CHARACTERIZATION OF BIOCHEMICAL AND PHYSIOLOGICAL ROLE OF LexA IN THE NITROGEN-FIXING CYANOBACTERIUM, *Anabaena* sp. STRAIN PCC7120

By

ARVIND KUMAR

LIFE01201404002

Bhabha Atomic Research Centre, Mumbai, India

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

October, 2018

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Arvind Kumar entitled "Characterisation of biochemical and physiological role of LexA in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Vinay Kumar	Chairman		
2.	Dr. (Smt.) Hema Rajaram	Guide/ Convener		
3.		Co-guide (if any)		
4.	Dr. Prasun Mukherjee	Member		
5.	Dr. Satyendra Gautam	Member		
6.	Dr. Anu Ghosh	Member		

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date:

Place: **Guide's Signature**

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Arvind Kumar

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Arvind Kumar

List of Publications arising from the thesis

Journal

1. "LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits *in vitro* pH-dependent and RecA-independent autoproteolytic activity", **Arvind Kumar**, Anurag Kirti and Hema Rajaram, *The Int. J Biochem. Cell Biol.*, **2015**, *59*: 84-93.

2. **"**Differential regulation of *ssb* genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120" Anurag Kirti, **Arvind Kumar** and Hema Rajaram, *J. Phycol.*, **2017,** *53*: 322-332.

3. "The SbcC and SbcD homologs of the cyanobacterium Anabaena sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair", Sarita Pandey, Anurag Kirti, **Arvind Kumar** and Hema Rajaram *Funct Integr Genomics.,* **2018,** *18*:357-367.

4. "Regulation of multiple abiotic stress tolerance by LexA in the cyanobacterium *Anabaena* sp. strain PCC7120", **Arvind Kumar**, Anurag Kirti and Hema Rajaram, *BBA Gene Regulatory Mechanisms.,* **2018,** Doi: 10.1016/j.bbagrm.2018.07.007

Conferences

1. **Arvind Kumar** and Hema Rajaram **(2014)** Biochemical characterization of LexA of N_2-fixing cyanobacterium *Anabaena* sp. strain PCC7120; Symposium on biotechnology and stress biology of Algae and cyanobacteria; OP-18 Pg. No70; held at BHU, Varanasi from February 24th to 26th, 2014.

2. **Arvind Kumar**, Hema Rajaram **(2015)** *In vitro* analysis of autoproteolytic activity of LexA of *Anabaena* PCC 7120; DAE-BRNS Life Sciences Symposium-2015 on "Advances in Microbiology of Food, Agriculture, Health and Environment"; held at Nabhikiya Urja Bhavan Auditorium, Anushaktinagar, BARC, Mumbai (Feb 3-5, 2015) (Poster Presentation), P51: page No. 143

Others

1. Gene sequences submitted: Kumar, A., Rajaram, H. and Apte, S. K. (2013) *Anabaena* sp. strain PCC7120 *lexA* gene complete sequence. "LexA regulon of *Anabaena* PCC7120." GenBank Accession No. KF269537.

Arvind Kumar

Dedicated to

Almighty Lord Shiva,

My family and Friends

ACKNOWLEDGEMENTS

On the outset, I would like to express my sincere and deepest gratitude to my guide Dr. Hema Rajaram for the continuous encouragement of my Ph.D study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I am also very grateful to Dr. S. K. Apte, Dr. S. K. Nayak and Dr. H. S. Misra for their support during this course of study.

Further, I would like to thank my doctoral committee members: Dr. Vinay Kumar (chairman), Dr. Prasun Mukherjee, Dr. S. K. Gautam and Dr. Anu Ghosh, for their insightful comments and encouragement, so also for the intensely intriguing questions which challenged me to widen my research from various perspectives.

My sincere thanks to Dr. Bhakti Basu for helping me in my proteomic studies and Dr. Gagan Deep Gupta for biophysical studies, and Mr. Nishad Srambikkal for PD Quest analysis. Their expertise has helped me overcome many a challenges.

It is also my pleasure to acknowledge and thank my fellow labmates and friends, Dr. Prashanth Raghavan, Dr. Anurag Kirti, Dr. Sarita Pandey, Mr. Akhilesh Potnis and Mr. Nilesh Kolhe who provided a fun-filled, positive and light environment to carry out my research work. Special thanks are due for Mr. Avindra Kadam for his prompt technical support. I would also like to thank all MBD colleagues who have directly or indirectly helped me during the research.

Finally, I must express my very profound gratitude to my caring parents and to my supportive wife for their unfailing cooperation and continuous encouragement throughout my years of study and the process of researching and writing this thesis.

Above all, Almighty Lord Shiva, who is always there by my side and showering His blessings on me.

Arvind Kumar

CONTENTS

Synopsis

Homi Bhabha National Institute

SYNOPSIS OF Ph. D. THESIS

1. **Name of the Student: Arvind Kumar**
2. **Name of the Constituent Institution: BARC**
3. **Enrolment No. : LIFE01201404002**
4. **Title of the Thesis: Characterization of biochemical and physiological role of LexA in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120**
5. **Board of Studies: Life Sciences**

SYNOPSIS

(Limited to 10 pages in double spacing)

Cyanobacteria are considered as most primitive, Gram-negative photosynthetic bacteria which originated about 3.5 billion years ago [1]. They have been exposed to several extreme conditions like heat, radiation, drought etc., and have shown the ability to adapt themselves to them [2]. The filamentous, heterocystous, nitrogen-fixing cyanobacterium, *Anabaena* sp. PCC7120 exhibits high resistance to both γ-radiation and desiccation stresses [3, 4]. The ability to stitch back damaged DNA during recovery from these stresses is suggestive of an efficient DNA repair system [4]. However, the proteins and mechanism involved in DNA repair has not been investigated enough to pinpoint the possible repair mechanisms in cyanobacteria. Recent work in our laboratory on

truncated single stranded DNA binding (SSB) proteins of *Anabaena* 7120 showed that over-expression of Alr0088, one of the two truncated SSB proteins, enhanced tolerance to UV and mitomycin C stresses, but decreased that to γ-radiation stress suggesting a possible role for this protein in DNA damage and repair [5]. The upstream region of *alr0088* was shown to have a LexA Box-like sequence [6] suggesting possible regulation by LexA.

LexA is one of the key regulators of DNA repair genes in bacteria and is known to regulate more than 30 genes involved in SOS response in *E. coli* [7]. Under normal growth conditions, LexA represses the SOS genes by binding to the LexA Box [8], present in the vicinity of their promoter region. Upon DNA damage, RecA interacts with single stranded DNA, gets activated and induces auto-cleavage of the Gly^{84}-Ala^{85} peptide bond in LexA resulting in derepression of expression of the regulated DNA repair genes [9].

The physiological role for *lexA* has been investigated in the unicellular cyanobacterium, *Synechocystis* PCC6803, where it was found to be involved in modulating expression of several genes involved in carbon metabolism, but not SOS-response genes [10]. RNA-Seq analysis of *lexA* mutant revealed altered expression of genes belonging to several functional categories, namely photosynthesis, hydrogenases, motility but none of the DNA metabolism genes were altered upon disruption of *lexA* gene in *Synechocystis* [11]. Computational analysis of the 33 sequenced cyanobacterial genomes for the presence of LexA proteins revealed that 6 do not possess LexA and while among the remaining, LexA proteins of *Synechocystis* PCC6803 and *Gleobacter violaceus* PCC7421 are non-cleavable [12], which was based on sequence information [10]. Unlike the reported transcriptional regulation of *lexA* in bacteria, that of *Synechocystis* PCC 6803 was found to occur both at post-transcriptional and post-translational levels [13].

In *Anabaena* 7120, LexA is coded by *alr4908*, and has been shown to bind LexA Box-like sequences upstream of *lexA* and *recA* genes of *Anabaena* 7120 [6]. It was also shown to regulate a few genes involved in C-metabolism, such as hydrogenases [14]. Based on the few proven regulated genes, several sequences for the proposed LexA-box in *Anabaena* were predicted [6]. Bioinformatic analysis using the sequence predicted by Mazon et al, revealed possible regulation of about 216 genes by LexA in *Anabaena* [14]. However, demonstration of the cleavage of LexA and the amino acids involved therein, variation in the LexA-boxes and the significance of the possible regulation of several genes by LexA in *Anabaena* were not answered. Hence, it was proposed to investigate the biochemical and physiological role of LexA in *Anabaena* 7120 to obtain an insight into the LexA protein and its function as possible a global regulator.

OBJECTIVES:

- Cloning of *lexA* gene from *Anabaena* 7120 into *E. coli* for over expression of protein: Purification of LexA protein for determination of oligomeric status, biophysical studies, assessment of auto-proteolytic activity in terms of RecA and pH-dependence, site-specific mutants to analyse the importance of specific residues in auto-cleavage, assessment of binding to LexA box and generation of polyclonal antibodies
- Expression analysis of *lexA* during stress at transcript and protein level.
- Generation of LexA overexpressing/*lexA* mutant strains of *Anabaena*, and if viable, assess their role in regulating stress tolerance.
- Proteomic approach to identify proteins regulated by LexA and analyse the regulation of some of the genes.

ORGANIZATION OF THESIS:

The thesis is divided into five chapters which include introduction (Chapter 1), materials and methods (Chapter 2) and the two chapters; Chapter 3 and Chapter 4 describing the findings in terms of results with relevant discussion. A summary (Chapter 5) of the major findings of the whole thesis work is given at the end followed by the references cited in the order of citation.

Chapter 1: Introduction

The first chapter provides an overview on DNA damage and repair, with greater emphasis on SOS response. The mechanism of regulation of SOS response by LexA has also been discussed. Structural aspects of LexA protein has been discussed with a view on its autoproteolytic cleavage mechanism. A brief introduction about cyanobacteria, and *Anabaena* 7120, in particular, covering the above aspects and a detailed review on the findings about the role and regulation of LexA in cyanobacteria has also been included. The chapter concludes with the scope and outline of the thesis work on LexA in *Anabaena* 7120.

Chapter 2: Materials and Methods

This chapter provides the details of the experimental procedures and techniques used in the present study. The different reagents, plasmids and strains used and constructed in the current study have been tabulated. In brief, the procedures explained are those for isolation of plasmid and chromosomal DNA, RNA and proteins. Procedures followed for cloning of DNA fragments and generation of recombinant bacterial strains, which includes PCR amplification of DNA fragments, restriction digestion, ligation and transformation and conjugation have been described in detail. Anlaysis of transcription through including Reverse- Transcriptase (RT)-PCR has also

been detailed. The protocols followed for overexpression and purification of proteins of interest, separation by SDS-PAGE and visualization as well as Western blot analysis have also been described. Proteome analysis by two-dimensional gel electrophoresis (2-D) and the identification of protein spots have been described in detail. The details of different biophysical and biochemical methods like gel filtration, circular dichroism (CD), dynamic light scattering (DLS), autoproteolytic activity of *Anabaena* LexA protein, DNA binding assay (EMSA) and promoter activity analysis using *gfpmutII* gene as reporter have also been included.

Chapter 3: Biophysical and biochemical characterization of *Anabaena* LexA protein

The LexA protein of *Anabaena* 7120, encoded by *alr4908*, is a 201 amino acid long and ~23 kDa protein exhibiting 54 % similarity with *E. coli* LexA protein. The ORF of *lexA* gene was PCR amplified and cloned in pET16b expression vector. The protein was overexpressed in *E. coli* BL21 (p*LysS*) (DE3) cells by inducing with 1 mM IPTG. The overexpressed LexA protein was purified using Ni-NTA chromatography under native conditions, and found to exist as a dimer of ~55 kDa with hydrodynamic diameter of 9.4 ± 3.4 nm as determined by DLS. CD analysis revealed that *Anabaena* LexA comprised of 20% β-sheet and 28% α-helix secondary structure elements. *Anabaena* LexA exhibited autoproteolytic cleavage under alkaline pH (i.e. pH 8.0 and above) conditions, but not in the presence of activated RecA proteins from *Anabaena* and *E. coli*. The extent of cleavage was monitored as a function of pH and temperature. Mutational analysis of *Anabaena* LexA protein revealed that the cleavage occurs at the peptide bond between Ala^{84} and Gly^{85} and requires the presence of active site residues Ser^{118} and Lys^{159} and a three amino acid region $^{86}GLI^{88}$ in the cleavage site region (CSR) for the cleavage. The GLI region was found to be conserved among all cyanobacterial LexA which are predicted to

encode the cleavable form of LexA. Of the two cleaved products, the 13 kDa product corresponding to the C-terminal part was found to be stable, while the 12 kDa N-terminal part of the protein was not detected, possibly due to susceptibility to proteolytic degradation. The expression analysis of *lexA* monitored under different DNA damaging conditions revealed induction at transcript level, but no change in protein levels or any cleavage when exposed to UV-radiation or mitomycin C stress. While upon exposure to desiccation as well as γ-irradiation stresses individually, expression of *lexA* was found to be repressed both at the transcript as well as the protein level. Regulation of *lexA* in response to stresses inducing DNA damage suggested a possible role for *Anabaena* LexA in the process of DNA repair or SOS response.

Chapter 4: Physiological role of LexA in *Anabaena*

In order to understand the physiological role of LexA in *Anabaena*, attempts were made to modulate the intrinsic expression of LexA through either insertional mutation or by constitutively overexpressing the protein. The *lexA* knockout strain of *Anabaena* was only partially segregated and unstable and hence, could not be used for further studies. Recombinant *Anabaena* strain overexpressing the LexA protein, *AnlexA*$^{+}$ was generated, wherein the *lexA* gene is transcriptionally fused with *gfpmutII* and both the genes expressed from *psbA1* promoter. An*lexA*$^{+}$ strain exhibited green fluorescence and high levels of the LexA protein. An*lexA*$^{+}$ cells exhibited growth similar to the vector control cells, AnpAM (*Anabaena* cells harbouring the plasmid pAM1956, used as the vector for overexpressing LexA) [15] under control conditions. When exposed to photosynthetic inhibitor, DCMU [3-(3, 4-dicholophyllphenyl)-1,1-dimethylurea)], it exhibited better growth than AnpAM cells. Among the DNA damage inducing stresses, exposure to mitomycin C did not result in any significant impact on growth or survival

upon overexpression of LexA, while An*lexA*$^+$ cells exhibited lower growth, survival and recovery compared to AnpAM cells when exposed to either UV or γ-radiation or desiccation stresses. An*lexA*$^+$ cells also exhibited lower tolerance to different forms of oxidative stresses, applied with H_2O_2, Methyl viologen (MV) or heavy metals [Cd(II) and As(V)]. The decreased oxidative stress tolerance was concomitant with decreased levels of KatB protein, MnSOD activity and increased accumulation of reactive oxygen species (ROS) in An*lexA*$^+$ cells. This suggested that LexA may be negatively regulating genes involved in DNA repair and oxidative stress alleviation. To investigate this, a combination of three approaches, namely 2-D proteomic study, transcript analysis and promoter studies, targeting specific genes was carried out.

Proteomic analysis of An*lexA*$^+$ and AnpAM cells revealed differential abundance for 47 protein spots, of which 30 showed decreased abundance and 17 with increased abundance in An*lexA*$^+$ cells. Of these, 23 protein spots could be identified conclusively, of which 15 were those with the decreased abundance and 8 with increased abundance in An*lexA*$^+$. One major groups of identified proteins belonged to the functional category of C-metabolism, which included those involved in light reaction, C-fixation and C-breakdown. Among the proteins constituting the phycobilisome, abundance of Cpc (phycocyanin) proteins was lower, that of Pec (phycoerythrins) same, while one of the allophycocyanins, ApcE showed higher abundance. The overall photosynthetic efficiency of *Anabaena*, however, remained unchanged upon overexpression of LexA. The increase in abundance of RuBisCo (RbcL) protein along with the decrease in that of TalB, GlgB and GalE would result in maintenance of higher intracellular C-levels, which could account for higher growth of these cells when exposed to DCMU. Among the proteins involved in oxidative stress alleviation, PetH and PrxA showed lower abundance, while that of AhpC was higher. Additionally, the abundance of DnaK, Ndk and GlnA was also

found to be lower in An*lexA*$^{+}$ cells. Other than Ndk, involved in DNA metabolism, no proteins involved in DNA repair could be detected in the proteomic analysis. In all the 23 identified protein spots corresponded to 11 different genes. Of the 11 genes, *apcE*, *rbcL* and *ahpC* genes were found to be transcriptionally upregulated by LexA, while the others were downregulated. The mode of regulation of these genes along with 6 DNA repair genes by LexA was investigated by EMSA studies to evaluate binding of LexA to their upstream regions and promoter studies to evaluate the regulation. LexA was found to bind to upstream regions of all the genes tested, with binding affinity in the range of 18 – 96 nM, with the exception of *ssb3*, which has been earlier shown to be regulated by FurA and not LexA [16]. Analysis of the activity of the corresponding promoters using *gfpmutII* as the reporter gene in *E. coli* cells overexpressing *Anabaena* LexA, confirmed direct regulation of the genes by LexA. A good agreement was found between proteomic, transcript and promoter studies. Thus, it can be concluded that these genes, belonging to different functional categories, are directly up- or down-regulated by LexA. Bioinformatic analysis of the tested promoter regions revealed the presence of the palindromic sequence AGT-$N_{4\text{-}11}$-ACT in 10 of the 17 gene promoters, and with one base variation in the remaining 7 genes. Mutational analysis confirmed the requirement of the palindrome for binding of LexA, while shortening of the space length from 8 to 4 did not affect the binding of LexA. The sequence was designated as the AnLexA box in *Anabaena* 7120. A short genome search revealed the presence of this element, in the vicinity of an AT-rich region, upstream of at least 40 genes in *Anabaena,* belonging to varied functional groups.

Chapter 5: Summary

The LexA protein of the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120, is encoded by ORF *alr4908. Anabaena* LexA existed as a dimer under native conditions, and exhibited pH-dependent and RecA-independent autoproteolytic cleavage at the peptide bond between Ala^{84} and Gly^{85}. The other amino acids found to be essential for cleavage were Ser^{118}, Lys^{159} and $^{86}GLI^{88}$ by mutational analysis. The absence of RecA-dependent cleavage at physiological pH may possibly be due to the absence of RecA interacting sites on *Anabaena* LexA protein, as well as low levels of RecA in *Anabaena*. Of all the bacterial LexA proteins characterised till date, that of *Anabaena* is the only one exhibiting no cleavage in the presence of activated RecA. This was reflected in the absence of cleavage even under *in vivo* conditions, when *Anabaena* cells were exposed to UV-radiation and mitomycin C stresses, unlike other bacteria.

The paradigm that LexA functioned only as a SOS response regulator in bacteria was challenged for the first time in the unicellular cyanobacterium, *Synechocystis* PCC6803, wherein it was shown to be involved in the regulation of C-metabolism genes [10], and thereafter even in the filamentous cyanobacterium, *Anabaena* PCC7120. Our studies revealed that overexpression of LexA helped *Anabaena* tolerate C-depletion through upregulation of carbon-fixation through RbcL and down-regulation of certain C-catabolism genes, TalB, GlgB, and GalE while maintaining the photosynthetic efficiency. The decreased tolerance to DNA damaging stresses as well as oxidative stress was found to be due to the down regulation of genes involved in DNA repair and oxidative stress alleviation. A consensus AnLexA box, having the sequence AGT-N_{4-11}-ACT, with a single base mismatch in a 4-base palindrome, was found to be present upstream of the regulated genes, and was designated as AnLexA box. The AnLexA box was found to be

present upstream of 6 genes of C-metabolism, 14 genes of DNA repair, 8 genes of oxidative stress alleviation, 4 genes of fatty acid biosynthesis, 2 transcriptional regulator genes, 2 metal responsive genes and 3 toxin-antitoxin genes through a short genome analysis, suggesting a global regulatory role for LexA in *Anabaena*, and possibly in cyanobacteria, in general. It is possible that during the course of evolution from the ancient cyanobacteria, which had to face extreme environmental conditions, to the modern day bacteria, the role of LexA became more narrow and restricted to that of SOS response regulator.

References:

1. Brock, T.D. **(1973)** pp. 487-500, **Algae**. Berkeley: Univ. California Press
2. Fay, P. **(1992) Microbiol. Rev.** 56, 340-373
3. Singh H, Fernandes T, Apte SK **(2010) J. Biosci**. 35: 427-434.
4. Singh H., Anurag K., Apte SK **(2013) Photosyn. Res.** 118: 71-81
5. Kirti A., Rajaram, H., Apte S.K. **(2013) Photosyn. Res**. 118: 147-154.
6. Mazon G., Lucena J.M., et al. **(2004) Mol. Gen. Genomics** 271: 40-49
7. Fernandez de Henestrosa AR, Ogi T, et al. (**2000) Mol Microbiol**. 35:1560-1572
8. Walker, G.C. **(1984) Microbiol. Rev.** 48: 60-93.
9. Janion C (**2008) Int J Biol Sci**. 4:338-344.
10. Domain F, Houot L, Chauvat F, Cassier-Chauvat C (**2004) Mol Microbiol**. 53:65-80
11. A. Kizawa, A. Kawahara, et al. **(2016) Front Microbiol**. 7: 193.
12. Li S., Xu M., Su Z. **(2010) BMC Genomics** 11: 527
13. P. Oliveira, P. Lindblad **(2011) J. Bacteriol.** 193: 3804-3814.
14. J. Sjöholm, P. Oliveira, P. Lindblad, **(2007) Appl. Environ. Microbiol**. 73:

5435-5446.

15. H. Rajaram, S.K. Apte, **(2010) Arch. Microbiol**. 192: 729-738.

16. A. Kirti, A. Kumar, H. Rajaram, **(2017) J. Phycol.** 53: 322-332

Publications in Refereed Journal:

a. Published

- **Arvind Kumar**, Anurag Kirti and Hema Rajaram **(2015)** LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits *in vitro* pH-dependent and RecA-independent autoproteolytic activity. **The Int. J Biochem. Cell Biol**. 59: 84-93
- Anurag Kirti, **Arvind Kumar** and Hema Rajaram **(2017)** Differential regulation of *ssb* genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120. **J. Phycol.** 53: 322-332
- Sarita Pandey, Anurag Kirti, **Arvind Kumar** and Hema Rajaram **(2018)** The SbcC and SbcD homologs of the cyanobacterium Anabaena sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair. **Funct Integr Genomics.** 18:357-367.
- **Arvind Kumar**, Anurag Kirti and Hema Rajaram, **(2018)** Regulation of multiple abiotic stress tolerance by LexA in the cyanobacterium *Anabaena* sp. strain PCC7120.**BBA Gene Regulatory Mechanisms.** Doi: 10.1016/j.bbagrm.2018.07.007

Other Publications:

a. Book/Book Chapter

b. Conference/Symposium

- **Arvind Kumar** and Hema Rajaram **(2014)** Biochemical characterization of LexA of N_2-fixing cyanobacterium *Anabaena* sp. strain PCC7120; Symposium on biotechnology and stress biology of Algae and cyanobacteria; OP-18 Pg. No70; held at BHU, Varanasi from February 24th to 26th, 2014.
- **Arvind Kumar**, Hema Rajaram **(2015)** *In vitro* analysis of autoproteolytic activity of LexA of *Anabaena* PCC 7120; DAE-BRNS Life Sciences Symposium-2015 on "Advances in Microbiology of Food, Agriculture, Health and Environment"; held at Nabhikiya Urja Bhavan Auditorium, Anushaktinagar, BARC, Mumbai (Feb 3-5, 2015) (Poster Presentation), P51: page No. 143

Signature of Student:

Date:

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1.	Dr. Vinay Kumar	Chairman		
2.	Dr. (Smt.) Hema Rajaram	Guide/ Convener		
3.		Co-guide (if any)		
4.	Dr. Prasun Mukherjee	Member		
5.	Dr. Satyendra Gautam	Member		
6.	Dr. Anu Ghosh	Member		

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACN	Acetonitrile
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BPB	Bromo phenol blue
Cb	Carbenicillin
CBB	Coomassie Brilliant Blue
CDD	Conserved Domain Database
CD	Circular Dichroism
Cm	Chloramphenicol
CSR	Cleavage Site Region
CTD	C-terminal domain
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea
dsDNA	Double stranded DNA
dsb	Double strand DNA Break
DTT	Dithiothreitol
DLS	Dynamic Light Scattering
EDTA	Ethylene Diamine Tetraacetic Acid
EMSA	Electrophoretic Mobility Shift Assay
IAA	Iodoacetamide
IEF	Isoelectric focussing
IPTG	Isopropyl-beta-D-thiogalactopyranoside
Kan	Kanamycin
kb	Kilo Base
kDa	Kilo Dalton
MALDI-ToF	Matrix Assisted Laser Desorption/Ionization-Time of Flight
MV	Methyl Viologen
NBT/BCIP	Nitro Blue Tetrazolium chloride/5-Bromo-4-chloro-3-indolyl phosphate
Ni-NTA	Nickel-Nitrilo triacetic acid

Neo	Neomycin
NTD	N-terminal domain
PAGE	Poly Acrylamide Gel Electrophoresis
PCC	Pasture Culture Collection
PCR	Polymerase Chain Reaction
PMSF	Phenyl Methyl Sulfonyl Fluoruide
SDS	Sodium dodecyl sulphate
SSB	Single stranded DNA Binding protein
ssb	Single strand DNA Break
ssDNA	Single stranded DNA
µM	Micromolar
nM	Nanomolar

CHAPTER 1

INTRODUCTION

LexA is an important negative regulator of expression of DNA repair proteins in bacteria, coordinating the response of cells to DNA damage.

1.1 DNA DAMAGE

DNA damage can be defined in numerous ways, depending on whether it is translated into a phenotypic or a genotypic variation. Chakarov defined DNA damage in simple terms as "*Any change in the DNA structure of physical or chemical nature, resulting in altered DNA molecule, different to original with respect to its chemical, physical, structural or functional properties*" (*1*)

1.1.1 *Factors causing DNA damage*

DNA damage can be caused by (i) endogenous factors, referring to several potentially toxic agents produced during normal cell metabolism, or (ii) exogenous factors i.e. due to external conditions. The endogenous factors, due to their intrinsic nature, result in a higher frequency of DNA damage events, and hence, have a more potent impact on DNA damage compared to external factors (*1*). The most common endogenous DNA damaging agent is the Reactive Oxygen Species (ROS), which refers to radical and non-radical oxygen species generated by the reduction of molecular oxygen, such as singlet oxygen atoms ($^1O_2$), hydroxyl radicals ($HO^{\cdot}$), superoxide radicals ($O_2^{\cdot -}$), hydrogen peroxide (H_2O_2) etc. **(Fig. 1.1)** (*2*). In general, the ROS scavenging mechanisms evolved by the cell helps in mitigating the ill effects of ROS generation by neutralising them. However, when there is an imbalance in the ratio of ROS levels to the available ROS scavenging defence mechanisms, it leads to oxidative stress.

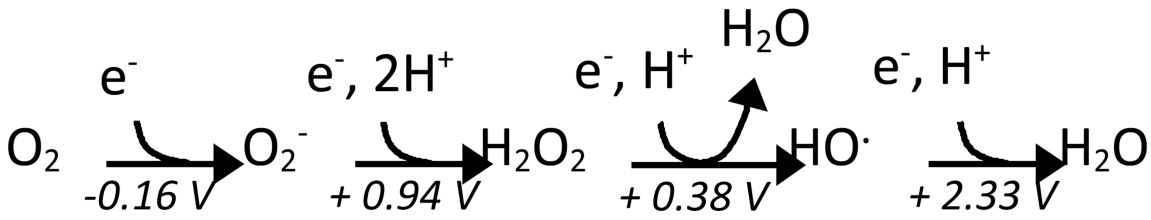


Fig. 1.1. Generation of reactive oxygen species (ROS) by univalent reduction of molecular oxygen (O_2) along with the standard reduction potentials is indicated. **[Adapted from Imlay, 2013] (*2*)**

The cellular defence mechanisms are of two types, namely (i) *enzymatic*: through the action of superoxide dismutase, catalase, peroxidases and peroxiredoxins, and (ii) *non-enzymatic*: which includes vitamins and pigments (*2-3*). The exogenous DNA damaging agents can be broadly classified into physical and chemical agents. Physical agents include ionizing and non-ionizing radiations that damage DNA (*4*). The source of radiation could either be (i) Natural: cosmic rays from sun, radioactive elements in the environment, soil and plants, or (ii) Artificial: radio pharmaceuticals and X-rays used for diagnosis and treatment. Chemical agents include (i) Heavy metals such as mercury, lead, arsenic, cadmium etc., which damage DNA through ROS generation, and (ii) Mitomycin C and their derivatives: 2,7-diaminomitosene (2,7-DAM) and 10-decarbamoyl mitomycin C (DMC) which induce formation of DNA adducts (*5*).

1.1.2 *Types of DNA damage*

DNA damage could involve (i) modification of nitrogenous bases through deamination, hydrolysis, alkylation or oxidation, (ii) induction of formation of DNA

adducts, or (iii) single or double stranded breaks in DNA. Deamination of nitrogenous base causes mismatch between DNA strands, while hydrolysis of the base causes depurination, and oxidation by ROS results in the conversion of purines to 8-oxopurines and pyrimidines to thymine glycol and uracil glycol (*6*). Oxidation of bases can also result in base substitution, the most common being that of GC to AT, and the conversion of cytosine to 5-hydroxyuracil (5-OHU) (*6*). Methylation of cytosine, a naturally occurring post-replication DNA modification process, generates 5-methyl cytosine (5-mC) which is highly prone to mutation and its spontaneous deamination results in the formation of thymine. The DNA repair machinery recognises G-T mismatch and replaces thymine with cytosine in order to restore the original DNA sequence, but under conditions when the repair is not efficient, it finally leads to mutations in the DNA (*7-8*). Exposure to UV-radiation induces the formation of DNA adducts such as 6-4 photoproducts (6-4PPs) and cyclobutane-pyrimidine dimers (CPDs) and their isomers, of which thymine dimers are very common (*9*). All these damages finally result in single or double DNA strand breaks (*1*). DNA damage through ionising radiations (IR) could either be a direct effect resulting in single and double strand breaks or indirect through the action of hydroxyl radicals generated by radiolysis of water. In general, it is thought that 80% of the DNA damage is indirect through the action of ROS, while 20% results from direct interaction of IR photons with DNA (*10*).

1.2 DNA REPAIR

The type of DNA damage dictates the genetic outcome of the damage and the cellular response to overcome it. The DNA lesions may be occurring continuously in response to growth or environmental factors, and thus every organism has developed a

very complex mechanisms of DNA repair to cater to the various forms of DNA damage (*11*). The major types of DNA repair are listed below:

1.2.1 *Direct Repair*

Direct repair of damaged DNA takes place through three major mechanisms: (i) UV-light induced photolesions by photolyases, (ii) O-alkylated DNA damages by AGTs, i.e. O^6-alkylguanine-DNA-alkyltransferases, and (iii) N-alkylated DNA damages by AlkB family dioxygenases. A schematic representation of these changes is shown in **Fig. 1.2** which has been adapted from published work (*12*). UV-radiation induces two types of lesions: (i) cyclobutane pyrimidines dimers (CPDs), and (ii) pyrimidine pyrimidones (6-4) photoproducts (6-4 PPs) **(Fig. 1.2A)**, which are repaired by specific photolyases (*12*). All photolyases bind the cofactor flavin adenine dinucleotide (FAD) and only when in the fully reduced $FADH^-$ form, they are rendered active and utilise the blue and near-UV light to reverse the UV-induced DNA damage (*13*). Transfer of electron from the photoexcited cofactor to CPD induces breakage of the two intradimeric bonds, thereby repairing the damage (*14*). On the other hand, repair of 6-4 PPs is more complex, since in addition to the breaking of the covalent bond between the two pyrimidines, the hydroxyl group also needs to be transferred from the 5' end to the 3' end of pyrimidine **(Fig. 1.2A)**. Several hypotheses have been postulated which govern the repair mechanism and these have been reviewed extensively (*15*). They involve the formation of either an oxetane intermediate (*16*) or a water molecule (*17*). The mechanism requires transfer of both electrons and protons involving the cofactor FADH and the His^{364} residue of the photolyase, resulting in the conversion of 6-4 PPs to two thymine dimers (*15*).

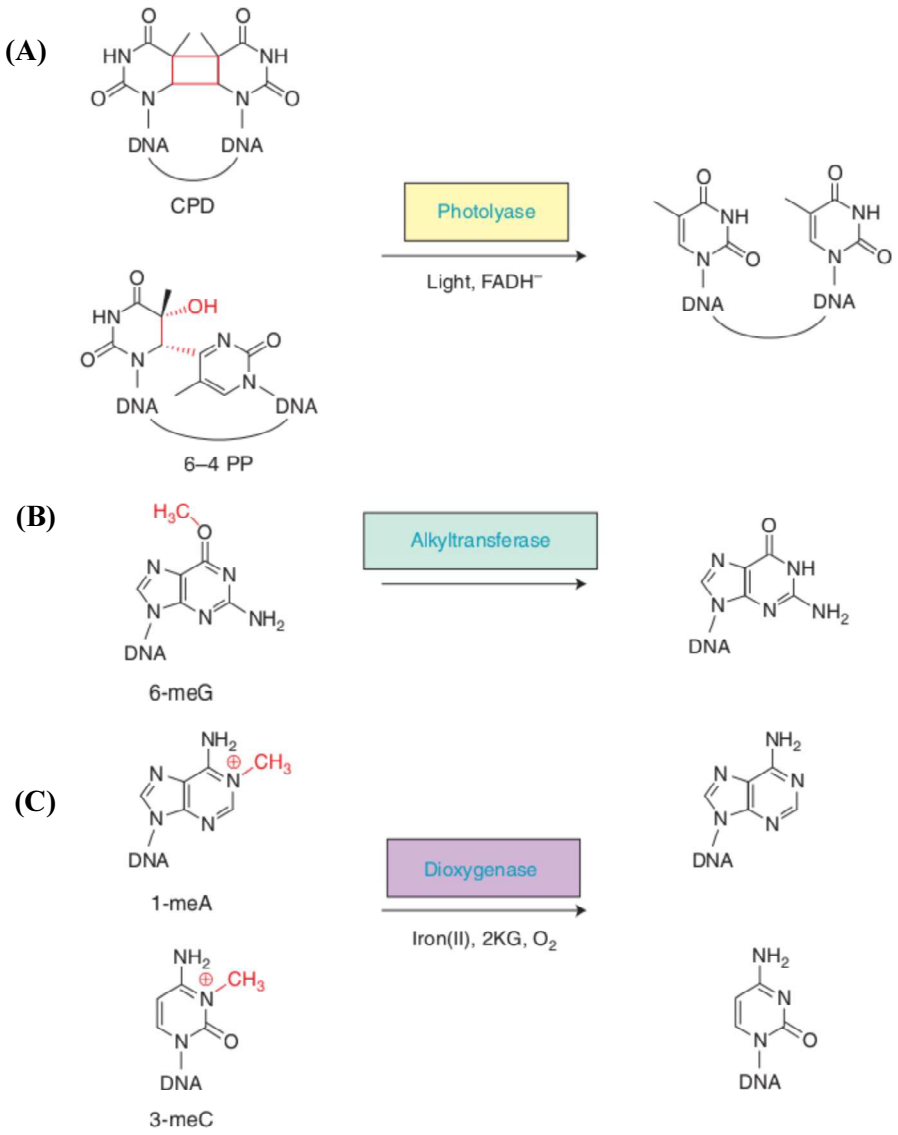


Fig. 1.2. Schematic representation of direct DNA repair, indicating the substrates, repair proteins, co-factors, and the ensuing repair product**. (A)** UV-damage, **(B)** O-alkylated damage, and **(C)** N-alkylated damage **[Adapted from Yi and He, 2013]**(***12***).

Repair of O^6-alkylguanine adduct is carried out by alkylguanine transferases (AGTs) as shown in **Fig. 1.2B**. The mechanism involves transfer of alkyl group from O^6-alkylguanine to the reactive cysteine residue of the AGT enzyme through an irreversible reaction as it involves covalent bonding which inactivates the enzyme (*18*). Alkyl adducts, N^1-methyladenine (1-meA) and N^3-methylcytosine (3-meC) (**Fig 1.2C**) are repaired by AlkB enzymes (*19*), through oxidative dealkylation, i.e. hydroxylation of the aberrant alkyl group, followed by cleavage of the C-N bond (*12*).

1.2.2 *Base Excision Repair*

Base excision repair (BER) system corrects most of the DNA damage caused by endogenous factors such as deaminations, oxidations, alkylations, depurinations and single strand breaks (*20*). These types of damages do not generally cause major distortions in the structure of DNA as they are usually restricted to the affected base. BER is mediated by several DNA glycosylases which catalyse the cleavage of the glycosidic bond linking the lesion to the DNA backbone (*21*). The mechanism of repair depends on whether the glycosylase is monofunctional or bi-/tri-functional as shown in **Fig. 1.3**. In brief, the DNA glycosylase recognises the DNA lesion and removes the modified base by hydrolysing the N-glycosidic linkage, leaving behind an AP (apurinic/apyrimidinic) site. The AP site is then targeted by AP endonucleases or the bi/tri-functional glycosylases, followed by filling up of the gap in the nucleotide chain by DNA polymerase and sealing of the ensued nick by DNA ligase (*11, 21*) .

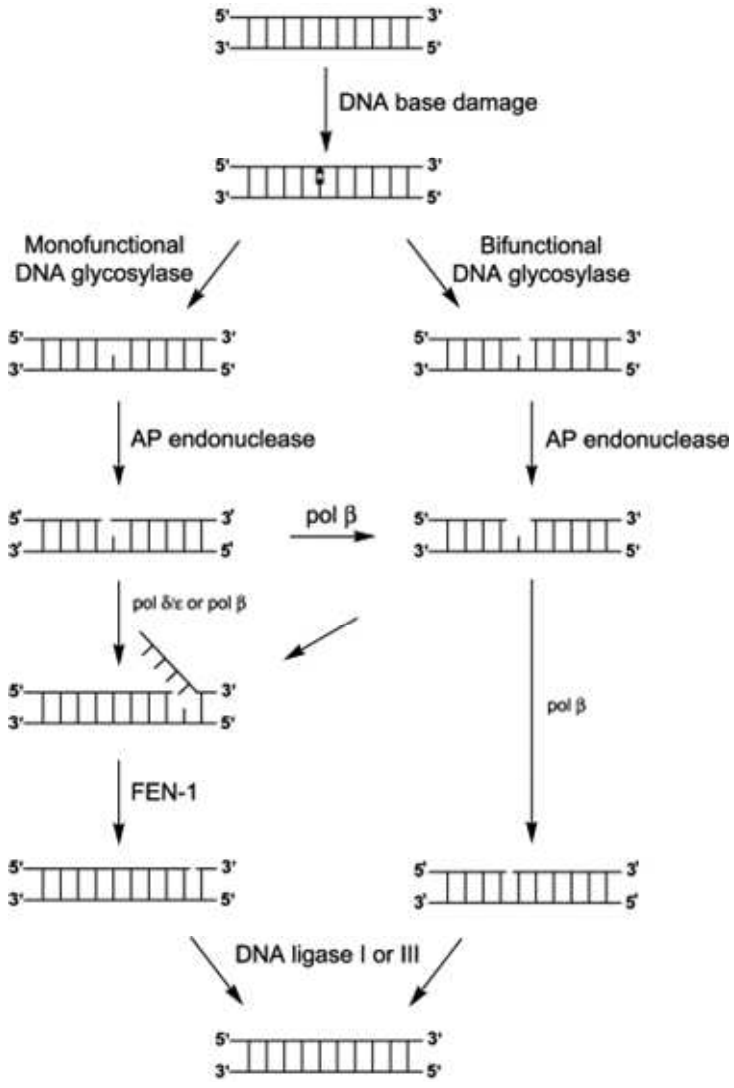


Fig. 1.3. Base excision repair pathway, indicating the different enzymes involved **[Adapted from Fromme and Verdine, 2004] (*21*).**

1.2.3 *Nucleotide Excision Repair*

Nucleotide excision repair (NER) system is the most conserved DNA repair system. It involves four steps: (i) damage recognition, (ii) formation of pre-incision complex, (iii) excision and repair synthesis, and (iv) release of post-incision complex followed by DNA ligation (*22*). There are two pathways for NER, namely Global Genome Repair (GGR) and Transcription-Coupled Repair (TCR), which differ mainly in the DNA detection step, as shown in **Fig. 1.4**. In GGR, the damage is detected across the genome directly by $UvrA_2B$ complex of UvrAB machinery, but in case of TCR the damage is encountered by RNA Polymerase during transcription which stalls at the damage site from where it is dislodged by transcription coupled release factors. This is followed by the recruitment of UvrAB machinery to the damaged site (*23*). The steps involving verification of the damage and its repair is same for both the pathways **(Fig. 1.4)**. Formation of the UvrB-DNA pre-incision complex is followed by recruitment of UvrC to the site. The nuclease activity of UvrC causes hydrolysis of the phosphodiester bonds on both 5' and 3' end of the lesion. The post-incision complex is removed by the helicase, UvrD and the gap filled by DNA Polymerase, followed by ligation by DNA Ligase (*11, 24*).

1.2.4 *Alternative Excision Repair*

UV-radiation induced damage including CPD and 6-4PP can also be repaired *via* alternative excision repair (AER) pathway, initiated by UV Damage Endonucleases (UVDE). The enzyme UVDE creates a single nick at immediate 5' end of the different UV-induced DNA lesions, followed by its removal through NER-independent UVDE-

mediated Excision Repair (UVER) pathway. Thereafter, the repair pathway is similar to that for BER or single strand break repair (*25*).

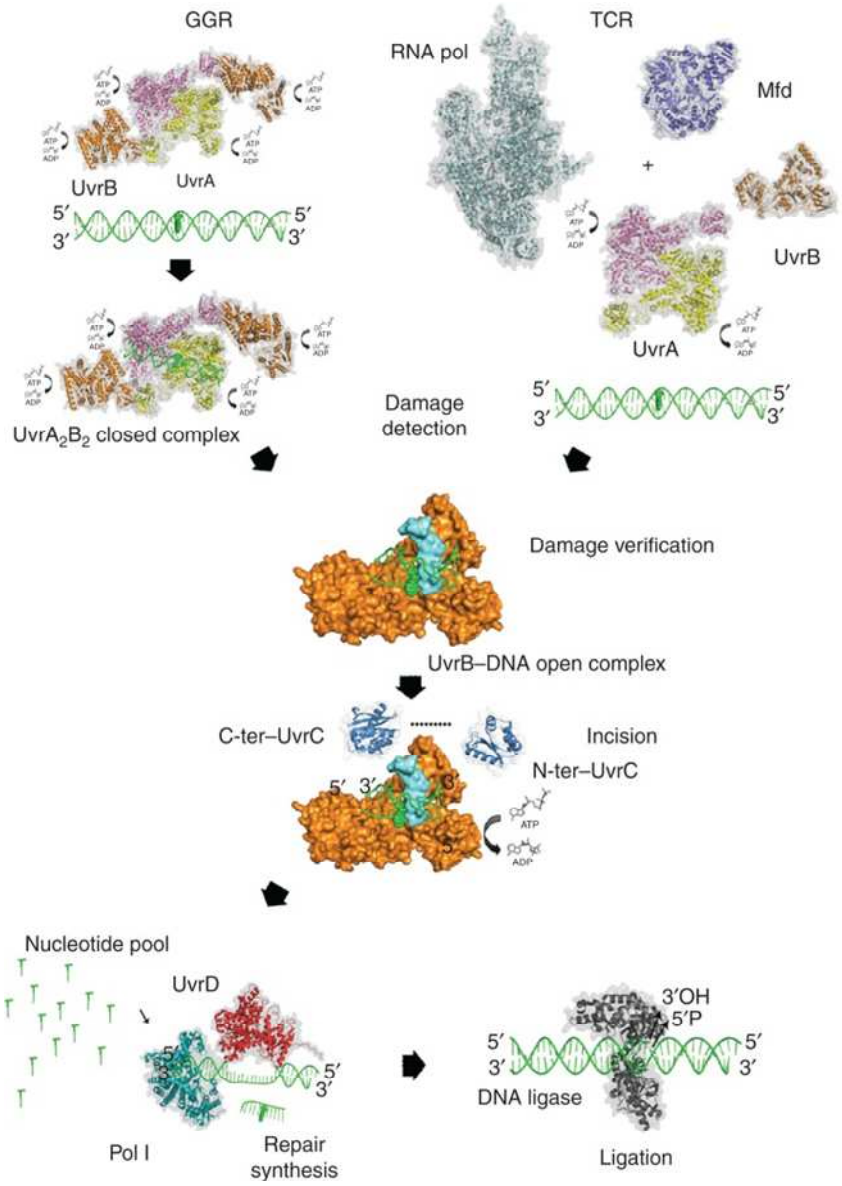


Fig. 1.4 Schematic representation of the prokaryotic NER pathway. GGR stands for Global Genome Repair and TCR stands for Transcriptional Coupled Repair. Involvement of UvrAB machinery, RNAP, DNA PolI, DNA ligase are indicated**[Adapted from Kisker et al. 2013] (*24*)**.

1.2.5 *Mismatch Repair*

Mismatch repair (MMR) is usually used for correction of the errors resulting during DNA replication, manifested as mismatched or unpaired bases, thus enhancing fidelity of DNA replication. This is carried out with the help of three proteins MutS, MutL and MutH comprising the MMR system (*26*). Damage recognition occurs by the binding of the MutS homodimer to the mismatched base (**Fig. 1.5**), followed by the recruitment of MutL homodimer, thereby stabilising it and finally activating the restriction endonuclease activity of MutH, which creates a nick (*27*). The nick on the unmethylated strand at GATC site is excised by DNA helicase and ssDNA-specific endonuclease and the gap is filled by DNA polymerase III and ligase (*28-29*) as shown in **Fig. 1.5**.

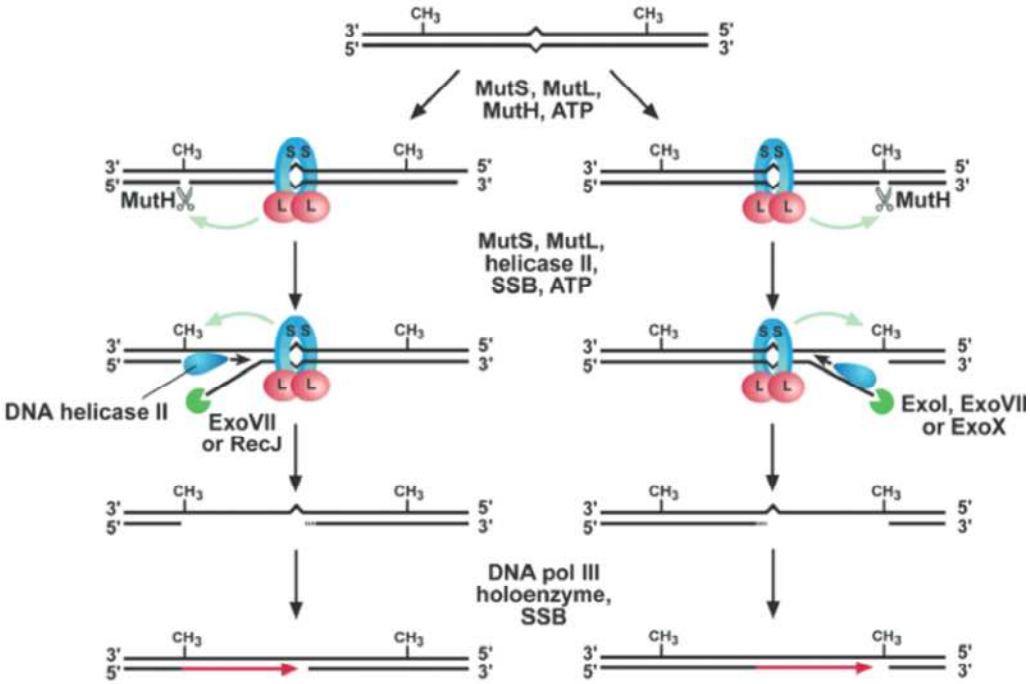


Fig. 1.5 Mechanism of *E. coli* methyl-directed mismatch repair **[Adapted from Iyer et al., 2006] (*30*).**

1.2.6 *Double strand breaks repair*

Double strand breaks (DSBs) generated as a result of different DNA damaging agents can result in loss of information and thus, could be lethal if not repaired efficiently before cell division. In bacteria, DSBs are repaired through either Homologous Recombination (**HR**) or Non-Homologous End Joining (**NHEJ**) (*31*).

Homologous recombination (HR) is a ubiquitous process across all organisms, though the mechanism may differ depending on the complexity of the organism. The repair of double strand breaks by HR involves five conserved steps as shown in **Fig. 1.6**. The different steps are (i) recognition of the DSB, followed by its processing to generate free 3' single stranded DNA (ssDNA) overhangs, which is carried out by the exonuclease RecBCD in *E. coli*, (ii) generation of nucleoprotein filament by loading of the recombinase RecA onto the generated ssDNA, (iii) homologous pairing of ssDNA with an intact homologous DNA segment, followed by strand invasion and formation of a crossover junction, (iv) DNA synthesis from the 3'-OH of the invading strand using homologous DNA as template. The intermediates, thus formed, are processed by a branch migration reaction which results in the formation of stable four-stranded DNA structures known as Holliday Junctions (HJs), and (v) endonucleolytic resolution of the HJs followed by sealing of the nick by ligase. This finally results in the formation of two intact repaired DNA molecules (*11, 32-33*).

Non-Homologous End Joining (NHEJ) pathway is relatively simple and is not dependent on the presence of homologous copy of the chromosome. This offers an advantage over HR, since a second copy of the chromosome need not be available for repair of DSBs. The process is initiated by Ku proteins which hold both the DNA strands

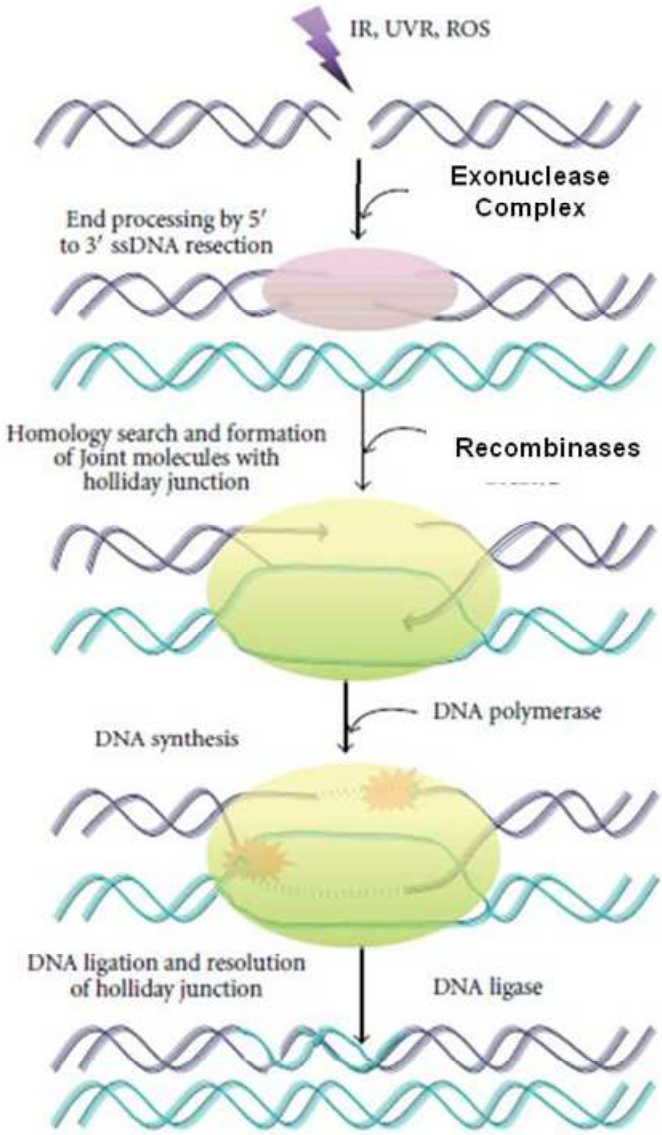


Fig. 1.6 Schematic representation of homologous recombination (HR) process for repair of DSBs **[Adapted from Rastogi et al., 2010] (*34*).**

together, followed by sealing by DNA ligase, LigD as shown in **Fig. 1.7**. This pathway can result in error-free repair only if ends are sealed directly. However, if the ends are acted upon by nucleases or polymerases, it can result in mutagenic repair (*35-36*) . **HR** has an advantage over **NHEJ** as it results in high fidelity DNA repair, but the major disadvantage is that it requires homologous DNA molecule to use as a template to carry out repair of the damaged copy of the DNA (*31*).

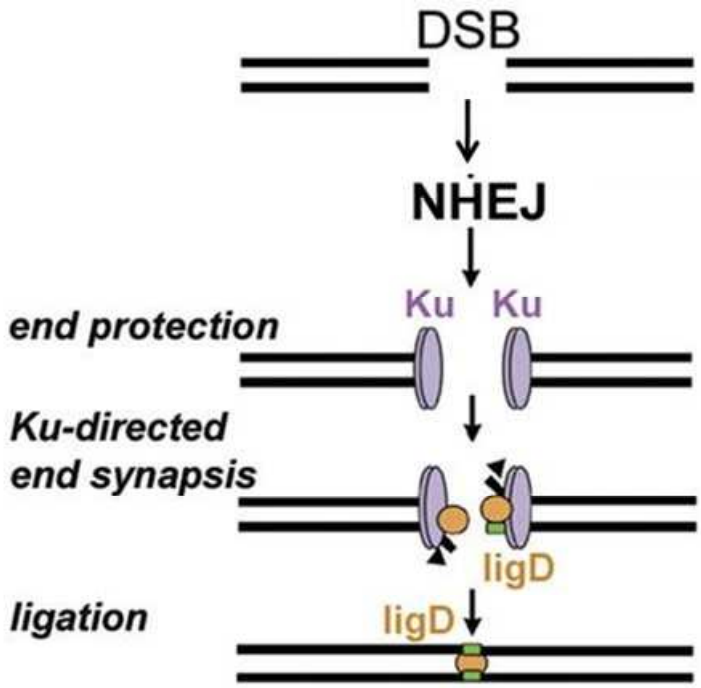


Fig. 1.7 Schematic representation of non homologous end joining pathway for repair of DSBs **[Adapted from Chayot et al., 2010] (*37*).**

In NHEJ deficient cells, another error prone DNA repair system was identified to repair DSBs, which was referred to as alt-NHEJ pathway. One of the forms of alt-NHEJ pathways is **Microhomology Mediated End Joining (MMEJ).** Herein small homologous sequences present at the broken ends align prior to end joining. This type of

repair has been mostly reported in eukaryotic systems, such as yeast and mammalian cells (*38-39*).

1.2.7 *Error prone DNA repair (SOS repair)*

Upon accumulation of massive DNA damages under specific physiological conditions, the SOS response is induced, which is an error-prone method of DNA repair. This topic has been dealt in detail below in **Section 1.3**.

1.3 SOS Response

Exposure to certain environmental changes can cause sudden increase in DNA damage in bacteria necessitating the need to evolve a DNA repair system in order to survive. The first experimental evidence of an inducible DNA repair system was demonstrated in early 1970s by Miroslav Radman, who coined the term "SOS Response" (*40*). SOS response involves interaction between two key proteins, LexA and RecA which governs its ability to repair DNA in a coordinated manner and was first described in detail in *Escherichia coli* (*41*). Under normal growth conditions when DNA damage is low, the need for repair is also low. Under these conditions, LexA exists as a dimeric protein and binds to the SOS (or LexA) box. The LexA box spans a 20 base pair region with the ends representing palindromic DNA sequence and is present in the upstream regulatory region of SOS responsive genes (*42*). Upon DNA damage, RecA is activated by the generated ssDNA, which in turn aids conformational change in LexA protein inducing auto-proteolytic cleavage, thereby resulting in the de-repression of SOS-responsive genes (*43*).

1.3.1 *Role of RecA in SOS response*

The DNA binding protein, RecA upon binding to single-stranded DNA (ssDNA), leads to the generation of nucleoprotein filaments (*44*). These filaments of RecA have dual functions: (i) catalysis of strand exchange reaction, a key step during homologous recombinational repair of DNA, and (ii) inducing the cleavage of LexA (*45*). During normal course of DNA replication, single stranded DNA binding (SSB) protein remains bound to the ssDNA near the replication fork. In order to displace the SSB from ssDNA RecA protein has to overcome the nucleation barrier. This does not happen during normal replication, since the ssDNA is available for a very short time which is not sufficient for RecA to displace SSB proteins (*45*). However when the DNA is damaged, two recombination pathways (RecFOR and RecBCD) come into play and enable RecA protein to overcome the barrier and displace the SSB proteins bound to the ssDNA. Of these two, the RecFOR system helps RecA protein to bind ssDNA generated due to DNA damage. In the RecBCD system, RecBC proteins help in the loading of RecA proteins onto the ssDNA generated after processing of double strand breaks (*46*). On the other hand, DinI, PsiB and $UmuD_2C$ proteins, negatively regulate the formation and activity of RecA-nucleoprotein filament (*47-50*). DinI protein specifically binds to RecA-nucleoprotein filaments, but not to free ssDNA or free RecA proteins, thereby inhibits the co-protease activity of RecA (*48*). PsiB protein binds to free RecA protein, preventing it from forming active RecA-nucleoprotein filaments (*49*). The $Umu(D')_2C$ protein complex binds to the deep groove of the helix turn of the RecA nucleoprotein thereby modifying its activity and preventing it from participating in the proteolytic cleavage of LexA (*47*).

1.3.2 *Triggers of the SOS Response*

The SOS response in bacteria can be triggered by either endogenous events or exogenous agents that cause accumulation of DNA damage and is considered as an adaptive response to DNA damage (*51*). The endogenous events result from cellular metabolic activities, such as stalled replication fork arising during replication/recombination events of DNA resulting in either single strand or double strand DNA breaks (*52*). Double stand breaks (DSBs), originating due to stalling of replication **(Fig. 1.8A),** are converted to ssDNA mostly by a RecBCD-dependent mechanism (*52*). At times these conversions also occur through a RecBCD-independent mechanism (*53*), which is of prime importance for microbes, such as cyanobacteria, which do not possess the *recBC* genes or their homologs. Stalling of replication can also result in the formation of ssDNA in the lagging strand, giving rise to RecA-dependent SOS response (*54*). Collision between transcription and replication can also result in the generation of ssDNA **(Fig. 1.8B)**. This usually occurs when bulky transcriptional and replication machinery progress in opposite directions and the bacterial cell is unable to stall the RNA polymerase (*55*). Under non-replicating conditions, stalling of RNA polymerase results in the formation of RNA-DNA hybrids which are known as R-loops, which in turn induce formation of DSBs (*56*) as shown in **Fig. 1.8**.

Reactive oxygen species (ROS) such as superoxide radical ($O_2^{\cdot-}$), hydrogen peroxide (H_2O_2), and the highly reactive hydroxyl radical ($HO^{\cdot}$), generated endogenously during respiration/photosynthesis or due to the presence of exogenous agents in the environment, also inflicts damage to the DNA resulting in the induction of SOS response (*57*). Superoxide radicals inducing the release of free Fe from the Fe-S clusters of proteins which can localize in the phosphodiester backbone of DNA. When present at high concentrations, it can induce the formation of highly reactive hydroxyl

radical ($HO^{\cdot}$) through Fenton reaction which induces DSBs by attacking the DNA backbone (*2*). ROS also targets the lipids and proteins leading to their denaturation, which in turn results in stalled replication or transcription, thereby inducing the SOS response (*52*). External oxidative stress-inducing agents such as H_2O_2 also results in the induction of SOS response through the generation of $HO^{\cdot}$ by Fenton reaction. This is mainly repaired by RecA-mediated homologous recombinational repair (*58*).

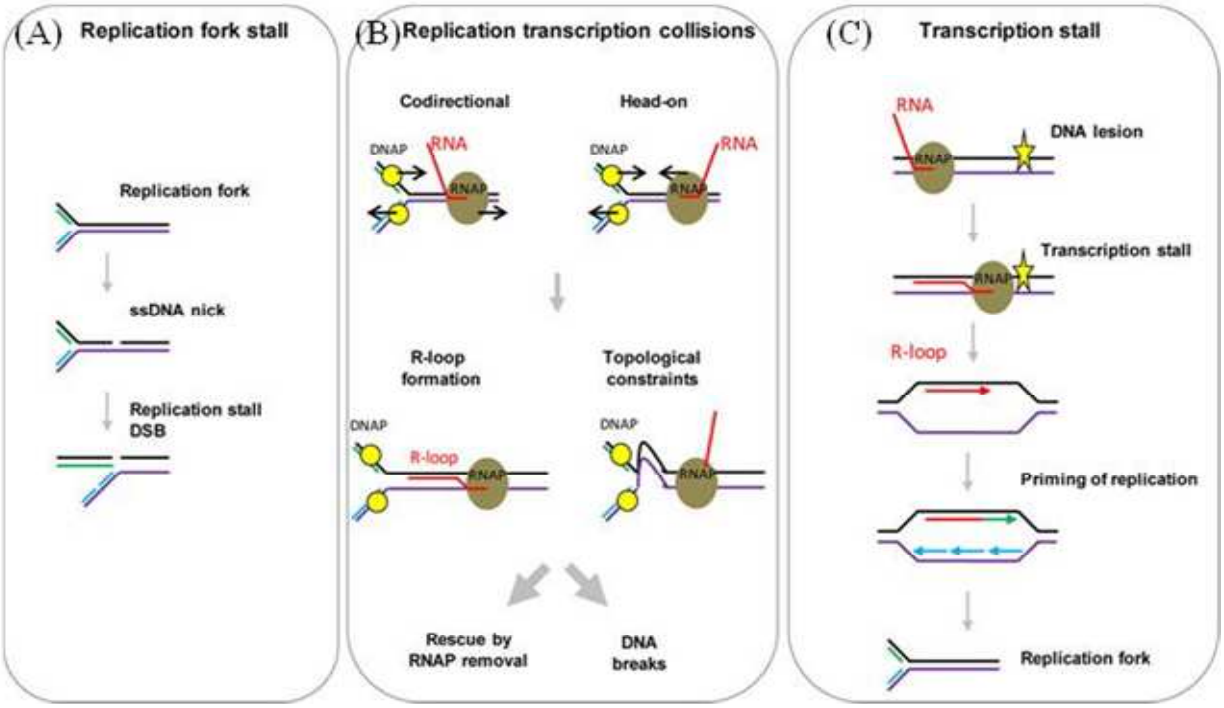


Fig. 1.8. SOS-induction through generation of ssDNA by; (a) Stalled replication fork (b) Collision between replication and transcription. (c) Stalling of transcription stalling **[Adapted from Baharoglu and Mazel, 2014]** (*52*)**.**

Other exogenous agents which can induce SOS response are (i) UV-irradiation through adduct formation followed by breaks in DNA (*9*), (ii) chemicals such as mitomycin C through the formation of DNA adducts (*59*), (iii) antibiotics such as

fluoroquinolones which inhibit DNA gyrase (*60*), trimethoprim which affects purine and pyrmidine synthesis (*61*), and ß lactams through inhibition of cell wall synthesis (*62*). Physical stressors, such as high pressure result in the generation of double stranded breaks though the activation of Mrr restriction endonuclease (*63*). In case of *Vibrio cholerae,* addition of sub-lethal level of non-genotoxic antibiotics such as aminoglycosides, tetracycline, and chloramphenicol have also been shown to result in the induction of SOS response (*64-66*), which was found to be due to the stalling of transcription (*67*).

1.3.3 *SOS response at the molecular level*

Generation of ssDNA gaps, which is central to the induction of SOS response, has been proposed to occur through two models, namely "Runaway Helicase Model" and "Lesion Skipping Model" (*68*). According to the **Runaway Helicase Model**, named due to the involvement of DnaB helicase, the differences arising in the leading and lagging strand due to stalled replication leads to the generation of single strand breaks in the DNA. When the replication is stalled due to the presence of the DNA lesion, DnaB helicase continues to unwind dsDNA ahead of the stalled replication fork, resulting in the generation of ssDNA. In case of the lagging strand, wherein the replication is discontinuous, the SSB proteins coat the generated ssDNA protecting it from the action of nucleases. A new RNA primer is generated with the help of DNA primase converting it to dsDNA (*69*). However, in the leading strand wherein DNA synthesis is continuous, due to the persistence of ssDNA for longer duration SSB proteins are dislodged by RecA, leading to the formation of RecA nucleoprotein complex, aided by RecFOR proteins (*69*). The discontinuous synthesis of DNA on lagging strand leads to the generation of ssDNA gaps as the DNA Pol III keeps hopping to the next primer and

continues DNA synthesis encountering DNA lesions at replication fork (*69*). Single stranded gaps in the leading strand are generated owing to two mechanisms. In the first, when replication and transcription are in the same direction, the replisome due to its higher speed, overtakes and displaces RNA polymerase, but retains the transcript which is then used to continue the DNA synthesis in the leading strand and beyond the stalled replication fork resulting in gaps. However, when the fork does not collapse, the DnaB of lagging strand may allow Primase to prime leading strand which allows replisome Pol III to continue DNA replication in the leading strand, thus leading to the formation of ssDNA gaps (*69*). According to the "**Lesion Skipping Model**", when replisome encounters DNA damage, it skips the DNA block and continues DNA synthesis, resulting in the generation of ssDNA gaps. If the replisome encounters heavy DNA damage, it will result in multiple ssDNA gaps which ultimately allow RecFOR to load RecA and induce SOS response. In order to allow efficient DNA repair under these conditions, the RecA protein acts as a switch to regulate replisome by promoting translesional DNA polymerases (TLS) over DNA Pol III. TLS replisome carries out replication at a slower speed than that by Pol III thereby giving sufficient time to repair DNA by other systems (*70*). Under normal growth conditions i.e. in the absence of DNA damage, the short ssDNA generated during the course of replication is coated by SSBs which inhibits translesional DNA polymerases.

1.3.4 *The induction order of SOS response*

Induction of SOS response shows a staggered and temporal approach which is considered as one of the most important characteristic feature of the SOS response (*71*) This is governed by strength of the promoter, its vicinity to the SOS-box as well as the binding affinity of LexA to the SOS-box (*46*). The first set of genes induced during SOS

response are *uvrA*, *uvrB* and *uvrD,* whose protein products along with the endonuclease UvrC are involved in nucleotide excision repair (NER) (*24*). The next line of defence for repair of the DNA lesions is to facilitate homologous recombination (HR) function through the induction of RecA and other proteins involved in HR which converts them to a double-stranded region, which in turn act as a substrate for NER (*72*). Further, in order to provide sufficient time for DNA repair to be completed, cell division is inhibited by inducing the inhibitor SfiA (*45*). The last of the genes to be induced, usually about 40 min after DNA damage in *E. coli*, are the low fidelity, error-prone repair DNA polymerases namely Pol II (*polB*), Pol IV (*dinB*), and Pol V (*umuC*, *umuD*) (*45*). These polymerases take part in translesional DNA synthesis which is necessitated when there is extensive DNA damage, which is not be repaired efficiently by NER and HR (*73*). Due to their low fidelity, these DNA polymerases operate in a less processive and error-prone manner which allows the irreparable DNA damages to be repaired, thus removing the blocks at replication and giving a chance to the cell to survive (*74*). However, the repair of damaged DNA is at the cost of introduction of some mutations, thus increasing the mutation rate in bacteria (*54*). In the end, in case DNA damage is not repaired by the above enzymes, the prophages and colicin plasmids are induced to induce cell lysis (*45*).

1.3.5 *SOS response: Multiple Roles*

SOS response is not only a means of handling DNA damage, but has a significant role to play in the induction of adaptive mutations, drug resistance, persister formation, biofilms and quorum sensing. The induction of DNA Polymerases Pol II, Pol IV and Pol V during SOS response plays a significant role in the accumulation of adaptive mutations. This was proven for Pol IV, encoded by *dinB* using the *lac* operon as the test system (*75*). Pol IV played a role in hypermutations, as indicated by the lower frequency

of hypermutations in adaptive *lac*$^{+}$ revertants in *E. coli* in Pol IV-defective strain (*76*), but not in the chromosomal mutation rate (*77*). It also aided Pol V ($UmuD_2C$) in inducing SOS mutator phenotype in *E. coli* (*78*). When subjected to stress, the different DNA polymerases; Pol I, Pol II and Pol III in *E. coli* cells compete with Pol IV at the site of damage, and with the levels of Pol IV being higher upon SOS induction, it tilts the usage in favour of Pol IV thus contributing to adaptive mutagenesis (*79*).

The introduction of mutations during SOS response has also been shown to play a role in the evolution of antibiotic resistance (*80*). This was found to be very critical in the development of resistance to ciprofloxacin, used for treating *Pseudomonas* infections, and thus, the inhibition of LexA activity has been suggested as a means to combat the growth of this pathogen (*81*). Induction of SOS response in *E. coli* has been shown to aid in the development of resistance to quinolones (*82*), fluoroquinolones (*83*) and ciprofloxacin, which is through the induction of the TisB toxin (*84*). Thus, targeting the SOS response could help in the reversion of the quinolone resistance (*85*). In *Staphylococcus,* the SOS response has been found to be induced in response to β-lactam antibiotics, thereby resulting in the spread of virulence factors across bacterial sp. (*86*), and also implicated in contributing to ciprofloxacin resistance (*87*). As a step towards fighting drug resistance, inhibition of RecA which is required for the activation of LexA proteolytic activity, has been considered as a potential candidate for blocking antibiotic resistance (*88*). Therefore, the SOS response through the induction of different genes, leads to the development of persistence and drug resistance, and can be viewed as a target for tacking drug resistance (*67*).

SOS response has also been implicated in the formation of biofilms and sporulation in bacteria. In general, biofilms are induced in response to DNA damaging agents and antibiotics, which are in turn known to induce SOS response. Using non-

cleavable LexA mutants and *lexA* deletion mutant in *Pseudomonas auregenosa,* direct involvement of LexA in biofilm formation was established (*89*). Expression of LexA also reduced the mobility of the organism, thus contributing to the biofilm formation (*90*). In *Clostridium difficile,* in addition to mobility and biofilm formation, the SOS response has been shown to be associated with sporulation also (*91*). Due to the effect on the mobility of bacteria, the SOS response also contributed to quorum sensing as shown for *E. coli* O157:H7 (*92*). Thus, the SOS response has multidimensional role in the life of bacteria depending on the demands of the environment.

1.4 Cleavage of LexA

The proteolytic digestion of LexA is the key to its function as a transcriptional regulator. The information for the cleavage resides in its amino acid sequence and is aided by various factors for conversion from a non-cleavable to a cleavable conformation as discussed in the following sections.

1.4.1 *Autodigestion of LexA*

The LexA and RecA proteins of *E. coli* are the two key regulatory proteins of the SOS response and are controlled through the proteolytic cleavage of LexA (*54*). Initially, the cleavage of LexA protein both *in vivo* and *in vitro* was considered as a result of the protease function of RecA protein (*93*) but, later on this concept of LexA cleavage underwent change when its cleavage was observed even under alkaline conditions in the absence of RecA (*94*). Thereafter, the cleavage of LexA was termed as auto-digestion (*95*). During auto-digestion at alkaline pH, the cleavage was found to occur exactly at the same site i.e. Ala^{84}-Gly^{85} as observed in the presence of RecA protein (*94*). This

suggested that the protease activity lies within the LexA protein, and the RecA protein may be playing an indirect stimulatory role as a co-protease rather than as an active and direct protease in inducing the cleavage of LexA (*94-97*).

1.4.2 *LexA Architecture*

E. coli LexA is a 198 amino acid long 24 kDa protein comprising of a N-terminal DNA binding domain (NTD) (1-69 amino acids) and a C-terminal catalytic domain (CTD) (75-198 amino acids) which are structurally very different and connected by a hydrophilic hinge region (*98-99*) as shown in **Fig. 1.9**. The dimeric state of native LexA protein was confirmed using both crystallographic and non-crystallographic techniques (*43*). The NTD contains three α-helices (α1, α2 and α3) and two anti-parallel β-strands (β1 and β2) (*43, 100*) as shown in **Fig. 1.9**. The α2 and α3 helices are involved in the formation of helix-turn-helix (HTH) DNA binding motif of the NTD of LexA protein (*100*).

The binding of LexA dimer to the SOS box occurs through classical winged helix-turn-helix motif (winged HTH), wherein HTH of each monomer of the LexA dimer interacts with two different but adjacent major groove of the DNA, while the wings (formed by strand-loop-strand structure of NTD) of both the NTD bind the same minor groove as shown in **Fig. 1.10**. This possibly accounts for the spacing between the two halves of SOS box being invariant in *E. coli* (*101*).

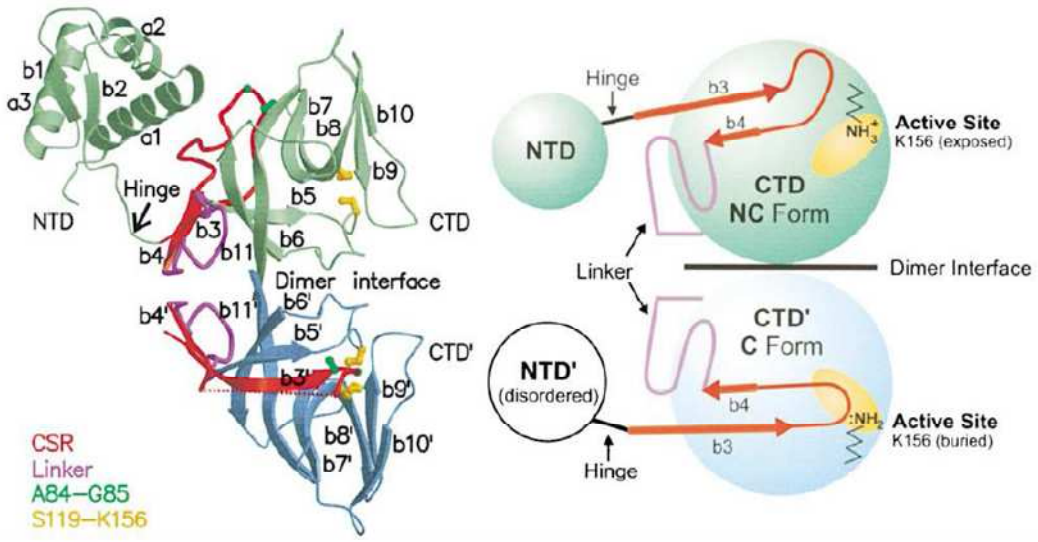


Fig. 1.9 Architecture of LexA-S119A dimer protein shown as (**A**) Ribbon model (**B**) Schematic model [**Adapted from Luo et al., 2001**] (*43*).

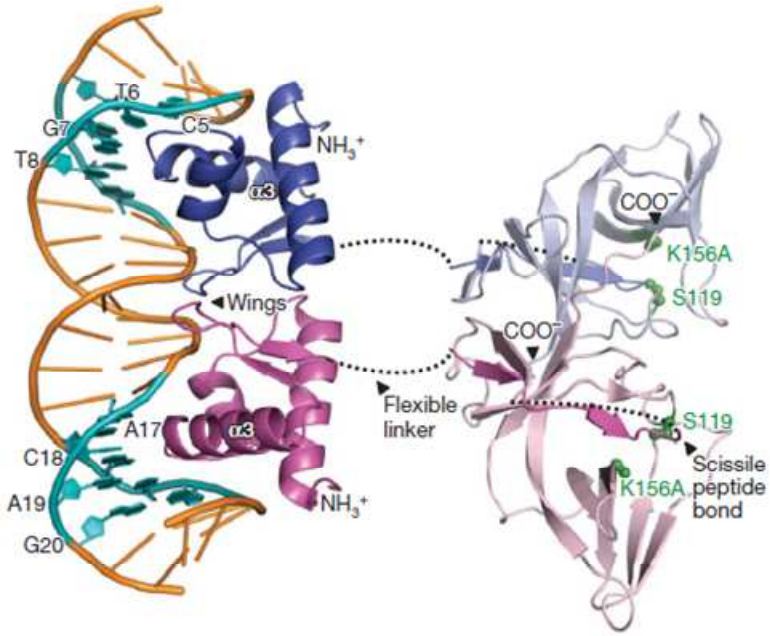


Fig. 1.10 Structure of *E. coli* LexA–DNA complex [**Adapted from Zhang et al., 2010**] (*101*).

The specific interaction with operator DNA is mainly through the Ser^{39}, Asn^{41}, Ala^{42}, Glu^{44} and Glu^{45} residues of α3 helix of HTH motif of NTD as revealed by both crystallographic (*43*) as well as NMR studies (*100, 102*). The CTD, on the other hand, exclusively consists of nine β-strands (β3 to β11) (*43*), sub-divided further into three regions: (i) cleavage-site region (CSR) (from Glu-75 to Tyr-98), (ii) linker region spanning Gln-99 to Asp-110, and (iii) the catalytic core region (from Phe-111 to Asn-198). The catalytic core region includes the proposed active site residues, Ser-118 which functions as a nucleophile, and Lys-156 which acts as a general base (*43*). The CTD also includes the dimerization interface, formed by the structural interaction of a portion of two loops (Gln^{99}-Asp^{110} and Ser^{116}-Gly^{128}) and the 11^{th} β-strand (i.e., β11) at the C-terminal (*43*), as shown in **Fig. 1.9**. Both the domains (NTD and CTD) of the LexA are connected through the hinge region, which is a short (Gln^{72} to Glu^{74}), flexible, highly hydrophilic region exposed on the surface (*43*), as shown in **Fig. 1.9**.

1.4.3 *Structure and Conservation of the Cleavage Site Region (CSR) across the LexA Superfamily*

The LexA protein can attain two conformations viz. cleavable (C) and non-cleavable (NC) depending on the structure of the cleavage site region (CSR). The structure of CSR is ordered and consists of two anti-parallel β-strands i.e. β3 and β4, connected with a small loop and closely associated with the catalytic core through the H-bonds between the 3^{rd} and the 5^{th} parallel β-strands (*43*). In its cleavable or C form, the CSR with its cleavage peptide bond (Ala^{84}-Gly^{84}) lies in the active site, adjacent to serine-lysine active site dyad, whereas in the non-cleavable or NC form, the cleavage site is approximately 20 Å away from the active site dyad (*43*), as shown in **Fig. 1.11**.

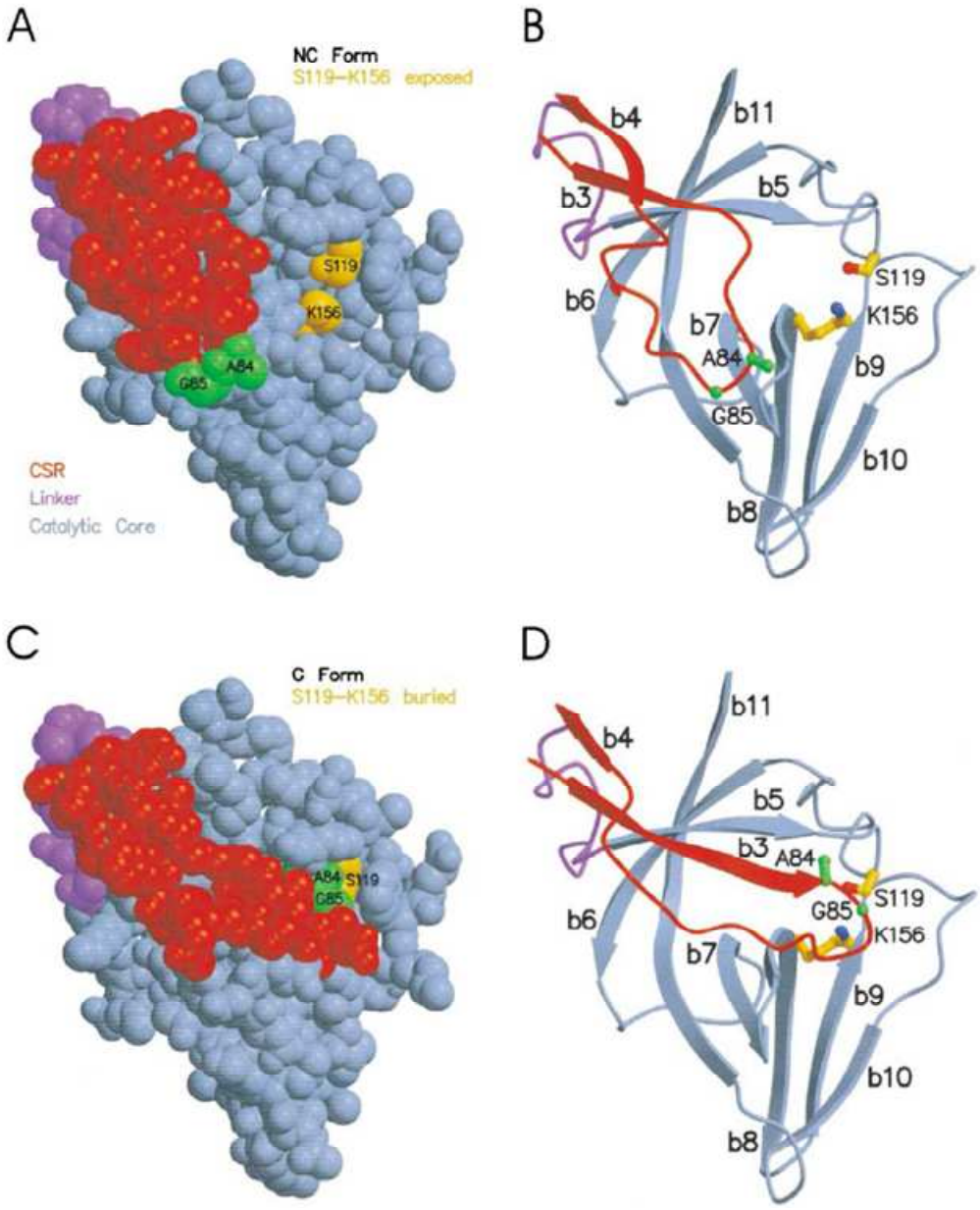


Fig. 1.11 Different conformational states of LexA. CTD of $\Delta_{1\text{-}67}$ S119A LexA mutant in NC conformation shown as (A) spacefilling and (B) ribbon model. CTD of $\Delta_{1\text{-}67}$ QM of LexA protein in C conformation depicted as (C) spacefilling and (D) ribbon model respectively [**Adapted from Luo et al., 2001**] (*43*).

The sequence of β3 strand of CSR is well conserved (**Fig.1.12)** which includes the highly conserved Pro[77] among the CSR of LexA Superfamily members. The CSR of all members of LexA Superfamily share a common structural motif which is known as the strand-loop-strand (SLS) motif of LexA. The hydrophobic surface which stabilizes CSR was also observed to be relatively conserved among LexA Superfamily members which is a indicative of that they all follow similar role in conversion of NC to C conformations of LexA (*43*). However, the rate of cleavage is different among the different members of LexA Superfamily, which is possibly in accordance with their individual biological roles (*103*).

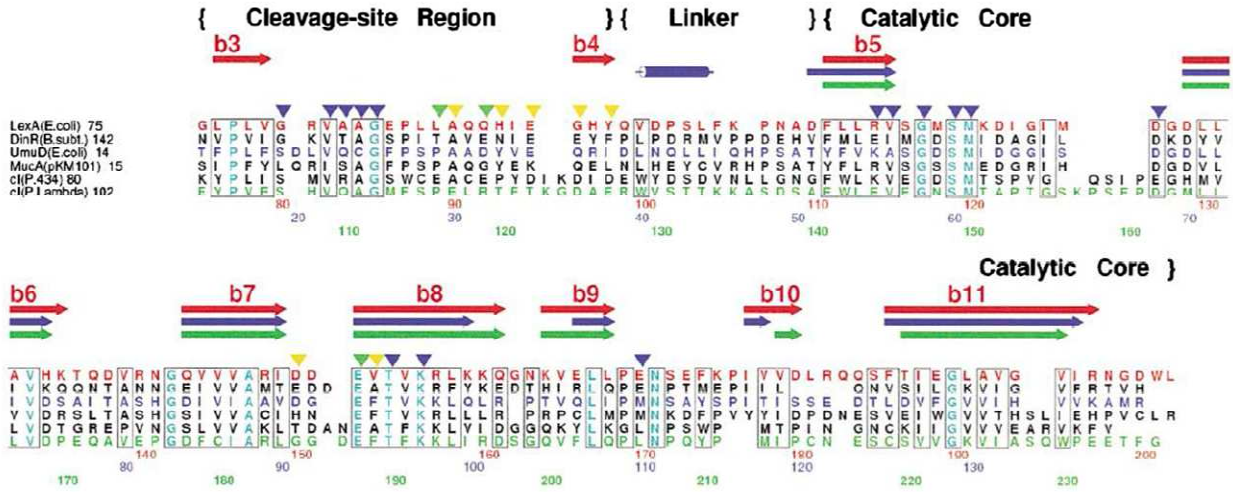


Fig. 1.12 Sequence alignment of the CTDs of six members of the LexA Superfamily, (LexA, DinR, UmuD, MucA, cl and λ cl [**Adapted from Luo et al., 2001**] (*43*).

1.4.4 *Conformational changes in LexA enabling its cleavage*

Specific residues identified as essential for the autoproteolytic activity of LexA are involved in inducing the conformational change required for exhibiting the activity. The β-hydroxyl group of Ser[119] forms a strong hydrogen bond with the ε-amino group of

Lys^{156} and this conformation of the side chains involving the Ser-Lys dyad is highly conserved among LexA and other proteins of the LexA Superfamily, such as UmuD` and λ-cI (*43*). The hydrophobic residues Met^{120} and Ile^{177} which form part of the active site interact with Lys^{156} through van der Waals forces, and are further buried by Ala^{84} to Pro^{87} from the CSR in the C-form. As shown in **Fig. 1.13**, in the C-form the Ser-Lys dyad is present at the end of the extended hydrophobic cleft, while Ala^{84} and Gly^{85} residues fit into the active site aided by the side chains of Val^{82} and Ala^{84} in complementing the hydrophobic nature of the pockets. Substitution of the side chains of Val to smaller side chains as in Ala or Gly or to polar ones, such as Ser or Thr renders the LexA incapable of auto-cleavage under alkaline pH conditions, but retains the ability to cleave under RecA-stimulated conditions (*104*).

Both the C and NC forms are highly stable, aided by 12 and 7 hydrogen bonds respectively between the CSR and the catalytic core. Interconversion between them requires only localized conformational change in terms of bond rotation in the region spanning Val^{79} to Glu^{95}. The substrate recognition cleft provides hydrophobic pockets to accommodate the side chains of Val^{82} and Ala^{84} in the C-form, while it is filled with ordered solvents in the NC form (**Figs. 1.11 and 1.13**). The binding of CSR to adjacent hydrophobic surface of the catalytic core helps to stabilize NC form, while the β-bulge formed by Val^{79} to Gly^{80} of β3-strand opposing Leu^{112} of β5-strand destabilize the C form. The protonation of the ε-amino group of Lys^{156} further destabilizes the C form. Thus, under normal *in vivo* conditions, the NC form of LexA is more favoured.

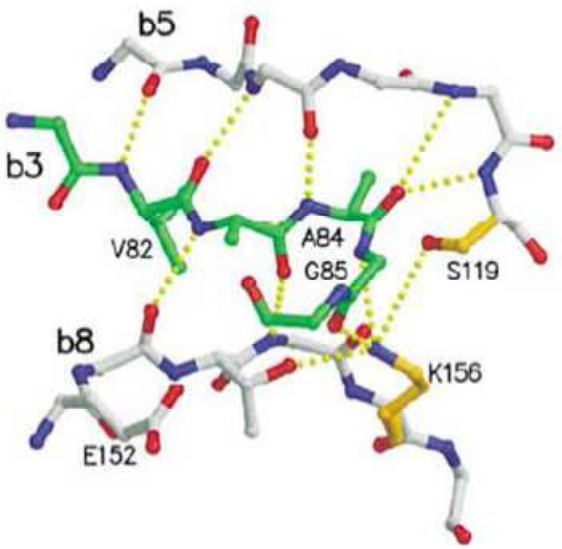


Fig. 1.13 Active site geometry of LexA with CSR in C conformation shown as Ball and stick model [**Adapted from Luo et al., 2001**] (*43*)

1.4.5 *Interaction of RecA with LexA: stimulation of cleavage*

The RecA co-protease could stimulate the cleavage of LexA under *in vivo* conditions, either by interacting directly with the cleavage mechanism or indirectly by favouring the C-form over the NC-form of LexA. Based on the crystallographic structure of LexA (*43*), direct involvement of RecA is ruled out as the active site in C-form of LexA as is fully buried, and thus the possibility for the RecA to reach the catalytic site negligible. Thus, the RecA may either be binding specifically to the C-form of LexA stabilising it or binds to both NC and C form equally well, while offering higher stability to the C-form rendering the conversion from the NC form to C form favourable (*105*). One of the subunits of the LexA dimer interacts with the deep helical groove of RecA nucleofilament through the DNA binding domain of LexA, through specific residues of

LexA, which have been identified as essential for interaction with RecA (*106*). Since, the RecA nucleofilament (activated RecA) is formed only when ssDNA is generated under DNA damage-inducing conditions, the entire process of activated RecA-mediated conversion from NC to C form occurs only under these conditions.

When tested under *in vitro* conditions, activated RecA needs to be added to the LexA to observe its cleavage under physiological pH (*107*). However, under alkaline pH, the OH^- ions de-protonates the protonated ε-amino group of Lys^{156}, thus stabilising the C-form. This activates the hydroxyl group of Ser^{118}, which then acts as a nucleophile targeting the peptide bond between Ala^{84} and Gly^{85}, resulting in the cleavage of LexA (**Fig. 1.14**). Thus, the mechanism of proteolytic cleavage of LexA remains the same irrespective of whether it is initiated by activated RecA or by alkaline pH.

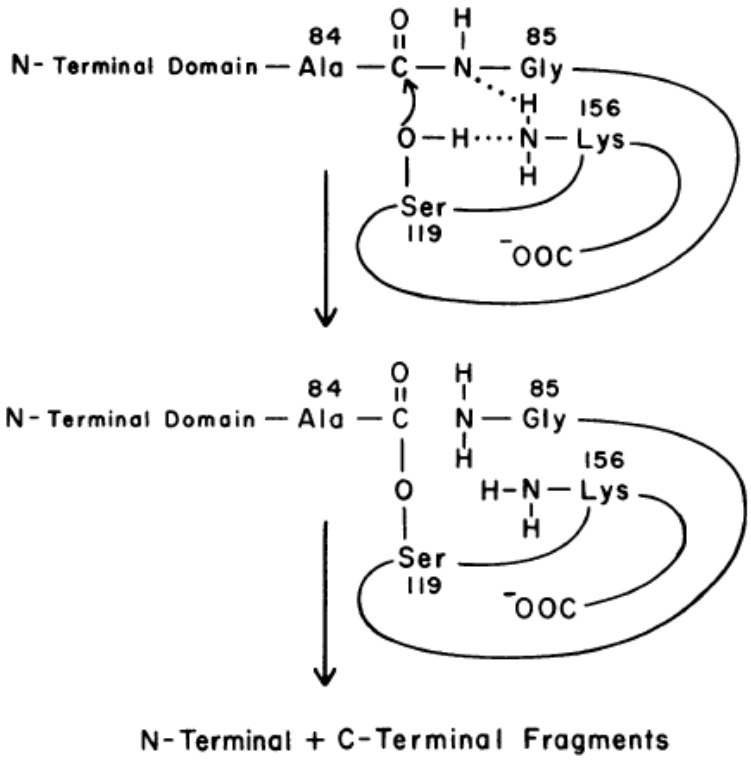


Fig. 1.14 Proposed mechanism for autoproteolytic cleavage of LexA repressor **[Adapted from Slilaty and Little, 1987] (*107*)**

1.5 Cyanobacteria

Cyanobacteria are considered to be the oldest, phylogenetically primitive and the most diverse group of Gram-negative bacteria, having originated around 3.5 billion years ago (*108-109*). Cyanobacteria are also known as Blue Green Algae (BGA) because of their bluish-green colour arising due to the presence of water soluble blue pigment phycocyanin and the green pigment chlorophyll *a* (*110*). Owing to similar architecture of their plasma membrane and thylakoid membrane, lipid composition, proteins organisation and similarities in several photobiochemical processes, they are considered to be linked to plants (*111-113*). Cyanobacteria are the only prokaryotes having the capability of oxygenic photosynthesis, and could possibly have contributed to the oxygenic atmosphere of Earth (*114-115*), and subsequently leading to the development of aerobic metabolism and allowing development of higher plants and animals (*116*). Cyanobacteria play a very important role in production of oxygen, and fixation of carbon dioxide (CO_2) and nitrogen (N_2 and NO_3) (*117*). They are thought to fix about 25 giga tonnes of carbon from CO_2 per year into energy biomass (*118*). In addition to these contributions, they have been known to thrive in diverse conditions, which include extreme conditions such as hot springs, deserts, frozen lakes etc (*119*).

A few of the cyanobacteria, in addition to exhibiting photosynthetic activity, also have the ability to fix atmospheric dinitrogen to ammonia with the help of the highly oxygen sensitive dinitrogenase enzyme (*120*). During the light reaction of photosynthesis, oxygen is evolved by photolysis of water, which is carried out by the oxygen evolving complex (OEC). This would, however, render nitrogen-fixation an untenable process. In order to overcome this problem, cyanobacteria have evolved two mechanisms to separate the two physiological processes, through space (spatial) or time

(temporal) (*120*). In temporal separation, photosynthesis occurs in light, while nitrogen fixation occurs in dark as is observed in unicellular *Gloecapsa* sp (*121*) and filamentous non-heterocystous, nitrogen-fixing cyanobacteria, *Oscillatoria* sp. (*122*). In spatial separation, both the physiological processes occur at physically separated spaces, through the formation of specialised cells known as heterocysts as observed in *Anabaena* sp., *Mastigocladus* sp. etc. (*123-124*). The heterocysts arise from vegetative cells through terminal differentiation (*125*). Their characteristic features include having less pigments, thick walls and absence of PSII. The absence of PSII, the site of oxygen evolution and the presence of thick walled envelop, composed of glycolipid and polysaccharides which reduces the diffusion of oxygen into the heterocyst from neighbouring vegetative cells, contribute to the maintenance of anaerobic interior. The heterocysts also show a high rate of respiration, compared to vegetative cells, which results in higher oxygen consumption. Heterocysts supply the fixed nitrogen to the vegetative cells and in turn take up carbohydrates from the vegetative cells to generate energy for various processes (*126*).

1.6 LexA in cyanobacteria

The very first papers on LexA in cyanobacteria were in the unicellular *Synechocystis* PCC6803 and *Synechococcus* PCC7942. By carrying out DNA microarray analysis of a LexA-depleted strain of *Synechocystis* PCC6803, involvement of LexA was found to be in the direct regulation of more than 50 genes which included 11 genes (7 down-regulated and 4 up-regulated) involved in carbon metabolism (*127*). In *Synechocystis* PCC6803, LexA was shown to be involved in the regulation of C-metabolism but not in that of DNA repair genes (*127-128*). LexA was found to be

essential for coping with inorganic carbon starvation as many of the genes involved in carbon uptake and assimilation were found to be regulated by LexA (*127*). The expression of *lexA* and *recA* gene transcripts was found to be negatively regulated by UV-C radiation (*127*), unlike *E. coli* where both the gene are induced by UV-C (*129*).

