CHAPTER 1

INTRODUCTION AND OBJECTIVES
1.1. Genetic engineering and recombinant products

The developments in *invitro* genetic manipulations, commonly known as genetic engineering have extended the range of potential fermentation products. Genes from higher organisms have been introduced into microbial cells such that the recipients are capable of synthesizing 'foreign' (or heterologous) proteins. This genetic engineering or r-DNA technology constitutes a powerful tool of modern biologists to manipulate, accurately the genetic component of the cell, the DNA, in order to load the cell with the desired characteristics. A wide range of microbial, animal, insect and plant cells have been used as hosts for the expression of numerous kinds of genes whose products find use in food, chemical, textile, cosmetic, agricultural and pharmaceutical industries. Many human and mammalian genes have been expressed in micro-organisms, thereby enabling large scale production of human proteins, which are used therapeutically.

The practical application of genetic engineering is to make the cell synthesize certain products not synthesized by it, in its wild type state, or to overproduce some of its endogenous metabolites. In the former case, the gene coding for the desired heterologous protein is introduced into the cell while in the latter case, the genetic manipulation is done to produce the proteins which will facilitate the overproduction of the desired metabolite. This is done by inserting the gene of interest along with the control sequences for gene expression on to a vector molecule e.g. phage or a plasmid which is then introduced into a suitable recipient host cell by transformation.

The majority of the early work on recombinant DNA technology was concentrated on the transport of genetic material into *E. coli* and this led to the adoption of the organism as the host of choice for the production of several heterologous proteins. Availability of a wide range of vectors, simple fermentation processes using cheap media, promising protein yields in the range of 1-5 g/L and well documented literature on its genetics, have made it possible to develop *E. coli* as a host for the production of recombinant proteins for therapeutic use like insulin, hGH, factor VIII etc. Despite these advantages, *E. coli* also presents several problems for the production of heterologous proteins. Very often the proteins formed are intracellular insoluble aggregates and lack post-translational modifications (Primrose, 1986; Zabriskie and Arcuri, 1986). This led researchers to develop other hosts for recombinant protein production.
The development of genetic tools for *Bacillus subtilis* has made it an attractive host for recombinant protein production. However ideal strains for industrial production have yet to be developed because they secrete large amounts of proteolytic enzymes which degrade the desired product. Food-grade bacteria such as *Lactococcus lactis* (Wells et al., 1993; Steidler et al., 1995) and *Corynebacterium glutamicum* (Billman-Jacobe et al., 1994; 1995) have also come out as potential hosts for heterologous proteins synthesis. Depending on the type of recombinant protein required to be produced, mammalian and plant cells have also been used (Smiley et al., 1989).

1.2. Yeast as a system for production

The English *yeast* or German *geast* has been invaluable for human mankind since the dawn of this century. The more familiar term ‘fermentation’ is also derived from the latin verb *fervere*, to boil, based on the action of yeast on extracts of fruit or malted grain which gives a bubbling appearance due to the production of carbon dioxide. The production of alcohol by the action of yeast on malt or fruit extracts has been carried out on a large scale for many years and was the first industrial process for the production of a microbial metabolite. Industrial microbiologists have now extended the term fermentation to describe any process for the production of any product by the mass culture of a micro-organism. In addition to its classical use as biomass, a low-value, high-volume product in food industries, yeast is now being used for very high-value, low-volume products. To achieve this purpose, culture conditions have also been modernized from wooden boxes with the traditional thermometer, hydrometer and heat exchanger to sophisticated computer controlled bioreactors with facility for the on-line analysis of a range of fermentation parameters.

In 1980s researchers in industrial and academic laboratories began to explore yeast as an alternate host for expressing heterologous proteins when the nature of the recombinant proteins became too complex to be expressed in a simple bacterium like *E. coli*. Studies have showed yeast to be an excellent host for the expression of foreign proteins and especially mammalian polypeptides with potential clinical applications (Kingsman et al., 1987). Recombinant proteins produced successfully using yeast as hosts include human interferon (Hitzeman et al., 1981), human fertility hormone (Lemontt et al., 1986), human epidermal growth factor (Brake et al., 1984),
interleukin-2 (Lemoine et al., 1985), human α-antitrypsin (Cabezon et al., 1984),
antigenic determinants for hepatitis B (Valenzuela et al., 1985) and rabies (Lathe et al., 1985), etc. Thus yeast fermentation has covered a long distance from its original use in alcohol and baking industry in the pre 1900 era to its present day use in the production of therapeutic/pharmaceutical products.

Yeasts offer a number of advantages as a system for expression of complex proteins. Recombinant systems employing this micro-organism occupy a niche between prokaryotes and eukaryotes. Unlike prokaryotic systems, their eukaryotic subcellular organization enables them to carry out many of the post-translational folding, processing and modification events (glycosylation, disulfide bond formation etc.) required to produce ‘authentic’ and bioactive mammalian proteins. In addition, they retain the advantages of a unicellular micro-organism, with respect to rapid growth on cheap media, high cell density culture and ease of genetic manipulation. Yeasts can be grown in a wide range of cultivation conditions (pH, Temperature). A number of heterologous proteins produced in *Saccharomyces cerevisiae* have either obtained commercial licenses or are undergoing clinical trials.

