in the activities of xylose catabolizing enzymes.

1.5.2.2.2 Protoplast fusion

Protoplast fusion is a useful technique for genetic recombination especially for microorganisms that are genetically uncharacterized. In the presence of a fusogenic agent such as polyethylene glycol, protoplasts are induced to fuse and form transient hybrids or diploids. It is presumed that re-assortment of chromosomes which leads to genetic recombination occurs during the hybrid state. Sphaeroplast fusion has been studied by a number of workers in order to improve ethanol production in yeasts (Gupthar, 1987;1992; Gupthar and Garnett, 1987; Johanssen *et al.*, 1985; Maleszka *et al.*, 1983; Pina *et al.*, 1986).

Studies have indicated that this technique may be employed for the construction of yeast strains with new capabilities for utilizing substrates (Tubb, 1985). Johanssen *et al.* (1985) employed this technique for the construction of polyploids of *C. shehatae*. They correlated small increases in ethanol production with increase in ploidy. Attempts to improve ethanol production

by protoplast fusion with strains of *P. stipitis* (Gupthar, 1987) and *C. shehatae* (Johannsen *et al.*, 1985) have, however, met with surprisingly little success.

1.5.2.2.3 Gene cloning, expression and characterization

1.5.2.2.3.1 Recombinant bacteria

Several bacteria, including *Escherischia coli* (Alterthum and Ingram, 1989; Lindsay *et al.*, 1994; Neale *et al.*, 1988; Ohta *et al.*, 1990; 1991a), *Erwinia chrysanthemi* (Tolan and Finn, 1987a), *Klebsiella oxytoca* (Ohta *et al.*, 1991b; Zhang *et al.*, 1995) and *K. planticola* (Feldmann *et al.*, 1989; Tolan and Finn, 1987b) have been candidates for metabolic engineering to render them suitable for the fermentation of xylose. The most promising recombinant bacteria for the fermentation of D-glucose and xylose are *E. coli* and *K. oxytoca.*

E. coli does not normally grow well on xylose. Anaerobic fermentation of sugars normally results in a range of products including lactate, acetate, succinate and formate with ethanol being a minor product (Gottschalk, 1979). Expression of the *PDC* gene from *Z. mobilis* diverts the pyruvate to ethanol. However, accumulation of acetaldehyde often occurs due to insufficient levels of the native ADH activity. *E. chrysanthemi* has a very high ethanol tolerance (up to 4%) and was thus considered a good candidate for recombinant work. Transconjugated *E. chrysanthemi* containing the *PDC* gene from *Z. mobilis* fermented xylose at a higher yield and produced less formate, acetate and lactate than the wild-type (Tolan and Finn, 1987a). The transconjugants, however, grew four times more slowly and tolerated only 2% ethanol.

The introduction of the genes encoding *PDC* and *ADH*II resulted in a more efficient conversion of xylose and other sugars to ethanol (Neale *et al.*, 1988). *E. coli* harbouring plasmid-borne *Z. mobilis PDC* and *ADH*b fermented xylose to ethanol at a yield of 0.49 g/g under optimal

conditions (Ohta *et al.*, 1990). Alterthum and Ingram (1989) obtained volumetric ethanol productivities of 1.4 g/l, 1.3 g/l and 0.64 g/l from glucose, lactose and xylose, respectively, with recombinant *E. coli* harbouring these *Z. mobilis* genes. Ethanol production at an efficiency of 94% of the theoretical maximum from hardwood hydrolysates was obtained by Rousseau and Lawford (1993) with recombinant *E. coli* carrying the plasmids constructed by Alterthum and Ingram (1989). Fermentation of newsprint hydrolysate by the same recombinant *E. coli* strain produced slightly lower yields of ethanol (Lawford and Rousseau, 1991) due to by-product formation. Padukone *et al.* (1995) characterized ethanol production in recombinant *E. coli* harbouring *Z. mobilis PDC* and *ADH*b on cellulose and xylose. Ethanol was a strong inhibitor of both fermentations with a maximum of 50 g/l at 37°C. Recombinant strains of *K. oxytoca* containing the above-mentioned genes on a plasmid were found to ferment xylose to ethanol more efficiently compared to strains reported previously (Ohta *et al.*, 1991b). A maximal volumetric productivity of 2.0 g/l/h was achieved which is double that for recombinant *E. coli*.

1.5.2.2.3.2 Recombinant yeasts

P. stipitis and *C. shehatae* are the best known pentose fermenting yeasts, but the yield and rate of production of ethanol are still not economically viable. The brewer's yeast *S. cerevisiae* is one of the most efficient producers of ethanol but cannot metabolize xylose to ethanol. The advantages of using *S. cerevisiae* as a host for expression of xylose-utilizing genes include its high ethanol tolerance (Brown *et al.*, 1981), its high resistance to inhibitors in lignocellulose-derived hydrolysates (Linden and Hahn-Hägerdal, 1989) and its well-documented use as a role model organism for genetic engineering. Since *S. cerevisiae* can utilize xylulose for ethanol production (Gong *et al.*, 1983; Wang *et al.*, 1980a), the construction of a xylose fermenting strain would involve the expression of either the bacterial gene encoding XI or the two yeast genes encoding

XR and XDH. *S. pombe* is another yeast which would be suitable for expression of xylose-utilizing genes since it has high XK activity (Deng and Ho, 1990).

Previous attempts to enable *S. cerevisiae* to utilize xylose involved cloning the xylose isomerase (*XYLA*) gene from several bacterial sources including *Actinoplanes missouriensis* (Amore and Hollenberg, 1989), *Ampullariella* spp. (Saari *et al.*, 1987), *C. thermohydrosulfuricum* (Dekker *et al.*, 1991), *C. thermosulfurogenes* (Moes *et al.*, 1996), *E. coli* (Batt *et al.*, 1986b; Ho *et al.*, 1983), *Streptomyces griseofuscus* (Kikuchi *et al.*, 1990) and *S. violaceoniger* (Drocourt *et al.*, 1988). The *E. coli XYLA* gene was expressed in *S. cerevisiae* under the control of the yeast *ADH* promoter (Sarthy *et al.*, 1987). Although the gene was translated, no activity was detected. The *XYLA* gene from *B. subtilis* was also expressed in *S. cerevisiae* under the control of two different promoters (*PDC1* and yeast invertase) but produced negligible XI activity (Amore *et al.*, 1989). The polypeptides in both cases were incorrectly folded into protein aggregates. Chan *et al.* (1989) obtained a more favourable result with the *E. coli XYLA* cloned in *S. pombe* but the XI activity and ethanol production were dependent on yeast extract, malt extract and peptone in the medium.

Walfridsson *et al.* (1996) reported the production of an active XI from *Thermus thermophilus* in *S. cerevisiae*. The gene was expressed under control of the yeast *PGK1* promoter. Recombinant XI had highest activity at 85°C. This was the first successful expression of a procaryotic *XYLA* gene in *S. cerevisiae*. The authors attributed the success to the relatedness between the two organisms. Ethanol, acetic acid and xylitol were formed during oxygen-limited fermentation.

Several researchers cloned both the *XYL1* and *XYL2* genes from three strains of *P. stipitis* (CBS5774 - Amore *et al.*, 1991; CBS 6054 - Hallborn *et al*, 1991 and CBS 5773 - Takuma *et al.*, 1991). *S. cerevisiae* transformed with the *XYL1* gene from *P. stipitis* yielded a hybrid strain with

XR activities comparable to that of *P. stipitis*. This strain was able to produce xylitol from xylose if the medium was supplemented with a co-substrate and close to 100% of the xylose was converted to xylitol and secreted into the medium (Hallborn *et al.*, 1991).

XYL2 has been cloned from *P. stipitis* (CBS 5774) and introduced into *S. cerevisiae* together with the *XYL1* gene (Kötter *et al.*, 1990). Both genes were expressed with their own regulatory elements and gave activities comparable with those of native enzymes from *P. stipitis*. The recombinant *S. cerevisiae* was capable of producing ethanol from xylose but with yields and rates of production that were very low. Tantirungkij *et al.* (1993) also co-expressed both *XYL1* and *XYL2* in *S. cerevisiae*. The lower efficiency of xylose metabolism by the recombinant strains was attributed to the fact that they were expressed under the control of their natural promoters which would be less effective than promoters indigenous to *S. cerevisiae*. Chen and Ho (1993) fused the structural *XYL1* gene to the *ADC1* promoter and xylulokinase terminator from *S. cerevisiae* and obtained an approximately 20-fold increase in expression of this gene. Tantirungkij *et al.* (1994a) obtained mutants after ethylmethylsulfonate (EMS) treatment of recombinant *S. cerevisiae* carrying *XYL1* and *XYL2* genes. They selected for rapid growth on xylose and showed that the fastest grower had elevated levels of XDH and XK activities and lower XR activity than the parental strain. Batch fermentation in limited aeration conditions with the mutant produced 1.6 times higher yield and an increased rate of production (2.7 times). These workers reported a further increase in yield but a decrease in rate in fed batch culture (Tantirungkij *et al.* 1994b).

Walfridsson *et al.* (1995) overexpressed the *TKL1* and *TAL1* genes in a *S. cerevisiae* recombinant harbouring the *XYL1* and *XYL2* genes in order to increase the flux through the PPP. This strain showed considerably enhanced growth on xylose compared to a strain harbouring only *XYL1* and *XYL2*. Their results indicated that the TAL level in *S. cerevisiae* was insufficient for

efficient utilization of PPP metabolites. The recombinant strain was capable of simultaneous consumption of glucose and xylose. However, xylose was utilized for growth and xylitol production and not for ethanol production. Decreased aeration resulted in impaired growth and increased xylitol formation.

A xylulokinase mutant was constructed by Stevis and Ho (1989). The *XYL3* gene from *S. cerevisiae* was then cloned by complementation of this mutant strain (Ho and Chang, 1989). Deng and Ho (1990) transformed *S. cerevisiae* with a high copy number plasmid carrying the *XYL3* gene under the control of the *TRP*5 promoter. They reported improved xylulose fermentation by the strain over-expressing XK (up to 20 fold). The recombinant strain fermented xylulose twice as fast as the parent strain.

Recently Ho *et al.* (1998) reported the construction of a recombinant *Saccharomyces* yeast capable of effective fermentation of glucose and xylose. The recombinant plasmids carried the *P. stipitis XYL1* and *XYL2* genes under the control of the yeast *ADC1* and pyruvate kinase (*PYK*) promoters, respectively, as well as the *XK* gene from *S. cerevisiae* under its own promoter control. The recombinant yeast effectively fermented xylose to ethanol and also used it for aerobic growth. The recombinant strain co-fermented most of the 8% glucose and 4% xylose to ethanol in 48 h whereas the parent strain could ferment glucose to ethanol but could not ferment xylose.

Attempts were also made to enhance the xylose fermenting capability of *P. stipitis* using recombinant DNA techniques. Ho *et al.* (1991) developed a plasmid-mediated transformation system for *P. stipitis*. Plasmid vectors containing the *S. cerevisiae* 2 micron (2μ) replicon and kanamycin resistance gene were introduced into *P. stipitis* and maintained as extrachromosomal elements. Selection for transformants was based on resistance to geneticin. *P. stipitis* was transformed by the $CaCl_2$-polyethylene glycol protoplast process or by direct electroporation.

Geneticin resistance was, however, not the ideal selection marker since numerous non-transformants also became resistant to geneticin.

Yang *et al.* (1994) developed a high efficiency transformation system for *P. stipitis* which included integrating and autonomously replicating plasmids. They obtained *ura* auxotrophs by selecting for resistance to 5-fluoroorotic acid which they identified as *ura3* mutants by transformation with *P. stipitis URA3*. The vectors contained an autonomously replicating sequence element (ARS2) and the *P. stipitis URA3* gene. Circular plasmids produced 600 - 8 600 URA^+ transformants/μg DNA by electroporation. Transformants contained ten copies of plasmid/cell. Vectors linearized within the *URA3* gene produced 5 200 - 12 500 transformants that were stable integrants. Integration was site specific for the genomic *URA3* in 20% of the transformants examined. *P. stipitis ura* auxotrophs transformed with the above-mentioned vectors containing the *P. stipitis XYL1* and *XYL2* genes did not produce more ethanol than did the parent strain (Jeffries - personal communication). Shi *et al.* (1999) disrupted the cytochrome c gene in *P. stipitis*. The mutant had a 50% lower growth rate on fermentable sugars but produced 21% higher ethanol yield than the parent strain (0.46 g/g ethanol from 8% xylose compared to 0.38 g/g).

Ideally, a xylose fermenting microorganism should have the following properties: (i) ability to use glucose and xylose simultaneously; (ii) ability to ferment glucose and xylose at equally high fermentation rates; (iii) absence of by-product formation; (iv) ability to ferment xylose aerobically; (v) no cell growth requirement during fermentation; (vi) no inhibition by ethanol produced or by acetic acid and other inhibitors present in lignocellulose-derived hydrolysates; (vii) ability to ferment at a low pH; and (viii) ability to ferment at high temperatures (Hahn-Hägerdal *et al.*, 1993). Recombinant bacteria are capable of fermenting xylose and glucose at the same rate but not simultaneously. Results from *K. oxytoca* in a well controlled system on a laboratory scale

showed promise. However, when using a lignocellulose-derived hydrolysate on an industrial scale as a substrate, inhibition of xylose utilization by glucose and the presence of acetic acid present must be taken into account. The use of high pH for bacteria also renders the fermentation more prone to contamination. The use of bacteria could also constitute a health hazard as microorganisms such as *K. oxytoca* are potential pathogens.

A recombinant strain of *S. cerevisiae* with good xylose fermenting properties in a lignocellulose-derived hydrolysate is desirable since it tolerates high ethanol concentrations and acetic acid. Other advantages for industrial purposes are that the yeast is easy to recycle, the risk of contamination is minimized because of fermentation at low pH, and it is a non-pathogenic microorganism. Results obtained to date with recombinant yeasts are not as promising as those for recombinant bacteria. Limitations encountered with recombinant *S. cerevisiae* include the non-specific transport of xylose through the membrane which might limit the fermentation rate, and/or the redox imbalance may be more pronounced than in natural xylose fermenting yeasts.

Cloning of a transport system with a higher affinity for xylose may enhance xylose transport (Kötter and Ciriacy, 1993). Changing the cofactor preference of XR by site-directed mutagenesis may overcome the redox imbalance during xylose fermentation . The disruption of the *ADH2* gene in *S. cerevisiae* or its equivalent gene in *P. stipitis* and other xylose fermenting yeasts might prevent oxidation of ethanol to acetaldehyde and acetic acid, thus increasing ethanol yield and ethanol tolerance (Hahn-Hägerdal *et al.*, 1993). Expressing the *P. stipitis XYL3* gene or overexpressing the *S. cerevisiae XYL3* gene may relieve the bottleneck in the flow of metabolites in the PPP.

1.6 SCOPE OF THIS STUDY

Previous attempts to genetically manipulate *S. cerevisiae* for ethanol production from

xylose were not successful (Kötter *et al.*, 1990; Tantirungkij *et al.*, 1993; Walfridsson *et al.*, 1997b). There are two possible reasons for these failures. Firstly, the *XYL1* and *XYL2* genes were expressed under the control of their own promoters and their expression in the heterologous host was constitutive. Secondly, all the experiments carried out thus far involved the expression of only the first two genes in xylose metabolism. It has been speculated that the *XYL3* gene is needed to "pull" xylose metabolism towards ethanol production (Deng and Ho, 1990; Kötter and Ciriacy, 1993).

The objective of this study, therefore, was to clone the xylose-metabolizing genes (*XYL1*, *XYL2* and *XYL3*) from *C. shehatae* and *P. stipitis* and to express them in *S. cerevisiae* in an attempt to enable this strain to ferment xylose more efficiently. Several cloning techniques including PCR cloning and the construction and screening of genomic libraries using heterologous gene probes will be attempted. Once these genes are isolated they will be fused to highly effective glycolytic promoters, viz., the inducible *ADH2* promoter and the *PGK1* promoter.

Activity of the heterologous proteins in their host will be confirmed by enzyme assays and fermentation experiments. The strategy of expressing the xylose metabolizing genes under the control of glycolytic and constitutive promoters should result in the expression of these genes without xylose induction. Furthermore, a recombinant *S. cerevisiae* strain harbouring all three genes should prove more efficient at fermenting xylose than the recombinant *S. cerevisiae* carrying only the first two genes in xylose metabolism.

CHAPTER TWO

CLONING OF XYLOSE-METABOLIZING GENES

2.1 INTRODUCTION

Improvement in product yield or process efficiency in industrial processes involving microorganisms can be achieved by genetic manipulation of the producing organisms. A wide range of techniques are available for yeast strain development, including recombinant DNA technology, rare mating and protoplast fusion. Recombinant DNA technology involves the cloning of specific genes of interest and transformation of the industrial strains with these genes.

The discovery of plasmids triggered the development of cloning procedures, making it possible for a defined fragment of DNA to be amplified and purified in milligram quantities (Crabb *et al.*, 1991). The method employed in cloning an individual gene is determined by the type of gene, the source of the DNA, what is already known about the gene (or gene product) and the focus of the cloning project. There are two steps to cloning: (i) physical cloning of the gene; and (ii) identification of clones of interest. Two approaches exist for cloning: (i) the production of DNA libraries; or (ii) direct cloning of specific DNA fragments. The ability to identify appropriate clones determines the cloning approach which includes (i) cloning by functional gene activity [Ratzkin and Carbon, 1977]; (ii) cloning by DNA homology [Grunstein and Hogness, 1975]; (iii) cloning by amino acid sequence data; (iv) cloning by linkage groups; and (v) cloning of cDNA [Hicks *et al.*, 1982].

DNA libraries are constructed by restricting genomic DNA and ligating the resultant fragments to a cloning vector. The vector system employed is either plasmid- or phage-derived, depending on the size of fragments generated. Fragments greater than 10 kb are not easily cloned into plasmid vectors. For fragments greater than 10 kb, a lambda (λ) or cosmid vector is the only

viable choice (Collins and Hohn, 1978). A second factor to consider would be the number of potential clones to be screened. Clark and Carbon (1976) derived a statistical formula to calculate the number of clones required to give a specific probability of finding the clone of interest in a library (Section 2.2.11). Because of the high packaging efficiency of λ and the ease of screening large numbers of plaques, λ vectors are often the first choice for library preparation from genomic DNA. Also, several λ-derived vectors are commercially available in the form of cloning kits.

Cloning by functional activity is limited to those cases where the gene is likely to be expressed in the host organism and the expression of the gene imparts some distinguishable characteristic to the host cell. This technique was used by Ho and Chang (1989) and Stevis *et al.* (1987) to clone the xylulokinase gene from *S. cerevisiae* and *P. tannophilus*, respectively, by complementation of *E. coli* and yeast mutations.

Grunstein and Hogness (1975) used DNA hybridization to develop a method for screening recombinant clones in plasmids. Cloning by DNA homology will only be possible if that gene is homologous to a previously cloned gene from another organism or if a synthetic oligonucleotide probe is generated based on the amino acid sequence of the cloned gene product. In some cases, neither a functional assay nor a known DNA sequence is available. If the amino acid sequence of the protein is known it is possible to synthesize a family of putative DNA sequences for the gene. Because of the degeneracy of the genetic code it is not possible to "back translate" an amino acid sequence into an unequivocal DNA sequence. Therefore, to guarantee that a probe will match the sequence, a pool of potential oligonucleotide probes must be synthesized which include all possible codon combinations.

Cloning by linkage groups is possible in organisms with well-defined genetic systems. Many genes may be expressed, but do not impart a directly selectable characteristic to the host organism. If previous mapping data show that the gene is linked to a selectable marker, an

approach may be designed which screens initially for cloned fragments carrying the linked selectable marker and subsequently for clones also carrying the non-selectable gene. The development of transposon systems has made it possible to generate strains with selectable markers randomly placed around the genome (Crabb *et al.*, 1991).

In prokaryotes, the functional transcript from which the protein is produced is, in most cases, the direct transcription product of the gene. In eukaryotes, the genomic DNA sequence has additional non-coding regions (introns) which are removed during maturation of the mRNA. In some cases, for expression of an active protein in heterologous hosts, genomic DNA clones are inappropriate and sequences based on the mature mRNA are required (Crabb *et al.*, 1991). In this case, total RNA is isolated and complementary DNA is synthesized based on the sequence of mRNA. A cDNA library can be generated and clones detected as described previously.

Several other cloning tools exist - the use of transposons, the use of integrative vectors and chromosome walking. Another technique that has become a popular tool is a polymerase chain reaction (PCR)-based method for the direct cloning of genomic sequences (Scharf *et al.*, 1986). Specific or degenerate primers can be used to amplify genomic DNA. PCR primers are synthesized from sequences derived from a gene of interest in a different organism that has already been cloned and sequenced. Cloning of the amplified product is facilitated by the use of special PCR-cloning vectors that are commercially available.

This chapter describes the following: (i) the cloning of the *XYL1* and *XYL2* genes from *P. stipitis* and the *XYL3* gene from *S. cerevisiae* using specific PCR primers; (ii) the construction of a genomic library for *C. shehatae*; (iii) the cloning of the *C. shehatae XYL1* gene; (iv) subcloning; and (v) sequencing of these four genes.

2.2 MATERIALS AND METHODS

2.2.1 GROWTH AND MAINTENANCE OF CULTURES

Yeast and bacterial strains used in the study are listed in Table 2.1. Yeast strains were grown at 30°C for two days on YPD medium (10 g yeast extract, 20 g peptone, 20 g glucose and 15 g agar per litre). Cultures were stored at 4°C and sub-cultured every four to six weeks. For long term storage, cultures were grown in YPD broth at 30°C and mixtures containing equal volumes of broth culture and sterile 30% glycerol were snap frozen in liquid nitrogen and stored at -70°C. Recombinant yeast strains were grown on synthetic complete (SC) medium:6.7 g yeast nitrogen base - lacking amino acids [Difco] and 15 g agar per litre supplemented with 20 g galactose or 20 g xylose or 18 g xylose + 2 g galactose as carbon source. For long term storage of recombinant strains, cultures were grown in SC broth and treated as described above.

E. coli strains were grown on YT agar (8 g tryptone, 5 g yeast extract, 5 g NaCl and 15 g agar per litre) at 37°C. Exponential phase broth cultures were used to prepare stocks as described above. Recombinant *E. coli* strains were selected on YT agar supplemented with ampicillin (100 μg/ml), 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal - 50 μg/ml) and isopropylthio-β-O-galactoside (IPTG - 1 mM).

2.2.2. ISOLATION OF GENOMIC DNA

Genomic DNA was isolated from *P. stipitis*, *C. shehatae* and *S. cerevisiae* using standard protocols (Rose *et al.*, 1990). Ten ml aliquots of YPD broth were inoculated with a colony of each culture, incubated overnight (O/N) at 30°C. Cells were pelleted by centrifugation and resuspended in 0.5 ml distilled water (dH_2O). The cell suspension was transferred to Eppendorf tubes and pelleted (15 800 × *g* for 20 s) in a microcentrifuge. The supernatant was decanted

TABLE 2.1 Yeast and bacterial strains used in this study

Strains	Characteristics/Genotype	Source
Yeasts		
Pichia stipitis		
PsY633	wild type	CSIR[a]
Candida shehatae		
CsY117A/1	wild type	CSIR[a]
CsY492	wild type	CSIR[a]
Cs051	wild type	T.W. Jeffries[b]
Saccharomyces cerevisiae		
Y294	α *leu*2-3,112,*ura*3-52 *his*3,*trp*1-289	W.H. van Zyl[c]
GPY55-15Bα	*leu*2-3,*leu*2-112,*ura*3-52,*trp*1-289,*his*4-519,*prb*1 *cir*°	B. Hahn-Hagerdal[d]
Bacteria		
Escherichia coli		
DH5αF′	*supE*44,Δ*lac*U169(ϕ80*lacZ*ΔM15)*hsdR*17,*recA*1,*endA*1, *gyrA*96,*thi*-1*relA*1	BRL[e]
VCS257	derivative of DP50*supF*	Stratagene[f]
SOLR	e14°(*mcr*A) 183, Δ(*mcr*CB-*hsf*dSMR-*mrr*)171, *sbc*C, *rec*B, *recJ*, *umu*C::Tn5(kan^r), *uvr*C, *lac*, *gyr*A96, *rel*A1, *thi*-1, *end*A1, λ^R, [F' *pro*AB, *lac*I^{q}ZΔM15] Su- (nonsuppressing)	Stratagene[f]
XL1 Blue MRF′	Δ(*mcr*A)183, Δ(*mcr*CB-*hsd*SMR-*mrr*)173, *end*A1, *sup*E44, *thi*-1, *rec*A1, *gyr*A96, *rel*A1, *lac*[F' *pro*AB, *lac*I^{q}ZΔM15, Tn10(tet^r)]	Stratagene[f]

[a] CSIR: Council for Scientific and Industrial Research, Pretoria, RSA.
[b] T.W. Jeffries, Institute for Microbial and Biochemical Technology, Madison, Wisconsin, USA.
[c] W.H. van Zyl, Department of Microbiology, University of Stellenbosch, Stellenbosch, RSA.
[d] B. Hahn-Hagerdal, Department of Applied Microbiology, Lund University, Sweden.
[e] BRL: Bethedsa Research Laboratories, Bethedsa, USA.
[f] Stratagene: La Jolla, California, USA.

and the tube vortexed to resuspend the pellets in the residual liquid. Two hundred μl of lysis buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris.HCl - pH 8.0, 1 mM EDTA), 200 μl of phenol:chloroform:isoamyl alcohol (25:24:1) and 300 mg acid- washed glass beads (425-600

μm) were added to each tube and the tubes were vortexed for 3 min. Two hundred μl TE buffer (10 mM Tris.HCl, pH 8.0; 1 mM EDTA) was added to each tube. The tubes were centrifuged for 5 min at 15 800 × *g* and the aqueous phase transferred to a new tube. Two phenol-chloroform extractions were performed to remove residual proteins by inverting the tubes several times, centrifuging and removing the upper aqueous layer. Genomic DNA was precipitated by the addition of two volumes of 100% ethanol. The tubes were centrifuged for 2 min and the pellets resuspended in 400 μl TE buffer containing 30 μg RNaseA and incubated at 37°C for 5 min. Genomic DNA was re-precipitated by adding 10 μl of 4 M ammonium acetate and 1 ml of 100% ethanol followed by centrifugation. The pellet was washed in 70% ethanol, air-dried and resuspended in 50 μl TE buffer.

2.2.3 POLYMERASE CHAIN REACTION AMPLIFICATION

2.2.3.1 Xylose reductase genes

PCR primers derived from the sequence of the *P. stipitis XYL1* gene (Amore *et al.*,1991) were designed in order to clone the corresponding gene from the genomic DNA of *P. stipitis* strain Y633 and *C. shehatae* strains CsY492 and CsY117A/1. The primer sequences included a site for the restriction enzyme, *Xho*I (CTC GAG), in order to facilitate the manipulation of the amplification product in *E. coli* and *S. cerevisiae*.

PCR primers, based on sequencing data of the genomic clone, were designed in order to amplify only the coding region of the *C. shehatae XYL1* gene. Primers included restriction sequences sites for the enzymes *Eco*RI (GAA TTC) and *Xho*I in order to facilitate cloning. Primers are listed in Table 2.2.

2.2.3.2 Xylitol dehydrogenase genes

The *XYL2* gene from *P. stipitis* was isolated by PCR amplification of genomic DNA. PCR primers specific for the *P. stipitis XYL2* gene were based on the sequence of the *P. stipitis XYL2* gene published by Kötter *et al.* (1990). Sequences for the restriction enzymes *Eco*RI and *Bgl*II (AGA TCT) were included in the primer sequence to facilitate cloning in *E. coli* and *S. cerevisiae*. PCR primers are listed in Table 2.2.

2.2.3.3 Xylulokinase gene

PCR primers were derived from the sequence of the *S. cerevisiae XYL3* gene obtained from a BLAST search of patents published under the patent cooperation treaty (International publication number: WO 95/13362). Primer sequences included sites for the restriction enzymes *Eco*RI and *Xho*I. PCR primers are listed in Table 2.2.

TABLE 2.2 PCR primers used in this study

Organism	Gene	Primers
C. shehatae	*XYL1*	Forward: 5' GATC*GAATTC*ATGAGCCCAAGCCCAATT 3'
		Reverse: 5' CTAG*CTCGAG*CAAGCACGCTTAAACGA 3'
P. stipitis	*XYL1*	Forward: 5' GACT*CTCGAG*CAATGCCTTCTATTAAGTTGAACTC 3'
		Reverse: 5' CTAG*CTCGAG*TTTCCTCTCTATAAAGCAACCTTC 3'
	XYL2	Forward: 5' GATC*GAATTC*ATGACTGCTAACCCTTCCTTGG 3'
		Reverse: 5' GACT*AGATCT*GCGGTTGACTTACTCAG 3'
S. cerevisiae	*XYL3*	Forward: 5' GATC*GAATTC*ATGTTGTGTTCAGTAAT 3'
		Reverse: 5' CTAG*CTCGAG*TTTCCAGTTCGCTTA 3'

DNA was amplified in 50 μl reaction mixtures. Each reaction contained 100 ng template DNA, 0.25 μM each primer, 500 μM each dNTP, 1 × *TaKaRa Taq* buffer, 2.5 U *TaKaRa Taq*

polymerase (TaKaRa Biomedicals) [for amplification of *C. shehatae XYL1* and *S. cerevisiae XYL3*] with a Hybaid OmniGene Thermal Cycler or 1 × *Taq* buffer supplemented with 0 - 3.75 mM $MgCl_2$ and 2.5 U *Taq* DNA polymerase (Boehringer Mannheim) [for the amplification of *P. stipitis XYL1* and *XYL2*] in a Biometra Trio Thermoblock TB1 (Biometra Biomedizinische Analytik). Thirty cycles of denaturation, annealing and primer extension were carried out for 1 min at 94°C, 55°C and 72°C, respectively.

PCR reaction mixtures were electrophoresed on 0.7% agarose gels in 1 × TAE buffer (40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA). PCR products were excised from the gels and purified with the QIAEX II Gel Extraction Kit (Qiagen) according to the manufacturer's instructions.

2.2.4 DIGESTION OF PCR PRODUCTS AND pBLUESCRIPT

Purified PCR products were digested with the appropriate restriction endonucleases and cloned into the plasmid cloning vector, pBluescript (pBS) cleaved with the corresponding restriction endonucleases. A modified pBS vector (*Hin*cII site replaced by *Bgl*II) obtained from W.H. van Zyl (Department of Microbiology, University of Stellenbosch, Stellenbosch), was used for the cloning of the *XYL2* genes.

All restriction endonucleases and their corresponding buffers were obtained from Boehringer Mannheim. A typical restriction reaction contained the following components in a total volume of 15 μl: ≈500 ng DNA, 1 × SuRE/Cut buffer and 5 U restriction endonuclease(s). Reactions were carried out at 37°C, unless otherwise stated by the manufacturer, for 60-90 min. When double digests were performed, the appropriate SuRE/Cut buffer was selected to enable both enzymes to digest at least 75% of the DNA within 1 h. If this was not possible, 2 h digestions were performed by cleaving in the lower salt buffer for 1 h, and adding the required

amounts of NaCl and Tris.HCl to ensure that reaction conditions for the second enzyme were optimal. When the vector was linearized with a single restriction enzyme, dephosphorylation of the 5′ ends was performed to prevent subsequent re-circularisation of the vector. For the dephosphorylation reaction, one unit of calf intestinal phosphatase (Boehringer Mannheim) was added to the restriction reaction mixture and incubated at 37°C for 1 h. Restriction digests of PCR products and pBS were electrophoresed and purified as described in Section 2.2.2.3.

2.2.5 LIGATIONS

PCR products were ligated to the linearized vector using standard protocols (Sambrook *et al.*, 1989). A typical ligation involving DNA species with sticky ends comprised the following in a total volume of 10 μl: 50 fmol vector and insert DNA, 1 × ligation buffer and 1 U T4 DNA ligase (Boehringer Mannheim). Ligations were carried out either at room temperature (RT) for 4 h or at 15°C O/N. The *C. shehatae XYL1* PCR product was purified, polished and ligated to the PCR-Script vector (PCR Cloning Kit - Stratagene) as described by the manufacturer.

2.2.6 TRANSFORMATION

An aliquot of the ligation mixtures was transformed into $CaCl_2$- competent or electro-competent *E. coli* cells.

2.2.6.1 Preparation of competent and electrocompetent cells

For the preparation of competent cells, 1 ml of an O/N culture of *E. coli* DH5αF′ was added to 29 ml of YT medium and incubated with shaking until the cell density reached 0.375 absorbance units at 590 nm. Cells were immediately placed on ice and kept cold in all subsequent steps. They were pelleted by centrifugation for 10 min at 10 000 × *g* at 4°C. The pellet was

gently washed in 10 ml cold $CaCl_2$ (50 mM $CaCl_2$, 10 mM Tris.HCl - pH 8.0), resuspended in 10 ml $CaCl_2$ solution and incubated on ice for 20 min and centrifuged. Cells were finally resuspended in 2.5 ml cold $CaCl_2$-glycerol solution (50 mM $CaCl_2$, 10 mM Tris.HCl - pH 8.0, 15% glycerol). One hundred μl aliquots were dispensed into Eppendorf tubes and incubated at 4°C for 24 h to increase the competency of the cells (Ausubel *et al.*, 1989), snap-frozen and stored at -70°C.

For electrocompetent cell preparation, 3 ml of an O/N culture of *E. coli* DH5αF′ was added to 100 ml of YT medium and incubated with shaking until the cell density reached 0.5 absorbance units at 600 nm. The culture was immediately placed on ice to stop further growth. Cells were pelleted by centrifugation for 10 min at 10 000 × *g* at 4°C, gently resuspended in 100 ml cold 10% glycerol and centrifuged. The pellet was resuspended successively in 50 ml, 5 ml and 2.5 ml cold 10% glycerol. Fifty μl aliquots were dispensed into Eppendorf tubes, snap-frozen and stored at -70°C.

2.2.6.2 Transformation and electroporation

Transformation was carried out by adding 3 μl of the ligation mixture to 100 μl competent cells, followed by gentle mixing and incubation for 30 min on ice to allow the DNA to bind to the cells. The cells were heat-shocked by incubating in a water bath at 42°C for 2 min and immediately placed on ice for 15 min. A 900 μl aliquot of YT broth was added to the tubes followed by an incubation for 1 h at 37°C. One hundred μl of this mixture was plated onto YT agar plates supplemented with ampicillin, X-Gal and IPTG and incubated O/N at 37°C.

One μl of the ligation mixture was diluted ten-fold in distilled water and 2 μl were electroporated into 40 μl electro-competent cells in 0.1 cm cuvettes using a Biorad Gene Pulser. The Gene Pulser was set at 25 μF and 1.8 kV and the pulse contoller at 200 Ω. Immediately after electroporation, 900 μl of SOC medium (20 g tryptone, 5 g yeast extract, 20 mM glucose, 0.5 g

NaCl, 2.5 mM KCl, 10 mM $MgCl_2$) was added to the cuvettes and the cells incubated for 1 h at 37°C and plated out as described above.

