SUMMARY

The presence of microflora, capable of utilizing various hydrocarbons as their sole source of carbon and energy is a well established fact. Hydrocarbons of almost all categories such as short chain alkanes, long chain alkanes, aromatic hydrocarbons and acrylic hydrocarbons have been reported to be utilized by these microorganisms. Among the gaseous hydrocarbon utilizers, only methanotrophs have been studied elaborately as far as the genetics and molecular biology is concerned. On the other hand metabolic pathways and their regulation among the higher hydrocarbon. (such as octane, tolune, xylene, naphthalene *etc.*) utilizers have also been well characterized. The small chain gaseous hydrocarbons such as ethane, propane and butane utilizers have not been studied in detail. The metabolic pathways, their regulatory mechanism and the different genes involved in it have not been worked out. The present study was undertaken with the objectives of having some insight into these aspects of a bacterium isolated from Gujarat (India) oil field.

The organism was termed as butanotroph because the lowest alkane it can utilize is butane. It can also utilize higher hydrocarbons (upto C_{10} were checked) as a sole source of carbon and energy but does not grow on lower hydrocarbons such as methane, ethane and propane. Various biochemical tests, GC content and antibiotic resistance profile of this organism helped in identifying it as a species of *Pseudomonas* and was designated as *Pseudomonas* sp. IMT37. Its generation time in minimal medium with butane as a sole source of carbon and energy was 10-12 hours. compared to 95 minutes in glucose. Glucose was found not to repress the induction of hydrocarbon utilization genes when glucose and butane were used together as substrates for growth.

A 4.9 kb fragment was cloned earlier from the genomic library of a propanotroph, *Pseudomonas* sp. IMT40. It encodes a 58 kDa protein, which gets induced by growth on propane or butane. This fragment was shown to be highly specific to hydrocarbon utilizers. (S. Kaul, 1992). It was used as a probe to compare the flanking region of this fragment in two organisms *Pseudomonas* sp. IMT40 (a propanotroph) and *Pseudomonas* sp. IMT37 (a butanotroph). The total DNA of the two organisms was digested with *Eco*RI, *Sal*I, *Hin*dIII, *Pst*I and *Xho*I. After Southern blotting and hybridization, the region around the 4.9 kb fragment was found to be exactly similar as far as the restriction sites was concerned.

A cosmid library of *Pseudomonas* sp. IMT37 was constructed in pHC79. Genomic DNA of the organism was partially digested with *Hin*dIII and ligated to the cosmid pHC79 which was also digested with *Hin*dIII and dephosphorylated. The ligated mixture was packaged and transfected as well as electroporated into *E.coli*. The libraries obtained by both the methods were screened using the 4.9 kb fragment as a probe. In the library obtained by electroporation a total of seven clones were obtained which hybridized with this fragment. These were designated as pRT1, pRT2, pRT3, pRT4, pRT5, pRT6 and pRT7. The insert size in these clones varied from 6 kb to 26 kb. These clones together covered a region of approximately 40 kb flanking the 4.9 kb fragment. Among these clones, pRT3 carried the corresponding region of 4.9 kb fragment which encodes the 58 kDa protein. Therefore, it was selected for further characterization. Two subclones- pRT3A carrying an insert of 2.3 kb and pRT3B carrying an insert of 3.7 kb were generated in pUC19. The subclone pRT3A carried the coding region for 58 kDa whereas pRT3B represented its upstream region. Detailed restriction endonuclease map of pRT3A was constructed and smaller subclones were generated. The 2.3 kb insert from pRT3A and its 300 bp upstream region in pRT3B were sequenced. The longest ORF found in the total sequence of 2606 bp was of 1512 bp, with the initiation codon at 502 bp. A slightly smaller ORF is also possible if ATG at 544 bp is considered as the initiation site.

