5.1 Introduction

The ecophysiology of most cyanobacteria till date lacks proper understanding, and furthermore, there is often no simple correlation between genotypic and phenotypic diversity (Crosbie et al., 2003). The availability of a repertoire of genetic systems now permits application of genetic technology to identify the mechanisms governing complex physiological activities like photosynthesis, nitrogen fixation and membrane biogenesis. Synthesis and assembly of phycobilisomes require exquisite regulation of the genes for phycobiliproteins and linker proteins as well as co-ordinate regulation of the heme biosynthetic pathway to provide phycocyanobilin and phycoerythrobilin chromophores needed for synthesis of α and β subunits of phycobilisomes. In cyanobacteria, this is achieved in part by organization of genes for phycobilisomes and associated linker proteins into polycistronic transcription units that are differentially expressed and processed such that the abundance of transcripts for components approximates their composition in phycobilisomes (Troxler et al., 1995). Genes encoding phycobilisome components have been isolated and studied from several cyanobacteria and red algae (Conley et al., 1986; Mazel et al., 1988; Hess et al., 1999; Apt et al., 2001; Yu et al., 2002). The phycobiliproteins are encoded by a gene family and exhibit varying degrees of sequence homology. The entire PC operon contains PC genes coding for two bilin subunits and three linker polypeptides (Neilan et al., 1995). Genes encoding the linker proteins have also recently been identified. The core linker gene is immediately downstream from the APC gene set in both *Anacystis nidulans* (Houmard et al, 1986) and *Fremyella diplosiphon* (Lomax et al., 1987) and is co-transcribed with the APC genes in the latter species. Linker genes have also been located downstream from the PC gene sets

in *Synechococcus* sp. Strain PCC 7002 (Lomax et al., 1987) and *Anabaena* sp. Strain PCC 7120 (Belknap, 1987). Füglistaller et al., (1985) have shown that these linker polypeptides show homology to each other and also contain sequence homologous to the phycobiliproteins, suggesting that all of the genes encoding PBS components have evolved from a single ancestral gene. The ancestral gene product appears to have some similarities with globin (Schirmer, 1985).

The isolation of genes encoding important structural proteins provides sequence information and a means to acquire further insights into structure function relationships. Sequence analysis at the nucleotide level will help in better understanding on the evolution of phycobiliproteins and to search the regulatory sequences of these genes for unique features. The aim of the present study is to isolate and characterize the genes responsible for the α and β subunits of the phycobiliproteins. According to the literature survey done, for three phycobiliproteins, PC, has been documented to be the most exploited due to its various applicative properties. Thus, PC gene was chosen for genetic studies. The presence of PE would confirm the chromatic adapting quality of cyanobacterial strains, so an attempt to trace and isolate PE gene was also done.

5.2 Materials and Methods

Materials:

The reagents used for electrophoresis, DNA isolation and PCR were ordered from Qbiogene Research Services, France. All other reagents used were of analytical grade available from commercial sources.

Methods:

5.2.1 *DNA Isolation*:

Three methods were used for sample preparation of PCR. In two of them, whole cell extracts were prepared by exposing the whole cells to some physical disruption methods, whereas in third one modified chloroform-phenol method was followed.

Method 1:

A very small amount of lyophilized cells were taken in 100 µl of autoclaved distilled water and 5 cycles of freezing and thawing* were carried out. Every step of freezing and thawing was carried out for 1½ min alternately. The cells were then spun for 5 s and the supernatant was used as the sample for PCR.

[*Samples were freezed in liquid nitrogen (-195°C) and thawed in the water bath maintained at 65°C]

Method 2:

A small amount of lyophilized cells were taken in 100 µl of autoclaved distilled water and 5 cycles of freeze and thaw were carried out for 1½ min each alternately. The conditions were maintained as mentioned above. The cell mass was then further exposed to bead beating for 2 min, at 5000 xg. The cells were then spun for 5 s and the supernatant used as the sample for PCR.

