CHAPTER 6

METAGENOMIC STUDIES

METAGENOME ISOLATION AND CAPTURING FUNCTIONAL ATTRIBUTES
SECTION-I

METAGENOMICS STUDIES FOR BACTERIAL DIVERSITY AND ITS AMENABILITY FOR FURTHER FUNCTIONAL ATTRIBUTES (ALKALINE PROTEASES)
6.1.1 INTRODUCTION

Metagenomics is an emerging approach based on the extensive analysis of the DNA of microbial communities in their natural environment. Metagenomics has been developed over the last several years to assess the genomes of the non-cultivable microbes towards better understanding of global microbial ecology and to trap vast biotechnological potential of a given habitat. The basic strategies encompass sequence and functional based approaches. Since it is widely accepted that the majority of the microbes are not cultivable, the not-yet-cultivated microbes represent a shear unlimited and intriguing resource for the development of novel genes, enzymes and other compounds for applications in biotechnology.

Studies on metagenomes have revealed vast scope of biodiversity in a wide range of environment, and new functional capacities of individual cells and communities, including complex evolutionary relationships between them (Kennedy and Marchesi, 2007). Microorganisms offer huge potential for new biocatalysts for industrial and commercial applications. Of late metagenomic based strategies have recently been employed as powerful tools to isolate and identify enzymes with novel biocatalytic activities from the unculturable component of microbial communities from various terrestrial and aquatic environmental niches. Besides, attention on the diversity and phylogeny of unculturable organisms is also being focused, as marine environment has enormous microbial biodiversity, yet to be explored (Kennedy and Marchesi, 2007; Kennedy et al., 2008; Purohit and Singh, 2009; Siddhpura et al., 2010).

Several metagenomic mega projects such as Sargasso Sea, Acid-mine drainage, Human-Microbial Gut are completed worldwide successfully. However, similar efforts have not been paid in context with saline habitats. The initial results hold significance in the light of the fact that although saline environments display enormous microbial biodiversity, it remains largely unexplored. The application of metagenomic strategies embraces great potential to study and exploit the enormous microbial biodiversity present within the saline habitats.

One of the hurdles in the way of metagenomics is the extraction of total environmental DNA (metagenome) from a given habitat. We have explored various
protocols, in terms of DNA purity, yield and humic acid content, for the isolation of metagenome from various saline soils of Gujarat, to substantiate its applications for further molecular biological work. Diversity based assessment has been elucidated on the basis of 16S rRNA amplicons – DGGE analysis (Molecular Fingerprinting Technique). Beside, the source would also provide a huge and comprehensive platform for capturing novel gene sequences. As an extension of our on-going work on haloalkaliphilic bacteria from the saline habitats of Coastal Gujarat, we have taken alkaline proteases as model system for the assessment of genetic diversity among these habitats by designing degenerate primers with the aid of bioinformatics tools. Successful cloning and expression of alkaline proteases revealed unidentified gene/s with interesting features.
6.1.2 MATERIALS AND METHODS

6.1.2.1 Environmental Soil Sampling and Storage
Two soil samples, designated as O.M.6.2 and O.M.6.5 were collected in September 2007 from Coastal region of Okha Madhi (Latitude 22.20 N, Longitude 70.05 E) Gujarat, India. They represent a typical saline soil with heavy deposition of salt. At the site of collection, a block of soil was removed and transported to laboratory in sterile plastic bags for storage at 4°C. Total DNA extraction and further analyses were carried out from these samples within 15 days.

6.1.2.2 Direct DNA extraction methods

A. Soft Lysis method

DNA extraction using Lysis Buffer
Soil sample (1g) was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150 rpm. The samples were re-extracted in 1ml of extraction buffer (100mM Tris HCl (pH-8.2); 100 mM EDTA (pH-8); 1.5 M NaCl). Supernatant were collected by low speed centrifugation (5000rpm) for 10min. A 4ml of Lysis buffer (20% (w/v) SDS; Lysozyme 1mg/ml; ProtinaseK 1mg/ml; N-lauroyl sarcosine 10mg/ml;1%(w/v) CTAB(Cetyltrimethyl ammonium bromide) was added and incubated at 65°C for 2 hours with vigorous shaking at every 15 min. Samples were centrifuged at10000 rpm for 10 min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20 min at 4°C. Again upper aqueous phase was extracted with equal volume of C: I (24:1) at 10000 rpm for 10 min at 4°C. DNA was precipitated by adding 1/10th volume of 7.5M potassium acetate and DNA was subsequently precipitated by adding 2 times of chilled ethanol. DNA precipitate was collected by centrifugation at 10000 rpm for 10min, air dried and suspended in 20-50 µl TE buffer.

B. Harsh methods

DNA extraction using bead beating method
Soil sample (1g) was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150 rpm. Re-extract the sample in 1ml of extraction
buffer. Supernatant were collected by low speed centrifugation (5000 rpm) for 10mins. Glass beads (5g) were added and the sample blended for 15min and incubated at 65°C for 2 hours. Samples were then centrifuged at 10000 rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20min at 4°C. Upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10000 rpm for 10 min at 4°C. DNA was precipitated by adding 1/10th volume of 7.5M potassium acetate and DNA was subsequently precipitated by adding 2 times of chilled ethanol. DNA precipitate was collected by centrifugation at 10000 rpm for 10min, air dried and suspended in 20-50 µl TE buffer.

DNA extraction using sonication treatment method
Soil sample (1g) was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12h with shaking at 150 rpm. The sample was re-extracted in 1ml of extraction buffer and the supernatant was collected by low speed centrifugation (5000rpm) for 10min. The supernatant was sonicated using a high intensity ultrasonic processor (Sartorius, India) with a standard 13mm horn solid probe for 3 pulses of 30 seconds each in a chilled ice bath. The sample was cooled in ice and repeatedly sonicated (6 cycles of 30 seconds) followed by incubation at 65°C for 10mins. Samples were then centrifuged at 10000rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20 min at 4°C. The upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10000 rpm for 10 min at 4°C. DNA was treated by adding 1/10 volume of 7.5M potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol. DNA precipitate was collected by centrifugation at 10000 rpm for 10min, air dried and suspended in TE buffer.

DNA extraction by a combination of Bead beating and Sonication treatment
Soil sample (1g) was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150rpm. The sample was re-extracted in 1ml of extraction buffer. Supernatant were collected by low speed centrifugation (5000rpm;10min) and sample was blended with glass beads (1g) for 15min followed by sonication using a high intensity ultrasonic processor (Sartorious) with a standard 13mm horn solid probe for 3 pulses of 30 seconds each in a chilled ice bath. The sample was cooled in ice and sonicated repeated (6 cycles of 30 seconds) and incubated at 65°C for 10 min.
Samples were centrifuged at 10000 rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20 min at 4°C. The upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10000 rpm for 10 min at 4°C. DNA was treated with 1/10 volume of 7.5M Potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol. DNA precipitate was collected by centrifugation (10000 rpm; 10min) and air dried before suspending in 20-50 µl TE buffer.

C. DNA extraction by Combination of Soft and Harsh Method

DNA extraction using bead beating combined with Lysis buffer treatment

Soil sample (1g) was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150 rpm. Re-extract the sample in 1ml of extraction buffer. Supernatant were collected by low speed centrifugation (5000rpm) for 10min. Glass beads (5g) were added and the sample blended for 5min and 15min. A 4ml of lysis buffer was added and incubated at 65°C for 2 h and shake vigorously at every 15 min. Samples were centrifuged at 10000rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20min at 4°C. The upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10000 rpm for 10 min at 4°C. DNA was treated with 1/10 volume of 7.5M potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol. DNA precipitate was collected by centrifugation (10,000 rpm; 10min) and air dried before suspending in 20-50 µl TE buffer.

DNA extraction using sonication treatment combined with lysis buffer

Duplicate 1g of soil sample was suspended in 10 ml of Extraction buffer and incubated at 37°C for 10-12 h with shaking at 150 rpm. Re-extract the sample in 1ml of extraction buffer. Supernatant were collected by low speed centrifugation (5000rpm) for 10min. The supernatants were sonicated using a high intensity ultrasonic processor (sartorious) with a standard 13mm horn solid probe for 3 pulses & 6 pulses of 30 seconds each in a chilled ice bath. A 4ml of lysis buffer was added and incubated at 65°C for 2 hours and shake vigorously at every 15 min. Samples were centrifuged at 10,000rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20min at 4°C. The
upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10,000 rpm for 10 min at 4°C. DNA was treated with 1/10 volume of 7.5M Potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol. DNA precipitate was collected by centrifugation (10,000 rpm; 10min) and air dried before suspending in 20-50 µl TE buffer.