2.2.7 PLASMID DNA ISOLATION

Transformants containing recombinant plasmids were selected by blue-white colony screening (Sambrook *et al.*, 1989). Plasmid DNA from these colonies was isolated using a modified version of the alkaline lysis method of Birnboim and Doly (1979).

A 1.5 ml aliquot of the bacterial culture was pelleted in an Eppendorf tube by centrifugation at 15 800 × *g* for 1 min. Pellets were resuspended in 100 μl solution A (25 mM Tris.HCl, 50 mM glucose, 10 mM EDTA, 100 μg/ml RNase). Two hundred μl solution B (0.2 M NaOH, 1% SDS) was added, the tubes inverted several times to mix the contents and incubated at RT for 5 min. During this step cell lysis, DNA denaturation and degradation of RNA occurs. Denatured chromosomal DNA and proteins were precipitated by the addition of 150 μl of cold solution C (3 M sodium acetate, pH 4.8) and incubation on ice for 5 min. Cell debris was pelleted by centrifugation for 10 min at 15 800 × *g*. The supernatant was transferred to a clean tube and plasmid DNA precipitated by the addition of two volumes cold 100% ethanol and incubation at -70°C for 15 min. Following centrifugation, DNA was washed with 70% ethanol, air dried and resuspended in 20 μl TE buffer.

The presence of cloned genes was verified by restriction endonuclease digestion and subsequent electrophoresis as described in Sections 2.2.4 and 2.2.3, respectively. For automated DNA sequencing, plasmid DNA was isolated using the High Pure Plasmid DNA Isolation Kit (Boehringer Mannheim) according to the maufacturer's instructions.

2.2.8 SUBCLONING

The recombinant plasmids containing cloned xylose-metabolizing genes were digested with several enzymes in order to construct restriction maps so that appropriate subclones could be constructed for sequencing. Two techniques were used to generate subclones: (i) dropout cloning - selected fragments were "dropped" out of the recombinant plasmid by appropriate restriction digests. The restriction sites were then filled-in (polished or blunt-ended) with Klenow enzyme and the blunt ends ligated to re-circularize the vector; and (ii) by isolating the fragments of interest after restriction digestion and ligating to linearized pBS. DNA fragments were polished in order to produce blunt ends. The reaction was carried out at 37°C for 30 min in a mixture containing the DNA reactants, 1 × restriction buffer, 0.05 mM dNTPs and 2 units of Klenow enzyme (Boehringer Mannheim). The enzyme was subsequently denatured by heating for 15 min at 65°C .

2.2.9 SEQUENCING

2.2.9.1 Manual sequencing

The nucleotide sequence of the *P. stipitis XYL1* and *XYL2* genes were determined by sequencing both strands of the subcloned inserts by the radioactive dideoxy chain termination method (Sanger *et al.*, 1977) using the T7 Sequencing Kit (Pharmacia Biotech Inc.). Annealing of sequencing primers to the templates and labelling and termination reactions were performed according to the manufacturer's instructions. DNA was radioactively labelled with [α^{32}P]dATP (> 1 000 Ci/mmol at 10 mCi/ml - Amersham International). Chain-terminated reaction products were loaded in 6% polyacrylamide sequencing gels (per 100 ml: 42 g urea; 10 ml 10 × TBE buffer; 15 ml 40% acrylamide solution [19:1 acrylamide:bisacrylamide]; 42.5 ml H_2O; 60 µl TEMED; 800 µl of 10% $(NH_4)_2SO_4$). Gels (0.35 mm) were formed between 40 cm × 45 cm

Sigmacote (Sigma)-treated glass plates and pre-run at 800 V at a maximum current strength of 40 mA in 1 × TBE buffer (45 mM Tris-borate, 1 mM EDTA). Electrophoresis was carried out at 55 W constant power in a Fisher Biotech Adjustable Sequencing Electrophoresis System (Fisher Scientific). Long and short runs were performed at 55 W for 5 and 2 h, respectively. Following electrophoresis, the gel was fixed by spraying with 10% acetic acid/10% methanol, transferred to a supporting sheet of filter paper (Whatman number 1) and dried for 45 min under a vacuum at 80°C in a gel dryer (Model 583 Gel Dryer, Bio-Rad Laboratories Inc.). Dried gels were exposed to X-ray film (Hyperfilm-MP, Amersham International) O/N at RT and the films were developed according to the manufacturer's instructions.

2.2.9.2 Automated sequencing

Automated sequencing was performed using the ALFexpress Sequencer (Amersham Pharmacia Biotech) at the DNA Sequencing Laboratory, Department of Microbiology, University of Cape Town. Chain terminating sequencing reactions (Sanger *et al.*, 1977) were performed using the Thermo Sequenase Fluorescent Labelled Primer Cycle Sequencing Kit (Amersham Pharmacia Biotech) and end-labelled fluorescent primers. All reactions were performed according to the manufacturer's instructions.

2.2.9.3 Processing of sequencing data

DNA sequencing data was initially processed at the DNA Sequencing Laboratory (University of Cape Town) using ALFwin version 1.1 software (Amersham Pharmacia Biotech). Raw sequencing data was then manually edited to remove DNA sequences adjacent to the sequencing primers that corresponded to the cloning vectors. Assembly of contiguous DNA sequences, identification of ORFs and translations into corresponding protein sequences were

performed using Vector NTI 5.0 (InforMax Inc.) and DNAmend 1.01 (Piet Jonas and Björn Maul, University of Greifswald, Germany). Similarity searches to DNA and protein databases were performed using the BLAST 2.0 algorithm (Altschul *et al.*, 1997). Alignments of DNA and protein sequences to homologs were performed using the CLUSTALW programme on the GenomeNet server. Protein alignments were edited and analysed using Genedoc version 2.4 (Nicholas and Nicholas, 1997).

2.2.10 CONSTRUCTION OF A GENOMIC LIBRARY

Construction of a genomic library involved isolation of genomic DNA, generation of appropriately-sized genomic fragments, their ligation to pre-digested Lambda Zap II vector arms and packaging of the ligation products into the λ capsid proteins and infection of an appropriate host strain to amplify the λ DNA. The preparation of genomic fragments and their ligation to the lambda vector is described in detail. The packaging reaction using the Gigapack II Packaging Extract kit (Stratagene) and subsequent titration and amplification of the library is described in the manufacturer's instruction booklet.

2.2.10.1 Partial digestion of genomic DNA

Genomic DNA was isolated from *C. shehatae* CS051 as described in Section 2.2.2. Thirty μg of genomic DNA was diluted in 900 μl 10 mM Tris.HCl (pH 8.0). One hundred μl of restriction buffer 4 (Promega) was added to the DNA which was carefully mixed and refrigerated for several hours to ensure proper distribution of buffer. Sixty μl of the DNA solution was added to an Eppendorf tube and 30 μl to another eight tubes. Two and a half units of restriction enzyme *Tsp*509 I (Promega) were added to the first tube and carefully mixed with the pipette tip. The enzyme was serially diluted in the nine tubes by sequential transfer of 30 μl of solution into

successive tubes. Tubes were incubated at 65°C for 1 h and the enzyme activity stopped by the addition of 3 μl of 0.5 M EDTA. An aliquot of each restriction reaction was electrophoresed in a 0.7% agarose gel to visualize the degree of cleavage. Fragments produced under these conditions were predominantly under 3 kb in size. A second set of restrictions was then set up as described above except that 60 μg of DNA was used and the restriction reactions were allowed to proceed for only 5 min before being stopped. An aliquot of the restriction products was electrophoresed as above and tubes containing DNA fragments within the 5 - 10 kb range were identified, their contents electrophoresed and the gel slices containing DNA fragments in this size range excised. Genomic DNA fragments were purified using the Geneclean Kit (BIO 101) according to the manufacturer's instructions. DNA concentration was determined using the Genequant RNA/DNA Calculator (Pharmacia).

2.2.10.2 Ligation of genomic fragments to pre-digested lambda arms

Two ligation reactions were set up: a test ligation using test insert supplied in the kit and a ligation using the isolated genomic fragments as inserts. Each reaction mixture consisted of 1.0 μl of λ Zap II prepared arms (1 μg), 0.5 μl of 10 × ligase buffer, 0.5 μl of 10 mM rATP (pH 7.5), 1 μl T4 DNA ligase, 1.6 μl test insert (0.4 μg) or 2.5 μl DNA fragments (≈250 ng) and H_2O to a final volume of 5 μl. The ligation reactions were incubated at 14°C O/N. To prevent ligation of multiple copies of the inserts, an equimolar ratio of genomic DNA and λ arms was used in the ligation reaction.

2.2.10.3 Packaging and titration of recombinant lambda DNA

The packaging efficiency was calculated by counting the number of plaques formed and applying the following formula:

$$\frac{\text{Number of plaques} \times \text{dilution factor} \times \text{total packaging volume (1 000 } \mu\text{l)}}{\text{Number of micrograms packaged} \times \text{number of microlitres plated}}$$

An efficient packaging reaction usually yields approximately 400 plaques on the 10^{-4} dilution plate according to the Gigapack II Plus Packaging Extract Instruction Manual. In this study, 380 plaques were obtained for this dilution. The efficiency of the packaging reaction was calculated to be 1.9×10^9 pfu/μg DNA. For the library the packaging efficiency was 1.5×10^6 pfu/μg DNA with 696 white (1.4×10^6 pfu/μg DNA) and 38 blue (7.6×10^4 pfu/μg DNA) plaques. Background plaques should be less than 1.5×10^5 pfu/μg DNA and white plaques should be 10 - 100 fold above background. The ratio of recombinant to blue plaques obtained was 18 which is between the desired ratios. The titre of the library was determined to be 3.8×10^5 pfu/μl.

2.2.11 SCREENING OF THE LIBRARY

In order to screen the library, a suitable host strain was infected and plaque lifts were made which were then screened by Southern hybridisation using the *P. stipitis XYL1* gene as a probe. In order to have a 99% chance of picking up a single-copy gene, approximately 8 600 plaques were screened. This was determined from the equation:

$$N = \ln(1 - P)/ \ln(1 - I/G)$$

where N = number of plaques to be screened, P = probability (99%), I = 7 500 (average insert size) and $G = 14 \times 10^6$ (haploid genome size of *S. cerevisiae*).

2.2.11.1 Labelling of probe

The *P. stipitis XYL1* gene was removed from the recombinant plasmid pBPX1 (Table 3.1) as a *Xho*1 fragment, purified as described previously and labelled using the non-radioactive DNA

labelling and detection kit [DIG] (Boehringer Mannheim).

DNA fragments were denatured in a boiling water bath for 10 min and immediately cooled on ice. A typical labelling reaction had the following components in a total volume of 20 μl: 1 - 10 μg freshly denatured DNA, 2 μl hexanucleotide mixture, 2 μl dNTP labelling mixture and 1 μl Klenow enzyme. These components were added to an Eppendorf tube on ice, mixed well, centrifuged briefly and incubated at 37°C for 20 h. The labelling reaction was stopped by the addition of 2 μl 0.5 M EDTA solution. Labelled DNA was precipitated with 2.5 μl LiCl and 75 μl cold 100% ethanol at -70°C for 30 min. Precipitated DNA was pelleted by centrifugation at 15 800 × *g* for 30 min, washed in 70% ethanol and vacuum dried. The labelled DNA was dissolved in 50 μl TE buffer and stored at -20°C.

2.2.11.2 Primary screen

NZY plates (150 mm) were seeded with 600 μl of OD_{600} = 0.5 XL1-Blue cells infected with 50 000 pfus. Plaques were allowed to form. Plates were then chilled for 2 h at 4°C to prevent the agar from sticking to the nylon membranes. Plaques were transferred onto the membranes (Magna nylon transfer membrane, Micron Separations Inc.) for 2 min. A needle was used to prick through the agar at strategic points for orientation. The membrane was denatured by submerging in 1.5 M NaCl /0.5 M NaOH for 2 min. Excess top agar was washed off at this stage by rubbing between gloved fingers. The membrane was neutralized for 5 min by submerging in 1.5 M NaCl/ 0.5 M Tris.HCl (pH 8.0), then rinsed for 30 s in 0.2 M Tris.HCl (pH 7.5) and 2 × SSC buffer (30 mM NaCl, 300 mM Na-citrate - pH 7.0). Membranes were air dried and the DNA crosslinked to the membranes for 3 min using a UV transilluminator (Wilber Lourmat).

Membranes were washed in 2 × SSC and pre-hybridized for 2 h at 60°C in 20 ml pre-hybridization solution (5 × SSC, 1% blocking reagent, 0.1% Na-lauroyl sarcosine, 0.02% SDS).

Pre-hybridization solution was replaced by 2.5 ml hybridization solution containing ≈30 ng of freshly denatured probe and incubated at 60°C O/N. Following hybridization, membranes were rinsed at RT with 50 ml 2 × SSC/0.1% SDS; washed twice at RT for 5 min in 3 × SSC/0.1% SDS and twice for 15 min at 60°C in 0.5 × SSC/0.1% SDS (low and high stringency washes, respectively).

Detection of hybridized DNA was performed immediately by an enzyme-linked immuno-assay using high-affinity antibody Fab fragments coupled to alkaline phosphatase and visualized using the chromogenic substrate nitroblue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indoyl-phosphate (X-phosphate). After signals appeared, membranes were washed for 5 min with 50 ml buffer 4 (TE buffer - pH 8.0) and stored in the same buffer solution.

2.2.11.3 Secondary screen

Membranes were oriented in relation to the plaque plates by lining up the needle pricks. Positive clones were identified by comparing the intensity of the coloured spots. Membranes were lined up against the master plates and an agar slice corresponding to a positive signal was punched out using a Pasteur pipette. Each slice was placed in an Eppendorf tube with 1 ml SM buffer and 20 μl chloroform and mixed. Clones were diluted and titred with host cells so that each plate would have ≈50 plaques. Plaque lifts were made from chilled plates as described previously. Hybridization and pre-hybridization were performed and membranes were compared to plates and the putative clones displaying intense signals isolated. In most cases, the positive plaques were well separated from the background plaques and a tertiary screen was not necessary.

2.2.11.4 *In vivo* excision of the pBS SK(-) phagemid from the Lambda Zap II vector

In vivo excision was performed using the ExAssist /SOLR system which is designed to allow efficient excision of the pBS SK(-) phagemid from the λ Zap II vector, while preventing problems associated with helper phage co-infection. The ExAssist helper phage contains an amber mutation that prevents replication of the phage genome in a non-suppressing *E. coli* strain such as SOLR cells. This allows only the excised phagemid to replicate in the host, removing the possibility of co-infection from the ExAssist helper phage. The method was followed according to the manufacturer's instructions.

2.2.11.5 Location of *C. shehatae XYL1* gene

Plasmid DNA mini preparations were performed for each clone as described previously (Section 2.2.7). Plasmid DNA was cleaved with *Eco*RI, *Bam*HI, *Hin*dIII, *Not*I, *Bgl*II and *Sty*I and electrophoresed. DNA fragments were transferred to a nylon membrane using a vacuum blotting unit (Omeg Scientific). Two pieces of filter paper (Whatman 4 MM) were cut to exactly the size of the agarose gel and the nylon membrane was cut 4 - 5 mm bigger than the gel. The filter paper and nylon membrane were pre-wet in dH_2O for 15 - 20 min. After placing the filter paper on the porous support of the vacuum blotter, the nylon membrane was centred on the filter paper. The gel was carefully placed on the membrane and parafilm was placed around the gel ensuring that at least 2 mm of the membrane was under the parafilm on all sides which contributed to a good vacuum seal. DNA was transferred from the gel onto the membrane by applying a vacuum of 15 mm Hg. Approximately15 - 20 ml of depurination solution (0.25 M HCl) was poured over the gel and left on for 20 min. This solution breaks the DNA into fragments, thus enhancing the transfer process. The excess depurination solution was removed and transfer solution (0.4 M

NaOH/0.6 M NaCl) was poured onto the gel and left on for 30 min. Both the depurination and transfer solutions were regularly topped up to prevent the gel from drying out. At the end of the procedure, the vacuum pump was turned off and the gel and filter paper were removed. The positions of the wells were marked onto the nylon membrane which was then air dried. DNA was UV cross-linked for 3 min and the membrane used immediately or sealed in a plastic bag and stored at 4°C.

Hybridization and detection of DNA fragments was performed as described in Section 2.2.11.2. The positive clones were selected, plasmid DNA isolated and cleaved with a selection of restriction endonucleases in order to generate subclones of the inserts. These were electrophoresed, blotted and Southern hybrizations were performed to locate the *XYL1* gene on as small a fragment of DNA as possible. Once this was achieved the DNA fragment carrying the gene was cloned and sequenced as described in Section 2.2.9.2.

2.3 RESULTS

2.3.1 *P. stipitis* XYLOSE REDUCTASE GENE

PCR amplification of the *P. stipitis XYL1* gene produced a single PCR product of approximately 1.0 kb (Fig. 2.1). An amplification product was obtained only when $MgCl_2$ was added to the reaction, the highest yield of the product was observed at a $MgCl_2$ concentration of 1.25 mM. The *P. stipitis XYL1* gene was cloned as a *Xho*I fragment into the *Xho*I site of pBS to produce pBPX1 (Table 3.1). This clone was restricted with several enzymes in order to construct a restriction map so that appropriate subclones could be constructed for sequencing. Three subclones were constructed either by dropout cloning (a 430 bp *Eco*RI fragment and a 278 bp *Kpn*I/*Sty*I fragment) or by ligating a fragment of interest (a 406 bp *Sty*I fragment) in pBS. Figure 2.2 shows a partial restriction map of the *XYL1* gene as well as the three DNA fragments

subcloned into pBS for sequencing. The nucleotide sequence of the *XYL1* gene (Appendix One) has 99% homology to the previously reported *P. stipitis XYL1* gene.

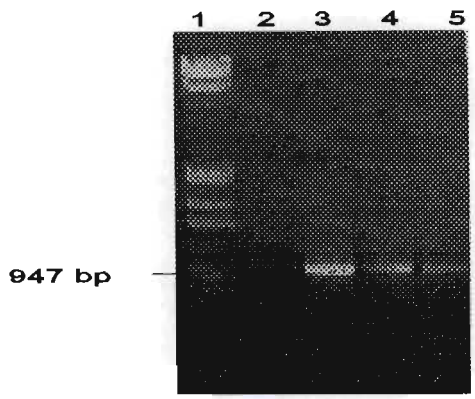


FIG. 2.1 PCR amplification product for the *P. stipitis XYL1* gene using several $MgCl_2$ concentrations. Lane 1: molecular weight marker (λ cleaved with *Eco*RI/*Hin*dIII); lane 2: no $MgCl_2$; lane 3: 1.25 mM $MgCl_2$; lane 4: 2.5 mM $MgCl_2$; lane 5: 3.75 mM $MgCl_2$.

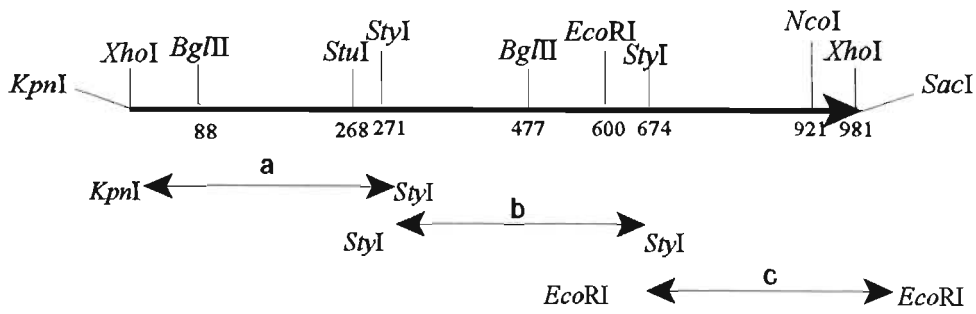


FIG. 2.2 Partial restriction map of the *P. stipitis XYL1* gene showing the three subclones (a, b and c) used for sequencing this gene.

2.3.2 *C. shehatae* XYLOSE REDUCTASE GENE

Primary screening of the genomic library with the *XYL1* gene probe of *P. stipitis* identified approximately 20 putative clones although a high background was produced on some of the membranes. The secondary screen yielded ten putative clones. Restriction profiles of plasmid DNA of the clones were generated by single enzyme digests: *Bam*HI, *Eco*RI, *Hin*dIII, *Not*I, *Sty*I and *Bgl*II (not shown) and double digests: *Eco*RI/*Bam*HI (not shown) and *Nco*I/*Bam*HI (Fig. 2.3). The sizes of the inserts of most of the putative clones were 4 kb and greater. Southern hybridizations performed on the double digests localized the putative gene on DNA fragments of approximately 4.8 kb. The clones in lanes 2 and 5 of Fig. 2.3 show clear positive signals. Clone 1 was subcloned as an *Nco*I/*Not*I and an *Nco*I fragment (Fig. 2.4) and sequenced. Sequencing data revealed that approximately 40 bp of the 5' coding region of the gene were located on the *Nco*I/*Not*I subclone while the rest of the gene was located on the *Nco*I fragment. PCR primers, based on the sequencing data of the genomic clone, were designed in order to amplify and clone only the coding sequence of the gene. Primers contained restriction sites to facilitate cloning.

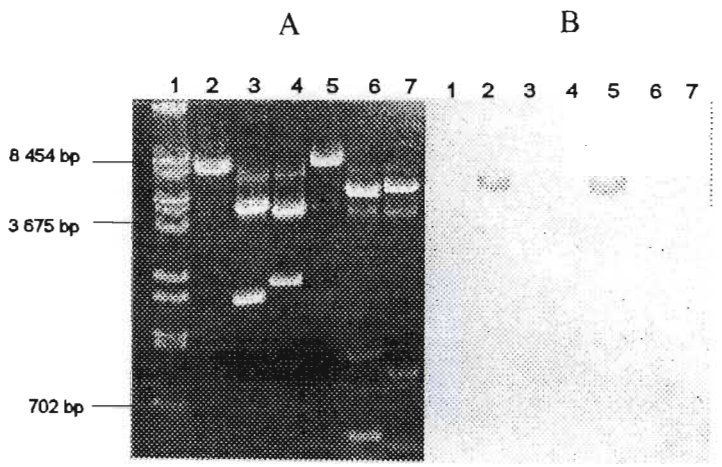


FIG. 2.3 A: *Eco*RI/*Bam*HI double digest of putative clones. B: Southern blot of restriction digest. Lane 1: Molecular weight marker (λ DNA cleaved with *Eco*RI/*Hin*dIII); lanes 2-7: putative *C. shehatae XYL1* genomic clones 1-6.

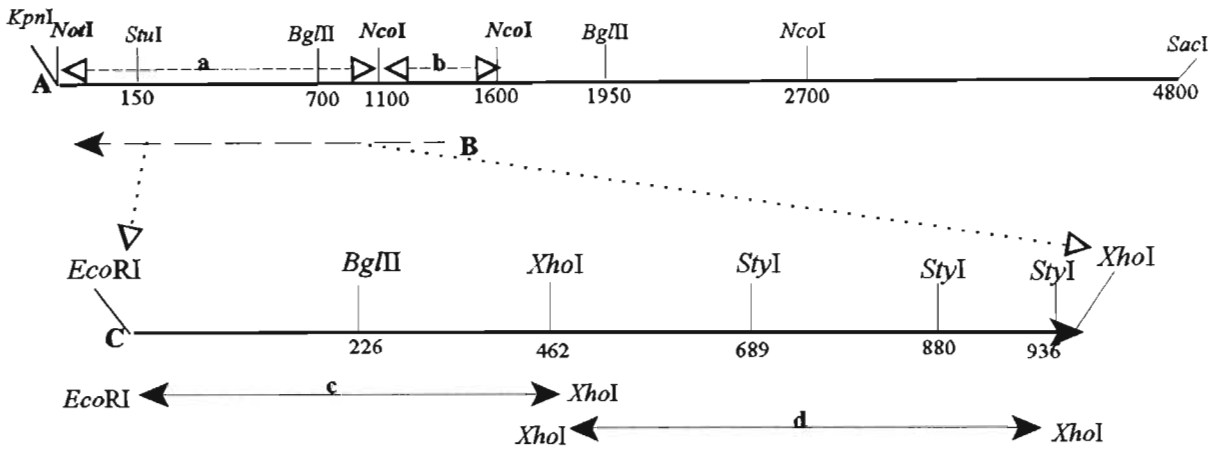


FIG. 2.4 Partial restriction map of the *Candida shehatae XYL1* gene. Line A (solid) represents the putative genomic clone, line B (dashed with arrow) represents the coding and non-coding regions of the gene and the direction of transcription and line C (solid with arrow) represents the coding region of the gene and the direction of transcription. The subclones *Not*I/NcoI (a) and *Nco*I (b) used to determine the DNA sequence of the genomic clone and *Eco*RI/*Xho*I (c) and *Xho*I (d) used to determine the sequence of the coding region of the gene are also shown.

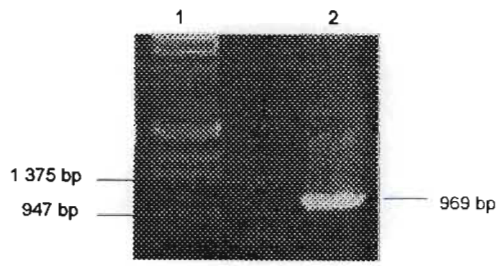


FIG. 2.5 PCR amplification product for the *C. shehatae XYL1* gene. Lane 1: molecular weight marker (λ DNA cleaved with *Eco*RI/*Hin*dIII); lane 2: *C. shehatae XYL1* gene.

PCR amplification of the plasmid DNA from the genomic clone produced a fragment of approximately 1 kb (Fig. 2.5). This DNA molecule was cloned into PCR-Script and sequenced.

Sequencing results revealed that the gene is 969 bp in length (Fig. 2.6). The 5' region of the gene contains several elements that are commonly found in yeast promoters. Candidates for

a Goldstein-Hogness box or TATA box are located upstream of the ATG codon (5' - TATAA) starting at nucleotides -114 (39) and -84 (69). Another promoter element, a 5' -CCAAT sequence, occurs at nucleotide -153 (1). A putative polyadenylation signal (AATAA) was found 151 bp downstream from the stop codon. The 3' non-coding region also contained several TAG sequences and a TAG...TACT...TTT motif. The amino acid sequence of the *C. shehatae* gene showed 89% identity and 94% homology with *C. tenuis*, 78% identity and 87% homology with *C. tropicalis* and 77% identity and 88% homology with *P. stipitis* (Fig. 2.7). Both *C. shehatae* and *C. tenuis* contain additional amino acids in the amino terminal sequence - five and four, respectively.

```
   1  CCAATGGTTA GCTCATCCTG GATGGGTGAA AATGGGGTAT ATAAGTCGGT
  51  GGATTCCCCC CAGATATCGT ATAAGCTGGA AAGAAGTATC AGTACCAGTC
 101  TACAGTTGTG TCTAATCCCA GCTTCTTTCT GTAATCTACT AACTACATCC
 151  ACAATGAGCC CAAGCCCAAT TCCAGCTTTC AAGTTGAACA ACGGCCTTGA
 201  AATGCCATCC ATCGGTTTCG GCTGTTGGAA GCTCGACAAA TCTACCGCCG
 251  CCGACCAGGT CTACAACGCC ATCAAGGCCG GTTACAGATT GTTCGACGGT
 301  GCCGAGGACT ACGGTAACGA ACAAGAAGTC GGTGAAGGTG TCAAGAGAGC
 351  CATCGACGAA GGTATTGTCA CCAGAGAGGA GATCTTCCTC ACCTCCAAGT
 401  TGTGGAACAA CTACCACGAC CCAAAGAACG TCGAAACCGC CTTGAACAAG
 451  ACCCTCAAGG ACCTTAAGGT CGACTACGTT GACTTGTTCT TGATCCACTT
 501  CCCAATTGCC TTCAAGTTTG TCCCAATCGA GGAGAAATAT CCACCAGGAT
 551  TCTACTGTGG TGACGGTGAC AACTTCGTCT ACGAAGACGT CCCAATCTTG
 601  GAGACCTGGA AGGCCCTCGA GAAGTTGGTC AAGGCCGGTA AGATTAGATC
 651  CATCGGTGTC TCCAACTTCC CAGGTGCTTT ACTCTTGGAC TTGTTCAGAG
 701  GTGCCACCAT CAAGCCTGCT GTTTTGCAAG TTGAGCACCA CCCATACTTG
 751  CAACAACCAA AGTTGATTGA GTACGCTCAA AAGGTCGGTA TCACTGTCAC
 801  CGCTTACTCT TCTTTCGGTC CTCAATCTTT CGTTGAGATG AACCAAGGTA
 851  GAGCTTTGAA CACCCCAACC TTGTTCGAAC ATGACGTCAT TAAGGCTATT
 901  GCTGCCAAGC ACAACAAAGT CCCAGCCGAG GTTTTGTTGA GATGGTCCGC
 951  TCAAAGAGGC ATAGCTGTCA TTCCAAAGTC TAACCTTCCA GAGAGATTAG
1001  TTCAAAACAG AAGTTTCAAC GACTTCGAGT TGACCAAGGA GGACTTTGAG
1051  GAAATCTCCA AGTTGGACAT CAACTTGAGA TTCAACGACC CATGGGACTG
1101  GGACAACATT CCAATCTTCG TTTAAGCGTG CTTGTTGCAG AAGAAGAAGA
1151  ATGCCGACAA AAATGCCGAA TCAATGGGTT CAATGTCATA ATTAGAAACT
1201  ATATATAAAT ACTATTAATT TAGCCCATTA TTTATAACCT GCTTTCTTTG
1251  CGTAGAACTA ACACCATAAG CCCAGTAATA ACCCGAAATA TCTCTGAAAA
1301  ATCCCAATTT TTTTTTCTT GATCTCTCTC ACCACATCAC CAGCATCATA
1351  ACAACACCAA AATCACCCCA TTCATCCATC ATCAACACCC ATAAAACAAT
1401  ATAACCATCA GATCAAACAA CCAAACAAAA TGCTTTTCGC CATACAAAAA
```

FIG. 2.6 Nucleotide sequence of the *C. shehatae XYL1* gene. Features highlighted include: non-coding sequences (in italics), coding sequencs, the 5' non-coding 'CCAA' and 'TATA' boxes (bold, underlined), the 3' non-coding transcriptional termination signal 'TAG-N-TATGT-N-TTT' (underlined) and the start (ATG), stop (TAA) and polyadenylation signals 'AATAA' (bold).

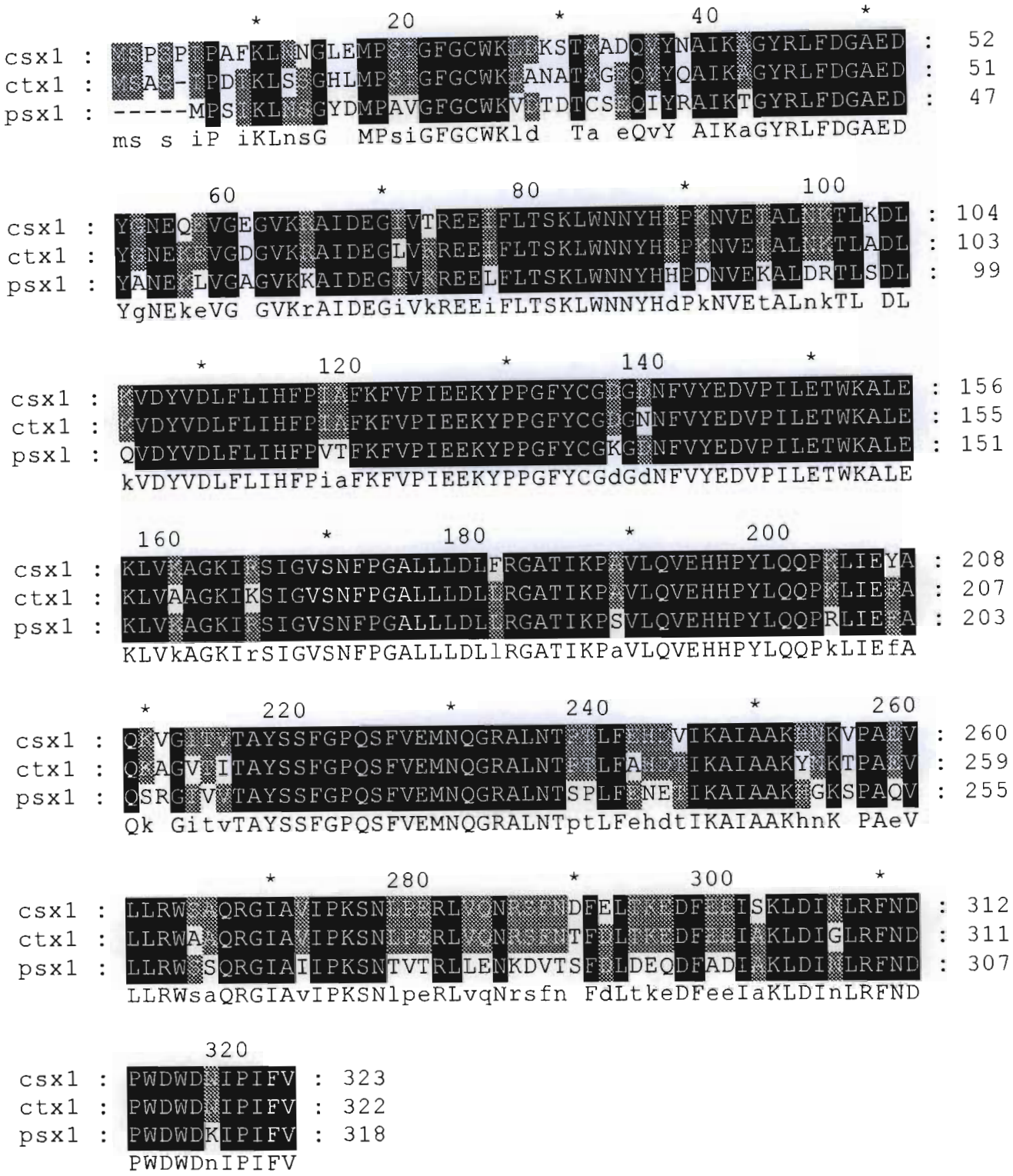


FIG. 2.7 Amino acid alignment of the *XYL1* genes from *C. shehatae* (csx1), *C. tenuis* (ctx1) and *P. stipitis* (psx1). Black boxes indicate homology among all the genes, grey boxes indicate homology between any two genes and dashes (-) indicate gaps in the sequence. Asterisks appear above every 20 amino acids starting from amino acid 10 and numbering appears after every 20 amino acids staring at amino acid 20.

