Chapter 1

Introduction and Review of literature

Introduction 1
Review of literature 3
1.1 Xylose fermenting organisms 3
1.2 Pathways for D-xylose assimilation 4
  1.2.1 Xylose metabolism in eubacteria, yeast and fungi 4
  1.2.2 Xylose metabolism in archaebacteria 6
1.3 Components of Yeast Xylose metabolism 7
  1.3.1 Xylose Reductase (XR) enzyme 8
  1.3.2 Xylitol Dehydrogenase (XDH) enzyme 11
1.4 Engineering yeasts for Xylose Metabolism 14
  1.4.1 Improvement of xylose uptake with transporters 15
  1.4.2 Engineering of xylose metabolizing proteins 16
  1.4.2.1 Engineering of Xylose reductase 16
  1.4.2.2 Engineering of xylitol dehydrogenase 17
  1.4.3 Approaches for rational engineering of metabolic pathways 18
1.5 Scope of the present study 19
  1.5.1 Scope of studying xylose metabolism in Debaryomyces hansenii 20
  1.5.2 Scope of studying xylose metabolizing enzymes in Pichia angusta 21
INTRODUCTION

Rapid exhaustion of the limited reserve of fossil fuel stocks, geopolitical issues and potential climate change due to global warming led us to search for an environment friendly sustainable alternative energy source. Plant biomass possesses huge potential as a source for the production of biofuels (Lynd et al., 1991) and the concept of using plants as alternative fuel source is rather old. In 1922 biologist A. J. Kluyver remarked: "Their gradual depletion and the resulting higher price of fossil raw materials will consequently entail that the chemical industry will gravitate more and more towards procuring its starting materials immediately from the present-day plant world...".

Large quantities of lignocellulosic material in the form of agricultural and forest residues are available in nature. Thus the incentive to energy recovery from agricultural biomass is an attractive option, as it can reduce environmental pollution and facilitate waste management. Over 150 billion tones of organic substances are photosynthesized annually which consists of three major constituents namely; cellulose, hemicelluloses and lignin with an average proportion of 4:3:3. Celluloses and hemicelluloses are embedded in the matrix of lignin. The celluloses and hemicelluloses are two major compounds that have the potential to be used as raw material for fermentation and thus production of bio-ethanol as an alternative source of fuel. Conversion of biomass to useable energy is not economical unless hemicellulose is used in addition to cellulose.

Cellulose is the bio-polymers of glucose where as hemicelluloses consists a main chain xylan backbone (β-1, 4 linkages) with short, highly branched, chains of sugars: mannose, arabinose, galactose, glucuronic acid, etc. Hemicelluloses contain five carbon sugars (D-xylose and L-arabinose) and six carbon sugars (D-galactose, D-glucose and D-mannose) and uronic acid. The principal components of hemicelluloses are aldo-pentoses (85-90%), which are either in pyranose or furanose form. The predominant form of aldo-pentose present here is, β-D-xylopyranose. Hydrolysis of xylan mostly releases D-xylose. It is the second most abundant sugar (after glucose) in nature and a product of photosynthesis and as such constitutes an inexhaustible renewable resource.

Xylose can be further metabolized by micro-organisms to produce ethanol and other value added products (like ethanol, sugar alcohols, solvents, and organic acids) and thus have enormous potential to be an alternative fuel source as well as a
commercially important molecule. Utilization of D-xylose for the commercial production of valuable chemicals like ethanol, xylitol, acetic acid, 2,3 butanediol, acetone, isopropanol and n-butanol using micro-organisms is thus important for enhancing the economic viability of lignocelluloses utilization. The xylose to ethanol (and other value added products) conversion is possible either by thermo-chemical or biological methods. Thermo-chemical method requires high temperature and pressure and produces a complex mixture of products. The biological conversion method using microorganism operates at lower temperature and produces specific products in high yields with fewer byproducts (Stewart et al., 1983). Fermentation involves simpler technology, and the by-products are mostly nontoxic, unlike those from chemical plants (Palson et al., 1981). For lignocellulose to be amenable to fermentation, it needs to undergo some physical, chemical or enzymatic pre-treatment that release its monomeric sugars, which then can be converted by a microorganism. Hemicellulose can be converted by some microorganisms to various value-added products such as ethanol, sugar alcohols, solvents, and organic acids (Lovitt et al., 1988). The nature and amount of conversion products depends on the type of pentose sugar, metabolic efficiency of the microorganism, and cultural conditions employed. While the fermentation technology employing yeast and bacteria to make chemicals from hexoses is well known, the ability of these organisms to ferment pentoses has been considered problematical (McCracken and Gong, 1983; Kurtzman et al., 1983). The economics of bioconversion would be more feasible if both hexose and pentose sugars present in lignocellulosic materials could be utilized with equal efficiency (Skoog and Hahn-Hagerdal, 1988). Only a handful of organisms can metabolize pentose sugars efficiently. But, the ever increasing demand for biofuel production around the world adds impetus to the search for more and more organisms that can be commercially used towards this end. Extremophilic microorganism that are capable of utilizing xylose are potential source for discovering new pathways, more robust enzymes and tools for undertaking metabolic engineering. To this end, enzymes of xylose metabolic pathway from highly halotolerant yeast Debaryomyces hansenii and thermotolerant yeast Pichia angusta have been studied in the thesis.
D-xylose is the second most abundant sugar in the plant biomass and possesses huge potential as a source for the production of bio-fuel. Efficient conversion of lignocelluloses to ethanol is the major challenge in energy biotechnology presently, as there are many bottle-necks for efficient xylose metabolism and product formation. So, there is a need to devise technologies for improved utilization and fermentation of xylose. Xylose is fermented by a handful of microorganisms and curiously, the best glucose fermenting yeast S. cerevisiae lacks the genes for uptake and metabolism of xylose. Xylose fermenting strains of S. cerevisiae have been constructed through metabolic engineering. Despite these efforts, there is a pressing need to look for better microbes either naturally occurring or created by engineering for this purpose. The focus of this review is to provide the current status of our understanding of xylose metabolism; the components and the pathways; the problem for efficient metabolism and approaches to overcome them.