In *Anabaena* 7120, LexA was found to bind its own promoter at "TAGTACTAATGTTCTA" sequence as well as the promoter of *recA* gene. On the basis of mutational analysis of each nucleotide in the predicted site, a consensus sequence for binding of *Anabaena* LexA was derived, represented as "RGTACNNNDGTWCB" (*130*). The consensus cyanobacterial LexA box revealed that these are closely related to the Gram-positive LexA binding site (CGAACRNRYGTTYC). A single base substitution (GTAC→GAAC) in the LexA binding sequence of *lexA* gene promoter of *Anabaena* 7120 enable the LexA protein of *B. subtilis* to bind to the *Anabaena lexA* gene promoter. The *Anabaena* LexA protein was also able to bind this altered sequence but with the lower efficiency (*130*).

The LexA was found to regulate the genes coding for bidirectional hydrogenase both in *Synechocystis* (*131*) and *Anabaena* (*132*). The *hox* operon in *Synechocystis* sp. PCC6803 comprised of five genes (*hoxEFYUH* genes) which was transcribed as a single operon. The purified *Synechocystis* LexA protein was shown to bind to the *hox* regulatory region through EMSA studies. At lower LexA concentration, only one band corresponding to LexA-DNA complex was observed, while at higher LexA concentration, two bands corresponding to LexA-DNA complexes were observed, leading to the speculation of the presence of two putative LexA binding sites in the regulatory region of *hox* operon (*131*). The binding of LexA to the upstream region of the *hox* operon was also confirmed independently by the reduction of hydrogenase activity in LexA depleted mutants of *Synechocystis* (*133*). This was the first report on the

role of LexA as a transcriptional activator, rather than as a repressor in any bacteria. Subsequently, another protein AbrB-like was identified as an activator of the bidirectional dehyrogenase of *Synechocystis*, in addition to that by LexA (*134*) . In *Anabaena* 7120, the *hox* genes are clustered in two separate operons, the first constituting *hoxE* (*alr0751*) and *hoxF* (*alr0752*) genes preceded by ORF *alr0750*, and the second operon constituting *hoxU* (*alr0762*), *hoxY* (*alr0764*) and *hoxH* (*alr0766*) genes along with two other ORFs *alr0761* and *alr0763* of unknown functions (*132*). *Anabaena* LexA was found to bind to the region upstream of both the operons of bidirectional hydrogenase (*132*). The sequence "GGTACTCTGGTTCG" present 63 to 76 bp upstream of the *alr0750* ORF was identified as the LexA-binding box and involved in the regulation of the *hoxEF* operon. Though a similar LexA box could not be detected upstream of the *hoxU* gene, the sequence "TTACACTTTAA" similar to the consensus LexA box of *B. bacteriovorus* (TTACNNNGTAA) was found (*135*). Though, the LexA could bind to this region upstream of the *hoxU* gene, the expression of LexA was unaltered under anaerobic conditions, while that of *hox* genes was induced, suggesting presence of other additional regulators of *hox* operon in *Anabaena* 7120 as well (*132*).

In the marine cyanobacterium *Lyngbya majuscula*, LexA was also found to be involved in the regulation of hydrogenase accessory genes (*136*), transcribed by a single operon, *hypFCDEAB* (*137*). The products of the genes of *hyp* operon were suggested to be involved in the maturation and regulation of hydrogenase in *L. majuscula* (*136*). The promoter region of the *hyp* operon was predicted to have binding sites for two cyanobacterial transcriptional regulators, namely NtcA and LexA, and subsequently proven by EMSA studies (*137*).

Besides the hydrogenases, the LexA in *Synechocystis* also regulated the expression of RNA helicase, *crhR*, which is a redox-sensitive gene, by repressing its expression (*128*). The *lexA* and *crhR* genes were divergently expressed by the binding of the LexA protein. In spite of this, the two proteins exhibited differential expression under different conditions (*128*). DNaseI foot printing analysis for identifying the LexA-binding site in the intergenic region between *crhR* and *lexA* genes, revealed a 31 nucleotide region from +108 to +139 in case of *crhR* gene and 34 nucleotides from -5 to +29 in case of *lexA* gene as the probable LexA-bindiing boxes for the two genes. The alignment of both the protected sequences of *crhR* gene and *lexA* gene revealed the presence of the 12 bp direct repeat "CTA-N_9-CTA" in the two sequences. The importance of conserved nucleotides in this direct repeat motif was confirmed through mutational analysis (*138*).

In recent times, LexA was shown to be involved in the regulation of expression of fatty acid biosynthesis genes in *Synechocystis* 6803 (*139*). *In vitro* studies revealed that, LexA bound to the promoter region of all fatty acid biosynthesis genes, namely *fabD, fabH, fabF, fabG, fabZ* and *fabI.* Compared to the LexA-depleted strain of *Synechocystis*, the transcript levels of *fabD, fabH, fabF* and *fabG* genes were found to be reduced in the wild type *Synechocystis*, while that of *fabZ* and *fabI* unchanged under normal growth conditions. This suggested that LexA acts as a repressor for the *fab* genes which are involved in the initiation of fatty acid biosynthesis (*fabD, fabH* and *fabF*) and the first reductive step in elongation cycle (*fabG*). Under N-depleted conditions, the transcript levels of *fabD, fabH* and *fabF* were lower in the wild type compared to the *lexA* mutant, while that of *fabG, fabZ* and *fabI* were higher under P-depletion condition. Thus under P-depleted condition, LexA acts as an activator of *fabG, fabZ* and *fabI* genes

in *Synechocystis* 6803. Thus, the LexA has been considered as a global transcriptional regulator of *fab* genes in *Synechocystis* (*139*).

LexA in *Synechocystis* 6803, in addition to being under transcriptional regulation was also found to be regulated at the post-transcriptional level, based on *lexA* transcript and LexA protein expression studies under different stress conditions, such as darkness and anaerobic conditions. In both the cases, an increase in the levels of the *lexA* transcripts, but no change in the protein levels was observed. 2-D analysis further indicated that LexA, predominantly localised in the cytoplasm, undergoes post-translational modifications (*140*).

Whole genome wide transcriptional profiles using DNA microarrays were carried out in two marine cyanobacteria *Synechococcus* WH8102 (open ocean marine cyanobacterium) and *Synechococcus* CC9311 (coastal marine cyanobacterium), following two DNA damaging agents exposure by mitomycin C and ethidium bromide. Both the *Synechococcus* sp. WH8102 and CC9311 have shown several differences in their transcriptional responses to both the agents but both have shown the induction of SOS response by upregulating the *recA*, *lexA* and *umuC* genes under mitomycin C treatment condition but not under ethidium bromide (*141*).

1.7 *Anabaena* sp. strain PCC7120

Anabaena sp. strain PCC 7120 is a filamentous, heterocystous cyanobacterium capable of fixing atmospheric nitrogen whose complete genome sequence is available (*142*). *Anabaena* PCC7120 genome consists of one chromosome and six plasmids, having a total size of 6.41Mb and the plasmids are α, β, γ, δ, ε and ζ of 408 kbp, 186 kbp, 101 kbp, 55 kbp, 40 kbp and 5.5 kbp respectively.

Anabaena PCC7120 chromosome has 5368 protein encoding genes, of which 2396 are with known function, 1519 genes were annotated as unknown and 1453 annotated as hypothetical proteins (*142*). Additionally, the chromosome has four structural RNA genes, four sets of rRNA genes and 48 tRNA genes. In *Anabaena*, GTG, TTG and ATT are also assigned as start codon in addition to the universal start codon ATG (*142*). The ORFs of *Anabaena* are denoted by three letters followed by a serial number in which the first letter represents the species name, **'a'** for *Anabaena*; the second letter represents length of the ORF i.e. if it is more than 100 codons, it is represented as '**l'** and if it is less than 100 codons, it is represented as '**s'**; and the last letter refers to the orientation of the ORF which is left '**l'** or right '**r'**.

Anabaena 7120 is amenable to genetic manipulation and the techniques such as conjugation (*143*), mutagenesis by homologous recombination (*144*) or conditionally lethal approach (*145*), and expression of exogenous genes by integrating in the chromosome (*146*) or on plasmid (*147*) are well established. This makes *Anabaena* 7120 a model system to study various gene functions in filamentous cyanobacteria.

1.8 Scope of the Thesis

The heterocystous cyanobacterium, *Anabaena* sp. strain PCC7120 displays high tolerance to radiation and desiccation stresses (*148*), which is thought to be a result of an efficient DNA repair system (*149*). LexA is one of the key regulators of DNA repair genes in bacteria and is known to regulate more than 30 genes involved in SOS response in *E. coli* (*42*). In *Anabaena* 7120, the *recA*, *lexA*, *uvrA* and *ssb* genes were shown to possess a LexA box-like sequence, which showed binding

affinity to LexA (*130*). This suggested possible regulation of at least some of the DNA repair genes of *Anabaena* by LexA, however its role as a SOS-response regulator, as in other bacteria, was not ascertained. In order to obtain an insight into the physiological role of LexA in *Anabaena*, work was initiated on the LexA protein. In *Anabaena* 7120, LexA is coded by the ORF *alr4908* (http://genome.microbedb.jp/cyanobase/*Anabaena*). This study emphasizes on the following investigations:

(i) Biochemical characterization of the *Anabaena* LexA protein

(ii) Characterization of physiological role of LexA

(iii) Identification of genes regulated by LexA and the determination of AnLexA-box

The thesis is composed in five chapters including **Chapter 1** as general introduction. The results have been compiled in two chapters **(Chapters 3 and 4)** and summarized in **Chapter 5** at the end, followed by the list of references. Brief details of different chapters are given below

Chapter 2 gives the details about the organisms used, growth conditions and abiotic stresses applied. It also provides details of the composition of various buffers used for isolation of nucleic acids and proteins, their electrophoretic separation and biochemical studies.

Chapter 3 describes the biophysical and biochemical characterization LexA proteins of *Anabaena* 7120. *Anabaena* LexA protein was overexpressed and purified from *E. coli* and used for biophysical and biochemical characterization in terms of its auto-proteolytic activity. Site-directed mutagenesis approach was adopted to generate LexA mutant proteins to identify the amino acids essential for proteolytic activity.

Abiotic stress-induced changes in the expression of the *lexA* gene in *Anabaena* have also been investigated.

Chapter 4 deals with studying the physiological response of overexpression of LexA in *Anabaena* during abiotic stresses, such as DNA damage-inducing, oxidative, C-starvation and heavy metal stresses. This chapter also gives the details of 2-D proteomic analysis of LexA overexpressing strain of *Anabaena* 7120. Through transcriptional analysis, promoter studies and DNA binding assays, the genes regulated by LexA has been identified and the AnLexA-Box has also been determined. Global role of LexA in *Anabaena* has been summarized in **Chapter 5**.

CHAPTER 2

MATERIALS

AND

METHODS

2.1 Bacterial Strains, Growth Measurement and Visualisation

2.1.1 *Organisms, culture media and growth conditions*

Different *E. coli* and *Anabaena* strains used during the course of this study are listed below in **Table 2.1**.

Table 2.1: List of bacterial strains used in this study

Strain	Characteristics	Source/Reference
***E. coli* strains**		
DH5α	F^- *recA*41 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 (rk^- mk^-) *supE*44 *relA*λ Δ*lacU*169	Lab collection
BL21(DE3)(p*lysS*)	F^- *ompT gal dcm lon hsd*S_B(r_B^- m_B^-) λ(DE3) p*lysS*(Cm^r)	Novagen
HB101	F^-mc^r Bm^r rhsdS20($r_B^-m_B^-$) *recA*13 *leu* B6ara-14 *pro*A2 *lac*Y1 *gal*K2 *xyl*-5 *mtl*-1 *rpsL*20 (Sm^R) *gln*V44 λ^-	Lab collection
***Anabaena* Strains**		
Anabaena sp. strain PCC7120 (*Anabaena* 7120)	Fresh water, filamentous, heterocystous, diazotroph	Lab collection
AnpAM	Nm^r, *Anabaena* 7120 harbouring the plasmid pAM1956	(*150*)

E. coli cells were grown in Luria-Bertani (LB) medium containing appropriate antibiotics **(Table 2.2)** at 37 °C either in liquid medium under shaking at 150 rpm in an orbital shaker or on LB solid agar (1.5%) plates under stationary conditions. *Anabaena* cultures were grown photoautotrophically in BG-11 liquid medium containing with 17 mM $NaNO_3$, pH 7.2 at 27 ± 2 °C under continuous illumination (30 μE) (*151*) with or without neomycin (10 μg mL^{-1} in liquid medium and 25 μg mL^{-1} in solid medium) either with constant shaking at 125 rpm or under stationary conditions.

Table 2.2: List of Antibiotics and other additives used in growth media

Antibiotics/additives	**Stock Concentration**	**Medium of preparation**	**Final Concentration**
Carbenicillin	100 mg mL^{-1}	Sterile Distilled water	100 μg mL^{-1}
Kanamycin	50 mg mL^{-1}	Sterile Distilled water	50 μg mL^{-1}
Chloramphenicol	34 mg mL^{-1}	100% distilled ethanol	34 μg mL^{-1}
Neomycin	50 mg mL^{-1}	Sterile Distilled water	10 or 25 μg mL^{-1}
IPTG	1 M	Sterile Distilled water	1 mM

2.1.2 *Abiotic stresses applied*

For experiments involving growth measurements during stress, three-day-old *Anabaena* cultures were inoculated in fresh BG-11, N^+ medium at 1 μg mL^{-1} chlorophyll *a* (Chl *a*) and incubated under different stresses as detailed below. For assessing the ability of *Anabaena* to survive abiotic stress, three-day-old *Anabaena* cultures were first

concentrated to 10 µg Chl *a* mL^{-1} and then subjected to different stresses. The wild type and recombinant *Anabaena* cultures were exposed to the following different abiotic stresses: (i) mitomycin C stress (0 – 5 µM) 30 min, (ii) UV stress (0 – 1.5 kJ m^{-2}), dose rate 5 J m^{-2} s^{-1}, (iii) 6 kGy ^{60}Co γ-radiation, dose rate 4.5 kGy h^{-1}, and oxidative stress with (iv) 0-0.5 mM H_2O_2, and (v) 0.2 µM methyl viologen (MV), (vi) C-starvation with 5 µM DCMU [3-(3,4-dichlorophenyl)-1,1-dimethylurea], and heavy metal stress with (vii) 0 -20 µM $CdCl_2$ and (viii) 0 – 50 mM $Na_2HAsO_4.7H_2O$. Desiccation stress was carried out by filtering the *Anabaena* cultures onto filter paper under vacuum, followed by placing the filter papers either in desiccators for 6 days for desiccation or in humid chamber for the corresponding control. After 6 days, *Anabaena* cells were scraped from the membrane and used for further analysis. In case of DNA damaging stresses the cultures were concentrated to 10 µg Chl *a* mL^{-1} before exposure to the stresses while in case of others the cultures were inoculated at 1 µg Chl *a* mL^{-1} in fresh BG-11 medium and then subjected to different stresses. For post-irradiation and post-desiccation recovery, the corresponding cultures were washed and inoculated in fresh medium at 1 µg Chl *a* mL^{-1} followed by incubation under shaking with 125 rpm with continuous illumination at 27 °C over a period of 8 days.

2.1.3 *Measurement of growth and cell survival*

Growth was measured in terms of Chl *a* content, which was estimated in methanolic extracts as described earlier (*152*). Survival after stress was measured in terms of colony forming units (cfu) on BG-11, N^+, Neo_{25} agar plates after 10 days of incubation under illumination at room temperature (RT). All experiments were performed as three independent events, taking duplicates each time, and the data

included in various figures is represented by the mean value and standard deviation of all these independent experiments put together.

2.1.4 *Preparation of competent cells of E. coli*

Overnight grown *E. coli* cells were inoculated in fresh LB medium at a dilution of 1:100 and allowed to grow at 37 °C with shaking at 140 rpm for up to 2 h or till the culture reached an OD_{600} mL^{-1} of 0.8. The culture was chilled on ice and then centrifuged at 3000 rpm for 5 min at 4 °C. Supernatant was discarded and the cell pellet resuspended gently in solution A (10 mM MOPS, pH 7, 10 mM RbCl) in half the initial culture volume and kept on ice for 10 min, followed by centrifugation at 3000 rpm for 7 min at 4 °C. The pellet was then resuspended gently in solution B (100 mM MOPS, pH 6.5, 50 mM $CaCl_2$, 10 mM RbCl) to the same volume as of solution A, kept on ice for 45 min, followed by centrifugation at 3000 rpm for 7 min at 4 °C. Pellet was resuspended in solution B containing 20 % glycerol to the $1/10^{th}$ to the culture volume. The cells were either used immediately for transformations or snap frozen in liquid nitrogen and stored in -70 °C for future use.

2.1.5 *Transformation of DNA into E. coli*

DNA (<20 ng for intact plasmid and <500 ng for ligation mix) to be transformed was mixed with competent *E. coli* cells and incubated on ice for 1 h followed by a brief heat shock at 42 °C for 90 sec. Cells were then quickly chilled on ice followed by the addition of 750-800 µL LB medium and cells allowed to recover at 37 °C with shaking at 140 rpm for 1 h. In case of plasmid DNA transformation, transformants were selected by plating 100 µL of the recovered cells whereas for transformation of ligation mix, the

whole recovered culture was harvested by centrifugation, resuspended in 100 µL of LB and plated on LB agar plates containing appropriate antibiotics. The different *E. coli* strains harbouring various plasmids are listed in **Appendix A**.

2.1.6 *Electroporation of plasmid DNA into Anabaena 7120*

Three-day-old culture of *Anabaena* 7120 **(Table 2.1)** was concentrated to 10 µg Chl*a* density mL^{-1} and washed three to four times with sterile stage II water. Recombinant Plasmid DNA (~1 µg) was added to 200 µL of the concentrated *Anabaena* 7120 cells (10 µg Chl*a* mL^{-1}) and incubated under illumination for 1 h at room temperature, followed by electroporation at 1200 V with 5 msec pulse. BG-11, N^+ liquid media of double strength (2X) was added to the electroporated cells followed by incubation with continuous illumination for 18 h at 25 ± 2 °C. The cultures were spread on small areas on membranes placed on BG-11, N^+, Neo_{25} agar plates and kept under illumination. The membrane was transferred to fresh plate every 3-4 days till the spread became clear and distinct colonies appeared. These colonies were inoculated in 100 µL BG-11, Neo_{10} medium and grown under stationary conditions and continuous illumination. The volume of the media was doubled every three days.

2.1.7 *Triparental conjugation in Anabaena 7120*

Introduction of plasmid DNA into *Anabaena* 7120 by triparental conjugation was achieved by following a 3-step procedure. In the first step, *E. coli* HB101 strain harbouring conjugal plasmid (pRL443) (*153*) was conjugated with HB101 strain carrying pRL623 plasmid (*153*) which codes for *Ava*I, *Ava*II and *Ava*III methylase, resulting in the generation of Cb^R, Cm^R *E. coli* HB101 strain harbouring both pRL443 and pRL623

plasmids **(Table 2.3)**. This strain was then conjugated with HB101 strain harbouring the cargo plasmid [pAM1956 (*154*)/pRL271 (*145*)] **(Table 2.3)** containing the gene of interest. The resultant Cb^R, Cm^R, Neo^R recombinant *E. coli* HB101 strain harbouring all three plasmids was then conjugated with 3-day-old *Anabaena* 7120.

Table 2.3: List of plasmids used in this study

Plasmid	**Characteristics**	**Source**
pBluescript (pBS)	Cb^r, multicopy plasmid with MCS and ColE1 origin of replication	Stratagene
pET16b	Cb^r, Expression vector with N-terminal His-tag	Novagen
pFPN	Cb^r, Kan^r, Integrative expression vector	(*146*)
pAM1956	Kan^r, promoterless vector with g*fpmut2* reporter gene	(*154*)
pBS*nptII*C	*nptII*C gene cloned in pBS plasmid at *Hinc*II and *Sma*I restriction site	(*150*)
pRL271	Cb^r, Cm^r, Er^r, cyanobacterial non replicating vector containing *sacB*	*(145)*
pRL443	Cb^r, Shuttle vector suitable for mobilizing plasmids to *Anabaena* sp. strain PCC 7120	(*153*)
pRL623	Cm^r, encodes methylases *Ava*IM, *Eco*47IIM, and *Ecot*22IM that target sites subject to restriction in *Anabaena* sp. strain PCC 7120	(*153*)

For conjugation of *E. coli* strains, mid-log phase *E. coli* cultures were washed with LB medium without any antibiotic to wash off all antibiotic used during growth and then mixed in a centrifuge tube in 1:1 ratio followed by incubation at 37 °C for 3 h to allow conjugation. After 3 h, 100 µL of these cells were spread on LB-agar plate containing appropriate antibiotic and incubated overnight at 37 °C. Recombinant colonies obtained were inoculated in LB medium for growth with appropriate antibiotics.

For conjugation between *E. coli* as donor and *Anabaena* 7120 strains as a recipient, mid-log phase cultures (~3 h for *E. coli* and 3 days for *Anabaena*) were washed with respective media without any antibiotic and then mixed in a glass tube in 1:1 ratio followed by incubation at 27 °C for at least 3 h under constant illumination to allow conjugation. After that the settled cells were spread on nitrocellulose membranes placed on BG-11, N^+ agar plate without any antibiotic and incubated O/N at 27 °C under illumination. After 24 h, the membranes were transferred on to fresh BG-11, N^+, Neo_{25} agar plates to facilitate selection of recombinant *Anabaena* colonies. The membranes were transferred repeatedly on to fresh BG-11, N^+, Neo_{25} agar plate after every three days until isolated colonies appeared on the membrane. In case of pRL*lexA*mut where the aim was to generate *lexA* knockout, the colonies thus obtained were further screened out for the double crossover event by placing the membrane having colonies on to BG-11, N^+, Neo_{25} agar plate containing 5% sucrose as described earlier (*145*). Recombinant *Anabaena* colonies were then inoculated in liquid medium for further analysis. Recombinant *Anabaena* strains thus obtained were sub-cultured on every third day to achieve faster and better segregation and subsequently the strains were maintained in liquid media containing appropriate antibiotics and were sub-cultured every week.

2.1.8 *Bright-Field and Fluorescence Microscopy*

Microscopy was carried out using either 40 X objective (600 X magnification) or 100 X objective with oil emersion (1500 X magnification). Excitation wavelength of 489 nm (λ_{ex} 489) and emission wavelength of 509 nM (λ_{em} 509) were used for fluorescence microscopy. Microscopy was done using Carl Zeiss Axioscop 40 microscope and images were captured with the help of charge-coupled device (CCD) axiocam MRc (Zeiss) camera and image stored on to the computer.

2.1.9 *Nitroblue tetrazolium (NBT) staining of Anabaena cells*

The NBT staining procedure was followed with modification to that of used for plant leaves (*155*). Three-day-old cultures of recombinant *Anabaena* strains were inoculated in fresh BG-11, N^+ medium at 1 μg Chl *a* mL^{-1} in the absence or presence of 0.2 mM H_2O_2. After 90 min of incubation under illumination at room temperature, 0.01 mM NBT solution (prepared in 50 mM Tris–HCl, pH 8.0) was added to 200 μL *Anabaena* cells, treated or untreated with H_2O_2. The cells were then kept under light for 20 min followed by visualisation under the microscope (Carl Zeiss Axioscope 40, Germany) in bright light field at 1500X magnification with oil-immersion.

2.2 Isolation of Nucleic Acids, Electrophoretic Separation and Cloning

2.2.1 *Isolation of plasmid DNA*

Isolation of plasmid DNA was carried out using Qiagen Mini Prep Kit (Germany) as per the manufacturer's protocol by using the alkaline lysis method. The plasmid DNA was stored in elution buffer provided with the kit.

Table 2.4 List of reagents used for nucleic acid purification

Reagent	Composition
$T_{10}E_1$	10 mM Tris-HCl, 1 mM EDTA, pH 8.0
$T_{10}E_{25}$	10 mM Tris-HCl, 25 mM EDTA, pH 8.0
TES	10 mM Tris-HCl, 25 mM EDTA, pH 8.0, 150 mM NaCl
1 X TBE	8.9 mM Trizma base, 8.9 mM Boric Acid, 0.2 mM EDTA
10X DNA Loading Dye	0.4% bromophenol blue, 0.4% Xylene cyanol, 50% Glycerol
2X RNA Loading Dye (10 ml)	64% Formamide (6.4 ml), 4.4% Formaldehyde (1.2 ml of 37% formaldehyde), 2X MOPS (2 ml of 10X), 8% Sucrose (0.4 ml of 50% sucrose solution), 0.2% Bromophenol Blue, 0.2% Xylene Cyanol

2.2.2 *Isolation of chromosomal DNA*

Ten mL culture of *Anabaena* 7120 having chl*a* density 10 μg mL^{-1} was centrifuged at 6000 rpm for 10 min at room temperature and the pellet was washed with 10 mL of 1 M NaCl followed by two washes with 20 mL of Tris EDTA Saline (TES) (10 mM Tris, pH 8.0, 20 mM EDTA and 150 mM NaCl). At each step centrifugation was carried out at 6000 rpm for 10 min and finally the pellet was resuspended in 10 mL of $T_{10}E_{25}$ buffer **(Table 2.4)**. The cells were lysed by incubating with 1 mL of lysozyme solution (2 mg mL^{-1} in $T_{10}E_1$ buffer) **(Table 2.4)** under gentle shaking conditions at 37°C for at least 1 h in order to break the filaments to 2-3 cell stage, which was observed under microscope. Post-lysis, freshly prepared 1.2 mL mix of sarcosyl/Proteinase K solution (130 mg lauryl sarcosine in 0.5 mL of 20 mg mL^{-1} Proteinase K solution in a total volume of 1.2 mL) was added and incubated further for 1 h at 37 °C. Equal volume of phenol equilibrated with 100 mM Tris-HCl, pH 8.0 was added to the solution, mixed gently and centrifuged at 5000 rpm. The aqueous layer was further purified with phenol:

chloroform mix followed by chloroform and isoamyl alcohol solution (24:1). Chromosomal DNA was spooled after gentle mixing of aqueous layer with $1/10^{th}$ volume of 3 M sodium acetate, pH 5.2 and 2 volumes of chilled absolute ethanol distilled, which was then washed with 70% ethanol, air dried, dissolved in $T_{10}E_1$ and stored in -20 °C in small aliquots.

2.2.3 *Isolation of RNA*

Total RNA was extracted from logarithmic phase cultures of wild type and recombinant *Anabaena* strains (~ 50 µg chl *a*) using MN RNA isolation Kit as per the manufacturer's manual. The RNA was quantified using Nanodrop (NANODROP-2000C, Thermo Scientific) at 260 nm. RNA sample was mixed with appropriate amount of 2X RNA loading dye **(Table 2.4)** and loaded on to the gel and resolved electrophoretically on 1.2% agarose gel, containing 0.05 µg mL^{-1} ethidium bromide prepared in 1X Tris-Borate-EDTA (TBE) buffer at 80 V. RNA was visualised using UV transilluminator (Vilber Lourmat, France).

2.2.4 *Restriction digestion and electrophoretic separation of DNA*

DNA was digested with the specified restriction endonucleases (2 U/µg of DNA) using compatible buffer at 37 °C for 3-4 h followed by electrophoretic separation on agarose gel. Digestion of DNA was checked by loading an aliquot mixed with 10X DNA loading dye **(Table 2.4)** on to 0.8-1.5% agarose gels (Sisco Research Laboratories, India) prepared in 1X Tris-Borate-EDTA (TBE) buffer and containing 0.05 µg mL^{-1} ethidium bromide. Electrophoresis was carried out at 80 V and DNA fragments were visualised using a UV transilluminator (Vilber Lourmat, France). DNA molecular mass ladders (1

kb or 100 bp) (Bangalore Genie, India) were used to calculate the molecular size of the various DNA fragments. DNA excised out from the gel was extracted using Qiagen Gel Extraction kit, Qiagen, Germany as per the manufacturer's protocol. Ligation of the restriction digested insert DNA with suitably digested vector DNA was carried out by taking 3-fold molar excess of the insert DNA in case of sticky end ligation and 5-fold excess of insert DNA in case of blunt end ligation, 1 U of T4 DNA ligase (Rapid DNA ligation Kit, Fermentas, USA) along with ligase buffer and incubated at 25 °C for 1 h.

2.3 PCR Amplification

2.3.1 *PCR amplification of DNA*

PCR amplification of DNA fragments of interest was carried out from genomic DNA using 1 μmole each of the corresponding specific forward and reverse primers, 100 μM dNTPs, and 1U Taq DNA Polymerase in the Taq Buffer provided (BRIT, Jonaki, Hyderabad, India). The mix was subjected to quick-denaturation at 94 °C for 4 min followed by 25-30 cycles of amplification. Each amplification cycle comprised of three consecutive steps (i) denaturation at 94 °C for 45 sec, (ii) annealing at 50 °C-68 °C (depending on the T_m of the primers used) for 45 sec to 1 min, and (iii) extension at 72 °C for varying time depending on the length of fragment to be amplified (approximately 1 min extension was used for amplifying 1 kb DNA fragment). Final extension was carried out at 72 °C for 10 min, followed by storage at 4 °C.

2.3.2 *Generation of point or deletion mutations using PCR-based approach*

For generating point or deletion mutation, site directed mutagenesis approach was adopted in which two step-PCR was carried out. A total of four primers were used, one each corresponding to the extreme end (of the complete DNA fragment to be cloned) forward and reverse primer, and the other set in the middle having the desired mutation. Each amplicon was generated using one extreme end primer and one primer with the desired mutation. The two amplicons generated for a single mutation were mixed and used as template for the amplification of complete gene sequence with desired mutation using extreme end primers. The single complete DNA fragment with the desired mutations was checked for the presence of mutation by sequencing prior to further use.

2.3.3 *Reverse Transcriptase (RT)-PCR*

Isolated RNA was first checked for the absence of DNA contamination by carrying out PCR in the presence of dNTPs, primers and Taq DNA Polymerase. In case amplification was observed, then the RNA was treated with DNaseI and the Taq-based PCR repeated again till no amplification was observed. This confirmed that there was no DNA contamination in the isolated RNA. RT-PCR was carried out using total RNA (10-20 ng) as a template with gene specific primers and one step RT-PCR reaction enzyme mix (Qiagen, USA). The master mix was incubated first for reverse transcription at 50 °C for 30 min which led to generation of cDNA, followed by 30 cycles of normal PCR. Amplification of 16s rDNA was taken as internal control for each sample.

2.4 Overexpression, Purification and Estimation of Proteins

2.4.1 *Protein extraction under denaturing conditions and estimation of protein content*

Reagents used for protein related work have been listed in **Table 2.5**.

Table 2.5: List of reagents used for protein work

Reagent	Composition
Gel Running Buffer	0.3% Trizma Base, 1.44% Glycine, 0.1% SDS
Transfer Buffer	0.125 M Trizma Base, 0.192 M Glycine
Tris Saline Buffer	0.05 M Trizma Base, 0.9% NaCl
Reaction Buffer	0.1 M Trizma Base, 0.1 M NaCl
Ni-NTA	Nickel-Nitrilo tri-acetate (Ni-NTA) slurry Qiagen, Germany
Ni-NTA binding buffer	0.1 M Tris pH 8, 0.3 M NaCl, 10 mM Imidazole
Ni-NTA wash buffer	Ni-NTA binding buffer containing varying concentrations of imidazole from 50 mM to 200 mM with 10 mM increments
1X Blocking Reagent	1% Blocking Reagent (Roche, Germany) in 1X MaNa Buffer
NBT/BCIP	NBT/BCIP (Roche, Germany) diluted 1:50 in 1X Reaction Buffer
Dialysis and Storage buffer	0.1 M Tris pH 7.5, 0.2 M NaCl, 5% Glycerol
Maleic Acid (MaNa) Buffer	0.1 M Maleic acid, 0.15 M NaCl
Coomassie Brilliant Blue (CBB) G-250	0.2% Coomassie Brilliant Blue G 250, 20% Glacial acetic acid, 40% Methanol
Destaining Solution-I	10% Glacial acetic acid, 50% Methanol
Destaining Solution-II	10% Glacial acetic acid, 10% Methanol, 2 % Glycerol
5X Laemmli's Buffer	0.5 M Tris-HCl, pH 6.8, 12.5% (w/v) 2-mercaptoethanol, 5% (w/v) SDS, 50% Glycerol, 0.1% Bromophenol Blue (BPB), 0.19% EGTA, 0.5% Sodium azide, 0.05% PMSF

Total protein from *E. coli* or *Anabaena* cells were extracted by resuspending the cell pellet in stage II water followed by the addition of equal volume of 2X Laemmli's buffer (*156*), mixing and boiling for 10 min in a water bath and then centrifugation at 10000 rpm for 10 min to remove unbroken cells and cell debris. The protein content was estimated in the supernatant using Lowry's method (*157*). In brief, the supernatant containing extracted protein (5-10 µL) was diluted to a final volume of 500 µL with distilled water followed by the addition of 50 µL of 10% deoxycholate (DOC), mixed and incubated at 25°C for 10 min. To this, 50 µL of 78% tri-chloroacetic acid (TCA) was added to precipitate the solubilised protein, followed by centrifugation at 10000 rpm for 10 min. The pellet was resuspended in 500 µL of distilled water, followed by the addition of 500 µL of Lowry's Reagent and incubated for 15 min at room temperature, followed by the addition of 250 µL of Folin-Ciocalteau's reagent and further incubation at room temperature in dark for 30 min. The intensity of the colour was estimated spectrophotometrically at 750 nm. Different concentrations of Bovine Serum Albumin (BSA)(Sigma Aldrich, USA) were used as controls and to plot a standard graph. From the standard graph, the protein concentrations of unknown samples were estimated.

The protein extracts were electrophoretically resolved on 12-15% SDS-polyacrylamide gels at 100 V till the dye front reached the end of the gel. The protein profile was visualized by staining the gel with Coomassie Brilliant Blue (CBB, G-250) staining solution for 30 min followed by destaining with Destaining Solution (DS)-I for 10-15 min and DS-II for 1 h to overnight.

2.4.2 *Overexpression and Purification of recombinant proteins*

For overexpression of proteins, the corresponding genes cloned in pET16b expression vector (**Table 2.3**) and introduced in *E. coli* BL21(*pLysS*)(DE3) (**Table 2.1**).

The corresponding overnight grown *E. coli* cultures were inoculated in LB Cm_{34} Cb_{100} using 1% inoculum and allowed to grow in orbital shaking at 140 rpm at 37 °C. Once culture reached an OD_{600} mL^{-1} of 0.6 to 0.8, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to these cultures to induce overexpression of protein of interest and continued to grow for the duration required (maximum of 3 h). Cells were harvested by centrifugation at 5000 rpm for 5 min, protein extracted, electrophoretically separated by SDS-PAGE, and visualised by staining with coomassie.

For purification of the overexpressed proteins using Ni-NTA chromatography column (Qiagen), the cell pellet, containing soluble protein, was resuspended in one-tenth volume binding buffer (10 mM imidazole, 300 mM NaCl, 100 mM Tris-HCl, pH 8.0), and lysed by sonication on ice using Branson Digital Sonifier at 25% amplitude, pulse ON 1 sec, pulse OFF 2 sec, for 5 min. The cell lysate was centrifuged at 14000 rpm for 45 min at 4 °C. The clear supernatant was allowed to bind to Ni-NTA resin (about 1 mL Ni-NTA slurry for 400 mL induced culture volume) at 4 °C overnight with constant shaking at 40 rpm on rocker. The column was washed with wash buffer (300 mM NaCl, 100 mM Tris-HCl, pH 8.0) with increasing concentrations of imidazole (20 – 150 mM). The protein was eluted in buffer containing 250 mM imidazole and finally with 1 M imidazole to remove any bound protein which was not eluted with 250 mM imidazole. All eluates were collected for analysis by SDS-PAGE. After the SDS-PAGE analysis the aliquots containing purified protein were transferred to a dialysis membrane and dialysed against dialysis buffer (200 mM NaCl, 100 mM Tris-HCl, pH 7.5, 5% glycerol) overnight at 4 °C. Dialysed protein was aliquoted and stored at -70°C till further use.

For overexpressed proteins which formed inclusion bodies, the corresponding cell pellets were lysed in Lysis Buffer containing urea (8 M urea, 10 mM Tris-HCl, 100 mM

$NaH_2PO_4.2H_2O$, pH 8.0). The lysate was centrifuged at 14000 rpm for 45 min at 25°C and the clear supernatant was mixed with Ni-NTA matrix and allowed to bind overnight at rocker at room temperature. Thereafter, it was loaded on to the column, given multiple washes with wash buffer (8 M urea, 10 mM Tris-HCl, 100 mM $NaH_2PO_4.2H_2O$, pH 6.3) and eluted in Elution Buffer (8 M urea, 10 mM Tris-HCl, 100 mM $NaH_2PO_4.2H_2O$, pH 4.5). All fractions were visualised as described above and the fractions having pure proteins were pooled and allowed to renature by dialysis in a step-wise manner decreasing the urea concentration in the dialysis buffer by 2 M every 2 h. Renaturation was monitored by fluorescence analysis.

2.4.3 *Western blotting and immuno-detection*

Post-electrophoretic separation of proteins by SDS-polyacrylamide gel, it was equilibrated in transfer buffer (25 mM Tris, 192 mM Glycine, pH 8.0) for 10 min at room temperature and electroblotted on to the nitrocellulose membrane at constant current of 400 mA for 2-4 h. Blot was then washed and stained with Ponceau in order to confirm the transfer of proteins on the membrane and check for equal protein loading. The blot was incubated in 1X blocking reagent, prepared in MaNa buffer (100 mM Maleic acid, 150 mM NaCl, pH 7.4) for 1 h and probed with specific polyclonal primary antibody raised in rabbit, appropriately diluted in 1X blocking reagent in MaNa buffer for 1 h to overnight. This was followed by three washes of 20 min each with MaNa Buffer and incubation with anti-rabbit IgG secondary antibodies conjugated with alkaline phosphatase (Roche Diagnostics, GmBH) in 1:3000 dilution for 1 h. The blot was again washed with MaNa buffer three times for 20 min each, rinsed in reaction buffer (100 mM Tris, 100 mM NaCl, pH 9.5) and colour development carried out using NBT/BCIP.

2.4.4 *In-gel SOD activity assay (Zymogram)*

Total protein was extracted from recombinant *Anabaena* cultures in 100 mM Tris-HCl pH 8.0 by freeze-thaw method under native conditions. Protein was mixed with loading buffer without reducing agent and loaded on to 12% native polyacrylamide gel and allowed to run till dye front reached the bottom at 4 °C. After the gel run, staining for SOD activity was carried out as per standard earlier described protocol for SOD zymogram (*158*). In brief, the gel was soaked in 1.225 mM NBT solution (prepared in 50 mM Tris–HCl, pH 8.0) for 45 min at room temperature which was followed by incubation of gel in Tris buffer (50 mM Tris-HCl, pH 8.0) containing 0.028 mM riboflavin and 28 mM TEMED for 30 min at room temperature. The gel was then exposed to light till clear and contrast white bands were observed.

2.4.5 *Two-Dimensional proteomic analysis of recombinant Anabaena strains*

For preparation of protein samples for 2-D electrophoresis, 3-day-old recombinant *Anabaena* cells [An*lexA*$^{+}$ and AnpAM (*150*)] (~200 µg Chl*a*) were harvested by centrifugation at 5000 rpm for 5 minutes. The cell pellet was resuspended in lysis buffer (1 mM Tris-HCl, pH 8.0 and 1 mM PMSF), four volumes of the pellet by vortexing. Lysis of cells was achieved by alternative cycles of freezing in liquid N_2 and thawing at 37 °C at least four times which was followed by centrifugation at 14000 rpm for 1 h at 4 °C, and the protein content in the cell free supernatant was estimated by Lowry reagent (Sigma, India). Protein extract was subjected to simultaneous DNase and RNase treatment (10 µg mL^{-1} each) for 1 h on ice to remove DNA and RNA

contamination from the protein samples as described earlier (*159*). The protein (~1 mg) was concentrated to 10 µL and solubilised in 80 µL rehydration buffer (8 M urea, 1 M thiourea, 4% CHAPS, 150 mM DTT, 2% IPG buffer and traces of bromophenol blue) for 1 h at room temperature as described earlier (*159*), followed by storage at -70 °C till further use.

IPG strips with non linear pH (3-10), 17 in length cm were rehydrated by layering 320 µL of rehydration buffer on to the strips paced in rehydration plate for atleast 15 h at room temperature. The solubilised protein was loaded in 100 µL cup placed at the cathode on to rehydrated IPG strips placed in IEF tray and IEF carried out with Protean Isoelectric Focussing Cell (Bio-Rad, India) at 20 °C. The second dimension electrophoresis was carried out by 14% SDS-PAG wherein IPG strips after IEF run were pasted on to the top of 14% polyacrylamide gel and then the proteins were resolved on the basis of their molecular mass. Each experiment was repeated three times, corresponding to three biological replicates, for each strain.

Gels were imaged using Dyversity 6-gel imager (Syngene, UK) with the help of Gene Snap software (Syngene, UK). PD-Quest (version 8.1.0, Bio-Rad) analysis was carried out. This resulted in the generation of a first level match set from the three biological replicates with a minimum correlation coefficient of 0.7. Detection and matching of spots on the replicate gels were carried out in an automated mode which was followed by manual editing. For the normalization of spot densities the local regression method was used and statistical analysis (Students t-test) was also performed on various protein spots. The protein spots with p-values ≤ 0.05 were considered to be significantly modulated exhibiting 1.2 fold or greater were considered as induced whereas -1.2 or lesser were considered as repressed in comparative analysis. The protein spots were excised out from the gel, washed first with miliQ water twice followed by destaining

with 50 µL solution of 50 mM NH_4HCO_3 and ACN (1:1). It was then washed with 50 µL ACN and again with 25 µL NH_4HCO_3 and 25 µL of ACN for 5 min each at room temperature and finally with 50 µL ACN at stationary for 5 min, followed by vacuum drying for 2 min. The reduction of protein spots was carried out with 50 µL of 50 mM NH_4HCO_3 containing 10 mM DTT for 45 min at 56 °C, followed by alkylation with 50 µL of 55 mM iodoacetoamide (IA) for 30 min in the dark at room temperature. The protein spots were again washed with 50 mM NH_4HCO_3 and ACN as described above. Tryptic digestion was carried out by the addition of 2 µL trypsin (20 ng/µL) in each tube containing gel particle having protein spot) followed by addition of 3 µL of trypsin buffer (25 mM NH_4HCO_3 and 9% ACN) in the same tube on ice of 30 min. After 30 min of incubation, 2 µL trysin buffer was added again to keep gel particle wet and incubated at 37°C for overnight or 16 h. Peptides were eluted from the gel particle using 0.1% trifluoroacetic acid (TFA) and ACN in three steps. In the first step, peptides were eluted by addition of 25 µL of 0.1% TFA, votexed for 15 min, followed by second elution of peptides by 25 µL of 0.1% TFA and ACN (1:1) after votexing for 15 min and finally with 25 µL of ACN after votexing for 15 min. The eluted peptides (75 µL) were vacuum dried to 7-10µL and stored in -70°C for further analysis. The eluted peptides were co-crystalized with α-cyano-4-hydroxycinnamic acid (TFA) and 30% ACN on a 384-well ground steel target plate (Bruker Daltonics, Germany). The MALDI ToF/ToF (Matrix-assisted laser desorption/ionization-Time of flight) UltraFlexIII Mass spectrometer was externally calibrated as per manufacturer's protocol. Analysis and acquisition of spectra using Flex Analysis software 3.0 (Bruker Daltonics) was carried out as described earlier (*159*).

The mass spectra corresponding to each spots were imported into the database search engine (BioTools v 3.1 connected to Mascot, version 2.2.04, Matrix Science).

Mascot search was carried out from NCBI non-redundant database (released Jan 2012 or later). Criteria chosen for homology search were (i) number of missed cleavages permitted 1 or 2, (ii) fixed modifications such as carbamidomethyl on cystine, (iii) variable modification of oxidation on methionine residue, (iv) peptide tolerance 100 ppm, (v) enzyme used trypsin, and (vi) peptide change setting at +1. The best score in each Mascot with a match with *Anabaena* 7120 protein was accepted as a successful identification. Protein identification was considered significant if at least two of the following three criteria were fulfilled: (i) Mascot score >60 ($p<0.05$), (ii) minimum match of 5 peptides, (iii) sequence coverage >20%.

2.4.6 *Electrophoretic Mobility Shift Assay (EMSA)*

DNA protein interaction as studied by electrophoretic mobility shift assay (EMSA). Linear DNA fragments (~50 ng) were incubated with specified concentrations of purified *Anabaena* LexA protein in DNA binding buffer (20 mM Tris-HCl pH 8.0, 1 mM $MgCl_2$, 100 mM KCl, 8 mM DTT, 80 μgmL^{-1} BSA and 4% sucrose) at 37 °C for 30 min. EMSA reaction was stopped by adding 2µL of 10X DNA loading dye. The reaction mixture was then resolved by 8-10% native PAGE, followed by staining of gel with SYBR Dye I and visualization under UV-transilluminator. ImageJ software was used to determine the amount of free and bound DNA densitometrically (*160*). Three independent experiments for each set of DNA-protein interaction were carried out for plotting the affinity graphs. The concentration of protein at which 50% DNA binding was observed was defined as the binding affinity of the protein to that particular DNA fragment.

2.4.7 *LexA autoproteolytic cleavage assay*

For analysis of the autoproteolytic cleavage of wild type and mutant *Anabaena* LexA proteins, the purified native proteins were incubated in Buffer A (100 mM Tris-HCl, specified pH ranging from 6.5-10, 200 mM NaCl, 1 mM EDTA, 12 mM DTT and 5% glycerol) at the specified temperature for up to 24 h. Autoproteolytic cleavage LexA under RecA-dependent condition was carried out by incubating LexA protein with purified RecA from either *Anabaena* 7120 or *E. coli* in Buffer B (20 mM Tris-HCl pH 7.5, 100mM NaCl, 2 mM DTT, 5 mM β-mercaptoethanol, 2 mM $MgCl_2$, 0.5 mM EDTA, 2.5% glycerol, 1 mM Thio ATP, 6 μg mL^{-1} M13 single stranded DNA) as described earlier (*107*) at 37 °C for 1h. Protein cleavage was visualised by 15% SDS-PAGE followed by CBB staining.

2.4.8 *Fluorimetric Analysis of GFP Expression*

Quantification of expression of GFP in *E. coli* cells was carried out using a Jasco FP 1500 spectrofluorimeter. The mid log-phage *E. coli* cells with or without stress were exposed to an excitation wavelength of 490 nm and the emission of fluorescence energy was recorded at wavelength of 510 nm, along with proper controls. The activity of promoter was calculated using GFP arbitrary units per μg Chl *a*.

2.5 Biophysical Techniques

2.5.1 *Size determination*

Molecular mass determination of native purified protein was carried out by gel filtration chromatography using Superdex HR200 column. The column was equilibrated with Tris-NaCl buffer (200 mM NaCl, 100 mM Tris-HCl, pH 7.5 and 5% glycerol) and standard graph obtained using the following proteins: Thioglobulin (669 kDa), Ferritin (440 kDa), Bovine Serum Albumin (67 kDa), Albumin (44.3 kDa), Carbonic Anhydrase (29 kDa). Molecular mass of the *Anabaena* LexA was calculated on the basis of the elution volume using the equation of standard graph.

2.5.2 *Circular Dichroism (CD) Analysis*

CD Spectra for purified protein (1 mg mL^{-1}) was recorded in the far UV wavelength region i.e. 260–200 nm using JASCO spectro-polarimeter (model J-810) equipped with Peltier thermostatic cell holder using 1 mm path-length quartz cuvette at 20 °C. Each spectrum was obtained by averaging three individual scans. A reference spectrum was recorded with storage buffer which was subsequently subtracted from protein sample spectrum. The observed spectrum was measured in ellipticity units (millidegree units). The spectrum was further deconvoluted to calculate the secondary structure elements of the protein using K2D3 web server (http://www.ogic.ca/projects/k2d3/) (*161*).

2.5.3 *Dynamic Light Scattering (DLS)*

Dynamic light scattering measures the intensity of scattered laser light by molecules in solution which is related to the hydrodynamic radius (Rh) of the molecules,

as in the Stokes–Einstein equation. Through DLS, different sizes of particles or aggregation status are recognized. In the current study, the DLS analysis was used to determine the hydrodynamic radii and percent polydispersity of proteins, using the Malvern Zetasizer nano ZS instrument. The protein solution was centrifuged at 14000 rpm for 15 minutes followed by filtration by 0.2 µm milipore filter before the analysis to remove any particulate matter. The analysis was carried out at 25 °C and the data analysis was done with the help Malvern Zetasizer software suite (http://www.malvern.com).

CHAPTER 3

BIOPHYSICAL

AND

BIOCHEMICAL

CHARACTERISATION OF

Anabaena LexA PROTEIN

3.1 INTRODUCTION

LexA is a transcriptional regulator of DNA repair genes in bacteria and is known to down-regulate more than 30 genes involved in SOS response in *E. coli* (*42*). Under normal growth conditions, LexA represses the SOS genes by binding to the LexA-Box [CTG-$(TA)_5$-CAG] (*54*), present in the vicinity of their promoter region. Upon DNA damage, RecA interacts with the generated single stranded DNA, gets activated and induces auto-cleavage of the Gly^{84}-Ala^{85} peptide bond in LexA. *In vivo*, LexA exists both in cleavable and non-cleavable form (*43*). In its non-cleavable form, the catalytic residues Lys^{156} and Ser^{119} are present on the surface of the protein, but upon interaction with activated RecA, they get buried inside coming close to the cleavage site. The hydroxyl group of serine acts as nucleophile which attacks on carbonyl carbon of sessile peptide bond between Ala^{84} and Gly^{85} and the unprotonated form of ε-amino group of Lys^{156} serves to activate the nucleophilic hydroxyl group of Ser^{119} (*107*). The N-terminal domain (NTD) of LexA contains the DNA binding domain, whereas the C-terminal domain (CTD) contains dimerization domain, cleavage site and active site and residues crucial for inducing cleavage, such as Glu^{92}, Ser^{119}, Lys^{156} and Leu^{89}. The cleavage of LexA inhibits its dimerisation, and also its ability to bind to DNA, thereby derepressing the expression of downstream SOS genes (*162*). The proteolytic degradation of the individual N-terminal domain (NTD) and the C-terminal domain after autoproteolytic cleavage by Lon and Clp proteases further ensures complete derepression of the SOS response system (*163-164*). However, when the LexA is bound to DNA, the autoproteolytic cleavage of LexA results in the separation of CTD from the NTD which remains bound to the SOS box but with lower affinity, resulting in incomplete repression (*165*).

Among the different cyanobacterial species, LexA was found to be present in 27 of the 33 cyanobacterial genomes analysed. Of these, the LexA of *Synechocystis* sp. strain PCC6803 and *Gleobacter violaceus* PCC7421 coded for a non-cleavable form of LexA (*43*). The stress-induced expression of the cyanobacterial *lexA* gene was also found to be different from that observed for other bacteria. In *Synechocystis* 6803, the *lexA* gene was not induced in response to DNA damage (*128*), while high light intensity regulated the expression of *lexA* in the marine cyanobacterial strains, *Synechococcus* sp. WH7803 and *Prochlorococcus marinus* PCC9511 (*166-167*). These data suggested a possible deviation of the activity of cyanobacterial LexA from other bacterial LexA proteins which has not been characterized till date. Keeping this in view the LexA protein from the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120, hereafter referred to as *Anabaena* 7120 was characterised biophysically and biochemically.

3.2 RESULTS

3.2.1 *Cloning, overexpression and purification of Anabaena LexA protein using E. coli as a host*

3.2.1.1 CLONING OF *Anabaena lexA* GENE

The ORF of chromosomal *lexA* gene was PCR amplified from the genomic DNA of *Anabaena* 7120 as a 0.6 kb DNA fragment using gene specific primers namely *lexA* forward and *lexA* reverse primer **(Table 3.1)**. The primers were designed based on the available genome sequence of *Anabaena* 7120 in the cyanobacterial database (http://genome.microbedb.jp/cyanobase/GCA_000009705.1) and corresponded to the start and end of the gene.

Table 3.1: Primers used in this study

Primers	Nucleotide Sequence*
lexA-OE-Fwd	5'-GCCGCCATATGGAACGCCTAACAGAA-3'
lexA-OE-Rev	5' -GAAGGATCCTCACATATAACCGCGCCA-3'
lexA-A84S-Fwd	5'-GGAGCGATCGCC**T**CAGGTGGTTTAATAGAA-3'
lexA-A84S-Rev	5'-TTCTATTAAACCACCTG**A**GGCGATCGCTCC-3'
lexA-G85S-Fwd	5'-GGAGCGATCGCCGCA**A**GTGGTTTAATAGAA-3'
lexA-G85S-Rev	5'-TTCTATTAAACCAC**T**TGCGGCGATCGCTCC-3'
lexA-S118A-Fwd	5'-GTAACTGGTGAC**GC**CATGATTGAAGA-3'
lexA-S118A-Rev	5'-TCTTCAATCATG**GC**GTCACCAGTTAC-3'
lexA-K159R-Fwd	5'-GGTAATACATTAA**G**ACGTTTTTATCGA-3'
lexA-K159R-Rev	5'-TCGATAAAAACGT**C**TTAATGTATTACC-3'
lexA-CTD-Fwd	5'-GGCCCATATGGGTAAAGCCCGAACAATTCGAGTTTT-3'
lexA-NTD-Rev	5'-GGCCGGATCCTTAACCTTCAGTCCATTCAATG-3'
lexA-E96Q-Fwd	5'-TCACTGATGCTGTC**C**AGCATATCGACTTT-3'
lexA-E96Q-Rev	5'-AAAGTCGATATGCT**G**GACAGCATCAGTGA-3'
*lexA*ΔGLI-Fwd	5'-GCGATCGCCGCAGGTGAACCATTCACTGATG-3'
*lexA*ΔGLI-Rev	5'-CATCAGTGAATGGTTCACCTGCGGCGATCGC-3'

*The underlined sequence corresponds to NdeI and BamHI restriction sites in the forward and reverse primers respectively. The nucleotides shown in bold and larger font corresponds to the mutated base in the primer used for generating point mutants.

A single amplicon corresponding to 0.6 kb was observed when both the primers were used **(lane F+R, Fig. 3.1A)**. No PCR amplification was observed when only one of the two primers were used, indicating that the observed 0.6 kb amplicon is a product of specific amplification **(Fig. 3.1A)**. The amplicon (0.6 kb) was purified using PCR-purification kit (Qiagen, India) and restriction digested with *Nde*I and *Bam*HI restriction enzymes. The resulting 0.6 kb *Nde*I/*Bam*HI fragment was ligated to the expression vector pET16b **(Table 2.4)**, digested with the same set of restriction enzymes. Upon transformation of the ligation mix into *E. coli* DH5α cells, several colonies were obtained on LB Cb_{100} plates which were screened for the presence of *Anabaena lexA* gene by colony PCR using *lexA* forward and *lexA* reverse primers. Further confirmation of the PCR positive clones was carried out by subjecting the plasmid DNA isolated from these clones to restriction digestion with *Nde*I and *Bam*HI enzymes which resulted in the release of 0.6 kb insert corresponding to *lexA* gene and 5.7 kb vector **(Fig. 3.1B)**. The recombinant plasmid was designated as pET*lexA* **(Table 3.2)**. The cloned *lexA* ORF was sequenced on both the strands and the sequence of *lexA* ORF was submitted to GenBank (Accession no. KF269537) **(Appendix B)**.

3.2.1.2 OVEREXPRESSION AND PURIFICATION OF *Anabaena* LexA PROTEIN

The recombinant plasmid pET*lexA* was transformed into *E. coli* BL-21(p*lysS*)(DE3) cells **(Table 2.1)**, to generate the recombinant strain designated as BL21(pET*lexA*) (**Appendix A**)**.** *Anabaena* LexA protein was overexpressed by inducing the exponentially growing BL-21(pET*lexA*) cells by the addition of 1 mM IPTG and continued incubation at 37 ºC for up to 3 h. A 25 kDa protein was found to be overexpressed within 1 h of induction with IPTG and continued accumulation up to 3 h of induction **(Fig. 3.2A)**.

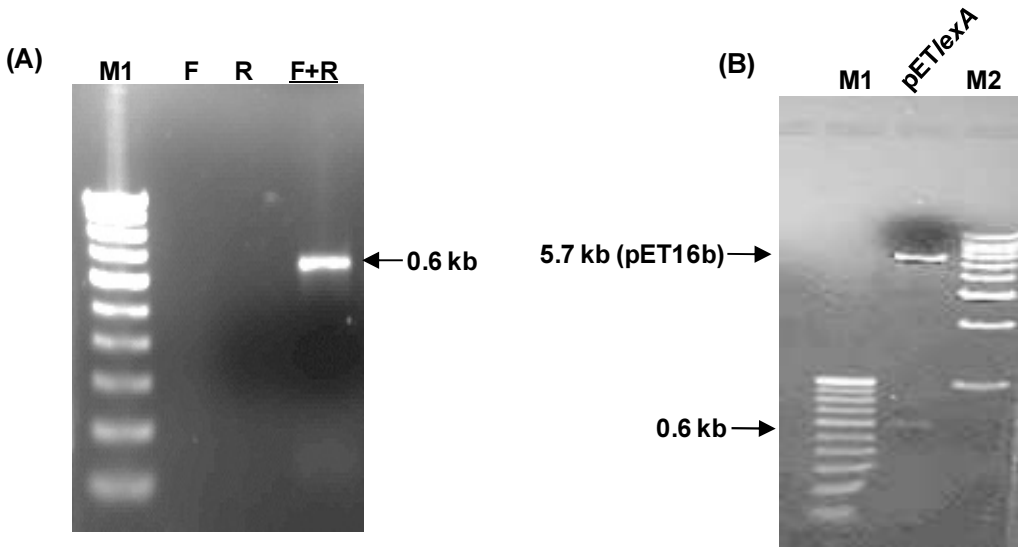


Fig. 3.1 PCR amplification and cloning of *lexA* gene of *Anabaena* 7120. (A) PCR amplification of *lexA* using genomic DNA of *Anabaena* 7120 as a template with gene specific primers. F and R indicate forward and reverse primers respectively. The 0.6 kb amplicon is indicated by an arrow. **(B)** Restriction digestion of pET*lexA* with *Nde*I and *Bam*HI restriction endonucleases. The two fragments corresponding to 5.7 kb pET16b vector and 0.6 kb insert are indicated with arrows. DNA was electrophoretically separated on 0.8% agarose gel and visualised under UV-transilluminator. M1 and M2 indicate 100 bp and 1 kb DNA ladders (Bangalore Genei, India) respectively.

Table 3.2: List of recombinant plasmids generated during this study

Plasmid	Characteristics	Source
pET*lexA*	0.6 kb, *lexA* gene cloned in pET16b at *Nde*I-*Bam*HI restriction sites	This study
pET*lex*A-NTD	0.18 kb, partial *lexA* gene corresponding to first 60 aa[@] cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pET*lexA*-CTD	0.45 kb, partial *lexA* gene corresponding to LexA from the 61st to 201st aa cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pET*lexA*-ΔGLI	0.6 kb *lexA* gene, with deletion of three amino acids (G^{86}LI^{88}) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pET*lexA*-A84S	0.6 kb *lexA* gene (A to S substitution at 84th aa) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pETlexA-G85S	0.6 kb *lexA* gene (G to S substitution at 85th aa) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pETlexA-E96Q	0.6 kb *lexA* gene (E to Q substitution at 96th aa) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pET*lexA*-S118A	0.6 kb *lexA* gene (S to A substitution at 118th aa) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study
pET*lexA*-K159R	0.6 kb *lexA* gene (K to R substitution at 159th aa) cloned in pET16b at *Nde*I-*Xho*I restriction sites	This study

lexA gene in each case has been amplified from genomic DNA of *Anabaena* 7120
@ aa refers to amino acids

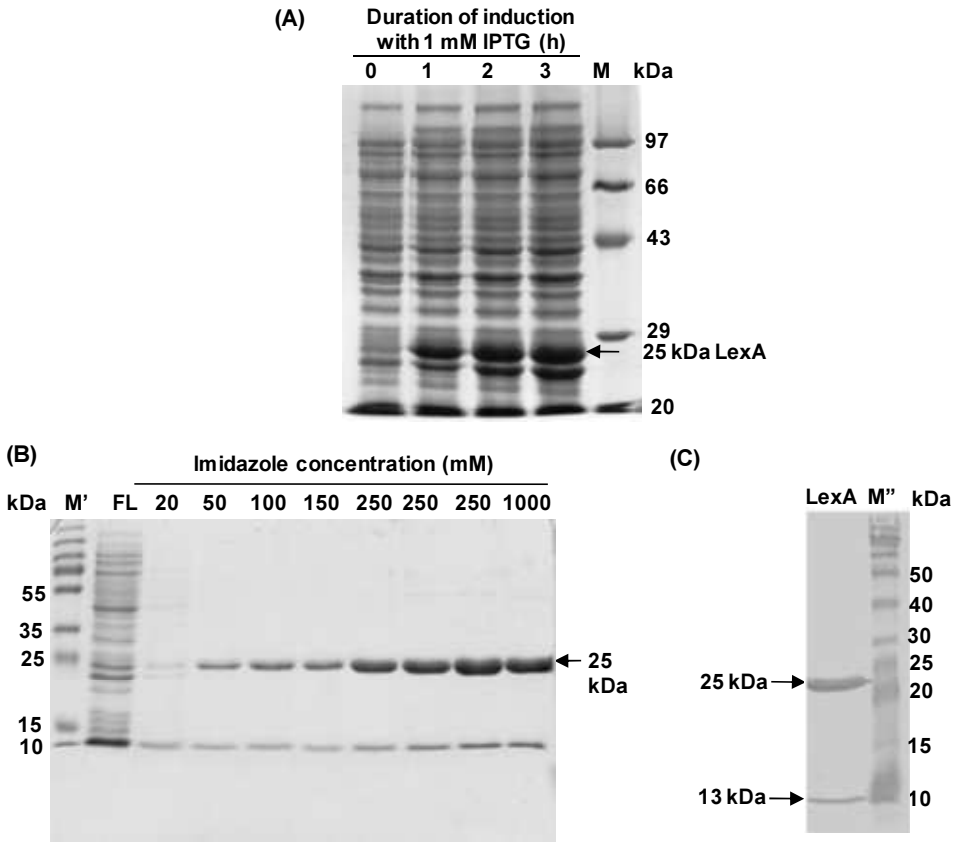


Fig. 3.2 Overexpression and purification of *Anabaena* LexA protein from *E. coli*. (A) Exponential phase (0.6 OD_{600} mL^{-1}) BL-21 cells harbouring pET*lexA* plasmid were induced with 1 mM IPTG for 3 h and samples were collected at the end of every hour. Total proteins from each sample extracted using 5X Laemmli's buffer. **(B) N**ative protein extract in 100 mM Tris-HCl buffer, pH 8.0 from IPTG induced BL21(pET*lexA*) cells were purified using Ni-NTA affinity chromatography with increasing concentration of imidazole (20 – 250 mM) as indicated and finally eluted using 1 M imidazole. **(C)** Western blotting and immunodetection of purified *Anabaena* LexA (AnLexA) protein using anti-AnLexA antibody. Electrophoretic separation of proteins was carried out on 15% SDS - polyacrylamide gel and either stained with Coomassie Brilliant Blue (CBB) dye for visualization or blotted onto nitrocellulose membrane for western blotting. The overexpressed LexA in (A), purified LexA in (B) and cross-reacting LexA proteins in (C) are indicated by arrows. Molecular mass of the protein bands corresponding to the different protein molecular mass standards used (M, M' and M") are indicated.

The overexpressed *Anabaena* LexA protein was purified from the total lysate by Ni-NTA affinity chromatography under native conditions using increasing concentrations of imidazole (20 – 1000 mM) in 100 mM Tris-HCl, pH 8.0, 300 mM NaCl buffer. Pure fractions of the 25 kDa LexA with the exception of co-eluting 13 kDa protein, was obtained at 150 mM or higher concentrations of imidazole **(Fig. 3.2B).** The intensity of the 13 kDa protein increased proportionally with that of the 25 kDa LexA. The purified fractions corresponding to 150 mM to 1 M imidazole were pooled and dialysed against 100 mM Tris-HCl, pH 7.5, 300 mM NaCl buffer to remove imidazole. The purified dialysed LexA protein was used for biophysical and biochemical characterization. For generating polyclonal antibody against *Anabaena* LexA, the 25 kDa protein was cut from the gel and pooled. The anti-AnLexA antibody, thus generated, cross-reacted with both 25 kDa and 13 kDa proteins **(Fig. 3.2C)**, indicating that the co-eluting 13 kDa protein is probably a proteolytically cleaved part of the full length 25 kDa *Anabaena* LexA.

3.2.2 *Biophysical characterization of Anabaena LexA*

3.2.2.1 OLIGOMERIC STATUS

The oligomeric status of purified *Anabaena* LexA protein under native conditions was determined by gel filtration chromatography using Superdex HR200 column. A single major peak was observed in the eluate **(Fig. 3.3A)**. Based on the standard graph and equation Y= - 0.193X + 7.545 for the matrix and column used, the peak corresponded to molecular mass of ~55 kDa **(Fig. 3.3A)**. This corresponds to approx twice the size of LexA monomer (~25 kDa), indicating that under native conditions *Anabaena* LexA exists in a dimeric state.

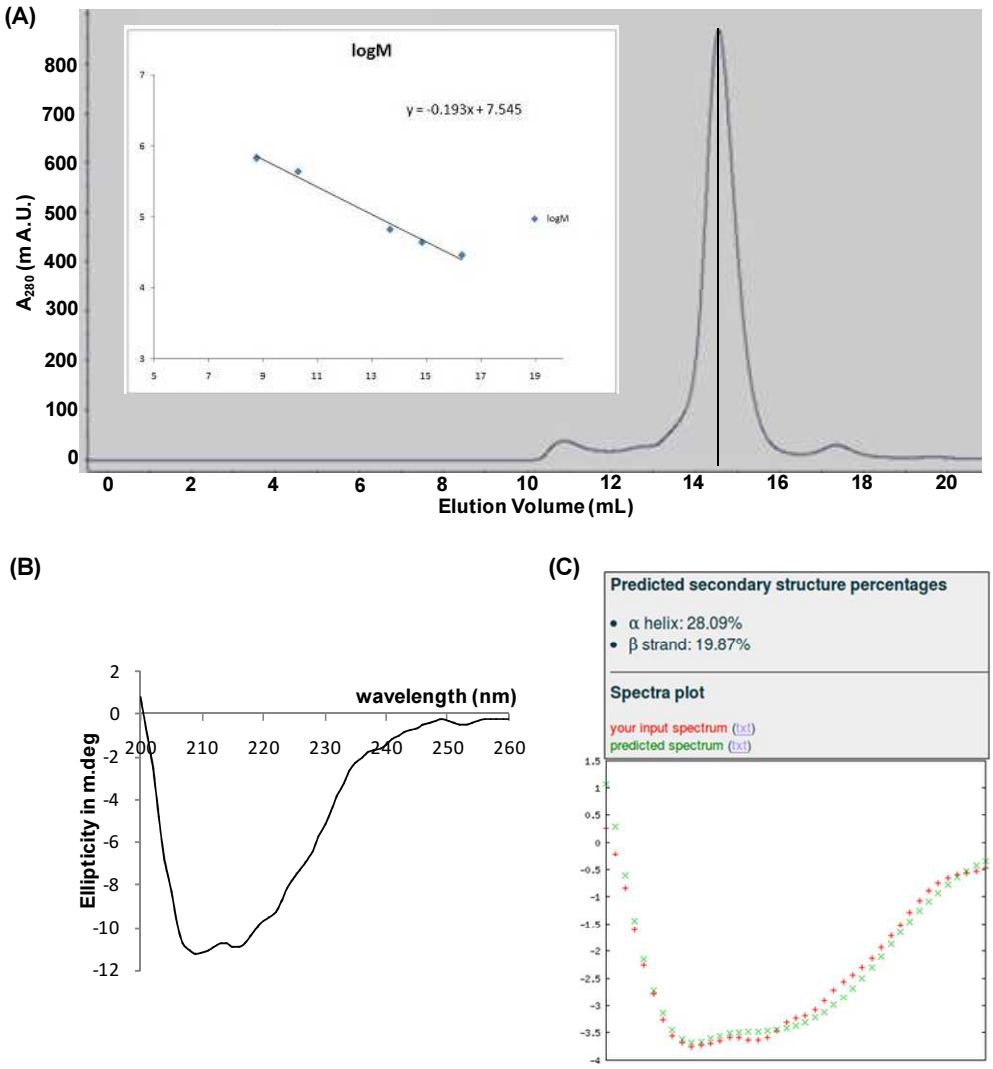


Fig. 3.3 Biophysical characterization of *Anabaena* LexA protein. **(A)** Determination of oligomeric status of LexA. The native purified *Anabaena* LexA protein was subjected to gel filtration using Superdex 200 HR column. The eluted protein peak was detected by measuring absorbance at 280 nm. The absorbance values are given in milli Arbitrary Units (mAUs). The standard graph used for molecular mass determination is shown in the inset. **(B)** Circular Dichroism analysis of purified LexA protein in the wavelength range of 200–260 nm. The CD spectrum is expressed in terms of ellipticity in m deg. **(C)** Estimation of elements of secondary structure of LexA protein using K2D3 web server by deconvoluting the CD spectrum.

3.2.2.2 SECONDARY STRUCTURE DETERMINATION

To determine the elements of the secondary structure of LexA, Circular Dichroism (CD) analysis was carried out using 0.1 mg mL^{-1} LexA in 50 mM Tris-HCl, pH 7.5 and 100 mM NaCl at 25 °C, in the far-UV region. The CD spectrum expressed in terms of ellipticity (m.deg) versus wavelength **(Fig. 3.3B)**, was deconvoluted with K2D3 server (http://www.ogic.ca/projects/k2d3/) (*161*) to estimate the contributing secondary structure elements. The server predicted the LexA protein to have ~28 % α-helices and ~20 % β-sheets, which indicated that the major portion of the protein structure is in the form of loop. The experimental spectrum matched well with the predicted spectrum **(Fig. 3.3C)**, suggesting a reasonable accuracy in the prediction of secondary structure elements.

3.2.2.3 MOLECULAR SIZE DETERMINATION OF LexA

The hydrodynamic size of purified LexA protein (100 μg mL^{-1}) under native conditions in 100 mM Tris-HCl, pH 7.5 and 300 mM NaCl buffer was determined by Dynamic Light Scattering (DLS). The small plateau in the correlation coefficient curve **(Fig. 3.4A)** indicated small hydrodynamic size of the protein. Data analysis revealed a major peak with 99.9% volume corresponding to hydrodynamic molecular diameter of 9.8 ± 3.5 nm **(Fig. 3.4B)**. Presence of a single peak suggests that the protein solution is monodispersive indicative of a single type of protein species in the solution.

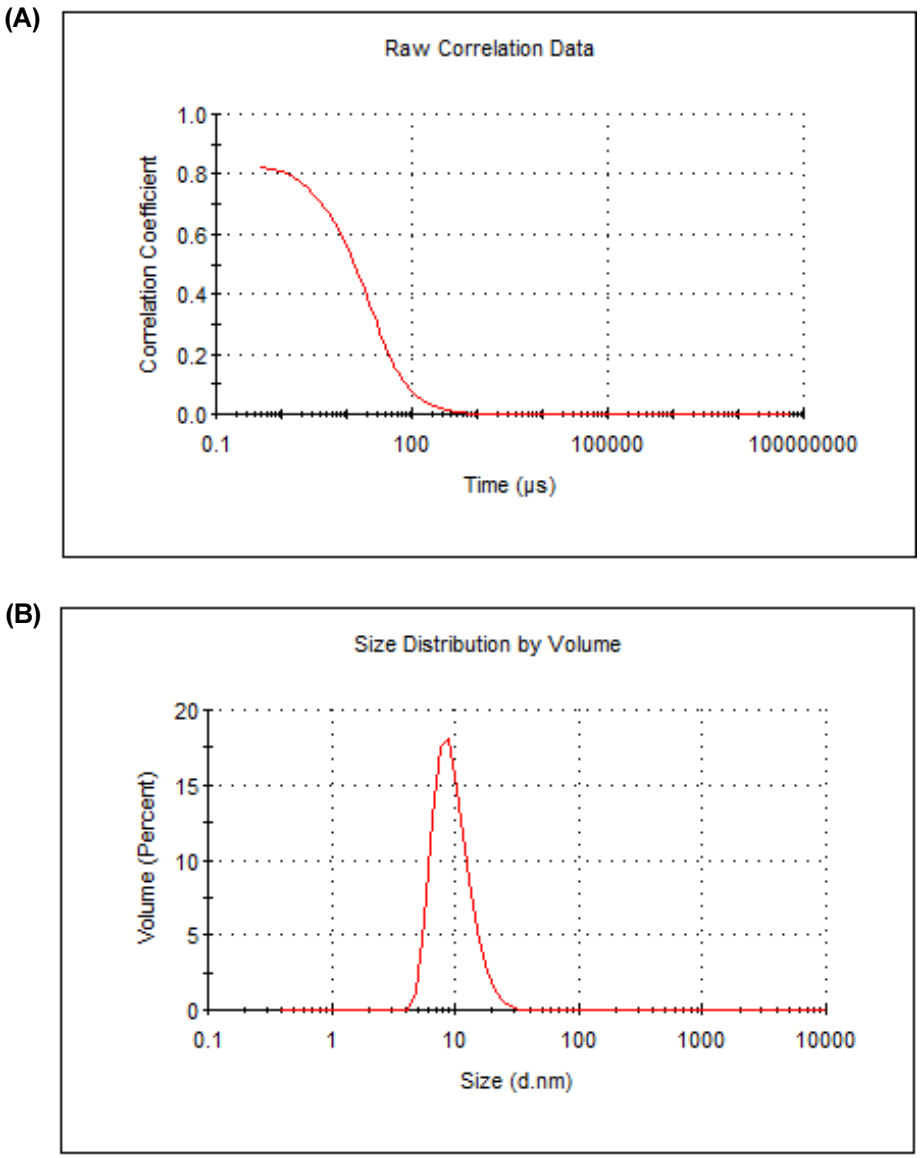


Fig. 3.4 Molecular size determination of LexA by Dynamic Light Scattering (DLS). **(A)** Correlation graph expressing variation of correlation coefficient with size of particles and indicates that result meets the quality criteria. **(B)** Dynamic light scattering profile of LexA. The hydrodynamic diameters for the native LexA protein expressed in terms of percent volume verses size.

3.2.3 *Biochemical characterization of Anabaena LexA protein*

3.2.3.1 BIOINFORMATIC ANALYSIS OF CONSERVED DOMAINS OF *Anabaena* LexA

Conserved domain analysis of *Anabaena* LexA using the CDD software (*168*) revealed the presence of a proteolytic domain, a DNA binding domain and a catalytic site typical of bacterial LexA proteins **(Fig. 3.5)**. This suggested that *Anabaena* LexA may undergo proteolytic cleavage and thus it was speculated that the co-eluting 13 kDa protein, which cross-reacted with anti-AnLexA antibody **(Fig. 3.2B, C)** is in all likelihood the proteolytically cleaved product of the 25 kDa *Anabaena* LexA. Further analysis was carried out to understand the nature of cleavage and the amino acids essential for cleavage.

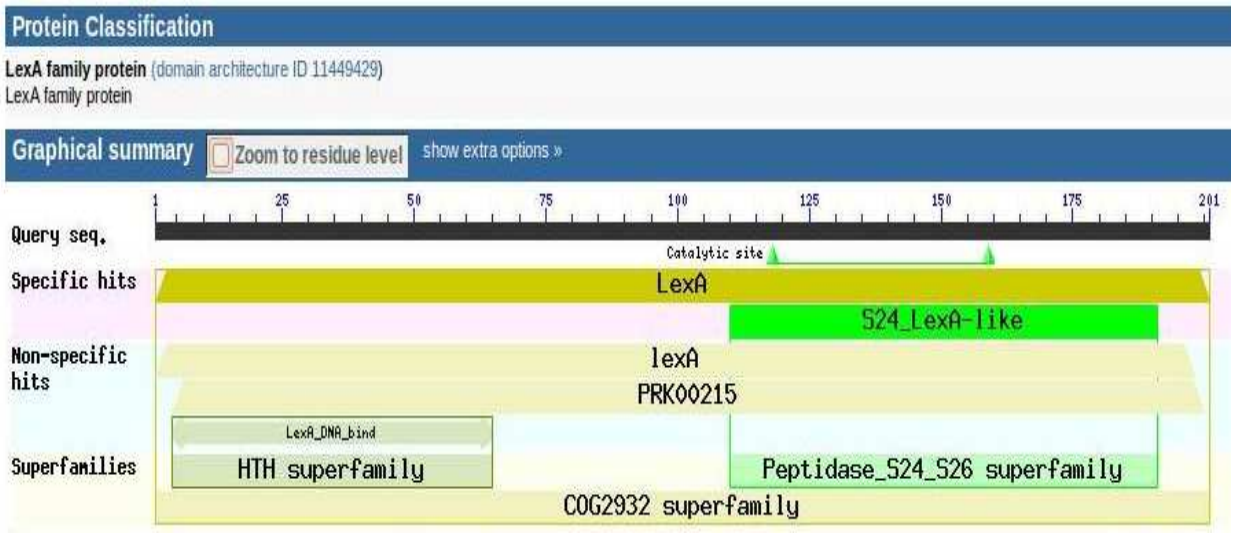


Fig. 3.5 Conserved domain analysis of *Anabaena* LexA protein. CDD software based prediction of domains corresponding to LexA and HTH, peptidase S24-S26 and COG 2932 superfamilies as indicated.

3.2.3.2 AUTOPROTEOLYTIC CLEAVAGE OF LexA AS A FUNCTION OF pH

Ni-NTA purified native *Anabaena* LexA was incubated in Tris-HCl buffer having pH ranging from 6.5 to 10.0 for up to 24 h. At the end of 24 h, in addition to analysing the cleavage, pH of the solution was also measured to confirm that there has been no change in pH during the course of the experiments. The ratio of the levels of the 13 kDa form to that of the 25 kDa form of LexA remained relatively unaltered over a 24 h period at pH 6.5 **(Fig. 3.6A)** and 7.5 **(Fig. 3.6B)** indicating that *Anabaena* LexA did not undergo any further cleavage in the pH range 6.5–7.5 at 37 °C. At pH 8.0, a gradual decrease in the ratio of 25 kDa to 13 kDa from 3.22 to 1.22 was observed over a 24 h period **(Fig. 3.6C)**. Upon increasing the pH further to 8.5 and higher, significant cleavage was observed **(Fig. 3.6 D-F)**. At pH 8.5, the levels of the 13 kDa form was over 2-fold of that of the 25 kDa LexA at the end of 24 h **(Fig. 3.6D)**, while at pH 9.5, 2-fold higher levels of 13 kDa compared to the 25 kDa was observed by 7 h itself **(Fig. 3.6E)**. By the end of the 24 h, the 25 kDa LexA was almost completely cleaved to the 13 kDa form at pH 9.5 **(Fig. 3.6E).** Upon further increase in the alkalinity of the buffer, i.e. at pH 10.0, almost 80% of the 25 kDa LexA was cleaved within 5 h **(Fig. 3.6F)**. Thus, *Anabaena* LexA exhibited autoproteolytic cleavage activity at alkaline pH. Though, the rate of cleavage increased at pH 10.0, the net protein content also decreased **(Fig. 3.6F)**, hence all further experiments pertaining to cleavage at alkaline pH were carried out at pH 9.5. At all the pHs tested, only a single cleaved product was obtained instead of the expected two i.e. 12 and 13 kDa. The cleaved product cross-reacted with anti-LexA antibody **(Fig. 3.2C)**, but not with anti-His antibody. Since, His-tag is at the N-terminal of the purified *Anabaena* LexA protein, it is speculated that only the C-terminal product (13 kDa) is stable and hence, detected upon cleavage.

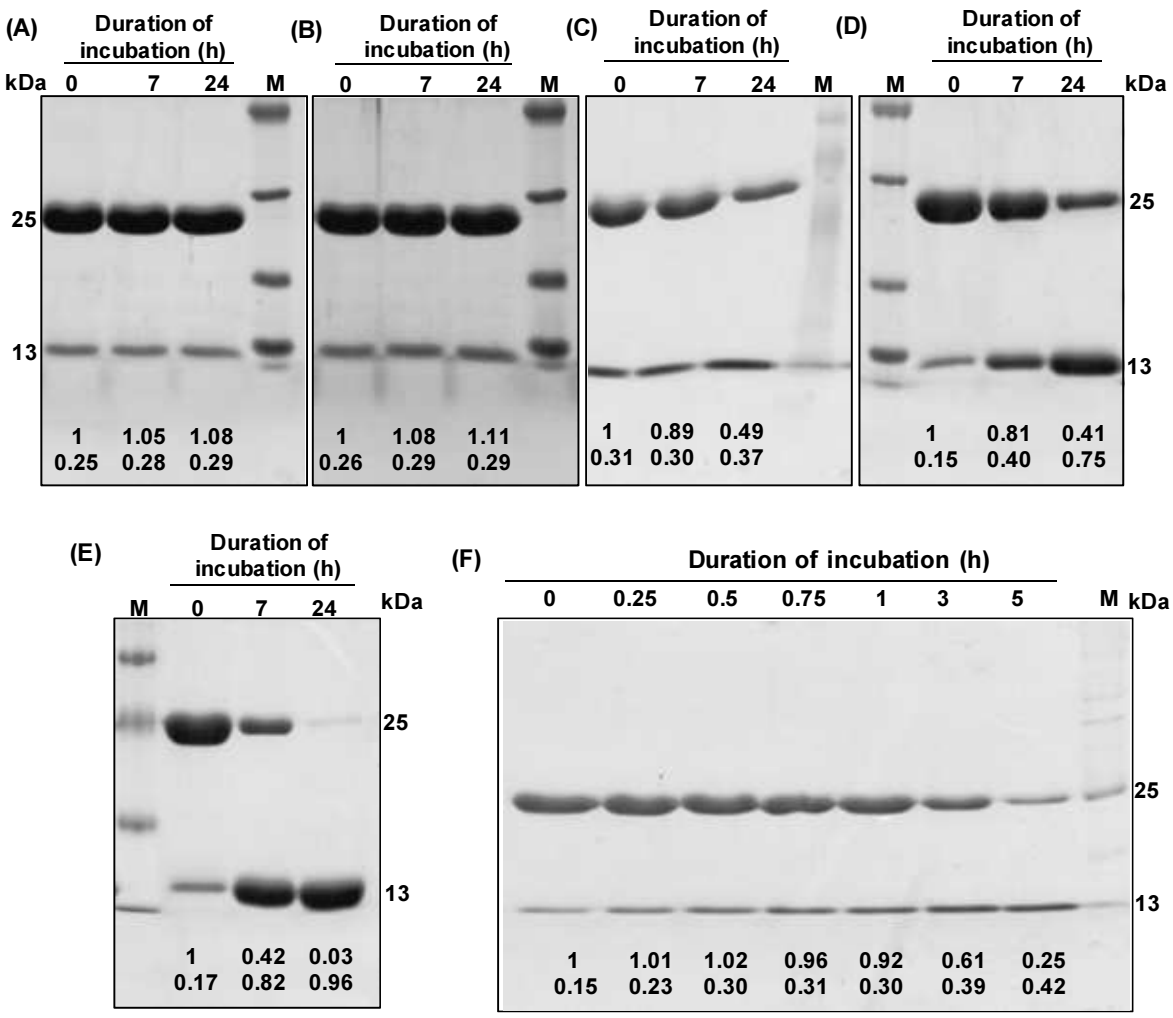


Fig. 3.6. Analysis of auto-proteolytic cleavage of *Anabaena* LexA protein as a function of pH. Purified native LexA protein (8 μM) was incubated in cleavage solution 'A' (100 mM Tris-HCl of different pH, 200 mM NaCl, 1 mM EDTA, 12 mM DTT and 5% glycerol). The different pH used for incubation are **(A)** 6.5, **(B)** 7.5, **(C)** 8.0, **(D)** 8.5, **(E)** 9.5 and **(F)** 10.0 for different duration, as indicated, at 37 °C. The two protein bands corresponding to LexA are indicated with their molecular masses. The intensities of 25 kDa (Top) and 13 kDa (Bottom), indicated below the gel, were calculated using Image J software (Schneider et al 2012) and the values mentioned are relative to that of the 25 kDa band at '0' time point. Other details are as described in the legend to Fig. 3.2.

3.2.3.3 EFFECT OF CALCIUM AND TEMPERATURE ON THE AUTOPROTEOLYTIC ACTIVITY OF *Anabaena* LexA

Presence of Ca^{2+} has been shown to augment autoproteolytic activity of *E. coli* LexA (*107*), hence the effect of Ca^{2+} on the autoproteolytic cleavage activity was analysed. At different time points tested, the cleavage of the 25 kDa to 13 kDa was found to be higher in presence of 10 mM Ca^{2+} compared to that in its absence at pH 9.5 **(Fig. 3.7A)**. In fact, complete cleavage was observed within 5 h in the presence of Ca^{2+} as against only about 70% cleavage in its absence **(Fig. 3.7A)**.

To assess the optimal temperature for autoproteolytic activity of *Anabaena* LexA at pH 9.5, cleavage was analysed over a temperature range of 4 – 60 °C for 3 h. No cleavage was observed at 4 °C, while at 25 °C it was very low **(Fig. 3.7B)**. In the temperature range of 30 °C – 60 °C, maximal amount of the cleaved 13 kDa form was observed at 37 °C. Since the net protein content varied at different temperatures, the relative levels of the 13 kDa cleavage product and the 25 kDa LexA at a specific temperature was calculated relative to the corresponding levels at 4 °C, followed by obtaining the ratio of these values at each temperature **(Fig. 3.7C)**. This revealed maximum cleavage at 46 °C, with at least 5-fold higher levels of the 13 kDa proteolytic product compared to the 25 kDa LexA **(Fig. 3.7C)**.

3.2.3.4 ANALYSIS OF ACTIVATED RecA-DEPENDENT CLEAVAGE OF *Anabaena* LexA

In vivo LexA is cleaved during DNA damage due to the presence of activated RecA which induces the autoproteolytic activity of LexA (*107*). Hence, the autoproteolytic cleavage of *Anabaena* LexA in the presence of activated *E. coli* or *Anabaena* RecA at physiological pH of 7.5 was monitored. Generation of activated *Anabaena*/*E. coli* RecA in the presence of M13 ssDNA was confirmed by increase in the

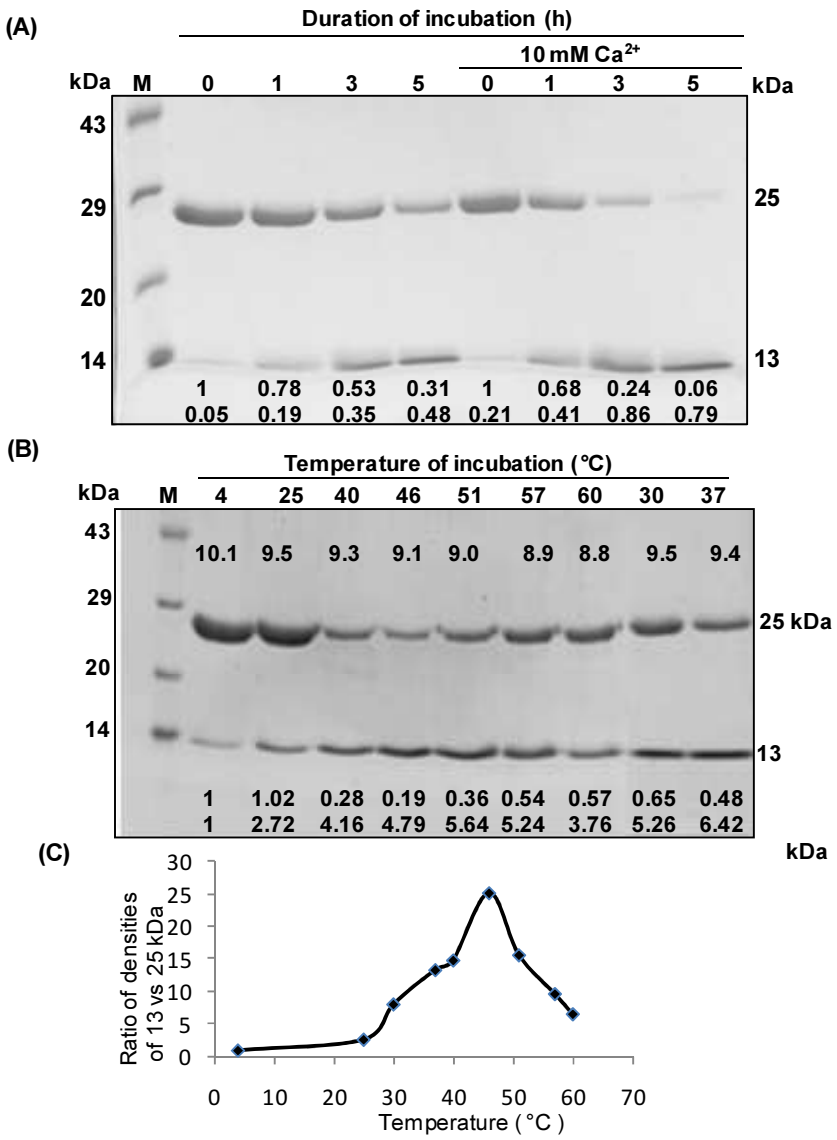


Fig. 3.7. Analysis of Ca^{2+} and temperature dependency of the auto-proteolytic cleavage activity of *Anabaena* LexA. (A) *Anabaena* LexA was incubated in 100 m M Tris-HCl at pH 9.5 buffer for up to 5 h in the absence or presence of 10 mM $CaCl_2$ at 37 °C. **(B)** LexA was incubated in 100 m M Tris-HCl, pH 9.5 buffer at temperature ranging from 4 °C to 60 °C for 3 h. The pH values of the Tris–HCl solution at the end of incubation at different temperatures is given on the top of each lane within the gel. The intensities of 25 kDa (Top)and 13 kDa (Bottom) protein bands have been calculated relative to time'0' in (A) and 4 °C in (B). **(C)** The graph of ratio of the intensities of the 13 kDa to 25 kDa protein bands against incubation temperature relative to the corresponding band at 4°C. Other details are as described in legend to Fig. 3.2.

ATPase activity of RecA upon activation. The change in absorbance at 820 nm due to released P_i upon interaction of RecA and ssDNA was assessed continuously for a period of 1 h at 37 °C. In the absence of ssDNA, the ATPase activity of *Anabaena* RecA was low resulting in only an increase of 0.2 units of absorbance over a 1 h period, whereas in the presence of M13 ssDNA, this increase was achieved within 2.5 min, and a steady increase in absorbance to over 1.1 units was observed over a period of 1 h **(Fig. 3.8A)**. Based on the standard graph, it was estimated that an increase of 1 unit of absorbance corresponds to release of 1 mM Pi. *Anabaena* LexA was incubated in Tris-HCl, pH 7.5 buffer in the presence of activated *Anabaena*/*E. coli* RecA and in the absence or presence of 10 mM $CaCl_2$ for 1 h at 37 °C. Addition of neither the commercially available *E. coli* RecA, activated in a similar manner, nor the laboratory isolated and activated *Anabaena* RecA, stimulated the cleavage of *Anabaena* LexA at pH 7.5 **(Fig. 3.8B)**. Even the addition of Ca^{2+} did not aid in stimulating the cleavage of *Anabaena* LexA in the presence of activated *Anabaena* RecA **(Fig. 3.8B)**. In order to check if the His-tag of LexA was interfering with its interaction with RecA and thereby inhibiting its cleavage, the His-tag was cleaved off from AnLexA using Factor Xa **(Fig. 3.9A)** and used for further analysis. Activated RecA had no effect on the cleavage of either LexA or LexAΔHis_6tag (without His-Tag) **(Fig. 3.9B)**, suggesting *Anabaena* LexA does not exhibit activated RecA-dependent cleavage.

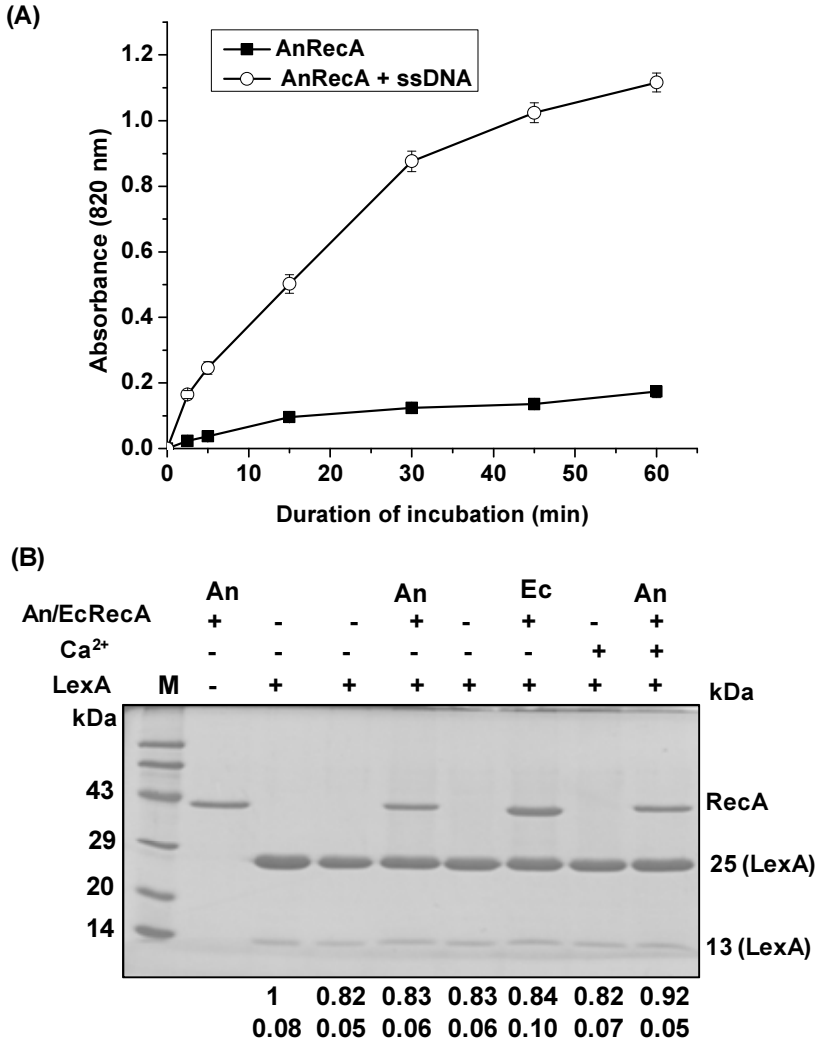


Fig. 3.8. Analysis of auto-proteolytic cleavage of *Anabaena* LexA in presence of activated RecA. (A) ATP hydrolysis by *Anabaena* RecA. RecA (30 μM) was incubated in Tris–HCl, pH 7.5 buffer with 5 mM ATP in the absence or presence of 18 μM M13 ssDNA at 37 °C for 1 h. An aliquot was taken out at the specified time points and assayed for the released amount of inorganic phosphate (Pi) as a function of absorbance at 820 nm. **(B)** LexA (17 μM) was incubated in cleavage solution 'B' (20 mM Tris–HCl pH 7.5, 100 mM NaCl, 2 mM DTT, 5 mM β-mercaptoethanol, 2 mM $MgCl_2$, 0.5 mM EDTA, 2.5% glycerol, 1 mM Thio ATP, 6 μg mL^{-1} M13 single stranded DNA) in the absence or presence of *Anabaena* or *E. coli* RecA (30 μM), pre-activated by the addition of M13 ssDNA and thio-ATP, at 37 °C for 1 h. Other details are as described in legend to Fig. 3.2.