2.3.3 XYLITOL DEHYDROGENASE GENE

A single PCR product of approximately 1.1 kb was obtained for the *P. stipitis XYL2* gene (Fig. 2.8). An amplification product was obtained only when $MgCl_2$ was added to the reaction. Optimal amplification was obtained at a $MgCl_2$ concentration of 1.25 mM. Fig. 2.9 shows a partial restriction map of the gene as well as subclones constructed for sequencing. Five subclones were constructed for manual sequencing. The nucleotide sequence of the gene (Appendix One) has 99% homology to that previously reported for *P. stipitis XYL2*.

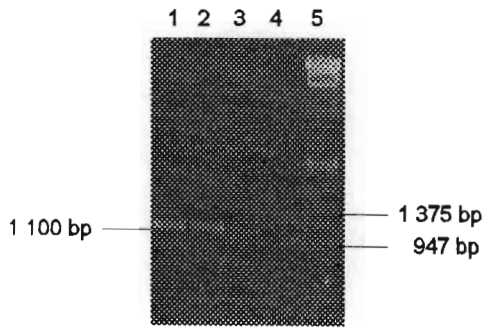


FIG. 2.8 PCR amplification product of the *P. stiptis XYL2* gene using various $MgCl_2$ concentrations. Lane 1: 1.25 mM $MgCl_2$; lane 2: 2.5 mM $MgCl_2$; lane 3: 3.75 mM $MgCl_2$; lane 5: molecular weight marker (λ DNA cleaved with *Eco*RI/*Hin*dIII).

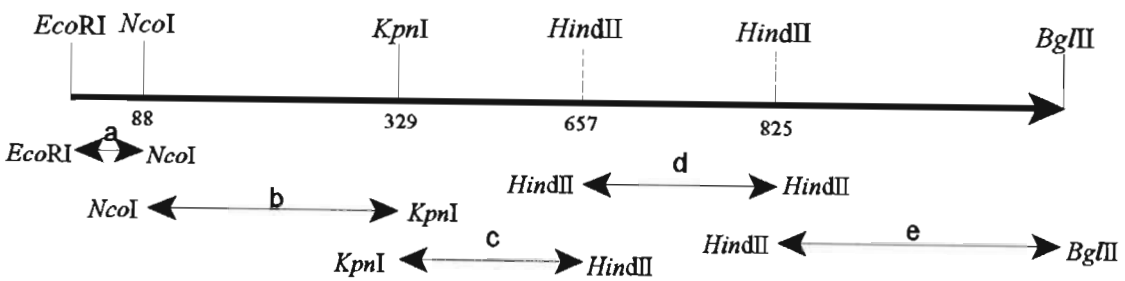


FIG. 2.9 Partial restriction map of the *P. stipitis XYL2* gene and the sub-cloning strategy used for sequencing. Five subclones (a-e) were constructed.

2.3.4. XYLULOKINASE GENE

A PCR product of approximately 1.8 kb was obtained for the *S. cerevisiae XYL3* gene (Fig. 2.10). A partial restriction map of this gene as well as the subclones (a-d) constructed for sequencing are indicated in Fig. 2.11. The nucleotide sequence of the gene (Appendix One) has 99% homology to the xylulokinase gene of *S. cerevisiae* as well as to an open reading frame on chromosome VII of *S. cerevisiae*.

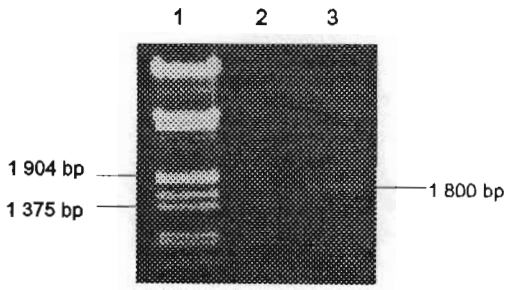


FIG. 2.10 PCR amplification product for the *S. cerevisiae XYL3* gene. Lane 1: molecular weight marker (λ DNA cleaved with *Eco*RI/*Hin*dIII); lane 3: *S. cerevisiae XYL3* gene.

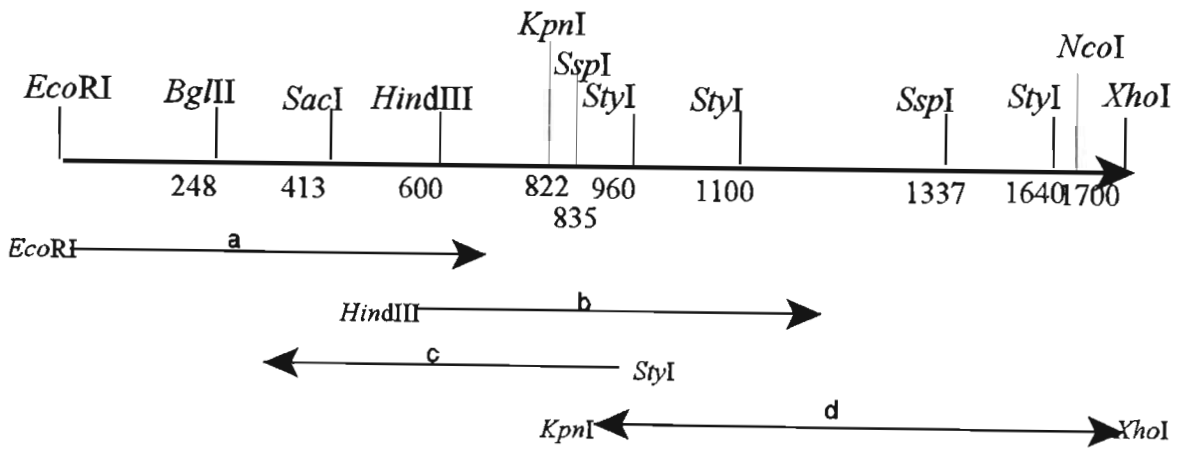


FIG. 2.11 Partial restriction map of the *S. cerevisiae XYL3* gene and the cloning strategy used for sequencing this gene. Four subclones (a-d) were constructed.

2.4 DISCUSSION

The nucleotide sequences of the *P. stipitis XYL1* and *XYL2* genes, the *C. shehatae XYL1* gene and the *S. cerevisiae XYL3* gene were determined and compared to DNA sequences available in the various databases accessible with the BLAST 2.0.6 algorithm (Altschul *et al.*, 1997). The nucleotide sequence of the *P. stipitis XYL1* gene has 99% homology to that reported by Amore *et al.* (1991) for *P. stipitis* (EMBL Accession Number X59465), differing by only six base pair substitutions. Substitutions occurred at positions 70 (G → A), 114 (T → C), 126 (C → T), 198 (C → T), 670 (T → A) and 943 (T →A). The nucleotide substitution at position 70 caused a change in amino acid - from a non-polar to a polar amino acid (Val to Thr). The other substitutions produced no amino acid changes (114, 126, 198 and 943) or an amino acid change with the same charge (Met instead of Lys) at position 670. The change from Val to Thr did not affect protein function (Chapter Three) and therefore does not appear to be in an essential part of the protein. The nucleotide substitutions may be due to differences between the two *P. stipitis* strains due to point mutations or could be due to mis-incorporation of nucleotides by the *Taq* DNA polymerase during amplification

PCR amplification of the *C. shehatae XYL1* gene was unsuccessful despite varying the reaction conditions. At very low annealing temperatures, multiple products were obtained. Therefore, a different strategy was attempted to clone this gene. This involved the construction of a genomic library of *C. shehatae* and screening the library using the *P. stipitis XYL1* gene as a probe. This led to the identification of a genomic clone containing the *C. shehatae XYL1* gene, the coding region of which was amplified using PCR, cloned and then sequenced.

The nucleotide sequence of the *C. shehatae XYL1* gene has 84% and 92% homology to that of *P. stipitis* (EMBL Accession Number X59465) and *C. tenuis* (GenBank Accession Number AF074484), respectively. The 5' and 3' non-coding regions contain several motifs

common to other eukaryotic genes. These include the two TATA boxes (TATAA at -114 and -84) as well as a CAAT sequence (CCAAT starting at nucleotide -153) in the 5' non-coding region of the gene and a polyadenylation signal (AATAA starting 151 bp downstream from the stop codon), several TAG sequences as well as a TAG...TACT...TTT motif in the 3' non-coding region. These sequences are also present in the 3' non-coding region of the *P. stipitis XYL1* (Amore *et al.*, 1991). This motif is similar to the TAG...TATGT...TTT sequence believed to be required for efficient transcription termination (Zaret and Sherman, 1982).

The *C. shehatae XYL1* gene contains a 969 bp open reading frame that starts with a Met codon and encodes a 323 amino acid polypeptide. The *P. stipitis* XR contains 318 amino acids (Amore *et al.*, 1991) and that of *C. tenuis* 322 amino acids (GenBank Accession Number AF074484). The *C. shehatae XYL1* gene shows highest homology to the *C. tenuis XYL1* gene with 89% identity and 94% homology at the protein level. It has 78%, 77% and 63% identity with the *C. tropicalis*, *P. stipitis* and *P. tannophilus XYL1* genes, respectively. Compared to *P. stipitis*, both *C. shehatae* and *C. tenuis* contain five and four additional amino acids at the amino terminal, respectively. The carboxy terminal of both proteins are almost identical. The presence of the additional sequences at the 5' end of the gene would explain why PCR amplification of the *C. shehatae XYL1* gene was unsuccessful using *P. stipitis*-specific primers. The 5' primer contained 24 nucleotides (the first 8 amino acids) that were specific for the 5' end of the *P. stipitis XYL1* gene. In this region, only three of the eight amino acids are identical for *C. shehatae* and *P. stipitis*. Therefore, this primer would not have annealed to the template DNA because of lack of homology. Even if it had annealed under low stringency conditions, the gene would have been truncated at the 5' end and may not have encoded a functional protein. The 3' primer would have annealed but for the PCR reaction to be possible both primers need to anneal to the template DNA.

Twenty two out of the first 23 amino acids at the amino terminal are identical to the 23 amino acids sequenced by Ho and co-workers (1990) indicating that the amino terminal of this protein is conserved among different strains of *C. shehatae*. The suggested cofactor binding site (IPKS) for aldose reductases was located between amino acids 273 and 276, 272 and 275, and 268 and 271 for *C. shehatae*, *C. tenuis* and *P. stipitis*, respectively. Yokoyama *et al.* (1995b) located the cofactor binding site between amino acids 274 to 277 for two aldose reductases from *C. tropicalis*. This indicates that the location of the cofactor binding site in *Candida* and *Pichia* species is conserved. The region downstream of this site - between 280 and 301 is thought to be responsible for cofactor binding specificity and shows relatively lower homology compared to the rest of the protein (Yokoyama *et al.*, 1995a). *C. shehatae* and *C. tenuis*, however, have almost identical sequences in this region (18 of 21 amino acids are identical). This could have implications for the cofactor utilization properties of the *Candida* and *Pichia* species. Both the *P. stipitis* and *C. shehatae* XRs have dual cofactor specificity. However, *P. stipitis* displays an almost equal preference (NADH activity is 72% that of NADPH) [Verduyn *et al.*, 1985b] for both cofactors compared to *C. shehatae* which shows a greater specificity for NADPH (NADH activity is 30% [Alexander *et al.*, 1988c] or 40% [Bruinenberg *et al.*, 1984; Ho *et al.*, 1990] that of NADPH). The *C. shehatae* and *C. tenuis* XR amino acid sequence in this region therefore displays greater specificity for NADPH while *P. stipitis* enzyme which has only eight out of 21 amino acids identical to those of *C. shehatae* and *C. tenuis* would display greater flexibility to utilize either cofactor.

The *P. stipitis XYL2* gene has 99% homology at the nucleotide level with that reported by Kötter *et al.* (1990), differing by only five base pair substitutions. Substitutions occurred at positions 553, 555, 794, 972 and 1 024. Only those at positions 555 (T → G) and 794 (T → C) resulted in an amino acid change from Gly → Trp and Ile → Thr, respectively. Although Gly is a

polar amino acid and Trp non-polar and Ile non-polar and Thr polar, these changes did not affect the activity of the enzyme (Chapter Three). Since both amino acids lie outside the putative NAD-binding domain described by Kötter *et al.* (1990), cofactor binding would not be affected. These amino acids do not appear to be essential for substrate binding since XDH activity is unaffected.

The *S. cerevisiae XYL3* gene has 99% homology at the nucleotide level with the open reading frame ORF YGR194c situated on chromosome VII of *S. cerevisiae* (EMBL Accession Number Z72979) and with the xylulokinase gene of *S. cerevisiae* (EMBL Accession Number X82408). The deduced amino acid sequence of the gene cloned in this study has 600 amino acids (including the first Met). One nucleotide substitution at position 597 (A → G) did not result in an amino acid change. The cloned gene had 100% identity with the reported *XYL3* gene from *S. cerevisiae* and the ORF on chromosome VII. This strongly suggests that the *XYL3* gene is located on chromosome VII.

Repeated attempts to clone the *XYL3* gene from *P. stipitis* and *C. shehatae* failed. Initial experiments involved PCR amplification using *S. cerevisiae*-specific primers and varying the amplification reaction conditions (Innis and Gelfand, 1990). When this proved unsuccessful, *P. stipitis* and *C. shehatae* genomic libraries were screened using the entire *S. cerevisiae XYL3* gene and smaller fragments of this gene as probes. Putative clones were only obtained at very low stringency conditions. However, partial sequencing did not confirm the presence of the xylulokinase gene in these clones. The *XYL3* genes from *E. coli* (Rosenfeld *et al.*, 1984), *P. tannophilus* (Stevis *et al.*, 1987) and *S. cerevisiae* (Ho and Chang, 1989) have been cloned by complementation in *E. coli* and *S. cerevisiae*. The low homology between the *XYL3* genes from different yeast and bacterial genera was further confirmed by the lack of cross-hybridization of the *E. coli*, *P. tannophilus* and *S. cerevisiae* genes as reported by Stevis *et al.* (1987). When

repeated attempts to clone either the *P. stipitis* or *C. shehatae XYL3* gene failed, it was decided to use the *S. cerevisiae XYL3* gene for further work in this study.

Alternative, but time-consuming, means of obtaining the *XYL3* gene from *P. stipitis* or *C. shehatae*, described below, were not pursued. The most likely route would have been by complementation of an *E. coli* strain capable of xylose metabolism in which the *XYLB* gene has been inactivated or mutated. Initially, this approach was investigated. The entire genome library *C. shehatae* in λ was converted to a double-stranded library in pBS. The *XYLB* gene was cloned from an *E. coli* strain by PCR amplification and a disruption cassette was constructed. Since a suitable strain was not available for the disruption study, this strategy was abandoned. Another approach would have been to purify the XK enzyme from either *P. stipitis* or *C. shehatae*, perform amino-terminal protein sequencing, generate the corresponding oligonucleotide probes and screen the respective genome libraries.

Three of the four genes cloned were obtained by PCR amplification. Although the *XYL1* and *XYL2* genes from *P. stipitis* are only approximately half the size of the *S. cerevisiae XYL3* gene, sequence alignments revealed greater variation between Genbank sequences and that obtained in this study for the *P. stipitis XYL1* and *XYL2* genes than the *S. cerevisiae XYL3* gene. There are two possible reasons for these differences. Firstly, the differences in sequence could be due to the fact that greater variation exists between these genes in the *P. stipitis* strains whereas the *XYL3* gene sequence is highly conserved among strains of *S. cerevisiae*. Alternatively, the differences could be due to mis-incorporations during the PCR amplification reaction or mis-reading of the nucleotide sequence generated by manual sequencing. The enzymes used for the amplification of the *P. stipitis XYL1* and *XYL2* genes was *Taq* polymerase (Boehringer Mannheim) and that used for the *XYL3* gene was the *TaKaRa Taq* (TaKaRa Biomedicals). Apparently, more mis-incorporations occurred during amplification reaction with

the *Taq* polymerase from Boehringer Mannheim. It is therefore likely that apparent sequence differences reported in our study compared to previous findings was due to the differing fidelities of the polymerases used in the PCR experiments. In order to establish whether it is indeed mis-incorporations during amplification that gives rise to the variation in gene sequence, the PCR would have to be repeated and the products sequenced and compared. If the same bases differ for the two strains then variation is most likely due to mutations. However, if variations occur at other nucleotides then it must be concluded that they are due to mis-incorporations by the polymerase enzymes during the amplification reaction. Whatever the cause of the sequence variations, they did not affect enzyme activity (Chapter Three).

All four xylose metabolizing genes cloned in this study were cloned into yeast expression vectors and used to transform a strain of *S. cerevisiae* in order to obtain a xylose-utilizing strain. These experiments are described in the subsequent Chapters.

CHAPTER THREE

CONSTRUCTION OF EXPRESSION VECTORS AND EXPRESSION OF XYLOSE METABOLIZING GENES IN *Saccharomyces cerevisiae*

3.1 INTRODUCTION

The majority of recombinant proteins produced in yeasts have been expressed using *S. cerevisiae* as a host. The preference for this organism is based on several important factors, viz., (i) its familiarity to molecular biologists; (ii) the wealth of knowledge available about its genetics and physiology; (iii) auxotrophic markers are well characterized; (iv) strong promoters have been identified; (v) a high copy number mitotically-stable plasmid vector has been identified; and (vi) extensive information about its safety as a GRAS organism through its use in the baking and brewing industries is available (Buckholz and Gleeson, 1991).

Specialized vectors that allow the efficient introduction of heterologous genes into the yeast cell have been developed. The foreign gene can either be integrated into the host genome in single or multiple copies or located on plasmids or other extrachromosomal DNA fragments (Marino, 1991). Extrachromosomal replicons (YEp) are based either on yeast autonomously replicating sequences (ARS) [Campbell, 1983] which function as origins of replication or on the 2μ circle of *S. cerevisiae* (Beggs, 1978). ARS vectors are present in multiple copies per cell (1 - 20) but are mitotically unstable. They can be stabilized by the addition of yeast centromeric sequences (CEN) but the copy number is then reduced to 1 - 2 per cell (Clark and Carbon, 1980). In practice, however, ARS vectors are rarely used for foreign gene expression (Romanos *et al.*, 1992).

Most yeast expression vectors are based on the multi-copy 2μ plasmid and contain sequences for propagation in *E. coli* and in yeasts. The 2μ plasmid is 6.3 kb in size and present

in most *Saccharomyces* strains at about 100 copies per haploid genome (Futcher, 1988; Murray, 1987). The plasmid encodes four genes: *FLP*, *REP1*, *REP2* and D; an origin of replication (ORI); the *STB* locus required *in cis* for stability; and two 599 bp inverted repeat sequences. *FLP* encodes a site-specific recombinase. Efficient segregation depends on having the *STB* locus in *cis* and the *REP1* and *REP2* gene products (Kikuchi, 1983). Although it confers no known phenotype and may be slightly disadvantageous to the host cell (Futcher and Cox, 1983; Mead *et al.*, 1986), the 2μ plasmid is stably inherited.

The simplest 2μ vectors contain the *ORI-STB* region, a yeast selectable marker and bacterial plasmid sequences and must be used in 2μ^+ host strains which supplies REP1 and REP2 proteins (Kikuchi, 1983). *ORI-STB* vectors are the most convenient for routine laboratory use due to their small size and ease of manipulation. They are ten-fold more stable than ARS plasmids. More complex 2μ-based shuttle vectors contain the *REP1* and *REP2* genes in addition to *ORI-STB* and can therefore be used in 2μ-free host strains (Armstrong *et al.*, 1989). They are more cumbersome than *ORI-STB* vectors but have greater stability and are more suitable for scale-up.

Chromosomal integration offers a much more stable alternative to episomal maintenance of foreign DNA. In *Saccharomyces*, integration occurs by homologous recombination (Orr-Weaver *et al.*, 1983). Integrating vectors (YIp) contain yeast chromosomal DNA sequences to target integration, a selectable marker and a bacterial replicon. Digestion at a unique restriction site within the homologous DNA promotes high efficiency transformation and targets integration. This single cross-over integration results in a duplication of the target sequence and the vector can 'pop out' by excisional recombination. Double homologous recombination (transplacement) results in a stable structure without duplication (Rothstein, 1983).

Several strategies based on integration into re-iterated chromosomal DNA sequences have

been used to generate stable multi-copy integrants (Kingsman *et al.*, 1985; Lopes *et al.*, 1989; 1990). Targeting of the ribosomal DNA (rDNA) cluster produces the best results in terms of copy number and expression (Lopes *et al.*, 1989; 1990). The transposable element *Ty*, which is present at 30 to 40 copies per genome in most *Saccharomyces* strains is another target for integration (Kingsman *et al.*, 1985).

For efficient transcription of foreign genes, they must be present on cloned cDNA downstream from yeast promoters. The first published example of such a promoter was the use of a 1 500 bp fragment upstream from the *ADH* gene for efficient expression of leucocyte α-interferon (Hitzeman *et al.*, 1981). At least three elements which regulate the efficiency and accuracy of transcription initiation are present in yeast mRNA promoters (Struhl, 1989), viz., upstream activation sequences (UASs), TATA elements and initiator elements. Many also contain elements involved in transcription repression.

The first promoters used were from genes encoding glycolytic enzymes, e.g., *ADH1* (Hitzeman *et al.*, 1981), *PGK* (Hitzeman *et al.*, 1983) or glyceraldehyde-3-phosphate dehydrogenase. At first, these were thought to be constitutive but were later found to be induced 20 - 30 fold by the addition of glucose (Tuite *et al.*, 1982). Glycolytic promoters are the most powerful promoters for *S. cerevisiae*, e.g., *PGK* mRNA accumulates to 5% of total mRNA (Romanos *et al.*, 1992). The *PGK* promoter has been studied in some detail. It extends over 500 bp and contains a complex UAS at -473 to -422, a heat shock regulatory element and other features contributing to accurate and efficient initiation (Kingsman *et al.*, 1985).

Glucose repression is a global system regulating the expression of a number of genes, including sugar fermentation genes, by the availability of glucose. Promoters regulated primarily by glucose repression include *ADH2*, *SUC2* (invertase) and *CYC1* (cytochrome C). The *ADH2* promoter is both powerful and tightly regulated and has been successfully used for foreign gene

expression (Price *et al.*, 1990). Since it is repressed over 100-fold by glucose, it can be used for efficient expression of toxic proteins. Other regulated promoter systems include: (i) a temperature-regulated system based on mating type control, (ii) the promoter of the *CUP1* gene which depends on the concentration of Cu^{2+} ions for induction, and (iii) phosphate regulated promoters (*PHO5*) [Romanos *et al.*, 1992].

Yeast transcriptional terminators are usually present in expression vectors for efficient mRNA 3'-end formation. Terminators of prokaryotic and higher eukaryotic genes are not normally active in yeast. Contrary to earlier ideas, it appears that yeast mRNAs follow the same pattern of termination, processing and polyadenylation of pre-mRNA as higher eukaryotes. In yeast, however, these processes are tightly coupled and occur within a shorter distance near the 3'-end of the gene (Butler *et al.*, 1990).

In addition to an expression cassette and 2μ sequences, shuttle vectors generally contain markers for selection in *S. cerevisiae* and *E. coli*. The most commonly used selectable markers for *E. coli* are genes that confer resistance to antibiotics such as ampicillin, tetracycline, kanamycin and chloramphenicol (Sambrook *et al.*, 1989). In yeast, auxotrophic selection markers are most commonly used. *LEU2*, *TRP1*, *URA3* and *HIS3* are used in corresponding mutant strains which are auxotrophic for leucine, tryptophan, uracil and histidine, respectively (Beggs, 1978; Rose *et al.*, 1984; Struhl and Davis, 1980; Tschumper and Carbon, 1980). For continued selection, recombinants are maintained on minimal media lacking the relevant nutrient. *URA3* and *LYS2* are particularly useful in that there are also methods for counter-selection of the marker.

Dominant selection markers are useful since they increase the range of host strains to be tested to include prototrophic and industrial strains of *S. cerevisiae*. Other antibiotic resistance markers used successfully in yeast are hygromycin B (Gritz and Davies, 1983) and chloramphenicol resistance (Hadfield *et al.*, 1986).

Interest in transformation of yeast cells probably began shortly after the phenomenon was reported in bacteria. Hinnen *et al.* (1978) obtained the first successful transformation of yeast by treating yeast protoplasts with foreign DNA in the presence of PEG. They used a chimeric ColE1 plasmid carrying the yeast *LEU2* gene to transform a yeast strain carrying a *leu2*$^-$ double mutation. At the same time, Beggs (1978) used the 2μ plasmid from yeast to transform *S. cerevisiae*. She constructed a chimeric plasmid from 2μ, the gene for tetracycline resistance and the yeast *LEU2* gene.

These first methods for the transformation of *S. cerevisiae* involved enzymatic removal of the cell wall to produce sphaeroplasts which could take up DNA on treatment with calcium and PEG. Later, a more convenient method was developed. Intact yeast cells were made competent by treatment with lithium ions (Ito *et al.*,1983). This method is now widely used although it gives low transformation frequencies. A variation, using DMSO (Hill *et al.*, 1991), increases frequency 25-fold. A third approach, electroporation, has also been used. A highly efficient method has been reported by Becker and Guarente (1991) and Meilhoc *et al.* (1990). These protocols require strict control of the growth phase of cells which involves many time-consuming steps of cell handling before electroporation. Grey and Brendel (1992) reported a rapid protocol which does not require preparation of cell cultures and produced a transformation efficiency of about 10% of the most efficient transformations that utilize cells of defined growth phase.

An important factor to consider in foreign gene expression is the frequent wide variation in the productivity of different transformants when 2μ vectors are used (Romanos *et al.*, 1992). This is apparently due to an unexplained stable variation in plasmid copy number between different transformants (Purvis *et al.*, 1987). It is therefore important to analyse a number of transformants when optimizing expression.

Gene expression is most frequently regulated at the level of transcription and it is generally

assumed that the steady-state mRNA level is a primary determinant of the yield of a foreign protein. In most cases, the yield of a foreign protein expressed in yeast using a yeast promoter has been much lower than the yield of the homologous protein using the same promoter (Romanos *et al.*, 1992).

In this study, two YEp plasmids were used to express the xylose-metabolizing genes described in the previous Chapter. This Chapter describes the construction of yeast expression vectors harbouring one, two and three xylose-metabolizing genes under the control of both expression cassettes and the introduction of the heterologous genes into a laboratory strain of *S. cerevisiae*. The biological activity of the recombinant proteins were confirmed by enzyme assays.

3.2 MATERIALS AND METHODS

3.2.1 CONSTRUCTION OF EXPRESSION VECTORS

Yeast expression vectors (pDLG1 and pJC1) were obtained from W.H. van Zyl (University of Stellenbosch, South Africa). Both plasmids are based on the yeast-*E. coli* shuttle vector YEp332 which contains the *URA3* selectable marker, but differ in their expression cassettes. pDLG1 (Fig. 3.1) contains the *ADH2* promoter/terminator (La Grange *et al.*, 1996) while pJC1 (Fig. 3.2) contains the *PGK* promoter/terminator cassette (Crous *et al.*, 1995). A polylinker containing unique *Eco*RI, *Bgl*II and *Xho*I sites separated the promoter/terminator sequences of the expression cassettes.

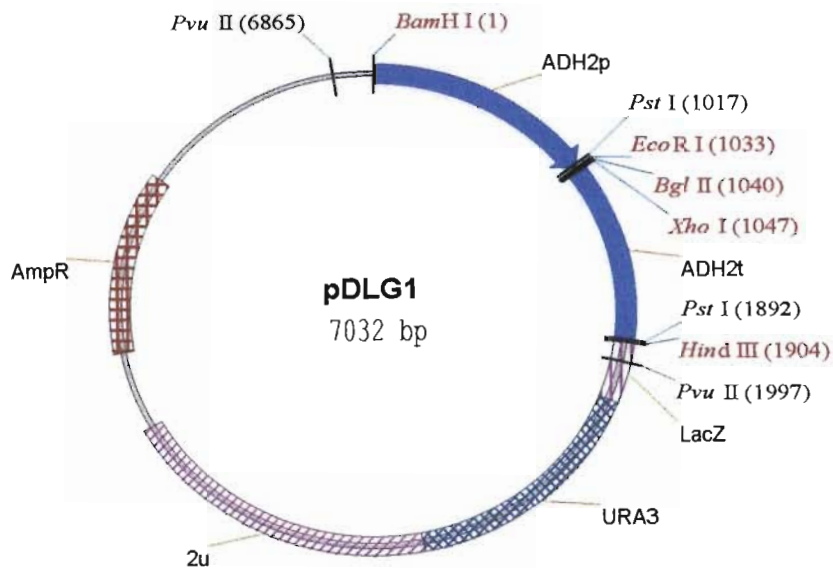


FIG. 3.1 Map of the yeast - *E. coli* shuttle vector pDLG1. The vector contains the 2μ origin of replication, ampicillin resistance gene (Amp^R), uracil selectable marker (*URA3*) and the *ADH2* promoter/terminator cassette (La Grange *et al.*, 1996).

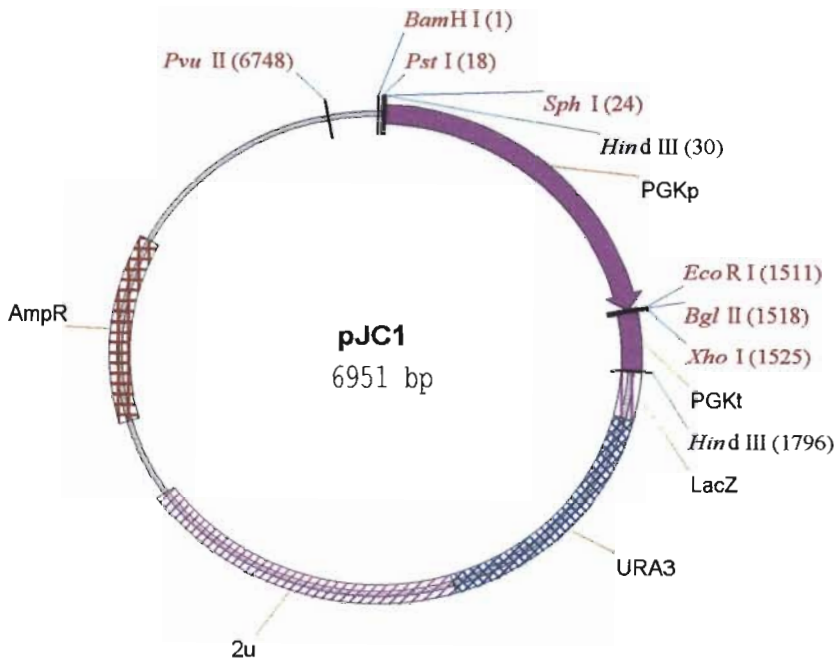


FIG. 3.2 Map of the yeast - *E. coli* shuttle vector pJC1. The vector contains the 2μ origin of replication, ampicillin resistance gene (Amp^R), uracil selectable marker (*URA3*) and the *PGK1* promoter/terminator cassette (Crous *et al.*, 1995).

3.2.1.1 Single gene constructs

Plasmids pRG6 (Fig. 3.3a) and pRG8 (Fig. 3.3b) were constructed by cloning the *XYL1* gene of *P. stipitis*, removed from plasmid pBPX1 (Table 3.1), as a *Xho*I fragment into the *Xho*I site of the expression cassette of pDLG1 and pJC1, respectively. Putative clones were cleaved with *Eco*RI to verify orientation of the gene in the expression cassette. A 600 bp fragment indicated that the gene was in the correct orientation for transcription. Plasmids pRG7 (Fig. 3.4a) and pRG9 (Fig. 3.4b) were constructed by cloning the *XYL2* gene of *P. stipitis* removed from pBPX2 (Table 3.1), as an *Eco*RI/*Bgl*II fragment into the *Eco*RI/*Bgl*II site of pDLG1 and pJC1, respectively. pRG16 (Fig. 3.5a) and pRG17 (Fig. 3.5b) were constructed by cloning the *XYL1* gene of *C. shehatae*, removed from plasmid pBCX1 (Table 3.1), as an *Eco*RI fragment into the *Eco*RI site of pJC1 and pDLG1, respectively. Orientation of the gene in the expression cassettes was verified by diagnostic digests with the restriction enzyme *Bgl*II. An 800 bp fragment was produced when the gene was in the correct orientation. Plasmids pRG14 (Fig. 3.6a) and pRG15 (Fig. 3.6b) were constructed by cloning the *XYL3* gene of *S. cerevisiae*, removed from plasmid pBSX3 (Table 3.1), as an *Eco*RI/*Xho*I fragment into the *Eco*RI/*Xho*I site of pJC1 and pDLG1, respectively. The yeast expression vectors used and constructed in this study are listed in Table 3.1. The protocol for ligation reactions was described in Chapter Two (Section 2.2.5).

3.2.1.2 Double gene constructs

In order to facilitate manipulation of the expression cassettes, all the single gene constructs were removed as *Bam*HI/*Hin*dIII, *Hin*dIII or *Pvu*II fragments from the yeast expression vectors and cloned into their respective sites in pBS. The expression cassettes and genes were removed as *Hin*dIII fragments from pRG8 and pRG16 and cloned into the *Hin*dIII site of pBS resulting in plasmids pBS8 and pBS16, respectively (Table 3.1). A *Hin*dIII and a *Pvu*II fragment from

pRG6 and pRG17 were cloned into the *Hin*dIII and *Sma*I sites of pBS yielding plasmids pBS6 and pBS17, respectively. The above-mentioned plasmids are constructs containing the *XYL1* genes from *P. stipitis* and *C. shehatae*. The *XYL2* gene from *P. stipitis* was removed with its expression cassette from pRG9 as a *Hin*dIII fragment and cloned into pBS8 and pBS16 linearized with *Hin*dIII and dephosphorylated (described below). The resulting plasmids were named pBS89 and pBS169, respectively. The *Hin*dIII fragment from pRG9 was polished (described below) and ligated into pBS17 (linearized with *Eco*RV and dephosphorylated) to produce plasmid pBS179. The *XYL3* gene was removed with its expression cassettes as 3.8 kb and 3.98 kb *Pvu*II fragments from pRG14 and pRG15 and cloned into the *Eco*RV site of pBS8 and pBS16 to yield pBS814, pBS815 and pBS1615, respectively.

Blunt-ended ligations were performed using the Rapid DNA Ligation Kit (Boehringer Mannheim) according to the manufacturer's instructions. Dephosphorylation of the 5' ends of DNA fragments used in the ligation reactions was carried out by adding one unit of calf intestinal phosphatase (Boehringer Mannheim) to a restriction reaction mixture and incubating at 37°C for a further 1 h. In order to produce DNA fragments with blunt ends, a mixture containing 20 μl DNA, 1 × SuRECut buffer H, 0.5 μM dNTPs and 2 units of Klenow enzyme (Boehringer Mannheim) was incubated at 37°C for 30 min.

