CHAPTER IV
DISCUSSIONS
4.1 GENERAL BIOLOGY OF THE PHAGE PIS136

4.1.1 Genome of the phage PIS136

The phage PIS136, which was isolated from a soil strain of *Saccharomonospora* sps. has a double stranded DNA (dsDNA) genome (Fig.1) of ~90-kb (Fig.2). The genome of the phage PIS136 is linear but circularly permuted and terminally redundant (Fig.4). A probable pac-site could also be seen. The presence of pac-site suggests a headful mechanism of packaging of DNA. The pac-site, which is usually not very large, initiates packaging for the first headful DNA molecule (Sternberg and Maurer, 1991).

Bacteriophages with large genome have been reported from *e.g.* *Streptomyces* DAH1, DAH2, DAH4, DAH5, DAH6 (Burke et al., 2001), *E. coli* e.g. T4 (Lewin, 1997), *Salmonella* e.g. phage P1 (Sternberg and Maurer, 1991). These phages are known to have a genome of ~100-kb or more. Similarly phages with linear but circularly permuted genome have also been reported *e.g.* Pg2 (Sladkova et al.,1979), SF1 (Chung et al.,1983) FP43 (Hahn et al.,1991). Since the genome of the phage PIS136 is linear and circularly permuted a cos-site was not expected. Results shown in (Fig.3) did confirm that genome of the phage PIS136 does not have a cos-site. Restriction digestions with several enzymes suggested that the genome of the phage is large. Therefore the genome of PIS136 was resolved by pulse field gel electrophoresis. The results clearly showed that the genome of the phage PIS136 is over 90-kb. Therefore, it can be summarized that the phage PIS136 has a large dsDNA genome of ~90-kb which is linear, circularly permuted, terminally redundant and is devoid of any cos-site.

A possible implication of the above properties of the phage PIS136 is that the phage PIS136 may be useful as a generalized transducing phage. Generalized transduction in eubacteria usually depends upon the absence of highly specific DNA sequence recognition during the formation of mature virions *i.e.* such phages usually package their DNA by headful mechanism, which generates some degree of terminal redundancy (Chater, 1986). However, the phage PIS136 has not been tested for transduction therefore it is not possible to make a definite statement.
The PIS136 genome is GC rich (69-71 mol percent, unpublished observation) which is similar to the large number of actinomycetes genera to which the phage PIS136 infects. Although most of the phages have GC composition matching with that of their host (Table B) however phages with a GC composition quite different from that of their host are known e.g. the phage FP22 has a GC content of ~49 mol percent in comparison to ~71 mol percent of its host, *Streptomycyes* sps. (Cox and Baltz, 1984). The impact of high GC content of the phage PIS136 could be seen in the presence of target sites for restriction enzymes. Those restriction enzymes which recognize large GC stretch cut the genome more frequently than the ones which recognize AT rich sequences. The distribution of GC bases is such that an octacutter like *NotI* has 3 target sites within a ~90-kb genome of the phage PIS136.

### 4.1.2 General properties of the phage PIS136

#### 4.1.2.1 Host range of the phage PIS136

The phage PIS136 has a very wide host range within the actinomycetes (Table A). Results described in Table A clearly show that the phage PIS136 crosses the generic barrier. Bacteriophage infection to a host cell is mainly controlled at two levels: (i) extracellular, in the form of phage specific receptor to which tail fibers of a phage attaches itself (Okanishi et al., 1968; Lomovskaya et al., 1977) (ii) intracellular, in the restriction-modification system of the host plays a role (Chater, 1978, 1979; Lomovskaya et al., 1980; Kruger et al., 1983). Bacteriophages overcome the restriction barrier of the host either by modifying their own genome by methylases (Kahmann, 1984; Schroeder et al., 1986; Scherzer et al., 1987; Kossykh et al., 1995; Kossykh et al., 1997) or by a better evolutionary mechanism where target sites for the host R-M systems are not present (Kruger et al., 1983; Cox et al., 1984).

