

LITERATURE REVIEW

Antibiotics

The term antibiotic was defined by Vullemin in 1889 and redefined by Wakesman in 1945 as chemical substances of microbial origin which exerts antimicrobial activity in small amounts. Though penicillin was the first antibiotic to be discovered, systematic search of antibiotics was first done by Gratia and Bath in 1924, which resulted in discovery of Actinomycetin in strains of *Actinomycetes*, also known as soil bacteria. *Streptomyces* belong to this group and produces 70% of antibiotics in therapeutic use. This makes *Streptomyces* industrially important (Alekhshun, 2005).

In the early 1940s, the industrialization of penicillin production was quickly followed by the successful isolation and development of a large number of antibiotics that have led to the major classes of antibiotics in use even to this day (Watve et al., 2001) namely, the tetracyclines, macrolides, aminoglycosides, cephalosporins, chloramphenicol, glycopeptides and rifamycins³ (Fig.1). As the majority of existing compounds originated from bacteria, the term 'antibiotics' has become almost synonymous with the more inclusive term of antibacterial agents (Yu et al., 1999).

Antibacterial drugs have been the most effective of all medicines. Their success is reflected by their continued use and the decrease in morbidity and mortality from bacterial infections over the past 50 years. In recent years, however, the increase in the number of multidrug-resistant bacteria has led to the prediction that we are reentering the pre-antibiotic era (Zahner et al., 1995). In reality, the situation will be far worse because today's bacterial strains are not only resistant to commonly available antibiotics, but more importantly may also have acquired virulence genes (Powers, 2003). As a result, even commonly occurring bacteria have been transformed into invasive and toxin-producing pathogens. This precarious setting has been worsened by lowering of antibacterial drug discovery efforts at large pharmaceutical companies.

A brief history of antibiotic development

The antibiotic era involved the work of several pioneering scientists. In 1929, Alexander Fleming, of St. Mary's Hospital in London, published a paper in the *British Journal of Experimental Pathology* describing the isolation of penicillin from *Penicillium* mold and its potential use in surgical dressings; Howard Walter Florey and colleagues at the University of Oxford later demonstrated penicillin's *in vivo* antibacterial efficacy. In 1935, Gerhard Domagk at I.G. Farben (Frankfurt main, Germany) synthesized the first synthetic antibacterial drug, Prontosil, whose active ingredient sulfanilamide (the prototype for all sulfa drugs) was later identified by the Pasteur Institute's Daniel Bovet. And in 1939, René Dubos in Oswald Avery's laboratory at the Rockefeller Institute for Medical Research in New York identified tyrothricin a mixture of the peptide antibiotics tyrocidin and gramicidin D from soil bacteria - widely regarded as the first antibiotic to be established as a therapeutic substance. In the early 1940s, the industrialization of penicillin production was quickly followed by the successful isolation and development of a large number of antibiotics that have led to most of the major classes of antibiotics in use even to this day, namely, the tetracyclines, lipopeptides, macrolides, aminoglycosides, cephalosporins, chloramphenicol, glycopeptides and rifamycins (Grunewald et al., 2004, Grunewald and Marahiel, 2006, Kopp et al., 2006, Miao et al., 2006).

Resistance against antibiotics

Infectious disease continues to be concern for drug industry, even after so much of vaccination program and use of antibiotics. Microorganisms resistant to antibiotics in therapeutic use arise and resistant determinants spread across different genera partly due to inappropriate management of antibiotic therapy and mainly due the fact that microorganisms develop resistance against any antibiotic which is introduced in due course of time. This is due to the fact that any cycle of infection involves large number of bacteria, rapid generation time and the intrinsic rate of mutation of about one in 10^7 - 10^9 which allows a pool of 10^{10} bacteria to have a mutation in a thousand loci, one of these mutation may confer resistance to an antibiotic allowing these bacteria to survive. This resistant bacterium multiplies to give rise to resistant population and spread of resistance

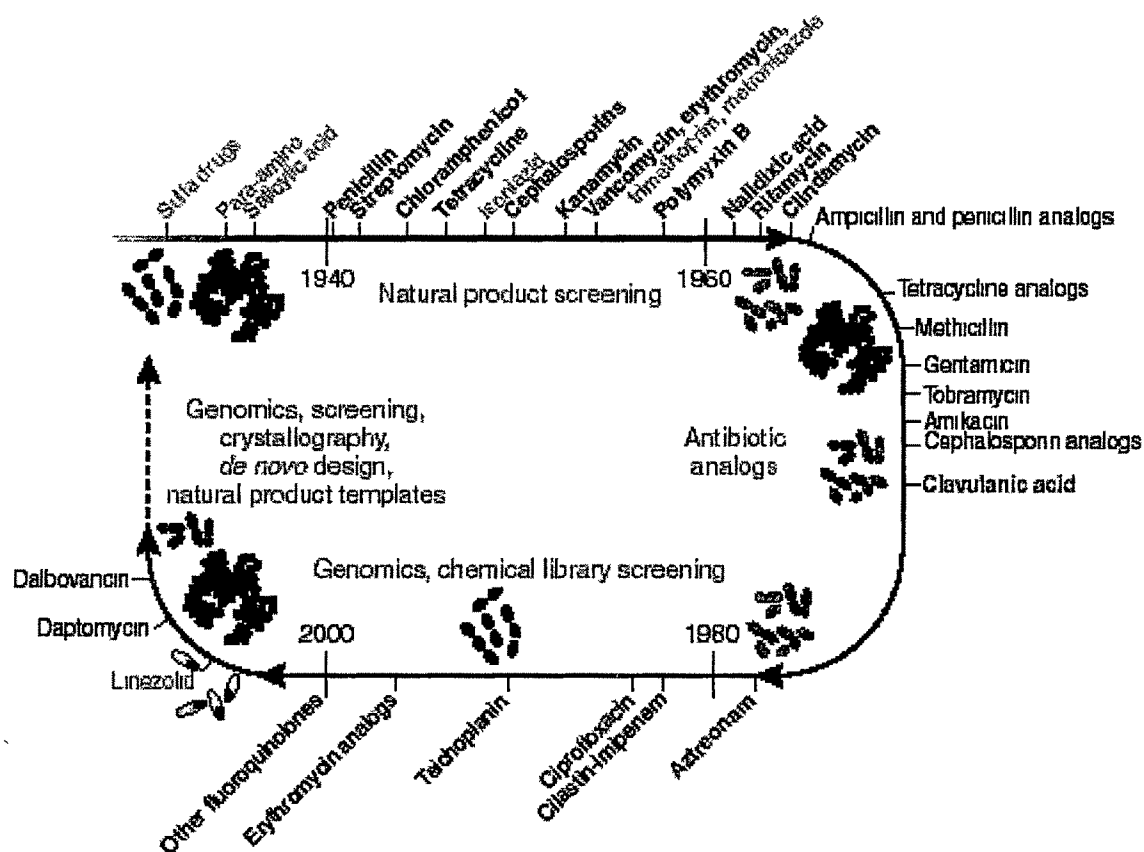


Fig.1: Antibiotic Discovery over the years.

genes to other genera by gene transfer mechanism. So the competition of looking for/ developing new bioactive compound by humans and acquiring resistance against bioactive compound by bacteria continues.

Just imagine life without antibiotics. It would be like it was 100 years ago, when pneumonia and tuberculosis were the most frequent causes of death, and the risk of infection turned a simple appendectomy into a dangerous operation. Luckily we do have antibiotics. However, they are becoming increasingly ineffective. Doctors are more and more frequently confronted with infections they cannot do anything about because the bacteria have become resistant. This has dire consequences for patients. Many end up living with a chronic infection for years on end, some are forced to become amputees and

yet others succumb to the infections. The crisis affects people in both industrialized and developing countries. In the US and the UK, the bug *Staphylococcus aureus* is wreaking havoc (Bischoff et al., 2005). Forty to fifty per cent of infections that people contract in hospitals are resistant to more than one antibiotic. The developing countries are groaning under the burden of tuberculosis, which claims the lives of two million victims throughout the world every year. The increase in multiresistant TB is especially alarming; treating it costs 100 times more than treating the regular form, making a cure unaffordable for many people in developing countries. Despite this, many pharmaceutical companies have stopped developing antibiotics. They see the financial risk as too big and potential profits too skimpy. This has led to very few new drugs for fighting bacterial infections being launched in recent years. A survey of 11 large pharmaceutical companies revealed that of 400 substances they were developing; only five were antibacterial drugs. What can be done about the resistance crisis? One thing is needed for sure: new drugs! And one of the important natural sources of new structures is mining it from microorganisms found in different niches.

Streptomyces

Streptomyces are Gram-positive, aerobic, filamentous soil bacteria that undergo morphological differentiation during their life cycle. They produce extensive branching vegetative (substrate) mycelium and aerial mycelium bearing chains of arthrospores. The *Streptomyces* are able to utilize a wide range of organic compounds as a carbon source, including complex biological materials, such as cellulose and lignin, and can also utilize an inorganic nitrogen source. *Streptomyces* are common in soil, but also found in composts, fodder and aquatic habitats. Due to their characteristic life cycle, they are good survivors under the fluctuating growth conditions predominating in nature.

On agar plates, they form lichenoid, leathery or butyrous colonies. The GC-content of the DNA is 69-78 % (Wright and Bibb, 1992). L-diaminopimelic acid is the characteristic compound present in the cell wall peptidoglycan of *Streptomyces* (Lechevalier and Lechevalier, 1970). They normally occur as spores, but in the presence of sufficient moisture and nutrients, the spores can germinate and form vegetative mycelium

(Williams et al., 1997). In response to environmental signals, such as shortage of nutrients or water, the process of differentiation begins, and spores resistant to desiccation and starvation are formed again. At the same time, the production of pigments and other secondary metabolites is initiated. Apart from diffusible pigments produced, mycelium and spores can also be pigmented. Antibiotics and other bioactive compounds are also synthesized as secondary metabolite (Wietzorrek and Bibb, 1997).

