Chapter 3

Differentiation of obligate and facultative alkaliphiles
3.0.0 Introduction

Central dogma of molecular biology includes replication, transcription and translation. The amino acid sequence of a particular protein encoded in mRNA appears through translation process. This process is carried out by ribosomes. Ribosomes are composed of protein and RNA molecules. Ribosomes constitute two subunits both in prokaryotes (30S, 50S) and eukaryotes (40S, 60S). In prokaryotes, 70S ribosomes are made from complexes of RNA and proteins. Ribosomes are divided into two subunits, one larger subunit (50S) & other smaller subunit (30S). The smaller subunit (30S) binds to the mRNA, while the larger subunit (70S) binds to the tRNA and the amino acids. The 30S ribosomal subunit has a 1542 nucleotide RNA subunit (16S) bound to 21 proteins.  

3.1.1. Structure of 16S rRNA

The primary structure of 16S rRNA is highly conserved. It consists of 1542 bases. The arrangement of the 16S rRNA creates a 5’ domain, central domain, 3’ major domain, and a 3’ minor domain. The 5’ domain consists of 19 double helices that make up the bulk of the body. The central domain of the rRNA generates the platform and is an elongated, curved structure of nine helices, with the junction of helices 20, 21, and 22 being at the heart of it. The 3’ major domain contains 15 helical elements and composes the head. Sequence analysis of the 16S ribosomal RNA (rRNA) gene has been widely used to identify bacterial species and perform taxonomic studies.  

3.1.2. 16S rRNA and hyper variable region for species level identification

Sequence analysis of the 16S ribosomal RNA (rRNA) gene has been widely used to identify bacterial species and perform taxonomic studies. Bacterial 16S rRNA genes generally contain nine “Hypervariable regions” that demonstrate considerable sequence diversity among different bacterial species and can be used for species identification. Moreover, species-specific sequences within a given hyper variable region constitute useful targets for diagnostic assays. Hypervariable regions are flanked by conserved stretches in most bacteria, enabling PCR amplification of target sequences using universal primers.
Fig. 3.1 Secondary structure of 16S ribosomal RNA of *Escherichia coli* deduced from sequence comparison and results chemical studies, part A (adapted from Stryer L. 1995). The tertiary structure of 16S RNA determined by X ray crystallography, part B (adapted from Dr. Bryn Weiser and Dr. Harry Noller). 16S ribosomal RNA of 1542 nucleotide molecule contains 5’ terminal domain (red), a central domain (green), 3’ terminal major and minor domains (blue).
Although 16S rRNA Hyper Variable Regions exhibit different degrees of sequence diversity, no single hypervariable region (HVR) is able to distinguish bacteria at species level. Briefly, these hypervariable regions can be described as follows: HVR V1, which corresponds to nucleotides 69-99, distinguish common pathogenic *Streptococcus* sp. and to differentiate between *Staphylococcus aureus* and coagulase negative *Staphylococcus* species. HVR V2, which consists of nucleotides 137-242, distinguishes among the common *Staphylococcal* and *Streptococcal* pathogens and among *Clostridium Haemophilus* and *Neisseria* species and more specifically for distinguishing among *Mycobacterial* species. HVR V3, nucleotides 433-497, was appeared to be better in distinguishing between the closely related enterobacteriaceae *K. pneumoniae* and *E. aerogenes*, and the SNP variation among different *Haemophilus* species. HVR V6, nucleotides 986-1043, appeared to be the best target region for assays designed to distinguish between *B. anthracis* and *B. cereus*. The other hypervariable regions include V4 (nucleotide 576–682), V7 (nucleotide 1117–1173), V8 (nucleotide 1243–1294), and V9 (nucleotide 1435-1465) are not much important with reference to species level identification.

