J. Indian Inst. Sci., Jan.–Feb. 1993, 73, 31–45.

Molecular biology of mycobacteria and mycobacteriophages

K. P. GOPINATHAN
Department of Microbiology and Cell Biology, and Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012.

Received on December 11, 1992.

Abstract

The salient features of the work related to the genetics and molecular biology of mycobacteria carried out by our group at the Department of Microbiology and Cell Biology for the past 20 years have been summarized here. The emphasis has mostly been on the characterization of the components of the synthetic machinery for macromolecules such as DNA, RNA and proteins in order to correlate a molecular basis for the 'slow growth' phenotype of mycobacteria, and to define possible targets for drug action. The molecular mechanisms of the action of existing antitubercular drugs and the development of resistance against these drugs by mycobacteria have also been investigated. Another significant contribution has been the isolation of a transducting mycobacteriophage, phage I3 and the detailed characterization of its physiology and biochemistry including genome organization. Investigations on bacteriophages have proved to be one of the most important approaches to understand the host, as seen from such parallel studies carried out on other bacterial genera like Escherichia or Salmonella. The establishment of conditions for transduction, transfection, and spheroplast formation, fusion and regeneration in order to provide a direct genetic access to mycobacterial research has also been a noteworthy contribution. More recently, the application of recombinant DNA methodologies have permitted the cloning and characterization of genes from mycobacteria and mycobacteriophages, revealing a variety of features ranging from the codon usages to the organisation of transcriptional promoter elements, and thus bringing us a step closer towards the understanding of the molecular mysteries of mycobacterial growth.

Key words: Aminoacyl tRNA synthetases, DNA synthesis, drug resistance, isoniazid, mycobacteriophage I3, pathogenicity, promoter consensus sequence, protein synthesis, RNA polymerase, streptomycin.

1. Background and rationale

The genus *Mycobacterium* encompasses a variety of clinically important pathogens such as *M. tuberculosis* and *M. leprae* responsible for causing tuberculosis and leprosy in human beings. These diseases continue to be major public health problems in India as well as many countries of the developing world. Further, the spectre of AIDS with concomitant mycobacterial infections has brought tuberculosis back into the western world where it was once considered to be well under control.

According to the recent WHO progress and evaluation report of the Tuberculosis Control Program (EB 87/4, November 1990), 1700 million people all over the world are infected with tuberculosis. Many cases are due to endogenous reactivation of the infection caused in the past rather than the result of recent infections; nevertheless, 8 million new cases of tuberculosis are added on each year and 2.9 million people die due to this disease, world wide. Leprosy is unique, and the fear, abhorence and social stigma associated with it lingers even today. This disease is dreaded not because it killed but because it left one alive, with extremely crippling mutilations. The prevalance of tuberculosis and leprosy is not due to the nonavailability of effective drugs but more due to the prolonged duration of chemotherapy and the general socioeconomic problems of the inflicted people. The treatment becomes more complicated due to the occurrence of the drug-resistant variants of the causative organism.

In this scenario, research in mycobacteria has gained considerable momentum in the past few years and today the mycobacterial species are amongst the most intensely studied bacterial pathogens. A wealth of information has accumulated on the physiology and biochemistry of mycobacteria. Some information has been gathered on the immunopathology of mycobacterial infections and the characterization of mycobacterial antigens. Genetics of mycobacteria has become one of the most advancing fields in bacterial genetics at present due to the application of recombinant DNA methodologies. However, at the time the work on genetics and molecular biology of mycobacteria was initiated at the Department of Microbiology and Cell Biology in the late 1960s and early 1970s, as a natural continuation of the earlier investigations on mycobacterial metabolism initiated during the mid-1950s and early 1960s, very few laboratories in the scientifically advanced countries were interested in this group of organisms. This was mainly due to the fact that: (i) mycobacterial infections were well under control and were no longer considered a major public health problem in these countries, and (ii) the organisms belonging to this genus were generally more difficult to grow in culture and were not amenable to most of the conventional techniques of bacterial genetics.

The earlier research work on mycobacteria from this Department related mostly to finding out the differences in intermediary metabolism between virulent and avirulent strains, in order to provide a molecular basis for the virulence of *M. tuberculosis*. Although quantitative differences in the operation of dissimilative pathways of energy metabolism or in the levels of individual enzymes between the virulent and avirulent strains were discernible from such studies, no direct correlation was attributable to the virulence or pathogenicity of mycobacteria. These aspects of research from here have been reviewed periodically[1,2] and therefore I shall confine only to those areas not covered previously.

2. Objectives and achievements

2.1. *Molecular biology of mycobacteria*

A basic understanding of the genetics and molecular biology of the causative organism

is essential for achieving effective control of the disease. An obvious starting point to study the molecular biology of mycobacteria is the phenomenon of slow growth. Why does *M. tuberculosis* grow so slowly and *M. leprae*, not at all *in vitro*? This characteristic is probably crucial to their success as parasites and has been possibly selected in order to facilitate their interaction with the eukaryotic hosts. The inert nature may also provide them with the selective advantage of overcoming the effects of bacteriostatic or bactericidal agents in the course of chemotherapy. Understanding the crucial cellular processes such as replication, transcription and translation, may prove to be potential targets for novel drug design. The bulk of our knowledge on prokaryotic gene regulation is derived from *E. coli* cells with a division time of 20–30 min. Are there fundamental differences in bacterium which divides only once in 24 hours (*M. tuberculosis*) or once every week or so (*M. leprae*)? Apart from the academic challenge of these problems, they may also provide novel perspectives relevant to the disease control. The work here has been mostly carried out with the virulent strain of tubercle bacillus *M. tuberculosis* $H_{37}R_v$ or its avirulent variant $H_{37}R_a$, and the saprophytic fast-growing species *M. smegmatis* as prototypes.

