Summary and Perspective

Summary and Perspective:

The work done in this thesis explains the Physical Mapping, Functional characterization and sequence analysis of Type II Polyketide Synthase genes of *S.flavisclerotícus*. The organism of this study, *S.flavisclerotícus*, was selected on the basis of following criteria – (i) it is lesser known species of *Streptomyces*, (ii) there is brevity of knowledge in the public domain of the literature of its production potential, (iii) the genetic potential was revealed in terms of production of polyketide compound in possessing the genes for aromatic type II polyketide compound, a finding borne by the genome hybridization to conserved PKS genes, (iv) importantly, *S. flavisclerotícus* also produces multiple bioactivities; thus the happenstance that the polyketide compound is produced is real.

Literature in Chapter I covers the work by different groups on exploiting genetic potential of different *Streptomyces* strains which are polyketide producers to generate novel molecules by 'mix and match' strategies. The science of 'mix and match' has been improved to the point that 'tailor made' products can be produced. Though the combinatorial potential of the type II aromatic PKS genes has been appreciated in the literature surveyed, it is a well accepted fact that the potential is larger for the type I PKS genes that biosynthesise reduced polyketides (see Introduction and references therein). Manipulation of the tailoring enzymes and engineering the polyketide backbone are the two levels at which the genetic engineering of the new cluster can be effected for production of new derivatives. This has been increasingly recognized as the complementary approach to combinatorial chemistry of small but complex biomolecules. The two approaches are important in the light of fact that there is dearth of the newer molecules, specifically antibiotics being produced and reinventing the same molecule in exhaustive screening programs for newer structures. The project pursued in the lab revolves around the characterization of the PKS II genes of *S. flavisclerotícus*.

The thesis work has been organized as under: In Chapter 1 we describe the successful construction of the genomic DNA library of *S.flavisclerotícus*. The library was constructed in to a *E.coli-Streptomyces* shuttle cosmid vector pKC505 and PKS^+ clones were selected on the basis of their being able to hybridize to most conserved genes for the Polyketide synthase cluster, the *actI* DNA from *S. coelicolor*. Four putative PKS

containing clones (1.23, 1.52, 2.19 and 1.1) were isolated and the overlap between them was worked out. The region spanned between these four clones was ~45 kb in size and was presumed to contain the putative PKS cluster, deducing the overlap among the four selected clones helped us in narrowing down to the PKS genes containing region. On the basis of the information of the overlap, the restriction map for the ~45 kb region was constructed.

Chapter 2 reassesses the restriction map of the PKS DNA which involved subcloning different restriction fragments. Southern hybridization of the restriction digested DNA of the different library clones helped locate the conserved PKS DNA, that of KSα and KSβ. Subclones were constructed (Chapter 3) that carried the KSα and KSβ region and were sequenced. The sequencing results indeed showed the presence of Ketosynthase (KSα) and Chain length factor (KSβ) but did not reveal the gene sequence of Acyl Carrier Protein (ACP). This feature is unusual in that KS, CLF and ACP makes the minimal PKS and is responsible for the synthesis of the basic polyketide backbone. With the exception of *dnr* cluster for daunomycin biosynthesis in *S. peucetius*, medermycin cluster from *Streptomyces* sp.AM-7161, chromomycin cluster from *S. griseus,* the genes for ketosynthase (KSα), chain length factor (KSβ) and acyl carrier protein (ACP) are not only together but transcriptionally coupled too.

In case the molecule is a glycoconjugate, it's a well established fact that the glycosylation of the polyketide backbone adds up to the potency of the molecule. In the PKS cluster of this organism too, we found the presence of the genes like TDP-4-6, glucose dehydratase and O-methyltransferase involved in the biosynthesis of sugar. TDP-4-6, glucose dehydratase is transcriptionally coupled with the KS and CLF genes and becomes the part of an operon. This indicates that the sugar biosynthesis starts alongside the polyketide backbone biosynthesis. In the present study we tried to understand the role of these genes in the polyketide biosynthesis by constructing the deletion and disruption mutant for these genes (Chapter 4).

The genetic similarity between *S.flavisclerotícus* and *S.griseus*, and that *S.griseus* produces chromomycin raises the possibility that molecule synthesised by *S.flavisclerotícus* could be related to chromomycin. The NMR, IR and MS studies

established the biomolecule to be the Chromomycin antibiotic (Namita Gupta, Ph. D. thesis 2007).

In an interesting experiment in 2002, Metsa- Ketala (Metsa- Ketala et.al., 2002) showed that the essential 613 bp region of the KS gene could be used better to exploit the phylogenetic relatedness between the two aromatic polyketide producing *Streptomyces* species. They have categorically mentioned there that the phylogenetic tree constructed by using the 120 bp of the γ-variable region of the 16S ribosomal DNA is not congruent with the one created by the KS gene. This way they have established that there has been the horizontal gene transfer of aromatic polyketide biosynthesis gene among the *Streptomyces* bacteria. Sequencing results obtained in the studies here bear on this fact.

The same explanation as above is applicable to the situation we encountered here. Sequence similarity of more than 90%.between the KSα and KSβ genes of *S.flavisclerotícus* and *S.griseus* is indicative of very close relatedness between the two; on the other hand, ribotyping results don't corroborate this fact. Horizontal gene transfer of chromomycin gene cluster between the two species of *Streptomyces* is the likely cause of inconsistent pattern of relatedness between the pair of species.