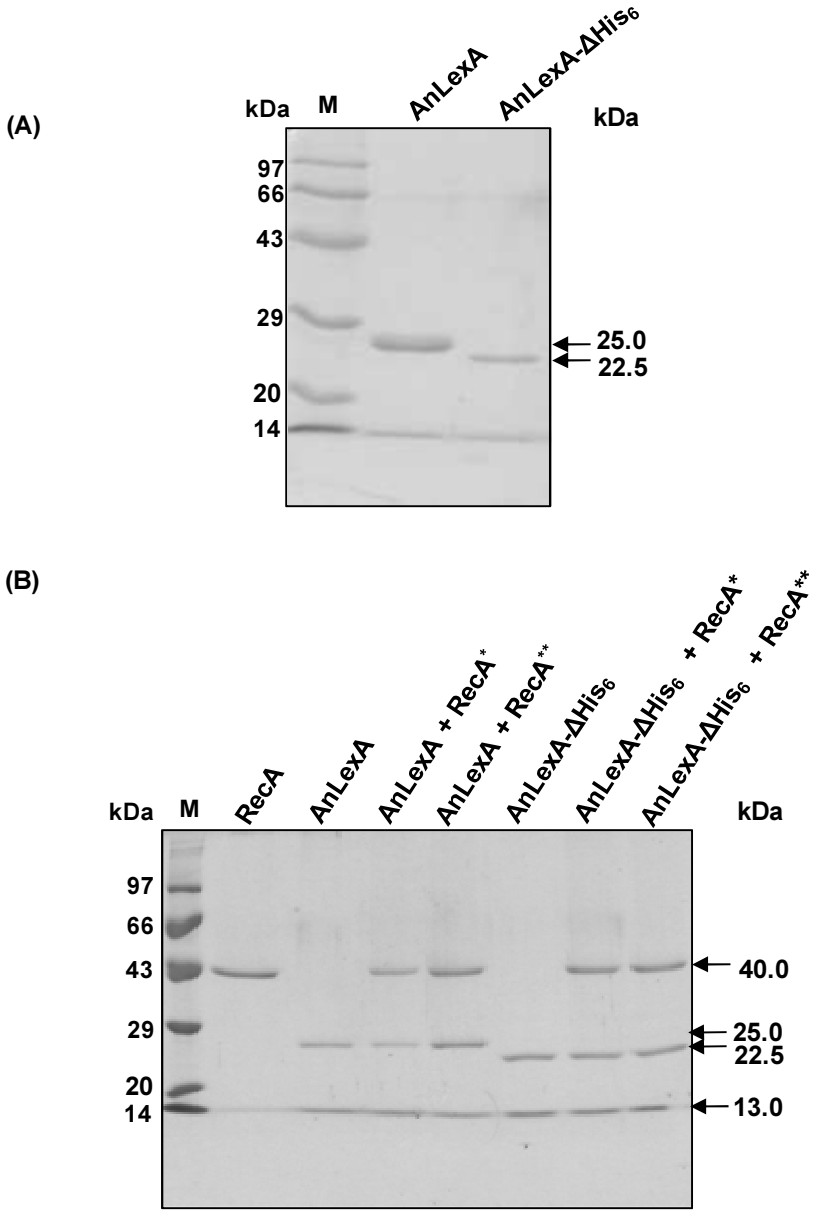


Fig. 3.9. Analysis of auto-proteolytic cleavage of *Anabaena* LexA without His-Tag in presence of activated RecA. (A) Purified *Anabaena* LexA (AnLexA) was treated with Factor-Xa to generate *Anabaena* LexA without His-tag (AnLexA-ΔHis$_6$). **(B)** Comparison of cleavage of AnLexA and AnLexAΔHis$_6$ in the presence of *Anabaena* RecA, indicated as RecA*, if activated for 5 min or as RecA** if activated for 1 h. The molecular mass of different proteins are indicated. Other details are as described in legend to Fig.3.2.

3.2.4 *Amino acid residues contributing to the autoproteolytic activity of Anabaena LexA*

3.2.4.1 COMPARISON OF AMINO ACID SEQUENCE OF *Anabaena* AND *E. coli* LexA PROTEINS

In order to check if the reasons for comparatively lower rate of autoproteolysis of *Anabaena* LexA and the absence of RecA-mediated proteolysis lies in the amino acid sequence of *Anabaena* LexA, it was compared with that of *E. coli* LexA protein. LexA proteins of *Anabaena* and *E. coli* showed an overall identity of 34% and a similarity of 54%, distributed through the length of the protein sequence and not restricted to few pockets **(Fig. 3.10)**. The comparison revealed the presence of cleavage site residues (A^{84}, G^{85}) at identical positions and active site residues ($S^{119/118}$ and $K^{156/159}$) at nearly identical positions in *Anabaena* LexA with respect to that in *E. coli* LexA **(Fig.3.10)**. This possibly accounted for the observed autoproteolytic cleavage at alkaline pH **(Figs. 3.6 and 3.7)**. However, *Anabaena* LexA possessed a longer cleavage site region (CSR), spanning from V^{78} to D^{99} (22 amino acids) compared to the 18 amino acid long CSR of *E. coli* LexA spanning from L^{78} to E^{95} **(Fig. 3.10)**. The difference in length was due to the presence of additional amino acids (G^{86}, L^{87} and I^{88}) in *Anabaena* LexA immediately after the cleavage site and presence of ^{91}FTD^{93} in place of ^{88}LL^{89} immediately after the conserved residues 'EP' **(Fig. 3.10)**. The regions in *E. coli* LexA protein which are predicted to be essential for interaction with activated RecA (*106*), and underlined in **Fig. 3.10** bear low homology with the corresponding region in *Anabaena* LexA, which possibly explains the inability of activated RecA to induce autoproteolytic cleavage of *Anabaena* LexA at physiological pH **(Figs. 3.8 and 3.9)**.

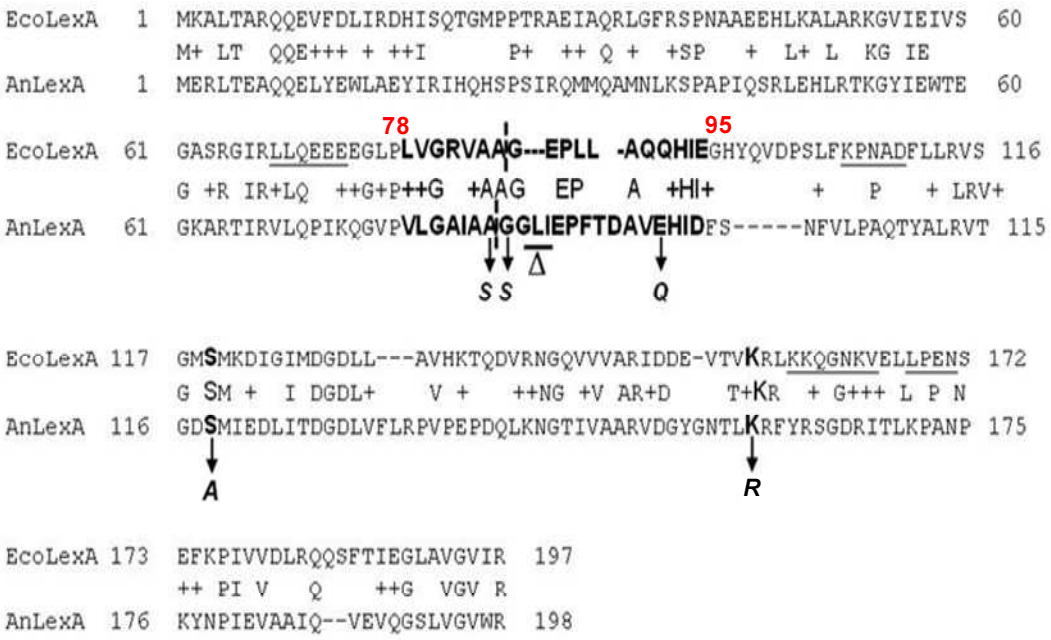


Fig. 3.10 Comparison of *Anabaena* LexA and *E. coli* LexA protein sequences. Amino acid sequence of LexA proteins of *E. coli* (EcoLexA) (Uniprot No. P0A7C2) and *Anabaena* 7120 (AnLexA) (Uniprot No. Q8YMM5) were aligned using NCBI BLASTp program. The middle line represents the identical and similar amino acids in the two proteins. The cleavage site region (CSR) is shown in bold and cleavage site i.e. A84–G85 by a vertical dotted line. For generation of point mutations in *Anabaena* LexA discussed in section 3.2.4.2, the amino acid substitutions are indicated by arrows and deletion of 'GLI' by 'Δ'. The numbers '78' and '95' shown in red colour represent the start and end of the predicted CSR respectively.

To check if the long CSR was exclusive to LexA of *Anabaena* 7120 or present in other cyanobacterial LexA proteins as well, the corresponding region was compared for LexA protein from 11 different cyanobacterial strains including *Anabaena* 7120. Comparison revealed great degree of similarity in the cleavage site region among the different cyanobacterial LexA proteins **(Fig. 3.11)**. Of the 11 sequences compared, the cleavage site residues A^{84}, G^{85} were present in all except *Synechocystis* 6803 and *Gloeobacter violaceus* LexA, which have been predicted to be non-cleavable (*169*), wherein they have been replaced with 'GG' and 'EE' respectively **(Fig. 3.11)**. In the remaining nine cyanobacterial LexA, the residues ^{86}GLI/V^{88} immediately after the cleavage site is well conserved. Another difference, was the substitution of Q^{92} in *E. coli* LexA, shown to be essential for RecA-dependent cleavage in *E. coli* LexA (*43*), by ^{96}E/D in most cyanobacterial LexA proteins **(Fig. 3.11)**. This may also account for the lack of *in vitro* RecA-dependent cleavage of *Anabaena* LexA.

3.2.4.2 EFFECT OF MUTATIONS IN LexA ON ITS AUTOPROTEOLYTIC CLEAVAGE ACTIVITY

Point mutations and deletion mutants of *Anabaena* LexA were generated to confirm the involvement of (i) the predicted cleavage and active site residues, and (ii) specific residues in the cleavage site region in autoproteolytic activity of LexA. The different point mutations carried out were that of the cleavage site residues [A^{84} to S^{84} (A84S) and G^{85} to S^{85} (G85S)], active site residues [S^{118} to A^{118} (S118A) and K^{159} to R^{159} (K159R)] and a residue shown to be involved in RecA-mediated cleavage of *E.* coli LexA E^{96} to Q^{96} (E96Q), as indicated in **Fig. 3.10**. Different deletion mutants generated included, deletion of (i) 3 residues ^{86}GLI^{88} in the CSR (LexA-ΔGLI), (ii) C-terminal residues ^{61}G to ^{201}M (LexA-NTD), and (iii) N-terminal residues ^{1}M-^{61}G (LexA-CTD). The primers used for the generation of the different mutants are listed in **Table 3.1**, and

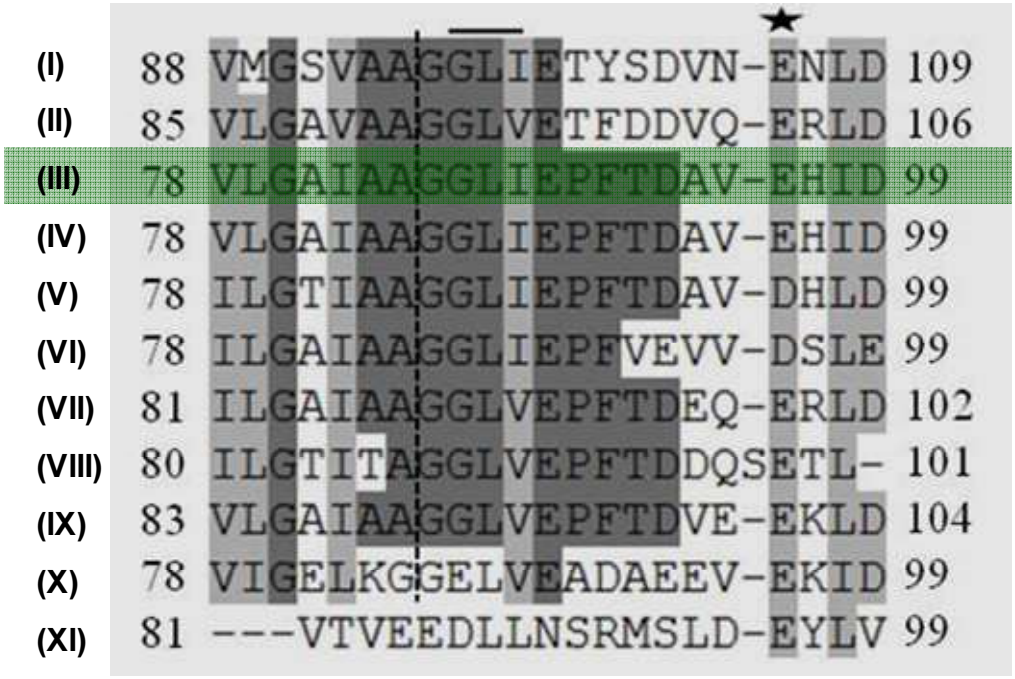


Fig. 3.11 Comparison of predicted CSR (Cleavage Site Region)region of different cyanobacterial LexA proteins. The predicted CSR (22 amino acids long) of different cyanobacterial LexA proteins was compared with respect to *Anabaena* 7120 LexA (III) using ClustalW program (www.genome.jp/tools/clustalw).The identical amino acids are shaded 'dark grey' and similar amino acids 'light grey'. The cleavage site is shown by a dotted vertical line and arrow marked on the top of the dotted line. The conserved ^{86}GLI/V^{88} residues indicated by a horizontal line of the top and ^{96}E/D residue by an asterisk. The numbers on either side of the amino acids sequence indicate the position of the amino acid at the extreme ends in the corresponding full length protein. The different cyanobacterial strains with LexA proteins Uniprot nos. written in square brackets used for analysis are: (I) *Prochlorococcus marinus* MIT9515 [A2BXW9], (II) *Synechococcus sp.* WH7803 [A5GME1], (III) *Anabaena* PCC7120 [Q8YMM5], (IV) *A. variabilis* ATCC29413 [Q3MB18], (V) *Nostoc punctiforme* PCC73102 [B21VN5], (VI) *Cyanothece sp.* ATCC 51142 [B1X0G0], (VII) *Microcystis aerugenosa* NIES843 [B0JPQ9], (VIII) *Synechococcus sp.* PCC7002 [B1XQ55], (IX) *Acaryochloris marina* AM1 [B0C8B3], (X) *Synechocystis sp* PCC6803 [P73722], (XI) *Gloeobacter violaceus* PCC7421 [Q7NMQ6].

the resulting recombinant plasmids generated by cloning the mutant *lexA* variants in pET16b are listed in **Table 3.2**. All the above generated LexA mutants could be overexpressed in *E. coli* BL21 cells **(Fig. 3.12A, 3.13A)**. With the exception of LexA-ΔGLI and LexA-NTD mutants, the others could be expressed in soluble fractions and purified under native conditions using Tris-HCl buffer **(Fig. 3.12B)**. The overexpressed LexA-ΔGLI and LexA-NTD proteins, which went into inclusion bodies, were purified by Ni-NTA affinity chromatography under denaturing conditions using 8 M urea **(Fig. 3.13 B, C).** Renaturation of the purified denatured LexA-ΔGLI and LexA-NTD proteins were carried out by stepwise dialysis against decreasing concentrations of urea (from 8 M to 0 M in steps of 2 M decrease). At the end of the renaturation process, no protein was detected in case of LexA-NTD **(Fig. 3.13B)**, suggesting that it may be degraded, while LexA-ΔGLI was detected **(Fig. 3.13C)**. Addition of protease inhibitor PMSF also could not aid in the recovery of renatured LexA-NTD. Taking advantage of the fact that no aromatic residues have been deleted in the LexA-ΔGLI mutant, and no change in structure predicted with this deletion, renaturation of the protein was assessed by analysing the fluorescence spectra and comparing it with that of native and denatured wild type LexA. The fluorescence spectrum of native wild type LexA and renatured LexA-ΔGLI were found to be identical **(Fig. 3.13D)**, while upon denaturation with 8 M urea, both exhibited an additional peak at 425 nm **(Fig. 3.13 D)**, indicating the LexA-ΔGLI protein was properly folded.

The cleavage site residue mutants, LexA-A84S and LexA-G85S did not exhibit any cleavage unlike the wild type LexA upon incubation in cleavage buffer at pH 9.5 for 5 h at 37 °C **(Fig. 3.14A)**. Mutation in the active site residues i.e. LexA-K159R and LexA-S118A also completely abolished the autoproteolytic cleavage activity of LexA **(Fig. 3.14A and B).** The LexA E96Q mutant exhibited hyper cleavage with three-fold

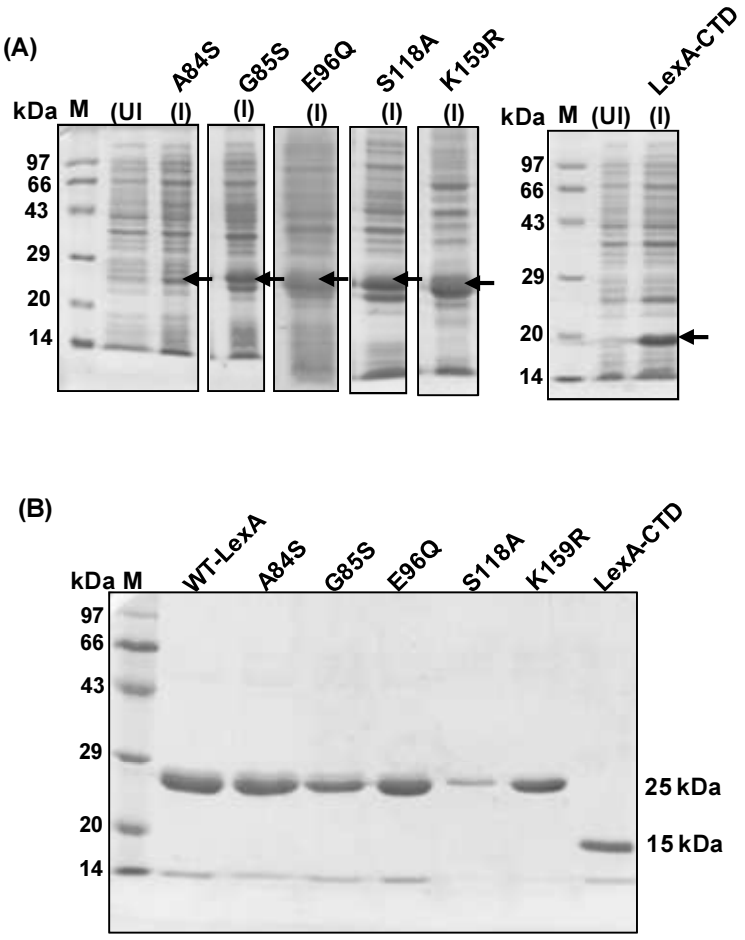


Fig. 3.12 Overexpression and purification of mutant *Anabaena* LexA proteins in *E. coli*. (A) Overexpression of different point mutants and N-terminal deletion mutants (CTD) of *Anabaena* LexA as indicated by arrows. **(B)** Ni-NTA affinity purified mutant LexA proteins, as indicated, under native conditions. Lane M indicate protein molecular marker which are indicated with molecular size. 'UI' indicates un-induced and 'I' indicates induced with 1 mM IPTG. Other details are as described in the legend to Fig. 3.2.

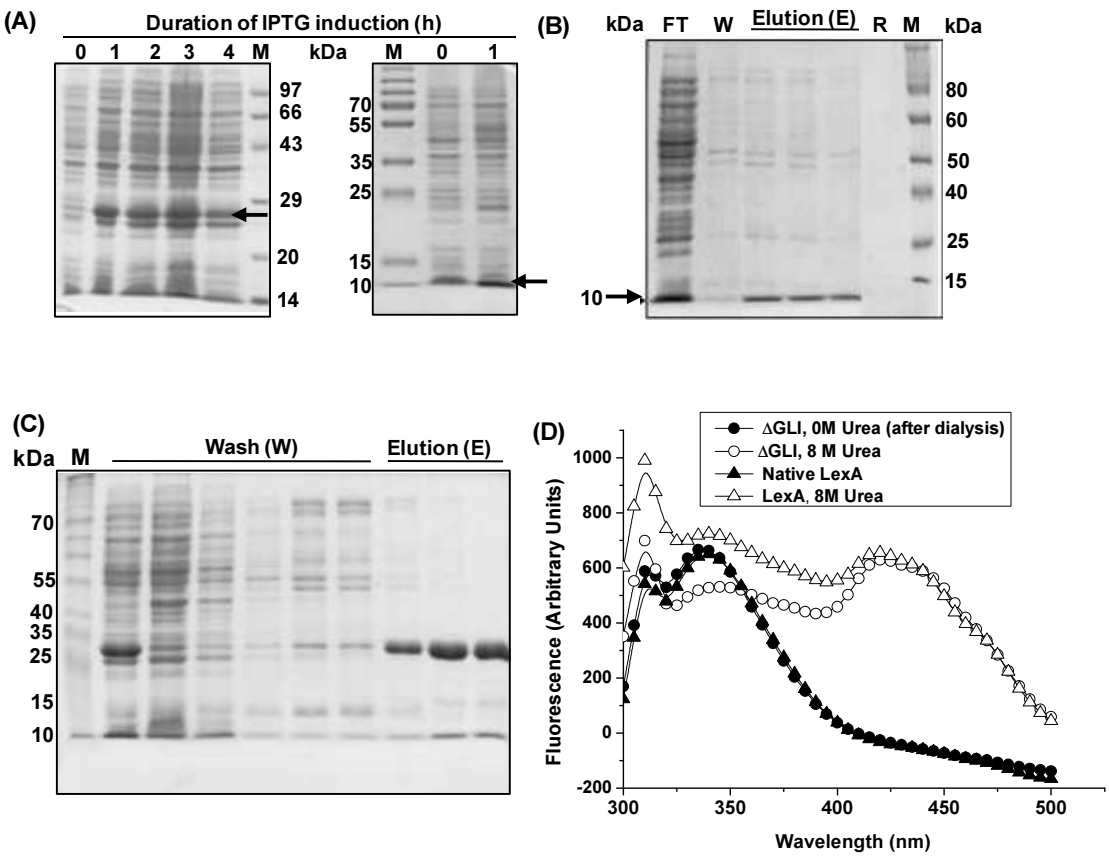


Fig. 3.13. Overexpression and purification of LexA-ΔGLI and LexA-NTD proteins of *Anabaena* 7120 from *E. coli* BL-21 cells. (A) Overexpression of LexA-ΔGLI (86-88) and LexA-NTD proteins by induction with 1 mM IPTG **(B)** LexA-NTD protein was purified under denaturing condition with 8M urea using Ni-NTA affinity chromatography. The eluted protein was dialysed against a decreasing concentration of urea in step wise manner from 8 M to 0 M and the final protein was loaded on to lane as indicated as 'R' (after renaturation). **(C)** Purification of LexA-ΔGLI protein using 8 M urea-Ni-NTA chromatography. Different lanes correspond to FT: flow through; W: washes and E: elutions. The different protein molecular weight markers used are loaded in lane 'M' and the molecular mass of different protein bands indicated. Other details as described in the legend to Fig.3.2. **(D)** Fluorescence emission spectrum (obtained after subtracting the spectra obtained with the corresponding buffer) of purified native and denatured WT LexA and LexA- ΔGLI proteins after excitation at 282 nm.

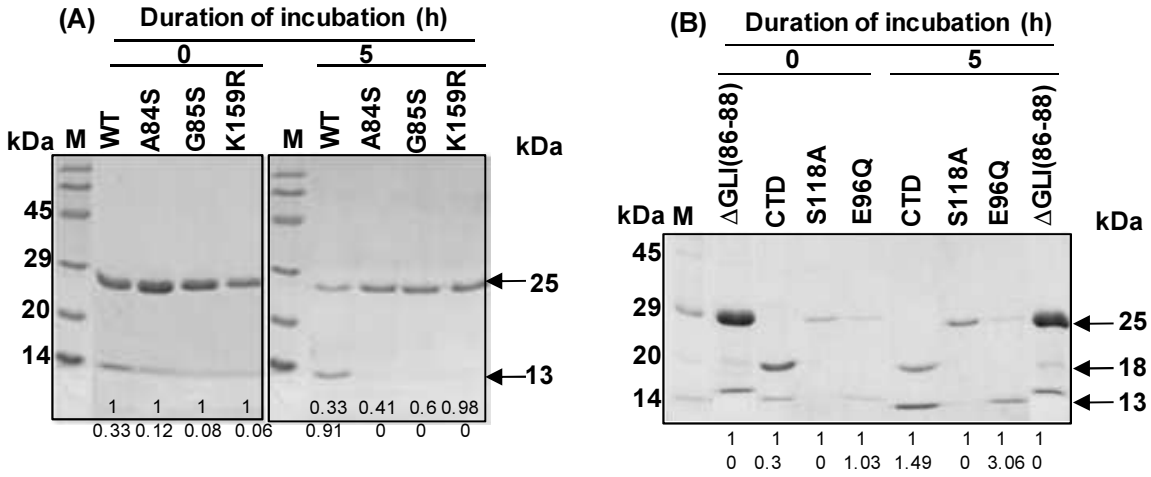


Fig. 3.14 Analysis of autoproteolytic cleavage of mutant LexA proteins of *Anabaena*. **(A)** Incubation of purified wild type and mutant LexA proteins in Tris-HCl, pH 9.5 at 37 °C for 5 h. Proteins used were **(A)** wild type LexA (WT), point mutants (A84S, G85S and K159R), In 0 h gel the intensity of the 13 kDa protein band relative to that of the 25 kDa protein is listed below the gel; while in the 5 h gel, intensity of the band relative to the corresponding band at '0 h' has been calculated and given below the gel. **(B)** Point mutants (S118A and E96Q) and deletion mutants (GLI and CTD). The intensities of the uncleaved and cleaved mutant LexA protein is given below the gel, with the intensity at time '0 h' for each mutant LexA protein individually taken as '1' and that at '5 h' determined relative to it. The molecular mass of the purified proteins and their cleaved products, if any, are indicated by arrows. Other details are as described in the legend to Fig. 3.2.

higher levels of the 13 kDa protein band compared to wild type LexA after 5 h incubation at 37 °C at pH 9.5 **(Fig. 3.13 B)**. Of the two deletion mutants of LexA which could be purified, LexA-CTD exhibited cleavage comparable to that of wild type LexA, whereas LexA-ΔGLI did not exhibit any cleavage upon incubation at 37 °C at pH 9.5 for 5 h **(Fig. 3.14 B).** In all cases only a single protein of 13 kDa was observed after cleavage of the cleavable forms of LexA. The 12 kDa protein fragment not detected corresponds to NTD, which under native conditions was shown to be unstable.

Lack of RecA-dependent cleavage of *Anabaena* LexA raised questions on whether the regulation of the protein activity is through autoproteolytic cleavage or through the regulation of gene expression under *in vivo* conditions in response to various abiotic stresses and normal conditions, which was investigated.

3.2.5 *Expression analysis of LexA under different DNA damaging stresses*

Being a known SOS-response regulator in general in bacteria, the expression of *Anabaena* LexA in response to various DNA damage inducing stresses, namely mitomycin C, UV- and γ-radiation and desiccation stresses, was analysed both at transcript and protein levels.

3.2.5.1 Effect of stresses inducing DNA adduct formation on the expression of *lexA* in *Anabaena*

Upon exposure to 1.5 kJ m^{-2} of UV-radiation, the *lexA* transcript levels increased several folds in *Anabaena* **(Fig. 3.15A)**. However, this was not translated into an increased levels of the LexA protein upon exposure to 0-1.5 kJ m^{-2} of UV-radiation **(Fig.**

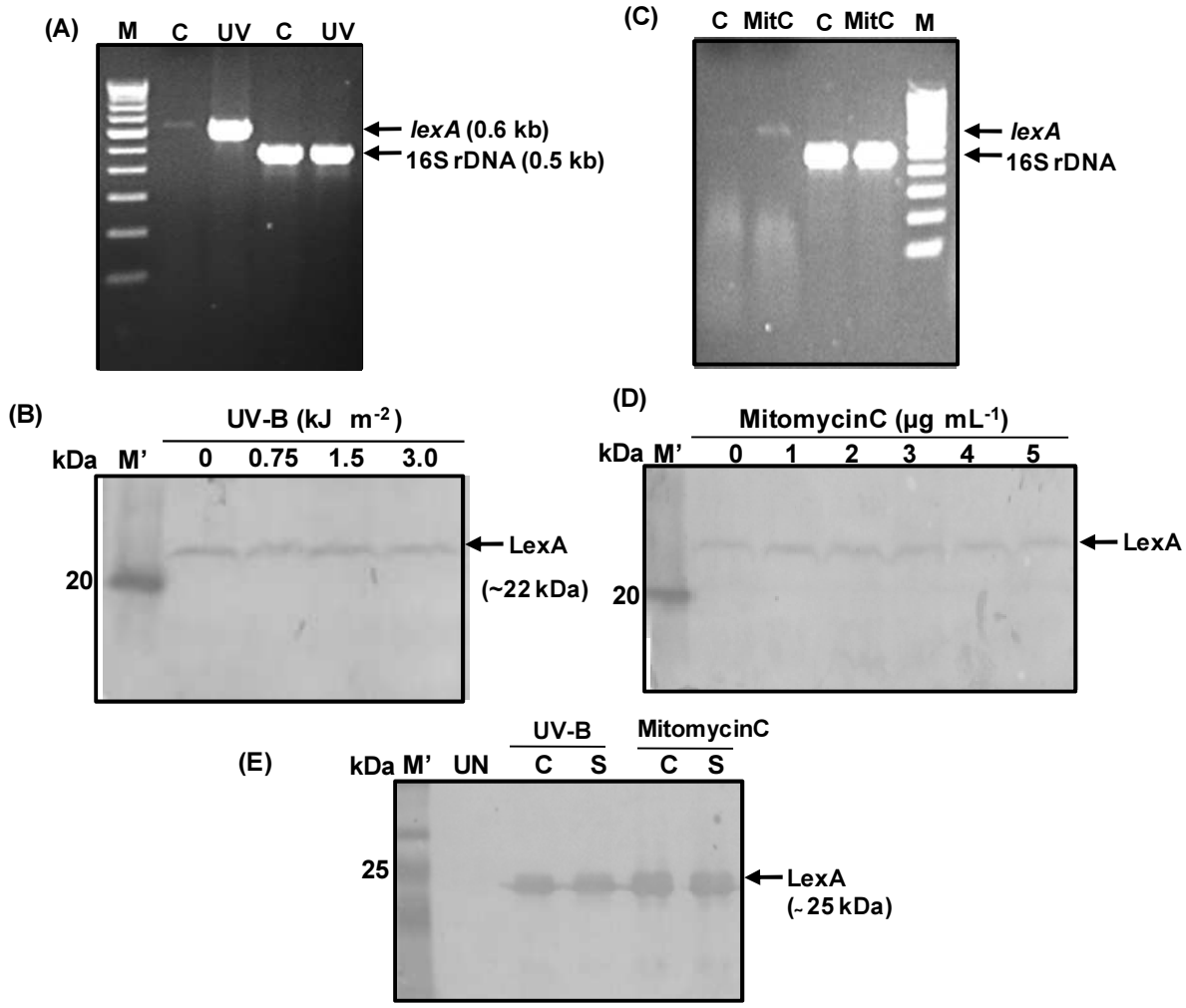


Fig. 3.15. Effect of UV-radiation and mitomycin C on the expression of *lexA*, and its autoproteolytic cleavage. Three-day-old *Anabaena* 7120 cultures were exposed to **(A and B)** UV-stress, or **(C and D)** mitomycin C stress. Levels of *lexA* and 16S rDNA (partial) transcripts in response to exposure to **(A)** 1.5 kJ m^{-2} UV-stress and **(C)** Mitomycin C (4 µg mL^{-1}) for 30 min. The amount of RNA used for *lexA* and 16S rDNA was 160 ng per lane . The amplified *lexA* and 16S rDNA transcripts are indicated by arrows, and Lane M corresponds to 100 bp DNA ladder. Levels of LexA protein determined by Western blotting and immunodetection using anti-LexA antibody in *Anabaena* 7120 in response to different doses of **(B)** UV-radiation, and **(D)** mitomycin C. **(E)** Effect of UV (50 J m^{-2}) and mitomycin C (2 µg mL^{-1}, 1 h) stress on the expression and cleavage of *Anabaena* LexA overexpressed in logarithmic phase cultures of *E. coli* by the addition of 1 mM IPTG. Different lanes are UN: uninduced, C; induced, control, S: induced, stress and M': prestained protein marker and the cross-reacting LexA protein is indicated.

3.15B). A ~3-fold increase in *lexA* transcript levels was observed upon exposure to *Anabaena* 7120 to 4 µg mitomycin C mL^{-1} for 30 min **(Fig. 3.15C)**, but there was no detectable change in the levels of the LexA protein upon exposure to 0-5 µg mitomycin C mL^{-1} for 30 min **(Fig. 3.15D)**. To check if this discrepancy was due to lack of translation of *lexA* or degradation of LexA during stress, *E. coli* cells overexpressing *Anabaena* LexA were exposed to UV- and mitomycin C stresses. No change in the levels of the LexA protein was observed post stress **(Fig. 3.15E)**, indicating that *Anabaena* LexA is not proteolytically degraded upon exposure to either UV or mitomycin C stresses. This suggested that the translation of the accumulated transcripts of *lexA* is inhibited by some yet to be identified process.

3.2.5.2 Effect of stresses inducing single and double strand breaks in DNA on the expression of *lexA* in *Anabaena*

The *lexA* transcript was not detected in *Anabaena* 7120 cells exposed to 6 kGy of γ-irradiation **(Fig. 3.16A)**, indicating not only complete repression of its transcription, but also degradation of the existing *lexA* mRNA. This was also reflected in the absence of the LexA protein in γ-irradiated cells of *Anabaena* 7120 **(Fig. 3.16B)**. Six days of desiccation also resulted in the repression of transcription of the *lexA* gene **(Fig. 3.16C)**, as well as the synthesis of the LexA protein **(Fig. 3.16D)**.

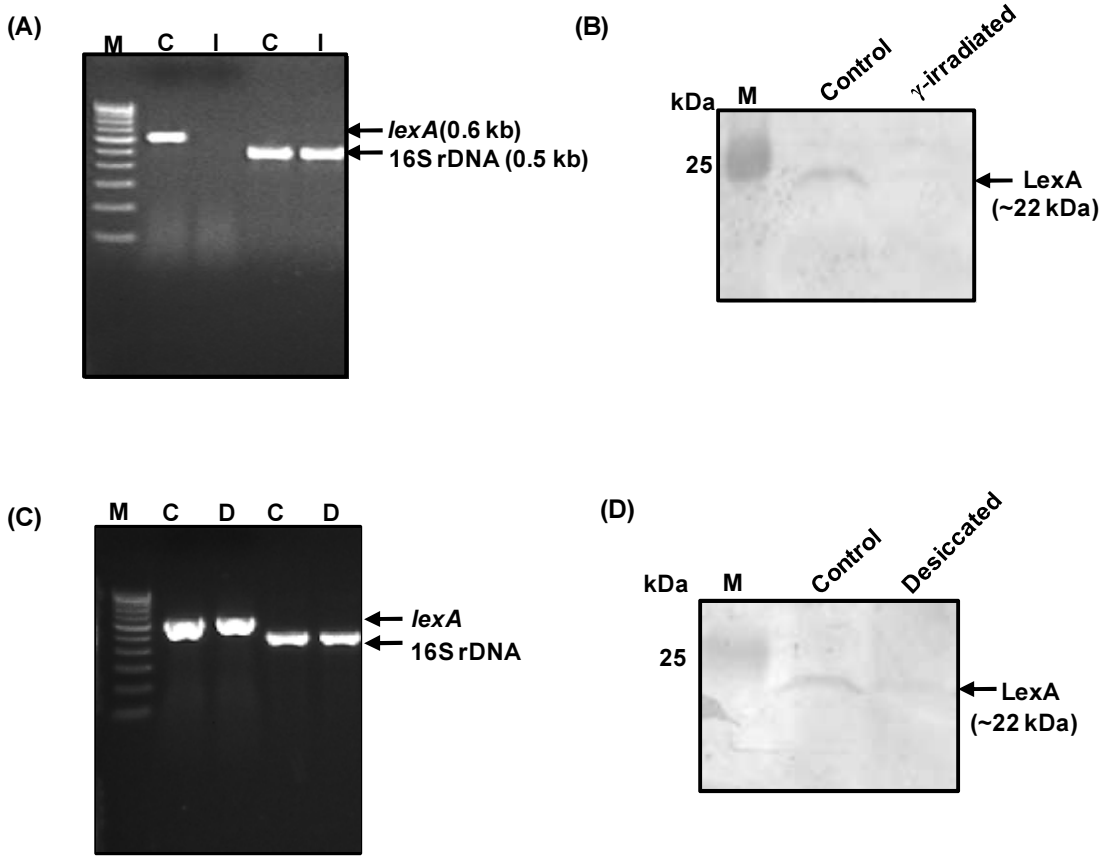


Fig. 3.16. Effect of γ-radiation and desiccation stresses on the expression of *lexA*

Three-day-old *Anabaena* 7120 cultures after concentration to 10 μg Chl *a* mL^{-1} were exposed to **(A and B)** 6 kGy of γ-radiation or **(C and D)** 6 days of desiccation. Levels of *lexA* and 16S rDNA (partial) transcripts in response to exposure to **(A)** γ-radiation and **(C)** desiccation. The amount of RNA used for *lexA* and 16S rDNA was 200 ng per lane . The amplified *lexA* and 16S rDNA transcripts are indicated by arrows, and Lane M corresponds to 100 bp DNA ladder. Levels of LexA protein determined by Western blotting and immunodetection using anti-LexA antibody in *Anabaena* 7120 in response to **(B)** γ-radiation, and **(D)** desiccation. The cross-reacting LexA protein is indicated. Other details are as given in legend for Fig. 3.15.

3.3 DISCUSSION

SOS-response in bacteria is triggered by autoproteolytic cleavage of LexA at Ala^{84}-Gly^{85}, which in turn is induced by activated RecA in response to DNA damage (*162*). In bacteria LexA regulates over 30 genes involved in DNA repair (*42*). The SOS-response, *per se*, has not been characterized in cyanobacteria, but the two key regulatory proteins of SOS response i.e. LexA and RecA are present across cyanobacterial species (*169*). Two cyanobacteria i.e. *Synechocystis* 6803 and *Gloeobacter* were predicted to possess non-cleavable form of LexA while all other sequenced cyanobacteria possessing LexA had a cleavable form (*169*). Our studies in this chapter focussed on the mechanism of autoproteolytic cleavage of *Anabaena* LexA, as no information was available on the cleavage of cyanobacterial LexA proteins.

In vitro studies revealed that *Anabaena* LexA exists as a dimer (~55 kDa) under native conditions **(Fig. 3.3A)** with a hydrodynamic size of 9.8 nm in diameter **(Fig. 3.4)**, having secondary structure elements represented as 28% α-sheets and 20% β-sheets at pH 7.5 **(Fig. 3.3B).** This indicated that the major portion of the protein is in the loop form, which would provide the flexibility to the protein to carry out required proteolytic function, by the action of residues in the C-terminal to the cleavage site residues in the N-terminal. As observed for other bacterial LexA proteins, *Anabaena* LexA also autoproteolytically cleaved at Ala^{84}-Gly^{85} peptide bond **(Fig. 3.14),** and the presence of Ser at 118^{th} position and Lys at 159^{th} position was found to be absolutely essential for the cleavage activity **(Fig. 3.14)**. The corresponding Ser and Lys residues in *E. coli* and other bacteria form the part of active sites and are involved in the nucleophilic attack required for the cleavage (*107*). The conversion of LexA from its non-cleavable to cleavable form occurs either through the interaction of activated RecA with LexA or at alkaline pH,

wherein OH^- acts as a nucleophile (*107*). Of the two modes of cleavage, *Anabaena* LexA exhibited cleavage only at alkaline pH i.e. pH 8.0 and above **(Fig. 3.6)** and not in the presence of activated RecA **(Fig. 3.8 and 3.9)**. In fact, the inability of RecA to induce cleavage of *Anabaena* LexA was observed both under *in vitro* **(Fig. 3.8 and 3.9)**, and *in vivo* conditions **(Fig. 3.15C, D and E)**. Comparison of *Anabaena* LexA protein sequence with that of *E. coli* **(Fig. 3.10)**, revealed significant differences in the stretches of amino acids predicted to be involved in interaction of LexA with RecA in *E. coli* (*106*). This could account for the absence of interaction between *Anabaena* LexA and RecA proteins and hence the absence of RecA-mediated cleavage by LexA.

Another distinct difference in the cleavage was in the number of products obtained. In case of *E. coli* LexA, two products of 12 and 13 kDa are obtained under *in vitro* conditions upon inducing the cleavage (*107*). On the other hand, *Anabaena* LexA consistently gave rise to a single product upon cleavage which corresponded to the 13 kDa CTD **(Fig. 3.6)**, based on size comparison with the LexA-CTD expressed independently **(Figs. 3.12A and B)**. In fact, *Anabaena* LexA-NTD could be overexpressed but not purified from *E. coli* **(Figs. 3.13A and B)** and was possibly extensively degraded. In *E. coli*, to achieve efficient deregulation by LexA, the N-terminal and the C-terminal domains are degraded by the action of Clp and Lon proteases respectively *in vivo* (*163-164*). On the other hand, the high proteolytic instability of the NTD of *Anabaena* LexA, which contains the DNA-binding domain, would be self sufficient in effecting deregulation of LexA-regulated genes in *Anabaena.*

In addition to the importance of A^{84}, G^{85}, S^{118} and K^{159} residues in autoproteolytic cleavage of *Anabaena* LexA, a few more amino acids, which are conserved across cyanobacteria, were also found to be essential for the cleavage. This included the residues '^{86}GLI^{88}' in the CSR and found exclusively in cyanobacteria, the deletion of

which resulted in complete loss of autoproteolytic activity **(Fig. 3.13B)**, indicating that these residues may be playing a role in managing the micro environment required for cleavage. Additionally the conserved 'Q^{92}' residue in *E. coli* LexA essential for cleavage by interaction with RecA (*43*) is substituted by 'E^{96}' in *Anabaena* LexA **(Fig. 3.10)**. The Q to E mutation rendered the *E. coli* LexA hypercleavable and did not require activated RecA for cleavage (*43*). This may also contribute to the observed RecA-independent cleavage of *Anabaena* LexA. The mutation of E to Q in *Anabaena*, however, did not decrease the autoproteolytic activity **(Fig. 3.14B),** as was expected based on the *E. coli* results *(43)*. This suggested that other factors, such as the absence of the RecA-LexA interacting sites may also be playing a significant role in rendering proteolytic activity of *Anabaena* LexA being RecA-independent. Cleavage of LexA was not observed under DNA damaging stresses in *Anabaena* **(Fig. 3.15)**, which could be due to the lack of RecA-mediated cleavage. The significance of cleavage of *Anabaena* LexA at alkaline pH in cellular physiology needs to be investigated. Though, LexA levels remain unchanged in response to agents inducing DNA adduct formation **(Fig. 3.15)**, the transcription and translation of LexA was shut down/repressed in *Anabaena* in response to agents inducing DNA strand breaks **(Fig. 3.16)**. This suggested that LexA may be acting as a repressor of the DNA repair genes in *Anabaena,* as observed for *E. coli* (*42*), which is essential for efficient DNA repair in radiation resistant *Anabaena.* Study on the physiological relevance of LexA in *Anabaena,* thus gains importance and is discussed in the next chapter.

CHAPTER 4

PHYSIOLOGICAL ROLE OF LexA IN *Anabaena*

4.1 INTRODUCTION

In cyanobacteria, studies on physiological role of LexA have been restricted to unicellular *Synechocystis* 6803, while in *Anabaena*, few genes regulated by LexA have been identified. Depletion of LexA in the partially segregated *lexA* mutant *Synechocystis* 6803, decreased its ability to survive under inorganic carbon starvation (*127*) and also affected the genes involved in phototactic movements, bidirectional hydrogenase and components of phycobilisome complexes (*170*). DNA microarrays analysis (*127*) and RNA-seq profiling (*170*) revealed genes involved in different metabolic processes, including fatty acid biosynthesis (*139*), but not DNA repair as being regulated by LexA in *Synechocystis* *lexA* mutant. LexA was also found to be involved in the regulation of *hupSL* operon and *hypFCDEAB* operon in *Lyngbya majuscule* (*136*). In *Anabaena* 7120, LexA was found to be involved in the regulation of the *hoxEFUYH* operon, coding for bidirectional hydrogenase (*132*) and few DNA repair genes (*130, 171*) confirmed by DNA binding and foot printing studies with regions upstream of a few of these genes (*169*). However, the LexA-binding site varied in terms of sequence as well as location (*130, 169*) and a consensus sequence encompassing all genes speculated to be regulated by LexA was not identified, instead a variation of the sequence predicted earlier (*130*) was used. This sequence was found to be present in 216 genes of *Anabaena* (*132*), but the position of the box was either upstream or downstream of the gene and the distance from the start codon also varied.

This chapter deals with obtaining an insight into the varied regulatory roles of LexA and a consensus LexA-binding site using the model system *Anabaena* 7120 for the study. A fully segregated recombinant strain of *Anabaena* overexpressing LexA, An*lexA*$^{+}$ was generated and comparative proteomic analysis of AnpAM and An*lexA*$^{+}$

was carried out. Regulation by LexA at transcription level was analysed and consensus AnLexA-Box defined. The physiological role of LexA was analysed by comparative abiotic stress tolerance studies of AnpAM and An*lexA*$^{+}$ strains.

4.2 RESULTS

4.2.1 *Generation of Anabaena lexA mutant strain*

For generating a recombinant *Anabaena* strain having an insertional mutation in the *lexA* gene, a construct was generated wherein the *lexA* gene was interrupted with *nptII* cassette. A schematic representation of the strategy used is shown in **Fig. 4.1**. The first step involved PCR amplification of (i) 1 kb *lexA*up region corresponding to 714 bp upstream to *lexA* gene and 279 bp of the 5'end of the *lexA* gene, and (ii) *lexA*dn region corresponding to 315 bp of the 3' end of the *lexA* gene and 611 bp further downstream to the *lexA* gene using *Anabaena* 7120 genomic DNA as the template DNA and specific primers as listed in **Table 4.1**.

Table 4.1: List of primer used for making LexA knockout

Primers	Nucleotide Sequence*	Amplicon
*lexA*up Fwd	GCGGCCGGATCCTGTAATTAATGCAC	*lexA*up
*lexA*up Rev	CGTGAGGCCGATATCAGTGAATGGTTCTAT	(0.99 kb)
*lexA*dn Fwd	GCCGCCGATATCGACTTTTCTAATTTCGTTTTACC	*lexA*dn
*lexA*dn Rev	GATCTGCTCGAGTTGTGACTGGAACCCATC	(0.93 kb)

*The underlined sequence represents the restriction site

The 1 kb *lexA*up DNA fragment **(Fig. 4.2A)** and the 1 kb *lexA*dn DNA fragment **(Fig. 4.2B)** were restriction digested and cloned in pBluescript vector II (pBS) **(Table 2.3)**

individually at *Bam*HI/*Eco*RV and *Eco*RV/*Xho*I restriction sites respectively as shown in **Step II** of **Fig. 4.1**. The resulting plasmid constructs were designated as pBS*lexA*up and

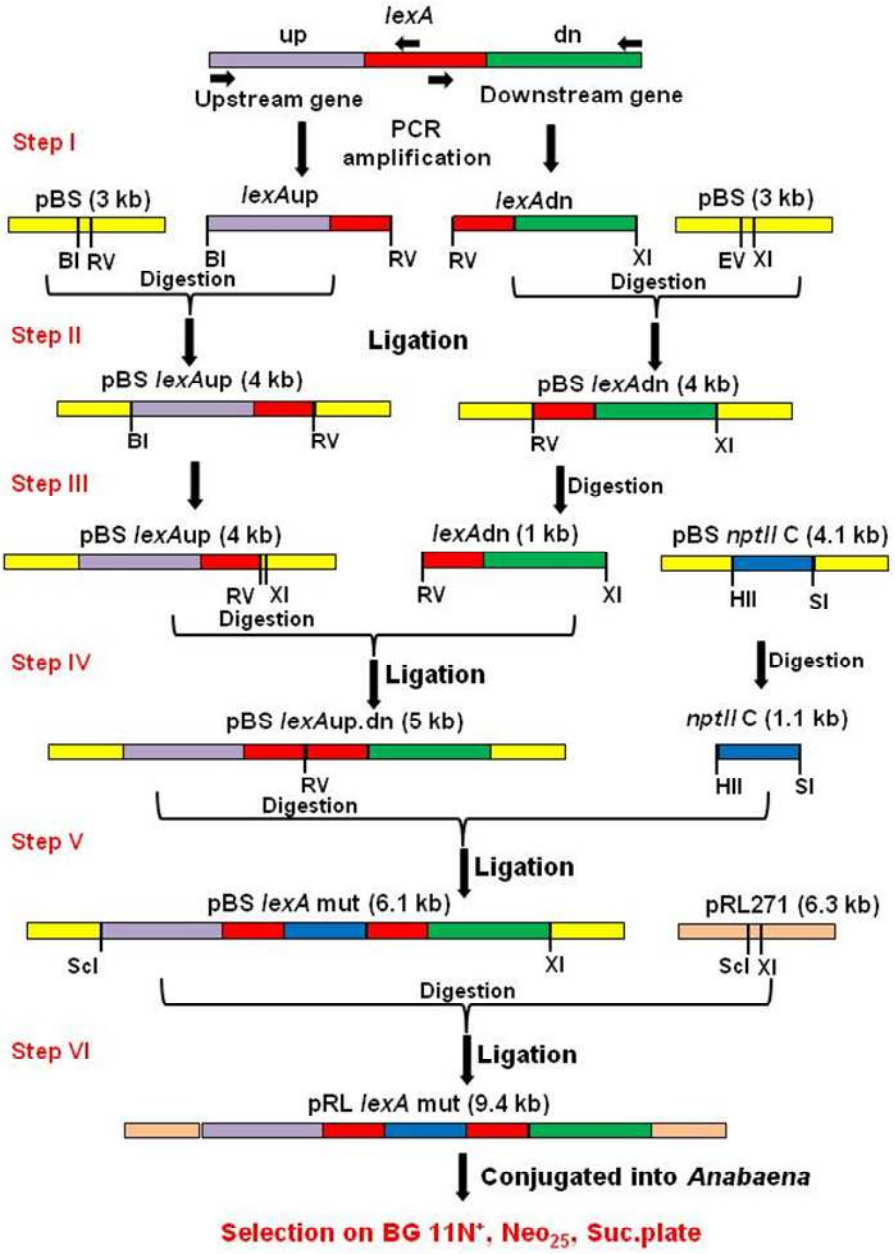


Fig. 4.1 Schematic representation of the strategy used for generation of *lexA* knockout mutant of *Anabaena* 7120. Different DNA fragment which include genes and plasmids as indicated, are shown as rectangular boxes in different colours. The different steps involved, indicated to the right of the block arrows. The different restriction enzymes and sites indicated are: BI: *Bam*HI, EV: *Eco*RV, XI: *Xho*I, HII: *Hinc*II, SI: *Sma*I and ScI: *Sac*I.

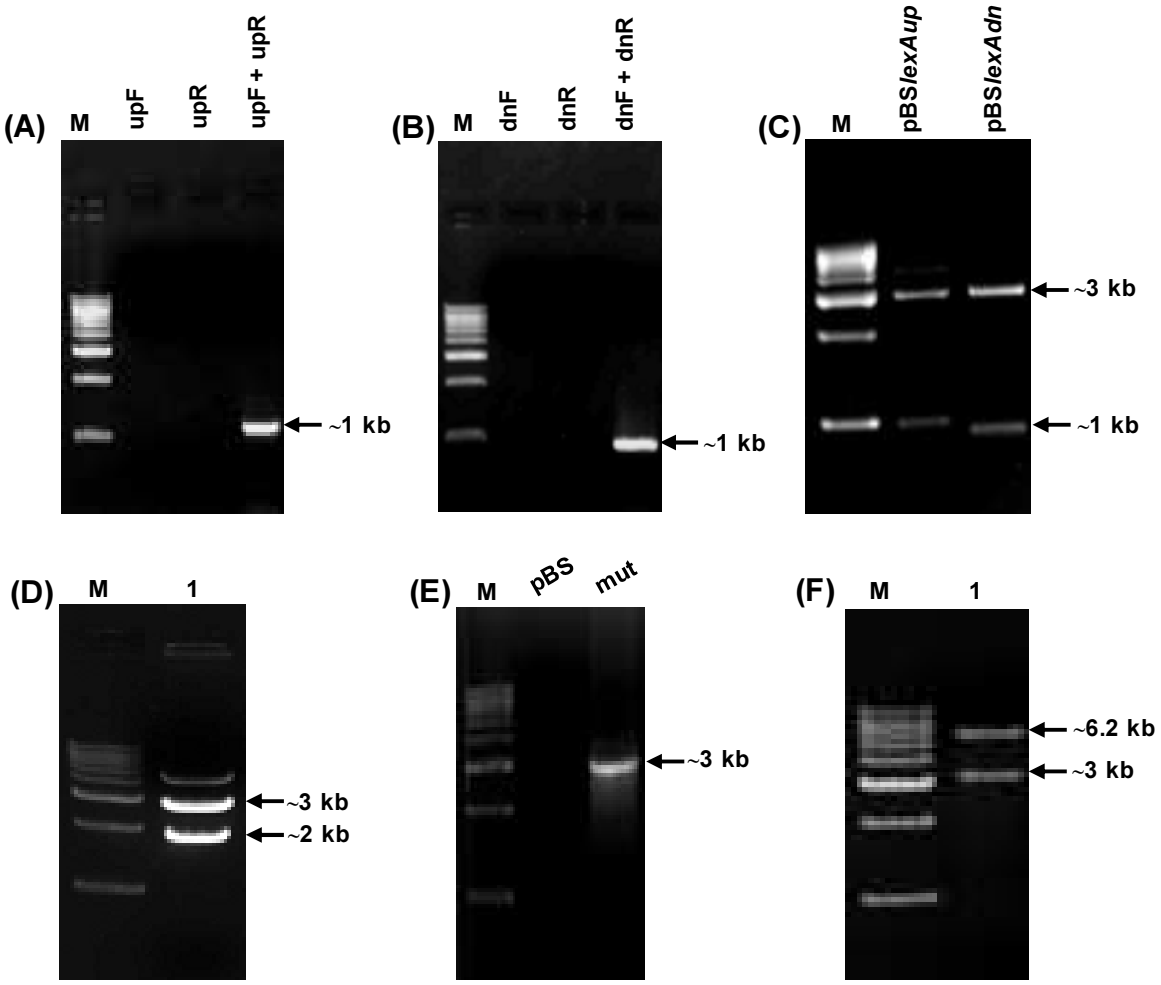


Fig. 4.2 Cloning of *lexA* mutant constructs. (A and B) PCR amplification of **(A)** *lexA*up and **(B)** *lexA*dn DNA fragments using specific primers; upF and upR refers to *lexA*up forward and reverse primers respectively, and dnF and dnR to *lexA*dn forward and reverse primers respectively. **(C and D)** Restriction digestion of recombinant plasmids **(C)** pBS*lexA*up with *Bam*HI and *Eco*RV, and pBS*lexA*dn with *Eco*RV and *Xho*I **(D)** pBS*lexA*up.dn with *Bam*HI and *Xho*I. **(E)** PCR amplification of plasmids pBS and pBS*lexA*mut (represented as mut) with *lexA*up forward and *lexA*dn reverse primers. **(F)** Restriction digestion of the recombinant plasmid pRL*lexA*mut (lane 1) with *Sac*I and *Xho*I enzymes. Lane M refers to 1 kb DNA ladder. Molecular sizes of different DNA fragments are indicated by arrows.

pBS*lexA*dn respectively **(Table 4.2)** and the cloning confirmed by restriction digestion of the corresponding plasmids, pBS*lexA*up and pBS*lexA*dn with *Bam*HI/*Eco*RV and *Eco*RV/*Xho*I restriction enzymes respectively resulting in the release of ~1 kb insert and a ~ 3 kb pBS vector DNA **(Fig. 4.2C)**.

Table 4.2: List of plasmids used to generate recombinant *Anabaena* strains

Plasmids	**Description**	**Source**
pBS*lexA*up	Cb^r, 0.99 kb *lexA*up cloned in pBS at *BamHI-EcoRV* restriction sites	This study
pBS*lexA*dn	Cb^r, 0.93 kb *lexA*dn cloned in pBS at *EcoRV-XhoI* restriction sites	This study
pBS*lexA*updn	Cb^r, 1.92 kb *lexA*updn cloned in pBS at *BamHI- XhoI* restriction sites	This study
pBS*lexA*mut	Cb^r, Kan^r, 1.1 kb *nptIIC* gene cassette cloned in pBS*lexA*updn at *EcoRV* restriction sites	This study
pRL*lexA*mut	Cb^r, Cm^r, Kan^r, ~3 kb lexAmut cloned in pRL271 at *SacI-XhoI* restriction sites	This study
pFPN*lexA*	Cb^r, Kan^r, 0.6 kb *lexA* cloned in pFPN at *NdeI-BamHI* restriction sites	This study
pAM*lexA*	0.87 kb *XmaI-SalI* fragment from pFPN*lexA*, cloned in pAM1956 vector	This study

The next step involved cloning of the 1 kb *Eco*RV/*Xho*I *lexAdn* fragment from pBS*lexA*dn, into pBS*lexA*up digested with the same set of restriction enzymes, resulting in the plasmid construct designated as pBS*lexA*up.dn **(Step IV, Fig. 4.1; Table 4.2)**. Restriction digestion of pBS*lexA*up.dn with *Bam*HI and *Xho*I resulted in the release of two DNA fragments of about 2 kb and 3 kb corresponding to *lexAupdn* insert and the pBS vector respectively **(Fig. 4.2D)**, thereby confirming the presence of insert in

pBS*lexA*up.dn. In **Step V**, the 1.1 kb *nptII*C cassette was excised out from pBS*nptII*C (*150*) **(Table 2.3)** by digesting with *Hinc*II and *Sma*I restriction enzymes and ligated to pBS*lexAup.dn*, linearized with *Eco*RV **(Fig. 4.1)**. PCR amplification of the resulting plasmid construct, designated as pBS*lexA*mut **(Table 4.2)**, with *lexAup* forward and *lexAdn* reverse primers, resulted in the amplification of a 3 kb DNA fragment **(Fig. 4.2E)** confirming the presence of the insert. In the last step **(Step VI, Fig. 4.1**) pBS*lexA*mut plasmid was digested with *Sac*I and *Xho*I restriction enzymes and the resulting ~3 kb fragment was ligated to pRL271 plasmid vector digested with the same set of enzymes. The resulting plasmid construct was designated as pRL*lexA*mut **(Table 4.2)**. Restriction digestion of pRL*lexAmut* with *Sac*I and *Xho*I restriction enzymes resulted in the release of the 6.3 kb vector and 3 kb insert DNA **(Fig. 4.2F)**, thus confirming the cloning.

The plasmid pBS*lexA*mut was introduced into *Anabaena* by electroporation as described in **Section 2.1.6**, while the plasmid pRL*lexA*mut was introduced into *Anabaena* 7120 by conjugation using *E. coli* strain HB101 and two helper plasmids pRL443 and pRL623 also maintained in HB101 as described in materials and methods **(Section 2.1.7)**. The recombinant *Anabaena* strains, wherein the chromosomal *lexA* gene is interrupted with *nptII* gene were selected on BG-11, N^+, Neo_{25} plates. Both the plasmids pBS*lexA*mut and pRL*lexA*mut act as suicide vectors in *Anabaena* as they fail to replicate, but the basic difference while obtaining Neo^r recombinant *Anabaena* strains is the ability to select double point crossovers selectively using the pRL*lexA*mut plasmid. Thus, in Neo^r exconjugants, obtained using pRL*lexA*mut plasmid wherein a single recombination event has taken place resulting in integration of the entire plasmid on to the chromosome, the *sacB* gene coding for levan sucrose, will be expressed and the recombinants would die on BG-11 plates containing sucrose. On the other hand, in case of double recombination event, only the *lexA* gene interrupted with *nptII* cassette would

be integrated, so *sacB* will not be expressed and cells would survive on BG-11 sucrose plates. However, no double point crossovers were obtained using this strategy, while with pBS*lexA*mut plasmid the segregation of the mutant was very poor and unstable. Thus, a viable and stable *Anabaena lexA* mutant strain could not be obtained. Hence, efforts were made to generate a recombinant *Anabaena* strain constitutively overexpressing the LexA protein.

4.2.2 *Generation of recombinant Anabaena strain overexpressing LexA protein*

The plasmid pAM*lexA* **(Table 4.2)** used for overexpressing LexA in *Anabaena* was generated as shown in **Fig. 4.3**.

In **Step I**, the plasmids pET*lexA* **(Table 3.2)** and pFPN (*146*) **(Table 2.3)** were digested with *Nde*I and *Bam*HI restriction enzymes. The 0.6 kb *lexA* gene released from pET*lexA* was ligated to the linearized pFPN plasmid **(Step II, Fig. 4.3)**. The positive colonies were screened by colony PCR (data not shown) using *lexA* forward and *lexA* reverse primers **(Table 3.1)**, and further confirmed by subjecting the plasmid DNA from PCR positive clone to restriction digestion with *Nde*I and *Bam*HI restriction enzymes, which resulted in the release of 0.6 kb *lexA* insert and 5 kb vector DNA **(Fig. 4.4A)**. The resulting plasmid construct was designated as pFPN*lexA* **(Table 4.2)**. The 0.8 kb *lexA* ORF along with *psbA1* promoter was excised out from pFPN*lexA* by digesting with *Sal*I *and Xma*I restriction enzymes and ligated to pAM1956 vector **(Table 2.3)** at same restriction site upstream to *gfpmutII* **(Steps III and IV, Fig. 4.3)**. The resulting plasmid construct was designated as pAM*lexA* **(Table 4.2)**.

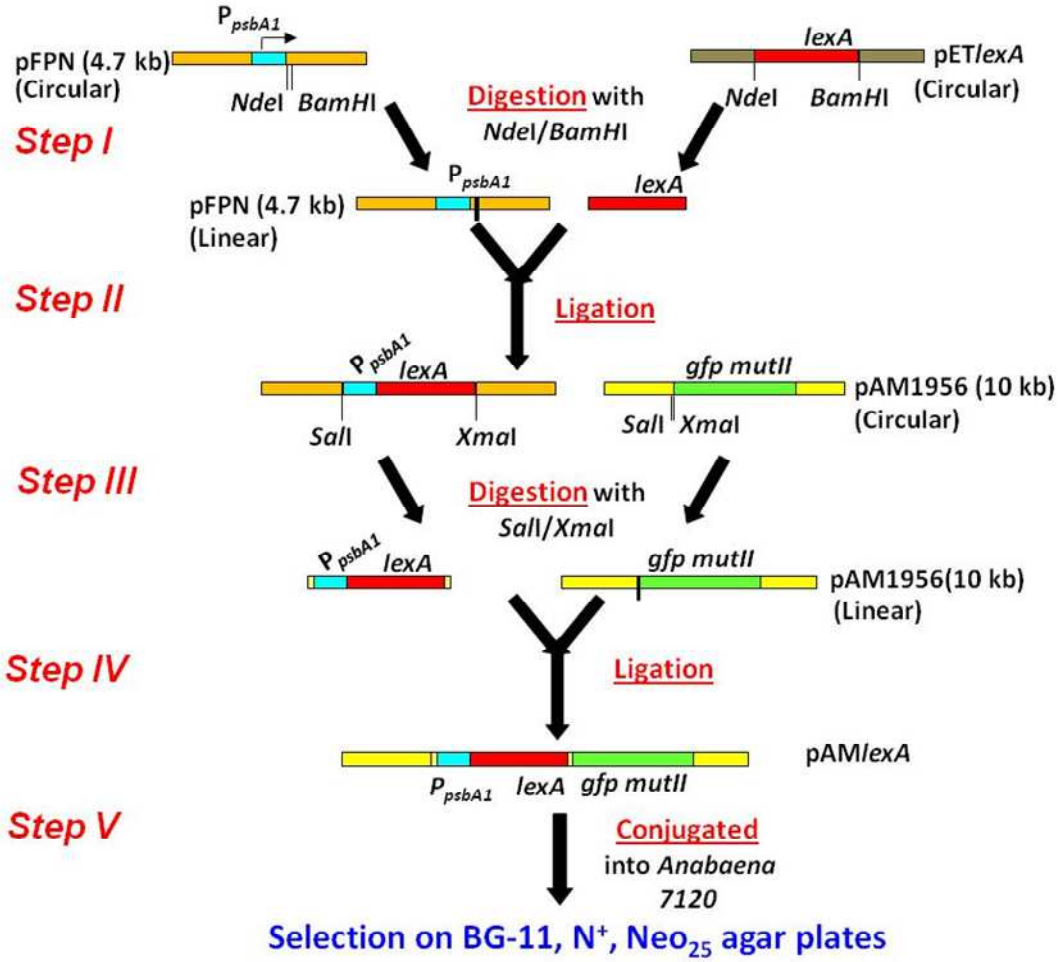


Fig. 4.3 Schematic representation of the strategy used for generation of LexA overexpressing strain of *Anabaena* 7120 (An*lexA*$^{+}$). Different DNA fragment which include genes and plasmids as indicated, are shown as rectangular boxes in different colours. The different steps involved are indicated to the right of the block arrows. The different restriction enzyme sites are also shown.

The cloning of *lexA* gene into pAM1956 was confirmed by plasmid PCR using *lexA* forward and *lexA* reverse primers, which resulted in the detection of a 0.6 kb amplicon **(Fig. 4.4B)**. The plasmid pAM*lexA* was introduced in *Anabaena* by triparental conjugation and the exconjugants were selected on BG-11, N^+, Neo_{25} plates by repeated sub-culturing to allow complete segregation. In the plasmid pAM*lexA*, the *lexA* gene is transcriptionally fused to *gfpmutII* (which codes for a stable Green Fluorescence Protein), resulting in *gfpmutII* being co-transcribed along with the *lexA* gene from *psbA1* promoter. The translation of the two ORFs i.e. *lexA* and *gfpmutII* was, however, independent.

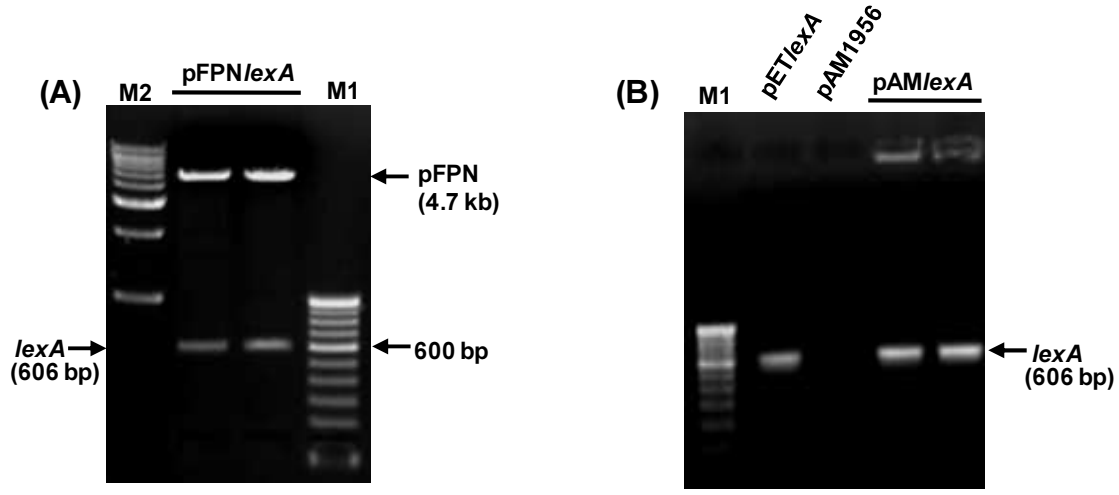


Fig. 4.4 Cloning of *lexA* in pFPN and pAM1956 plasmids. (A) Restriction digestion of two recombinant plasmids corresponding to pFPN*lexA*1 with *Nde*I and *Bam*HI restriction enzymes. **(B)** PCR amplification of *lexA* from two recombinant plasmids corresponding to pAM*lexA* using *lexA* primers. M1 and M2 indicate 100 bp and 1 kb DNA ladder respectively. Molecular sizes of different DNA fragment were indicated by arrows.

GFP being a stable protein and visualisable under fluorescence microscope (λ_{ex} 470 nm, λ_{em} 508 nm), the expression of GFP in the generated recombinant *Anabaena* strains was used to monitor the extent of segregation in different cells of the filament as well as different filaments when under the fluorescence microscope. For all experiments, a Neo^r vector control *Anabaena* strain AnpAM (*150*) (**Table 2.1**) was used, wherein the recombinant *Anabaena* strain harbours the empty vector pAM1956. Wild type *Anabaena* as well as vector control AnpAM exhibited red fluorescence **(Fig. 4.5A)** due to the presence of chlorophyll while the recombinant *Anabaena* overexpressing LexA, designated as An*lexA*$^+$ exhibited orange to green fluorescence **(Fig. 4.5A)**, dependent upon the extent of GFP expression. Western blotting and immunodetection using anti-AnLexA antibody revealed higher levels of LexA protein in case of An*lexA*$^+$ compared to AnpAM cells **(Fig. 4.5B)**, thereby confirming overexpression of both LexA and GFP proteins in An*lexA*$^+$. Overexpression of LexA did not cause any visible change in the gross morphological features as observed under bright light microscope **(Fig. 4.6A)**. When allowed to grow under normal growth conditions in BG-11, N^+ media, the growth of An*lexA*$^+$ strain was comparable to that of vector control AnpAM cells as indicated by similar increase in the chlorophyll *a* content over a 7-day period **(Fig. 4.6B)**. Thus, constitutive expression of LexA had no bearing on the growth of *Anabaena* under normal growth conditions. But being a regulatory protein, it could play a role in the modulation of response to various abiotic stresses. This was assessed by comparing the growth (in terms of chlorophyll *a* content) and survival (in terms of colony forming units after 10 days) of An*lexA*$^+$ and AnpAM cells upon exposure to various abiotic stresses.

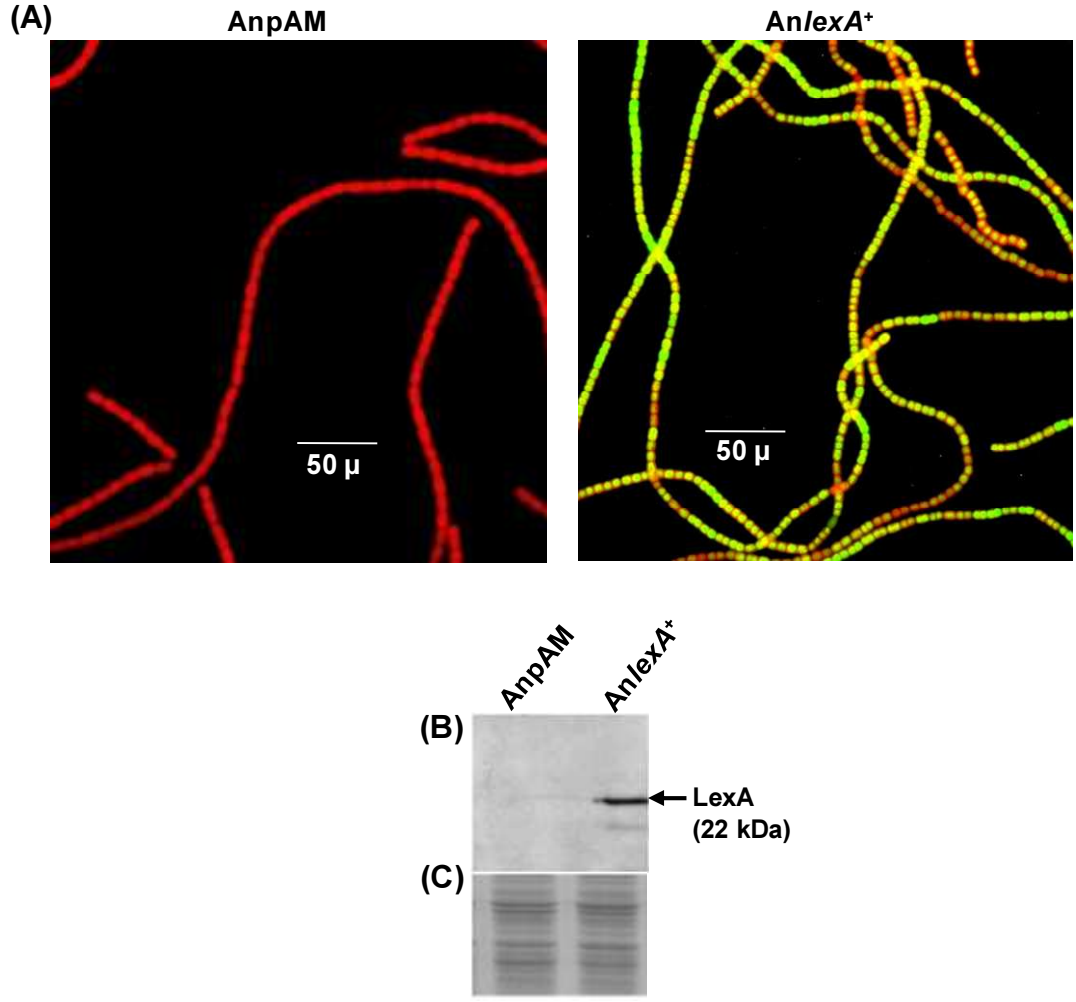


Fig. 4.5 Generation of recombinant An*lexA*$^+$ strain. (A) Fluorescence micrographs (600X magnification) of recombinant *Anabaena* strains, AnpAM (vector control) and An*lexA*$^+$ filaments (LexA overexpressing strain). The fluorescence was observed using Hg-Arc lamp with microscope (excitation 470 nm, emission 508 nm). **(B)** Western blotting and immunodetection of LexA in whole cell lysate of AnpAM and An*lexA*$^+$ cells. Polyclonal antibodies generated against purified *Anabaena* LexA protein was used for immunodetection. The cross-reacting 22 kDa LexA protein is indicated. **(C)** Coomassie stained gel corresponding to the blot to in (B).

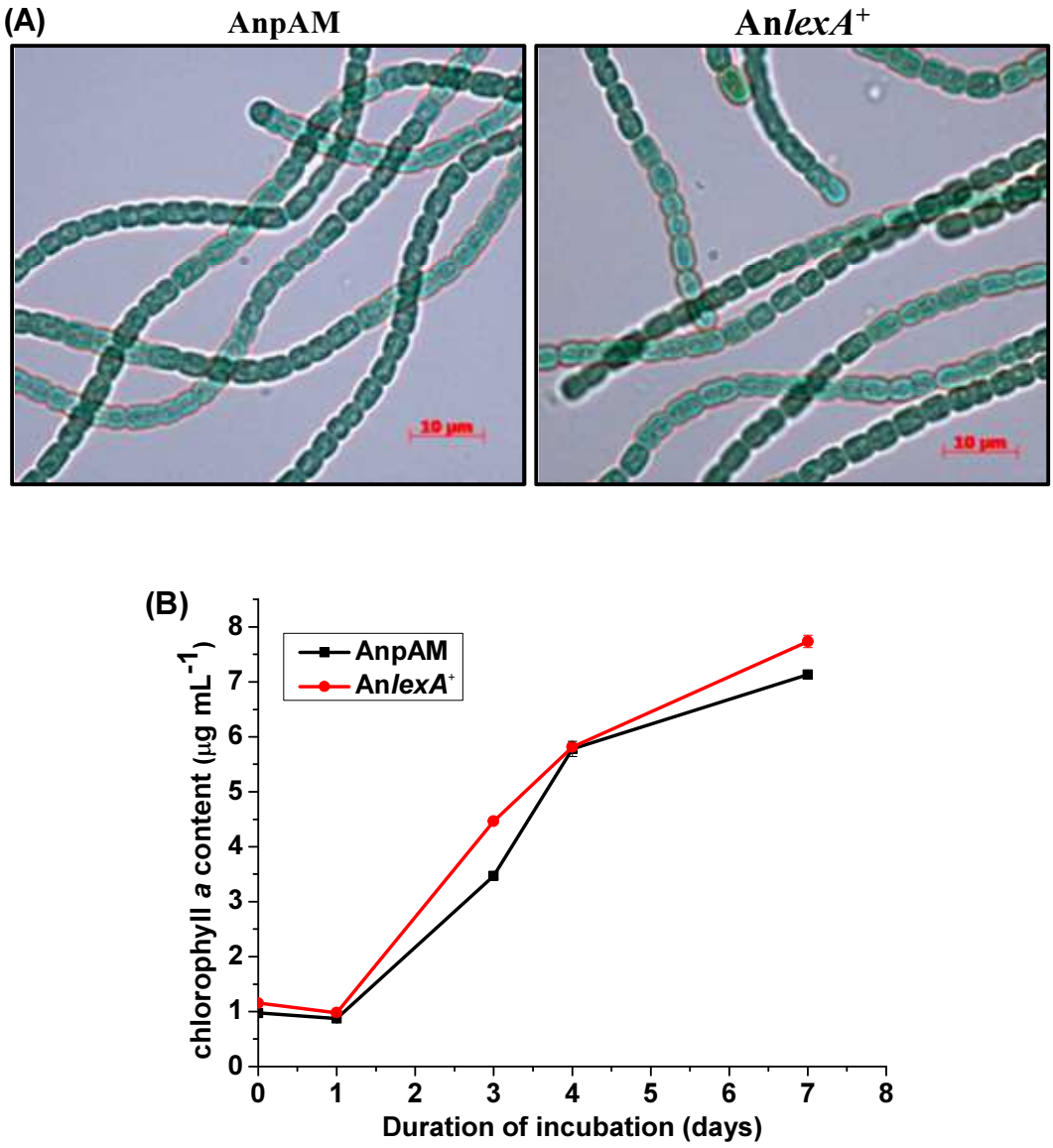


Fig. 4.6 Morphological and growth comparison of AnpAM and An*lexA*$^+$ strains. **(A)** Bright light micrographs (1500 X magnification) of AnpAM and An*lexA*$^+$ cells. **(B)** Growth of AnpAM and An*lexA*$^+$ under normal growth conditions over a 7-day period, measured in terms of chlorophyll *a* content.

4.2.3 *Effect of LexA overexpression on growth and stress tolerance of Anabaena*

The different abiotic stresses applied included C-starvation, exposure to different DNA damage and oxidative stress inducing agents. The choice of testing C-starvation stress was based on earlier results in the unicellular cyanobacterium, *Synechocystis* wherein LexA-depletion resulted in decreased tolerance to C-starvation (*127*). The known increase in sensitivity of *E. coli* to DNA damage inducing agents upon rendering the LexA non-cleavable (*172-173*) formed the basis of analysing the response of An*lexA*$^{+}$ strains to DNA damage inducing agents. Being photosynthetic organisms, they are continuously exposed to intrinsic oxidative stress, and hence, the possible role of LexA in modulating tolerance to extrinsic oxidative stresses was also assessed.

4.2.3.1 EFFECT OF DCMU-INDUCED CARBON STARVATION ON LexA OVEREXPRESSING *Anabaena* CELLS

Inorganic C-starvation in photosynthetic organism can be induced by inhibiting photosynthesis. This can be carried out by either exposing the cells to DCMU [3-(3,4-dicholophyllphenyl)-1,1-dimethylurea)], an inhibitor of electron flow from PSII to PSI in the Z-scheme (*174*) or by keeping the cultures in dark. Since, the overexpression of LexA protein is through a light inducible promoter, *psbA1*, keeping the cultures in dark to inhibit photosynthesis would also result in the inhibition of expression of LexA. Hence, DCMU was used at concentration of 5 μM to induce C-starvation. Under normal growth conditions, the photosynthetic efficiency of An*lexA*$^{+}$ was found to be 7.1 ± 0.2 μmol O_2 mg Chl a^{-1} min^{-1}, comparable to that of AnpAM cells, estimated as 7.4 ± 0.1 μmol O_2 mg Chl a^{-1} min^{-1}. Three-day-old cultures of AnpAM and An*lexA*$^{+}$ were freshly

inoculated at chlorophyll *a* density of 1 µg mL^{-1} and then subjected to 5 µM DCMU at 27 °C under continuous illumination for 3 days. At the end of 2 days, AnpAM cultures bore a bleached look, while the An*lexA*$^{+}$ cells were greener **(Fig. 4.7A)**. AnpAM cells exhibited continuous decrease in chlorophyll content over 3 days of DCMU treatment, while that of An*lexA*$^{+}$ remained nearly constant under these conditions **(Fig. 4.7B)**. When compared with their respective unstressed control, the survival of An*lexA*$^{+}$ strain was estimated to be 29% as against only 11% for AnpAM **(Fig. 4.7B)**. This indicated that LexA overexpression aids *Anabaena* cells to overcome induced C-starvation, and thus beneficial under these conditions.

4.2.3.2 DNA-DAMAGE INDUCING STRESSES: MODULATION OF TOLERANCE

DNA damage can be induced using agents such as mitomycin C and UV-radiation which result in the formation of DNA adducts, or by exposing the cells to γ-radiation or desiccation, which introduce single and double strand breaks in DNA.

UV-radiation and Mitomycin C stresses:

Three-day-old culture of An*lexA*$^{+}$ and AnpAM were concentrated to 10 µg Chl *a* mL^{-1}, exposed to different doses of UV-radiation (0 - 3.0 kJ m^{-2}) or mitomycin C (0 - 5 µg mL^{-1}) stresses, followed by analysis of survival in terms of colony forming units (CFUs). Survival of An*lexA*$^{+}$ cells was comparable to that of AnpAM under control growth conditions. However, upon exposure to 0.75 kJm^{-2} of UV-radiation, An*lexA*$^{+}$ cells showed only 34% survival compared to AnpAM which was further reduced to 10% at 1.5 kJm^{-2} **(Fig. 4.8A)**. When compared to their respective controls, the survival upon exposure to 0.75 kJ m^{-2} and 1.5 kJ m^{-2} of UV-radiation was found to be 20.4% and 4.1% respectively for An*lexA*$^{+}$ cells as against 50.2% and 26.2% for AnpAM cells for the

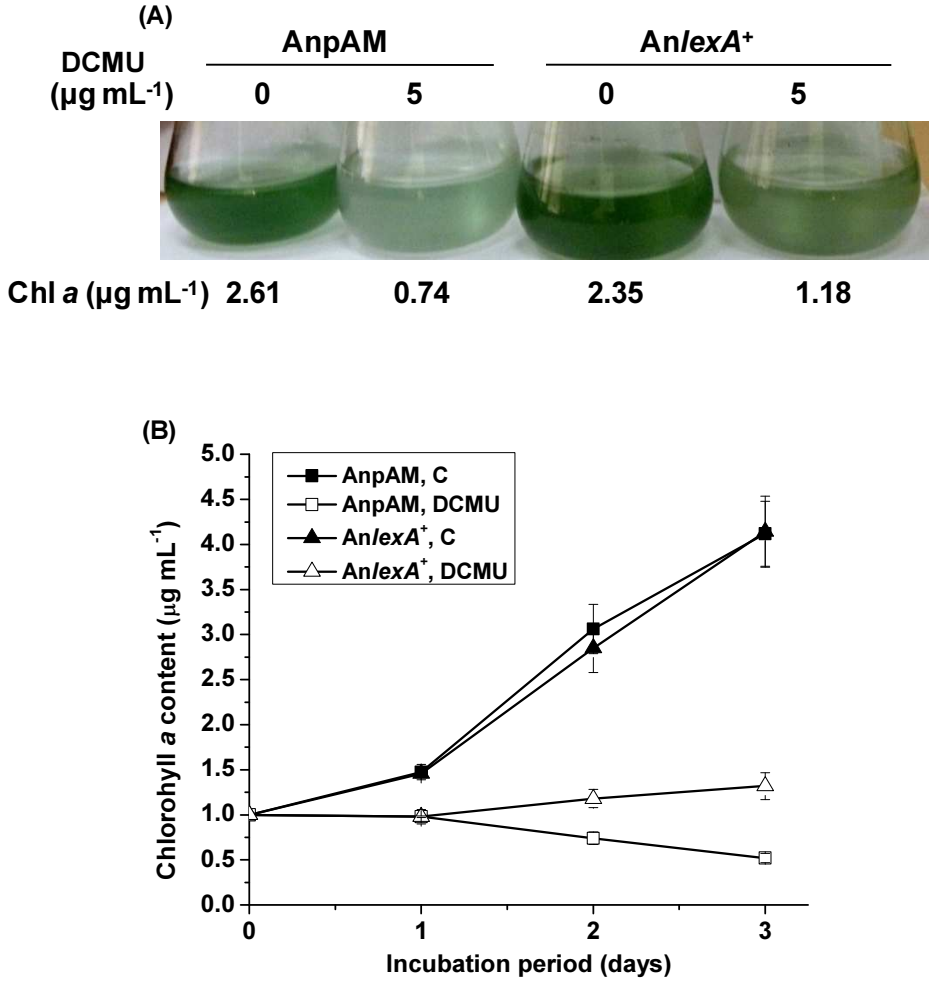


Fig. 4.7 Effect of LexA overexpression on tolerance of *Anabaena* to DCMU induced C-starvation. Three-day-old cultures of recombinant *Anabaena* strains, AnpAM and An*lexA*$^+$ were inoculated in fresh BG-11, N^+ medium at 1 µg Chl *a* mL^{-1}, followed by exposure to 0 and 5 µg DCMU mL^{-1}. **(A)** Image of 2-day-old *Anabaena* cultures, as indicated, in conical flasks and the corresponding Chl *a* readings. Images were captured with Nikon D5100 (DSLR) camera, Japan. **(B)** Growth of the above cultures measured in terms of Chl *a* content over a 3 day period.

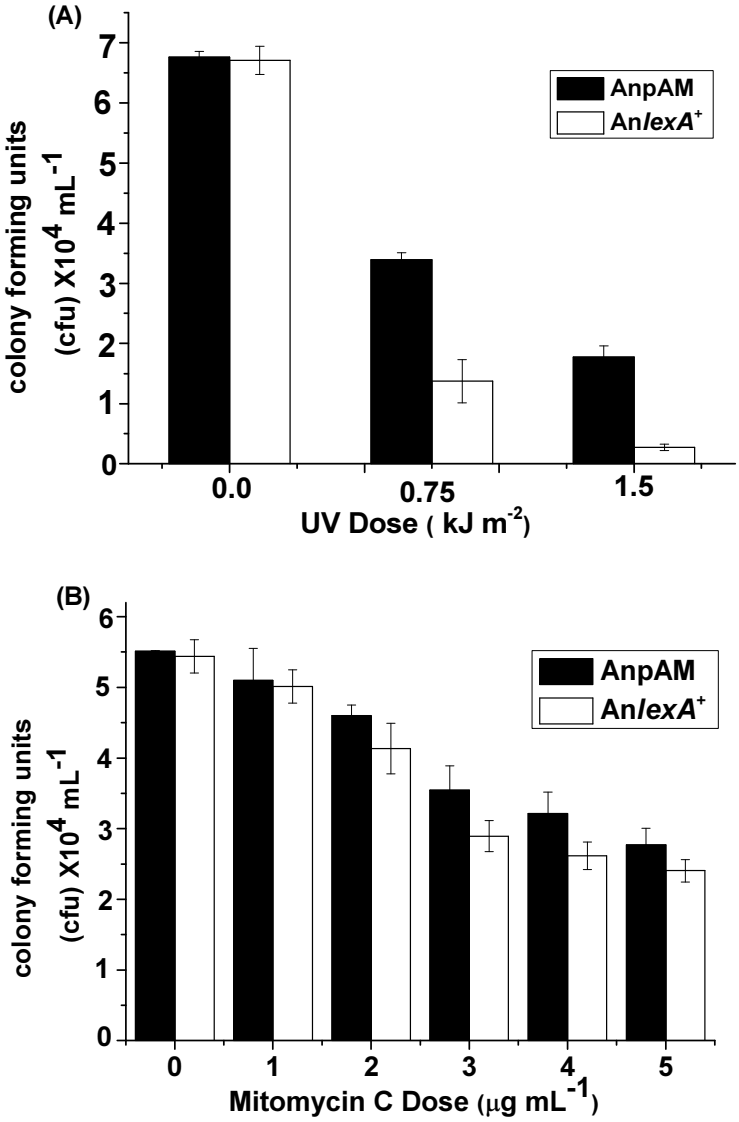


Fig. 4.8 Growth of AnpAM and An*lexA*$^+$ cultures exposed to UV-radiation and Mitomycin C. The recombinant *Anabaena* cultures, AnpAM and An*lexA*$^+$ (10 µg Chl *a* density mL^{-1}), subjected to **(A)** 0-5 µg mL^{-1} mitomycin C for 30 min, followed by plating on BG-11, N$^+$ agar plates and **(B)** UV-radiation (0 - 1.5 kJ m^{-2}) on BG-11, N$^+$ agar plates. Plates were incubated under continuous illumination and control growth conditions for 10 days and colonies were counted.

corresponding UV-exposure **(Fig. 4.8A)**. Exposure to different concentrations of mitomycin C, however, did not result in any significant change in the tolerance of An*lexA*$^{+}$ with respect to that of the AnpAM cells **(Fig. 4.8B)**.

γ-radiation and desiccation stresses

Three-day-old concentrated cultures (10 µg Chl *a* mL^{-1}) of An*lexA*$^{+}$ and AnpAM were exposed to 6 kGy γ-radiation or 6 days of desiccation for which the corresponding controls used were sham-irradiated and 6 days in humid chamber. Upon exposure to 6 kGy of γ-radiation, survival of An*lexA*$^{+}$ was found to be 27.2% as against 59% for AnpAM cells **(Fig. 4.9A)**. These cells were allowed to recover in fresh growth media for 8 days. While the chlorophyll *a* content of irradiated AnpAM cells increased from 1 to 3 µg mL^{-1} over an 8-day period, that of An*lexA*$^{+}$ decreased marginally to 0.8 µg mL^{-1} **(Fig. 4.9B)**. On day 8 of post-irradiation recovery, compared to their respective unirradiated controls, the growth was found to be 17.7% for An*lexA*$^{+}$ compared to 48.33% for AnpAM **(Fig. 4.9B)**. The 2.7-fold lower recovery of An*lexA*$^{+}$ was distinctly visible in the cultures in the conical flasks **(Fig. 4.9C)**. Upon exposure to 6 days of desiccation, An*lexA*$^{+}$ cells fared very poorly exhibiting only 16% survival compared to 6-day desiccated AnpAM cells. When compared to their respective humid chamber controls, the survival was found to be 8.9% for An*lexA*$^{+}$ as against 53.4% for AnpAM (**Fig. 4.10A**). Upon post-desiccation recovery, AnpAM cells attained 3.9 µg Chl *a* mL^{-1} density from the initial density of 1 µg mL^{-1} on day 8 (**Fig. 4.10B**). On the other hand, the chlorophyll *a* content of An*lexA*$^{+}$ showed a continuous decrease till day 4 of recovery, followed by resumption of growth resulting in 1.9 µg Chl *a* density mL^{-1} at the end of 8 days (**Fig. 4.10B**). Compared to their respective controls, the recovery of

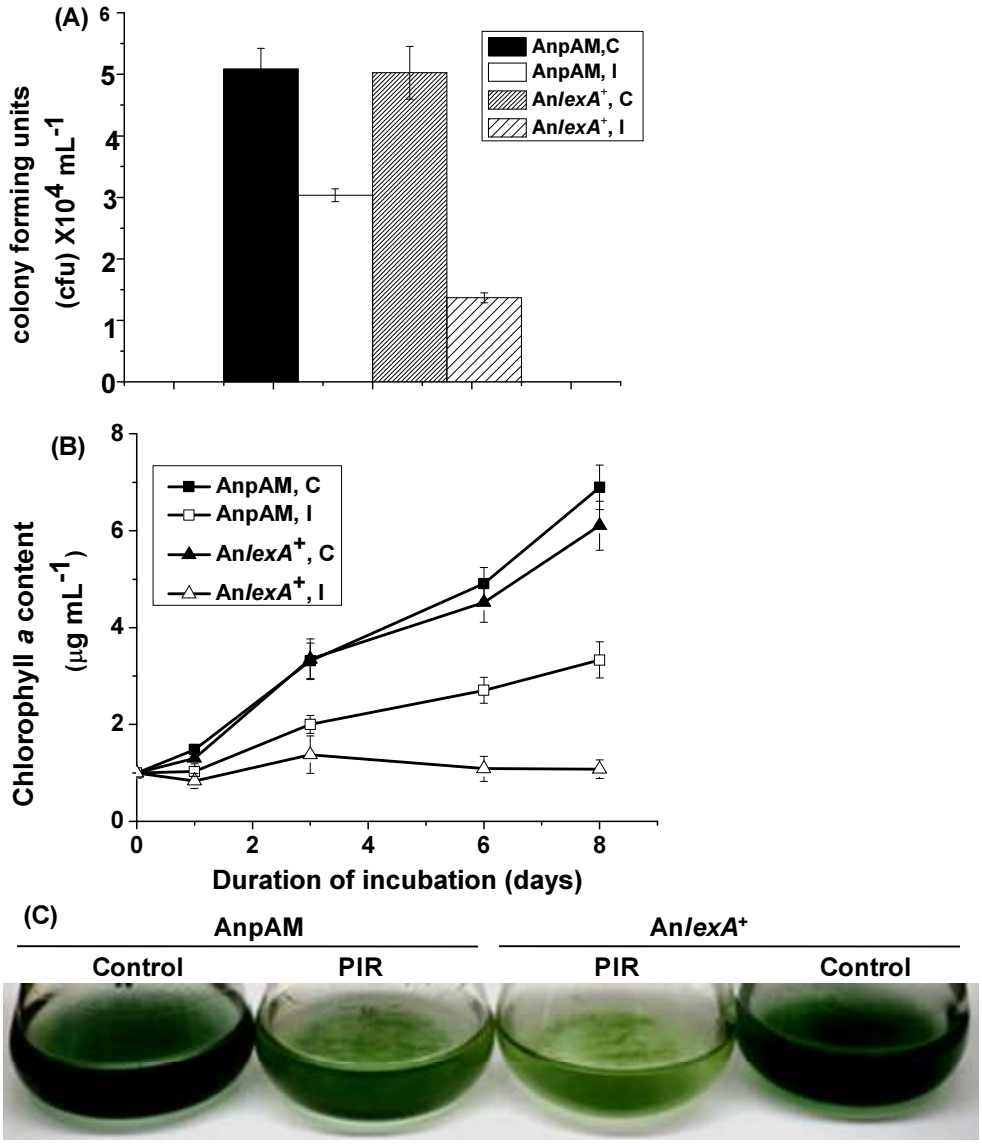


Fig. 4.9 Effect of LexA overexpression on tolerance of *Anabaena* to γ-irradiation. Three-day-old cultures of recombinant *Anabaena* strains, AnpAM and An*lexA*$^+$ (10 µg Chl *a* density mL^{-1}) were subjected to **(A)** ^{60}Co γ-irradiation (6 kGy) 'I' compared with corresponding sham-irradiated control 'C', followed by plating and incubation under constant illumination for 10 days for optimal growth of distinct colonies. **(B)** Post-irradiation recovery (PIR) was determined by inoculating the stressed and unstressed samples in fresh growth medium at 1 µg Chl *a* ml^{-1} and allowed to grow under control growth conditions, followed by measuring Chl *a* content over a period of 8 days. **(C)** Image of AnpAM and An*lexA*$^+$ cultures in flasks after 6 days of post-irradiation recovery (PIR) with their corresponding controls.