Streptomyces research group is seeking knowledge of *Streptomyces* genetics, molecular biology and physiology, which will provide background to their use for applied objectives as well as yielding a wealth of fundamental knowledge of prokaryotic genetics with wide implications (Alvarez, 1996).

Secondary metabolites

Microbial secondary metabolites are substances that are not needed for the growth or other essential processes in the cell (Vining, 1990). Secondary metabolites are mainly produced by microbial genera inhabiting soil and undergoing morphological differentiation (Wang et al., 2003), such as actinobacteria, bacilli and fungi (Vining, 1990). There are over 23,000 known microbial secondary metabolites, 42 % of which are produced by actinobacteria, 42 % by fungi, and 16 % by other bacteria (Lamrani et al. 2000). The *Streptomyces* are very potent producers of secondary metabolites. Out of the approximately 10,000 known antibiotics, 45-55 % is produced by *Streptomyces* (Demain 1999, Lamrani et al. 2000). One of the important classes of antibiotics is polyketide. They are next in line to penicillin as far as importance as antibiotics are concerned, but as far as number of drugs of medical importance is concerned polyketides are 2/3rd in number as compared to 1/3rd being rest of the compounds together.

Polyketides

Microorganisms make a wealth of unusual metabolites that have a secondary role in the organism's ontogeny, such as self-defense, aggression, or even communication, as the need arises (Anderson, 2001). These compounds often have biological activity valuable to humankind. Polyketides are such a group of secondary metabolites, exhibiting remarkable diversity both in terms of their structure and function. Polyketide natural products are known to possess a wealth of pharmacologically important activities, including e.g. antibacterials (streptomycin, tetracycline, chloramphenicol), antifungal (nystatin), antiviral (tunicamycin), antiparasitic (ivermectin), immunosuppressive (rapamycin), antitumor (actinomycin, mitomycin C, anthracyclines), enzyme inhibitory (clavulanic acid), diabetogenic (bafilomycin, streptozotocin) these are a few more of the thousands of polyketides discovered so far (Wang et al., 2000). These metabolites are ubiquitous in distribution and have been reported from organisms as diverse as bacteria, fungi, plants, insects, dinoflagellates, mollusks and sponges. The wide spectrum of activity of polyketides makes them economically, clinically and industrially the most sought after secondary metabolite. Polyketides are very diverse in structure and may be divided into four Classes, aromatics (e.g., doxorubicin and tetracycline), macrolides (e.g., erythromycin and rapamycin), polyethers (e.g., monensin and salinomycin), and polyenes (e.g., amphotericin and candicidin), the last three are grouped together as complex polyketides. Some examples of the myriad chemical structures a polyketide can take:

Although the structures of polyketides vary enormously, they are all synthesized, in their initial stages, by a mechanism that is very similar to fatty acid biosynthesis: simple acyl precursors such as acetyl and malonyl units are condensed in a sequential fashion to give a long carbon chain, catalyzed by the polyketide synthase (Hopwood, 1993). Polyketide synthase (PKS) are a family of enzymes that catalyses the biosynthesis of structurally diverse and pharmaceutically important class of natural products just described, the polyketides. They are divided into two classes following the convention of fatty acid synthases (FASs), according to their enzyme architecture and gene organization. Type I PKSs are multifunctional proteins consisting of domain for individual enzyme activities and have been found in bacteria as well as in fungi and plants. Type II PKSs are multienzyme complexes consisting of discrete proteins that are largely monofunctional

and have so far only been found in bacteria. Antibiotics other than polyketide (Dandliker et al., 2003, Demain, 1999) like glycopeptide (Bister et al., 2003, Borisova et al., 2006, Donadio et al., 2005, Weist et al., 2002) and lipoproteins (Coeffet et al., 2006), Long – chain amino acid (Brady and Clardy 2000), azoxy antibiotic (Garg et al., 2002), orthosomycin antibiotic (Weitnauer et al., 2001) are also important.

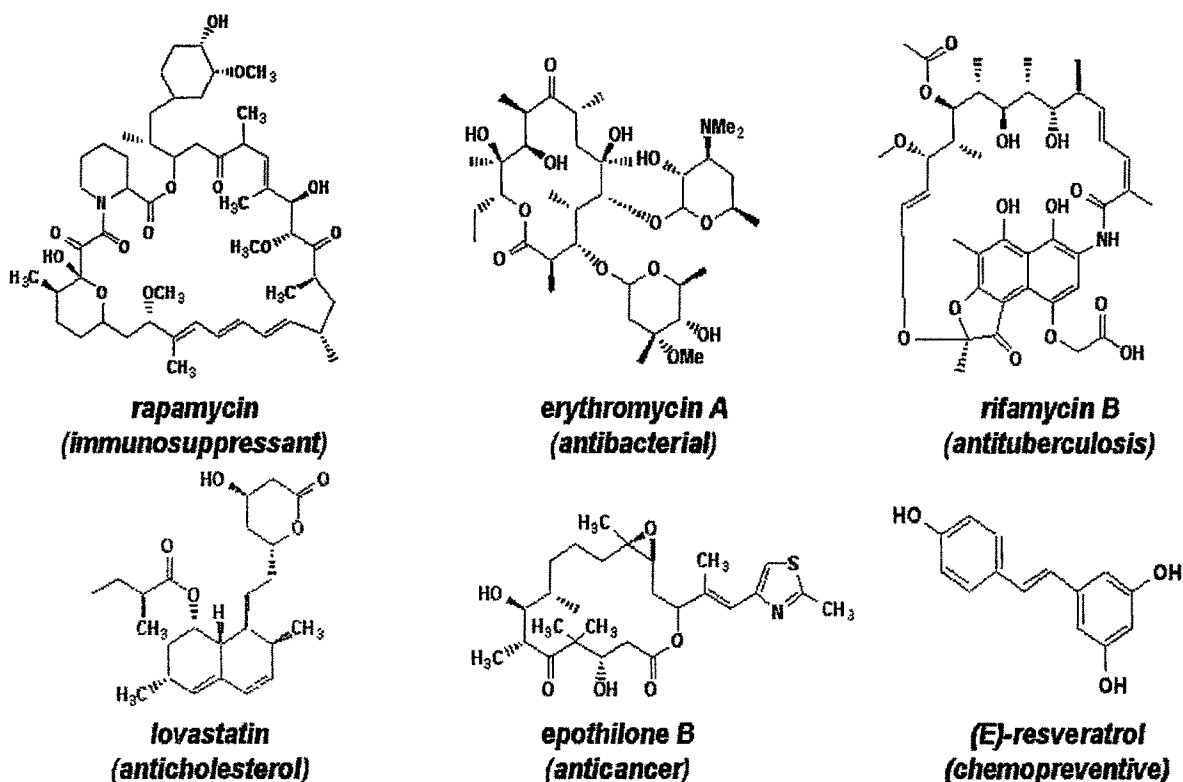


Fig.2: Polyketide Structures

The differences between polyketide and fatty acid biosynthesis (and also among diverse polyketide biosyntheses) are in the number and type of acyl precursors used, the extent and position of keto-group reductions, and the cyclization pattern of the products (subsequent post-PKS modifications add to the structural variety observed). Thus fatty acid and polyketide biosyntheses are mechanistically related, and often the same precursor molecules are used. The first PKS cluster to be studied was from *Streptomyces*

coelicolor A3(2) (Adamidis et al., 1990). It is the genetically most studied member of the genus, synthesizes at least two aromatic polyketides at different stages during its development cycle on solid media. The blue pigmented antibiotic, actinorhodin is synthesized just as the vegetative mycelial cells begin to differentiate and enter the reproductive phase of development (Bibb, 2005). The other, grey-pigmented polyketide is synthesized near the end of colony differentiation and is associated with the spore walls. The *act* and *whiE* PKS gene sets (*act* for actinorhodin and *whiE* for spore pigment, so called because the mutants have white spores) have been cloned and sequenced (Bartel et. al., 1990). The essential members of both gene sets encode the three subunits that make up the minimal-PKSs, so called because they are all that are required for assembly of the polyketide chain (McDaniel et. al, 1994). The *act* and *whiE* minimal-PKSs are very similar to each other and to those involved in synthesis of aromatic polyketides from related *Streptomyces* spp. and hence the minimal PKS from actinorhodin (*actI*) have been used as probe to identify carriers of PKS.

The deduced products of the minimal- PKS genes also have a moderate similarity to subunits of the *Escherichia coli* FAS. Thus, the FASs and PKSs are related not only in their mechanism of carbon chain assembly but also in their primary sequences (Arjo and Dannert, 2003). It seems probable that they have evolved from a common origin after an early gene duplication. The minimal- PKS subunits are a β -ketoacyl carrier protein synthase, required for condensation of the acyl units to form the growing polyketide chain; a polyketide chain-length factor which has a high degree of homology to the ketoacylsynthase but lacks the active site cysteine residue; and an acyl carrier protein (ACP), which serves a dual purpose in receiving the malonyl extender units prior to condensation and in holding the growing carbon chain after each condensation (O'Hare et. al., 2006).

Recently, an enzyme subunit in *S. coelicolor* has been identified which is potentially required for both fatty acid and polyketide biosynthesis within the same organism (Emmanuel et al., 2003). This is the malonyl coenzyme A (CoA):ACP acyltransferase, thought to be responsible for charging the ACP subunit of the FAS and both PKSs prior

to condensation (Baltz, 1998, 2006). (Similar observations have been made in *Streptomyces glaucescens* (Bechthold et al., 1995). The malonyltransferase appears to provide a tangible link between the FAS and the PKSs that are expressed within the same organism. The subunits of the minimal-PKSs, on the other hand, appear to be prevented from forming any functional link between the synthases: a mutation in any of the *act* minimal-PKS subunit genes results in loss of actinorhodin production and is not complemented by the naturally expressed equivalent subunits from either the *fab* (encoding the FAS) or *whiE* gene set.