**DNA extraction using bead beating combined with lysis buffer containing 30% PEG.**

1g of soil sample was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150rpm. Re-extract the sample in 1ml of extraction buffer. Particles (supernatant) were collected by low speed centrifugation (5000rpm) for 10min. Glass beads (5g) were added and the sample blended for 15min. A 4ml of lysis buffer containing 30%PEG was added and incubated at 65°C for 2 hours and shake vigorously at every 15 min. Samples were centrifuged at 10000rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10000 rpm for 20min at 4°C. The upper aqueous phase was again extracted with equal volume of C: I (24:1) at 10,000 rpm for 10 min at 4°C. DNA was treated with 1/10 volume of 7.5M potassium acetate and subsequently precipitated by adding 2 volumes of chilled ethanol. DNA precipitates were collected by centrifugation (10,000 rpm; 10min) and air dried before suspending in 20-50 µl TE buffer.

**DNA extraction using small and big beads together with lysis buffer**

Duplicate 1g of soil sample was suspended in 10 ml of extraction buffer and incubated at 37°C for 10-12 h with shaking at 150rpm. Re-extract the sample in 1ml of extraction buffer. Particles (supernatant) were collected by low speed centrifugation (5000 rpm) for 10min. Small glass bead (0.5g) & big glass beads (0.5g) were added and the sample blended for 5min &15min. A 4ml of lysis buffer was added and incubated at 65°C for 2 hours and shake vigorously at every 15 min. Samples were centrifuged at 10000 rpm for 10min at 4°C. The upper aqueous phase was extracted with equal volume of P: C: I (25:24:1) at 10,000 rpm for 20 min at 4°C. Again upper aqueous phase was extracted with equal volume of C: I (24:1) at 10,000 rpm for 10 min at 4°C. DNA was treated with 1/10 volume of 7.5M potassium acetate and subsequently precipitated by adding 2 volumes
of chilled ethanol. DNA precipitate was collected by centrifugation (10,000 rpm; 10min) and air dried before suspending in 20-50 µl TE buffer.

6.1.2.3 Determination of purity and yield of DNA

Co-extracted humic acids are the major contaminant when DNA is extracted from soil. These compounds absorb at 230nm, DNA at 260nm and protein at 280nm. To evaluate the purity of the extracted environmental DNA (eDNA), absorbance ratios at 260nm/230nm (DNA/humic acid) and 260nm/280nm (DNA/protein) were determined.

6.1.2.4 Gel Electrophoresis

DNA extracts (10µl) from each method were mixed with 5µl loading buffer and analyzed on 0.8% agarose gels using TAE as electrophoresis buffer. Gels were stained with ethidium bromide and gel photographs were scanned and analyzed by syngene Gene Genius Bio-imaging system. A DNA marker (DNA Ruler-Middle range, Merk life sci, India) was included in each run.

6.1.2.5 PCR amplification of 16S rRNA gene

The DNA preparations described above were used as a template to amplify a DNA fragment encoding 16S rRNA gene. The reaction mixture preparation and amplification protocol was as described in detail in Materials and Method section of Chapter-3.

6.1.2.6 Denaturing Gradient Gel Electrophoresis

Denaturing Gradient Gel Electrophoresis was performed according to Muyzer and Smalla (1998). 50µl of 16S rRNA amplicon were subjected to increasingly higher concentrations of urea and formamide which act as a chemical denaturant (20- 50%). The amplicon migrated through polyacrylamide gel containing denaturants at constant voltage; 200V for 1 hour followed by 30V for 10 hour. Gels were visualized and analyzed by Syngene Gene Genius Bio-imaging system after staining with ethidium bromide (5µg/ml). The extracted DNA was assessed by PCR amplification of 16S rRNA region followed by DGGE analysis.
6.1.2.7 PCR primer designing and amplification of alkaline protease gene/s

Amplification of extracellular alkaline proteases directly from soil was carried out by designing degenerate primers using CODEHOP method (Timothy et al., 2003). The designing of primers is as described in detail in materials and method section of chapter 5. The primer set that yielded the specific amplified product is as follows: SPS-5 forward 5’-ggagctcgccgagagacgactatatgtc-3’ and reverse 5’-ggagctgagctgagacgtc-3’; SPS-6 forward 5’-ggagctcgccgagagacgactatatgtc-3’ and reverse 5’-ggagctgagctgagacgtc-3’; SPS-7 forward 5’-gctctggagctggagaagcgc-3’ and reverse 5’-gtgatggctgagctggagaagcgc-3’.

The DNA preparations obtained in the present study were used as template to amplify region/s coding alkaline protease. The reaction mixture preparation and amplification protocol for protease gene amplification from soil DNA was as described in detail in Materials and Method section of Chapter-5.
6.1.3 RESULTS AND DISCUSSION

6.1.3.1 Isolation of Total DNA (Metagenome/Environmental DNA)
We have assessed and compared various methods for the extraction of total environmental DNA from saline soil of coastal Gujarat and optimized them for the quality, yield and PCR amplification ability. The applicability of various methods for the extraction of total DNA using small quantity of soil sample was explored. Total DNA was isolated from the two samples collected from Okha Madhi site by soft lysis, bead beating, sonication and by combination of these methods. In comparison to these methods, we also attempted the extraction of total DNA by Clean Gene Isolation kit (Mo Bio laboratories) and also tried several other approaches.

DNA extraction from saline soil in the present study had three folds objectives; lysis of representative microbes within the sample, obtaining high molecular weight intact DNA and removal of inhibitors from the extracted DNA for subsequent molecular manipulations. Therefore, various methods were examined for DNA extraction towards fulfilling these objectives. We have developed an improved method for isolating total metagenomic DNA from saline soil through which intact and unsheared DNA, amenable for further molecular biology work was obtained.

6.1.3.2 Purity and Yield assessment on the basis of spectrophotometer and agarose gel electrophoresis
The extracted DNA was assessed for purity and yield on the basis of absorbance ratios at 260/230 nm (DNA/Humic acids) and 260 nm/280 nm (DNA/protein) (Table 6.1.3.1). High ratio of 260/230 nm indicated the purity of extracted DNA with respect to humic acid contamination, whereas high 260/280 nm ratio was an indicative of the purity with respect to protein contamination. DNA samples were analyzed using 0.8% agarose gel using λDNA/Hind III digest (Merk life science, India) as marker and smart ladder (0.2-10 kbp) (invitrogen). There was no noticeable variation in the quality and quantity of DNA on the basis of agarose gel patterns (Fig. 6.1.3.1). The striking features of the extraction methods highlights that sonication alone was not suitable for DNA extraction from O.M.6.5, as humic acids content was not reduced and the purity and concentration of the extracted DNA did not compare favorably with other methods.
Fig. 6.1.3.1: Isolation of metagenomic DNA by various methods for saline soil sample O.M.6.2 and O.M.6.5. Lane 1: Lambda DNA /HindIII Marker (Banglo Genei), Lane 2: Soft Lysis, Lane 3: Bead Beating, Lane 4: Bead beating+Lysis, Lane 5: Sonication+Lysis Lane 6: Sonication Lane 7: Sonication+Bead Beating

Fig. 6.1.3.2: Agarose gel electrophoresis of the Total Genomic DNA. Lane 1: DNA ruler (Middle range marker, Merk life sciences, India), Lane 2: lysis buffer treatment (sample-6.2), Lane-3: Lysis Buffer Treatment (sample-6.5), Lane 4: Bead Beating only (sample-6.2); Lane 5: Bead Beating only (sample-6.5); Lane 6: Bead Beating + Lysis Buffer Treatment (sample-6.2), Lane 7: Bead Beating + Lysis Buffer Treatment (sample-6.5); Lane 8: Bead Beating + sonication treatment (sample-6.2); Lane 9: Bead Beating + sonication treatment (sample-6.5); Lane 10: lysis buffer + sonication treatment (sample-6.2); Lane 11: lysis buffer + sonication treatment (sample-6.5); Lane 12: sonication treatment (sample-6.2); Lane 13: sonication treatment (sample-6.5) (Fig.6.1.3.2).

However, the method based on sonication yielded better results with another sample, O.M.6.2 (Fig.6.1.3.2). The differential results by sonication may reflect on the fact
that the matrix of the concerned habitat might be causing a barrier against the extraction and lysis of the cell. Spectrophotometric assessment revealed that there were quite similarity in purity and yield of retrieved DNA from both sites. Comparative analysis indicated that the soft lysis and bead beating method yielded pure form of DNA from both the sites. Contradictorily, sonication method was not suitable for both the samples as the yield of DNA was very low when compared to other methods. Combinations of the above methods were quite encouraging as bead beating combined with lysis buffer treatment yielded pure DNA in good quantity as compared to bead beating and lysis buffer method independently, from both the site. On the other hand, sonication in combination with lysis buffer treatment did not emerge as a better alternate, as compared to their independent outcomes.

Bead beating emerged as equally effective method for the extractions of quality DNA in appreciable quantity from both environmental samples. Standardization of method, for 5 mins and 15 mins, on the basis of assessment was found that agitation for 15 mins gives optimum concentration, similarly size of bead also have a great influence on yield of DNA. The major factor, responsible for choosing bead size and time is type and characteristic of soil material. In present study, moderate size beads gave better result as compared to its counterpart. On the other hand, combinations of different mechanical methods lead to decreased concentration and purity of the extracted DNA with excessive shearing.