1.1 Xylose fermenting organisms

A variety of bacteria, yeast, and filamentous fungi naturally ferment xylose to ethanol or other value added products like xylitol etc. (Jeffries, 1983; Toivola et al., 1984; Skoog and Hahn-Hägerdal 1988). The applicability of yeasts to convert carbohydrate to ethanol has been studied extensively however only a few yeast species have been reported so far that can ferment pentoses. More than 200 species of yeast are able to grow on xylose while only a handful of them are able to ferment xylose to ethanol. With respect to xylose fermentation, attention has been focussed principally on Pachysolen tannophilus, Candida shehatae and Pichia stipitis. Other than these organisms, xylose fermenting ability was investigated in other mesophilic yeasts; Clavispora, Brettanomyces, different Candida species e.g. C. tenuis, C. tropicalis, C. utilis, C. blankii, C. parapsilosis and species of Debaryomyces viz, D. nepalensis, D. polymorpha, D. hansenii etc. Among filamentous fungi, Rhizopus sp, Fusarium oxysporum, Neurospora crassa, Monilia sp., Paecilomyces sp have been found to have potential for fermenting both glucose and xylose. N. crassa, F. oxysporum and Monilia sp have been found to convert other sugars in hemicellulose complex (D-glucose, D-mannose, D-galactose, sucrose and xylitol) in addition to xylose. Thus, these fungi are
used in direct conversion of cellulose/ hemicelluloses to ethanol and acetic acid in a single step. Yeasts and fungi have limitation in anaerobic conversion of xylose to ethanol, where bacteria excel and readily convert xylose to a variety of products in absence of oxygen. *Bacillus polymyxa, B. macerans, Aerobacter hydrophila, Klebsiella pneumoniae* are some mesophilic strains which shows the advantage over yeast in fermentation of pentoses. Bacteria have shorter generation rate thus show high conversion rate and they can carry out fermentation without crucial need of oxygenation. Thermophilic bacteria can also ferment a wide range of substrates including cellulose, hemicelluloses, pectin etc. Different species of *Clostridium* e.g. *C. thermocellum, C. thermosulfurogenes* etc., *Thermoanaerobacter ethanolicus* are some examples of promising pentose fermenting organisms. In addition to low oxygen requirement and high metabolic activity, the most advantageous feature of thermophilic bacteria is the ability to carry out fermentation at high temperature thus preventing the risk of contamination.

### 1.2 Pathways for D-xylose assimilation

Few pathways have been discovered for xylose catabolism which varies from one class of organism to other. Most of them proceed via pentose phosphate pathway while there are some exceptions which follow different routes. A brief overview of the pathways is presented below.

#### 1.2.1 Xylose metabolism in eubacteria, yeast and fungi

Major pathways for xylose metabolism are similar in bacteria, yeast and fungi. However significant diversities exist among them with respect to xylose transport, genetic regulation of the pathway, co-factor requirements of the enzymes and the generation of the final products from pyruvate. After the transport into the cell, most hexoses are readily phosphorylated while xylose undergoes few biochemical steps before phosphorylation. The steps before phosphorylation differ in bacteria and yeasts (Fig 1). In yeast and fungi, xylose is reduced to xylitol with the help of NADPH/ NADH dependent xylose reductase (XR) enzyme (aldose reductase EC 1.1.1.21). The xylitol in turn is oxidised to xylulose by NAD⁺ dependent xylitol dehydrogenase (XDH). XDH (EC 1.1.1.9) belongs to NAD dependent medium-chain dehydrogenase (MDR) family and the polyol dehydrogenase (PDH) subfamily (Nordling et al., 2002;
Riveros-Rosas et al., 2003). Most bacteria (except Enterobacteria, Corynebacteria, and Brevibacteria which uses the XR-XDH enzyme system similar to yeast) directly converts xylose into xylulose. This is performed by a single enzymatic step involving xylose isomerase (XI) (D-xylose ketol-isomerase, EC 5.3.1.5) (Hofer et al., 1971). Xylulose thus produced is phosphorylated to form xylulose 5 phosphate by the enzyme xylulose kinase (XK) (EC 2.7.1.17) before entering into the Pentose Phosphate (PP) Pathway. After phosphorylation, ethanol production is believed to proceed through Pentose Phosphate and Embden-Meyerhof-Parnas (EMP) pathways. Within the PP cycle the xylulose 5 phosphate is metabolised to glyceraldehydes 3 phosphate and fructose 6 phosphate. Then these compounds are converted to pyruvate via EMP Pathway. The pyruvate is finally converted to ethanol. In the overall fermentative reaction, 3 moles of xylose are required to produce 5 moles of ethanol. The stoichiometry of the reaction (neglecting the NADPH balance) is shown in the following equation.

\[
3 \text{ Xylose} + 5 \text{ ADP} + 5\text{P}_i = 5 \text{ Ethanol} + 5 \text{ CO}_2 + 5\text{ATP} + 5\text{H}_2\text{O}
\]

Fig 1: Xylose metabolism pathway in bacteria, yeast and fungi.
Two alternative pathways for xylose metabolism, known as the phosphoketolase and oxidative bypasses, also exist in yeast and fungi (McCacken and Gong 1983; Jeffries, 1990). These pathways may also be present in many bacteria (Gottschalk et al., 1986). Phosphoketolase bypass carries out the cleavage of either D-xylulose-5-phosphate or D-fructose-6-phosphate to form glyceraldehyde-3-phosphate or D-erythrose-4-phosphate plus acetyl phosphate.

\[
\text{Mg}^{++} \quad \text{D-xylulose-5-phosphate} \quad \text{phosphate} \quad \text{glyceraldehyde-3phosphate} \quad \text{acetyl phosphate} \quad \text{H}_2\text{O} \\
\text{TPP}
\]

Glyceraldehyde-3-phosphate enters glycolysis and ultimately produces ethanol. Phosphoketolase is known to be present in P. tannophilus, Rhodotorula graminis, Rhodotorula glutinis, C. tropicalis, Candida humicola etc and it has been speculated that this pathway is important for ATP generation, particularly under anaerobic conditions (Evans and Ratledge 1984; Jeffries, 1990). Phosphoketolase is also the key enzyme used by lactic acid bacteria under fermentation condition (Gottschalk et al., 1986). The oxidative bypass oxidizes fructose-6-phosphate to ribulose-5-phosphate (ribulose-5-P) and carbon dioxide through the intermediate 6-phospho-gluconate (Bruinenberg et al., 1983; Evans and Ratledge 1984; Jeffries, 1990).