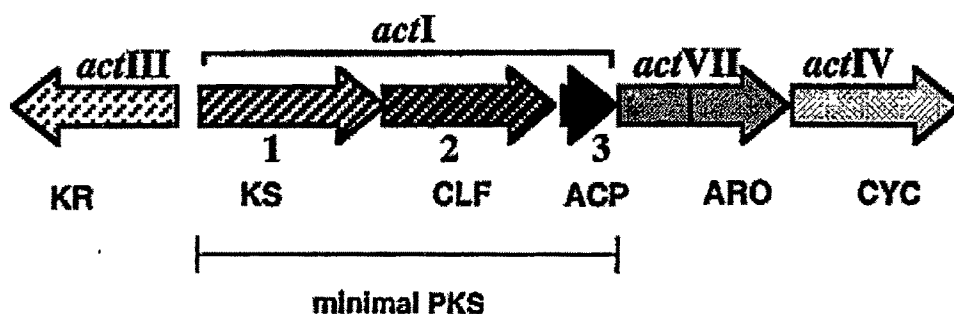


Fig.3: Minimal PKS

Advances in the genetic manipulation of *Streptomyces* species enabled the identification and cloning of entire biosynthetic gene clusters which contain gene encoding polyketide synthases. Subsequently combinatorial expression of these genes resulted in assignment of specific catalytic function to the individual genes.

Synthesis of polyketides

Polyketides are synthesized by sequential reactions catalyzed by a collection of enzyme activities called polyketide synthases (PKSs). These are large multienzyme protein complexes that contain a coordinated group of active sites (Weissman and Leadlay, 2005). The biosynthesis occurs in a stepwise manner from simple 2-, 3-, 4-carbon building blocks such as acetyl-CoA, propionyl CoA, butyryl-CoA and their activated derivatives (Bradley and Hertweck, 2001), malonyl-, methylmalonyl- and ethylmalonyl-CoA.

Various domains of polyketide synthase:

Component domains of polyketides consist of acyl-transferases (AT) for the loading of starter, extender and intermediate acyl units; acyl carrier proteins (ACP) which hold the growing macrolide as a thiol ester; β -keto-acyl synthases (KS) which catalyse chain

STARTER AND EXTENDER UNITS

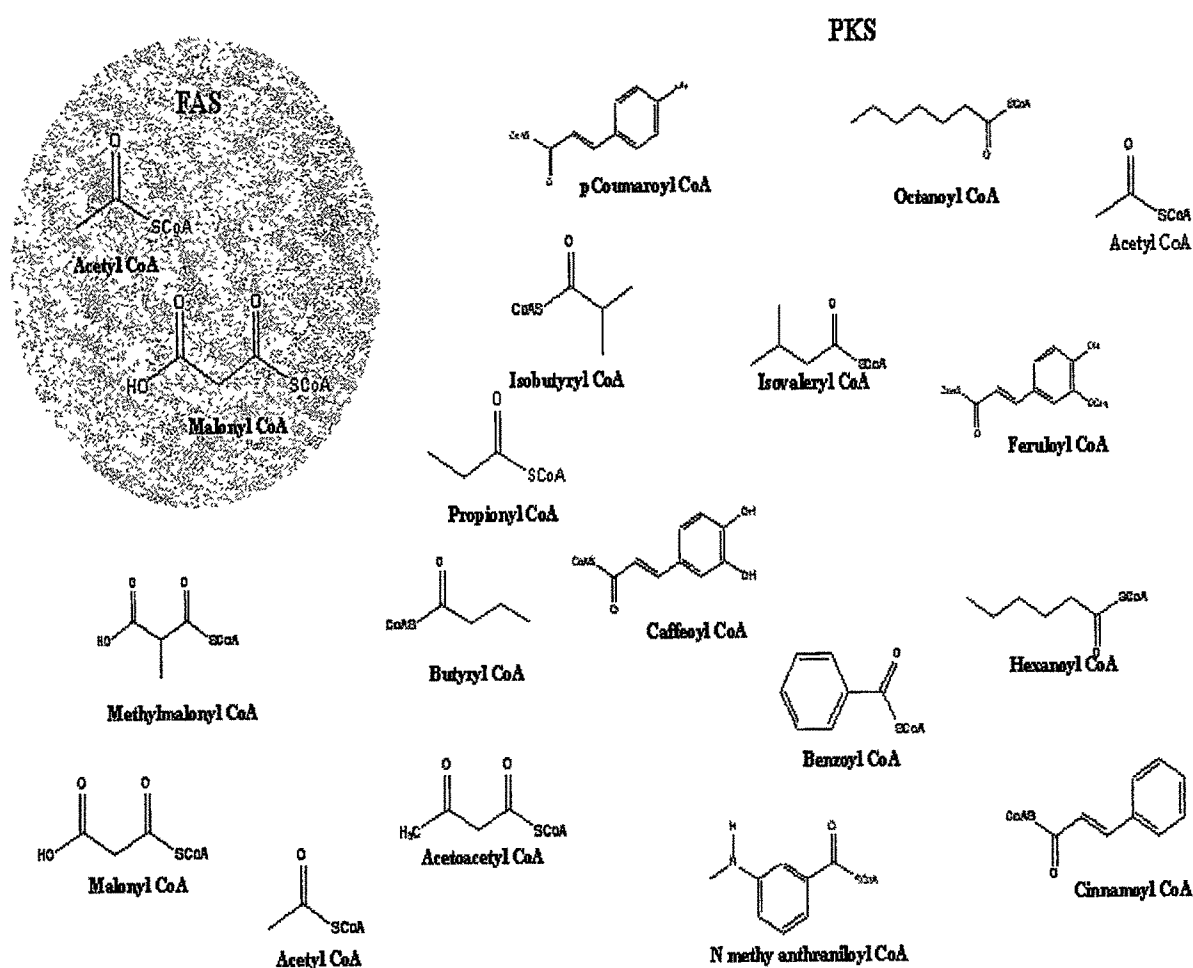
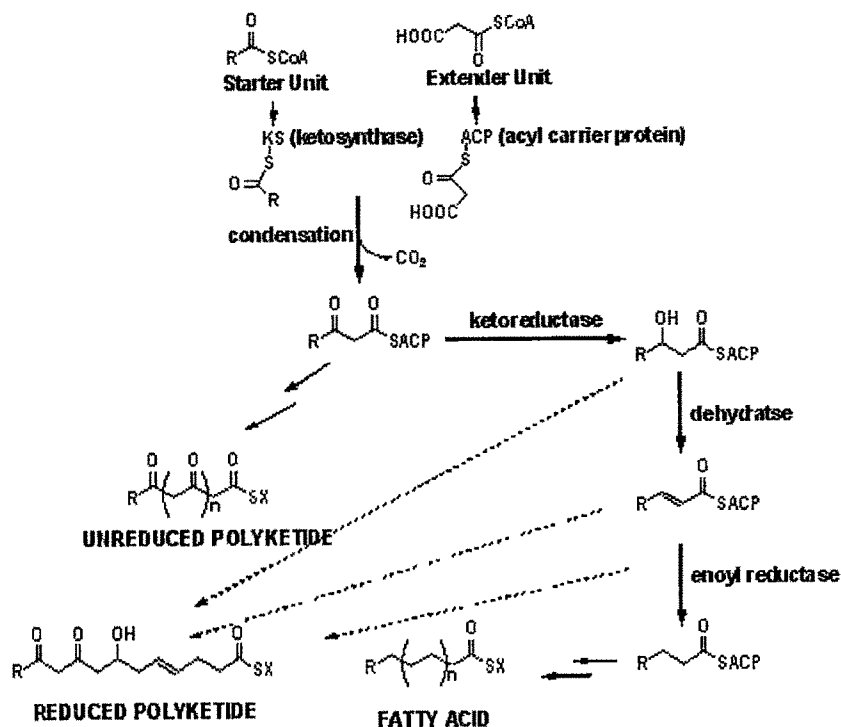


Fig.4: Structure of polyketide precursors.

CHEMISTRY OF POLYKETIDE SYNTHASES ROLES OF VARIOUS DOMAINS



extension; β -keto reductases (KR) responsible for the first reduction to an alcohol functionality; dehydratases (DH) which eliminate water to give an unsaturated thiolester;

Fig. 5. Role of various domains in polyketide synthases.

enoyl reductases (ER) which catalyse the final reduction to full saturation; and finally a thiolesterase (TE) to catalyse macrolide release and cyclisation.

Types of Polyketide Synthases (PKSs)

At least three architecturally different types of PKSs have been discovered in the microbial world.

Type I systems consist of very large multifunctional proteins which can be either processive (for example the unique modular systems responsible for synthesis of macrolides like erythromycin, rapamycin, rifamycin etc.) or iterative (for example the lovastatin nonaketide synthase). Iterative Type I synthases are analogous to vertebrate fatty acid synthases. These are typically involved in the biosynthesis of fungal polyketides such as 6-methylsalicylic acid and aflatoxin. These PKSs are large multidomain proteins carrying all the active sites required for polyketide biosynthesis.

The iterative Type II systems consist of complexes of mono-functional proteins exemplified by the actinorhodin PKS from *Streptomyces coelicolor* (Bentley et al., 2002, 2004.). In these synthases, active sites are distributed among several smaller, typically monofunctional polypeptides. Type II synthases catalyse the formation of compounds that require aromatization and cyclization, but not extensive reduction or reduction/dehydration cycles. These PKSs are analogous to bacterial Fatty Acid Synthases and are involved in the biosynthesis of bacterial aromatic natural products such as actinorhodin, tetracenomycin and doxorubicin.

Type III polyketide synthases are responsible for the synthesis of chalcones and stilbenes in plants and polyhydroxy phenols in bacteria. Chalcone synthase like proteins are comparatively small proteins with a single polypeptide chain and are involved in the biosynthesis of precursors for flavonoids. Unlike all other PKSs, these proteins do not have a phosphopantetheinyl (P-Pant) arm on which the growing polyketide chains are tethered.

