Introduction

Introduction:

To address the worsening problem of antibiotic resistance in bacteria, there is an urgent need to develop new antibiotic. Comparative genomics and molecular genetics are being applied to produce the lists of essential new targets for screening of new compounds. In recent decade the effectiveness of the antibiotic has drastically decreased and there is a tremendous rise in the microbial drug resistance.

S. aureus is the most significant of these pathogens. It causes community and hospital acquired infections and is known to cause high mortality rate. Vancomycin has been used as the last resort against the *M*ethicillin *r*esistant *Staphylococcus aureus* (MRSA) since 1980s. *V*ancomycin *I*ntermediate *Staphylococcus aureus* (VISA) has been known since 1996 (Hiramatsu 1997), but the emergence of highly resistance strain (VRSA) is almost considered as the significant marker of the end of initial period of antibiotic. The other serious treatment problem includes Multidrug resistance in the Tuberculosis, *V*ancomycin *R*esistance *E*nterococci (VRE), resistance owing to extended spectrum β- lactamase in *Enterobacteriacae* and *P. aeruginosa* and Penicillin resistance in *S. pneumoniae* (Hughes 2003).

I. *Streptomyces*, Antibiotics and its regulation:

Streptomyces spp. is unusual among bacteria in growing as mycelial colonies with sporulating aerial hyphae. They are very important as the source of most of the major antibiotics. Pioneering work by David Hopwood in the 1950s and 1960s established *Streptomyces coelicolor* A3(2) as the model system for the genus. In the 1970s, plasmids were discovered and characterized, and used both to establish conditions for transformation and in the subsequent development of cloning vectors. Protoplasts were used in both transformation and highly efficient cell fusion. In the 1980s, the early cloning of resistance genes from antibiotic-producing strains was followed by the cloning of antibiotic biosynthetic gene clusters, and the development of general methods and probes for the cloning of such clusters from diverse species. Analysis of these gene sets led to wide-ranging inferences about the biosynthesis of the important polyketide class of antibiotics, and to the production of hybrid antibiotics, and then, in the last decade, to more sophisticated combinatorial biosynthesis of designer

molecules. In parallel, David Hopwood's work has also provided a crucial platform for studies of the regulation of the morphological and physiological differentiation that is manifested by sporulating antibiotic-producing colonies. Moreover the complete genome sequence of *S .coelicolor* has provided lot of insight in understanding and exploiting the area of *Streptomyces* biochemistry and combinatorial chemistry (Chater 1999).

The complex morphogenesis of the bacterial genus *Streptomyces* has made this genus a model prokaryote for study of multicellular differentiation, and its ability to produce a wide variety of secondary metabolites has made it an excellent supplier of biologically active substances, including antibiotics. As of now, of all the antibiotics known today, over 70% have been contributed by the *Streptomyces* species. They are the gram positive filamentous bacteria which generally reside in the soil. Just like fungi they show the complex morphological differentiation. On agar medium one or more substrate hyphae formed form a germinating spore branch frequently and grow rapidly by cell wall extension at the hyphal tips. Subsequently aerial hyphae emerge by reuse of DNA, Proteins and storage compounds, many cells in the substarte hyphae thus lyse and die. When apical growth of aerial hypha stops, in contrast to substarte mycelium, septa are formed at regular intervals along the hyphae to form many unigenomic compartments within a sheath composed of elongated hollow or grooved elements, finer fibrillar elements and amorphous material. The sporulation septa consist of two membrane layers separated by a double layer of cell wall material, which permits the eventual separation of the adjacent spores. Spore chains usually consist of many tens of spores. The aerial spores thus formed are resistant to heat and lysozyme digestion. *Streptomyces* strains are hence known as "boundary organisms" between prokaryotes and eukaryotes. (Elliot et. al., 2004)

Another characteristic of the genus *Streptomyces* is their ability to produce a wide variety of secondary metabolites, including antibiotics and biologically active compounds. Secondary metabolism is sometimes termed as "physiological" differentiation because it occurs during the idiophase after the main period of rapid vegetative growth and assimilative metabolism. The importance of *Streptomyces* could be gauged by not only its ability to produce the secondary metabolites such as antibiotics, immunosuppresants and enzyme inhibitors but producing industrial

enzymes such as glucose isomerase and transglutaminase make them versatile as well. Moreover, a variety of enzymes for secondary metabolite formation have become members of artificial gene clusters for the biosynthesis of the novel compounds by so called combinatorial biosynthesis.

Some of the aspects of the genetics of this antibiotic producing organism provide an insight in to the morphology and lifecycle of this complex organism. After spore germination, they grow to form aseptate multinucleate branching hyphae which can form a network to exploit a nutritive substrate. After depletion of the resource they subsequently erect aerial mycelia on which spores are elaborated prior to dispersal. This complex phenomenon must be achieved by the special genetic system. Like every other prokaryote, the genome of this organism constitute the chromosome and other genetic elements such as plasmids, phages and transposons. All these components interact with each other giving rise to associations of varying permanence. This genome constitution and interaction set them apart from *E. coli.* One of the major differences is in the DNA base composition of the organism. The *Streptomyces* spp. shows the 73% G+C composition in comparison to 50% in *E. coli.* The reason for this kind of composition is still unclear but one of the ideas is that it may confer enhanced UV resistance important in protecting the genomes of aerial mycelia and spores. One of the implications is also the different codon usage; one of the example where it is exploited is to ensure correct developmental timing of translation as genes necessary for vegetative growth lack the rare TTA leucine codon. (Benteley et. al., 2002)

The restriction modification system in these organisms is not as extensive as in *E. coli.* Though at least two species, *S. lividans* and *S. avermitilis* shows low level of novel and site specific modification to guanine, structure of the base remains to be determined but the modifying activity which acts post replicatively requires much more complex and extensive DNA substrate as compared to conventional DNA modification enzymes. This modification may influence the aspects of DNA metabolism such as gene expression and DNA repair in the same manner as host methylation in *E. coli* (Ikeda et. al., 2003).

It is being considered that the *Strpetomyces* genome is evolved by duplication. It is a linear molecule of 8 Mb size making it to be largest among the prokaryotes. The ends of the chromosome are characterized by long terminal inverted repeats (TIRs) which vary in length between 30 and 60 kb depending on the species. Terminal proteins covalently attached to the free 5' ends are thought to prime DNA synthesis, perhaps by formation of a complex of polymerase and the terminal protein bearing a priming nucleotide, just like adenovirus replication. Terminal protein primed synthesis is likely to be facilitated by other proteins that specifically interact within the TIRs. The first 170 bp of this TIR are found to be conserved in different species arranged as seven palindromes: this configuration may stimulate formation of DNA secondary structure important for stability of the terminal sequences, protein binding and correct processing of terminal sequences during replication. The *S. coelicolor* genome has been sequenced. It is assumed that the organism contain > 7000 proteins which is 20% more than unicellular eukaryote *Sacchromyces cerevisiae*. (Shwecke et. al.,1995).

Secondary metabolic pathways generally arose from the primary metabolic pathways by gene duplication followed by differentiation. That is the reason why some of the steps in the biosynthesis pathways of secondary metabolites are common and mechanistically similar to the primary pathway. Moreover, it often happens that the primary and secondary pathways use the same pool of small precursors. This suggests that some enzymatic activities are shared up to the point at which this pathways diverge. *Streptomyces* and *actinomycetes* are the prolific source of novel secondary metabolites with a range of biological activity which could be considered very versatile. The production of secondary metabolites by the gram positive mycelia *Streptomyces* generally coincides, or slightly precedes, the development of aerial hyphae in surface grown culture. The genes for the production of individual secondary metabolites are arranged in clusters that vary in size from a few to over 100 kb (Benteley et al., 2002).

Almost all of these mentioned clusters contain pathways specific regulatory genes whose expression frequently depends on the genes that are required for the production of several secondary metabolites made by the strain. Some of these genes

specially the *bld* genes are also needed for the formation of aerial hyphae and spores. (Chater et. al., 2003).

In the regulation of the secondary metabolites; anitbiotics is a major area of research. Though numerous genes that affect the antibiotic production have been identified as well as characterized, the exact networking between the environmental factors and the antibiotic production has not been completely understood. Uguru et. al., (2005), have described the identification and characterization of a transcription factor, designated AtrA, that regulates transcription of *actII*-ORF4, the pathway-specific activator of the actinorhodin biosynthetic gene cluster in *S. coelicolor*. The disruption of the *atrA* gene causes the decrease in the antibiotic (actionrhodin) production though it is not attached to the antibiotic producing gene cluster but it didn't affect the production level of undecylprodigiosin. This implicates the pathway specificity of the activator. This particular activator Atr A also binds in vitro to the promoter of *strR* which is unrelated to the *act II-* ORF 4 and the final regulator of the streptomycin production in *Streptomyces griseus*. This suggests that the pathways specific activators of the antibiotic production of different *Strpetomyces* may share evolutionary conserved components (Uguru et. al., 2005). Yamanaka et. al., (2005) have shown that there is interspecific stimulatory event between the *Streptomyces* species for the antibiotic production as well as for the morphological differentiation which is mediated by the diffusion of the metabolites at high frequency. *Streptomyces* and related organisms produces a siderophore like desferrioxamine E which is such stimulatory compound . They have shown this similar kind of effect on the *S.tanahensis* by this desfroxamine E produced by *S.griseus* (Yamanaka et. al., 2005).