### 1.2.2 Xylose metabolism in archaebacteria

Knowledge about pentose metabolism in archa has come up only with recent studies (Johnsen and Schonheit, 2004; Brouns et al., 2006; Johnsen et al., 2009). D-xylose and L-arabinose are known to support the growth of few achaeal species eg, Sulfolobus solfataricus and some halophiles like, Haloarcula marismortui, Haloferax volcanii. Some recent reports about H. marismortui and H. volcanii indicated that these archaenal species follow a pathway entirely different from that of bacteria and yeast (Johnsen et al., 2004 and 2009). In this pathway, D-xylose is oxidised to D-xylonolactone involving an inducible xylose dehydrogenase (Fig 2). The xylonolactone forms D-xylonate which is dehydrated into 2-keto-3-deoxyxylonate and is further degraded into α-ketoglutarate via an intermediate step. As the α-ketoglutarate is an intermediate in TCA cycle, thus it enters into the cycle and produce pyruvate. This pyruvate in turn can produce ethanol under anaerobic conditions. This pathway shows similarities to
proposed oxidative pentose degradation pathways to $\alpha$-ketoglutarate in few bacteria, e.g. *Azospirillum brasilense* and *Caulobacter crescentus*, and in the archaeon *Sulfolobus solfataricus*.

![Diagram of xylose catabolism pathway](image)

**Figure 2**: Archaeal xylose catabolism pathway (Adapted from Johnsen et al, 2009).

### 1.3 Components of Yeast Xylose metabolism

Long association of yeast with alcohol fermentation made them preferred organism for commercial xylose fermentation. XR and XDH are the principal components of the yeast xylose metabolism, and their product xylulose is routed via pentose phosphate pathway which is found virtually in all cellular organisms. These enzymes have been the focus of biochemical and structural studies due to the fact that, xylose is highly abundant in nature but is under-utilized as a fermentable sugars. Efficiency of xylose fermentation is limited by the properties of these enzymes. They are potential tool for constructing industrially useful strains for xylose fermentation through metabolic engineering. An overview about the enzymes involved in xylose metabolism of yeast is described here.
1.3.1 Xylose Reductase (XR) enzyme

XR catalyses the NADPH/NADH dependent reduction of xylose to xylitol, which is the first step of xylose metabolic pathway. XRs from yeast belong to AKR2 family of aldo-keto reductase (AKR) superfamily. This superfamily consists of varieties of members which are divided into 15 families. The AKRs are (α/β)₈ barrel proteins, approximately 280 - 350 residues in length, which bind nicotinamide cofactor without a Rossmann-fold motif. Although the majority of AKRs are monomeric proteins, the AKR2, AKR6 and AKR7 family may form multimers. They are present in a wide variety of organisms that includes mammals, amphibians, plants, yeast, protozoa and bacteria. These enzymes act on a diverse range of substrates, including aliphatic and aromatic aldehyde, monosaccharide, steroid, prostaglandin, polycyclic aromatic hydrocarbon and isoflavonoid (Jez et al., 2001; Hyndman et al., 2003). Multiple sequence alignment of the AKRs reveals the presence of three conserved regions - an N-terminal region (LxxxGxxxPxxGxG), an active site region (GxxxxDxAxxY, which contains the conserved Asp and Tyr), and a third region also including the active site (LxxxxxxxDxxxxH, which contains the conserved His). The active site lysine residue is also invariant. In addition, numerous other conserved residues are scattered throughout the sequences (Hyndman et al., 2003).

The XR from P. stipitis was the first yeast homologue that was cloned (Amore et al., 1991). Subsequently XR was also cloned few other yeast species (Yokohama, et al., 1995, Billard et al., 1995, Bolen et al., 1995, Neuhauser et al., 1997 and Lee et al., 2003). Deletion of the gene, encoding XR, renders the yeast cells unable to grow in a xylose containing medium (Voronovovsky et al., 2005). XR activity is dependent on the presence of co-factor. XRs from fungi (N. crassa, C. tropicalis and C. utilis) mostly showed a strict requirement for NADPH while some of them e.g. from P. stipitis and C. tenuis function with either NADPH or NADH. Pachysolen tannophilus contains two distinct XRs. One of them shows dual co-enzyme specificity, while the second one is specific for NADPH. The ability of some yeast to ferment xylose anaerobically has been attributed to the dual NADPH and NADH co-factor specificity of their xylose reductases (Bruinenberg et al., 1984). The xylose reductase of C. parapsilosis is very unique as it is solely NADH specific (Lee et al., 2003). Like many AKR enzymes, XR shows broad in vitro substrate specificity; aside from the canonical substrate they are also capable of reducing a variety of aldehydes. XRs from yeast are mostly dimers with
molecular weight of the subunits ranging from 33 to 40 kDa. Yeast XRs exhibit Michaelis–Menten kinetics with respect to both the substrate and cofactor. Like members of AKR families, XRs from yeast also show an ordered bi bi mechanism with the co-factor binding first to the enzyme followed by the substrate (Rizzi et al., 1988). Comparison of sequences with aldose reductase of human has revealed that, yeast XRs also possess signature sequences of aldo-keto reductase family (Lee et al., 1998). In human aldose reductase few residues e.g. Asp43, Tyr48, Lys77, His110 and Lys262 have been found to be important for catalysis and co-factor binding. These residues are also conserved in all yeast XRs. Despite these similarities, yeast XRs are quite distinct from human aldose reductase on following aspects. Majority of mammalian aldose reductases are monomeric while the yeast XRs are mostly dimeric and rarely monomeric. They also differ in their co-factor requirements. Yeast XRs are versatile – capable of using either NADPH or NADH or both as cofactor while the mammalian aldose reductases are strictly NADPH dependent.