Extremophiles are those bacteria which survive in extreme climatic and environmental conditions. Based on pH, two extremophiles were grouped in two distinct categories: Acidophiles (survive in acidic conditions) and alkaliphiles (survive in acidic conditions). These bacteria not only just survive in acidic or alkaline environment but also optimally work in these conditions. Some psychrophilic bacteria can sustain life in extreme cold environment in Antarctica and also they can optimally grow in refrigerators in kitchen room at 4°C. Extremophiles thrive in uncommonly harsh environments. Some extreme thermophiles live in temperatures ranging from 45°C to 122°C and halophiles dwell in environments with high salt content and most importantly antibiotic resistant life forms which flourish despite antibiotics. Antibiotics resistance is an important area of study because many diseases are developing resistance to antibiotics. The driving force of newer antibiotic search and research is the resultant phenomenon of antibiotic resistance of microbes. Casey R. Richardson hypothesized that the prokaryotic extremophiles can be characterized by their 16S rRNA sequential characteristics. He has
examined to justify his hypothesis based on four base content and pattern of their arrangement in 16S rRNA. He examined nucleotide and dinucleotide frequencies, adenosine and uracil (A+U) and guanine and cytosine (G+C) content, adjusted base pairing propensity (Pb), adjusted base pair distance (dD), adjusted Shannon entropy (dQ), and minimum free energy (MFE)\(^{102}\). Secondary structure and G+C content have been shown to characterize the living environment of bacteria.\(^{103}\) 16S rRNA has also been shown to play a part in a bacteria’s fitness.\(^{104}\)

Since Brock et al. isolated and named the extreme thermophile *Thermus aquaticus* from Yellowstone National Park\(^ {105}\) around the world various bacterial strains have been isolated from many hydrothermal areas with water temperatures higher than about 55 °C and pH ranging from neutral to alkaline. On the basis of the phylogenetic data of 16S rRNA, the growth temperature, polar lipid pattern and the hydroxy fatty acid composition, Nobre et al. proposed the new genus *Meiothermus* reclassified from the genus *Thermus*\(^ {106}\), Rainey and da Costa proposed ‘Thermales ord. nov.’ and ‘Thermaceae fam. nov.’\(^ {107}\)

### 3.1.3. Obligate and facultative alkaliphiles

The alkaliphiles are those bacteria, which grow optimally or very well around pH 9 and above pH 9; but cannot near neutral pH value of 7.0. There are two types of alkaliphiles as: obligate alkaliphiles, which grow above pH 9.0 but do not near neutral pH, and facultative alkaliphiles, which grow from pH 7 to pH 10.

As mentioned above some of the closely related species e.g. *B. anthracis* and *B. cereus* can be differentiated based on specific hypervariable region V6, this regions was first time reported for analysis of obligate alkaliphiles by Ntougias and Russel in 2001. They reported that obligate bacteria can be grouped together based on secondary structure of 16S rRNA hypervariable region V6.\(^ {108}\) Based on their 16S rRNA sequences, the three obligately alkaliphilic bacteria, isolated from wash waters of edible olive production, present almost identical genotypic characteristics, since their phylogenetic topology and 16S rRNA secondary structure are almost the same. Comparative studies of 16S rRNA sequence revealed that the strains represent a new alkaliphilic linkage in the order *Bacillales*, belonging to the *Carnobacterium/Aerococcus*-like spectrum. Ntougias and
Russel proposed that the strains should be assigned to a new genus and species, *Alkalibacterium olivoapovliticus*. They have the same G-C content in their DNA and the level of hybridisation of the DNA between all pairs of the three bacteria is >88%, a value well above that of 70% which is generally regarded as being the cut-off point for distinguishing different species. Therefore, the genotypic characteristics indicate that the three obligate alkaliphiles are different strains of the same species. However differentiation of obligate and facultative alkaliphiles were not reported or incompletely understood.