A very wide host range of the phage PIS136 within actinomycetes suggests that in the case of the phage PIS136 the adsorption specificity or extracellular barriers may not be a limiting factor. The phage PIS136 infected *Streptomycyes achromogenes* subspecies *achromogenes* (DSM40028), *Streptomycyes albus* G and *Streptomycyes canescens* (ISP5001). These three species are known to produce at least one of the restriction enzymes *SacI, SalI* and *ScaI* respectively. The phage PIS136 has more than
20 sites each for *SacI* and *SalI* and has 4 sites for *ScaI* (Fig.3, 7 and 14). Therefore, in case of the phage PIS136 the possibility of having a very wide host range by avoiding the target sites for a host encoded restriction-modification system can also be ruled out.

The phage PIS136 DNA does not appear to be methylated (Fig.5) therefore it would be prone to host encoded restriction endonucleases. The question arises that how does the phage PIS136 overcome the host restriction barrier? One possible mechanism by which PIS136 DNA may overcome the host's restriction-modification system is by encoding DNA modifying enzymes of its own. Expression of the phage encoded DNA modifying enzymes is regulated in such a way that unless needed, the presumed phage mediated modification system is silent. Various reports have shown that phages T1 (Scherzer *et al.*, 1987), T2 (Kossykh *et al.*, 1997), T4 (Kossykh *et al.*, 1995; Kossykh *et al.*, 1997) and Mu (Allet *et al.*, 1975; Kahmann, 1984) do encode their own methyl-transferases. Hence it is possible that the phage PIS136 may be encoding its own DNA modification machinery to survive in wide range of host strains (so far part of an ORF encoding dCMP methylase has been sequenced by K.L. Pathi.

A strain of *Saccharomonospora* PA136 which is a natural lysogen of the phage PIS136, produces two types of colonies: (a) produced green pigment (PA136(G)/PIS136(G) (b) did not produce green pigment (PA136(W)/PIS136(W) (Fig.9). The host range of the phage PIS136(W) isolated from the non-pigment colonies was found to be rather narrow (unpublished observation). Total protein profile of the phage PIS136(W) and PIS136(G) did not show any difference (Fig.10) therefore we concentrated on the genome of the phage PIS136. To our surprise, the *ClaI* digest of the phage PIS136(W) and PIS136(G) showed differential restriction pattern (Fig.11).

### 4.1.2.2 Inversion of a DNA fragment in the phage PIS136

The differential restriction pattern for two different populations of a same phage has also been reported in the phage Mu (van de Putte *et al.*, 1980; Kamp *et al.*, 1981), P1 and P7 (Iida *et al.*, 1982; Rozsa *et al.*, 1995). *Neisseria gonorrhoeae*
Discussion

shows antigenic variation in pili formation. The pili in *N. gonorrhoeae* is encoded by a pilin gene, *pilE*. The *pilE* shows antigenic variation by unidirectional transfer of DNA sequences from one of the several silent pilin gene copies (pilS) by recombination to replace variable sequence within the *pilE* gene. The recombination event changes the size of *pilE* gene, which shows differential restriction digestion pattern in comparison to normal *pilE* gene (Meyer *et al.*, 1982; Mehr and Seifert, 1998). In all the above cited examples of the differential restriction pattern, of two different populations of a phage or bacteria, has been found to be due to DNA inversion where a part of genome changes its orientation. The change in orientation of a part of the genome changes some properties of the genes present in that particular stretch of DNA. The DNA inversion in the phage Mu, P1, P7 and is mediated by a group of closely related enzymes called invertase. 'Gin' is an invertase of the phage Mu, inverts the orientation of a DNA fragment, 'G-region', by acting on a 34-bp inverted repeat located on the boundaries of the 'G'-region. Other invertase also function by a similar mechanism (Plasterk *et al.*, 1984).