Several members of the AKR superfamily have been structurally characterized, which includes, human and porcine aldehyde reductase (El-Kabbani and Flynn, 1995), Chinese hamster ovary reductase (Ye et al., 2000), rat 3-hydroxysteroid/dihydriodiol dehydrogenase (Hoog et al., 1994), Corynebacterium 2,5-diketo-D-gluconic acid reductase A (Khurana et al., 1998, Sanli et al., 2001) etc. Among yeast xylose reductases (AKR2), the crystal structure of apo and holo form of the enzyme from C. tenuis (AKR2B5) has been determined up to 2.2 Å (Kavanagh et al., 2002). In addition to the native protein, structures of few CtXR mutants (N309D, K274R, H114A) are also been solved. These structural studies had helped to elucidate the details of the mechanism of the enzyme activity, co-factor and substrate binding.

From the structural analysis of these different members of the superfamily, it can be said that all AKRs share a common (α/β)8 -barrel three-dimensional structure and the outer region of the structure is lined by eight α-helices that run anti-parallel to the central β-sheets. An eight-stranded parallel β-barrel sheet comprises the central region of the structure and includes the site of catalysis. The catalytic and cofactor-binding residues are highly conserved among the AKRs whereas the residues defining the substrate-binding site are less conserved. All of the AKRs are dependent on a nicotinamide adenine dinucleotide cofactor for redox chemistry, and they retain similar details within the cofactor-binding site. The NAD(P)H cofactor binds in an extended
conformation within a crevice that extends from the outer edge of the protein to the inner central barrel and it does not involve a Rossmann-fold motif. Kinetic analyses have revealed that AKRs follow an ordered bi–bi reaction mechanism in which NAD(P)H cofactor binds first and leaves at the end (Kubiseski et al., 1992). The release of cofactor is the rate-limiting step in the reaction cycle (Grimshaw et al., 1990). Crystallographic studies of certain AKRs have demonstrated that a conformational change appears necessary for NAD(P)H cofactor to gain entrance into, and to exit out of, the active site (Rondeau et al., 1992, Wilson, 1992, Borhani et al., 1992, Sanli et al., 2001). A structural comparison of the apo- and holo forms of AKR2B5 suggests that the NADPH cofactor is a key structural component in re-organizing both the active-site residues and the substrate-binding pocket to produce a catalytically competent form of the enzyme (Kavanagh et al., 2002).

The arrangement of the catalytic residues is similar in most aldo-keto reductases, indicating that the catalytic mechanism is conserved within the AKR superfamily (Khurana et al., 1998; El-Kabbani et al., 1998; Wilson et al., 1992; Jez et al., 1997). The overall catalysis constitutes a two-step process in which NAD(P)H first binds to the apo-enzyme, and is followed by binding of the substrate to the holo-enzyme. The 4-pro-R hydrogen of the cofactor nicotinamide ring C4 is transferred as a hydride ion directly to the re face of a carbonyl carbon of the substrate. Concomitantly a proton is transferred to the carbonyl oxygen of the substrate to complete the reduction reaction. The reduced product is released, followed by the release of oxidized NAD(P) cofactor (Harrison et al., 1994).

The crystal structure of CtXR shows, the active site is composed of side chains from Tyr52, Asp47, and Lys81 that form a catalytic triad and the tyrosine function as a general acid (Kavanagh et al., 2002). The Lys-81 makes a contact with Tyr 52 and thus lowers down the pKa of Tyr52 to facilitate proton transfer. Lys-81 in turn engages in a salt link with the Asp47 side chain carboxylate. His114 is believed to serve in orienting the substrate carbonyl, a function analogous to His110 in human aldose reductase (Bohren et al., 1994). The co-factor binding pocket is conserved and NADPH binds in the anti conformation like other AKR members. By comparing apo- and holo structures, loop7 is found to become ordered upon cofactor binding.

In absence of crystal structure with substrate, modelling experiments were performed to gain understanding the aldehyde substrate binding. This revealed that, C4
and the carbonyl oxygen hydrogen-bonded to the phenolic oxygen of Tyr52 and Ne2 of His114. In this orientation, the aldehyde proton points toward Trp24, and the re-face of the carbonyl is positioned toward and parallel with the nicotinamide. The side chains of residues Asp51 and Asn310 are subsequently available to bind the xylose C2 hydroxyl. Xylose C2 hydroxyl forms a hydrogen bond to Nδ of Asn310 which suggests that the C2 hydroxyl acts primarily as an acceptor of hydrogen bonds.

The substrate binding site of CtXR has also been probed by modelling and site directed mutagenesis approach (Kratzer et al., 2006). Trp23, Asp50 and Asn309 have been shown to be the main components of pentose specific substrate binding and recognition. It has been further suggested that Trp23 has a conserved role in determining aldehyde-versus-ketone substrate selectivity. Replacement of Asn309 with alanine or aspartic acid disrupts the hydrogen bonding with the substrate at C-2(R) hydroxyl group which is thought to be a prime contact for the screening aldose substrates and crucial to transition state binding. Like other AKR members CtXR also shows high substrate promiscuity and it can reduce different aldehydes (R-CHO) in vitro where ‘R’ may diverge from an uncharged polyhydroxylated side chain to hydrophobic aromatic or aliphatic side chains. The CtXR crystal structure also shows a substrate binding pocket that is lined by hydrophobic residues and very few polar residues are available to make hydrogen bonding partners.

Crystal structure of wild type CtXR bound to NADP and NAD has showed that, Glu227 confers NADH specificity by making interaction with 2'-OH and 3'-OH groups of the nicotinamide, when the 2' phosphate group is lacking in the co-factor. Thus, Glu227 is conserved among dual specific AKRs. On the other hand, side chain of Lys274 forms a strong hydrogen bond with the 2' phosphate group and confers NADP specificity. So, Lys274 is conserved among all the members of AKR2 family except in NADH specific C. parapsilosis where it shows Lys to Arg replacement (Leitgeb et al., 2005).