Chromomycin is a DNA minor groove binding antibiotic (Chakraborty et.al., 2001) which is used as a dye. It belongs to aureolic class of compounds, a glycocojugate too, which includes other compounds like mithramycin, olivomycins, chromocyclomycin, UCH9, and durhamycin A, possessing DNA binding properties and described in the literature since 1950. The aureolic acids are neoplastic antibiotics that act against gram-positive bacteria and also stop the proliferation of tumor cells. In the presence of Mg^{2+}, these compounds inhibit replication and transcription processes by interacting with G-C-rich regions in the minor groove of DNA.

The glycans are water soluble molecules and are considered to be very potent molecules as most of the anticancer drugs are glycosidic in nature (Vladimír Kren* et.al., 2001). The initial studies of the partially purified extract on the primary cell lines revealed that this compound is highly cytotxic in nature (see Appendix 1). Given that the glycoconjugate antibiotic is chromomycin, it is not unexpected that it is scored positive in vitro cell lines cytotoxicity assay.

Mendez-Salas group in Oviedo, Spain in collaboration with Rohr group in Kentucky, USA, sequenced and published the whole cluster of chromomycin synthase gene cluster from *S. griseus* sub *griseus* (Mendez et al, 2004) but the minimal PKS sequence determined by us, in cluster from *S. flavisclerotícus,* was submitted to Pubmed prior to the release of sequence from *S. griseus* by them. Sequencing of complete ~45 kb region was carried out and it was found that the gene organization of the *S.flavisclerotícus* and chromomycin producing *S.griseus* is astoundingly same and genes' similarity index high for the two species.

Though the chromomycin's existence is several decades old, genetic studies of its biosynthesis, regulation are being attempted only recently. Programming of the PKS genes for anguicycline, tetracycline, isochromanequinine, anthracycline class of compounds is being understood, PKS genes of aureolic class of compounds were however not available till recently. Furthermore, the cluster is a rich source of glycosyltransferse activities and C- and O-methyl transferase functions. Studies on specificity of these enzyme systems could be exploited for generation of the modified polyketides.

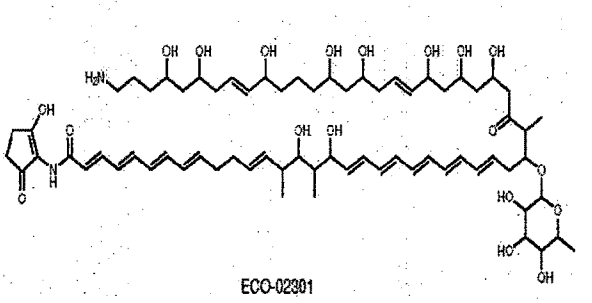
Genomics of the bacteria has a great deal of impact in the antibiotic discovery too With the improvement in the understanding of the *Streptomyces* genetics and with better molecular biology techniques available, it has become easy to find out the new candidates for the antibiotic production.

The recent advances in the whole genome sequencing of microorganisms have revealed some important facts. The discovery that there are far more biosynthetic gene clusters than there are currently known metabolites for a given organism, suggests that the biosynthetic potential for natural products in microorganisms is open to be tapped. For example, the genomes of 294 microorganism have be been sequenced and annotated in the National Centre for Biotechnology Information genome project, and a remarkable outcome of this effort is the understanding that the number of genes that are expected to be involved in secondary metabolite production dramatically outnumbers the amount of known secondary metabolites An extreme example of this is in the cyanobacterium *Nostoc punctiforme.* The genome of this organism harbors 22 genes that encode

probable polyketide synthases or NRPSs, although only one of these has been related to a secondary metabolite.

Survey of 11 secondary metabolite producers— including *Streptomyces coelicolor*, and *Streptomyces avermitilis*, unveiled a total of 118 NRPS and polyketide synthase clusters, only 14 of which had been assigned to known non-ribosomal peptides or polyketides at the time the genome sequences were reported.

A case in point of genomics aiding antibiotic discovery is that of genome scanning and bioinformatic analysis of the bicyclomycin-producer *Streptomyces aizunensis* NRRL B-11277 has revealed a 35 open-reading-frame (ORF) gene cluster and predicted putative structure .of the molecule it could synthesise. Fermentation of this strain under 50 different culture conditions led to isolation and characterization of ECO-02301 by researchers at Ecopia Biosciences (Saint Laurent, QC, Canada), which has a molecular mass of 1,298 (C70H109N2O20) and shows activity against a broad spectrum of pathogenic mycoses, including drug-resistant strains of *Candida albicans*



That only 1% of the microbial community is estimated to have been cultivated in the lab, implies that lying ahead is a vast biodiversity of natural products in microorganism that remains to be exploited

.

Efforts to expand the range of bacteria that can be tapped for antibiotic research are being facilitated by several strategies: expanded conventional culturing approaches, novel culture methods, heterologous DNA-based methods and metagenomics. The last in the list of strategies, metagenomics is quite recent. This refers to an attempt to capture

DNA from the environment (from the so-called 99% unculturable majority) and use it in heterologous expression systems

Appendix 1.

ACDSF#	Compounds	Cell line	Activity Status
347	JP1	Colo205	10
349	JP1	Colo205	10
355	JP1	Hop62	NA
356	JP1	SiHa	10
366	JP1	Colo205	10
372	JP1	Hop62	10
374	JP1	Hop62	10
380	JP1	SiHa	10
387	JP1	SiHa	10
396	JP1	PC-3	10
404	JP1	MCF7	10
406	JP1	MCF7	10

10 microgram /ml concentration of the partially purified compound from *S.flavisclerotícus* was used compound on the 5 cell lines we used. These cell lines are of 5 different tissues of origin which indicates a broad spectrum activity of the compound under study