It is being considered that reduction in growth rate is an important signal for triggering secondary metabolism. Theory of antibiotic production triggering by the highly phosphorylated guanosine (p)ppGpp has gained a considerable ground. Chakraborty et al in 1997 has observed that the ribosome attached ppGpp synthetase (Rel A) is required for the antibiotic production under nitrogen limiting condition in *Streptomyces coelicolor* A3(2) (Chakraburty et.al., 1997). Though its unclear whether ppGpp was directly involved in promoting transcription of antibiotic

biosynthetic genes or it was an indirect consequence of a reduction in growth rate prompted by ppGpp-mediated inhibition of rRNA synthesis. Hesketh and Bibb (2001) produced a more direct evidence of its involvement in the antibiotic production. They used a modified *rel A* gene to induce the ppGpp synthesis in the *S. coelicolor*; without detectably reducing in the growth rate, the transcription of the *act II* – orf 4, the pathways specific gene for actinorhodin production occurred under these conditions (Hesketh et. al.,2001).

ppGpp plays a direct role in activating the transcription of pathway-specific regulatory genes for antibiotic production. Both morphological development and secondary metabolism are simultaneously under the control of various nutritional environments such as carbon, nitrogen, phosphorous and trace elements. Chemical signaling molecule like γ- butyrolactone ring also controls both morphological and physiological differentiation of *Streptomyces* spp. The pioneer work of Khokhlov et. al., (1967) on an autoregulatory factor (Factor A, 2- iso capryloyl- 3R-hydroxymethyl-γ- butyrolactone) which induces both sporulation and streptomycin biosynthesis in a mutant of *Streptomyces grisues* revealed an exact link between secondary metabolism and morphological differentiation. A factor and its derivatives in various *Streptomyces* spp. are autoregulators or microbial hormone that can switch on morphological differentiation or secondary metabolites production, or both. The role of A factor which requires *afs A* for its synthesis, in regulating the onset of streptomycin production has been largely resolved. (Natsume et. al.,2004; Yamazaki et.al., 2004; Kato et. al., 2004). Binding of A factor to its cytoplasmic binding protein ArpA releases the latter from the *adpA* promoter, allowing *adpA* transcription. AdpA is required for the activation of transcription of *strR*, the pathways specific regulatory genes for streptomycin production and for the expression of their members of the *adp A* regulon, some of which are required for the morphological differentiation. Adp A appears to be the only Arp A- dependant gene for both secondary metabolism and morphological differentiation, and to play a role in modulating the A factor synthesis once the γ- butyrolactone has fulfilled its function in triggering both the processes (Kato et.al., 2004).

Identification of *adpA* as a key activator gene for both secondary metabolism and morphological differentiation in *S. griseus* has been elucidated in the figure below.

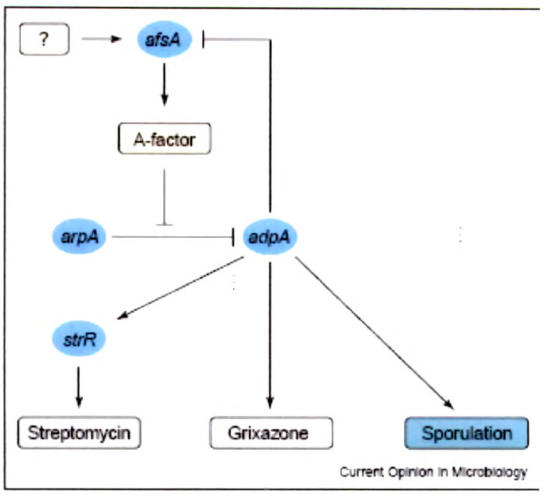

Figure I.1: The A-factor regulatory cascade of *Streptomyces griseus*. A-factor is detectable in the culture medium just before the onset of streptomycin production. The signal(s) (shown as '?') that trigger its synthesis, mediated in some manner by AfsA, are not known.

Just like A factor and γ- butyrolactone, there are lesser known family of the secondary metabolite regulatory family, recently a PI factor was identified from the *Streptomyces natalensis* which is elicitor of the anti fungal glycosylated polyene pimaricin (Recio et.al., 2004).

Similarly many regulatory proteins belong to the SARP family. These transcriptional activators contain a winged helix turn helix motif towards their N termini that is also found in the Omp R family of proteins, and some of the SARP family recognize the heptameric repeats within the promoter regions of the genes that they regulate. (Wietzorrek et.al., 1997; Widdick et. al., 2005; Lombo et. al.,1999; Sheldon et. al., 2002). They have been found to be associated with the secondary metabolite gene clusters that encode aromatic polyketide clusters e.g mithramycin, doxorubicin and medermycin (Pang et. al., 2004; Sheldon et. al., 2002; Ichinose et. al., 2003). Some ribosomally and non ribosomally synthesised peptides e. g. cinnamycin (Widdick et. al., 2005; Ryding 2002), undecylprodiginines (Cerdeno et.al., 2001), Type I polyketides (Takano et. al., 2005; Sun et. al., 2003; Oliynyk et. al., 2003), β- lactams e.g. carbapenem thienamycin (Nunez et. al., 2003), and Azoxy compounds e.g.

valanimycin (Garg et. al., 2002). These SARP family of regulatory proteins are the characteristic of actinomycests class of organisms. In the *S.coelicolor* and some of the Type I polyketide producing atnitbiotics, another class of transcriptional regulators are found which are known as the LAL family (Large ATP- binding regulators of Lux R family) of transcriptional regulators. These proteins contain the N-terminally located nucleotide triphosphate (NTP) binding motif and a C terminal helix turn helix motif characteristic of the Lux R family of DNA binding proteins. Just like SARP family these types of proteins are also confined to the *actinomycetes*. Though very little is known about this family of proteins in regulating the onset of antibiotic production, deletion or mutation in the NTP binding domain of PikD abolished the pikromycin production in *Streptomyces venezuela*. This family of proteins regulates the production of antibiotics like candidicine (Campelo et. al., 2002) and geldanamycine (Rascher et. al., 2003).

II. History and definition of Polyketide:

In 1907 Collie coined the term Polyketide. The very first definition of the polyketide given was: a compound with polyketomethylene groups, $(CH_2 - CO)_n$. Such material were said to be containing the multiple keten groups (Collie 1907). This also included the compounds derived from the polyketomethylene structures, for example by addition or loss of the water or by decarboxylation. Loss of water typically produced the cyclization of the carbon skeleton and therefore made further possible extensive modification. It was finally established that the central carbon skeleton of polyketides was formed by iterative decarboxylative condensations of malonic acid (or substituted maloninc acid). The pharmaceutical importance of polyketide has been increasingly appreciated and exploited in modern times as summarized by Cane (Cane 1997) the polyketides comprise a significant fraction, not only of the total number of microbial metabolites which have been identified with the physiological activities, but of the much smaller number which have found the greatest commercial application" This includes the antibacterials (e.g Erythromycin, tetracycline, and tylosin), immunosuppressants [e.g. mycophenolic acid, rapamycin and tacrolimus (FK506)], anticancer agents (e.g. daunomycin), antifungal agents (e.g. amphotericin and griseofulvin), cholesterol lowering agents (e.g. Lovastatin), and veterinary

products (e.g. avermectin and monensin). Diacetylacetone was also converted to acetyldihydorxy dimethylnapthalene. In describing the diacetylone – naphthalene condensation Collie led the foundation for his Polyketide biosynthetic hypothesis by writing as follows: " Polymerization and condensation are probably the two chief types of change which are instrumental in forming many of the multitudinous natural compounds, condensation being usually the outcome of the union of carbon atoms in consequence of the elimination of water" and he further noted that "it is well known how easily the acetyl group present in aldehyde and acetone condenses with itself." The 20 the century marked the importance of Polymerization in natural products biosynthesis and emphasized the role for simple and well understood chemical reactions such as dehydration, decarboxylation, condensation and oxidation/reduction. In the mid of the 20 the century Robert Robinson recalled the work of Collie and came up suggesting the biosynthetic role of polymerization and noted that "So far as we can see now the chemists of the Future must concentrate on the study of Polymerization". That small units polymerizes in to more complex structures (e.g. Polyketide and Fatty acid formation), and on a larger scale in to macromolecules (e.g. nucleic acids, polysaccharides, and proteins), is now a dominant theme in Biochemistry and Molecular Biology.

By the mid of the 20 th century the role of acetic acid in both primary and secondary metabolism was very well established and acetyl – coenzyme A was already identified as the biologically active compound. During that time only Birch and Donovan proposed a "polyacetate hypothesis" they got interested in the natural product structure determination and the newly generated topic of biosynthesis. Collie in his lecture at the American chemical society described the idea of biosynthesis and polyketones. Birch started working on camnospermol from camnosperma brevipetiolata which contains an obvious oleic acid side chain that is linked to a phenolic unit by –CH2-. At that time fatty acids were well known to be derived from the acetate, he proposed a continuation of the head to tail condensation leading to the formation of the aromatic ring. Birch constructed a structure which was possible β-polyketo acid side chain that could then be schematically cyclized. The polyketones formed by the head to tail linkage of the acetate units were proposed to cyclize by an aldol reaction(-Ch2-+-CO-) or by acylation (Ch2-+-CO-OH) to phenols Tetraacetic acid could form orsellinic acid or 6-methylsalicylic acid (6- MSA) via an aldol

reaction, or the phlorglucinol xanthoxylin by an acyl reaction. The modification of carboxylic acid was also considered thus decarboxylation of orsellinic acid would produce orcinol, and 6-MSa would lead to m-cresol. Compounds with fused ring systems could be derived from polyacetate compounds (Birch 1995).

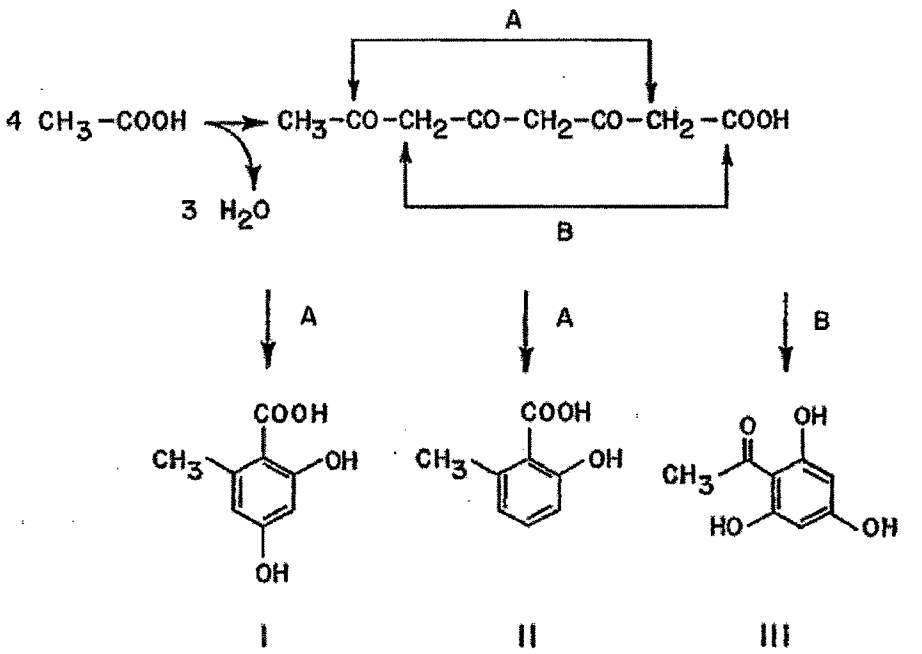

Figure I.2: The Birch & Donovan polyacetate hypothesis. Describing aldol type cyclization.