TABLE 3.1 Bacterial plasmids used in this study

Plasmid	Relevant genotype	Source or reference
Vectors		
pBS	2.98 kb cloning vector, Amp^R, *lacZ*	Stratagene
pBS(J)	2.98 kb cloning vector, Amp^R, *lacZ*	W.H. van Zyl*
PCRscript	2.96 kb cloning vector, Amp^R, *lacZ*	Stratagene
Constructs		
pBPX1	*PsXYL1* gene in pBS	This study
pBCX1g	*CsXYL1* -genomic clone in pBS	This study
pBCX1	*CsXYL1* gene in PCRscript	This study
pBPX2	*PsXYL2* gene in pBS(J)	This study
pBSX3	*ScXYL3* gene in PCRscript	This study
pBS6	$ADH2_P$-*PsXYL*1-$ADH2_T$ in pBS	This study
pBS7	$ADH2_P$-*PsXYL2*-$ADH2_T$ in pBS	This study
pBS8	$PGK1_P$-*PsXYL*1-$PGK1_T$ in pBS	This study
pBS9	$PGK1_P$-*PsXYL2*-$PGK1_T$ in pBS	This study
pBS14	$PGK1_P$-*ScXYL3*-$PGK1_T$ in pBS	This study
pBS15	$ADH2_P$-*ScXYL3*-$ADH2_T$ in pBS	This study
pBS16	$PGK1_P$-*CsXYL*1-$PGK1_T$ in pBS	This study
pBS17	$ADH2_P$-*CsXYL*1-$ADH2_T$ in pBS	This study
pBS814	$PGK1_P$-*PsXYL*1-$PGK1_T$ $PGK1_P$-*ScXYL3*-$PGK1_T$ in pBS	This study
pBS815	$PGK1_P$-*PsXYL*1-$PGK1_T$ $ADH2_P$-*ScXYL3*-$ADH2_T$ in pBS	This study
pBS169	$PGK1_P$-*CsXYL*1-$PGK1_T$ $PGK1_P$-*PsXYL2*-$PGK1_T$ in pBS	This study
pBS1615	$PGK1_P$-*CsXYL*1-$PGK1_T$ $ADH2_P$-*ScXYL3*-$ADH2_T$ in pBS	This study
pBS1715	$ADH2_P$-*CsXYL*1-$ADH2_T$ $ADH2_P$-*ScXYL3*-$ADH2_T$ in pBS	This study

* Stratagene: La Jolla, California, USA
*W.H. van Zyl: University of Stellenbosch, Stellenbosch, South Africa

Table 3.2 Yeast vectors used in this study

Vector	Relevant Genotype	Source
pDLG1	Amp^R *URA3 ADH2*$_{PT}$	La Grange *et al.* (1996)
pJC1	Amp^R *URA3 PGK1*$_{PT}$	Crous *et al.* (1995)
pRG6	Amp^R *URA3 ADH2*$_P$-*PsXYL1*-*ADH2*$_T$	This study
pRG7	Amp^R *URA3 ADH2*$_P$-*PsXYL2*-*ADH2*$_T$	This study
pRG8	Amp^R *URA3 PGK1*$_P$-*PsXYL1*-*PGK1*$_T$	This study
pRG9	Amp^R *URA3 PGK1*$_P$-*PsXYL2*-*PGK1*$_T$	This study
pRG14	Amp^R *URA3 PGK1*$_P$-*ScXYL3*-*PGK1*$_T$	This study
pRG15	Amp^R *URA3 ADH2*$_P$-*ScXYL3*-*ADH2*$_T$	This study
pRG16	Amp^R *URA3 PGK1*$_P$-*CsXYL1*-*PGK1*$_T$	This study
pRG17	Amp^R *URA3 ADH2*$_P$-*CsXYL1*-*ADH2*$_T$	This study
pRG16159	Amp^R *URA3 PGK1*$_P$-*CsXYL1*-*PGK1*$_T$ *ADH2*$_P$-*ScXYL3*-*ADH2*$_T$ *PGK1*$_P$-*PsXYL2*-*PGK1*$_T$	This study
pRG16914	Amp^R *PGK1*$_P$-*CsXYL1*-*PGK1*$_T$ *PGK1*$_P$-*PsXYL2*-*PGK1*$_T$-*PGK1*$_P$-*ScXYL3*-*PGK1*$_T$	This study
pRG17157	Amp^R *URA3 ADH2*$_P$-*CsXYL1*-*ADH2*$_T$ *ADH2*$_P$-*ScXYL3*-*ADH2*$_T$ *ADH2*$_P$-*PsXYL2*-*ADH2*$_T$	This study

3.2.1.3 Triple gene constructs

Double gene constructs were removed from the corresponding pBS vectors either as *Apa*I/*Not*I, and *Cla*I/*Not*I fragments and blunt-ended or as *Pvu*II fragments and ligated into the yeast - *E. coli* shuttle vectors. A 6.4 kb *Apa*I/*Not*I fragment containing the *PGK1*$_P$-*CsXYL1*-*PGK1*$_T$ *ADH2*$_P$-*ScXYL3*-*ADH2*$_T$ cassettes was blunt-ended and ligated to the *Sma*I site of pRG9

to create pRG16159 (Fig. 3.7). A 5.3 kb *Cla*I/*Not*I fragment containing the *PGK1*$_P$-*CsXYL1*-*PGK1*$_T$ *PGK1*$_P$-*PsXYL2*-*PGK1*$_T$ cassette was blunt-ended and ligated to the *Sma*I site of pRG14 to create pRG16914 (Fig. 3.8). A 6.9 kb *Pvu*II fragment containing the *ADH2*$_P$-*CsXYL1*-*ADH2*$_T$ *ADH2*$_P$-*ScXYL3*-*ADH2*$_T$ cassette was cloned into the *Sma*I site of pRG7 to create pRG17157 (not shown). Orientation of the different gene cassettes in relation to each other was determined by restriction endonuclease analysis with either *Eco*RI or *Xho*I.

3.2.2 TRANSFORMATION

The single-gene and triple-gene expression vector constructs described in Section 3.2.1.1 were transformed into *S. cerevisiae* Y294 by the LiOAc/DMSO method (Hill *et al.*, 1991) or by electroporation (Grey and Brendel, 1992). Transformants were selected on SC medium lacking uracil (SC^{-URA}). The carbon source was either xylose or xylitol supplemented with galactose.

3.2.3 ENZYME ASSAYS

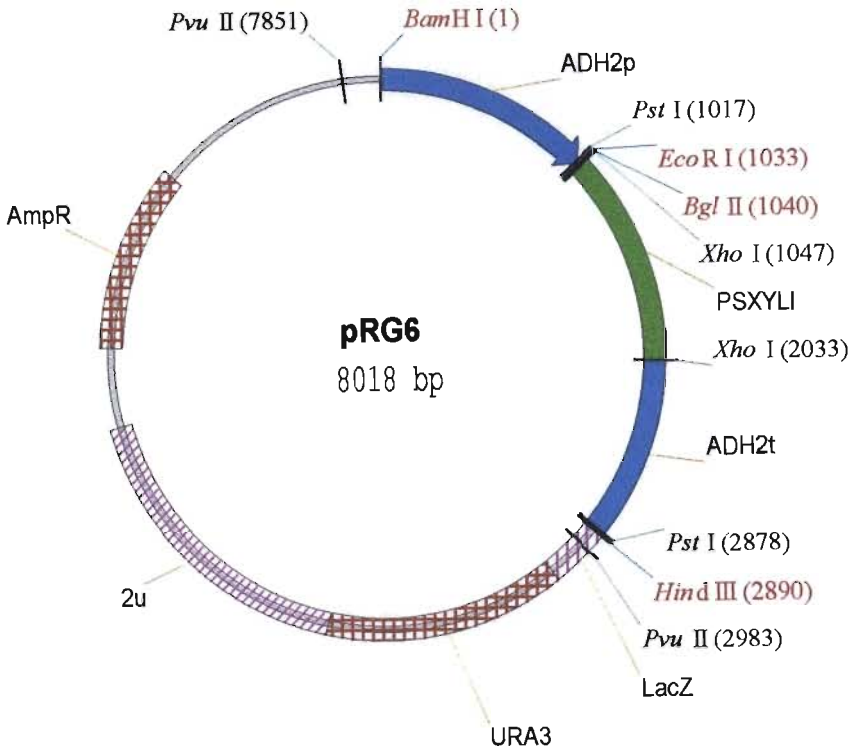
Cell extracts were prepared from parents and transformants. Cells were grown in SC^{-URA} medium, harvested by centrifugation, washed in sterile distilled water and resuspended in buffer (0.1 M potassium phosphate, pH 7.2). Three hundred μl of 0.1 M 2-mercaptoethanol and 0.3 g acid washed glass beads were added to cells which were then disrupted by vortexing. XR and XDH were assayed for according to the method of Smiley and Bolen (1982) and XK as described by Shamanna and Sanderson (1979). Absorbance values were corrected for background absorbance readings due to substances present in the cell extracts as well as absorbance fluctuations due to non-specific oxidation and reduction reactions. Protein concentrations were determined according to the method of Bradford (1976).

3.3 RESULTS

3.3.1 EXPRESSION VECTORS

Putative positive clones were verified after plasmid DNA isolations and cleavage with restriction endonucleases. Digests with the restriction endonucleases *Eco*RI or *Xho*I confirmed the orientation of the genes and expression cassettes in the vectors. Schematic diagrams of the expression vectors constructed in this study are depicted in Figs. 3.3 to 3.8. Expression vectors pRG6 (8 018 bp) and pRG8 (7 937 bp) containing the *P. stipitis XYL1* gene under the control of the *ADH2* and *PGK1* promoter/terminator cassettes, respectively, are represented in Fig. 3.3. The corresponding vectors bearing the *C. shehatae XYL1* genes, pRG17 (Fig. 3.4a) and pRG16 (Fig. 3.4b) were larger by 24 bp. pRG7 (8 132 bp) carries the *P. stipitis XYL2* gene controlled by the *ADH1* promoter/terminator cassette and pRG8 (8 051 bp) the same gene under the control of the *PGK1* promoter/terminator cassette (Fig. 3.5). The *S. cerevisiae XYL3* vectors, pRG15 and pRG14, are 8 832 bp and 8 751 bp in size, respectively (Fig. 3.6). The triple gene expression vectors are much larger in size. pRG16159 is 14 516 bp (Fig. 3.7) and carries the *S. cerevisiae XYL3* gene controlled by the *ADH2* promoter/terminator cassette sandwiched between the *C. shehatae XYL1* gene and *P. stipitis XYL2* gene controlled by the *PGK1* promoter/terminator cassette. The direction of transcription is the same for all three genes. The second triple gene construct, pRG16914 (Fig. 3.8), is slightly smaller (14 375 bp) and contains all three xylose metabolizing genes under the control of the *PGK1* promoter/terminator cassette. The genes follow the order of appearance in the xylose metabolic pathway and are transcribed in the same direction.

a



b

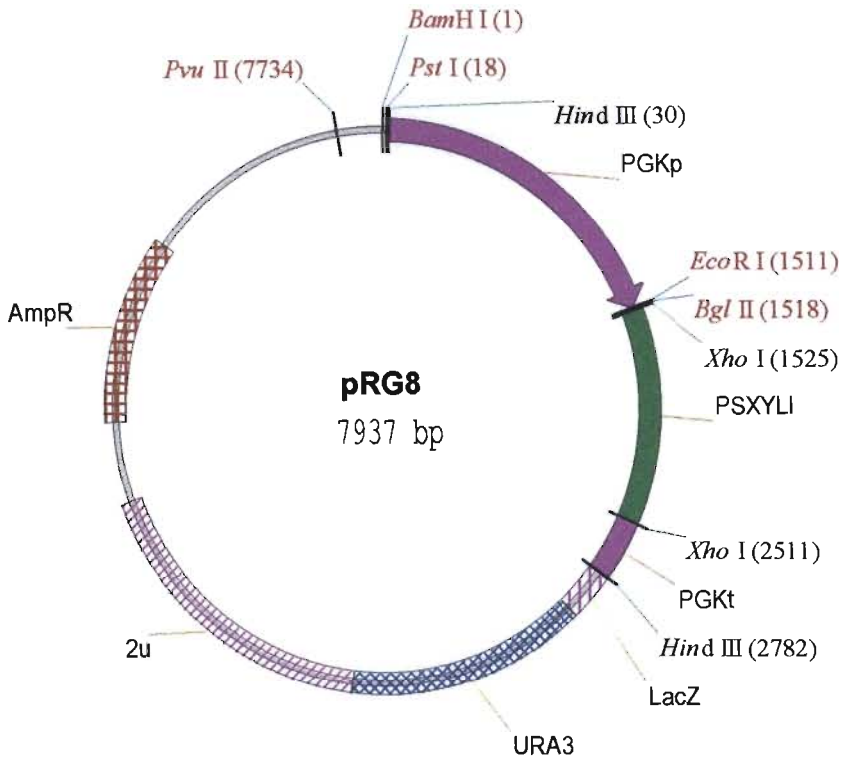
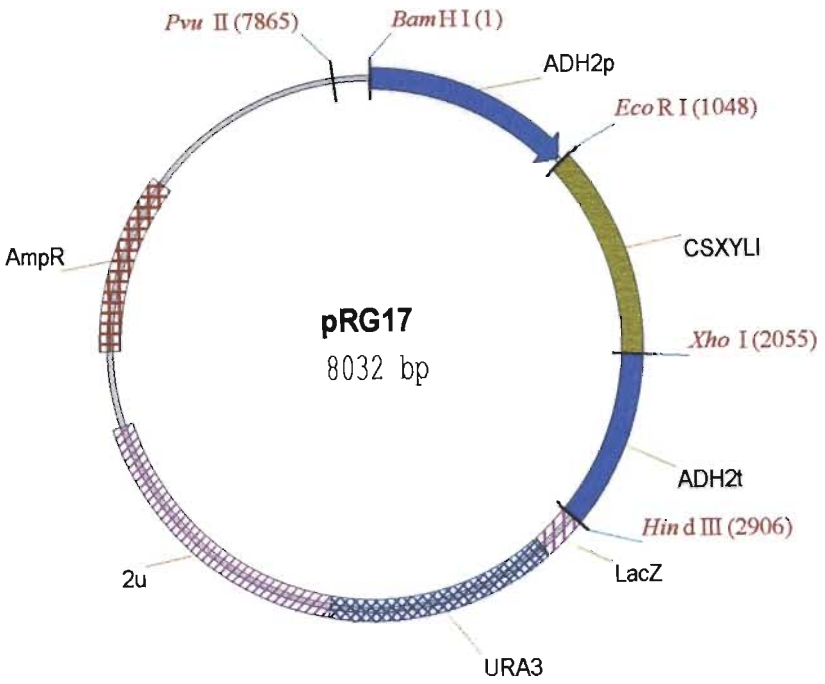


Fig. 3.3 Maps of the expression vectors for the *P. stipitis XYL1* gene under the control of the (a) *ADH2* (pRG6) and (b) *PGK1* (pRG8) promoter/terminator cassettes.

a



b

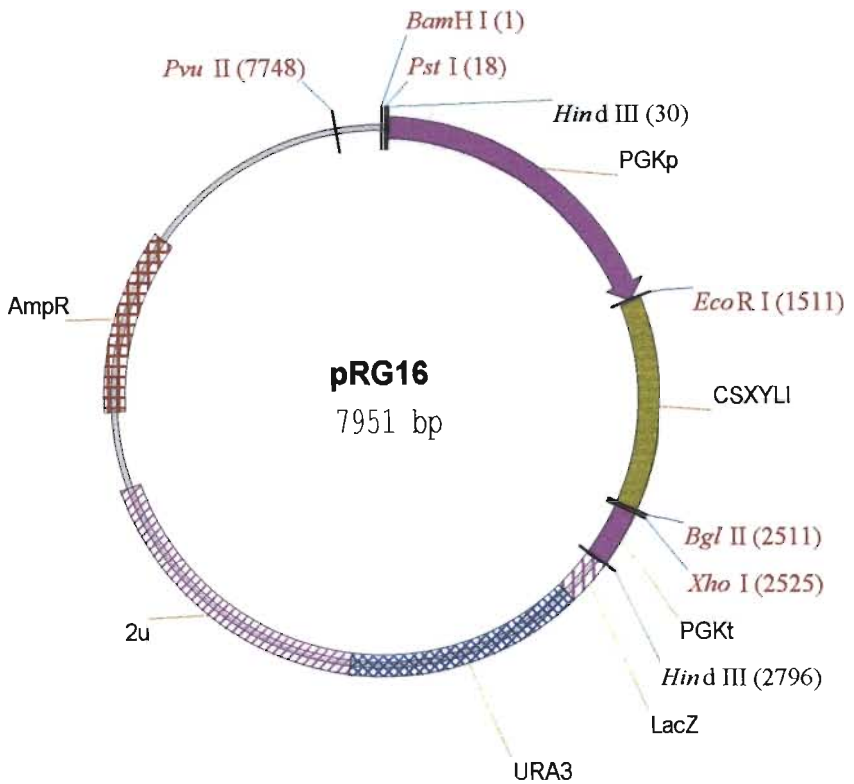
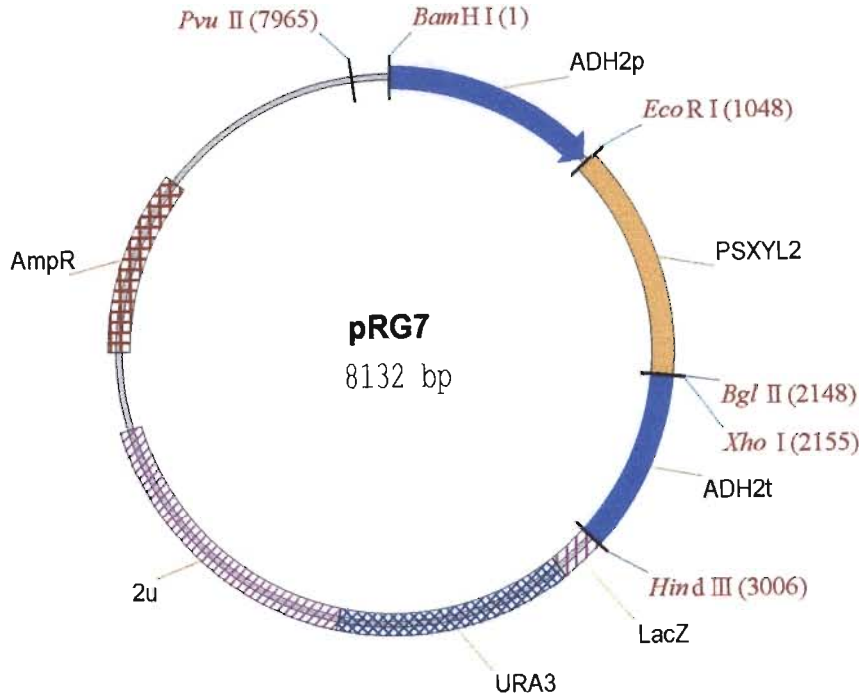


Fig. 3.4 Schematic diagrams of the expression vectors for the *C. shehatae XYL1* gene under the control of the (a) *ADH2* (pRG17) and (b) *PGK1* (pRG16) promoter/terminator cassettes.

a



b

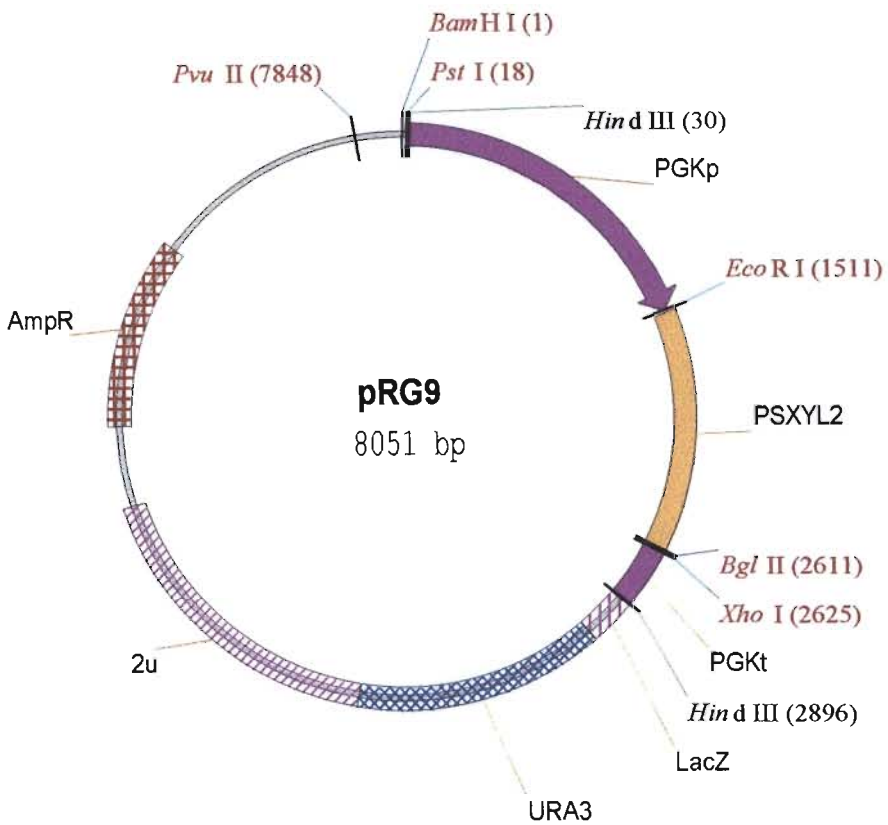
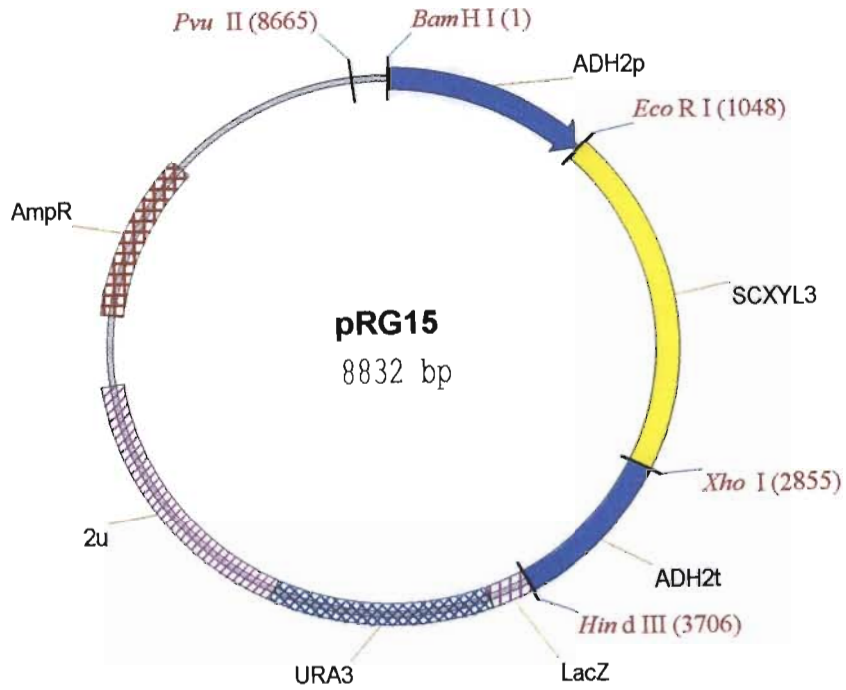


Fig. 3.5 Schematic diagrams of the expression vectors for the *P. stipitis XYL2* gene under the control of the (a) *ADH2* (pRG7) and (b) *PGK1* (pRG9) promoter/terminator cassettes.

a



b

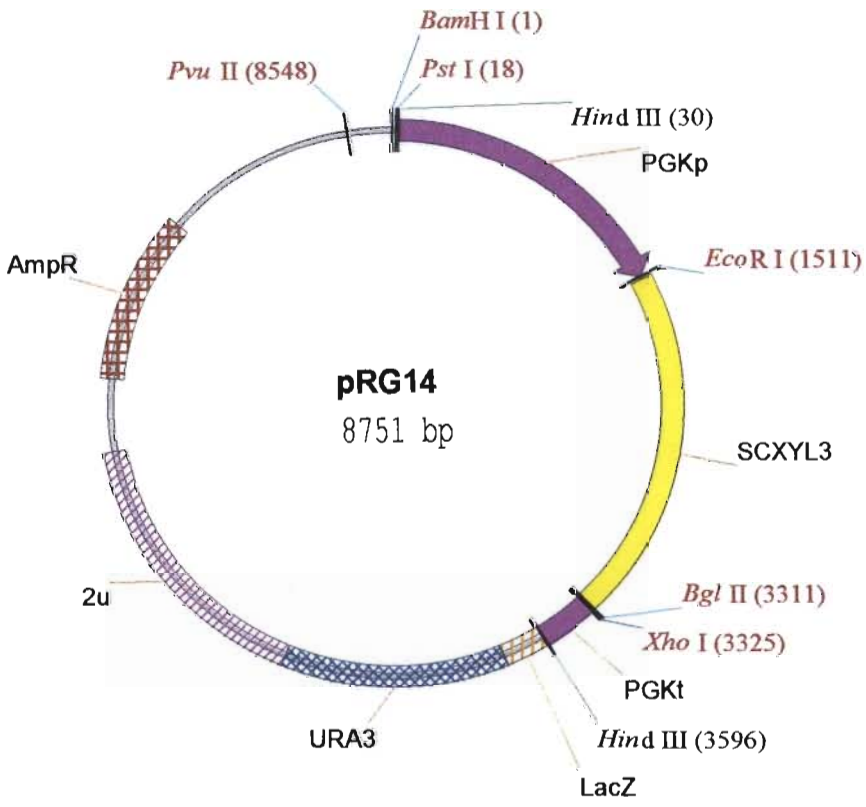


Fig. 3.6 Schematic diagrams of the expression vectors for the *S. cerevisiae XYL3* gene under the control of the (a) *ADH2* (pRG15) and (b) *PGK1* (pRG14) promoter/terminator cassettes.

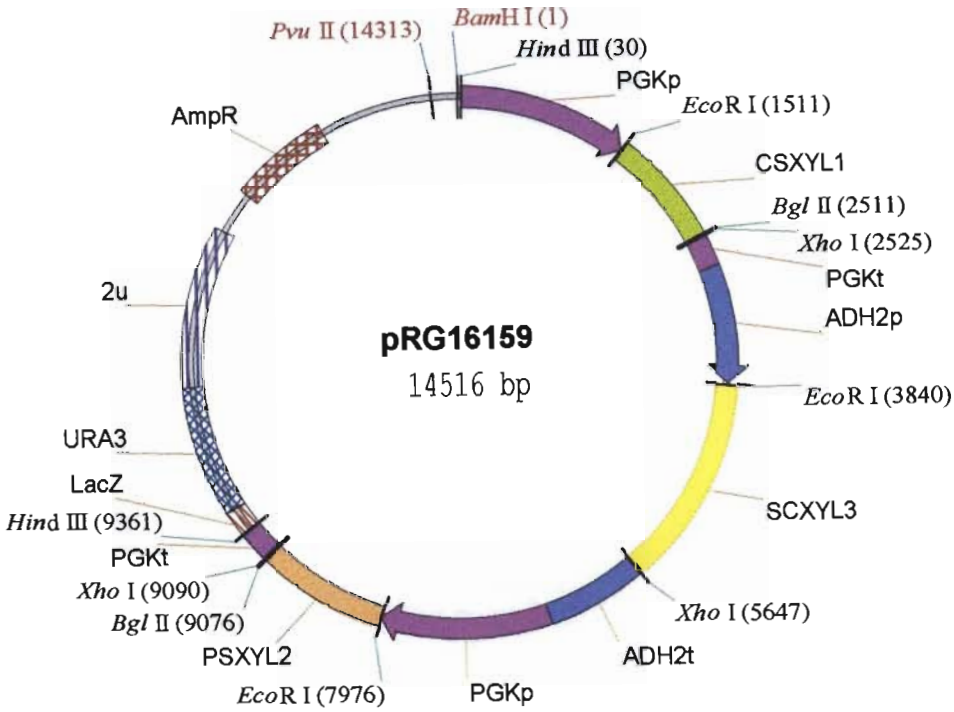


Fig. 3.7 Map of expression vector containing the xylose metabolising gene cassettes. An *ApaI*/*Not*I fragment containing the *C. shehatae XYL1* genes under control of the *PGK1* promoter/terminator cassette and the *S. cerevisiae XYL3* gene under control of the *ADH2* promoter/terminator cassette was blunt-ended and cloned into the *Sma*I site of pRG9.

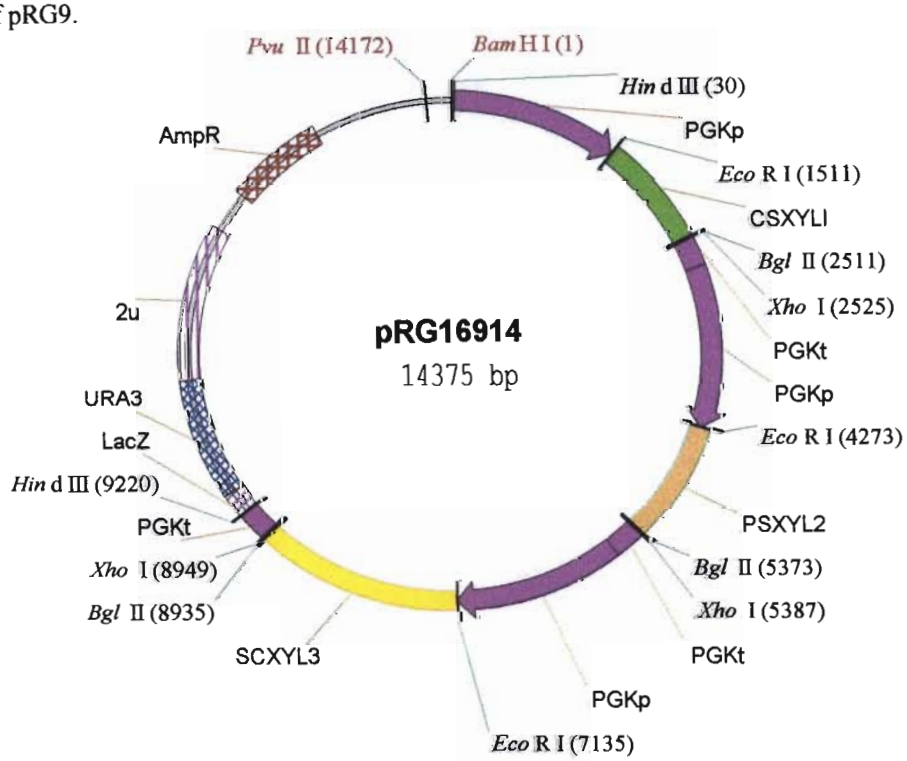


Fig. 3.8 Map of expression vector containing the xylose metabolising gene cassettes. A 5.3 kb *Cla*I/*Not*I fragment containing the *C. shehatae XYL1* and *P. stipitis XYL2* genes under control of the *PGK1* promoter/terminator cassettes was blunt-ended and ligated to the *Sma*I site of pRG14 to create pRG16914.

3.3.2 HETEROLOGOUS GENE EXPRESSION

For all three enzymes assayed, higher specific activity was obtained when the genes were under the control of the *PGK1* promoter than the *ADH2* promoter. The recombinant strain Y294:pRG16 displayed the highest activity with 15 and 9.45 U/mg with NADPH and NAD, respectively (Table 3.3). The next highest activity was obtained for *P. stipitis* - 7.64 and 10.48 U/mg with NADH and NADPH, respectively. The XR activity of *P. stipitis* with either cofactor was higher than that of *C. shehatae* by approximately 1 U/mg. Untransformed *S. cerevisiae* has almost a hundred-fold lower activity with NADPH than *P. stipitis* and *C. shehatae*. The activity with NADH was even lower - almost 250-fold.

TABLE 3.3 Xylose reductase activity of untransformed and recombinant strains of *S. cerevisiae*

Strain	Vector	Specific Activity (U/mg*)		
		NADH	NADPH	NADH/NADPH
P. stipitis PsY633	-	7.64	10.48	0.73
C. shehatae CsY051	-	5.72	9.39	0.61
S. cerevisiae Y294	-	0.004	0.11	0.036
S. cerevisiae Y294	pRG6	ND**	2.95	ND**
S. cerevisiae Y294	pRG8	6.32	8.67	0.73
S. cerevisiae Y294	pRG16	9.45	15	0.63
S. cerevisiae Y294	pRG17	0.17	0.29	0.59

*Units = U/mg (μmoles of converted substrate per mg protein per ml)
**ND = Not detected

The XDH activity was slightly higher in the recombinant strain Y294:pRG9 compared to *P. stipitis* (Table 3.4). The recombinant XDH activity with the *PGK1* promoter was almost ten-

fold that with the *ADH2* promoter. XK activities of the untransformed and recombinant strain Y294:pRG15 were similar (Table 3.5). XK activity controlled by the *PGK1* promoter was over five times higher than that of the untransformed strain or the recombinant strain Y294:pRG15.

TABLE 3.4 Xylitol dehydrogenase activity of untransformed and recombinant strains of *S. cerevisiae*

Strain	Vector	Specific Activity (U/mg*)
P. stipitis PsY633	-	13.79
S. cerevisiae GPY155-15β	-	0.13
S. cerevisiae GPY155-15β	pRG7	1.86
S. cerevisiae GPY155-15β	pRG9	14.89

*Units = U/mg (μmoles of converted substrate per mg protein per ml)

TABLE 3.5 Xylulokinase activity of untransformed and recombinant strains of *S. cerevisiae*

Strain	Vector	Specific activity (U/mg*)
S. cerevisiae Y294	-	0.68
S. cerevisiae Y294	pRG14	4.45
S. cerevisiae Y294	pRG15	0.86

*Units = U/mg (μmoles of converted substrate per mg protein per ml)

The XR activities of both Y294:pRG16159 and Y294:pRG16914 strains were slightly lower (Table 3.6) than the recombinant strains Y294:pRG16 harbouring only the gene coding for XR (Table 3.3). The XK activity on the other hand was very similar for the recombinant strains Y294:pRG15 and Y294:pRG16159. XK activity under control of the *PGK1* promoter is 25% higher in the recombinant strain (Y294:pRG16914) expressing all three genes than in the strain

(Y294:pRG14) expressing XK alone.

Table 3.6 Xylose reductase and xylulokinase activities of untransformed and recombinant strains of *S. cerevisiae*

Strain	Vector	XR Activity (U/mg*)		XK Activity (U/mg*)
		NADH	NADPH	
S. cerevisiae Y294	-	0.004	0.11	0.68
S. cerevisiae Y294	pRG16159	11.7	13	0.8
S. cerevisiae Y294	pRG16914	10.5	14.1	5.68

Units = U/mg (μmoles of converted substrate per mg protein per ml)

3.4 DISCUSSION

The easiest fragments to clone carry non-complementary termini generated by digestion with two different restriction enzymes. This strategy was followed whenever possible in the construction of the expression vectors. Construction of the single gene expression cassettes in the yeast expression vector was fairly simple since the PCR products were designed for unidirectional cloning into the vectors pDLG1 and pJC1. Even in the cases where single enzymes were used to generate the inserts and vectors, the resultant cohesive termini were not difficult to ligate. The 5'-phosphate groups of the linearized vectors were removed with alkaline phosphatase to prevent re-circularization and thus increase ligation efficiency.