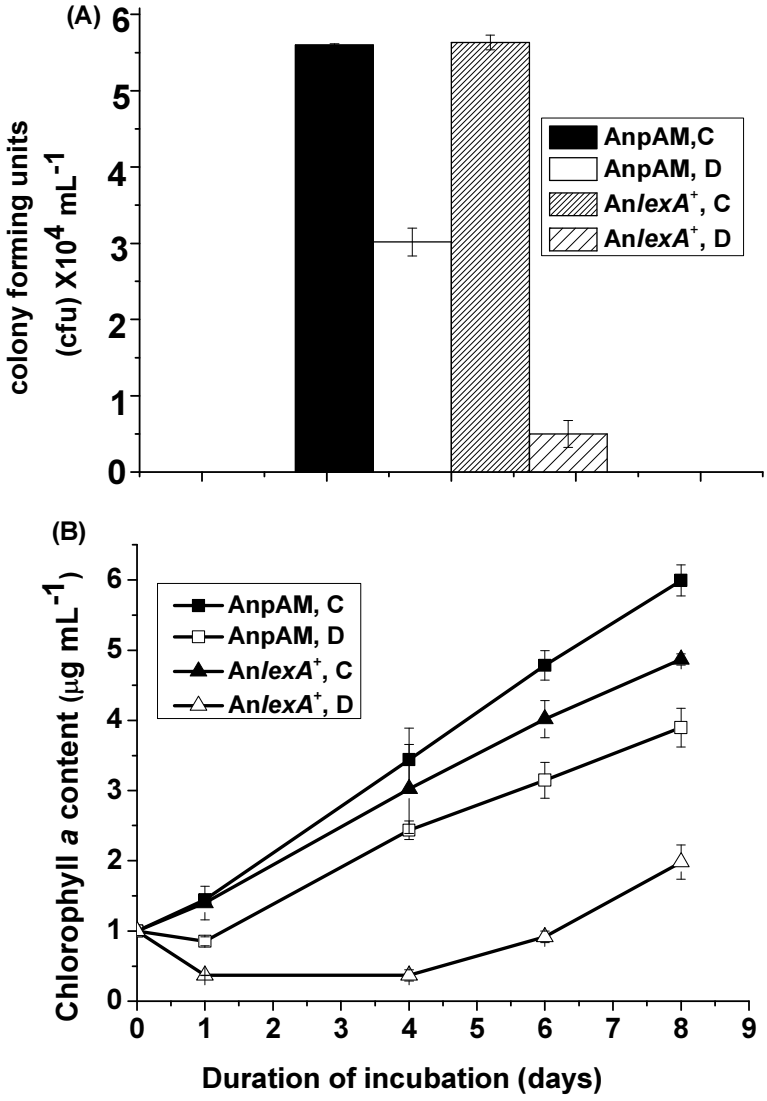


Fig. 4.10 Effect of LexA overexpression on tolerance of *Anabaena* to desiccation stress. Three-day-old cultures of AnpAM and An*lexA*$^{+}$ cells (10 μg Chl *a* density mL^{-1}) were subjected to **(A)** six days of desiccation 'D' compared with corresponding humid chamber control 'C', followed by plating and incubation under constant illumination for 10 days for optimal growth of distinct colonies. **(B)** Post-desiccation recovery (PDR) was determined by inoculating the stressed and unstressed samples in fresh growth medium at 1 μg Chl *a* ml^{-1}, and allowed to grow under control growth conditions, followed by measuring Chl *a* content over a period of 8 days.

AnpAM was 64.9% as against 40.6% for An*lexA*$^{+}$ cells (**Fig. 4.10B**). Thus, overexpression of LexA hampered recovery of *Anabaena* post DNA damage.

4.2.3.3 EFFECT ON OXIDATIVE STRESS TOLERANCE

The recombinant *Anabaena* cultures were exposed to different forms of oxidative stresses which included exposure to H_2O_2, Methyl Viologen (MV, which induces superoxide radicals) and heavy metals [Cd(II) and As(V)].

H_2O_2-mediated oxidative stress

Freshly inoculated An*lexA*$^{+}$ and AnpAM cultures (1 µg Chl *a* mL^{-1}) were exposed to different concentrations of H_2O_2 (0-0.4 mM) for 2 days. As observed visually in microtitre plates, no green colouration was detected in An*lexA*$^{+}$ exposed to 0.3 mM and higher concentrations of H_2O_2, while AnpAM cells exhibited faint green colour even after exposure to 0.4 mM H_2O_2 **(Fig. 4.11A)**. The qualitative difference in chlorophyll *a* content of AnpAM and An*lexA*$^{+}$ cells was apparent even at 0.2 mM H_2O_2 **(Fig. 4.11A)**. When expressed in terms of chlorophyll *a* content of the corresponding unstressed cultures, the percent growth upon exposure to 0.2 mM and 0.3 mM of H_2O_2 at the end of 2 days of An*lexA*$^{+}$ cells was found to be 55% and 1.4% respectively, while that of AnpAM was 97.6% and 76% respectively **(Fig. 4.11B)**. This indicated that LexA overexpression rendered *Anabaena* cells sensitive to H_2O_2-mediated oxidative stress. LexA overexpression also resulted in accumulation of more ROS under normal growth conditions as well as when treated with 0.2 mM H_2O_2 as apparent upon NBT-staining, wherein more number of black granules were observed in An*lexA*$^{+}$ cells compared to AnpAM cells under control as well as H_2O_2-treated conditions **(Fig. 4.12A)**. One of the important enzymes involved in detoxification of peroxides, Catalase B (KatB), exhibited

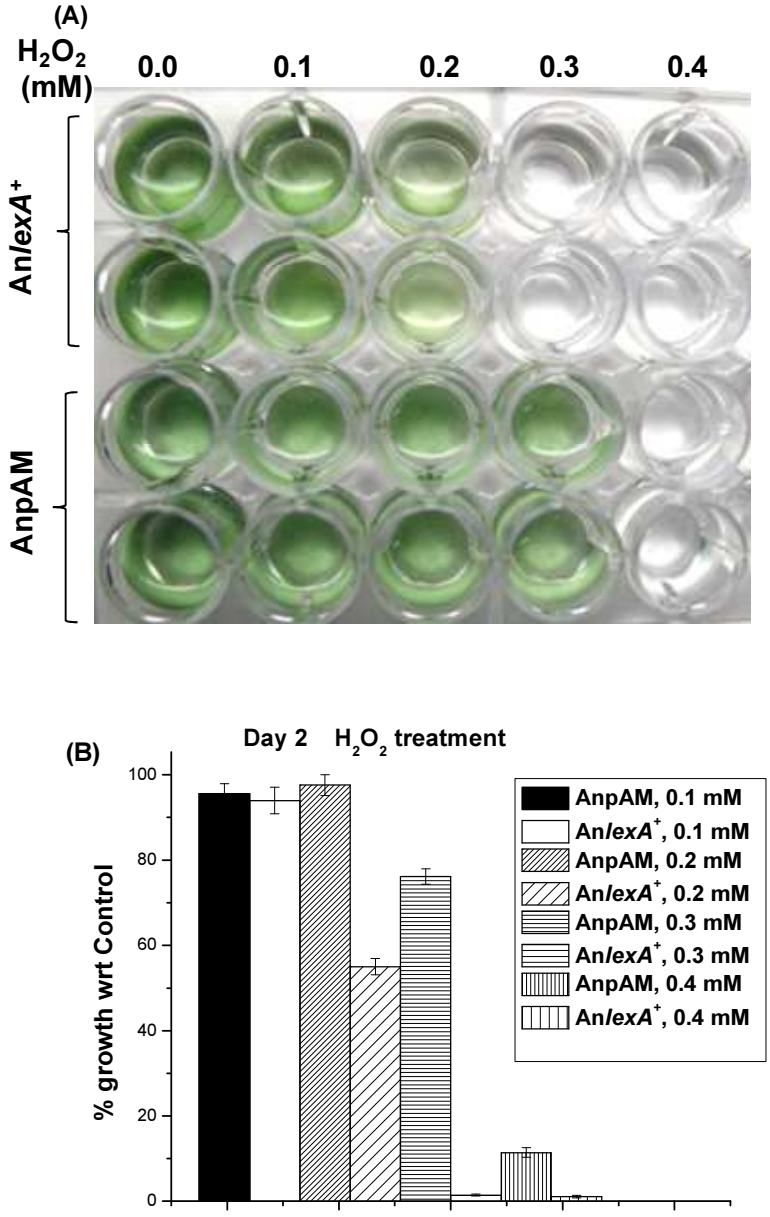


Fig. 4.11 Effect of LexA overexpression on tolerance of *Anabaena* to H_2O_2-induced oxidative stress. Freshly inoculated (at 1 µg Chl *a* mL^{-1}) cultures of AnpAM and AnlexA^+ were exposed to 0- 0.4 mM H_2O_2 for 2 days in microtitre plates. **(A)** Image of the cultures in microtitre plates captured with Nikon D5100 (DSLR) camera, Japan. **(B)** Growth (Chl *a* content) upon exposure to different concentration of H_2O_2 on day 2, expressed as percentage of that under control (0 mM H_2O_2) conditions.

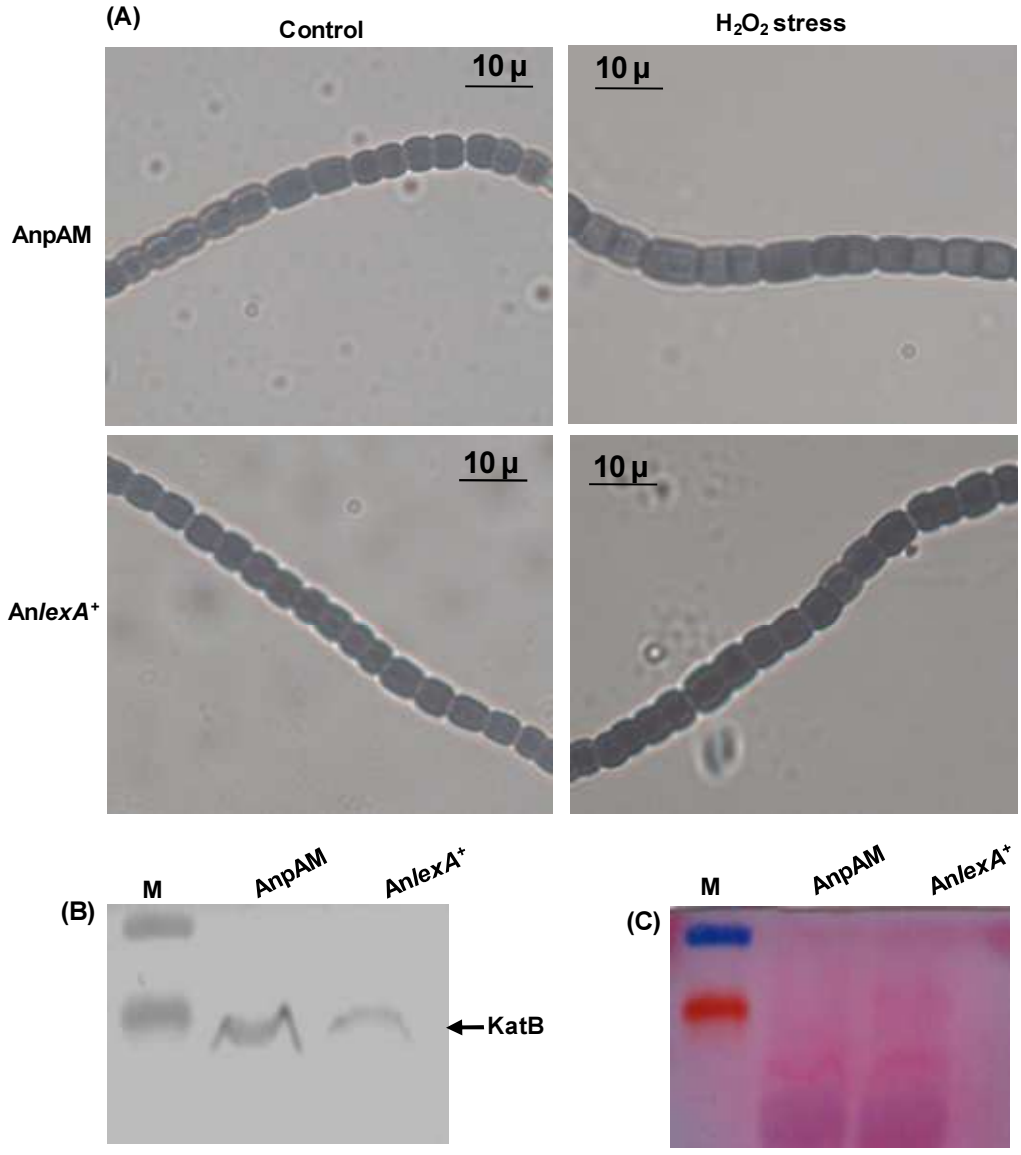


Fig. 4.12 Effect of LexA overexpression on accumulation of ROS and KatB expression in *Anabaena*. (A) Bright field microphotographs (1000X-magnification) of NBT-stained cells of AnpAM and An*lexA*$^+$ under control conditions and after 90 min exposure to 0.2 mM H_2O_2. Black dots indicate ROS accumulation. **(B)** Western blotting and immunodetection of KatB in 3-day-old AnpAM and An*lexA*$^+$ cells grown under control conditions. Proteins (100 µg/lane) were electrophoretically separated by 15% SDS-PAGE, blotted on to nitrocellulose membrane and immunodetected with anti-KatB antibody protein. The cross-reacting 26 kDa KatB protein is indicated by an arrow. Lane 'M' refers to pre-stained molecular weight marker. (C) Ponceau stain of the blot shown in 'B'.

lower levels in LexA-overexpressing strains of *Anabaena* **(Fig. 4.12B)**, indicating that *Anabaena katB* is negatively regulated by LexA, which may be contributing to the observed sensitivity of An*lexA*$^{+}$ cells to H_2O_2 stress.

Oxidative stress induced by Methyl Viologen (MV)

When exposed to 0.2 μM MV for 3 days, An*lexA*$^{+}$ cells was visually found to be less green than AnpAM cells **(Fig. 4.13A)**. While the AnpAM cells exhibited a slow growth with increase in chlorophyll *a* content from 1 μg mL^{-1} to 1.7 μg mL^{-1} over a 3-day period, that of An*lexA*$^{+}$ decreased to 0.8 μg mL^{-1} **(Fig. 4.13B)**. Compared with their respective unstressed controls An*lexA*$^{+}$ exhibited only 27.3% growth as against 48.5% in case of AnpAM on day 3 **(Fig. 4.13B)**. Susceptibility to MV-induced oxidative stress usually correlates with decrease in the activity of the superoxide dismutase (SOD) enzymes in the organism. SOD zymogram analysis of 3-day-old cultures grown under control conditions revealed negligible MnSOD activity in An*lexA*$^{+}$ cells compared to AnpAM cells, while that of FeSOD was similar in the two strains **(Fig. 4.13C)**. Thus, the presence of only FeSOD in An*lexA*$^{+}$ allowed it to combat internal oxidative stress due to photosynthetic and respiratory electron transport chain, but was unable to dismutate the additional superoxides generated during MV stress, resulting in cell death in case of An*lexA*$^{+}$ cells upon exposure to MV.

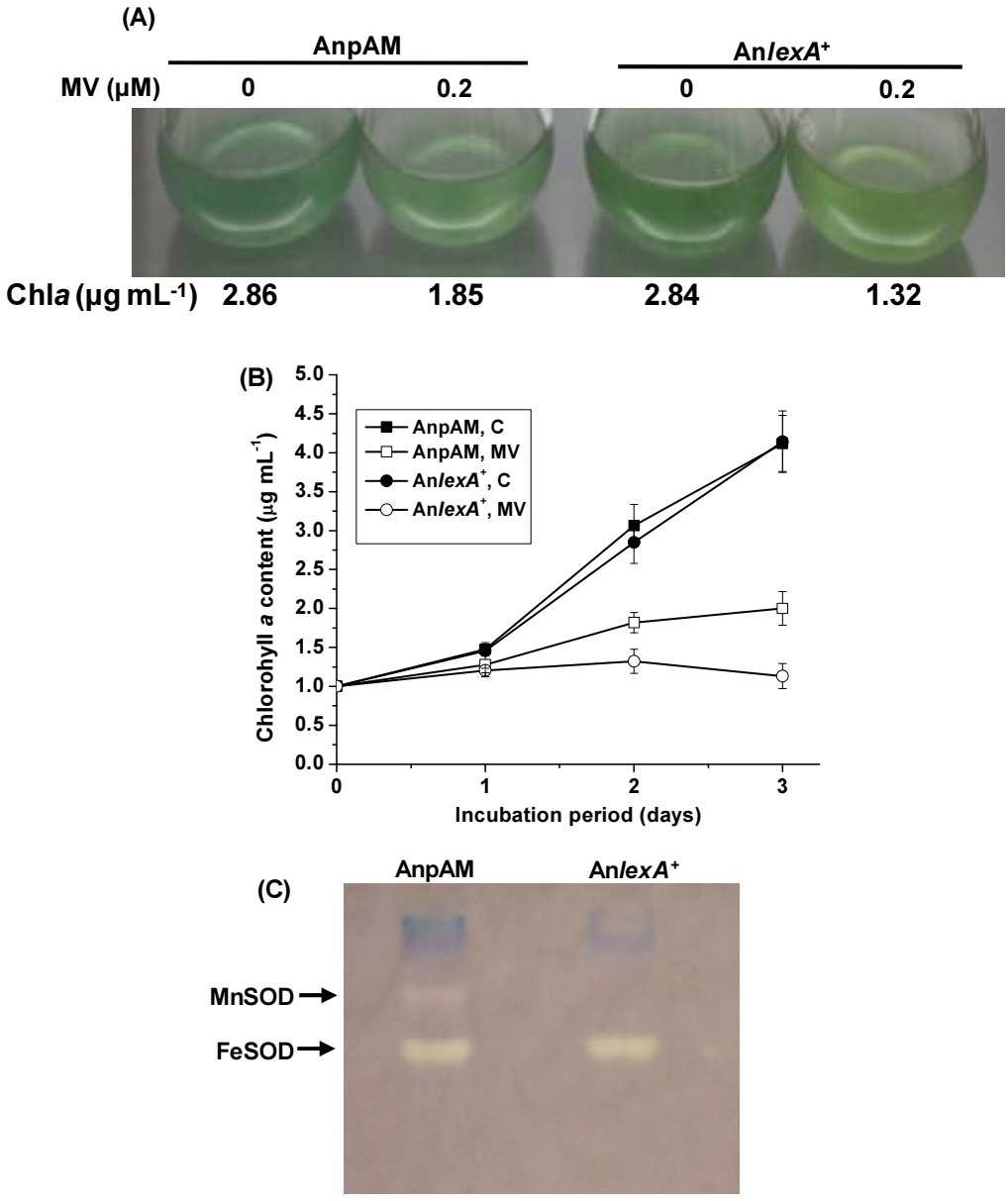


Fig. 4.13 Effect of LexA overexpression on tolerance of *Anabaena* to Methyl Viologen (MV)-induced oxidative stress. AnpAM and An*lexA*$^+$ freshly inoculated in BG-11, N^+ medium at 1 µg Chl *a* mL^{-1} were exposed to 0 or 0.2 µM MV. (A) Image of 2-day-old *Anabaena* cultures, as indicated, in conical flasks and the corresponding Chl *a* readings. (B) Growth of control and MV-treated cultures measured in terms of Chl *a* content over a 3 day period. (C) SOD zymogram of 3-day-old cultures of AnpAM and An*lexA*$^+$ grown under control conditions. The total proteins extracted under native conditions were separated on 8% native PAG, followed by staining for SOD activity using NBT. The MnSOD and FeSOD activity bands are indicated by arrows.

Heavy metal induced oxidative stress

Heavy metals are also known to induce oxidative stress (*175*). The effect of LexA overexpression on modulation of tolerance of *Anabaena* was investigated in response to two heavy metals, Cd(II) and As(V). Freshly inoculated An*lexA*$^{+}$ and AnpAM cultures (1 µg Chl*a* mL^{-1}) were exposed to either 0-20 µM $CdCl_2$ [Cd(II)] or 0-50 mM $Na_2HAsO_4.7H_2O$ [As(V)]. A distinct decrease in growth of An*lexA*$^{+}$ compared to AnpAM was visible upon exposure to 10 µM Cd(II) on day 3 **(Fig. 4.14A)**. Compared to their respective unstressed controls, An*lexA*$^{+}$ cells exhibited only 12% growth as against 39% by AnpAM cells on day 3 of exposure to 10 µM [Cd(II)] **(Fig. 4.14B)**. At concentrations beyond 10 µM Cd(II), neither An*lexA*$^{+}$ nor AnpAM cells exhibited any survival. An*lexA*$^{+}$ cells were found to be less green than AnpAM cells, when exposed to 10- 50 mM As(V) **(Fig. 4.15A)**. Exposure to 10 mM As(V), decreased the growth of AnpAM to 29% and at higher concentrations till 50 mM, the growth was in the range of 20 to 22% **(Fig. 4.15B)**. On the other hand, An*lexA*$^{+}$ exhibited 15% to 18% growth on day 3 upon exposure to 10 – 50 mM of As(V) **(Fig. 4.15B)**. Thus, the increased expression of LexA decreased the ability of *Anabaena* cells to tolerate heavy metal stress.

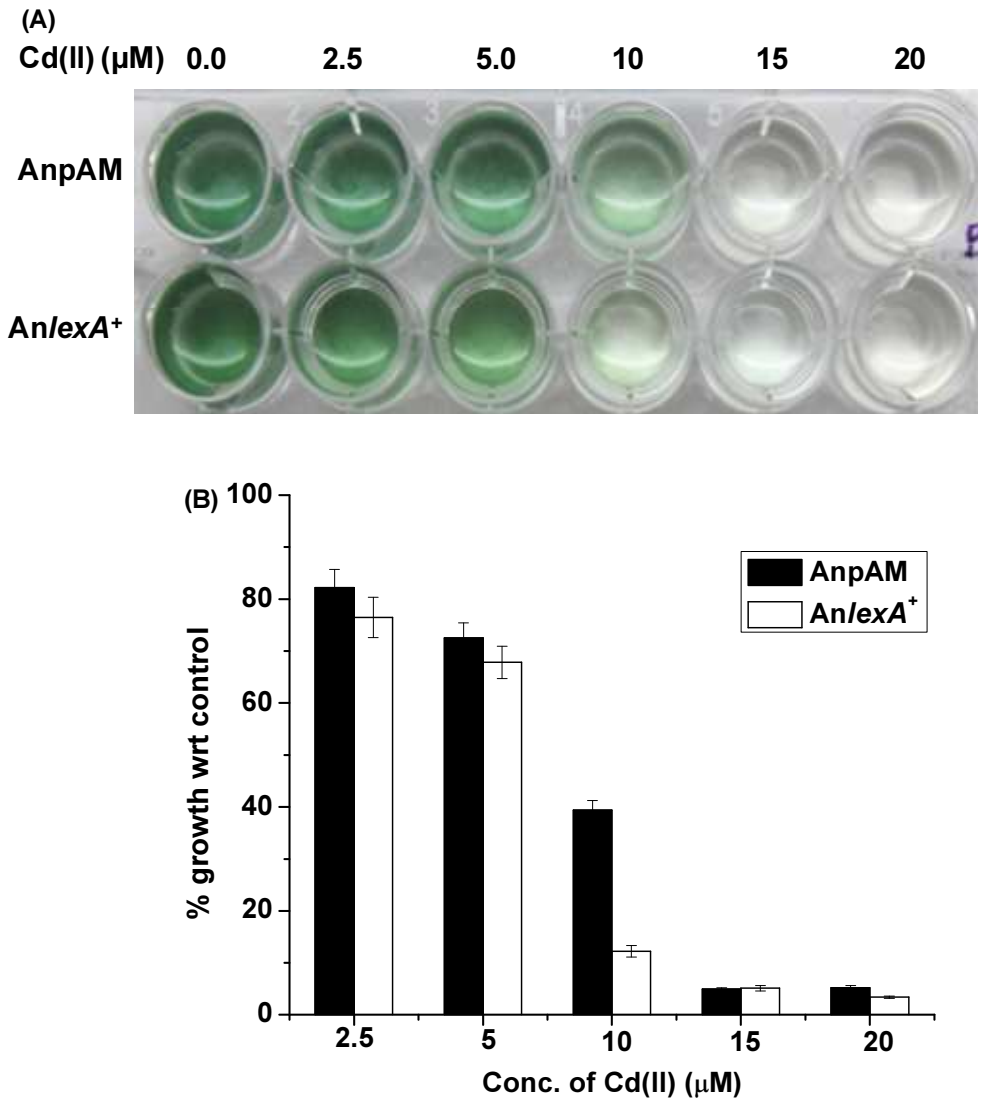


Fig. 4.14 Comparative tolerance of AnpAM and An*lexA*$^+$ strains to heavy metal Cd(II). AnpAM and An*lexA*$^+$ at 1 µg Chl *a* mL^{-1} were exposed to different concentrations of CdCl2 (0-20 µM). **(A)** Image of cultures in microtitre plates exposed to Cd(II) for 3 days. **(B)** Growth (measured in terms of Chl *a* content) on day 3 of stress compared to the unstressed control after exposure to different concentrations of $CdCl_2$, expressed in percentage.

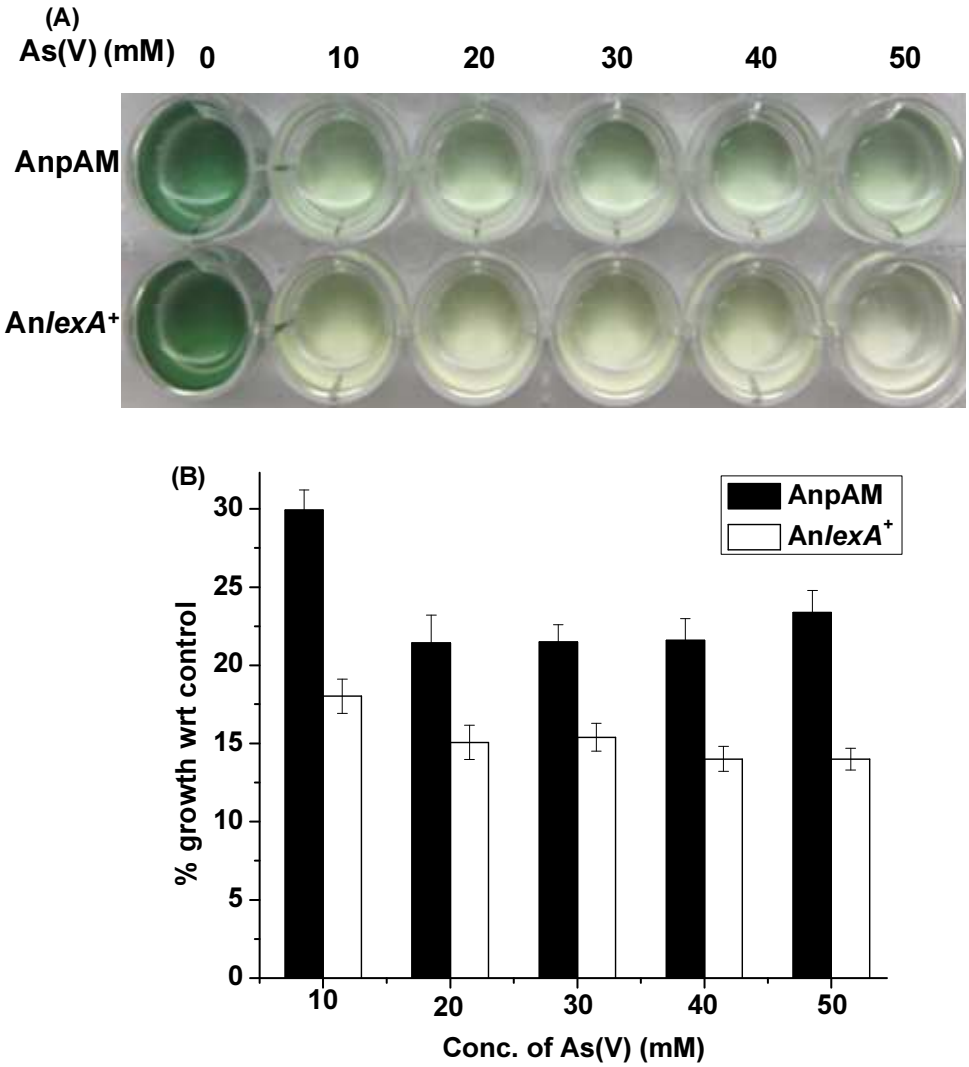


Fig. 4.15 Growth of AnpAM and An*lexA*$^+$ cells exposed to As(V). *Anabaena* cultures, AnpAM and An*lexA*$^+$ (1 µg Chl *a* mL^{-1}) were exposed 0-50 mM $Na_2HAsO_4.7H_2O$ in microtitre plates for 3 days. **(A)** Image of cultures in microtitre plates exposed to As (V) for 3 days. **(B)** Growth was measured in terms of Chl *a* content on day 3 for all cultures and expressed as percentage of that of unstressed control.

4.2.4 *Analysis of proteomic changes upon overexpression of LexA in Anabaena*

Modulation of abiotic stress response of *Anabaena* upon overexpression of LexA, suggested possible modulation of the levels of stress responsive proteins by LexA. This was assessed by two-dimensional (2D) proteomic analysis of the two recombinant *Anabaena* strains, AnpAM and An*lexA*$^{+}$ grown under normal growth conditions **(Fig. 4.16)**. A total of 119 spots were detected by PD Quest 2D analysis software, common to both the strains and observed in all three gel replicates for each strain. Out of the 119 protein spots, 47 protein spots exhibited altered expression i.e. a fold change of 1.2 or higher in all three replicates. Of these, 30 spots showed decreased abundance (fold change ≤ 1.2) in An*lexA*$^{+}$ compared to AnpAM, indicated with red arrow heads in **Fig. 4.16A**. The remaining 17 protein spots exhibited increased abundance (fold change ≥ 1.2) in An*lexA*$^{+}$, indicated with green arrow heads in **Fig. 4.16B**. The 47 protein spots were excised out from the corresponding gels, digested with trypsin after appropriate treatments and subjected to MALDI-ToF-ToF based protein identification as described in **Section 2.4.5**. A representative figure corresponding to one of the identified proteins is shown in **Fig. 4.17**. The mascot score represents the negative logarithm of the probability of observed match that is random event. The peptides shown in red are peptides matched from the database sequence. These criteria are used to determine the significance attributed to the identity of the proteins. A total of 23 proteins could be identified, of which 15 were those with the decreased abundance in An*lexA*$^{+}$ **(Fig. 4.16A)** and 8 with increased abundance in An*lexA*$^{+}$ (**Fig. 4.16B**). In case of multiple protein spots showing the same identification, the protein name was suffixed with a hyphen and a number, e. g. CpcG-1, CpcG-2 etc.

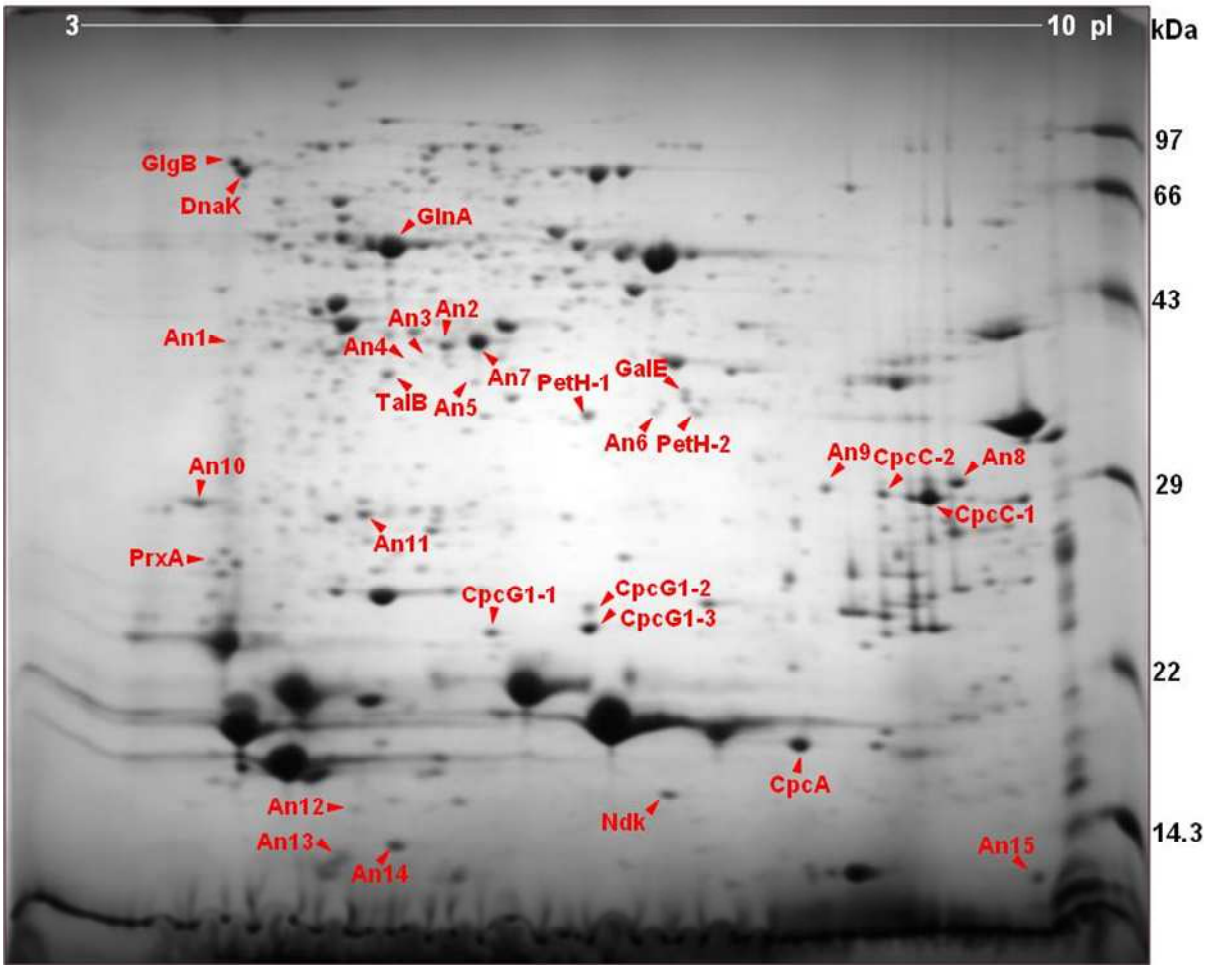


Fig. 4.16 Representative proteome profiles of recombinant *Anabaena* strains, AnpAM and An*lexA*$^{+}$. Total cellular proteins (1 mg) of 3-day-old cultures of AnpAM and *AnlexA*$^{+}$ were resolved first on on IPG strips (pH 3-10 NL, 17 cm) by Iso-electric Focussing (IEF), followed by 14% SDS-PAGE. Proteins spots were visualized by CBB staining. **(A)** Proteins with decreased abundance are marked with red arrow heads in the proteome profile of AnpAM cells.

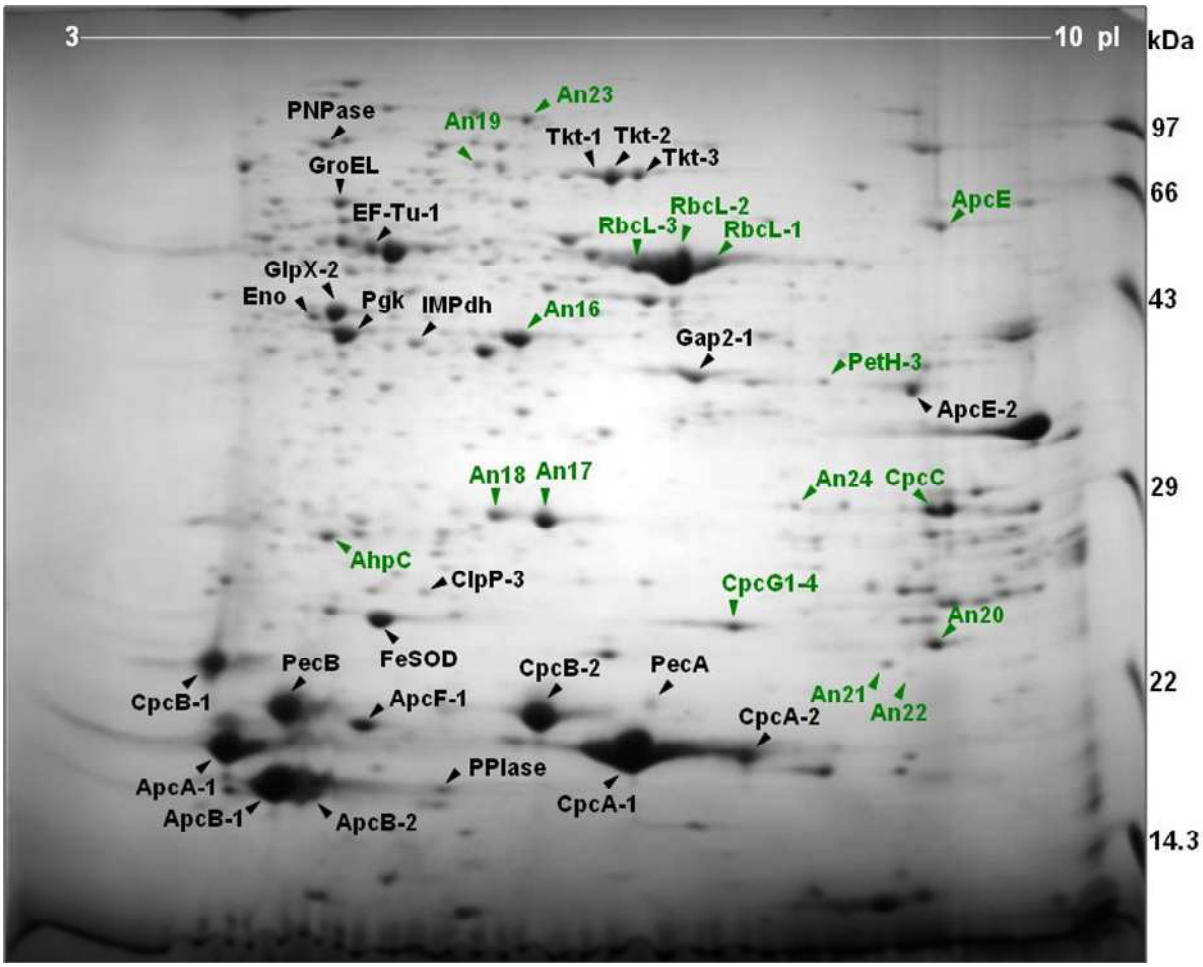
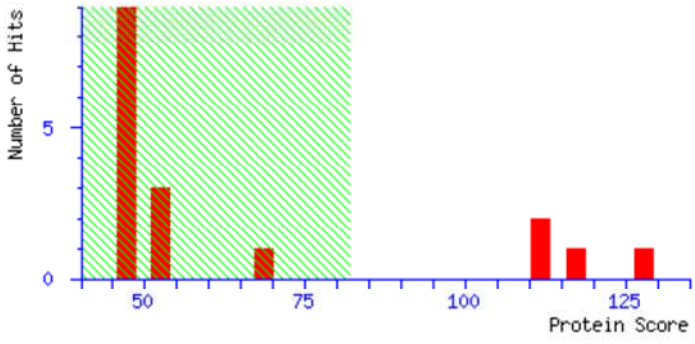


Fig. 4.16 (B) Proteins with increased abundance upon overexpression of LexA are marked with green arrow heads in the proteome profile of An*lexA*+ cells. The proteins indentified during the course of this study are indicated by the protein names while the unidentified proteins are indicated by numbers prefixed with 'An'. The protein spots marked with black arrow heads are those showing no significant change in abundance upon LexA overexpression and identified on the basis of earlier published data (*159, 176-177*)

(A) Mascot Score Histogram



(B)

gi|499307485 **Mass:** 48979 **Score:** 114 **Expect:** 3.2e-05 **Matches:** 18

ferredoxin--NADP(+) reductase [Nostoc sp. PCC 7120]

(C)

Protein sequence coverage: 45% & Matched peptides shown in ***bold red***.

```
  1 MSNQGAFDGA ANVESGSRVF VYEVVGMRQN EETDQTNYPI RKSGSVFIRV
 51 PYNRMNQEMQ RITRLGGKIV SIQTVSALQQ LNGRTTIATV TDASSEIAKS
101 EGNGKATPVK TDSGAKGFAK PPAEEQLKKK DNKGNTMTQA KAKHADVPVN
151 LYRPNAPFIG KVISNEPLVK EGGIGIVQHI KFDLTGGNLK YIEGQSIGII
201 PPGVDKNGKP EKLRLYSIAS TRHGDDVDDK TISLCVRQLE YKHPESGETV
251 YGVCSTYLTH IEPGSEVKIT GPVGKEMLLP DDPEANVIML ATGTGIAPMR
301 TYLWRMFKDA ERAANPEYQF KGFSWLVFGV PTTPNILYKE ELEEIQQKYP
351 DNFRLTYAIS REQKNPQGGR MYIQDRVAEH ADELWQLIKN EKTHTYICGL
401 RGMEEGIDAA LSAAAAKEGV TWSDYQKDLK KAGRWHVETY
```

(D)

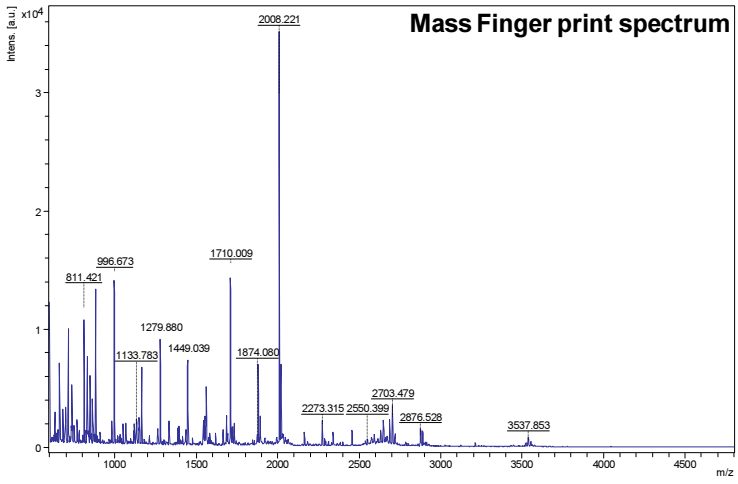


Fig. 4.17 Protein identification of Ferrodoxin NADP (+) Reductase, (PetH-1) by MALTI-ToF-ToF analysis. **(A)** Mascot Score Histogram **(B)** Concise protein summary **(C)** Protein Sequence coverage **(D)** Peptide mass finger print of the identified protein

Table 4.3: Proteins of recombinant *Anabaena* strains, AnpAM and An*lexA*$^{+}$ identified by MALDI ToF/ToF mass spectrometry

Sr. no.	Swiss Prot ID	Annotated ORF No	Description (Abbreviation used in the manuscript)	Score	Seq. coverage	Theoretical MW/pI	Observed MW/ pI	Fold Change ± SD	*p* value
Photosynthesis-Light Reaction									
1	P80559	*alr0020*	Phycobilisome core-membrane linker protein (ApcE)	126	24%	126.981/9.41	62.3/8.95	2.826 ± 0.151	0.003
2	P07121	*alr0529*	Phycocyanin A subunit (CpcA-3)	68	62%	17.508/6.82	16.6/8.1	-1.532 ± 0.057	0.049
3	P07123	*alr0530*	Phycocyanin associated Rod linker protein (CpcC-1)	96	40%	32.910/9.33	30.9/8.9	-1.409 ± 0.409	0.036
4	P07123	*alr0530*	Phycocyanin associated Rod linker protein (CpcC-2)	104	44%	32.910/9.33	30.9/8.6	-2.499 ± 0.376	0.0414
5	P07123	*alr0530*	Phycocyanin associated Rod linker protein (CpcC-3)	89	37%	32.910/9.33	30.9/8.8	2.148 ± 0.348	0.001
6	P29986	*alr0534*	Phycobilisome rod-core linker protein (CpcG1-1)	112	41%	31.915/9.11	22/6.1	Not detected in An*lexA*$^{+}$	-
7	P29986	*alr0534*	Phycobilisome rod-core linker protein (CpcG1-2)	78	33%	31.915/9.11	23.3/6.75	-7.306 ± 0.599	0.005
8	P29986	*alr0534*	Phycobilisome rod-core linker protein (CpcG1-3)	96	37%	31.915/9.11	22.3/6.75	-1.332 ± 0.093	0.031
9	P29986	*alr0534*	Phycobilisome rod-core linker protein (CpcG1-4)	93	35%	31.915/9.11	23.6/7.4	1.411 ± 0.111	0.035
Carbon fixation									
10	P00879	*alr1524*	Ribulose bisphosphate carboxylase/oxygenase (RbcL-1)	137	35%	53.411/6.25	56.2/7.3	2.303 ± 0.269	0.003
11	P00879	*alr1524*	Ribulose bisphosphate carboxylase/oxygenase (RbcL-2)	129	34%	53.411/6.25	56.2/7.1	1.386 ± 0.095	0.029
12	P00879	*alr1524*	Ribulose bisphosphate carboxylase/oxygenase (RbcL-3)	174	36%	53.411/6.25	56.2/6.9	1.247 ± 0.079	0.017
Carbon metabolism: Glycolysis and Pentose phosphate pathway									

13	P58561	*all2563*	Transaldolase B (TalB)	73	23%	36.256/5.19	42.0/5.4	-1.297 ± 0.024	0.023
14	Q8YN57	*all4713*	UDP-glucose 4-epimerase (GalE)	54	50%	10.519/9.26	40.3/7.3	-3.139 ± 0.295	0.001
15	Q44528	*all0875*	Alpha gluconotransferase (GlgB)	110	33%	64.022/4.79	72.7/4.3	-12.186 ± 0.112	0.003
Oxidative Stress alleviation									
16	P58558	*all4121*	Ferredoxin NADP(+) Reductase (PetH-1)	114	45%	48.979/7.66	36.2/6.7	- 1.673 ± 0.159	0.049
17	P58558	*all4121*	Ferredoxin NADP(+) Reductase (PetH-2)	53	28%	48.979/7.66	35.6/7.4	Not detected in An*lexA*$^+$	-
18	P58558	*all4121*	Ferredoxin NADP(+) Reductase (PetH-3)	82	28%	48.979/7.66	38.0/8.25	Not detected in AnpAM	-
19	Q8YP05	*alr4404*	Peroxiredoxin (AhpC)	77	39%	23.856/4.92	29.6/5.05	1.579 ± 0.061	0.004
20	Q8YNC5	*alr4641*	Peroxiredoxin (PrxA)	127	59%	22.731/4.87	26.4/4.2	Not detected in An*lexA*$^+$	-
Transcription, translation and protein folding									
21	Q8YW74	*alr1742*	DnaK (DnaK)	145	35%	68.037/4.84	70.7/4.4	-1.257 ± 0.048	0.046
22	Q8YRP2	*alr3402*	Nucleoside diphosphate kinase (Ndk)	70	36%	16.594/5.95	14.7/7.2	-1.299 ± 0.078	0.049
Amino acid biosynthesis									
23	P00964	*alr2328*	Glutamate-ammonia ligase (GlnA)	153	47%	53.445/5.28	58.0/5.4	-1.280 ± 0.038	0.005
Unidentified proteins									
24	An1						46.4/4.3	Not detected in An*lexA*$^+$	-
25	An2						45.7/5.8	Not detected in An*lexA*$^+$	-
26	An3						43.8/5.7	-16.934 ± 1.527	0.03
27	An4						43.2/5.0	-19.689 ± 2.031	0.001

28	An5						41.4/5.9	Not detected in An*lexA*$^{+}$	-
29	An6						28.6/7.0	-3.941 ± 0.545	0.001
30	An7						45.7/5.9	-1.403 ± 0.084	0.008
31	An8						32.2/9.2	-1.489 ± 0.046	0.05
32	An9						31.7/8.25	Not detected in An*lexA*$^{+}$	-
33	An10						30.4/4.1	-12.127 ± 0.458	0.002
34	An11						29.6/5.25	-1.284 ± 0.125	0.041
35	An12						14.1/5.2	-1.307 ± 0.056	0.037
36	An13						12.4/5.1	-1.447 ± 0.040	0.046
37	An14						12.7/5.5	-1.327 ± 0.059	0.015
38	An15						11.9/9.5	-39.348 ± 4.181	0.050
39	An16						47.7/6.1	1.553 ± 0.147	0.044
40	An17						30.9/6.25	Not detected in AnpAM	-
41	An18						30.4/5.9	Not detected in AnpAM	-
42	An19						71.7/5.92	1.571 ± 0.274	0.049
43	An20						22.3/8.9	Not detected in AnpAM	-
44	An21						21.4/8.5	Not detected in AnpAM	-
45	An22						20.8/8.7	Not detected in AnpAM	-
46	An23						70.1/6.2	1.570 ± 0.219	0.002
47	An24						30.9/7.9	Not detected in AnpAM	-

The raw data corresponding to all identified proteins are presented in **Appendix C**. The unidentified protein spots have been given a serial number prefixed with An. A few (25) protein spots which exhibited no change in abundance in AnpAM and An*lexA*$^{+}$ cells were identified on the basis of earlier published results of *Anabaena* 7120 proteome carried out under identical conditions in the same laboratory (*159, 176-177*) and these have been indicated in black **(Fig. 4.16B)**. The average fold-change of the 47 protein spots, across all triplicates, along with standard deviation and p-value is given in **Table 4.3**.

4.2.4.1 PROTEINS INVOLVED IN C-METABOLISM

Among the 23 identified protein spots, 15 corresponded to those involved in C-metabolism. Of these, 9 corresponded to proteins with role in photosynthetic light reaction and 3 each in carbon-fixation and carbon-catabolism. Among the proteins involved in light reaction of photosynthesis, phycobilisome core-membrane linker protein (ApcE) and one spot each of phycocyanin (CpcC) i.e. CpcC-3 and rod-linker proteins (CpcG1) i.e. CpcG1-4 **(Fig. 4.18)** exhibited 2.83, 2.15 and 1.44-fold increase in abundance respectively, while allopyhcocyanin (CpcA-3), two protein spots corresponding to phycocyanin (CpcC-1, CpcC-2) and two protein spots corresponding to rod-linker protein CpcG1-2, CpcG1-3 showed 1.5, 1.41, 2.49, 7.31, 1.33-fold decrease in abundance respectively. Protein spot corresponding to CpcG1-1 was detected only in AnpAM cells **(Fig. 4.18, left panel)** and indicated as not detected in An*lexA*$^{+}$. Among other proteins involved in light reaction, 2 spots corresponding to CpcA (A-1 and A-2), CpcB (B1 and B2), ApcA-1, ApcB-1, ApcB-2, ApcE-2, PecA and PecB did not show any change in abundance **(Fig. 4.16B)**. Three protein spots corresponding to RuBisCo large subunit (RbcL) were identified, and designated as RbcL-1, RbcL-2, and RbcL-3.

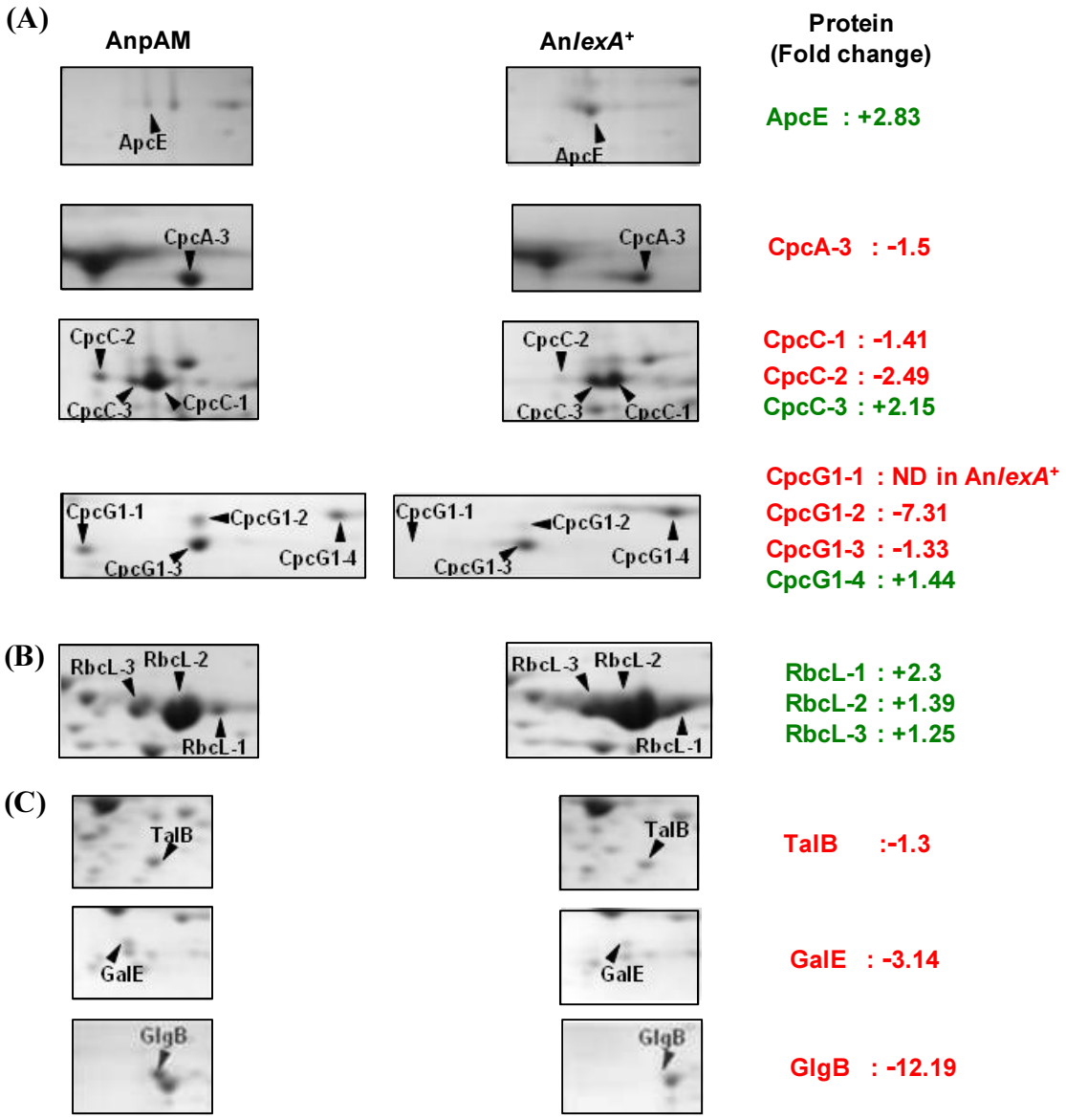


Fig. 4.18 Effect of LexA overexpression on change in abundance of proteins belonging to different functional groups. The proteins showing increased abundance in An*lexA*$^{+}$ are shown in green, and those with decreased abundance in red. The corresponding fold change in abundance is indicated by '+' and '-' respectively. Change in the abundance of **(A)** proteins involved in light reaction of photosynthesis: ApcE, CpcA, CpcC and CpcG1. **(B)** RbcL involved in carbon fixation **(C)** Proteins involved in carbon catabolism: TalB, GalE and GlgB. ND stands for 'not detected'.

The abundance of these protein spots was found to be 2.3, 1.39 and 1.25-fold higher respectively in An*lexA*$^{+}$ compared to AnpAM cells **(Fig. 4.18)**. Among the other identified protein spots, three corresponded to proteins involved in carbon-catabolism. These were transaldolase B (TalB), UDP-galactose-4'-epimerase (GalE) and alpha-glucanotransferase (GlgB). They exhibited 1.3, 3.14 and 12.19-fold decrease in abundance respectively upon overexpression of LexA **(Fig. 4.18C)**. Six protein spots, identified based on the earlier published data (*159, 176-177*), Enolase (Eno), Transketolase (Tkt), Fructose 1,6-bisphosphatase (GlpX), phosphoglycerate kinase (Pgk) and glyceraldehyde 3-phosphate dehydrogeanse-2 (Gap-2), which play a role in C-catabolism, did not exhibit any significant change upon overexpression of LexA **(Fig. 4.16B)**.

4.2.4.2 PROTEINS INVOLVED IN OXIDATIVE STRESS TOLERANCE

Among the identified proteins, three spots corresponded to ferredoxin-NADP(+) reductase (PetH). Of these, PetH-1 exhibited 1.67-fold decrease upon overexpression of LexA **(Fig. 4.18D),** while PetH-2 was detected only in AnpAM cells **(Fig. 4.16A, 4.18 D)** and PetH-3 in An*lexA*$^{+}$ **(Fig. 4.16B, 4.18D)**. Protein spot corresponding to peroxiredoxin (PrxA) was detected only in AnpAM cells **(Fig. 4.18D)**, while that identified as AhpC exhibited 1.58-fold increase in abundance upon overexpression of LexA **(Fig. 4.18D)**. The abundance of the FeSOD protein, identified based on earlier results, were found to be similar in AnpAM and An*lexA*$^{+}$ cells **(Fig. 4.16B)**.

4.2.4.3 PROTEINS INVOLVED IN OTHER CELLULAR FUNCTIONS

The remaining three identified protein spots corresponded to DnaK, nucleoside diphosphate kinase (Ndk) and glutamate-ammonia ligase (GlnA), which are known to be

involved in different cellular metabolic processes. The protein spots DnaK, Ndk and GlnA exhibited 1.26, 1.3 and 1.28-fold decrease in abundance respectively upon overexpression of LexA **(Fig. 4.18E)**.

4.2.5 *Analysis of LexA-mediated gene regulation*

The 23 identified protein spots **(Table 4.3)** corresponded to 11 different genes, as some of the gene products were detected as multiple spots, possibly owing to processing of the synthesised protein resulting in change of molecular mass or pI or both. The differential abundance of the multiple protein spots, coded by the same gene, in AnpAM and An*lexA*$^{+}$ cells, could not throw clarity on whether the gene is up or down-regulated by LexA. To overcome this problem, the transcript levels of the genes coding for the identified proteins were compared. Since, Ndk (Nulceoside diphosphate kinase) which is involved in the conversion of nucleoside diphosphate to triphosphate was identified to be down-regulated, one of the genes involved in the generation of nucleoside diphosphates was also included for analysis. The gene included for analysis was *tdk*, coding for thymidylate kinase, which is an important enzyme in regulating DNA synthesis (*178*).

The primer pairs used for transcript analysis of different genes along with the expected size of the insert is listed in **Table 4.4**. For all sets, amplification of 16S rDNA was used as an internal control to ascertain that the amount of RNA taken for amplification from AnpAM and An*lexA*$^{+}$ cells are same. Of the different genes analysed, increase in the transcript levels was observed for *ahpC* and *apcE* (**Fig. 4.19A and B**) indicating that the observed increase in the abundance of ApcE and AhpC were due to transcriptional up-regulation. Overexpression of LexA in An*lexA*$^{+}$ resulted in decrease in the transcript levels of *cpcA*, *galE, talB, dnaK, glnA, ndk*, and *tdk* (**Fig. 4.19C-I**).

Table 4.4: List of primer pairs used for reverse transcriptase analysis

Primers	Nucleotide Sequence	Amplicon
glgB Fwd	ATGGCAAAGCCCATTGAA	*glgB* (400 bp)
glgB Rev	GGTCATCTTCTCCACCAG	
talB Fwd	GCGAGAAATGACTGTAGTAG	*talB* (388 bp)
talB Rev	GGCAATTTTAATCAAAACCCG	
cpcA Fwd	ACCGAAGCAATTGCAGCTGCTG	*talB* (474 bp)
cpcA Rev	CTAGCTGAGAGCGTTGATAGCG	
apcE Fwd	ATGAGTGTTAAGGCGAGT	*apcE* (407 bp)
apcE Rev	CCACTGGGAGATGTATTA	
rbcL Fwd	ATGTCTTACGCTCAAACG	*rbcL* (400 bp)
rbcL Rev	ATGCGCGTAATGCTTTAA	
prxA Fwd	ATGTCCATCACCTACGGA	*prxA* (612 bp)
prxA Rev	TTACACAGCAGCGAAGTA	
ahpC Fwd	ATGGCTCTCCGTCTTGGT	*ahpC* (630 bp)
ahpC Rev	AGGTTGAGGAGTTAACCGC	
ndk Fwd	GTGGGTTTGAAGTTCCTC	*ndk* (354 bp)
ndk Rev	TTACTCGTGCAACCAAGG	
galE Fwd	AAAGCCCAGCATTTTGGT	*galE* (391 bp)
galE Rev	AACAGTTTTTGGCACTCC	
petH Fwd	ATGTCTAATCAAGGTGCT	*petH* (406 bp)
petH Rev	TGTTGCCTTTGTTGTCTT	
tdk Fwd	ATGGGTGGCAGATTCATT	*tdk* (636 bp)
tdk Rev	CTACAGCCATTGCTTGAG	
dnaK Fwd	ATGGCAAAAGTAGTTGGAATTGAC	*dnaK* (407 bp)
dnaK Rev	ATCTTGGTGAAACTGTAACCCA	
glnA Fwd	ATGACAACCCCACAAGAAGTCT	*glnA* (400 bp)
glnA Rev	TTCTTTGGCCCTGAAGCTGAAT	
16S rDNA Fwd	AAAACGGAGAGTTTGATCCTGGCTCA	16S rDNA (500 bp)
16S rDNA Rev	GATAACGCTTGCATCCTCCGTATTA	

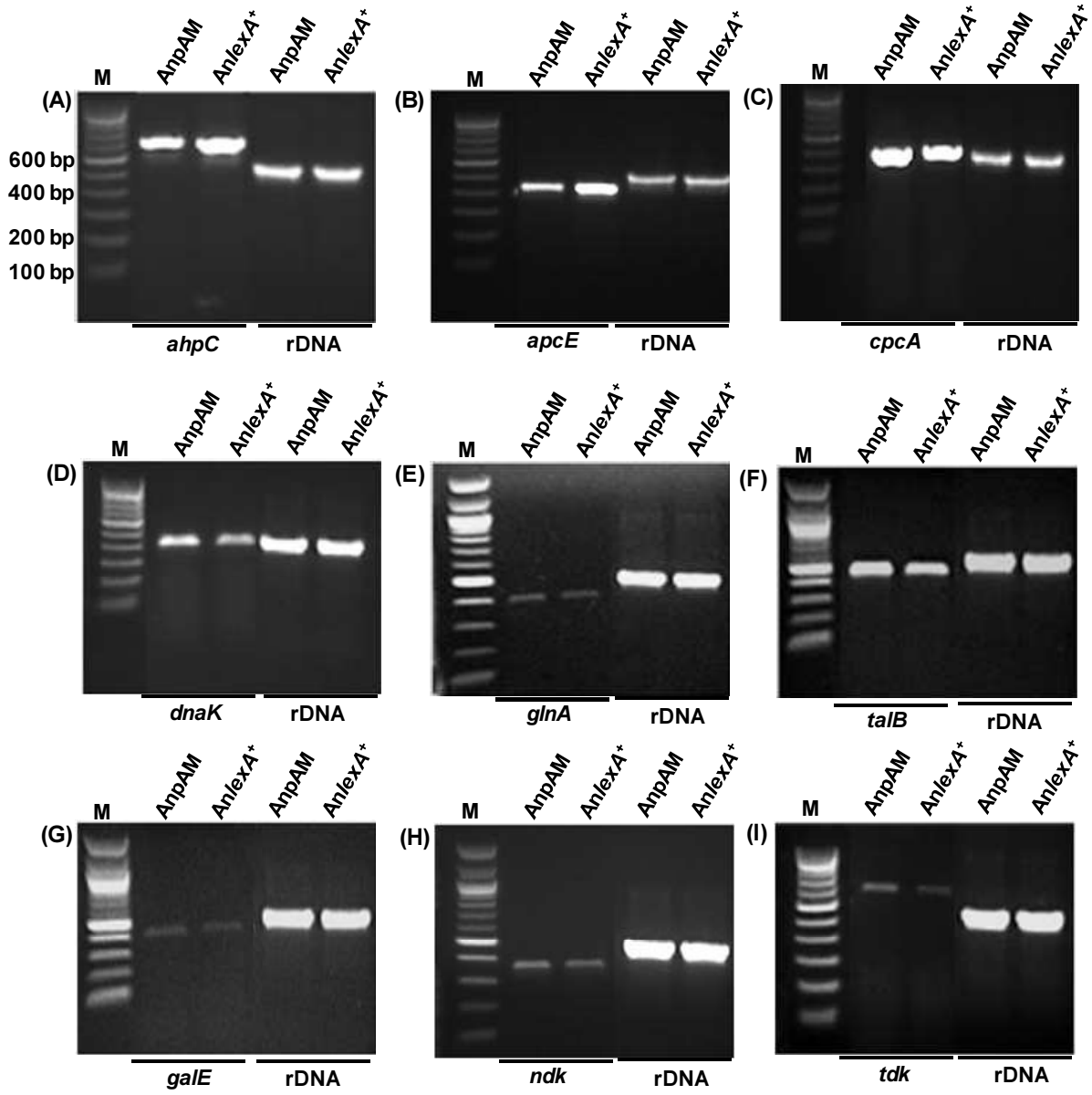


Fig. 4.19 Effect of LexA overexpression on transcription of select genes of *Anabaena*. Total RNA isolated from three-day-old cultures of AnpAM and An*lexA*$^{+}$ were checked for DNA contamination. DNA-free RNA was subjected to one step Reverse Transcriptase PCR using gene specific primers. Amplification with 16S rDNA gene primers was used as internal control. Lane M refers to 100 bp DNA ladder (Bangalore Genei, India). The different gene transcripts analysed corresponded to (A) *ahpC*, (B) *apcE*, (C) *cpcA*, and (D) *dnaK* (E) *glnA*, (F) *talB*, (G) *galE*, (H) *ndk* and (I) *tdk*.

The observed decrease in transcript levels of these genes, thus accounts for the decrease in the abundance of the corresponding proteins observed in An*lexA*$^{+}$ compared to AnpAM cells (**Fig. 4.16 and 4.18**). The transcript level of *cpcA* was analyzed in place of *cpcB* as they are the part of single operon *cpcBACDEG* (*179*), with *cpcA* being the second gene of operon.

Having confirmed the transcriptional regulation of the genes corresponding to proteins identified by MALDI-ToF analysis, and *tdk* by LexA, the next step was to ascertain if the regulation is a direct regulation or indirect through other transcriptional activators/repressors. Since, direct regulation by LexA in bacteria is known to be through the binding of LexA to the LexA-box present in the vicinity of the promoter region of the genes, thereby repressing the promoter activity (*162*), the binding of *Anabaena* LexA to the upstream regulatory region (promoter) of these genes, and the regulation of their promoter activity using a reporter gene (*gfpmutII*) was analysed. For these experiments, 5 additional genes with possible role in DNA repair, but not detected by proteomic studies, *viz ssb1*, *ssb2*, *ssb3*, *recA* and *lexA* were also included. Inclusion of these genes was based on earlier studies. Binding of LexA to *ssb1*, *ssb2*, *recA* and *lexA* promoters had been demonstrated earlier in *Anabaena* (*130, 171*), while *ssb3* promoter has been used as a negative control, as it has been shown to be regulated by FurA and not LexA (*171*).

The different promoter regions were amplified using the primers listed in the **Table 4.5**. Binding of *Anabaena* LexA to the promoter regions was assessed by Electrophoretic Mobility Shift Assay (EMSA). *Anabaena* LexA exhibited binding to all promoters tested except *ssb3* promoter **(Fig. 4.20)**. In order to determine the binding affinity of LexA to each of these promoters, the binding was assessed over a concentration range of 0 – 180 nM LexA **(Fig. 4.20)**. Amount of free and bound DNA

was estimated densitometrically using Image J software (*160*) and averaged over three sets of EMSAs for each promoter DNA fragment-LexA pair.

Table 4.5: List of primer pairs used to amplify the upstream regulatory regions (promoter) of different genes

Primers	Nucleotide Sequence	Restriction site	Amplicon
P_{glgB} Fwd	CGAGCTCTCATTTTTCCCCGAATATTTG	*SacI*	P_{glgB} (261 bp)
P_{glgB} Rev	GGGGTACCGCTAGTATTAAGCATTTTC	*KpnI*	
P_{galE} Fwd	CGAGCTCATATTCCTAGAATAAGTC	*SacI*	P_{galE} (275 bp)
P_{galE} Rev	GGGGTACCAAAATACTCCATGATTT	*KpnI*	
P_{cpcB} Fwd	CGAGCTCGCAACTTATTTATTCACAAT	*SacI*	P_{cpcB} (300 bp)
P_{cpcB} Rev	GGGGTACCGGACTTTAATCTCCTAATT	*KpnI*	
P_{apcE} Fwd	CGAGCTCAGAGAACCAAAGCAAATT	*SacI*	P_{apcE} (308 bp)
P_{apcE} Rev	GGGGTACCCTGATTTGTAAAACTCCC	*KpnI*	
P_{rbcL} Fwd	CGAGCTCGAAGGTCGTGAAGGGATA	*SacI*	P_{rbcL} (294 bp)
P_{rbcL} Rev	GGGGTACCCCTTCCAAGATGTCACTC	*KpnI*	
P_{prxA} Fwd	CGAGCTCTTTGCTACTCAAAATCGT	*SacI*	P_{prxA} (199 bp)
P_{prxA} Rev	GGGGTACCTAACTTAATTCTCCTTCA	*KpnI*	
P_{ahpC} Fwd	CGAGCTCTAACACAATATTAGTGCC	*SacI*	P_{ahpC} (286 bp)
P_{ahpC} Rev	GGGGTACCATTCCCTGTAGTTAAGTA	*KpnI*	
P_{ndk} Fwd	CGAGCTCCCCACGAATAAACAATTT	*SacI*	P_{ndk} (251 bp)
P_{ndk} Rev	GGGGTACCAAGGGTAAAACCTTTTGT	*KpnI*	
P_{galE} Fwd	CGAGCTCGATTTACTCTCAAATCTCC	*SacI*	P_{galE} (296 bp)
P_{galE} Rev	GGGGTACCAGTTTTTCCTTTTGAGTT	*KpnI*	
P_{petH} Fwd	CGAGCTCTCAAGACTAATTGCTTTT	*SacI*	P_{petH} (284 bp)
P_{petH} Rev	GGGGTACC TCCGATCTTTAACTTAA	*KpnI*	
P_{tdk} Fwd	CGAGCTCAACAGGCAAATCGATTTT	*SacI*	P_{tdk} (296 bp)
P_{tdk} Rev	GGGGTACCTTTTTATGTTCGCCGAT	*KpnI*	
P_{dnaK} Fwd	CGAGCTCAGACAGAGTGACAGCAG	*SacI*	P_{dnaK}

P_{dnaK} Rev	GGGGTACCCCATTTTAGGTGCTGG	*KpnI*	(272 bp)
P_{glnA} Fwd	CGAGCTCTTGGTTGCACTACGCACCA	*SacI*	P_{glnA} (317 bp)
P_{glnA} Rev	GGGGTACCTTGTTACTCCTTCTCTGCC	*KpnI*	
P_{lexA} Fwd	CGAGCTCAGCCCATCGTGGT	*SacI*	P_{lexA} (280 bp)
P_{lexA} Rev	GGGGTACCAAAATAATAGGGAGGATA	*KpnI*	
P_{recA} Fwd	CGAGCTCGTATAGCTAGGTTA	*SacI*	P_{recA} (301 bp)
P_{recA} Rev	GGGGTACCTTCCTGCCTCTAAAAT	*KpnI*	
P_{ssb1} Fwd	GCGCCGGAGCTCGGTATTTGCTGTACCGAG	*SacI*	P_{ssb1} (400 bp)
P_{ssb1} Rev	GATCGGTACCCTGCCTTATCCTTTAGTACATAAGTACT	*KpnI*	
P_{ssb2} Fwd	GCAACGAGCTCCAACAGATGTTTCTCCT	*SacI*	P_{ssb2} (400 bp)
P_{ssb2} Rev	GCCCGGGTACCAACGCGAGAATTGAT	*KpnI*	
P_{ssb1}M0Rev	GCCTTATCCTTT***AGT***ACATAAGT***ACT***		P_{ssb1}M0
P_{ssb1}M1Rev	GCCTTATCCTTT***AGT***ACATAAGT***GTG***		P_{ssb1}M1
P_{ssb1}M2Rev	GCCTTATCCTTT***CAG***ACATAAGT***ACT***		P_{ssb1}M2
P_{ssb1}M3Rev	GCCTTATCCTTT***AGT***ACAT***ACT***		P_{ssb1}M3

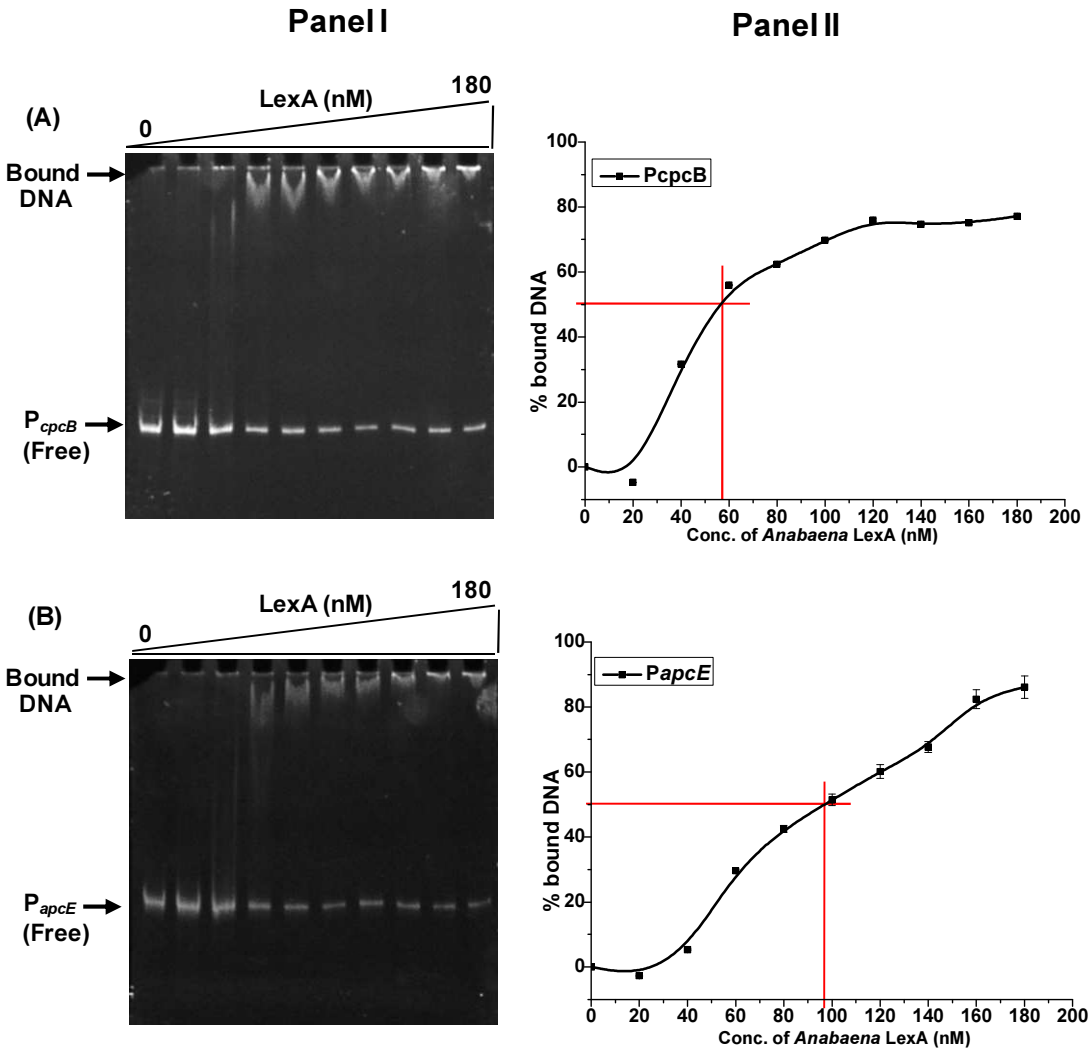


Fig. 4.20 Interaction of *Anabaena* LexA to upstream regulatory (promoter) region of specific genes of *Anabaena* 7120. Electrophoretic mobility shift assays (EMSA) of promoter regions were carried out with increasing concentrations (0-180 nM) of the purified *Anabaena* LexA protein. The DNA-protein mix after incubation were electrophoretically separated on 8% native polyacrylamide gel, followed by staining with SYBR Dye I as shown in Panel I. The free and bound DNA fragments are indicated by arrows. The % bound DNA was determined by densitometric analysis of free and bound DNA using ImageJ software. The average of three replicate gels has been plotted as a function of LexA concentration to determine the binding affinity as hown in Panel II. Different promoter fragments used are indicated as $P_{abc,}$ wherein P stands for amplified promoter DNA and *abc* is the gene whose promoter has been amplified. The different genes are **(A)** *cpcB* (*alr0528*), **(B)** *apcE* (*alr0020*)

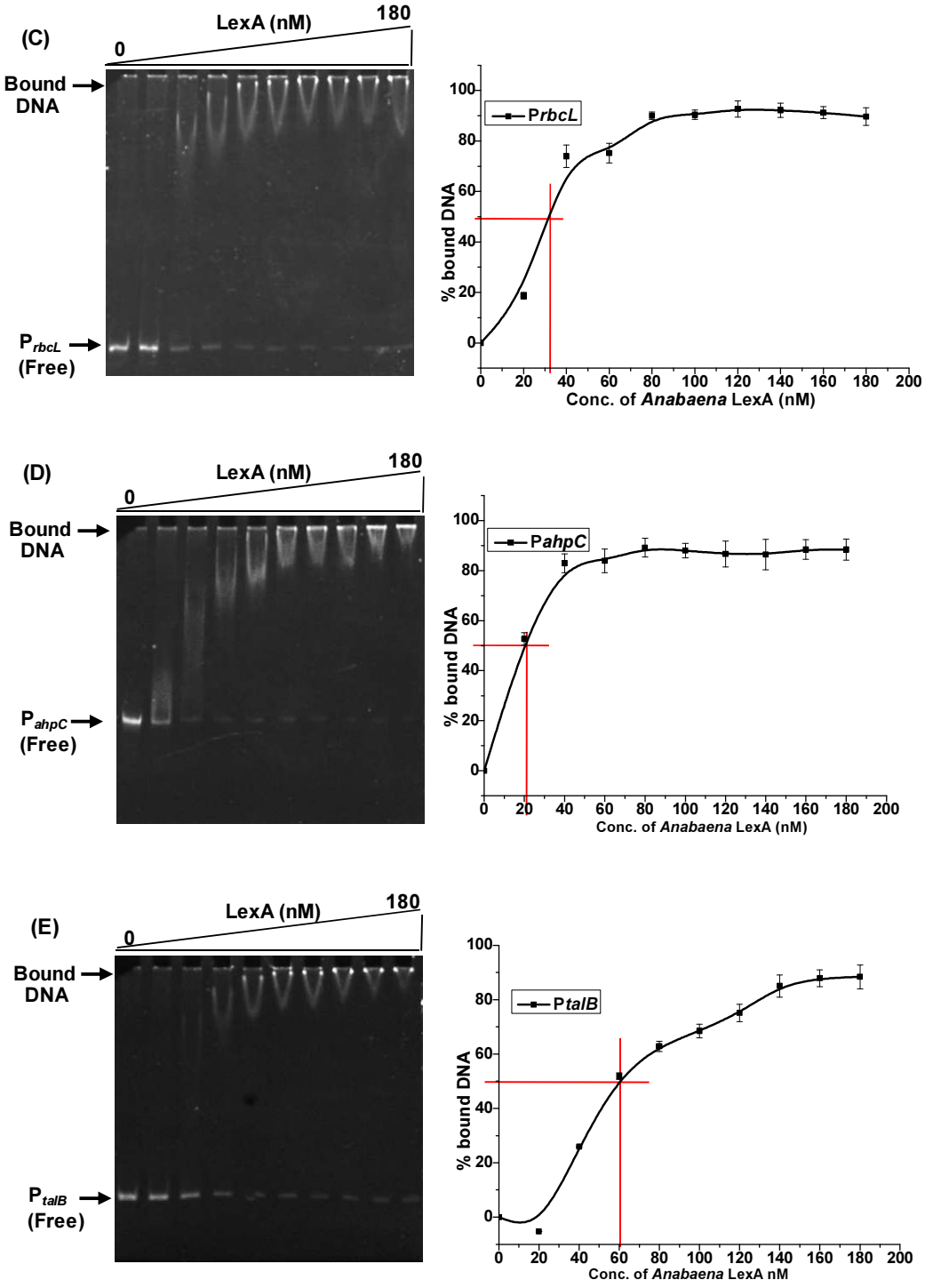


Fig. 4.20 Promoters of **(C)** *rbcL* (alr1524) operon, P_{rbcL}; **(D)** *ahpC* (alr4404) gene, P_{ahpC}; **(E)** *talB* gene (all2563), P_{talB}

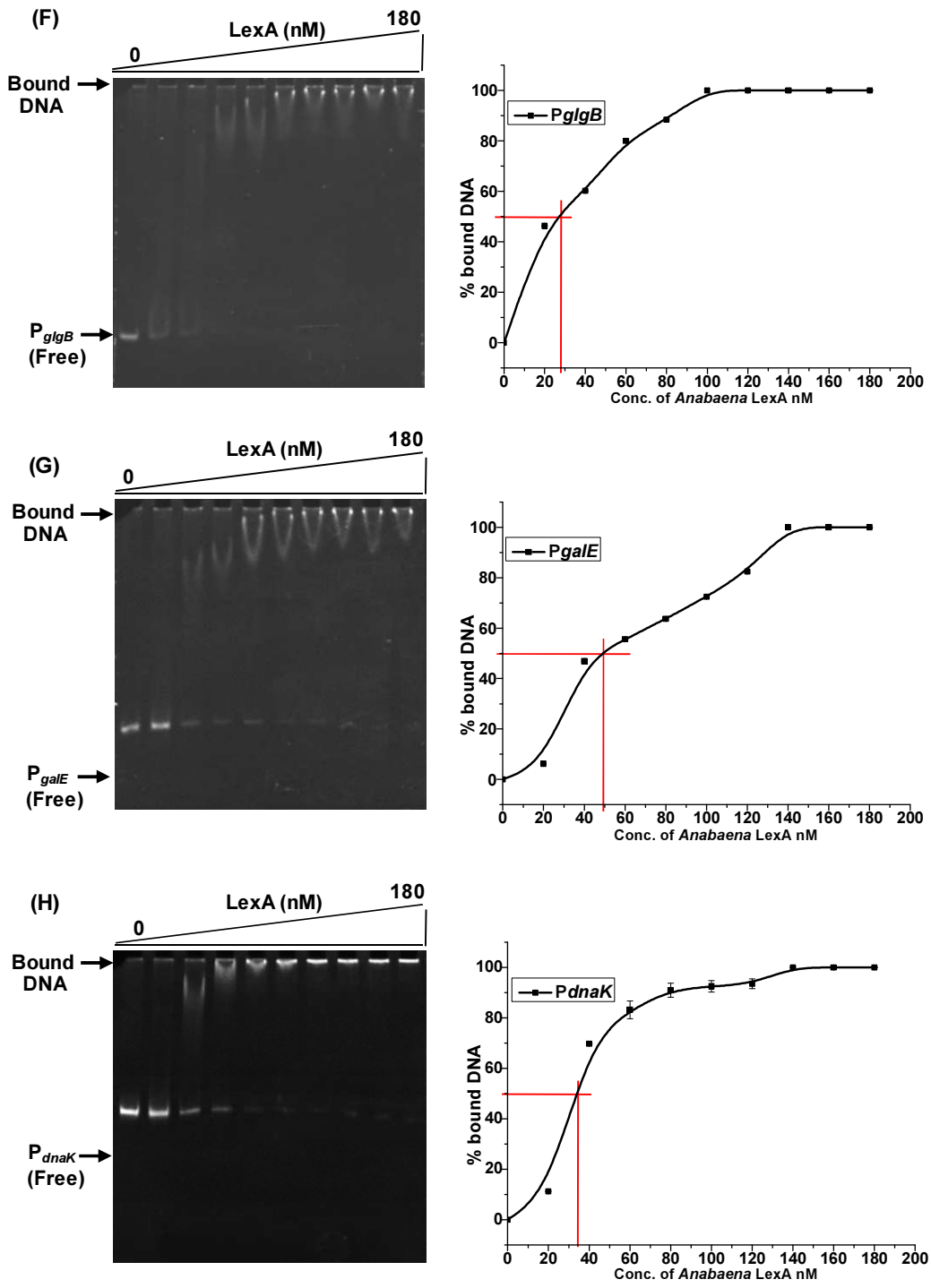


Fig. 4.20 Promoters of **(F)** *glgB* (*all0875*) gene, P_{glgB}; **(G)** *galE* gene (*all4713*), P_{galE}; **(H)** *dnaK* (*alr4712*) gene, P_{dnaK}.

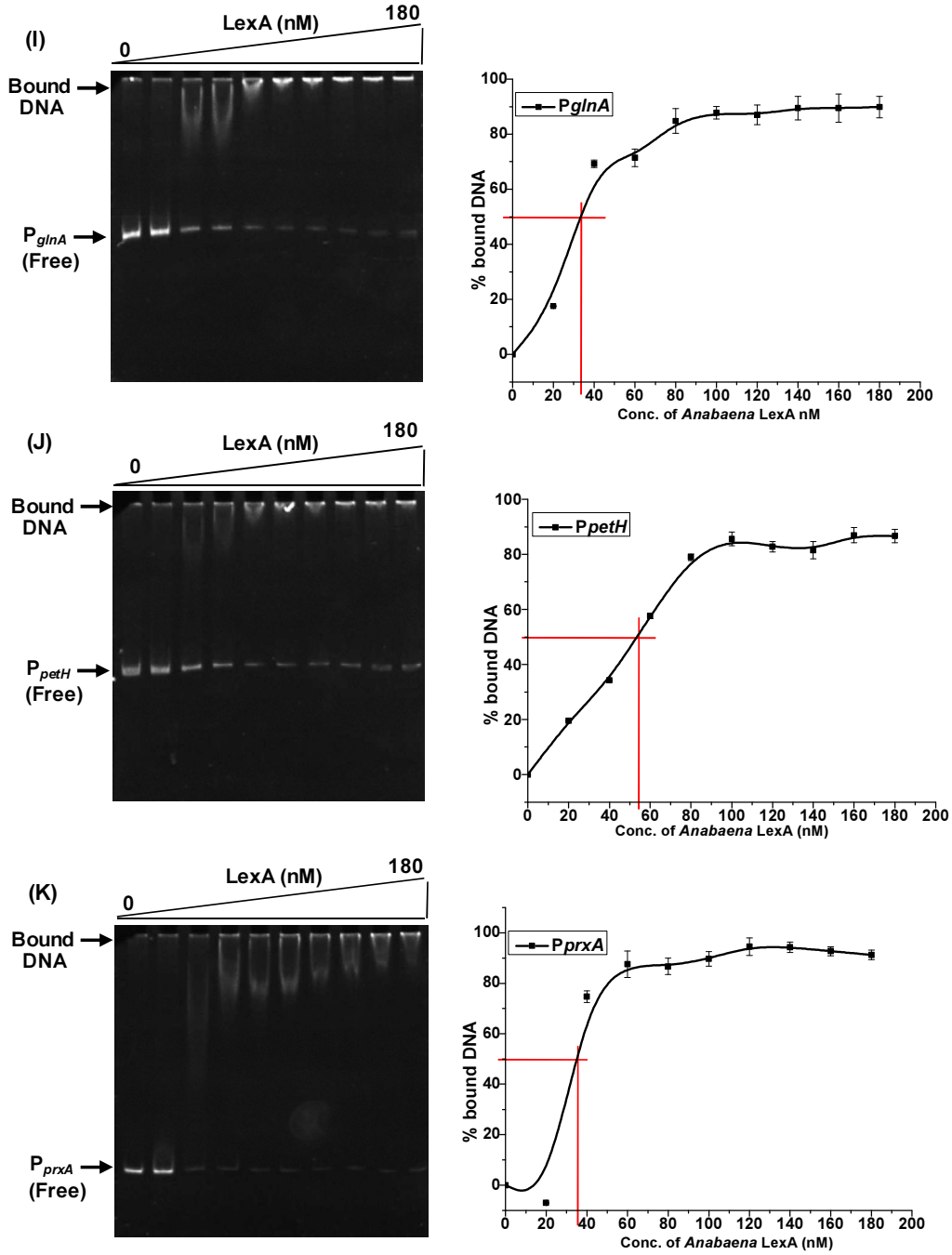


Fig. 4.20 Promoters of **(I)** *glnA* (*alr2328*) gene, P_{glnA}; **(J)** *petH* gene (*all4121*), P_{petH}; **(K)** *prxA* (*alr4641*) gene, P_{prxA}

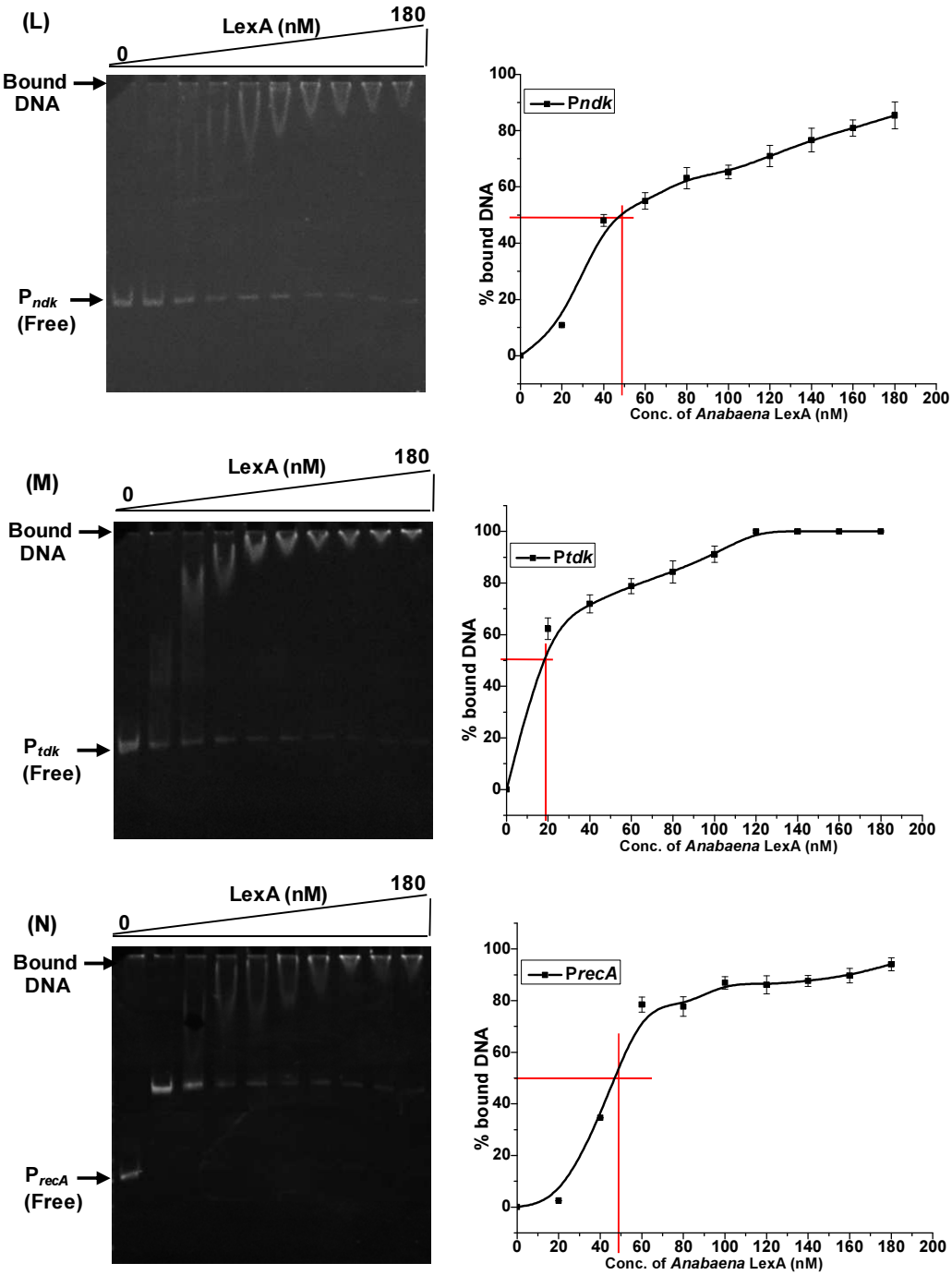


Fig. 4.20 Promoters of **(L)** *ndk* (*alr3402*) gene, P_{ndk}; **(M)** *tdk* (*all4708*) gene, P_{tdk}; **(N)** *recA* (*all3272*) gene, P_{recA}

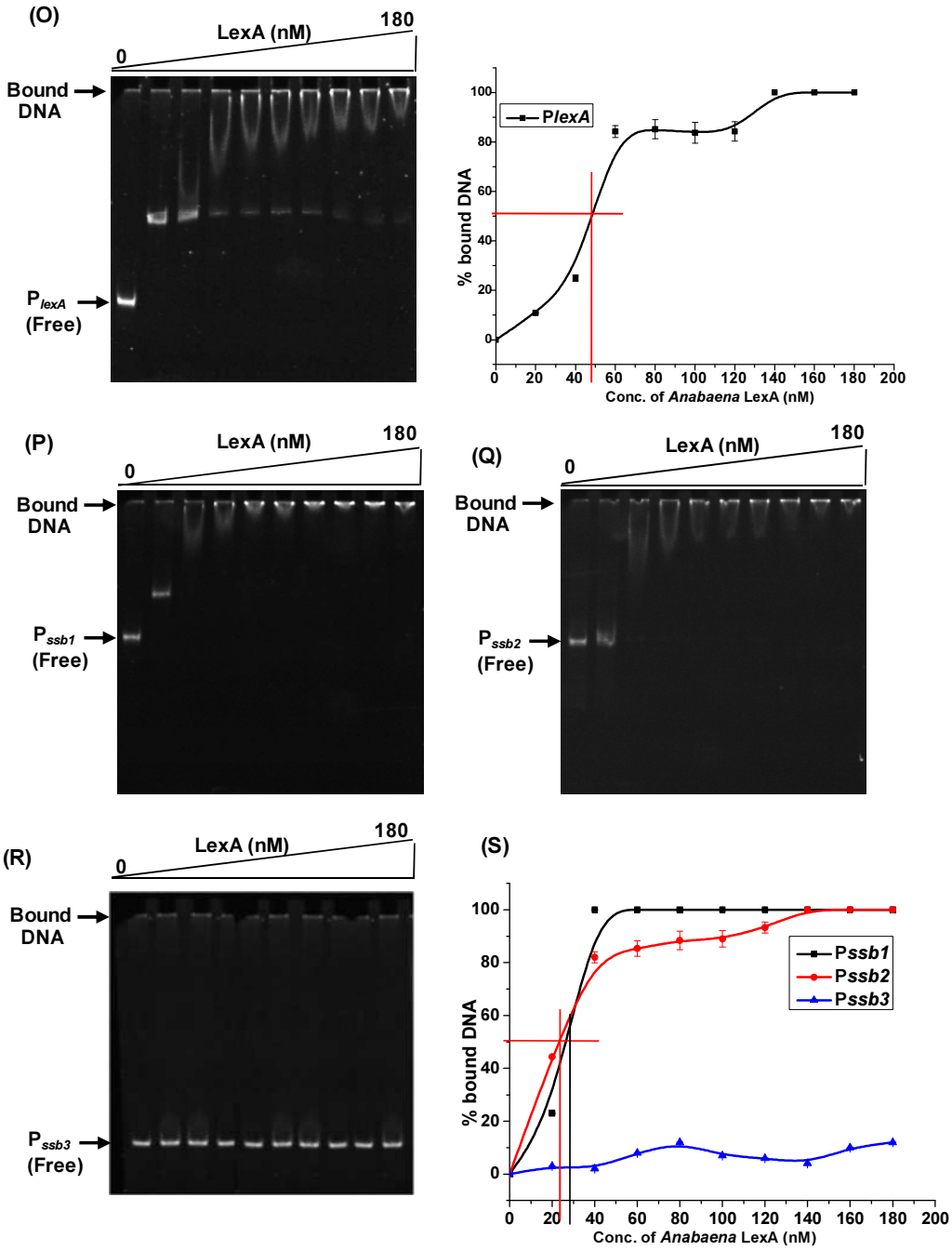


Fig. 4.20 Promoters of **(O)** *lexA* (*alr4908*) gene, P_{lexA}; **(P)** *ssb1* (*alr0088*) gene, P_{ssb1}; **(Q)** *ssb2* (*alr7579*) gene, P_{ssb2}; **(R)** *ssb3* (*all4779*) gene; and **(S)** binding affinity graph of all three *ssb* genes.

