Heterologous expression of *act*I homologus polyketide synthase (PKS) genes of *S. flaviscleroticus*.

### Introduction

Heterologous expression of foreign biosynthetic genes in a host that is amenable to genetic manipulation is a powerful approach for exploring the small-molecule repertoire of the bacterial world whose successful application requires that the biosynthetic, resistance and regulatory genes reside on a contiguous stretch of DNA and that the host has an appropriate metabolic and genetic background. Genes for the complete pathways on contiguous stretches of DNA are commonplace for bacterial natural products. Encouragingly, attempts have been successful with engineered hosts (for example, *E. coli* and *Pseudomonas putida*) (Mutka et.al., 2006; Pfeifer et.al., 2001 b) expressing drug candidates genes and genes for other small molecules. Further efforts and progress on more metabolically versatile hosts is likely (Lanen & Shen, 2006; Wenzel & Muller, 2005).

Success with the heterologous expression of DNA has its own potential benefits and difficulties. For example, (i) heterologous expression facilitates the discovery of biosynthetic gene clusters through subcloning and mutagenesis, (ii) of accessing uncultured microbes for antibiotics - to date, all reported small molecules from metagenomic libraries of captured biosynthetic pathways have relied on heterologous expression as a result, mismatches particularly that of background metabolism, codon usage and promoter structure between the host and the captured pathway may limit the range of small molecules that can be discovered.

The conventional practice of using of compatible host (a surrogate *Streptomyces* strain) to heterologously express new peptide and polyketide antibiotics (from producer *Streptomyces*) have been successful using genetic engineering or combinatorial biosynthesis (Kim et al., 1995). Both PKS I and PKS II have been used for the combinatorial biosynthesis for the last 25 years (see review of literature).

P175

## Understanding polyketide synthase:

Initial approach to understanding PKS involved cloning of entire cluster or their fragments into mutants blocked in a step of polyketide synthesis, looking for complementation of mutation and then finding genes for other step(s) in the complementing fragment. Another approach was cloning library of DNA fragments from antibiotic producer strain into sensitive host (usually a derivatives of *S. lividans*), selecting resistant clones and seeking biosynthesis genes by their being linked to the resistant gene(s), a finding common in nature; there are very exceptions to this linkage.

In the past 10-15 years cloning of polyketide antibiotic biosynthesis genes has become easy because of availability of *act*I and *act*III probes which code for conserved ketosynthase and ketoreductase functions respectively. The clones carrying the PKS genes hybridize to the library of genomic DNA; these clones are then expressed in heterologous host or in cell free system in isolation or in combination with other PKS genes to generate hybrid PKSs. Individual components of PKS are also studied either in cell free system or native host.

Successes with the heterologous expression of DNA provide glimpse of the potential benefits and difficulties. For example, (i) heterologous expression facilitates the discovery of biosynthetic gene clusters through subcloning and mutagenesis; (ii) of accessing uncultured microbes for antibiotics - to date, all reported small molecules from captured biosynthetic pathways have relied on heterologous expression as a result, mismatches particularly that of background metabolism, codon usage and promoter structure between the host and the captured pathway may limit the range of small molecules that can be discovered.

Heterologous expression of foreign biosynthetic genes in a host that is amenable to genetic manipulation is a powerful approach for exploring the small-molecule repertoire of the bacterial world whose successful application requires that the biosynthetic, resistance and regulatory genes reside on a contiguous stretch of DNA and that the host

has an appropriate metabolic and genetic background. Complete pathways on contiguous stretches of DNA are commonplace for bacterial natural products. Encouragingly, attempts have been successful with engineered hosts (for example, E coli and *Pseudomonas putida*) (Mutka et al., 2006; Pfeifer et al., 2001b) expressing drug candidates genes and genes for other small molecules. Further efforts and progress on more metabolically versatile hosts is likely (Lanen & Shen, 2006; Wenzel & Muller, 2005).

Polyketide clusters of many *Streptomyces* sp. have been expressed in heterologous host either in isolation or in combination with the endogenous compounds produced by the host. As a corollary thus, heterologous expression of additional type II polyketide biosynthetic genes from uncharacterized *Streptomyces* sp. would help in producing novel metabolites by combinatorial biosynthesis.