Soft lysis method proved best for O.M.6.2 and O.M.6.5, as it yielded higher concentration and eliminated humic acid to significant extent. Modification of soft lysis method with PEG (30%), was also tried for better yield, however encouraging results were not observed. The best quality of DNA was obtained by employing combination of soft lysis with bead beating method, while its combination with sonication was not as good in terms of DNA yield. The DNA preparations were considered for molecular biology applications, as obtained environmental DNA should not only satisfy the yield and purity criteria but it should also be amenable for further work (Desai and Madamwar, 2007).

As, we evaluated DNA extraction methods to identify a procedure that results in high molecular weight DNA that is relatively free from contaminants and maximizes detectable diversity. Bead beating treatment, an easy to perform method, is based on
the ballistic disintegration of the cells, where the results depend upon the time of agitation and bead size. The efficiency of cell disruption and consequently, the damage to the DNA strands during sonication mainly depends on the energy input. Even under optimized conditions, harsh treatment may result in shearing of high molecular weight DNA, low yields and small fragment sizes. This method may have a possibility of introducing a bias in microbial community analysis.

The enzymatic method relied on the proteinase K and lysozyme digestion of microbial cells to release DNA, while the treatment of soil with surfactants and chelating agents resulted into removal of inhibitors and prevented chemical flocculation with minimal loss of DNA yield. PVPP treatment was employed for removing traces of humic acid in DNA sample. Surprisingly, although literature reports that better quality of DNA is resulted after treatment, noticeable difference was not seen in our sample. A combination of mild bead beating and enzymatic lysis treatment emerged as the most successful protocol for recovering higher yields and inhibitor free DNA from saline soil sample. In order to obtain pure form of DNA in an easiest way, a modification of soft lysis method with cesium chloride and Ethidium Bromide was also attempted, in this case although the pure form of DNA was found, yield was quite less, which certainly limited its further applications.

Although, based on the spectroscopic analysis, humic acid was detectable to varying extent; the extracted DNA preparations were amenable for further molecular biology work. This finding appears to be a favorable observation in comparison to some reports in literature where humic acid strongly inhibited the DNA application in molecular biology (Kauffmann et al., 2004; Santosa 2001; Desai and Madamwar, 2007). The quality of the extracted metagenome is of prime importance in metagenomics, as the DNA should be suitable to proceed for molecular biological applications such as molecular diversity and functional genomics (Rajendhran and Gunasekaran, 2008).

6.1.3.3 PCR amplification of 16S rRNA gene

The environmental DNA extracted by all the above mentioned methods was used as template for PCR amplification. Amplification of the 16S rRNA gene (approximately-1.5kb) directly from undiluted DNA samples (Okha Madhi) indicated the high purity of DNA (Fig. 6.1.3.3, 6.1.3.4, 6.1.3.5).
Fig. 6.1.3.3: 16S rRNA amplification of Total DNA isolated by various method by using eubacterial universal primer Left Panel: 16S rRNA PCR of environmental sample using isolation methods Bead beating + Sonication Method, Lane 1: 0.2-10Kb ladder, Lane 2: Site 6.2 (Ta=52.4), Lane 3: Site 6.2 (Ta=55.7), Lane 4 : Site 6.2(Ta=56.9), Lane 5 : Site 6.5(Ta=52.4), Lane 6 : Site 6.5 (Ta=55.7), Lane 7 : Site 6.5 (Ta=56.9).

Right panel : (Bead Beating Method) Lane 1: 0.2-10Kb ladder, Lane 2: Site 6.2 (Ta=52.4), Lane 3: Site 6.2 (Ta=55.7), Lane 4 : Site 6.2(Ta=56.9), Lane 5 : Site 6.5(Ta=52.4), Lane 6 : Site 6.5 (Ta=55.7), Lane 7 : Site 6.5 (Ta=56.9)

Fig. 6.1.3.4: 16S rRNA Amplification from Total DNA of sample O.M.6.2 (Ta-64°C) Lane 1, smart ladder 0.2-10 kbp ladder (invitrogen); Lane 2, Lysis treatment; Lane 3, Soft Lysis + Bead Beating; Lane 4, Soft Lysis +Sonication; Lane 5, Bead beating; Lane 6, Sonication; Lane 7, Sonication+ Bead Beating. 16S rRNA Amplification from Total DNA of sample O.M.6.5 (Ta-64°C) Lane8, Lysis treatment; Lane 9, Soft Lysis +Bead Beating; Lane 10, Soft Lysis +Sonication; Lane 11, Bead beating; Lane 12, Sonication; Lane 13, Sonication+ Bead Beating
Fig. 6.1.3.5: 16S rRNA Amplification from Total DNA of sample O.M.6.2 (Ta- 62.5°C) Lane 1, Broad range ruler (Merk life science, India); Lane 2, Lysis treatment; Lane 3, Soft Lysis + Bead Beating; Lane 4, Soft Lysis +Sonication; Lane 5, Bead beating; Lane 6, Sonication; Lane 7, Sonication+ Bead Beating.

16S rRNA Amplification from Total DNA of sample O.M.6.5 (Ta- 62.5°C) Lane 8, Lysis treatment; Lane 9, Soft Lysis +Bead Beating; Lane 10, Soft Lysis +Sonication; Lane 11, Bead beating; Lane 12, Sonication; Lane 13, Sonication+ Bead Beating; Lane 14: Positive control

Fig. 6.1.3.6 16S rRNA Amplification from Total DNA of sample O.M.6.5 (Ta- 62.5°C) Lane 1, Broad range ruler (Merk life science, India); Lane 2, Lysis treatment; Lane 3, Soft Lysis + Bead Beating; Lane 4, Soft Lysis +Sonication; Lane 5, Bead beating; Lane 6, Sonication; Lane 7, Sonication+ Bead Beating

16S rRNA Amplification from Total DNA of sample O.M.6.5 (Ta- 62.5°C) Lane 8, Lysis treatment; Lane 9, Soft Lysis + Bead Beating; Lane 10, Soft Lysis +Sonication; Lane 11, Bead beating; Lane 12, Sonication
Total DNA preparations extracted by chemical lysis and bead beating method from the samples of both sites were used as template for PCR amplification of 16S rRNA gene. Amplification was successfully carried out in all the gradient range of temperatures selected for annealing by gradient PCR (Fig. 6.1.3.3, 6.1.3.4, 6.1.3.5, 6.1.3.6). Intense amplified band of 1.5 kb from the saline soil sample O.M.6.3 and O.M.6.5 was observed from both the sites. However, the intensity of amplicon varied at different T\textsubscript{a} used for profile generation (Fig. 6.1.3.6). In general, good amplification was observed at all the tested temperature conditions.

**6.1.3.4 Denaturing Gradient Gel Electrophoresis**

To gauge the utility of extracted DNA in molecular fingerprinting methods, especially in microbial ecology studies, the amplified 16S rRNA DNA was subjected to denaturing gradient gel electrophoresis. In DGGE at threshold concentrations of denaturant, different sequences of DNA, presumably from different bacteria, denatured resulting in a pattern of bands. As revealed in Fig 6.1.3.7, the DGGE band patterns of 16S rRNA amplified from different DNA samples obtained by various extraction protocols, were quite comparable. The observation was also reflected with the extracted DNA from different sample sites. Therefore, differences in DGGE banding pattern suggested that there was not much bias generated from DNA extraction procedures (Fig. 6.1.3.7). The diversity in the banding profile, however, revealed population heterogeneity and differences in both samples. Marker DNA (smart ladder, 10kbp) in denaturing gel did not generate bands according to standards (Fig. 6.1.3.7).

The described methods could allow the use of large scale preparations providing greater probability of detecting genes present in low abundance in the soil environment. These methods would be applicable to more challenging and heavily contaminated soils; therefore, microbial biodiversity assessment can now be more readily assessed and useful sequences could be retrieved.
Fig.6.1.3.7 DGGE (Denaturing gradient gel electrophoresis) of 16S rRNA amplicons

Upper panel, DGGE Analysis (Urea and Formamide as denaturant) of the PCR amplified product from Total DNA of sample O.M.6.1.2. Lane 1, smart ladder 0.2-10kbp ladder (invitrogen); Lane 2, Lysis treatment; Lane 3, Soft Lysis + Bead Beating; Lane 4, Soft Lysis +Sonication; Lane 5, Bead Beating; Lane 6, Sonication; Lane 7, Sonication+Bead Beating

Lower panel, DGGE Analysis (Urea and Formamide as denaturant) of the PCR amplified product from Total DNA of sample O.M.6.5. Lane 1, Lysis treatment; Lane 2, Soft Lysis + Bead Beating; Lane 3, Soft Lysis +Sonication; Lane 4, Beadbeating; Lane 5, Sonication; Lane 6, Sonication+BeadBeating; Lane 7, smart ladder 0.2-10kbp ladder (Invitrogen).