Attempts to clone a second expression cassette into the yeast expression vector were not successful. The ligation efficiency was relatively good but numerous products were obtained that were bigger in size than the original vector but smaller than the predicted size of the vector containing a double-gene cassette. Since 2μ-based plasmids could be present at an average copy

number ranging between 10 and 40 per cell, this phenomenon could be due to rearrangements between copies of the vector. Rose and Broach (1991) reported rearrangements between endogenous 2μ plasmids and pJDB219, a vector containing the entire 2μ circle, the *LEU2* allele and unique cloning sites.

A second cloning strategy, viz., cloning the different genes with their expression cassettes into pBS and then transferring the two-gene and three-gene constructs into the shuttle vector proved more successful since three different triple-gene constructs were obtained. However, when the numerous ligation reactions that were performed were taken into account, even these reactions were not very efficient. Very few restriction sites were available for DNA manipulations since most of the enzyme sites on the polylinker were also present in one or more of the genes to be cloned. Expression cassettes were removed using the appropriate enzymes and their termini were then blunt-ended. DNA fragments with recessed 3' termini were filled-in using the Klenow fragment of DNA polymerase I. Linearized vectors were generated with enzymes producing blunt ends. Blunt-ended ligations are the most difficult of the three types of ligations but these ligations were rendered even more difficult since the manipulation steps (dephosphorylation and filling-in) further decrease the ligation frequency. The three triple-gene constructs obtained were produced with vector:insert termini ratios of 1:3 in the ligation reactions. This is the recommended ratio for maximum monomeric plasmid:foreign DNA chimeras (Sambrook *et al.*, 1989). However, for greater success with blunt ended ligations involving large insert sizes, a vector: insert ratio of 1:30 is required (Jacqui Goodwin, Roche Diagnostics, personal communication).

The insert size of the triple gene constructs (between 9.3 kb and 10.5 kb), as well as the shuttle vector (5.13 kb), made ligations difficult. The mechanical damage of the triple-gene cassettes during the purification steps after endonuclease digestion and agarose gel electrophoresis may have also contributed to poor ligation efficiency, since the purification was performed using

silica beads. This was further aggravated by the fact that they were blunt-ended ligations. Even with the aid of a ligation kit, which contained increased concentrations of the ligase enzyme (five times higher than normally used) and incorporated polyethylene glycol in the ligation buffer, which serves to concentrate the DNA species or bring the DNA termini closer together, these reactions were still very difficult.

Both the fragment size and the particular sequences cloned into a YEp vector can influence stability and copy number. As a general rule of thumb, the smaller the YEp plasmid, the higher the stability and copy level. In addition, insertion of a gene whose presence is deleterious to the host strain yields diminished stability and reduced copy levels, even under growth conditions selective for retention of the plasmid (Rose and Broach, 1991).

After transfer of the recombinant plasmids into *S. cerevisiae*, the expression of the genes were confirmed by enzyme assays. Instead of verifying positive transformants by Southern hybridizations and investigating heterologous protein synthesis by SDS polyacrylamide gel electrophoresis, this technique provides a direct indication of whether the transferred genes are functional or not. The general trend observed was that specific enzyme activities were higher when the genes were under the control of the *PGK1* promoter than the *ADH2* promoter. All recombinant strains demonstrated higher XR activity than the untransformed strains. The specific XR activity of the recombinants harbouring the *C. shehatae* gene controlled by the *PGK1* promoter had twice the activity of that harbouring the *P. stipitis* gene controlled by the same promoter. This trend was observed for both cofactors. All recombinant strains had higher specific activity with NADPH than NADH. NADH/NADPH ratios were 0.73 for *P. stipitis* and the Y294:pRG8 transformant, 0.61 for *C. shehatae* and 0.63 and 0.59 for the transformants carrying the *C. shehatae* gene under the control of the *PGK1* and *ADH2* promoter/terminator cassettes, respectively. The NADH/NADPH ratios obtained in this study are higher than those reported in

the literature (Alexander *et al.*,1988c - 30% or Bruinenberg *et al.*, 1984 and Ho *et al.*, 1990 - 40%). This discrepancy may be attributed to differences in XR activities in different strains. Amore *et al.* (1991) reported a variation in NADH/NADPH ratios between 0.60 and 0.92 for different cell extracts of *P. stipitis* and 0.66 and 0.96 for different *S. cerevisiae* recombinant strains harbouring the *P. stipitis XYL1* gene.

The level of production of the recombinant enzymes is not only dependent on the promoter that controls gene expression and efficiency of translation but is also influenced by the copy number and stability of the vector. The 2μ origin present in the vectors used in this study results in an average copy number of 10 to 40 per cell but with a large variation from cell to cell (Futcher and Cox, 1984). This difference in copy number could explain why the *P. stipitis XYL1* controlled by the *ADH2* promoter had approximately 20 times the activity of the *C. shehatae XYL1*. Although XR activity was slightly higher in *P. stipitis* than *C. shehatae*, the activity of the recombinant XR from *C. shehatae* under the control of the *PGK1* promoter was double that of the recombinant XR from *P. stipitis*. A plausible explanation for this phenomenon would be that the transcription from the promoter varies for both genes due to a difference in transcription efficiency. The *C. shehatae XYL1* gene was inserted into the *Eco*RI site within the expression cassette while the *P. stipitis XYL1* gene was inserted into the *Xho*I site. The *Eco*RI recognition sequence, GAA TTC, probably provides a better 5' region (TATA box) than the *Xho*1 recognition sequence, CTC GAG. Moreover, since the *Eco*RI site is 5' to the *Xho*I site in the expression cassette, cloning in the former site will bring the ATG codon of the gene closer to the promoter sequences, possibly in a more desirable position than cloning in the *Xho*I site would. Another reason for the different levels of expression could be that the copy number of the recombinant vectors varied in the host strain. This is also supported by the fact that various clones from the different recombinant strains produced varying enzyme activities. A comparison of transcriptional

activity and enzyme activity will allow one to establish the level at which gene expression is controlled for these heterologous genes..

The *XYL2* gene expression vector was initially transformed into *S. cerevisiae* Y294. However, non-specific dehydrogenase activity in this strain was very high and it was very difficult to obtain recombinant activity readings. Therefore, *S. cerevisiae* GPY155-15β, a strain with lower non-specific dehydrogenase activity was used. XDH activity under control of the *PGK1* promoter in the recombinant strain was slightly higher than that in *P. stipitis*. In comparison, XDH activity was over ten-fold higher when under the control of the *PGK1* promoter than the *ADH2* promoter.

S. cerevisiae strains possess a chromosomal copy of the *XYL3* gene (Ho and Chang, 1989). XK activity in *S. cerevisiae*, however, is very low (almost thirtyfold lower) compared to that in *P. stipitis* and *C. shehatae* (Deng and Ho,1990; Flanagan and Waites, 1992). The recombinant strains therefore contain two types of the gene - a single chromosomal copy and copies of the gene on the vector. The XK activity controlled by the *ADH2* promoter was only slightly higher than that of the parental strain while that controlled by the *PGK1* promoter was five-fold higher. The low recombinant enzyme activity values obtained when the genes were controlled by the *ADH2* promoter was due to the strains being grown in glucose. The *ADH2* promoter is strongly repressed by glucose and the low levels of enzyme were probably produced once glucose was depleted in the medium. The *PGK1* promoter on the other hand is induced by glucose and therefore the enzymes levels are higher than those controlled by the *ADH2* promoter.

Once it was established that the xylose-metabolizing genes did indeed encode functional proteins that were efficiently processed by the recombinant strains into active proteins, the expression cassettes were combined to produce vectors containing all three xylose-metabolizing genes. XK activities were higher in the recombinant strain co-expressing all three genes than the recombinant strain expressing only XK. The presence of other promoter/terminator sequences

apparently produced a positive effect on the *XYL3* gene. A similar result was reported by Walfridsson *et al.* (1997) with sequences in the *ADH1* promoter enhancing the activity from the *PGK1* promoter. Once the activity of the individual enzymes encoded by triple-gene expression vectors were established, several transformants were screened for enzyme activity. The most likely candidates, displaying high enzyme activity, were then chosen for fermentation experiments for the conversion of xylose to ethanol. These experiments will be discussed in the following Chapter.

CHAPTER FOUR

XYLOSE METABOLISM BY RECOMBINANT *Saccharomyces cerevisiae*

4.1 INTRODUCTION

The construction of a xylose-utilizing *S. cerevisiae* strain has been the focus of research since 1986. Initial attempts that involved cloning bacterial xylose isomerase genes in *S. cerevisiae* (Amore *et al*, 1989; Moes *et al.*, 1996; Sarthy *et al.*, 1987) were unsuccessful. Several researchers cloned the *P. stipitis XYL1* gene (Amore *et al.*, 1991; Takuma *et al.*, 1991) and *XYL2* gene (Hallborn *et al.*, 1991; Kötter *et al.*, 1990; Tantirungkij *et al.*, 1993) and expressed them in *S. cerevisiae*. Although the recombinant *S. cerevisiae* strain harbouring both genes produced a level of enzyme activity comparable to that of *P. stipitis in vitro*, growth on xylose was slow and ethanol production low (Kötter *et al.*, 1990). Although the genes were expressed under the control of their native promoters which are inducible by aldoses, their activity was constitutive in *S. cerevisiae*. The authors attributed this weak expression to the absence of xylose-specific activator proteins in *S. cerevisiae*. However, they found that expression of these two genes in a *S. cerevisiae* strain with a five-fold higher XK activity, significantly improved both growth on xylose and xylose utilization.

Tantirungkij and co-workers (1994a) isolated mutants after EMS treatment of a recombinant *S. cerevisiae* strain (carrying *P. stipitis XYL1* and *XYL2*). The mutants had a higher ratio of XDH to XR activity as well as higher XK activity compared to the parent strain. Mutants also showed an improved yield and production rate of ethanol compared to the parent strain. Kötter and Ciriacy (1993) compared xylose utilization and product formation in *S. cerevisiae XYL1/XYL2* transformants and *P. stipitis*. They concluded that conversion of xylose to ethanol in *S. cerevisiae* was limited by cofactor imbalance and by an insufficient capacity of xylulose

conversion by the PPP. This is in agreement with the results of Tantirungkij *et al.* (1994a; 1994b) who demonstrated that strains with higher XK activity were better producers of ethanol from xylose.

Since the original promoters of the *XYL1* and *XYL2* genes were not active in *S. cerevisiae*, the next step towards improving heterologous expression of these two genes were to express them under the control of strong glycolytic promoters. Walfridsson *et al.* (1995) expressed both genes under the control of the *ADH1* and *PGK1* promoters. When *XYL1* was expressed under the control of the *ADH*1 promoter and *XYL2* was expressed under the control of the *PGK1* promoter, XR and XDH activities were 0.5 and 30.1 U/mg, respectively. Activities of 4.2 and 5.8 U/mg were obtained when *XYL1* was expressed under the control of the *PGK1* promoter and *XYL2* under the control of the *ADH1* promoter, respectively. Depending on the choice of promoter, different ratios of XR and XDH activities were produced in the transformants. Although XR and XDH were expressed constitutively and xylulose was metabolized further, the transformants showed limited growth on xylose and ethanol was still not produced. Overexpressing *TAL* in *XYL1/XYL2* transformants increased the growth rate on xylose by increasing the flux through the PPP. With decreasing oxygenation, biomass formation was reduced but xylitol formation increased.

Oxygen-limited fermentation of a glucose/xylose mix improved the rate of xylose utilization (Walfridsson *et al.*, 1997). Both glucose and xylose were consumed simultaneously although the rate of xylose utilization decreased once glucose was depleted. Two reasons were suggested for this phenomenon. Firstly, that xylose transport was limited due to the low xylose concentration and secondly, that fermentation of xylose alone does not produce sufficient intermediary metabolites for the induction of key ethanologenic enzymes.

Xylose transport into the recombinant strains is the first step of xylose utilization, whether

the desired end-product is xylitol or ethanol. Therefore, the effectiveness of xylose transport is crucial in the rate of product formation and yield of the final product. Several sugars affect the rate of transport of xylose. In *S. cerevisiae* xylose is transported *via* the facilitated diffusion system which also transports glucose, fructose, mannose and galactose (Busturia and Lagunas, 1986; Leao and van Uden, 1982; van Zyl *et al.*, 1989). Depending on growth conditions and strains, the transport system has various affinity constants for the different sugars, with the highest affinity for glucose (Bisson and Fraenkel, 1984; Does and Bisson, 1989) and the lowest for xylose (Busturia and Lagunas, 1986; van Zyl *et al.*, 1993). It is therefore expected that, at high concentrations, these sugars may inhibit xylose uptake. Glucose inhibited xylose conversion almost completely (Meinander and Hahn-Hägerdal, 1997a) while fructose and mannose, in agreement with their lower affinities had a lower effect. Maltose did not affect xylose transport since it is transported by a specific transporter.

Recombinant *S. cerevisiae* expressing the *XYL1* gene only, cannot grow on xylose as a sole carbon source but would be able to convert xylose to xylitol in the presence of a co-substrate (Hallborn *et al.*, 1994; Meinander *et al.*, 1994). A *S. cerevisiae* strain expressing the *XYL1* gene converted xylose to xylitol with a 1:1 yield, because xylitol cannot be metabolized due to the absence of XDH (Hallborn *et al.*, 1991). A co-substrate like glucose is needed for regeneration of the reduced cofactor required in the reaction and to supply metabolic energy required for growth and maintenance (Meinander and Hahn-Hägerdal, 1997a). Several factors have been identified as being important for the rate of xylose conversion by *XYL1*-expressing *S. cerevisiae*, including: (i) the rate of xylose uptake; (ii) the XR activity; and (iii) the supply of reduced co-factors (Meinander and Hahn-Hägerdal, 1997b; Meinander *et al.*, 1996).

During the course of this study, Ho *et al.* (1998) reported the conversion of xylose to ethanol by a recombinant *Saccharomyces* spp. strain carrying the *XYL1*, *XYL2* and *XYL3* genes.

These workers compared the expression of *XYL1* in *S. cerevisiae* firstly under the control of the *ADH1* promoter and secondly using the 5' noncoding sequence of *P. stipitis XYL1* from -1 to -50 followed by the *ADH1* promoter. Both constructs conferred similar xylose fermentation capabilities to their host strain. They also expressed *P. stipitis XYL2* and *S. cerevisiae XYL3* genes under the control of the *PYK* promoter. The *Saccharomyces* spp. strain was transformed with vectors bearing these three gene constructs and transformants were capable of co-fermentation of glucose and xylose to ethanol in 48 h. The recombinant strains could also effectively ferment xylose in the absence of glucose with very little xylitol production.

It is clearly evident that the over-expression of *XYL3*, in conjunction with the first two genes of xylose metabolism in *S. cerevisiae*, overcame the problem of slow flux through the PPP. This chapter compares the conversion of xylose to xylitol by recombinant *S. cerevisiae* strains carrying the *XYL1* genes from *C. shehatae* and *P. stipitis*. The fermentation of xylose to ethanol by recombinant *S. cerevisiae* strains containing the first three genes in xylose metabolism is also described. This study differs from others in that the heterologous activity of the *XYL1* gene from *C. shehatae* has not been reported previously. Valuable information may be obtained by comparing the heterologous enzyme activities of the *P. stipitis* and *C. shehatae* XR enzymes which have a differing cofactor preference in their native strains as well as their effectiveness in converting xylose to xylitol in the host strains.

4.2 MATERIALS AND METHODS

4.2.1 SHAKE-FLASK FERMENTATION

Shake-flask fermentations were performed to demonstrate (i) xylitol production by the *XYL1* transformants; and (ii) ethanol production by the recombinants carrying all three xylose-metabolizing genes. Inocula were prepared in two stages. A loopful of culture was inoculated

into 5 ml SC^{-URA} medium containing 2% glucose, incubated O/N at 30°C at 200 rpm, subsequently transferred to 50 ml fresh medium in a 250 ml Erlenmeyer flask and incubated. After 24 h, cells were pelleted by centrifugation at 4 000 × *g* for 10 min and the pellets resuspended in SC^{-URA} medium. For xylitol and ethanol production, 50 ml of fresh SC^{-URA} medium containing 50 g/l xylose and either 10 g/l or 20 g/l glucose, galactose, or maltose in 125 ml screw-cap Erlenmeyer flasks were used. Fermentations were carried out in triplicate at 30°C with shaking at 200 rpm. Aliquots of 500 µl were sampled every 8 h. A 100 µl aliquot was removed for biomass determination (Section 4.2.2) and the remaining samples were centrifuged at 8 000 × *g* for 30 min and the supernatants stored at -20°C until required for analyses (Section 4.3.3).

4.2.2 BIOMASS DETERMINATION

Absorbance measurements at 600 nm were performed on serial dilutions of a 24 h culture of *S. cerevisiae*. One ml aliquots of each dilution were filtered onto pre-weighed Millipore type HA filters (0.45µm), dried in an oven at 105°C, allowed to cool in a dessicator and their masses determined. A dry cell mass versus absorbance standard graph was drawn. Subsequently, the absorbance of the fermentation samples was measured at 600 nm and dry cell mass determined from this curve.

4.2.3 ANALYSIS OF SUGARS AND ETHANOL

Frozen fermentation samples were allowed to thaw, centrifuged to remove particulate matter and analyzed using a high pressure liquid chromatography (HPLC) system equipped with a refractive index detector (Shimadzu) and an autosampler (Shimadzu). Glucose, xylose, mannose, xylitol, ethanol and glycerol were separated using two columns. The first, a Sugar-Pak I column (Waters) was maintained at 90°C, the mobile phase was 50 ppm calcium titriplex

dihydrate maintained at 50°C. The second column was a Resex 8μ, 8% calcium monosaccharide column (Phenomenex). This column was maintained at 80°C and 100 ppm calcium titriplex dihydrate maintained at 50°C was used as the mobile phase. The flow rate was maintained at 0.5 ml/min, an injection volume of 20 μl was used and the run time was 25 min. A Waters guard column was used to protect the column since the samples were not filtered before injection. Results were recorded by a Shimadzu Class VP Data Chromatography Workstation.

4.3 RESULTS

4.3.1 XYLITOL PRODUCTION

All the strains produced similar biomass profiles for the different sugars utilized (Fig. 4.1). Both glucose and galactose utilization as co-substrates produced similar profiles while growth on maltose was slower. The total biomass approximately doubled for the untransformed strain and the recombinant strain Y294:pRG8 during the fermentations on glucose and galactose. When glucose was the co-substrate, all strains produced biomass most rapidly during the first 8 h. After this period, which corresponded to glucose depletion, biomass remained constant. All the strains showed only a slight increase in biomass on maltose (an increase of approximately one third of the inoculated values - Fig. 4.1).

Both the wild type and recombinant strains took up xylose. Xylose uptake by the recombinant strains was slightly more efficient with between 12 and 20 g/l xylose transported into the cells over the 70 h fermentation period. Xylose uptake by Y294:pRG16 was most efficient when glucose was the co-substrate (Fig. 4.2c). Xylose transport was very rapid in the first 16 h of the fermentation while the co-substrate was still present in the medium (Fig. 4.2 b, 4.2c). After 32 h, xylose uptake decreased significantly for both the recombinant and wild type strains with all the media types. Glucose was consumed most rapidly - it was totally depleted within

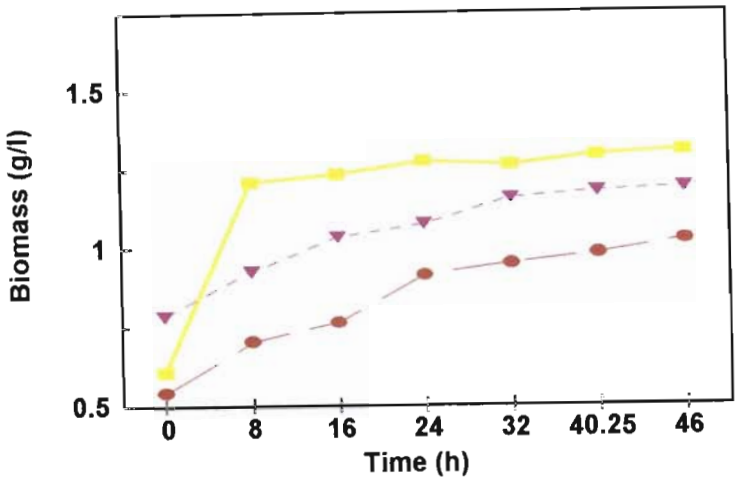
8 h (Fig. 4.2).

Although 10 g/l xylose was taken up by the untransformed *S. cerevisiae* cells, only a small amount of xylitol (up to 2 g/l) was formed in all the media types. The recombinant strains, on the other hand, converted xylose to xylitol with yields that were almost theoretical (1:1). Xylitol production was most rapid in the first 8 h of fermentation when glucose was the co-substrate (Figs.4.2b, 4.2c). The wild type strain produced xylitol slowly and steadily throughout the fermentation (Fig. 4.2a). When glucose was the co-substrate, 3 - 5 g/l ethanol was accumulated during the first 8 h by Y294 (Fig. 4.2a) and Y294:16 (Fig. 4.2c) and in 16 h by Y294:pRG8 (Fig. 4.2b). The ethanol was gradually utilized thereafter. Low concentrations of ethanol were produced when maltose and galactose were the co-substrates. Glycerol was also produced in very small amounts by all strains.

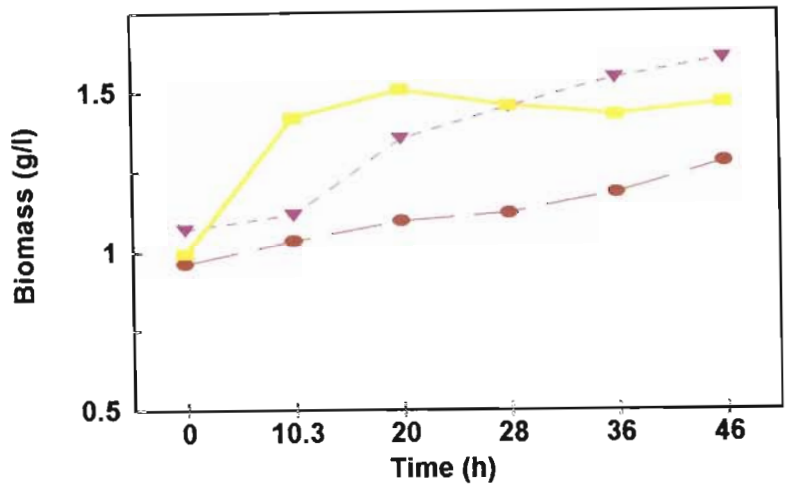
It was not possible to differentiate between galactose and xylose using the SugarPak column since both sugars displayed identical retention times when this column was used. Final xylitol concentrations, however, were lower for galactose than those obtained on glucose and for Y294:pRG8 (8 g/l - Fig. 4.3b) and Y294:pRG16 (13 g/l - Fig. 4.3c), respectively. The recombinant strain carrying the *C. shehatae XYL1* gene produced higher xylitol concentrations (Fig. 4.3 c). Glycerol formation in the recombinant strains was lower than in the untransformed strain (Fig. 4.3).

Maltose was utilized slowly compared to the other two substrates. At the end of the fermentation, residual maltose was present for all strains. Y294:pRG8 and Y294:pRG16 produced 9 and 10 g/l xylitol in this medium. Production occurred at a slower rate than for the other two sugars. There appeared to be a "lag" in the first 8 h of the fermentation for the Y294:pRG8 transformant. During this period, xylose uptake was slow, maltose did not appear to be taken up and xylitol formation was slow (Fig. 4.4b). The other recombinant strain did not display this effect (Fig. 4.4c).

a



b



c

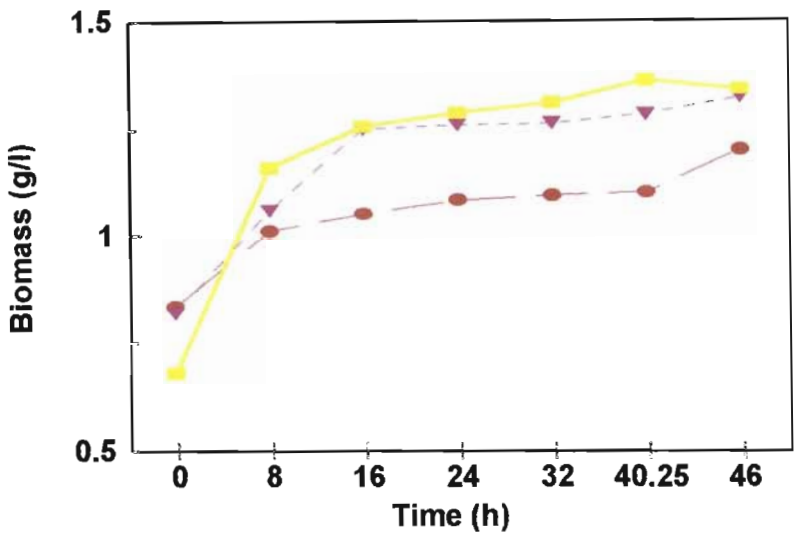
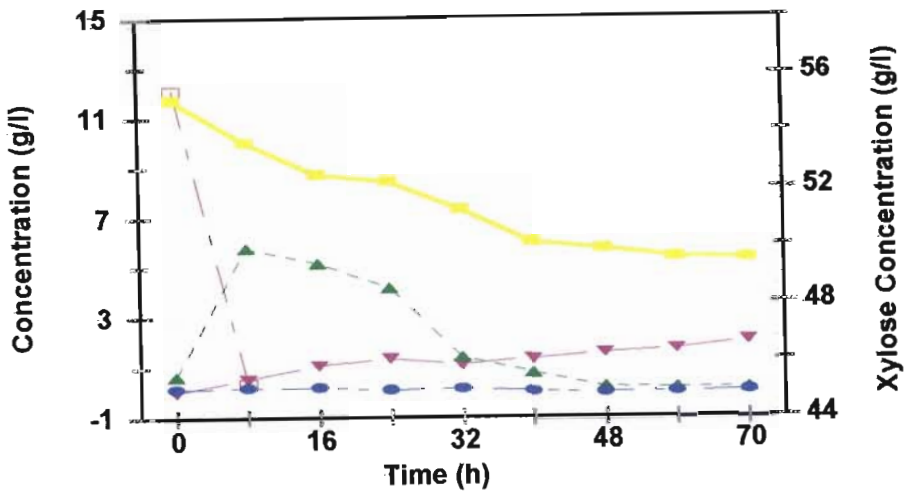
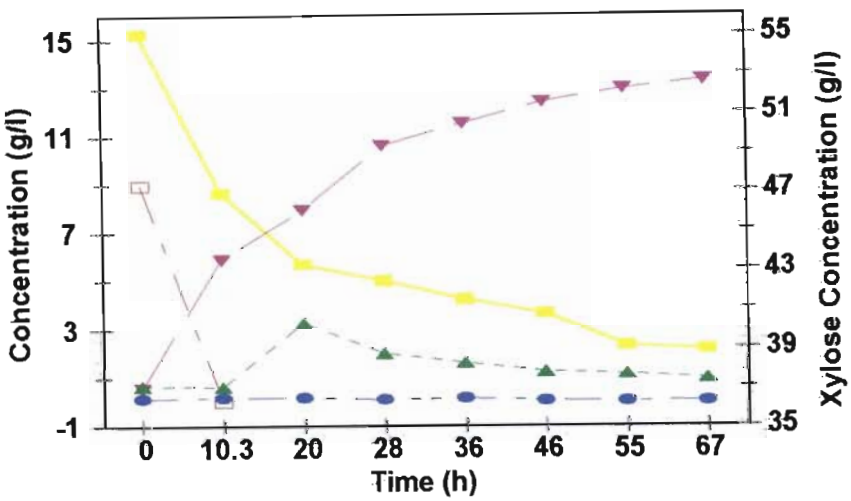


FIG. 4.1 Biomass profiles of shake-flask xylose fermentations for xylitol production by (a) *S. cerevisiae* Y294 and recombinant strains carrying the (b) *P. stipitis XYL1* gene and the (c) *C. shehatae XYL1* gene. Co substrates were glucose (■), galactose (▼) and maltose (●).

a



b



c

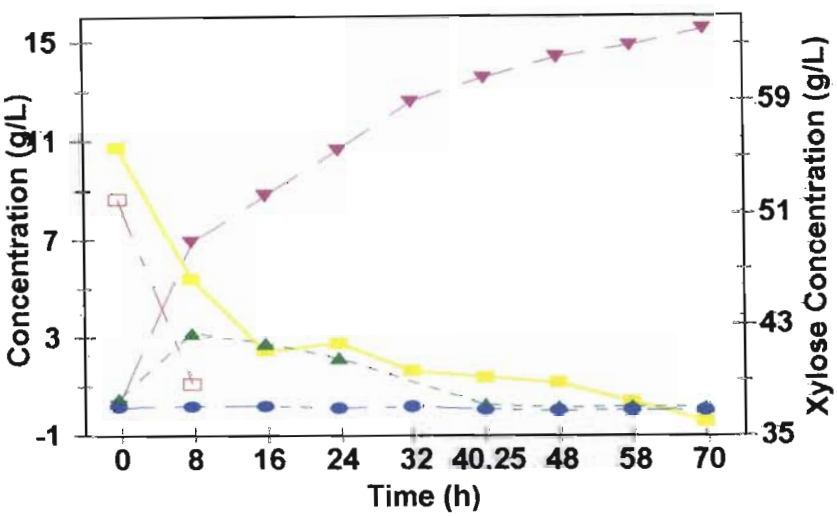


FIG 4.2 Shake-flask xylose fermentation profiles of xylitol production by (a) *S. cerevisiae* Y294 and recombinant strains carrying the (b) *P. stipitis XYL1* gene and the (c) *C. shehatae XYL1* gene with glucose as the co-substrate. The utilization of xylose (■) and glucose (□) and the production of xylitol (▼), ethanol (▲) and glycerol (●) were monitored.

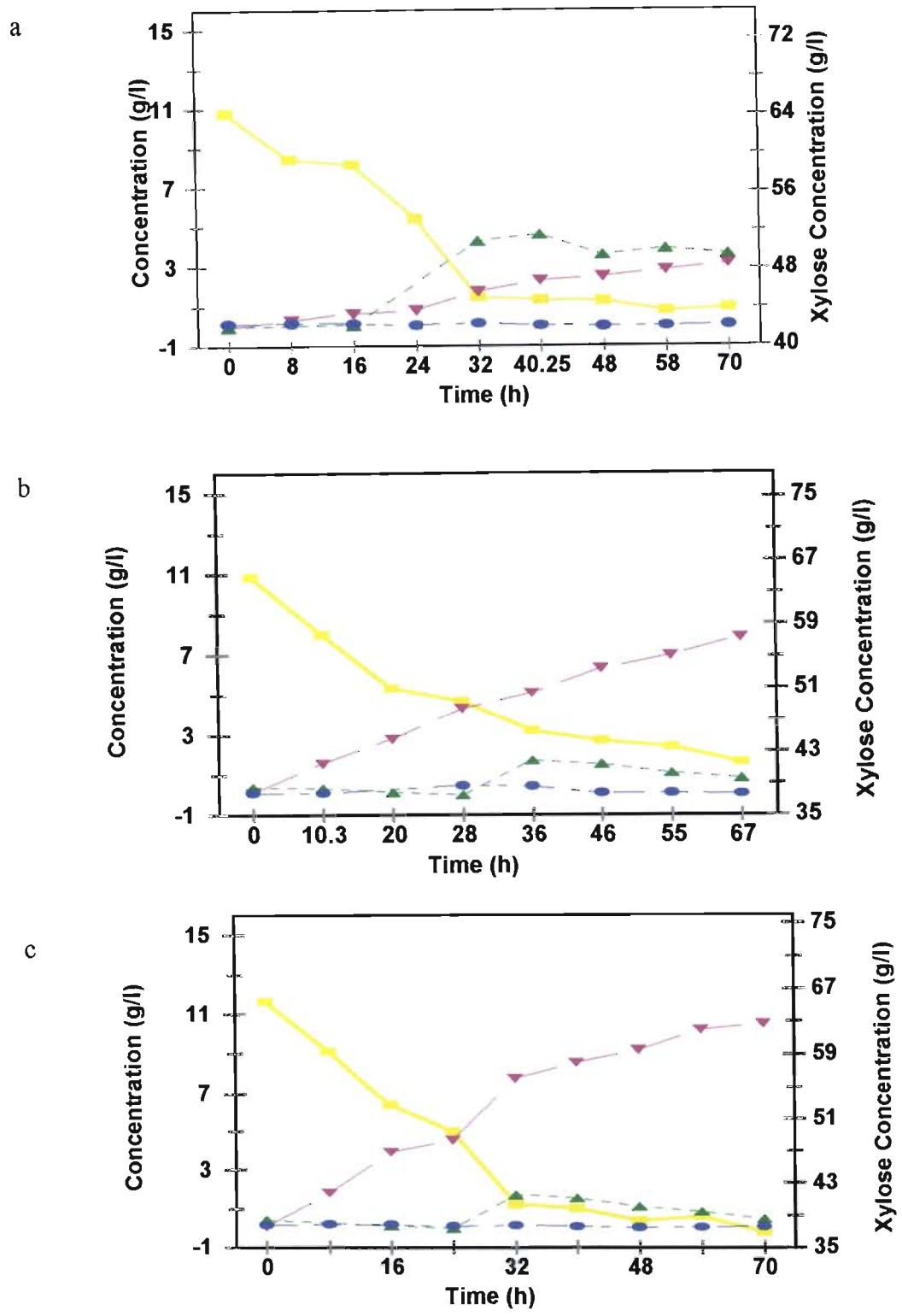
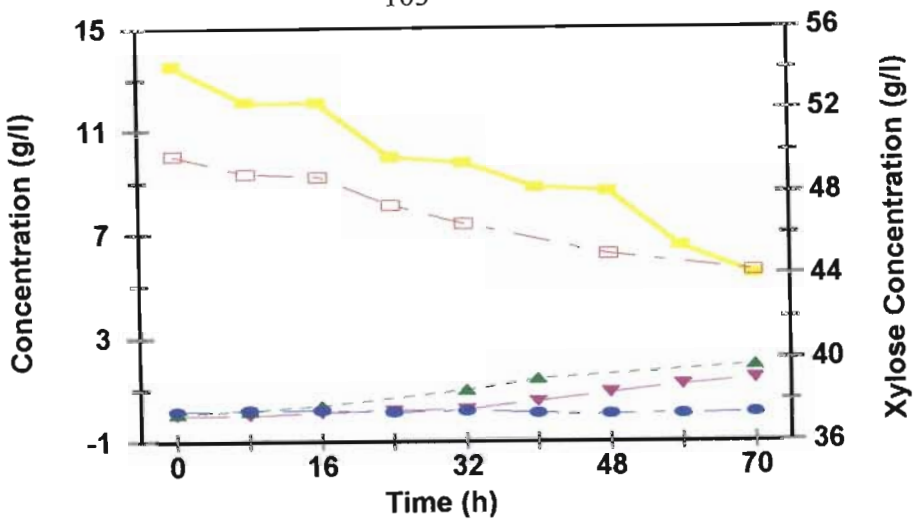
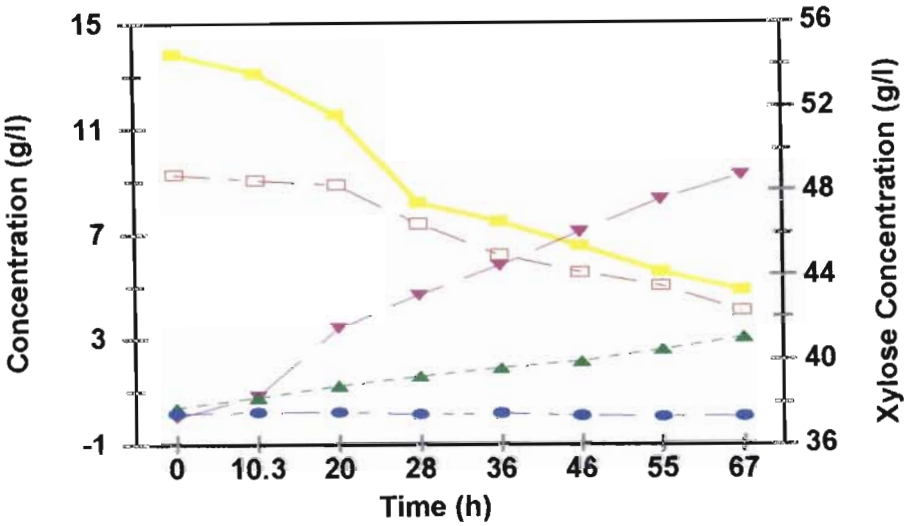


FIG 4.3 Shake flask xylose fermentation profiles for xylitol production by (a) *S. cerevisiae* Y294 and recombinant strains carrying the (b) *P. stipitis XYL1* gene and the (c) *C. shehatae XYL1* gene with galactose as the co-substrate. The utilization of xylose (■) and the production of xylitol (▼), ethanol (▲) and glycerol (●) were monitored. The xylose concentration also includes galactose.

a



b



c

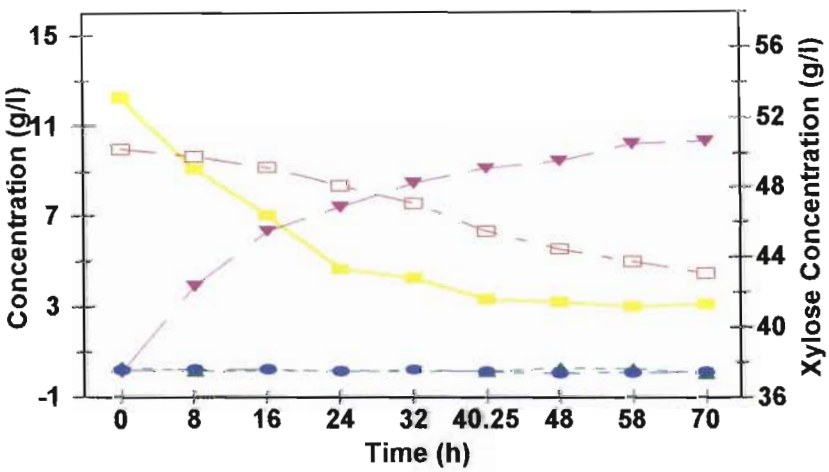


FIG 4.4 Shake flask xylose fermentation profiles for xylitol production by (a) *S. cerevisiae* Y294 and recombinant strains carrying the (b) *P. stipitis XYL1* gene and the (c) *C. shehatae XYL1* gene with maltose as the co-substrate. The utilization of xylose (■) and glucose (□) and the production of xylitol (▼), ethanol (▲) and glycerol (●) were monitored.