DNA based prescreening suggested that *S. flaviscleroticus* is a producer organism of type II aromatic polyketide, as its genomic DNA hybridized to *act*I probe representing conserved region in PKS cluster and encodes ketosynthase. This chapter describes expression of *act*I homologous PKS genes of *S. flaviscleroticus* cloned in shuttle cosmid vector pKC505 (Fig.3.1). The 40 kb genomic DNA of the producer strain is present on a set of eight overlapping cosmid clones. The cosmid set was expressed in each of the three heterologus hosts, viz., *S. lividans, S coelicolor,* and *S parvulus.* The former two produce polyketide, actinorhodin and the latter strain does not produce any endogenous polyketide. The expression studies would define the cosmid clone that carries complete complement of the PKS genes necessary for the production of *S flaviscleroticus* specific molecule and the nature of bioactivity associated with it. The heterologus expression studies would corroborate the results of gene disruption studies, which indicated the polyketide genes to be functional. Transformants of *S. lividans* and *S. coelicolor* carrying complete PKS cluster of *S. flaviscleroticus* would also be analyzed for new polyketide structures.

### Expression of set of eight cosmid clones of S. flaviscleroticus:

The cosmid used to clone PKS DNA of *S. flaviscleroticus* is *E coli* - *Streptomyces* bifunctional shuttle cosmid vector, pKC505 (Fig. 3.1). It is a derivative of *Streptomyces* plasmid SCP1\* minus its stability function (Kinashi et al., 1991); has pBR322 replicon for its maintenance in *E. coli*, and has gene aac(3)IV for apramycin-resistance as marker for selection and maintenance in both hosts.

### **Choice of Heterologous host:**

Two cosmid DNA were selected at random for transformation into each of the three hosts. Following results describe the need for experimenting with the three hosts and assesses suitability of each of the host for the purpose of heterologous expression studies.

Two of the cosmid DNAs namely, #2.19 and #1.51 were each used in transformation experiment. As expected, *S. lividans*, and *S. parvulus* could be transformed with the DNA isolated from DH5 $\alpha$ , a methylase positive strain, whereas DNA for transformation in *S coelicolor* had to be propagated through a methylase mutant of *E. coli* (ET12567), as this host restricts the methylated DNA.

There were two problems encountered during heterologous expression studies and analysis of the transfromants for production of new metabolites. First one pertained to the hosts *S. lividans*- and *S. coelicolor*-specific endogenous antibiotic actinorhodin's  $R_f$  overlapping that of polyketide antibiotic of *S. flaviscleroticus* Second was that in each of the three hosts, the cosmid transformants lost the antibiotic marker in the absence of selection pressure, suggesting that the cosmids were segregationally unstable in all the three hosts tested.

Each of the two problems -

1. Background Noise due to endogenous antibiotic and 2. Instability of the cosmid clones was dealt in the following manner.

- i) Deleting the actinorhodin biosynthesis genes and
- ii). Integrating the cosmid clones into chromosome of the heterologous host.



Fig. 3.1: Shuttle cosmid vector pKC505.

...

•••

1. Background Noise: *S. lividans* produces actinorhodin very poorly under our lab conditions; thus the organic extract of *S. lividans* did not reveal bioactive spots on TLC developed using *M. luteus* as test bacteria. However, for reasons unclear to us, #2.19 and #1.51 transformants produced large amount of actinorhodin and the R<sub>f</sub> of both *S flaviscleroticus* specific antibiotic and actinorhodin on TLC was almost same, making it impossible to analyze the biomolecule produced by the heterologous genes. An observation similar to this finds mention in Hopwood et.al, (1995) in which *S. lividans* and its derivatives transformed with heterologous DNA have been stated to be stimulated for synthesis of endogenous antibiotic, probably as a part of "shock" response. In light of these facts it was anticipated that transformation of the host with other cosmid clones may also enhance the production of its endogenous antibiotic, actinorhodin. Thus it is imperative that the 'clean' host be constructed for expression of the heterologous genes.