2.1.1. *DNA replication and RNA synthesis in mycobacteria*

On the premise that the slow growth of *M. tuberculosis* reflects retarded rates in macromolecular synthesis and a consequent step down of metabolic machinery, the rates of DNA and RNA chain elongations in this organism have been investigated[3,4]. The time required for a DNA replication fork to traverse the chromosome from the origin of replication to the terminus, designated as C period, is a fundamental parameter of bacterial growth. Assuming a single replication origin and a bidirectional replication fork, for *M. tuberculosis* $H_{37}R_v$ growing in a synthetic medium, the C period was 620 min as compared to 105 min for the fast-growing *M. smegmatis* or 40–55 min for different *E. coli* strains. This corresponded to a DNA chain elongation rate of 3200 nucleotides per min in *M. tuberculosis* as against 39000–54000 nucleotides per min in *E. coli*, 13–17 times slower than the latter. However, the DNA chain elongation in the fast-growing *M. smegmatis* at 32400 nucleotides per min was comparable to *E. coli*, although the doubling time of the former was about 2–3 times slower. The rate of RNA chain elongation in *M. tuberculosis* at the rate of 240 nucleotides per min is also considerably slower as against 2400 nucleotides per min seen in *E. coli*. Whether these slow chain growth rates primarily or exclusively contribute to the slow-growing nature of *M. tuberculosis* is not established yet. The intrinsic catalytic activities of DNA polymerase I and RNA polymerase from *M. tuberculosis* are also of lower efficiencies compared to their counterparts from *E. coli*[5,6]. The mycobacterial RNA polymerases are multisubunit complexes comprising two α, and one each of β and β' subunits as well as the σ factor[6,7] resembling the prototype prokaryotic RNA polymerases.

2.1.2. *Protein synthesis and the action of streptomycin in M. tuberculosis*

Streptomycin, though no longer used as extensively for the treatment of tuberculosis as in the past due to its high toxicity on prolonged administration, gained prominence

in clinical medicine as a powerful antituberculous drug. This antibiotic has been established to be a potent inhibitor of prokaryotic protein synthesis at the level of 30S ribosomes using *E. coli* as the target organism[8]. Streptomycin inhibited the protein synthesis in whole cells and cell-free extracts of the drug-sensitive strains of *M. tuberculosis*[9]. Based on the effect of the drug on the *in vitro* protein-synthesizing system, it has been shown that the site of streptomycin sensitivity resides on the ribosome but the site of resistance can either be the ribosome or the bacterial membrane. In the latter case, drug resistance is reversible and under conditions of altered permeability, streptomycin could still prove to be lethal to those strains. A detailed analysis of the mechanism of inhibition of protein synthesis by streptomycin has shown that the drug breaks down the polypeptide chain initiation complex (comprising 30S ribosome–fmet tRNA–mRNA) rather than by blocking the complex formation itself[10]. These studies lend support to the hypothesis that streptomycin acts by preventing chain extension when added to a protein-synthesizing system. The higher sensitivity of *M. tuberculosis* to streptomycin compared to *E. coli* was at least partially attributable to the lower population of ribosomes within the mycobacterial cells.

2.1.3. *Transfer RNA and aminoacyl tRNA synthetases*

Extensive studies have been carried out[11–19] on tRNA and the aminoacyl tRNA synthetases from *M. tuberculosis* and *M. smegmatis*, as they constitute important components of the protein synthetic machinery. Fractionation of tRNA preparations from *M. smegmatis* on a variety of chromatographic columns has revealed the typical prokaryotic distribution of the isoacceptor species[11]. There were at least two isoacceptor species for methionine, of which only one could be formylated. The enzyme transformylase (Met $tRNA^{met}$ transformylase) was also present in the mycobacterial extracts. Subsequently, the complete nucleotide sequence of the initiator tRNA from *M. smegmatis* has been determined[12]; a striking feature of this tRNA sequence is the absence of the highly conserved ribothymidine.

The studies on aminoacyl tRNA synthetases from mycobacteria have revealed some interesting features and have contributed to the understanding of the complex kinetic mechanisms followed by this class of enzymes. The synthetases extensively characterized include met–tRNA synthetase[13], val- and ile-tRNA synthetases[14–17], and arg– and lys–tRNA synthetases[18,19], from *M. smegmatis* and *M. tuberculosis*.

2.1.4. *Mechanism of aminoacylation reactions*

The aminoacyl tRNA synthetases catalyze the esterification of tRNAs *via* a reaction in which three substrates (ATP, aminoacid [AA], and tRNA) are converted into three end products (AMP, PPi and AA–tRNA). The fidelity in protein synthesis during the decoding of genetic message is totally dependent on the error-free esterification of the cognate amino acid and tRNA species by the synthetases. How do these

enzymes discriminate between the different species of tRNA without bias to the isoacceptor species for a given amino acid-specific tRNA, whose structures are so closely related? For these reasons the mechanism of aminoacyl tRNA synthetases has proved to be a favourite topic of molecular enzymology.