Modular polyketide synthases

Modular PKSs constitute a unique class of Type I polyketide synthases. Each of these proteins consists of multiple active domains organised into modules. Each module is responsible for the construction of a carbon-carbon bond, via the decarboxylative condensation of a ketide extender unit with the growing polyketide chain, followed by a programmed reductive cycle. (Reid et al., 2003). In addition there is a loading module in most PKSs for obtaining the starter unit at the front of module 1 and a thioesterase

responsible for unloading the product at the end of the last module. The biosynthesis of 6-deoxyerythronolide B (6-dEB) has become the textbook example for modular polyketide synthases. In *Saccaropolyspora erythrae*, it is converted to the biologically active erythromycin (Hutchinson, 1999) by hydroxylation followed by addition of sugars. This example has been shown in figure 6 below.

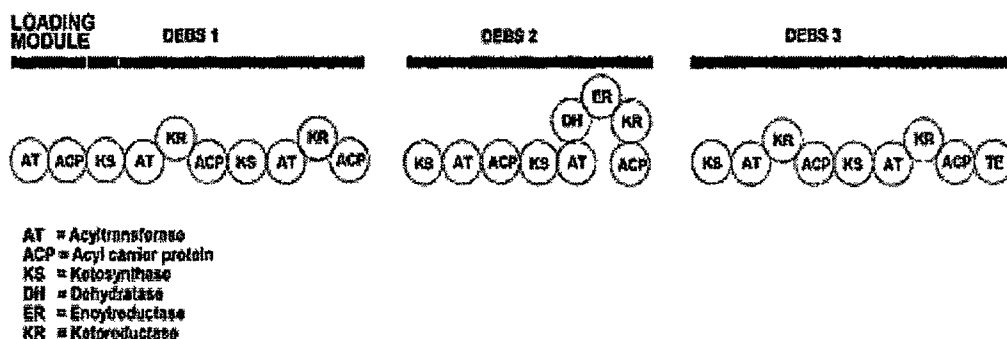
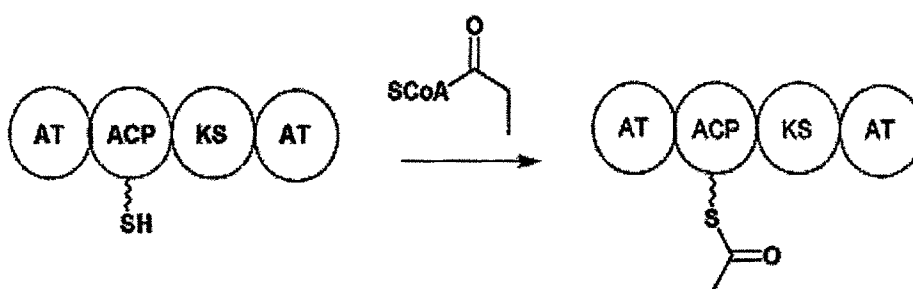


Fig.6: Modular PKS genes of Erythromycin biosynthesis.

First, a loading module consisting of an acyltransferase (AT) selects a propionate from propionyl Co-A and transfers the propionyl group to an acyl carrier protein (ACP).



The propionyl group is then transferred to a ketosynthase domain (KS). Subsequently, the polyketide chain is extended by condensation with methylmalonate (from methylmalonyl CoA) pre-loaded on the ACP domain of an extension module.

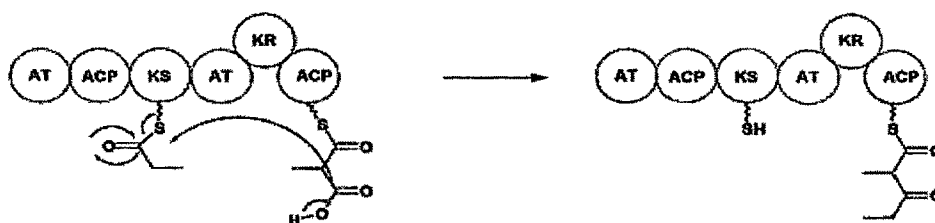
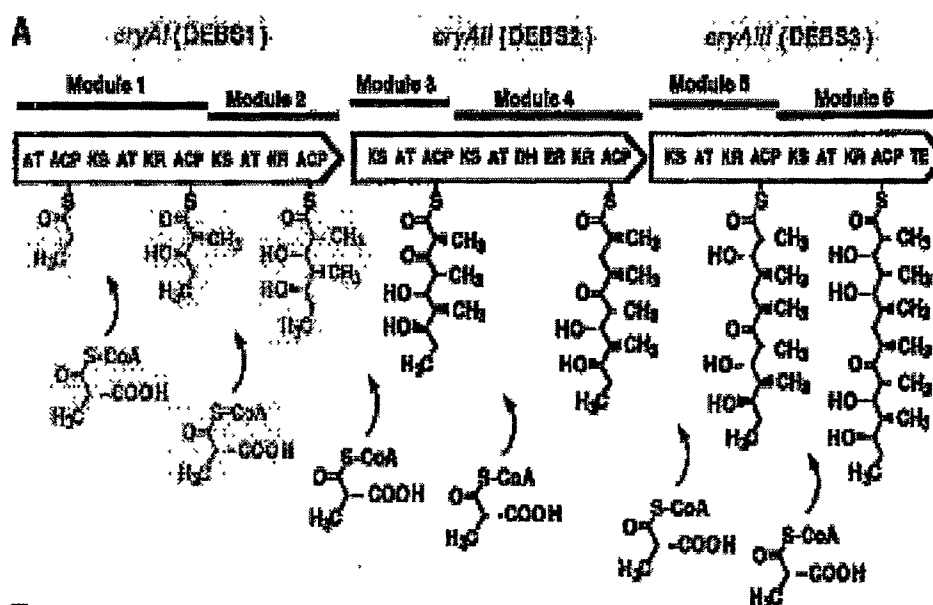


Fig.7: The Erythromycin PKS: Primary organization of the proteins



This process continues along the PKS and other domains such as the ketoreductase (KR) dehydratase (DH) and enoyl reductase (ER) domains can reduce each carbonyl group accordingly. Because the erythromycin PKS has one loading module and six extension modules, the result is a heptaketide chain. The polyketide is then released from the enzyme by a thioesterase domain and post PKS enzymes such as glycosyl and methyltransferases complete the biosynthesis.(Menzella et al., 2005)

Linker Hypothesis

Proposed first by Prof. Khosla (Khosla et al., 1999), this hypothesis states that linkers, which can be defined as the amino acid stretches joining various domains in a modular polyketide synthase, play a crucial role in the establishment of structural and functional assembly of these multimodular proteins (Katz, 1997). It is believed that these dynamic linkers establish communication by directing the correlated movements of various domains. Two categories of linkers have been postulated: linkers that connect covalently connected modules (intra-polypeptide linkers e.g., between modules 1 and 2 of DEBS) and linkers between modules that are present on two different polypeptides (inter-polypeptide linkers e.g., between modules 2 and 3 of DBES) (Menzella et al., 2006). The following figure explains the concept of linkers in a typical modular system:

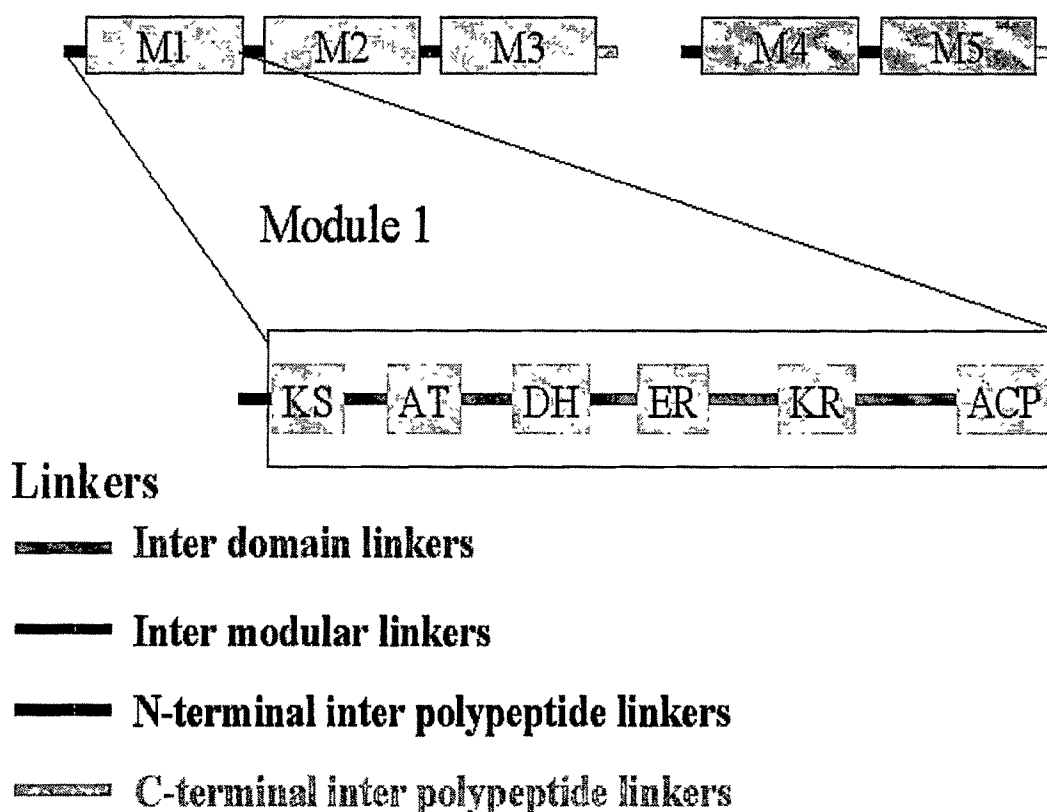


Fig. Linker Hypothesis

Source of antibiotics

Natural products research was a mainstay of the pharmaceutical industry for decades, and it was the key contributor to the 'golden age' of antibiotics discovery, the period from the 1940s to the 1960s when the vast majority of new antibiotic classes were introduced. Members of the actinomycetes group of bacteria, including the genus *Streptomyces*, have been a particularly fecund source of novel chemistry because of their extensive secondary metabolism, which enables them to compete against other microbes in the soil. Even today, new naturally occurring antibiotics are still being discovered - though not necessarily developed from natural product libraries that are decades old (Gerdes et al., 2002).