i) Deleting the actinorhodin genes: Actinorhodin non producing mutant of S. lividans was constructed by deleting the actinorhodin biosynthesis genes so that it lacks the ability to produce the endogenous pigment antibiotic. For this, cosmid clone 9E12 (Redenbach et al., 1996) containing the actinorhodin genes, was obtained from the lab of M. Redenbach, and gene replacement vector was constructed using this construct (Fig.3.2). Cosmid 9E12 DNA was digested by restriction enzyme, Sall and subjected to intramolecular ligation. This step removes all but  $\sim$  5-6 kb DNA, generates unique Sall site which is flanked by truncated actinorhodin genes of 2-3 (Fig.3.2 & 3.3). A double crossover recombination between the act sequences on the plasmid and the corresponding sequences in the chromosome places the deletion from the plasmid into the chromosome as shown in Fig.3.4. The transformants (integrants) were obtained using selection pressure of kanamycin. The integrants were cultured in absence of kanamycin spores were collected and plated at dilution of  $10^{-4}$  and  $10^{-5}$  on R4 plates in the absence of kanamycin. The white colonies (mutant phenotype) were checked for their ability/inability to produce actinorhodin. The mutants were Kan<sup>s</sup> unselected, and did not produce pigmented antibiotic actinorhodin (Fig. 3.5).



Fig. 3.2: Schematic drawing of construction of act deletion vector.



Fig. 3.3: RE digestion analysis of gene knockout vector, 9E12 $\Delta$ sal: (lane1), Sall digestion of 9E12 $\Delta$ sal; (lane 2), Sall digestion of 9E12 DNA; (lane4), double digestion of 9E12 $\Delta$ sal with Sall and EcoRI and (lane3),marker DNA  $\lambda$ H.



Fig. 3.4: Schematic drawing of homology dependent two-crossover recombination to place the deletion of *act* genes into the genome of *Streptomyces lividans*.

...

•

Out of 10,000 colonies that were screened, 10 had mutant phenotype. The mutant phenotype is a result of double crossover where the vector integrates into the chromosome in the first crossover.

Instability of the cosmid clones: The cosmid DNA of #2.19 and #1.51 were transformed into each of the three hosts and instability of the transformed DNA was a common problem to all of them. The instability is two fold and refers to (i) inability of the plasmid to be stably maintained by the cells in the absence of antibiotic selection and (ii) one region and excises out in the second crossover event using the homology of the other. Unexpectedly, however these mutants started producing undecylprodigiosin, the second non polyketide antibiotic it is known to produce in more amounts, but since its Rf is different from Rf of S. flaviscleroticus specific compound, we could still proceed with heterologous expression studies loss/rearrangement of the insert DNA in the proportion of cells maintaining the plasmid. The instability was measured as (i) loss of plasmidborne apramycin resistance marker in high proportion of cells amplified in the absence of selection pressure (ii) difference in the restriction digestion pattern of the plasmid before and after the retrieval in E. coli from the cells of the heterologous host. Instability of the insert DNA was mainly observed as deletion of the insert DNA than as its rearrangement. #2.19 plasmid DNA isolated from S lividans transformants recovered in E. coli strain DH5a was analyzed by restriction digestion (four independent clones were used) - insert derived fragments were completely absent and vector derived fragment were unidentifiable with those of pKC505 (Fig 3.6); such instability of pKC505 was reported earlier in S. lividans (Richardson et.al, 1987).

#### 2. Integration the cosmid clones into chromosome of the surrogate host:

In establishing the phage lysogenic state, several temperate phages including *Streptomyces* phage,  $\varphi$ C31, *int* (integrase) and *att* (attachment) sequences play a very important role. Integrase is the only viral protein required for the integration reaction. Integration requires recognizing specific loci on the bacterial and phage DNA called attachment (*att*) sites. The attachment site on the bacterial chromosome is called *att*B and



Fig. 3.5: Morphological features and color produced by S. lividans (WT) and act deletion mutant of S. lividans (M).



Fig. 3.6: Instability of the cosmid and the insert DNA in the surrogate host, S. *lividans*: (lane 1,2,3,4), BglII restriction digestion of #2.19 DNA retrieved from four independent transformants of DH5 $\alpha$ ; (lane 5), BglII digestion of #2.19 DNA; and lane 6,  $\lambda$ H.

that in the phage *att*P. The locus is defined by mutations that prevent integration of phage; it is occupied by prophage in lysogenic strains.