Another example of DNA inversion is the expression of type I fimbriae in *E. coli*. In *E. coli* the inversion of *fimS*, the gene which controls the expression of type I fimbriae, is unique. Inversion of *fimS* is mediated by two site-specific recombinases, which are integrases, named as FimB and FimE. Both the integrases invert *fimS* by recombination between two 9-bp inverted repeat bordering the *fimS* gene (Kawala *et al.*, 1991; Guo *et al.*, 1997).

The Clal digest of the phage PIS136(G) and PIS136(W) showed (Fig.11) that an 2.33 kb [C(i)] fragment was missing from the PIS136(G). Subsequent studies proved that the fragment C(i) is represented as fragments C6 and C10 in the PIS136(G) (Fig.13). Our studies also proved that all the three fragments *i.e.* C(i), C6 and C10 are present in both PIS136(G) and PIS136(W) but their concentration varied between PIS136(G) and PIS136(W). The fragment C(i) was a major component of PIS136(W) genome while fragments C6 and C10 were the major components of PIS136(G) genome. The inversion fragment of the PIS136 did show some homology with MuG (Fig.14) however the DNA sequence of fragments C6, C10 and the flanking region so far has not revealed the presence of an invertase. Rather the DNA sequence of the fragment C6 was the C-terminus of a DNA polymerase α-subunit.
Discussion

gene and C10 fragment showed the presence of two ORFs which have homology to helicases and DnaJ (Table 4, Fig.10). Upstream to the DNA polymerase α-subunit gene, a 408 amino acids long integrase is present (Unpublished work of K.L. Pathi).

4.1.2.3 DNA polymerase α-subunit of the phage PIS136

Many of the bacteriophages such as φC31 (Smith et al.,1999), Bxb1 (Mediavilla et al.,2000), D29 (Ford et al.,1998), L5 (Hatfull et al.,1993), M2 (Matsumata et al.,1989) PRD1 (Jung et al.,1987), T5 (Laevitt et al.,1989), T3 and T7 (Patel and Loeb, 2001) are known to encode their own DNA polymerase I and some of the accessory factors also. Similarly the phages such as RB69 (Wang et al.,1997), φ29 (Escarmis and Salas,1982), T4 (Lamm et al.,1988), T6 (Patel and Loeb, 2001) are known to encode their own DNA polymerase II and some of the accessory factors. However, only the phage SPBc2 (Lazarevic et al. unpublished, with GenBank accession number NP_0466885) is known to encode a putative DNA polymerase α-subunit, which is a part of DNA Pol III holoenzyme.

More than half (666 aminoacids) of the DNA polymerase α-subunit of PIS136 was a part of the attP fragment (Unpublished work of K.L. Pathi). However, the entire ORF codes for a 1206 amino acids long protein. Both DNA and protein sequence showed very high homology with eubacterial DNA polymerase α-subunit (Fig.22 and 21). Surprisingly, the highest degree of homology was found with *Mycobacterium leprae* and *M. tuberculosis* H37Rv rather than with *Streptomyces coelicolor* A3(2). Both the alignments (Fig. 21 and 22) and the phylogenetic tree (Fig.23) showed that although DNA polymerase α-subunit of the phage PIS136 originated from *M. leprae* and *M. tuberculosis* H37Rv, it definitely has diverged. At present it is difficult to suggest the impact of divergence. However, it is obvious from the alignment data (Fig.21) that the DNA polymerase α-subunit of the phage PIS136 has undergone deletions, additions and substitutions, while still maintaining the overall properties of the DNA polymerase α-subunit such as the conserved phosphoesterase domain (php) at the N-terminus and OB-fold domain at the C-terminus (Fig.20). How did the phage PIS136 acquire DNA polymerase α-subunit, which is a part of DNA pol III complex, where all the other known actinophages have
DNA pol I or DNA pol II which do not have conserved phosphoesterase and OB-fold domains? Assuming that the DNA polymerase α-subunit of the phage PIS136 is a result of the horizontal gene transfer then neither *Mycobacterium leprae* nor *Mycobacterium tuberculosis* is a soil bacteria. Therefore, it is tempting to suggest that having its own DNA polymerase α-subunit which is closer to both very high G+C containing *Steptomyces* (73 mol percent) and relatively low G+C containing *Mycobacterium* (65+ mol percent) is advantageous for the phage. At present it is not known whether the phage PIS136 encodes other proteins of the replication machinery or not.