From the 1960s to the 1970s, the emergence of antibiotic resistant bacteria and their spread into the hospital environment added urgency and impetus to the search for new antibiotic compounds (Davies, 1994, Labischinski, 2001, McDevitt and Rosenberg, 2001). Most of the low-hanging antibacterial natural products fruit had already been picked from microbial fermentations, and after the late 1960s, screening programs were mostly rediscovering existing antibiotics. It thus appeared that the all 'magic bullets' had mostly been found.

Bacteria and fungi offer impressive arrays of secondary metabolites whose biosynthetic genes are typically clustered and thus readily amenable to genetic manipulation (Emmanuel et al., 2003). Combinatorial biosynthesis involves the exploitation of nature's synthetic capabilities in a combinatorial 'mix-and-match' fashion by interchanging natural product biosynthesis genes to create unnatural gene combinations (Bradley et. al., 2002, Cerdano et al., 2001). Resulting recombinant microorganisms synthesize novel or 'hybrid' metabolites due to the effect of new enzymes on the metabolic pathway (Hopwood et al., 1985). Consequently, novel molecules never encountered in nature can be generated and represent an important new source of chemicals for use in drug discovery (Cane et al. 1998; Hutchinson 1998). Pioneering studies by Hopwood, Khosla and coworkers on the actinorhodin polyketide synthase PKS in *Streptomyces coelicolor*

demonstrated the extraordinary utility of this approach in generating molecular diversity and in dissecting gene functions of complex biosynthetic processes (McDaniel et al. 1993, 1995). Numerous Type II iterative (or aromatic) PKSs have since been sequenced and employed in designing an assortment of hybrid aromatic metabolites (Hopwood 1997; Rawlings 1999). Each PKS contains a 'minimal' set of core subunits (the two β -ketoacyl:ACP synthase subunits KS α and KS β , an acyl carrier protein (ACP), and possibly a malonyl-CoA:ACP transacylase), which is required for *in vivo* polyketide biosynthesis. Additional PKS subunits, including ketoreductases, cyclases, and aromatases, are responsible for the processing of the nascent chain to form specific polyaromatic products. Nature only provides a few different core polyketide structures by varying the combination of the gene sets. Although well over 500 multicyclic aromatic polyketides have been characterized from actinomycetes, they belong to just a few common structural types that include benzoisochromanequinones, tetracyclines, angucyclines, anthracyclines, tetracenomycins and aureolic acids (Figure 1). Structural diversity within aromatic polyketides is dictated more so by tailoring enzymes during post-PKS events. Consequently, combinatorial biosynthesis with type II PKSs offers limited options for true molecular diversity (McDaniel et al, 1995). The majority of metabolites engineered with type II PKS genes represent a single structural class: namely, planar carbocyclic aromatics carrying pendent pyrone rings. Genetic engineering a diverse series of polyketides from a single recombinant unnaturally expressing the minimal *whiE* PKS from *Streptomyces coelicolor* on elucidating how metabolic diversity is naturally achieved by '*S. maritimus*' in relation to a diverse series of polyketides known as the enterocins and wailupemycins was demonstrated by Moore and Piel, 2000.

In each of these type II PKS systems, metabolic diversity is largely achieved (both naturally and unnaturally) by the absence of cyclases or aromatases (McDaniel et al., 1999). These enzymes stabilize the minimal PKS enzyme complex to ensure the proper assembly and folding of the highly labile polyketide intermediate, and their absence leads to the production of diverse series of polyketide products. The generation of diverse sets of metabolites from a single pathway (either engineered or natural) is pertinent to one of

the goals of combinatorial biosynthesis, and that is to extend natural product structural diversity for drug discovery.

Genetic Engineering for new antibiotic structure

Engineering molecular diversity with *whiE* (spore pigment) PKS genes

Two type II PKS gene clusters are present in the *S. coelicolor* genome, *act* (actinorhodin) (Fernández-Moreno et al. 1992) and *whiE* (spore pigmentation, Figure 2) (Davis & Chater 1990, Chakraborty and Bibb, 1997). As the chemical structure of the *whiE*-encoded spore pigmentation product is unknown, expression studies with the *whiE* genes were conducted to shed some light on the chemical nature of the corresponding spore pigment(s) (Yu et al. 1998; Shen et al. 1999). A series of expression plasmids containing the minimal *whiE* PKS (ORFs III-V) with and without various arrangements of the three *whiE* cyclases (ORFs II, VI, and VII) were constructed. These plasmids were introduced into the engineered host *S. coelicolor* YU105 (Yu & Hopwood 1995) from which the natural chromosomal copies of the *act* and *whiE* gene clusters were deleted. Analysis of the resultant pigmented polyketides indicated that the *whiE* spore pigment is probably a polycyclic dodecaketide that first cyclizes between carbons nine and 14 (Yu et al. 1998). Unfortunately, further structural information on the natural composition of the spore pigment was elusive. The results from this series of expression experiments, although missing its initial target, demonstrated that type II PKSs can be nearly as artful as its modular type I PKS cousins in creating polyketide molecular diversity (Shen et al. 1999). A large array of mono- and polycyclic aromatic polyketides of various sizes and shapes was uniquely produced by single recombinant carrying the minimal *whiE* PKS. Well over 30 polyketides were produced and analyzed by liquid chromatography-mass spectrometry (LC-MS), and the structures of eight were fully elucidated (Figure 3). This genetically engineered library is the largest and most complex of its kind described to date from a type II PKS system. Although dodecaketides (24 carbons) predominated the mixture, characterized products were observed as small as heptaketides (14 carbons). Each of the six identified dodecaketides carries quite different ring systems, especially that of TW93h

which possesses a unique 2,4-dioxadamantane moiety. This novel feature represents a new structural class of polyketides not described to date and clearly supports the claim that engineered biosynthesis is capable of generating novel chemotypes. The wide assortment of chain lengths and cyclization regiochemistries among the minimal *whiE* PKS products clearly indicates that the minimal PKS does not alone dictate chain length or the first cyclization event. This is in contrast to other systems previously studied, including the *act* and tetracenomycin minimal PKSs, whose principal recombinant products reflect the chain length of the parent molecule (Hopwood 1997). The minimal PKS, at least in the case of the *whiE* system, does not tightly control the size and shape of growing poly- β -carbonyl intermediate. Rather, the minimal PKS must rely on the stabilizing effects of the entire enzyme complex in order to ensure that the chain reaches its full length and correctly cyclizes. For instance, when the minimal *whiE* PKS was accompanied with the *whiE*-ORFVI cyclase, only single polycyclic product was formed (Figure 3) (Yu et al. 1998). This observation suggests that cyclases act as "chaperones" (Hutchinson 1999) in molding the highly labile poly- β -carbonyl intermediate into single product and away from the numerous metabolic options that are spontaneously available (Chandran et al., 2006).

The first natural type II PKS gene cluster, *enc* genes that is devoid of cyclases and whose product profile reflects this natural absence, thus confirming that the deletion of cyclases is the molecular basis for the generation of truncated and spontaneously cyclized polyketide products. The absence of cyclases from the minimal PKS enzyme complex either unnaturally (as in the engineered expression of the minimal *whiE* PKS) or naturally (as in the *enc* gene set) is in large part the basis for the observed molecular spontaneity of these systems. Large series of polyketide products were clearly evident in both complimentary, yet distinct, systems, which represent the most versatile engineered and natural type II PKSs investigated to date (Moore 1999). Furthermore, the small size and novel features of the newly established *enc* biosynthetic gene cluster provide the foundation for engineering hybrid expression systems with more typical type II PKS gene sets in the generation of diverse ranges of novel compounds for use in drug discovery. Finally, the work with the *enc* cluster represents the first expression of a marine bacterial

natural product in a terrestrial heterologous host (Moore 1999) and serves to validate that marine bacterial natural products can be produced in non-marine expression systems (Chopra et al., 2002, Corates et al., 2002).

Genetic Engineering and Polyketides

Drug discovery in the past 30 years is becoming a process which involves multidisciplinary science. It requires multi forked approaches which not only helps in generating new low molecular weight compounds but also helps in improving the potential of the old compounds. To date natural products have been important source of screening material and they are further enriched by genetically engineered bacteria that may produce novel metabolites. Other approaches like chemical synthesis, combinatorial chemistry, computational methods involving molecular graphics etc. help in further improving the newly discovered bioactive compounds.

There are various approaches to engineer microorganisms to produce new chemical compounds, one of them is combinatorial biosynthesis (Arjo Boer and Claudia, 2003). Combinatorial biosynthesis has been especially successful with bacterial polyketide synthases (PKSs), enzymes that, in essence, polymerize simple fatty acids into a myriad of chemical structures called “polyketides”. Genetic manipulation of type I and type II polyketide synthase genes has been successfully used and in addition genetic pathway of aminoglycoside and oligopeptide may also be manipulated to give novel metabolites.

The polyketides discovered so far number into the thousands (more than 7000) and are usually categorized on the basis of their structures. The remarkable structural diversity is a result of combinatorial possibilities of arranging modules of PKS genes containing various catalytic domains, the sequence and number of modules and the post PKS enzymes that are closely associated with PKS. The diversity is further enhanced by a megasynthase which is a hybrid of PKS and Non Ribosomal Peptide Synthase (NRPS). Leinamycin produced by *S. atroolivaceus* is a peptide – Polyketide hybrid having remarkable antitumor activity (Cheng et al., 2002). Recombining cloned PKS genes from different biosynthetic pathways to create new pathways defines combinatorial

biosynthesis; the process starts with a PKS gene cluster cloned into *E. coli*. PKS genes from different clusters then replace selected genes. Ligating together PKS genes from different parental clusters to generate newer combination of genes and hence novel bioactive compounds can also create new gene clusters.

