Introduction

1. INTRODUCTION

Zymomonas mobilis, a Gram-negative bacterium is considered as promising alternative to yeast for industrial ethanol production. It was first isolated from the African palm wine, the Mexican *pulque*, the European cider and beer. It metabolizes sugars to ethanol and carbon dioxide through Entner-Doudoroff pathway. The advantages of *Z. mobilis* over *Saccharomyces cerevisiae* include amenability to genetic manipulations, higher sugar uptake (3-4 times), higher ethanol yield (93-96 % theoretical yield), higher ethanol tolerance (13% w/v), lower biomass production, non-requirement of controlled addition of oxygen during the fermentation. But, it has the ability to utilise only a few sugars such as glucose, fructose, and sucrose (Gunasekaran *et al.,* 1986).

The complete genome sequence of *Z. mobilis* ZM4 [NC_006526] was first reported by Seo *et al* (2005). The whole genome consists of chromosome of 2,056,416 bp with G+C content of 46.33%. The total number of ORFs is reported to be 1,998. Among them, 1,346 ORFs are annotated with their functions, 258 ORFs are predicted putative coding sequences for general functions and 394 ORFs are unknown hypothetical genes.

Z. mobilis exhibits high level of resistance to various ß-lactam antibiotics. There have been no detailed studies on the ß-lactam resistance-determinants in this bacterium. *Z. mobilis* shows resistance to an array of chemicals that have the capability to inhibit most of the bacteria (Bochner *et al.,* 2010). The available genome sequence of *Z. mobilis* ZM4 [NC_006526] provides an opportunity to understand the antibiotic resistance genes. The genome of *Z. mobilis* ZM4 (NC_006526) carries sequences encoding four Beta-lactamases (BLA), three Beta-lactamase Domain-

containing Proteins (BDP) and three Penicillin binding proteins (PBP) (Seo *et al.,* 2005). These sequences are present in three other sequenced genomes of *Z. mobilis* strains namely *Z. mobilis* NCIMB 11163, *Z. mobilis* subsp. *mobilis* lectotype ATCC 10988 and *Z. mobilis* subsp. *pomaceae* lectotype ATCC 29192. Even though the antibiotic resistance has been reported in *Z. mobilis*, the genetic and physiological basis of the ß-lactam resistance is poorly understood.

Z. mobilis has been known to change its morphology and formed as flocs during continuous culture (Fein *et al.,* 1983). The basis of floc formation in *Z. mobilis* is not known so far. There have been previous reports of floc-forming bacteria having synthesized extracellular cellulosic fibrils that help in flocculation and cell adhesion to surfaces (De Boks and Van Eybergen, 1981). There were no detailed studies and previous reports on the production of cellulose in *Z. mobilis*. The available genome sequence of *Z. mobilis* ZM4 provides us the opportunity to study the genes responsible for cellulose production in *Z. mobilis*. The genome of Z. *mobilis* ZM4 (NC_006526) also carry sequences encoding cellulose synthase catalytic subunit (BcsA: YP_162818), cellulose synthase regulator protein (BcsB: YP_162819), cellulose synthase operon C domain-containing protein (BcsC: YP_162820) and cellulase (CelA: YP_162821) (Seo *et al.,* 2005). These sequences are identical to three other genes in the sequenced genomes of *Z. mobilis* strains. Even though the floc formation has been reported in *Z. mobilis*, the role in cellulose production in floc formation has not been studied so far.

One of the major limiting factors for the usage of *Z. mobilis* in the industrial production of ethanol is the narrow utilizable substrate range and requirement of clean substrate with simple sugars. Moreover, during pre-genomic period there are no reports on the genes in *Z. mobilis* for lactose metabolism. However, later the genome

of Z. *mobilis* ZM4 was found to carry sequence encoding beta-galactosidase (YP_162639) (Seo *et al.,* 2005) which is also present in three other sequenced genomes of *Z. mobilis* strains.

Various bioinformatics tools, softwares, servers have been developed to predict the structure and function of enzymes. These tools use an extensive variety of algorithms to predict the properties of proteins. The precision of the bioinformatics tools has been improving. The large quantity of data generated by genome sequencing projects has to be analysed. Bioinformatics approach is decisive in analysis of functional genomics data. Data clustering, primary component analysis, artificial neural networks (NN), support vector machines (SVM) are very useful tools for data analysis.

To study ß-lactam resistance, floc formation/cellulose production and lactose metabolism in *Zymomonas mobilis* through experimental and bioinformatics approach, the following objectives are proposed:

- Computational and functional analysis of beta-lactam resistance in *Zymomonas mobilis*
- Computational analysis of cellulose synthase operon and functional analysis of the cellulose production in *Zymomonas mobilis*
- Computational analysis and functional characterization of beta-galactosidase in *Zymomonas mobilis*