In most cases saturation of binding was observed by 60-80 nM LexA. The slopes of the graph between percent bound DNA and LexA concentration was distinct for each of the EMSA (**Fig. 4.20**), indicating some variation in their binding affinity. Binding affinity was calculated as the concentration of LexA at which 50% binding is achieved. The calculated binding affinity was found to be in the range of 18 to 96 nM (**Table 4.5, Fig. 4.20**), with highest affinity for *tdk* promoter (18.0 ± 0.5 nM) (**Table 4.5, Fig. 4.20M**), and least for *apcE* promoter (96.0 ± 3.0 nM) (**Table 4.5, Fig. 4.20B**). With the binding of LexA to the promoter regions established, it was required to be ascertained if both up- and down-regulation of the different genes was a result of direct regulation by *Anabaena* LexA or required other *Anabaena* proteins for modulating its regulator activity. Competition experiments were carried out for the binding of *Anabaena* LexA to the three *ssb* gene promoters. Of the three *ssb* promoters, LexA binds only to the *ssb1* and *ssb2* gene promoters, but not *ssb3* promoter **(Fig. 4.20 P-R)**. When 20 nM LexA was pre-incubated with P_{ssb2} followed by challenge with P_{ssb1}, partial displacement of the LexA bound to P_{ssb2} was observed resulting in the formation of P_{ssb1}-LexA complex in addition to P_{ssb2}-LexA complex **(lane 8, Fig. 4.21A)**. When LexA was pre-incubated with P_{ssb3}, only P_{ssb1}-LexA complex was detected **(lane 9, Fig. 4.21A)** as P_{ssb3} does not bind to LexA.

In order to determine the activity of the different promoters, the corresponding DNA fragments used for EMSA analysis were cloned in pAM1956 and the generated plasmids (**Table 4.6**) transformed into *E. coli* cells, BL21(pET16b) and BL21(pET*lexA*) individually. All *E. coli* cells harbouring the different plasmid constructs, pAM-P_{abc}:*gfpmutII* (where *abc* refers to the different genes as listed in **Table 4.6**) exhibited green fluorescence when observed under fluorescence microscope, indicating that the cloned promoters were functional in *E. coli* resulting in GFP. The strength of individual

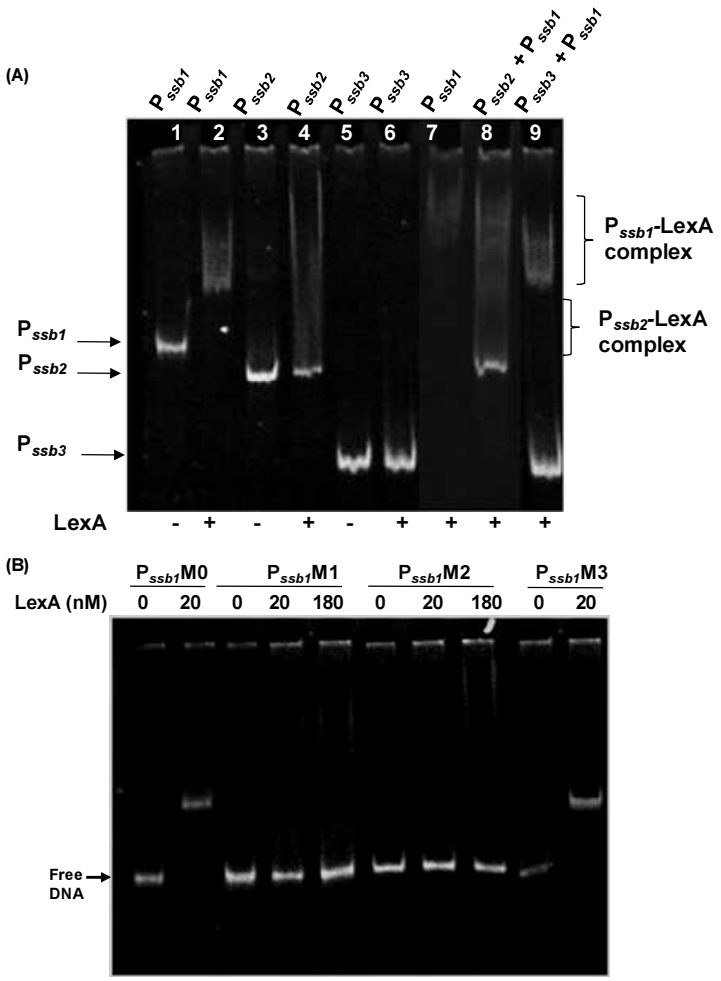


Fig. 4.21 Confirmation of AnLexA-box. (**A**) Competition EMSA of *Anabaena ssb1* promoter fragment P_{ssb1} (lanes 2 and 7) with P_{ssb2} (lane 8) and P_{ssb3} (lane 9) in the presence of LexA. For the competition experiment LexA was first incubated with either P_{ssb2} (lane 4) or P_{ssb3} (lane 6) followed by the addition of P_{ssb1} and continued incubation for 30 min. Lanes 1, 3 and 5 correspond to P_{ssb1}, P_{ssb2} and P_{ssb3} fragments respectively in the absence of LexA protein. (**B**) EMSA of mutated LexA binding sites of P_{ssb1} with *Anabaena* LexA. The ~400 bp promoter region of *ssb1* (P_{ssb1}) was PCR amplified using P_{ssb1}Fwd and one of the four variants of P_{ssb1} Rev primers; wherein M0 refers to intact LexA binding site (AGT-N_x-ACT), M1 has a mutation in the left arm (AGT to CAC), M2 has a mutation in the right arm (ACT to CTG); and in M3, the intergenic region is shortened from 8 bases to 4 while retaining the sequence of the left and right arm. EMSA was performed with 20 nM LexA for P_{ssb1} M0 and P_{ssb1}M3 DNA fragments and with 20 nM and 180 nM for P_{ssb1}M1 and P_{ssb1}M2 DNA fragments.

Table 4.6 Recombinant plasmids generated during promoter studies

Plasmids	Description	Source
pAM-P$_{petH}$	0.284 kb *petH* promoter fragment cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{glgB}$	0.261 kb *glgB* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{apcE}$	0.308 kb *apcE* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{ndk}$	0.251 kb *ndk* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{rbcL}$	0.294 kb *rbcL* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{ahpC}$	0.286 kb *ahpC* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{cpcB}$	0.3 kb *cpcB* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{dnaK}$	0.272 kb *dnaK* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{glnA}$	0.317 kb *glnA* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{recA}$	0.284 kb *recA* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{tdk}$	0.296 kb *tdk* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{galE}$	0.296 kb *galE* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{talB}$	0.275 kb *talB* promoter cloned in pAM1956 at *SacI-KpnI* restriction site	This study
pAM-P$_{ssb1}$	0.4 kb *ssb1* promoter cloned in pAM1956 at *SacI-KpnI* sites	(*171*)
pAM-P$_{ssb2}$	0.4 kb *ssb2* promoter cloned in pAM1956 at *SacI-KpnI* sites	(*171*)

promoters in these strains was assessed by measuring GFP fluorescence. The activity of the individual promoters was similar in BL21(pET16b) and BL21(pET*lexA*) cells in the absence of IPTG, and the average of their values taken as the strength of the respective promoter in *E. coli*. The strength of the different promoters showed variation from ~840-15000 arbitrary units OD^{-1}**(Table 4.7)**, with the *ahpC* being the weakest and *prxA* the strongest promoter in *E. coli*. This is a measure of extent of similarity of these promoters but not a true measure of their relative promoter activity in *Anabaena*. The activity of *ahpC*, *apcE*, and *rbcL* promoters showed an increase within 1 h of induction of *Anabaena* LexA in *E. coli* **(Table 4.7)**. Upon addition of IPTG to the cells, the activity of the promoters showed no significant variation in BL21(pET16b) cells, but varied over a 1 h period in BL21(pET*lexA*) cells. Activity of *rbcL* promoter increases 1.5-fold after 30 min of induction of LexA, and decreased thereafter, while that of *ahpC* and *apcE* showed a steady increase in the duration of induction of LexA, with them exhibiting 2.9- and 1.7-fold increase in activity response over a 1 h period **(Table 4.7)**. This indicated that the increased abundance of RbcL, AhpC and ApcE is a direct consequence of the up-regulation of their promoters by LexA. However, role of additional *Anabaena* proteins in further up-regulating their expression cannot be ruled out. Activity of the remaining 14 gene promoters tested was found to decrease by 0.21-0.7-fold upon overexpression of LexA (**Table 4.7**). Thus, the down-regulation of these genes directly contributes to the decrease in the abundance of corresponding proteins observed during proteome analysis (**Fig. 4.16, 4.18**). However, the magnitude of change in promoter activity was not found to be same as that of the protein. This could be due to post-transcriptional control or possibility of other regulator also functioning simultaneously with LexA in *Anabaena*.

Table 4.7: Binding Affinity of *Anabaena* promoters to purified *Anabaena* LexA and promoter activity as a function of LexA expression in *E. coli*

Promoter	Binding Affinity (nM)	Promoter activity in *E. coli*[#] (AU)	Fold change on LexA induction	
			0.5 h	1 h
P_{ahpC}	20.0 ± 1.0	842.10 ± 54.48	1.07 ± 0.04	2.88 ± 0.07
P_{apcE}	96.0 ± 3.0	4287.04 ± 244.01	1.33 ± 0.10	1.67 ± 0.05
P_{rbcL}	32.0 ± 1.5	1296.47 ± 43.03	1.50 ± 0.03	1.14 ± 0.01
P_{cpcB}	55.5 ± 2.5	2637.62 ± 155.12	0.42 ± 0.02	0.63 ± 0.06
P_{galE}	49.0 ± 3.6	3183.05 ± 97.04	0.58 ± 0.04	0.55 ± 0.04
P_{glgB}	27.5 ± 0.5	1149.81 ± 102.11	0.27 ± 0.01	0.38 ± 0.01
P_{talB}	60.0 ± 4.1	1808.53 ± 88.5	0.93 ± 0.07	0.21 ± 0.01
P_{dnaK}	34.5 ± 1.5	6099.29 ± 245.35	0.75 ± 0.03	0.33 ± 0.03
P_{glnA}	34.0 ± 2.1	5375.01 ± 208.32	0.58 ± 0.04	0.45 ± 0.02
P_{ndk}	49.0 ± 3.5	6430.46 ± 198.63	0.38 ± 0.03	0.31 ± 0.07
P_{petH}	54.5 ± 3.0	2094.93 ± 172.05	1.03 ± 0.07	0.73 ± 0.07
P_{prxA}	34.5 ± 2.5	15181.76 ± 874.32	0.66 ± 0.07	0.53 ± 0.06
P_{lexA}	49.5 ± 4.5	7500.19 ± 590.32	0.62 ± 0.02	0.33 ± 0.06
P_{recA}	46.5 ± 2.1	8532.25 ± 3451.15	0.47 ± 0.04	0.28 ± 0.03
P_{ssb1}	26.5 ± 2.3	4376.36 ± 98.32	0.70 ± 0.02	0.55 ± 0.02
P_{ssb2}	23.0 ± 1.9	3880.8 ± 235.12	0.62 ± 0.04	0.45 ± 0.02
P_{tdk}	18.0 ± 0.5	4385.55 ± 256.21	0.71 ± 0.03	0.43 ± 0.04

The amplified promoter regions were ligated to pAM1956 to generate the different promoter constructs, P_{abc}, wherein *abc* refers to the different genes. Activity of the different promoters was assessed by transforming the different promoter constructs in *E. coli* cells, BL21(p*lysS*)(DE3) harbouring either pET16b or pET*lexA*.

4.2.6 *Elucidation of the LexA-binding site in Anabaena*

Several probable LexA-binding sites have been predicted earlier for cyanobacteria. These include RGTACN$_3$DGTWCB (*130, 169*), CTA-N$_9$ (AT-rich)-CTA (*138*) and RRTACRNNYGTWYK (*166*). However, none of these sequences were detected in all the 17 *Anabaena* gene promoters taken up for the present study and shown to be directly regulated by LexA. With the binding affinity of *Anabaena* LexA to the promoter regions tested being in the same range i.e. 18-96 nM **(Table 4.7)**, they were expected to have similar LexA binding sites. Analysis of the upstream regions of these revealed the presence of the palindromic sequence AGT-N$_x$-ACT in 10 of the 17 gene promoters, where is spacer X varied from 4-11, the sequence was present **(Table 4.8)**. In the remaining 7 genes a single base variation was detected in the left arm in 3 cases and in the right arm in 4 cases. In all the 7 cases, a minimum of four base palindrome including the mismatch was observed, and hereafter retained as a criteria for prediction of this site in other genes. The predicted AnLexA-Box was present within 300 bases upstream of the start codon ATG and an AT region, which could function as a probable -10 region, was present within 50 bases of the predicted box. Close proximation of the LexA binding site and the -10 region would allow for better regulation of transcription. In order to confirm that the predicted sequence is the AnLexA-Box, three mutations in the predicted AnLexA-Box of promoter of *ssb1* gene were carried out, one each in the left and in the right arm to disrupt the palindrome and the other in the spacer, by decreasing it in length from 8 to 4 bases, designated as P$_{ssb1}$M1, P$_{ssb1}$M2 and P$_{ssb1}$M3 respectively, whereas full length promoter was designated as P$_{ssb1}$M0. These were generated by PCR amplification of the cloned *ssb1* promoter (pAM-P$_{ssb1}$::*gfpmutII*) using P$_{ssb1}$ Fwd primer in combination with either P$_{ssb1}$ M0Rev, P$_{ssb1}$ M1Rev, P$_{ssb1}$ M2Rev or P$_{ssb1}$ M3Rev **(Table 4.7)** primers respectively.

Table 4.8: Probable LexA binding site in *Anabaena* promoters

S. No.	Gene Symbol (annotation)	Probable LexA-binding site (AGT-N_x-ACT)			Probable -10 region	
		Sequence	Spacer region (x: no. of bases)	No. of bases upstream of ATG	Sequence	No. of bases upstream of ATG
Downregulated genes						
1	*glgB (all0875)*	**AGT**ataaa**ACT**	5	94	TAATAA	**51**
2	*talB (all2563)*	**AGT**cttaacaa**ACT**	8	32	TAAAAT	**19**
3	*galE(all4713)*	**AGT**agttgtca**ACT**	8	28	TAAATT	**39**
4	*petH (all4121)*	**AGTT**agaaggtg**AACa**	10	198	TTATTAA	**230**
5	*prxA (alr4641)*	**ACAGaT**cgtaatca**AACTGT**	10	160	TAATCA	**152**
6	*dnaK (alr1742)*	**AGTA**ttaa**TtCT**	6	292	TAAATT	**278**
7	*glnA (alr2328)*	**AGgTG**tgtcat**CAACT**	10	94	TATTTC	**100**
8	*cpcB (alr0258)*	**AAGT**ggcaaga**AaTT**	7	134	TAAAAT	**170**
9	*ndk (alr3402)*	**AGT**taattgttg**ACT**	9	166	TAAACT	**185**
10	*lexA (alr4908)*	**TAGT**actaatgt**tCTA**	8	48	TATTTT	**55**
11	*recA (all3272)*	**AGTA**tatctgttc**TACT**	11	67	TTAATT	**108**
12	*tdk (all4708)*	**AGT**gttcaca**ACT**	7	90	TAAAAAT	**51**
13	*ssb1 (alr0088)*	**AGTA**cttatg**TACT**	8/6	28	TAATAA	**33**
14	*ssb2 (alr7579)*	**TAAGT**catt**ACTTA**	4	236	TATAAT	**224**
Upregulated genes						
1	*apcE (alr0020)*	**AGT**atgt**ACT**	4	225	GAAAAT	**190**
2	*rbcL (alr1524)*	**AGT**aaaagcgtta**ACT**	10	218	ACTTAT	**205**
3	*ahpC (alr4404)*	**TtGT**taaaaacc**ACTA**	8	98	TTAAGA	**115**

The consensus palindromic sequence '**AGT', 'ACT'** is written in bold and underlined. The base mismatch in this region is written in small bold caps. Additional bases contributing to the palindromic sequence are in bold and caps.

The binding of LexA was tested for all mutated *ssb1* promoter fragments **(Fig. 4.21B)**. The disruption of either left arm or right arm of the predicted LexA-Box, resulted in complete loss of binding LexA to the *ssb1* promoter while decrease in the spacer region did not affect the binding **(Fig. 4.21B)**. In case of the up-regulated genes, *apcE*, *rbcL* and *ahpC*, the same palindromic sequence was observed, with additional GAAA or GTTT sequences in the vicinity of the palindrome. The role of this sequence in gene regulation has, however, not been investigated.

Thus, the mutational studies confirmed the sequence "AGT-N_x-ACT" with variable spacer as the LexA binding site for *Anabaena* and, hereafter designated as AGT-N_{4-11}-ACT as the AnLexA box allowing for a one base mismatch provided it is present only in one of the two arms within a minimum of 4 bp palindrome. Presence of the AnLexA box was investigated upstream of genes under the functional category of DNA repair, C-metabolism and oxidative stress alleviation. In addition to these, few other genes corresponding to those suggested to be possibly regulated by LexA in *Anabeana* 7120 earlier (*132*) and fatty acid biosynthesis genes shown to be regulated by LexA in *Synechocystis* (*139*) were also included. As shown in **Table 4.9**, 14 genes belonging to category of DNA repair genes were found to have AnLexA box in the upstream regulatory region. This included *uvrA*, *recG,* topoI, *alr3988* (*sbcC*-like), *all4463* (*sbcD*-like), *recF*, *recO*, *recR*, helicase, *alr3199* and *alr3200* (**Table 4.9**). Regulation of *uvrA* (*130*) and *alr3988* (*sbcC*) and *all4463* (*sbcD*) (*180*) genes by *Anabaena* has already been proven experimentally, validating our prediction. Among those involved in oxidative stress alleviation, seven genes shown to possess the AnLexA-box were *katA*, *katB* (coding for catalase, involved in H_2O_2 removal), *sodA* (superoxode dismutase), glutaredoxin, dioxygenase, *petC* (component of the cytochrome b6-f complex, involved in the electron transfer) and FeS-binding protein **(Table 4.9)**. AnLexA-box was also

found to be present upstream of *hoxU*, *hoxEF* and *hupS* genes **(Table 4.9)** earlier shown to be up-regulated by LexA both in *Anabaena* and *Synechocystis* (*131-132*). In case of *hoxEF*, the AnLexA box overlapped with the sequence predicted earlier (*132*). AnLexA-box was also detected upstream of *fab* genes of *Anabaena* 7120 **(Table 4.9)**, whose homologues were shown to be regulated by LexA in *Synechocystis* 6803 (*139*). Additionally, the metal responsive regulator gene, *smtB* and a few other transcriptional regulators were found to possess AnLexA-Box upstream to the start codon **(Table 4.9)**.

Table 4.9: List of genes with the probable LexA-binding site in their upstream regulatory region

S. No.	Gene annotation (Symbol)	AnLexA-Box (AGT-N_x-ACT)			Predicted seq in Sjoholm et al., 2007[#]	No. Bases From ATG
		Sequence	Spacer region (x: no. of bases)	No. of bases upstream of ATG		
DNA repair genes						
1	*alr3716 (uvrA)*	**AGT**attattaaaa**ACT**	10	50	aGTACtattGTTCt	-84
2	*all4790- recG*	**AGT**tttttg**ACT**	6	168	gGTACaaatGTACg	-45
3	*alr2780 (TopoI)*	**AGT**ttctcccct**ACT**	9	318	gGTATcgttGTACt	-276
4	*alr3988 (sbcC)*	**AGT**gcatctgt**ACT**	8	52		
5	*all4463(sbcD)*	**AGTTG**aaaa**CAACT**	8	53		
6	*all3374(recF)*	**AGTTT**t**AAAtT**	5	50	aGTACagaaGTTCt	+654
7	*alr4175(recO)*	**TTtGT**cttc**ACTAA**	4	69		
8	*alr4977 (recR)*	**CAGTTT**gtga**AAtCTG**	8	71		
9	*all2951 (helicase)*	**AGTC**ttgttgt**GtCT**	9	218	aGAATtagtGTACc	+33
10	*alr2351 (radC)*	**TAGTTGCAtCTA**	4	36		
11	*alr3200* (DNase)	**AAtT**ggtata**ACTT**	6	49		
12	*alr3199* (DNase)	**AGT**taagaatccg**ACT**	10	38		
13	*alr2104* (methyl transferase)	**AGT**acataagt**ACT**	8	59	aGTACataaGTACt	-59
14	*all0061* (site sp. DNA methyl transferase)	**AGaA**aaa**TACT**	5	83	aGTACttttGTTCc	-869

C-metabolism genes						
15	*all0688 (hupS)*	**AGT**ggcaaaaca**ACT**	9	115	aGAACcagaGTTCc	-246
16	*alr0750 -hoxEF*	**AGT**tgagggt**ACT**	7	149	gGTACtctgGTTCg	-142
17	*alr0762 (hoxU)*	**AGT**tcctcaa**ACT**	7	90		
18	*alr1041 (fbp)*	**AGTACTTAAGTACT**	8	175	aGTACttaaGTACt	-175
19	*alr1095* (GAPDH)	**AGT**taacagttg**ACT**	9	55	gGTATcgctGTACg	+687
20	*all0475 (short chain DH)*	**AGT**tgattgttg**ACT**	9	74	gGTACtccaGTACg	+528
Fatty acid biosynthesis genes						
21	*alr0239* (*fabH)*	**AGT**aaatg**ACT**	5	55		
22	*alr3343 (fabF)*	**AAGT**caag**AtTT**	4	140		
23	*alr1894 (fabG)*	**TAGT**tcctca**ACgA**	6	88		
24	*alr2271* (*fabZ)*	**TAGcT**ttgagt**AACTA**	6	78		
Oxidative stress alleviation genes						
25	*alr3090 (katB)*	**AGT**ttctggtttg**ACT**	10	127		
26	*all0070* (*sodA*)	**AGT**aata**ACT**	4	369		
27	*asl3860 (glutaredoxin)*	**AGT**ttataact**ACT**	8	77	aGTACcaaaGTTCt	-46
28	*all3866 dioxygenase*	**AGT**aacaaatgaca**ACT**	11	46	gGTACttgtGTTCc	-234
29	*asr4942- FeS binding protein*	**AGT**actgcca**ACT**	7	45	aGTACtatgGTACg	-53
30	*alr0998 (katA)*	**GAGT**tttttgag**tCTC**	8	324	gGAACaggaGTACt	+630
31	*all1512 (petC)*	**AGTC**tcttaaagg**GAtT**	11	45	gGTACtgtaGTACg	-311
32	*asl0401*(nitro-reductase)	**AGTAC**ttttt**GTACT**	8	71	aGTACttttGTACt	-71
Metal responsive genes						
33	*all7621 (smtB)*	**AGTT**atca**AtCT**	4	21		
34	*alr0831 (smtB)*	**AGT**gggt**ACT**	4	162		
Toxin-antitoxin genes						
35	*alr4239* toxin secretory pr	**AGT**aaaa**ACT**	4	58	gGTACttttGTTCg	-241
36	*asr0757 (mazE)*	**AGT**aaaa**ACT**	4	60		

37	*all0754 (vapC)*	**TAGTATATACcA**	4	53		
Transcriptional regulators						
38	*alr1194* (two-component regulator)	**AGT**gag**AACT**	5	429	aGAACtttaGTTCt	-425
39	*alr2481* (two-component regulator)	**TAGT**tctt**ACTA**	4	257	aGAACtttaGTTCt	-265
40	*alr1044* (transcriptional regulator)	**TAGT**aattcga**tCTA**	7	53	gGTATttatGTACt	-472

4.3 DISCUSSION

LexA, a negative regulator of more than 30 genes which are part of SOS response (*42*) has been shown to modulate response to DNA damage in bacteria (*162*). However, in addition to the LexA non-cleavable mutants being susceptible to UV-radiation and mitomycin C in *E. coli* and several other bacteria (*172-173*), the partial *lexA* mutants have also been shown to modulate cold shock response in a deep-sea bacteria, *Shewanella piezotolerans* WP3 (*181*) and C-starvation in *Synechocystis* (*127*). Various studies on LexA in cyanobacteria revealed its predominant role in C-metabolism (*131-132, 170*), regulation of *uvrA*, *recA* (*130*) and *ssb* among DNA repair genes (*171*), fatty acid biosynthesis genes (*139*). This suggested possible multifunctional role for LexA in cyanobacteria. To get a better insight into the possible role of LexA in the N_2-fixing cyanobacterium *Anabaena*, attempts were made to generate *lexA* insertional mutant and LexA overexpressing strains of *Anabaena*. Unlike in *Synechocystis*, wherein the partially segregated *lexA* mutant was stable (*127*), that of *Anabaena* was highly unstable. On the other hand, constitutive overexpression of LexA did not hamper growth under normal physiological conditions **(Fig. 4.6)**, resulting in the generation of a fully segregated A*nabaena* strain overexpressing LexA, An*lexA*$^{+}$.

Transcriptomic based approach in *Synechocystis* revealed a plethora of genes involved in multitude of functions to be regulated by LexA (*127, 170*). However, unlike a single consensus sequence for LexA box in *E. coli* (*54*), a great deal of variation was observed in case of cyanobacteria (*130, 138, 166, 169*). A consensus sequence was built around these variants and using possible degeneracy in the sequence, several *Anabaena* genes were found to possess the probable LexA-binding region upstream or downstream of the gene (*132*). The major drawback, however, was that no mechanism of up and down-regulation by LexA could be postulated, which could explain the ability of LexA to regulate the identified genes (*132*), irrespective of the location of the box with respect to the transcriptional or translation start sites of the genes. An attempt was made to fill this lacuna by identifying and confirming the most probable consensus LexA box using a platform of 15-20 genes of *Anabaena*, regulated by LexA. The observed modulation of different abiotic stress responses of *Anabaena* upon overexpression of LexA **(Fig. 4.7-4.15)** confirmed that LexA would indeed be regulating several genes belonging to different functional groups either directly or indirectly through other repressors or activators. A combination of proteomic approach **(Figs. 4.16-4.18, Table 4.3)**, transcriptomic **(Fig. 4.19)**, EMSA **(Fig. 4.20)** and promoter activity analysis **(Table 4.7)**, helped identify 14 genes down-regulated and 3 genes up-regulated by LexA directly. These belonged to different functional categories which is in line with the observations in *Synechocystis* RNA-sequence analysis (*170*).

Analysis of the DNA sequence of all 17 promoters, which included 6 DNA repair genes, revealed the presence of the common palindromic sequence AGC-N_{4-11}-ACT with a maximum of one bp mismatch in a 4-base palindrome, designated as AnLexA Box. The importance of the palindromic sequence and allowance of variation in the space region was confirmed by mutation analysis **(Fig. 4.21B)**. The sequence was found

upstream of 40 more genes belonging to different functional categories **(Table 4.9)**, such as DNA repair genes (14 genes), C-metabolism (6 genes), fatty acid biosynthesis (4 genes), oxidative stress alleviation (8 genes), metal responsive (2 genes), toxin-antitoxin (3 genes) and transcriptional regulators (3 genes). In all cases the AnLexA Box was in close vicinity of an AT-rich region, a probable -10 region, allowing for better regulation of transcription. Identification of these genes possessing the AnLexA Box along with those identified experimentally through proteomic and transcript studies complement the observed changes in stress tolerance upon overexpression of LexA in *Anabaena*.

The major functional group of genes found to be regulated by LexA were those involved in DNA repair and metabolism, which encompassed a total of 21 genes/proteins **(Tables 4.3, 4.8 and 4.9)**. Of these, only Ndk was detected during proteomic studies (**Fig. 4.18E**). Of the two proteins involved in DNA metabolism, Ndk (nucleoside diphosphate kinase) is involved in maintaining a pool of ribose and deoxyribose triphosphates for RNA and DNA synthesis by catalysing transfer of γ-phoshphate from nucleoside triphosphate to precursor corresponding diphosphates (*182*). Tdk (Thymidylate kinase), is involved in the conversion of deoxythymidine monophosphate (dTMP) to deoxythymidine diphosphate (dDMP) (*183*). Thus, the down-regulation of at least two genes involved in synthesis of precursors for DNA synthesis or replication repair would affect the ability of the cells overexpressing LexA to repair its damaged DNA. Among the listed DNA repair genes, predicted based on the studies in *E. coli*, role of only four proteins in DNA repair has been demonstrated in *Anabaena*. These include Alr0088 (SSB1) (*184*), Alr3200 (*185*), Alr3988 (SbcC) and All4463 (SbcD), which have been shown to be down-regulated by LexA (*180*). Thus, the down-regulation of genes coding for DNA repair proteins would result in lower repair of damaged DNA, thus contributing to the observed increased sensitivity of An*lexA*$^{+}$ cells to DNA damage

inducing agents **(Fig. 4.9 and 4.10)**. This is in complete congruence with the known role for bacteria LexA in SOS response

The 2nd major functional category of LexA-regulated genes included those involved in C-metabolism which included photosynthetic light reaction, C-fixation and C-catabolism. The light reaction is carried out by phycobilisome complex, which is composed of three pigment proteins i.e. phycocyanin (Cpc), phycoerythrin (Pec), allophycocyanin (ApcA,B and Apc) and two linker proteins i.e. rod-linker protein (CpcG) and phycobilisome core membrane-linker protein (ApcE), by harvesting of light and transfer the energy to the photosystems (*186-187*). Of these, the expression of *apcE* was up-regulated by LexA in *Anabaena* (**Fig. 4.19B**), resulting in its increased abundance, while *cpcBACDE* operon was down-regulated by LexA **(Fig. 4.19C)**, accounting for the observed decrease in the total abundance of the CpcC and CpcG proteins **(Fig. 4.18A)**. A few protein spots corresponding to PecA, PecB, CpcB, ApcB and ApcA did not exhibit change in their abundance **(Fig. 4.16B)**. Of these, the AnLexA-box in the vicinity of AT-rich region was not detected upstream of *pecAB* and *apcAB* operons. Photosynthetic efficiency of An*lexA*$^{+}$ under control conditions was found to be 7.1 ± 0.2 μmol O_2 mg Chl a^{-1} min^{-1}, which is comparable to that observed for AnpAM cells (7.4 ± 0.1 μmol O_2 mg Chl a^{-1} min^{-1}). Variaitons in the proportion of the pigments constituting the light harvesting phycobillisomes result in cyanobacteria adjusting their light-harvesting capabilities through the phenomenon of complementary chromatic acclimation (*188*). LexA overexpression increased the abundance of RbcL protein in *Anabaena* **(Fig. 4.18B)** through positive regulation by LexA through the AnLexA-box **(Fig. 4.20C, Table 4.7)**. In the *lexA* mutant of *Synechocystis*, the *rbcS* transcript levels were found to be decreased (*170*). The *rbcL* and *rbcS* genes have been shown to be co-transcribed in both unicellular (*189*) and filamentous cyanobacteria (*190*). Thus, it can be

stated that the corresponding operon is up-regulated by LexA in cyanobacteria. The *rbcL* and *rbcS* genes code for the large and small subunits of RuBisCo [Ribulose 1, 5-Bisphosphate Carboxylase/oxygenase], is involved in the assimilation of CO_2, and its activity is coupled with that of light reaction of photosynthesis (*191*). With increased RuBisCo levels and unaffected photosynthetic efficiency in An*lexA*$^+$ cells, higher rate of C-fixation is expected in these cells. Three proteins involved in breakdown of sugar, Transaldolase B (TalB), UDP-4-galactose epimerase (GalE) and probable α-glucanotransferase (GlgB), showed decreased abundance in An*lexA*$^+$ cells **(Fig. 4.18C)**, which may have an overall effect in reducing C-catabolism. This in combination with the expected increase in C-fixation would allow maintaining higher C-levels in *Anabaena*. This was reflected in the ability of An*lexA*$^+$ cells to tolerate inorganic C-starvation induced by DCMU better than the AnpAM cells **(Fig. 4.7)**. This also correlated well with the observed sensitivity to inorganic C-starvation in LexA-depleted cells of *Synechocystis* (*127*).

The 3rd major group of proteins showing differential expression on the face of overexpression of LexA or possessing the AnLexA-box were those involved in alleviation of oxidative stress. The AnLexA Box was found upstream of *sodA*, coding for MnSOD and *kat* genes coding for catalase, and negatively by LexA accounting for the decreased MnSOD activity **(Fig. 4.13C)** and catalase levels **(Fig. 4.12B)** in An*lexA*$^+$ cells. This concomitant with decreased abundance of PetH and PrxA, not compensated by increase in that of AhpC **(Fig. 4.18D)** further decreases the ability of An*lexA*$^+$ cels in detoxifying the accumulated ROS during oxidative stress resulting in the observed lower tolerance of these cells to different forms of oxidative stresses **(Figs. 4.11-4.15)**. *Anabaena* cells responds to exposure to heavy metals by modulating their protein expression. In addition to modulation of levels of proteins involved in photosynthesis of

C-metabolism and other general stress proteins, a few proteins involved in metal transport or reduction are also regulated. In response to arsenic stress, the major arsenic related proteins regulated are *alr1105* and *all0195* (*192-193*); while *all3940* and *ahpC* genes were found to be involved in cadmium stress management (*194-195*). Of these only AhpC was detected during proteomic studies in An*lexA*$^{+}$ and shown to have AnLexA-box in the upstream regulatory regions. But, the AnLexA-box was found upstream of *all7621* and *alr0831* genes **(Table 4.9)** which code for SmtB-homologs, which functions as a negative transcriptional regulator of metallothionin, SmtA (*196*). Thus the LexA may be indirectly regulating the expression of metal responsive genes, through the regulation of the repressor SmtB-like protein. The AnLexA-box was also found upstream of the three transcriptional regulators, *alr1194*, *alr2481*, and *alr1044* (**Table 4.9**). Thus, LexA could be involved in indirect regulation of several other genes as well.

Recently, fatty acid biosynthesis genes of *Synechocystis* 6803 were shown to be down-regulated by LexA (*139*). The corresponding genes of *Anabaena* 7120, *alr0239* (*fabH*), *alr3343* (*fabF*), *alr1894* (*fabG*) and *alr2271* (*fabZ*) were also found to posses the AnLexA-box **(Table 4.9)** and may be down-regulated by LexA. Fatty acids and lipids play a very important role in stress management in cyanobacteria (*197*) and thus the down-regulation of fatty acid biosynthesis genes upon overexpression of LexA would result in decrease in tolerance to various abiotic stresses.

Based on the proteomic, transcript, promoter and bioinformatic studies, the AGT-$N_{4\text{-}11}$-ACT sequence (AnLexA-box) is a strong contender for the LexA box in *Anabaena* encompassing the earlier predicted LexA-boxes (*130, 132, 169*). The sequence AGT-N_X-ACT has also been found upstream of the genes shown to be regulated by LexA in *Synechocystis* (*131, 133, 139*), within their predicted regulatory region. Thus, the

sequence may have greater implications as LexA-binding box across cyanobacterial sp., though this needs to be investigated thoroughly. The presence of AnLexA-box upstream of genes belonging to varied functional groups, indicating a global regulatory role- both direct and indirect through other regulators. Thus, the functional designation of LexA from being a SOS-response regulator has been expanded to a general global regulator of stress response in cyanobacteria. Cyanobacteria being ancient organisms, it is possible that the multi-gene regulatory role of LexA was lost during the course of evolution to modern day bacteria and became restricted to only SOS-response.

CHAPTER 5

SUMMARY

Bacteria are continuously exposed to endogenous and exogenous agents that can cause damage to DNA and if unrepaired, can be lethal. One of the first responses to DNA damage is the SOS response, which is negatively regulated at the transcriptional level by LexA protein, across almost all eubacterial groups. In *E. coli*, it has been found to regulate expression of more than 30 genes involved in DNA repair (*42*). Under normal growth conditions, when DNA damage is low, LexA in dimeric form binds to the SOS box (LexA box) present in the vicinity of the promoter region of the SOS-regulated DNA repair genes. This results in repression of expression of these genes at transcriptional level. When DNA is damaged, it generates single stranded DNA (ssDNA) which interacts with RecA protein resulting in the formation of active RecA nucleofilaments (*54*). The activated RecA upon interaction with LexA induces conformational changes in the protein, converting it from non-cleavable to cleavable form (*43*), and leading to the autoproteolytic cleavage of LexA (*107*). The LexA protein comprises of two domains, namely N-terminal domain (NTD) spanning from 1- 69, and C-terminal domain (CTD) spanning from 75 to 198 (*43*). DNA binding domain is present in NTD, while the dimerization domain as well as cleavage site (Ala^{84}-Gly^{85}) and residues of active site residues (Ser^{119}, Lys^{156}) are present in the CTD of LexA. Proteolytic cleavage of LexA involves the breaking of the peptide bond between Ala^{84} and Gly^{85}, and the resulting NTD and CTD of the protein are further degraded by proteases (*163-164*). This leads to derepression of the down-stream regulated genes.

LexA protein has also been found in the ancient organism, cyanobacteria which inhabited the earth over 3.5 billion years ago (*108*). Of the 33 cyanobacteria whose genome sequence is available, only 6 do not possess the LexA protein (*169*). Majority of studies on LexA in cyanobacteria has been restricted to the unicellular *Synechocystis*

PCC 6803 and the filamentous nitrogen-fixing *Anabaena* PCC 7120. The LexA protein of *Synechocystis* was reported to be non-cleavable (*169*), and shown to function as an activator of C-metabolism genes (*133*) but not in DNA repair (*128*). On the other hand, LexA protein of *Anabaena* possessed all residues required for cleavage and thus predicted to be cleavable and found to regulate expression of few DNA repair genes such as *recA*, *ssb* and *uvrA* (*130, 171*) as well as bidirectional hydrogenase (*132*). The LexA box in cyanobacteria showed a great deal of variation (*130*) and was distinct from that for *E. coli* (*54*). Based on the predicted sequence of the LexA box in *Anabaena*, bioinformatic analysis revealed its presence in regions upstream, downstream or within 216 genes (*132*). However, these were not experimentally verified and the position of the LexA box with respect to the transcriptional or translational start site varied drastically. The genes predicted to be regulated by LexA belonged to different functional groups in *Anabaena* (*132*). This was similar to the observation in *Synechocystis*, wherein transcriptomic studies on partial *lexA* deletion mutant revealed up- or down-regulation of genes belonging to varied functional groups (*127*). These studies revealed a different or atypical role for cyanobacterial LexA as a regulatory protein. However, a few questions remained unanswered: (i) does cyanobacterial LexA exhibit autoproteolytic cleavage similar to that of *E. coli* ? (ii) how does LexA carry out its role as regulator ?, and (iii) the physiological role of LexA in cyanobacteria. To answer these questions, a detailed investigation was carried out on LexA in the cyanobacterium *Anabaena* PCC 7120.

The LexA protein of *Anabaena* 7120 is encoded by *alr4908. Anabaena* LexA was overexpressed in *E. coli* by cloning the gene in pET16b expression vector. The purified *Anabaena* LexA protein was found to exist as a single species of about 55 kDa under native conditions based on gel filtration studies. This indicated that the protein exists as a dimer, with an estimated hydrodynamic diameter of 9.4 ± 3.5 nm based on

dynamic light scattering (DLS) studies. However, on SDS-polyacrylamide gels, a 13 kDa protein was found to be co-eluted with the monomeric 25 kDa LexA. Immunodetection with the antibodies raised against the 25 kDa LexA revealed the 13 kDa protein to be LexA as well. Based on the molecular mass, it was likely to be the C-terminal cleaved product of LexA and this gave the first indication that autoproteolytic cleavage does occur in cyanobacterial LexA. Analysis of autoproteolytic cleavage of *Anabaena* LexA as a function of pH, revealed that the cleavage occurred at pH 8.0 and above under *in vitro* conditions. However, under i*n vivo* conditions, no cleavage of *Anabaena* LexA was observed under physiological pH of 7.5 either in *E. coli* overexpressing *Anabaena* LexA or in *Anabaena* PCC7120. Cleavage of LexA under physiological conditions is aided by activated RecA. However, when tested under *in vitro* conditions, activated RecA from neither *Anabaena* nor *E. coli* could induce the cleavage of *Anabaena* LexA at pH 7.5. This could possibly be due to the absence of RecA interacting sites on *Anabaena* LexA protein and may have necessitated by the low levels of RecA in *Anabaena* which would not be sufficient for inducing the cleavage of LexA under physiological conditions. The rate of cleavage of LexA was observed to increase with increase in pH from 8.0 to 10.0 and increase in temperature above 25 °C till 46 °C. The cleavage was aided further in presence of calcium (Ca^{2+}).

Point mutants of LexA protein were generated to determine the residues involved in its cleavage. The results revealed that *Anabaena* LexA is cleaved between Ala^{84} and Gly^{85} and required the presence of active site residues Ser^{118} and Lys^{156} and a three amino acid stretch ($^{86}GLI^{88}$) in the cleavage site region (CSR). The GLI region was found to be highly conserved among the CSR of different cyanobacterial LexA protein. For assessing the physiological role of LexA in *Anabaena*, first the expression of LexA protein and the corresponding transcripts in response to various DNA damage inducing

agents was monitored in *Anabaena*. Under different DNA damaging conditions, though induction of *lexA* was observed at the transcript level, there was no change in protein levels nor any cleavage observed for LexA upon exposure of *Anabaena* 7120 to UV-radiation or mitomycin C stress. This indicated regulation at both transcriptional and post-transcriptional levels of the *lexA* gene, possibly to ensure lower levels of the protein so that the expression of DNA repair genes is not hampered. When exposed to stresses inducing DNA strand breaks such as desiccation and γ-irradiation stresses, the expression of *lexA* was repressed both at the transcript as well as the protein level. Regulation of *lexA* in response to stresses inducing DNA damage suggested a possible role for *Anabaena* LexA in the process of DNA repair. Since, the classical SOS response involves cleavage of LexA in response to DNA damage, which is not observed in *Anabaena*, the existence of the SOS-response *per se* in classical terms in *Anabaena* may be debated.

A recombinant strain of *Anabaena*, constitutively overexpressing LexA (*AnlexA*$^{+}$) was generated to understand the role of LexA in *Anabaena*. Unlike the generation of a viable partially segregated deletion mutant of *lexA* in unicellular *Synechocystis* (*127*), in *Anabaena* these were not viable. The An*lexA*$^{+}$ cells did not exhibit any defect in growth or morphology compared to vector control Neo^{r} recombinant *Anabaena* (AnpAM) cells (*Anabaena* cells harbouring pAM1956 plasmid) as well as wild type *Anabaena* 7120 under control conditions. However, upon being subjected to different abiotic stresses, An*lexA*$^{+}$ showed differential growth and survival compared to the AnpAM cells. All comparisons were carried out under N-supplemented conditions as the An*lexA*$^{+}$ strain was not stable under N_2-fixing conditions for more than two generations. When exposed to photosynthesis inhibitor, DCMU [3-(3, 4-dicholophyllphenyl)-1, 1-dimethylurea)], An*lexA*$^{+}$ cells exhibited better growth than AnpAM cells. However, the survival as well

as the tolerance of An*lexA*$^+$ cells decreased significantly compared to AnpAM cells when exposed to DNA damage inducing stresses such as, UV, γ-radiation or desiccation or different forms of oxidative stresses, applied with H_2O_2, Methyl viologen (MV) or heavy metals [Cd(II) and As(V)]. This suggested that the overexpression of LexA probably reduced the expression of proteins involved in the repair of damaged DNA and alleviation of oxidative stress. Two of the major proteins involved in the oxidative stress management in *Anabaena* are superoxide dismutase (*147*) and catalase (*198*). Zymogram and Western blotting-immunodetection analysis revealed lower levels of MnSOD activity and KatB protein in An*lexA*$^+$ compared to AnpAM cells both under control and oxidative stress conditions. This resulted in higher accumulation of reactive oxygen species (ROS) in An*lexA*$^+$ cells compared to AnpAM observed microscopically contributing to lower oxidative stress tolerance of An*lexA*$^+$ cells.

Comparative 2-D proteomic analysis of An*lexA*$^+$ and AnpAM cells revealed altered abundance for 47 protein spots, which included 30 with decreased abundance and 17 with increased abundance in An*lexA*$^+$ compared to AnpAM under control growth conditions. Twenty three protein spots could be identified conclusively, which included 15 corresponding to those showing decreased abundance and 8 exhibiting increased abundance. These corresponded to 11 different genes, of which three *apcE*, *rbcL* and *ahpC* genes coded for proteins showing increased abundance. These could be grouped into two major functional categories: C-metabolism and oxidative stress alleviation. None of the genes or proteins with a role in DNA repair could, however be identified. The functional category of C-metabolism included those involved in (i) light reaction wherein CpcC and CpcG proteins showed decreased abundance while ApcE showed increased abundance, but no effect on the overall photosynthetic efficiency, (ii) C-fixation, wherein RbcL showed increased abundance, and (iii) C-catabolism, wherein

GlgB, TalB and GalE showed decreased abundance. The increase in C-fixation and decrease in its breakdown would enhance the intracellular storage of carbon and thus help the An*lexA*$^{+}$ cells survive DCMU stress better than AnpAM. Among the oxidative stress alleviation proteins, only AhpC showed increased abundance, while decreased abundance of PrxA and PetH along with MnSOD and KatB contributed to the observed lower oxidative stress tolerance An*lexA*$^{+}$ cells. The other three proteins identified were (i) DnaK which functions as a molecular chaperone, (ii) Ndk in nucleotide biosynthesis, and (iii) GlnA in amino acid biosynthesis, all of which showed decreased abundance in An*lexA*$^{+}$ cells affecting their overall tolerance to abiotic stresses in general.

To assess the role of LexA in the regulation of expression of the identified proteins, the transcriptional regulation of the corresponding genes along with 5 DNA repair genes (*recA, lexA, ssb1, ssb2 and tdk*) by LexA was investigated through (i) transcript analysis, (ii) EMSA studies to evaluate binding of LexA to their upstream regions, and (iii) promoter activities in *E. coli* in the presence or absence of *Anabaena* LexA. Comparative transcript analysis of these genes indicated up-regulation of the gene corresponding to proteins showing increased abundance and down-regulation of those with decreased abundance as well as the chosen DNA repair genes. This indicated that *Anabaena* LexA also functions at transcriptional level in regulating the expression of genes, as is the case for other bacteria. LexA was found to bind to upstream regions of all the genes tested, with the exception of *ssb3*, and having binding affinity in the range of 18 – 96 nM. The *ssb3* promoter was used as a negative control for all analysis, as *ssb3* has been shown to be regulated by FurA and not LexA (*171*). A similar range of binding affinity suggested a similar LexA box. The up- and down-regulation of the promoter activities of different genes tested upon overexpression of *Anabaena* LexA in *E. coli* correlated well with the observed changes in proteomic and transcript studies.

Bioinformatic analysis of upstream regions of these 17 genes was carried out to build a consensus LexA box which was found to be an imperfect palindromic sequence [AGT-N_{4-11}-ACT] in 10 of the 17 gene promoters, and with one base variation in the remaining 7 gene promoters. The importance of palindrome of LexA box was confirmed by mutational analysis of the arms using *ssb1* promoter wherein LexA failed to bind to the promoter upon disruption of the palindrome. However, when spacer region was shortened from 8 bases to 4 bases, the binding of LexA to the DNA fragment was unaffected. This indicated that the palindrome is essential while the spacer is flexible in case of AnLexA-box. The presence of longer hinge region in *Anabaena* LexA between the NTD which has DNA binding domain and the CTD which has the dimerization domain possibly gives the protein enough flexibility to bind to a varied spacer region. The sequence was designated as the AnLexA-box in *Anabaena* 7120. A short genome search revealed the presence of this element, in the vicinity of an AT-rich region, upstream of at least 40 genes in *Anabaena,* which included 6 genes belonging to the functional group of C-metabolism, 14 genes which possibly involved in DNA repair, 8 genes contributing to the oxidative stress alleviation, 4 genes shown to be involved in fatty acid biosynthesis, 3 transcriptional regulator genes, 2 metal responsive genes and 3 toxin-antitoxin genes. These results from our study suggest a global regulatory role for LexA in *Anabaena* 7120 as depicted in **Fig. 5.** This can possibly be extended to cyanobacteria in general, as the same has been speculated for unicellular *Synechocystis* as well where it has been shown to regulate genes involved in C-metabolism (*127*), fatty acid biosynthesis (*139*) and other functional categories (*170*), but not DNA repair (*128*). The non-cleavable form of *Synechocystis* LexA could possibly be the reason why LexA is not regulating DNA repair genes in *Synechocystis*, but does so in *Anabaena* where LexA is cleavable. It can be speculated that during the course of evolution from the

ancient cyanobacteria which existed under extreme environmental conditions, to the modern day bacteria living under moderate environmental conditions, the role of LexA narrowed down, restricting it to that of SOS response regulator, involved in regulating DNA repair genes only.

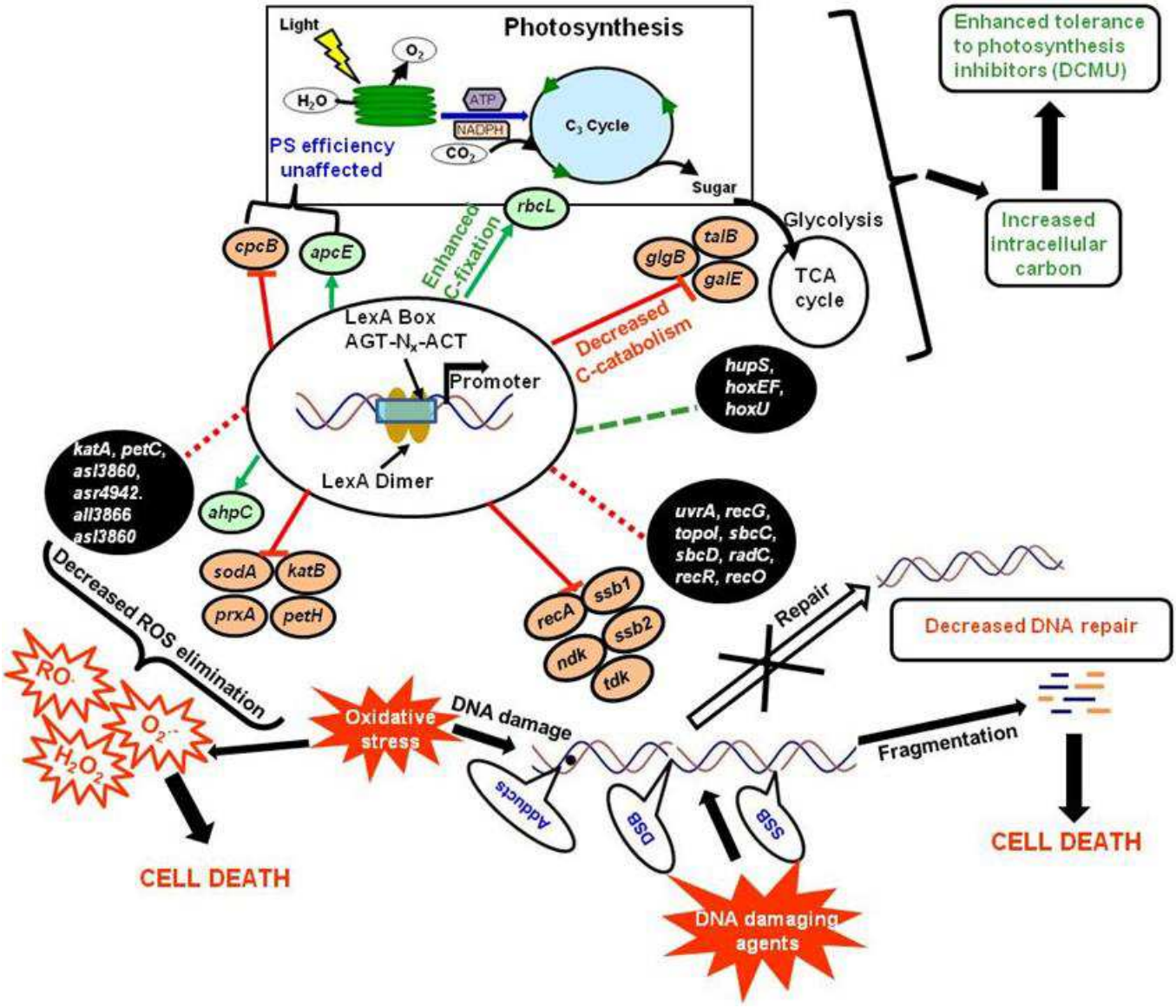


Fig. 5 Global regulation by LexA in *Anabaena* 7120. Red colour indicates down regulation of genes and green colour indicates up-regulation of genes. Dotted lines indicate probable regulation by LexA. DSB stands for double strand breaks and SSB stands for single stranded breaks. Green colour arrows and red colour arrows indicate positive regulation and negative regulation by LexA respectively.

REFERENCES

1. Chakarov, S., Petkova, R., Russev, G. C., and Zhelev, N. (2014) DNA damage and mutation. Types of DNA damage, *BioDiscovery 11*, e8957.
2. Imlay, J. A. (2013) The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium, *Nat Rev Microbiol 11*, 443-454.
3. Slupphaug, G., Kavli, B., and Krokan, H. E. (2003) The interacting pathways for prevention and repair of oxidative DNA damage, *Mutat Res 531*, 231-251.
4. Hakem, R. (2008) DNA-damage repair; the good, the bad, and the ugly, *Embo j 27*, 589-605.
5. Bargonetti, J., Champeil, E., and Tomasz, M. (2010) Differential toxicity of DNA adducts of mitomycin C, *J Nucleic Acids 2010*.
6. Thiviyanathan, V., Somasunderam, A., Volk, D. E., Hazra, T. K., Mitra, S., and Gorenstein, D. G. (2008) Base-pairing properties of the oxidized cytosine derivative, 5-hydroxy uracil, *Biochem Biophys Res Commun 366*, 752-757.
7. Holliday, R., and Grigg, G. W. (1993) DNA methylation and mutation, *Mutat Res 285*, 61-67.
8. Casadesus, J. (2016) Bacterial DNA Methylation and Methylomes, *Adv Exp Med Biol 945*, 35-61.
9. Sinha, R. P., and Hader, D. P. (2002) UV-induced DNA damage and repair: a review, *Photochem Photobiol Sci 1*, 225-236.
10. Confalonieri, F., and Sommer, S. (2011) Bacterial and archaeal resistance to ionizing radiation, *Journal of Physics: Conference Series 261*, 012005.
11. Morita, R., Nakane, S., Shimada, A., Inoue, M., Iino, H., Wakamatsu, T., Fukui, K., Nakagawa, N., Masui, R., and Kuramitsu, S. (2010) Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems, *J Nucleic Acids 2010*, 179594.

12. Yi, C., and He, C. (2013) DNA repair by reversal of DNA damage, *Cold Spring Harb Perspect Biol 5*, a012575.

13. Muller, M., and Carell, T. (2009) Structural biology of DNA photolyases and cryptochromes, *Curr Opin Struct Biol 19*, 277-285.

14. Brettel, K., and Byrdin, M. (2010) Reaction mechanisms of DNA photolyase, *Curr Opin Struct Biol 20*, 693-701.

15. Liu, Z., Wang, L., and Zhong, D. (2015) Dynamics and mechanisms of DNA repair by photolyase, *Phys Chem Chem Phys 17*, 11933-11949.

16. Zhao, X., Liu, J., Hsu, D. S., Zhao, S., Taylor, J. S., and Sancar, A. (1997) Reaction mechanism of (6-4) photolyase, *J Biol Chem 272*, 32580-32590.

17. Maul, M. J., Barends, T. R., Glas, A. F., Cryle, M. J., Domratcheva, T., Schneider, S., Schlichting, I., and Carell, T. (2008) Crystal structure and mechanism of a DNA (6-4) photolyase, *Angew Chem Int Ed Engl 47*, 10076-10080.

18. Scharer, O. D. (2003) Chemistry and biology of DNA repair, *Angew Chem Int Ed Engl 42*, 2946-2974.

19. Falnes, P. O., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*, *Nature 419*, 178-182.

20. Wallace, S. S. (2014) Base excision repair: a critical player in many games, *DNA Repair (Amst) 19*, 14-26.

21. Fromme, J. C., and Verdine, G. L. (2004) Base excision repair, *Adv Protein Chem 69*, 1-41.

22. Van Houten, B. (1990) Nucleotide excision repair in *Escherichia coli*, *Microbiol Rev 54*, 18-51.

23. Kamarthapu, V., and Nudler, E. (2015) Rethinking transcription coupled DNA repair, *Curr Opin Microbiol 24*, 15-20.

24. Kisker, C., Kuper, J., and Van Houten, B. (2013) Prokaryotic nucleotide excision repair, *Cold Spring Harb Perspect Biol 5*, a012591.

25. Yasui, A. (2013) Alternative excision repair pathways, *Cold Spring Harb Perspect Biol 5*.

26. Yang, W. (2000) Structure and function of mismatch repair proteins, *Mutat Res 460*, 245-256.

27. Ban, C., and Yang, W. (1998) Structural basis for MutH activation in *E.coli* mismatch repair and relationship of MutH to restriction endonucleases, *Embo j 17*, 1526-1534.

28. Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA, *Nature 407*, 703-710.

29. Putnam, C. D. (2016) Evolution of the methyl directed mismatch repair system in *Escherichia coli*, *DNA Repair (Amst) 38*, 32-41.

30. Iyer, R. R., Pluciennik, A., Burdett, V., and Modrich, P. L. (2006) DNA mismatch repair: functions and mechanisms, *Chem Rev 106*, 302-323.

31. Wigley, D. B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB, *Nat Rev Microbiol 11*, 9-13.

32. Ayora, S., Carrasco, B., Cardenas, P. P., Cesar, C. E., Canas, C., Yadav, T., Marchisone, C., and Alonso, J. C. (2011) Double-strand break repair in bacteria: a view from *Bacillus subtilis*, *FEMS Microbiol Rev 35*, 1055-1081.

33. Lenhart, J. S., Schroeder, J. W., Walsh, B. W., and Simmons, L. A. (2012) DNA repair and genome maintenance in *Bacillus subtilis*, *Microbiol Mol Biol Rev 76*, 530-564.

34. Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., and Sinha, R. P. (2010) Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair, *J Nucleic Acids 2010*, 592980.

35. Shuman, S., and Glickman, M. S. (2007) Bacterial DNA repair by non-homologous end joining, *Nat Rev Microbiol 5*, 852-861.

36. Symington, L. S., and Gautier, J. (2011) Double-strand break end resection and repair pathway choice, *Annu Rev Genet 45*, 247-271.

37. Chayot, R., Montagne, B., Mazel, D., and Ricchetti, M. (2010) An end-joining repair mechanism in *Escherichia coli*, *Proc Natl Acad Sci U S A 107*, 2141-2146.

38. Sfeir, A., and Symington, L. S. (2015) Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway?, *Trends Biochem Sci 40*, 701-714.

39. Pannunzio, N. R., Li, S., Watanabe, G., and Lieber, M. R. (2014) Non-homologous end joining often uses microhomology: implications for alternative end joining, *DNA Repair (Amst) 17*, 74-80.

40. Radman, M. (1975) SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis, *Basic Life Sci 5a*, 355-367.

41. Little, J. W., and Mount, D. W. (1982) The SOS regulatory system of *Escherichia coli*, *Cell 29*, 11-22.

42. Fernandez De Henestrosa, A. R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J. J., Ohmori, H., and Woodgate, R. (2000) Identification of additional genes belonging to the LexA regulon in *Escherichia coli*, *Mol Microbiol 35*, 1560-1572.

43. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. (2001) Crystal structure of LexA: a conformational switch for regulation of self-cleavage, *Cell 106*, 585-594.

44. Singleton, S. F., and Xiao, J. (2001) The stretched DNA geometry of recombination and repair nucleoprotein filaments, *Biopolymers 61*, 145-158.

45. Kuzminov, A. (1999) Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda, *Microbiol Mol Biol Rev 63*, 751-813.

46. Michel, B. (2005) After 30 years of study, the bacterial SOS response still surprises us, *PLoS Biol 3*, e255.

47. Rehrauer, W. M., Bruck, I., Woodgate, R., Goodman, M. F., and Kowalczykowski, S. C. (1998) Modulation of RecA nucleoprotein function by the mutagenic UmuD'C protein complex, *J Biol Chem 273*, 32384-32387.

48. Yasuda, T., Morimatsu, K., Kato, R., Usukura, J., Takahashi, M., and Ohmori, H. (2001) Physical interactions between DinI and RecA nucleoprotein filament for the regulation of SOS mutagenesis, *Embo j 20*, 1192-1202.

49. Petrova, V., Chitteni-Pattu, S., Drees, J. C., Inman, R. B., and Cox, M. M. (2009) An SOS inhibitor that binds to free RecA protein: the PsiB protein, *Mol Cell 36*, 121-130.

50. Galkin, V. E., Britt, R. L., Bane, L. B., Yu, X., Cox, M. M., and Egelman, E. H. (2011) Two modes of binding of DinI to RecA filament provide a new insight into the regulation of SOS response by DinI protein, *J Mol Biol 408*, 815-824.

51. Shinagawa, H. (1996) SOS response as an adaptive response to DNA damage in prokaryotes, *Exs 77*, 221-235.

52. Baharoglu, Z., and Mazel, D. (2014) SOS, the formidable strategy of bacteria against aggressions, *FEMS Microbiol Rev 38*, 1126-1145.

53. Pennington, J. M., and Rosenberg, S. M. (2007) Spontaneous DNA breakage in single living *Escherichia coli* cells, *Nat Genet 39*, 797-802.

54. Walker, G. C. (1984) Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*, *Microbiol Rev 48*, 60-93.

55. Boubakri, H., de Septenville, A. L., Viguera, E., and Michel, B. (2010) The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units in vivo, *Embo j 29*, 145-157.

56. Wimberly, H., Shee, C., Thornton, P. C., Sivaramakrishnan, P., Rosenberg, S. M., and Hastings, P. J. (2013) R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*, *Nat Commun 4*, 2115.

57. Rodriguez-Beltran, J., Rodriguez-Rojas, A., Guelfo, J. R., Couce, A., and Blazquez, J. (2012) The *Escherichia coli* SOS gene dinF protects against oxidative stress and bile salts, *PLoS One 7*, e34791.

58. Konola, J. T., Sargent, K. E., and Gow, J. B. (2000) Efficient repair of hydrogen peroxide-induced DNA damage by *Escherichia coli* requires SOS induction of RecA and RuvA proteins, *Mutat Res 459*, 187-194.

59. Paz, M. M., Kumar, G. S., Glover, M., Waring, M. J., and Tomasz, M. (2004) Mitomycin dimers: polyfunctional cross-linkers of DNA, *J Med Chem 47*, 3308-3319.

60. Pohlhaus, J. R., and Kreuzer, K. N. (2005) Norfloxacin-induced DNA gyrase cleavage complexes block *Escherichia coli* replication forks, causing double-stranded breaks in vivo, *Mol Microbiol 56*, 1416-1429.

61. Lewin, C. S., and Amyes, S. G. (1991) The role of the SOS response in bacteria exposed to zidovudine or trimethoprim, *J Med Microbiol 34*, 329-332.

62. Miller, C., Thomsen, L. E., Gaggero, C., Mosseri, R., Ingmer, H., and Cohen, S. N. (2004) SOS response induction by beta-lactams and bacterial defense against antibiotic lethality, *Science 305*, 1629-1631.

63. Aertsen, A., and Michiels, C. W. (2005) Mrr instigates the SOS response after high pressure stress in *Escherichia coli*, *Mol Microbiol 58*, 1381-1391.

64. Baharoglu, Z., and Mazel, D. (2011) Vibrio cholerae triggers SOS and mutagenesis in response to a wide range of antibiotics: a route towards multiresistance, *Antimicrob Agents Chemother 55*, 2438-2441.

65. Baharoglu, Z., Krin, E., and Mazel, D. (2013) RpoS plays a central role in the SOS induction by sub-lethal aminoglycoside concentrations in *Vibrio cholerae*, *PLoS Genet 9*, e1003421.

66. Zgur-Bertok, D. (2013) DNA damage repair and bacterial pathogens, *PLoS Pathog 9*, e1003711.

67. Baharoglu, Z., Babosan, A., and Mazel, D. (2014) Identification of genes involved in low aminoglycoside-induced SOS response in *Vibrio cholerae*: a role for transcription stalling and Mfd helicase, *Nucleic Acids Res 42*, 2366-2379.

68. Indiani, C., and O'Donnell, M. (2013) A proposal: Source of single strand DNA that elicits the SOS response, *Front Biosci (Landmark Ed) 18*, 312-323.

69. Yeeles, J. T., and Marians, K. J. (2011) The *Escherichia coli* replisome is inherently DNA damage tolerant, *Science 334*, 235-238.

70. Indiani, C., Patel, M., Goodman, M. F., and O'Donnell, M. E. (2013) RecA acts as a switch to regulate polymerase occupancy in a moving replication fork, *Proc Natl Acad Sci U S A 110*, 5410-5415.

71. Friedman, N., Vardi, S., Ronen, M., Alon, U., and Stavans, J. (2005) Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria, *PLoS Biol 3*, e238.

72. Kowalczykowski, S. C. (2015) An Overview of the Molecular Mechanisms of Recombinational DNA Repair, *Cold Spring Harb Perspect Biol 7*.

73. Tippin, B., Pham, P., and Goodman, M. F. (2004) Error-prone replication for better or worse, *Trends Microbiol 12*, 288-295.

74. Napolitano, R., Janel-Bintz, R., Wagner, J., and Fuchs, R. P. (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis, *Embo j 19*, 6259-6265.

75. McKenzie, G. J., Lee, P. L., Lombardo, M. J., Hastings, P. J., and Rosenberg, S. M. (2001) SOS mutator DNA polymerase IV functions in adaptive mutation and not adaptive amplification, *Mol Cell 7*, 571-579.

76. Tompkins, J. D., Nelson, J. L., Hazel, J. C., Leugers, S. L., Stumpf, J. D., and Foster, P. L. (2003) Error-prone polymerase, DNA polymerase IV, is responsible for transient hypermutation during adaptive mutation in *Escherichia coli*, *J Bacteriol 185*, 3469-3472.

77. Kuban, W., Jonczyk, P., Gawel, D., Malanowska, K., Schaaper, R. M., and Fijalkowska, I. J. (2004) Role of *Escherichia coli* DNA polymerase IV in in vivo replication fidelity, *J Bacteriol 186*, 4802-4807.

78. Kuban, W., Banach-Orlowska, M., Schaaper, R. M., Jonczyk, P., and Fijalkowska, I. J. (2006) Role of DNA polymerase IV in *Escherichia coli* SOS mutator activity, *J Bacteriol 188*, 7977-7980.

79. Hastings, P. J., Hersh, M. N., Thornton, P. C., Fonville, N. C., Slack, A., Frisch, R. L., Ray, M. P., Harris, R. S., Leal, S. M., and Rosenberg, S. M. (2010) Competition of *Escherichia coli* DNA polymerases I, II and III with DNA Pol IV in stressed cells, *PLoS One 5*, e10862.

80. Cirz, R. T., Chin, J. K., Andes, D. R., de Crecy-Lagard, V., Craig, W. A., and Romesberg, F. E. (2005) Inhibition of mutation and combating the evolution of antibiotic resistance, *PLoS Biol 3*, e176.

81. Cirz, R. T., O'Neill, B. M., Hammond, J. A., Head, S. R., and Romesberg, F. E. (2006) Defining the Pseudomonas aeruginosa SOS response and its role in the global response to the antibiotic ciprofloxacin, *J Bacteriol 188*, 7101-7110.

82. Qin, T. T., Kang, H. Q., Ma, P., Li, P. P., Huang, L. Y., and Gu, B. (2015) SOS response and its regulation on the fluoroquinolone resistance, *Ann Transl Med 3*, 358.

83. Dorr, T., Lewis, K., and Vulic, M. (2009) SOS response induces persistence to fluoroquinolones in *Escherichia coli*, *PLoS Genet 5*, e1000760.

84. Dorr, T., Vulic, M., and Lewis, K. (2010) Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*, *PLoS Biol 8*, e1000317.

85. Recacha, E., Machuca, J., Diaz de Alba, P., Ramos-Guelfo, M., Docobo-Perez, F., Rodriguez-Beltran, J., Blazquez, J., Pascual, A., and Rodriguez-Martinez, J. M. (2017) Quinolone Resistance Reversion by Targeting the SOS Response, *MBio 8*.

86. Maiques, E., Ubeda, C., Campoy, S., Salvador, N., Lasa, I., Novick, R. P., Barbe, J., and Penades, J. R. (2006) beta-lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in *Staphylococcus aureus*, *J Bacteriol 188*, 2726-2729.

87. Cirz, R. T., Jones, M. B., Gingles, N. A., Minogue, T. D., Jarrahi, B., Peterson, S. N., and Romesberg, F. E. (2007) Complete and SOS-mediated response of *Staphylococcus aureus* to the antibiotic ciprofloxacin, *J Bacteriol 189*, 531-539.

88. Alam, M. K., Alhhazmi, A., DeCoteau, J. F., Luo, Y., and Geyer, C. R. (2016) RecA Inhibitors Potentiate Antibiotic Activity and Block Evolution of Antibiotic Resistance, *Cell Chem Biol 23*, 381-391.

89. Gotoh, H., Kasaraneni, N., Devineni, N., Dallo, S. F., and Weitao, T. (2010) SOS involvement in stress-inducible biofilm formation, *Biofouling 26*, 603-611.

90. Chellappa, S. T., Maredia, R., Phipps, K., Haskins, W. E., and Weitao, T. (2013) Motility of *Pseudomonas aeruginosa* contributes to SOS-inducible biofilm formation, *Res Microbiol 164*, 1019-1027.

91. Walter, B. M., Cartman, S. T., Minton, N. P., Butala, M., and Rupnik, M. (2015) The SOS Response Master Regulator LexA Is Associated with Sporulation, Motility and Biofilm Formation in *Clostridium difficile*, *PLoS One 10*, e0144763.

92. Sperandio, V., Torres, A. G., Giron, J. A., and Kaper, J. B. (2001) Quorum sensing is a global regulatory mechanism in enterohemorrhagic *Escherichia coli* O157:H7, *J Bacteriol 183*, 5187-5197.

93. Little, J. W., Edmiston, S. H., Pacelli, L. Z., and Mount, D. W. (1980) Cleavage of the *Escherichia coli* LexA protein by the recA protease, *Proc Natl Acad Sci U S A 77*, 3225-3229.

94. Little, J. W. (1984) Autodigestion of lexA and phage lambda repressors, *Proc Natl Acad Sci U S A 81*, 1375-1379.

95. Slilaty, S. N., Rupley, J. A., and Little, J. W. (1986) Intramolecular cleavage of LexA and phage lambda repressors: dependence of kinetics on repressor concentration, pH, temperature, and solvent, *Biochemistry 25*, 6866-6875.

96. Little, J. W. (1991) Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease, *Biochimie 73*, 411-421.

97. Lin, L. L., and Little, J. W. (1989) Autodigestion and RecA-dependent cleavage of Ind- mutant LexA proteins, *J Mol Biol 210*, 439-452.

98. Hurstel, S., Granger-Schnarr, M., Daune, M., and Schnarr, M. (1986) In vitro binding of LexA repressor to DNA: evidence for the involvement of the amino-terminal domain, *Embo j 5*, 793-798.

99. Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., and Baker, T. A. (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals, *Mol Cell 11*, 671-683.

100. Fogh, R. H., Ottleben, G., Ruterjans, H., Schnarr, M., Boelens, R., and Kaptein, R. (1994) Solution structure of the LexA repressor DNA binding domain determined by 1H NMR spectroscopy, *Embo j 13*, 3936-3944.

101. Zhang, A. P., Pigli, Y. Z., and Rice, P. A. (2010) Structure of the LexA-DNA complex and implications for SOS box measurement, *Nature 466*, 883-886.

102. Knegtel, R. M., Fogh, R. H., Ottleben, G., Ruterjans, H., Dumoulin, P., Schnarr, M., Boelens, R., and Kaptein, R. (1995) A model for the LexA repressor DNA complex, *Proteins 21*, 226-236.

103. Kim, B., and Little, J. W. (1993) LexA and lambda Cl repressors as enzymes: specific cleavage in an intermolecular reaction, *Cell 73*, 1165-1173.

104. Shepley, D. P., and Little, J. W. (1996) Mutant LexA proteins with specific defects in autodigestion, *Proc Natl Acad Sci U S A 93*, 11528-11533.

105. Giese, K. C., Michalowski, C. B., and Little, J. W. (2008) RecA-dependent cleavage of LexA dimers, *J Mol Biol 377*, 148-161.

106. Kovacic, L., Paulic, N., Leonardi, A., Hodnik, V., Anderluh, G., Podlesek, Z., Zgur-Bertok, D., Krizaj, I., and Butala, M. (2013) Structural insight into LexA-RecA* interaction, *Nucleic Acids Res 41*, 9901-9910.

107. Slilaty, S. N., and Little, J. W. (1987) Lysine-156 and serine-119 are required for LexA repressor cleavage: a possible mechanism, *Proc Natl Acad Sci U S A 84*, 3987-3991.

108. Brock, T. D. (1973) Lower pH limit for the existence of blue-green algae: evolutionary and ecological implications, *Science 179*, 480-483.

109. Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent, T., Goodwin, L., Nolan, M., Davenport, K. W., Han, C. S., Rubin, E. M., Eisen, J. A., Woyke, T., Gugger, M., and Kerfeld, C. A. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing, *Proc Natl Acad Sci U S A 110*, 1053-1058.

110. Stanier, R. Y., and Cohen-Bazire, G. (1977) Phototrophic prokaryotes: the cyanobacteria, *Annu Rev Microbiol 31*, 225-274.

111. Stewart, A. C., and Bendall, D. S. (1981) Properties of oxygen-evolving photosystem-II particles from Phormidium laminosum, a thermophilic blue-green alga, *Biochem J 194*, 877-887.

112. Walsby, A. E. (1986) Prochlorophytes: Origins of chloroplasts, *Nature 320*, 212.

113. Los, D. A., and Murata, N. (1999) Responses to cold shock in cyanobacteria, *J Mol Microbiol Biotechnol 1*, 221-230.

114. William Schopf, J. (2011) The paleobiological record of photosynthesis, *Photosynth Res 107*, 87-101.

115. Hamilton, T. L., Bryant, D. A., and Macalady, J. L. (2016) The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans, *Environ Microbiol 18*, 325-340.

116. Schopf, J. W. (1975) Precambrian Paleobiology: Problems and Perspectives, *Annual Review of Earth and Planetary Sciences 3*, 213-249.

117. Zehr, J. P. (2011) Nitrogen fixation by marine cyanobacteria, *Trends Microbiol 19*, 162-173.

118. Jansson, C., and Northen, T. (2010) Calcifying cyanobacteria--the potential of biomineralization for carbon capture and storage, *Curr Opin Biotechnol 21*, 365-371.

119. Wynn-Williams, D. D. (1990) Ecological aspects of Antarctic Microbiology, *Advances in Microbial Ecology 11*, 71-146.

120. Stewart, W. D. (1980) Some aspects of structure and function in N2-fixing cyanobacteria, *Annu Rev Microbiol 34*, 497-536.

121. Gallon, J. R., LaRue, T. A., and Kurz, W. G. (1974) Photosynthesis and nitrogenase activity in the blue-green alga *Gloeocapsa*, *Can J Microbiol 20*, 1633-1637.

122. Stal, L. J. a. K., W.E. (1987) Temporal separation of nitrogen fixation and photosynthesis in the filamentous, non-heterocystous cyanobacterium *Oscillatoria* sp, *Arch. Microbiol. 149*, 76-80.

123. Wolk, C. P. (1996) Heterocyst formation, *Annu Rev Genet 30*, 59-78.

124. Gugger, M. F., and Hoffmann, L. (2004) Polyphyly of true branching cyanobacteria (Stigonematales), *Int J Syst Evol Microbiol 54*, 349-357.

125. Haselkorn, R. (1978) Heterocysts, *Annual Review of Plant Physiology 29*, 319-344.

126. Zhang, C. C., Laurent, S., Sakr, S., Peng, L., and Bedu, S. (2006) Heterocyst differentiation and pattern formation in cyanobacteria: a chorus of signals, *Mol Microbiol 59*, 367-375.

127. Domain, F., Houot, L., Chauvat, F., and Cassier-Chauvat, C. (2004) Function and regulation of the cyanobacterial genes lexA, recA and ruvB: LexA is critical to the survival of cells facing inorganic carbon starvation, *Mol Microbiol 53*, 65-80.

128. Patterson-Fortin, L. M., Colvin, K. R., and Owttrim, G. W. (2006) A LexA-related protein regulates redox-sensitive expression of the cyanobacterial RNA helicase, crhR, *Nucleic Acids Res 34*, 3446-3454.

129. Walker, G. C. (1996) The SOS Response of *Escherichia coli* in *Escherichia coli* and salmonella : Cellular and molecular biology, *eds Neidhardt F.C., Curtiss III R.,*

Ingraham J.L., Lin E.C.C., Low K.B., Magasanik B., Reznikoff W.S., Riley M.,Schaechter M., Umbarger H.E.(ASM Press, Washington, D.C.), pp 1400–1416.

130. Mazon, G., Lucena, J. M., Campoy, S., Fernandez de Henestrosa, A. R., Candau, P., and Barbe, J. (2004) LexA-binding sequences in Gram-positive and cyanobacteria are closely related, *Mol Genet Genomics 271*, 40-49.

131. Oliveira, P., and Lindblad, P. (2005) LexA, a transcription regulator binding in the promoter region of the bidirectional hydrogenase in the cyanobacterium *Synechocystis* sp. PCC 6803, *FEMS Microbiol Lett 251*, 59-66.

132. Sjoholm, J., Oliveira, P., and Lindblad, P. (2007) Transcription and regulation of the bidirectional hydrogenase in the cyanobacterium Nostoc sp. strain PCC 7120, *Appl Environ Microbiol 73*, 5435-5446.

133. Gutekunst, K., Phunpruch, S., Schwarz, C., Schuchardt, S., Schulz-Friedrich, R., and Appel, J. (2005) LexA regulates the bidirectional hydrogenase in the cyanobacterium *Synechocystis* sp. PCC 6803 as a transcription activator, *Mol Microbiol 58*, 810-823.

134. Oliveira, P., and Lindblad, P. (2008) An AbrB-Like protein regulates the expression of the bidirectional hydrogenase in *Synechocystis* sp. strain PCC 6803, *J Bacteriol 190*, 1011-1019.

135. Campoy, S., Salvador, N., Cortes, P., Erill, I., and Barbe, J. (2005) Expression of canonical SOS genes is not under LexA repression in *Bdellovibrio bacteriovorus*, *J Bacteriol 187*, 5367-5375.

136. Ferreira, D., Leitao, E., Sjoholm, J., Oliveira, P., Lindblad, P., Moradas-Ferreira, P., and Tamagnini, P. (2007) Transcription and regulation of the hydrogenase(s) accessory genes, hypFCDEAB, in the cyanobacterium *Lyngbya majuscula* CCAP 1446/4, *Arch Microbiol 188*, 609-617.

137. Leitao, E., Pereira, S., Bondoso, J., Ferreira, D., Pinto, F., Moradas-Ferreira, P., and Tamagnini, P. (2006) Genes involved in the maturation of hydrogenase(s) in the nonheterocystous cyanobacterium *Lyngbya majuscula* CCAP 1446/4, *International Journal of Hydrogen Energy 31*, 1469-1477.

138. Patterson-Fortin, L. M., and Owttrim, G. W. (2008) A *Synechocystis* LexA-orthologue binds direct repeats in target genes, *FEBS Lett 582*, 2424-2430.

139. Kizawa, A., Kawahara, A., Takashima, K., Takimura, Y., Nishiyama, Y., and Hihara, Y. (2017) The LexA transcription factor regulates fatty acid biosynthetic genes in the cyanobacterium *Synechocystis* sp. PCC 6803, *Plant J 92*, 189-198.

140. Oliveira, P., and Lindblad, P. (2011) Novel insights into the regulation of LexA in the cyanobacterium *Synechocystis* sp. Strain PCC 6803, *J Bacteriol 193*, 3804-3814.

141. Tetu, S. G., Johnson, D. A., Varkey, D., Phillippy, K., Stuart, R. K., Dupont, C. L., Hassan, K. A., Palenik, B., and Paulsen, I. T. (2013) Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species, *Front Microbiol 4*, 232.

142. Kaneko, T., Nakamura, Y., Wolk, C. P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A., Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M., and Tabata, S. (2001) Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120, *DNA Res 8*, 205-213; 227-253.

143. Wolk, C. P., Vonshak, A., Kehoe, P., and Elhai, J. (1984) Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria, *Proc Natl Acad Sci U S A 81*, 1561-1565.

144. Golden, J. W., Carrasco, C. D., Mulligan, M. E., Schneider, G. J., and Haselkorn, R. (1988) Deletion of a 55-kilobase-pair DNA element from the chromosome during heterocyst differentiation of *Anabaena* sp. strain PCC 7120, *J Bacteriol 170*, 5034-5041.

145. Cai, Y. P., and Wolk, C. P. (1990) Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences, *J Bacteriol 172*, 3138-3145.

146. Chaurasia, A. K., Parasnis, A., and Apte, S. K. (2008) An integrative expression vector for strain improvement and environmental applications of the nitrogen fixing cyanobacterium, *Anabaena* sp. strain PCC7120, *J Microbiol Methods 73*, 133-141.

147. Raghavan, P. S., Rajaram, H., and Apte, S. K. (2011) Nitrogen status dependent oxidative stress tolerance conferred by overexpression of MnSOD and FeSOD proteins in *Anabaena* sp. strain PCC7120, *Plant Mol Biol 77*, 407-417.

148. Singh, H., Fernandes, T., and Apte, S. K. (2010) Unusual radioresistance of nitrogen-fixing cultures of *Anabaena* strains, *J Biosci 35*, 427-434.

149. Singh, H., Anurag, K., and Apte, S. K. (2013) High radiation and desiccation tolerance of nitrogen-fixing cultures of the cyanobacterium *Anabaena* sp. strain PCC 7120 emanates from genome/proteome repair capabilities, *Photosynth Res*.

150. Rajaram, H., and Apte, S. K. (2010) Differential regulation of groESL operon expression in response to heat and light in *Anabaena*, *Arch Microbiol 192*, 729-738.

151. Castenholz, R. W. (1988) [3] Culturing methods for cyanobacteria, In *Methods in Enzymology*, pp 68-93, Academic Press.

152. Mackinney, G. (1941) Absorption of light by chlorophyll solutions., *J Biol Chem. 140*, 315-322.

153. Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E., and Wolk, C. P. (1997) Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120, *J Bacteriol 179*, 1998-2005.

154. Yoon, H. S., and Golden, J. W. (1998) Heterocyst pattern formation controlled by a diffusible peptide, *Science 282*, 935-938.

155. Kumar, D., Yusuf, M. A., Singh, P., Sardar, M., and Sarin, N. B. (2013) Modulation of antioxidant machinery in alpha-tocopherol-enriched transgenic *Brassica juncea* plants tolerant to abiotic stress conditions, *Protoplasma 250*, 1079-1089.

156. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680-685.

157. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) Protein measurement with the Folin phenol reagent, *J Biol Chem 193*, 265-275.

158. Chung, M. H., Chen, M. K., and Pan, S. M. (2000) Floral spray transformation can efficiently generate *Arabidopsis* transgenic plants, *Transgenic Res 9*, 471-476.

159. Panda, B., Basu, B., Rajaram, H., and Kumar Apte, S. (2014) Methyl viologen responsive proteome dynamics of *Anabaena* sp. strain PCC7120, *Proteomics 14*, 1895-1904.

160. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis, *Nat Methods 9*, 671-675.

161. Louis-Jeune, C., Andrade-Navarro, M. A., and Perez-Iratxeta, C. (2012) Prediction of protein secondary structure from circular dichroism using theoretically derived spectra, *Proteins 80*, 374-381.

162. Janion, C. (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*, *Int J Biol Sci 4*, 338-344.

163. Little, J. W. (1983) The SOS regulatory system: control of its state by the level of RecA protease, *J Mol Biol 167*, 791-808.

164. Neher, S. B., Flynn, J. M., Sauer, R. T., and Baker, T. A. (2003) Latent ClpX-recognition signals ensure LexA destruction after DNA damage, *Genes Dev 17*, 1084-1089.

165. Little, J. W., and Hill, S. A. (1985) Deletions within a hinge region of a specific DNA-binding protein, *Proc Natl Acad Sci U S A 82*, 2301-2305.

166. Blot, N., Mella-Flores, D., Six, C., Le Corguille, G., Boutte, C., Peyrat, A., Monnier, A., Ratin, M., Gourvil, P., Campbell, D. A., and Garczarek, L. (2011) Light history influences the response of the marine cyanobacterium *Synechococcus* sp. WH7803 to oxidative stress, *Plant Physiol 156*, 1934-1954.

167. Kolowrat, C., Partensky, F., Mella-Flores, D., Le Corguillé, G., Boutte, C., Blot, N., Ratin, M., Ferréol, M., Lecomte, X., Gourvil, P., Lennon, J.-F., Kehoe, D. M., and Garczarek, L. (2010) Ultraviolet stress delays chromosome replication in light/dark synchronized cells of the marine cyanobacterium *Prochlorococcus marinus* PCC9511, *BMC Microbiology 10*, 204.

168. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures, *Nucleic Acids Res 45*, D200-d203.

169. Li, S., Xu, M., and Su, Z. (2010) Computational analysis of LexA regulons in Cyanobacteria, *BMC Genomics 11*, 527.

170. Kizawa, A., Kawahara, A., Takimura, Y., Nishiyama, Y., and Hihara, Y. (2016) RNA-seq Profiling Reveals Novel Target Genes of LexA in the Cyanobacterium *Synechocystis* sp. PCC 6803, *Front Microbiol 7*, 193.

171. Kirti, A., Kumar, A., and Rajaram, H. (2017) Differential regulation of ssb genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120(1), *J Phycol 53*, 322-332.

172. Mount, D. W., Low, K. B., and Edmiston, S. J. (1972) Dominant mutations (lex) in *Escherichia coli* K-12 which affect radiation sensitivity and frequency of ultraviolet lght-induced mutations, *J Bacteriol 112*, 886-893.

173. Lin, L. L., and Little, J. W. (1988) Isolation and characterization of noncleavable (Ind-) mutants of the LexA repressor of *Escherichia coli* K-12, *J Bacteriol 170*, 2163-2173.

174. Lazar, D., Brokes, M., Naus, J., and Dvorak, L. (1998) Mathematical modelling of 3-(3',4'-dichlorophenyl)-1,1-dimenthylurea action in plant leaves, *J Theor Biol 191*, 79-86.

175. Mittler, R. (2002) Oxidative stress, antioxidants and stress tolerance, *Trends Plant Sci 7*, 405-410.

176. Panda, B., Basu, B., Acharya, C., Rajaram, H., and Apte, S. K. (2017) Proteomic analysis reveals contrasting stress response to uranium in two nitrogen-fixing *Anabaena* strains, differentially tolerant to uranium, *Aquat Toxicol 182*, 205-213.

177. Panda, B., Basu, B., Rajaram, H., and Apte, S. K. (2015) Comparative proteomics of oxidative stress response in three cyanobacterial strains native to Indian paddy fields, *J Proteomics 127*, 152-160.

178. Jong, A. Y., and Campbell, J. L. (1984) Characterization of Saccharomyces cerevisiae thymidylate kinase, the CDC8 gene product. General properties, kinetic analysis, and subcellular localization, *J Biol Chem 259*, 14394-14398.

179. Belknap, W. R., and Haselkorn, R. (1987) Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena, Embo j 6*, 871-884.

180. Pandey, S., Kirti, A., Kumar, A., and Rajaram, H. (2018) The SbcC and SbcD homologs of the cyanobacterium *Anabaena* sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair, *Funct Integr Genomics 18*, 357-367.

181. Jian, H., Xiong, L., He, Y., and Xiao, X. (2015) The regulatory function of LexA is temperature-dependent in the deep-sea bacterium *Shewanella piezotolerans* WP3, *Front Microbiol 6*, 627.

182. Almaula, N., Lu, Q., Delgado, J., Belkin, S., and Inouye, M. (1995) Nucleoside diphosphate kinase from *Escherichia coli*, *J Bacteriol 177*, 2524-2529.

183. Cui, Q., Shin, W. S., Luo, Y., Tian, J., Cui, H., and Yin, D. (2013) Thymidylate kinase: an old topic brings new perspectives, *Curr Med Chem 20*, 1286-1305.

184. Kirti, A., Rajaram, H., and Apte, S. K. (2013) Characterization of two naturally truncated, Ssb-like proteins from the nitrogen-fixing cyanobacterium, *Anabaena* sp. PCC7120, *Photosynth Res 118*, 147-154.

185. Raghavan, P. S., Gupta, G. D., Rajaram, H., and Kumar, V. (2016) Characterization of a DUF820 family protein Alr3200 of the cyanobacterium *Anabaena* sp. strain PCC7120, *J Biosci 41*, 589-600.

186. Liu, L. N., Chen, X. L., Zhang, Y. Z., and Zhou, B. C. (2005) Characterization, structure and function of linker polypeptides in phycobilisomes of cyanobacteria and red algae: an overview, *Biochim Biophys Acta 1708*, 133-142.

187. MacColl, R. (1998) Cyanobacterial phycobilisomes, *J Struct Biol 124*, 311-334.

188. Kehoe, D. M., and Gutu, A. (2006) Responding to color: the regulation of complementary chromatic adaptation, *Annu Rev Plant Biol 57*, 127-150.

189. Shinozaki, K. a. S., M. (1985) Genes for the large and small subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase constitute a single operon in a cyanobacterium *Anacystis nidulans* 6301, *Mol. Gen. Genet. 200*, 27-32

190. Nierzwicki-Bauer, S. A., Curtis, S. E., and Haselkorn, R. (1984) Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* 7120, *Proc Natl Acad Sci U S A 81*, 5961-5965.

191. Campbell, W. J., and Ogren, W. L. (1990) Electron Transport through photosystem I Stimulates Light Activation of Ribulose Bisphosphate Carboxylase/Oxygenase (Rubisco) by Rubisco Activase, *Plant Physiol 94*, 479-484.

192. Pandey, S., Shrivastava, A. K., Rai, R., and Rai, L. C. (2013) Molecular characterization of Alr1105 a novel arsenate reductase of the diazotrophic cyanobacterium *Anabaena* sp. PCC7120 and decoding its role in abiotic stress management in *Escherichia coli*, *Plant Mol Biol 83*, 417-432.

193. Pandey, S., Shrivastava, A. K., Singh, V. K., Rai, R., Singh, P. K., Rai, S., and Rai, L. C. (2013) A new arsenate reductase involved in arsenic detoxification in *Anabaena* sp. PCC7120, *Funct Integr Genomics 13*, 43-55.