The aminoacylation reactions are conventionally represented as a two-step process, (i) and (ii) to give the total reaction (iii).

$$\text{(i)}\quad AA + ATP \rightleftharpoons AA - AMP + PPi$$

$$\text{(ii)}\quad AA - AMP + tRNA \rightleftharpoons AA - tRNA + AMP$$

$$\text{(iii)}\quad AA + ATP + tRNA \rightleftharpoons AA - tRNA + AMP + PPi.$$

The enzymatic mechanism for aminoacyl tRNA synthetases has been controversial as to whether the reaction proceeds in a two-step, sequential manner [(i) + (ii)] to give the total reaction (iii), or in a single step, concerted mode[20]. While 17 out of 20 of the aminoacyl tRNA synthetases could catalyze the partial PPi exchange reaction independent of the tRNA, the synthetases for arg–, glu– and gln required the presence of tRNA to carry out the partial reaction. No enzyme-adenylate intermediate could be isolated for the latter enzymes, and therefore a one-step pathway (iii) for tRNA changing was proposed.

Pure preparations of val–, ile–, arg– and lys–tRNA synthetases from *M. smegmatis* have been exploited thoroughly to investigate the mechanism of aminoacylation reaction[14–19]. Detailed kinetic analyses have now established that the aminoacylation proceeds either by the concerted mode or the sequential mechanism in the case of val– and ile–tRNA synthetases depending on the internal concentrations or availability of cations such as spermine and magnesium[14–17]. Likewise, the initial velocity and product inhibition studies with arg– and lys–tRNA synthetases from *M. smegmatis* has revealed a *rapid-equilibrium random ter ter mechanism* for both these enzymes[18,19]. Although these two synthetases belong to two different classes (arg–tRNA synthetase requiring cognate tRNA to catalyze the ATP–PPi exchange reaction), they follow a common reaction mechanism.

2.2. *Action of antitubercular drugs and the development of drug resistance by mycobacteria*

Prominent amongst the antitubercular drugs in use have been isoniazid (INH), streptomycin, rifampicin, PAS, ethambutal, pyrazinamide and ethionamide, the last three being used mostly as second-line drugs. Despite the availability of these effective drugs, the clinical treatment and eradication of tuberculosis have proved to be difficult mainly due to the extended duration of chemotherapy, which in turn has been warranted by the slow-growing nature of the causative organism. The prolonged and intermittent exposure of the tubercle bacillus in the patients leads to frequent development of drug resistance by the organism, which complicates the treatment. The existence of drug-resistant variants of the tubercle bacillus has proved to be a clinical problem, automatically warranting the development and use of second-line drugs. In this context, the understanding of the molecular basis of drug action on mycobacteria attains special significance.

The action of streptomycin on the tubercle bacillus has been established by us[9, 10] to be on the protein synthesis at the level of 30S ribosomes. Rifampicin acts at the level of DNA-dependent RNA polymerase and the extreme sensitivity of the mycobacterial RNA polymerase to rifampicin correlates well with the high efficiency of the drug against mycobacteria[6,7]. Development of drug resistance at both these targets could be readily achieved in the laboratory. The exploitation of the RNA polymerase as a potential target for designing drugs against pathogenic mycobacteria is still a distinct possibility[21]. The action of ethionamide, an active, tuberculostatic, second-line drug which plays an important role in the retreatment of tuberculosis resistant to one or more classical drugs has also been examined by us in some detail[22–24]. The antitubercular activity of this drug is more pronounced *in vivo* in infected animals because the antibacterial action is exerted through the activated metabolite, ethionamide sulfoxide.

2.2.1. *Isoniazid (INH)*

This powerful antituberculous agent has occupied a unique position in the treatment of tuberculosis because of its extreme effectiveness. The drug brings about complete inhibition of growth of *M. tuberculosis in vitro* at 0·01–0·1 μg/ml. The special features of the action of isoniazid are: (i) it is highly specific in its action against mycobacteria, and (ii) is effective at very low concentrations. The action of isoniazid has been reported to be mediated through the blocking of a variety of metabolic pathways including the biosynthesis of nucleic acids, proteins and lipids. The two most intriguing questions on the action of INH, however, are the specificity of the drug against mycobacteria (and why it is not effective against other microorganisms) and the ready development of resistance against the drug (despite the presence of multiple targets identified for exerting its effects). Analysing the genetic and biochemical aspects associated with the development of isoniazid resistance by mycobacteria, we have proposed[25, 26] a unified model invoking a peroxidase in the 'uptake' of the drug by the organism. This multifunctional peroxidase interacts with the drug and transports it into the cells. The inability of the organism to take up isoniazid renders it resistant to the drug. The resistance/sensitivity to INH by mycobacteria is determined by one genetic locus, which codes for the peroxidase, responsible for the uptake of the drug[27]. Once inside the cell, the drug exerts the antibacterial action by blocking one or more of the proposed biochemical steps. The specificity of isoniazid action could be attributed to the presence of this unique peroxidase capable of transporting the drug into the bacterial cell. By mycobacteriophage I3-mediated transduction of INH sensitivity or resistance markers resulting in a concomitant gain or loss of three enzyme activities (catalase, peroxidase and Y enzyme) as well as the capacity for the uptake of the drug, we had provided the first-ever genetic evidence for the molecular basis of INH action or the development of INH resistance by mycobacteria[27]. Very recently, these results have been reproduced by others[28], making use of the recombinant DNA approach with a cloned DNA fragment harbouring the catalase/peroxidase gene from *M. tuberculosis*.