[Glass Bead specification: Glasperlen Nr.31/41, 0.1-0.11 µm, NV.854,140/0,B.Brown Biotech, INV. GMBH, Melsugen]

Method 3:

A small amount of lyophilized cells were taken, 500 µl of TESC buffer [10 mM Tris (pH 8.0), 100 mM EDTA (pH 8.0), 1.5 M NaCl, 1% CTAB] was added and the sample

was homogenized. Five cycles of freezing and thawing were carried out for 1½ min each, alternately. Bead beating was then performed for 2 min at 5000 xg. Further, 500 µl of TESC buffer, 92 µl of 10% SDS, 5 µl of proteinase 'K' (20 mg/ml) was then added and the cells were incubated at 50°C for 20 min. Equal volume of phenol was then added and mixed properly by inverting for 10 min. Cells were centrifuged and supernatant was taken for further processing. The phenol phase was re-extracted with TESC. Both the upper aqueous layers were pooled together. 1 vol. of phenol:chloroform (1:1) was then added to it and centrifuged at 8000 xg for 10 min. Upper aqueous layer was collected again and extracted twice with 1 vol. of chloroform. Upper aqueous layer was again collected and treated with cold 0.6 vol. isopropanol and kept for 30 min at room temperature (23±1°C). for DNA precipitation. The sample was spun for 45 min at 7000 xg, pellet collected and washed with 70% cold ethanol. Finally, the pellet was vacuum dried, redissolved in 100 µl of distilled water and used as template for PCR.

5.2.2 *PCR Programs and Primers*:

Different PCR protocols and different set of primers were tried out for the isolation of genes responsible for phycobiliprotein production. The different PCR programs used are listed below:

Program Nr1: PC amplification

(93°C for 5 min) x 1 cycle
(93°C for 1 min, 55°C for 1 min, 72°C for 3 min) x 35 cycles
(72°C for 8 min) x 1 cycle

Primer set 1

Forward Primer : **5′ – GGA GAT AAG TCC ATG TTT GA–3′**
Reverse Primer : **5′–CTA GCT TAG GGC GTT GAT CGC–3′**
(Yu et al., 2002)

Program Nr.2: PC amplification

(94°C for 5 min) x 1 cycle
(94°C for 20 s, 50°C for 30 s, 72°C for 1 min) x 40 cycles
(72°C for 8 min) x 1 cycle

Primer set 2

Forward Primer : **5′–GGC TGC TTG TTT ACG CGA CA–3′**
Reverse Primer : **5′–CCA GTA CCA CCA GCA ACT AA–3′** (Neilan et al., 1995)

Program Nr.3: PE amplification

(94°C for 4 min) x 1 cycle
(94°C for 45 s, 50°C for 1 min, 72°C for 1 min) x 4 cycles
(94°C for 45 s, 55°C for 1 min, 72°C for 1 min) x 25 cycles
(72°C for 5 min) x 1 cycle

Primer set 3

Forward Primer : **5′–TA(CT) CCT AAC CGT CG(CT) (AT)T(GT) GCT GC–3′**
Reverse Primer **: 5′–GC(AG) CGT TG(AG) AT(AG) GAA CCT TGT AC–3′**
(Beard et al., 1999)

Working solution of primers:

Lyophilized primers were dissolved in autoclaved distilled water to get a final concentration of 1 μg/μl stock and working concentration of 100 ng/μl.

Reaction Volume:

Red Taq Polymerase : 2.5 μl/50 μl
Buffer : 5.0 μl/50 μl
Primer F : 2.5 μl/50 μl
Primer R : 2.5 μl/50 μl
dNTP mix : 4.0 μl/50 μl
BSA (2%) : 5.0 μl/50 μl
DNA : x
[x = aliquoted in two volumes (1 μl and 10 μl)]

The volume was made up with double distilled water.