6.1.3.5 Alkaline protease gene amplification

Total DNA extracted by chemical lysis method from both the sample of Okha Madhi (O.M.6.2 and 6.5) were used as template for alkaline protease gene amplification. Selection of template for amplification was based on the results of quantification and
purity. All the sets of primers designed for alkaline proteases were used for amplification. Amplicons of varied size and concentration were obtained by using different sets of primer. Intense amplified bands were obtained by using combination of SPS-5, SPS-6 and SPS-7. Range of amplicons, were obtained of different size SPS-5 gave 0.5kb product for O.M.6.2 and 0.7 kb for O.M.6.5, the size of the product is quite less as compared to the size of alkaline protease judged from literature. SPS-5F and SPS-6R gave 1kb product for O.M.6.2 and for O.M. 6.5 no products were obtained (Fig.6.1.3.8). Similar, results were also obtained for SPS-5F and SPS-7R; on the basis of this it could be judged that availability of alkaline proteases in the sample O.M.6.5 would be less as compared to its counterpart soil sample. SPS-6 and combination of SPS6F and 5R generated no product. SPS-7F and SPS-7R gave product size of 1.2kb with O.M.6.5 while partial product of 0.5kb, 0.7kb, 1.1kb and 2.8kb was obtained with same primer combination (Fig. 6.1.3.9). However, SPS-6F and SPS-7R gave 1kb product with O.M.6.2 (Fig.6.1.3.9). Different sizes of bands were obtained after amplification procedure, which were quiet interesting for generating profile of alkaline proteases (Fig.6.1.3.9).

Amplification of varied size of products were obtained by SPS-7, as the primer were designed by using degenerate primer designing tool-CODEHOP (Fig.6.1.3.9). In this procedure, the chances of getting different types of proteases are quiet higher as compared to primer designed on the basis of individual known sequences. As, a whole CODEHOP primer under the optimized sets of protocol, would certainly allow to capture proteases sequences present in the particular saline soil at a given habitat. Along, with functional attributes this would also give us some of the information on diversity of proteases, particularly alkaline proteases. Further, for O.M.6.2, nucleotide sequence explored by chromosome walking method (Fig.6.1.3.10). On the basis of nucleotide sequence and ORF prediction, amino acid sequence was predicted (Fig.6.1.3.10). Alkaline protease sequence could be further explored to use as a marker trait for identification, knowledge driven process of saline soil of Okha Madhi (Gujarat, India).

While there is no doubt that multiple bands in Okha Madhi site could be due to the annealing of primer at multiple sites within the template, as this question could be better addressed by sequencing of the amplicons. To address, this ambiguity, PCR
products were run on low melting agarose (low EEO). Sample were gel eluted and sequenced by using SPS-6F and SPS-6R primer by using chromosome walking method. Complete sequence was determined and subjected to Mega 4.0 for phylogenetic determination. Sequence was found most homologus to protease type of family. All the characteristic features related to physico-chemical properties and structure details were elucidated.

Along the same line, it was further analyzed and confirmed by sequencing of dominant bands that multiple bands generated by a primer within a single reaction, is not the result of binding of primer at several position in same gene. The amplification profile was found to be reproducible from both the sites, a fact which could be explained on the basis of equal distribution of organism producing alkaline protease gene within a particular habitat and a good concentration of template DNA. Results of 16S rRNA PCR amplification and banding profiles visualized in DGGE provided the evidence for expediency of the DNA extraction protocol in studies related to molecular diversity (Ercolini, 2004). In view of the heterogeneity of the environmental samples, it is quite obvious that the extraction procedures would have to be case specific and hence need to be optimized for different soil samples (Santosa, 2001; Vereshchagin and Kostornova, 2008; Purohit and Singh, 2009; Siddhpura et al., 2010). However, the methods described in the present study appear to have wide applicability in investigating molecular diversity and exploring functional genes from the total DNA.

In a nutshell, given that the majority of natural products are of microbial origin, and that the vast majority of microbial genomes are yet to be explored, it’s quite logical that microbial metagenomes harbour a great economic potential. Due to their huge but largely unexplored diversity and history as sources of commercially valuable molecules with agricultural, chemical, industrial and pharmaceutical applications, marine environments would be among the most common habitats to explore from metagenomics view point (Morrissey et al., 2010). Improved functional screening methods would potentially provide a means to discover new variants of functions of interest.

With the possibilities to access vast genetic resources in different ecosystems, the unlimited realms of microbial diversity would slowly but steadily lead to new
knowledge and novel biotechnological avenues. However, the usual challenge of heterologus gene expression needs to be addressed to turn metagenomic technologies into commercial successes, particularly in applications where bulk enzyme or product have to be produced at viable cost (Kennedy et al., 2007).

The goals of researchers venturing into the microbial metagenome vary from directed product discovery to total community characterization and assessment of the phylogenetic complexity of the environments. Metagenomics has redefined the concept of a genome, and accelerated the rate of gene discovery. The potential for application of metagenomics to biotechnology seems endless. Metagenomics, together with in-vitro evolution and high-throughput screening technologies would provide unprecedented opportunities to bring new generation of biomolecules into various fields, besides adding to new knowledge in our understanding on biotic and abiotic interactions in ecosystems.

**Fig. 6.1.3.8: Functional attributes of alkaline proteases:** Amplification profile of alkaline proteases gene by different sets of combination of primer for O.M.6.2 and O.M.6.5: Sets of primers: SPS5F and SPSR; SPS5F and 6R; SPS5F and 7R; SPS6F and 6R; SPS6F and SPS5R; SPS6F and SPS7R; SPS7F and SPS7R; SPS7F and SPS6R; SPS7F and SPS5R
Fig 6.1.3.9: **PCR amplification of alkaline proteases genes**

(A) PCR amplification of alkaline protease gene by SPS7F and SPS7R  **Lane 1**: 62°C-O.M.6.2;  **Lane 2**: 63.5°C-O.M.6.2;  **Lane 3**: 60.1°C-O.M.6.5;  **Lane 4**: 62.3°C-O.M.6.5;  **Lane 5**: Low range DNA Ruler (3000bp), (Merk Life Science).

(B) PCR amplification of alkaline protease gene by SPS5F and SPS6R  **Lane 1**: Low range DNA Ruler (3000bp) (Merk life science);  **Lane 2**: 62°C-O.M.6.2;  **Lane 3**: 63.2°C-O.M.6.2;  **Lane 4**: 60.1°C-O.M.6.5;  **Lane 5**: 62.3°C-O.M.6.5;  **Lane 6**: 63.5°C-O.M.6.5

(C) PCR amplification of alkaline protease gene by SPS6F and SPS6R  **Lane 1**: 62°C-O.M.6.2;  **Lane 2**: 63.2°C-O.M.6.2;  **Lane 3**: 60.1°C-O.M.6.5;  **Lane 4**: 62.3°C-O.M.6.5;  **Lane 5**: Medium range DNA Ruler (5000bp)(Merk Life Science)

(D) PCR amplification of alkaline protease gene by SPS5F and SPS5R  **Lane 1**: DNA marker;  **Lane 2**: 63.2°C-O.M.6.2;  **Lane 3**: 60.1°C-O.M.6.5;  **Lane 4**: 62.3°C-O.M.6.5;  **Lane 5**: 62°C-O.M.6.2
Fig. 6.1.3.10: Partial nucleotide sequence analysis of O.M.6.2 alkaline proteases by chromosome walking method and partial amino acid sequence prediction of O.M.6.2 alkaline proteases by reverse translate tool (ExPASY).
SECTION-II

CAPTURING OF ALKALINE PROTEASES FROM SALINE SOIL METAGENOME: A CULTURE INDEPENDENT APPROACH
6.2.1 INTRODUCTION

An enormous variety of different biocatalysts or other functional products can be theoretically obtained using DNA extracted from a given environmental sample. By fragmenting total DNA from an alkaline marine sample, cloning it into an expression vector, and screening for protease/esterase/lipase activity in an easily cultivable host strain, 120 new enzymes were discovered, falling into 21 protein families (Miller, 2000). As, smaller cloned fragments are created; further necessitate larger gene banks required for a comprehensive and comparable coverage of the genetic informations (Kennedy et al., 2008). During the past five years, cloning of genes from the metagenome has become the most popular tool for cultivation-independent enzyme discovery, leading to the recovery of a range of new biocatalysts by academic and commercial institutions (Kennedy and Marchesi, 2007; Kennedy et al., 2008).
6.2.2 MATERIALS AND METHODS

6.2.2.1 Amplification, cloning procedures, expression analysis and one-step purification of alkaline protease enzyme
Amplification of metagenomics DNA was carried as described in section 6.1.2.2.1.8 of this chapter. Cloning procedures, expression analysis and one step procedures were carried out as described in cloning and over-expression of haloalkaliphilic isolates in chapter-5.