In 1950 with the help of isotopic tracer analysis in the intermediary metabolism, all the major pathways by which small precursors were incorporated in to large macromolecules were elucidated. Fatty acid Biosynthesis which shares the mechanistic similarity with the Polyketide synthesis, was shown to occur by sequential condensations of activated three carbon malonyl units with subsequent decarboxylation (Bentley 1962). In 1959 Lynen extended the concept of an acetate plus polymalonate condensation to the plant polyketide eleutherinol of eleutherinol from 8 malonyl- coA units but he could not emphasise on the importance of "starter unit" in the synthesis. After this with the help of same technique complete 'head to

tail' linkage of acetate units for the formation of 6-MSA from *Penicillium griseofulvin* and orsellinc acid from lichens were experimentally verified by feeding the organism with sodium [1- 14 C] acetate (Hutchinson and Fujii, 1995).

The general mechanism for the Polyketide construction was thus conceived to be the extension of the single starter unit R'-CH_2-CO-S-X, with elongation by R-CH(COOH)-CO-S-X extender units; where X-SH is normally either Co A or an Acyl Carrier Protein (ACP). In the process of extension, CO2 is lost. The basic reaction is as follows.

$$CH_3-CO-SX + n-1\ CH_2R''-CO-SX \rightarrow H-(CH_2-CO)_n-SX + n-1\ CO_2 + n-1\ X-SH$$

"Starter" "Extender"

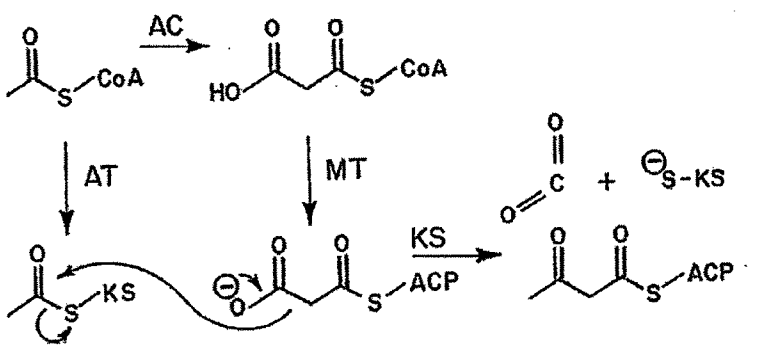

Figure. I.3: Initial reactions in fatty acid and polyketide biosynthesis. Acetyl-coenzyme A is converted to malonyl-coenzyme A by acetyl-coenzyme A carboxylase (AC). Acetyltransferase (AT) transfers an acetyl group to the SH group of the ketosynthase (KS). Malonyl transferase (MT) transfers a malonyl group from malonyl-CoA to acyl carrier protein (ACP). The acetyl-KS and malonyl-ACP yield acetoacetyl-ACP under the influence of KS; CO_2 and S–KS are eliminated in this decarboxylative condensation.

Mechanistically this process is the attack of a nucleophile, -CH- CO-SX, formed from the malonate component. Hence finally the Polyketides could be described in the biosynthetic terms. The formal definition of the Polyketide was proposed: "Polyketides are secondary metabolites with carbon chains formed biosynthetically

by extension of a "starter" unit with –CHR-CO- "extender" units where R= H or an alkyl group; the alkyl group is frequently either CH3 or C2H5. The starter unit can range from a simple structure, e.g. acetyl-CoA or propionyl-CoA, to a more complex material, e.g. C6–C3 and C7 units derived from the shikimate pathway or C7-N units probably derived from an aminoshikimate pathway. The extender units usually originate from a malonic acid or substituted malonic acid thioester formed with the SH group of acyl carrier protein (ACP); the ACP esters in turn are derived from CoA esters. Insertion of the extender unit into the chain requires a condensation reaction involving decarboxylation of the malonate structure. Some plant polyketides are formed by pathways not requiring the use of ACP. Polyketide chains may contain –CHR–CO– units in which R –esters- H at some positions and R-esters and alkyl group at others. The initially formed chains can retain all of the introduced CO groups, thus having the general structure R'–(CHR–CO)n–S–X, where R' derives from a starter; alternatively, one or more of the CO groups can be transformed enzymatically. Possible enzyme activities include: ketoreductases (KRs), –CO– to –CHOH–; dehydratases (DHs), –CHOH–CH2– to –CH=CH–; and enoylreductases (ERs), –CH=CH– to –CH2–CH2–. In many cases, chains undergo cyclization. The enzymes building the initial chain or backbone are termed polyketide synthases PKSs. Further post-PKS "tailoring" processes are very common; they include hydration, dehydration, oxidation, reduction, decarboxylation, methylation, glycosidation addition of isopentenyl groups, and expansion or contraction of ring systems" (Benetley et. al., 1999).

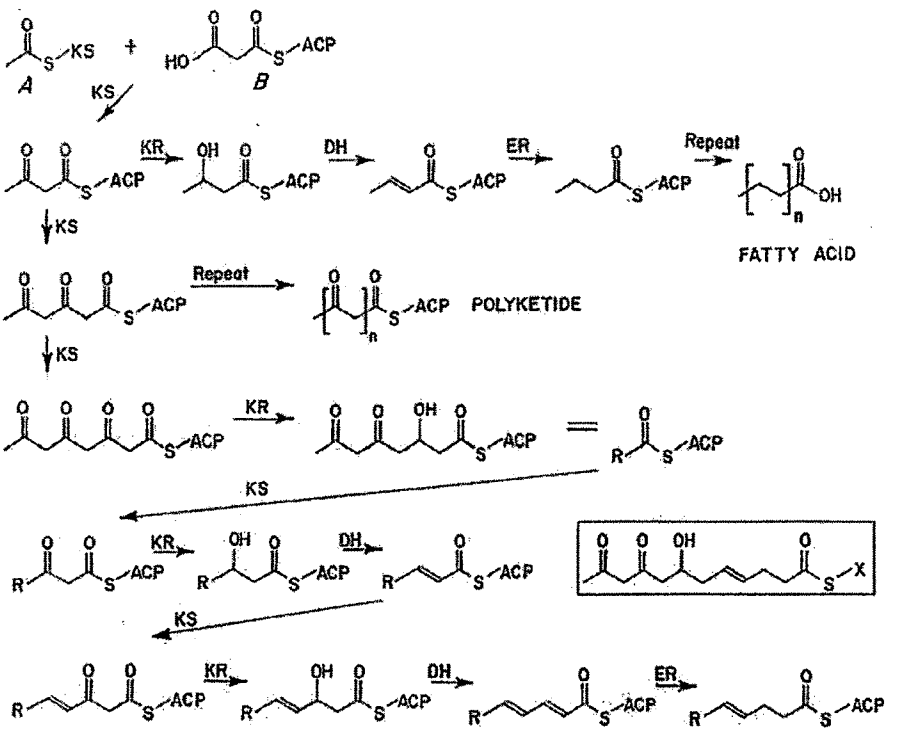

Figure. I.4: Biosynthesis of fatty acids and polyketides. Enzyme activities are abbreviated as follows: KS, ketosynthase; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; ACP, acyl carrier protein; *A*, the acetyl derivative of the KS enzyme; and *B*, malonyl-ACP. The figure depicts the sequential reactions used for the biosynthesis of fatty acids and an initial polyketide (polyketomethylene); for the former the enzymes used are KS + KR + DH + ER while for the latter only the KS enzyme activity is used.

III. Polyketides as secondary metabolites:

The original meaning of the polyketides is the compound with the polyketomethylene groups, $(CH_2\text{-}\ CO)_n$; such materials were said to contain "multiple keten groups" (Collie et. al., 1907).

Polyketides are one of the largest classes of secondary metabolites of which many are biologically active and covers a very wide pharmacological role (Strohl et. al., 1997; Katz and Donadio 1993) The polyketides are found in bacteria, plants and fungi and represents one of the largest group of natural products (Hutchinson 1988). Structurally they are divided in to four groups 1) aromatics e.g. doxorubicin and tetracycline. 2) Macrolide e.g. Erythromycin, rapamycin 3) Polyethers e.g. Monensin and salinomycin. 4) Polyenes e.g. amphotericine and candidicin. In vivo Polyketides result from the sequential condensations of short fatty acid biosynthesis and are catalysed by polyketide synthases (Lal et.al., 1996; Baltz et. al., 1997). The major thrust now a days in polyketide research is on elucidating the structure function relation of this complex multifunctional complex and to tap the potential of generating new polyketide libraries via combinatorial biosyntheis with engineered PKS (Hopwood 1993; Khosla et. al., 1996). Based on their molecular genetics, the biosynthetic systems of polyketides are generally classified into two groups modular (Type I) and aromatic (Type II) Polyketide Synthases. The third group, Type III polyketide synthases are distributed mainly in plants as chalcone snthase and stilbene synthases (Moore et. al., 2001). Type III PKS is found in microorganisms too. Here are the two characteristic examples of the different types of polyketides.

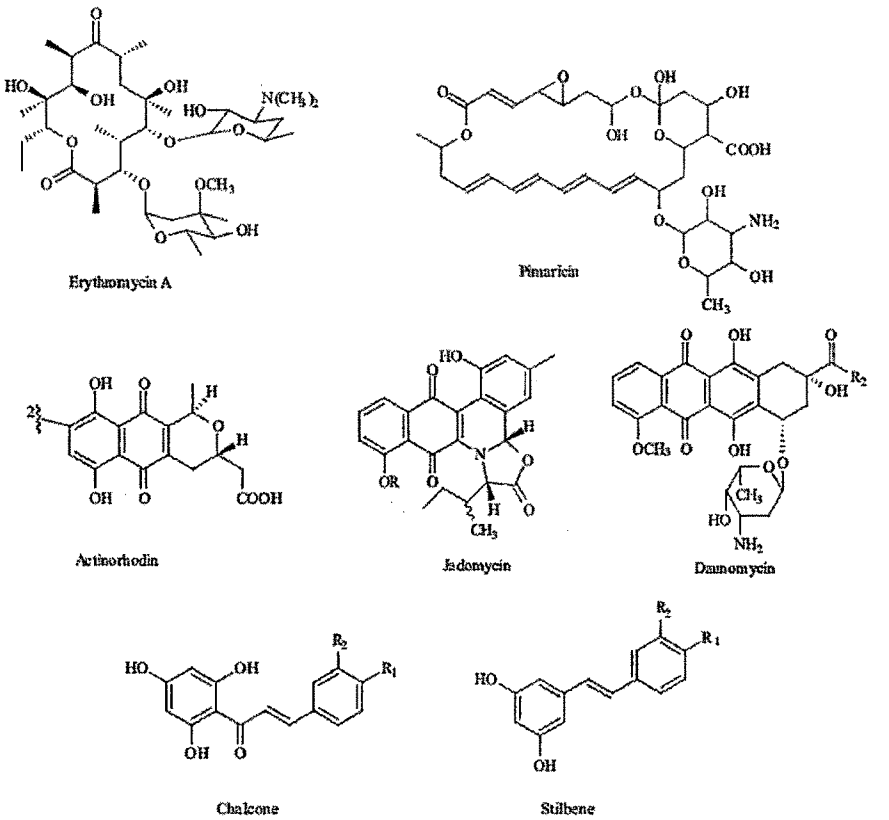

Figure. I.5: The structures of polyketides made by different types of polyketide synthases: erythromycin and pimaricin (type I); actinorhodin, jadomycin and daunomycin (type II); chalcone and stilbene (type III).