1.3.2 Xylitol Dehydrogenase (XDH) enzyme

In the xylose catabolic pathway, oxidation of xylitol to xylulose is carried out by NAD⁺ linked xylitol dehydrogenase enzyme which belongs to medium chain alcohol dehydrogenases (MDR) super-family. This superfamily has both Zn containing and non-Zn containing families. Xylitol dehydrogenases come under the Zn containing
polyol dehydrogenase family (PDH) (Nordling et al., 2002). They have the amino acid signature motif \([GHE]xx[G]xxxx[G]xx[V]\) and a typical alcohol dehydrogenase ‘Rossmann fold’ pattern composed of NAD\(^+\) and zinc ion binding domains. The PDH family is composed of 12 subfamilies. Their characterized members contain zinc, show dehydrogenase or reductase activities, bind to NAD(H), and are generally present in cytosol (Riveros-Rosas et al., 2003). Cys, His and Glu that co-ordinate catalytic Zn atom are conserved in most PDH (Nordling et al., 2002). The mammalian medium-chain alcohol dehydrogenases are dimeric, whereas several fungal and bacterial enzymes have been found to be tetramers.

General structure of a medium chain alcohol dehydrogenase was studied by Eklund and Ramaswamy (2008) from human, horse liver and bacterial MDR and a coherent picture was presented by them. The structure of the alcohol dehydrogenases can be distinctly divided into two domains, co-enzyme binding domain and catalytic domain which are separated by a long deep cleft. The co-enzyme binding domain is located at the N-terminal domain while the C-terminal domain contains the catalytic domain. The catalytic domain has a complicated arrangement with three \(\beta\)-structural motifs and a few helices.

Alcohol dehydrogenases (ADHs) of mammalian origin belong to MDR family and they normally contain two zinc atoms per subunit, one of these is a structural atom liganded by four cysteine residues (Drum et al., 1967). The structural zinc atom stabilizes a long loop excursion from the sheet structure. The other zinc atom is present at the active site and is essential for catalysis. It is coordinated either by two cysteines and one histidine or by one cysteine, glutamic acid and a histidine residue. Active site zinc coordinated by Glu has been observed generally in tetrameric ADHs, while it has not been observed in most dimeric ADH structures. The fourth position coordinates the substrate or a water molecule. The yeast XDHs, characterised so far only has catalytic Zn ion and no structural Zn was reported.

The cofactor binding has been studied from the crystal structure of horse liver dehydrogenase and NAD(H) binding was observed in an extended conformation across the C-terminal edge of the parallel \(\beta\)-sheet in a way that the adenosine half is positioned in a cleft at the surface of the domain and the nicotinamide half pushes itself deep in the protein at the active site cleft. The adenosine ribose is bound with double hydrogen bonds to an aspartate, constituting a common feature of dehydrogenases with
preference for NAD⁺ over NADP⁺. With NADP⁺, the aspartate would create repulsion of the extra phosphate group due to charge and space, thus explaining the coenzyme preference of the enzyme (Eklund et al., 1984). Several structures of horse liver ADH have been determined to a resolution better than 2Å. The best resolution so far is for the complex with NADH, DMSO and two inhibitors: dimethyl sulfoxide and isobutyramide which was determined at 1.0 Å resolution (Meijers et al., 2007). The nicotinamide ring in this complex is puckered in a twisted boat conformation. A puckered conformation of reduced nicotinamide ring of NADH has also been observed in another high resolution structure (Venkataramaiah and Plapp., 2003). The puckering of the nicotinamide ring was predicted to be important for hydride transfer.

MDR family of enzymes undergo large conformational changes upon binding of the coenzyme and this was first demonstrated with horse liver alcohol dehydrogenase (Branden et al., 1977). Since crystal structure of yeast alcohol dehydrogenase is not available, the catalytic mechanism of them was postulated mainly from the studies on horse liver alcohol dehydrogenase. Upon binding of the coenzyme, the protein becomes closed around the nicotinamide ring at the active site and one surface of the nicotinamide ring makes contact with the zinc coordinated cysteine sulphur atom. The oxygen atom of the substrate binds directly to the zinc atom and the rest of the substrate becomes close to the nicotinamide of the coenzyme (Eklund et al., 1983). The catalytic mechanism of these enzymes is quite well studied. The net removal of two hydrogen atoms from the substrate proceeds via proton loss from the substrate to form Zn-bound alkoxide ion and final hydride transfer occurs from alkoxide to NAD⁺ to form NADH. Role of a proton relay system involving few residues at the active site has also been proposed in this hydride transfer (Aggarwal et al., 2000). However, an interesting variation of this well accepted mechanism has been suggested recently for glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei* which is also a MDR family enzyme (Baker et al., 2009). This involves a dynamic coordination of active site Zn atom with the substrate, bulk water, and few conserved amino acid residues. The water molecule coordinated to Zn atom (a strong Lewis acid) possibly exists as hydroxide ion and acts as a base to trap the proton released from the substrate during the formation of alkoxide. Amino acid residues equivalent to Glu155 appeared to play a vital role in coordinating the Zn atom at the active site as well as in the catalysis.
Chapter 1: Introduction and Review of literature

The gene encoding XDH has been cloned as well as the enzyme was characterized in few yeast and fungal species e.g. *P. stipitis*, *Aspergillus oryzae*, *A. adeninivorans* etc. All of them are mesophilic enzymes and show optimum activity at 25-35°C. Mutants of *C. tropicalis* having deletion in XDH gene showed impaired growth on medium containing D-xylose as a sole carbon source (Ko et al., 2006).