2.2.2. β-*Lactam antibiotics*

β-Lactam antibiotics kill the target bacteria by three completely different mechanisms:

rapid lysis, production of spherical cells and filamentation, or a combination of these. The demonstration that penicillin binds covalently to the enzymes which it inhibits has provided a method for the detection of penicillin-sensitive enzymes as penicillin-binding proteins (PBPs). All the bacteria examined so far contain between four and seven PBPs in their cytoplasmic membrane, and their inactivation can result in cell death. The PBPs are generally enzymes involved in cell wall biosynthesis, such as transglycosylases and transpeptidases (cell wall elongation), peptidoglycan insertions, formation of cross wall at cell division, D-alanine carboxypeptidase (maturation of peptidoglycan), etc. Despite the fact that PBPs have been identified in mycobacterial species, they are generally resistant to β-lactams. The permeability barrier of the antibiotics to the cell as well as the presence of β-lactamases contribute to the intrinsic resistance in mycobacteria to penicillins. The β-lactamases are generally intracellular in location and found in the extracellular medium primarily due to the autolysis of mycobacteria during growth[29]. At least in some strains, it appears that the β-lactamase could be coded by both chromosomal and extrachromosomal genes. In their biochemical properties, the mycobacterial β-lactamases resemble both gram-positive and gram-negative TEM enzymes[30, 31]. Treatment of *M. smegmatis* SN2 with plasmid-curing agents such as acridine orange results in a decreased resistance to penicillin, suggesting the presence of plasmids in this species. Using a gentle lysis method, the presence of plasmids could be demonstrated in several fast-growing species of mycobacteria[32], but the presence of the β-lactamase gene on the plasmids is not yet established.

2.3. *Molecular biology of mycobacteriophage I3*

Although the operation of classical methods of genetic exchanges has been demonstrated in mycobacteria, our grasp on the molecular genetics of this genus has lagged behind until the application of the recombinant DNA methodologies has come into vogue in recent times. Prior to this, a method established for the mycobacterial spheroplast isolation, fusion and regeneration by us[33, 34] provided an opportunity to explore the genetic exchanges, to a limited extent at least. Bacteriophages have contributed immensely towards our present-day knowledge in genetics and molecular biology. In this context, the isolation of the transducing mycobacteriophage I3 from Bangalore soil[35] has attained tremendous significance. This phage has been utilised to transduce auxotrophic- and drug-resistance markers in the genetic analysis of *M. smegmatis*[27, 35, 36]. Nearly 250 mycobacteriophages have been reported in literature but the extensive investigation on the biology of phage I3 conducted here makes it one of the best characterized mycobacteriophages. However, very recently, the 50 kbp genome of mycobacteriophage L5 is reported to have been sequenced completely (G. Hatful, personal communication).

Morphologically, mycobacteriophages are similar to those found in other bacterial genera, and most of them show the typical head and tail structures. Mycobacteriophage I3 has a large hexagonal head (80 ± 10 nm) and a long tail (80 ± 10 nm) which includes a base plate, collar, contractile tail sheath and tail cannaliculus

resembling the tail structure of T-even phages[37] Phage I3 tail can contract up to 60% and so far it is the only mycobacteriophage with a contractile tail, thus belonging to Bradleys type A.

2.3.1. *Growth properties*

Most of the mycobacteriophages are polyvalent, *i.e.*, they grow on more than one species of mycobacteria. In conformity with the growth patterns of hosts, mycobacteriophages in general are slow growing. Phage I3 has a growth cycle comprising a latent period of 180 min, an eclipse period of 120 min and a rise period of 60 min; the burst size usually varies between 30 and 60. Tween 80 (polyoxyethylenesorbitan monooleate), a nonionic surface-active agent used in the growth medium of mycobacteria to prevent clumping, inhibits the propagation of mycobacteriophage I3. This inhibition is at the level of the DNA injection from the phage to the host bacterium while the detergent has no influence on the adsorption between them[38]. The DNA injection process takes about 20 min to complete and is totally dependent on Ca^{++} ions[39]. An '*in vitro*' DNA injection system was subsequently set up to analyse this process at a molecular level[40]. The model proposed for DNA injection process in phage I3 infection involves the contraction of tail sheath on contact with the cell wall/cell membrane of the host in presence of Ca^{++} ions and with concomitant hydrolysis of ATP due to the presence of an associated ATPase activity. The ATP molecules for this process in the extracellular environment are made available by the phage itself, which carries nearly 150 molecules of NTP per particle[41].

2.3.2. *Biochemical properties*

In addition to nucleic acids and proteins, mycobacteriophages also contain lipids, which render them sensitive to organic solvents. Although phage I3 is relatively resistant to inactivation by chloroform, a lipid content of 16% (phospholipids and neutral lipids make up 69 and 31% respectively) has been demonstrated[42]. The presence of lipids in I3 has been confirmed by electron microscopy which showed a membranous bilayer structure around the phage head and tail. The lipids appear to be essential for maintaining the structural integrity and infectivity of the phage.

The structural proteins of phage I3 can be resolved into more than 30 bands on polyacrylamide gel electrophoresis. More refined analysis using antibodies raised against the total phage proteins in combination with the Western blot techniques have revealed the presence of at least 35–36 individual protein bands[43].

2.3.3. *Genomic characterization*[44–46]

All the mycobacteriophages have linear, double-stranded DNA and so far, the best characterized genome of any mycobacteriophage is that of phage I3. The GC content of mycobacteriophage I3 DNA is 67% as determined from thermal melting analysis, buoyant density in CsCl gradient and by direct base composition analysis. High-resolution melting of I3 DNA reveals that the base distribution is random, and there is

no indication of the presence of modified bases to any significant extent. The restriction cleavage patterns for a variety of restriction endonucleases on this 140 kbp genomic DNA have been determined and a partial restriction map of the genome has been constructed. The genome being circularly permuted, the packaging of DNA during phage maturation is inferred to follow the 'headful' mechanism.