Even though the polyketides show a lot of natural structural diversity, the pharmaceutical industry are still on a lookout for new bioactive polyketide products. One method to achieve this is by combinatorial biosynthesis. The polyketides are the ideal candidates for that. The first target of the molecular genetic studies among the *Streptomyces* products was actinorhodin in the 1980s. This antibiotic represents the group of isochromane quinones that are biosynthesized by a polyketide pathway. Malpartida et. al in 1987 identified the *actI* homologous regions in *Streptomyces* spp producing aromatic polyketides; *actI* contains a PKS part of actinorhodin biosynthesis which is knows as the 'minimal PKS (Malpartida et. al. 1987).

Product type	Product	Strain	PKS genes	Accession number
Anthraquinones:	Actinorhodin	*S. coelicolor*	*act*	X63449
	Frenolicin/ nanaomycin	*S. roseofulvus*	*fren*	L26338
	Granaticin	*S. violaceoruber*	*gra*	X16144
	Griseusin	*S. griseus*	*grs*	X77865
	Naphthocyclinone	*S. arenae*	*ncn*	AF098965
Anthracyclines:	Daunomycin	*S. peucetius*	*dnrI/dps*	L35560
	Daunomycin	*Streptomyces* sp. C5	*dauI/dps*	L34880
	Nogalamycin	*S. nogalater*	*sno*	AJ224512
	Aclacinomycin	*S. galilaeus*	*akn*	AF257324
Anthracycline-like:	Tetracenomycin	*S. glaucescens*	*tcm*	M80674
Tetracyclines:	Oxytetracycline	*S. rimosus*	*otc*	Z25538
Aureolic acids:	Mithramycin	*S. argillaceus*	*mtm*	X89899
Angucyclines:	Jadomycin	*S. venezuelae*	*jad*	L33245
	Urdamycin	*S. fradiae*	*urd*	X87093
	Landomycin	*S. cyanogenus*	*lan*	AF080235
	unknown	*Streptomyces* sp. PGA64*	*pga*	AY034378
Others:	Pradimicin	*Actinomadura hibisca***	*prm*	D87924
	Ardacin	*Kibdelosporangium aridum***	*ard*	L24518
Spore pigments:	unknown	*S. coelicolor*	*whiE*	X55942
	unknown	*S. halstedii*	*sch*	L05390
	unknown	*S. curacoi*	*cur*	M33704
Unknown:	unknown	*S. cinnamonensis****	*mon*	Z11511
	unknown	*Saccharopolyspora hirsuta*****	*hir*	M98258
	unknown	*S. griseus* subsp. *griseus*****	*non*	AF074603
	unknown	*S. melaci*		AB021222
	unknown	*S. albus******	*sal*	-

Table. I.1: Cloned and Sequenced Aromatic PKS Clusters from *Actinomycetes*. Accession Numbers sequence of Min PKS.

This led to the idea of cloning a new cluster by using this *actI* as probe for hybridization. It was generally observed that the whole set enzymes involved in the biosynthesis are usually set in the contiguous DNA sequence in the host's genomic DNA. This way almost 26 gene clusters for aromatic polyketides have been cloned. They are shown in the table I.1.

As shown in above table, the polyketides are a versatile class of antibiotics and they have been subdivided in to several groups based upon their chemical nature, structure, their starter and extender units and range. Like anthraquinones, which are the derivatives of the anthracene used as dyes as well as antitumour agents, the anthracyclines and anthracylines like compounds mainly includes the antitumour agents like daunomycin, doxorubicin, and anticancer agents like tetracenomycin etc, they are based upon amine and tetra- hydro- naphthacene- dione. Then there are

aureolic acids like mithramycin, chromomycin which are strong antineoplastic agents. The angucyline like compounds are the derivatives of the anthracyclines like landomycin, jadomycin. The biosynthetic pathway of aromatic polyketides starts by deacarboxylative condensation of simple carboxylic acid residues. The starter units are mostly the acetate units like their extenders but there are some exceptions as well, like a propionate is used as the starter unit for the synthesis of the anthracyclines like aclacinomycines and daunomycine (Strohl et.al.,1997), butyrate for frenolicin (McDaniel et. al., 1993) and malonamide for oxytetracycline biosynthesis (McDowall et. al., 1991).

The enzyme reaction for the first condensation reactions is minimal PKS, including a condensing enzyme KSα, its homologue KSβ and an acyl carrier protein (ACP). The close similarity between the genes for minimal PKS derived from different *Streptomyces* suggests common stem genes for a wide variety of aromatic polyketides (Metsä-Ketelä et. al., 1999).

This minimal PKS genes forms the basic structure for any aromatic polyketide. Furthermore additional enzymes such as ketoreductases, an aromatase, cyclases and oxygenases, contribute to form a stable, correctly folded structures. Modifying reactions too occur post polyketide synthesis such as hydroxylation, O-metylation and glycosylation from an enormous source of different sugar residues, imparting huge structural variety for these compounds.

IV. Polyketide synthase (PKS):

Polyketides are produced from the microorganism through a specialized metabolism which is a variant of fatty acid synthesis. They are not only structurally and sequentially similar but also mechanistically similar. This similarity helped in understanding the polyketide biochemistry. There are two major pathways believed to be followed by the Polyketide synthase machinery; in one polyketides are formed by enzymes complexes consisting of four or seven monofunctional proteins in which the β- carbonyl groups of the intermediates resulting from the condensation of acetate residues are largely not reduced, and cyclisation of the intermediates produces

aromatic compounds. The intermediates in the other model are formed by multifunctional enzymes in which each of the initial condensation products is processed through reduction; reduction and dehydration or reduction dehydration and further reduction to produce highly reduced compounds from acetate, propionate and butyrate residues. Expression of the genes encoding each type of Polyketide synthase or their mutant forms has provided much information about the underlying biochemistry and in some cases resulted in the formation of novel natural products. Although early isotope labeling experiments clearly demonstrated that FASs and PKSs use similar substrates, it is the recent cloning of PKS genes and the biochemical characterization of PKS enzymes that have provided a mechanistic explanation of how PKSs achieve the vast structural diversity during polyketide biosynthesis by varying the similar biosynthetic reactions of FASs. Thus, unlike fatty acid biosynthesis, in which the β-ketone group of the growing fatty acid intermediate 1 undergoes full reduction to a methylene group 2, 3 during each cycle of elongation (pathway A in Figure I.5), the β-ketone group of the growing polyketide intermediate 4 could either be left untouched, leading to aromatic polyketides (pathway B in Figure I.5.), or be subjected to no, partial, or full reduction , depending on a given cycle of elongation, leading to macrolides, polyethers,or polyenes (pathway C in Figure I.5.). The latter forms the mechanistic basis for grouping macrolides, polyethers, and polyenes together as complex or reduced polyketides (Hutchinson and Fujii 1995).

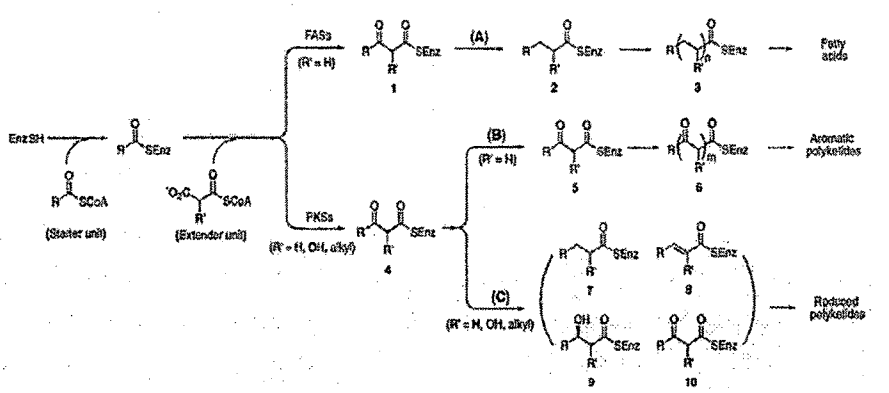

Figure. I.6: Biosynthetic pathways showing the requirement of starter unit and extender unit and the processing of the elongating intermediates: (A) for fatty acids, (B) for aromatic polyketides, and (C) for reduced polyketides

Type I : Micorbial PKS are of two types, type I and type II. Type I or modular Polyketide synthases consists of giant multifunctional proteins to make polyoxygenated aliphatic compounds from several different types of acyl- Coenzyme A substrates. The modular PKSs have led to the most effective genetic engineering route to produce structural variants of polyketides that are important therapeutic drugs, like the antibacterial erythromycin A or experimental agents such as 17-demethoxygeldanamycin (17-AAG) that currently is undergoing clinical trials as an antitumor drug (McDaniel et al. 2005).

A modular PKS is a massive complex of large, multifunction proteins. Within each protein are one or more "modules", each with different combinations of domains that function like the constituent biochemical activities of fatty acid synthases to catalyze a single cycle of polyketide chain elongation and modification.

6-Deoxyerythronolide B synthase (DEBS) is the PKS that forms the backbone of the erythromycins and is encoded by the three genes, *eryAI, eryAII eryAIII.*

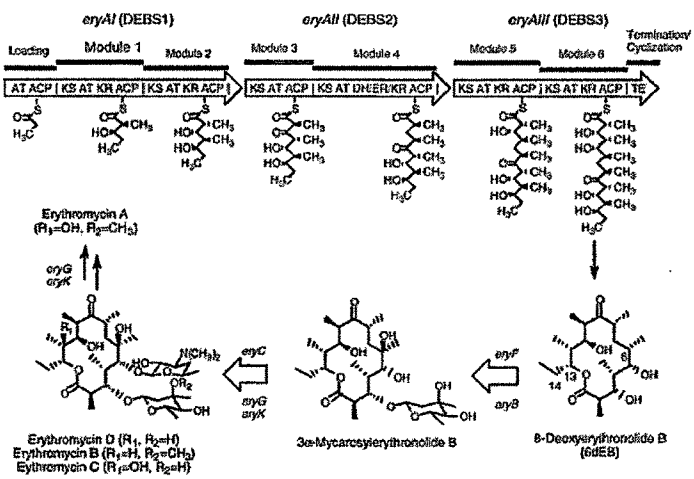

Figure. I.7: Illustration of the mechanism of the type I modular PKS involved in the biosynthesis of 6dEB.