As most of the earlier 16S rRNA studies of Lonar Lake bacteria were reported with the aim of molecular identification, large amount these data were not explored for analysis of stability of 16S rRNA in high salt stress condition or in extreme alkaline conditions. Mechanisms of working model of ribosome or polyribosomes at high salt and alkaline conditions or haloalkaline conditions outside bacteria cell environment need deep interest of research to explore this line of research. Specific domain in 16S rRNA or its specific hypervariable region or regions may play direct or indirect roles in maintaining stability of ribosomes during protein synthesis. Initial results of complete 16S rRNA sequence of obligate and facultative alkaliphiles prompted us to carry our further details analysis of available hypervariable regions. Attempts were made to initiate few steps in this direction as brief part of this thesis, by correlating growth pattern studies at alkaline conditions with phylogenetic tree of known obligate and facultative alkaliphile. In the present study attempts were made to distinguish or categorize adaptive nature of bacteria based on pH as one of the parameter. In alkaliphiles if the outer pH of bacterial cell is 10.5 then inside of cytoplasm pH is just 2 to 2.5 units less than outer pH. With this 8.0 to 8.5 pH inside the bacterial cell, there might be pressure on ribosomes for protein synthesis at alkaline pH. Some of the crucial region for RNA folding needs to study at different pH to understand this. Some of the hypervariable regions may play important role in maintaining 3D structure of ribosome. Comparative studies of acidophiles, neutrophiles and alkaliphiles may suggest some clue for differentiation of acidophiles and alkaliphiles. Further detail analysis of hypervariable regions may differentiate obligate and facultative alkaliphiles. The purpose of this chapter is not
species level identification but differentiation of obligate and facultative alkaliphile base on hypervariable region.

3.2.0 Materials and method

3.2.1 Growth pattern of alkaliphiles on solid nutrient medium

After initial identification in second chapter following strains were selected: *Aquiflexum* species DL6 strain, *Kocuria* species DL strain, *Bacillus badius* D1 and *Lysinibacillus* species DL15 strain. All these strains were grown at different pH conditions from pH 7.0 to pH 12.0 (Detailed protocol is mentioned in chapter 2).

**Bioinformatics tools include the softwares, and databases:** NCBI database, ClustalW, BLAST2, BLAST, MEGA 4.1, and NCBI database (Detailed protocol is mentioned in chapter 2).

3.2.2.0 16S rRNA sequence analysis of Lonar Lake alkaliphiles available in NCBI database

Initial phylogenetic analysis was carried out on complete or available length of sequence of 16S rRNA. Briefly, the sequences were downloaded from NCBI GenBank database using accession number, DQ026060, AB188090, CP001878, X76446, X76440 for facultative alkaliphiles, while EF103128, AB270706 for obligate alkalophilic bacteria. Lonar lake strains from present study with following accession number JN595811, JF812063, HQ015711, HM439779 were considered for obligate and facultative alkaliphile studies. Phylogenetic analysis was carried out at mentioned in the chapter 2. Briefly, sequences were aligned by ClustalW and output was saved in FASTA format. Then, phylogenetic tree analyses of 16S rRNA as well as hypervariable regions were performed by the neighbor-joining method with bootstrap values of 1000 replicates generated using the program MEGA 4.1.

3.2.2.1 Hypervariable region derivation from *Escherichia coli*

16S rRNA hypervariable region studies are proposed and established for identification of any bacteria on the basis of each hypervariable region within 16S rRNA. All the 16S ribosomal RNA sequences were compared with *E. coli* as reference.
strain (Ac. N0. J01859.1) for finding their hypervariable regions. Hypervariable regions were identified based on two sequence comparative study i.e. NCBI BLAST2. Hypervariable regions V2, V3 and V6 are most important for molecular detection of most of the bacteria at species level. Phylogenetic analysis of entire 16S region was carried out by MEGA 4.1.