Agarose Gel Electrophoresis

1% agarose gel was prepared in 1x TAE buffer (pH 8.0). DNA was stained with ethidium bromide which was added to the TAE buffer prior to gel solidification with the concentration of 0.5 μg/ml. It was then observed and photographed under UV light. Initially the voltage was maintained at 100 V and when the DNA was entirely out of the well, the voltage was reduced to 60 V.

[Stock solution: 50x TAE buffer : 242 g Tris, 57.1 ml glacial acetic acid, 0.5 M EDTA (100 ml), pH adjusted to 8.0. The volume is made up to 1 L.

Working solution: 1x TAE] (Sambrook et al., 1989).

5.2.3 *Gel extraction and gene sequencing*:

The bands with the appropriate size were cut with the help of sterile scalpel and further extracted by Gel Extraction kit (Qiagen). The extracted DNA was dissolved in 100 μl of double distilled water and later 10 μl was again loaded on a thin agarose gel to check the purity and concentration of the isolated DNA.

The amplified and purified DNA fragments were sent for commercial sequencing to the University of Giessen, Germany.

5.3 Results

5.3.1 *DNA Isolation*:

DNA analysis or manipulation, usually requires that DNA is isolated and purified to a certain extent. Three different methods were employed for DNA isolation, two of them involved the use of whole cell extracts, where the cells were disrupted mechanically and physically, and the extract was used as sample for PCR. The third method was a modification of the chloroform-phenol method, where DNA was isolated eliminating all

other interfering substances. *A. indica* responded well with all the three methods, even the DNA sample obtained after freezing and thawing was sufficiently good enough for further analysis. In case of *L. limnetica* and *P. tenue*, it was required to isolate pure DNA, eliminating all other interfering substances.

5.3.2 *PCR amplifications*:

PCR is the first step towards DNA analysis. Permutation and combination of different PCR programs along with different DNA concentrations in the reaction volume were employed to get the maximum yield of target DNA. Primers used in program Nr. 1 did not work with *L. limnetica* and *P.tenue*, but showed an intense band of amplified PC gene of ~1000 bp from DNA sample of *A. indica* (Fig.5.1a). Program Nr. 2 with different primers worked very well with all the three strains. It yielded a band of ~700 bp for both the marine strains (Fig.5.1b). Program Nr. 3 was employed for amplification of PE gene. No amplification was observed in *P. tenue* and *A. indica*, but a faint band was observed in *L. limnetica* of ~500 bp (Fig.5.1c). But DNA sample of a non-chromatically adapting strain run along with other samples also showed a light band.

5.3.3 *Gene Sequencing*:

Gene sequencing of the amplified PC gene fragment of *A. indica* yielded a fragment of 767 bp and that of *P. tenue* yielded a fragment of 360 bp. In case of *L. limnetica* the two sequences of PC fragment were checked for overlapping sequences and the final fragment deciphered was 610 bp.

The later part of the graphical data of the PC sequence of the marine strains, the peaks merged and so the nucleotide sequence could not be decoded. Therefore, though the gel observations showed ~700 bp fragment, only 610 bp in case of *L. limnetica* and 310 bp