It appears that a functional DNA polymerase α-subunit is very important to maintain the wide host range of the phage PIS136. It is clear from the results of the present study (Fig.24) that change in orientation due to the 9-bp inverted repeats at positions 54 to 62 and 2207 to 2215 results into repositioning of *Clal* site from 1981 to 283 which ultimately leads to the truncation of the ORF encoding DNA polymerase α-subunit. The truncated ORF encodes 645 amino acids long protein in place of 1206 amino acids long protein. After inversion a new *Clal* fragment of 2.333-kb appears in the PIS136 (W) digests, which was named as C(i) fragment, and has a narrow host range. Therefore it appears that the DNA polymerase α-subunit plays a major role in maintaining the wide host range of the phage PIS136. One of the most important and novel contribution of present study is the demonstration that apart from the adsorption, host restriction-modification systems and phage mediated modifications, replication of the phage itself could be a major contributor in controlling the host range of a phage.

It has been shown that the DNA polymerase α-subunit is also a part of the fragment harboring *attP* region (Fig.18). Immediately upstream to the DNA polymerase α-subunit, an ORF coding for integrase was found (unpublished work of K.L. Pathi). We have also suggested that the DNA polymerase α-subunit may have a major role in controlling the host range of the phage PIS136. Therefore, it appears that integration, replication and the regulation of host range may be an integrated process.
4.1.2.4 The phage PIS136 is a mutator phage

Selection of the four lysogens of *A. mediterranei*, AR4, AR14, AR17 and AR35, was based upon the pigment production (Fig. 25). Hybridization of genomic DNA of the AR4, AR14, AR17 and AR35 with the 'attP' region clearly suggested that the phage PIS136 has integrated at random sites (Fig. 26). Which suggests that the page PIS136 is a mutator phage like the phage Mu (Taylor, 1962).

4.2 THE PHAGE PIS136 AND MODULATION OF RIFAMYCIN BIOSYNTHETIC PATHWAY

The phage PIS136 infects a large number of actinomycetes genera. The strain *Amycolatopsis mediterranei* MTCC17 is one of those members of the group to which the phage PIS136 can infect. The strain of *A. mediterranei* MTCC17 produces rifamycin B. August *et al.* (1998) sequenced the entire rifamycin biosynthetic pathway and showed that the rifamycin is synthesized by a modular Type I polyketide synthase. The rifamycin gene cluster has 10 modules which use 3-amino-5-hydroxybenzoic acid as a starter molecule. It has been demonstrated (Katz and Donadio, 1993; Hutchinson and Fujii, 1995) that the enzyme of the modular Type I polyketide synthases (PKSs) are highly flexible therefore it is not difficult either to replace a starter molecule (Rinehart, 1977; Omura *et al.*, 1980; Denoya *et al.*, 1995; Jacobsen *et al.*, 1997) or replace a module of one pathway with a module from another pathway (Hopwood *et al.*, 1985; Epp *et al.*, 1989; Weber *et al.*, 1991; Donadio *et al.*, 1993; McDaniel *et al.*, 1993; Neimi *et al.*, 1994; Ruan *et al.*, 1997; Sezonov *et al.*, 1997; Bohn *et al.*, 1998; Ritsema *et al.*, 1998; Tropf *et al.*, 1998; McDaniel *et al.*, 1999; Tang *et al.*, 2000; Wilkinson *et al.*, 2000; Xue *et al.*, 2000).