**16S rRNA of *E. coli* reference strain (Ac. N0. J01859)***

AAATTGAAGAGTTTATGCTGGCCATGAGTTAGACCTGGGCAAGGCTCACAAAAACATTGCAAGTGACGGT
AACAGGAAGAGCTTGTCTTCTTGACAGATGCGGTGCGCGAGCTGAGATGATTGCTGAGTGCAGAGAGGCT
GAGGGGATACATTGCAAGAAACGGTATTATATGCAAGCTCGCAAGGAGAGGCAAGAAACGGTATTATAT
GCCTTACCGATGGCAAGATTACGCTTACTTACCGACACGCGCGCGGAGCCTAGCTTCTTACGGCAG
ATCCCTAAGCTGTCGAGAGGAGTAGACCCACACCTGGTAAGCCGTCGACGAGC
GAGGGGATAACGTGCAAGAAACGGTATTATATGCAAGCTCGCAAGGAGAGGCAAGAAACGGTATTATAT
GCCTTACCGATGGCAAGATTACGCTTACTTACCGACACGCGCGCGGAGCCTAGCTTCTTACGGCAG
16S rRNA of *E. coli* reference strain (Ac. N0. J01859.1)

V1:69 – 99
GTAACAGGAAGAAGCTTGTCTTATGCTGGCCATGAGTTAGACCTGGGCAAGGCTCACAAAAACATTGCAAGTGACGGT

V2:137–242
GATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGACCT

V3:433 – 497
TACTTTCAGCGGGGAGGAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGA

V4:576 – 682
GCACGCCGCGGCTGTTATGTTAAGTCTGAGATGTGAATTTCCCGCTACGACCGCAAGGGACTCGGGATGG

V5:822 – 879
CGACCTTGGAGGTGTGCTCTTGAGGCGGATGG

**Hypervariable regions (V1 to V9) in 16sRNA** (Chakravorty J. et. al. 2007)

V1: 69 – 99
GTAACAGGAAGAAGCTTGTCTTATGCTGGCCATGAGTTAGACCTGGGCAAGGCTCACAAAAACATTGCAAGTGACGGT

V2: 137–242
GATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGACCT

V3: 433 – 497
TACTTTCAGCGGGGAGGAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGA

V4: 576 – 682
GCACGCCGCGGCTGTTATGTTAAGTCTGAGATGTGAATTTCCCGCTACGACCGCAAGGGACTCGGGATGG

V5: 822 – 879
CGACCTTGGAGGTGTGCTCTTGAGGCGGATGG

109
3.2.2.2 Phylogenetic analysis

Phylogenetic analysis was carried out by MEGA 4.1. The well established and reported obligate and facultative alkaliphilic bacterial sequences were obtained from NCBI nucleotide database and bootstrap consensus trees inferred from 1000 replicates was selected to represent the evolutionary history of 16S rRNA sequence analysis of alkaliphiles and for each hypervariable region analysis.
3.3.0 Results and Discussion

Identification of facultative and obligate alkaliphiles
Facultative and the obligate alkaliphiles were differentiated on the basis of their pattern of growth in the range of pH 7.0 to pH 12.0 i.e. neutral to alkaline medium. Bacteria which grow in the range of pH 7.0 to pH 12.0 or pH 7.0 to pH 10.0 were considered as facultative alkaliphiles. Bacteria which could not grow below pH 9.0 were considered as obligate alkaliphiles. Bacteria isolated from alkaline source, acidic source and neutral source were also identified as mentioned in table 3.1.

Table 3.1: Differentiation of obligate and facultative alkaliphiles based on growth