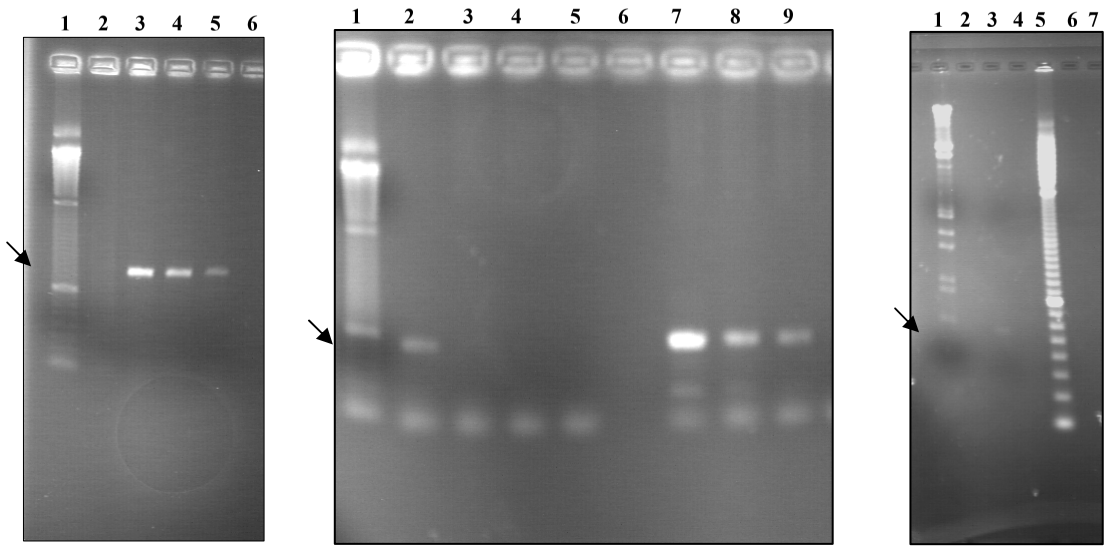

Fig 5.1 Ethidium bromide stained gel showing amplification products.
(a) Lane 1 : 100 bp ladder, Lane 3: PC fragment from *A. indica* (DNA isolation method 1) ,
Lane 4: PC fragment from *A. .indica* (DNA isolation method 2) , Lane 5: PC fragment from *A.indica* (DNA isolation method 3)
Arrow indicates 1000 bp.
(b) Lane 1: 100 bp ladder, Lane 2: PC fragment from *L. limnetica* Lane 7-9 PC fragment from *P. tenue.* Arrow indicates 700 bp.
(c) Lane 1: λ DNA marker, Lane 3: PE fragment of *L. limnetica*, Lane 5: 100 bp ladder. Arrow indicates 500 bp.

fragment in *P. tenue* could be deciphered. Similarly, in case of *A. indica* though gel visualization showed a fragment of ~1000 bp only 767 bp fragment was decoded.

There are about 70 different gene sequences of different *Arthrospira* spp. deposited and available at National Center for Biotechnology Information [NCBI] (http://www.ncbi.nlm.nih.gov/) and out of them 26 are of different parts of PB operon. 53 different gene sequences of the genus *Lyngbya* are deposited at NCBI and only one of them is of PC gene fragment of *Lyngbya aestuarii* PCC7419 (Acc. No. AJ401187). In case of *Phormidium* genus, though there are 68 different gene sequence deposits at NCBI, yet, no report of any part of PB operon was found. Sequence alignment of PC gene sequence obtained from all the three studied cyanobacterial strains was carried out with the known sequence of cyanobacteria which have been deposited at NCBI. The comparative analysis was done using the program **Blastn** available on the site http://www.ncbi.nlm.nih.gov/. Blast results also show that the sequence of PC fragment of *A. indica* shares a high similarity % (95-97%) with many different members of *Arthrospira* as well as *Spirulina* genera (Table 5.1).

Table 5.1 Similarity % of *A. indica* as compared with known sequences.