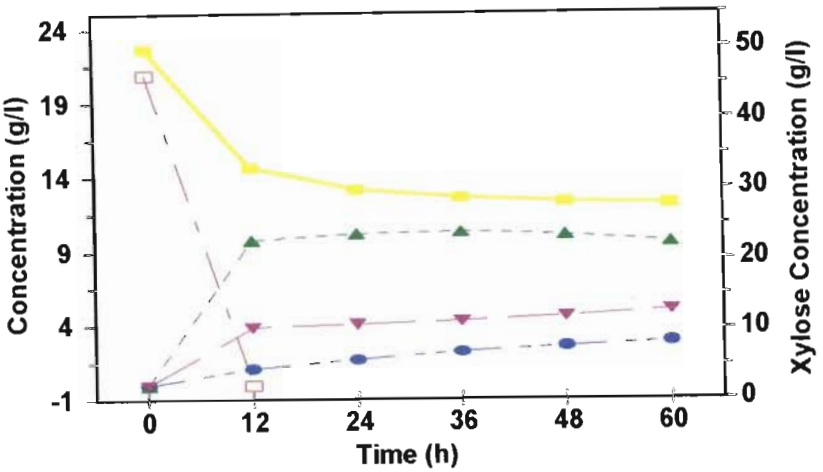
4.3.2 ETHANOL PRODUCTION

All three strains produced ethanol during fermentation on glucose and xylose. However, only the recombinant strains produced ethanol from xylose. Both Y294 and Y294:pRG16159 produced approximately 10 g/l ethanol in the first 12 h. Y294:pRG16914 produced approximately 12 g/l in the same time interval, the additional 2 g/l ethanol was produced from xylose fermentation. Overall, Y294:pRG16159 produced only slightly more ethanol than the untransformed strain with final ethanol concentrations being 13 g/l and 10.5 g/l, respectively. Y294:pRG16914 on the other hand showed an ethanol yield of approximately 18.6 g/l, almost double that of the untransformed strain. Glycerol and xylitol were also produced by all three strains. Y294:pRG16159 produced substantial amounts of xylitol, followed by Y294:pRG16914 and the untransformed strain.

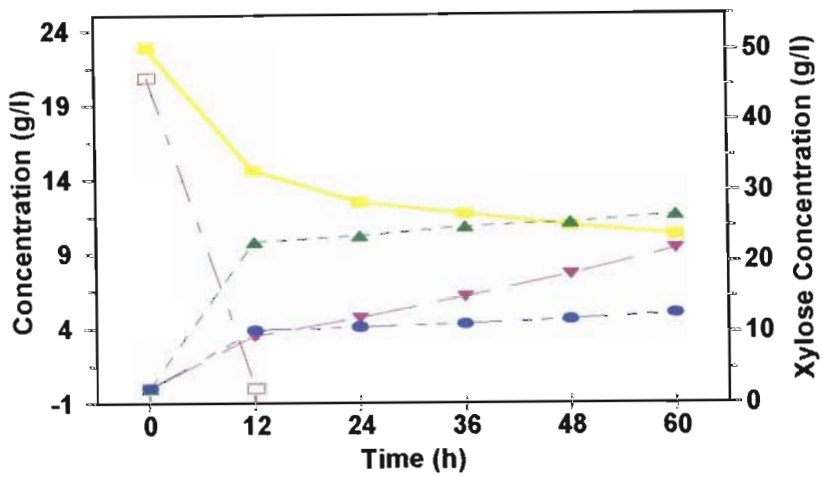
4.3.3 GROWTH ON XYLOSE

The untransformed strain was not capable of aerobic growth on xylose as a sole carbon source. Its biomass remained constant throughout the 28 h sampling period (Fig. 4.6). Both recombinant strains were capable of growth on xylose. Y294:pRG16914 produced more than double the biomass of Y294:pRG16159 (Fig. 4.6).

a



b



c

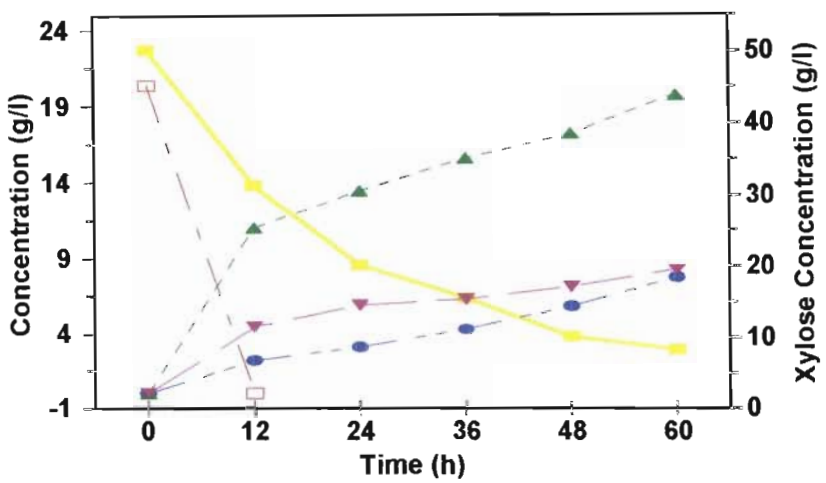


FIG. 4.5 Shake flask xylose fermentation profiles for ethanol production by (a) *S. cerevisiae* Y294 and recombinant strains (b) Y294:pRG116159 and (c) Y294:pRG16914 with glucose as the co-substrate. The utilization of xylose (■) and glucose (□) and the production of xylitol (▼), ethanol (▲) and glycerol (●) were monitored.

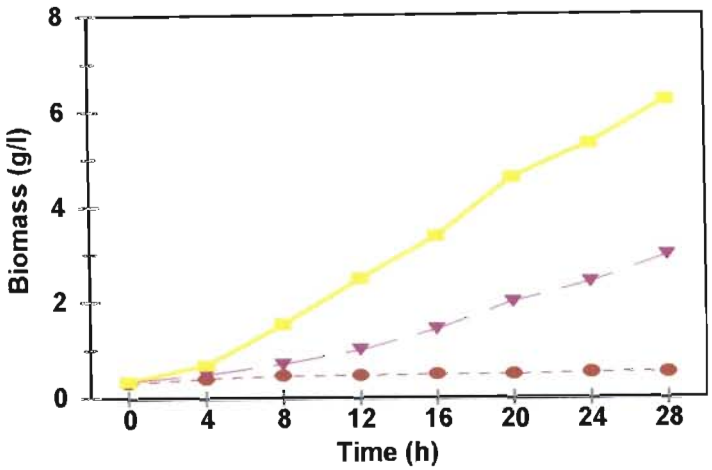


Fig. 4.6 Growth of *S. cerevisiae* Y294 (●) and recombinant strains Y294:pRG116159 (▼) and Y294:pRG16914 (■) on xylose as the sole carbon source.

4.4 DISCUSSION

The recombinant strains carrying only the *XYL1* genes cannot metabolize xylose. A co-substrate is therefore required for the generation of reduced co-factors [NAD(P)H] required by XR for growth and maintenance. Since the enzymes for further metabolism of xylitol are absent in *S. cerevisiae*, all the xylitol formed is excreted into the medium. The recombinant strains converted xylose to xylitol with yields close to the expected theoretical yields. Hallborn *et al.* (1991) and Thestrup and Hahn-Hägerdal (1995) obtained similar results by over-expressing the *P. stipitis XYL1* gene in *S. cerevisiae.* Xylitol production was most rapid and efficient when glucose was the co-substrate for both recombinant strains. Y294:pRG8 and Y294:pRG16 produced 0.83 and 0.86 g/g xylitol when glucose was the co-substrate. Final xylitol concentrations were lower with galactose and maltose. Galactose and xylose displayed identical retention times with the SugarPak column used for HPLC analyses. It was, thus, not possible to

distinguish between these sugars or establish whether or not all the galactose was consumed during the fermentation. The final xylitol concentrations obtained were higher with glucose as co-substrate than galactose. Meinander and Hahn-Hägerdal (1997b) reported higher yields with galactose than glucose. Their studies were performed using a different strain of *S. cerevisiae* harbouring the *P. stipitis XYL1* gene in strict anaerobic batch and fed-batch fermentations. Variations in yield can be attributed to the different conditions (simple shake flask fermentations that were not strictly anaerobic) as well as the different strains used. Another possibility to consider is that the recombinant genes were expressed under the control of the *PGK1* promoter/terminator cassette. This promoter was considered to be a strongly expressed constitutive promoter but it has been found to be induced up to 20-fold by glucose (Chambers *et al.*, 1995). Therefore, in medium containing glucose, the recombinant gene will be expressed more efficiently than in media containing galactose or maltose. Xylose was, therefore, converted to xylitol at a faster rate due to increased levels of XR. Meinander and Hahn-Hägerdal (1997a) demonstrated that a 20-fold difference in XR activity resulted in an approximately 2-fold difference in the rate of xylitol production. Maltose was consumed slowly compared to glucose by all three strains tested. This may be attributed to poor utilization of maltose by this strain of *S. cerevisiae*.

Published data (Meinander and Hahn-Hägerdal, 1997a; 1997b) on transport interactions between xylose and different co-substrates suggest that the co-substrates inhibit xylose conversion by inhibiting transport. Therefore, when these co-substrates are present in high concentrations, the transport of xylose probably controls the rate of xylose conversion. However, these inhibitory effects were not observed in this study. The ratio of co-substrate to xylose used was 1:5 whereas in most other studies either equal or higher amounts of xylose was used. Xylose and glucose are transported by the same transport system (facilitated diffusion) which has a higher affinity for

glucose (Busturia and Lagunas, 1986; van Zyl *et al.*, 1989). Therefore, at high glucose concentrations, xylose transport would be inhibited. Since competitive inhibition is mutual, in high xylose concentrations glucose transport will be inhibited or slower. This is evident in the present study where glucose and xylose are transported simultaneously. Xylitol production is rapid while glucose is metabolized and reduced cofactors are supplied for xylitol formation. Once glucose is depleted, xylitol formation continues at a much slower rate and reduced cofactors may be supplied during this phase by metabolism through the PPP and probably ethanol oxidation. It has already been established that ethanol is unable to support xylitol formation in recombinant *S. cerevisiae* under anaerobic conditions (Meinander *et al.*, 1994). However, during ethanol oxidation accompanied by simultaneous glucose oxidation with concomitant ATP formation, reduction equivalents are generated (Thestrup and Hahn-Hägerdal, 1995), making xylitol formation possible.

Glycerol is normally produced under anaerobic metabolism by *S. cerevisiae* in order to regenerate NAD^+ (Oura, 1977) by oxidizing excess NADH. Relatively low levels of glycerol were formed during this study. This is because excess NADH was most likely utilized by the XR enzyme to produce xylitol. In addition, the low glycerol levels can be attributed to the fact that substrate metabolism was not carried out under strict anaerobic conditions. Thus, the higher respiratory activity could explain the lower glycerol levels detected. The *C. shehatae XYL1* gene utilized either NADH or NADPH with a slight preference for the NADPH cofactor. NADPH could be regenerated during the formation of acetate during co-substrate metabolism (Meinander and Hahn Hägerdal, 1997b; Thestrup and Hahn Hägerdal, 1995). Acetate levels were not quantified in this study but there is compelling evidence that it was produced since the fermentation medium became acidic during the course of the fermentation.

Since high co-substrate concentrations inhibit xylose transport and have a negative effect

on the yield of xylitol, it is important that fermentations be performed at very low co-substrate concentrations. A bioreactor with a continuous feed of glucose or galactose at a very low concentration for growth, regeneration of the reduced cofactor [NAD(P)H] and supply of metabolic maintenance energy for the conversion of xylose to xylitol would be most effective for optimal production of xylitol by the recombinant strains. Meinander *et al.* (1996) studied the effect of a heterologous reductase on the redox balance of recombinant *S. cerevisiae* in anoxic chemostat culture at two different dilution rates. They found that the glycerol flux decreased and the acetate and carbon dioxide fluxes increased at both dilution rates. These results were attributed to the redox perturbation caused by the consumption of reduced cofactors in the XR-catalyzed reaction. In this study, regeneration of NAD, partly through XR instead of glycerol production, resulted in low levels of glycerol. In addition, xylose reduction activated those pathways which produce reduced cofactors such as acetate formation and the PPP. These results were further substantiated by Liden *et al.* (1996) who transformed the *XYL1* gene from *P. stipitis* into a mutant of *S. cerevisiae* incapable of glycerol production. The transformed strain was capable of anaerobic glucose conversion in the presence of added xylose indicating that the XR reaction can fulfill the role of the glyceraldehyde-3-phosphate dehydrogenase reaction as a redox sink.

The recombinant strains carrying all three xylose-metabolizing genes produced ethanol from xylose. The significant difference between the two strains was that the plasmid copy of *S. cerevisiae XYL3* gene was controlled by different promoters. In strainY294:pRG16159, the gene was controlled by the *ADH2* promoter which is repressed in the presence of glucose. This effect was evident in the fermentation experiments containing xylose and glucose. Most of the ethanol produced was due to the metabolism of glucose. Only a small quantity of ethanol was produced from xylose fermentation. Xylose was converted to xylitol and glycerol, at the expense

of ethanol. The diversion from the fermentative pathway can be directly attributed to low XK activity. The influence of XK in fermentation activity was demonstrated by Tantirungkij *et al.* (1994a) who obtained higher ethanol yields from a mutated *S. cerevisiae* recombinant strain expressing *P. stipitis XYL1* and *XYL2* genes that displayed elevated XK activity. This idea is further re-inforced by the study of Deng and Ho (1990) who observed improved fermentation of xylulose by over-expressing XK in *S. cerevisiae*.

The recombinant strain Y294:pRG16914 which was identical to the Y294:pRG16159 transformant, except for the XK expression cassette, was capable of simultaneous ethanol production from xylose and glucose. All the xylose-metabolizing genes were under the control of the *PGK1* promoter which meant that the expression of the xylose-metabolizing genes was induced in the presence of glucose, ensuring high levels of these heterologous proteins and efficient xylose metabolism. The difference in the xylose fermentation capablilties of these two recombinant strains can be directly linked to the difference in XK levels. In conditions where the glucose to xylose concentration was low, both sugars are taken up simultaneously and ethanol production was due to co-fermentation of both sugars. Meinander and Hahn-Hägerdal (1997a) established that glucose and xylose are only taken up simultaneously under glucose-limited conditions and attributed this to competition for the transport system. Thus, initial ethanol production was due to co-fermentation of both sugars. At higher glucose concentrations, xylose transport is inhibited and glucose is taken up first and fermented into ethanol. Once glucose concentrations are lower, xylose is also transported into the cells and both sugars are fermented. Approximately 20 g/l ethanol was produced from 70 g/l sugar (20 g/l glucose and 50 g/l xylose).

The only other report of ethanol production by a recombinant *S. cerevisiae* strain carrying xylose-metabolizing genes was by Ho *et al.* (1998). They obtained 60 g/l ethanol with a recombinant *Saccharomyces* strain from 155 g/l sugar (90 g/l glucose and 45 g/l xylose). The

high ethanol yield is due to a large extent to the high glucose concentration that was utilized. As observed in this study (20 g/l ethanol from 20 g/l glucose and 50 g/l xylose), most of the ethanol was due to glucose fermentation. Also, both sugars were fermented simultaneously in our study, since glucose levels were not inhibitory to xylose uptake. In the study by Ho *et al.* (1998), however, glucose levels were high and exercised an inhibitory effect on xylose transport. It was utilized first and xylose utilization occurred once glucose concentrations reached levels that were no longer inhibitory. The strain utilized by Ho *et al.* (1998) had two advantages over the strain used in this study. Firstly, it possessed superior fermentation capabilities and achieved the stated ethanol levels in 24 h. For the laboratory strain used in our study therefore, fermentation rates were lower. Secondly, the strain constructed by Ho *et al.* (1998) was stable even in a rich medium (without selection) whereas the strain constructed by us was stably maintained only under selection.

The laboratory strain constructed in our study, although not an ideal fermentative organism did manage to produce approximately 20 g/l ethanol from 70 g/l sugars (20 g/l glucose and 50 g/l xylose). Since almost 10 g/l xylose was present at the end of the fermentation, the yield was therefore almost 0.46 g ethanol/g sugar consumed. This yield is above the minimum yield of 0.4 g/g described as the prerequisite for xylose metabolizing organisms for economically viable production of ethanol from xylose on a commercial scale. However, the final ethanol concentration was low and the fermentation time too long. Ideally, for commercial application, a xylose-fermenting yeast should produce 50- 60 g/l ethanol within 36 h with a yield of at least 0.4 g/g (Jeffries, 1985). Higher initial sugar concentrations would certainly help in this respect, since glucose would then be fermented rapidly to correspondingly high concentrations of ethanol. The use of a more efficient fermentative host strain for the vector pRG16914 will also ensure quicker rates of fermentation.

Nissen and Nielsen (1999) proposed two strategies to enhance ethanol production by recombinant strains of *S. cerevisiae*. Firstly, elimination of glycerol production by converting NADH to NADPH. To do this, they expressed the transhydrogenase gene in a strain mutated in the glycerol biosynthetic pathway. Secondly, they overexpressed the ammonia assimilation pathway. They found that the second strategy was most effective in that glycerol formation was reduced by 50% and ethanol production increased by 8%.

The xylose-metabolizing genes on plasmid pRG16914 effectively enabled *S. cerevisiae* to convert xylose to ethanol. Since all the genes are controlled by the *PGK1* promoter their expression will not be negatively influenced by glucose. This has implications if the recombinant strains are to be used in the fermentation of lignocellulosic materials which consist of a variety of sugars, the major sugars being xylose and glucose. Future work to realize the full potential of this expression vector will entail choosing more efficient fermentative strains as hosts. Mutated strains carrying altered genotypes that enhance the formation of ethanol would be desirable. Furthermore, to alleviate the need to maintain selective pressure on the recombinant strains, the xylose gene cassette will be integrated into the chromosome of the selected strains, preferably at multiple sites. Kim *et al.* (1999) obtained a recombinant *S. cerevisiae* strain that had at least 40 copies of the *P. stipitis XYL1* gene integrated at chromosomal delta sequences. This strain was more efficient in long-term non-selective culture than the strain with the episomal plasmid. The rDNA genes have already been successfully targeted for the expression of other genes and will be the focus for our integration experiments in the future.

This chapter as well as the preceding ones have provided a detailed description of the different steps pursued to obtained the goals stated in Chapter One. The final chapter that follows provides an overview of the results obtained in the preceding chapters as well as proposing research and application of the genetically engineered strains constructed thus far (both in our

study as well as by other groups), their implications with respect to energy, economic and environmental issues.

CHAPTER FIVE

GENERAL DISCUSSION

The two major sugars in lignocellulosic material are glucose and xylose. Technology for the conversion of glucose to ethanol has long been established and superior organisms have been isolated or constructed for the efficient bioconversion of this sugar. The fermentation of xylose has been the focus of research for the past 20 years. Naturally occurring xylose-fermenting bacteria, fungi and yeasts have been in existence for centuries. *P. stipitis* and *C. shehatae* are the best known naturally occurring xylose-fermenting yeasts but they are unsuitable for large scale commercial production of ethanol since they have a relatively low tolerance to ethanol (Ligthelm *et al.*, 1988b; Skoog and Hahn-Hägerdal, 1990; Skoog *et al.*, 1992a).

Attempts to improve the xylose fermentative capabilities of yeasts and bacteria included classical genetic techniques, mutation and recombinant DNA technology. The latter technique has proved to have the greatest potential since its approach is direct. Bacteria have been easier to manipulate than yeasts. Genetically engineered xylose-fermenting bacteria have been in existence for almost a decade and as early as 1992, a company in the United States of America has employed a genetically modified bacterium to convert agricultural residues to ethanol (Wyman, 1999). Although bacteria can convert ethanol faster and with higher yields than yeasts, their use is not always desirable.

Saccharomyces spp. have traditionally been used for the fermentation of glucose-based feedstocks to ethanol. They cannot, however, convert xylose to ethanol or use it for growth. In the past 15 years, considerable effort has been made to genetically modify traditional glucose-fermenting yeasts, particularly *S. cerevisiae*, by recombinant DNA techniques. Initially, the xylose isomerase gene was cloned into yeasts but these efforts were unsuccessful because the genes

encoding bacterial XI were incapable of directing the synthesis of an active enzyme.

In the past decade, efforts to genetically engineer *S. cerevisiae* have focussed on expressing the *P. stipitis XYL1* and *XYL2* genes (Kötter and Ciriacy, 1993; Tantirungkij *et al.*, 1994b; Walfridsson *et al.*, 1995). These recombinant yeasts could still not effectively convert xylose to ethanol. Studies involving the over-expression of the *S. cerevisiae* gene encoding XK implicated the importance of this enzyme in the metabolic pathway of xylose to ethanol (Deng and Ho, 1990; Tantirungkij *et al.*, 1994a). Two important conclusions were gleaned from these experiments. Firstly, the expression of these genes should be under the control of strong promoters that would not be negatively influenced in the presence of substrates used for fermentation. Secondly, all three xylose-metabolizing genes would be most likely required for efficient conversion of xylose to ethanol. These conclusions were validated in our studies. Although the *C. shehatae* XR has a different cofactor preference than that from *P. stipitis*, it effectively converted xylose to xylitol. Two other laboratories have constructed recombinant *S. cerevisiae* strains co-expressing *P. stipitis XYL1* and *XYL2* genes and *S. cerevisiae XYL3* (Ho, 1999; Ho *et al.*, 1998; Roca *et al.*, 1999). Similar results to those in our study (Govinden *et al.*, 1999) were obtained by these authors and the overwhelming conclusion was that ethanol formation from xylose by *S. cerevisiae* requires over-expression of XK in conjunction with XR and XDH.

The recombinant *S. cerevisiae* strains capable of xylose fermentation can be further improved by several techniques: including selection of better fermentative hosts and genetic manipulations, e.g., integration into the chromosome for genetic stability, creation of mutants in specific metabolic pathways (Ho *et al.*, 1998; Nissen and Nielsen, 1999) in order to fine tune ethanol production from xylose. However, these techniques may only moderately improve the rate of ethanol production. A crucial factor affecting xylose metabolism, especially in recombinant

S. cerevisiae strains that has as yet not been resolved, is the transport of xylose. In *S. cerevisiae*, xylose transport is by facilitated diffusion via the uptake system that also transports glucose. In fermentation media where glucose is at a higher concentration, xylose uptake is inhibited, thus increasing fermentation time of mixed sugars. The xylose-fermenting yeasts *P. stipitis* and *C. shehatae* have two transport systems for the transport of xylose. In the case of *P. stipitis*, both are active transport systems and *C. shehatae* has a facilitated diffusion system as well as an active transport system. The expression of the active uptake system from either of these yeasts may improve the rate of production of ethanol from xylose by recombinant *S. cerevisiae*. However, with highly fermentative species (Ho *et al.*, 1998), transport does not appear to be a limiting feature since ethanol production was fairly rapid even though xylose utilization occurred only after glucose levels were no longer inhibitory.

It is now opportune for these recombinant stains to be tested in the environment that they were targeted for initially, i.e., they need to be tested in scaled-up or industrial fermentors using the different lignocellulosic materials, industrial waste streams and other effluents as their carbon source. However, one has to strike a balance between fine tuning ethanol production by further genetic manipulations and simply using one of the strains already available but optimising process parameters for that strain and the particular lignocellulose or industrial effluent to be used as carbon source. An important point to consider is that the different recombinant strains may produce yields and rates of ethanol production that differ for the various lignocellulose materials or industrial effluents. It may, therefore, be beneficial to use these strains in conditions where they perform best rather than to try to obtain one super-fermenter that will work in every situation. Agricultural residues and forestry products will contain different sugars in different ratios. Industrial effluents will contain various substances with differing degrees of toxicity to the recombinant strains.

In order for this technology to have its full impact on the world, industry has to also adapt to the use of ethanol as an energy source. Brazil for example, is the leading user of ethanol as a fuel source (Nissen and Nielsen, 1999). Other developing and developed countries need to switch over to this energy source. The motor industry needs to adapt engines, if need be, to burn mixes containing larger proportions of ethanol, or ideally, ethanol only.

One of the major arguments against ethanol as an energy source is that it is more expensive to produce than fossil fuels. Most lignocellulosic materials, especially agricultural residues, are scattered over a wide area in the fields where they are produced and have to be gathered and transported to sites for conversion to ethanol. Although the cost of the raw material is nominal, transportation costs are prohibitive. One solution to this problem would be to select areas that produce large amounts of a similar type of agricultural product and construct a small fermentation plant in this vicinity so as to alleviate transportation costs.

When it comes to exploiting the full potential of industrial effluents, e.g., spent sulphite liquors or pulp mill effluents, in order to minimize costs, (i) waste-holding tanks could be converted to fermentation vessels; or (ii) fermentation vessels could be added on or (iii) built on site as the final phase in the process, to avoid transportation costs. In addition to ethanol production from lignocellulosics, there are other products e.g., xylitol, furfural and other solvents that have commercial value that can be produced as secondary products or downstream from the major product which will remain the production of ethanol.

The quest of the molecular biologist is fast reaching its close in the adventure involving the genetic engineering of bacteria and yeast species. Recombinant *S. cerevisiae* strains have been constructed that are now capable of ethanol fermentation from xylose. Parallel progress has been made in improving methods to release the sugars from lignocellulosic materials, thus increasing the economic feasibility of the "green refineries" (Girio *et al.*, 1999). It therefore, seems likely

that the energy problems of the planet have been resolved. In fact, the planet may be saved by the cleaner burning ethanol which will result in a reduction of pollution levels and an alleviation of the greenhouse effect.

REFERENCES

Alcorn, M.E. and C.C. Griffin. 1978. A kinetic analysis of D-xylose transport in *Rhodotorula glutinis. Biochim Biophys. Acta.* **510**:361-371.

Alexander, M.A., T.W. Chapman and T.W. Jeffries. 1988a. Continuous xylose fermentation by *Candida shehatae* in a two-stage reactor. *Appl. Microbiol. Biotechnol.* **17**:221-229.

Alexander, M.A., T.W. Chapman and T.W. Jeffries. 1988b. Xylose metabolism by *Candida shehatae* in continuous culture. *Appl. Microbiol. Biotechnol.* **28**:478-486.

Alexander, M.A., V.W. Yang and T.W. Jeffries. 1988c. Levels of pentose phosphate pathway enzymes from *Candida shehatae* grown in continuous culture. *Appl. Microbiol. Biotechnol.* **29**:282-288.

Alexander, N.J. 1990. Characterization of a respiratory-deficient mutant of *Pachysolen tannophilus. Curr. Genet.* **17**: 493-497.

Alexander, N.J. and R.W. Detroy. 1983. Alcohol production by mitochondrial mutants and wild type *Saccharomyces cerevisiae. Biotechnol. Lett.* **5**:165-168.

Alterthum, F. and L.O. Ingram. 1989. Efficient ethanol production from glucose, lactose and xylose by recombinant *Escherichia coli. Appl. Microbiol. Biotechnol.* **55**:1943-1948.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389-3402.

Amore, R. and C.P. Hollenberg. 1989. Xylose isomerase from *Actinoplanes missouriensis*: primary structure of the gene and the protein. *Nucleic Acids Res.* **17**:7515.

Amore, R., P. Kötter, C. Kuster, M. Ciriacy and C.P. Hollenberg. 1991. Cloning and expression in *Saccharomyces cerevisiae* of the NAD(P)H-dependent xylose reductase encoding gene (*XYL1*) from the xylose-assimilating yeast *Pichia stipitis*. *Gene* **109**:89-97.

Amore, R., M. Wilhelm and C.P. Hollenberg. 1989. The fermentation of xylose: an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. *Appl. Microbiol. Biotechnol.* **30**:351-357.

Armstrong, K.A., T. Som, F.C. Volkert, A. Rose and J.R. Broach. 1989. Propagation and expression of genes in yeast using 2-micron circle vectors. pp. 165-192. *In* Bar, P.J.; A.J. Brake and P. Valenzuela (eds.) Yeast Genetic Engineering, Butterworths.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl. 1989. Current protocols in molecular biology, Vol. 2. Greene Publishing Associates and Wiley-Interscience, New York.

Barnett, J.A. 1976. Utilization of sugars by yeasts. *Adv. Carbohyd. Chem. Biochem.* **32**:125-134.

Batt, C.A., S. Carvallo, D.D. Easson, M. Aked and A.J. Sinskey. 1986a. Direct evidence for a xylose metabolic pathway in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **28**:549-553.

Batt, C.A., E. O'Neill, S.R. Novak, J. Ko and A. Sinskey. 1986b. Hyperexpression of *Escherichia coli* xylose isomerase. *Biotechnol. Prog.* **2**:140.

Becker, D.M. and L. Guarente. 1991. High efficiency transformation of yeast by electroporation. pp. 182-187. *In*: C. Guthrie and G.R. Fink (eds.), Methods in Enzymology, Vol. 194, Guide to yeast genetics and molecular biology. Academic Press Inc., San Diego.

Beggs, J.D. 1978. Transformation of yeast by a replicating hybrid plasmid. *Nature* **275**:104-109.

Biesterveld, S., S.J.W.H. Oude Elferink, A.J.B. Zehnder and A.J.M. Stams. 1994. Xylose and glucose utilization by *Bacteroides xylanolyticus* X5-1 cells grown in batch and continuous culture. *Appl. Environ. Microbiol.* **60**:576-580.

Birnboim, H.C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.

Bisson, L.F. and D.G. Fraenkel. 1984. Expression of kinase-dependent glucose uptake in *Saccharomyces cerevisiae. J. Bacteriol.* **159**:1013-1017.

Bolen, P.L. and R.W. Detroy. 1985. Induction of NADPH-linked D-xylose reductase and NAD-linked xylitol dehydrogenase activities in *Pachysolen tannophilus* by D-xylose, L-arabinose, or D-galactose. *Biotechnol. Bioeng.* **27**:302-307.

Bolen, P.L., K.A. Roth and S.N. Freer.1986. Affinity purifications of aldose reductase and xylitol dehydrogenase from the xylose-fermenting yeast *Pachysolen tannophilus. Appl. Environ. Microbiol.* **52**:660-664.

Bothast, R.J., B.S. Dien and N.N. Nichols. 1999. Conversion of corn fiber to fuel ethanol: applications of metabolic engineering technologies. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, O14.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.

Brown, S.W., S.G. Oliver, D.E.F. Harrison and R. Righelato. 1981. Ethanol inhibition of yeast growth and fermentation: differences in magnitude and complexity of the effect. *Eur. J. Appl. Microbiol. Biotechnol.* **11**:151-155.

Bruinenberg, P.M., P.H.M. de Bot, J.P. van Dijken and W.A. Scheffers. 1983. The role of redox imbalances in the anaerobic fermentation of xylose by yeasts. *Appl. Microbiol. Biotechnol.* **18**:287-292.