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Growth condition</th>
<th>Bacteria source</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Kocuria species</em> DL strain</td>
<td>pH &gt;9.0 and not &lt;9.0</td>
<td>Lonar Lake</td>
<td>Obligate alkaliphile</td>
</tr>
<tr>
<td><em>Aquiflexum species</em> DL6 strain</td>
<td>pH &gt;9.0 and not &lt;9.0</td>
<td>Lonar Lake</td>
<td>Obligate alkaliphile</td>
</tr>
<tr>
<td><em>Bacillus badius</em> DL1</td>
<td>pH 7.0 to 10</td>
<td>Lonar Lake</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Lysinibacillus</em> DL15 strain</td>
<td>pH 7.0 to 12.5</td>
<td>Lonar Lake</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Stenotrophomonas</em> DL16 strain</td>
<td>pH 7.0 to 12.0</td>
<td>Lonar Lake</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Pseudomonas species</em> DL17 strain</td>
<td>pH 7.0 to 12.0</td>
<td>Lonar Lake</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Stenotrophomonas</em> DL18 strain</td>
<td>pH 7.0 to 12.0</td>
<td>Lonar Lake</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Acinetobacter junii</em> strain DL2</td>
<td>pH 7.0</td>
<td>Mula River</td>
<td>Neutrophile</td>
</tr>
<tr>
<td><em>Oceanimonas denitrificans</em> DL3</td>
<td>pH 7.0</td>
<td>Mula River</td>
<td>Neutrophile</td>
</tr>
<tr>
<td><em>Exiguobacterium species</em> DL4 strain</td>
<td>pH 7.0</td>
<td>Mula River</td>
<td>Neutrophile</td>
</tr>
<tr>
<td><em>Bacillus subtilis</em> DL5</td>
<td>pH 7.0</td>
<td>Mula River</td>
<td>Neutrophile</td>
</tr>
<tr>
<td><em>Microbacterium arborescens</em> D6</td>
<td>pH 4.0 to 7.0</td>
<td>Citrus limon canker</td>
<td>Facultative Acidophile</td>
</tr>
<tr>
<td><em>Lysinibacillus sphaericus</em> DL8</td>
<td>pH 7.0 to 9.0</td>
<td>Hen’s stool sample on egg shell surface</td>
<td>Facultative alkaliphile</td>
</tr>
<tr>
<td><em>Bacillus pumilus</em> DL12</td>
<td>pH 7.0 to 9.0</td>
<td>Hen’s stool sample on egg shell surface</td>
<td>Facultative alkaliphile</td>
</tr>
</tbody>
</table>
*Lysinibacillus species* DL15 strain on different pH conditions

Fig. 3.2 Facultative alkaliphile, *Lysinibacillus species* DL15 strain: Growth in the range of pH 7.0 to 12.0
Stenotrophomonas species DL18 strain on different pH conditions

Fig. 3.3: Facultative alkaliophile Stenotrophomonas species DL18 strain: Growth in the range of pH 7.0 to 12.0
Phylogenetic studies based on complete sequenced of 16S rRNA

Fig. 3.4: Phylogenetic tree based on entire 16S rRNA Facultative alkaliphile (clade I) and obligate alkaliphile (clade II)

Based on accepted properties of obligate and facultative alkaliphiles (highly cited and widely accepted obligate and facultative alkaliphiles) were considered for alkaliphile differentiation studies. The facultative alkaliphiles can be grouped in clade I with 100% bootstrap trees but the obligate alkaliphiles were not grouped in distinctive clade as compared to facultative one. Since our finding of microbiological studies show Kocuria DL and Aquiflexum DL6 were obligate alkaliphiles, these result prompted further analysis based on each hepervariable.
Derivation of hypervariable region of \textit{Aquiflexum} species DL6 strain from \textit{Escherichia coli} EcoRRD strain

\textbf{Query:} \textit{Aquiflexum} DL6; \textbf{Subject:} \textit{E. coli} J01859.1 (ECORRD strain)

ID: lcl28381Length: 1541Number of Matches: 1

Related Information;
Range 1: 240 to 1423 Graphics Next Match Previous Match First Match

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
& Score & Expect & Identities & Gaps & Strand & Frame \\
\hline
& 784 bits(424) & 0.0 & 949/1201(79\%) & 41/1201(3\%) & Plus/Plus & \\
\hline
\end{tabular}
\end{table}