6.2.2.2 DNA Sequencing, In-silico analysis and 3D structure modeling
Plasmids were re-retrieved from positive clones and sequenced from both ends, using standard T7 promoter and terminator sequence which is on a flanking region of insert by chromosome walking method (Merk Life sciences, India). The amino acid sequence was deduced using CLC main workbench (Daintith, 2004). Sequence homologies and three dimensional structures of serine protease were modeled using the online I-TASSER server as described in detail in Chapter-5.

6.2.2.3 Nucleotide sequence accession number
The DNA sequence of the protease gene cloned and studied in present work was submitted in the GenBank database under the accession number HM219181 with its characteristic properties in native and recombinant system.
6.2.3 RESULT AND DISCUSSION

6.2.3.1 Cloning confirmation

Positive clones exhibiting resistance towards ampicillin (30µg/ml) were selected for plasmid isolation. For cloning confirmation, protease gene from environmental sample was sequenced by custom based service of Merk life science, India. Phylogenetically, protease sequence was found homologous to protease sequence of *Bacillus* sp. by the neighbor-joining method clustering strategy in Mega 4.0(www.megasoftware.net) (Tamura, 2007).

6.2.3.2 Functional analysis of recombinant clone

PCR-based cloning methods are being employed to recover novel enzymes. In most cases, degenerate primers are used, hybridizing with conserved regions that preferentially are located close to the extremities of the target genes (Liles *et al.*, 2008; Ni *et al.*, 2009). The characteristic features of native and recombinant enzymes were studied; interestingly, we noticed that recombinant clones have maintained their nascent properties with higher specific activity.

In detail, functional attributes of metagenomic clone was judged on gelatin agar plate. The zone of clearance was quite comparable with our studies on other reported haloalkaliphilic bacteria, halotolerant actinomycetes and other recombinant clones (Thumar and Singh, 2007; Dodia *et al.*, 2008a and b; Joshi *et al.*, 2008; Thumar and Singh, 2009; Singh *et al.*, 2010a and b; Purohit and Singh, 2011). Similar results were observed on recombinant clones studied for over-expression and characterization from the same soil sample.

6.2.3.3 Protein solubilization

An enormous variety of different biocatalysts or other functional products can be theoretically obtained using DNA extracted from a given environmental sample. However, to check for functional clone; within a metagenome DNA, by designing primer based approach is to search needle in a haystack (Handelsman, 2004; 2005; 2008). However, in our present studies it has been observed that, protein was able to solubilize. Along the same line, it was quiet interesting that to get alkaline protease as a signature sequence from a total DNA. Although, equally challenging is to get active
enzyme from a soil DNA. However, if draw a comparative picture of over-expression profile and enzyme activity of recombinant clones of haloalkaliphiles and metagenome derived clone; it is general trend seen that enzyme studied in current studies is far more sensitive as compared to cultivable alkaline proteases.

6.2.3.4 Effect of temperature
For, a soil derived recombinant clone, as compared to results of chapter-5, growth temperature of 27°C was optimum for secretion of recombinant protein on gelatin plate and SDS-PAGE; however, level of expression was also found satisfactory at other parameters. However, obviously growth as expected was definitely higher at 37°C as compared to 27°C.

6.2.3.5 Effect of IPTG induction
As described in chapter-5, levels of induction have a profound impact on expression analysis as transcriptions of genes are controlled by T7 strong promoter in pET21a+ (Novagen, Madisen, USA). At 1mM IPTG induction, higher amount of enzyme was produced as compared to 3mM which was evidently seen on SDS-PAGE. Level of induction, however, did not have significant effect on growth of host cell.

6.2.3.6 Synergistic effect of IPTG induction and temperature
Synergistic effect of temperature and IPTG was checked both at best combination i.e. 27°C and 1mM IPTG and 37°C and 3mM IPTG, to check the effect of best factors on protein solubilization (Fig. 6.2.3.1). According to literature study, it is generally observed that at low level of temperature and induction subsequent level of enzyme is over-expressed (Singh et al., 2009; Yan et al., 2009; Xu et al., 2009), similar results were also observed in the present studies; i.e. 27°C and 1mM IPTG; high level of protein was expressed (Fig. 6.2.3.1). However, in our studies subsequent amount of enzyme was also expressed at other mentioned parameter. The SDS-PAGE profile and determination of proteolytic activities patterns bear that there was no activity at basal level; however after four hours of induction, there was gradual increase of the target protein in soluble and insoluble fraction.
Fig. 6.2.3.1: **Synergistic effect of induction:** Effect of inducer (IPTG) was checked on growth of cells and enzyme production of O.M.A18 (Colony no: 1 and 2) and O.M. E12(Colony no: 3 and 4); where effect of 1mM IPTG at 37°C; 1mM IPTG at 27°C; 3mM IPTG at 37°C; 3mM IPTG at 27°C

6.2.3.7 Proteolytic activity assay

A result of SDS PAGE was quite comparable with the proteolytic assay in terms of its activity. No basal level activity was seen in uninduced sample, with increase in time significant amount of activity was monitored after 6 hours of induction (Fig. 6.2.3.3).

6.2.3.8 Purification of protein

To facilitate purification, the recombinant alkaline proteases, which carries His-tag at its C-terminal in pET 21a+ was exploited. Purification was achieved at its homogeneity by one step chromatography; by using 50mM imimidazole concentration. Complete purification was evident from SDS-PAGE and specific activity of 6765.76 with fold purification of 6.41 and high yield (Table 6.2.3.1). Profound amount of enzyme was noticed in soluble and insoluble fractions. Result holds significance as, within one step of purification significant amount of enzyme is produced. Along this line, characteristic features of alkaline proteases derived from haloalkali philies and metagenome clone would be quite similar; evident from similar purification strategies.
Fig. 6.2.3.2: Effect of IPTG and temperature on growth and enzyme secretion: SDS PAGE of comparative effect of temperature and IPTG induction on growth and secretion of alkaline protease enzyme.

**Left panel Soluble fraction:** Lane 1: Protein molecular weight marker (3500-205000Da); Lane 2: 27°C; 1 mM (0 hr); Lane 3: 27°C; 1 mM (2 hr); Lane 4: 27°C; 1 mM (4 hr); Lane 5: 37°C; 3 mM (6 hr); Lane 6: 37°C; 3 mM (24 hr). **Right panel Insoluble fractions:** Lane 1: Protein molecular weight marker (3500-205000Da); Lane 2: 27°C; 1 mM (0 hr); Lane 3: 27°C; 1 mM (2 hr); Lane 4: 27°C; 1 mM (4 hr); Lane 5: 27°C; 1 mM (4 hr); Lane 6: 37°C; 3 mM (6 hr); Lane 7: 37°C; 3 mM (24 hr).

These results are quite encouraging and interesting in light of verity that, purification was achieved in one step, and substantial level of enzyme was obtained in simple bacterial system. The approximate calculated size of protein is estimated to be around 30 KDa, which is quite comparable to our bioinformatics based prediction and our study on the same protein secreted from several haloalkaliphilic bacterium (Fig.6.2.3.2).
Table 6.2.3.1: One step purification of enzyme O.M.6.2 by affinity chromatography

<table>
<thead>
<tr>
<th>Enzyme Preparations</th>
<th>Activity (U/ml)</th>
<th>Total activity (U)</th>
<th>Protein (mg/ml)</th>
<th>Total protein (mg)</th>
<th>Specific Activity (U/mg)</th>
<th>Yield</th>
<th>Purification fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombinant fraction</td>
<td>216</td>
<td>1728</td>
<td>0.204</td>
<td>1.632</td>
<td>1054.94</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>Purified enzyme</td>
<td>751</td>
<td>1502</td>
<td>0.111</td>
<td>0.222</td>
<td>6765.76</td>
<td>86.92</td>
<td>6.41</td>
</tr>
</tbody>
</table>

6.2.3.10 Characterization of enzyme

Effect of Temperature and pH

The pH profile for the isolate was quite broad and they were able to grow and secrete protease at pH, 7-10 (Fig. 6.2.3.3A). Although, the optimum pH for enzyme secretion was pH-7, there was only marginal difference in enzyme secretion from pH-7 and 9; enzyme was not able to maintain its activity at 10. On the basis of our study on moderate saline habitats from coastal region of Gujarat, from last 15 years, we have not come across enzyme; which is not catalytically active at higher alkaline range; infact we have several reports of alkaline protease active at pH-9, 10. Similar, contradictory results were obtained in studying effect of temperature; the organisms were able to secrete enzyme efficiently at 37°C, however sparse activity was seen at 50°C, while enzyme was completely denatured at 60°C(Fig.6.2.3.3B). These results are equally intriguing, as we have several isolates isolated from same soil sample O.M.6.2; which are active upto 50°C-90°C (i.e., Oceanobacillus iheyensis O.M.E12 (EU680960); Haloalkaliphilic bacterium O.M.A18 (EU680961). Infact, it was found that an enzyme activity increases in magnitudes with increase in temperature to its optimum. Similar trend is not noticed in our present studies.
Fig. 6.2.3.2: Effect of pH (7, 8, 9, 10) and Temperature (35, 50, 60, 70) on recombinant alkaline proteases.