The enzymes that govern the assembly of polyketides by *Streptomyces*, are receiving increasing attention as access to them has improved through molecular genetic methods. An explosion of discoveries and technological innovations has enhanced the capacity to "mutate" the structure of natural products. Empirical gene fusion approaches has led to the biosynthesis of diverse unnatural natural products revealing the versatility and combinatorial potential of PKSs. With the advancement in the understanding of the genetics of antibiotic production by the *Streptomyces* and related filamentous bacteria, begun by David Hopwood (Hopwood, 1999) and other scientists nearly 20 years ago of the polyketide machinery, it is possible to rationally alter the PKS genes (Shen, 2000).

Combinatorial potential of TypeI PKS

It is generally accepted that the combinatorial potential of type I polyketide is considerably greater than that of aromatic polyketide. Naturally occurring complex polyketide shows lots of chemical diversity. There is lot of scope for generating new structures using type I synthase, due to following aspects.

1. Characteristic one- to- one correspondence between active sites and product structure, which allows specified alterations to be made at desired sites.
2. For initiating the condensation reaction, wide range of primer units are available, like aliphatic, aromatic, alicyclic and heterocyclic monomers.
3. Possibility of genetic manipulation of the degree of β -keto reduction following each condensation reaction.
4. Manipulation of the size of the polyketide product by deleting or repositioning of enzymatic domains. (Donadio et al., 1991). The mutant PKS has desired number of active sites.

Using the above mentioned biosynthetic engineering of PKS for erythronolide, novel metabolites have been generated. The initial block due to size of modular gene cluster has

been overcome by using multiplasmid approach, where different subunit of PKS genes are cloned on the compatible vector and then desired mutation are incorporated. Combination of these vectors is introduced in heterologous host. Modifications involving tailoring enzymes are also done in this manner. Combinatorial library of over 50 novel polyketides have been prepared by systematic modification of DEBS PKS using this approach. (Reeves et al., 2004)

Combinatorial Potential of Type II PKS

Aromatic polyketide biosynthesis involves iterative use of few active sites of multisubunit complex. For combinatorial biosynthesis of type II PKS, it is necessary to elucidate the function and specificities of various enzymes. For example, manipulation of early steps requires understanding of specificity of the enzyme with respect to the choice of starter unit, the chain length of the molecule, regiospecific ketoreduction and cyclization (Tsoi and Khosla, 1995). The first cluster to be identified in this group was actinorhodin (*act*) cluster from *S. coelicolor*, followed by tetracenomycin (*tcm*) from *S. glaucescens* and oxytetracyclin (*otc*) from *S. rimosus*. The presumptive positions of the DNA encoding the KS and KR functions within the actinorhodin gene cluster had already been deduced and so these DNA fragments (carrying the so-called *actI* and *actIII* 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 cross hybridization 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.

The *actI* and *actIII* probes were hybridized to Southern blots of restriction digests of total genomic DNA of 25 actinomycetes, including 18 polyketide producers and seven nonproducers: hybridizing bands were revealed in most, but not all, of the producers, and

were absent from most, but not all, of the nonproducers. These results indicated correlation between the presence of, especially in *actinomycete*, DNA sequences that cross-hybridized with one or both of the *act* PKS probes and production of polyketides by that strain, suggesting that the probes could indeed be used to isolate further PKS genes. This was successfully accomplished for granaticin from *Streptomyces violaceoruber*. Proof that the hybridizing sequences isolated from this strain indeed encoded the granaticin PKS was obtained by a gene disruption experiment. In this technique a fragment of a gene is introduced into the wild-type strain and a homologous crossover occurs between it and the resident copy of the same DNA to generate a nonproducing mutant (Figure 4)

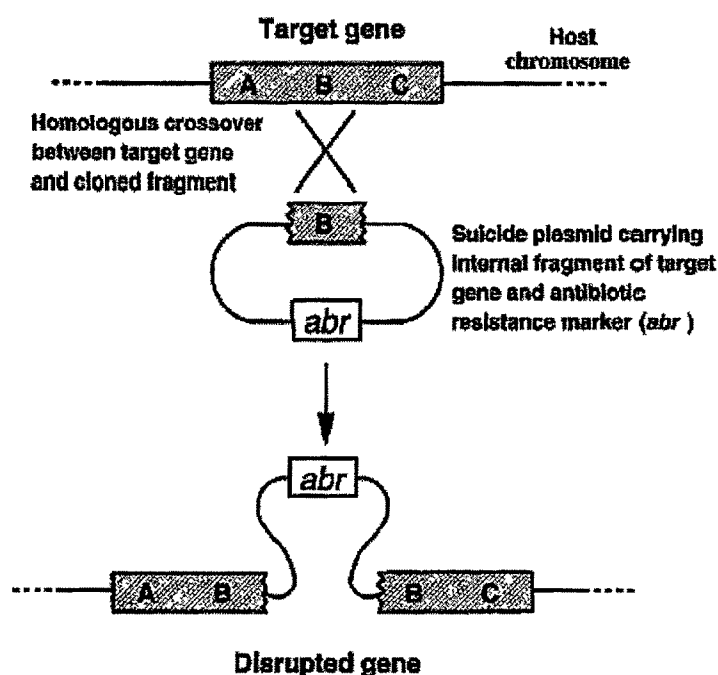


Fig. 9: Gene disruption in one step

The *act* probes were now used to isolate genes for PKS from different organisms. Initially these attempts were successful. However, not all of the genes isolated using these probes actually encoded the PKS. For example polyether, monensin PKS genes from the producer, *S. cinnamonensis*, was identified using *act* I probe, but their disruption failed to interfere with monensin biosynthesis. Similar work has also been reported for peptide synthetase gene of *A. mediterranei* DSM5908, balhimycin producer (Pelzer et al.,

1997). This could be explained as by now the first example of a modular PKS, for biosynthesis of the macrolide erythromycin by *Saccharopolyspora erythraea*, had been discovered, and the *act* PKS probes failed to hybridize to the DNA encoding the erythromycin PKS; doubtless they did not hybridize to the monensin PKS genes either. Presumably genes for a PKS for an unidentified aromatic polyketide had been isolated from *S. cinnamonensis*. A second complication was revealed by the isolation and sequencing of DNA that complemented a mutation (called *whiE*) that prevents biosynthesis of the gray-brown spore pigment of *S. coelicolor* A3(2), resulting in a white spore phenotype. The chemical nature of the pigment was unknown, but the sequence of the complementing DNA left little doubt that it represented PKS genes for an aromatic polyketide, and indeed the *act* and *whiE* PKS genes cross-hybridized, so that the *actI* probe revealed two bands when hybridized to *S. coelicolor* genomic DNA. Probably reflecting a similar situation, use of the *act* probes to try to isolate the PKS for the simple aromatic polyketide curamycin from *Streptomyces curacoii* yielded DNA that may well encode the PKS for a polyketide spore pigment rather than the antibiotic (Hopwood, 1997).

The use of combinatorial biosynthesis was first demonstrated when segments of the *act* biosynthetic gene cluster or the whole cluster, was transferred into the *Streptomyces* that produce medermycin or granaticin (dihydrogranaticin) (Hopwood et al 1985; Ichinose, et al 1998.). In the later experiments the first indication to disprove that specific protein-protein interaction might have evolved to ensure that a set of subunits could work efficiently together to form a functional synthase and that it would not necessarily be a barrier to efficient mixing and matching of PKS subunits was provided by the complementation of a mutation in the *actIII* (KR) gene of *S. coelicolor* by the homologous DNA from the granaticin gene cluster (Malpartida, et. al. 1987). A reciprocal experiment was later reported by (Bartel et al., 1990), who transformed a KR negative mutant of *S. galileus* that produced 2-Hydroxy aklavinone with the *actIII* gene and restored aklavinone production. The approach was further demonstrated by production of a novel polyketide, aloesaponarin II by wild type *S. galileus*. Although nature has provided aromatic polyketide known to us which belong to few common

structural type, about 500 multicyclic aromatic polyketides have been characterized to date, biodiversity may be engineered with the help of combinatorial biosynthesis of type II PKS. A library of more than 30 polyketides was produced by combinatorial biosynthesis of type II PKS (Moore and Piel, 2000). Identification and characterization of new PKS II clusters specially with novel properties like macrotetrolide which operates non iteratively and catalyses both C-C and C=O formation acting directly on Acyl Co-A substrate would enhance the potential of combinatorial biosynthesis. Also combinatorial biosynthesis of hybrid polyketide and ribosomal peptide would help generate novel bioactive compound. Since the starter unit in this case is limited, novel starter units may be attached to PKS to yield new metabolites (Lanen and Shen, 2006).

One very important feature of both modular PKS and type II PKS which makes their cloning and expression in heterologous host is existence of PKS genes (together with transcriptional regulators and self resistance genes) as clusters (Hopwood, 1997). The cloning and heterologous studies are an important aspect of combinatorial biosynthesis as it helps to understand function of each gene and also in studying and developing recombinant metabolites, some of them may be novel in nature.

Requirement of Heterologous Expression

Complex biochemical interactions may be studied in surrogate host of same or different genus as recombinant expression system. The accuracy of the cloned DNA's ability to encode correct function may be assessed precisely in heterologous host. Influence of pathways other than the one to be studied may be avoided and the authenticity of the biochemical pathway may be assessed on one hand and on the other hand matching and mixing of the genes involved in the production of similar compounds may lead to new compounds.

Genetic engineering along with heterologous expression involves exchange of genes encoding structural features among closely related compounds. Existing molecules of both polyketide and oligopeptide antibiotics have been improved/ altered using the above approach.

Heterologous Expression for new bioactive compounds

Heterologous expression involves mixing matching the genes of two related antibiotic producer strains and may yield biologically relevant structure (Flatman et al., 2006). This approach has been able to provide new leads, for example Glycosyl transferases are important for genetic manipulation to generate structural diversity (Walsh et al., 2003) and they are substrate specific and are used from different sources to substitute / add different types of sugars / deoxy sugars and their derivatives to aglycone polyketide, or aminoglycosides. Methyl transferases are also target for improving / redistributing the biological function and generate new structures (McDaniel et al., 1993, Eustaquio et al., 2005).