\textbf{Features:}

\begin{description}
\item[Query 186] GGATTAGCTAGTTGGCGGGGTAACGGCCCACCAAGGCGACGATCCTTAGGGTTCTGAGA 245
\item[Sbjct 240] GGATTAGCTAGTTGGCGGGGTAACGGCCCACCAAGGCGACGATCCTTAGGGTTCTGAGA 299
\item[Query 246] GGAAGGTCCCCCACCAGCCTGAGTACGGGCGAGACTCTACGGGAGGCGACGATAG 305
\item[Sbjct 300] GGAAGGTCCCCCACCAGCCTGAGTACGGGCGAGACTCTACGGGAGGCGACGATAG 359
\item[Query 366] TACCGGTTGTAAGACTCTTATAC--G--GGAA--GA--AAAGAACAAGTGGGGTTG--AAAT 418
\item[Sbjct 420] TACCGGTTGTAAGACTCTTATAC--G--GGAA--GA--AAAGAACAAGTGGGGTTG--AAAT 478
\item[Query 419] TGCCGGTA--CCGTAGTGAATAAGACACCGGCTAACTCGGTCGAGACCCGCGGTTAATACGG 477
\item[Sbjct 479] TGCCGGTA--CCGTAGTGAATAAGACACCGGCTAACTCGGTCGAGACCCGCGGTTAATACGG 537
\item[Query 478] AGGGTGCGACGCTTCGACCAGTGCTTATGGTGTKTTAAAAGGCGGTCGCCAGGGCGCTTATAGT 537
\item[Sbjct 538] AGGGTGCGACGCTTCGACCAGTGCTTATGGTGTKTTAAAAGGCGGTCGCCAGGGCGCTTATAGT 597
\item[Query 538] CAGCGGTGAAATACCTCCGGCTCAACCGAGGGGGTGGCTGTTATGATACTGGTTTGGATGTG 597
\item[Sbjct 598] CAGCGGTGAAATACCTCCGGCTCAACCGAGGGGGTGGCTGTTATGATACTGGTTTGGATGTG 657
\item[Query 598] CCGTCCTACGTACATGGAATTCTCGGCTAAGCTGATTGTTYAGTGGACGTGATGCAAACTACCACCTAAGAACGAACGCTTGAGT 657
\item[Sbjct 658] CCGTCCTACGTACATGGAATTCTCGGCTAAGCTGATTGTTYAGTGGACGTGATGCAAACTACCACCTAAGAACGAACGCTTGAGT 717
\item[Query 658] CCGTAGGCGAAGCTTGGCAGGTAGTCCGAGGGGGGTTAGCGGAGAGCAGCGAGATGAAAGGCTTGAGT 716
\item[Sbjct 718] CCGTAGGCGAAGCTTGGCAGGTAGTCCGAGGGGGGTTAGCGGAGAGCAGCGAGATGAAAGGCTTGAGT 776
\item[Query 717] GCCAACGGATGATTAGATACCTCGTGTAGTCCGAGGGGGTTAGCGGAGAGCAGCGAGATGAAAGGCTTGAGT 774
\item[Sbjct 777] GCCAACGGATGATTAGATACCTCGTGTAGTCCGAGGGGGTTAGCGGAGAGCAGCGAGATGAAAGGCTTGAGT 835
\item[Query 775] T--CCTTATTGGGTAGTGAGGGGCA--AGCGAAAGCGTTAAGGT--GATCCACCTGGGGAGTACG 831
\item[Sbjct 836] T--CCTTATTGGGTAGTGAGGGGCA--AGCGAAAGCGTTAAGGT--GATCCACCTGGGGAGTACG 893
\end{description}
Hypervariable region V4 of *Aquiflexum* species DL6 strain derived from *E.coli*

```
GTGCGCAAGGCGGCTGTAAGTCAAGGGGTTGAAATACCTCCGCTCAACCGGAGGGGTGCC
GTTGAATGTGGCGGCTGTAAGTCAAGGGGTTGAAATACCTCCGCTCAACCGGAGGGGTGCC
```

All hypervariable regions were derived from all the selected and isolated alkaliphilic strains.
Analysis of alkaliphiles from Lonar Lake available in NCBI nucleotide database based on each hypervariable region

Fig. 3.5: Phylogeny for Region V3 does not include studied obligate alkaliphiles in clade of obligate alkaliphile. As the position of Aquiflexum DL6 is in between facultative and obligate alkaliphiles, this region in not significant.