1.4 Engineering yeasts for Xylose Metabolism

Although yeast and bacteria are capable of fermenting xylose, the former is preferred in commercial fermentation due to some of its advantageous characteristics like: larger sizes, thicker cell walls, low pH tolerance, less stringent nutritional requirements, and greater resistance to contamination. Ethanolic fermentation of xylose is either absent or extremely slow in most yeast species those are capable of xylose utilisation, (Toivola et al., 1984). Only a few yeast species, namely strains of *P. tannophilus*, *P. stipitis*, *C. shehatae*, *Pichia segobiensis* and some *C. tenuis* strains are able to ferment xylose slowly under anaerobic conditions (Bruinenberg et al., 1984; Toivola et al., 1984), but none of these yeasts is able to grow under anaerobic conditions on either xylose or glucose (Visser et al., 1990). Among yeasts, *S. cerevisiae* is the most thoroughly investigated and exploited in the past for food and alcoholic beverage production and it is considered to be the premium fermentative yeast. *S. cerevisiae* is not able to metabolise xylose, although some strains that are able to co-utilise it along other substrates or obtained ability to grow on xylose at extremely slow rate under aerobic conditions have been reported (van Zyl et al. 1989; Attfield and Bell et al., 2006). Thus, several attempts have been made to engineer *S. cerevisiae* with the introduction of XR and XDH genes from other xylose metabolizing yeast species like *P. stipitis* (Jeppsson et al., 2005; Watanabe et al., 2007). Moreover, to improve the xylose uptake in these strains, xylose transporter genes from other species have been used (Saloheimo et al., 2007; Leandro et al., 2008; Runquist et al., 2009). However, in spite of these efforts, fermentation of xylose remained significantly less efficient than that of glucose by these recombinant strains due to a major bottle neck of the oxido-reductase pathway. A redox imbalance is found to exist due to the difference in co-factor specificities of the XR and XDH enzymes. Generally the catabolic reactions in cells are associated with NAD+ /NADH while the anabolic reactions are associated with NADP+ /NADPH.
system. However in xylose metabolism both cofactor pairs are needed for catabolic reaction, as XR and XDH shows different co-factor specificity. The XR is chiefly NADPH dependent while XDH uses NAD\(^+\). Thus the NADP\(^+\) produced by the first enzymatic reaction, is not utilized as a co-factor by the second enzyme, XDH, which rather uses NAD\(^+\) for reaction. In the second enzymatic step, the NAD\(^+\) is reduced to NADH that can be recycled only through oxidative respiration. Thus in micro-aerobic fermentation condition, there is a buildup of both xylitol and NADH due to imbalance of NAD\(^+\)/NADH ratio in the cell. However, another reason for co-factor imbalance was described to be varying expression of XR and XDH enzymes in cells. Level of expression of enzymes varies from organism to organism. If XR expression is higher than XDH, then it leads to overproduction of xylitol byproduct, accumulation of NADH and inefficient production of ethanol. In order to alleviate cofactor imbalance, protein as well as metabolic engineering have been attempted. Most of the strain improvement approach for ethanol production from xylose focuses on improving sugar uptake and initial assimilation steps.

1.4.1 Improvement of xylose uptake with transporters
Enhancing the xylose uptake ability by \textit{S. cerevisiae} is an important factor for efficient ethanol production. \textit{S. cerevisiae} takes up xylose very poorly through hexose transporters (HXT family members, HXT 1 to 7). They have low affinity (10-100 times less) for xylose than glucose (Hamacher et al., 2002). The hexose transporters are not well regulated or specific enough to facilitate uptake of pentose sugars, moreover these hexose transporters show glucose repression and can only take up xylose in absence of glucose (Sedlak and Ho, 2004). Katahira et al. (2008) has attempted to express Sut1 hexose transporter of \textit{P. stipitis} (which also transports other mono-saccharides including xylose) into \textit{S. cerevisiae} and observed enhancement in xylose fermentation to ethanol. Glucose/xylose facilitator (Gxfl) as well as glucose/xylose proton symporter (Gxs1) from \textit{Candida intermedia} have been cloned and characterised recently (Leandro et al., 2006). Expression of this Gxfl in \textit{S. cerevisiae} increased the efficiency of the xylose uptake in the recipient strain (Runquist et al., 2009). However, when Gxs1 was expressed together with Gxfl or other hexose facilitators, symport activity was perturbed (Leandro et al., 2008). This might be due to controlled expression of energy-driven transporters in yeast. The xylose transporter Trxl1 of
Chapter 1: Introduction and Review of literature

filamentous fungus *Trichoderma reesai* has also been introduced into *S. cerevisiae* with the effect of elevated xylose uptake rate (Saloheimo et al., 2007).

1.4.2 *Engineering of xylose metabolizing proteins*

Anaerobic alcoholic fermentation of xylose without building up of byproducts is possible only under the conditions of identical co-factor specificity of XR and XDH. A few protein engineering approaches have been taken to change the co-factor specificity of these enzymes.

1.4.2.1 *Engineering of Xylose reductase*

Yeast XR contains a conservative NADPH-binding motif (Lys270–Ser271–Asn272–Thr273) and site-specific mutagenesis revealed that, lysine-270 in *P. stipitis* participates directly in binding xylose and NADPH. Thus, there have been several attempts to mutate this residue to change the co-factor specificity: (1) in the mutant K270M, there was a substantial reduction in the affinity to NADPH and increase in affinity with NADH were observed. In addition to this, $K_m$ towards xylose was also increased (Kostrzynska et al, 1998) (2) K270R mutant possesses reduced affinity to NADPH, though the affinity to NADH remains unchanged in comparison to wild-type XR. (3) In K270G mutant, the activity was substantially reduced with both the cofactor NADH/NADPH. In addition to lysine other residues of the cofactor binding motif were also mutated in *P. stipitis*. S271A and N272D mutants showed higher affinity to NADPH than to NADH. On the other hand, in T273S mutant overall activity of the enzyme was reduced, by analogy with the K270G (Watanabe et al., 2007).