This translates into a protein of mol.wt. 54 kDa. There were two putative ribosomal binding sites (RBS) at -14 (AGG CGG) and at -16 (AGC AGG) which are very similar to the AGGAGG consensus sequence of *E.coli* and GGAGA sequence in many RBS of *Pseudomonas*. Apparently there was no promoter like sequence upstream to the longer ORF. The other initiation codon at 544 bp makes a protein of 52.3 kDa. The putative RBS here is CGGGCA at -8 bp. The sequence at -14 (TGTC) and -22 base (TGGC) upstream of this ORF shows homology to the -12 (TTGCT) and -24 (TGGC) consensus sequences of *Pseudomonas* promoters. Nucleic acid and protein sequence homology search of this region showed no significant homology with any sequence in the database. The sequence, therefore, appears to be unique. This sequence was also compared with about 500 bp region from another organism - a propanotroph, *Pseudomonas* sp. IMT40. The sequence was exactly similar in both organism indicating that it is highly conserved among these hydrocarbon utilizers.

In order to understand the function of the protein encoded by the ORF in pRT3A, it was disrupted by putting 1.3 kb 'kan' cassette in it. The vector carrying the disrupted

gene was electroporated into the host organism *Pseudomonas* sp. IMT37. Since the vector pUC19 is unable to replicate in *Pseudomonas*, 'kan' resistant colonies of *Pseudomonas* would, in theory, appear only if the disrupted gene from the vector replaces the wild type gene in the host organism by homologous recombination. The vector carrying the disrupted gene may also get inserted in the wild type gene if a single cross over event occurs. Both type of 'kan' resistant transformants were obtained in this case. This was confirmed by digesting the total DNA of the transformants with *Eco*RI and doing Southern blotting and hybridization using the 4.9 kb fragment and 1.3 kb 'kan' cassette as probes. In all 43 'kan' resistant transformants were obtained out of which 5 lost their ability to grow on hydrocarbons (C_4 to C_{10}) although they were able to grow on their respective alcohols. This indicate that the ORF encodes a protein which is part of the multicomponent monooxygenase enzyme involved in the conversion of hydrocarbons to their respective alcohols. This was also confirmed by doing the whole cell monooxygenase assay with one of 'kan' resistant transformant (37.1K) unable to grow on hydrocarbons. No detectable monooxygenase was found in this transformant. With similar approach the disrupted gene was also put in Pseudomonas sp. IMT40 a propanotroph. In this case also the 'kan' resistant transformants obtained were found unable to grow on all hydrocarbons tested (C_3 to C_{10}) but could grow on their respective alcohols. This proves that this sequence is conserved among the hydrocarbon utilizers checked and plays an important role in the first step of the hydrocarbon utilization pathway.

Another approach of isolating differentially expressed genes was used in identifying the genes induced by the growth on hydrocarbons. The total RNA from glucose and butane grown cells was isolated. Using random arbitrary primers, cDNA was prepared and amplified using the random arbitrarily primed PCR (RAP) technique. The differentially expressed bands were eluted from the gel and cloned into pUC19 after reamplification. A total of seven clones designated as pRS2B to pRS2H were obtained. By cross hybridization of the four clones pRS2D, pRS2G, pRS2F and pRS2H with other clones and the 40 kb cloned region of *Pseudomonas* sp. IMT37, it was found that these are independent. Two of these clones - pRS2D and pRS2G were sequenced also. Nucleic acid homology search revealed no significant homology in the database. Northern hybridization of these clones with the total RNA isolated from LB, glucose and butane grown cells shows these to be hydrocarbon specific sequences. To get deeper insight into the role of these sequences in hydrocarbon utilization, complete genes from the library of *Pseudomonas* sp. IMT37 have to be cloned and analyzed further. The isolation of mutants

defective at different steps of hydrocarbon utilization and the attempt to do their complementation by the cloned region of this organism will prove to be a step forward in the direction of having better understanding about the genetics and molecular biology of the butane utilization pathway.