The *int-att* system of  $\phi$ C31 phage was cloned into the cosmid DNA used in this study for their stable integration into the chromosome of *Streptomyces*. For this purpose, plasmid called # 2' (Fig. 3.7) was constructed. A 2.4 kb PstI fragment from partial digest of the plasmid, pSET152 (Bierman et al., 1992) was cloned into the plasmid pBR322 (Bolivar et.al, 1977). This DNA contains attP and intC of  $\varphi$ C31 phage, and oriT DNA from broad host range R4 replicon for conjugal transfer of DNA in cis. In each of the cosmid clones, int-att DNA from the plasmid #2' was cloned by in vivo recombination (Fig. 3.8). For this, the plasmid, #2' was transformed into E. coli DH5 $\alpha$  (recA<sup>-</sup>) containing each of the cosmids. Wild type phage  $\lambda$  was grown on double transformant which harbored both #2' and cosmid. ori region homology between the two plasmids allowed for cointegrate formation by lambda Red pathway of recombination. The frequency of transduction of cointegrate in DH5 $\alpha$  was 10<sup>-5</sup>/pfu as against 10<sup>-1</sup>/pfu for transduction of only cosmid DNA. Cointegrate of one of the cosmid clones #1.23, was analyzed by restriction enzyme digestion. HindIII cleaves #2' DNA to give 0.8- and 0.9 kb fragment and does not cleave #1.23 DNA, whereas Bg/II digests insert DNA of #1.23 but not the plasmid #2' (Fig. 3.9). The BglII digests of the cointegrate plasmid yields #1.23 specific bands whereas HindIII digestion gives #2' specific bands indicating that cointegrate did arise as a result of fusion of these plasmid during recombination.

Since transformation of *S. coelicolor* requires unmethylated DNA, cointegrate cosmids were processed through ET12567, methylase-negative mutant of *E. coli*. Passing the 50 kb DNA through ET12567 was almost impossible when introduced by transformation because when the methylated DNA (isolated normally from *E. coli* like DH5 $\alpha$ ) enters the methylase negative mutant, ET12567, there is block to further replication of the plasmid after the first round of replication that generates hemi-methylated DNA (Kaguni J. M., 2006). Very few molecules of DNA may overcome this block, however the problem of introduction of large size cosmid DNA was reduced in the case of the transduction by  $\lambda$ .



Fig. 3.7: Construction of vector #2'.



Fig. 3.8: *In vivo* cloning by homologous recombination of the *intC-attP* DNA of φC31 phage into the cosmid.



Fig. 3.9: Proof of cointegrate formation: BglII restriction digestion of DNA of – (lane 1), #2.19/#2'; (lane 2), #2.19; (lane 3), #2'; (lane 4), λH marker; HindIII restriction digestion of DNA of – (lane 5), #2.19; (lane 6), #2.19/#2'; (lane 7), #2'.

# Site-specific integration of PKS cosmid cointegrates (cosmid/#2') in $\triangle act S.$ lividans, S. coelicolor and S. parvulus:

The *act* deletion mutant of *S. lividans* produced another antibiotic undecylprodigiosin in large amounts but since its  $R_f$  was different from that of *S. flaviscleoriticus* specific biomolecule, *S. lividans* was used as a host to transform the cosmid clones now modified to integrate into the chromosome. However the cointegrate clones harboring *att*P and *int* from phage  $\varphi$ C31 did not integrate into chromosome of *S lividans*, as shown by the loss of apramycin (cosmid) marker. Also, as mentioned above the insert DNA on cosmid was prone to rearrangement precluding episomal analysis of the insert DNA expression. Result of restriction digestion analysis of the retrieved DNA (of the plasmid in *E. coli*, DH5 $\alpha$ ) was same as that described in Fig. 3.6.

Nevertheless, the cointegrate cosmid did integrate stably into the chromosome of S *coelicolor*, as here the cosmid marker was not lost when amplified in absence of antibiotic pressure. The cointegrate cosmid did stably integrate in the genome of S. *parvulus* - the inoculum from the transformation plate could be purified easily on the plate supplemented with antibiotic and also that the cells retained the plasmid marker in the absence of selection, indicating that our strategy of integrating the clones into the chromosome was practical and without flaw. For transformation of rest of the clones, *S. parvulus* was used since there is no barrier to accepting methylated/unmethylated DNA.