It is of particular importance for production of heterologous proteins that protein secretion in yeast takes place via a multi-component secretory apparatus, within which disulfide bond formation, proteolytic maturation of prohormones N- and O-linked glycosylation and other post-translational modifications occur (Schekman and Novick, 1982). However, *Saccharomyces cerevisiae* has been found to have certain limitations as a host for heterologous protein expression. Product yields are usually low; except for a few notable exceptions; yields of heterologous proteins reach a maximum of 1-5% of total protein even with a strong promoter (Mellor et al., 1985). Also there are several reports on hyperglycosylation of secreted glycoproteins in *Saccharomyces cerevisiae* (Moir and Dumais, 1987) which often cause dramatic differences in immunogenicity, diminished activity and decreased serum retention of the recombinant protein. Many *Saccharomyces cerevisiae* secreted proteins are not found free in the culture medium, but are retained within the periplasmic space in a
cell-associated form (Smith et al., 1985; Zsebo et al., 1986; Jigami et al., 1986). This can lead to problems of purification as well as yields of purified product.

1.3. Role of plasmids in production of recombinant proteins

Plasmids are mostly circular double-stranded DNA molecules although recently some examples of linear plasmids have also been reported (Dani and Zakian, 1983). A plasmid is able to replicate independently and is not physically linked with the chromosome. It must, therefore, be transferred to daughter cells during cell division, and can be lost from the host cell due to either defective partitioning or unequal distribution among the daughter cells. Plasmids often carry genes, which are not part of the central metabolism of the cell such as antibiotic resistance, antibiotic production, sugar degradation or heavy metal resistance. They vary in size from below 1 kb to over 300 kb. They are relatively small when compared with the *Saccharomyces cerevisiae* genome, which is approximately 13,800 kb (Lauer et al., 1977).

With the development of genetic engineering, plasmids along with phages have become important cloning vectors in molecular biology. In the simplest case, a chimeric plasmid is constructed by carrying out restriction enzyme digestion of the plasmid and the DNA sequences to be inserted. These are ligated using DNA ligase. Many plasmids are currently available which carry the promoter and terminator sequences from the host, which are redesigned to meet the different requirements. Both promoter and terminator sequences are required for high level transcription and optimal RNA accumulation (Bitter et al., 1987). Using these expression cassettes it is possible, in theory, to obtain high expression levels from any piece of inserted DNA, which codes for a protein.

1.4. Overproduction of proteins in recombinant organisms and plasmid stability

Often the aim of genetic engineering is to construct a micro-organism that makes large quantities of a commercially important protein. In simple terms, an organism that makes ten times as much protein as another organism will cut costs of fermentation ten times, everything else being equal. The recombinant organisms
performance on a commercial scale is dependent on the proportion of cells that carry the recombinant plasmid (Imanaka and Aiba, 1981). It has been demonstrated that the rate of emergence of plasmid-free segregants increases with expression of cloned gene(s). This problem of plasmid instability, therefore, has to be addressed properly before going in for the scale-up of the recombinant product. Since the DNA sequences coding for the target protein are tagged on the plasmids, the loss of the plasmid from the microbial culture would mean the loss of the foreign gene sequence and hence reduced recombinant protein yield. The plasmid-free cells, which grow at the expense of cells containing the plasmid usually, have less metabolic burden associated with their growth since they do not have to synthesize the foreign protein. They therefore start overtaking the plasmid containing cells and thus divert the nutrient availability away from the recombinant protein synthesis thereby affecting the product yield.

It has been observed that the synthesis of a cloned gene product places an additional stress on the cell (Seo and Bailey, 1985) and the use of autonomously replicating expression plasmids has been found to result in poor plasmid stability during production runs. To minimize the deleterious effects, inducible promoters are employed and the time over which product synthesis occurs, is regulated to achieve a partial separation of growth and expression phases in the culture. However, even using an inducible promoter, plasmid stability has been found to drop significantly (Da Silva and Bailey, 1991) resulting in an expression system that may be difficult to scale up to production volumes. Instability is particularly severe when the product being expressed has some toxicity (Shuster, 1989; Hinnen et al., 1989).