Fig. 3.6: Hypervariable region V4 can differentiate obligate and facultative alkaliphile isolated in this study. These results are consistent with near to complete 16S rRNA sequence analysis for differentiation of facultative (clade I) and obligate alkaliphiles (clade II)
**Fig. 3.7:** Region V5: *Kocuria DL*, along with all other obligate alkaliphiles were grouped in clade I, which belongs to clade of facultative alkaliphile. So this is not significant region.

**Fig. 3.8:** Region V6: *Kocuria DL* an obligate alkaliphile was grouped in Clade I which is of facultative alkaliphile. This region is not significant.
Isolation and identification of these bacteria were based on culture studies and molecular studies of 16S rRNA. Bacteria were isolated on the basis of pH and identified on the basis of 16S rRNA include *Kocuria sp DL* (HM439779.1), *Aquiflexum sp. DL6* (JF812063.1), *Bacillus badius sp D1* (HQ015711.1), *Lysinibacillus sp DL15* (JN595811.1). On the basis of their growth condition, they were classified as obligate alkaliphiles and facultative alkaliphiles. Obligate alkaliphiles, which grow at pH around 10.0 and not below pH 9.0, are *Kocuria sp DL* (HM439779.1) and *Aquiflexum sp. DL6* (JF812063.1) while facultative, which grow between neutral to alkaline range are *Bacillus badius sp D1* (HQ015711.1) (pH 7.0 to pH 10) and *Lysinibacillus sp DL15* (JN595811.1) (pH 7.0 to pH 12.5). The phylogeny based on the 16S rRNA suggests two different categories of alkaliphiles i.e. obligate and facultative alkaliphiles can be distinguished based on 16S rRNA sequence.

**Fig. 3.9:** Distinct differentiation can not be made from Region V7. Both the obligate alkaliphiles were grouped under obligate alkaliphile. But clade of obligate alkaliphile is not distinct as V4 shows.
All samples were tested on neutral to alkaline medium on nutrient agar (pH 7.0 to 12.0). At pH 7.0 only those samples which were very close to completely sequence bacteria and potentially non hazardous were further selected for ATP synthase studies. Strain D1, facultative alkaliphile and strain DL6 were used for MFC experiment and strain DL15 and strain DL18, both were facultative alkaliphiles were used in ATP synthase studies. All other remaining bacterial strains isolated from Lonar as well as from Non Lonar Lake sources were used for differentiation of obligate and facultative alkaliphile by 16S rRNA hypervariable region studies supported by growth conditions.

**Table 3.2: GC content of hypervariable region V4**

<table>
<thead>
<tr>
<th>Alkaliphiles</th>
<th>Type</th>
<th>V4 region GC content</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Kocuria DL</em></td>
<td>Obligate</td>
<td>62.6%</td>
</tr>
<tr>
<td><em>Aquiflexum DL6</em></td>
<td>Obligate</td>
<td>55.1%</td>
</tr>
<tr>
<td>A. collagenimarina</td>
<td>Obligate</td>
<td>53.3%</td>
</tr>
<tr>
<td><em>Natronocella acetinitrilica</em></td>
<td>Obligate</td>
<td>52.9%</td>
</tr>
<tr>
<td>Lysinibacillus sp DL15</td>
<td>Facultative</td>
<td>52.3%</td>
</tr>
<tr>
<td><em>Bacillus badius D1</em></td>
<td>Facultative</td>
<td>57.9%</td>
</tr>
<tr>
<td><em>Bacillus okhensis</em></td>
<td>Facultative</td>
<td>56.1%</td>
</tr>
<tr>
<td><em>Bacillus oshimensis</em></td>
<td>Facultative</td>
<td>57.0%</td>
</tr>
<tr>
<td><em>Bacillus pseudofirmus</em></td>
<td>Facultative</td>
<td>55.1%</td>
</tr>
<tr>
<td><em>Bacillus gibsonii</em></td>
<td>Facultative</td>
<td>55.1%</td>
</tr>
<tr>
<td><em>Bacillus clausii</em></td>
<td>Facultative</td>
<td>57.9%</td>
</tr>
</tbody>
</table>