DEBS catalyzes formation of 6-deoxyerythronolide B (6dEB) by the successive condensation of one propionyl and six 2-methylmalonyl molecules in their activated

Coenzyme A (CoA) thioester form. Each of the three subunits of DEBS have two extender modules, containing the activities needed for one cycle of polyketide chain elongation, as illustrated by the structures of the six enzyme bound intermediates in above figure. In addition, the first module is preceded by a loading didomain for the starter unit, and the last is followed by a thioesterase domain for product release and cyclization. Every extender module contains a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP) domain that together catalyze a two-carbon extension of the chain. In DEBS, the AT domains of extender modules are specific for 2-methylmalonyl- CoA, while the AT in the loading module uses propionyl-CoA. After each two-carbon unit condensation, the oxidation state of the β-carbon is either retained as a ketone (module 3) or modified to a hydroxyl, methenyl, or methylene group by the presence of a ketoreductase (KR) (module 2), a KR + a dehydratase (DH), or a KR + DH + an enoyl reductase (ER) (module 4), respectively. In effect, the AT specificity and the types of catalytic domains within a module serve as codes for the structure of each two-carbon unit; the order of the modules in a PKS specifies the sequence of the distinct two-carbon units, and the number of modules determines the length of the polyketide chain. Variations in the acyl-CoA substrates used by a modular PKS, the number of domains within a module, and the number of modules in the PKS are responsible for establishing the first set of structural characteristics of the polyketide, including the chirality of hydroxyl- and alkyl-bearing carbon centers. After this, the kinds of biochemical transformations the compound produced by the PKS undergoes, such as glycosylation or oxidation, are dictated by the "tailoring enzymes" that establish the final structure. Consequently, engineering a microorganism to produce novel polyketides can involve altering only the PKS genes or the tailoring genes as well. The 14 membered macrolides are exemplified by erythromycin, for which the complete gene sequence has been determined (Corte's, et. al., 1990; Donadio et. al., 1991; Bevitt et. al., 1992; Summers, et.al., 1997; Gaisser et.al., 1997; 1998).

The overall DEBS like architecture of the PKS is conserved in the molecules like oleandomycin, and megalomycin with an exception inpicromycin (PicPKS) in which the last two modules are contained on the two separate proteins. Thus the gene architecture is almost the same still the variation seen in the structure of oleandomycin megalomycin and picromycin is attributed to the use of different

starter units and in the mechanism of their loading domains (Shah et. al., 2000; Xue et. al., 1998).

The polyketide macrolactone are further modified by a series of glycosylation and oxidation. Synthesis of 6dEB is followed by C6- hydroxylation by the product of eryF to yield erythronolide B. Addition of the sugar L-mycarose (via thymidine diphosphate mycarose) yields 3-O-α- mycasoryl erythronolideB, and addition of D-desosamine (via TDP – desosamine) yields erythromycin D. The two sugars are produced by independent sets of genes designated *eryB* (mycarose) and *eryC* (Desosamine), which flank the PKS genes in the cluster. The final steps of erythromycin biosynthesis are hydroxylation of erythromycin D to yield erythromycin C by a second P450 enzyme encoded by *ery K* and O-methylation of the macrosyl residue by the *ery G* product to yield the cladinosyl moiety of erythromycin A and B. Similarly the 16 membered macrolactone too are classified in to three groups on the basis of their structure and composition. While the basic gene organization is conserved among all only last few differs in the number of modules. The main difference is in the composition of the extender units utilized. They vary from malonyl Co A, methyl malonyl Co A to crotonyl Co A (Bisang et. al., 1999; Mc Daniel et. al., 2005).

Similarly ansamycins are related to the macrolides biosynthetically but differ in the choice of starter unit (3- amino-5-hydroxybenzoic acid (AHBA)) and lack of glycosylation. Formation of macrolactam between the terminal carboxyl and 3 amino group of AHBA, instead of amcrolactone as discussed earlier, results in a characteristics "basket with handle" molecular conformation (Rinehart 1976) e.g Rifamycin, geldanamycin, herbmycin, and ansamitocin are the few ansamycins.

V. Aromatic Polyketide synthases

V.1: Fungal Aromatic PKS:

Fungal aromatic PKSs are iterative type I enzymes. 6-Methylsalicyclic acid synthase (6MSAS), the first microbial PKS purified (Shen 2000), is a classical example of

fungal PKSs, though the 6MSAS gene has only been cloned recently from *Penicillium patulum,* homologous 6MSAS gene, *atX*, has also been cloned from *Aspergillus terreus* with the KS domain of the 6MSAS of *P. patulum* as a probe. Sequence determination of the 6MSAS gene from both *P. patulum* revealed a single orf that consists of characteristic KS, AT, KR, and ACP domains, confirming the type I nature of 6MSAS (Fujii et. al., 1996).

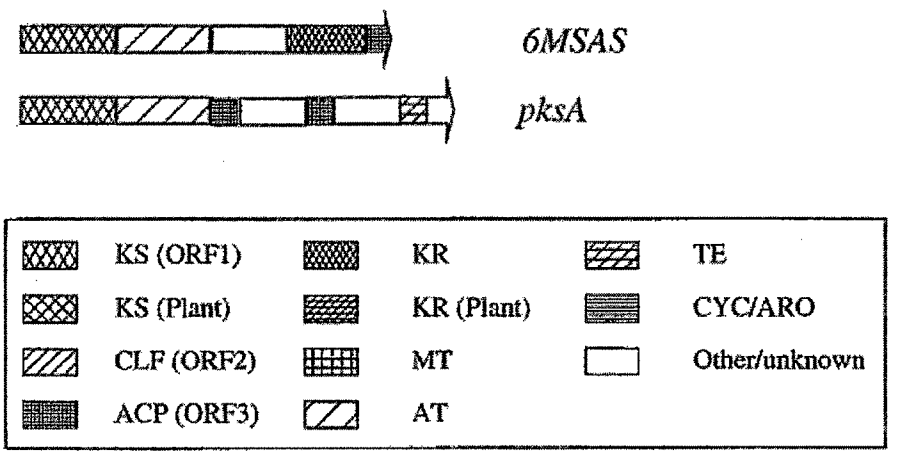

Figure. I.8: digrammetic presentation of the two fungal aromatic PKS clusters

Since the synthesis of 6-methylsalicyclic acid requires three steps of condensation, 6MSAS must have utilized the KS, AT, and ACP activities iteratively. The PKS genes, *pksA* and *stcA*, for norsolorinic acid, an octaketide intermediate of aflatoxin (AF) and sterigmatocystin (ST) biosynthesis, were also determined recently from *Aspergillus parasiticus* and *Aspergillus nidulans*], respectively. The *pksA*, also known as *pksL1*, and *stcA* genes are highly homologous and consist of only one set of characteristic KS, AT, and ACP domains, along with an additional thioesterase (TE) domain, reinforcing the notion that fungal PKSs are iterative type I enzymes (Fig. I.7). Interestingly, two functionally distinct FAS genes have been identified that are involved in primary and secondary metabolism, respectively. In *A. nidulans*, mutants of FAS genes for secondary metabolism, *stcJ* and *stcK*, grew normally but cannot synthesize ST, Similarly, in *A. parasiticus*, the two specialized FASs were identified as *fas-1A* and *fas-2A* that showed high sequence homology to the yeast FAS*b* and FAS*a* subunits, respectively, and insertional inactivation of *fas-1A* gave

mutants that were unable to incorporate acetate into AF pathway. These results strongly suggested the functional requirement for FAS in the biosynthesis of the C-6 starter unit that is further elongated by PksA and StcA during the biosyntheses of Aflatoxin and Sterigmatocystein, respectively (Brown et. al., 1996; Mahanti et. al., 1996). Other striking features of fungal PKS genes include the apparent lack of ARO or CYC in all known fungal PKSs, The TE domain is known to off-load the fully processed polyketide product from the noniterative type I PKS in reduced polyketide biosynthesis but has not been found in all other known aromatic PKSs. (Aggarwal et. al., 1995; Cortes et. al.,1995)

V.2: Type III Polyketide Synthase

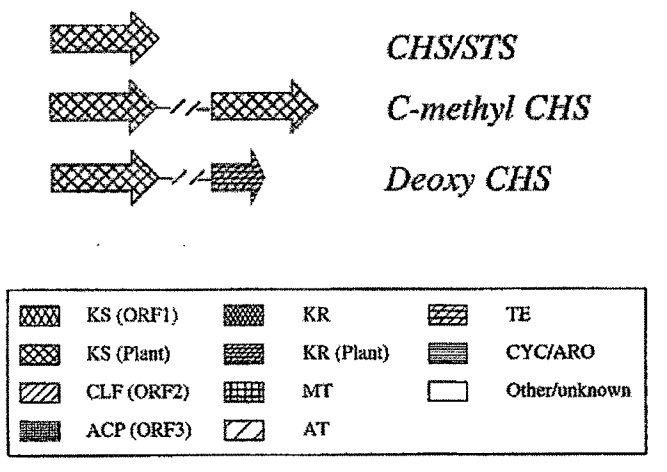

Figure. I.9: digrammetic presentation of TypeIII PKS clusters

Type III polyketides belong to the highly diverse family of polyketide natural products. Type III polyketides (as shown in figure), such as chalcone, stilbene, resveratrol, and pterostilbene, recently received much attention due to their biological activities. Abundant in colorful fruits such as grape and blueberry, these polyketides play beneficial medical roles as antioxidants, platelet aggregation inhibitors, anti-inflammatories and anti-cancer agents (O'Hagan et. al., 1995). Recently, they have been found to bind the PPAR, a family of proteins involved in lowering cholesterol

and other blood fats. Resveratrol present in red wine ("the French paradox") is beneficial to cardiovascular health (Monaghan et. al., 1990). Type III PKSs are structurally simple homodimeric enzymes that catalyse repeated chain elongation between a Co A linked starter unit (usually an aromatic CoA) and acetyl units (derived from Malonyl Co A). Following chain estension, the linear polyketide intermediate is cyclized in the same active site cavity. Chalcone and stilbene are both produced by the same chain elongation reaction, which involves the coupling of p-coumaroyl- CoA with the three malonyl CoAs (as shown in figure)