Above two are the examples of gene products involved in modification of the backbone structures. Both type I and type II are target for modification of backbone structures. In the recent years there is a massive increase in the number of cloned and sequenced genes. In the light of this there is a great need for robust heterologous gene expression methods as the proteins encoded by these numerous genes have an important impact not only in basic biochemical and biophysical research but also practical use of these proteins as pharmaceuticals, animal health products, and industrial enzymes.

Earlier, the ultimate goal of heterologous gene expression was to produce the desired protein in bulk quantities but over the past 20 years there has been a growing interest in the potential for using the intrinsic metabolic activity of proteins in heterologous hosts. Multiple genes are often co-expressed in metabolic engineering to interweave their activities among the genes as well as among functional protein networks in the host. Examples of metabolic engineering include the bacterial biosynthesis of indigo (Ensley et al., 1983) and the conversion of 3-dehydroshikimic acid, a key intermediate in aromatic amino acid metabolism, into a variety of value-added products such as vanillin (Li and Frost, 1998). In Recent times it has been used in biosynthesis of natural products.

Even though lots of effort has gone into the development of efficient toolbox for heterologous-expression, the production of a desired protein in correct form and bulk

quantities still remains unpredictable (Wenzel and Muller, 2005). Although there is no universal solution for heterologous protein production, it is desirable that hosts should be simple to use, should have excellent growth characteristics, and sufficient genetic tools must be available to manipulate it. Few prokaryotic and eukaryotic hosts such as *Escherichia coli*, *Saccharomyces cerevisiae*, Sf9 insect cells, and Chinese hamster ovary (CHO) cells have emerged as model hosts (Klarsfeld, 1994). In these hosts the cellular machinery responsible for protein synthesis, posttranslational modification, protein folding, trafficking, and degradation is well studied. In addition to these features good fermentation protocols especially where cell growth may be decoupled from recombinant gene expression is also important. Even though well developed fermentation protocol may be available for heterologous host and the protein to be expressed there exists a balance between product quality and quantity. High levels of gene expression are many times accompanied by inclusion body formation, increased amino acid mis-incorporation, incomplete or inaccurate posttranslational modification, reduced recombinant cell line stability, and the metabolic burden on the host as a result of excess protein production.

Heterologous Expression and PKS

In an exhaustive review by Pfeifer and Khosla (2001), the advantages and difficulties with different types of heterologous hosts are assessed. The past decade has not only seen heterologous expression of bioactive natural-products, it has also seen rapid growth in understanding and manipulating polyketide biosynthesis, which is one of the important groups of natural products. The complex multistep catalytic processes have been reasonably well studied in genetically friendly heterologous hosts.

With the increased development of polyketide synthases (PKSs) research, number of sequenced and cloned clusters of PKS increased since polyketides are an important class of compounds for biological use, genetic engineering of PKS genes developed as an important field in natural product biosynthesis and drug development (Li and Heide, 2005).

Since polyketide biosynthesis is a complex process and involves a multienzyme assembly of relatively large size (100 to 10,000 kDa), reconstitution of polyketide synthases in heterologous host requires that the multienzyme assembly be functionally expressed, posttranslational modification of PKS be adequately carried out, substrate for polyketide synthesis and post- PKS modification be available in the host in adequate quantity and the host be protected against toxicity of biosynthetic products. High G-C content of many PKS genes, especially those coming from the *actinomycetes*, where the G-C content typically exceeds 70%, is yet another hindrance. Also polyketides are produced as secondary metabolite and the understanding of regulatory and metabolic features associated with the transition from primary to secondary metabolism in any organism is not well understood (Martin, 2004).

In spite of the above problems being there, there are certain features associated with polyketide biosynthesis which makes them particularly well suited for heterologous expression. Firstly, polyketides are remarkably similar in primary sequence and in turn to tertiary structures, so if a good heterologous host is developed for one of the PKS, it is likely that it would work for others too. Secondly, PKS genes exist as clusters along with transcriptional regulators and self resistance genes, which makes not only their isolation easier but heterologous expression simpler especially in bacteria where multigene operons may be constructed. Thirdly, polyketides are produced as secondary metabolites suggesting that they may be produced in two stages fermentation where cell growth may be decoupled from product fermentation. Fourthly, the precursors required for polyketide synthesis are quiet few for eg. Acetyl- CoA, Propionoyl CoA, malonyl CoA and methylmalonyl CoA, which may be derived via multiple metabolite routes from exogenously available carbon source, thus offering a range of option for heterologus host. Lastly, since for every new PKS new microbes need not be developed, lateral transfer of pathway of interest into a well developed host, both for overproducing parent natural product and for developing novel analogous via combinatorial biosynthesis has been possible.

Factors influencing Heterologous Polyketide Production

Reconstitution studies of PKS have shown that they are soluble, cytosolic multienzyme system which does not require any intracellular substrate or organelle to maintain its activity. *In vitro* mixing of individually purified protein or heterodimer also reconstitutes the PKS activity. Some special cellular machinery is at times required to enhance the PKS activity. When the PKS genes introduced in foreign cellular host are able to express efficiently, production of polyketides is supported in presence of required substrate (Khosla et al., 1999). Some important points for polyketide production in heterologous host are:

- 1. Posttranslational modification of PKS:** Phosphopantetheinyl transferases (PPTase) are the enzymes which are known to exist in all organisms except most probably in Archaea which do not synthesise fatty acid or polyketides. This enzyme catalyses the transfer of covalently attaching thiol moiety of 4'-phosphopantetheinyl group to conserved serine residue in polypeptide Acyl carrier protein (ACP). This is a posttranslational modification and the group is derived from intracellular CoASH. The members of PPTase family have different configuration and substrate specificity for eg. *E.coli* has three distinct PPTase each as individual ORF and pantetheinylates fatty acid ACP with selectivity for its substrate; in *S.cerevisiae* the PPTase is a distinct subunit. The PPTase of *B. subtilis* is most tolerant and effectively modifies ACPs from all PKS subclasses and also from related peptidyl carrier protein and aryl carrier protein domains from nonribosomal peptide synthases (NRPS). Therefore for heterologous gene expression, choice of partner PPTase is an important consideration from both substrate specificity and gene regulation perspective. It is important to choose compatible host and expression system with regard to PPTase (Pfeifer and Khosla, 2001).
- 2. Substrate availability:** The task of metabolic engineering of polyketide is simplified by the fact that many natural products of polyketides are synthesized by subset of acetylCoA, propionylCoA, malonylCoA, methylmalonylCoA, few others require substrate like isobutyrylCoA, isovalerylCoA, ethylmalonylCoA, propylmalonyl CoA,

hydroxymalonylCoA and its methylated counterpart methoxymalonylCoA. Some of the substrates are chiral like hydroxymalonylCoA, here the corresponding acyltransferases exhibit strict stereospecificity. The α -carboxylated substrates eg. MethylmalonylCoA and malonylCoA are source of chain extender units whereas neutral substrate such as acetylCoA are source of primer units for polyketide chain synthesis.

Many PKS require only subset of the four substrates mentioned above and there are multiple routes known to exist for many of these precursors, for example, there are 4 pathways to methylmalonylCoA in bacteria. The broad specificity of certain enzyme like malonylCoA synthase can be used to synthesize numerous α -carboxylated CoA thioesters from their corresponding exogenously supplied 1,3 dicarboxylic acids (Zhang et al., 1999) The preferred substrate is tailored easily because of the modular feature of the PKS cluster. When uncommon substrates like hydroxymalonyl CoA or 3-amino-5-hydroxybenzoic acid are required, they are generally encoded as part of PKS gene cluster, and are cloned in the heterologous host along with the PKS genes. Since *actinomycetes* are known to produce polyketide, heterologous host from the same group would be a good choice. The complete genome sequence of *S. coelicolor* is known and it has been used as heterologous host for many metabolic studies, with good success (Hopwood, 2003). Further orthologues of many known precursor biosynthetic enzymes have been identified in its genome; functional genomic approach to investigate the metabolome of *S.coelicolor* may yield many more enzymes (McDaniel et al, 2005). Also many hybrid natural products are derived from tandem action of PKS and NRPS (Sieber and Marahiel, 2005). Carboxylic acid like p-aminobenzoic acid, 3-amino-5-hydroxybenzoic acid, cyclohexenoyl carboxylic acid and some α and β aminoacids are required by NRPS. Some activated forms (eg. CoA thioesters) of free acids are activated in situ by ATP dependent adenylation domains and these domains are present as intrinsic components of NRPS modules (Sussmuth and Wohlleben, 2004). The supply of NADPH and ATP must be coordinately regulated with polyketide biosynthesis by the host.

3. **Other Intracellular Factors:** Co-incubation of active PKS and their substrate is sufficient for polyketide production in heterologous host, but efficiency may depend on the cellular environment. Though dependence of PKS folding on any chaperons have not been shown but PKS activity may be affected by factors which affect the stability of >50kb transcript, translational processivity of >30kb ORFs, translational coupling of ORFs encoding subunits that must assemble in 1:1 stoichiometric ratios, co-translational and posttranslational folding and degradation. Folding and subsequent quaternary assembly of PKSs may be efficient in presence of some chaperons but no mutants defective in PKS have been shown to map to putative chaperone genes, nor have chaperone like gene been identified in PKS clusters. The fact that PKS have been heterologously expressed over a wide range of microbial hosts like *E coli*, *Streptomyces*, *Aspergillus* and Yeast indicates that chaperons are not essential for proper folding of Polyketides.
4. **Transmembrane Transporters:** Bioactive polyketide are potential cytotoxic agents, so they must be exported out of cell. Generally putative export proteins, mostly like ATP binding cassette transporter homologues are found associated with PKS cluster. Their mechanism and selectivity is not well understood. *S. coelicolor* accumulates actinorhodin in intracellular space when gene encoding the actinorhodin exporter is inactivated. *S. coelicolor* genome encodes other transporter proteins which transport and renders the organism tolerant to many antibacterial compounds. These transporter proteins are present even in mutants lacking actinorhodin cluster as the mutants are able to tolerate many novel polyfunctional aromatic compounds by secreting these compounds out of the cell suggesting them to be different from actinorhodin exporter.
5. **Post PKS modification:** The product synthesized by PKS might require some 'tailoring' enzyme to produce biologically active polyketide. These include cyclases, group transferases eg. C-, O-, and N-methyltransferases, glycosyl transferases (Zhang et al., 2006) and acyltransferases, NADP(H) or FAD(H) dependent oxidoreductases and cytochrome P₄₅₀ type oxygenases. Genes encoding these enzymes are present

adjacent to PKS genes and are cloned associated with them. Sometimes cosubstrates like TDP-deoxy sugars maybe required by glycosyltransferases associated with polyketide pathways. TDP-deoxy sugars are themselves product of multistep pathways. These genes are also present in the PKS cluster only.