Organism	Similarity %
Arthrospira platensis FACHBOUQDS6 (Accession No. AY244671.1)	97
Spirulina subsulsa FACHB351 (Accession No. AY244667.1)	97
Arthrospira maxima OUQDSM (Accession No. AY244672.1)	97
Arthrospira platensis FACHB439 (Accession No. AY244669.1)	95
Spirulina maxima (Accession No. AF441177.1)	95

```
AI  32  gcgaaatgctgagtacagccaaatcgatgctctgagccaaatggttgctgaaagcaaca 91
        |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
AP  22  gcgaaatgctgagtacagctcaaatcgatgctctgagccaaatggttgctgaaagcaaca 81

AI 92  aacgtttggattctgttaaccgcattaccagcaacgcttccaccattgtttccaacgctg 151
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 82  aacgtttggattctgttaaccgcattaccagcaacgcttccaccattgtttccaacgctg 141

AI 152 ctcgttctttgttcgcagagcaaccccaactgattgctcccggtggaaacgcctacacca 211
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 142 ctcgttctttgttcgcagagcaaccccaactgattgctcccggtggaaacgcctacacca 201

AI 212 gccgtcgtatggctgcttgcttgcgtgacatggaaatcatcctgcgctatgtaacctacg 271
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 202 gccgtcgtatggctgcttgcttgcgtgacatggaaatcatcctgcgctatgtaacctacg 261

AI 272 ctgtgtttgctggcgatgcaagtgttctcgaagatcgttgcttgaacggtttgcgtgaaa 331
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 262 ctgtgtttgctggcgatgcaagtgttctcgaagatcgttgcttgaacggtttgcgtgaaa 321

AI 332 cttacctggctttgggaactcccggttcctccgttgctgtcggtgttggcaaaatgaaag 391
       |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
AP 322 cttacctggctttgggaactcccggttcttccgttgctgtcggtgttggcaaaatgaaag 381

AI 392 aagctgctctggcgatcgttaacgatcccgcaggtatcactcctggtgattgtagcgctt 451
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 382 aagctgctctggcgatcgttaacgatcccgcaggtatcactcctggtgattgtagcgctt 441

AI 452 tggcttcagaaatcgctggttacttggaccgtgctgctgctgcagtttcctaattaagca 511
       |||||||||||||||||||||||| |||||||||||||| |||||||||||||| |||||
AP 442 tggcttcagaaatcgctggttactttgaccgtgctgctgcagcagtttcctaatcaagca 501

AI 512 gctccatagcatatcacaattcaaacagtttagctgaagtctaagttatttgacttctct 571
       | |||||||||||| |||||| |||||||||||||||||||||||| | | ||||||| |
AP 502 gatccatagcatataacaattgaaacagtttagctgaagtctaagtgactggacttctgt 561

AI 572 tggttacctaatttttggtaaaccaatcgggagataactcgagaatgaaaaccccccctaa 631
       | |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
AP 562 ttgttacctaatttttttgtaaaccaatcgggagataactcgagaatgaaaaccccccctaa 621

AI 632 ccgaagcagtttctatcgctgattcccaaggtcgtttcctaagcagcaccgaaatccaag 691
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AP 622 ccgaagcagtttctatcgctgattcccaaggtcgtttcctaagcagcaccgaaatccaag 681

AI 692 tggctttgggccgttttcgtcaagccaaagctggtctggaagctgctaaagcttggacct 751
       | ||||| ||||||||||||||||||||||||||||||||||||||||||||| |||||
AP 682 tagctttggccgttttcgtcaagccaaagctggtctggaagctgctaaagcttgacct 741

AI 752 ccaaatctgatagtct 767
       | ||| ||||||||||
AP 742 ctaaagctgatagtct 757
```

Fig 5.2 Blast results of Sequence alignment of PC fragment of *A. indica* (767 bp) with *A. platensis* FACHBOUQDS6 (857 bp). Organism names are indicated in two letter code, AI - *A. indica*, AP- *A. platensis* Identities = 717/736 (97%) The dissimilarities in the sequence are indicated in red.

```
1   tggcgatccagtgtctccacgatcttgcttgaacggcctccgtgaaacctacgttgctct 60
61  tggcactcctggtgcttccgtagcagctggtgttgagaagatgaaagaagctgcagttgc 120
121 aatcgttaacgatcctgcaaacatcactcctgagtgattgcagctctctcgtttccgaga 180
181 tcggcagctactttgatctcgctgctgcagctgttgcttagttttaggcagcttctcagt 240
241 gcgttttacgcacacacactctctaacaaccaaggaccaacgtttgtacattaattagga 300
301 gatatttgcgatgaaaactcctctgactgaagccagtagctgctgctgatactcaaggtc 360
361 gtttcctcaggcaccagtgaaatgcaggccgctattggtcgatctgcgtcaggctcagac 420
421 tgggtgtagaagcatggcgaaggcgttggacttcccaagtctgactccttgatgctggtg 480
481 ctgctcaagctgtagacaacaagtgccccgtacaccacgtcaaatgcaggagcctaacta 540
541 cgcatctaccgaagtcggcaagcaaaagtgtgcgcgtgacatcggttactacttgcgcat 600
601 ggatacctac 610
```