Integrity of insert DNA: In anticipation of the results that cosmids without *int-att* DNA underwent rearrangements, it was essential to ensure that the transformed DNA in S *parvulus* was without any rearrangement. The site specific recombination of the cosmid cointegrates renders the DNA stable in the host. Though at low but measurable frequency, the reversal of integration causes looping out of the cosmid from the chromosome. Due to this process some amount of cosmid DNA is always present in episomal form. If total DNA from the *Streptomyces* transformant is used to transform E coli competent cells, these episomal plasmids can be recovered. Total DNA from #1.1 and #1.23 transformants was isolated and transformed into competent cells of DH5 $\alpha$ . The restriction enzyme digestion pattern of retrieved plasmid from E. coli transformants and

the original plasmid was same (Fig.3.10), indicating that the insert DNA did not contain any rearrangement.

**Expression of the set of eight PKS clones in** *S. parvulus*: The cosmid cointegrates were individually transformed into *S parvulus* by PEG mediated protoplast transformation protocol.  $\varphi$ C31's *int-att* integrating system ensured their stable integration. The organic solvent extract of three transformants of each clone was analyzed by separation on TLC using 1D solvent system 9.3:0.7::CHCl<sub>3</sub>: MeOH. The *S. flaviscleroticus* specific yellow UV fluorescent spots (Fig.3.11) and the associated bioactivity were not present in any of the transformants. However the transformants of #1.23 and #1.51 (Fig 3.12) showed resistance against *S. flaviscleroticus* specific yellow fluorescent compound, indicating the presence of resistance genes for the polyketide compound in these two plasmids and their ability to be expressed in the heterologous host. The position of the resistance gene(s) is validated by the phenotype exhibited by the #1.1 transformants of *S parvulus* which are sensitive to the *S. flaviscleroticus*-specific yellow UV fluorescent compound (Fig. 3.12). The study of the overlap between different clones (Fig. 3.13) thus places the gene for resistance in the 11 kb *EcoR*I fragment clearly outside the limit of DNA present in #1.1 and being common to #1.23 and #1.51.

In light of this we can infer that all the genes required to code for biosynthesis of *S*. *flaviscleroticus* specific polyketide compound has not been cloned in the individual cosmid DNA from the set selected.

**Pair-wise transformation of cosmid DNAs in S.** *parvulus*: The set of eight cosmid clones together span 40 kb DNA. This complete 40 kb DNA was assembled in the heterologous by the following method.

S parvulus transformant cells containing cointegrate DNA, # 1.23x2' integrated at its *att-int* site was transformed with the cointegrate DNA of #1.1x10 containing integrating system of pSAM (Sezonov et.al., 1998) and hygromycin resistance marker (Fig.3.14). The transformants selected on apramycin + hygromycin would have #1.23x2' integrated at the *att* site of S parvulus using phage  $\varphi$ C31's integrating system and #1.1x10 at *att* site using pSAM's integrating system (Kieser and Hopwood, 1991).



Fig. 3.10: Veracity of the PKS insert in the integrated state in the chromosome of surrogate host: Comparison of RE digestion pattern of retrieved cosmid (from the integrant) and its original version by BgIII restriction enzyme digestion of DNA of - lane 1, #1.23 original; lane 2, #1.23 retrieved; lane 4, #1.1 original; #1.1 retrieved; lane 3, Marker DNA λ HE.



Fig. 3.11: TLC of crude extracts of *S. parvulus*/cosmid integrant and *S. flaviscleroticus* (FS) cultures.

By the strategy above we were able to assemble 40 kb DNA in *S* parvulus though the integrity of the insert DNA in each of the cosmid after integration was not determined. It was presumed to be intact and coherent in the light of the results obtained with single cosmid construct integrated in the chromosome with the *int-att* system of *Streptomyces* phage. These transformants were analyzed for *S*. *flaviscleroticus* specific yellow fluorescent bioactive compound on 1D and 2D TLC, again to find that the compound was not produced. The insert in the cosmid DNA was insufficient to direct synthesis of *S flaviscleroticus* specific polyketide compound suggesting that though the cosmid clones picked up in colony hybridization experiment together span 40 kb DNA, the complete cluster has not been picked up. This result is not surprising as there could be several reasons for it, like PKS clusters at times are larger than 40 kb, or some genes for producing polyketide structure may be unlinked.