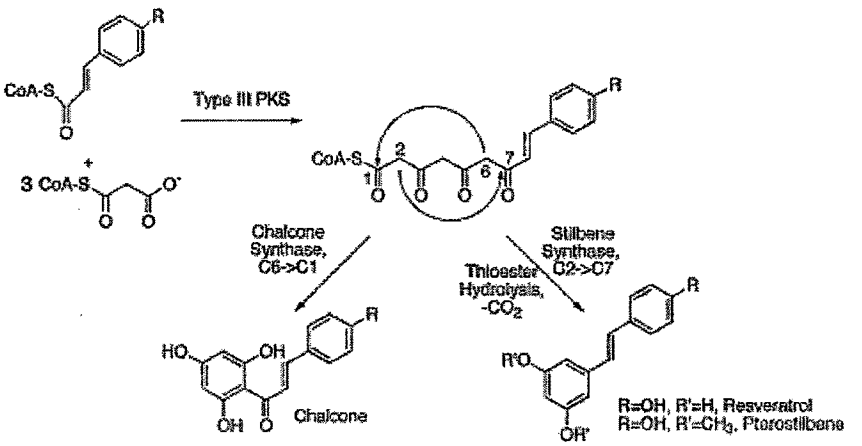

Figure. I.10: Biosynthetic mechanism of Chalcone and stilbene synthase

Subsequently, differential aldol- cyclization of the same linear polyketide intermediate, either C6- C1 or C2- C7, results in chalcone or stilbene respectively (as shown in figure). During the past decades, the combined studies of the stilbene and chalcone synthases spanning the structural and functional aspects of the enzymes have been very well studied. (Shen, 2000)

V.3 Aromatic Polyketide synthase Type II:

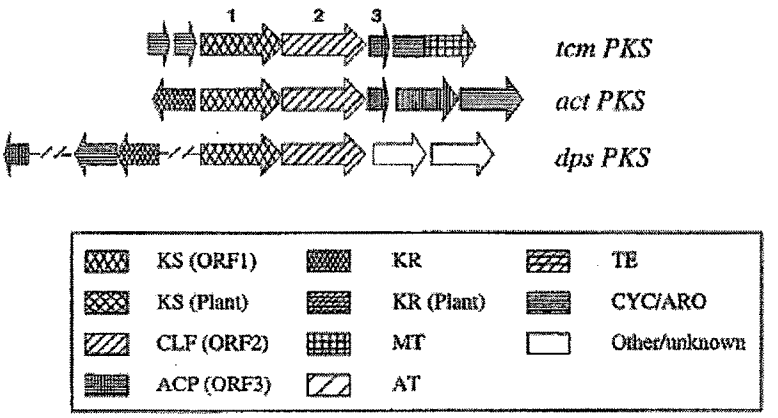

Figure. I.11: digrammetic presentation of the TypeII (aromatic) PKS clusters

Type II PKS is involved in synthesizing several clinically used drugs are aromatic polyketides, such as the antitumor agents daunorubicin/doxorubicin and the antibacterial antibiotic oxytetracycline etc. (O'Hagan 1991) Type II PKSs minimally consist of four subunits: ketoacyl synthase (KSα), ketoacyl synthase (KSβ), also known as chain length factor), acyl carrier protein (ACP), and malonyl CoA:ACP acyltransferase (MAT). A PKS catalyzes the biosynthesis of the polyketide backbone from acyl-CoA precursors, which usually are activated by a MAT for loading to the holo-ACP, and then condensed iteratively with a unique chain starter unit, via a series of decarboxylative Claisen condensation reactions until the specified chain length is reached. Aromatases and cyclases then catalyze the regiospecific folding and cyclization of the nascent polyketide chain to form the linear, angular, or discoid aromatic ring systems (O'Hagan 1992). The latter are then further customized by a host of tailoring enzymes that perform modifications such as oxidation, reduction, methylation, and glycosylation to yield the final polyketides. These modifications introduce a large amount of structural diversity, and many of the resulting products have interesting biological activities. (Rawlings 1997). This polymerization is almost similar to the fatty acid biosynthesis. Unlike fatty acid biosynthesis, the β-carbonyl of a growing polyketide chain may be left unreduced or converted to hydroxyl, enoyl,

or methylene functionalities. The level of ketoreduction, as well as the initial cyclization pattern of the full-length polyketide chain, is determined by the PKS (Khosla et al. 1993).

To obtain the nascent polyketide backbone, a subset of three proteins, the ketosynthase/ putative acyltransferase (KS/AT), chain length factor (CLF), and acyl carrier protein (ACP), is essential for polyketide synthesis and comprises the "minimal" PKS (Khosla, et. al., 1994).

The mechanisms by which bacterial aromatic PKSs synthesize a highly labile poly-β-ketone intermediate of precise chain length, and guide it toward a regiospecifically reduced and cyclized product, represent an exciting challenge in multifunctional enzymology and biomolecular engineering. Moreover, mutagenesis and heterologous expression of recombinant bacterial aromatic PKSs have provided some insight into the functions and molecular recognition features of the different protein components of the PKS and have laid the groundwork for the combinatorial biosynthesis of "unnatural" natural products (Khosla 1995).

Furthermore, cell-free systems for in vitro synthesis of aromatic polyketides have been developed, making it possible to purify individual components of a functional PKS complex. (Hutchinson 1993). The latter successfully led to the in vitro reconstitution of both the actinorhodin (Act) (Matharu et. al., 1998) and tetracenomycin (Tcm) PKSs (Carreras and Khosla 1998) from individually purified components, setting the stage to address the molecular basis of the structure and function relationship of type II aromatic PKS.

V.3.1: Phosphopantetheinyl Transferase:

Both the extender unit of malonyl CoA and the growing poly-β-ketone intermediates are covalently tethered to the ACP subunit of a bacterial PKS in an acyl thioester linkage to the phosphopantetheinyl moiety during polyketide biosynthesis (Hopwood et. al., 1990; Hopwood 1997). The latter prosthetic group is introduced posttranslationally by the phosphopantetheinyl transferase (PPTase), which transfers

the 4'- phosphopantetheine moiety of CoA to the highly conserved Ser residue of apo-ACP, converting into holo-ACP.

V.3.2: Acyl Carrier Protein:

ACP is a central component of type II PKS and is involved in possibly all reactions of bacterial aromatic polyketide biosynthesis. The starter unit (before its transfer to the KS subunit), the extender units, the growing poly- β-ketone intermediates, as well as the full length linear poly- β -ketone product, are covalently bound to ACP in a thioester linkage to the terminal sulfhydryl of the 4'-phosphopantetheine prosthetic group. These ACP thioesters must have been recognized as substrates by type II PKSs and other associated enzymes such as KR, CYC, or ARO to build and process the poly-β-ketone intermediates into aromatic polyketides. ACP are necessary for the function of a bacterial type II PKS as either deletion or introduction of a frame shift mutation into the ACP gene of the *act* cluster in *S. coelicolor* resulted in mutants whose polyketide production was completely abolished (Khosla et. al., 1993; Khosla et. al., 1992).

V.3.3: β-Ketoacyl Synthase:

Functional analysis of the PKS KS (also called KS α, (Bao et. al, 1998), encoded by *actI*ORF1, *tcmK*, or their homologs, was primarily based on the extensive knowledge of the FAS KS enzymes, which have been well characterized biochemically and genetically (Wakil 1989; Boom and Cronan 1989). Like fatty acid biosynthesis, KS α catalyzes the decarboxylative condensation between the growing poly-β-ketone intermediates and malonyl-ACP. The acyl group of acyl-ACP is first transferred to the Cys residue at the active site of the KS α, resulting in a thioester; malonyl-ACP is then decarboxylated to generate a carbanion that nucleophilically attacks the thioester to complete one cycle of elongation (Helariutta et. al., 1995; Ehmann and Schafer 1988)

V. 3.4: Cyclase and aromatase:

The genetic studies have identified a number of putative polyketide AROs (Rajgarhia and Sttrohl 1997; Alvarez et.al., 1996) and CYCs (Fernandez-Moreno et. al., 1992;

Grim et. al., 1994) these studies failed to reveal any mechanistic insights for these enzymatic aldol or Claisen condensations. (ARO and CYC are often used interchangeably, and there is little biochemical data to separate them into two subgroups of enzymes (Fu et. al., 1994; McDaniel et. al., 1994; 1995; Sherman et. al., 1991; Lomovskaya et. al., 1998). Biochemical characterizations of these enzymes are difficult because the presumed substrates, whether they are linear or monocyclic poly- *b*-ketones, are not directly available. Tcm F2 CYC (Shen and Hutchinson 1993) and TcmN (Shen and Hutchinson 1996) are the only polyketide CYCs that have been purified and characterized so far.

Tcm F2 CYC catalyzes the cyclization of tetracenomycin intermediates a reaction that closely resembles the intramolecular aldol and Claisen condensations involved in the synthesis of polycyclic aromatic polyketides.

TcmN is a bifunctional protein with the N-terminus catalyzing the cyclization of the Tcm PKS-bound linear decaketide to one of the tetracenomycin intermediate and the C-terminus catalyzing the O-methylation of Tcm D3 to Tcm B3 (Shen and Hutchinson 1996)

V.3.5: Tailoring enzymes:

Many natural products from this family require post-polyketide synthase modifications including glycosylation, alkylation and oxidation/reduction to be fully active (Rix 2002). The attachment of the saccharide moieties to the macrocyclic and aromatic backbone adds up to the versatility and potency of these molecules, change in the nature and extent of the sugar decoration alters the activity and even the specificity of these antimicrobial agents (Fu et.al., 2003). Glycan alteration is one of the most common resistance mechanisms in the intracellular inactivation of polyketide antibiotics and is used by the Gram-positive *streptomycetes* that produce these molecules as a protective mechanism from the action of their endogenous antibiotics (Quiros et. al., 1998). Of which the enzymes involved in the sugar biosynthesis are of very much importance. This synthesis mainly starts along side the polyketide backbone synthesis and often is the part of the operon. The enzymes mainly involved here are the ones involved in the deoxysugar synthesis and transefer

of the syntheised glycan moieties or the eznymes involved in the terminal modification or oxidation reduction.