Fig 5.3 Gene sequence of PC fragment of *L. limnetica* (610 bp).

```
  1 tgcctagtatacangaaggctggatttagcactttatcgagatctnaggggcgcgttact 60
 61 catcccgattggtacactgaaactcatagctaatacgacctacgaggttgagaagcagga 120
121 aatcctatggtgtgaaccaggcaatctccctgtcatacctcgataggaggttgctgagga 180
181 gcaggatgtcaactacgcctggacttgctcgatatcagagtgtcgcttttgactgtacga 240
241 gacccaacggttaggtaactgacttagagcgcgctatgacgcgcccgtcgaacgctcgag 300
301 tcgcaggtgactaggaatacgagcgatgaagatcgccatatgcgtagggagctccggtcg 360
```

Fig 5.4 Gene sequence of PC fragment of *P. tenue* (360 bp).

```
Ly 25   ttgcttgaacggcctccgtgaaacctacgttgctcttggcactcctggtgcttccgtagc 84
        ||||||||||||  | ||||||||| ||| | ||| | || ||||| ||| | ||||| ||
AI 309  ttgcttgaacggtttgcgtgaaacttacctggctttgggaactcccggttcctccgttgc 368


Ly 85   agctggtgttgagaagatgaaagaagctgcagttgcaatcgttaacgatcctgcaaacat 144
         |  |||||||  || ||||||||||||||  | || |||||||||||||| |||   ||
AI 369  tgtcggtgttggcaaaatgaaagaagctgctctggcgatcgttaacgatcccgcaggtat 428


LY 145  cactcctgagtgattgcagctctctcgtttccgagatcggcagctactttgatctcgctg 204
        |||||||| ||||||| ||| || | | ||| || ||||   | ||||| || |  ||||
AI 429  cactcctg-gtgattgtagcgctttggcttcagaaatcgctggttacttggaccgtgctg 487


LY 205  ctgcagctgtt 215
        |||| || |||
AI 488  ctgctgcagtt 498
```

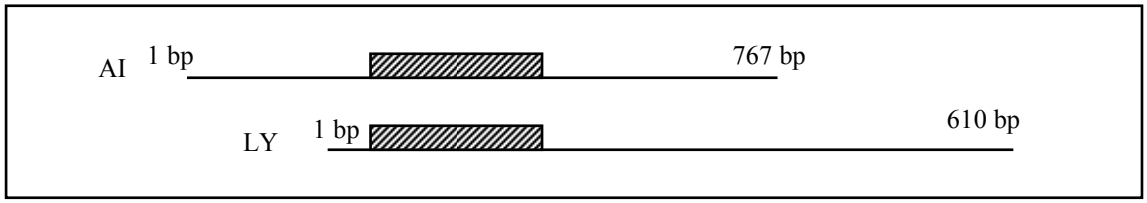

Fig. 5.5 Sequence alignment: PC fragment of *L. limnetica* with *A. indica*. Organism names are indicated in two letter codes LY-*L. limnetica* AI- *A. indica*. Identities 146/191 (76%). The dissimilarities in the sequence are indicated in red.

Fig.5.2 shows the pairwise sequence alignment of *Arthrospira indica* with *Arthrospira platensis* FACHBOUQDS6. In case of *Lyngbya limnetica* and *Phormidium tenue*, since there were no adequate depositions for comparison, Blastn yielded no significant results. The sequence of the PC fragment of *L. limnetica* and *P. tenue* is as shown in Fig.5.3 and Fig.5.4, respectively.