**Discussion:** Given the easy manipulation of *S. lividans* host, we attempted construction of 'clean' mutant derivative that does not produce actinorhodin in order to reduce 'contamination' by the host specific endogenous antibiotics. This apparently did not help much for the host, *S. lividans* which produced enhanced amounts of undecylprodigiosine, and methylenomycin in the absence of actinorhodin. Secondly, the cosmid clones having the integrating system did not help stabilize cointegrate DNA in *S lividans* chromosome for the reason we do not understand. The apramycin marker present in the cosmid clones was still lost with same frequency as in the case of cosmids that did not have *int-att* DNA, and the transformants could not be analysed for expression



Fig. 3.12: Localization of the resistance gene for the *S. flaviscleroticus* specific polyketide compound in the PKS cluster: III, purified fraction III of *S. flaviscleroticus* organic extract; P, crude extract of *S. parvulus*. A), lawn of *S. parvulus*/#1.51; B), lawn of *S. parvulus*/#1.23; C), lawn of *S. parvulus*/#1.1; D), lawn of *S. flaviscleroticus*; E), lawn of *S. parvulus*.

....

• • •



Fig.3.13: Restriction map of the ~40Kb DNA with respect to EEco RI and Bg:Bg/II and the limits of the insert DNA in the different cosmid clones.

The Blue colour filled bar in clone #2.19 shows the deleted region from the chromosome of the PKS mutant.

DP Glucose Dehydratase, Beta Keto-acyl Synthase Chain Length Factor
O-methyl transferase, NDP-Glucose synthase Cyclase, Monooxygenase

11



Fig. 3.14: *In vivo* cloning by homologous recombination of the *int-attP* DNA of plasmid pSAM into the cosmid.

of cloned DNA. The cosmid neither integrated nor stably remained episomal. Besides, the plasmid marker and insert were lost from the cells in absence of selection pressure.

The cosmids with integrating system did integrate in the genome of other two hosts namely S coelicolor and S. parvulus, indicating lack of obvious flaw in the strategy. Though we constructed act mutant of S. coelicolor, production of two other antibiotic namely undecylprodegeosine and calcium dependant antibiotic, methylenomycin was often enhanced as in the case of S. lividans (see above). Also TLC profiles of recombinant and untransformed strains often overlapped, introducing ambiguity in the interpretation of results of bioautogram. Though derivatives of S coelicolor and S. lividans with the deletion of the genes for all the antibiotic production have been constructed, availability of those strains was limited. There was another factor which had to be considered in case of S. coelicolor, that regards strong restriction system exhibited by S. coelicolor which required passing every time the 50 kb cosmid DNA through methylase mutant of E. coli. It was one of the main reasons behind not selecting S. coelicolor as heterologous host. On the other hand, in S parvulus, not only did the cosmid clones integrate stably, the host did not restrict the methylated DNA. If we could express the cloned DNA in S. parvulus, in terms of S. flaviscleroticus specific compound being produced, then it would have been worthwhile going through the arduous task of passing the 50 kb DNA through ET12567 and transform the cosmid DNA into actinorhodin producing S coelicolor so as to look for hybrid products.

The set of eight clones carrying overlapping PKS genes of *S* flaviscleroticus could be integrated into the *S*. parvulus chromosome only after providing the cosmids with integrating system of phage  $\varphi$ C31. None of the cosmid clone however, could specify synthesis of bioactive compound specified by the cloned genes of *S* flaviscleroticus in spite of the fact that each cosmid contains 20-25 kb insert DNA, a range which may be sufficient to encompass complete gene cluster of a typical type II PKS (see literature review).

Integrity of the cosmid clones was confirmed by restriction digestion pattern of original cosmid being same as that of retrieved cosmid from *S. parvulus* transformants.

Interestingly, the transformants of #1.51, and #1.23 cosmid cointegrate were resistant to the polyketide produced by *S. flaviscleroticus* indicating the presence of self resistance genes.

In order to construct the complete cluster, end sequences of cosmid #1.1 and #1.23 clones may be used as probe, to select hybridizing clones from genomic DNA library which may complement the present clones to synthesize the polyketide compound defined by PKS genes of *S* flaviscleroticus. Even if some of the intermediates were produced in the transformants generated in our work, we could not assign bioactivity to them and which needed sophisticated techniques to identify and characterize them.