Preface

Advances made in the fields of molecular biology and genetic engineering have revolutionized our understanding of life processes in genomic details. These developments have opened up new routes not only to achieve excellence in the efforts to improve the quality of life and environment, but also to develop more effective strategies to gain understanding of the fundamental aspects of molecular biology as regulation of gene expression, RNA processing, to name just a few.

Conventional plant breeding provided the only major route for crop improvement in terms of incorporating various desirable traits. Long, arduous and cumbersome as this route is, the advent of molecular approaches opened up new vistas for incorporation of gene(s) directly and rapidly. The Molecular Biology Laboratory, JNU, New Delhi, where the present work was carried out had earlier embarked upon a project aimed at improving the quality of crop plants in terms of their amino acid composition. This work culminated in the cloning of a seed specific protein gene AmAJ, from *Amaranthus hypochondriacus*, which encodes an albumin with a well balanced amino acid composition and high levels of the essential amino acids such as lysine, tryptophan and sulphur containing amino acids corresponding well with the World Health Organization standards for optimal human nutrition (Raina and Datta, 1992). Added to this unique feature was another novelty that the protein was nonallergenic in nature. Further, our laboratory has recently reported successful introduction and expression the *AmAJ* coding sequence in potato tubers in an attempt to realize the objective of developing economically important crop plants with improved nutritional quality. The study demonstrated the feasibility of using the *AmAJ* gene as a donor in genetic engineering programs to improve the nutritive value of other non-seed and grain crops (Chakraborty *et al.*, 2000).

Synchronously, an effort was initiated to improve the nutritional quality of certain non-conventional food supplements such as yeast using *AmAJ* gene as a donor. Earlier trials in this direction failed to express the gene in *Saccharomyces cerevisiae* for reasons not understood clearly. At this juncture, it was prudent to think about the feasibility of replacing
the host micro-organism with another one having proven ability for heterologous gene expression and that was *Schizosaccharomyces pombe*. In recent years it has been possible to produce foreign proteins, especially those derived from higher eukaryotes, at a high level using *S. pombe* expression systems. High-level production, easy manipulability, and low cost, among other factors, make these expression systems attractive not only for fundamental research but also for industrial uses. *S. pombe* serves as a model system for the expression of heterologous proteins from higher eukaryotes because of its resemblance with various higher eukaryotic intracellular events and also for reasons of phylogeny (Moreno *et al.*, 1991; Giga-Hama, 1997a). The prospect of investigating how a plant seed albumin gene encoding a high pool of essential amino acids would confront a heterologous environment appeared highly exciting. A drive for over-expression of the protein was thought to improve the strains from the nutritional standpoint and it may eventually bring forth another promising animal feed supplement. The over-expressed protein in its purified form further pledges to be a source of nutrition for human consumption. The work carried out in this direction have been presented in the chapter-1 of the thesis.

Studies conducted over the years on *S. pombe* revealed it as one of the best experimental models for cell cycle related studies, besides for studies related to chromosome structure, signal transduction, sexual differentiation and others (Russell, 1989). Further, *S. pombe* has successfully been used as a model system for the expression of heterologous proteins from higher eukaryotes because it resembles higher eukaryotes in various aspects, and quite notable one of them is, post translational modification. A pertinent question was raised whether parallels can be drawn between the co- and post-transcriptional events of *S. pombe* and those of higher eukaryotes and if so, the former can be used as an alternative system for these studies as well. Studies on different co- and post-transcriptional events, like, splicing, transcription termination, RNA 3' end maturation and 5' capping, have gained momentum in yeast as well as mammalian system besides a few other hosts. These studies have revealed many of their mechanistic details. However, little is known about these events in plants. In an attempt to study the feasibility of using of *S. pombe* as a model system for plant pre-mRNA splicing studies, an intron containing gene was introduced in *S. pombe*. The study was motivated by the finding of Kaufer *et al.*, 1985, that *S. pombe* correctly excises a mammalian
RNA transcript intervening sequence with similar cis-elements as required in the heterologous host. Earlier, our laboratory had reported isolation of the genomic clone of AmAJ, the coding sequence of which was interrupted by an intron of size ~1.5 kb (Biswas, 1997). Moreover, interestingly enough, various S. pombe cis-elements required for splicing had their matching counterparts in the gene. Hence, the study was conducted using AmAJ genomic DNA as a model gene. However, the cells expressing the genomic DNA were growing unusually slowly in comparison to the cells expressing the cDNA clone. The occurrence of a premature termination event in the intronic region resulting in a short but stable transcript gave a new direction to the study. Now the focus of interest was transiently shifted to find out why the cells expressing AmAJ genomic DNA had a retarded growth and to identify the factor responsible for it. Findings towards this end have been documented in the chapter-2 of the thesis.

The original line of investigation was then again picked up. It was of much interest to find out how the event of transcription termination as such had taken place. As background information, certain mechanistic aspects of transcription termination in S. pombe for three of its genes were available though there was no universal consensus among them (Hansen et al., 1998; Aranda and Proudfoot, 1999). Subsequently, different elements present in this heterologous piece of DNA that were recognised in S. pombe and which brought forth transcription termination and 3' end maturation, were dissected and their relevance was studied. These findings are embodied in the chapter-3 of the thesis.

Studies on transcription termination event revealed occurrence of read-through transcripts spanning the whole gene. A question was raised whether there was any splicing event taking place on the residual read-through transcripts that escaped 3' end maturation at an earlier site. Distinct evidences emanating from the study is presented in the chapter-4 of the thesis. This study convincingly showed that plant intronic sequences were, indeed, recognised in S. pombe, though with certain variability. These minor aberrations can be attributed to an acceptable degree of flexibility of the splicing mechanism that is conserved throughout despite divergence in life forms. Further, the chapter discusses preliminary success scored in the development of a cell free in vitro splicing extract from S. pombe cells. It was of interest
for quite long because the development of an in vitro splicing system would accelerate splicing studies in *S. pombe*.

Findings in these chapters impart extra significance to *AmA1* genomic DNA and its behaviour in *S. pombe* in terms of a natural occurrence of both the very important co-transcriptional events, i.e., transcription termination and splicing. Subsequently it was observed that the spatial arrangements of the cis-elements required for both the important processes in the gene warrant a competition between two otherwise coupled processes, making them mutually exclusive. Chapter-3 and 4 of the thesis present a discussion on these two coupled processes in the light of the findings and explores a number of questions, like, why transcription termination event took precedence over splicing in *S. pombe* and why the event of premature transcription does not occur in plants. Chapter-5 summarizes the whole work embodied in the thesis. General experimental protocols used in the studies have been described in the appendices attached towards the end of the thesis.

On the whole the work embodied in the thesis addresses the entire gamut of issues relating to plant gene expression in *S. pombe* system. The work establishes *S. pombe* as a model system for plant seed protein gene expression and as an alternative system for studying RNA processing. The findings have far reaching significance both in the basic as well as applied aspects of these key areas of molecular biology.