194. Mishra, Y., Chaurasia, N., and Rai, L. C. (2009) AhpC (alkyl hydroperoxide reductase) from *Anabaena* sp. PCC 7120 protects Escherichia coli from multiple abiotic stresses, *Biochem Biophys Res Commun 381*, 606-611.

195. Narayan, O. P., Kumari, N., and Rai, L. C. (2010) Heterologous expression of *Anabaena* PCC 7120 all3940 (a Dps family gene) protects *Escherichia coli* from nutrient limitation and abiotic stresses, *Biochem Biophys Res Commun 394*, 163-169.

196. Morby, A. P., Turner, J. S., Huckle, J. W., and Robinson, N. J. (1993) SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene smtA: identification of a Zn inhibited DNA-protein complex, *Nucleic Acids Res 21*, 921-925.

197. Singh, S. C., Sinha, R.P.and Hader, Donat-P. . (2002) Role of Lipids and Fatty Acids in Stress Tolerance in Cyanobacteria, *Acta Protozool. 41*, 297 – 308.

198. Banerjee, M., Raghavan, P. S., Ballal, A., Rajaram, H., and Apte, S. K. (2013) Oxidative stress management in the filamentous, heterocystous, diazotrophic cyanobacterium, *Anabaena* PCC7120, *Photosynthesis Research 118*, 59-70.

List of Publications

1. **Arvind Kumar**, Anurag Kirti and Hema Rajaram (2015) LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits *in vitro* pH-dependent and RecA-independent autoproteolytic activity. **The Int. J Biochem. Cell Biol.** 59: 84-93
2. Anurag Kirti, **Arvind Kumar** and Hema Rajaram (2017) Differential regulation of *ssb* genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120. **J. Phycol.** 53: 322-332
3. Sarita Pandey, Anurag Kirti, **Arvind Kumar** and Hema Rajaram (2018) The SbcC and SbcD homologs of the cyanobacterium Anabaena sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair. **Funct Integr Genomics.** 18:357-367.
4. **Arvind Kumar**, Anurag Kirti and Hema Rajaram, (2018) Regulation of multiple abiotic stress tolerance by LexA in the cyanobacterium *Anabaena* sp. strain PCC7120. **BBA Gene Regulatory Mechanisms.** Doi: 10.1016/j.bbagrm.2018.07.007

Reprints of Published articles

The International Journal of Biochemistry & Cell Biology 59 (2015) 84–93

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits *in vitro* pH-dependent and RecA-independent autoproteolytic activity

Arvind Kumar, Anurag Kirti, Hema Rajaram*

Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India

ARTICLE INFO

Article history:
Received 10 July 2014
Received in revised form
27 November 2014
Accepted 8 December 2014
Available online 15 December 2014

Keywords:
Anabaena
Alr4908
Autoproteolytic cleavage
LexA
RecA

ABSTRACT

The LexA protein of the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120 exhibits a RecA-independent and alkaline pH-dependent autoproteolytic cleavage. The autoproteolytic cleavage of *Anabaena* LexA occurs at pH 8.5 and above, stimulated by the addition of Ca^{2+} and in the temperature range of 30–57 °C. Mutational analysis of *Anabaena* LexA protein indicated that the cleavage occurred at the peptide bond between Ala-84 and Gly-85, and optimal cleavage required the presence of Ser-118 and Lys-159, as also observed for LexA protein of *Escherichia coli*. Cleavage of *Anabaena* LexA was affected upon deletion of three amino acids, ^{86}GLI. These three amino acids are unique to all cyanobacterial LexA proteins predicted to be cleavable. The absence of RecA-dependent cleavage at physiological pH, which has not been reported for other bacterial LexA proteins, is possibly due to the absence of RecA interacting sites on *Anabaena* LexA protein, corresponding to the residues identified in *E. coli* LexA, and low cellular levels of RecA in *Anabaena*. Exposure to SOS-response inducing stresses, such as UV-B and mitomycin C neither affected the expression of LexA in *Anabaena* nor induced cleavage of LexA in either *Anabaena* 7120 or *E. coli* overexpressing *Anabaena* LexA protein. Though the LexA may be acting as a repressor by binding to the LexA box in the vicinity of the promoter region of specific gene, their derepression may not be *via* proteolytic cleavage during SOS-inducing stresses, unless the stress induces increase in cytoplasmic pH. This could account for the regulation of several carbon metabolism genes rather than DNA-repair genes under the regulation of LexA in cyanobacteria especially during high light induced oxidative stress.

1. Introduction

LexA is one of the key regulators of DNA repair genes in bacteria and is known to regulate more than 30 genes involved in SOS response in *Escherichia coli* (Fernandez de Henestrosa et al., 2000; Janion, 2008). In addition to induction of SOS response by direct DNA damage, it was also observed upon infection by λ phage in *E. coli* involving the host proteases (Schlacher and Goodman, 2007). Under normal growth conditions, LexA represses the SOS genes by binding to the LexA (or SOS) Box, an imperfect 16 bp palindromic sequence $CTGTN_8ACAG$ (Walker, 1984), in the vicinity of their promoter region. Upon DNA damage, RecA interacts with the unwound single stranded DNA, gets activated and induces auto-cleavage of the Ala^{84}–Gly^{85} peptide bond in LexA. This inhibits dimerisation of LexA and also prevents its binding to DNA, thereby derepressing the SOS genes (Janion, 2008). LexA has two domains, the N-terminal domain (NTD) possessing the DNA binding site, and the C-terminal domain (CTD) comprising of both the dimerisation site and the active site for cleavage. The auto-catalytic cleavage of LexA is controlled by several residues in the C-terminal domain, namely L89, Q92, S119, E152 and K156. It can also be achieved at alkaline pH i.e. >10.5 in the absence of activated RecA (Giese et al., 2008; Little, 1984; Luo et al., 2001). While the autocleavge of LexA separates the two domains, the NTD is still capable of binding to the SOS box and retains some LexA repressor function (Little and Hill, 1985). Further proteolytic degradation of the NTD by ClpXP proteolytic complex (Neher et al., 2003) and the CTD by Lon protease (Little and Gellert, 1983) ensures complete derepression of the LexA regulon in *E. coli*.

Though LexA is known to be generally involved in regulation of DNA repair in bacteria, this has been challenged in several organisms including the unicellular cyanobacterium, *Synechocystis* sp. strain PCC6803, wherein it is involved in modulating expression of several genes involved in carbon metabolism (Domain et al., 2004). Additionally, the *Synechocystis* PCC6803 *lexA* gene is not induced by DNA damage (Patterson-Fortin et al., 2006). The corresponding LexA protein lacks two of the three residues required for proteolytic activity of LexA (Li et al., 2010). Thus, it possibly could not function to repress expression of DNA repair genes as has been observed for

* Corresponding author. Tel.: +91 22 25593213.
E-mail address: hemaraj@barc.gov.in (H. Rajaram).

http://dx.doi.org/10.1016/j.biocel.2014.12.003
1357-2725/

other prokaryotes (Patterson-Fortin et al., 2006). In the marine unicellular cyanobacterial strains, *Prochlorococcus marinus* PCC9511 and *Synechococcus* sp. WH7803, several SOS-response genes are up-regulated during high light including *lexA* and *recA*, which are co-expressed (Blot et al., 2011; Kolowrat et al., 2010), contradicting the well established succession of LexA and RecA expression in other bacteria. Also, the expression pattern of LexA is influenced by high light intensity, but not by UV-B radiation in *P. marinus* (Kolowrat et al., 2010). Whole genome transcriptional profiling of two marine *Synechococcus* strains in response to DNA damaging agents also reveal a possible role for LexA in regulating *recA*, *lexA* and *umuC* genes (Tetu et al., 2013). Computational analysis of LexA in cyanobacteria indicated that of the 33 sequenced cyanobacterial genomes, six do not posses LexA and among those which posses LexA, *Synechocystis* PCC6803 and *Gloeobacter violaceus* PCC7421 are the only two to code for a non-cleavable LexA, while the others including *Anabaena* sp., code for a theoretically cleavable form of LexA (Li et al., 2010).

In the heterocystous filamentous, nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120, hereafter referred to as *Anabaena* 7120, LexA is coded by *alr4908* (http://genome.microbedb.jp/cyanobase/*Anabaena*). *In vitro* studies have revealed the binding of purified *Anabaena* LexA to putative LexA-Box upstream of both *recA* and *lexA* genes (Mazón et al., 2004), but its interaction with RecA in terms of regulating its role as a repressor has not been dealt with. This paper addresses the autoproteolytic cleavage of *Anabaena* LexA as a function of RecA and pH *in vitro* and by DNA-damage inducing stresses *in vivo*. The data revealed that *Anabaena* LexA is autoproteolytically cleaved at Ala^{84}–Gly^{85} peptide bond, but only at alkaline pH (≥8.5), and not at physiological pH even in the presence of activated RecA. The cleavage of *Anabaena* LexA was not observed upon exposure of either *E. coli* overexpressing *Anabaena* LexA or *Anabaena* 7120 to DNA-damage inducing stresses. The low levels of RecA detected in *Anabaena* and the absence of RecA-binding sites in *Anabaena* LexA could be responsible for the absence of RecA-dependent cleavage of LexA in *Anabaena*.

2. Materials and methods

2.1. Organisms and growth conditions

E. coli cells were grown in LB medium at 37 °C under shaking condition (150 rpm). Antibiotics [34 μg chloramphenicol mL^{-1} (Cm_{34}), 50 μg kanamycin mL^{-1} (Kan_{50}), or 100 μg carbenicillin mL^{-1} (Cb_{100})] were added to the LB medium, as and when required. *Anabaena* 7120 cells were grown in BG-11 liquid medium, pH 7.0 (Castenholz, 1988) with17 mM $NaNO_3$ under stationary conditions and continuous illumination (30 μE m^{-2} s^{-1}) at 27 ± 2 °C. MitomycinC (mitC) stress was applied by exposing 3-day-old *Anabaena* 7120 cultures (10 μg chl *a* density mL^{-1}) to 1–5 μg mitC mL^{-1} for 30 min or logarithmic phase cultures of *E. coli* to 2 μg mitC mL^{-1} for 1 h. UV stress was applied by exposing 3-day-old *Anabaena* cultures (10 μg chl *a* mL^{-1}) to 0.75–3 kJ m^{-2} and *E. coli* cells to 50 J m^{-2} of UV-B (dose rate 7 J m^{-2} s^{-1}).

2.2. Generation of plasmid constructs

The wild type *lexA* and its variants, and *recA* were PCR amplified from *Anabaena* 7120 genomic DNA using specific primers (Table 1), 1 μM dNTPs and 1 U Taq DNA polymerase (BRIT, Mumbai) in Taq buffer. In case of the complete *lexA* or *recA* genes or deletion of either C-terminal sequence (NTD) or N-terminal sequence (CTD) of *lexA*, single step PCR was carried out. For point mutations or deletion from internal gene sequence, a two step PCR was carried out.

Table 1
Primers used in this study.

Primers	Nucleotide sequence
lexA-OE-Fwd	5′-GCCGCCATATGGAACGCCTAACAGAA-3′
lexA-OE-Rev	5′-GAAGGATCCTCACATATAACCGCGCCA-3′
recA-OE-Fwd	5′-GGCCATATGGCTATCAATACCGAT-3′
recA-OE-Rev	5′-GAAGGATCCTTATTCCTCTTCTAC-3′
lexA-A84S-Fwd	5′-GGAGCGATCGCCtCAGGTGGTTTAATAGAA-3′
lexA-A84S-Rev	5′-TTCTATTAAACCACCTGaGGCGATCGCTCC-3′
lexA-G85S-Fwd	5′-GGAGCGATCGCCGCAaGTGGTTTAATAGAA-3′
lexA-G85S-Rev	5′-TTCTATTAAACCActTGCGGCGATCGCTCC-3′
lexA-S118A-Fwd	5′-GTAACTGGTGACgcCATGATTGAAGA-3′
lexA-S118A-Rev	5′-TCTTCAATCATGgcGTCACCAGTTAC-3′
lexA-K159R-Fwd	5′-GGTAATACATTAAGACGTTTTTATCGA-3′
lexA-K159R-Rev	5′-TCGATAAAAACGTcTTAATGTATTACC-3′
lexA-CTD-Fwd	5′-GGCCCATATGGGTAAAGCCCGAACAATTCGAGTTTT-3′
lexA-NTD-Rev	5′-GGCCGGATCCTTAACCTTCAGTCCATTCAATG-3′
lexA-E96Q-Fwd	5′-TCACTGATGCTGTCcAGCATATCGACTTT-3′
lexA-E96Q-Rev	5′-AAAGTCGATATGCTgGACAGCATCAGTGA-3′
lexAΔGLI-Fwd	5′-GCGATCGCCGCAGGTG(Δ)AACCATTCACTGATG-3′
lexAΔGLI-Rev	5′-CATCAGTGAATGGTT(Δ)CACCTGCGGCGATCGC-3′
lexA-Pro-Fwd	5′-GGCCGGTACCAGCCCATCGTGGT-3′
lexA-Pro-Rev	5′-CCGGGAGCTCAAAATAATAGGGAGGATA-3′
recA-Pro-Fwd	5′-GGCCGGTACCGTATAGCTAGGTTA-3′
recA-Pro-Rev	5′-GGCCGAGCTCTTCCTGCCTCTAAAAT-3′
alr0088-Pro-Fwd	5′-GCGCCGGAGCTCGGTATTTGCTGTACCGAG-3′
alr0088-Pro-Rev	5′-GATCGGTACCCTGCCTTATCCTTTAGTACATAAGTACT-3′

The underlined sequence in the Fwd and Rev primer indicate *Nde*I and *Bam*HI restriction sites respectively. The change in nucleotide sequence for generating LexA mutants is written in small caps. (Δ) indicates deletion of 9 bases nucleotides corresponding to GLI.

Two amplicons (supplementary Table S1) were generated for each mutant and a nested PCR was carried out to get the complete gene with the desired mutation. The different amplicons (supplementary Table S1) were digested with *Nde*I and *Bam*HI restriction endonucleases, individually ligated to the expression vector pET16b at identical restriction sites and the ligation mixture transformed into competent *E. coli* DH5α cells. The different plasmid constructs generated are listed in Table 2, and were maintained in *E. coli* DH5α cells (supplementary Table S2). The DNA inserts of all plasmids were sequenced in both the strands to confirm that no error has arisen due to PCR amplification as well as to confirm introduced mutations. The sequence of the insert of pET*lexA* was submitted to GenBank (GenBank Accession no. KF269537).

2.3. Overexpression and purification of wild type and LexA mutant proteins

The different plasmid constructs corresponding to wild type or mutant *lexA* or wild type *recA* in pET16b were transformed into *E. coli* BL21(plysS) (supplementary Table 2), and overexpression of the corresponding proteins was achieved by induction with 1 mM IPTG at 37 °C for 3 h. The proteins were purified using Ni-NTA affinity chromatography with a step-wise imidazole gradient (20 mM to 1 M) in 100 mM Tris–HCl, pH 8.0, 200 mM NaCl buffer. After initial washes with 20, 50, 100 and 150 mM imidazole, pure proteins were eluted with 250 mM and 1 M imidazole. All fractions were visualised after electrophoretic separation on 12% or 15% SDS-polyacrylamide gel (SDS-PAGE) followed by staining with Coomassie Brilliant Blue (CBB) G-250. Fractions containing pure proteins were pooled, imidazole removed by dialysis and proteins stored in buffer containing 100 mM Tris–HCl, pH 7.5, 200 mM NaCl, 1 mM EDTA and 5% glycerol. The proteins, LexA-NTD and Lex-ΔGLI upon overexpression in *E. coli*, were insoluble and detected in inclusion bodies. These were purified under denaturing conditions with 8 M urea using Ni-NTA affinity chromatography. The corresponding cell pellets were lysed in lysis buffer (8 M urea, 10 mM Tris–HCl, 100 mM NaH_2PO_4, pH 8.0) and eluted in elution buffer (8 M urea,

Table 2
Plasmids used in this study[a].

Vector	Characteristics	Source
pET16b	Expression vector with N-terminal His-Tag	Novagen
pET*lexA*	0.6 kb, *lexA* gene cloned in pET16b at *Nde*I–*Bam*HI restriction sites	This study
pETrecA	1.07 kb, recA gene cloned in pET16b at *Nde*I–*Bam*HI restriction sites	This study
pET*lexA*-NTD	0.18 kb, partial *lexA* gene corresponding to first 60 aa[b] cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*-CTD	0.45 kb, partial *lexA* gene corresponding to *lexA* from the 61st to 201st aa cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*ΔGLI	0.6 kb *lexA* gene, with deletion of three amino acids (G86LI88) cloned in pET16b at *Nde*I–*Xho*I *restriction sites*	This study
pET*lexA*-A84S	0.6 kb *lexA* gene (A to S substitution at 84th aa) cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*-G85S	0.6 kb *lexA* gene (G to S substitution at 85th aa) cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*-E96Q	0.6 kb *lexA* gene (E to Q substitution at 96th aa) cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*-S118A	0.6 kb *lexA* gene (S to A substitution at 118th aa) cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study
pET*lexA*-K159R	0.6 kb *lexA* gene (K to R substitution at 159th aa) cloned in pET16b at *Nde*I–*Xho*I restriction sites	This study

[a] *lexA* gene in each case has been amplified from genomic DNA of *Anabaena* 7120.
[b] Amino acids.

10 mM Tris–HCl, 100 mM NaH_2PO_4, pH 4.5). All fractions were visualised as described above and the fractions having pure proteins were pooled and allowed to renature by dialysis in a step-wise manner decreasing the urea concentration in the dialysis buffer by 2 M every 2 h. Prior to analysis the protein content was estimated with Lowry reagent (Lowry et al., 1951).

2.4. Size determination

Molecular mass determination of native purified protein was carried out by gel filtration chromatography using Superdex HR200 column. The column was equilibrated with Tris–NaCl buffer and standard graph obtained using the following proteins: thioglobulin (669 kDa), ferritin (440 kDa), bovine serum albumin (67 kDa), albumin (44.3 kDa), and carbonic anhydrase (29 kDa). Molecular mass of the *Anabaena* LexA was calculated on the basis of the elution volume using the equation of standard graph.

2.5. LexA autoproteolytic cleavage assay

For analysing the autoproteolytic cleavage of wild type or mutant LexA proteins at different pH (6.5–9.5) or temperatures (4–60 °C), the purified native proteins were incubated in cleavage solution 'A' (100 mM Tris–HCl of different pH, 200 mM NaCl, 1 mM EDTA, 12 mM DTT and 5% glycerol) at the specified temperature for up to 24 h. The pH of the Tris–HCl solution was adjusted to the desired value prior to the addition of LexA, and also checked at the end of the experiment (using a pH paper) to check for any significant change in pH during the course of the experiment, and the measured pH, if deviating from pH 9.5, is indicated in the corresponding figure. In case of RecA-dependent cleavage of LexA, LexA was incubated with purified RecA from either *Anabaena* 7120 or *E. coli* in cleavage solution 'B' (20 mM Tris–HCl pH 7.5, 100 mM NaCl, 2 mM DTT, 5 mM β-mercaptoethanol, 2 mM $MgCl_2$, 0.5 mM EDTA, 2.5% glycerol, 1 mM Thio ATP, 6 μg mL^{-1} M13 single stranded DNA) as described earlier (Slilaty and Little, 1987) at 37 °C for 1 h. *Anabaena* RecA was purified as described in section 2.3 and *E. coli* RecA was commercially obtained from New England Biolabs, UK. RecA (30 μM) was incubated with 18 μM M13 ssDNA in Tris–NaCl (ingredient of cleavage solution B) and incubated at 37 °C for 5 min or 1 h prior to addition of other reagents of the solution and 17 μM LexA. Protein cleavage was visualised by PAGE followed by CBB staining. Activation of RecA upon addition of ssDNA was assessed in terms of its ATPase activity. RecA (30 μM) in Tris–HCl, pH 7.5 buffer was incubated with ATP (5 mM) in the absence or presence of M13 ssDNA (18 μM) at 37 °C for up to 1 h. Inorganic phosphate (P_i) released upon hydrolysis of ATP was estimated spectrophotometrically at 820 nm as described earlier (Fiske and Subbarow, 1925). In brief, 30% of the test solution was mixed with 70% of the reaction mix (10% ascorbic acid and 0.42% ammonium molybdate in 1:6 ratio) and incubated at 45 °C for 20 min prior to the measurement of absorbance at 820 nm.

2.6. Western blotting and immuno-detection

Purified LexA was used to generate specific polyclonal antibodies in rabbit. Proteins were extracted from *Anabaena* 7120 and *E. coli* cells grown under control conditions or after stress in Laemmli's buffer and separated by 15% SDS-PAGE, followed by transfer of proteins onto nitrocellulose membrane. Immunodetection was done using anti-LexA, anti-RecA or anti-Alr0088 antibodies in 1:1000 dilution, followed by secondary anti-rabbit IgG antibody coupled to alkaline phosphatase and colour developed using NBT-BCIP.

2.7. Electrophoretic mobility shift assay (EMSA)

The different DNA fragments (*lexA*, *recA* and *alr0088* promoters) used for EMSA were PCR amplified from *Anabaena* 7120 chromosomal DNA using specific primers (Table 1) and labelled with digoxigenin (DIG) using Random primer labelling Kit (Roche Diagnostics, Germany GmbH) at 37 °C for 1 h. The unincorporated dNTPs and hexanucleotide primers were removed using Qiagen nucleotide removal kit (Qiagen, USA). The labelled DNA fragments were incubated with increasing concentration of purified LexA protein (0.75–4.5 μg) in binding buffer (20 mM Tris–HCl, pH 8, 1 mM $MgCl_2$, 100 mM KCl, 8 mM DTT, 80 μg mL^{-1} BSA, 40% sucrose) at 37 °C for 30 min and reaction was stopped by adding 2 μl of 10× DNA loading dye containing EDTA. The reaction mixture was separated electrophoretically on a 6% native PAGE in 1X TBE buffer at 15 mA for 2 h at 4 °C, followed by electroblotting onto a nylon membrane. The UV-crosslinked blot was probed with anti-DIG antibody and developed using NBT-BCIP.

2.8. Reverse-transcriptase (RT)-PCR

Total RNA was extracted from logarithmic phase cultures of *Anabaena* 7120 (10 μg chlorophyll *a* mL^{-1}) using MN RNA isolation kit as per manufacturer's manual. The RNA was quantified spectrophotometrically at 260 nm. To ensure that the RNA carries no DNA contamination, PCR was carried out with total RNA using gene specific primers (Table 1) and Taq DNA polymerase. Total RNA, devoid of any DNA contamination, was used as template for RT-PCR with the gene (*lexA*, *recA*, *alr0088*) specific primers (Table 1) and one step-RT PCR enzyme mix (Qiagen, USA). The master mix was incubated first with reverse transcriptase at 50 °C for 30 min leading to the generation of c-DNA using gene specific reverse primer, followed by 30 cycles of normal PCR (denaturation at 94 °C for 15 min, denaturation at 94 °C for 45 s, annealing temperature was 52 °C for *alr0088* and *recA*, and 62 °C for *lexA*, elongation at 72 °C and following final extension at 72 °C for 10 min).

3. Results

3.1. Bioinformatics based comparison of cyanobacterial and E. coli LexA proteins

Comparison of the amino acid sequences of *E. coli* and *Anabaena* LexA indicated an overall identity of 34% and similarity of 54%, which was distributed through the length of the protein sequence (Fig. 1A). The identified cleavage site region (CSR) of *E. coli* LexA is 18 amino acid long, extending from L-78 to E-95 (Fig. 1A), which is fairly conserved across most bacterial LexA proteins. In case of *Anabaena* LexA, the corresponding region is longer by 4 amino acids (V-78 to D-99), due to the presence of additional amino acids (G^{86}, L^{87} and I^{88}) (Fig. 1B). The presence of cleavage site residues (A^{84} and G^{85}) and active site residues ($S^{119/118}$ and $K^{156/159}$) at nearly identical positions in the two proteins suggests that *Anabaena* LexA may be autocleavable. However, the regions in *E. coli* LexA predicted as essential for interaction with activated RecA (underlined in Fig. 1A) bear low homology with the corresponding regions in *Anabaena* LexA raising doubts on RecA-mediated cleavage of *Anabaena* LexA.

The comparison of the residues of the cleavage site region among the different cyanobacterial LexA proteins indicated great degree of similarity (Fig. 1B). Of the 11 sequences compared, the cleavage site residues A^{84}, G^{85} were present in all except *Synechocystis* LexA, wherein it was 'GG' (Fig. 1B). Another significant identity was the presence of the three extra amino acids '^{86}GLI/V^{88}' in the CSR in 9 of the 11 cyanobacterial LexA proteins. The two LexA proteins, which did not have this sequence, were the LexA proteins of *Synechocystis* and *Gloeobacter*, which have been predicted to be non-cleavable (Li *et al.*, 2010). The Q^{92} residue shown to be essential for RecA-dependent cleavage in *E. coli* RecA, its substitution resulting in enhanced RecA-independent cleavage (Luo *et al.*, 2001), is substituted by ^{96}E/D in most cyanobacterial LexA proteins (Fig. 1B), which may influence the RecA-dependent cleavage of cyanobacterial LexA proteins.

3.2. Analysis of autoproteolytic cleavage of Anabaena LexA

The Ni-NTA purified LexA was detected as a 25 kDa protein along with a significant amount of a 13 kDa protein (supp. Fig. S1A). Both these proteins cross-reacted with the polyclonal antibody raised against the 25 kDa protein (lane W, supp. Fig. S1A). The native *Anabaena* LexA protein was eluted in a single fraction in a HPLC based gel-filtration chromatography with Superdex 200HR column. The major peak (Fig. 1C) corresponded to 55 kDa based on the standard graph equation (inset in Fig. 1C), suggesting a dimeric form. A minor peak corresponding to about 25 kDa was also detected to the right of the major peak (Fig. 1C).

The 25 kDa *Anabaena* LexA protein (co-eluted with the cleaved 13 kDa protein) isolated under native conditions did not undergo any further cleavage in the pH range 7.5–8.0 at 37 °C for up to 24 h (Fig. 2A), but upon increasing the pH to 8.5 and higher, significant cleavage was observed (Fig. 2A and B). At pH 9.5, LexA was almost completely cleaved to the 13 kDa form after 24 h, while after 7 h, the extent of cleavage was similar to that observed after 24 h at pH 8.5 (Fig. 2A). Though, the rate of cleavage increased at pH 10, the net protein content decreased (Fig. 2B), so for all further experiments pH 9.5 was used. The autoproteolytic cleavage was enhanced further by the addition of 10 mM Ca^{2+}, resulting in almost complete cleavage at 5 h (Fig. 2C), and occurred over a temperature range of 30–57 °C (Fig. 2D). The variation in the pH over this temperature range was from 10.1 at 4 °C to 8.8 at 60 °C (Fig. 2D). The pH being above 8.5 in all cases, autoproteolytic activity of LexA would not be inhibited. The autoproteolytic cleavage of *Anabaena* LexA at pH 7.5 in the presence of activated RecA, based on an earlier published protocol (Slilaty and Little, 1987), was monitored. Generation of activated *Anabaena* RecA in the presence of M13ssDNA was monitored in terms of the change in ATPase activity of RecA, measured as increase in absorbance at 820 nm due to release of P_i. The interaction was monitored over a period of 1 h at 37 °C. In the absence of ssDNA, the ATPase activity of *Anabaena* RecA was low resulting in only an increase of 0.2 in absorbance over a 1 h period. On the other hand in the presence of M13 ssDNA, this increase was achieved within 2.5 min, and a steady increase in absorbance to over 1.1 units was observed over a period of 1 h (Fig. 1E). Based on standard graph, an increase of 1 unit of absorbance corresponds to 1 mM P_i released. Addition of neither the commercially available *E. coli* RecA, nor the laboratory isolated RecA from *Anabaena*, stimulated the cleavage of *Anabaena* LexA at pH 7.5 (Fig. 2F). Increasing the activation time of RecA with ssDNA to 1 h prior to the addition of LexA also did not result in the cleavage of AnLexA at pH 7.5 (data not shown). In order to check if the His-tag of LexA was interfering with its interaction with RecA, the His-tag was cleaved off from AnLexA using Factor Xa. RecA-dependent cleavage at pH 7.5 was not observed for AnLexAΔHis-tag either (data not shown).

3.3. Effect of mutations on autoproteolytic cleavage of Anabaena LexA

To analyse the role of major residues predicted to be involved in proteolytic cleavage, specific LexA mutants were generated. These included (i) two cleavage site mutants i.e. A84S and G85S; (ii) two active site mutants i.e. S118A and K159R; (iii) E96Q corresponding to a residue involved in RecA-independent cleavage; (iv) a three amino acid deletion mutant in the cleavage site region (CSR) i.e., $\Delta G^{86}LI^{88}$; and (v) domain specific mutants i.e., NTD with only the first 60 amino acids of LexA and CTD comprising of LexA from the 61st amino acid till end. All the mutant proteins, except LexA-ΔGLI and LexA-NTD were obtained as soluble proteins and were purified under native condition similar to the procedure employed for wild type LexA. LexA-ΔGLI and LexA-NTD were isolated with 8 M urea and then renatured. Mutation in either of the cleavage sites, ^{84}A or ^{85}G resulted in low levels (7.4–10.7%) of the 13 kDa polypeptide compared to wild type LexA (24.8%) antibody at time '0′, and both the 25 kDa full length and the 13 kDa polypeptide were found to cross-react with anti-LexA (Fig. 3A). However, upon incubation at 37 °C at pH 9.5 for 5 h, the 13 kDa polypeptide was not detected in the mutant LexA proteins, while in the WT LexA it was over 50% (Fig. 3A). The disappearance of the 13 kDa polypeptide in the mutant LexA proteins upon incubation could be due to non-specific proteolytic degradation. This indicated that the maximum probability of cleavage of AnLexA occurs between Ala-84 and Gly-85. Mutation in one of the catalytic residues, K159R decreased the cleavage to about 6% at '0 h' and to undetectable levels after 5 h at 37 °C (Fig. 3A). Mutation in another catalytic residue, S118A completely abolished the cleavage of the mutant LexA protein (Fig. 3B). On the other hand, E96Q mutation resulted in hypercleavage with a 3-fold higher levels of the 13 kDa polypeptide compared to the 25 kDa LexA protein after 5 h of incubation at pH 9.5 (Fig. 3B). Of the three deletion mutants, LexA-CTD exhibited cleavage comparable to wild type LexA (WT) (Fig. 3A and B), while LexA-ΔGLI was not cleaved (Fig. 3B). In order to ensure that the lack of cleavage of LexA-ΔGLI was not due to improper folding, the refolding of the protein was monitored at each step of reduction of urea concentration by analysing the fluorescence spectrum after excitation at 282 nm, and compared it with that of wild type LexA of *Anabaena*. The fluorescence spectrum of the native wild type LexA and renatured LexA-ΔGLI were identical, while upon denaturation with 8 M urea, both exhibited an additional peak at 425 nm (Fig. 3C).

Cleavage of the 25 kDa LexA should result in the release of a 12 kDa (corresponding to the N-terminal region up to Ala-84, and associated with the N-terminal His-tag) and a 13 kDa protein

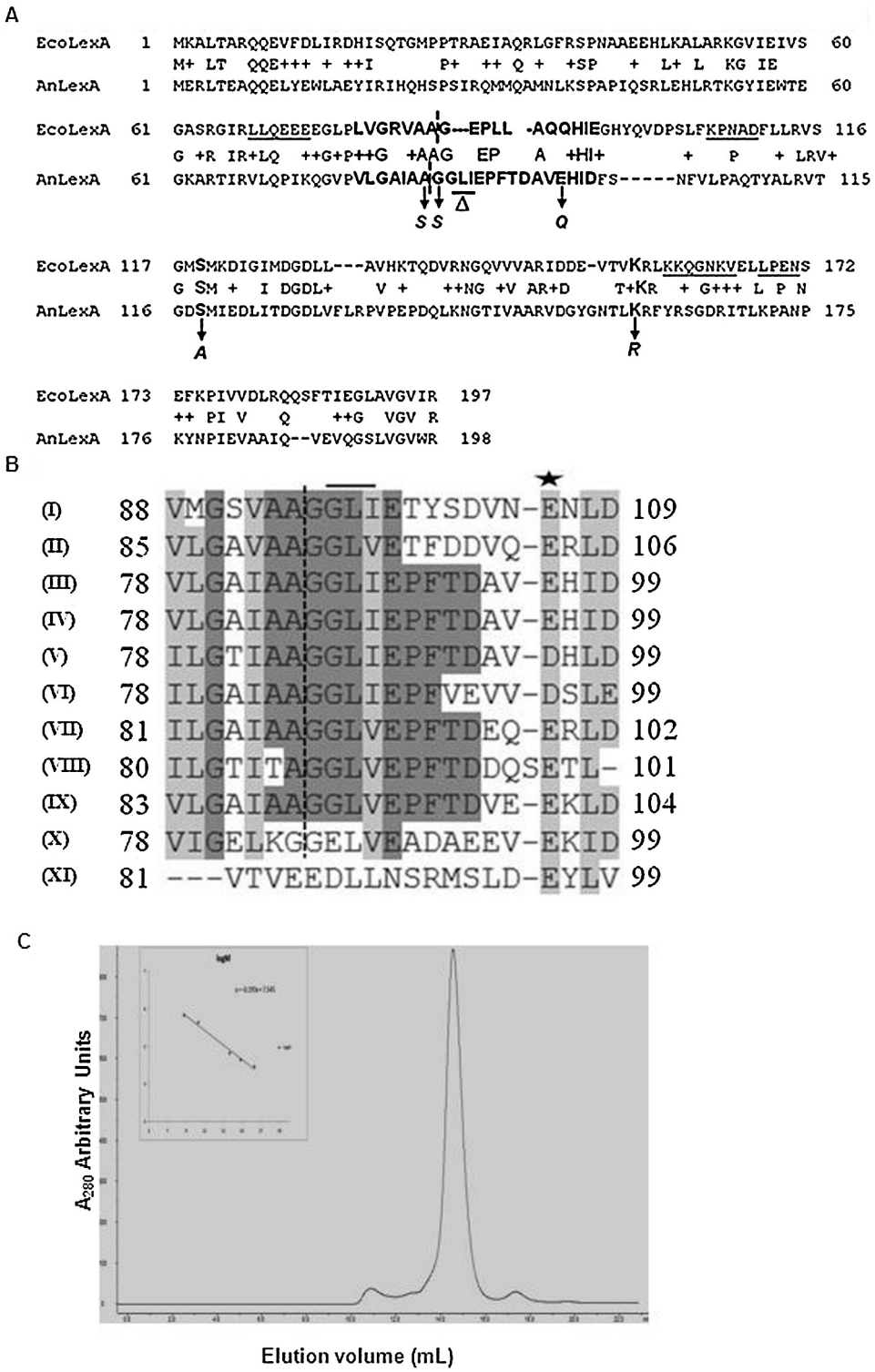


Fig. 1. Comparison of cyanobacterial and *E. coli* LexA proteins (A) Alignment of amino acid sequence of LexA proteins of *E. coli* (EcoLexA) (Uniprot no. P0A7C2) and *Anabaena* 7120 (AnLexA) (Uniprot no. Q8YMM5). The middle line represents the identical and similar amino acids in the two proteins. The cleavage site region (CSR) is shown in bold and cleavage site i.e. A^{84}–G^{85} by a vertical dotted line. The amino acids substituted in the point mutations of LexA generated are indicated by arrows and deletion of 'GLI' by Δ. (B) Comparison of the predicted cleavage site region (22 amino acids) of different cyanobacterial LexA proteins using ClustalW program (www.genome.jp/tools/clustalw). The identical amino acids are shaded 'dark grey' and similar amino acids 'light grey'. The cleavage site is shown by a dotted vertical line, the conserved ^{86}GLI/V by a horizontal line of the top and ^{96}E/D by an asterisk. The numbers on either side of the amino acids sequence indicate the position of the amino acid at the extreme ends in the corresponding full length protein. The different cyanobacterial LexA proteins along with their Uniprot nos. written in square brackets used for analysis are: (I) P9515_14231 (*Prochlorococcus marinus* MIT9515) [A2BXW9], (II) SynWH7803_1680 (*Synechococcus* sp. WH7803) [A5GME1], (III) alr4908 (*Anabaena* PCC7120) [Q8YMM5], (IV) Ava_2198 (*A. variabilis* ATCC29413) [Q3MB18], (V) Npun_F4481 (*Nostoc punctiforme* PCC73102) [B21VN5], (VI) cce_1899 (*Cyanothece* sp. ATCC 51142) [B1X0G0], (VII) MAE38110 (*Microcystis aerugenosa* NIES843) [B0JPQ9], (VIII) SYNPCC7002_A1849 (*Synechococcus* sp. PCC7002) [B1XQ55], (IX) AM1_3948 (*Acaryochloris marina* AM1) [B0C8B3], (X) sll1626 (*Synechocystis* sp PCC6803) [P73722], (XI) gll0709 (*Gloeobacter violaceus* PCC7421) [Q7NMQ6]. (C) Determination of oligomeric status of LexA. The native *Anabaena* LexA was subjected to gel filtration using Superdex 200 HR column and the eluted protein peak obtained by measuring absorbance at 280 nm. The standard graph used for molecular mass determination is shown in the inset.

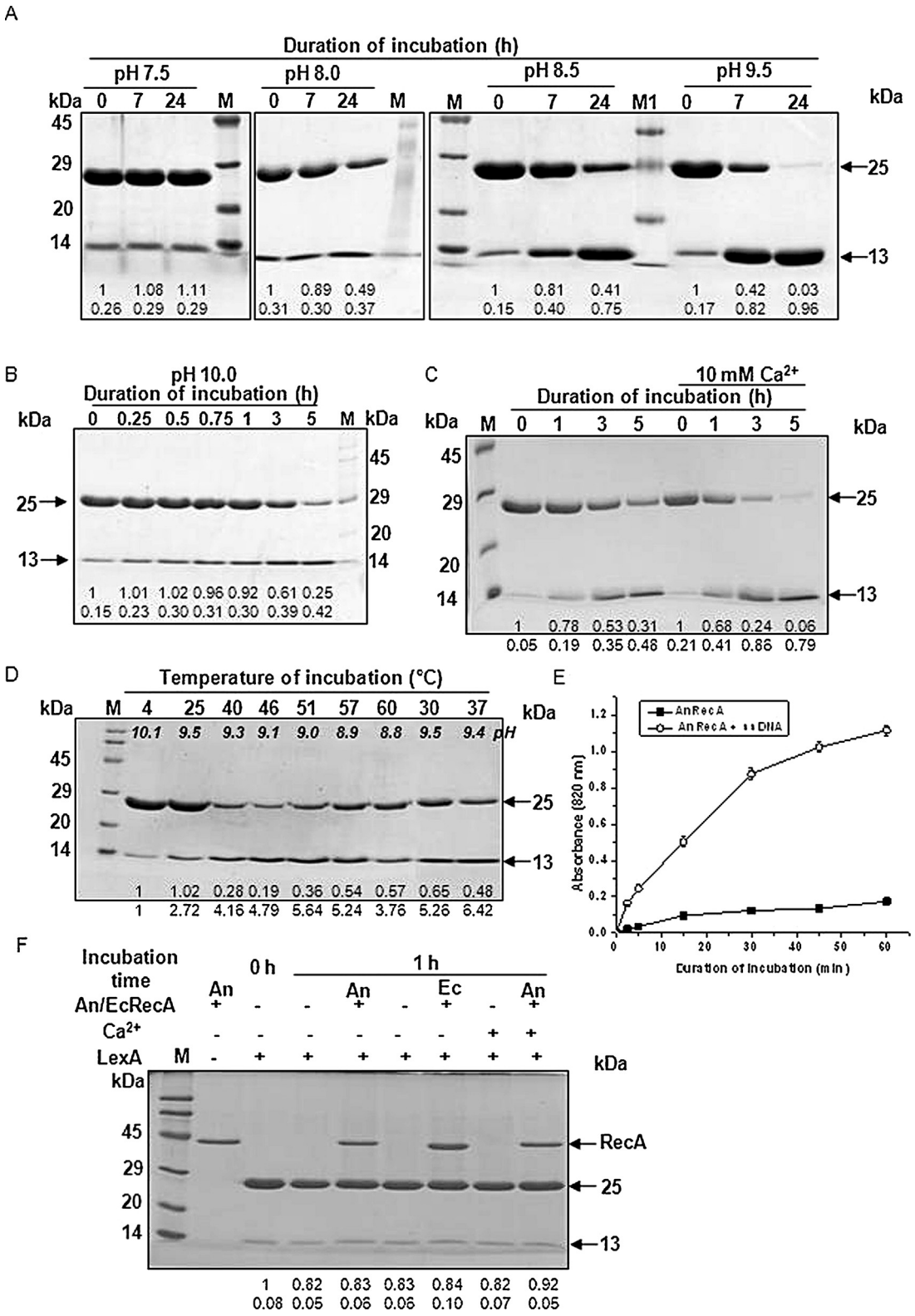


Fig. 2. Analysis of auto-proteolytic cleavage of *Anabaena* LexA. (A-D) Purified native *Anabaena* LexA protein (8 μM) was incubated in cleavage solution 'A' with (A) pH ranging from 7.5 to 9.5 for up to 24 h at 37 °C, (B) pH 10 for up to 5 h at 37 °C, (C) pH 9.5 at 37 °C for up to 5 h in the absence or presence of 10 mM Ca^{2+}, and (D) pH 9.5 solution for 3 h at temperature ranging from 4 °C to 60 °C. The pH values of the Tris–HCl solution at different temperatures is given at the top end of each lane. (E) ATP hydrolysis by *Anabaena* RecA. RecA (30 μM) in Tris–HCl, pH 7.5 buffer was incubated with 5 mM ATP in the absence or presence of 18 μM M13ssDNA at 37 °C for 1 h. An aliquot was taken out at the specified time points and assayed for the amount of P_i released as a measure of absorbance at 820 nm. (F) LexA (17 μM) was incubated in cleavage solution 'B' in the absence or presence of *Anabaena* or *E. coli* RecA (30 μM), pre-activated by the addition of M13ssDNA and thio-ATP, at 37 °C for 1 h. The different proteins visualised on polyacrylamide gel are indicated by arrows along with their molecular mass. Molecular mass of the protein marker used (M) are also indicated. The intensity of the 13 and 25 kDa protein bands relative to the intensity of the 25 kDa band at zero time, taken as '1′ is given below Figs. A–C and F. In 'D', the intensity of the 13 kDa and 25 kDa protein bands relative to the corresponding band at 4 °C has been shown. All the ratios stated are the mean of three independent experiments and the standard deviation ranged between 8 and 11%.

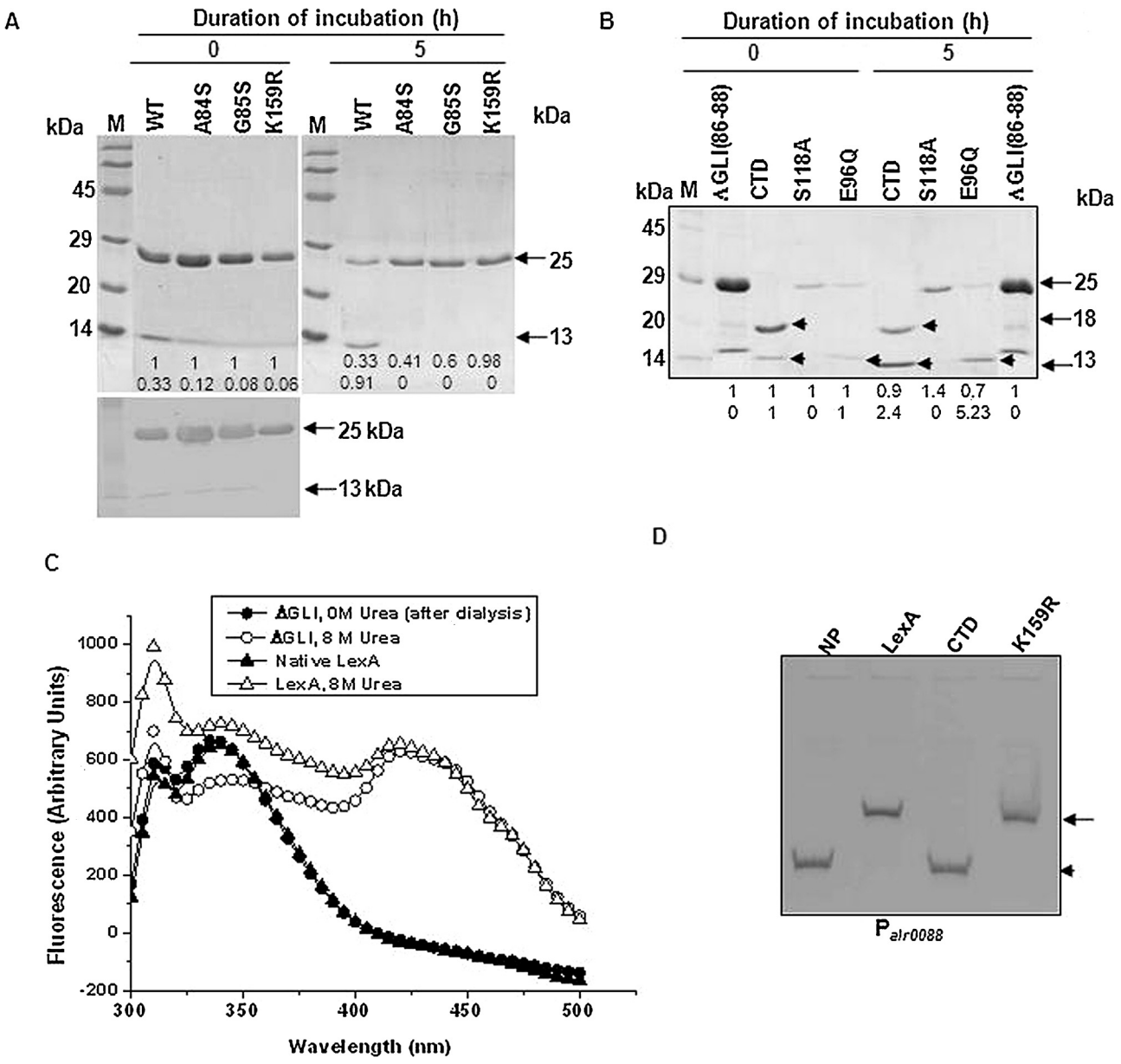


Fig. 3. Analysis of autoproteolytic cleavage of mutant LexA proteins of *Anabaena*. Incubation of purified wild type and mutant LexA proteins in cleavage solution 'A', pH 9.5 at 37 °C for 5 h. Proteins used were (A) wild type LexA (WT), point mutants (A84S, G85S and K159R), The intensity of the 13 kDa protein band relative to that of the 25 kDa protein is listed below the Coomassie stained gel in 'A' at '0 h'. In case of the '5 h' time point in 'A', the intensity of the band relative to the corresponding band at '0 h' has been given below the gel. Western blotting and immunodetection of the samples of '0′ time point in 'A' using anti-LexA antibodies is shown below the corresponding Coomassie-stained gel. The cross-reacting proteins have been indicated. (B) Point mutants (S118A and E96Q) and deletion mutants (ΔGLI and CTD). The intensities of the uncleaved and cleaved mutant LexA protein is given below the gel, with the intensity at time '0 h' for each mutant LexA protein individually taken as '1′ and that at 5 h determined relative to it. The molecular mass of the purified proteins and their cleaved products, if any, are indicated by arrows. (C) Fluorescence emission spectrum (obtained after subtracting the spectra obtained with the corresponding buffer) of purified native and denatured WT LexA and LexA-ΔGLI proteins after excitation at 282 nm. Denaturation was carried out with 8 M urea. (D) Electrophoretic mobility shift assay (EMSA) of DIG-labelled ~400 bp promoter fragment of *alr0088* ($P_{alr0088}$) in the presence of 0.5 μM of WT or mutant LexA proteins (CTD and K159R) followed by electrophoretic separation on polyacrylamide gel and visualisation. The free DNA is indicated by an arrowhead and bound DNA by arrows.

(corresponding to the C-terminal region from G-85). However, only the 13 kDa protein was detected in all experiments (Figs. 2 and 3A and B). Western blotting and immunodetection with either anti-*Anabaena* LexA antibody or anti-His antibody did not detect the 12 kDa protein. The protein could not be detected by MALDI-ToF either (data not shown), indicating that it is probably degraded. The protein corresponding to LexA-NTD could be overexpressed in *E. coli*, extracted and eluted under denaturing conditions in the presence of 8 M urea, but was not detected after renaturation by step-wise removal of urea (supp. Fig. 1B). No precipitate of the protein was detected either (data not shown).

3.4. Binding of Anabaena LexA to LexA box of Anabaena promoters

The native LexA was tested for its ability to bind DNA using ~400 bp putative promoter regions of *alr0088*, which codes for SSB (Kirti et al., 2013; 2014), *recA* and *lexA*, all of which have a putative LexA box (supp. Fig. S2A). Decrease in the mobility of all three promoter fragments was observed in the presence of the purified *Anabaena* LexA protein (supp. Fig. S2B–D), indicating that the isolated protein is active in terms of its DNA binding ability. This binding was unaffected if the cleavage of LexA was inhibited

by mutation of the catalytic site residue (K159R), but was completely inhibited upon the deletion of the NTD as observed using the promoter of *alr0088* ($P_{alr0088}$) as the target DNA (Fig. 3D). This confirmed that the DNA-binding activity resides in the NTD in case of *Anabaena* LexA as well.

3.5. Expression of alr0088, lexA and recA in Anabaena 7120

The expression of *alr0088*, *recA* and *lexA* genes whose promoters have a LexA box was monitored at both transcript and protein level. Total RNA isolated from *Anabaena* 7120 did not yield any product with Taq DNA polymerase and gene specific primers (Fig. 4A), indicating no DNA contamination. The RNA was then subjected to one step reverse trancriptase (RT)-PCR using gene specific primers (Table 1). Amplicons of near equal intensity of sizes 0.36 kb and 0.6 kb were obtained for *alr0088* and *lexA* respectively with 0.5 μg total RNA (Fig. 4A), while an amplicon of 1 kb and with 25% lower intensity was obtained in case of *recA* with 1 μg total RNA (Fig. 4A), indicating much lower transcription of *recA* compared to the other two genes. This was also reflected in the expression of the corresponding proteins under control growth conditions in *Anabaena* 7120, analysed by Western blotting and immunodetection with the corresponding antibodies. While the 13 kDa Alr0088 (~10 ng) was detected with 40 μg of total protein, LexA and RecA were detected (~10–15 ng) only when 350 μg and 400 μg respectively of total protein was used (Fig. 4B). The titre of each antibody used was determined using different dilutions of the antibody and fixed amount of the corresponding pure protein followed by quantification of the concentration of the pure protein cross-reacting with 1:1000 dilution of each of the antibody. Based on these experiments, the cross-reactivity of anti-LexA, anti-RecA and anti-Alr0088 antibodies with their respective pure proteins was found to be similar, and thus could be used for comparing the relative protein concentration in *Anabaena* cell extracts. The lower translation of LexA compared to Alr0088 in spite of similar levels of transcript suggests a possible regulation at post-transcriptional or translational level for *lexA* in *Anabaena*. *Anabaena* 7120 was subjected to UV-B and mitomycin C stresses both of which are known to induce SOS response. Its survival decreased to ~49% and ~1.2% upon exposure to 0.75 kJ m^{-2} and 3 kJ m^{-2} respectively. In case of mitomycin C stress given for 30 min, a gradual decrease in the cell viability was observed with increasing concentration of mitomycin C, it being ~80%, ~61% and ~42% after exposure to 1, 4 or 5 μg mit C mL^{-1} respectively (data not shown). Neither the level of LexA nor its proteolytic cleavage of LexA was observed when *Anabaena* 7120 was exposed to either mitomycin C or UV-B stress (Fig. 4C and D). However, even in *E. coli* wherein RecA expression increases and RecA gets activated in response to UV-B and mitomycin C stresses, the over-expressed *Anabaena* LexA did not undergo any proteolytic cleavage (Fig. 4E).

4. Discussion

Regulation of SOS-response in bacteria occurs through the repressor protein LexA, which gets inactivated by autoproteolytic cleavage at Ala^{84}–Gly^{85} bond mediated by activated RecA upon DNA damage (Janion, 2008). Though the SOS response has not been well documented in cyanobacteria, most cyanobacteria possess LexA protein (Li et al., 2010). Comparison of the sequences of different cyanobacterial LexA proteins revealed that two of the LexA proteins including that of *Synechocystis* 6803 represented a non-cleavable form (Li et al., 2010), speculated to be due to the absence of Ser^{119}, which is replaced by Asp (Mazón et al., 2004). Additionally, the *lexA* of *Synechocystis* 6803 was not involved in SOS response, but in regulating genes involved in carbon metabolism, RNA helicase and bidirectional hydrogenase (Domain et al., 2004; Li et al., 2010). The unlikelihood of the *Synechocystis* LexA to undergo the auto-cleavage in response to DNA damage may have resulted in the LexA adopting a different functional role as against the known role as a SOS response regulator in other bacteria (Janion, 2008). On the other hand, in *Anabaena* 7120 which possesses a theoretically cleavable form of LexA capable of binding to LexA boxes of *recA* and *lexA* gene promoters (Mazón et al., 2004, supp. Fig. S2), it is also involved in the transcriptional regulation of a bidirectional hydrogenase, which is speculated to have a putative LexA-Box in the promoter region (Sjöholm et al., 2007).

In vitro studies revealed that the LexA protein of *Anabaena* is indeed cleavable at the Ala-Gly bond, in the presence of Ser and Lys as active site residues (Fig. 3A and B), but it occurs only at alkaline pH of 8.5 and above (Fig. 2). Low levels of the 13 kDa polypeptide in LexA-A84S and LexA-G85S mutants at time '0' (Fig. 3A) could be on account of the possible cleavage between ^{83}A and ^{84}A in LexA-G85S mutant and between ^{85}G and ^{86}G in LexA-A84S mutant. Low probability of such cleavages has been reported recently for LexA from *Psuedomonas aeruginosa* (Mo et al., 2014). The disappearance of this polypeptide after 5 h of incubation could be due to proteolytic activity. In *E. coli*, both the NTD and CTD of LexA have shown to be proteolytically degraded by ClpXP and Lon protease systems (Little and Gellert, 1983; Neher et al., 2003). However, the extent of degradation of the two domains by proteases is not uniform, with NTD showing higher susceptibility (Cohn et al., 2011). A similar situation may also be true for *Anabaena* LexA resulting in preferential degradation of its NTD resulting and hence is not detected in pure form (Figs. 2 and 3). Our results also indicated the importance of three amino acid residues $^{86}GLI^{88}$ in the autocleavage of *Anabaena* LexA at alkaline pH (Fig. 3B), since LexA lacking the three residues (LexA-ΔGLI) did not undergo any cleavage in spite of the presence of A-G and S^{118} and K^{159} residues. This sequence (^{86}GLI/V) is conserved across all cyanobacterial LexA proteins predicted to be cleavable and absent in LexA of *Synechocystis* and *Gloeobacter* (Fig. 1B), both of which code for a non-cleavable LexA. Individual point mutations in the three residues would provide insight into whether the increase in length of the CSR due to these residues is crucial or the three individual residues itself are crucial for the autocleavage of *Anabaena* LexA, at alkaline pH. Another significant deviation from *E. coli* LexA was the presence of E/D residue at the 96th position (number corresponds to that in AnLexA) in all cyanobacterial LexA proteins (Fig. 1B). In *E. coli*, mutation of ^{92}Q to E resulted in a hypercleavable form of LexA (Luo et al., 2001), while the reverse was found to be true for the corresponding amino acid in AnLexA (Fig. 3B).

However, at physiological pH *Anabaena* LexA did not undergo cleavage irrespective of the absence or presence of RecA (Fig. 2F), which could be attributed to the absence of the predicted RecA interacting sites on *Anabaena* LexA (Fig. 1A). In *E. coli*, the RecA interacts with LexA at specific residues, which stimulates/stabilises change in conformation of the LexA from a non-cleavable form to a cleavable form (Kovačič et al., 2013). The low expression of *recA* in *Anabaena*, both at the transcriptional and at the protein levels (Fig. 4A and B), may have contributed to the absence of RecA-mediated cleavage in *Anabaena*. This would indicate that the SOS response may not be following the canonical RecA-LexA mediated pathway even in *Anabaena*, as has been reported earlier for *Synechocystis*. The only gene with possible role in SOS response studied in *Anabaena* 7120 is *alr0088*, which codes for a functional SSB (Kirti et al., 2014), and exhibits a marginal increase in expression in response to UV-B and mitomycin C stresses (Kirti et al., 2013). However, the protein levels of Alr0088 are

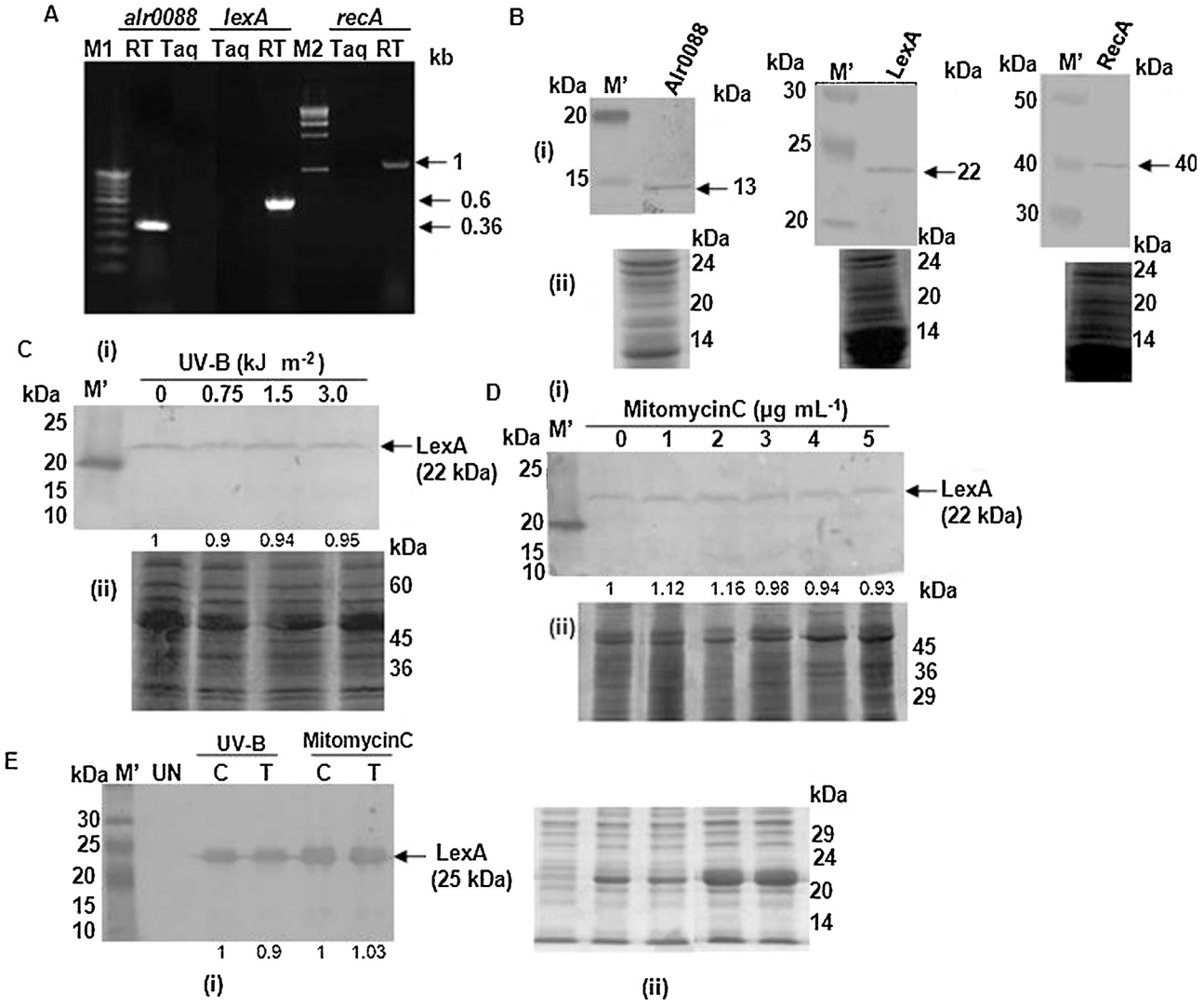


Fig. 4. Expression analysis of LexA in *Anabaena*. (A and B) Analysis of expression of *alr0088*, *lexA* and *recA* in *Anabaena* 7120 at (A) transcript level using total RNA isolated (0.5 µg for *alr0088* and *lexA*, and 1 µg for *recA*) from *Anabaena* 7120. PCR amplification was carried out using one-step reverse transcriptase (RT) or Taq DNA polymerase (Taq) using gene specific primers. The transcripts visualised on agarose gels along with their sizes are indicated by arrows. (B) Protein level using total protein extracts from *Anabaena* 7120 and immunodetected with the corresponding antibody. The cross-reacting protein along with the molecular mass is indicated by an arrow. Amount of protein used per lane was 40 µg for Alr0088, 350 µg for LexA and 400 µg for RecA. The Western blot is shown in Panel (i) and the corresponding Coomassie stained gel in Panel (ii). (C and D) Three-day-old *Anabaena* 7120 were exposed to different doses of (C) UV-B (dose rate 7 J m^{-2} s^{-1}), and (D) mitomycin C as indicated. Proteins were extracted, separated electrophoretically (350 µg per lane) and immunodetected with anti-LexA antibody. (E) Logarithmic phase cultures of *E. coli* BL21 (pET*lexA*) were induced with 1 mM IPTG for 30 min at 37 °C followed by exposure to 2 µg mL^{-1} mitomycin C for 1 h and 50 J m^{-2} of UV-B individually. The different lanes correspond to protein extracts from cells uninduced with IPTG (UN), induced but not exposed to stress i.e. control (C) and both induced and exposed to stress i.e. treatment (T). The cross-reacting LexA protein is indicated. Molecular mass of the pre-stained protein marker used (M′) is indicated. The intensity of lexA band during stress with respect to that under control conditions is indicated. The values are mean of three independent experiments and the standard deviation was in the range of 6–8%.

about 9–10-fold higher than that of LexA in the cell extracts of *Anabaena* 7120 based on immunodetection assays (Fig. 4B), and LexA itself neither undergoes cleavage nor any change in expression under these conditions (Fig. 4C–E). Thus, it is possible that the regulation by LexA, even in case of *alr0088* may not be following the typical de-repression arising due to auto-cleavage of LexA. Thus, it may be speculated that *Anabaena* LexA also may not be regulating DNA repair gene, as has been reported for *Synechocystis*, in spite of exhibiting alkaline pH-dependent cleavage.

However, this does not rule out *in vivo* autocleavage of *Anabaena* LexA, which could occur during such abiotic stresses which cause increase in cytoplasmic pH. Oxidative stress arising due to high light intensity has been shown to modulate the expression of LexA, RecA and several SOS response genes in the cyanobacteria, *Synechococcus* and *Prochlorococcus* (Blot et al., 2011; Kolowrat et al., 2010). Though a direct correlation indicating increase in pH during oxidative stress has not been shown in cyanobacteria or other bacteria, reports on alkaline pH inducing oxidative stress and SOS response has been reported (Follmann et al., 2009; Maurer et al., 2005; Schuldiner et al., 1986).

Thus, compared to other bacteria, the regulation of active LexA levels, which could in turn regulate expression of other genes, is different in cyanobacteria. The lack of sufficient levels of RecA may have resulted in evolving a mechanism of a RecA-independent cleavage of LexA. This would also help in achieving a larger number of gene substrates for regulation, other than those required for repair of DNA damage, encompassing regulation of a wide range of abiotic stresses. Cyanobacteria, being an ancient organism, it is speculated that RecA-mediated cleavage in other bacteria may have evolved much later.

Acknowledgements

The authors wish to thank Dr. Prashanth Raghavan, MBD, BARC for providing *Anabaena recA* plasmid (pET*recA*) and the corresponding anti-RecA antibody and Dr. Gagandeep Gupta, SSPD, BARC for HPLC-based size determination.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biocel.2014.12.003.

References

Blot N, Mella-Flores D, Six C, Le Corguillé G, Boutte C, Peyrat A, et al. Light history influences the response of the marine cyanobacterium *Synechococcus* sp. WH7803 to oxidative stress. Plant Physiol 2011;156:1934–54.

Castenholz RW. Culturing of cyanobacteria. Methods Enzymol 1988;167:68–93.

Cohn TM, Kjelgaard P, Frees D, Penedes JR, Ingmer H. Clp-dependent proteolysis of the LexA N-terminal domain in *Staphylococcus aureus*. Microbiology 2011;157:677–84.

Domain F, Houot L, Chauvat F, Cassier-Chauvat C. Function and regulation of the cyanobacterial genes lexA, recA and ruvB: LexA is critical to the survival of cells facing inorganic carbon starvation. Mol Microbiol 2004;53:65–80.

Fernandez de Henestrosa AR, Ogi T, Aoyagi S, Chafin D, Hayes JJ, Ohmori H, et al. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. Mol Microbiol 2000;35:1560–72.

Fiske CH, Subbarow. The calorimetric determination of phosphorus. J Biol Chem 1925;66:375–400.

Follmann M, Ochrombel I, Krämer R, Trötschel C, Poetsch A, et al. Functional genomics of pH homeostasis in *Corynebacterium glutamicum* revealed novel links between pH response, oxidative stress, iron homeostasis and methionine synthesis. BMC Genomics 2009;10:621, http://dx.doi.org/10.1186/1471-2164-10-621.

Giese KC, Michalowski1 CB, Little JW. RecA-dependent cleavage of LexA dimers. J Mol Biol 2008;377:148–61.

Janion C. Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. Int J Biol Sci 2008;4:338–44.

Kirti A, Rajaram H, Apte SK. Characterization of two naturally truncated, Ssb-like proteins from the nitrogen-fixing cyanobacterium, *Anabaena* sp. PCC7120. Photosynth Res 2013;118:147–54.

Kirti A, Rajaram H, Apte SK. The hypothetical protein 'All4779', and not the annotated 'Alr0088' and 'Alr7579' proteins, is the major typical single-stranded DNA binding protein of the Cyanobacterium, *Anabaena* sp. PCC7120. PLOS ONE 2014., http://dx.doi.org/10.1371/journal.pone.0093592.

Kolowrat C, Partensky F, Mella-Flores D, Le Corguillé G, Boutte C, Blot N, et al. Ultraviolet stress delays chromosome replication in light/dark synchronized cells of the marine cyanobacterium *Prochlorococcus marinus* PCC9511. BMC Microbiol 2010., http://dx.doi.org/10.1186/1471-2180-10-204.

Kovačič L, Paulič N, Leonardi A, Hodnik V, Anderluh G, Podlesek Z, et al. Structural insight into LexA–RecA* interaction. Nucleic Acids Res 2013;41:9901–10.

Little JW, Autodigestion of LexA and phage lambda repressors. Proc Natl Acad Sci U S A 1984;81:1375–9.

Little JW, Gellert M. The SIS regulatory system: control of its state by the level of RecA protease. J Mol Biol 1983;167:791–808.

Little JW, Hill SA. Deletions within a hinge region of a specific DNA-binding protein. Proc Natl Acad Sci U S A 1985;82:2301–5.

Li S, Xu M, Su Z. Computational analysis of LexA regulons in cyanobacteria. BMC Genomics 2010., http://dx.doi.org/10.1186/1471-2164-11-527.

Lowry OH, Rosebrough NJ, Farr AL. Randall RJ protein measurement with the Folin phenol reagent. J Biol Chem 1951;193:265–75.

Luo Y, Pfuetzner RA, Mosimann S, Paetzel M, Frey EA, Cherney M, et al. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. Cell 2001;106:585–94.

Maurer LM, Yohannes E, Bondurant SS, Radmacher M, Slonczewski JL. pH regulates genes for flagellar motility, catabolism and oxidative stress in *Escherichia coli* K-12. J Bacteriol 2005;187:304–19.

Mazón G, Lucena JM, Campoy S, Fernandez de Henestrosa ARF, Candau P, Barbé J. LexA-binding sequences in gram positive and cyanobacteria are closely related. Mol Gen Genomics 2004;271:40–9.

Mo CY, Birdwed LD, Kohli RM. Specificity determinants for autoproteolysis of LexA, a key regulator of bacterial SOS mutagenesis. Biochemistry 2014;53:3158–68.

Neher SB, Flynn JM, Sauer RT, Baker TA. Latenet ClpX-recognition signals ensure LexA destruction after DNA damage. Genes Dev 2003;17:1084–9.

Patterson-Fortin LM, Colvin KR, Owttrim GW. A LexA-related protein regulates redox-sensitive expression of the cyanobacterial RNA helicase, *crhR*. Nucleic Acids Res 2006;34:3446–54.

Schlacher K, Goodman MF. Lessons from 50 years of SOS. DNA-damage-induced mutagenesis. Nature Rev Mol Cell Biol 2007;8:587–94.

Schuldiner S, Agmon V, Brandsma J, Cohen A, Friedman E, et al. Induction of SOS functions by alkaline intracellular pH in *Escherichia coli*. J Bacteriol 1986;168:936–9.

Sjöholm J, Oliveira P, Lindblad P. Transcription and regulation of the bidirectional hydrogenase in the cyanobacterium *Nostoc* sp. strain PCC 7120. Appl Environ Microbiol 2007;73:5435–46.

Slilaty SN, Little JW. Lysine-156 and serine-119 are required for LexA repressor cleavage: a possible mechanism. Proc Natl Acad Sci U S A 1987;84:3987–91.

Tetu SG, Johnson DA, Varkey D, Phillippy K, Stuart RK, Dupont CL, et al. Impact of DNA damaging agents on genome-wide transcriptional profiles in two marine *Synechococcus* species. Front Microbiol 2013;4:232.

Walker GC. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol Rev 1984;48:60–93.

J. Phycol. **53,** 322–332 (2017)

DOI: 10.1111/jpy.12500

DIFFERENTIAL REGULATION OF *SSB* GENES IN THE NITROGEN-FIXING CYANOBACTERIUM, *ANABAENA* SP. STRAIN PCC7120[1]

Anurag Kirti, Arvind Kumar, and Hema Rajaram[2]

Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India

***Anabaena* sp. PCC7120 possesses three genes coding for single-stranded DNA-binding (SSB) protein, of which *ssb1* was a single gene, and *ssb2* and *ssb3* are the first genes of their corresponding operons. Regulation of the truncated *ssb* genes, *ssb1* (*alr0088*) and *ssb2* (*alr7559*), was unaffected by N-status of growth. They were negatively regulated by the SOS-response regulatory protein LexA, as indicated by the (i) binding of *Anabaena* LexA to the LexA box of regulatory regions of *ssb1* and *ssb2*, and (ii) decreased expression of the downstream *gfp* reporter gene in *Escherichia coli* upon co-expression of LexA. However, the full-length *ssb* gene, *ssb3* (*all4779*), was regulated by the availability of Fe^{2+} and combined nitrogen, as indicated by (i) increase in the levels of SSB3 protein on Fe^{2+}-depletion and decrease under Fe^{2+}-excess conditions, and (ii) 1.5- to 1.6-fold decrease in activity under nitrogen-fixing conditions compared to nitrogen-supplemented conditions. The requirement of Fe^{2+} as a co-factor for repression by FurA and the increase in levels of FurA under nitrogen-deficient conditions in *Anabaena* (Lopez-Gomollon et al. 2007) indicated a possible regulation of *ssb3* by FurA. This was substantiated by (i) the binding of FurA to the regulatory region of *ssb3*, (ii) repression of the expression of the downstream *gfp* reporter gene in *E. coli* upon co-expression of FurA, and (iii) negative regulation of *ssb3* promoter activity by the upstream AT-rich region in *Anabaena*. This is the first report on possible role of FurA, an important protein for iron homeostasis, in DNA repair of cyanobacteria.**

***Key index words:* *Anabaena*; AT-rich; DNA-damage-inducing stress; FurA; LexA; promoter activity; SSB**

***Abbreviations*: bp, base pair; Cm, chloramphenicol; dsDNA, double-stranded DNA; DTT, dithiothreitol; EDTA, ethylene diamine tetraacetic acid; GFP, green fluorescent protein; Kan, kanamycin; kGy, kilo Gray; Neo, neomycin; N, nitrogen status; RDRM, radiation/desiccation response motif; RT-PCR, reverse transcriptase-polymerase chain reaction; SSB, single-stranded DNA-binding protein; ssDNA, single-stranded DNA; TBE, Tris-Borate EDTA**

Single-stranded DNA-binding (SSB) proteins play a crucial role during DNA replication, repair, and recombination by protecting the single-stranded DNA (ssDNA) generated upon the unwinding of the duplex double-stranded DNA (dsDNA) until the ssDNA regions are utilized by downstream proteins involved in these pathways (Shereda et al. 2008). In addition to the requirement of SSB throughout the growth period of bacteria, enhanced amount of the protein would be required during logarithmic phase when replication is high or after DNA damage to facilitate repair. This necessitates regulation of *ssb* gene expression in bacteria. In *Escherichia coli*, the expression of the *ssb* gene is controlled by three promoters, one of which is inducible by DNA damage (Brandsma et al. 1985). The *ssb* gene is localized divergently with respect to *uvrA*, both being regulated by the same LexA box (Brandsma et al. 1983). Similar organization of the *ssb* gene is found in several gram-negative bacteria, but not gram-positive bacteria. In gram-positive bacteria, such as *Bacillus subtilis*, the main *ssb* gene is flanked by *rpsF* and *rpsR*, transcribed from the *rpsF* promoter (Lindner et al. 2004). However, the regulation of the *rpsF* promoter is not under the control of DinR, the equivalent of LexA in *B. subtilis* (Lindner et al. 2004). Based on the organization of the different *ssb* genes across bacterial genomes, they have been grouped under four classes (Lindner et al. 2004). Those having the gene organization as *rpsF–ssb–rpsR* have been grouped together in Class I (correspond to those bacteria having multiple *ssb* paralogs) or Class II (bacteria with single *ssb* gene). Class III corresponded to those with *uvrA–ssb* gene organization, while all other bacteria wherein the *ssb* gene organization did not fall into any of the above three classes was grouped together in Class IV. The *ssb* gene of cyanobacteria including *Anabaena* PCC7120 is placed in Class IV (Lindner et al. 2004).

The expression of *ssb* in response to DNA damage was found to be relatively slow compared to other genes, both in *E. coli* (Perrino et al. 1987) and *B. subtilis* (Lindner et al. 2004). In *E. coli* and several other bacteria, RecA is activated upon DNA damage. The activated RecA mediates the activation of the auto-proteolytic activity of the repressor LexA. The resulting cleavage of LexA renders it inactive, thus ensuring de-repression of the SOS regulon (Janion 2008). Another mode of regulation of

[1]Received 15 September 2015. Accepted 23 September 2016.
[2]Author for correspondence: e-mail hemaraj@barc.gov.in.
Editorial Responsibility: D. Lindell (Associate Editor)

SSB expression observed in *E. coli* was at the translational level, wherein SSB specifically bound to its own mRNA, thereby preventing its translation (Shimamoto et al. 1987). In *Deinococcus radiodurans*, which does not possess a LexA-like protein, the increased expression of SSB upon exposure to mitomycin C or γ-radiation was through two conserved 17 bp radiation/desiccation response (RDRM) motifs (Ujaoney et al. 2010).

In the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120, hereafter referred to as *Anabaena* 7120, the upstream region of *alr0088*, coding for truncated SSB interacts with the LexA protein (Kumar et al. 2015). Unlike other bacterial LexA proteins, which exhibit auto-cleavage in the presence of activated RecA (Janion 2008), *Anabaena* LexA exhibits auto-cleavage only at alkaline pH and not in the presence of activated RecA (Kumar et al. 2015). This raised a question about the regulation of the *ssb* genes by LexA in *Anabaena*. *Anabaena* 7120 has three genes coding for SSB, two of these, (i.e., *alr0088* and *alr7579*), code for truncated SSBs having only the N-terminal OB-fold, while *all4779* codes for a full-length SSB (Kirti et al. 2014). Of these, *alr0088* and *all4779* are localized on the chromosome and *alr7579* on beta-plasmid (http://genome.microbedb.jp/cyanobase/Anabaena). The two truncated *ssb* genes exhibited less than a 1.5-fold increase in the corresponding protein levels in response to mitomycin C and UV-B stresses (Kirti et al. 2013). This article deals with the localization of the *ssb* gene promoters of *Anabaena* and their regulation in response to DNA damage–induced stresses. All three *ssb* genes of *Anabaena* were found to be regulated by their upstream regions, with LexA acting as a repressor for the truncated *ssb* genes, *alr0088* (*ssb1*) and *alr7579* (*ssb2*) and FurA as a repressor for *all4779* (*ssb3*).

METHODS

Organism and growth conditions. *Escherichia coli* cells were grown in Luria-Bertani medium at 37°C with constant shaking at 150 rpm, in the presence of appropriate antibiotics, i.e., 34 μg chloramphenicol · mL^{-1} (Cm_{34}), or 50 μg kanamycin · mL^{-1} (Kan_{50}) or 100 μg carbenicillin · mL^{-1} (Cb_{100}). *Anabaena* strains, both wildtype (*Anabaena* PCC7120), and the recombinant strains were maintained and grown under axenic conditions in BG-11 liquid medium (Castenholz 1988), with (BG-11, N^+) or without (BG-11, N^-) 17 mM $NaNO_3$, in the presence of 10 μg · mL^{-1} neomycin (Nm_{10}) for the recombinant strains, under nitrogen-fixing conditions under continuous illumination (30 μmol photons · m^{-2} · s^{-1}) and shaking (100 rpm) at 27 ± 2°C. Growth was measured in terms of chlorophyll *a* content in methanolic extracts (Mackinney 1941). For selection of recombinant *Anabaena* strains, neomycin was used at 10 μg · mL^{-1} (Nm_{10}) in liquid or 25 μg · mL^{-1} (Nm_{25}) in solid BG-11, N^+ media.

Three-day-old nitrogen-fixing or nitrogen-supplemented cultures of different recombinant *Anabaena* strains were concentrated to 10 μg chlorophyll *a* density · mL^{-1} and subjected to 5 μg · mL^{-1} mitomycin C for 30 min, 6 kGy of ^{60}Co γ-irradiation (dose rate 4.5 kGy · h^{-1}), or 6 d of desiccation. In case of irradiation stress, the corresponding control (sham-irradiated) was kept in the dark under the same conditions as the irradiated sample except for exposure to γ-irradiation. In case of desiccation stress, the concentrated culture was filtered for 1 min through vacuum onto a membrane and placed on a petridish in a desiccator containing fused $CaCl_2$ for 6 d under illumination. In the corresponding control (humid chamber), the membranes were placed on agar plates (without any growth medium), which was incubated in a chamber containing water. For experiments involving different Fe^{2+} concentrations, 3 d old cultures grown under control conditions were washed five times with 20-fold excess of Fe^{2+} free or 2-fold higher concentration (2×) Fe^{2+} BG-11, N^- media and inoculated in the corresponding media, i.e., Fe^{2+}-free (depletion) or Fe^{2+}-rich (excess) media at 1 μg chl *a* density per mL and grown for 3 d. The 3 d old cultures were repeatedly subcultured in their respective generations up to four times.

Protein extraction and immunodetection of SSB3 protein. Protein was extracted from 3 d old cultures of *Anabaena* grown under different conditions using Laemmli's buffer, electrophoretically separated (40 μg protein per lane) on 12% SDS-PAGE, and blotted on to nitrocellulose membrane. The SSB3 protein was immunodetected using anti-SSB3 polyclonal antibody as described earlier (Kirti et al. 2014).

Generation of recombinant Anabaena *strains harboring different plasmid constructs.* The different promoter fragments corresponding to *ssb1*, *ssb2*, and *ssb3* genes of *Anabaena* 7120 were PCR amplified from the genomic DNA using specific primers as indicated (Table 1) and ligated to pAM1956 (Table 2) at *Sac*I/*Kpn*I restriction sites. The resulting plasmid constructs containing 400 bp P_{ssb1}, 400 bp P_{ssb2}, 350 bp P_{ssb3}, and 150 bp P_{ssb3s} promoter fragments were designated as pAM-P_{ssb1}::*gfpmutII*, pAM-P_{ssb2}::*gfpmutII*, pAM-P_{ssb3}::*gfpmutII*, and pAM-P_{ssb3s}::*gfpmutII*, respectively (Table 2). These plasmids were individually introduced into *Anabaena* by triparental conjugation, ex-conjugants selected on BG-11, N^+, Neo_{25} plates, and screened for GFP expression, as described earlier (Raghavan et al. 2011), resulting in the recombinant strains, An(pAM-P_{ssb1}::*gfpmutII*), An(pAM-P_{ssb2}::*gfpmutII*), An(pAM-P_{ssb3}::*gfpmutII*), and An(pAM-P_{ssb3s}::*gfpmutII*) strains (Table S1 in the Supporting Information).

Reverse transcriptase-PCR analysis. To assess if the isolated RNA carries any DNA contamination, PCR was carried out with total RNA using specific gene primers and Taq DNA polymerase. Only if no amplicons were obtained, the RNA was used for reverse transcriptase (RT)-PCR using total RNA (0.5 μg) as a template with gene-specific primers (Table 1) and one-step RT-PCR enzyme mix (Qiagen, USA). The master mix was incubated first for reverse transcription at 50°C for 30 min leading to the generation of c-DNA, followed by 30 cycles of normal PCR (denaturation at 94°C for 15 min, denaturation at 94°C for 45 s, annealing at 52°C for gene-specific primers of *ssb1* and *ssb2* and 54°C for *ssb3*, elongation at 72°C and following final extension at 72°C for 10 min). Amplification of genomic DNA with Taq DNA polymerase for each set of gene primers was kept as a positive control.

Overexpression and purification of Anabaena *LexA and* Anabaena *FurA proteins.* The different plasmid constructs (pET*lexA* and pET*furA*) were individually transformed into *E. coli* BL21(p*lysS*)(DE3). Overexpression of the corresponding proteins, coded by insert cloned, was achieved upon incubation of logarithmic phase cultures with 1 mM IPTG at 37°C for up to 3 h. The proteins were purified using Ni-NTA affinity chromatography with step-wise increase of imidazole

TABLE 1. (a) Primers used in this study. (b) Amplicons used in this study.

(a) Primers	Nucleotide sequence	Restriction site
P_{ssb1}Fwd	5′-GCGCCGGAGCTCGGTATTTGCTGTACCGAG-3′	*Sac*I
P_{ssb1}Rev	5′-GATCGGTACCCTGCCTTATCCTTTAGTACATAAGTACT-3′	*Kpn*I
P_{ssb2}Fwd	5′-GCAACGAGCTCCAACAGATGTTTCTCCT-3′	*Sac*I
P_{ssb2}Rev	5′-GCCCGGGTACCAACGCGAGAATTGAT-3′	*Kpn*I
P_{ssb3}Fwd	5′-GAAGAGCTCTCTGTCAATAAAGTGGT-3′	*Sac*I
P_{ssb3s}Fwd	5′-GAAGAGCTCGCCAACTATTAAATATTAAGTAT-3′	*Sac*I
P_{ssb3}Rev	5′-GCCCGGGTACCTAAATCCTCCTGACTGTA-3′	*Kpn*I
78 R	5′-TCGCCACCCCCGGTATAGTTGCA-3′	
79 F	5′-GGCCATATGAACAGCTGTGTTTTA-3′	
79 R	5′-GGCGGATCCTAAAATGGAATATCGTC-3′	
80 F	5′-CCCTATATTACAATGTTGCCTC-3′	
*furA*Fwd	5′-GGCCATATGACTGTCTACACAAAT-3′	*Nde*I
*furA*Rev	5′-GGCGGATCCTAAAGTGGCATGAGCGC-3′	*Bam*HI

(b) Amplicons	Primer set used	Size (kb)
P_{ssb1}	P_{ssb1} Fwd and P_{ssb1} Rev	0.4
P_{ssb2}	P_{ssb2} Fwd and P_{ssb2} Rev	0.4
P_{ssb3}	P_{ssb3} Fwd and P_{ssb3} Rev	0.34
P_{ssb3s}	P_{ssb3s} Fwd and P_{ssb3} Rev	0.15

concentration (20 mM–1 M). All fractions were visualized by electrophoretic separation on 15% SDS-PAGE after staining with Coomassie Brilliant Blue G-250. Fractions containing pure proteins were pooled, and imidazole was removed by dialysis. LexA protein was stored in buffer containing 100 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM ethylene diamine tetraacetic acid (EDTA), and 5% glycerol, while FurA protein was stored in a buffer containing 100 mM Tris-HCl pH 8.0, 300 mM NaCl. Prior to analysis, the protein content was estimated by using Lowry reagent.

Electrophoretic mobility shift assay. The different amplicons (Table 1B; 50 ng) used for electrophoretic mobility shift assay (EMSA) were incubated with specified concentrations of purified LexA or FurA proteins in binding buffer (20 mM Tris pH 8.0, 1 mM $MgCl_2$, 100 mM KCl, 8 mM DTT, 80 $\mu g \cdot mL^{-1}$ bovine serum albumin, 40% sucrose) at 37°C for 30 min, and reaction was stopped by adding 2 μL of 10× DNA loading dye containing EDTA. The reaction mixture was separated electrophoretically on an 8% native polyacrylamide gel in 1× Tris-Borate-EDTA (TBE) buffer at 15 mA for 2 h at 4°C, followed by staining with SYBR Dye I for 30 min and visualized under UV transilluminator.

Analysis of promoter activity in terms of GFP expression. The activity of different *ssb* promoters was quantitatively estimated under control growth conditions as well as after the specified duration of stress in *E. coli* and *Anabaena* cells in terms of GFP expression using a JASCO FP6500 spectrofluorometer. The excitation and emission wavelengths used were 490 and 510 nm, respectively. In case of *E. coli* cells, BL21(p*lys*S) cells harboring the plasmid constructs, pET*lexA* or pET*furA* (Table S1), overnight grown cultures were inoculated in fresh LB medium at 0.05 OD_{600} per mL and allowed to grow until OD_{600} of 0.6 was reached. The expression of LexA or FurA was induced by the addition of 1 mM IPTG, and samples taken for measurement of promoter activity at different time points as indicated. In case of recombinant *Anabaena* strains bearing different promoter constructs as well as the vector control, AnpAM (Table S1), 3 d old cultures subjected to different stresses were used. GFP fluorescence was expressed as arbitrary units per OD_{600} for *E. coli* and per μg chl *a* for *Anabaena*. The values presented are average of three sets of technical replicates, each set comprising of two biological replicates.

TABLE 2. Plasmids used in this study.

Plasmid	Characteristics	Source
pET16b	Cb^r, Expression vector with N-terminal His-tag	Novagen
pET*furA*	Cb^r, 0.45 kb *furA* gene cloned in pET16b at *Nde*I, *Bam*HI restriction sites	This study
pET*lexA*	Cb^r, 0.6 kb *lexA* gene cloned in pET16b at *Nde*I, *Bam*HI restriction sites	Kumar et al. (2015)
pAM1956	Kan^r, promoterless, replicative vector with g*fpmutII* as a reporter gene	Yoon and Golden (1998)
pAM-P_{ssb1}::*gfpmutII* (pAM-P_{ssb1})	0.4 kb P_{ssb1} fragment cloned in pAM1956 at *Sac*I-*Kpn*I restriction site	This study
pAM-P_{ssb2}::*gfpmutII* (pAM-P_{ssb2})	0.4 kb P_{ssb2} fragment cloned in pAM1956 at *Sac*I-*Kpn*I restriction site	This study
pAM-P_{ssb3}::*gfpmutII* (pAM-P_{ssb3})	0.34 kb P_{ssb3} fragment cloned in pAM1956 vector at *Sac*I-*Kpn*I restriction site	This study
pAM-P_{ssb3s}::*gfpmutII* (pAM-P_{ssb3s})	0.15 kb P_{ssb3s} fragment cloned in pAM1956 vector at *Sac*I-*Kpn*I restriction site	This study

RESULTS

Localization of putative promoter regions of Anabaena ssb *genes.* The genes coding for truncated SSBs of *Anabaena* 7120, "*alr0088*" and "*alr7579*" has been designated as *ssb1* and *ssb2*, respectively, and "*all4779*" which codes for a full-length SSB has been designated as *ssb3*. Based on the information available on genome organization on cyanobase, the orientation of *ssb1* is found to be opposite to that of its flanking genes, i.e., *all0087* (coding for MreB) and *all0089* (coding for hypothetical protein; Fig. 1a), indicating that it is a monocistronic gene, and hence transcribed by its upstream region. In case of

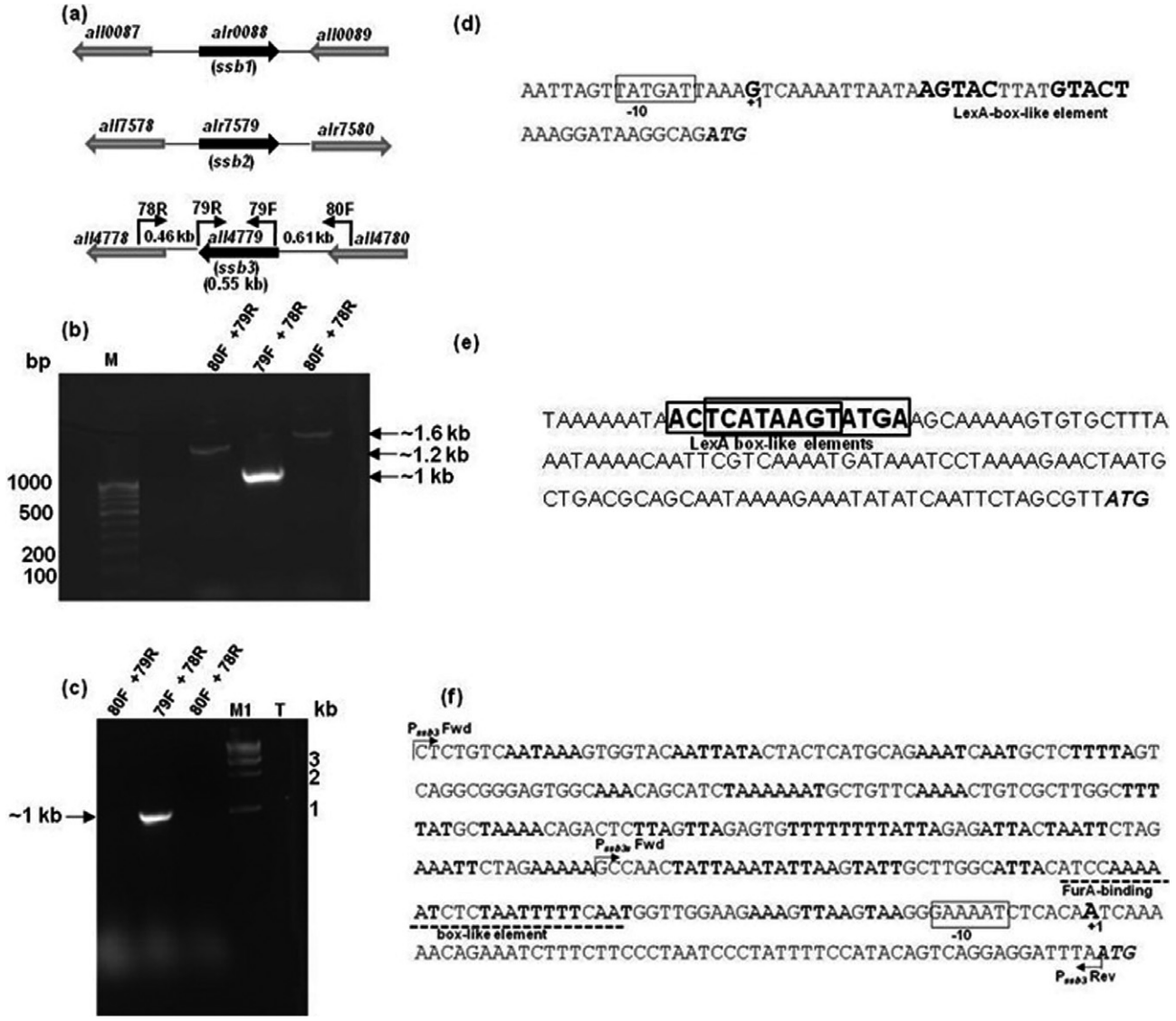


FIG. 1. Probable promoter regions of *Anabaena ssb* genes. (a) Genomic organization of *Anabaena ssb1* (*alr0088*), *ssb2* (*alr7579*), and *ssb3* (*all7579*) genes with respect to its immediate neighbors on either side. The direction of arrow indicates the direction of transcription of the gene. The approximate position of the four primers (78R, 79R, 79F, and 80F) used for RT-PCR analysis has been indicated in *ssb3*. (b) PCR amplification of DNA fragments from *Anabaena* genomic DNA using the primer pairs 80F+79R, 79F+78R, and 80F+78R and Taq DNA Polymerase. The amplified products are indicated by arrows along with the product size. (c) Analysis of co-transcription of *ssb3* and its adjacent genes using 0.2 μg *Anabaena* total RNA as the template, primer sets as indicated and one-step reverse transcriptase. The amplified ~1 kb fragment is indicated by an arrow. The PCR products obtained in (b) and (c) were electrophoretically separated on 0.8% agarose gel stained with ethidium bromide and visualized using UV transilluminator. (d–f) DNA sequence of part of the putative promoter regions of the *Anabaena* (d) *ssb1*, (e) *ssb2*, and (f) *ssb3* genes designated as P_{ssb1}, P_{ssb2}, and P_{ssb3}, respectively. The predicted −10 and −35 regions are underlined. The predicted LexA-box-like element in P_{ssb1} and P_{ssb2} are indicated in bold and larger font and the A/T-rich tract in P_{ssb3} in bold. The probable FurA-binding site in P_{ssb3} is underlined with a dotted line. The start codon "ATG" is bold and italicized. The DNA molecular weight markers used were 100 bp DNA ladder (Bangalore Genie, India) in "b" and 1 kb DNA ladder (Bangalore Genei, India) in "c."

ssb2, the upstream gene, i.e., *all7578* (coding for unknown protein) is transcribed in the opposite direction, while the downstream gene, i.e., *alr7580* (coding for unknown protein) is transcribed in the same direction as *ssb2* (Fig. 1a). Thus, irrespective of whether *ssb2* and the gene(s) downstream to it are part of the same operon or not, *ssb2* will be transcribed from its immediate upstream region. However, in case of *ssb3* the flanking genes, i.e., *all4778* (coding for ABC transporter) and *all4780* (coding for probable monooxygenase) are transcribed in the same direction as *ssb3* (Fig. 1a), necessitating localization of the promoter regulating *ssb3* expression. Based on the protein coded by the flanking genes, genomic organization of *ssb3* in addition to that of *ssb1* and *ssb2* falls in the variant

category as defined earlier (Lindner et al. 2004), confirming their categorization in "Class IV" of bacterial *ssb* gene organization.