The genomic DNA of I3 exhibits certain unusual features. For instance, when denatured with alkali, heat, formamide or DMSO, the DNA breaks down to heterogenous-sized, single-stranded fragments smaller than the expected intact unit genomic length due to the presence of DNA strand interruptions. These interruptions on I3 DNA are host-independent, and their location is random.

The nature of the random single-strand interruptions on phage I3 DNA has been characterized in detail and proved to be in the form of short gaps. About 13–14 gaps are present per DNA molecule and the average length of the gap is about 10 nucleotides. Nevertheless, the genomic DNA is biologically active. Such random interruptions on the genomic DNA could have arisen due to the limitations in the postreplicational repair processes. Although the significance of these interruptions on the genomic DNA of phage I3 is not known, it is hypothesized that their presence would facilitate the compaction and packaging of the genomic DNA into the phage head, considering the high GC content and the consequent rigidity of the backbone structure of the DNA. This interrupted genomic organization may also have a bearing on the recombinational status of the host organism.

The genomic characterization has subsequently been extended to the other mycobacteriophages isolated in our laboratory[47]. The restriction patterns of phage I1 and I5 DNAs are similar to that of I3 DNA but the DNA of phage I8, a lytic phage, differed from them in many respects. Phage I8 DNA is much smaller (~ 34 kbp), has a GC content of 54% and does not harbour any single-strand interruptions. Based on restriction endonuclease cleavage patterns, and hybridization analyses, the DNAs of phages I1, I3 and I5 have been shown to be homologous and indistinguishable but entirely different from phage I8.

A few conditional lethal and plaque morphology mutants of phage I3 have been isolated. Making use of the ts and clear phaque-forming mutants, by two- and three-factor crosses, a recombination map of the phage I3 chromosome has been constructed[48].

2.3.4. *Metabolism in mycobacteriophage-infected cells*

During intracellular development, a temporal sequence of events is seen following phage I3 infection[49,50]. RNA and protein synthesis precede the commencement of DNA replication. Being a temparate phage, I3 does not degrade the host DNA upon infection. Several enzyme activities are altered in phage I3-infected *M. smegmatis* and most significantly, there is a tremendous increase in the RNA polymerase activity. Judicious use of the antibiotic rifampicin, a potent inhibitor of RNA polymerase, has shown that in the case of phage I3 development, the host RNA polymerase is

continuously utilised for phage development. The exploitation of rifampicin–Sepharose affinity matrix for the purification of RNA polymerases yielded homogenous preparations of the enzyme in a two-step purification protocol[7]. The subunit structure of RNA polymerase from phage-infected cells is altered due to the presence of an additional polypeptide subunit of 79 kDa. The template specificity of the enzyme from phage I3-infected cells is considerably different and it could utilise the phage DNA template ten times more efficiently. The continued involvement of DNA gyrase in phage I3 DNA replication and transcription processes during the entire phage development period has also been established[51].

2.3.5. *Transfection, transduction and genome mapping*

Transformation of bacteria by isolated phage DNA resulting in the production of viable phages is called transfection. Although the transformation efficiency in mycobacteria is of low order, transfection of sensitized cells of *M. smegmatis* SN2 by phage I3 DNA could be achieved[52]. The efficiency of this system though not comparable to the transfection of *E. coli* has potential to be exploited for carrying out gene transfer in mycobacteria. However, more efficient gene transfer systems utilising electroporation and mutant strains of *M. smegmatis* which are more prone to transfection are currently available.

The process of phage-mediated gene transfer in bacteria is termed transduction. Because of the limited number of genes transferred, transduction is especially useful in fine structure mapping of bacterial genome. Phage I3 has been shown to transduce auxotrophic markers[35], or the determinants for resistance and susceptibility to isoniazid[27] and resistance to streptomycin[36] and thus behaves as a generalized transducing phage.

2.3.6. *Cloning of phage I3 genome*

A genomic library of the phage I3 DNA has been constructed in *E. coli* using different plasmid vectors such as pBR322, pTZ and pUC series of vectors[43]. Such libraries could be exploited for mapping and sequencing I3 DNA segments which have otherwise proved to be difficult due to the presence of random single-stranded interruptions on the phage genome. The features of genomic organization (monocistronic/polycistronic), the nature of transcriptional and translational initiation and termination signals and the patterns of codon usage in mycobacteria could be deciphered from these studies. Alternatively, strategies have also been adopted to clone the transcriptional promoter sequences from phage I3 by using direct promoter selection vectors (unpublished). Several such promoter elements derived from phage I3 genome as well as some of the phage structural protein genes have now been sequenced to generate the mycobacterial 'promoter consensus' sequences. These sequences differ noticeably from the prototype-promoter consensus generated from *E. coli*. Since most of the reported promoter structures in other organisms are 'AT' rich, can one expect the presence of some unusual features in mycobacterial promoters considering the high GC contents of their DNA, which in turn might explain the lower transcriptional efficiency consequently leading to slow growth, a characteristic feature of mycobacteria.