**Discussion**

Based on availability of full and partial 16S rRNA sequence of selected reference strains from NCBI, analysis of hypervariable regions V3, V4, V5, V6 and V7 was carried out. The phylogenetic analysis shows that the V4 region can distinguish between facultative and obligate alkaliphiles. The facultative alkaliphiles were shown in one clad with significant value of 87% bootstrap trees in V4 region based phylogeny, while 60% bootstrap trees clustered facultative alkaliphiles in clad in V7 based phylogeny. Therefore, V4 could best distinguish the facultative alkaliphiles from obligate alkaliphiles. GC content of V4 hypervariable region is shown in table 3.2. One of the obligate alkaliphile Kocuria species DL strain shows high GC content among all alkaliphiles. From whole genome of one of reference strain it is know that the overall
genome is GC rich and for *Kocuria rhizophila* it is > 70%. Alkaliphiles are either not available or incompletely understood, however one of the prominent studies on thermopiles may explain requirement of high GC bonding for stability of 16S rRNA secondary structure. Very few studies were explored based on gene sequence of 16S rRNA and cataloging of particular group of organism. In one of the study on thermopiles conducted by Chen C et al 2003 family thermaceae was proposed for a 38 strains of bacteria with high GC bonding in secondary structure of 16S rRNA. They categorized 38 different species in 18 Genuses in four categories i.e. Thermus, Meiothermus, Marinothermus and Oceanothermus. One of most famous and widely applied bacteria i.e. *Thermus aquaticus* for it thermostable enzyme *Taq* polymerase, widely used in all biotechnology laboratories around the world for PCR, is also grouped in this category. Although this bacterium was grouped under genus *Thermus* based only on sequence of 16S rRNA, property of its thermostable enzyme, *Taq* polymerase, completely proved why it was included in genus *Thermus*.

In brief, attempts were made to differentiate alkaliphiles on the basis of hypervariable regions equally supported by microbial growth studies at different pH conditions. Out of nine hypervariable regions, V4 was found to be the most significant in the differentiation of obligate and facultative alkaliphiles. More specifically, the two distinct clades were observed for differentiation of alkaliphiles based on V4 region of 106 nucleotides, better than differentiation based on that of 16S rRNA of 1.5 kb. Data analysis of individual hypervariable region of specific group of alkaliphilic bacteria suggests interesting results with hypervariable region V4. It is known that central domain of 16S rRNA comprises this hypervariable region, which starts after most important region i.e. 20\(^{th}\), 21\(^{st}\) and 22\(^{nd}\) helices of the 16S rRNA, i.e. heart of the molecule as mentioned in figure 1.1 and figure 3.1. This hypervariable region V4 (nucleotide numbers 576 to 682, *E. coli* numbering system), along with other factors, (although it is very premature to say at this stage) may prove significant in exploring role in rRNA stability during protein synthesis at comparatively high alkaline pH.

It is reported that the phylogenetic analysis using 16S rDNA is based on the comparison of similarity of the 16S rDNA molecules. When comparing 16S rRNA sequence data, the
main problem one encounters is how to evaluate sequence divergence and how to derive conclusions about the relatedness of organisms. As the selection criteria used in the present study was very strict about growth conditions and overall bacterial response at alkaline pH carefully understood from original research papers then sequences were downloaded for studies, obligate and facultative alkaliphiles was differentiated into obligate and facultative alkaliphiles on briefly the basis on complete or available 16S rRNA sequence and completely differentiated on the basis of hypervariable region V4.