6. **Important hosts for Heterologous polyketide production:** Choice of heterologous host depends on objective of heterologous expression. It may be carried out for the following reasons.
 1. **Overproduction of target natural products:** Many polyketides are derived from organism which are difficult to culture or is produced in small quantities. Transferring the biosynthetic genes in physiologically characterized heterologous host helps in overproduction.
 2. **Combinatorial biosynthesis:** Many wild type strains producing the bioactive compounds are not amenable to genetic manipulation or biochemical analysis; in such cases a heterologous host might prove to be better for combinatorial biosynthesis.
 3. **Lateral transfer of PKS genes** might lead to important bioactive compound or library of derivatives in heterologous host. For example, production of clonal library of small molecule ligand and target receptor within the same cell could facilitate the design of selectable system for ligand optimization.

Among the important host *S. coelicolor* is an ideal host for heterologous production of polyketides. It belongs to *Streptomyces* genus and its biology has been worked out in detail. It produces two polyketides, one is actinorhodin and the other is *WhiE* spore pigment. A clean host CH999 is constructed by eliminating the background noise due to the actinorhodin pathway, genetically deleting the entire actinorhodin gene cluster via homologous recombination and replaced with *ermE* marker gene. A low copy shuttle vector pRM5 carrying *actI* –*actIII* bidirectional promoter and *actII*-ORF4 gene encoding activator for *actI*-*actIII* promoter has also been engineered. Polyketide

genes cloned under this expression system allow polyketide production in the same manner as secondary metabolites in wide range of *actinomycetes* strain including *S. coelicolor*, *S. lividans*, *S. parvulus* and *S. erythraea*. Many polyketide products has been produced using this host- vector system like products derived from frenolicin, tetracenomycin, oxytetracyclin, erythromycin etc. clusters. Actinorrhodin cluster

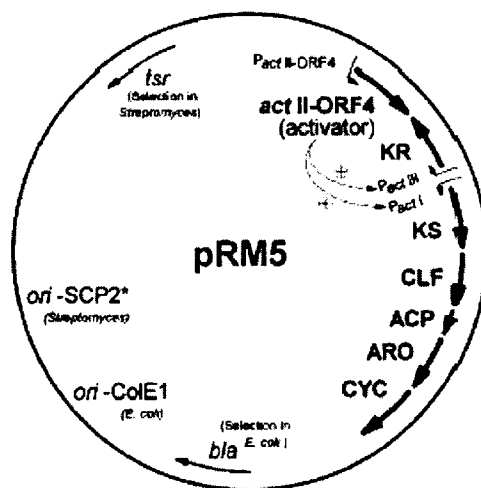


Fig.10: Plasmid pRM5

lacking derivative of *S.lividans*, K4-114 has also been used as heterologous host. Both type II and modular PKS have been manipulated using *S. coelicolor* as host and “unnatural” natural products have been produced. PKS proteins are produced at approximately 1% total cellular protein levels and polyketides are produced in the range of 1-100mg/l. of culture. One of the limitations using *S.coelicolor* as host is that it only uses malonyl CoA derived building blocks as substrate so if heterologous expression of polyketide requires other building blocks, the heterologous expression would suffer. These problems are being overcome by understanding the pathway, for example genome sequencing has identified putative methylmalonylCoA mutase subunit. The biosynthesis of α -carboxylated CoA thioesters is a redundant process, hence there are at least three different acylCoA carboxylase enzymes encoded in *S. coelicolor* and different alleles have different substrate preferences (eg. Acetyl-CoA

versus propionyl-CoA). The organism may also express heterologous genes devoted to substrate production, for eg. malonyl-CoA synthase and dicarboxylate transporter protein from *Rhizobium trifolii* are essential for production of methylmalonyl-CoA which in turn is converted to erythronolide. The enzyme malonylCoA synthase activates a wide range of 1,3 dicarboxylic acids into corresponding CoA thioesters and hence is useful in intracellular generation of substrates for polyketide biosynthesis. Efforts are on to understand and maximize gene expression mainly concerned with secondary metabolism in this host so that both expression of PKSs and synthesis of polyketides of useful cluster are benefited. Since promoters are important for heterologous expression, *actI*, *ermE*, *tipA* has been on trial in *S. coelicolor*. *actI* promoter has been found useful in both expression level and induction of onset of stationary phase in which polyketides are synthesized. Exact signal which leads to maximum induction of these promoter systems is not well understood, so the PKS proteins are synthesized only during short period. For unknown reasons high copy vectors are unstable. Vectors with wide range of copy number are yet to be developed. Till date, expression of PKS genes in *S. coelicolor* is a challenging task but still some clusters have been successfully expressed.

***E. coli*:** It is a significantly different heterologous host from those that naturally produce polyketide. Judicious control of temperature, medium composition and other induction condition yields substantial levels (1-5% of total cellular proteins) of correctly folded protein for PKSs with molecular mass less than 200KDa. Coexpression of Sfp PPTase gene in plasmid borne chromosomal format leads to stiochometric pantetheinylation of soluble PKS proteins in *E. coli* (Ikemura, 1981). High G+C content of *actinomycetes* PKS genes leads to inappropriate bias in codon usage. Using host strain that contain extra copies of rare AGG (arginine) and CCC (proline) helps in expression of genes from such organism (Knight et al., 2001). Reconstitution of heterogonous substrate generation pathway has to be done as biosynthesis of malonyl-CoA is under tight control and other substrates such as propionylCOA are produced under poorly understood condition. Development of two-stage, high- cell-density fermentation protocols, where growth phase may be

decoupled from PKS gene expression and polyketide production has helped in volumetric productivities for both PKS and polyketide. Coexpression of Sfp and 6-methylsalicylic acid synthase genes in *E. coli* leads to production of 6-methylsalicylic acid. Polyketide production was enhanced with the condition which favored production of malonyl-CoA. *E. coli* produces acetyl-, malonyl-, propionyl- and probably methylmalonyl-CoA and their levels may be increased by alteration of pathway. Also, acetyl-CoA carboxylase and malonylCoA carboxylase enzymes from *S. coelicolor* when reconstituted in *E. coli* forms malonyl and methylmalonyl-CoA. Both the above mentioned carboxylase genes have been cloned and overexpressed in *E. coli* and provides sufficient substrate for polyketide synthesis. Although *E. coli* is relatively new heterologous host in use, it has promising opportunity for harnessing the biosynthetic capability of polyketide pathway. 6-deoxyerthronolideB has been produced in *E. coli* (Kealey et al., 1998, Haller et al., 2000).

Other Hosts

Other hosts which may be considered are from *actinomycetes* group to which *S. coelicolor* belongs and are among the major producer of polyketide. They have an added advantage of having naturally occurring polyketide biosynthetic pathway and therefore are better suited for expression of PKS clusters. *Actinomycetes* like *S. lividans*, *S. glaucescens* and *Saccharopolyspora erythraea* have been used as heterologous host (Ziermann and Betlach, 1999). *S. erythraea* has been mutagenised to evolve into high producer of polyketide (50-100 times) more than wild type and the genetic basis of overproduction is not in the PKS cluster but in some other loci and is multigenic (Thompson et al., 1982) If the mechanism of overproduction is identified and translated to model host like *S. coelicolor* and *E. coli*, it would be useful for many polyketides (Gaisser et. al., 2000, Peiru et al., 2005).

Bacteria other than *actinomycetes* also produce polyketide like myxobacteria produces soraphean and epothilone, *Pseudomonas*, Pseudomonic acid, and *Mycobacteria*, coronatine. *Mycobacterium tuberculosis* has many uncharacterized

PKS pathways and *Mycobacterium ulcerans* produces potent immunosuppressant, mycolactone. Strains like *Myxococcus xanthus*, *Pseudomonas putida* and *Mycobacterium smegmatis* are well characterized genetically but their potential as heterologous host is not yet explored (Kealy et al., 1998). Filamentous fungi also produce polyketides. *Aspergillus nidulans* has been successfully used as heterologous host for lovastatin biosynthesis. An added advantage of fungal host would be that they are able to splice introns so eukaryotic PKS genes may be expressed in them. *S. cerevisiae* has been classified as GPRS host, since it has highly evolved pathway of fatty acid synthesis but does not have polyketide pathway of its own, with the introduction of some special feature of polyketide pathway and genes for high lovastatin production from *A. nidulans*, it may be used as one of the good heterologous host. Plants produce polyketides like chalcones, stilbenes and coumarins and their chloroplast provide environment like bacterium, well suited for fatty acid biosynthesis. Since plant genetic engineering is well developed, PKS genes may be introduced in them whose product may directly protect the plant (Gerth et al., 1994).

In choosing a heterologous host where multiple proteins have to be expressed, lot of care has to be taken where apart from good growth in the lab, and it should have cellular machinery robust enough to produce lot of protein and also be retained. Also once a heterologous host is developed in one genera, it should be possible to use them for lateral transfer of PKS genes so that we are able to get better combinatorial potential directing us to novel functional molecules.