The sequence of PC fragment of all the selected strains were aligned with each other to check the similarity % among them. The program used here was **b12seq**, which is available on the site http://www.ncbi.nlm.nih.gov/BLAST/. It showed no significant

similarity between fragment of *A. indica* and *P. tenue*, *L. limnetica* and *P. tenue*. The results also indicated that, the initial part of the *L. limnetica* seq (25-215 bp) matched with the central part of *A. indica* (309- 498 bp) sequence as shown in Fig.5.5.

5.4 Discussion

DNA encodes all the information needed to specify the structure of every protein that the cell can produce. The structure in turn regulates the function of the protein. Isolation and characterization of PB genes was an important part of the attempt towards studying the regulation of PB production.

Of the three different methods employed for DNA isolation, the chloroform-phenol method seems to work for all the three selected strains. But whole cell PCR worked well with *A. indica*, where in the cells were exposed only to a physical treatment of freeze and thaw. This method did not work well with the two marine strains, which indicates that there may be some structural differences in the cell wall of fresh water and marine strains. Hypothetically it also can be said that BSA added to the reaction mix is more effective in case of *A. indica* than the marine strains. Different programs had to be employed for amplifying the PC gene. The primers which worked for *A. indica* did not yield any positive amplification results with *L. limnetica* and *P. tenue*. This highlights the possible variation in the genetic sequence of PB operon in different strains. Also the primers used in Program Nr.1 were very specific for *Spirulina* genus and since the two are closely related, worked well for *A. indica*, but not for *L. limnetica* and *P. tenue*. This hints towards the differences in PC gene even within a family. Program Nr.2 yielded results in all the three strains. These primers may not be very specific towards a particular genera. The positive amplification result in all the three strains supports the claims of

Neilan et al., (1995) of the presence of completely conserved regions within the PB operon. Primers of program Nr.2 amplifies part of the functional subunits flanking the IGS (Intergenic spacer region) thus, yielded a smaller fragment of ~700 bp. Primers used for PE amplification may not be very specific, as they showed amplification product in a non-chromatically adapting cyanobacteria, which does not possess PE.

Gene sequencing yielded a sequence of 767 bp long fragment of PC gene in *A. indica*, 610 bp fragment in *L. limnetica* and 360 bp fragment in *P. tenue*. The DNA sequence reveals the extent and locations of PC polymorphisms. There is a striking similarity between the gene sequence of PC from *A. indica*, and other cyanobacteria from the same genera. A very high % identity (95-97%) was observed when the sequence was compared with that of *A. platensis* FACHBOUQOS6, consequently exhibiting closeness between the two organisms.

Though there are many sequences of *Phormidium* and *Lyngbya* genera deposited at NCBI, yet not many reports of PC gene sequence are documented. Lack of availability of specific primers for the PB operon could be one of the reason for the limited information available in case of these strains. NCBI search also reveals fewer sequence depositions in case of the genera *Lyngbya* and *Phormidium* as compared to the genus *Arthrospira*, which could be due to the problems faced in isolation and maintainance of these thin filamentous marine cyanobacterial strains.

Another fact which is striking is that though the amplification products in case of *L. limnetica* and *P. tenue* were obtained by the same set of primers, yet no similarity is observed among the PC gene sequences of these two. Therefore, possibly the sequences would be very different in the two selected marine oscillatorian strains. This indicates

that even if phenotypical characters exhibit high % of identity, it may not be complemented genotypically. 76% identity was observed between *A. indica* and *L. limnetica.* Sequence alignment among the three showed a 76% identity between the freshwater strain, *A. indica* and the marine strain, *L. limnetica*. This result stands in favour of the belief that all PB genes arose from the duplication a single ancestral gene (Troxler et al., 1981).

.