The transcription of all three *ssb* genes under control growth conditions in *Anabaena* 7120 was confirmed by RT-PCR (Fig. S1 in the Supporting Information), using total RNA as the template and gene-specific primers (Table 1). In order to identify the transcription unit of *ssb3*, two sets each of forward and reverse primers were designed. Of the two forward primers, 80F (Table 1) corresponded to a position internal to *all4780*, ~0.61 kb upstream from the translational start site of *ssb3* (Fig. 1A), and 79F (Table 1) corresponded to the start of *ssb3* gene (Fig. 1a; Table 1). Of the two reverse primers, 79R (Table 1) corresponded to the end of the *ssb3* gene ORF, while 78R (Table 1) was internal to the *all4778* gene, ~0.46 kb from the translational stop site of *ssb3*. All the three sets of primers tested, i.e., (i) 80F and 79 R, (ii) 79F and 78R, and (iii) 80F and 78R could successfully amplify the expected size DNA fragments, i.e., 1.2 kb, 1 kb, and 1.6 kb, respectively, from *Anabaena* genomic DNA (Fig. 1b). For RT-PCR analysis, the total RNA of *Anabaena* was amplified by one-step RT-PCR using the same three sets of primers. However, a product (~1 kb) was obtained only with the primer set 79F and 78R (Fig. 1c). This indicated that of the three genes, *ssb3* and *all4778* probably form a bicistronic operon, co-transcribed from *ssb3* promoter. Thus, all three *ssb* genes of *Anabaena* are regulated by their upstream regions.

Regulation of Anabaena ssb *genes by LexA and FurA.* A ~400-bp region upstream to each of the three *ssb* genes, considered as a putative promoter region was analyzed. The −10 and −35 regions in the upstream regions of the three *ssb* genes were identified by differential RNAseq methodology (Mitschke et al. 2011). The −35 region was not well defined, and a probable corresponding region was predicted for *ssb1* and *ssb3*, but not *ssb2* gene promoter (Fig. S2 in the Supporting Information), with a spacing of 17 bp and 15 bp from the predicted −10 region for P_{ssb1} (Fig. S2a) and P_{ssb3} (Fig. S2c), respectively. These, however, need to be confirmed by primer extension studies. A well-defined LexA-box-like element was detected in P_{ssb1} (Fig. 1d, Fig. S2a) agreeing with the consensus LexA box (Mazon et al. 2004). P_{ssb2}, however, did not a have a typical LexA box, but two overlapping regions with closely placed inverted repeat elements (Fig. 1e, Fig. S2b) could function as a potential LexA-box-like element. P_{ssb3} was found to be AT-rich and two regions of 20–25 bp bore homology to the consensus FurA box (Gonzalez et al. 2012), with the central 11 bp region being over 80% identical (Fig. S2c). Of these two, the most probable FurA box has been marked (Fig. 1f). Since AT-rich region may play a role in gene regulation, a smaller upstream region (154 bp) devoid of the AT-rich tract and designated as P_{ssb3s} was also generated.

The interaction of the three probable regulatory regions with *Anabaena* LexA and FurA proteins and their consequent regulation were analyzed by (i) EMSA of the three promoters with LexA and (ii) promoter activity in *E. coli* upon co-expression of either LexA or FurA. The mobility of P_{ssb1} and P_{ssb2}, but not P_{ssb3}, was retarded in the presence of LexA (Fig. 2a), in accordance with the presence of LexA box in P_{ssb1} and P_{ssb2}, but not P_{ssb3} (Fig. 1, d–f). A competition experiment was carried out between (i) P_{ssb1} and P_{ssb2} and (ii) P_{ssb1} and P_{ssb3}. LexA was incubated with either P_{ssb2} or P_{ssb3} for 30 min, followed by the addition of P_{ssb1} for another 30 min. EMSA analysis revealed that P_{ssb1} was able to partially displace P_{ssb2} (lane 8, Fig. S3 in the Supporting Information) from LexA–P_{ssb2} complex (lane 4, Fig. S3), indicating higher binding affinity of P_{ssb1} to LexA. In the presence of P_{ssb3s}, the retarded fragment (lane 9, Fig. S3) observed corresponded to that of P_{ssb1}–LexA complex (Fig. S3), indicating that P_{ssb3s} does not interfere with the binding. However, movement of both P_{ssb3} and P_{ssb3s} was retarded with 0.25 μM FurA, but did not affect that of P_{ssb1} and P_{ssb2} even at 5 μM concentration (Fig. 2b). To confirm that the binding of LexA or FurA to upstream regulatory sequences of *Anabaena ssb* genes was not merely due to the presence of repeat elements, EMSA was also carried out using a DNA fragment with several direct and inverted repeats, such as *Anabaena groESL* promoter (Rajaram and Apte 2010). Neither LexA nor FurA exhibited binding to the *groESL* promoter fragment (data not shown), thus confirming that the binding of LexA and FurA was not unspecific.

The expression of the SSB3 protein in *Anabaena* was assessed as a function of the availability of Fe^{2+} in the growth medium. Three-day-old *Anabaena* 7120 cultures were washed five times with 20-fold excess of Fe^{2+}-free, normal or 2× Fe^{2+}-containing BG-11 media and inoculated at 1 μg chl *a* density per mL in the respective media. The subculturing in the respective media was carried out after every 3 d. Protein was extracted from 3 d old cultures at the 1st and 3rd subculturing and probed for the expression of SSB3. The levels of the SSB3 protein increased after 1st subculturing under Fe^{2+}-depleted conditions (lane 1, Panel I, Fig. 2c), and a significant difference was observed after the 3rd subculture (lane 1, Panel II, Fig. 2c). However, the SSB3 protein levels decreased in the presence of excess Fe^{2+} in the growth media (lane 3, Panels I and II, Fig. 2c), compared with the control cultures (lane 2, Panels I and II, Fig. 2c).

Activity of the upstream regulatory regions of *ssb* genes (P_{ssb}) of *Anabaena* was analyzed fluorimetrically using *gfpmutII* as a reporter gene. The constructs used has been shown schematically in Figure S3a. In *E. coli* BL21(p*lysS*)(DE3) cells, the activity of P_{ssb3} was maximum, while P_{ssb2} exhibited the least activity (Table 3). Since the comparison of promoter activity has been carried out in the

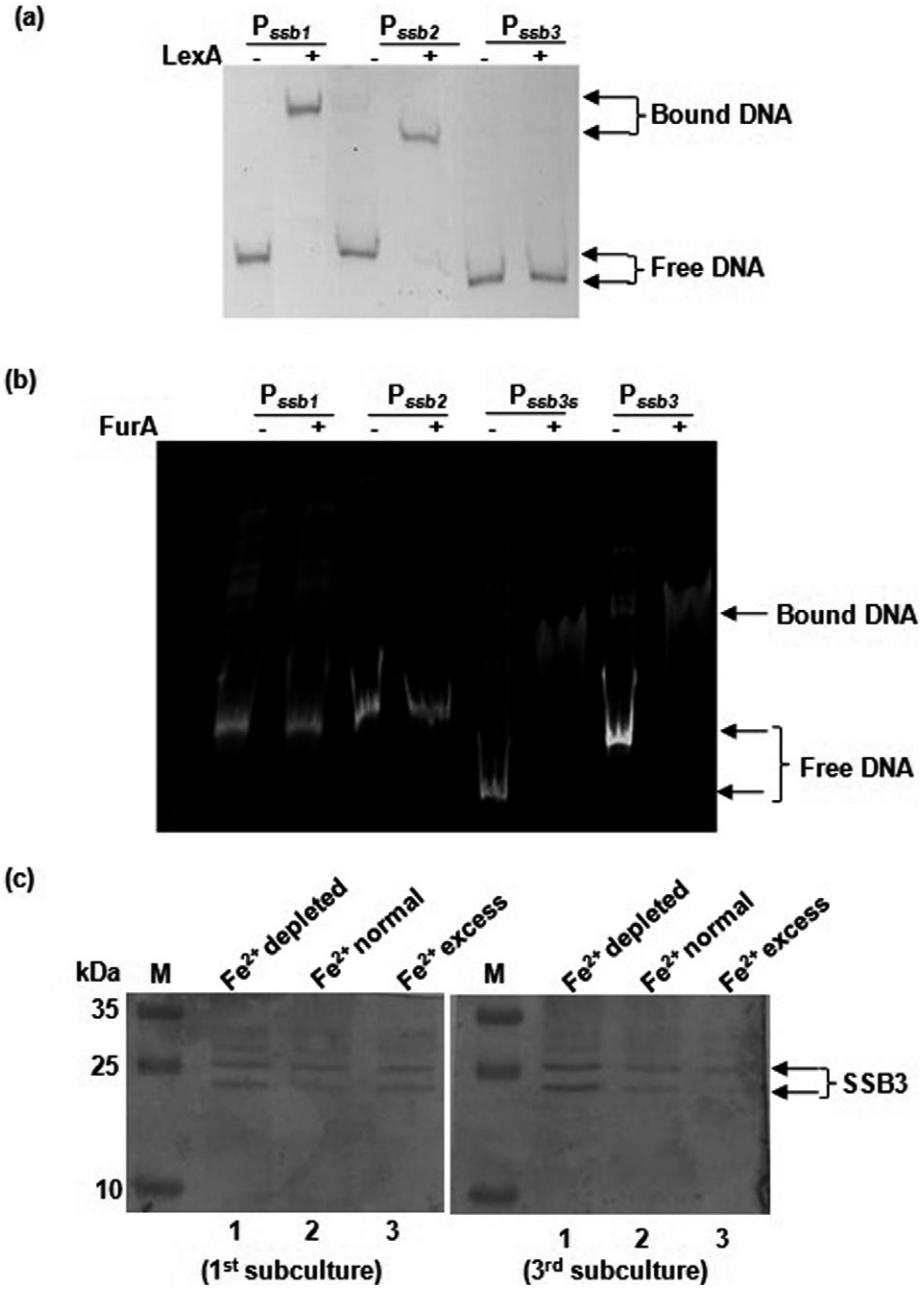


FIG. 2. Binding of regulatory proteins to *Anabaena ssb* upstream regulatory regions and expression of SSB3 as a function of Fe^{2+}-availability. (a and b) Electrophoretic mobility shift assay (EMSA) of different promoter regions of *Anabaena ssb* gene in the presence of (a) 0.5 μM *Anabaena* LexA protein and ~400 bp P_{ssb1} and P_{ssb2} and ~350 bp P_{ssb3} fragments, and (b) 2 μM FurA with P_{ssb1} and P_{ssb2} and 0.25 μM FurA with P_{ssb3} and ~150 bp P_{ssb3s}. The DNA–protein mix after interaction was separated electrophoretically on 8% polyacrylamide gel, stained with SYBR green and visualized using UV-transilluminator. The free and bound DNA fragments are indicated by arrows. (c) Proteins were extracted from 3 d old cultures of *Anabaena* PCC7120 grown under Fe^{2+}-depleted (lane 1), normal Fe^{2+} (lane 2) or 2-fold excess Fe^{2+} (lane 3) after the 1st subculture or 3rd subculture in the corresponding media. Electrophoretic separation was achieved by 12% SDS-PAGE, followed by electroblotting on to nitrocellulose membrane and immunodetection with anti-SSB3 antibody. The cross-reacting proteins (21–22 kDa) are indicated by arrows. Lane M corresponds to prestained protein marker (Bangalore Genei, India). The molecular mass of the proteins are indicated.

absence or presence of co-expressing, LexA or FurA, which is from pET16b-based vector, the control readings listed include pET16 as the second plasmid in the cells along with the promoter construct. The activity of only the promoter construct was similar in the absence or presence of pET16b (data not shown). The activity of the promoterless vector, pAM1956, was low and was not affected upon overexpression of either LexA or FurA (Table 3). The overexpression of LexA for 1 h resulted in 1.7-fold decrease in the activity of both P_{ssb1} and P_{ssb2}, but did not affect the activity of either P_{ssb3} or P_{ssb3s} (Table 3). The activity of P_{ssb1} and P_{ssb2} decreased 5-fold within 3 h of induction of LexA (data not shown). However, overexpression of FurA had no effect on the activity of either P_{ssb1} or P_{ssb2}, but decreased that of P_{ssb3} and P_{ssb3s} by 2.02- and 1.22-fold, respectively (Table 3).

N-status-dependent expression of ssb *promoters in* Anabaena *under different stress conditions.* Binding of FurA to P_{ssb3} suggests possible regulation of *ssb3* by the N-status of the growth media, since FurA is known to be up-regulated upon nitrogen-deprivation, through NtcA (Lopez-Gomollon et al. 2007). In order to assess this, comparative analysis of the activity of P_{ssb1}, P_{ssb2}, P_{ssb3}, and P_{ssb3s} in the absence or presence of combined nitrogen in growth media was carried out. Recombinant *Anabaena* strains

harboring one of the four promoter constructs or the empty vector were generated by introducing the corresponding plasmids into *Anabaena* by conjugation. AnpAM cells harboring the empty vector pAM1956 exhibited red fluorescence under microscope (Ex: 470 nm) due to the presence of chl *a* (Fig. S3b). Of the other recombinant *Anabaena* strains harboring the promoter constructs, only An (pAM-P_{ssb3}::*gfpmutII*) exhibited red fluorescence, while for the others, it ranged from orange to green (Fig. S3b), depending on the expression of the promoter transcribing the *gfpmutII* gene. The green fluorescence of GFP has to mask the red fluorescence of chl *a*, thus resulting in the differential coloration. Since in at least two instances in *Anabaena*, wherein proteins with similar functions were coded by two distinct genes, such as Hsp60 (Rajaram and Apte 2008) and SOD (Raghavan et al. 2011) were found to be differentially expressed in response to N-status, the expression of the three *ssb* gene promoters was also analyzed as a function of N-status.

A distinct difference in the activity of the promoters in *Anabaena* and *E. coli* was the lack of activity of P_{ssb3} in *Anabaena* (Figs. 3, 4 and S3) unlike maximal activity in *E. coli* (Table 3). The deletion of the AT-rich region in P_{ssb3s}, however, restored the promoter activity, indicating a role of the AT-rich region in *Anabaena* in regulating the expression of the downstream gene. A quantitative assessment of the activity was carried out as a function of N-status and in the absence or presence of DNA-damage-inducing stress conditions. Under control growth conditions, the activity of P_{ssb1} and P_{ssb2} was ~1.1- to 1.2-fold lower under N-supplemented conditions compared to nitrogen-fixing conditions, while that of P_{ssb3} remained low under both conditions (Figs. 3 and 4). However, the activity of P_{ssb3s} was 1.5- to 1.6-fold higher under nitrogen-supplemented conditions, compared to nitrogen-fixing conditions (Figs. 3, b and d; 4, b and d), indicating a direct or indirect regulation due to N-status.

Of the three DNA-damage-inducing stresses tested (i.e., mitomycin C, γ-radiation and desiccation), mitomycin C did not affect the expression of any of the promoters, irrespective of the absence or presence of combined nitrogen in growth medium (Figs. 3, a and b; 4, a and b). Exposure to 6 kGy of γ-radiation enhanced the expression of both P_{ssb2} and P_{ssb3s}, the increase for P_{ssb2} being 2.25- and 1.7-fold under nitrogen-fixing (N^-) and N-supplemented (N^+) conditions, respectively (Figs. 3a and 4a), while that of P_{ssb3s} increased 1.56- to 1.6-fold irrespective of the N-status (Figs. 3b and 4b). Desiccation stress for 6 d enhanced expression of only P_{ssb3s} by 1.5- and 1.4-fold under N^- and N^+ conditions, respectively (Figs. 3b and 4b). During postirradiation recovery (PIR), the activity of P_{ssb1} decreased by 1.63- and 2.2-fold on day 3 and day 5 under N^- conditions (Fig. 3c) and by 1.42- to 1.46-fold under N^+ conditions (Fig. 4c). A similar decrease was observed for P_{ssb2} activity as well, which on day 3 and day 5 of PIR decreased by 1.4- and 1.79-fold, respectively, under N^- (Fig. 3c) and by 1.47- and 2.29-fold under N^+ (Fig. 4c) conditions. However, the activity of P_{ssb3s} increased by 1.6- and 1.3-fold on day 3 and day 5, respectively, under N^--conditions (Fig. 3d) and 1.2- to 1.3-fold under N^+ conditions (Fig. 4d). During postdesiccation recovery (PDR), the activity of P_{ssb1} decreased by 1.4-fold under N^- conditions on day 3 (Fig. 3c), while it increased by 2.68- and 2.34-fold on day 3 under N^- (Fig. 3c) and N^+ (Fig. 4c), respectively, for P_{ssb2}, and by 1.25- and 1.34-fold on day 3 and day 5 under N^- conditions for P_{ssb3s} (Fig. 4d).

DISCUSSION

Several modes of regulation of *ssb* genes have been observed in bacteria, but the most common one is that by LexA, predominantly found in gram-negative bacteria, wherein the *ssb* gene is part of the SOS regulon. The presence of multiple promoters

TABLE 3. Activity of *Anabaena ssb* promoters in *Escherichia coli* BL21(p*lysS*)(DE3) cells.

Plasmids	pAM1956+ pET16b	pAM1956+ pET*lexA*	pAM1956+ pET*furA*
Uninduced	18.56 ± 2.32	16.35 ± 1.61	17.56 ± 2.11
1 mM IPTG, 1 h	20.63 ± 1.86	20.23 ± 2.36	19.89 ± 1.84
Plasmid(s)	pAM-P_{ssb1} + pET16b	pAM-P_{ssb1} + pET*lexA*	pAM-P_{ssb1} + pET*furA*
Uninduced	316.52 ± 29.31	322.32 ± 26.35	326.53 ± 28.31
1 mM IPTG, 1 h	320.23 ± 26.34	185.65 ± 16.53	305.65 ± 16.53
Plasmid(s)	pAM-P_{ssb2} + pET16b	pAM-P_{ssb2}+ pET*lexA*	pAM-P_{ssb2} + pET*furA*
Uninduced	106.53 ± 8.69	98.56 ± 9.12	102.68 ± 8.96
1 mM IPTG, 1 h	98.68 ± 8.43	56.85 ± 4.69	98.69 ± 7.89
Plasmid(s)	pAM-P_{ssb3} + pET16b	pAM-P_{ssb3}+ pET*lexA*	pAM-P_{ssb3} + pET*furA*
Uninduced	456.39 ± 38.29	473.56 ± 40.53	483.27 ± 40.12
1 mM IPTG, 1 h	460.35 ± 40.19	489.56 ± 50.21	239.56 ± 11.23
Plasmid(s)	pAM-P_{ssb3s} + pET16b	pAM-P_{ssb3s}+ pET*lexA*	pAM-P_{ssb3s} + pET*furA*
Uninduced	449.35 ± 49.35	488.31 ± 44.35	479.56 ± 44.39
1 mM IPTG, 1 h	458.15 ± 48.36	460.31 ± 50.31	390.31 ± 20.31

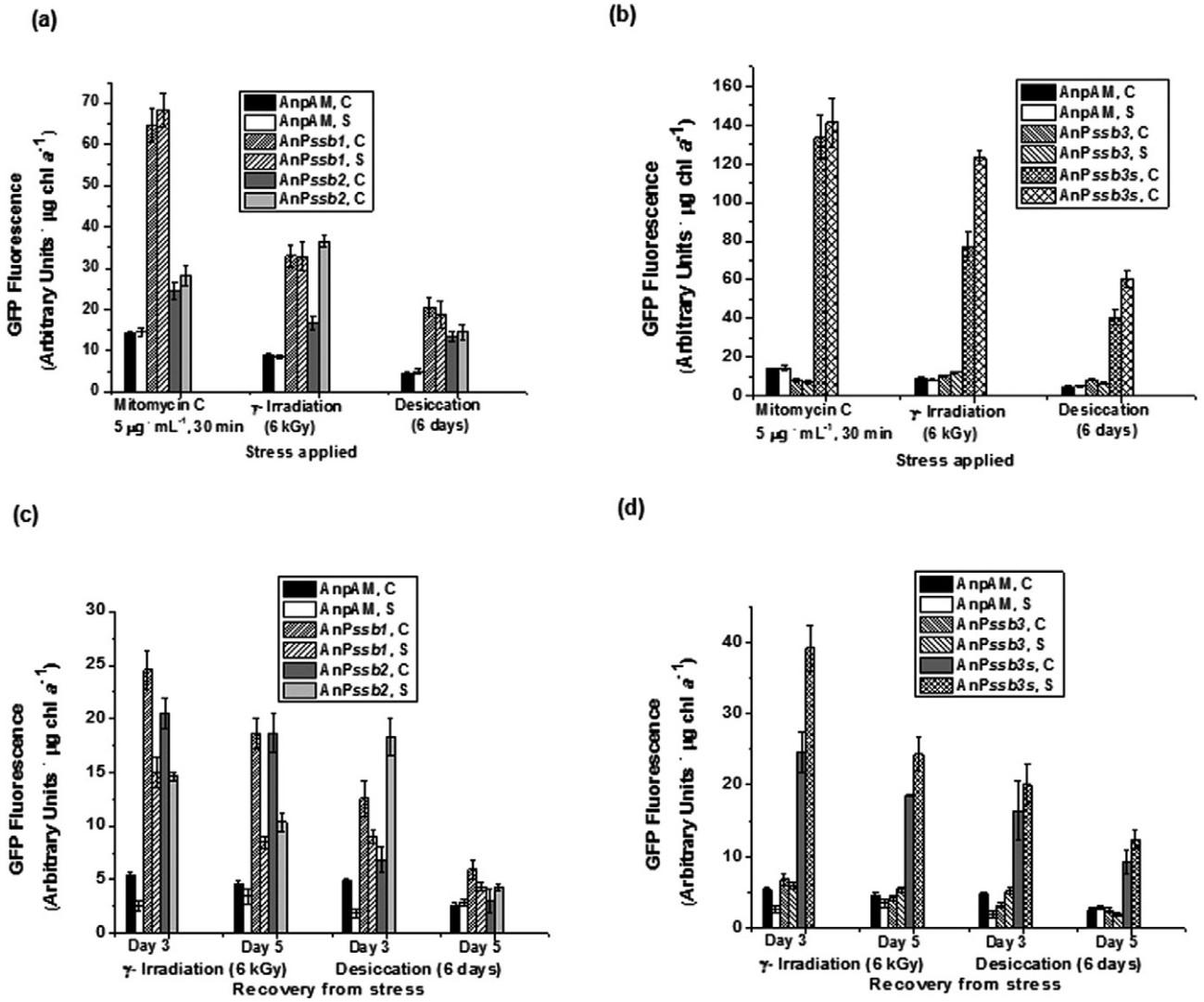


FIG. 3. Expression of *ssb* promoters in nitrogen-fixing cultures of *Anabaena*. Three-day-old recombinant *Anabaena* cultures harboring the empty vector, pAM1956, or one of the four *ssb* promoter constructs, as indicated, grown under nitrogen-fixing conditions were concentrated to 10 µg chl $a \cdot mL^{-1}$ and (a and b) exposed to mitomycin C (4 µg $\cdot$ mL^{-1}) for 30 min, 6 kGy of ^{60}Co γ-radiation or 6 d of desiccation. (c and d) Cells subjected to γ-irradiation or desiccation and their respective controls were washed and inoculated into fresh medium and allowed to recover. The activity of *Anabaena ssb* promoters was assessed immediately after the applied stress or corresponding control and during postirradiation or postdesiccation recovery. GFP Fluorescence was measured at 510 nm after excitation at 490 nm. Promoter activity was assessed as arbitrary GFP units per µg chl *a*. "S" in the graph inset refers to the stress applied (mitomycin C, γ-irradiation, or desiccation), and "C" to their corresponding control (control, sham-irradiated, or humid chamber).

for *ssb* in *E. coli* (Brandsma et al. 1983), with only one repressed by LexA (Brandsma et al. 1985), ensures basal-level expression of SSB, and at the same time enhances it upon exposure to DNA damage, upon derepression due to inactivation of LexA (Janion 2008). Of the three genes coding for SSB in the nitrogen-fixing cyanobacterium, *Anabaena* 7120, probable LexA-box-like elements, was predicted upstream of the truncated *ssb* genes, i.e., *alr0088* (*ssb1*) and *alr7579* (*ssb2*), while a possible FurA-binding region and AT-rich tract characterized the upstream region of *ssb3* (*all4779*; Fig. 1), which codes for the full-length SSB (Kirti et al. 2014). Involvement of LexA in regulating *ssb1* and *ssb2* gene promoters and FurA in *ssb3* gene promoter specifically was ascertained by (i) EMSA studies, wherein LexA bound specifically to *ssb1* and *ssb2*, but not *ssb3* promoter region (Fig. 2a), while the converse was true with FurA (Fig. 2b), and (ii) decrease in the activity of only *ssb1* and *ssb2* promoters (P_{ssb1} and P_{ssb2}) upon overexpression of LexA, and that of only *ssb3* promoters (P_{ssb3} and P_{ssb3s}) upon overexpression of FurA in *E. coli* (Table 3). In addition, overexpression of LexA in *Anabaena* also caused decrease in the transcript as well as protein levels corresponding to *ssb1* and *ssb2*, but not *ssb3* (A. Kumar, unpublished results), further confirming regulation of only *ssb1* and *ssb2* by LexA. The

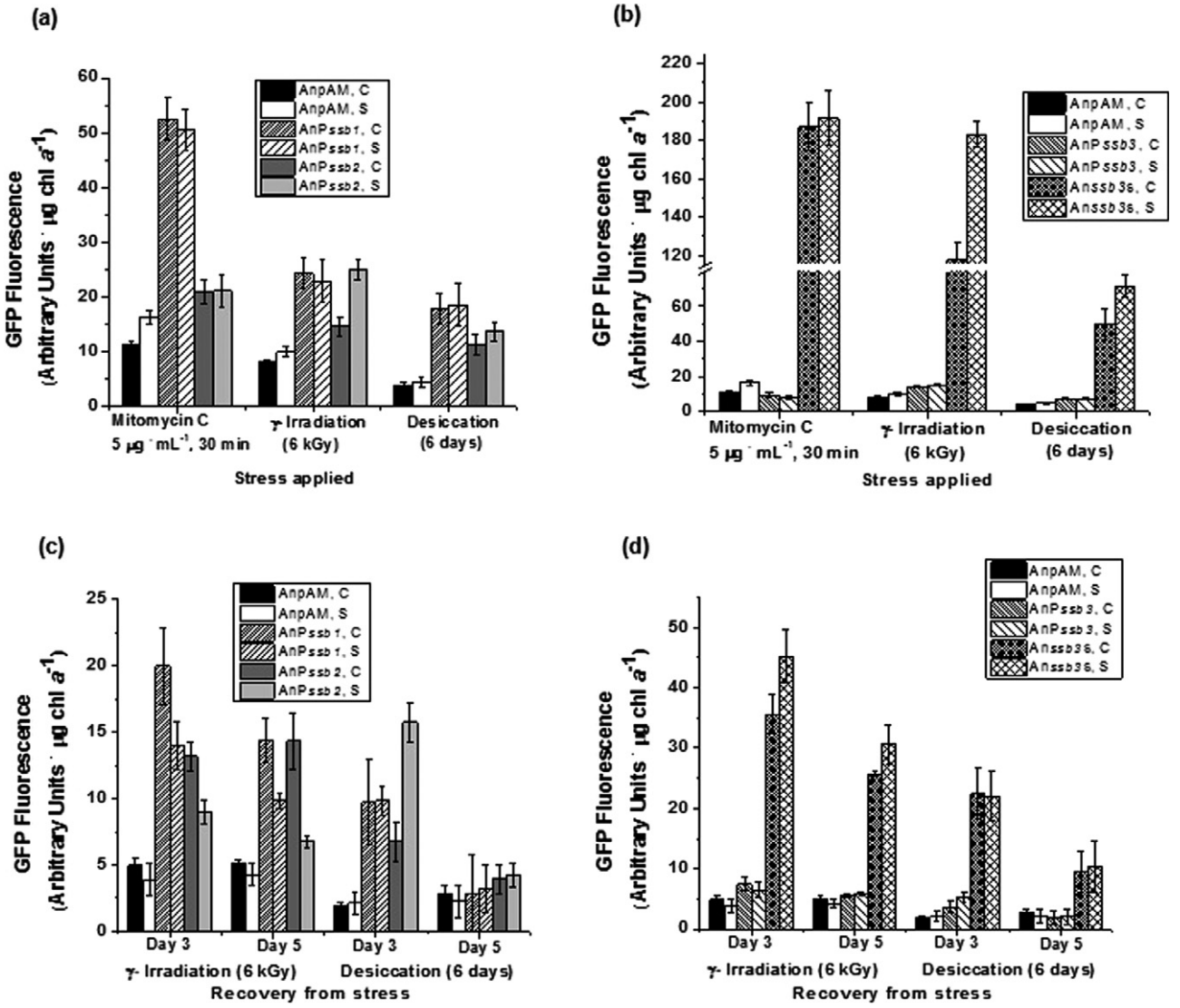


FIG. 4. Expression of *ssb* promoters in nitrogen-supplemented cultures of *Anabaena*. Different recombinant *Anabaena* cultures, as indicated, grown under nitrogen-supplemented conditions for 3 d were subjected to (a and b) mitomycin C, γ-radiation, or desiccation stress, or (b) allowed to recover from γ-radiation or desiccation stress, as described in Figure 4. Activity of the different promoters was assessed in terms of GFP fluorescence as described in Figure 3.

presence of *ssb2* on the beta-plasmid is suggestive of a horizontal transfer and could account for the lack of a consensus *Anabaena* LexA box in its upstream region.

In response to mitomycin C stress, which causes DNA adduct formation and activation of RecA in bacteria, no significant change in expression of any of the *ssb* promoters was observed (Figs. 3, a and b; 4, a and b). The absence of activated RecA-dependent cleavage of *Anabaena* LexA both *in vitro* and *in vivo* (Kumar et al. 2015) accounts for the lack of change in expression of P_{ssb1} and P_{ssb2}. The low levels of expression of LexA in *Anabaena* (Kumar et al. 2015), and the comparatively lower binding affinity of *Anabaena* LexA to the two promoters, as indicated by the higher concentration of the protein required for mobility shift (Fig. 2a), ensure that the two genes are not repressed in *Anabaena*, since both are essential based on the nonviability of their deletion mutants (Kirti et al. 2013). With the exception of two stress conditions, i.e., immediately after γ-irradiation and 3 d of PDR, the expression pattern of P_{ssb1} and P_{ssb2} was similar (Figs. 3, a and c; 4, a and c) varying only in fold-change, indicating that the regulation of the two genes by LexA follows a similar pattern. In case of the enhanced expression of P_{ssb2} observed in the two exceptional cases, it needs to be seen if certain other yet to be identified factors is involved. The repression of expression of P_{ssb1} during PIR (Figs. 3c and 4c) is necessitated for cell survival in line with the observed decrease in radiation tolerance of *Anabaena* upon

overexpression of SSB1 (Kirti et al. 2013). Since, no significant change in the level of LexA was observed during PIR in *Anabaena* (data not shown), the repression could be due to *in vivo* inactivation of LexA under these conditions, but this needs to be ascertained. The expression of the two promoters was not significantly affected by the absence or presence of combined nitrogen under control conditions, but the fold-change in expression was marginally higher under nitrogen-fixing conditions (Figs. 3, a and c; 4, a and c). This indicated that N-status does not affect the expression of *ssb1* and *ssb2* in *Anabaena*, which correlates well with N-status not having any effect on the expression of LexA either.

The absence of LexA-based regulation of *ssb3* of *Anabaena*, coding for the full-length SSB, is a deviation from what is prevalent in bacteria, with the exception of *D. radiodurans*, wherein it is regulated by a repeat element, RDRM (Ujaoney et al. 2010). The corresponding element is absent in *ssb3* promoter region, which is instead characterized by the presence of AT-rich tract and a FurA-binding-like region (Fig. 1f). However, neither SSB3 (All4779) nor any other protein involved in DNA repair was identified by a proteomic approach using FurA-overexpressing strain of *Anabaena* PCC7120 (Gonzalez et al. 2011). Though proteomic approach can help identify majority of the regulated proteins, a few may remain undetected due to low expression. The binding of FurA to the *ssb3* promoter (Fig. 2b) and its down-regulation in *E. coli* overexpressing FurA (Table 3) and up-regulation in the absence of the co-repressor Fe^{2+} in *Anabaena* (Fig. 2c) are suggestive of regulation of *ssb3* by FurA. Another interesting point is the effect of combined nitrogen availability on the activity of the *ssb3* promoter, which was found to be lower under nitrogen-fixing conditions by 1.5- to 1.6-fold compared to the nitrogen-supplemented conditions (Figs. 3, b and d; 4, b and d). This is supported by the observed up-regulation of the FurA repressor upon nitrogen deprivation, possibly through NtcA (Lopez-Gomollon et al. 2007). FurA acts as a typical repressor requiring Fe^{2+} as a co-repressor for binding effectively to the iron boxes (Pallares et al. 2014). Under Fe^{2+}-deficient conditions, the release of Fe^{2+} from FurA results in the dissociation of the FurA from the iron boxes thereby derepressing the expression of the downstream genes. The enhanced expression of SSB3 protein in *Anabaena* under Fe^{2+}-depleted conditions and decreased expression under Fe^{2+}-excess conditions (Fig. 2c) lend credence to the possible regulation by FurA. The repression of gene expression is also facilitated by DNA bending (Pallares et al. 2014), which can be induced by the presence of AT-rich tracts (Hizver et al. 2001), and could thus account for significantly lower promoter activity of P_{ssb3} (comprising of AT-rich regions) and P_{ssb3s} (devoid of AT-rich tract; Figs. 3, b and d; 4, b and d). The repression owing to the AT-rich tract has been observed only in the stand alone promoter, i.e., with the plasmid construct P_{ssb3}::*gfpmutII* in *Anabaena* (Figs. 3, b and d; 4, b and d; S3), but not in the genomic context, wherein the SSB3 (All4779) protein is expressed under control growth conditions (Kirti et al. 2014). The reason for this discrepancy is not known, but could involve the availability of the AT-rich tract for bending. Thus, in *Anabaena*, expression of the truncated *ssb* genes is regulated by the known SOS response regulator, LexA, while that of the full length *ssb* is most probably regulated by the Fe^{2+}-dependent regulator, FurA.

Brandsma, J. A., Bosch, D., Backendorf, C. & van de Putte, P. 1983. A common regulatory region shared by divergently transcribed genes of the *Escherichia coli* SOS system. *Nature* 305:243–5.

Brandsma, J. A., Bosch, D., de Ruyter, M. & van de Putte, P. 1985. Analysis of the regulatory region of the *ssb* gene of *Escherichia coli*. *Nucl. Acids Res.* 13:5095–109.

Castenholz, R. W. 1988. Culturing methods of cyanobacteria. *Methods Enzymol.* 167:68–93.

Gonzalez, A., Bes, M. T., Peleato, M. L. & Fillat, M. F. 2011. Unravelling the regulatory function of FurA in *Anabaena* sp. PCC7120 through 2-D DIGE proteomic analysis. *J. Proteomics* 74:660–71.

Gonzalez, A., Bes, M. T., Valladeras, A., Peleato, M. L. & Fillat, M. F. 2012. FurA is the master regulator of iron homeostasis and modulates the expression of tetrapyrrole biosynthesis genes in *Anabaena* sp. PCC7120. *Environ. Microbiol.* 14:3175–87.

Hizver, J., Rozenberg, H., Frolow, F., Rabinovich, D. & Shakked, Z. 2001. DNA bending by an adenine-thymine tract and its role in gene regulation. *Proc. Natl. Acad. Sci. USA* 98:8490–5.

Janion, C. 2008. Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int. J. Biol. Sci.* 4:338–44.

Kirti, A., Rajaram, H. & Apte, S. K. 2013. Characterization of two naturally truncated, Ssb-like proteins from the nitrogen-fixing cyanobacterium, *Anabaena* sp. PCC7120. *Photosyn. Res.* 118:147–54.

Kirti, A., Rajaram, H. & Apte, S. K. 2014. The hypothetical protein All4779 and not the annotated 'Alr0088' and 'Alr7579' protein is the major typical single-stranded DNA binding protein of the cyanobacterium, *Anabaena* sp. PCC7120. *PLoS ONE* 9:e93592.

Kumar, A., Kirti, A. & Rajaram, H. 2015. LexA protein of cyanobacterium Anabaena sp. strain PCC7120 exhibits in vitro pH-dependent and RecA-independent autoproteolytic activity. *Int. J. Biochem. Cell Biol.* 59:84–93.

Lindner, C., Nijland, R., van Hartskamp, M., Bron, S., Hamoen, L. W. & Kuipers, O. P. 2004. Differential expression of two paralogous genes of *Bacillus subtilis* encoding single-stranded DNA binding protein. *J. Bacteriol.* 186:1097–105.

Lopez-Gomollon, S., Hernandez, J. A., Wolk, C. P., Paleato, M. L. & Fillat, M. F. 2007. Expression of FurA is modulated by NtcA and strongly enhanced in heterocysts of *Anabaena* sp. PCC7120. *Microbiol.* 153:42–50.

Mackinney, G. 1941. Absorption of light by chlorophyll solutions. *J. Biol. Chem.* 140:315–22.

Mazon, G., Lucena, J. M., Campoy, S., Fernandez de Henestrosa, A. R.. Candau, P. & Barbe, J. 2004. LexA-binding sequences in Gram-positive and cyanobacteria are closely related. *Mol. Gen. Genet* 271:40–9.

Mitschke, J., Vioque, A., Haas, F., Hess, W. R. & Muro-Pastor, A. M. 2011. Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc. Natl. Acad. Sci. USA.* 108:20130–5.

Pallares, M. C., Marcuello, C., Botello-Morte, L., Gonzalez, A., Fillat, M. F. & Lostao, A. 2014. Sequential binding of FurA from *Anabaena* sp. PCC 7120 to iron boxes: exploring regulation at the nanoscale. *Biochim. Biophys. Acta* 1844:623–31.

Perrino, F. W., Rein, D. C., Bobst, A. M. & Meyer, R. R. 1987. The relative rate of synthesis and levels of single-stranded DNA binding protein during induction of SOS repair in *Escherichia coli*. *Mol. Gen. Genet.* 209:612–4.

Raghavan, P. S., Rajaram, H. & Apte, S. K. 2011. Nitrogen status dependent oxidative stress tolerance conferred by overexpression of MnSOD and FeSOD proteins in *Anabaena* sp. strain PCC7120. *Plant Mol. Biol.* 77:407–17.

Rajaram, H. & Apte, S. K. 2008. Nitrogen status and heat-stress-dependent differential expression of *cpn60* chaperonin gene influences thermotolerance in the cyanobacterium *Anabaena*. *Microbiol.* 154:317–25.

Rajaram, H. & Apte, S. K. 2010. Differential regulation of *groESL* operon expression in response to heat and light in *Anabaena*. *Arch. Microbiol.* 192:729–38.

Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. & Keck, J. L. 2008. SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* 43:289–318.

Shimamoto, N., Ikushima, N., Utiyama, H., Tachibana, H. & Horie, K. 1987. Specific and cooperative binding of *E. coli* single-stranded DNA binding protein to mRNA. *Nucl. Acids Res.* 15:5241–50.

Ujaoney, A. K., Potnis, A. A., Kane, P., Mukhopadhyaya, R. & Apte, S. K. 2010. Radiation desiccation response motif-like sequences are involved in transcriptional activation of the *Deinococcal ssb* gene by ionizing radiation but not by desiccation. *J. Bacteriol.* 192:5637–44.

Yoon, H. S. & Golden, J. W. 1998. Heterocyst pattern formation controlled by a diffusible peptide. *Science* 282:935–8.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. Detection of transcripts corresponding to *Anabaena* 7120 *ssb* genes. Total RNA was isolated from *Anabaena* 7120 cells grown under control conditions in nitrogen-fixing medium. Purified RNA (200 ng) was subjected to either one-step reverse transcriptase (RT)-PCR or Taq Polymerase (Taq)-based PCR using gene specific primers for (a) *ssb1* and *ssb2* and (b) *ssb3*. The positive control was genomic DNA PCR amplified using Taq DNA polymerase using the same set of primers and the negative control was PCR amplification of total RNA with Taq DNA polymerase. The amplified fragments are indicated by arrows along with the approximate molecular mass. Lane M corresponds to 100 bp DNA ladder (Bangalore Genei, India).

Figure S2. DNA sequence of the upstream regions of *Anabaena ssb* genes. DNA sequence of regions upstream of (A) *ssb1*, (B) *ssb2* and (C) *ssb3* genes. The probable promoter-like elements comprising of −35 and −10 regions were predicted on the basis of Mitschke et al. (2011). The nearest match to the −10 region are boxed and marked as −10 and the corresponding +1 base indicated. A typical −35 region could not be identified; the closest match with a reasonable spacing from −10 region is shown in gray box. The LexA-box-like elements are shown in bold in (a) and in bold enclosed in two overlapping boxes in (b). (c) The two possible FurA-binding box like are underlined with dotted lines and the most probable one is indicated. The start codon is shown in bold and italics. The positions of different primers used are indicated.

Figure S3. EMSA of LexA binding with multiple DNA fragments. The DNA fragments corresponding to upstream regulatory regions of *Anabaena ssb* genes, i.e., P_{ssb1} (lanes 1, 2, and 7), P_{ssb2} (lanes 3 and 4), and P_{ssb3s} (lanes 5 and 6) were incubated with 0.5 μM LexA (lanes 2, 4, and 6) for 30 min or 1 h (lane 7) and at 37°C. Lanes 1, 3, and 5 correspond to DNA fragments in the absence of proteins. Lanes 8 and 9 correspond to competition experiment, wherein LexA was incubated with either P_{ssb2} (lane 8) or P_{ssb3} (lane 9) for 30 min at 37°C, followed by the addition of P_{ssb1} and further incubation for 30 min. Electrophoretic separation was carried out on 6% polyacrylamide gel prepared in 0.5× TBE.

Figure S4. Generation of recombinant *Anabaena* stains harboring different promoter constructs. (a) Schematic representation of the promoter constructs, pAM-P_{ssb}::*gfpmutII*. The vector used for cloning, pAM1956 is shown as a black rectangle, the *ssb* gene promoters as white rectangles cloned at *Sac*I and *Kpn*I restriction sites, and the reporter gene, *gfpmutII* as a gray rectangle. (b) Fluorescence microphotograph (600× magnification) of recombinant *Anabaena* strains harboring either the empty vector, pAM1956 (AnpAM) or *ssb* promoter constructs namely pAM-P_{ssb1}::*gfpmutII* (AnP*ssb1*), pAM-P_{ssb2}::*gfpmutII* (AnP*ssb2*), pAM-P_{ssb3}::*gfpmutII* (AnP*ssb3*), and pAM-P_{ssb3s}::*gfpmutII* (AnP*ssb3s*). Fluorescence micrographs were obtained using Hg-Arc lamp (excitation 470 nm, emission 508 nm).

Table S1. Bacterial strains used in this study.

Functional & Integrative Genomics (2018) 18:357–367
https://doi.org/10.1007/s10142-018-0599-7

ORIGINAL ARTICLE

The SbcC and SbcD homologs of the cyanobacterium *Anabaena* sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair

Sarita Pandey[1] · Anurag Kirti[1] · Arvind Kumar[1,2] · Hema Rajaram[1,2]

Received: 15 January 2018 / Revised: 23 February 2018 / Accepted: 27 February 2018 / Published online: 9 March 2018

Abstract

The ubiquitous SbcCD exonuclease complex has been shown to perform an important role in DNA repair across prokaryotes and eukaryotes. However, they have remained uncharacterized in the ancient and stress-tolerant cyanobacteria. In the cyanobacterium *Anabaena* sp. strain PCC7120, SbcC and SbcD homologs, defined on the basis of the presence of corresponding functional domains, are annotated as hypothetical proteins, namely Alr3988 and All4463 respectively. Unlike the presence of *sbcC* and *sbcD* genes in a bicistronic operon in most organisms, these genes were distantly placed on the chromosome in *Anabaena*, and found to be negatively regulated by LexA. Both the genes were found to be essential in *Anabaena* as the individual deletion mutants were non-viable. On the other hand, the proteins could be individually overexpressed in *Anabaena* with no effect on normal cell physiology. However, they contributed positively to enhance the tolerance to different DNA damage-inducing stresses, such as mitomycin C and UV- and γ-radiation. This indicated that the two proteins, at least when overexpressed, could function independently and mitigate the damage caused due to the formation of DNA adducts and single- and double-strand breaks in *Anabaena*. This is the first report on possible independent in vivo functioning of SbcC and SbcD homologs in any bacteria, and the first effort to functionally characterize the proteins in any cyanobacteria.

Keywords Cyanobacteria · DNA damage · DNA repair · Radioresistance · SbcC · SbcD

Introduction

The photosynthetic, nitrogen-fixing filamentous cyanobacterium, *Anabaena* sp. strain PCC7120, hereafter referred to as *Anabaena* 7120, exhibits high tolerance to γ-radiation (LD_{50} of 6 kGy). This has been attributed to a robust DNA repair system (Singh et al. 2010, 2013). The mechanism of DNA repair is, however, not well defined in *Anabaena*, and in fact, in cyanobacteria as a whole. In *Anabaena* 7120, three single-stranded DNA-binding (SSB) proteins (Kirti et al. 2013, 2014) and a DNase, Alr3200 (Raghavan et al. 2016) have been shown to play a role in influencing its radiotolerance. Of the three SSBs, overexpression of the truncated SSB, Alr0088 (SSB1) decreased tolerance to γ-radiation, while that of the full length SSB, All4779 (SSB3) enhanced it (Kirti et al. 2013, 2014). Overexpression of Alr3200 resulted in increased DNA degradation post-irradiation resulting in extreme radiosensitivity (Raghavan et al. 2016). While SSB is central to several DNA repair pathways, the pathway involving Alr3200, a protein unique to cyanobacteria, is not known. Among the other known major bacterial DNA repair proteins, the expression of *recA* in *Anabaena* 7120 is low (Kumar et al. 2015), while homologs of RecB, RecC, SbcA, and SbcB proteins are absent in this organism. However, homologs of SbcC and SbcD proteins have been annotated in several cyanobacteria, including *Anabaena variabilis*, corresponding to gene numbers Ava_1709 (SbcC) and Ava_3331 (SbcD) (https://genome.microbedb.jp/cyanobase). The corresponding homologs in *Anabaena* 7120,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10142-018-0599-7) contains supplementary material, which is available to authorized users.

✉ Hema Rajaram
hemaraj@barc.gov.in

[1] Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India

[2] Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai 400094, India

namely Alr3988 and All4463 exhibiting over 95% homology with their counterparts in *A. variabilis*, are however annotated as hypothetical proteins. In general, SbcC and SbcD exhibit functions in conjunction with SbcA and RecBC, with mutations in *sbcC* and *sbcD* along with that in *sbcB* combine to co-suppress the defects in recombination, DNA repair, and cell viability associated with *recBC* mutations (Gibson et al. 1992). However, another radioresistant organism, *Deinococcus radiodurans* lack *recB*, *recC*, *sbcA*, and *sbcB* genes, but possess the *sbcDC* operon (Rocha et al. 2005). A correlation between radioresistance and lack of RecBC and SbcAB and presence of SbcCD has not been documented till date, to the best of our knowledge.

In bacteria, SbcC and SbcD proteins interact to form a large complex of 1.2 MDa (Connelly et al. 1997), wherein, the ATPase activity is associated with the SbcC subunit and nuclease activity with the SbcD subunit (Connelly and Leach 1996; Connelly et al. 1998). The SbcCD protein complex plays a positive role in (i) DNA replication (Connelly et al. 1998; Graumann and Knust 2009), (ii) repair of breaks made a result of restriction endonuclease activity (Cromie and Leach 2001), (iii) DNA inter-strand cross links induced by mitomycin C (Mit C) and (iv) double-strand breaks (DSBs) in *Bacillus subtilis* (Mascarenhas et al. 2006), and (v) DNA metabolism and repair in *Deinococcus radiodurans* (Hu et al. 2007; Kamble and Misra 2010). Deletion of *sbcC* and *sbcD* genes results in increased sensitivity to DNA damage (Mascarenhas et al. 2006; Hu et al. 2007; Thoms et al. 2008; Chen et al. 2007). In the case of *Staphylococcus aureus*, it has been reported that the *sbcD* and *sbcC* genes are involved in DNA repair as well as in repression of type 5 capsule production (Chen et al. 2007).

In bacteria, the *sbcC* and *sbcD* genes are usually present in the form of *sbcCD* operon as shown for *E. coli*, *B. subtilis*, *S. aureus*, and *D. radiodurans* (Connelly et al. 1997; Mascarenhas et al. 2006; Chen et al. 2007; Kamble and Misra 2010). However, the corresponding genes are located distantly on the chromosome in cyanobacteria, which includes *Anabaena* 7120. In bacteria, DNA repair genes constitute the SOS response and have been shown to be regulated by LexA (Janion 2008). The LexA repressor in the dimeric form binds to the LexA-box present in the proximity of the promoter, thereby repressing the transcription of the downstream gene. Upon DNA damage, the activated RecA induces the autoproteolytic cleavage activity of LexA, thereby derepressing the genes under its control (Janion 2008). However, in *E. coli*, the *sbcDC* operon does not form part of the SOS regulon (Darmon et al. 2007). This paper deals with evaluating the regulation and role of the two proteins in DNA repair, since their functional role has not been deciphered in cyanobacteria.

Materials and methods

Organism, growth conditions, and abiotic stresses

E. coli cells were grown in Luria–Bertani (LB) medium in the presence of appropriate antibiotics [50 μg kanamycin mL^{-1} (Kan_{50}), 100 μg carbenicillin mL^{-1}(Cb_{100}), or 34 μg chloramphenicol (Cm_{34})] at 37 °C and shaking at 150 rpm. Axenic cultures of *Anabaena* strains were grown in BG-11 liquid medium, pH 7.0 (Castenholz 1988) in the presence of 17 mM $NaNO_3$ (BG-11, N^+), under stationary conditions with continuous illumination (30 μE m^{-2} s^{-1}) at 27 °C ± 2 °C. Antibiotic neomycin was added at 12.5 μg mL^{-1} in liquid medium and 25 μg mL^{-1} on agar plates for the Neo^r recombinant *Anabaena* strains. Growth was assessed in terms of chlorophyll *a* (Chl *a*) content as described earlier (Mackinney 1941), and cell survival on BG-11, N^+ agar plates after incubation under illumination for 10 days.

Three-day-old cultures of *Anabaena* were concentrated to 10 μg Chl *a* mL^{-1} and subjected to (i) mitomycin C stress (0–5 μg mL^{-1}) for 30 min, (ii) UV-radiation stress (0–1.5 kJ m^{-2} s^{-1} at a dose rate 5 J m^{-2} s^{-1}), and (iii) 6 kGy of ^{60}Co γ-radiation stress (dose rate 1.8 kGy h^{-1}). Oxidative stress was applied to freshly inoculated *Anabaena* (3 μg Chl *a* mL^{-1}) with 0–0.7 mM H_2O_2. Tolerance to UV and mitomycin C stresses were analyzed by spot test analysis, wherein 100 μL of the concentrated cultures (unstressed or exposed to different concentrations of Mit C) were spread on to Millipore membrane filter (HAWP02500) placed on BG-11, N^+, Neo_{25} plates, followed by incubation under continuous illumination for 10 days. In case of UV-stress, the filters spotted with unstressed cultures were exposed to different doses of UV. Recovery of cells post-mitomycin C and post-irradiation (PIR) was also monitored by inoculating the stressed/unstressed cultures in fresh medium at 1 μg Chl *a* mL^{-1} and grown for a period of 7 days. All experiments were carried out using three technical and two biological replicates.

Generation of recombinant *Anabaena* strains

The ORF corresponding to *alr3988* and *all4463* were PCR amplified from the genomic DNA of *Anabaena* 7120 using gene-specific primers (Supplementary Table 1a). The 1.2-kb *Nde*I/*Bam*HI *all4463* fragment was ligated to pFPN vector (Chaurasia et al. 2008) resulting in the plasmid construct, pFPN*all4463* (Supplementary Table 1b). The *all4463* gene along with *psbA1* promoter was introduced into pAM1956 (Yoon and Golden 1998) as described earlier (Raghavan et al. 2011) to generate the construct pAM*all4463* (Supplementary Table 1b).

Since *alr3988* had an internal *Nde*I restriction site, a different approach was taken to generate the pAM*alr3988* construct. This has been schematically shown in Supplementary Fig. S1. In

brief, the *psbA1* promoter was PCR amplified from pFPN using specific primers (Supplementary Table 1a) and cloned in TA vector (Supplementary Table 1b). The ~0.2-kb *Sal*I/*Bam*HI P_{psbA1} fragment from the clone in TA vector was ligated to pBS plasmid (Supplementary Table 1b), to generate the construct pBS-P_{psbA1} (Supplementary Table 1b). The pBS-P_{psbA1} plasmid was digested with *Nco*I and *Bam*HI restriction enzymes and ligated with ~3-kb *alr3988* DNA fragment (complete ORF) digested with the same set of enzymes, to generate the construct, pBS-P_{psbA1}*alr3988* (Supplementary Table 1b). The ~3.2-kb *Sal*I/*Kpn*I fragment from pBS-P_{psbA1}*alr3988* was cloned into pAM1956 to generate the final construct pAM-P_{psbA1}*alr3988* (Supplementary Table 1b, Supplementary Fig. S1). The cloned *alr3988* and *all4463* genes have been sequenced from both the strands and the sequence submitted to GenBank (accession nos. MG356788 and KY212884 respectively).

The 474 and 138 bp region upstream of *alr3988* and *all4463* genes respectively were PCR amplified from *Anabaena* 7120 genomic DNA using the specified primers (Supplementary Table 1a), and designated as $P_{all4463}$ and $P_{alr3988}$ respectively. The ~0.47-kb $P_{alr3988}$ and ~0.14-kb $P_{all4463}$ fragments were digested with *Sac*I/*Kpn*I restriction enzymes and ligated to pAM1956 to generate the constructs, pAM-$P_{all4463}$ and pAM-$P_{alr3988}$ respectively (Supplementary Table 1b), wherein the promoters are transcriptionally fused to the downstream *gfpmutII* reporter gene present in pAM1956. All PCR-amplified fragments after cloning were sequenced from both the strands and the DNA sequence matched with that in the database (https://genome.microbedb.jp/cyanobase). All plasmids were maintained in *E. coli* cells, DH5α and HB101 (Supplementary Table 1c).

The four plasmids, pAM*all4463*, pAM- P_{psbA1}*alr3988*, pAM-$P_{all4463}$, and pAM-$P_{alr3988}$, were conjugated into *Anabaena* as described earlier (Raghavan et al. 2011) and the exconjugants selected on BG-11, N^+, Neo_{25} plates by repeated subculturing. The corresponding fully segregated recombinant *Anabaena* strains were designated as An*all4463*$^+$, An*alr3988*$^+$, An(pAM-$P_{all4463}$), and An(pAM-$P_{alr3988}$) respectively (Supplementary Table 1c).

Reverse transcriptase (RT)-PCR analysis

Total RNA from three-day-old *Anabaena* cultures was isolated using MACHEREY-NAGEL (MN) RNA isolation kit, as per the manufacturer's protocol. To rule out DNA contamination, PCR amplification of the isolated RNA was carried out using Taq DNA polymerase and corresponding gene primers (Supplementary Table 1a). The RNA was considered to be DNA-free only when no amplification was observed with Taq DNA polymerase, and thereafter used for reverse transcriptase (RT)-based analysis. Total RNA (20 ng) was used as template for carrying out reverse transcription using one-step RT-PCR kit (Qiagen, USA) as per the manufacturer's protocol. The 16 s rDNA primers (Supplementary Table 1a) were used as internal controls for all reactions.

Electrophoretic mobility shift assay

The promoter DNA fragments, $P_{alr3988}$ and $P_{all4463}$ (Supplementary Table 1a) (50 ng) were incubated with purified *Anabaena* LexA (0–180 nM) in binding buffer (20 mM Tris-HCl pH 8.0, 1 mM $MgCl_2$, 100 mM KCl, 8 mM DTT, 80 μg mL^{-1} BSA, 4% sucrose) at 37 °C for 30 min. Reaction was stopped by adding 10X DNA loading dye containing EDTA. Electrophoretic mobility shift assay (EMSA) of the reaction mix was carried out and visualized as described earlier (Kirti et al. 2017).

Analysis of promoter activity in terms of green fluorescent protein expression

The activity of $P_{alr3988}$ and $P_{all4463}$ promoters was quantitatively estimated under different growth conditions in *E. coli* and *Anabaena* cells harboring the plasmids pAM-$P_{alr3988}$ and pAM-$P_{all4463}$, in terms of green fluorescent protein (GFP) expression using JASCO FP6500 spectrofluorometer. The excitation and emission wavelengths used were 490 and 510 nm respectively. In case of *E. coli* cells, BL21(*plysS*) (DE3) cells harboring either of the plasmids pET16b or pET*lexA* (Supplementary Table 1B) were co-transformed individually with pAM-$P_{alr3988}$ and pAM-$P_{all4463}$ plasmids. O/N-grown cultures were inoculated in fresh medium and grown at 37 °C till $O.D._{600}$ of 0.1 was achieved, followed by the addition of 1 mM IPTG and continued incubation for about 2 h. GFP fluorescence was measured after every half hour and expressed as arbitrary units (A.U.) OD_{600}^{-1}. For *Anabaena* cells, strains AnpAM, An (pAM-$P_{alr3988}$), and An(pAM-$P_{all4463}$) were exposed to different stress as described in the results, and GFP fluorescence expressed as A.U. μg Chl a^{-1}.

Results and discussion

In the genome data base of *Anabaena* 7120 (http://genome.microbedb.jp/cyanobase/GCA_000009705.1), Alr3988 and All4463 are annotated as hypothetical proteins, though they exhibited 97 and 98% identity with the annotated SbcC (Ava_1709) and SbcD (Ava_3331) proteins of *A. variabilis* ATCC29413 respectively (Supplementary Fig. S2a, S2b). Analysis of the protein sequence for the presence for conserved domains using NCBI-CDD software (Marchler-Bauer et al. 2015) also revealed the presence of conserved domains corresponding to SbcC in Alr3988 (Supplementary Fig. S3a) and SbcD in All4463 (Supplementary Fig. S3b). This suggested that the two hypothetical proteins may correspond to SbcC and SbcD in *Anabaena* 7120. Since the proteins have

not been characterized in any cyanobacteria, antibodies against cyanobacterial SbcC and SbcD-like proteins were not available. Use of heterologous antibodies also did not aid in the detection of the corresponding proteins in *Anabaena* 7120 (data not shown). Hence, transcript analysis was carried out to assess if the genes are expressed in *Anabaena* 7120 using gene-specific (3988RT and 4463RT) primers (Supplementary Table S1a), and 16S rDNA as the internal control. Transcripts corresponding to *alr3988* and *all4463* were detected (Fig. 1a) in vector control AnpAM cells, indicating that the genes are expressed and hence, expected to be functional. The levels of *alr3988* and *all4463* were found to be relatively lower than that of 16S rDNA in *Anabaena* (Fig. 1a). Further investigation into the expression of these genes in *Anabaena* was carried out by analyzing the region immediately upstream of the corresponding genes.

Analysis of transcriptional regulatory regions of *alr3988* and *all4463*

The upstream 474 and 138 bp region of *alr3988* and *all4463* genes respectively (corresponding to the intergenic regions between them and their preceding gene) were analyzed for the presence of promoter-like regions, based on the earlier predicted promoter sequences for *Anabaena* 7120 (Mitschke et al. 2011) and possible regulatory boxes. A region similar to the predicted LexA-box of *Anabaena* (Mazón et al. 2004; Kirti et al. 2017) was detected upstream of *alr3988* (Fig. 1b) and *all4463* (Fig. 1c), 52 and 53 bases respectively upstream of the translational start site "ATG" (Fig. 1b, c). A probable −10 region (underlined) was also detected in the vicinity of the predicted LexA-box (boxed) (Fig. 1b, c). The predicted LexA-box for the two promoters is in good correlation with the sequence of the *Anabaena* LexA-box arrived at by consensus among more than 15 promoters regulated by it, and confirmed by mutational analysis followed by EMSA (Kumar et al., unpublished results). This suggested a possible regulation of the two genes by LexA. To carry out further anlaysis, the upstream regions of *alr3988* and *all4463*, hereafter designated as $P_{alr3988}$ and $P_{all4463}$, were individually cloned in promoterless vector pAM1956, and the resulting constructs were designated pAM-$P_{alr3988}$ and pAM-$P_{all4463}$ (Fig. 1d).

Upon incubation of the $P_{alr3988}$ and $P_{all4463}$ fragments with purified *Anabaena* LexA, a significant shift in their mobility

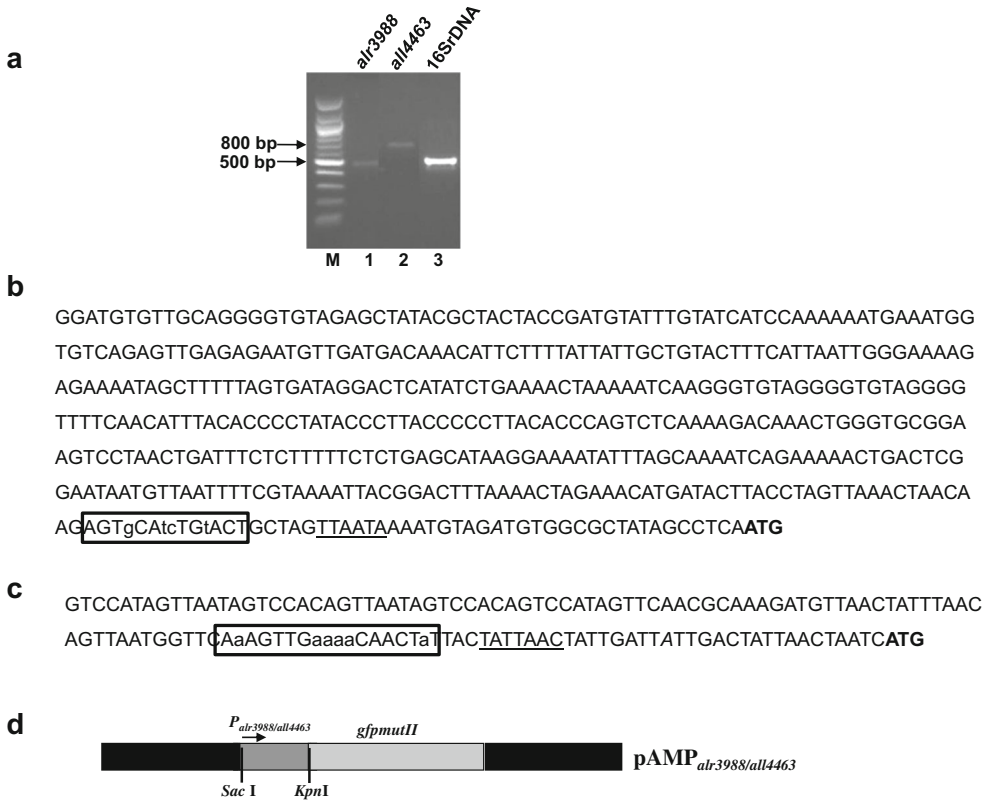


Fig. 1 Analysis of transcript and promoter region of *alr3988* and *all4463*. **a** Transcript levels of *alr3988* and *all4463* in AnpAM, analyzed by reverse transcriptase (RT)-PCR. Total RNA (DNA-free) isolated from AnpAM cells (20 ng for *alr3988* and *all4463* and 4 ng for 16SrDNA) was subject to RT-PCR using forward (Fwd) and reverse (Rev) primer sets corresponding to (lane 1) 3988RT, (lane 2) 4463RT, and (lane 3) 16SrDNA. Lane M corresponds to 100 bp DNA ladder. **b** and **c** DNA sequence of the upstream region of **b** *alr3988* ($P_{all3988}$) and **c** *all4463* ($P_{all4463}$). The probable LexA-like box is shown in a box, wherein the perfect palindromic base pair are in upper case and the mismatch sequence in small case. The probable −10 region is underlined. The translational start site "ATG" is in bold. **d** Schematic representation of the promoter construct pAM-$P_{alr3988/all4463}$

was observed with increasing concentrations of LexA (Fig. 2a, b). The binding affinity, based on the plot between % DNA bound and LexA concentration (Fig. 2c, d), was estimated as 66 ± 2.1 nM and 57 ± 1.3 nM respectively for $P_{alr3988}$ and $P_{all4463}$. To assess if the binding of LexA to the two promoters translates into regulation by LexA, the promoter constructs pAM-$P_{alr3988}$ and pAM-$P_{all4463}$ were transformed into *E. coli* BL21(p*lysS*) cells harboring pET*lexA*, wherein expression of *Anabaena* LexA in *E. coli* is induced by the addition of IPTG, and compared with that in BL21(pET16b) cells. The basal activity of $P_{alr3988}$ and $P_{all4463}$ in control *E. coli* BL21(pET16b) cells was found to be 4583.33 ± 286.32 and 4085.526 ± 253.62 A.U. OD_{700}^{-1} respectively irrespective of the absence or presence of IPTG. The promoter activity was similar in uninduced *E. coli* BL21(pET*lexA*) cells, but upon induction with 1 mM IPTG, the promoter activity decreased by 0.587 ± 0.02 and 0.551 ± 0.02-fold for $P_{alr3988}$ and $P_{all4463}$ respectively. Thus, *Anabaena* LexA acts as a repressor protein regulating the expression of *alr3988* and *all4463*.

The plasmids pAM-$P_{alr3988}$ and pAM-$P_{all4463}$ were also introduced into *Anabaena* to generate An(pAM-$P_{alr3988}$) and An(pAM-$P_{all4463}$) strains (Supplementary Table S1c) respectively. Both the promoters $P_{alr3988}$ and $P_{all4463}$ were found to be active in *Anabaena*, exhibiting comparable promoter activity (79.3 ± 8.8 and 60.2 ± 1.9 A.U. μg Chl a^{-1} for $P_{alr3988}$ and $P_{all4463}$ respectively), correlating well with the near equal transcript levels observed for the two genes (Fig. 1a). The activity of these promoters remained unaffected upon exposure to mitomycin C, H_2O_2 as well as γ-radiation stresses, with the exception of that of $P_{all4463}$ which increased 1.9-fold (to 114.4 ± 31.4 A.U. μg Chl a^{-1}) upon exposure to 6 kGy of γ-radiation. Levels of LexA in *Anabaena* remain unaffected by exposure to mitomycin C and UV-radiation (Kumar et al. 2015), but decreased upon exposure to γ-radiation (Kumar et al., unpublished results). Hence, the possibility of regulators in addition to LexA in *Anabaena* cannot be ruled out, as the two gene promoters show differential regulation under abiotic stress. Unlike in most other bacteria, wherein *sbcC* and *sbcD* form part of the bicistronic *sbcDC* operon (Makarova et al. 2001), the corresponding identified homologs in *Anabaena* (*alr3988* and *all4463*) and other cyanobacteria are distantly placed in the chromosome and under the regulation of their own promoters. Additionally, the *sbcDC* operon is a

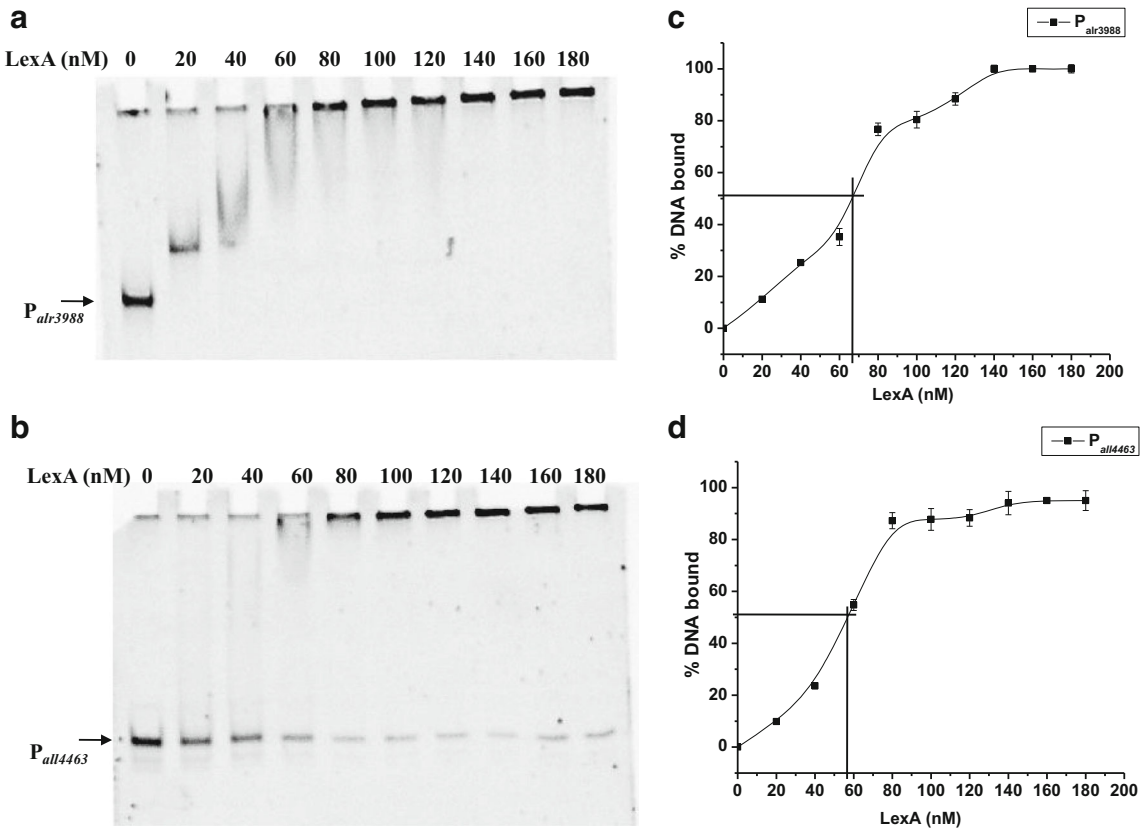


Fig. 2 Electrophoretic mobility shift assay (EMSA). **a** EMSA of different concentrations of *Anabaena* LexA protein (0–180 nM) with 50 ng of $P_{alr3988}$ and **b** EMSA of different concentrations of *Anabaena* LexA protein (0–180 nM) with 50 ng of $P_{all4463}$. The free DNA and the shifted DNA bands are indicated by arrows in the left and right respectively. **c** and **d** Graph depicting % bound DNA with increasing concentration of LexA over three replicates for $P_{alr3988}$ and $P_{all4463}$ respectively. The concentration of LexA at which 50% binding was observed is the binding affinity

constitutively expressed operon and does not form part of the LexA-controlled SOS regulon (Darmon et al. 2007), unlike the LexA-based regulation detected for the corresponding homologs in *Anabaena*. This stands out as a distinct feature of cyanobacterial *sbcC/sbcD*-like genes, and hence their functional role was investigated in *Anabaena* 7120.

Generation of recombinant *Anabaena* strains overexpressing Alr3988 and All4463 proteins individually

Functional studies on Alr3988 and All4463 in *Anabaena* were attempted by generating recombinant *Anabaena* strains with individual gene deletion mutants and/or by overexpressing the two proteins individually. The individual gene mutations in *alr3988* and *all4463* in *Anabaena* were found to be non-viable, indicating that the genes may be essential. This raised a question as to whether the two proteins (i) need to interact in *Anabaena* also, as observed for *E. coli* (Connelly et al. 1997) for efficient functioning, or (ii) have individual functions which are essential for normal cell growth. To answer these questions, recombinant *Anabaena* strains overexpressing the two proteins individually were generated. In the recombinant *Anabaena* strains, An*alr3988*$^+$ and An*all4463*$^+$ harboring the plasmids pAM*alr3988* and pAM*all4463* respectively, the two individual genes and *gfpmutII* were co-transcribed from the light-inducible *psbA1* promoter (Fig. 3a). The expression of the genes in the two strains was confirmed by the presence of green fluorescing filaments (indicative of expression of GFP from *gfpmutII* gene) of An*alr3988*$^+$ and An*all4463*$^+$ observed under the fluorescence microscope (excitation λ 470 nm) unlike Chl *a*-based red fluorescence of the vector control AnpAM (Rajaram and Apte 2010) cells (Fig. 3b). The orange fluorescence observed in some of the cells is due to inefficient masking of the chlorophyll fluorescence by GFP fluorescence.

Reverse transcriptase (RT)-PCR-based analysis revealed higher transcript levels of *alr3988* and *all4463* in An*alr3988*$^+$ and An*all4463*$^+$ respectively compared to that in AnpAM cells (Fig. 3c, d), thus confirming that both the genes are being transcribed in the two strains. The level of the internal control 16S

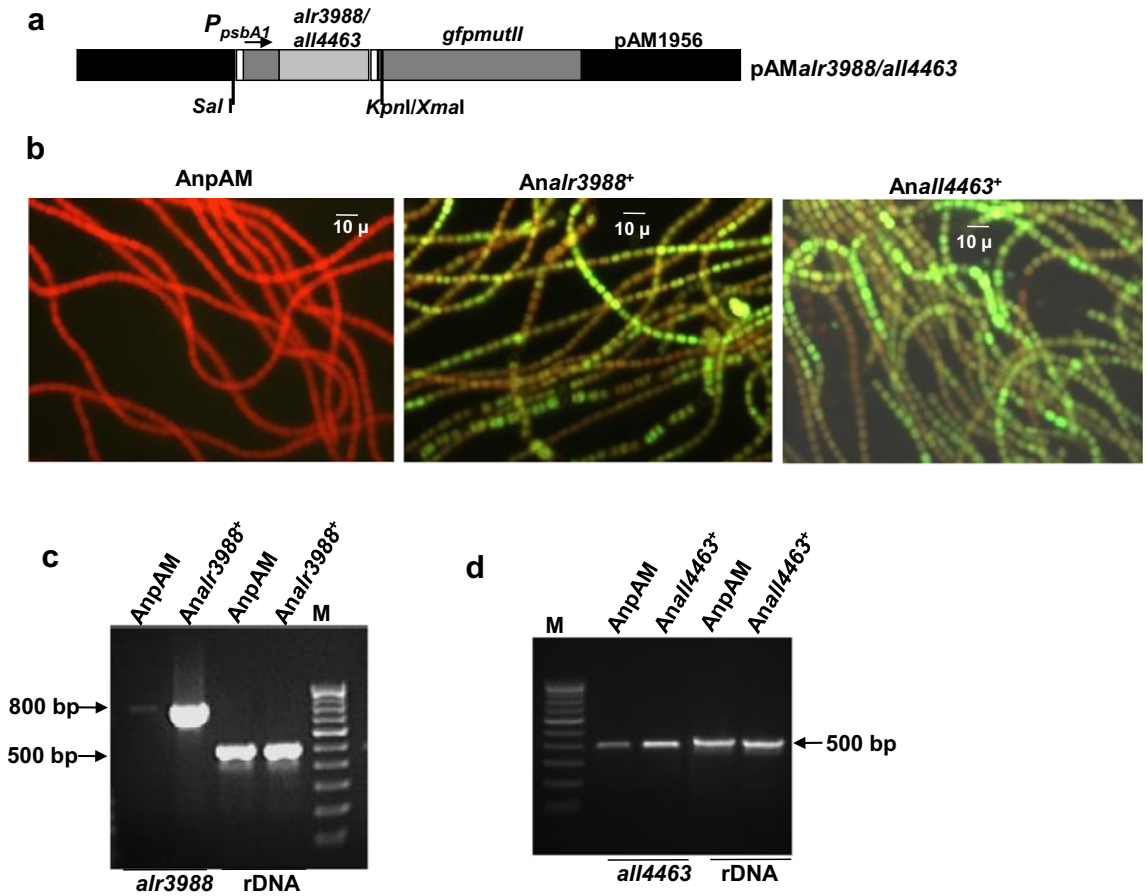


Fig. 3 Overexpression of Alr3988 and All4463 proteins individually in *Anabaena*. **a** Schematic diagram of the plasmid constructs, pAM(*alr3988*/*all4463*) used for overexpression of the corresponding proteins. The restriction endonuclease sites used are indicated. The P_{psbA1} promoter and *gfpmutII* genes are shaded in dark gray and the cloned gene in light gray. The basal vector used pAM1956 is shown in black. **b** Fluoromicrograph of AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ strains (×600 magnification). Red color indicates chlorophyll fluorescence and green fluorescence is due to GFP. **c** Transcript levels of the corresponding genes and rDNA in An*alr3988*$^+$ and An*all4463*$^+$ compared to AnpAM. The sizes of the partial transcripts are indicated. M refers to 100 bp DNA ladder, Bangalore Genei, India

rDNA was the same in all the RNA (Fig. 3c, d). Though the GFP levels were similar in An*alr3988*$^+$ and An*all4463*$^+$ cells, the increase in transcript level of *alr3988* in An*alr3988*$^+$ (Fig. 3c) was several folds higher than that of *all4463* in An*all4463*$^+$ cells (Fig. 3d). Since transcription is from the same promoter for the (*alr3988/all4463*)/*gfpmutII* transcriptional fusion, it is possible that *Anabaena* is regulating the levels of *all4463* mRNA by some other means for healthy growth of cells. The expression of *all4463* transcript in An*alr3988*$^+$ and *alr3988* transcript in An*all4463*$^+$ was similar to that in AnpAM cells (data not shown), and thus overexpression of one protein had no effect on the expression of the other protein. Expression of the corresponding proteins could not be confirmed by western blot analysis as specific antibodies against the individual proteins were not available and the available heterologous antibodies did not cross react with them (data not shown).

Alr3988 and All4463 play a role in protecting *Anabaena* from DNA damage

Since bacterial SbcC and SbcD proteins are known to be involved in DNA repair, the role of the corresponding *Anabaena* 7120 homologs, Alr3988 and All4463, in DNA repair was assessed by exposing the recombinant strains, An*alr3988*$^+$ and An*all4463*$^+$ to DNA damage-inducing stresses and measuring their survival/viability in comparison to that of the vector control recombinant *Anabaena* strain AnpAM (Rajaram and Apte 2010). The different stresses applied included exposure to (i) mitomycin C and UV-radiation which induce generation of DNA adducts (Sinha and Häder 2002; Bargonetti et al. 2010) and (ii) γ-irradiation and H_2O_2-mediated oxidative stress, which result in the generation of single and double-strand breaks (Bresler et al. 1984; He and Häder 2002).

Spot test analysis was carried out to assess response of AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ to mitomycin C (0–5 μg mL^{-1}, 30 min) and UV-B (0–1.5 kJ m^{-2}) radiation stresses. Of the three stains, An*all4463*$^+$ exhibited high survival even after exposure to 5 μg Mit C mL^{-1} cells, while only a few colonies were detected for An*alr3988*$^+$, and none for AnpAM cells (Fig. 4a). An*alr3988*$^+$ exhibited significant survival up to 3 μg Mit C mL^{-1}, while AnpAM showed low survival even at 2 μg Mit C mL^{-1} (Fig. 4a). When allowed to recover in liquid media after exposure to 5 μg Mit C mL^{-1} for 30 min, compared to its unstressed control, An*all4463*$^+$

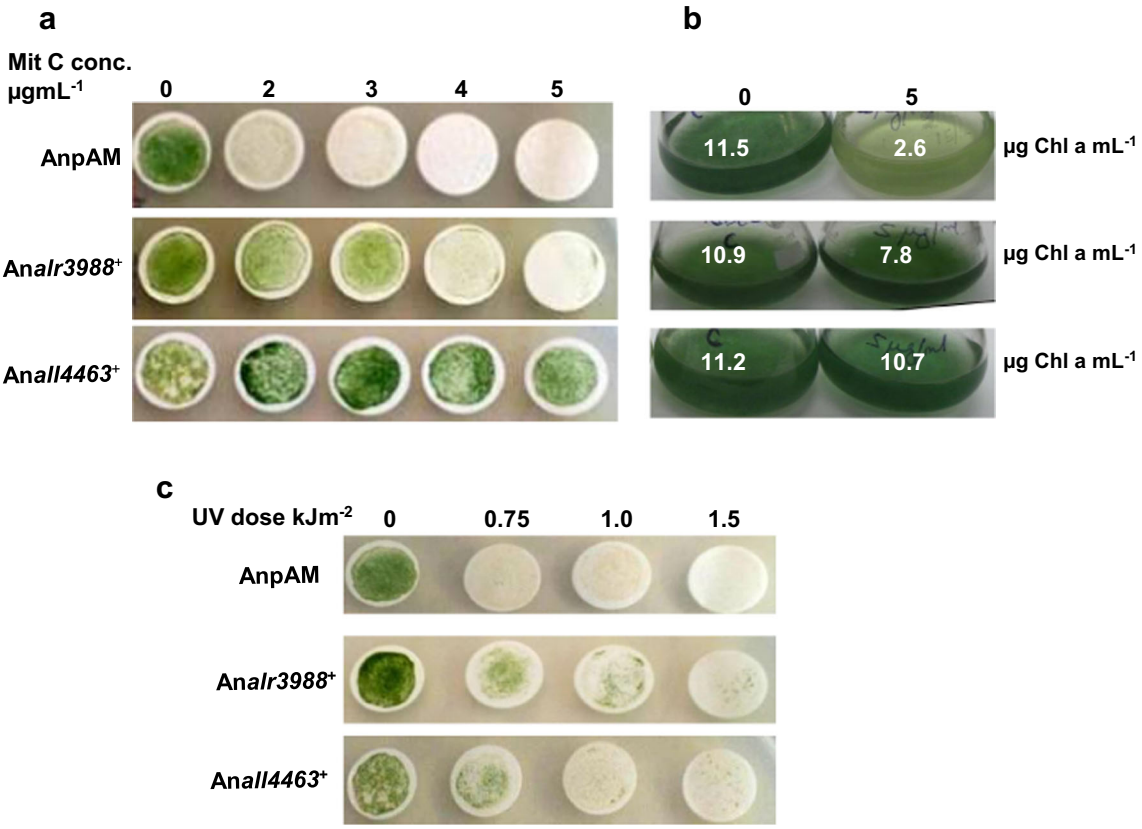


Fig. 4 Effect of overexpression of Alr3988 and All4463 on tolerance of *Anabaena* to mitomycin C (Mit C) and UV-radiation stresses. Three-day-old cultures of AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ were concentrated to 10 μg Chl *a* density mL^{-1} and exposed to **a** different concentration of mitomycin C (0–5 μg ml^{-1}) for 30 min or **c** different doses of UV (0–1.5 kJ m^{-2} at a dose rate of 5 J m^{-2} s^{-1}). Aliquots (100 μL) were spotted on the HAWP membrane and placed on BG-11, N^+, Neo_{25} plates and incubated under illumination for 10 days. **b** Recombinant *Anabaena* strains exposed to 0 and 5 μg Mit C mL^{-1} for 30 min were washed and inoculated in fresh BG-11, N^+, $Neo_{12.5}$ medium at 1 μg Chl *a* density ml^{-1} and grown for 7 days. Pictures have been taken using Nikon D5100 camera

exhibited 95.53 ± 1.31% recovery at the end of 7 days, as against 71.66 ± 1.09% for An*alr3988*$^{+}$ and 22.61 ± 0.14% for AnpAM cells (Fig. 4b), indicating maximum recovery upon overexpression of All4463.

When exposed to different doses of UV-radiation (0–1.5 kJ m^{-2}), An*alr3988*$^{+}$ and An*all4463*$^{+}$ showed higher survival than AnpAM at all the doses tested (Fig. 4c). The survival of An*alr3988*$^{+}$ was found to be higher than that of An*all4463*$^{+}$ at higher doses of UV-radiation (Fig. 4c), which could be due to the comparatively higher expression of Alr3988 in An*alr3988*$^{+}$ compared to that of All4463 in An*all4463*$^{+}$ cells based on transcript data (Fig. 3c, d). The increased tolerance of *Anabaena* upon overexpression of either Alr3988 or All4463 is in congruence with the observed increased mitomycin C sensitivity of the Deinococcal *sbcC* and *sbcD* mutants (Bentchikou et al. 2007) and *sbcC* mutant of *B. subtilis* (Mascarenhas et al. 2006), and

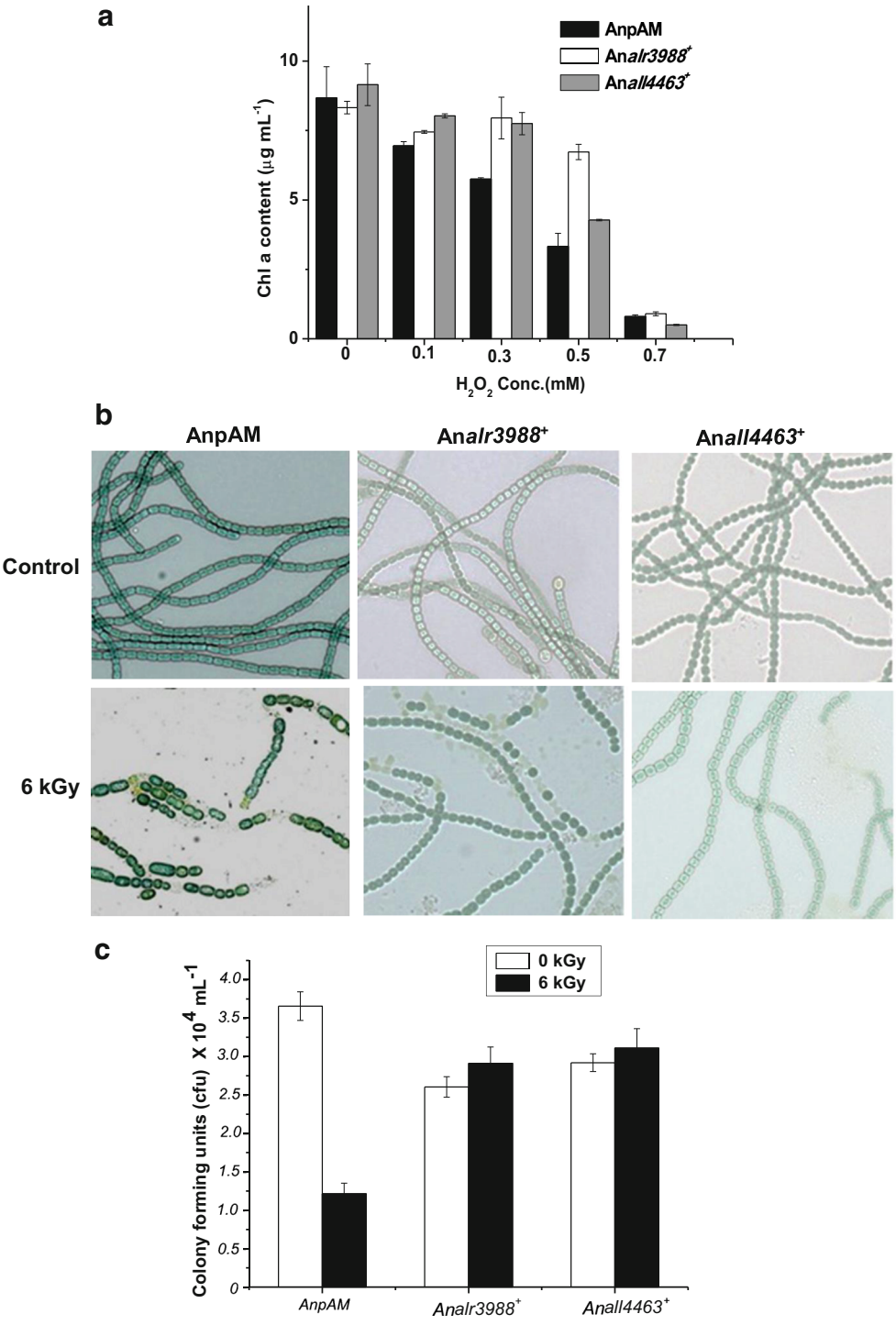


Fig. 5 Effect of overexpression of Alr3988 and All4463 on survival of H_2O_2 stress and γ-irradiated *Anabaena* cells. **a** Three-day-old cultures of AnpAM, An*alr3988*$^{+}$, and An*all4463*$^{+}$ were concentrated to 3 µg Chl *a* density mL^{-1} and exposed to different concentrations of H_2O_2 (0–0.7 mM), and growth measured in terms of Chl *a* content at the end of 3 days of H_2O_2 stress. **b** Three-day-old recombinant *Anabaena* cultures (AnpAM, An*alr3988*$^{+}$, and An*all4463*$^{+}$) were concentrated to 10 µg Chl *a* density mL^{-1} followed by exposure to 6 kGy of γ-irradiation. Bright-light microphotographs (×600 magnification) of the three strains under control conditions and immediately after irradiation "Irr." **c** Control (0 kGy) and irradiated (6 kGy) cultures were plated on to BG-11, N^{+}, Neo_{25} plates and colony forming units (cfu) counted after 10 days of incubation under illumination. All values pertain to mean of three technical and two biological replicates

increased UV sensitivity of *sbcCD* mutant of *D. radiodurans* and *E. coli* (Kamble and Misra 2010; Thoms et al. 2008). Thus, both Alr3988 and All4463 function in *Anabaena* to assist in the repair DNA inter-strand cross links either directly or indirectly, a role shown to be performed by SbcCD in other bacteria (Mascarenhas et al. 2006; Bentchikou et al. 2007; Thoms et al. 2008; Kamble and Misra 2010).

The recombinant *Anabaena* strains overexpressing Alr3988 and All4463 were also challenged with 0–0.7 mM H_2O_2 for 3 days or 6 kGy γ-radiation. The growth of these strains and the vector control recombinant *Anabaena* strain, AnpAM under control conditions, i.e., 0 mM H_2O_2 were similar (Fig. 5a). When exposed to different doses of H_2O_2, both An*alr3988*$^+$ and An*all4463*$^+$ showed better growth than AnpAM cells (Fig. 5a). Upon exposure to 0.5 mM H_2O_2, An*alr3988*$^+$ and An*all4463*$^+$ exhibited 80.8 and 46.7% growth compared to their respective controls as against 38.3% for AnpAM cells on day 3. Thus, compared to AnpAM cells, An*alr3988*$^+$ and An*all4463*$^+$ exhibited 2.03- and 1.3-fold higher growth on day 3 of H_2O_2 stress (Fig. 5a). Earlier reports, however, did not specify a role for SbcCD proteins in management of oxidative stress per se (Kamble and Misra 2010).

As seen in Fig. 5b, all three recombinant *Anabaena* strains, AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ formed long filaments under control growth conditions. However, immediately after exposure to 6 kGy of ^{60}Co γ-irradiation, AnpAM cells were found to be highly fragmented, while An*lar3988*$^+$ showed lower fragmentation, and An*all4463*$^+$ negligible fragmentation (Fig. 5b). This was reflected in the survival of the cells post-irradiation, with An*alr3988*$^+$ and An*all4463*$^+$ cells exhibiting survival comparable to their unstressed controls, as against only 33% survival for AnpAM cells (Fig. 5c). The sham-irradiated cultures of AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ showed similar growth recovery in terms of chlorophyll content (Fig. 6a). During post-irradiation recovery, measured in terms of Chl *a* content, the lag phase was found to be longer for AnpAM cells compared to An*alr3988*$^+$ and An*all4463*$^+$ (Fig. 6a). After 7 days of PIR compared to that of AnpAM cells, the growth of An*alr3988*$^+$ and An*all4463*$^+$ was

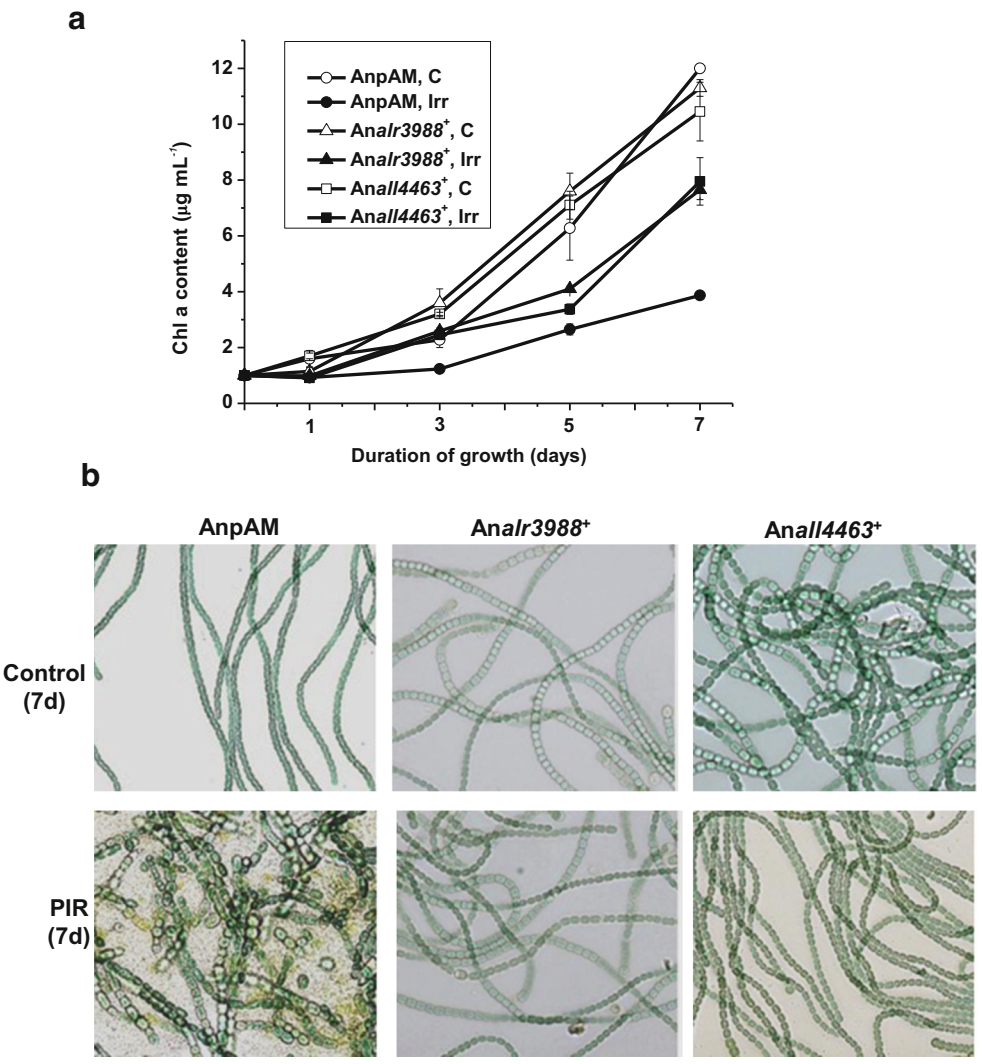


Fig. 6 Recovery of irradiated *Anabaena* cultures. Recombinant *Anabaena* strains, AnpAM, An*alr3988*$^+$, and An*all4463*$^+$ exposed to irradiation or corresponding control were washed and inoculated in fresh growth medium at 1 μg Chl *a* density mL^{-1} and grown under control conditions for 7 days. **a** Growth of the cultures measured in terms of Chl *a* content over a period of 7 days. **b** Bright-light microphotographs after 7 days of recovery of control and irradiated "PIR" cultures

found to be 2.0- and 2.1-fold higher respectively (Fig. 6a). The difference in the recovery was clearly visible in the morphological state of the filaments of the three strains. Even after 7 days of PIR, lot of debris was observed in case of AnpAM cultures and the filaments were also short (Fig. 6b). On the other hand, the filaments of An*alr3988*$^{+}$ and An*all4463*$^{+}$ after 7 days of PIR were found to be as healthy and long as that of the unstressed controls (Fig. 6b). Unlike the observed higher tolerance of An*alr3988*$^{+}$ compared to An*all4463*$^{+}$ upon exposure to Mit C, UV (Fig. 4), and H_2O_2 (Fig. 5a) stresses, the tolerance was similar for the two strains in response to γ-radiation stress. The observed increase in the promoter activity of *all4463* in response to γ-radiation would result in higher expression of the chromosomal *all4463* gene, which in combination with the overexpressed All4463 in An*all4463*$^{+}$ could be contributing to sufficient All4463 in these cells resulting in comparable tolerance to An*alr3988*$^{+}$, wherein transcription of *alr3988* is very high (Fig. 3c). In *Sulfolobous acidocaldarius*, MRE11 (ortholog of SbcD) has been shown to interact with Rad50 (SbcC homolog) and HerA bipolar helicase and play an active role in the repair of DNA damage induced by gamma rays (Quaiser et al. 2008). Likewise, the *sbcC/D* mutants of *D. radiodurans* and *B. subtilis* were found to be sensitive to γ-radiation (Mascarenhas et al. 2006; Bentchikou et al. 2007; Kamble and Misra 2010). Thus, functionally, Alr3988 and All4463 proteins of *Anabaena* also support the repair of γ-radiation-induced DNA damage as has been observed for bacterial SbcCD proteins.