3. Future perspectives

The ultimate objective of any research on pathogenic organisms should be directed to control the disease either by prevention or by effective (efficient) chemotherapy and finally eradicate it from the face of the earth. A basic understanding of the genetics and molecular biology of the disease-causing organism is essential for the rational approaches in chemotherapeutic designs, in deciphering the mechanism of drug action and the molecular basis of the development of drug resistance. Molecular geneticists have spent decades in understanding *Escherichia coli* to have the vast amount of knowledge. Surely reproducing all that with mycobacteria is not called for. Can we, however, use that knowledge to identify those aspects of mycobacterial genetics and biology which are most relevant to their pathogenicity or which might prove to be most useful for their elimination from the infected host. In the light of the past experiences, and the major objectives and goals defined, what are the future perspectives in mycobacterial research here in the Institute or elsewhere in the country. Given the importance of mycobacterial infections in public health management in the country, I have identified and listed below some potential areas for future research without bias or prioritisation.

3.1. *Drug design and drug targets*

The identification of suitable and unique target sites on the pathogen constitutes a major component of designing effective drugs, to be followed up by logical design based on computer-aided molecular modelling. The distinctive features of the mycobacterial cell wall organization and the consequent existence of unique metabolic pathways for their synthesis and assembly should provide excellent opportunities for genetic and biochemical analysis and offer target sites for rational drug design. The accessibility of drugs to the target site is warranted as an important criterion.

3.2. *Pathogenesis by mycobacteria*

The complex phenomenon of mycobacterial pathogenicity is virtually uninvestigated. The intriguing features of virulence range from the invasiveness of pathogens, to their capacity to survive within the host evading the host defence mechanism and how the pathological lesions or symptoms are caused. At present, the information on how the ingested mycobacteria is killed by macrophages, or eliminated from the host is also very limited.

3.3. *Vaccines against mycobacteria*

Protective immunity and pathological immunity are two faces of mycobacterial infections. The involvement of the same or different sets of T cell populations of the host is critical in the above process. To find a 'single' protective antigen derived from the pathogen may be nearly impossible and therefore identification of a group of 'disease associated' antigens will be a safer bet. While developing new and more effective vaccines against mycobacterial infections, these aspects are to be considered. Besides, the genetic predisposition of the individual plays a crucial role in susceptibility to disease. The development of BCG as a multipotent vaccine may hold great potential.

3.4. *Diagnostics*

Sensitive, specific, reliable, easy to perform diagnostic tests for early diagnosis of mycobacterial infections, based on immunological methods and recombinant DNA (RNA) probes are currently available. The introduction of PCR techniques has increased the sensitivity of detection from 1–5 bacteria. However, the question remains whether the presence of such a small number of bacteria is indicative of clinical infection. The diagnostic probes already developed elsewhere could be easily adapted here without much research efforts. Moreover, determining the viability of the detected bacteria in the host is an important feature to be considered.

3.5. *Molecular biology of mycobacteria*

A basic understanding of the genetics and molecular biology of the causative organism is an essential requirement for understanding the disease or its elimination. The fascinating features of mycobacteria are the extremely slow-growing nature (long division times) and the capacity to stay dormant within a host while retaining the ability to switch back to active division. These features are of significance in mycobacterial evasion of immune mechanisms of the host or drug action, as well as in the development of drug resistance. Molecular and genetic aspects of mycobacteria are now accessible for experimentation and a genetic approach to understand the virulence is feasible.

3.6. *Genome sequencing*

Is the sequencing of the complete genome of mycobacteria (be it *M. tuberculosis*, *M. leprae* or any other pathogenic species) worthwhile? The technology is now available in the world but is still being improved. The physical 'contig' maps of *M. tuberculosis* and *M. leprae* have already been constructed and cosmid/phage libraries have become available for distribution. However, financial and material inputs are high and the availability of trained manpower is essential. The basic information would ultimately reveal the 'unique' and 'unusual' features of mycobacteria. Utmost care should be taken to choose the right strain of mycobacteria for sequencing. Perhaps the genomic sequencing of mycobacteria has already commenced elsewhere.

Acknowledgements

The research work from the author's laboratory has been supported by grants from the Department of Science and Technology, the Department of Atomic Energy, and the Department of Biotechnology, Government of India. Research in the Microbiology and Cell Biology Department is also supported by the COSIST and Special Assistance Grants from the University Grants Commission. The author is thankful to the World Health Organization, Geneva, for providing substantial information on the current world status of research on mycobacteria. Technical assistance from Ms Rekha Ramaswamy in carrying out the work from the author's lab is gratefully acknowledged.