VI. Cloning and sequencing of the Genes:

Soon after methods for gene cloning in *Streptomyces* species were published in 1980 (Bibb et. al., Thomson, C.J.; et. al., Suarez, J. E.; et. al., 1980) it became possible to isolate genes for antibiotic biosynthesis by a variety of procedures.(Hopwood et. al., 1983 and 1986) One of the first approaches involved the shotgun cloning of random fragments of DNA from a wild-type strain into a mutant blocked at a step in the biosynthesis and looking for a restoration of antibiotic production. A second early approach depended on emerging evidence for close linkage between genes for self resistance to an antibiotic and one or more of the biosynthetic genes. From these early results a most important generalization soon emerged: that in *Streptomycetes*, and by implication in other bacteria too, all of the biosynthetic genes needed to make a particular antibiotic from primary metabolites occur together in a single cluster, and that one or more genes for antibiotic self-resistance are also to be found there (Chater and Bruton 1985). A striking early demonstration of this was the cloning of the entire cluster of genes (the *act* genes) for biosynthesis of the pigmented benzoisochromanequinone polyketide actinorhodin on a 35 kb fragment of chromosomal DNA from the producer *Streptomyces coelicolor* A3(2), and their expression in the actinorhodin-sensitive and nonproducing *Streptomyces parvulus*, which thereupon produced actinorhodin without killing itself (Malpartida and Hopwood 1984). Establishment of this generalization which has continued to hold in all examples known to dates provided two general routes to the isolation of potentially complete sets of genes for the biosynthesis of polyketides by *actinomycetes*: (1) cloning of a library of fairly large fragments of wild-type DNA into any available mutant blocked in a step of antibiotic biosynthesis, looking for complementation of the mutation, and then finding genes for the other steps of the pathway on the complementing fragments; and (2) cloning a library of DNA fragments from an antibiotic producer into a sensitive surrogate host (usually a derivative of *Streptomyces lividans* 66, which is a convenient, generally antibiotic-sensitive and easily manipulated strain), selecting resistant clones, and seeking

biosynthetic genes linked to the resistance gene on the cloned DNA. By these procedures, the complete sets of biosynthetic genes for two further aromatic polyketides were isolated: the anthracycline, tetracenomycin (*tcm*) from *Streptomyces glaucescens*, (Motamedi and Hutchinson 1987) and oxytetracycline (*otc*) from *Streptomyces rimosus* (Butler et. al., 1989). The availability of cloned DNA carrying the biosynthetic gene clusters for these three aromatic polyketides led to a test of the idea Malpartida and Hopwood (1984) that the sequences of different PKS genes, which are presumed to have diverged from a common ancestor (discussed in section VI), might be sufficiently conserved for a DNA fragment for one synthase to be used as a probe to isolate genes for others. The presumptive positions of the DNA encoding the polyketide KS and KR functions within the actinorhodin gene cluster had already been deduced (Malpartida and Hopwood 1986) and so these DNA fragments (carrying the so-called *act*I and *act*III genes, respectively) could be used as hybridization probes against restriction digests of the *tcm* and *otc* cloned DNA (Malpartida et. al., 1987). Not only did strong crosshybridization occur with the KS probe, but it recognized regions of the *tcm* and *otc* DNA clusters that had already been identified by complementation analysis as candidates for carrying the PKS genes. The KR probe hybridized to a second segment of the *otc* gene cluster, but not to any part of the cloned *tcm* DNA. This latter result was significant because tetracenomycin is one of the few polyketides that arise without any of the keto groups of the nascent carbon chain being reduced.

Here are the few examples of the cloning and sequencing of the aromatic gene cluster from *Streptomyces*

host	polyketide	PKS genes	cloning strategy	nucleotide sequence accession number
S. coelicolor	actinorhodin	*act*	Complementation	M19536
S. rimosus	oxytetracycline	*otc*	Resistance*	Z25538
S. glaucescens	tetracenomycin	*tcm*	Complementation	M80674
S. violaceoruber	granaticin	*gra*	*act* probe	X16144
S. coelicolor	spore pigment	*whiE*	*act* probe*	X55942
S. peucetius	daunorubicin	*dps*	*act*/*tcm* probes	L35560
S.cinnamonensis	unknown	*mon*	*act* probe	Z11511
S. halstedii	spore pigment	*sch*	*act* probe	L05390
S. curacoi	spore pigment	*cur*	*act* probe	M33704
Sac. hirsuta	unknown	*hir*	*act* probe	M98258
S. roseofulvus	nanaomycin/frenolicin	*fren*	*act* probe	L26338
S. griseus	griseusin	*gris*	*act* probe	X77865
S. venezuelae	jadomycin	*jad*	*act* probe	L33245
S. sp. C5	daunorubicin	*dau*	*act* probe	L34880
Kib. aridum	unknown	*ard*	*act* probes	L24518
S. fradiae	urdamycin	*urd*	*tcm*/*act* probes	X87093
S. nogalater	nogalamycin	*sno*	*act* probe	Z48262
S.argillaceus	mithramycin	*mtm*	*act* probe	X89899

*Complementation

Table I.2: Genetic Engineering of Polyketide Synthase for Novel Aromatic Polyketides

The production of the novel polyketide antibiotic mainly depends upon the gene organization in the PKS cluster and the specificity of the various components of the PKS cluster. Understanding of the *Streptomyces* genetics and the development of the molecular biological techniques and tools to further explore the genetic potential of the PKS cluster paved the pathway of PKS engineering leading to the combinatorial biosynthesis of the new polyketides. The metabolic pathway engineering and the structural diversity obtained thus are the two significant achievements as far as the field of combinatorial chemistry is concerned. To achieve this diversity there is one

thing which if of prime importance is the host vector expression and chain length plays a very important role. The ideal requirements here are the genetically well characterized host in which the heterologous genes can be easily introduced, expressed and then post translationally modified properly. In addition, a strong promoter is often preferred to ensure efficient expression of the recombinant PKS, and an inducible or regulated promoter should have the added benefit of minimizing the potential toxicity of heterologous proteins to the host. The *S. coelicolor*-pRM5-CH999-based host-vector system, developed by Hopwood and co-workers (Garwin et. al., 1980; Summers et. al., 1995a), is one of the best systems used in engineered biosynthesis. *S. coelicolor* CH999 is derived from *S. coelicolor* A3(2) by deleting the entire *act* gene cluster to reduce endogenous polyketide biosynthesis. pRM5, a *Streptomyces-E. coli* shuttle vector, contains a divergently paired *actI* and *actIII* promoter whose activity is developmentally regulated by the ActII-ORF4 protein. More recently, Ziermann and Betlach (Biothechniques, 1999) have engineered several *S. lividans* equivalents of *S. coelicolor* CH999, such as *S. lividans* K4–114 and K4–155 (Hopwood 1997), which can be transformed with pRM5-based plasmids with superior frequency. Similarly to obtain the particular chain length the KS α/KS β from the different polyketide biosynthetis cluster can be used. KS α/KS β from act minimal PKS(Hopwood 1993; Garwin et. al., 1980; Summers et. al., 1995a), frenolicin [Hopwood 1993; Leadlay and McDaniel 1997), and graniticin has been used to produce the engineered synthesis of octaketide, of which KS α/KS β from the minimal PKS of frenolicin is the only one known so far to produce a nonaketide (Hopwood 1993; Kao et. al., 1996; Leadlay and McDaniel 1997). For the production of decaketides, one has the choice of KS α/KS β from the minimal PKS of dauxorubicin (Kao et. al., 1994a; Leadlay et. al., 1994; Pieper 1995), jadomycin (Leadlay et. al., 1994), mithramycin (Hu et. al., 1994; Kao et. al., 1994b; Oliynyk et. al., 1996) or tetracenomycin (Hopwood 1993; Alvarez et. al., 1996; Leadlay et. al., 1994).

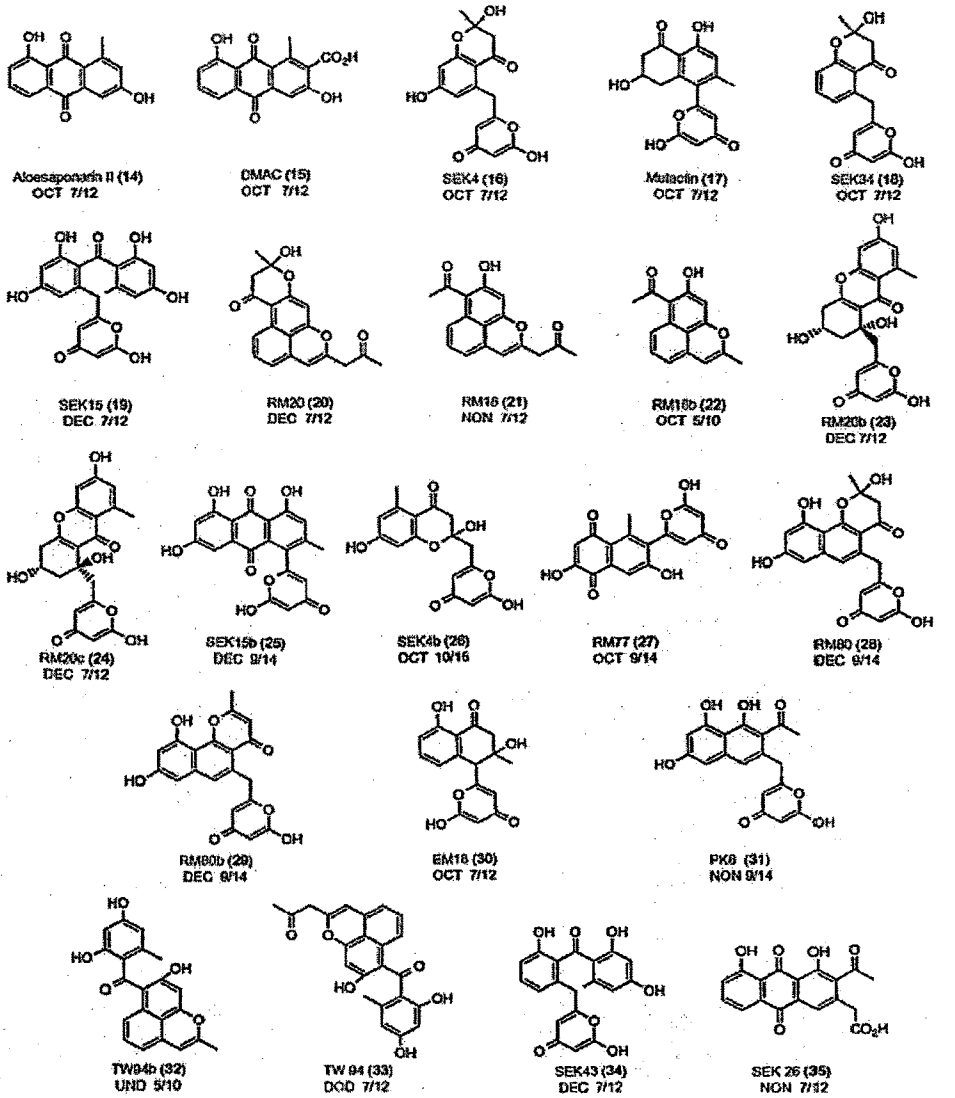

Figure. I.12: Novel polyketides produced by recombinant *S. coelicolor* CH999 strains containing plasmids related to pRM5 and carrying genes encoding various combinations of type II PKS subunits from the same or different *Streptomyces* species. Under the name of each metabolite is given the number of acetyl residues needed to build the carbon chain (OCT, octaketide; NON, nonaketide; DEC, decaketide; UND, undecaketide; DOD, dodecaketide), and the carbon atoms (counting from the carboxyl end) between which the first ring-closure occurs.