Effect of lysine mutations were studied in *S. cerevisiae* as some of the mutants showed decreased affinity for NADPH. Jeppssion et al., (2005) constructed a recombinant strain of *S. cerevisiae* that contained K270M mutant together with XDH from *P. stipitis* and over-expressed endogenous XK. This strain showed an increased yield of ethanol from xylose and reduced quantity of xylitol. Another mutant K270R that showed reduced NADPH preferring activity was also transformed in *S. cerevisiae* that express the native XDH from *P. stipitis*. This resultant strain exhibited a reduction in xylitol build-up in comparison to the strain containing the native XR of *P. stipitis*, (Watanabe et al., 2007).
The importance of the Lys and Asn residues in the NADPH binding motif was further reinforced by mutating them in *H. polymorpha* orthologue (Dmytruk et al., 2008). By site-specific mutagenesis, the Lys341 was replaced by arginine and Asn343 by aspartic acid. It was shown that the affinity of the modified XR to NADH is higher than that of the native XR. The level of fermentation of xylose to ethanol in strains that contained the modified XR was also elevated. The XR from *C. parapsilosis* has been observed to possess an affinity to NADH that is 100 times greater than its affinity to NADPH. The gene *XYLI* from *C. parapsilosis* was cloned and expressed in *C. tropicalis*. The recombinant strains obtained were characterized by an elevated level of production of ethanol from xylose (Lee et al., 2003).

**1.4.2.2 Engineering of xylitol dehydrogenase**

Similar to XR, attempts have been made to shift the cofactor specificity of XDH from NAD\(^+\) to NADP\(^+\). The coenzyme binding domain of XDH possess a \(\beta\)-\(\alpha\)-\(\beta\) motif centred around a highly conserved Gly-\(X\)-Gly-\(X\)-\(X\)-Gly sequence (where \(X\) is any amino acid). The ability of dehydrogenases to discriminate between NAD\(^+\) and NADP lies in the amino acid sequence of this \(\beta\)-\(\alpha\)-\(\beta\) motif. The primary determinant of NAD\(^+\) specificity is the presence of an aspartate residue, which forms double hydrogen bonds to both the 2'- and 3'-hydroxyl groups in the ribosyl moiety of NAD\(^+\) and induces negative electrostatic potential to the binding site (Baker et al., 1992). Commonly, this residue in NADP\(^+\) -dependent dehydrogenases is replaced by a smaller and uncharged residue such as Gly, Ala, and Ser, accompanied by the presence of an arginine residue that forms a positive binding pocket for the 2'-phosphate group of NADP.

Metzer and Hollenberg (1995) had engineered NAD\(^+\) binding site of PsXDH by site directed mutagenesis. Mutation of Asp207 alone or in combination with Asp210 to glycine although resulted in nine fold increments in the \(K_m\) for NAD\(^+\) but a substantial reduction in the xylitol dehydrogenase activity was observed in these mutants. In a separate study, NADP\(^+\) recognition sequence (GSRPVC) of an ADH from thermotolerant bacteria was also introduced in PsXDH by replacing analogous sequence (Metzer and Hollenberg., 1995). The resultant mutant enzyme was able to use both NAD\(^+\) and NADP\(^+\) as cofactor with equal apparent \(K_m\) values.

In a recent study, an extensive mutagenesis of the NAD\(^+\) binding site of PsXDH was carried out to reverse the coenzyme specificity toward NADP\(^+\) (Watanabe et al.,
2005). A single amino acid change e.g. D207A, I208R, F209S, N211R yielded a mutant with 5–48 fold increase activity in the presence of NADP. These mutants, however still retained a sufficiently high activity with NAD\(^+\). Interestingly, a triple mutant (D207A/I208R/F209S) or quadruple mutant (D207A/I208R/F209S/N211R) showed ~4500-fold increase in the catalytic activity with NADP. These mutated PsXDH were also used to construct better strains of \textit{S. cerevisiae} for xylose fermentation (Watanabe et al., 2007). The thermostability of PsXDH mutants was also engineered favourably by introducing additional structural zinc atoms (Watanabe et al., 2005; Annaluru et al., 2007).

1.4.3 Approaches for rational engineering of metabolic pathways

In addition to these protein engineering efforts, the xylose metabolic pathway has been modified rationally to achieve higher yield of ethanol from xylose fermentation. Approaches like, transcription profiling, proteomics and metabolite profiling allow the identification of global cellular effects of the genetic modifications at the level of gene expression, proteins, metabolites and metabolic fluxes. This information can be utilised to identify new targets for genetic manipulation and redesign of metabolic pathways for an improved phenotype.

Metabolic modeling is one of the systematic approaches that can guide metabolic engineering. Quantitative metabolic models can be built by calculating stoichiometric balances or data inferred from the labelling of intermediate metabolites. For example, Grotkjaer et al., (2005) experienced inefficient fermentation of recombinant \textit{S. cerevisiae} from xylose to ethanol due to the imbalance of redox cofactors. Metabolic flux models guided them to delete the gene encoding NADP\(^+\) dependent glutamate dehydrogenase (GDH1) and to over-express NAD\(^+\) dependent GDH2 to shift reductant demand. This increased the specific activity of XR with NADH, and boosted the ethanol production by 25%.

In addition to modelling, the flux analysis of a metabolic pathway could be done to calculate the partitioning of metabolites among the various pathways using stoichiometric metabolic matrices based on known reactions. The metabolic flux is the rate of turnover of metabolites through a metabolic pathway or an enzyme and fluxes reflect closely the metabolic state of the cell. Thus, flux analysis can accurately predict the amount of product formation. Flux analysis through PP pathway during D-xylose
metabolism has revealed that, limiting speed of this step may be beneficial to the efficiency of fermentation to ethanol. To reduce xylitol formation during xylose fermentation, Jeppsson et al., (2002) blocked PP pathway in the recombinant *S. cerevisiae* strain through the deletion of NADPH producing GND1 (6-phosphogluconate dehydrogenase) and ZWF1 genes (glucose -6-phosphate dehydrogenase) individually. Decreasing the phosphoglucone isomerase activity by 90% also lowered the PP pathway flux. All these modifications resulted in lower xylitol accumulation and higher ethanol yield than in the control strains. They also showed a reduced rate of xylose consumption. The low xylose fermentation rate is probably due to limited NADPH-mediated xylose reduction. Metabolic flux modelling of ZWF1 deleted strain confirmed that the NADPH-producing PP pathway was blocked and that reduction of xylose was mediated only by NADH, leading to a lower rate of xylose consumption. These results indicate that xylitol production is strongly connected to the flux through the oxidative part of the PP pathway (Jeppsson et al., 2002). Microarray based global gene expression analysis was also applied to compare strains of *S. cerevisiae* engineered for xylose fermentation. Such studies reveal that glucose-repression is an important metabolic bottleneck on efficient conversion of xylose to ethanol (Sonderegger et al., 2004).