**Thermostability of enzyme**

The thermal stability of the recombinant enzyme was assessed for 12 hours at temperatures, 37, 50 and 60°C, and pH 8. Broad range of activity difference was noticed with respect to elevated temperature. At, 37°C around 80% of activity was retained while with increase in temperature to 60°C, more than 90% of the activity was lost. With prolonged incubation up to 2 hours, only 30% of total activity was maintained at 37°C and marginal activity at 50°C. While, total loss of activity was observed in higher range of temperature, i.e. 60°C (Fig. 6.2.3.4). Further, on extending the time of incubation to 24 hour, enzyme was completely denatured at set temperature range (Fig. 6.2.3.4). From our current study, it’s apparent from the results that haloalkaliphilic nature of enzyme is maintained. Although, the resistant nature of enzyme is comparably very low as compared to extracellular alkaline proteases from haloalkaliphilic bacterial systems as reported in our literature (Patel *et al.*, 2005a and b; Patel *et al.*, 2006 a and b; Dodia *et al.*, 2008a and b; Joshi *et al.*, 2008; Thumar and Singh, 2007; Thumar and Singh, 2009) and our studies on characterization of recombinant proteases. On the basis of these studies; it is quite logical to suggest that the stability of proteases could be due to their genetic adaptability to carry out their biological activity at a higher temperature in haloalkaliphiles. But, the characteristic feature of enzyme is quite different as compared to our own studies and also available
literature. However, more detailed information on enzyme confirmation and structure could be explored by studying 3D structure and enzyme energetic of novel proteases.

**Effect of NaCl on enzyme stability**

Effect of NaCl on enzyme secretion was studied by incubating the reaction mixtures supplemented with various concentrations of NaCl (0-3M) (Fig. 6.2.3.5). Enzyme was found to be more sensitive in presence of salt, in all gradient of NaCl concentration around 25-40% of activity was reduced within an hour of incubation. With increase in time to 2 hours; at 4M concentration, complete activity was lost; however at 1, 2, and 3M concentration around 50% activity was retained which was completely lost at 12 hours of incubation. As discussed in our earlier results; these observations are also quite more sensitive than our similar studies to other recombinant enzymes; this itself provokes as a unique characteristic of this enzyme. In similar studies done with haloalkaliphilic organism; it was a general trend observed; with increase in NaCl concentration upto threshold there is increase in activity (Dodia *et al.*, 2008a; Joshi *et al.*, 2008). However, similar results are not noticed in present studies.

**Fig. 6.2.3.4: Thermostability profile of recombinant enzyme**: Thermostability of enzyme was characterized after different hours (1, 2, 3, and 24) of incubation at (37°C—); (50°C -■-), (60°C -▲-).
Fig. 6.2.3.5: Stability of NaCl : Stability of NaCl; where (1M♦-); (2M ■-), (3M▲-) was checked on recombinant alkaline proteases after different hours of incubation

**Urea denaturation**

Maximum amount of enzyme was able to maintain its active confirmation in the presence of chemical denaturant urea in narrow temperature; 37°C. However, enzyme was also able to maintain its marginal activity at 50°C for 30 mins, however total loss of activity was observed at both the temperature with further increase in temperature. (Fig.6.2.3.6). Percent residual activity was related to zero hour enzyme activity as 100%.

Fig. 6.2.3.6: Urea Denaturation profile: Effect of chemical denaturant urea was checked on enzyme after different time interval (37°C (-♦-), 50°C (-■-)).


6.2.3.11 *In-silico* analysis of protease gene

A salient feature of protein sequence was analyzed; by reverse translating the nucleotide sequence using translate tool (ncbi.nlm.nih.gov.in). The instability index (II) is computed to be 39.57. This classifies the protein as stable. Aliphatic index: 42.94. Grand average of hydropathicity (gravy): -0.747 theoretical PI/MW:5.15/46683.44. Predicted N-terminal sequence for; O.M.6.2 clone"MRQSLKVMVLSTVALLFMANPAAGSEEKKEYLIVVEPEEVSAQSVEES YDVDVIHEFEEIPVIHAELTKKELKKKDPNVKEHPAGA.On the basis of data, as similar to results of O.M.A18 and O.M.E12 described earlier in chapter-4 and 5, we can predict that structure of enzyme is quite stable, a fact which is strongly, reflected by our experimental data on thermal stability and resistance against chemical denaturation.

6.2.3.12 Hydropathy and 3D structure determination

The hydropathy profile of the nucleotide sequence of O.M.6.2 protease, showed increased presence of hydrophobic residues (Fig. 6.2.3.7A). We further predicted that enzyme contained serine amino acids at its active site by I-TASSER tool (Fig.6.2.3.7B), which was supported by our findings on the inhibitor studies, where polymethyl sulfonyl chloride (PMSF) strongly inhibited the enzyme activity. The stability of *in-silico* structure was predicted by Ramachandran plot-PROCHECK expasy tool, plotting psi vs. phi value; where it was analyzed that molecule was stable in its confirmation (6.2.3.8). To judge the relatedness, known alkaline proteases sequences were aligned with metagenomic O.M.6.2 alkaline proteases The phylogenetic position in constructed tree is shown in Result and Discussion section of chapter-7.Phylogentic analysis of recombinant metagenomics enzyme confirmed its 100% homology with extracellular protease sequence gene using Mega 4.0.
**Fig.6.2.3.7A:** Hydropathy analysis for O.M.6.2 protease according to Kyte and Doolittle. On the plot, a positive peak indicates a probability that the corresponding polypeptide fragment is hydrophobic (a negative peak indicates a probable hydrophilic segment).

**Fig.6.2.3.7B:** 3D structure prediction of O.M.6.2 protease enzyme by I-TASSER structure prediction tool
Overall, the results discussed in section II of chapter-6, are quite novel with viewpoint that enzyme studied is metagenomic in nature and characteristic feature reflected are quite different than our earlier studied enzymes from same saline soil sample. Further exploration of enzyme from soil; would address heterogeneity or diversity of soil sample. Identification of such unique properties could also be served as marker properties of proteases. Availability, of active recombinant protein; possessing ability to work under moderate to harsh condition would be of significance important from biotechnological standpoint. Enzyme could also be further explored for its commercial applications and for studying structural and functional properties of serine alkaline proteases.
CHAPTER 7

COMPARATIVE ANALYSIS OF NATIVE, RECOMBINANT AND METAGENOMIC ALKALINE PROTEASES WITH RESPECT TO THEIR TOLERANCE AGAINST ORGANIC SOLVENTS
7.1 INTRODUCTION

We have discussed in detail regarding characteristics, applications and commercial exploitation of proteases from the haloalkaliphilic bacteria in our previous chapters. Among them, one of the characteristic feature of proteases is that they are among the most valuable catalysts used in food, pharmaceutical and detergent industries as they hydrolyze peptide bonds in aqueous environments, while synthesize peptide bonds under microaqueous conditions (Ogino et al., 2001). In addition to proteolytic activity of protease, its application in organic synthesis has generated significant interest (Meos et al., 1993; Klibanov, 2001; Bordusa, 2002; Diego et al., 2007; Karan and Khare, 2010).

Organic solutions are ideal for synthesis reactions since the solubility of polar substrates increases in solutions supplemented with organic solvents and the stability of the alkaline proteases in organic solvents would be an attractive feature of the biocatalysis. Along with stability, it favors reversal of thermodynamics equilibrium over hydrolysis and decreases microbial contamination. Among the major applications of protease-catalyzed reaction is the synthesis of dipeptides, such as kyotorphin (Tyr-Arg) precursors (Meos et al., 1993; Sareen et al., 2004 a, b).

In recent years, several new proteases that are able to maintain stability and activity in organic solvents have been discovered (Patil et al., 2008; Reza et al., 2008 a and b; Hamid et al., 2011). There are many approaches to capture the non-aqueous biocatalytic potential; the most obvious being the microorganisms from extreme environments or contaminated areas enriched with various organic solvents (Hamid et al., 2011). The property of tolerating organic solvents makes these bacteria better candidates for exploiting solvent-stable enzymes. However, alternatively, genetic engineering and molecular biology could be considered as one of the approaches, to transfer solvent resistant gene/s. Genetic engineering is instrumental in opening new opportunities for the construction of genetically modified microbial strains with selected enzymes properties.

The demand for potentially useful proteases with specific properties continues to stimulate the search for new sources of organic-solvent tolerant proteases. With
diversity viewpoint, most of the reported solvent tolerant strains belong to genera; *Pseudomonas*, *Bacillus* and *Arthrobacter*. However, exploration of haloalkaliphilic bacteria with such unique characteristic feature is in infancy. Therefore, studies on the solvent tolerant nature of haloalkaliphilic bacteria from moderate saline habitats may open new arena for the basic research in non-aqueous enzymology.