As described above, analysis of the polyketides produced by many recombinants carrying different combinations of PKS subunits led to a series of conclusions about

PKS programming. This retrospective analysis could be converted into a prospective set of "design rules" with enough predictive power to generate novel molecules "to order", and at least three have already been made: SEK43 (**34**) and SEK26 (**35**),(McDaniel et. al., 1995) as well as PK8 already mentioned (Kramer et. al., 1996). The design rule is basically made to understand the relevance of the heterologous proteins in the biosynthesis of the new compound while following a particular set of process. In an early attempt to identify the minimum set of polyketide synthase (PKS) components required for in vivo biosynthesis of aromatic polyketides, Khosla et. al in 1994 did an experiment in which the combinations of genes encoding subunits of three different aromatic PKSs--act from *Streptomyces coelicolor* A3(2) (an actinorhodin producer), fren from *Streptomyces roseofulvus* (a frenolicin and nanaomycin producer), and tcm from *Streptomyces glaucescens* (a tetracenomycin producer)--were expressed in a recently developed *Streptomyces* host-vector system. The "minimal" components (ketosynthase/putative acyltransferase, chain length-determining factor, and acyl carrier protein) were produced with and without a functional polyketide ketoreductase and/or cyclase, and the polyketide products of these recombinant strains were structurally characterized. Several previously identified polyketides were isolated in addition to two previously unidentified polyketides, dehydromutactin and SEK 15b (Khosla et. al., 1994). The structural feature here includes the carbon chain lengh, ketoreduction cyclizations – terminal as well as the first ring and second ring cyclisation as well as the aromatization. These processes are governed by the different components of the PKS cluster and each has been assigned the particular task. The change in the component can be achieved by either changing the host or bringing in the engineered heterologous proteins. These can then result in the variation in the structure of the molecule.

VII. Role of tailoring enzymes in the biosynthesis and their role in synthesizing new molecules:

With the advanced molecular biology techniques available and the increased understanding of the *Streptomyces* genetics more and more attempts are being made to generate the novel compounds. The approaches are also varying with the increase

in the understanding of the mechanistic properties of the components involved in the polyketide biosynthesis. As we discussed earlier there are set rules for the synthesis of the polyketide molecule and the chain length, first cyclisations, second cyclisations and the aromatizations play a big role in the biosynthesis. The engineering of these enzymes was also looked for obtaining the modification in the polyketide molecule.

The work done by Moore in 1999 can be sighted here which exploits the role of cyclase in producing the novel antibiotic The single recombinant expressing the *Streptomyces coelicolor* minimal *whiE* (spore pigment) polyketide synthase (PKS) is uniquely capable of generating a large array of well more than 30 polyketides, many of which, so far, are novel to this recombinant. The characterized polyketides represent a diverse set of molecules that differ in size (chain length) and shape (cyclization pattern). According to them the minimal PKS enzyme complex must rely on the stabilizing effects of additional subunits (i.e., the cyclase *whiE*-ORFVI) to ensure that the chain reaches the full 24 carbons and cyclizes correctly. This dramatic loss of control implies that the growing polyketide chain does not remain enzyme bound, resulting in the spontaneous cyclization of the methyl terminus. This way they found that the dodecaketide TW93h possesses a unique 2,4-dioxaadamantane ring system and represents a new structural class of polyketides with no related structures isolated from natural or engineered organisms, thus supporting the claim that engineered biosynthesis is capable of producing novel chemotypes (Shen et. al.,1999).

Other that these enzymes, there are several other enzymes which are involved in the terminal modification of the polyketide molecule which includes the enzymes involved in the glycan biosynthesis and its modification. The glycosylation of the polyketide molecule as discussed earlier provides the specificity and imparts the potency to the molecules. The enzymes involved here are glycosyl transfrases, NDP-glycosyl dehydratases, and methyl transferases etc. The sugar decoration of the polyketides is a very complex but important phenomenon which goes hand in hand with the polyketide backbone synthesis.

Rohr and Mendez in 2006 published a research article extensively describing the involvement of the four glycosyl transferases in the synthesis of the anti tumor Chromomycin A3. Chromomycin consists of a tricyclic aglycone with two aliphatic side chains and two O-glycosidically linked saccharide chains, a disaccharide of 4-O-acetyl-D-oliose (sugar A) and 4-O-methyl-D-oliose (sugar B), and a trisaccharide of D-olivose (sugar C), D-olivose (sugar D), and 4-O-acetyl-L-chromose B (sugar E). The chromomycin gene cluster contains four glycosyltransferase genes (cmmGI, cmmGII, cmmGIII, and cmmGIV). They independently inactivated these genes which generated the mutants C60GI, C10GII, C10GIII, and C10GIV. Mutants C10GIV and C10GIII produced the known compounds premithramycinone and premithramycin A1, respectively, indicating the involvement of CmmGIV and CmmGIII in the sequential transfer of the sugar residues in the triassacharide chain and Mutant C10GII produced two new tetracyclic compounds lacking the disaccharide chain at the 8 position, named prechromomycin A3 and prechromomycin A2; this way a new biosynthetic pathway was proposed for the glycosylation events in chromomycin A3 biosynthesis (Menendez et. al., 2006).

In the recent findings the 'hybrid glycosylation' has also been tried and attempted successfully. This is a process for producing a hybrid glycosylated product by transferring one or more sugar moities to an aglycone template, the process comprising : (a) constructing a plasmid based gene cassette which contains nucleic acid encoding sugar synthesis genes sufficient to direct the synthesis of a specific activated sugar wherein at least some of said sugar synthesis genes in the cassette are flanked by restriction sites; (b) transformation of microorganism host cells: (i) with said plasmid based gene cassette; and (ii) with nucleic acid encoding a glycosyltransferase (GT) ; and, (c) providing an aglycone template to the transformed microorganism, or allowing its endogenous generation, so that the GT transfers one or more sugar moieties to the aglycone template in order to produce a hybrid glycosylated product.

Work done by Pageni et. al., (2008) suggests a successful engineering of a deoxysugar pathway to generate novel hybrid macrolide derivatives; here two sugar biosynthetic cassette plasmids were used to direct the biosynthesis of a deoxyaminosugar. The pOTBP1 plasmid containing TDP-glucose synthase (desIII),

TDP-glucose-4 4-aminotransferase (gerB) and 3,5-epimerase (orf9) was transformed, 6-dehydratase (desIV), and glycosyltransferase (desVII/desVIII) was constructed and transformed into *S. venezuelae* YJ003, a strain in which the entire gene cluster of desosamine biosynthesis is deleted. The expression plasmid pOTBP3 containing in into *S. venezuelae* YJ003- OTBP1 to obtain *S. venezuelae* YJ003-OTBP3 for the production of 4-amino-4,6-dideoxy-L-glucose derivatives. The crude extracts obtained from *S. venezuelae* ATCC 15439, *S. venezuelae* YJ003, and *S. venezuelae* YJ003-OTBP3 were further analyzed by TLC, bioassay, HPLC, ESI/MS, LC/MS, and MS/MS. The results of our study clearly shows that *S. venezuelae* YJ003-OTBP3 constructs other new hybrid macrolide derivatives including 4-amino-4,6-dideoxy-L-glycosylated YC-17, methymycin, novamethymycin, and pikromycin from a 12-membered ring aglycon (10-deoxymethynolide, 1) and 14-membered ring aglycon (narbonolide, 2) (Pageni et. al., 2008).

Recently the studies are being done to understand the mechanism of tailoring enzymes like NDP- glycosylases, glycosyltransferases and ketoreuductases (Baerga-Oritz et. al., and O'Hare et.al., 2006). One such example is the TypeI coded Spinosyn which is gyclosylated antibiotic which contains O-methylrhamnose as sugar residue which is required for its bioactivity. Sheehan et al., in 2006, engineered the loading module of the spinosyn PKS by replacing it with the avermectin and erythromycin loading modules, and fed the recombinants a variety of carboxylic acid starter analogs. This approach generated new molecules.

Reeves et. al., in 2004, constructed a hybrid PKS operon containing the first two subunits from the chalcomycin PKS, *chmGI-II*, and coupled it with the last three subunits from the spiramycin PKS, *srmGIII-V*. To assure functional protein-protein interaction between ChmGII and SrmGIII, they replaced the ChmGII C-terminal interpeptide docking site with that of SrmGII. They inserted the strong tylosin PKS promoter *tylGIp* in front of the hybrid PKS pathway in a pSET152-derived vector and introduced it into the phi C31 *attB* site of a strain deleted for the *tylGI-V* PKS genes.

Several derivatives of erythromycin and two glycopeptides are in current clinical use, and three glycopeptides are in clinical development. Many pharmaceutical companies

have invested years of product development activities in macrolides , aromatic polyketides and glycopeptides, so the hurdles for developing additional molecules from these classes are very high. Daptomycin, by contrast, is the first acidic lipopeptide approved for clinical use, so the lipopeptide class offers an opportunity for molecular engineering coupled with medicinal chemistry to generate second- and third-generation molecules.

The present study was also carried out keeping in view the need for new antibiotics which requires the new clusters for mixing and matching of the newly found genes with older ones in order to generate new molecules.

VIII. About peptide antibiotics:

Peptaibols and the related peptaibiotics are linear, amphipathic polypeptides. More than 300 of these secondary metabolites have been described to date. These compounds are composed of 5-20 amino acids and are generally produced in microheterogeneous mixtures. Peptaibols and peptaibiotics with unusual amino acid content are the result of non-ribosomal biosynthesis. Large multifunctional enzymes known as peptide synthetases assemble these molecules by the multiple carrier thiotemplate mechanism from a remarkable range of precursors, which can be N-methylated, acylated or reduced. Peptaibols and peptaibiotics show interesting physico-chemical and biological properties including the formation of pores in bilayer lipid membranes, as well as antibacterial, antifungal, occasionally antiviral activities, and may elicit plant resistance. The three-dimensional structure of peptaibols and peptaibiotics is characterized predominantly by one type of the helical motifs alpha-helix, 3(10)-helix and beta-bend ribbon spiral (Pienkowski, et. al., 2005).