We have demonstrated that the phage PIS136 has a variable host range. We have also demonstrated that the phage gives two populations: One with wide host range and the other with narrow host range, which are genetically identical. Therefore, the variation in the host range is likely to be due to variable expression of some of the genes. In the present study, it was thought that instead of *in vitro* manipulation of any antibiotic biosynthetic pathway, one can modulate the antibiotic biosynthetic pathway *in vivo* by using the inversion property of the phage PIS136. Therefore, *A. mediterranei* MTCC17 was infected with the phage PIS136 and four
lysogens were selected. All the four lysogens produced different color. Therefore it was assumed that the metabolic products of the four lysogens: AR4, AR14, AR17 and AR35 will be different. The ethylacetate extract of the fermentation broth of these lysogens were separated on TLC. All the four lysogens produced several compounds which were specific for that particular lysogen (Fig. 27). The result suggests that the phage integrates at random sites in the host genome and alters metabolic pathways. Integration at random sites in the host genome and then alteration/mutation of metabolic pathways appears to be similar to the properties of the phage Mu (Taylor, 1963). However, the present the mechanism of integration of the phage PIS136 is not known.

Assuming that lysogens AR4, AR14, AR17 and AR35 may produce some novel compounds, bioactivity of these lysogens was tested against the rifampicin resistant *E. coli* strains CSH108 and MTCC 1583. The lysogen AR17 (Fig. 28) did show activity against the rif*E. coli*. Since the *A. mediterranei* MTCC17 is not known to produce any other antibiotic but rifamycin, it is tempting to suggest that the lysogen AR17 may be producing a derivative of rifamycin which is active against rif*E. coli*. However, one cannot rule out the possibility that a new but yet unknown pathway was activated by the phage PIS136 and the killing of rif*E. coli* may be related to the product of that pathway and not to rifamycin.

Various metabolites produced by AR4, AR14, AR17 and AR35 were purified. The lysogen AR17 produced many compounds out of which two metabolites were identified phenylacetic acid (Fig.29) and 4-hydroxybenzoic acid (Fig.30). It has been reported that the 3-hydroxybenzoic acid and 3-5-dihydroxybenzoic acid can act as a starter molecule in place of the natural starter molecule, 3-amino-5-hydroxybenzoic acid (AHBA). However, whenever 3-hydroxybenzoic acid or 3,5-dihydroxybenzoic acid act as a starter molecule, the polyketide chain never grows beyond tetraketide stage (Hunziker et al., 1998; Hu et al., 1999).

NMR spectra of two compounds AR17(7) (Fig.31) and AR14-7(3) (Fig.32) which were produced by AR17 and AR14 respectively, showed the presence of a long polyketide chain (longer than the one reported in literature) over a starter molecule which is not a 3-amino-5-hydroxybenzoic acid i.e. natural starter molecule.
It has been reported that the 3-amino-5-hydroxybenzoic acid is synthesized by an alternate shikimic acid pathway (August et al., 1998). The results described in the present study suggest that either the shikimic acid or the alternate shikimate pathway is altered. The alteration in biosynthesis is leading to accumulation of various benzene derivatives: phenylacetic acid, 4-hydroxybenzoic acid, and may be some more. However, it is clear that the lysogens have failed to produce AHBA. It appears that some of the benzene derivatives are getting diverted to the rifamycin biosynthetic pathway and due to the plasticity of the PKS, the polyketide chain is growing over unnatural loading molecules. The incorporation of the unnatural loader molecules in the rifamycin biosynthetic pathway may lead to novel/hybrid rifamycin derivatives. It appears that the accumulation of AR17(7) and AR14-7(3) metabolites is the result of the growth of the polyketide chain over some benzene derivative which is not 3-amino-5-hydroxybenzoic acid. Since these metabolites are produced in very small amounts, that the polyketide chain has grown longer than what has been reported so far.