Plasmid stability studies in yeast *Saccharomyces cerevisiae* are important because of high degree of segregational instability. Plasmid instability in yeasts is different from *E. coli* because yeast replicates by budding rather than by binary division as is the case with *E. coli*. In this vegetative reproduction, replication occurs by the production of a daughter cell as a bud on the mother cell. The bud grows and eventually separates from the mother cell at division. The buds usually arise on the shoulders and at the long axis of the cells. This type of budding, referred to as 'multilateral budding' is characteristic of Saccharomyces. Division in *Saccharomyces*


Saccharomyces cerevisiae is asymmetrical, the daughter cell being smaller at birth than the mother cell (Hartwell and Unger, 1977). Since there is size requirement before initiation of division in Saccharomyces cerevisiae (Johnston et al., 1977) the daughter cell must first grow to attain this size, whilst the mother cell is already at this critical size. Thus the cell cycle time for daughter cell is longer than the mother cell (Lord and Wheals, 1981). The degree of asymmetry at division, i.e., the size of the daughter cell at birth relative to the mother cell, is growth rate dependent (Lord and Wheals, 1980). The faster the cells are growing the lower the degree of asymmetry (Lord and Wheals, 1980; 1981; Thompson and Wheals, 1980).

In addition, the first bud is usually detached when the mother cell initiates a second bud. However, sometimes buds remain attached so that the mother cell and first daughter cell may produce additional buds. In such instances, this results into large clusters or chain of cells. Therefore this causes a non-uniform distribution of plasmids among the daughter cells, in contrast to E. coli where binary division ensures an orderly distribution of plasmids among the daughter cells.

Another significant dissimilarity between yeasts and E. coli is the difference in the rate of overtaking of plasmid containing cells by the plasmid free cells, which depends on the ratio of their specific growth rates. Since the recombinant protein forms a very small fraction of the total protein in yeasts, this does not put a serious metabolic burden on the plasmid containing cells. Thus the difference between plasmid-bearing and plasmid-free cells is negligible (Zhang et al., 1997a). Therefore, it is the emergence of plasmid free cells, and not overtaking in the culture medium, which is critical for the stability of recombinant yeast.

Thus the problem of plasmid shedding in recombinant yeast cultures is more severe. Although a lot of information is available on yeast bioprocess technology, the problem of plasmid instability is not fully solved.

1.5. Objectives

While a lot of work has been done to modify and improve Saccharomyces cerevisiae as an expression system at the genetic level, biochemical engineers are also trying to
improve the productivity by optimizing different bioprocess parameters. Bioprocess optimization with respect to a recombinant cell refers to optimizing a metabolite or foreign protein yield as well as the economy of the process. These strategies have been used systematically to optimize protein expression in *E. coli*.

The present study was designed to address some aspects of plasmid stability in *Saccharomyces cerevisiae*.

**Approach to the work: Use of genetic reporters**

The role of genetic reporters in rapid quantitative evaluation of physiological events without extraordinary effort or cost is well known. Three reporter genes, coding for firefly luciferase, β-galactosidase and chloramphenicol acetyltransferase (CAT) are commonly used because their corresponding enzymes are easy to assay and are either absent from non-transformed cells or present only at low levels.

In the present work, the recombinant yeast carrying genes for β-galactosidase and luciferase, *lacZ* and *luc* respectively, have been used as model systems for expression and plasmid stability studies.

Thus our objectives were —

- To clone the genes for β-galactosidase and firefly luciferase in a suitable yeast expression vector and transforming an appropriate *Saccharomyces cerevisiae* host by these recombinant plasmids.

- To study the expression and plasmid stability levels in selective as well as non-selective media in batch and continuous cultures.

- To determine the kinetic parameters which affect the plasmid stability in non-selective media.

- To test in co-culture the competitiveness of the recombinant cell in different compositions of the non-selective media.

- To design bioprocess strategies in order to obtain high stability and expression of cloned gene(s).

- To study the differential stability of recombinant plasmids, which contained genes of different sizes (*lacZ* or *luc*) from different sources (prokaryote and eukaryote) in yeast.
To develop mathematical models to predict cell growth, product formation and plasmid stability in different cultivation techniques and compare these model simulations with the experimental data.

These studies would help to provide guidelines in finding solutions to plasmid stability and expression problem in other recombinant yeast fermentations involving heterologous proteins similar in nature and size.

1.6. Significance of present work

The recombinant protein optimization and plasmid stability studies in yeast Saccharomyces cerevisiae at the bioprocess level are significant in the light of the successful completion of Saccharomyces cerevisiae Genome Sequencing Project (Goffeau et al., 1996). The genome sequence of the yeast Saccharomyces cerevisiae has provided the first complete inventory of the working part of an eukaryotic cell. Now the agencies/network like EUROFAN (European functional analysis network) are entrusted with the responsibility to elucidate the function of each and every one of the almost 4000 novel protein encoding genes discovered in the yeast genome. It is expected that these studies will lead to the isolation of numerous interesting genes whose products would be found useful for humans. Once these genes are characterized for any kind of important biological activity, the genetic and biochemical engineers will attempt to overproduce these proteins using r-DNA technology. Although the level of expression is not very high in yeast, the requirement of post-translational modifications for complex proteins, secretion, non-toxicity, codon-bias etc. will force researchers to use yeast based expression systems.

Therefore, the present work will assist researchers in understanding and solving the problem of plasmid instability and thereby improve the level of expression in other recombinant yeast fermentations.