References

1. RAMAKRISHNAN, T., MURTHY, P. S. AND GOPINATHAN, K. P. — Intermediary metabolism of mycobacteria, *Bact. Rev.*, 1972, **36**, 65–108.
2. RAMAKRISHNAN, T. — Twentyfive years of research on tuberculosis at the Indian Institute of Science, Bangalore, *Curr. Sci.*, 1983, **52**, 398–402.
3. HIRIYANNA, K. T. AND RAMAKRISHNAN, T. — Deoxyribonucleic acid replication time in *Mycobacterium tuberculosis* $H_{37}R_v$, *Arch. Microbiol.*, 1986, **144**, 105–109.
4. HARSHEY, R. M. AND RAMAKRISHNAN, T. — Rate of ribonucleic acid chain growth in *Mycobacterium tuberculosis* $H_{37}R_v$, *J. Bact.*, 1977, **129**, 616–622.
5. HIRIYANNA, K. T. AND RAMAKRISHNAN, T. — Purification and properties and DNA polymerase from *Mycobacterium tuberculosis* $H_{37}R_v$, *Biochim. Biophys. Acta*, 1981, **652**, 274–282.
6. HARSHEY, R. M. AND RAMAKRISHNAN, T. — Purification and properties of DNA-dependent RNA polymerase from *Mycobacterium tuberculosis* $H_{37}R_v$, *Biochim. Biophys. Acta*, 1976, **432**, 49–59.
7. NAGARAJA, V. AND GOPINATHAN, K. P. — Transcriptional specificity after mycobacteriophage I3 infection, *J. Biosci*, 1987, **11**, 167–179.
8. SHAILA, M. S. AND GOPINATHAN, K. P. — Mechanism of streptomycin action–A review, *J. Scient. Ind. Res.*, 1974, **33**, 467–474.
9. SHAILA, M. S., GOPINATHAN, K. P. AND RAMAKRISHNAN, T. — Protein synthesis in *Mycobacterium tuberculosis* $H_{37}R_v$ and the effect of streptomycin in streptomycin-sensitive and streptomycin-resistant strains, *Antimicrob. Agents Chemother.*, 1973, **4**, 205–213.
10. SHAILA, M. S., GOPINATHAN, K. P. AND RAMAKRISHNAN, T. — Mechanism of streptomycin action in *Mycobacterium tuberculosis* $H_{37}R_v$, *J. Indian Inst. Sci.*, 1975, **57**, 301–316.
11. DEOBAGKAR, D. N. AND GOPINATHAN, K. P. — Studies on transfer RNA from mycobacteria, *Can. J. Microbiol.*, 1978, **24**, 693–702.
12. VANI, B. R., RAMAKRISHNAN, T., KUCHINO, Y. AND NISHIMURA, S. — Nucleotide sequence of initiator tRNA from *Mycobacterium smegmatis*, *Nucleic Acids Res.*, 1984, **12**, 3933–3936.
13. DEOBAGKAR, D. N. AND GOPINATHAN, K. P. — Two forms of methionyl tRNA synthetase from *Mycobacterium smegmatis*, *Biochem. Biophys. Res. Commun.*, 1976, **71**, 939–951.
14. NATARAJAN, V. AND GOPINATHAN, K. P. — Purification and properties of valyl tRNA synthetase from *Mycobacterium smegmatis*, *Biochim. Biophys. Acta*, 1979, **568**, 253–263.
15. NATARAJAN, V. AND GOPINATHAN, K. P. — Purification and properties of valyl tRNA synthetase reaction by polyamines, *J. Biosci.*, 1979, **1**, 357–367.
16. NATARAJAN, V. AND GOPINATHAN, K. P. — Stimulation of valyl- and isoleucyl tRNA synthetase from *Mycobacterium tuberculosis*, *Indian J. Biochem. Biophys.*, 1980, **17**, 330–334.
17. NATARAJAN, V. AND GOPINATHAN, K. P. — Mechanism of aminoacylation of tRNA: Influence of spermine on the kinetics of aminoacyl tRNA synthesis by isoleucyl- and valyl tRNA synthetases from *Mycobacterium smegmatis*, *Biochim. Biophys. Acta*, 1981, **654**, 94–101.
18. CHAR, S. AND GOPINATHAN, K. P. — Purification and kinetic mechanism of lysyl-tRNA synthetase from *Mycobacterium smegmatis* SN2, *Biochem. Int.*, 1985, **11**, 893–902.

19. CHAR, S. AND GOPINATHAN, K. P. — Arginyl-tRNA synthetase from *Mycobacterium smegmatis* SN2: purification and kinetic mechanism, *J. Biochem.*, 1986, **100**, 349–357.

20. GOPINATHAN, K. P. — Aminoacyl-tRNA synthetases. In *Transfer RNA and other soluble RNAs* (J. D. Cherayil, ed.), 1990, pp 103–127, CRC Press.

21. RAMAKRISHNAN, T. — RNA polymerase as the target for designing drugs against pathogenic mycobacteria, *Proc. Indo–UK Symp. Leprosy*, Central Jalma Institute for Leprosy, ICMR, Agra, India, 1986, pp 254–261.

22. PREMA, K. AND GOPINATHAN, K. P. — An oxygenase from guinea pig liver which catalyzes sulphoxidation, *Biochem. J.*, 1974, **143**, 613–624.

23. PREMA, K. AND GOPINATHAN, K. P. — Metabolism of ethionamide, a second line antitubercular drug, *J. Indian Inst. Sci.*, 1976, **58**, 16–27.

24. PREMA, K. AND GOPINATHAN, K. P. — Distribution, induction and purification of the enzyme, a monooxygenase, catalyzing sulfoxidation of drugs, *Biochem. Pharmac.*, 1976, **25**, 1299–1303.

25. GAYATHRI DEVI, B., SHAILA, M. S., RAMAKRISHNAN, T. AND GOPINATHAN, K. P. — The purification and properties of peroxidase from *Mycobacterium tuberculosis* $H_{37}R_v$ and its possible role in the mechanism of action of isonicotinic acid hydrazide, *Biochem. J.*, 1975, **149**, 187–197.

26. GAYATHRI DEVI, B., RAMAKRISHNAN, T. AND GOPINATHAN, K. P. — Studies on an enzyme reacting with isoniazid from *Mycobacterium tuberculosis* $H_{37}R_v$, *Proc. Indian Acad. Sci. B*, 1974, **80**, 240–252.

27. GOPINATHAN, K. P. — Specificity of isoniazid action against mycobacteria and a possible mechanism for the development of drug resistance, *Curr. Sci.*, 1981, **50**, 216–217.