In this chapter, we selected several organic solvents on the basis of their Log pOW and analyzed their effect on the native and recombinant proteases from O.M.A18 and O.M.E12 strains. The studies were then compared with a metagenomically derived alkaline protease from the saline habitat. Further, effect of physico-chemical parameters such as temperature, pH and NaCl were analyzed.
7.2 MATERIALS AND METHODS

7.2.1 Bacterial strains
Haloalkaliphilic bacteria, *Oceanobacillus iheyensis* O.M.A18 (EU680961) and *Haloalkaliphilic bacterium* O.M.E12 (EU680960) were isolated as earlier described in detail in chapter 2 and Purohit and Singh (2011).

7.2.2 Recombinant clones
Construction of recombinant clones of O.M.A18 and O.M.E12 is described in detail in chapter-5 and construction of metagenomic clone O.M.6.2 is described in chapter-6 was used for solvent studies. Sequencing of clones was done as described in detail at appropriate place in chapter-5 and 6.

7.2.3 Organic Solvents
Glycerol, Xylene, Methanol, n-Hexane, Acetone and Chloroform, with log P<sub>ow</sub> values as 1.07, 3, 0.82, 0.25, 0.2 and 1.9 respectively, were obtained from Merk chemicals (India).

7.2.4 Effect of organic solvents on enzyme catalysis
Protease activities were measured in a reaction mixture (Hagihara, 1958) with varied concentrations of 10-30% (v/v) of above mentioned solvents. Controls for each set were also carried out simultaneously.

7.2.5 Effect of organic solvents on enzyme stability
The solvent stability was studied by incubating the enzymes in different solvents (Methanol, Glycerol and Hexane) at 10% (v/v). The aliquotes of enzyme preparations were withdrawn at regular intervals for 15 hours and the residual enzyme activities were measured.

7.2.6 Effect of pH on enzyme catalysis
Effect of pH on native and recombinant proteases were examined by carrying out enzyme assay at different pH in the presence of (10% (v/v)) Hexane, using buffers systems (20mM): phosphate (pH-7), Tris-HCl (pH 8), NaOH-Borax (pH 9) and Glycine - NaOH (pH 10).
7.2.7 Effect of NaCl on enzyme activity
To assess the influence of NaCl and organic solvent, Hexane in conjunction, the reaction mixtures were supplemented with 1-4M NaCl and protease assay was carried out at 37°C with 10% (v/v) of the solvents.

7.2.8 Effect of Temperature on proteases catalysis
The temperature profile for protease activity was examined in the presence of Hexane by incubating the assay reaction mixtures at different temperatures, 30-60°C. The proteases activity was determined as mentioned above.

7.2.9 Multiple sequence alignment and phylogenetic determination
Multiple sequence alignment of three nucleotide sequences (recombinant O.M. A18, recombinant O.M.E12, metagenomic clone) were constructed by using CLUSTALW (www.ebi.ac.uk/Tools/maa/clustalw2/) (Thompson et al., 1994); data were further interpreted by CLC workbench to analyze conserved residue and consensus pattern (Dainith, 2004). A phylogenetic tree was constructed of aligned three enzymes studied in present work. The phylogenetic relatedness of these enzymes was also established with other proteases sequences of haloalkaliphilic organisms by the Neighbor-Joining method clustering strategy in Mega 4.0 (www.megasoftware.net) (Tamura et al., 2007).
7.3 RESULTS AND DISCUSSIONS

While there are many studies on the solvent resistant strains, the sensitivity of the organisms and their enzymes from saline habitat are scarce (Reza et al., 2009; Karan and Khare, 2011). There are several reports where effect of solvent is checked on enzyme activity (Thumar and Singh, 2007a; Reza et al., 2008) Similiarly, construction of recombinant clones and to study similar aspects are also studied (Reza et al., 2009; Hamid et al., 2011). However, to the best of our knowledge, study on comparison of catalytic efficiency of native and recombinant preparations is not found in literature. In Chapter-7, we have studied varied enzyme preparations; native, recombinant and metagenomic clone in a comparative manner.

7.3.1 Native, Recombinant and Metagenomic derived alkaline enzymes

Alkaline proteases were purified to its homogeneity from both the selected haloalkaliphilic bacterial strains using Phenyl Sepharose 6FF described in detail in, Purohit and Singh (2011) and Chapter 4. The construction of recombinant clones and purification of recombinant enzymes is described in detail in chapter 5 and 6. Sequence analysis of recombinant clones and the amino acid sequence prediction and analysis of its physico chemical properties were carried out as described at appropriate place in detail in chapter 5 and 6.

7.3.2 Catalysis of alkaline proteases in organic solvents

The five enzyme preparations; two native, two recombinants and one metagenomic, were studied with respect to their alkaline protease activities in the presence of various organic solvents described in Materials and Methods.

As regard to the effect of ethanol on the catalysis, for O.M.A18 native enzyme, 73% of the residual activity was retained at 5% ethanol, which on further increase in ethanol to 30%, reduced to 16.33%. Compared to native, the recombinant enzyme was relatively more sensitive, with 62 % loss of the residual activity at 5%. The activity was substantially reduced at 10% solvent. Further, enzyme was completely denatured at 30% ethanol (Fig.7.3.1).
In general, the native O.M.E₁₂ protease was more sensitive towards organic solvents as compared to its O.M.A₁₈ counterpart. The activity reduced to 52% of total activity with 10% methanol, while only 11.56% of the residual activity was evident at 20% solvent, followed by a total loss at 30%. The recombinant O.M.E₁₂ enzyme was relatively more sensitive than the native enzyme. For metagenome derived protease, only 21.5% residual activity was observed at 5% methanol (Fig. 7.3.1).

Around half of the activity of O.M.A₁₈ native enzyme was lost in the presence of 10% glycerol, with a complete loss at 30%. On the other hand, with reference to O.M.E₁₂ enzyme, 60% loss in enzyme activity at 10% and complete denaturation at 20% glycerol was evident. While, for recombinant O.M.A₁₈ enzyme, only one-tenth residual activity was apparent at 10% glycerol, with a total loss at 20%. The metagenomically derived protease maintained 12% of the residual activity at 10% glycerol.

The trends of enzyme responses in acetone and chloroform were quite similar for both native enzymes. At 5% acetone and chloroform, 35-40% loss in activities was observed for O.M.A₁₈ and O.M.E₁₂ native enzymes, leading to a total at 30% solvents. For recombinant enzymes, trends were quite different with different solvents. The O.M.A₁₈ recombinant protease maintained 30% of the residual activity at 5% of acetone and chloroform, while at 20% (v/v) solvents, nearly total loss of the activity was evident. The O.M.E₁₂ recombinant enzyme had 20-30% of the residual activity at 5% acetone and chloroform, and with increase in solvent to 20%, the activity reduced to 50%. The metagenomically derived enzyme was highly sensitive to acetone and chloroform (Fig.7.3.1) (Table 7.3.1, Table 7.3.2). The solvent tolerance of the enzymes was greater towards hexane and xylene for native O.M.A₁₈ and O.M.E₁₂ proteases. These enzymes retained approximately 70-80% activity at 5% hexane and xylene. However, the residual activities of both native enzymes reduced to 30-35% at 20% hexane. The recombinant enzymes, O.M.A₁₈ and O.M.E₁₂ had similar trends. The maximum activity at 40% residual level was apparent with 20% solvent (Fig.7.3.1), (Table 7.3.1, Table 7.3.2).

As described in review of literature, the log P values which are less than 4 are considered extremely toxic, as required water molecules on the enzyme surface are easily replaced with solvents (Ogino et al., 2001; Gupta et al., 2006b). In brief,
native proteases retained substantial catalysis at lower concentrations of both hydrophobic and hydrophilic solvents. Interestingly, at increased solvent concentrations, comparatively better activities were evident with hydrophobic solvents. Such trends of enzymatic efficiency for native enzymes are frequently sited in literature (Reza et al., 2008; Reza et al., 2009).

7.3.3 Effect of organic solvents on enzyme stability

The stability of enzymes was determined with the solvents: methanol, glycerol and hexane. The solvents were selected on the basis of trends displayed for enzyme catalysis. The enzymes were incubated with 10% (v/v) solvent up to 15 hours to monitor stability (Fig. 7.3.2).

O.M.A₁₈ native enzyme in methanol, at zero hours itself, it lost 43% residual activity compared to its control. The time required for native O.M.A₁₈ to reduce to 50% of its initial activity was 3 hours (Table 7.3.3, Table 7.3.4). For O.M.E₁₂ native enzyme, the enzyme activity was lost by 10-15% at every 3 hours of incubation and after 15 hours, 30% of the residual activity was recorded (Fig. 7.3.2). Native O.M.E₁₂ enzyme was resistant against solvents as half life in methanol was 9 hours (Table 7.3.3, Table 7.3.4). In comparison, recombinant O.M.A₁₈ enzyme in methanol had 70% of the residual activity after 6 hours of incubation. For recombinant O.M.E₁₂, trend was quite similar to native enzyme (Table 7.3.3, Table 7.3.4).