Bruinenberg, P.M., P.M.H. de Bot, J.P. van Dijken and W.A. Scheffers. 1984. NADH-linked aldose reductase: the key to anaerobic alcoholic fermentation of xylose by yeasts. *Appl. Microbiol. Biotechnol.* **19**:256-260.

Buckholz, R.G. and M.A.G. Gleeson. 1991. Yeast systems for the commercial production of heterologous proteins. *Bio/Technol.* **9**:1067-1071.

Busturia, A. and R. Lagunas. 1986. Catabolite inactivation of the glucose transport system in *Saccharomyces cerevisiae. J. Gen. Microbiol.* **132**:379-385.

Butler, J.S., P.P. Sadhale and T. Platt. 1990. RNA processing *in vitro* produces mature 3' ends of a variety of *Saccharomyces cerevisiae* mRNAs. *Mol. Cell. Biol.* **10**:2599-2605.

Campbell, J.L. 1983. Yeast DNA replication. pp. 109-146. *In* Setlow, J.K and A. Holleander (eds.) Genetic engineering Principles and Methods. Vol. 5, Plenum Press, New York.

Chakravorty, M., L.A. Veiga, M. Bacila and B.L. Horecker. 1962. Pentose metabolism in *Candida* II. The diphosphopyridine nucleotide-specific polyol dehydrogenase of *Candida utilis. Biol. Chem.* **237**:1014-1020.

Chambers, A., E.A. Packham and I.R. Graham. 1995. Control of glycolytic gene expression in the budding yeast (*Saccharomyces cerevisiae*). *Curr. Genet.* **29**:1-9.

Chan, E.C., P.P. Ueng, K.L. Eder and L.F. Chen. 1989. Integration and expression of the *Escherichia coli* xylose isomerase gene in *Schizosaccharomyces pombe. J. Ind. Microbiol.* **4**:409-418.

Chen, Z. and N.W.Y. Ho. 1993. Cloning and improving the expression of *P. stipitis* xylose reductase gene in *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **39/40**:135-147.

Chiang, L.C. and S.G. Knight. 1960. Metabolism of D-xylose by moulds. *Nature* **188**:79-81.

Chiang, L.C., C.S. Gong, L.F. Chen and G.T. Tsao. 1981a. D-xylulose fermentation to ethanol by *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **42**:284-289.

Chiang, L.C., H.Y. Hsiao, P.P. Ueng, L.F. Chen and G. Tsao. 1981b. Ethanol production from xylose by enzymic isomerization and yeast fermentation. *Biotechnol. Bioeng. Symp.* **11**:263-274.

Cho, J-Y. And T.W. Jeffries. 1998. *Pichia stipitis* genes for alcohol dehydrogenase with fermentative and respiratory functions. *Appl. Environ. Microbiol.* **64**:1350-1358.

Ciriacy, M. 1975. Genetics of alcohol dehydrogenase in *Saccharomyces cerevisiae*. I. Isolation of *adh* mutants. *Mutation Res.* **29**:315-326.

Cirillo, V.P. 1968. Relationship between sugar structure and competition for sugar transport in baker's yeast. *J. Bacteriol.* **95**:603-611.

Clark, L. and J. Carbon. 1976. A colony bank containing synthetic *Col*E1 hybrids representative of the entire *E. coli* genome. *Cell* **9**:91-99.

Clark, L. and J. Carbon. 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* **257**:504-509.

Collins, J. and B. Hohn. 1978. 'Cosmids: A type of plasmid gene cloning vector that is packageable *in vitro* in bacteriophage heads. *Proc. Natl. Acad. Sci.* **75**:4242-4246.

Crabb, W.D., E. Ferrari and B. Schmidt. 1991. Tools and strategies for cloning studies. pp. 365-388. *In* Modern Microbial Genetics. U.N. Streips and R.E. Yasbin (eds) Wiley-Liss Inc., New York.

Crous, J.M., I.S. Pretorius and W.H. van Zyl. 1995. Cloning and expression of an *Aspergillus kawachii* endo-1,4-β-xylanase gene in *Saccharomyces cerevisiae. Curr. Genet.* **28**:467-473.

David, J.D. and H. Weismeyer. 1970. Control of xylose metabolism in *Escherichia coli. Biochem. Biophys. Acta.* **208**:201-497.

Dekker, K., H. Yamagata, K. Sakaguchi and S. Udaka. 1991. Xylose (glucose) isomerase gene from the thermophile *Clostridium thermohydrosulfuricum*: cloning, sequencing and expression in *Escherichia coli. Agricultural Biological Chem.* **55**:221-227.

Dekker, R.F.K. and W.A. Lindner. 1979. Bioutilization of lignocellulosic waste materials: A Review. *South Afr. J Sci.* **75**:65-71

Deng, X.X. and N.W.Y. Ho. 1990. Xylulokinase activity in various yeasts including *Saccharomyces cerevisiae* containing the cloned xylulokinase gene. *Appl. Biochem. Biotechnol.* **24/25**:193-199.

Ditzelmüller, G., E.M. Kubicek-Pranz, M. Rohr and C.P. Kubicek. 1985. NADPH-specific and NADH-specific xylose reductase is catalyzed by two separate enzymes in *P. tannophilus. Appl. Microbiol. Biotechnol.* **22**:297-299.

Ditzelmüller, G., C.P. Kubicek, W. Wohrer and M. Rohr. 1984a. Xylitol dehydrogenase from *P. tannophilus. FEMS Microbiol. Lett.* **25**:195-198.

Ditzelmüller, G., C.P. Kubicek, W. Wohrer and M. Rohr. 1984b. Xylose metabolism in *P. tannophilus:* purification and properties of xylose reductase. *Can. J. Microbiol.* **30**:1330-1336.

Does, A.L. and L.F. Bisson. 1989. Characterization of xylose uptake in the yeast *Pichia heedii* and *Pichia stipitis. Appl. Environ. Microbiol.* **55**:159-164.

Does, A.L. and L.F. Bisson. 1990. Isolation and characterization of *Pichia heedii* mutants defective in xylose uptake. *Appl. Environ. Microbiol.* **56**:3321-3328.

Drocourt, D., S. Bejar, T. Calmels, J.P. Reynes and G. Tiraby. 1988. Nucleotide sequence of the xylose isomerase gene from *Streptomyces violaceoniger. Nucleic Acids Res.* **16**: 933-937.

du Preez, J.C., M. Bosch and B.A. Prior. 1986a. The fermentation of hexose and pentose sugars by *Candida shehatae* and *Pichia stipitis. Appl. Microbiol. Biotechnol.* **23**:228-233.

du Preez, J.C., M. Bosch and B.A. Prior. 1986b. Xylose fermentation by *Candida shehatae* and *Pichia stipitis* - effects of pH, temperature and substrate concentration. *Enzyme Microb. Technol.* **8**:360-364.

du Preez, J.C. and B.A. Prior. 1985. A quantitative screening of some xylose-fermenting yeast isolates. *Biotechnol. Lett.* **7**:241-246.

du Preez, J.C., B.A. Prior and A.M.T. Monteiro. 1984. The effect of aeration on xylose fermentation by *Candida shehatae* and *Pachysolen tannophilus*, a comparative study. *Appl. Microbiol. Biotechnol.* **19**:261-266.

du Preez, J.C. and J.P. van der Walt. 1983. Fermentation of D-xylose to ethanol by a strain of *Candida shehatae. Biotechnol. Lett.* **5**:357-362.

du Toit, P.J., S.P. Olivier and P.L. van Biljon. 1984. Sugar cane bagasse as a possible source of fermentable carbohydrate. I. Characterization of bagasse with regard to monosaccharide, hemicellulose and amino acid composition. *Biotechnol. Bioeng.* **26**: 1071-1078.

Emodi, A. 1978. Xylitol: its properties and food applications. *Food Technol.* **32**:20-32.

Evans, C.T. and C. Ratledge. 1983. A comparison of the oleaginous yeast, *Candida curvata*, grown on different carbon sources in continuous and batch culture. *Lipids.* **18**:623-629.

Evans, C.T. and C. Ratledge. 1984. Induction of xylose-5-phosphate phosphoketolase in a variety of yeasts grown on D-xylose: the key to efficient xylose metabolism. *Arch. Microbiol.* **139**:48-52.

Feldmann, S.; G.A. Sprenger and H. Sahm. 1989. Ethanol production with a pyruvate-formate-lyase mutant of *Klebsiella planticola* carrying a pyruvate decarboxylase gene from *Zymomonas mobilis*. *Appl. Microbiol. Biotechnol.* **31**:152-157.

Flanagan, T. and M.J. Waites. 1992. Purification and characterization of D-xylulokinase from the pentose-fermenting yeast *Pichia stipitis* NCYC 1541. *Enzyme Microb. Technol.* **14**: 975-979.

Futcher, A.B. 1988. The 2μm circle plasmid of *Saccharomyces cerevisiae*. *Yeast* **4**:27-40.

Futcher, A.B. and B.S. Cox. 1983. Maintenance of the 2μm circle plasmid in populations of *Saccharomyces cerevisiae*. *J. Bacteriol.* **154**:612-622.

Futcher, A.B. and B.S. Cox. 1984. Copy number and stability of 2μm-circle-based artificial plasmids of *Saccharomyces cerevisiae*. *J. Bacteriol.* **157**:283-290.

Gancedo, J.M. and R. Lagunas. 1973. Contribution of the pentose phosphate pathway to glucose metabolism in *Saccharomyces cerevisiae*: a critical analysis on the use of labelled glucose. *Plant Sci. Lett.* **1**:193-200.

Girio, F.M., F. Carvalheiro and P. Esteves. 1999. Lignocellulosic wastes fractionation using hydrothermel based treatments for xylo-oligosaccharide production. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, O5.

Girio, F.M., M.A. Peito and M.T. Amaral-Collaco. 1989. Enzymatic and physiological study of D-xylose metabolism in *Candida shehatae*. *Appl. Microbiol. Biotechnol.* **32**:199-204.

Gong, C.S., T.A. Claypool, L.D. McCracken, C.M. Mann, P.P. Ueng and G.T. Tsao. 1983. Conversion of pentoses by yeasts. *Biotechnol. Bioeng.* **25**:85-102.

Gong, C.S., M.R. Ladisch and G.T. Tsao. 1981a. Production of ethanol from wood hemicellulose hydrolysates by a xylose-fermenting yeast mutant, *Candida* sp. XF 217. *Biotechnol. Lett.* **3**:657-662.

Gong, C.S., L.D. McCracken and G.T. Tsao. 1981b. Direct fermentation of D-xylose to ethanol by a xylose fermenting yeast mutant *Candida* sp. XF 217. *Biotechnol. Lett.* **3**: 245-250.

Govinden, R., B. Pillay, D. Pillay and W.H. van Zyl. 1999. Xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, O11.

Gottschalk, G. 1979. Bacterial Metabolism. Springer Verlag, New York.

Grey, M. and M. Brendel. 1992. A ten minute protocol for transforming *Saccharomyces cerevisiae* by electroporation. *Curr. Genet.***22**:335-336.

Gritz, L. and J. Davies. 1983. Plasmid encoded hygromycin B resistance: the sequence of hygromycin B phosphotransferase and its expression in *Escherichia coli* and *Saccharomyces cerevisiae. Gene* **25**:179-188.

Grootjen, D.R.J.; R.G.J.M. van der Lans and K.Ch.AS.M. Luyben. 1990. Effects of the aeration rate on the fermentation of glucose and xylose by *Pichia stipitis* CBS 577. *Enzyme Microb. Technol.* **12**:20-23.

Grunstein, M. and D. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proceed. Natl. Acad. Sci.* **72**:3961-3965.

Gupthar, A.S. 1987. Construction of series of *Pichia stipitis* strains with increased DNA content. *Curr. Genet.* **12**:605-610.

Gupthar, A.S. 1992. Segregation of altered parental properties in fusions between *Saccharomyces cerevisiae* and the D-xylose fermenting yeasts *Candida shehatae* and *Pichia stipitis. Can. J. Microbiol.* **38**:1233-1237.

Gupthar, A.S. and H.M. Garnett. 1987. Hybridization of *Pichia stipitis* with its presumptive imperfect partner *Candida shehatae. Curr. Genet.* **12**:199-204.

Hadfield, C., A.M. Cashmore and P.A. Meacock. 1986. An efficient chloramphenicol-resistance marker for *Saccharomyces cerevisiae* and *Escherichia coli. Gene* **45**:149-158.

Hahn-Hägerdal, B., J. Hallborn, H. Jeppsson, L. Olsson, K. Skoog and M. Walfridsson. 1993. Pentose fermentation to alcohol. pp. 231-290. *In* J.N. Saddler (ed.). Bioconversion of forest and agricultural plant residues. CAB International, Wallingford, United Kingdom.

Hahn-Hägerdal, B., H. Jeppsson, K. Skoog and B.A. Prior. 1994. Biochemistry and physiology of xylose fermentation by yeasts. *Enzyme Microb. Technol.* **16**:1-11.

Hallborn, J., M.F. Gorwa, N. Meinander, M. Penttila, S. Keränen and B. Hahn-Hägerdahl. 1994. The influence of co-substrate and aeration on xylitol formation by recombinant *Saccharomyces cerevisiae* expressing the *XYL1* gene. *Appl. Microbiol. Biotechnol.* **42**:326-333.

Hallborn, J., M. Walfridsson, U. Airaksinen, H. Ojamo, B. Hahn-Hägerdal, M. Penttilä and S. Keränen. 1991. Xylitol production by recombinant *Saccharomyces cerevisiae*. *Bio/Technol.* **9**:1090-1096.

Hicks, J.B., J.N Strathern, A.J.S Klar and S.L. Dellaporta. 1982. Cloning by complementation in yeast: The mating type genes.pp. 219-248 *In* Genetic engineering. Principles and methods, Vol. 4. J.K. Setlow and A. Hollaender (eds), New York, Plenum.

Hill, J., K.A. Ian, G. Donald and D.E. Griffiths. 1991. DMSO-enhanced whole-cell yeast transformation. *Nucleic Acids Res.* **19**:5791.

Hinnen, A., J.B. Hicks and G.R. Fink. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. USA.* **75**:1929-1933.

Hitzeman, R.A., F.F. Hagie, H.L. Levine, D.W. Goedel, G. Ammerer and B.D. Hall. 1981. Expression of human gene for interferon in yeast. *Nature* **293**:717-723.

Hitzeman, R.A.; D.W. Leung; L.J. Perry; W.J. Kohr; H.L. Levine and D.W. Goedel. 1983. Secretion of human interferons by yeast. *Science* **219**:620-625.

Ho, N.W.Y. 1999. Our successful metabolic engineering of *Saccharomyces* yeasts for effective cofermentation of glucose and xylose to ethanol. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, O4.

Ho, N.W.Y. and S.F. Chang. 1989. Cloning of yeast xylulokinase gene by complementation of *E. coli* and yeast mutations. *Enzyme Microb. Technol.* **11**:417-421.

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

Ho, N.W.Y., F.P. Lin, S. Huang, P.C. Andrews and G.T. Tsao. 1990. Purification, characterization and amino acid terminal sequence of xylose reductase from *Candida shehatae*. *Enzyme Microb. Technol.* **12**:33-39.

Ho, N.W.Y., D. Petros and X.X. Deng. 1991. Genetic transformation of xylose-fermenting yeast *Pichia stipitis*. *Appl. Biochem. Biotechnol.* **28/29**:369-375.

Ho, N.W.Y., S. Rosenfeld, P. Stevis and G.T. Tsao. 1983. Purification and characterization of the D-xylose isomerase gene from *Escherichia coli*. *Enzyme Microb. Technol.* **5**: 417-420.

Ingram, L. O. and J.B. Doran. 1995. Conversion of cellulosic materials to ethanol. *FEMS Microbiol. Rev.* **16**:235-241.

Innis, M.A. and D.H. Gelfand. 1990. Optimization of PCRs. pp. 3-12. *In* M.A. Innis, D.H. Gelfand, J.J. Sninsky and T.J. White (eds.) PCR protocols: a guide to methods and application. Academic Press, Inc, San Diego, Calif.

Ito., H., Y. Fukuda, K. Murata and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153**:163-168.

James, A.P. and D.M. Zahab. 1982. A genetic system for *Pachysolen tannophilus*, a pentose-fermenting yeast. *J. Gen. Microbiol.* **128**:2297-2301.

James, A.P. and D.M. Zahab. 1983. The construction and genetic analysis of polyploids and aneuploids of the pentose-fermenting yeast *Pachysolen tannophilus*. *Gen. Microbiol.* **129**: 2489-2494.

Janda, S., A. Kotyk and R. Tauchova. 1976. Monosaccharide transport in the yeast *Rhodotorula glutinis*. *Arch. Microbiol.* **111**:151-154.

Jeffries, T.W. 1981a. Conversion of xylose to ethanol under aerobic conditions by *Candida tropicalis*. *Biotechnol. Lett.* **3**:213-218.

Jeffries, T.W. 1981b. Fermentation of xylulose to ethanol using xylose isomerase and yeasts. *Biotechnol. Bioeng.* **11**:315-324.

Jeffries, T.W. 1984a. Mutants of *Pachysolen tannophilus* showing enhanced rates of growth and ethanol formation from D-xylose. *Enzyme Microb. Technol.* **6**:254-258.

Jeffries, T.W. 1984b. Unstable petite and grande variants of *Candida shehatae*. *Biotechnol. Lett.* **6**:777-782.

Jeffries, T.W. 1985. Comparison of alternatives for the fermentation of pentoses to ethanol by yeasts. pp. 231-252. *In* Lowenstein, M. (ed.) Energy applications of biomass: Proceedings of the National Meeting on Biomass Research and Development for Energy Applications; 1984. October 1-3; Arlington VA. New York, NY: Elsevier Applied Science Publishers.

Jeffries, T.W. and C.P. Kurtzman. 1994. Strain selection, taxonomy and genetics of xylose-fermenting yeasts. *Enzyme Microb. Technol.* **16**:922-932.

Jeppsson, H., N.J. Alexander and B. Hahn-Hägerdal. 1995. Existence of cyanide insensitive respiration in the yeast *Pichia stipitis* and its possible influence on product formation during xylose utilization. *Appl. Environ. Microbiol.* **61**:2596-2600.

Johanssen, E., L. Eagle and G. Bredenhann. 1985. Protoplast fusion used for the construction of presumptive polyploids of the D-xylose fermenting yeast *Candida shehatae*. *Curr. Genet.* **9**:313-319.

Karczewska, H. 1959. Some observations on pentose utilization by *Candida tropicalis. Compt. Rend. Lab. Carlsberg.* **31**:251-258.

Kikuchi, Y. 1983. Yeast plasmid requires a *cis*-acting locus and two plasmid proteins for its stable maintenance. *Cell* **35**:487-493.

Kikuchi, T, Y. Itoh, T. Kasumi and C. Fukazawa. 1990. Molecular cloning of the *xylA* gene encoding xylose isomerase from *Streptomyces griseofuscus* S-41: primary structure of the gene and its products. *Agricultural Biological Chem.* **54**:2469-2472.

Kilian, S.G. and N. van Uden. 1988. Transport of xylose and glucose in the xylose fermenting yeast *Pichia stipitis. Appl. Microbiol. Biotechnol.* **27**:545-548.

Kim, Y.S., S.Y. Kim, J.H. Kim and S.C. Kim. 1999. Xylitol production using recombinant *Saccharomyces cerevisiae* containing multiple xylose reductase genes at chromosomal delta sequences. *J. Biotechnol.* **67**:159-171.

Kingsman, S.M., A.J. Kingsman, M.J. Dobson, J. Mellor and N.A. Roberts. 1985. Heterologous gene expression in *Saccharomyces cerevisiae. Biotechnol. Genet. Eng. Revs.* **3**:377-416.

Klapatch, T.R., D.A.L. Hogsett, S. Baskaran, S. Pal and L.R. Lynd. 1994. Organism development and characterization for ethanol production using thermophilic bacteria. *Appl. Biochem. Biotechnol.* **45/46**:209-223.

Kleinzeller, A. and A. Kotyk. 1967. Transport of monosaccharides in yeast cells and its relationship to cell metabolism. p33. *In* Aspects of yeast metabolism. A.K. Mills and H. Krebs (eds.) Philadelphia. F.A. Davis Co.

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

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

Kotyk, A. 1967. Properties of the sugar carrier in baker's yeast. II. Specificity of transport. *Folia Microbiol.* **12**:121-131.

La Grange, D.C., I.S. Pretorius and W.H. van Zyl. 1996. Expression of a *Trichoderma reesei* β-xylanase gene (*XYN2*) in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **62**: 1036-1044.

Lachke, A.H. and T.W. Jeffries. 1986. Levels of enzymes of the pentose phosphate pathway in *Pachysolen tannophilus* Y-2460 and selected mutants. *Enzyme Microb. Technol.* **8**: 353-359.

Lam, V.M.S., K.R. Daruwalla, P.J.F. Henderson and M.M.C. Jones-Mortimer. 1980. Proton linked D-xylose transport in *Escherichia coli*. *J. Bacteriol.* **143**: 396-402.

Laplace J.M., J.P. Delgenes; R. Moletta and J.M. Navarro. 1992. Alcoholic glucose and xylose fermentations by the coculture process: compatibility and typing of associated strains. *Can. J. Microbiol.* **38**:654-658.

Laplace J.M., J.P. Delgenes, R. Moletta and J.M. Navarro. 1993a. Effects of culture conditions on the co-fermentation of a glucose and xylose mixture to ethanol by a mutant of *Saccharomyces diastaticus* associated with *Pichia stipitis*. *Appl. Microbiol. Biotechnol.***39**:760-763.

Laplace J.M., J.P. Delgenes, R. Moletta and J.M. Navarro. 1993b. Ethanol production from glucose and xylose by separated and co-culture process using high cell density systems. *Process Biochem.* **28**:519-525.

Lastick, S.M., A. Mohaghegi, M.P. Tucker and K. Grohmann. 1990. Simultaneous fermentation and isomerization of xylose to ethanol at high xylose concentration. *Appl. Biochem. Biotechnol.* **24/25**:431-439.

Lawford, H.G. and J.D. Rousseau. 1991. Fuel ethanol from hardwood hemicellulose hydrolysate by genetically engineered *E. coli* B carrying genes from *Zymomonas mobilis*. *Biotechnol. Lett.* **13**:191-196.

Leao, C and N. van Uden. 1982. Effects of ethanol and other alkanols on the glucose transport system of *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **24**:2601-2604.

Lee, H., A.P. James, D.M. Zahab, G. Mahmourides, R. Maleszka and H. Schneider. 1986. Mutants of *Pachysolen tannophilus* with improved production of ethanol from D-xylose. *Appl. Environ. Microbiol.* **51**:1252-1258.

Liden, G., M. Walfridsson, R. Ansell, M. Anderlund, L. Adler and B. Hahn-Hägerdal. 1996. A glycerol-3-phosphate dehydrogenase-deficient mutant of *Saccharomyces cerevisiae* expressing the heterologous *xyl*1 gene. *Appl. Environ. Microbiol.* **62**:3894-3896.

Ligthelm, M.E., B.A. Prior and J.C. du Preez. 1988a. The effect of respiratory inhibitors on the fermentative ability of *Pichia stipitis*, *Pachysolen tannophilus* and *Saccharomyces cerevisiae* under various conditions of aerobiosis. *Appl. Microbiol. Biotechnol.* **29**:67-71.

Ligthelm, M.E., B.A. Prior and J.C. du Preez. 1988b. The oxygen requirements of yeasts for the fermentation of D-xylose and D-glucose to ethanol. *Appl. Microbiol. Biotechnol.* **28**: 63-68.

Ligthelm, M.E., B.A. Prior, J.C. du Preez and V. Brandt. 1988c. An investigation of D-[13C] xylose metabolism in *Pichia stipitis* under aerobic and anaerobic conditions. *Appl. Microbiol. Biotechnol.* **28**:293-296.

Linden, T. and B. Hahn-Hägerdal. 1989. Fermentation of lignocellulose hydrolysates with yeasts and xylose isomerase. *Enzyme Microb. Technol.* **11**:583-589.

Linden, T., J. Peetre and B. Hahn-Hägerdal. 1992. Isolation and characterization of acetic acid tolerant galactose fermenting strains of *Saccharomyces cerevisiae* from a spent sulphite liquor fermentation plant. *Appl. Environ. Microbiol.* **58**:1661-1669.

Lindsay, S.E., R.J. Bothast and L.O. Ingram. 1994. Improved strains of recombinant *Escherichia coli* for ethanol production from sugar mixes. *Appl. Microbiol. Biotechnol.* **42**:

Linko, Y.Y., H. Kautola, S. Uotila and P. Linko. 1986. Alcoholic fermentation of D-xylose by immobilized *Pichia stipitis. Biotechnol. Lett.* **8**:47-52.

Lohmeier-Vogel, E., K. Skoog, H. Vogel and B. Hahn-Hägerdal. 1989. ^{31}P nuclear magnetic resonance study of the effect of azide on xylose fermentation by *Candida tropicalis. Appl. Environ. Microbiol.* **55**:1974-1980.

Lopes, T.S., G.J. Hakaart, B.L. Koerts, H.A. Raue and R.J. Planta. 1990. Mechanism of high-copy-number integration of pMIRY-type vectors into the ribosomal DNA of *Saccharomyces cerevisiae. Gene* **105**:83-90.

Lopes, T.S., J. Klootwijk, A.E. Veenstra, P.C. van der Aar, H. van Heenhuizen, H.A. Raue and R.J. Planta. 1989. High-copy-number integration into the ribosomal DNA of *Saccharomyces cerevisiae. Gene* **79**:199-206.

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

Lutsorf, U. and R. Megnet. 1968. Multiple forms of alcohol dehydrogenase in *Saccharomyces cerevisiae. Arch. Biochem. Biophys.* **126**:933-944.

Maleszka, R., A.P. James and H. Schneider. 1983. Ethanol production from various sugars by strains of *Pachysolen tannophilus* bearing different numbers of chromosomes. *J. Gen. Microbiol.* **129**:2495-2500.

Maleszka, R. and H. Schneider. 1982. Fermentation of D-xylose, xylitol and D-xylulose by yeasts. *Can. J. Microbiol.* **28**:360-363.

Maleszka, R. and H. Schneider. 1984. Involvement of oxygen and mitochondrial function in the metabolism of D-xylulose by *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.* **228**:22-30.

Marino, M.H. 1991. Expression systems for heterologous protein production. *In Pichia* yeast expression system. Prospectus, The Phillips Petroleum Company, Bartlesville, Oklahoma.

Mead, D.J., D.C.J. Gardner and S.G. Oliver. 1986. The yeast 2μ plasmid: strategies for the survival of selfish DNA. *Mol. Gen. Genet.* **205**:417-422.

Meilhoc, E., J.M. Masson and J. Tessie. 1990. High-efficiency transformation of intact yeast cells by electroporation. *Bio/Technol.* **8**:223-227.

Meinander, N. and B. Hahn-Hägerdal. 1997a. Fed-batch xylitol production with two recombinant *Saccharomyces cerevisiae* strains expressing *XYL1* at different levels, using glucose as cosubstrate: a comparison of production parameters and strain stability. *Biotechnol. Bioeng.* **54**:391-399.

Meinander, N. and B. Hahn-Hägerdal. 1997b. Influence of cosubstrate concentration on xylose conversion by recombinant, *XYL1*-expressing *Saccharomyces cerevisiae*: a comparison of different sugars and ethanol as cosubstrates. *Appl. Environ. Microbiol.* **63**:1959-1964.

Meinander, N. B. Hahn-Hägerdal, M. Linko, P. Linko and H. Ojamo. 1994. Fed-batch xylitol production with recombinant *XYL1*-expressing *Saccharomyces cerevisiae* using ethanol as a cosubstrate. *Appl. Microbiol. Biotechnol.* **42**:334-339.

Meinander, N., G. Zacchi and B. Hahn-Hägerdal. 1996. A heterologous reductase affects the redox balance of recombinant *Saccharomyces cerevisiae*. *Microbiol.* **142**:165-172.

Miller, J.H. 1983. Mutational specificity in bacteria. *Annu. Rev. Genet.* 7:215-238.

Mishra, P. and A. Singh. 1993. Microbial pentose utilization. *Adv. Appl. Microbiol.* **39**:91.

Moes, C.M., I.S. Pretorius and W.H. van Zyl. 1996. Cloning and expression of the *Clostridium thermosulfurogenes* D-xylose isomerase gene (*xyl*A) in *Saccharomyces cerevisiae*. *Biotechnol. Lett.* **18**:269-274.

Murray, J.A.H. 1987. Bending the rules: the 2μ plasmid of yeast. *Mol. Microbiol.* **1**:1-4.

Murtagh, J.E. 1986. Fuel ethanol production - the U.S. experience. *Process Biochem.* **21**:61-65.

Neale, A.D,; R.K. Schopes and J.M. Kelly. 1988. Alcohol production from glucose and xylose using *Escherichia coli* containing *Zymomonas mobilis* genes. *Appl. Microbiol. Biotechnol.* **29**:162-167.

Neirinck, L.G., R. Maleszka and H. Schneider. 1984. The requirement of oxygen for incorporation of carbon from D-xylose and D-glucose by *Pachysolen tannophilus*. *Arch. Biochem. Biophys.* **228**:13-21.

Neirinck, L.G., C.S. Tsai, J.L. Labelle and H. Schneider. 1985. Xylitol as a carbon source for growth and ethanol production by *Pachysolen tannophilus*. *Can. J. Microbiol.* **31**:451-455.

Nicholas, K.B. and H.B. Nicholas Jnr. 1997. Genedoc: a tool for editing and annotating multiple sequence alignments.

Nigam, J.N., R.S. Ireland, A. Margaritas and M.A. Lachance. 1985. Isolation and screening of some yeasts that ferment D-xylose directly to ethanol. *Appl. Environ. Microbiol.* **50**: 1486-1489.

Nissen, T. and J. Nielsen. 1999. Metabolic engineering of *Saccharomyces cerevisiae* for the improvement of ethanol production. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, O9.

Ohta, K., F. Alterthum, and L.O. Ingram. 1990. Effects of environmental conditions on xylose fermentation by recombinant *Escherichia coli. Appl. Environ. Microbiol.* **56**:463-465.

Ohta, K, D.S. Beall, J.P. Meijia, K.T. Shanmugam and L.O. Ingram. 1991a. Genetic improvement of *Escherichia coli* for ethanol production: chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. *Appl. Environ. Microbiol.* **57**:893-900.

Ohta, K, D.S. Beall, J.P. Meijia, K.T. Shanmugam and L.O. Ingram. 1991b. Metabolic engineering of *Klebsiella oxytoca* strain M5A1 for ethanol production from xylose and glucose. *Appl. Environ. Microbiol.* **57**:2810-2815.

Orr-Weaver, T.L.; J.W. Szostak and R.J. Rothstein. 1983. Genetic application of yeast transformations with linear and gapped plasmids. *Methods Enzymol.* **101**:228-245.

Oura, E. 1977. Reaction products of yeast fermentations. *Process Biochem.* **12**:19-35.

Padukone, N., K.W. Evans, J.D. Mcmillan and C.E. Wyman. 1995. Characterization of recombinant *Escherichia coli* ATCC 11303 (pLO1 297) in the conversion of cellulose and xylose to ethanol. *Appl. Microbiol. Biotechnol.* **43**:850-855.

Paquin, C.E. and V.M. Williamson. 1986. Ty insertions account for most of the spontaneous antimycin A resistance mutations during growth at 15C of *Saccharomyces cerevisiae* lacking ADHI. *Mol. Cell. Biol.* **6**:70-79.

Parekh, S.R., S. Yu and M. Wayman. 1986. Adaptation of *Candida shehatae* and *Pichia stipitis* to wood hydrolysates for increased ethanol production. *Appl. Microbiol. Biotechnol.* **25**:300-304.

Pepper, T. and P.M. Olinger. 1988. Xylitol in sugar-free confections. *Food Technol.* **10**:98-106.

Persson, B., J. Hallborn, M. Walfridsson, B. Hahn-Hägerdal, S. Keränen, M. Pentillä and H. Jornvall. 1993. Dual relationships of xylitol and alcohol dehydrogenases in families of two protein types. *FEBS Lett.* **324**:9-14.

Pettersen, R.C. 1984. The chemical composition of wood. *Adv. Chem. Ser.* **207**:57-126.

Pilgrim, D. and E.T. Young. 1987. Primary sequence structure requirements for correct sorting of the yeast mitochondrial protein ADHIII to the yeast mitochondrial matrix space. *Mol. Cell. Biol.* **7**:294-304.

Pina, A., I.L. Calderon and T. Benitez. 1986. Intergeneric hybrids of *Saccharomyces cerevisiae* and *Zygosaccharomyces fermentati* obtained by protoplast fusion. *Appl. Environ. Microbiol.* **51**:995-1003.

Price, V.L., W.E. Taylor, W. Clevenger, M. Worthington and E.T. Young. 1990. Expression of heterologous proteins in *Saccharomyces cerevisiae* using *ADH2* promoter. *Methods Enzymol.* **185**:308-318.

Prior, B.A., M.A. Alexander, V. Yang and T.W. Jeffries. 1988. The role of alcohol dehydrogenase in the fermentation of D-xylose by *Candida shehatae* ATCC 22984. *Biotechnol. Lett.* **10**:37-42.

Purvis, I.J., A.J.E. Bettany, L. Loughlin and A.J.P. Brown. 1987. Translation and stability of an *Escherichia coli* β-galactoside mRNA expressed under the control of pyruvate kinase sequence in *Saccharomyces cerevisiae. Nucleic Acids Res.* **15**:7963-7974.

Ratzkin, B. and J. Carbon. 1977. Functional expression of cloned yeast DNA in *Escherichia coli. Proceed. Natl. Acad. Sci.* **74**:487-491.

Richard, P., M.H. Toivari and M. Penttila. 1999. Evidence that the gene YLR070c of *Saccharomyces cerevisiae* encodes a xylitol dehydrogenase. *FEBS Lett.* **457**:135-138.

Rizzi, M., P. Erlemann, N.A. Bui-Thanh and H. Dellweg. 1988. Xylose fermentation by yeasts. 4. Purification and kinetic studies of xylose reductase from *Pichia stipitis. Appl. Microbiol. Biotechnol.* **29**:148-154.

Rizzi, M., K. Harwart, N.A. Bui-Thanh, P. Erlemann and H. Dellweg. 1989a. A kinetic study of the NAD-xylitol dehydrogenase from the yeast *Pichia stipitis. J. Ferm. Bioeng.* **67**:25-30.

Rizzi, M., K. Harwart, P. Erlemann, N.A. Bui-Thanh and H. Dellweg. 1989b. Purification and properties of the NAD-xylitol dehydrogenase from the yeast *Pichia stipitis. J. Ferm. Bioeng.* **67**:20-24.