Evolutionary engineering approach is another powerful technique that has been utilized for constructing *S. cerevisiae* strains for efficient xylose fermentation. *S. cerevisiae* strain expressing xylose isomerise gene from anaerobic rumen fungus *Piromyces sp.* showed slow xylose fermentation. However, after continuous cultivation of this in an anaerobic, xylose-limited chemostat, an improved strain was obtained (Kuyper et al., 2005).

### 1.5 Scope of the present study

So far the *S. cerevisiae* is the most exploited yeast that has been modified with engineered proteins and altered metabolic pathways. In spite of various modifications, the dilemma of low ethanol yield could not be overcome completely in every case. This circumstance led us to search and explore for more organisms and characterize enzymes from various sources. As the enzymes of the yeasts are essential components for xylose metabolism thus they are important tool for metabolic as well as protein engineering. In this thesis, the xylose metabolizing properties of two unique yeasts,
halotolerant *D. hansenii* and thermotolerant *Pichia angusta* have been explored through detail study of their enzymes.

### 1.5.1 Scope of studying xylose metabolism in *Debaryomyces hansenii*

*D. hansenii* is an osmotolerant, cryotolerant, halotolerant, marine yeast and it was originally isolated from hyper-saline environments such as seawater (Norkans, 1966) and concentrated brines (Onishi, 1963). It has become increasingly important during the last few years as a model of osmotolerance of eukaryotic micro-organisms, due to the diversion of its metabolism towards the production of glycerol. An important characteristic of this yeast is its ability to accumulate and tolerate significant amounts of Na\(^+\) when incubated or grown in the presence of high Na\(^+\) concentrations (Norkrans and Kylin, 1969; Prista et al., 1997; Thome et al., 1998; Sharma et al., 2005; Aggarwal et al., 2005). On the other hand, Alder et al (1980) studied that, *D. hansenii*, which is osmotolerant yeast, may be a very attractive microorganism for polyol production. According to their study, the intracellular levels of polyols in *D. hansenii* were markedly enhanced by high salinity, the dominant solutes being glycerol in log phase cells and arabinitol in stationary phase cells. Beside the halotolerance and osmotolerant properties, it is a xylose assimilating yeast which has seldom been mentioned. *D. hansenii* is capable of converting xylose in to xylitol (main product) and ethanol under semi-aerobic conditions (Amaral-Collaco et al., 1989). *D. hansenii* was found to be an important xylitol producer exhibiting a xylitol/ethanol ratio above four and a carbon conversion of 54% for xylitol (Roseiro et al, 1991). Thus, one of attractive biotechnological potential of *D. hansenii* is its use in the production of xylitol from D-xylose or wood hydrolysate (Girio et al., 1996, 2000; Parajo et al., 1996). The bioconversions yields reported are similar or higher than values reported for other typical xylitol producers such as, *P. guilliermondii*, *C. tropicalis*, *C. boidini* and *C. parapsilosis*. Thus, to explore efficient xylitol producing system, we need to characterise the xylose metabolic pathway of *D. hansenii*. In addition to that, studying the xylose metabolic pathway of *D. hansenii* will help us to understand a new system which could highlight the virtues and limitations of the components of this pathway in this yeast, for further applications. Another advantage of studying the *D. hansenii* enzyme system may be metabolic engineering in heterologous system with the xylose metabolising genes of *D. hansenii* to improve the bio-ethanol production. Identification
Chapter 1: Introduction and Review of literature

and molecular as well as biochemical characterization of the components of this pathway is therefore an important step towards to explore this system.

1.5.2 Scope of studying xylose metabolizing enzymes in Pichia angusta

The thermotolerant methylotrophic yeast *Pichia angusta* (syn. *Hansenula polymorpha*) was isolated by Levine and Cooney (1973), which was shown to grow optimally in mineral supplemented methanol (0.5%) media at 45°C. *P. angusta* is a thermotolerant, methylotrophic yeast with tremendous biotechnological potential (Kunze et al., 2009). The advantage of *P. angusta* as compared to other mesophilic yeasts is that, it is able to carry out fermentation of glucose, cellobiose and xylose to ethanol at higher temperature (45-48°C) (Ryabova et al., 2003). Except *Kluyveromyces marxianus* (Banat et al., 1996) this is the only thermotolerant yeast known so far that naturally ferments xylose to ethanol at higher temperature. This property is advantageous for developing simultaneous saccharification and fermentation process which combines enzymatic hydrolysis of pretreated lignocellulose by cellulases and hemicellulases with fermentation of the produced hexoses and pentoses (Ryabova et al., 2003; Dmytruk et al., 2008a). A number of recent studies have been conducted with this yeast, which are focused on fermentation of xylose at higher temperature (Dmytruk et al., 2008a and 2008b; Ishchuk et al., 2008). *P. angusta* is reported to possess single gene encoding XR and two genes encoding XDH and substantial reduction of total cellular XDH activity was observed by deleting these genes (Dmytruk et al., 2008b). These enzymes from this thermotolerant yeast have not been characterised so far and furthermore such study will also reveal their suitability as tools for undertaking metabolic engineering in heterologous system. Thus, in the present study, biochemical characterization of the enzymes involved in the xylose metabolism in two extremophilic yeasts, *D. hansenii* and *P. angusta* are undertaken with the following broad objectives:

(1) Molecular cloning and characterization of *xylose reductase (XR)* enzyme from *Debaryomyces hansenii*.
(2) Molecular cloning and characterization of *xylitol dehydrogenase (XDH)* enzyme from *D. hansenii*.
(3) Identification, molecular cloning and characterization of the *xylitol dehydrogenases* from *P. angusta*. 