In general, for hexane and xylene trends were quite similar to methanol. With increase in incubation time, as the concentration of the solvents increased, the enzyme stability significantly decreased. With respect to metagenomic clone, around 40% of the activity was maintained between 0-9 hours of incubations. Overall, stability of n-hexane was highest as compared to other two solvents (Fig. 7.3.2), (Table 7.3.3, Table 7.3.4).

In general, proteases are relatively more stable in n-hexane, a hydrophobic solvent than hydrophilic solvents: methanol and glycerol. Almost similar stability trends for proteases in the presence of various organic solvents have been reported by others (Ogino et al., 1995; Gupta et al., 2006a).
Fig. 7.3.1: Effect of solvents on enzyme catalysis of native O.M.A₁₈ and O.M.E₁₂, recombinant O.M.A₁₈ and O.M.E₁₂ and metagenomics sample O.M.6.2. Organic solvents are; Glycerol, Xylene, n-Hexane, Methanol, Acetone, Chloroform.
Table 7.3.1: Comparative analysis of effect different solvents with range of solvent concentration on enzyme preparations

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Methanol</th>
<th>Glycerol</th>
<th>Acetone</th>
<th>Chloroform</th>
<th>n-hexane</th>
<th>Xylene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration (%)</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>O.M.A18 native enzyme</td>
<td>73.3</td>
<td>52.86</td>
<td>12.45</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O.M.A18 recombinant enzyme</td>
<td>100</td>
<td>73.45</td>
<td>39.73</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.3.2: Solvent concentrations (%) required to reduce enzyme catalysis to its half.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Methanol</th>
<th>Glycerol</th>
<th>Acetone</th>
<th>Chloroform</th>
<th>n-hexane</th>
<th>Xylene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration (%)</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>O.M.E2 native enzyme</td>
<td>21.5</td>
<td>10.56</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>O.M.E2 recombinant enzyme</td>
<td>100</td>
<td>38.88</td>
<td>11.56</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Methanol</th>
<th>Glycerol</th>
<th>Acetone</th>
<th>Chloroform</th>
<th>n-hexane</th>
<th>Xylene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration (%)</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>O.M.6.2 metagenome</td>
<td>100</td>
<td>28.55</td>
<td>11.56</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Fig. 7.3.2: Stability of enzyme in the presence of organic solvent at different time interval (hours)
Table 7.3.3: Effect of organic solvents of different concentration on enzyme stability

<table>
<thead>
<tr>
<th>Solvents</th>
<th>O.M.A₁₈ native enzyme</th>
<th>O.M.A₁₈ recombinant enzyme</th>
<th>O.M.E₁₂ native enzyme</th>
<th>O.M.E₁₂ recombinant enzyme</th>
<th>O.M.E₆.₂ metagenome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (hours)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Control</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Methanol</td>
<td>57.3</td>
<td>50.3</td>
<td>43.9</td>
<td>34.8</td>
<td>30.5</td>
</tr>
<tr>
<td>Glycerol</td>
<td>54.3</td>
<td>49.6</td>
<td>39.8</td>
<td>33.6</td>
<td>28.4</td>
</tr>
<tr>
<td>n-Hexane</td>
<td>62.5</td>
<td>58.6</td>
<td>51.6</td>
<td>46.8</td>
<td>40.2</td>
</tr>
</tbody>
</table>

Table 7.3.4: Stability of enzyme in the presence of organic solvent: Time required by enzymes to reduce its activity to 50%.

<table>
<thead>
<tr>
<th>Solvent concentration (%)</th>
<th>O.M.A₁₈</th>
<th>O.M.E₁₂</th>
<th>O.M.E₆.₂ metagenome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Native</td>
<td>Recombinant</td>
<td>Native</td>
</tr>
<tr>
<td>Methanol</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Glycerol</td>
<td>3</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>n-Hexane</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

7.3.4 Influence of pH on enzyme catalysis

Effect of pH was determined on all five recombinant enzyme preparations in the presence of varying concentrations of Hexane. With increase in solvent concentration, the enzyme activity was profoundly reduced at different pH. However, extent of the activity loss varied with pH. The trends of enzyme activity at different concentrations were assessed. pH profile were quite similar for both, native and recombinant enzymes. Native and recombinant O.M.A₁₈ enzymes had maximum activity at pH-8. With increase in pH, residual activities marginally reduced. However, the difference in activity profile was evidently revealed at 10% solvent concentration. While with
higher solvent concentrations, the differential effect of pH was less evident at pH-8, 9 and 10. With 30% solvent, around 30-35% of the residual activity was maintained at all pH (Fig.7.3.3).

With increase in solvent concentrations, there was gradual decrease in enzyme activity. Even at lower pH, significant residual activity was observed. There was no significant difference in the activity at pH-9 with 10 and 20% solvent for the native enzyme. For alkaline protease clone generated from soil sample, most favorable pH was 8. There the residual activity was significantly lost (Fig.7.3.3).

**7.3.5 NaCl effect on enzyme catalysis**

In presence of NaCl there was not much difference in trend revealed at pH-7 and 9 on catalysis in presence of n-hexane. With increasing solvent concentrations in the presence of NaCl, there was reduction in residual activities. For O.M.E\(_1\)2 native and recombinant enzymes, the residual activity was reduced to 50% at 1M NaCl. With, further increase in salt concentration to 2M, there was marginal increase in activity (Fig.7.4.4). For O.M.A\(_1\)8 enzymes, the observed trends were quite similar to O.M.E\(_1\)2, where although there was significant loss of activity at 1M NaCl, with increase in salt concentration to 2M, the percent residual activity was enhanced (Fig.7.4.4). With reference to metagenomic clone, the reduction of activity was similar to above discussed trends of native and recombinant clones; however increase in activity with enhanced salt concentration to 2M was not noticed in metagenomic clone.

**7.3.6 Effect of Temperature in enzyme catalysis**

Temperature profiles of all enzyme preparations were studied in the presence of n-Hexane, where it was clearly indicated that optimum temperature was 37°C. For, native and recombinant O.M.A\(_1\)8; with increased solvent concentrations, there was loss in activity. With increase in solvent concentration to 30%, around half of the residual activity was maintained compared to 10% of its residual activity. With reference to pH profile, there was not much change in percent residual activity of native, recombinant and metagenomic enzyme preparations. This clearly indicates the efficacy of recombinant clones. For O.M.E\(_1\)2, both native and recombinant enzyme displayed similar trends to O.M.A\(_1\)8 (Fig.7.3.5). The optimum temperature of metagenomic clone was 37°C. The loss of activity was quite significant as compared to earlier studied enzyme systems (Fig.7.3.5).
7.3.7 Phylogenetic identification

Nucleotide sequence of recombinant enzymes; O.M.A18, O.M.E12 and metagenomically derived protease were identified. The consensus amino acid sequence was deduced from the DNA sequence. On the basis of phylogenetic tree constructed by neighbor joining (NJ) method, using Mega 4.0, it was revealed from the inter node that O.M.A18 was closely related to uncultured bacterium O.M.6.2. O.M.E12 and both this described sequences are diverged from nodes of constructed tree. To get more insight into phylogenetic relatedness, the sequences in the present studies were aligned with known protease sequences available in the database. The enzyme sequences were quite diverse in its sequence similarity; they were more closely related to enzyme sequences of other halophilic organisms as compared to enzyme studied in present case which were of the haloalkaliphilic organism isolated from the same saline soil (Fig.7.3.6, Fig.7.3.7).
Fig.7.3.3: Effect of pH on enzyme catalysis in presence of Hexane with varied solvent concentrations (0-30%).
Fig. 7.3.4: Effect of NaCl on enzyme catalysis in presence of Hexane with varied solvent concentrations (0-30%).
Fig. 7.3.5: Effect of Temperature on enzyme catalysis in presence of Hexane with varied solvent concentrations (0-30%).
Fig. 7.3.6: Snapshot of Multiple sequence alignment of alkaline proteases sequences of O.M.A18, O.M.E12 and metagenomics sample O.M.6.2 by Clutal W and CLC workbench.
Fig. 7.3.7: Phylogenetic analysis of sequences by Neighbor joining method using Mega 4.0. **Upper Panel:** Phylogenetic relatedness of O.M.A18, O.M.E12 **Lower Panel:** Phylogenetic relatedness of several alkaline proteases
Although, there is much advancement in the field of molecular biology and biochemistry, only sparse number of archeal and bacterial haloalkaliphilic proteases have been reported as an organic solvent-stable protease (Diego et al., 2007). With this above objective; the behavior of alkaline proteases in its native and recombinant counterpart as well as functional attribute based metagenome alkaline protease were studied in presence of solvent. To, the best of our knowledge, this would be among the few reports dealing with comparative study of enzyme preparations, particularly organic solvents.