Rizzi, M., C. Klein, C. Schulz, N.A. Bui-Thanh and H. Dellweg. 1989c. Xylose fermentation by yeasts. 5. Use of ATP balances for modeling oxygen-limited growth and fermentation of the yeast *Pichia stipitis* with xylose as carbon source. *Biotechnol. Bioeng.* **34**:509-514.

Roca, C, B. Hahn-Hägerdal and L. Olsson. 1999. Characterization of a *Saccharomyces* cerevisiae strain containing the XYL-genes on different mixtures of sugars. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, A14.

Romanos, M.A., C.A. Scorer and J.J. Clare. 1992. Foreign gene expression in yeast: a review. *Yeast* **8**:423-488.

Rose, M.D. and J.R. Broach. 1991. Cloning genes by complementation in yeast. *In* Guide to yeast genetics and molecular biology, vol. 194, 195-230, C. Guthrie and G.R. Fink (eds.), Academic Press Inc., San Diego, California.

Rose, M., P. Grisafi and D. Botstein. 1984. Structure and function of the yeast *URA3* gene: expression in *Escherichia coli. Gene* **29**:113-124.

Rose, M.D., F. Winston and P. Hieter. 1990. Methods in yeast genetics: a laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Rosenberg, S.L. 1980. Fermentation of pentose sugars to ethanol and other neutral products by microorganisms. *Enzyme Microb. Technol.* **2**:185-193.

Rosenfeld, S., P. Stevis and N.W.Y. Ho. 1984. Cloning and characterization of the *XYL* genes from *E. coli. Mol. Gen. Genetics* **194**:410-415.

Rosillo-Calle, F. and D.O. Hall. 1987. Brazilian alcohol: food versus fuel? *Biomass* **12**:97-128.

Rothstein, R.J. 1983. One-step gene disruption in yeasts. *Methods Enzymol.* **101**:202-211.

Rousseau, J.D. and H.G. Lawford. 1993. Fermentation of newsprint prehydrolysate by an ethanologenic recombinant *Escherichia coli. Biotechnol. Lett.* **15**:505-510.

Saari, G.C., A.A. Kumar, G.H. Kawasaki, M.Y. Insley and P.J. O'Hara. 1987. Sequence of the *Ampullariella* sp strain 3876 gene coding for xylose isomerase. *J. Bacteriol.* **169**:612-618.

Saddler, J.N. 1999. Biotechnology for the conversion of lignocellulosics to ethanol: from research to deployment. *Proc. IEA Bioenenergy Workshop: Biotechnology for the conversion of lignocellulose*, P2.

Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**:5463-5467.

Sarthy, A.V., B.L. McConaughy, Z. Lobo, J.A. Sundstrom, C.E. Furlong and B.D. Hall. 1987. Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **53**:1996-2000.

Schepers, H.J., S. Bringer-Meyer and H. Sahm. 1987. Fermentation of D-xylose to ethanol by *Bacillus macerans. Zeitschrift fur Naturforschung* **42c**:401-407.

Schneider, H.P. 1989. Conversion of pentoses to ethanol by yeasts and fungi. *Crit. Rev. Biotechnol.* **9**:1-40.

Schneider, H.P., Y. Wang, Y.K. Chen and R. Maleszka. 1981. Conversion of D-xylose into ethanol by the yeast *Pachysolen tannophilus. Biotechnol Lett.* **3**:89-92.

Senac, T. and B. Hahn-Hagerdal. 1990. Intermediary metabolite concentrations in xylulosee- and glucose-fermenting *Saccharomyces cerevisiae cells. Appl. Environ. Microbiol.* **56**: 120-126.

Shamanna, D.K. and K.E. Sanderson. 1979. Uptake and catabolism of D-xylose in *Salmonella typhimurium* LT2. *J. Bacteriol.* **139**:64-70.

Shi, N.Q., B. Davis, F. Sherman, J. Cruz and T.W. Jeffries. 1999. Disruption of the cytochrome c gene in the xylose-utilizing yeast *P. stipitis* leads to a higher ethanol production. *Yeast* **15**:1021-1030.

Singh, A. and P. Mishra. 1995. Genetic improvement of pentose fermenting microorganisms *In* Microbial Pentose Utilization: Current applications in biotechnology progress in industrial microbiology vol. 33, pp. 351-352. Elsevier Science. B.V., Amsterdam, The Netherlands.

Skoog, K. and B. Hahn-Hägerdal. 1988. Xylose fermentation. *Enzyme Microb. Technol.* **10**: 66-80.

Skoog, K. and B. Hahn-Hägerdal. 1990. Effect of oxygenation on xylose fermentation by *Pichia stipitis. Appl. Environ. Microbiol.* **56**:3389-3394.

Skoog, K., H. Degn, B. Hahn-Hägerdal, J.P. Jacobsen and H.S. Jacobsen. 1992a. Ethanol reassimilation and ethanol tolerance in *Pichia stipitis* CBS 6054 as studied by ^{13}C nuclear magnetic resonance spectroscopy. *Appl. Environ. Microbiol.* **58**:2552-2558.

Skoog, K., H. Jeppsson and B. Hahn-Hägerdal. 1992b. The effect of oxxygenation on glucose fermentation with *Pichia stipitis. Appl. Biochem. Biotechnol.* **35**:369-375.

Slininger, P.J., P.L. Bolen and C.P. Kurtzman. 1987. *Pachysolen tannophilus*: properties and process considerations for ethanol production from D-xylose. *Enzyme Microb. Technol.* **9**:5-15.

Slininger, P.J., R.J. Bothast, M.R. Okos and M.R. Ladisch. 1985. Comparative evaluation of ethanol production by xylose-fermenting yeasts presented with high xylose concentrations. *Biotechnol. Lett.* 7:431-436.

Slininger, P.J., R.J. Bothast, J.E. van Cauwenberge and C.P. Kurtzman. 1982. Conversion of xylose to ethanol by the yeast *Pachysolen tabnnophilus. Biotechnol. Bioeng.* **24**:371-384.

Slininger, P.J., L.E. Branstrator, R.J. Bothast, M.R. Okos and R. Ladisch. 1991. Growth, death and oxygen uptake kinetics of *Pichia stipitis* on xylose. *Biotechnol. Bioeng.* **37**: 973-980.

Slininger, P.J., L.E. Branstrator, J.M. Lomont, B.S. Dien, M.R. Okos, R. Ladisch and R.J. Bothast. 1990. Stoichiometry and kinetics of xylose fermentation by *Pichia stipitis. Biochemical Eng.* **589**:25-40.

Smiley, K.L. and P.L. Bolen. 1982. Demonstration of D-xylose dehydrogenase in *Pachysolen tannophilus. Biotechnol. Lett.* **4**:607-610.

Spencer, J.F.T. and D.M. Spencer. 1983. Genetic improvements of industrial yeasts. *Ann. Rev. Microbiol.* **37**:121-142.

Spencer-Martins, I. 1994. Transport of sugars in yeasts:implications in the fermentation of lignocellulosic materials. *Biores. Technol.* **50**:51-57.

Sreenath, K.S. and T.W. Jeffries. 1987. Batch and membrane-assisted cell recycling in ethanol production by *Candida shehatae. Biotechnol. Lett.* **9**:293-298.

Stevis, P.E.; J.J. Huang and N.W.Y. Ho. 1987. Cloning of the *Pachysolen tannophilus* xylulokinase gene by complementation in *Escherichia coli. Appl. Environ. Microbiol.* **53**: 2975-2977.

Struhl, K. 1989. Molecular mechanisms of transcriptional regulation in yeast. *Ann. Rev. Biochem.* **58**:1051-1077.

Struhl, K. and R.W. Davis. 1980. A physical, genetic and transcriptional map of the yeast *HIS*3 gene of *Saccharomyces cerevisiae. J. Mol. Biol.* **136**:309-332.

Suikho, M.L. and K. Poutanen. 1984. D-xylulose fermentation by free and immobilized *Saccharomyces cerevisiae* cells. *Biotechnol Lett.* **6**:189-194.

Takuma, S., N. Nakashima, M. Tantirungkij, S. Kinoshita, H. Okada, T. Seki and T. Yoshida. 1991. Isolation of xylose reductase of *Pichia stipitis* and its expression in *Saccharomyces cerevisiae. Appl. Biochem. Biotechnol.* **28**:327-340.

Tantirungkij, M.; T. Izuishi; T. Seki and T. Yoshida. 1994a. Fed-batch fermentation of xylose by a fast-growing mutant of xylose-assimilating recombinant *Saccharomyces cerevisiae. Appl. Microbiol. Biotechnol.* **41**:8-12.

Tantirungkij, M.; N. Nakashima; T. Seki and T Yoshida. 1993. Construction of xylose-assimilating *Saccharomyces cerevisiae. J. Ferment. Bioeng.* **75**:83-88.

Tantirungkij, M., T. Seki and T. Yoshida. 1994b. Genetic improvement of *Saccharomyces cerevisiae* for ethanol production from xylose. *Annals. N.Y. Acad. Sciences* 138-147.

Thestrup, H.N. and B. Hahn-Hägerdal. 1995. Xylitol formation and reduction equivalent generation during anaerobic xylose conversion with glucose as cosubstrate in recombinant *Saccharomyces cerevisiae* expressing the *xyl1* gene. *Appl. Environ. Microbiol.* **61**:2043-2045.

Toivola, A., D. Yarrow, E. van den Bosch, J.P. van Dijken and W.A. Scheffers. 1984. Alcoholic fermentation of D-xylose by yeasts. *Appl. Environ. Microbiol.* **6**:1221-1223.

Tolan, J.S. and R.K. Finn. 1987a. Fermentation of D-xylose and L-arabinose to ethanol by *Erwinia chrysanthemi. Appl. Environ. Microbiol.* **53**:2033-2038.

Tolan, J.S. and R.K. Finn. 1987b. Fermentation of D-xylose to ethanol by genetically modified *Klebsiella planticola. Appl. Environ. Microbiol.* **53**:2039-2044.

Tschumper, G. and J. Carbon. 1980. Sequence of a yeast DNA fragment containing a chromosomal replicator and the *TRP*1 gene. *Gene* **10**:157-166.

Tubb, R.S. 1985. Genetics of ethanol-producing microorganisms. *In* Critical Reviews in Biotechnology, pp. 241-261. G.G. Stewart and I. Russel (eds.). CRC press Inc., Boca Raton, Florida.

Tuite, M.F., M.J. Dobson, N.A. Roberts, R.M. King, D.C. Burke, S.M. Kingsman and A.J. Kingsman. 1982. Regulated high efficiency expression of human interferon-alpha in *Saccharomyces cerevisiae. EMBO J.* **1**:603-608.

Ueng, P.P., C.A. Hunter, C.S. Gong and G.T. Tsao. 1981. D-xylulose fermentation in yeasts. *Biotechnol. Lett.* **3**:315-320.

van Zyl, C, B.A. Prior, S.G. Kilian and J.L.F. Kock. 1989. D-xylose utilization by *Saccharomyces cerevisiae. J. Gen. Microbiol.* **135**:2791-2798.

van Zyl, C, B.A. Prior, S.G. Kilian and E.V. Brandt. 1993. Role of D-ribose as cometabolite in D-xylose metabolism by *Saccharomyces cerevisiae. Appl. Environ. Microbiol.* **59**: 1487-1494.

Verduyn, C., J. Frank Jnr., J.P. van Dijken and W.A. Scheffers. 1985a. Multiple forms of xylose reductase in *Pachysolen tannophilus* CBS 4044. *FEMS Microbiol. Lett.* **30**:313-317.

Verduyn, C., R. van Kleef, J. Frank Jnr., H. Schreuder, J.P. van Dijken and W.A. Scheffers. 1985b. Properties of the NAD(PH)-dependent xylose reductase from the xylose-fermenting yeast *Pichia stipitis. Biochem. J.* **226**:669-677.

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

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

Walfridsson, M., J. Hallborn, M. Penttilä, S. Keränen and B. Hahn-Hägerdal. 1995. Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the *TKL*1 and *TAL*1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Appl. Environ. Microbiol.* **61**:4184-4190.

Wang, P.Y., B.F. Johnson and H. Schneider. 1980a. Fermentation of D-xylose by yeasts using glucose isomerase in the medium to convert D-xylose to D-xylulose. *Biotechnol. Lett.* **2**: 273-278.

Wang, P.Y., C. Shopsis and H. Schneider. 1980b. Fermentation of a pentose by yeasts. *Biochem. Biophys. Res. Commun.* **94**:248-254.

Wiegel, J. and L.G. Ljungdahl. 1986. The importance of thermophilic bacteria in biotechnology. *CRC Crit. Rev. Biotechnol.* **3**:39-107.

Wierenga, R.K., P. Terpstra and W.G.J. Hol. 1986. Prediction of the occurrence of the ADP-binding βαβ-fold in protein using an amino acid sequence fingerprint. *J. Mol. Biol.* **187**: 101-107.

Wyman, C.E. 1999. Commercial ethanol production from agricultural residues. *Proc. IEA Bioenergy Workshop: Biotechnology for the conversionof lignocellulose*, O2.

Yang, V.W. and T.W. Jeffries. 1990. Purification and properties of xylitol dehydrogenase from the xylose-fermenting yeast *Candida shehatae*. *Appl. Biochem. Biotechnol.* **26**:197-206.

Yang, V.W., J.A. Marks, B.P. Davis and T.W. Jeffries. 1994. High-efficiency transformation of *Pichia stipitis* based on its *URA3* and a homologous autonomous replication sequence *ARS2*. *Appl. Environ. Microbiol.* **60**:4245-4254.

Yokoyama, S-I, Y. Kinoshita, T. Suzuki, K. Kawai, H. Horitsu and K. Takamizawa. 1995a. Cloning and sequencing of two D-xylose reductase genes (*xyrA* and *xyrB*) from *Candida tropicalis*. *J. Ferment. Bioeng.* **80**:603-605.

Yokoyama, S-I, T. Suzuki, K. Kawai, H. Horitsu and K. Takamizawa. 1995b. Purification, characterization and structure analysis of NADPH-dependent D-xylose reductases from *Candida tropicalis*. *J. Ferment. Bioeng.* **79**:217-223.

Yu, S., H. Jeppsson and B. Hahn-Hägerdal. 1995. Xylulose fermentation by *Saccharomyces cerevisiae* and xylose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* **44**:314-320.

Yu, S., M. Wayman and S.K. Parekh. 1987. Fermentation to ethanol of pentose containing spent sulphite liquor. *Biotechnol. Bioeng.* **29**:1144-1150.

Zaret, K.S. and F. Sherman. 1982. DNA sequence required for efficient trancription termination in yeast. *Cell* **28**:563-573.

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

APPENDIX ONE

Pichia stipitis XYL1 GENE

```
PsY633      ATGCCTTCTATTAAGTTGAACTCTGGTTACGACATGCCAGCCGTCGGTTTCGGCTGTTGG
PsCBS5773   ATGCCTTCTATTAAGTTGAACTCTGGTTACGACATGCCAGCCGTCGGTTTCGGCTGTTGG   60
            ************************************************************

PsY633      AAAGTCGACATCGACACCTGTTCTGAACAGATCTACCGTGCTATCAAGACCGGCTACAGA
PsCBS5773   AAAGTCGACGTCGACACCTGTTCTGAACAGATCTACCGTGCTATCAAGACCGGTTACAGA  120
            *********.**************************************** ******

PsY633      TTGTTTGACGGTGCCGAAGATTACGCCAACGAAAAGTTAGTTGGTGCCGGTGTCAAGAAG
PsCBS5773   TTGTTCGACGGTGCCGAAGATTACGCCAACGAAAAGTTAGTTGGTGCCGGTGTCAAGAAG  180
            ***** ******************************************************

PsY633      GCCATTGACGAAGGTATTGTCAAGCGTGAAGACTTGTTCCTTACCTCCAAGTTGTGGAAC
PsCBS5773   GCCATTGACGAAGGTATCGTCAAGCGTGAAGACTTGTTCCTTACCTCCAAGTTGTGGAAC  240
            ***************** ******************************************

PsY633      AACTACCACCACCCAGACAACGTCGAAAAGGCCTTGAACAGAACCCTTTCTGACTTGCAA
PsCBS5773   AACTACCACCACCCAGACAACGTCGAAAAGGCCTTGAACAGAACCCTTTCTGACTTGCAA  300
            ************************************************************

PsY633      GTTGACTACGTTGACTTGTTCTTGATCCACTTCCCAGTCACCTTCAAGTTCGTTCCATTA
PsCBS5773   GTTGACTACGTTGACTTGTTCTTGATCCACTTCCCAGTCACCTTCAAGTTCGTTCCATTA  360
            ************************************************************

PsY633      GAAGAAAAGTACCCACCAGGATTCTACTGTGGTAAGGGTGACAACTTCGACTACGAAGAT
PsCBS5773   GAAGAAAAGTACCCACCAGGATTCTACTGTGGTAAGGGTGACAACTTCGACTACGAAGAT  420
            ************************************************************

PsY633      GTTCCAATTTTAGAGACCTGGAAGGCTCTTGAAAAGTTGGTCAAGGCCGGTAAGATCAGA
PsCBS5773   GTTCCAATTTTAGAGACCTGGAAGGCTCTTGAAAAGTTGGTCAAGGCCGGTAAGATCAGA  480
            ************************************************************

PsY633      TCTATCGGTGTTTCTAACTTCCCAGGTGCTTTGCTCTTGGACTTGTTGAGAGGTGCTACC
PsCBS5773   TCTATCGGTGTTTCTAACTTCCCAGGTGCTTTGCTCTTGGACTTGTTGAGAGGTGCTACC  540
            ************************************************************

PsY633      ATCAAGCCATCTGTCTTGCAAGTTGAACACCACCCATACTTGCAACAACCAAGATTGATC
PsCBS5773   ATCAAGCCATCTGTCTTGCAAGTTGAACACCACCCATACTTGCAACAACCAAGATTGATC  600
            ************************************************************

PsY633      GAATTCGCTCAATCCCGTGGTATTGCTGTCACCGCTTACTCTTCGTTCGGTCCTCAATCT
PsCBS5773   GAATTCGCTCAATCCCGTGGTATTGCTGTCACCGCTTACTCTTCGTTCGGTCCTCAATCT  660
            ************************************************************

PsY633      TTCGTTGAAATGAACCAAGGTAGAGCTTTGAACACTTCTCCATTGTTCGAGAACGAAACT
PsCBS5773   TTCGTTGAATTGAACCAAGGTAGAGCTTTGAACACTTCTCCATTGTTCGAGAACGAAACT  720
            *********:**************************************************

PsY633      ATCAAGGCTATCGCTGCTAAGCACGGTAAGTCTCCAGCTCAAGTCTTGTTGAGATGGTCT
PsCBS5773   ATCAAGGCTATCGCTGCTAAGCACGGTAAGTCTCCAGCTCAAGTCTTGTTGAGATGGTCT  780
            ************************************************************

PsY633      TCCCAAAGAGGCATTGCCATCATTCCAAAGTCCAACACTGTCCCAAGATTGTTGGAAAAC
PsCBS5773   TCCCAAAGAGGCATTGCCATCATTCCAAAGTCCAACACTGTCCCAAGATTGTTGGAAAAC  840
            ************************************************************

PsY633      AAGGACGTCAACAGCTTCGACTTGGACGAACAAGATTTCGCTGACATTGCCAAGTTGGAC
PsCBS5773   AAGGACGTCAACAGCTTCGACTTGGACGAACAAGATTTCGCTGACATTGCCAAGTTGGAC  900
            ************************************************************

PsY633      ATCAACTTGAGATTCAACGACCCATGGGACTGGGACAAGATTCCAATCTTCGTCTAA
PsCBS5773   ATCAACTTGAGATTCAACGACCCATGGGACTGGGACAAGATTCCTATCTTCGTCTAA  957
            ******************************************:*************
```

Pichia stipitis XYL2 GENE

```
PsY633      ATGACTGCTAACCCTTCCTTGGTGTTGAACAAGATCGACGACATTTCGTTCGAAACTTAC
PsCBS5773   ATGACTGCTAACCCTTCCTTGGTGTTGAACAAGATCGACGACATTTCGTTCGAAACTTAC   60
            ************************************************************

PsY633      GATGCCCCAGAAATCTCTGAACCTACCGATGTCCTCGTCCAGGTCAAGAAAACCGGTATC
PsCBS5773   GATGCCCCAGAAATCTCTGAACCTACCGATGTCCTCGTCCAGGTCAAGAAAACCGGTATC  120
            ************************************************************

PsY633      TGTGGTTCCGACATCCACTTCTACGCCCATGGTAGAATCGGTAACTTCGTTTTGACCAAG
PsCBS5773   TGTGGTTCCGACATCCACTTCTACGCCCATGGTAGAATCGGTAACTTCGTTTTGACCAAG  180
            ************************************************************

PsY633      CCAATGGTCTTGGGTCACGAATCCGCCGGTACTGTTGTCCAGGTTGGTAAGGGTGTCACC
PsCBS5773   CCAATGGTCTTGGGTCACGAATCCGCCGGTACTGTTGTCCAGGTTGGTAAGGGTGTCACC  240
            ************************************************************

PsY633      TCTCTTAAGGTTGGTGACAACGTCGCTATCGAACCAGGTATTCCATCCAGATTCTCCGAC
PsCBS5773   TCTCTTAAGGTTGGTGACAACGTCGCTATCGAACCAGGTATTCCATCCAGATTCTCCGAC  300
            ************************************************************

PsY633      GAATACAAGAGCGGTCACTACAACTTGTGTCCTCACATGGCCTTCGCCGCTACTCCTAAC
PsCBS5773   GAATACAAGAGCGGTCACTACAACTTGTGTCCTCACATGGCCTTCGCCGCTACTCCTAAC  360
            ************************************************************

PsY633      TCCAAGGAAGGCGAACCAAACCCACCAGGTACCTTATGTAAGTACTTCAAGTCGCCAGAA
PsCBS5773   TCCAAGGAAGGCGAACCAAACCCACCAGGTACCTTATGTAAGTACTTCAAGTCGCCAGAA  420
            ************************************************************

PsY633      GACTTCTTGGTCAAGTTGCCAGACCACGTCAGCTTGGAACTCGGTGCTCTTGTTGAGCCA
PsCBS5773   GACTTCTTGGTCAAGTTGCCAGACCACGTCAGCTTGGAACTCGGTGCTCTTGTTGAGCCA  480
            ************************************************************

PsY633      TTGTCTGTTGGTGTCCACGCCTCCAAGTTGGGTTCCGTTGCTTTCGGCGACTACGTTGCC
PsCBS5773   TTGTCTGTTGGTGTCCACGCCTCCAAGTTGGGTTCCGTTGCTTTCGGCGACTACGTTGCC  540
            ************************************************************

PsY633      GTCTTTGGTGCTTGGCCTGTTGGTCTTTTGGCTGCTGCTGTCGCCAAGACCTTCGGTGCT
PsCBS5773   GTCTTTGGTGCTGGTCCTGTTGGTCTTTTGGCTGCTGCTGTCGCCAAGACCTTCGGTGCT  600
            ************ * ********************************************

PsY633      AAGGGTGTCATCGTCGTTGACATTTTCGACAACAAGTTGAAGATGGCCAAGGACATTGGT
PsCBS5773   AAGGGTGTCATCGTCGTTGACATTTTCGACAACAAGTTGAAGATGGCCAAGGACATTGGT  660
            ************************************************************

PsY633      GCTGCTACTCACACCTTCAACTCCAAGACCGGTGGTTCTGAAGAATTGATCAAGGCTTTC
PsCBS5773   GCTGCTACTCACACCTTCAACTCCAAGACCGGTGGTTCTGAAGAATTGATCAAGGCTTTC  720
            ************************************************************

PsY633      GGTGGTAACGTGCCAAACGTCGTTTTGGAATGTACTGGTGCTGAACCTTGTATCAAGTTG
PsCBS5773   GGTGGTAACGTGCCAAACGTCGTTTTGGAATGTACTGGTGCTGAACCTTGTATCAAGTTG  780
            ************************************************************

PsY633      GGTGTTGACGCCACTGCCCCAGGTGGTCGTTTCGTTCAAGTTGGTAACGCTGCTGGTCCA
PsCBS5773   GGTGTTGACGCCATTGCCCCAGGTGGTCGTTTCGTTCAAGTTGGTAACGCTGCTGGTCCA  840
            ************* ************************************************

PsY633      GTCAGCTTCCCAATCACCGTTTTCGCCATGAAGGAATTGACTTTGTTCGGTTCTTTCAGA
PsCBS5773   GTCAGCTTCCCAATCACCGTTTTCGCCATGAAGGAATTGACTTTGTTCGGTTCTTTCAGA  900
            ************************************************************

PsY633      TACGGATTCAACGACTACAAGACTGCTGTTGGAATCTTTGACACTAACTACCAAAACGGT
PsCBS5773   TACGGATTCAACGACTACAAGACTGCTGTTGGAATCTTTGACACTAACTACCAAAACGGT  960
            ************************************************************


PsY633      TGAGAAAATGCTCCAATTGACTTTGAACAATTGATCACCCACAGATACAAGTTCAAGGAC
PsCBS5773   AGAGAAAATGCTCCAATTGACTTTGAACAATTGATCACCCACAGATACAAGTTCAAGGAC  1020
            :************************************************************
```

```
PsY633      GCTATTGAAGCCTACGACTTGGTAAGAGCCGGTAAGGGTGCTGTCAAGTGTCTCATTGAC
PsCBS5773   GCTATTGAAGCCTACGACTTGGTCAGAGCCGGTAAGGGTGCTGTCAAGTGTCTCATTGAC  1080
            ***********************.************************************

PsY633      GGCCCTGAGTAA
PsCBS5773   GGCCCTGAGTAA                                                  1092
            ************
```

Saccharomyces cerevisiae XYL3 GENE

```
ScY294      ATGTTGTGTTCAGTAATTCAGAGACAGACAAGAGAGGTTTCCAACACAATGTCTTTAGAC
ScX3        ATGTTGTGTTCAGTAATTCAGAGACAGACAAGAGAGGTTTCCAACACAATGTCTTTAGAC    60
            ************************************************************

ScY294      TCATACTATCTTGGGTTTGATCTTTCGACCCAACAACTGAAATGTCTCGCCATTAACCAG
ScX3        TCATACTATCTTGGGTTTGATCTTTCGACCCAACAACTGAAATGTCTCGCCATTAACCAG   120
            ************************************************************

ScY294      GACCTAAAAATTGTCCATTCAGAAACAGTGGAATTTGAAAAGGATCTTCCGCATTATCAC
ScX3        GACCTAAAAATTGTCCATTCAGAAACAGTGGAATTTGAAAAGGATCTTCCGCATTATCAC   180
            ************************************************************

ScY294      ACAAAGAAGGGTGTCTATATACACGGCGACACTATCGAATGTCCCGTAGCCATGTGGTTA
ScX3        ACAAAGAAGGGTGTCTATATACACGGCGACACTATCGAATGTCCCGTAGCCATGTGGTTA   240
            ************************************************************

ScY294      GAGGCTCTAGATCTGGTTCTCTCGAAATATCGCGAGGCTAAATTTCCATTGAACAAAGTT
ScX3        GAGGCTCTAGATCTGGTTCTCTCGAAATATCGCGAGGCTAAATTTCCATTGAACAAAGTT   300
            ************************************************************

ScY294      ATGGCCGTCTCAGGGTCCTGCCAGCAGCACGGGTCTGTCTACTGGTCCTCCCAAGCCGAA
ScX3        ATGGCCGTCTCAGGGTCCTGCCAGCAGCACGGGTCTGTCTACTGGTCCTCCCAAGCCGAA   360
            ************************************************************

ScY294      TCTCTGTTAGAGCAATTGAATAAGAAACCGGAAAAAGATTTATTGCACTACGTGAGCTCT
ScX3        TCTCTGTTAGAGCAATTGAATAAGAAACCGGAAAAAGATTTATTGCACTACGTGAGCTCT   420
            ************************************************************

ScY294      GTAGCATTTGCAAGGCAAACCGCCCCCAATTGGCAAGACCACAGTACTGCAAAGCAATGT
ScX3        GTAGCATTTGCAAGGCAAACCGCCCCCAATTGGCAAGACCACAGTACTGCAAAGCAATGT   480
            ************************************************************

ScY294      CAAGAGTTTGAAGAGTGCATAGGTGGGCCTGAAAAAATGGCTCAATTAACAGGGTCCAGA
ScX3        CAAGAGTTTGAAGAGTGCATAGGTGGGCCTGAAAAAATGGCTCAATTAACAGGGTCCAGA   540
            ************************************************************

ScY294      GCCCATTTTAGATTTACTGGTCCTCAAATTCTGAAAATTGCACAATTAGAACCAGAGGCT
ScX3        GCCCATTTTAGATTTACTGGTCCTCAAATTCTGAAAATTGCACAATTAGAACCAGAAGCT   600
            ***********************************************************.***

ScY294      TACGAAAAAACAAAGACCATTTCTTTAGTGTCTAATTTTTTGACTTCTATCTTAGTGGGC
ScX3        TACGAAAAAACAAAGACCATTTCTTTAGTGTCTAATTTTTTGACTTCTATCTTAGTGGGC   660
            ************************************************************

ScY294      CATCTTGTTGAATTAGAGGAGGCAGATGCCTGTGGTATGAACCTTTATGATATACGTGAA
ScX3        CATCTTGTTGAATTAGAGGAGGCAGATGCCTGTGGTATGAACCTTTATGATATACGTGAA   720
            ************************************************************

ScY294      AGAAAATTCAGTGATGAGCTACTACATCTAATTGATAGTTCTTCTAAGGATAAAACTATC
ScX3        AGAAAATTCAGTGATGAGCTACTACATCTAATTGATAGTTCTTCTAAGGATAAAACTATC   780
            ************************************************************

ScY294      AGACAAAAATTAATGAGAGCACCCATGAAAAATTTGATAGCGGGTACCATCTGTAAATAT
ScX3        AGACAAAAATTAATGAGAGCACCCATGAAAAATTTGATAGCGGGTACCATCTGTAAATAT   840
            ************************************************************
```

```
ScY294      TTTATTGAGAAGTACGGTTTCAATACAAACTGCAAGGTCTCTCCCATGACTGGGGATAAT
ScX3        TTTATTGAGAAGTACGGTTTCAATACAAACTGCAAGGTCTCTCCCATGACTGGGGATAAT  900
            ************************************************************

ScY294      TTAGCCACTATATGTTCTTTACCCCTGCGGAAGAATGACGTTCTCGTTTCCCTAGGAACA
ScX3        TTAGCCACTATATGTTCTTTACCCCTGCGGAAGAATGACGTTCTCGTTTCCCTAGGAACA  960
            ************************************************************

ScY294      AGTACTACAGTTCTTCTGGTCACCGATAAGTATCACCCCTCTCCGAACTATCATCTTTTC
ScX3        AGTACTACAGTTCTTCTGGTCACCGATAAGTATCACCCCTCTCCGAACTATCATCTTTTC  1020
            ************************************************************

ScY294      ATTCATCCAACTCTGCCAAACCATTATATGGGTATGATTTGTTATTGTAATGGTTCTTTG
ScX3        ATTCATCCAACTCTGCCAAACCATTATATGGGTATGATTTGTTATTGTAATGGTTCTTTG  1080
            ************************************************************

ScY294      GCAAGGGAGAGGATAAGAGACGAGTTAAACAAAGAACGGGAAAATAATTATGAGAAGACT
ScX3        GCAAGGGAGAGGATAAGAGACGAGTTAAACAAAGAACGGGAAAATAATTATGAGAAGACT  1140
            ************************************************************

ScY294      AACGATTGGACTCTTTTTAATCAAGCTGTGCTAGATGACTCAGAAAGTAGTGAAAATGAA
ScX3        AACGATTGGACTCTTTTTAATCAAGCTGTGCTAGATGACTCAGAAAGTAGTGAAAATGAA  1200
            ************************************************************

ScY294      TTAGGTGTAAATTTTCCTCTGGGGGAGATCGTTCCTAGCGTAAAAGCCATAAACAAAAGG
ScX3        TTAGGTGTAAATTTTCCTCTGGGGGAGATCGTTCCTAGCGTAAAAGCCATAAACAAAAGG  1260
            ************************************************************

ScY294      GTTATCTTCAATCCAAAAACGGGTATGATTGAAAGAGAGGTGGCCAAGTTCAAAGACAAG
ScX3        GTTATCTTCAATCCAAAAACGGGTATGATTGAAAGAGAGGTGGCCAAGTTCAAAGACAAG  1320
            ************************************************************

ScY294      AGGCACGATGCCAAAAATATTGTAGAATCACAGGCTTTAAGTTGCAGGGTAAGAATATCT
ScX3        AGGCACGATGCCAAAAATATTGTAGAATCACAGGCTTTAAGTTGCAGGGTAAGAATATCT  1380
            ************************************************************

ScY294      CCCCTGCTTTCGGATTCAAACGCAAGCTCACAACAGAGACTGAACGAAGATACAATCGTG
ScX3        CCCCTGCTTTCGGATTCAAACGCAAGCTCACAACAGAGACTGAACGAAGATACAATCGTG  1440
            ************************************************************

ScY294      AAGTTTGATTACGATGAATCTCCGCTGCGGGACTACCTAAATAAAAGGCCAGAAAGGACT
ScX3        AAGTTTGATTACGATGAATCTCCGCTGCGGGACTACCTAAATAAAAGGCCAGAAAGGACT  1500
            ************************************************************

ScY294      TTTTTTGTAGGTGGGGCTTCTAAAAACGATGCTATTGTGAAGAAGTTTGCTCAAGTCATT
ScX3        TTTTTTGTAGGTGGGGCTTCTAAAAACGATGCTATTGTGAAGAAGTTTGCTCAAGTCATT  1560
            ************************************************************

ScY294      GGTGCTACAAAGGGTAATTTTAGGCTAGAAACACCAAACTCATGTGCCCTTGGTGGTTGT
ScX3        GGTGCTACAAAGGGTAATTTTAGGCTAGAAACACCAAACTCATGTGCCCTTGGTGGTTGT  1620
            ************************************************************

ScY294      TATAAGGCCATGTGGTCATTGTTATATGACTCTAATAAAATTGCAGTTCCTTTTGATAAA
ScX3        TATAAGGCCATGTGGTCATTGTTATATGACTCTAATAAAATTGCAGTTCCTTTTGATAAA  1680
            ************************************************************

ScY294      TTTCTGAATGACAATTTTCCATGGCATGTAATGGAAAGCATATCCGATGTGGATAATGAA
ScX3        TTTCTGAATGACAATTTTCCATGGCATGTAATGGAAAGCATATCCGATGTGGATAATGAA  1740
            ************************************************************

ScY294      AATTGGGATCGCTATAATTCCAAGATTGTCCCCTTAA
ScX3        AATTGGGATCGCTATAATTCCAAGATTGTCCCCTTAA                         1777
            ************************************
```