28. ZHANG, Y., HEYM, B., ALLEN, B., YOUNG, D. AND COLE, S. — The catalase–peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*, *Nature*, 1992, **358**, 591–593.

29. CHOUBEY, D. AND GOPINATHAN, K. P. — Factors affecting the synthesis and distribution of β-lactamase in *Mycobacterium smegmatis* SN2 cultures, *Curr. Microbiol.*, 1986, **13**, 103–106.

30. CHOUBEY, D. AND GOPINATHAN, K. P. — Characterization of β-lactamase from *Mycobacterium smegmatis* SN2, *Curr. Microbiol.*, 1986, **13**, 171–175.

31. CHOUBEY, D. AND GOPINATHAN, K. P. — Characterization of β-lactamase from *Mycobacterium butyricum* ATCC19979, *Biochem. Int.*, 1986, **12**, 207–214.

32. GOPINATHAN, K. P., REDDY, A. B., CHOUBEY, D. AND SADHU, C. — Phages, plasmids and transfection of mycobacteria, *Proc. Indo–UK Symp. Leprosy*, Central Jalma Institute of Leprosy, ICMR, Agra, India, 1986, pp 270–286.

33. SADHU, C. AND GOPINATHAN, K. P. — A rapid procedure for the isolation of spheroplasts from *Mycobacterium smegmatis*, *FEMS Microbiol. Lett.*, 1982, **15**, 19–22.

34. SADHU, C. AND GOPINATHAN, K. P. — Expression of drug resistance markers by spheroplasts of *Mycobacterium smegmatis* after fusion, *J. Genet.*, 1988, **67**, 67–73.

35. SUNDAR RAJ, C. V. AND RAMAKRISHNAN, T. — Transduction in *Mycobacterium smegmatis*, *Nature*, 1970, **228**, 280–281.

36. SAROJA, D. AND GOPINATHAN, K. P. — Transduction of isoniazid sensitivity/resistance in *Mycobacterium smegmatis*, *Antimicrob. Agents Chemother.*, 1973, **4**, 643–645.

37. KOZLOFF, L. M., SUNDAR RAJ, C. V., RAO, R. N., CHAPMAN, V. A. AND DELONG, S. — Structure of a transducting mycobacteriophage, *J. Virol.*, 1972, **9**, 390–393.

38. GADAGKAR, R. AND GOPINATHAN, K. P. Inhibition of DNA injection from mycobacteriophage I3 by Tween-80, *Virology*, 1978, **91**, 487–488.

39. NAGARAJA, V. AND GOPINATHAN, K. P. Requirement of calcium ions in mycobacteriophage I3 DNA injection and propagation, *Arch. Microbiol.*, 1980, **124**, 249–254.

40. KARNIK, S. S. AND GOPINATHAN, K. P. Possible involvement of a calcium-stimulated ATP hydrolysing activity associated with mycobacteriophage I3 in the DNA injection process, *J. Virol.*, 1980, **33**, 969–975.

41. KARNIK, S. S. AND GOPINATHAN, K. P. Presence of nucleoside triphosphates and calcium associated with mycobacteriophage I3, *Arch Microbiol.*, 1981, **130**, 50–53.

42. GOPE, M. L. AND GOPINATHAN, K. P. Presence of lipids in mycobacteriophage I3, *J. Gen. Virol.*, 1982, **59**, 131–138.

43. RAMESH, G. R. AND GOPINATHAN, K. P. Cloning and characterization of mycobacteriophage I3 structural proteins and promoter elements (manuscript in preparation).

44. REDDY, A. B. AND GOPINATHAN, K. P. Presence of random single-strand gaps in mycobacteriophage I3 DNA, *Gene*, 1986, **44**, 227–234.

45. REDDY, A. B. AND GOPINATHAN, K. P. Characterization of genomic DNA of mycobacteriophge I3, *Curr. Microbiol.*, 1987, **14**, 241–245.

46. REDDY, A. B. AND GOPINATHAN, K. P. Existence of single-strand interruptions in the genomic DNA of mycobacteriophage I3, *FEMS Microbiol Lett.*, 1986, **37**, 163–167.

47. REDDY, A. B. AND GOPINATHAN, K. P. Characterization of mycobacteriophage I8 and its unrelatedness to mycobacteriophages I1, I3 and I5, *J. Gen. Virol.*, 1987, **68**, 949–956.

48. SADHU, C., GOPE, M. L. SADHU, K. AND GOPINATHAN, K. P. Isolation and characterization and mapping of temperature-sensitive mutants of mycobacteriophage I3, *J. Biosci.*, 1985, **7**, 39–47.

49. GOPINATHAN, K. P., SADHU, C. AND NAGARAJA, V. Interaction between phage and host chromosomes. In *Current trends in life sciences, XII Biological Macromolecules Symp.*, 1984, pp 89–96

50. NAGARAJA, V. AND GOPINATHAN, K. P. Macromolecular synthesis in mycobacteriophage I3 infected cells, *Mol. Biol. Rep.*, 1981, **8**, 11–15.

51. NAGARAJA, V. AND GOPINATHAN, K. P. Involvement of DNA gyrase in the replication and transcription of mycobacteriophage I3 DNA, *FEBS Lett.*, 1981, **127**, 57–62.

52. KARNIK, S. S. AND GOPINATHAN, K. P. Transfection of *Mycobacterium smegmatis* SN2 with mycobacteriophage I3 DNA, *Arch. Microbiol.*, 1983, **136**, 275–280.