Conclusion

The hypothetical proteins, Alr3988 and All4463, of *Anabaena* 7120 possess domains homologous to bacterial SbcC and SbcD proteins. When overexpressed individually, both the proteins enhance the ability of *Anabaena* to tolerate exposure to DNA damage-inducing agents. A striking difference, compared to other bacterial SbcCD systems, is that near equimolar expression of the two proteins simultaneously was not required for the participation of either Alr3988 or All4463 in enhancing the ability of *Anabaena* to recover from induced DNA damage. Since both the proteins are able to contribute directly or indirectly to the repair of DNA adducts as well as single and double-strand breaks, it is speculated that they may be involved in the same DNA repair pathway, but may not be functioning as a complex; however, this needs to be experimentally substantiated. Since individual gene mutants were not found to be viable, it indicates that Alr3988 and All4463 cannot fully compensate for each other, but as long as one of the two proteins is being expressed at basal levels, overexpression of the other enhances the capability of the DNA repair machinery of *Anabaena*. Based on the functional studies, it is proposed to annotate *alr3988* and *all4463* as *sbcC* and *sbcD* respectively in *Anabaena* 7120. This is the first report assigning a functional role for SbcC and SbcD-like proteins in cyanobacteria.

Funding information Sarita Pandey thanks SERB, New Delhi for Young Scientist Scheme (Start up research grant, YSS/2015/000306) for funding and MBD, BARC for providing research facilities.

Compliance with ethical standards

Conflict of interest statement The authors declare that they have no conflict of interest.

References

Bargonetti J, Champeil E, Tomasz M (2010) Differential toxicity of DNA adducts of mitomycin C. J Nucleic Acids 2010:698960. https://doi.org/10.4061/2010/698960

Bentchikou E, Servant P, Coste G, Sommer S (2007) Additive effects of SbcCD and PolX deficiencies in the in vivo repair of DNA double-strand breaks in *Deinococcus radiodurans*. J Bacteriol 189:4784–4790

Bresler SE, Noskin LA, Suslov AV (1984) Induction by gamma irradiation of double-strand breaks of *Escherichia coli* chromosomes and their role in cell lethality. Biophys J 45:749–754

Castenholz RW (1988) Culturing methods of cyanobacteria. Methods Enzymol 167:68–93

Chaurasia AK, Parasnis A, Apte SK (2008) An integrative expression vector for strain improvement and environment applications of nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC7120. J Microbiol Methods 73:133–141

Chen Z, Luong TT, Lee CY (2007) The sbcDC locus mediates repression of type 5 capsule production as part of the SOS response in *Staphylococcus aureus*. J Bacteriol 189:7343–7350

Connelly JC, deLeau ES, Okely EA, Leach DRF (1997) Overexpression, purification and characterization of the SbcCD protein from *Escherichia coli*. J Biol Chem 272:19819–19826

Connelly JC, Kirkham LA, Leach DRF (1998) The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. Proc Natl Acad Sci U S A 95:7969–7974

Connelly JC, Leach DR (1996) The sbcC and sbcD genes of *Escherichia coli* encode a nuclease involved in palindrome inviability and genetic recombination. Genes Cells 1:285–291

Cromie GA, Leach DR (2001) Recombinational repair of chromosomal DNA double-strand breaks generated by a restriction endonuclease. Mol Microbiol 41:873–883

Darmon E, Lopez-Vernaza MA, Helness AC, Borking A, Wilson E, Thacker Z, Wardrope L, Leach DRF (2007) SbcCD regulation and localization in *Escherichia coli*. J Bacteriol 189:6686–6694

Gibson FP, Leach DR, Lloyd RG (1992) Identification of *sbcD* mutations as cosuppressors of recBC that allow propagation of DNA palindromes in *Escherichia coli* K-12. J Bacteriol 174:1222–1228

Graumann PL, Knust T (2009) Dynamics of the bacterial SMC complex and SMC-like proteins involved in DNA repair. Chromosom Res 17:265–275

He YY, Häder D (2002) Reactive oxygen species and UV-B: effect on cyanobacteria. Photochem Photobiol Sci 1:729–736

Hu Y, Ma C, Tian B, Lin J, Hua Y (2007) Functional analysis of the *sbcD* (dr1291) gene of the extremely radioresistant bacterium *Deinococcus radiodurans*. Chin Sci Bull 52:2506–2513

Janion C (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. Int J Biol Sci 4:338–344

Kamble VA, Misra HS (2010) The SbcCD complex of *Deinococcus radiodurans* contributes to radioresistance and DNA strand break repair in vivo and exhibits Mre11–Rad50 type activity in vitro. DNA Repair 9:488–494

Kirti A, Kumar A, Rajaram H (2017) Differential regulation of ssb genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120. J Phycol 53:322–332

Kirti A, Rajaram H, Apte SK (2013) Characterization of two naturally truncated, Ssb-like proteins from the nitrogen-fixing cyanobacterium, *Anabaena* sp. PCC7120. Photosynth Res 118:147–154

Kirti A, Rajaram H, Apte SK (2014) The hypothetical protein 'All4779', and not the annotated 'Alr0088' and 'Alr7579' proteins, is the major typical single-stranded DNA binding protein of the cyanobacterium, *Anabaena* sp. PCC7120. PLoS One 9:e93592. https://doi.org/10.1371/journal.pone.0093592

Kumar A, Kirti A, Rajaram H (2015) LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits in vitro pH-dependent and RecA-independent autoproteolytic activity. Int J Biochem Cell Biol 59:84–93

Mackinney G (1941) Absorption of light by chlorophyll solutions. J Biol Chem 140:315–322

Makarova S, Aravind L, Wolf YI, Tatusov RL, Minton KW, Koonin EV, Daly MJ (2001) Genome of extremely radiation resistant bacterium, *Deinococcus radiodurans* viewed from the perspectives of comparative genomics. Microbiol Mol Biol Rev 65:45–79

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geeret RC et al (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:222–226

Mascarenhas J, Sanchez H, Tadesse S, Kidane D, Krisnamurthy M, Alonso JC, Graumann PL (2006) *Bacillus subtilis* SbcC protein plays an important role in DNA inter-strand cross-link repair. BMC Mol Biol 16:7–20

Mazón G, Lucena JM, Campoy S, Fernandez de Henestrosa ARF, Candau P, Barbé J (2004) LexA-binding sequences in gram positive and cyanobacteria are closely related. Mol Gen Genomics 271:40–49

Mitschke J, Vioque A, Haas F, Hess WR, Muro-Pastor AM (2011) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. Proc Natl Acad Sci U S A 108:20130–20135

Quaiser A, Constantinesco F, White MF, Forterre P, Elie C (2008) The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon *Sulfolobus acidocaldarius*. BMC Mol Biol 9:25. https://doi.org/10.1186/1471-2199-9-25

Raghavan PS, Gupta GD, Rajaram H, Kumar V (2016) Characterization of a DUF820 family protein Alr3200 of the cyanobacterium *Anabaena* sp. strain PCC7120. J Biosci 41:589–600

Raghavan PS, Rajaram H, Apte SK (2011) Nitrogen status dependent oxidative stress tolerance conferred by overexpression of MnSOD and FeSOD proteins in *Anabaena* sp. strain PCC7120. Plant Mol Biol 77:407–417

Rajaram H, Apte SK (2010) Differential regulation of *groESL* operon expression in response to heat and light in *Anabaena*. Arch Microbiol 192:729–738

Rocha EPC, Cornet E, Michel B (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. PLoS Genet 1(2):e15

Singh H, Anurag K, Apte SK (2013) High radiation and desiccation tolerance of nitrogen-fixing cultures of the cyanobacterium *Anabaena* sp. strain PCC7120 emanates from genome/proteome repair capabilities. Photosynth Res 118:71–81

Singh H, Fernandes T, Apte SK (2010) Unusual radioresistance of nitrogen fixing cultures of *Anabaena* strains. J Biosci 35:427–434

Sinha RP, Häder DP (2002) UV-induced DNA damage and repair: a review. Photochem Photobiol Sci 1:225–236

Thoms B, Borchers I, Wackernagel W (2008) Effects of single-strand DNases ExoI, RecJ, ExoVII, and SbcCD on homologous recombination of *recBCD*$^+$ strains of *Escherichia coli* and roles of SbcB15 and XonA2 ExoI mutant enzymes. J Bacteriol 190:179–192

Yoon HS, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. Science 282:935–938

BBA - Gene Regulatory Mechanisms 1861 (2018) 864–877

Contents lists available at ScienceDirect

BBA - Gene Regulatory Mechanisms

journal homepage: www.elsevier.com/locate/bbagrm

Regulation of multiple abiotic stress tolerance by LexA in the cyanobacterium *Anabaena* sp. strain PCC7120

Arvind Kumar[a,b], Anurag Kirti[a], Hema Rajaram[a,b,*]

[a] Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India
[b] Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Trombay, Mumbai 400094, India

ARTICLE INFO

Keywords:
Anabaena
LexA
Proteomics
Regulator
Stress tolerance

ABSTRACT

The paradigm of involvement of LexA in regulation of only SOS-response in bacteria through the down-regulation of DNA repair genes was challenged in the unicellular cyanobacterium, *Synechocystis* PCC6803, wherein it was originally shown not to be associated with DNA repair and later also involved in management of carbon-starvation through up-regulation of C-metabolism genes. In the filamentous cyanobacterium, *Anabaena* sp. strain PCC7120, global stress management role for LexA and a consensus LexA-binding box (AnLexA-box) has been established using a LexA-overexpressing recombinant strain, An*lexA*$^{+}$. High levels of LexA rendered *Anabaena* cells sensitive to different DNA damage and oxidative stress-inducing agents, through the transcriptional down-regulation of the genes involved in DNA repair and alleviation of oxidative stress. LexA overexpression enhanced the ability of *Anabaena* to tolerate C-depletion, induced by inhibiting photosynthesis, by up-regulating genes involved in C-fixation and down-regulating those involved in C-breakdown, while maintaining the overall photosynthetic efficiency. A consensus LexA-binding box, AnLexA-box [AGT-N_{4-11}-ACT] was identified upstream of both up- and down-regulated genes using a subset of *Anabaena* genes identified on the basis of proteomic analysis of An*lexA*$^{+}$ strain along with a few DNA repair genes. A short genome search revealed the presence of AnLexA box in at least 40 more genes, with functional roles in fatty acid biosynthesis, toxin-antitoxin systems in addition to DNA repair, oxidative stress, metal tolerance and C-metabolism. Thus, *Anabaena* LexA modulates the tolerance to multitude of stresses through transcriptional up/down-regulation of their functional genes directly by binding to the AnLexA Box present in their promoter region.

1. Introduction

In *Escherichia coli* wherein LexA has been extensively investigated, it regulates over 40 genes of the SOS-regulon [1], through its binding to the SOS (or LexA)-box present in the vicinity of the promoter region of the regulated gene [2]. In addition to the known palindromic SOS-box, a few unconventional target sites have also been identified [3]. Upon DNA damage, RecA gets activated and mediates the auto-proteolytic cleavage of LexA, resulting in de-repression of the downstream genes [4]. In addition to SOS response, LexA was shown to modify the cold shock response in the deep-sea bacterium, *Shewenella piezotolerance* [5], and contributed to the persistence of *Listeria monocytogenes* [6] and *Streptococcus mutans* [7]. LexA has also been associated with increased biofilm formation and decreased sporulation in *Clostridium difficile* [8].

In the unicellular cyanobacterium, *Synechocystis* sp. PCC6803 (hereafter referred to as *Synechocystis* 6803), the LexA-depleted mutant (i) exhibited decreased ability to cope with inorganic carbon limitation [9], and (ii) up- and down-regulation of transcripts of genes involved in photosynthesis and C-metabolism [10]. Involvement of LexA in the regulation of C-metabolism genes in cyanobacteria was shown for the first time in *Synechocystis* 6803, wherein the *hoxEFUYH* operon coding for the bidirectional hydrogenases was found to be regulated by it [11,12]. All the five genes formed a single transcript [12], with the transcriptional start site located 168 bp upstream of the *hoxE* ATG [11,12]. Two promoter-like regions were identified upstream of the operon, and one of the proteins binding to it was identified as LexA [11]. Using purified LexA protein from *Synechocystis*, it was shown that the protein indeed binds to the identified LexA box in the promoter region [12], thereby regulating the *hoxEFUYH* operon directly [13]. It was proposed that the LexA binds both the promoter regions which were distantly located and possibly interact through loop formation to regulate the *hox* operon [14]. However, the operon was not found to be exclusively regulated by LexA, but was also found to be regulated additionally by an AbrB-like protein [15]. Subsequently, the *hox* operon

* Corresponding author at: Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400085, India.
E-mail address: hemaraj@barc.gov.in (H. Rajaram).

https://doi.org/10.1016/j.bbagrm.2018.07.007
Received 2 May 2018; Received in revised form 15 July 2018; Accepted 19 July 2018
Available online 26 July 2018

[16] and *hoxW* gene were shown to be regulated by LexA even in *Anabaena* 7120 through binding to the site 'TAGTAGTTATG-TAATN$_{12}$-TAGCTT' [17], which showed resemblance to the LexA box motif RGTACN$_3$DGTWCB predicted by Mazon [18]. In the cyanobacterium, *Lyngbya majiscula*, LexA was shown to transcriptionally regulate the hydrogenase accessory genes, *hypFCDEB* [19]. Other than the regulation of C-metabolism genes in *Synechocystis* 6803, the involvement of LexA in the regulation of transcription of RNA helicase, CrhR has been extensively studied [20,21]. The *Synechocystis* 6803 LexA was shown to bind to a conserved sequence, CTA-N$_9$-CTA, present upstream of the *lexA* and *crhR* genes [20,21]. The *lexA* and *recA* genes were not found to be induced by DNA damage, and the *lexA* and *crhR* genes were divergently expressed i.e. under different conditions [20]. Subsequently, it was shown for the first time that LexA was not associated with regulation of DNA repair genes, but with the redox regulator CrhR in *Synechocystis* 6803 [20]. LexA was recently shown to be also associated with the regulation of fatty acid biosynthesis genes in *Synechocystis* PCC6803 [22]. LexA also regulated a few DNA repair genes, including *recA*, *ssb* in *Anabaena* 7120 [23,24]. This indicated diverse roles for LexA across cyanobacteria. Cyanobacterial LexA also differed from other bacterial LexA, which exhibit RecA-dependent cleavage [4], in being either non-cleavable as in case of *Synechocystis* 6803 [23], or exhibiting a pH-dependent and RecA-independent cleavage as observed for *Anabaena* 7120 [25]. In *Synechocystis* 6803, expression of LexA was found to be regulated both at transcriptional as well as at the post-transcriptional level [26]. Additionally, the LexA-box predicted for cyanobacteria varies across species and genes. Based on DNA footprinting and mobility shift assay studies, it varied from (i) RGTACN$_3$DGTWCB [18,23] and (ii) direct repeat CTA-N$_9$(AT-rich)-CTA, upstream of *lexA* and *crhR* genes of *Synechocystis* 6803 [21], and (iii) RRTACRNNYGTWYK in *Synechococcus* sp. WH7803 [27]. Whole genome search in *Anabaena* 7120 for RG(T/A)A(C/T) sites separated by 3 bases, revealed the presence of probable LexA-binding sites for 216 genes, either in the upstream region or within the gene [16]. This differed from the known *E. coli* LexA binding site, CTG-N$_{10}$-CAG present in the vicinity of a promoter region [2].

Cyanobacteria being very ancient microbes, the role of LexA may have been first defined in these organisms and then evolved to a more specific role in bacteria, such as *E. coli*, which evolved later. To address this issue as well as define a consensus LexA-binding box across all regulated genes, recombinant *Anabaena* strain overexpressing LexA (An*lexA*$^+$) was generated. Using a partial proteomic approach in combination with promoter, transcript and bioinformatic analysis, consensus LexA-box was defined. The role of LexA in regulating expression of different genes was found to reflect on growth studies carried out under different abiotic stress conditions, and suggested a role for LexA in oxidative stress, DNA damage-repair and C-homeostasis management.

2. Materials and methods

2.1. Generation of recombinant Anabaena strain

The 0.6 kb *Nde*I/*Bam*HI *lexA* gene of *Anabaena* 7120 from pET*lexA* [25] was cloned into pFPN [28] followed by cloning into pAM1956 [29] as shown in Supplementary Fig. S1. The resulting plasmid, pAM-*lexA* was introduced into *Anabaena* by triparental conjugation as described earlier [30] and detailed in Supplementary Methods, to generate the recombinant strain An*lexA*$^+$. The plasmids and strains used in the course of work are listed in Supplementary Table S1.

2.2. Growth conditions and abiotic stresses applied

Axenic cultures of AnpAM [31] and An*lexA*$^+$ were grown in BG-11 medium with 17 mM $NaNO_3$ (BG-11, N$^+$), pH 7.0 [32] with 10 μg neomycin (Neo) mL^{-1} at 27° ± 2°C under continuous illumination (30 μE m^{-2} s^{-1}) and stationary conditions (without any agitation or shaking). For experiments involving growth during stress, 3-day-old *Anabaena* cultures were inoculated in fresh medium at 1 μg chlorophyll *a* (Chl *a*) mL^{-1} and incubated under the specified conditions over a period of 3–8 days. Growth was measured in terms of Chl *a* content estimated in methanolic extracts as described earlier [33]. For survival experiments, stress was applied to three-day-old *Anabaena* cultures concentrated to 10 μg Chl *a* mL^{-1}. Survival after stress or their corresponding control was measured in terms of colony forming units (cfu) on BG-11, N$^+$, Neo$_{25}$ agar plates after 10 days of incubation under illumination.

The different stresses applied were (i) mitomycin C (0–5 μM) for 30 min, (ii) UV (0–1.5 kJ m^{-2}), dose rate 5 J m^{-2} s^{-1}, (iii) 6 kGy ^{60}Co γ-radiation, dose rate 4.5 kGy h^{-1}, and (iv) 6 days of desiccation as described earlier [16] to cultures concentrated to 10 μg Chl *a* mL^{-1}. Oxidative stress with (v) 0–0.5 mM H_2O_2, and (vi) 0.2 μM methyl viologen (MV), (vii) C-starvation with 5 μM DCMU [3-(3,4-dichlorophenyl)-1,1-dimethylurea], and heavy metal stress with (viii) 0–20 μM $CdCl_2$ and (ix) 0–50 mM $Na_2HAsO_4.7H_2O$ to cultures inoculated at 1 μg Chl *a* mL^{-1}. For experiments involving post-irradiation and post-desiccation recovery, the corresponding cultures were washed and inoculated in fresh medium at 1 μg Chl *a* mL^{-1} followed by incubation under normal growth conditions. All experiments were performed as three independent events, taking duplicates each time, and the data included in various figures is represented by the mean value and standard deviation of all these independent experiments put together.

2.3. Proteomic analysis of recombinant Anabaena strains

Protein samples for proteomic analysis were prepared by freeze-thaw method as described earlier [34], and DNA, RNA-free protein (1 mg) was then used for separation first on IEF carried out on Protean Isoelectric Focussing Cell (Bio-Rad, India), followed by 2nd dimensional resolution by 14% SDS-PAGE. The gel was imaged using Dyversity 6-gel imager (Syngene, UK), and match set generated using PD-Quest (version 8.1.0, Bio-Rad) with a minimum correlation of 0.6. Thirty spots from AnpAM and seventeen from An*lexA*$^+$ gels were excised and treated as described earlier [34] prior to analysis by MALDI-ToF/ToF (Matrix-assisted laser desorption/ionization-Time of flight) UltraFlexIII Mass spectrometer. The spectra were analysed using Flex Analysis software 3.0 (Bruker Daltonics). The mass spectra were imported into the database search engine (BioTools v 3.1 connected to Mascot, version 2.2.04, Matrix Science) and Mascot search were conducted from NCBI non-redundant database (released Jan 2012 or later) or SwissProt data base (released Jan 2012 or later) as described earlier [34]. The best score in each Mascot with a match with *Anabaena* 7120 protein was accepted as a successful identification. Protein identification was considered significant if at least two of the following three criteria were fulfilled: (i) Mascot score > 60 ($p < 0.05$), (ii) minimum match of 5 peptides, (iii) sequence coverage > 20%. The procedure followed has been described in more detail in Supplementary Methods.

2.4. Electrophoretic mobility shift assay (EMSA)

The DNA fragments (amplicons) corresponding to putative promoters of the specified genes were PCR amplified from the genomic DNA of *Anabaena* 7120, and have been listed in Supplementary Table S3A. The DNA sequence of all the amplicons were confirmed by DNA sequencing. EMSA using ~50 ng DNA and specified concentrations of purified *Anabaena* LexA protein was carried out as described earlier [24]. The gel was stained with SYBR Dye I and visualized under UV-transilluminator. The amount of free and bound DNA was determined densitometrically using ImageJ software [35]. Three independent experiments for each set of DNA-protein interaction was evaluated for plotting the graph between %DNA bound and concentration of LexA

used. The concentration of LexA resulting in 50% binding was defined as the binding affinity of LexA to that particular DNA fragment.

2.5. Fluorimetric analysis of promoter activity

The different amplicons (Supplementary Table S1A) were digested with *Sac*I/*Kpn*I and ligated to pAM1956 (Supplementary Table S1B) at the same restriction sites resulting in different plasmid constructs, designated as $pAMP_{abc}$ (wherein *abc* corresponds to the different genes) listed in Supplementary Table S1B. These were transformed into *E. coli* BL21(p*lysS*)(DE3) cells harbouring either pETl6b or pET*lexA* (Supplementary Table S1C). The overnight grown *E. coli* cultures (in LB medium with 35 μg chloramphenicol mL^{-1}, 100 μg carbenicillin mL^{-1} and 50 μg kanamycin mL^{-1} at 37 °C) were inoculated in fresh medium and grown till $O.D._{750}$ of 0.1 was reached. IPTG (1 mM) was added to all the cultures, and Green Fluorescent Protein (GFP) expression from the putative promoters was measured at time 0, 0.5 and 1 h after the addition of IPTG using a JASCO FP6500 spectrofluorometer (λexcitation: 490 nm and λemission: 510 nm). The values presented are average of three sets of technical replicates, each set comprising of two biological replicates.

2.6. Reverse transcriptase (RT)-PCR analysis

RNA was isolated from 3-day-old cultures of AnpAM and An*lexA*$^+$ (~50 μg Chl *a*) using MN (Macherey-Nagel) RNA Isolation kit, Germany as per the manufacturer's protocol. Total RNA was quantified using Nanodrop (NANODROP-2000C, Thermo Scientific). The presence of DNA contamination, if any, in the isolated RNA was checked by carrying out TaqDNA Polymerase-based PCR amplification using 16SrDNA. RNA was used for RT-PCR analysis only after ascertaining that there was no DNA contamination. RT-PCR from total RNA (10 ng) was carried out using specific gene primers (Supplementary Table S1A) using One-Step RT-PCR Kit, Qiagen, as per the manufacturer's protocol with annealing temperature varying from 45°C to 54°C depending on the gene to be amplified.

2.7. Western blotting and immunodetection and zymogram

Proteins were extracted from the recombinant *Anabaena* strains, electrophoretically separated on 15% SDS-Polyacrylamide gel, followed by blotting on to nitrocellulose membrane and immunodetection with anti-*Anabaena* KatB antibody [36] as described earlier [24]. Superoxide dismutase activity was carried out by *in-gel* zymogram assay as described earlier [30].

2.8. Measurement of ROS

The reactive oxygen species (ROS) content in 3-day-old recombinant *Anabaena* strains was measured using the fluorescent dye, Dichloro dihydrofluorescein diacetate (DCHFDA) [37]. *Anabaena* cells (1 μg Chl *a* mL^{-1}) were incubated with 10 μM DCHFDA for 20–25 min in dark at 25 °C with shaking, immediately followed by measurement of fluorescence (λ_{ex} 490 nm, λ_{em} 520 nm) using Multimode microplate reader, Model no. synergy H1 hybrid reader, Biotek. Staining of *Anabaena* cells with nitroblue tetrazolium (NBT) prepared in 50 mM Tris–HCl, pH 8.0 was carried out as described earlier [38].

3. Results and discussion

In cyanobacteria, the physiological role for LexA till date has been reported only for combating inorganic C-starvation [9], through the regulation of several bidirectional hydrogenases in different cyanobacterial species [11–17,19], and the lack of regulation of DNA repair genes in *Synechocystis* [20], while in most bacteria, such as *E. coli*, it is known to affect DNA repair through the SOS regulon [1,4]. Though, there are reports of LexA regulating a few DNA repair genes in *Anabaena*, such as *recA*, *uvrA*, *ssb* [18,22,23], and more recently *sbcC* and *sbcD*-like genes [39], the physiological role of LexA in response to DNA damage has not been shown. Cyanobacteria being an ancient organism, it is possible that the LexA might have had a different or wider regulatory role in terms of combating abiotic stresses. To evaluate the possible physiological implications of LexA in terms of tolerance to different abiotic stresses in cyanobacteria, *Anabaena* was chosen as the subject of study. To assess the possible global regulation by LexA in *Anabaena*, attempts were made to generate *lexA* insertional mutant and LexA overexpressing strains of *Anabaena*, using the *nptII* gene as a selection marker in both the cases. Though, several Neo^r colonies corresponding to An*lexA*$^-$ mutant were obtained, the extent of segregation of the mutated *lexA* gene in the filaments was very low. Even in *Synechocystis*, only a partially segregated *lexA* mutant was obtained [9], but this was found to be stable, possibly due to its unicellular nature. On the other hand, in the filamentous *Anabaena*, no two cultures corresponding to the partially segregated *lexA* mutant exhibited similar growth characteristics. Hence, further work was not continued in this recombinant strain. On the other hand, LexA overexpressing recombinant strain of *Anabaena* was found to be completely segregated and stable. The recombinant *Anabaena* strain overexpressing LexA, An*lexA*$^+$ was generated using the plasmid pAM*lexA*, wherein the *lexA* and *gfpmutII* genes are transcriptionally fused and under the control of light-inducible promoter of the *psbA1* (*alr4866*) gene (Supplementary Fig. S1A). An*lexA*$^+$ exhibited (i) uniform expression of GFP in all cells, as against only chlorophyll-based red fluorescence observed for the vector control, AnpAM cells (Supplementary Fig. S1B), and (ii) high levels of the 22 kDa LexA protein (Supplementary Fig. S1C).

3.1. Modulation of abiotic stress tolerance of Anabaena upon overexpression of LexA

Effect of LexA overexpression on tolerance of *Anabaena* to five different stresses, namely (i) C-starvation, exposure to (ii) DNA damage-inducing agents, (iii) H_2O_2, (iv) herbicide paraquat (methyl viologen, MV), and (v) heavy metals was investigated. The choice of C-starvation and DNA damage-inducing stresses was based on earlier reports of involvement of LexA in modulating tolerance to inorganic C-starvation [9], and lack of regulation of DNA repair genes [20,21] in *Synechocystis* as against that under DNA damage-inducing stresses in *E. coli* [2,4,40], and to understand if the role of LexA in *Anabaena* was different. Being a photosynthetic organism, the ability to cope with external reactive oxygen species (ROS)-inducing agent such as H_2O_2, and being a nitrogen-fixer, soil contaminants such as paraquat and heavy metals (Cd, As), which also induce oxidative stress were also included for analysis. All comparisons of stress tolerance were carried out between the vector control AnpAM [31], and LexA overexpressing An*lexA*$^+$ recombinant *Anabaena* strains.

3.1.1. Exposure to DCMU

Inhibition of photosynthesis was carried out by exposing the cells to DCMU, and not by keeping the cultures in dark, as that would have inhibited overexpression of LexA as well. When exposed to 5 μM DCMU, An*lexA*$^+$ cells appeared greener with a chlorophyll *a* content of 1.18 μg mL^{-1} as against 0.74 μg mL^{-1} for AnpAM cells (Fig. 1A), and exhibiting better growth than AnpAM (Fig. 1B). Upon exposure to DCMU, AnpAM cells showed a steady decline in chlorophyll content by day 3, form an initial value of 1 μg mL^{-1} to 0.52 ± 0.05 μg mL^{-1} indicating cell death, while that of An*lexA*$^+$ showed a marginal increase to 1.32 ± 0.15 μg mL^{-1} (Fig. 1B). Recovery of these cells after exposure to DCMU was carried by washing both the stressed and control cultures and inoculating them in fresh medium at 1 μg chl *a* mL^{-1} and grown for 7 days under control conditions. At the end of 7 days, the unstressed cultures of AnpAM and An*lexA*$^+$ showed similar recovery (12.5 ± 1.1 μg Chl *a* mL^{-1}), while the DCMU-exposed cells of AnpAM

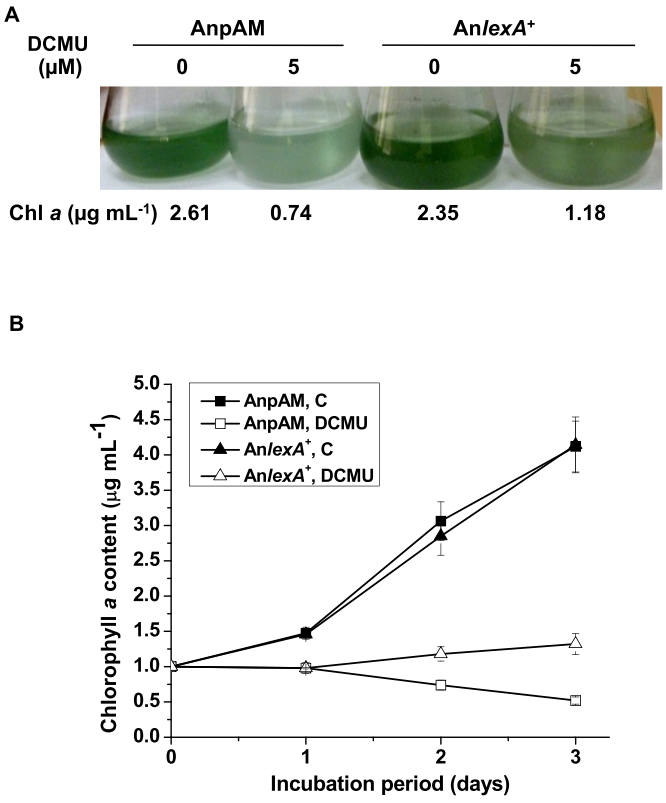


Fig. 1. Effect of LexA overexpression on tolerance of *Anabaena* to DCMU-induced carbon starvation stress.
Three-day-old cultures of recombinant *Anabaena* strains, AnpAM and An*lexA*$^+$ were inoculated in fresh BG-11, N^+ medium at 1 μg Chl*a* mL^{-1} and exposed to 5 μM DCMU. (A) Image of culture flasks on day 2. (B) Growth measured in terms of Chl *a* content upon exposure to 5 μM DCMU compared with that under control conditions.

recovered to 8.4 ± 0.4 μg Chl *a* mL^{-1} and An*lexA*$^+$ cells to 10.3 ± 0.6 μg Chl *a* mL^{-1}. This indicated that overexpression of LexA not only enabled the cells to overcome inhibition by DCMU, but also recover better when reverted to control growth conditions.

3.1.2. DNA damage-inducing stresses

Modulation of tolerance of *Anabaena* to different DNA damage-inducing stresses, such as mitomycin C, UV- and γ-radiation and desiccation, upon overexpression of LexA was assessed. While, An*lexA*$^+$ cells did not exhibit any significant change in tolerance to exposure to mitomycin C compared to AnpAM cells (Fig. 2A), the survival upon exposure to 0.75 kJ m^{-2} and 1.5 kJ m^{-2} of UV-radiation decreased to 20.4% and 4.1% respectively for An*lexA*$^+$ cells compared to 50.2% and 26.2% for AnpAM cells under corresponding stresses (Fig. 2B). Upon exposure to 6 kGy γ-radiation or 6 days of desiccation, both of which induce breaks in DNA, the survival decreased to 27.2% and 8.9% respectively for An*lexA*$^+$ as against 59% and 53.4% for AnpAM (Figs. 2C,D). An*lexA*$^+$ cells also showed lower post-stress recovery, which on day 8 was found to be 17.7% post-irradiation (Fig. 2E) and 40.6% post-desiccation (Fig. 2F), the corresponding values for AnpAM cells being 48.33% and 64.9% (Figs. 2E,F). This is in conformation with the observed impairment of DNA repair in *E. coli lexA* mutants [40].

3.1.3. H_2O_2-mediated oxidative stress

To assess if overexpression of LexA modulates the inherent and induced levels of ROS in *Anabaena*, the total ROS was estimated using DCHFDA. Under control growth conditions, the ROS levels in An*lexA*$^+$ cells were found to be 2.12 ± 0.11-fold higher compared to AnpAM cells, while upon exposure to 0.2 mM H_2O_2, the corresponding fold-difference was 1.64 ± 0.08. This was apparent even on NBT-staining of the recombinant *Anabaena* cells, wherein more number of black granules were observed in the cells of An*lexA*$^+$ compared to AnpAM cells both under control conditions and upon exposure to H_2O_2 stress (Fig. 3A), suggesting lower tolerance of An*lexA*$^+$ to oxidative stress. An*lexA*$^+$ cells showed visibly lower chlorophyll content upon exposure to different concentrations of H_2O_2, the difference becoming more apparent at 0.2 mM H_2O_2, and no growth detected in An*lexA*$^+$ at 0.3 mM H_2O_2 (Fig. 3B). Quantitatively assessed in terms of Chl *a* content, An*lexA*$^+$ cells exhibited 55% and 1.4% growth respectively compared to 97.6% and 76% for AnpAM cells when exposed to 0.2 mM and 0.3 mM of H_2O_2 for 3 days (Fig. 3C), indicating that LexA overexpression is rendering the cells sensitive to H_2O_2-mediated oxidative stress. Since, catalase is one of the important enzymes for detoxification of H_2O_2, the levels of KatB were compared in the two recombinant *Anabaena* strains. The levels of the 26 kDa KatB protein was found to be ~30% lower in An*lexA*$^+$ cells compared to that in AnpAM cells (Fig. 3D) suggesting repression of the expression of *Anabaena katB* by LexA.

3.1.4. Methyl viologen (MV) and heavy metal [Cd(II) and As(V)] stresses

Upon exposure to 0.2 μM MV, An*lexA*$^+$ cells exhibited no growth, in terms of Chl *a* content, as against 2-fold growth for AnpAM cells over a period of 3 days, which translated to 27.3% growth as against 48.5% for AnpAM cells compared with their respective unstressed controls (Fig. 4A). SOD zymogram analysis of 3-day-old control grown cultures revealed about negligible MnSOD (SodA) activity in An*lexA*$^+$ cells compared to AnpAM cells, while the FeSOD (SodB) activity remained unchanged (Fig. 4B), suggesting possible regulation of *sodA*, but not the *sodB* gene by LexA.

In the presence of 10 μM Cd(II), An*lexA*$^+$ cells exhibited 12% growth on day 3 as against 39% for AnpAM cells (Fig. 4C), while that at 10 mM As(V) was 18% and 29% respectively for An*lexA*$^+$ and AnpAM cells (Fig. 4D). The growth of An*lexA*$^+$ was consistently lower than that of AnpAM cells even at higher concentrations of As(V). This suggested possible negative regulation of heavy metal mitigating proteins by LexA, or could be due to the indirect effect of oxidative stress caused by the presence of heavy metals.

3.2. Gene regulation by LexA in Anabaena

To understand if the modulation in stress tolerance by overexpression of LexA in *Anabaena* was due to regulation of the corresponding stress-responsive genes by LexA, a combined approach involving proteomic, transcript and gene promoter analysis was taken.

3.2.1. Proteomic changes in Anabaena upon over-expression of LexA

Comparative proteomics of An*lexA*$^+$ and AnpAM cells under control growth conditions revealed distinct differences (Supplementary Fig. S2). PD Quest 2D analysis software detected 119 spots which were common to both the strains in all three gel replicates. Of these, 30 spots showed decreased abundance (fold change ≤ 1.2) (indicated with red arrow heads in Supplementary Fig. S2A), and 17 with increased abundance (fold change ≥ 1.2) (indicated with green arrow heads in Supplementary Fig. 2B) upon overexpression of LexA. The proteins showing no change in levels and identified using earlier data are indicated by black arrow heads (Supplementary Fig. S2B). The average fold-change of the 30 protein spots, across all triplicates, along with statistical analysis i.e. standard deviation and *p*-value is given in Supplementary Table S2. Some of the protein spots could be detected only in one of the two gels, i.e. either AnpAM or An*lexA*$^+$, and hence have been designated as not-detected in the other strain. Of these, 15 and 8 spots respectively could be identified conclusively (Supplementary Table S2, Supplementary MS data), and have been categorized on the basis of their functions in Fig. 5. Several of them belonged to the functional category of photosynthesis and C-metabolism. Phycocyanin (Cpc), phycoerythrins (Pec), allophycocyanin

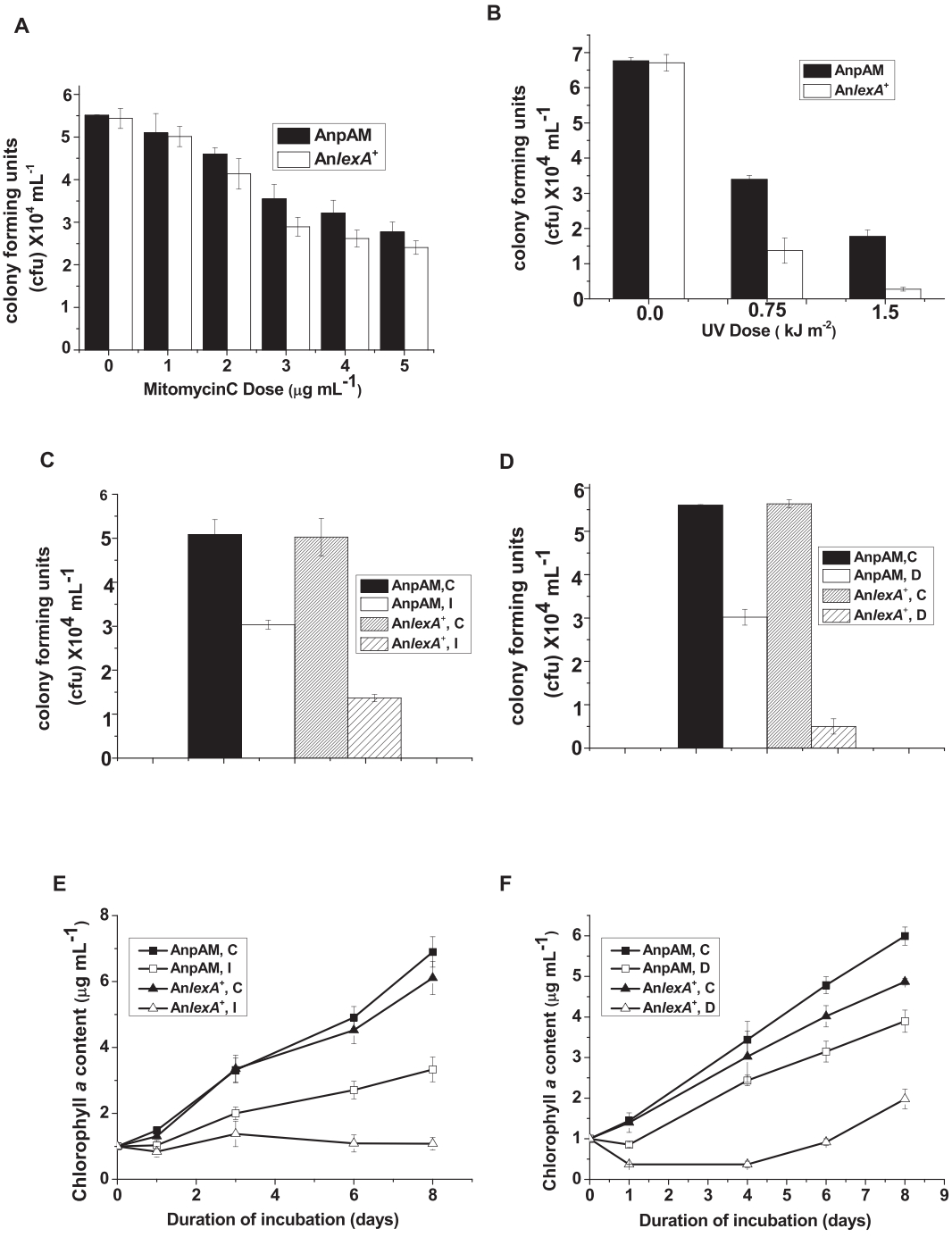


Fig. 2. Effect of LexA overexpression on tolerance of *Anabaena* to DNA-damage inducing stresses.
Three-day-old cultures of recombinant *Anabaena* strains, AnpAM and An*lexA*$^+$ were concentrated to 10 μg Chl *a* density mL^{-1} and subjected to (A) Different doses of mitomycin C (0–5 μg mL^{-1}) for 30 min., followed by plating on BG-11, N$^+$ agar plates, (B) Different doses of UV-radiation (0–1.5 kJ m^{-2}) on BG-11, N$^+$ agar plates, (C) ^{60}Co γ-radiation (6 kGy) 'I' compared with corresponding sham-irradiated control 'C', (D) Six days of desiccation 'D' compared with corresponding humid chamber control 'C'. The plates were incubated under constant illumination for 10 days for optimal growth of distinct colonies. (E) Post-irradiation and (F) post-desiccation recovery was determined by inoculating the stressed and control samples in fresh growth medium at 1 μg Chl *a* mL^{-1}, and grown under control growth conditions followed by measurement of Chl *a* content over a period of 8 days.

(ApcA,B), rod-linker proteins (CpcG) and phycobilisome core-membrane linker protein (ApcE) constitute the phycobilisome complex responsible for harvesting of light during photosynthesis [41]. The relative levels of the phycobiliproteins, which absorb light at different wavelengths (Pec: 570 nm, Cpc: 620 nm and Apc: 650 nm)and involved in transfer of energy define the light harvesting range of cyanobacteria and the photosynthetic efficiency [42]. Among these, ApcE and one spot each of CpcC (Cpc-3) and CpcG1 (CpcG1-4) showed 2.83, 2.15 and 1.44-fold increase in abundance respectively (Fig. 5A), while CpcA-3, CpcC-1, CpcC-2, CpcG1-2 and CpcG1-3 showed 1.5, 1.41, 2.49, 7.31, 1.33-fold decrease in abundance (Fig. 5A). The protein spot corresponding to CpcG1-1 was detected only in AnpAM cells (Fig. 5A). The abundance of PecA, PecB, ApcA, ApcB and CpcB proteins were found to be similar in AnpAM and An*lexA*$^+$ cells (Supplementary Fig. S2B). The photosynthetic efficiency measured using Fluorescence Monitoring System, Hansatech Instruments, UK was found to be similar for

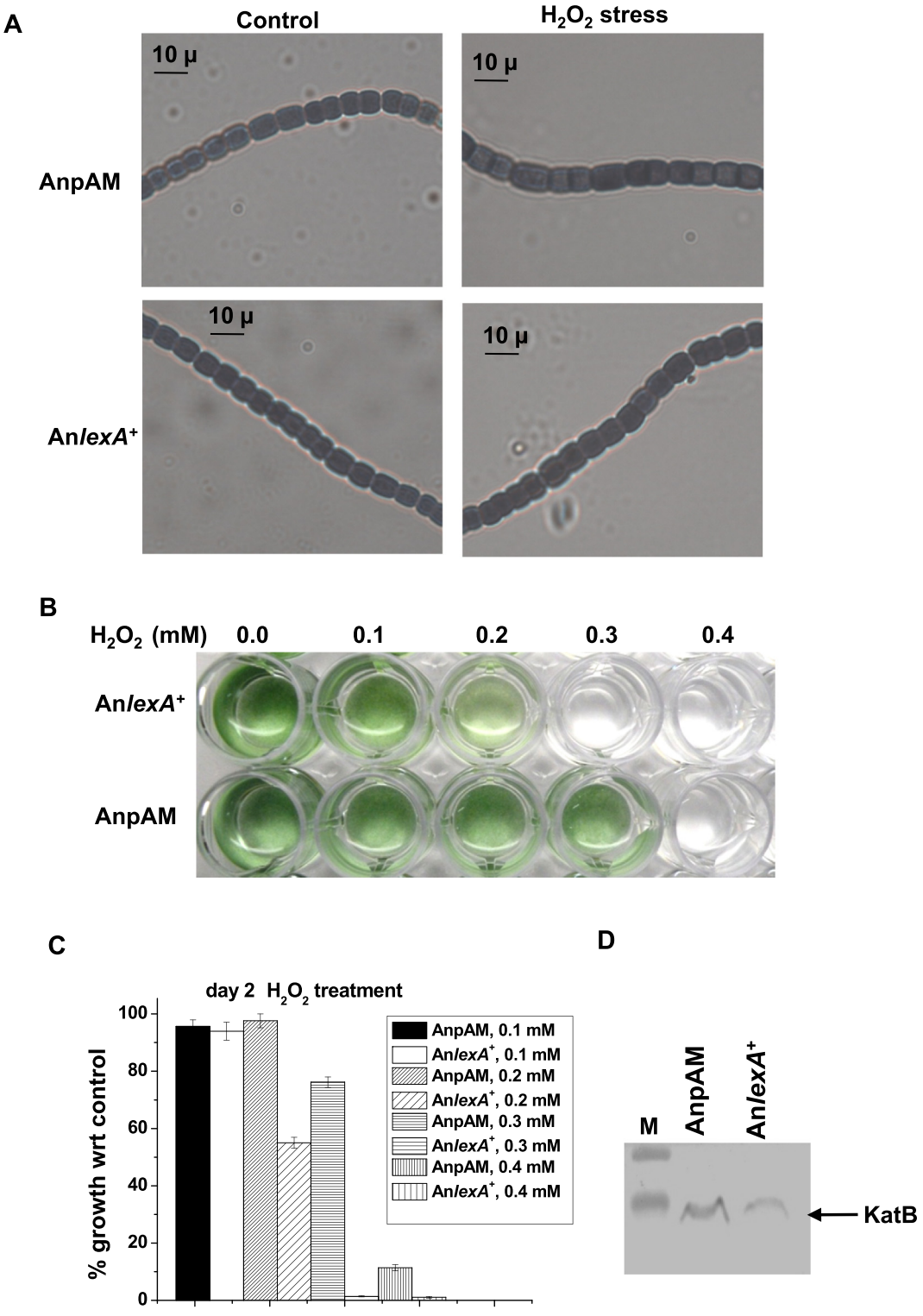


Fig. 3. Effect of LexA overexpression on tolerance of *Anabaena* to H_2O_2-mediated oxidative stress.
Three-day-old cultures of recombinant *Anabaena* strains, AnpAM and An*lexA*$^+$ were inoculated in fresh BG-11, N^+ medium at 1 µg Chl*a* mL^{-1} and exposed to different stresses in microtitre plates as indicated. Growth was measured in terms of chlorophyll *a* (Chl *a*) content and expressed either in absolute values or as % of that under control growth conditions. (A) Bright field microphotographs (1000 ×-magnification) of NBT-stained cells of AnpAM and An*lexA*$^+$ cultures under control conditions and after 90 min exposure to 0.2 mM H_2O_2. (B) Image of the cultures exposed to H_2O_2 and corresponding control captured with Nikon D5100 (DSLR) camera, Japan. (C) Growth of cells exposed to different concentrations of H_2O_2 (0.1–0.4 mM) expressed as percent of control (0 mM H_2O_2). (D) Western blotting and immunodetection of KatB in 3-day-old control grown AnpAM and An*lexA*$^+$ cells. Protein (100 µg/lane) were electrophoretically separated by 15% SDS-PAGE followed by blotting on to nitrocellulose membrane and immunodetection with antisera against *Anabaena* KatB protein. The cross-reacting 26 kDa KatB protein in the two lanes is indicated by an arrow. Lane 'M' refers to pre-stained molecular weight marker.

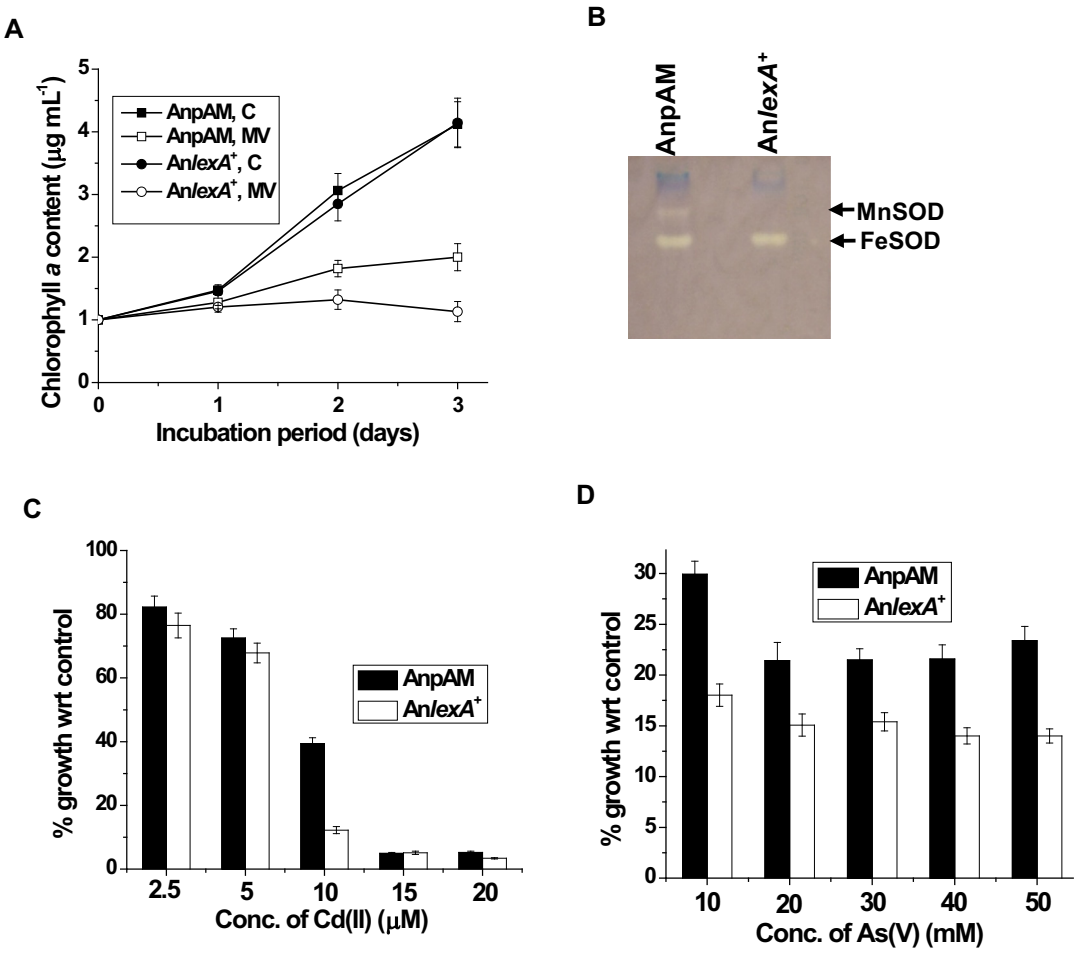


Fig. 4. Effect of LexA overexpression on tolerance of *Anabaena* to MV and heavy metal- stresses.
(A) Growth of cells exposed to 0 and 0.2 µM methyl viologen (MV). (B) SOD zymogram of 3-day-old cultures of AnpAM and An*lexA*$^+$. The MnSOD and FeSOD activity bands are indicated by arrows, (C and D) Percent growth of cells compared to of that under control conditions upon exposure to different concentrations of (C) Cd(II) (0–20 µM) in the form of $CdCl_2$, and (D) As(V) (0–50 mM) in the form of $Na_2HAsO_4.7H_2O$. Other details are as given in the legend to Fig. 3.

An*lexA*$^+$ (7.1 ± 0.2 µmol O_2 mg Chl a^{-1} min^{-1}) and AnpAM (7.4 ± 0.1 µmol O_2 mg Chl a^{-1} min^{-1}) cells. All three protein spots corresponding to RbcL (RbcL-1, RbcL-2 and RbcL-3), which codes for RuBisCo involved in carbon fixation, showed 2.3, 1.39 and 1.25-fold increase in abundance respectively (Fig. 5A), while TalB, GalE and GlgB, which are involved in C-catabolism, exhibited 1.3, 3.14 and 12.19-fold decrease in abundance respectively (Fig. 5B). The levels of other proteins involved in C-catabolism i.e. Enolase (Eno), Transketolase (Tkt), Fructose 1,6-bisphosphatase (GlpX), and phosphoglycerate kinase (Pgk) did not show any significant change upon overexpression of LexA (Supplementary Fig. S2B). The increase in the abundance of RuBisCo involved in C-fixation with a concomitant decrease in that of a few C-catabolism/breakdown proteins (Fig. 5) allowed An*lexA*$^+$ cells to maintain higher intracellular levels of carbon, thereby helping the cells survive inhibition of photosynthesis by DCMU better than the AnpAM cells (Fig. 1). This correlates well with the observed decrease in tolerance to inorganic C-starvation in LexA-depleted mutants of *Synechocystis* [9].

The second largest among the identified groups belonged to oxidative stress alleviation proteins. This included three spots corresponding to Ferredoxin reductase (PetH) and two peroxiredoxins, AhpC and PrxA. Of these, PetH-1 showed 1.67-fold decrease in abundance, while AhpC showed 1.58-fold increase in abundance (Fig. 5C). Protein spots corresponding to PetH-2 and PrxA were detected only in AnpAM gels, while that corresponding to PetH-3 only in An*lexA*$^+$ (Fig. 5C). The FeSOD protein levels were found to be unchanged in An*lexA*$^+$ cells (Supplementary Fig. S2B), correlating well with the unchanged FeSOD activity in An*lexA*$^+$ cells compared to AnpAM (Fig. 4B). The lower abundance of PrxA, PetH in An*lexA*$^+$ cells (Fig. 5C) along with the observed reduction in the levels of KatB (Fig. 3D) and activity of MnSOD (Fig. 4B) results in the higher sensitivity of An*lexA*$^+$ cells to oxidative stress (Figs. 3,4). Though, the abundance of AhpC, which functions as a peroxidase, increased in An*lexA*$^+$ cells, the increase, was probably not enough to offset the decrease in the abundance of other ROS scavenging proteins, resulting in An*lexA*$^+$ cells exhibiting sensitivity to oxidative stress. The remaining three identified protein spots were chaperone DnaK, Nucleoside diphosphate kinase (NdK) and Glutamate ammonia ligase (GlnA) (Supplementary Fig. S1A, Supplementary Table S1), which exhibited 1.26, 1.3 and 1.28-fold decrease in levels in An*lexA*$^+$ cells (Fig. 5D, Supplementary Table S1). No protein with possible involvement in DNA repair could be identified with the proteomics approach, which was not surprising or unique to *Anabaena*, since differential expression of DNA repair genes was not observed by RNA profiling in LexA-depleted mutant of *Synechocystis* either [10]. This, however, does not preclude them from being affected by modulation in LexA levels, since the An*lexA*$^+$ cells exhibited high sensitivity to DNA damage inducing agents (Fig. 2). Hence, a few genes with potential role in DNA repair were also chosen for further analysis.

3.2.2. Analysis of LexA-mediated gene regulation

Transcript analysis was carried out for a few select genes among those identified by proteomic analysis and with potential role in DNA

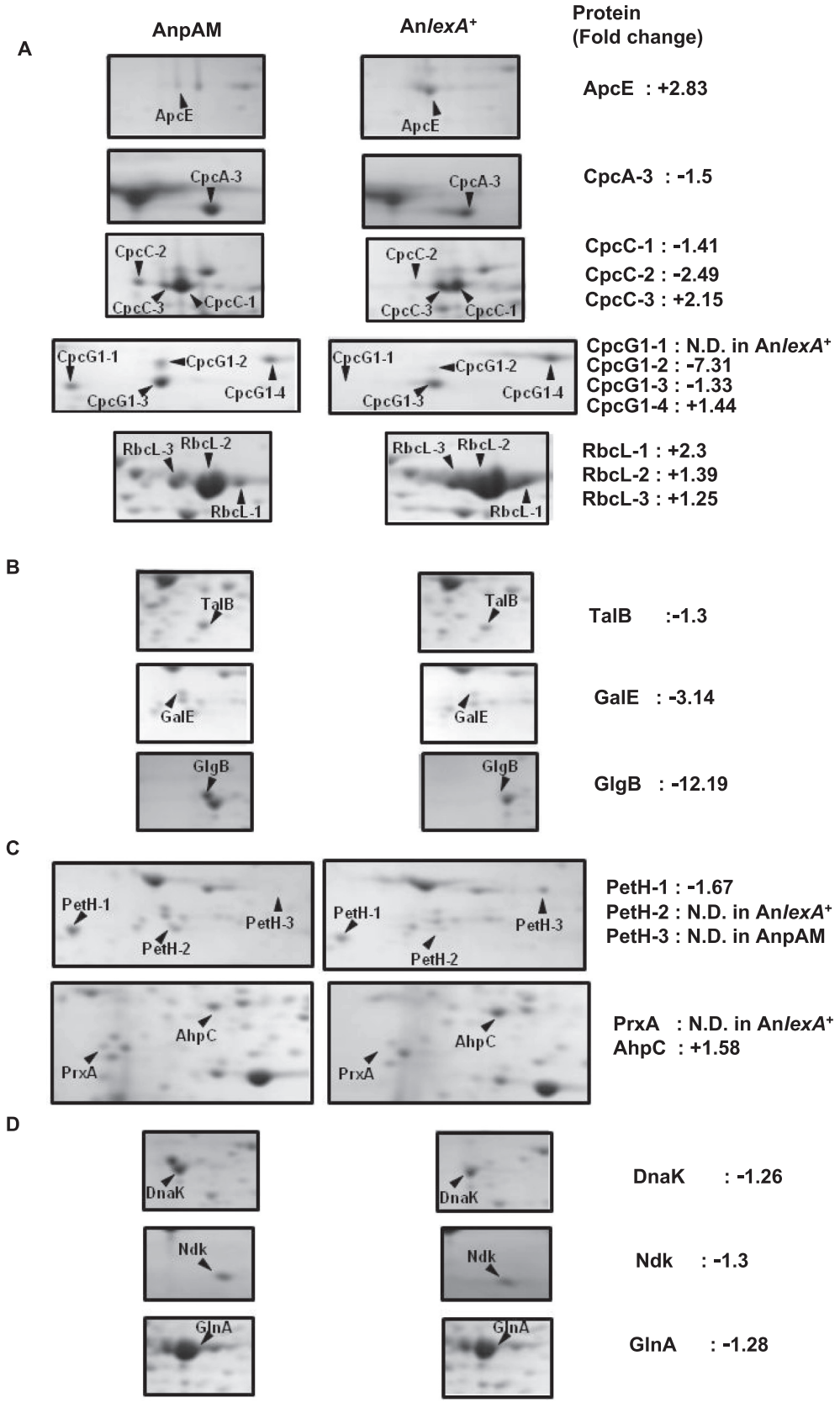


Fig. 5. LexA-overexpression induced changes in the abundance of proteins involved in different physiological functions.
(A) Photosynthesis and Carbon fixation, (B) Carbon metabolism, (C) Oxidative stress alleviation, (D) Transcription/protein folding/amino acid synthesis. The proteins showing increased abundance in An*lexA*$^{+}$ are shown in green, and those with decreased abundance in red. The corresponding fold change in abundance is indicated by '+' and '−' respectively. N.D. indicates not-detected in the corresponding gel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

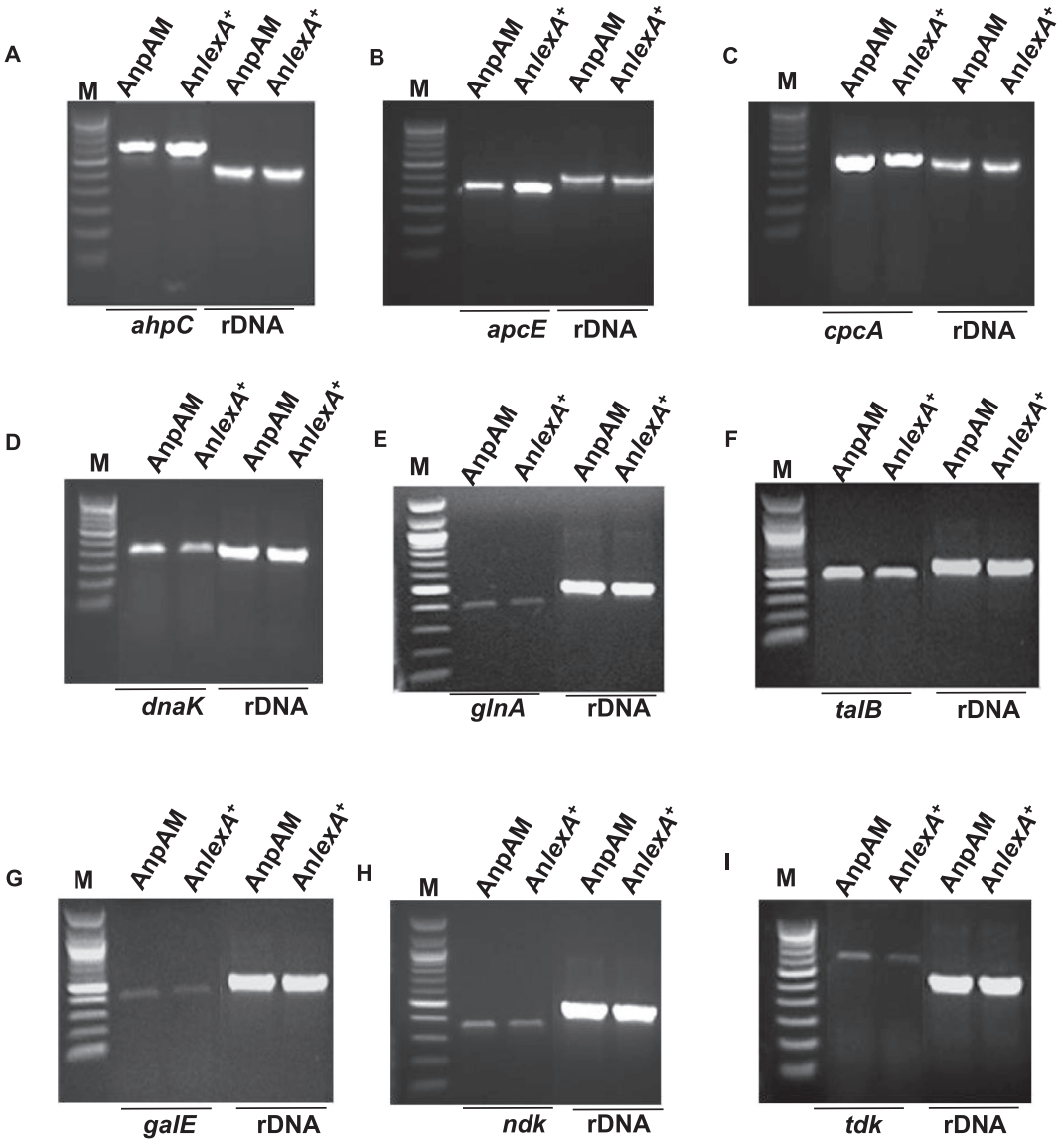
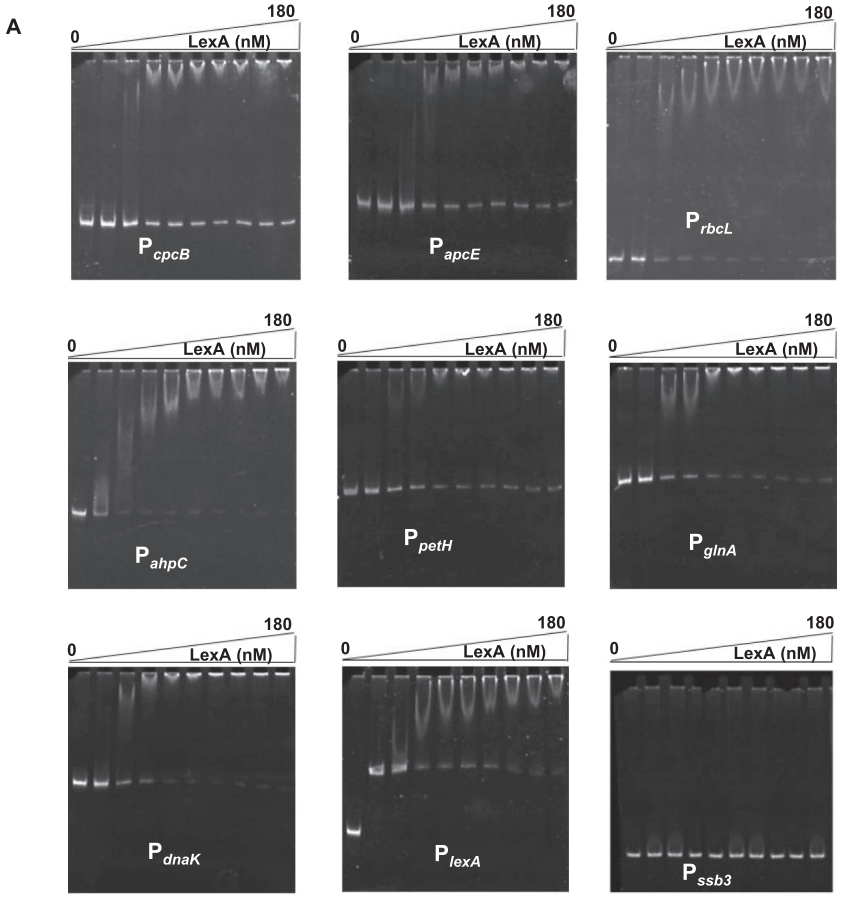


Fig. 6. Effect of LexA overexpression on transcript analysis of select genes of *Anabaena*. Total RNA isolated from 3-day-old AnpAM and An*lexA*$^+$ cells (devoid of DNA contamination) were subjected to Reverse Transcriptase PCR using gene specific primers corresponding to (A) *ahpC*, (B) *apcE*, (C) *cpcA*, (D) *dnaK*, (E) *glnA*, (F) *talB*, (G) *galE*, (H) *ndk*, and (I) *tdk*. The internal control used was 16SrDNA gene primers. Five-fold lower RNA was used for 16SrDNA RT-amplification. Lane M refers to 100 bp DNA ladder (Bangalore Genei, India).

repair. Transcript levels of both *apcE* and *ahpC* were found to be higher in An*lexA*$^+$ cells compared to AnpAM (Figs. 6A,B) indicating that the observed increase in the abundance of ApcE and AhpC (Figs. 5A,C) were due to transcriptional up-regulation. The decrease in transcript levels of *cpcA*, *dna*K, *glnA*, *talB*, *galE*, *ndk* and *tdk* in An*lexA*$^+$ (Fig. 6C–I) directly contributed to the observed decrease in the abundance of the corresponding proteins in these cells (Fig. 5).

To ascertain if the observed transcriptional up- and down-regulation of the different genes in An*lexA*$^+$ cells was directly through the regulatory action of *Anabaena* LexA, (i) Electrophoretic Mobility Shift Assays (EMSA) to determine binding of LexA protein to the upstream regions of these genes, and (ii) assessment of the regulation of their promoter activity by *Anabaena* LexA in *E. coli* were carried out. A 200–400 bp region upstream to these genes and also a few genes with possible role in DNA repair were amplified for analysis by EMSA and cloned in promoterless vector, pAM1956 for testing their promoter activity. EMSA studies revealed binding of LexA to all the promoters tested, with the exception of *ssb3* promoter (Fig. 7A), which has earlier been shown to be regulated by FurA and not LexA [24]. Densitometric analysis of the free and bound DNA was carried out for each DNA-LexA interaction (EMSA) studies in triplicates and plotted as a function of bound DNA (in %) against LexA concentration (in nM) (Supplementary Fig. S3). The binding affinity of LexA to the different promoters was comparable and found to be in the range of 18–96 nM (Table 1). Competition experiments using *ssb1*, *ssb2* and *ssb3* promoters for assessing binding with LexA revealed that (i) low concentration of P_{ssb1} is able to bind LexA even in the presence of excess P_{ssb3}, and (ii) and can partially displace P_{ssb2} from P_{ssb2}-LexA complex [24]. The activity of the different promoters, analysed on the basis of GFP expression, was evaluated in *E. coli* overexpressing *Anabaena* LexA. The activity of *rbcL*, *apcE* and *ahpC* promoters increased upon overexpression of LexA in *E. coli* (Table 1) confirming that the increase in the abundance of these proteins and transcripts was due to the up-regulation of their promoter activity directly by LexA. All other gene promoters corresponding to the proteins showing decreased abundance, showed decrease in their promoter activity upon overexpression of LexA in *E. coli* (Table 1). The correlation between promoter activity, transcript and protein studies indicated a direct role for *Anabaena* LexA in regulating these genes.



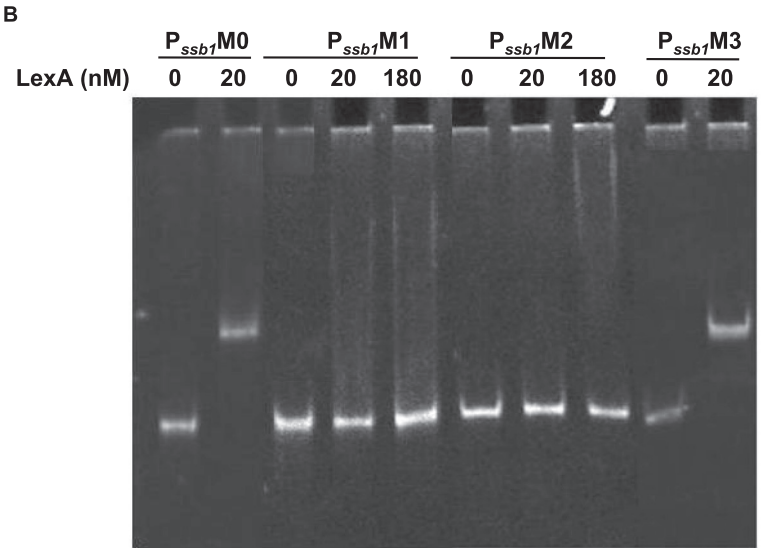


Fig. 7. Interaction of *Anabaena* LexA with upstream regulatory (promoter) region of *Anabaena* genes based on EMSA studies.

(A) Electrophoretic mobility shift assays (EMSA) of selected promoter regions, indicated in the bottom end of each gel, were carried out with increasing concentrations (0–180 nM) of the purified *Anabaena* LexA protein. The lower band in each gel corresponds to free DNA and the upper band to the bound DNA. (B) EMSA of mutated LexA binding sites of P_{ssb1} with *Anabaena* LexA. The ~400 bp promoter region of *ssb1* (P_{ssb1}) was PCR amplified using P_{ssb1}Fwd and one of the four variants of P_{ssb1} Rev. primers; wherein M0 refers to intact LexA binding site (AGT-N_x-ACT), M1 has a mutation in the left arm (AGT to CAC), M2 has a mutation in the right arm (ACT to CTG) and in M3, the intergenic region is shortened to 4 bases from 8, while retaining the sequence of the left and right arm. The DNA protein mix after interaction was separated by 8% PAGE and stained with SYBR green. EMSA was performed with either 1–180 nM LexA for different promoters as indicated in (A) or with 20 nM LexA for P_{ssb1} M0 and P_{ssb1}M3 DNA fragments and with 20 nM and 180 nM for P_{ssb1}M1 and P_{ssb1}M2 DNA fragments in (B). The DNA-protein mix after interaction was electrophoretically separated on 8% polyacrylamide gel followed by staining with SYBR Dye I. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Thus, unlike in other bacteria, LexA in cyanobacteria acts both as a transcriptional activator and a repressor as shown for both *Anabaena* 7120 in this study and *Synechocystis* 6803 earlier [10], and involved in the regulation of functionally diverse genes instead of being restricted to DNA repair genes only.

3.3. Elucidation of LexA-box in Anabaena

Another significant difference was in the sequence of the predicted LexA-box in cyanobacteria [18,21,23,27], which was distinct from that of *E. coli* [2], but no single consensus sequence across all genes regulated by LexA has been reported so far. The probable LexA-binding site of fatty acid biosynthesis genes of *Synechocystis* 6803, also showed variation from the earlier predicted consensus LexA-box [22]. The most likely of the predicted LexA-boxes was an imperfect inverted repeat element [18], which was found to be present upstream or downstream of 216 genes in *Anabaena* 7120, suggesting possible regulation of these genes by LexA [16]. The position of this LexA-box was, however, found

Table 1
Binding affinity and activity of *Anabaena* promoters as a function of LexA expression in *E. coli*.

S. No.	Promoter	Binding Affinity (nM)	Promoter activity in *E. coli* (AU)	Fold change on LexA induction	
				0.5 h	1 h
1	P_{ahpC}	20.0 ± 1.0	842.10 ± 54.48	1.07 ± 0.04	2.88 ± 0.07
2	P_{apcE}	96.0 ± 3.0	4287.04 ± 244.01	1.33 ± 0.10	1.67 ± 0.05
3	P_{rbcL}	32.0 ± 1.5	1296.47 ± 43.03	1.50 ± 0.03	1.14 ± 0.01
4	P_{cpcB}	55.5 ± 2.5	2637.62 ± 155.12	0.42 ± 0.02	0.63 ± 0.06
5	P_{galE}	49.0 ± 3.6	3183.05 ± 97.04	0.58 ± 0.04	0.55 ± 0.04
6	P_{glgB}	27.5 ± 0.5	1149.81 ± 102.11	0.27 ± 0.01	0.38 ± 0.01
7	P_{talB}	60.0 ± 4.1	1808.53 ± 88.5	0.93 ± 0.07	0.21 ± 0.01
8	P_{dnaK}	34.5 ± 1.5	6099.29 ± 245.35	0.75 ± 0.03	0.33 ± 0.03
9	P_{glnA}	34.0 ± 2.1	5375.01 ± 208.32	0.58 ± 0.04	0.45 ± 0.02
10	P_{ndk}	49.0 ± 3.5	6430.46 ± 198.63	0.38 ± 0.03	0.31 ± 0.07
11	P_{petH}	54.5 ± 3.0	2094.93 ± 172.05	1.03 ± 0.07	0.73 ± 0.07
12	P_{prxA}	34.5 ± 2.5	15,181.76 ± 874.32	0.66 ± 0.07	0.53 ± 0.06
13	P_{lexA}	49.5 ± 4.5	7500.19 ± 590.32	0.62 ± 0.02	0.33 ± 0.06
14	P_{recA}	46.5 ± 2.1	8532.25 ± 3451.15	0.47 ± 0.04	0.28 ± 0.03
15	P_{ssb1}	26.5 ± 2.3	4376.36 ± 98.32	0.70 ± 0.02	0.55 ± 0.02
16	P_{ssb2}	23.0 ± 1.9	3880.8 ± 235.12	0.62 ± 0.04	0.45 ± 0.02
17	P_{tdk}	18.0 ± 0.5	4385.55 ± 256.21	0.71 ± 0.03	0.43 ± 0.04

The promoters were cloned in pAM1956 and designated as pAM-P_{abc}.

to be highly variant, not only in terms of the distance from the translational start site of the gene, but also in terms of whether it was present upstream or downstream of the corresponding gene. This necessitated a need for a re-look into the possible consensus LexA-box in cyanobacteria. This was attempted in *Anabaena* 7120 using the small subset of 17 genes shown to be regulated by LexA during this study. When screened for the presence of DNA sequences corresponding to the earlier identified cyanobacterial LexA-box sequence [18,21,23,27], only a partial match was found and that too only for few of the 17 genes in the intergenic region.

Based on similar binding affinities, one would expect the binding site for LexA on these promoters to be similar in terms of DNA sequence. Sequence analysis of these regions revealed the presence of a palindromic sequence AGT-N_x-ACT in 10 of the 17 gene promoters, and with one base variation in the remaining 7 genes (Table 2A). A single base variation was detected in the left arm in 3 cases and in the right arm in 4 cases. However, in all the 7 cases, a minimum of four base palindrome including the mismatch was observed, and hereafter retained as a criteria for prediction of this site in other genes. The spacer region between the inverted repeat sequence varied between 4 and 11 bases (Table 2A). The predicted LexA-box was present in the vicinity of a probable −10 region, which would allow for better regulation of transcription. In order to validate the predicted sequence for binding to LexA, the left and right arm of *ssb1* promoter were mutated (P_{ssb1}M1/M2) to disrupt the palindrome, and the spacer region was reduced from 8 to 4 bases (P_{ssb1}M3) and the binding of LexA to the mutated *ssb1* promoter fragments tested. Fig. 7B showed that the disruption of the palindrome completely abolished the binding of LexA to the promoter region, while decrease in the spacer region did not affect the binding. This confirmed that the predicted sequence for the LexA-box in *Anabaena* is a functional box. In case of the up-regulated genes, *apcE, rbcL* and *ahpC*, the same palindromic sequence was observed. The only difference was the presence of a GAAA or GTTT sequence in the vicinity of the palindrome, however, the significance of this sequence has not been investigated. The ability of *Anabaena* LexA to function as an activator for these gene promoters by enhancing their promoter activity even in *E. coli* (Table 1), not just in *Anabaena*, suggests role for a *cis*-element within the upstream regulatory region in deciding if the *Anabaena* LexA once bound to the AnLexA-box will function as a repressor or an activator while regulating the expression of the downstream gene. The sequence "AGT-N_x-ACT" along with the allowed one-base mismatch has been hereafter designated as "AnLexA-box".

A short genome search for AnLexA-box, in the vicinity of an AT-rich region and within the intergenic region with respect to its preceding gene, revealed the presence of this box upstream of at least 40 genes in *Anabaena* (Table 2B), broadly classified into seven functional groups. A few gene members of three of the functional groups, namely DNA repair, C-metabolism and fatty acid biosynthesis, have earlier been shown to be regulated by LexA in either *Synechocystis* or *Anabaena*. Among the probable DNA repair genes, AnLexA-box was found upstream of *uvrA*, *recG,* topoI, *alr3988* (*sbcC*-like), *all4463* (*sbcD*-like), *recF*, *recO*, *recR*, *radC*, helicase, DNases *alr3199* and *alr3200*, and two methyl transferases *alr2104* and *all0061* (Table 2B). Of these, binding of LexA to *uvrA* promoter [18], and *alr3988* and *all4463* genes, resulting in their negative regulation [39] has been reported earlier. Most of these genes form part of the SOS regulon in bacteria [4], thus implicating a role for *Anabaena* LexA in SOS response, as observed for other bacteria. Six genes with different functions in C-metabolism, namely *hoxU*, *hoxEF*, *hupS*, *fbp*, *gap-2* and *all0475* possessed the AnLexA-box in their upstream intergenic region (Table 2). Of these up-regulation of bidirectional hydrogenase and hydrogen uptake genes *hoxU*, *hoxEF* and *hupS* by LexA has been shown earlier both in *Anabaena* and *Synechocystis* [11,14,16], further ascertaining the validity of the defined AnLexA-box. Recently, the fatty acid (FA) biosynthetic genes of *Synechocystis* 6803 were shown to be negatively regulated by LexA [22]. Fatty acids and lipids have been shown to play an important role in stress tolerance in cyanobacteria [43]. Among the FA biosynthesis genes, *alr0239* (*fabH*), *alr3343* (*fabF*), *alr1894* (*fabG*) and *alr2271* (*fabZ*) genes of *Anabaena* also possessed the AnLexA-box in their upstream regulatory regions (Table 2B), and their down-regulation would result in decreased fatty acid synthesis, thereby negatively affecting their stress tolerance.

Eight genes with probable or proven roles in oxidative stress alleviation, namely *katA*, *katB*, *sodA*, glutaredoxin, dioxygenase, *petC*, FeS-binding protein and *asl0401* were also shown to possess the AnLexA-box in the upstream region (Table 2B). The observed down-regulation of KatB (Fig. 3D) and MnSOD (SodA) (Fig. 4B) in An*lexA*$^+$ could thus be through the binding of LexA to the AnLexA-box found upstream of the *katB* and *sodA* genes (Table 2B). The observed down-regulation of *petH*, *prxA*, *katB* and *sodA* genes along with that speculated for the eight oxidative stress alleviation genes having AnLexA-box upstream to their translational start site (Table 2B), would result in accumulation of reactive oxygen species (ROS) upon exposure to different forms of oxidative stresses, eventually leading to cell death. The presence of AnLexA-box upstream of *all7621*, *alr0831*, which code for metalloregulator protein, SmtB, a known regulator of metal-inducible operons [44], could contribute to the lowering of heavy metal tolerance

Table 2
AnLexA-box upstream of *Anabaena* genes.

2A. Genes whose regulation by LexA has been validated experimentally

S. no.	Gene symbol (annotation)	AnLexA-box (AGT-N_x-ACT)[a]			Probable -10 region	
		Sequence	Spacer region (x: no. of bases)	No. of bases upstream of ATG	Sequence	No. of bases upstream of ATG
Downregulated genes						
1	*glgB (all0875)*	**AGT**ataaa**ACT**	5	94	TAATAA	**51**
2	*talB (all2563)*	**AGT**cttaacaa**ACT**	8	32	TAAAAT	**19**
3	*galE(all4713)*	**AGT**agttgtca**ACT**	8	28	TAAATT	**39**
4	*petH (all4121)*	**AGTT**agaaggtg**AACa**	10	198	TTATTAA	**230**
5	*prxA (alr4641)*	**ACAGaT**cgtaatca**AACTGT**	10	160	TAATCA	**152**
6	*dnaK (alr1742)*	**AGTA**ttaa**TtCT**	6	292	TAAATT	**278**
7	*glnA (alr2328)*	**AGgTG**tgtcat**CAACT**	10	94	TATTTC	**100**
8	*cpcB (alr0258)*	**AAGT**ggcaaga**AaTT**	7	134	TAAAAT	**170**
9	*ndk (alr3402)*	**AGT**taattgtt**gACT**	9	166	TAAACT	**185**
10	*lexA (alr4908)*	**TAGT**actaatgt**tCTA**	8	48	TATTTT	**55**
11	*recA (all3272)*	**AGTA**tatctgttc**TACT**	11	67	TTAATT	**108**
12	*tdk (all4708)*	**AGT**gttcaca**ACT**	7	90	TAAAAAT	**51**
13	*ssb1 (alr0088)*	**AGTA**cttatg**TACT**	8/6	28	TAATAA	**33**
14	*ssb2 (alr7579)*	**TAAGT**catt**ACTTA**	4	236	TATAAT	**224**
Upregulated genes						
1	*apcE (alr0020)*	**AGT**atgt**ACT**	4	225	GAAAAT	**190**
2	*rbcL (alr1524)*	**AGT**aaaagcgtta**ACT**	10	218	ACTTAT	**205**
3	*ahpC (alr4404)*	**TtGT**taaaaacc**ACTA**	8	98	TTAAGA	**115**

2B. *Anabaena* genes possessing AnLexA-box, identified bioinformatically

S. No.	Gene annotation (Symbol)	AnLexA-Box (AGT-N_x-ACT)			Predicted seq in Sjoholm et al., 2007[b]	No. Bases From ATG
		Sequence	Spacer region (x: no. of bases)	No. of bases upstream of ATG		
DNA repair genes						
1	*alr3716 (uvrA)*	**AGT**attattaaaa**ACT**	10	50	aGTACtattGTTCt	−84
2	*all4790- recG*	**AGT**ttttg**ACT**	6	168	gGTACaaatGTACg	−45
3	*alr2780 (TopoI)*	**AGT**ttctcccct**ACT**	9	318	gGTATcgttGTACt	−276
4	*alr3988 (sbcC)*	**AGT**gcatctgt**ACT**	8	52		
5	*all4463(sbcD)*	**AGTTG**aaaa**CAACT**	8	53		
6	*all3374(recF)*	**AGTTT**t**AAAtT**	5	50	aGTACagaaGTTCt	+654
7	*alr4175(recO)*	**TTtGT**cttc**ACTAA**	4	69		
8	*alr4977 (recR)*	**CAGTTT**gtga**AAtCTG**	8	71		
9	*all2951 (helicase)*	**AGTC**ttgttgt**GtCT**	9	218	aGAATtagtGTACc	+33
10	*alr2351 (radC)*	**TAGTTGCAtCTA**	4	36		
11	*alr3200* (DNase)	**AAtT**ggtata**ACTT**	6	49		
12	*alr3199* (DNase)	**AGT**taagaatccg**ACT**	10	38		
13	*alr2104* (methyl transferase)	**AGT**acataagt**ACT**	8	59	aGTACataaGTACt	−59
14	*all0061* (site sp. DNA methyl transferase)	**AGaA**aaa**TACT**	5	83	aGTACttttGTTCc	−869
C-metabolism genes						
15	*all0688 (hupS)*	**AGT**ggcaaaaca**ACT**	9	115	aGAACcagaGTTCc	−246
16	*alr0750 -hoxEF*	**AGT**tgagggt**ACT**	7	149	gGTACtctgGTTCg	−142
17	*alr0762 (hoxU)*	**AGT**tcctcaa**ACT**	7	90		
18	*alr1041 (fbp)*	**AGTACTTAAGTACT**	8	175	aGTACttaaGTACt	−175
19	*alr1095* (GAPDH)	**AGT**taacagttg**ACT**	9	55	gGTATcgctGTACg	+687
20	*all0475 (short chain DH)*	**AGT**tgattgttg**ACT**	9	74	gGTACtccaGTACg	+528
Fatty acid biosynthesis genes						
21	*alr0239 (fabH)*	**AGT**aaatg**ACT**	5	55		
22	*alr3343 (fabF)*	**AAGT**caag**AtTT**	4	140		
23	*alr1894 (fabG)*	**TAGT**tcctca**ACgA**	6	88		
24	*alr2271 (fabZ)*	**TAGcT**ttgagt**AACTA**	6	78		
Oxidative stress alleviation genes						
25	*alr3090 (katB)*	**AGT**ttctggtttg**ACT**	10	127		
26	*all0070 (sodA)*	**AGT**aata**ACT**	4	369		
27	*asl3860 (glutaredoxin)*	**AGT**ttataact**ACT**	8	77	aGTACcaaaGTTCt	−46
28	*all3866 dioxygenase*	**AGT**aacaaatgaca**ACT**	11	46	gGTACttgtGTTCc	−234
29	*asr4942- FeS binding protein*	**AGT**actgcca**ACT**	7	45	aGTACtatgGTACg	−53
30	*alr0998 (katA)*	**GAGT**tttttgagt**CTC**	8	324	gGAACaggaGTACt	+630
31	*all1512 (petC)*	**AGTC**tcttaaagg**GAtT**	11	45	gGTACtgtaGTACg	−311
32	*asl0401*(nitro-reductase)	**AGTAC**ttttt**GTACT**	8	71	aGTACttttGTACt	−71
Metal responsive genes						
33	*all7621 (smtB)*	**AGTT**atca**AtCT**	4	21		
34	*alr0831 (smtB)*	**AGT**gggt**ACT**	4	162		

(*continued on next page*)

Table 2 (*continued*)

2B. *Anabaena* genes possessing AnLexA-box, identified bioinformatically

S. No.	Gene annotation (Symbol)	AnLexA-Box (AGT-N_x-ACT)			Predicted seq in Sjoholm et al., 2007[b]	No. Bases From ATG
		Sequence	Spacer region (x: no. of bases)	No. of bases upstream of ATG		
Toxin-antitoxin genes						
35	*alr4239* toxin secretory pr	**AGT**aaaa**ACT**	4	58	gGTACttttGTTCg	−241
36	*asr0757 (mazE)*	**AGT**aaaa**ACT**	4	60		
37	*all0754 (vapC)*	**TAGTATATACcA**	4	53		
Transcriptional regulators						
38	*alr1194* (two-component regulator)	**AGT**gag**AACT**	5	429	aGAACtttaGTTCt	−425
39	*alr2481* (two-component regulator)	**TAGT**tctt**ACTA**	4	257	aGAACtttaGTTCt	−265
40	*alr1044* (transcriptional regulator)	**TAGT**aattcga**tCTA**	7	53	gGTATttatGTACt	−472

[a] The consensus LexA-binding box has been designated as AnLexA-box. The palindromic sequence is written in bold caps, with the sequences 'AGT' and 'ACT' underlined. The base mismatch in the Agt-N_x-ACT region is written in small bold caps, and the spacer region in small caps.

[b] Some of the genes identified to be possessing the AnLexA-box were also earlier identified [16] as being possibly regulated by LexA based on the presence of the imperfect palindrome "GT/AAC/T-N_4-GTA/TC" upstream or downstream of the gene.

observed in An*lexA*$^+$ cells compared to AnpAM (Figs. 4C,D). This could possibly be achieved by upregulating the levels of All7621 and Alr0831, which in turn would repress the expression of the metal mitigating proteins. Additionally, the AnLexA-box was also found upstream of three genes each comprising the toxin-antitoxin (TA) system, and transcriptional regulators (Table 2B). Of these the TA system genes are known to influence stress tolerance in bacteria [45], while by regulating other transcriptional regulators, LexA would be indirectly involved in regulating host of other responses. Thus, in true sense, LexA of *Anabaena* would function as a global regulator of gene expression.

4. Conclusion

Thus, we can conclude that in *Anabaena*, the LexA protein functions both as an activator and a repressor by binding to a palindromic AnLexA-box (AGT-$N_{[4–11]}$-ACT), present upstream of at least 57 genes belonging to varied functional groups. Based on the reported data for other cyanobacterial strains, such as *Synechocystis*, the global regulatory role of LexA may be true across cyanobacteria, thereby manipulating the response to different abiotic stresses ranging from oxidative stress to DNA damage, and C-limitation to heavy metal tolerance. It could also function as a non-specific modulator of stress response through regulation of fatty acid synthesis, toxins and other transcriptional regulators. It is speculated that during the course of evolution, the role of LexA may have become more defined and restricted to SOS-regulon in modern day bacteria.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbagrm.2018.07.007.

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgements

The authors wish to thank Dr. Bhakti Basu for help rendered in the use of MALDI-ToF/ToF MS for proteomic analysis.

Conflict of interest

The authors report no conflict of interest.

References

[1] A.R. Fernandez De Henestrosa, T. Ogi, S. Aoyagi, D. Chaffin, J.J. Hayes, H. Ohomori, R. Woodgate, Identification of additional genes belonging to the LexA regulon in *Escherichia coli*, Mol. Microbiol. 35 (2000) 1560–1572.
[2] G.C. Walker, Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*, Microbiol. Rev. 48 (1984) 60–93.
[3] J.T. Wade, N.B. Reppas, G.M. Church, K. Struhll, Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites, Genes Dev. 19 (2005) 2619–2630.
[4] C. Janion, Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*, Int. J. Biol. Sci. 4 (2008) 338–344.
[5] H. Jian, L. Xiong, Y. He, X. Xiao, The regulatory function of LexA is temperature-dependent in the deep-sea bacterium *Shewanella piezotolerans* WP3, Front. Microbiol. 6 (2015) 627.
[6] S. van der Veen, S. van Schalkwijk, D. Molenaar, W.M. de Vos, T. Abee, M.H.J. Wells-Bennik, The SOS response of *Listeria monocytogenes* is involved in stress resistance and mutagenesis, Microbiologica 156 (2010) 374–384.
[7] V. Leung, D. Ajdic, S. Koyanagi, C.M. Lévesque, The formation of *Streptococcus mutans* persisters induced by the quorum-sensing peptide pheromone is affected by the LexA regulator, J. Bacteriol. 197 (2015) 1083–1094.
[8] B.M. Walter, S.T. Cartman, N.P. Minton, M. Butala, M. Rupnik, The SOS response master regulator LexA is associated with sporulation, motility and biofilm formation in *Clostridium difficile*, PLoS One 10 (2015) e0144763.
[9] F. Domain, L. Houot, F. Chauvat, C. Cassier-Chauvat, Function and regulation of the cyanobacterial genes *lexA, recA* and *ruvB*: LexA is critical to the survival of cells facing inorganic carbon starvation, Mol. Microbiol. 53 (2004) 65–80.
[10] A. Kizawa, A. Kawahara, Y. Takimura, Y. Nishiyama, Y. Hihara, RNA-seq profiling reveals novel target genes of LexA in the cyanobacterium *Synechocystis* sp. PCC 6803, Front. Microbiol. 7 (193) (2016).
[11] K. Gutekunst, S. Punpruch, C. Schwarz, S. Schuchardt, R. Schulz-Friedrich, J. Appel, LexA regulates the bidirectional hydrogenase in the cyanobacterium *Synechocystis* sp. PCC 6803 as a transcription activator, Mol. Microbiol. 58 (2005) 810–823.
[12] P. Oliveira, P. Lindblad, LexA, a transcription regulator binding in the promoter region of the bidirectional hydrogenase in the cyanobacterium *Synechocystis* sp. PCC 6803, FEMS Microbiol. Lett. 251 (2005) 59–66.
[13] T.K. Antal, P. Oliveira, P. Lindblad, The bidirectional hydrogenases in the cyanobacterium *Synechocystis* sp. strain PCC6803, Int. J. Hydrog. Energy 31 (2006) 1439–1444.
[14] P. Oliveira, P. Lindblad, Transcriptional regulation of the cyanobacterial bidirectional Hox-hydrogenase, Dalton Trans. 45 (2009) 9990–9996.
[15] P. Oliveira, P. Lindblad, An AbrB-like protein regulates the expression of the bidirectional hydrogenase in *Synechocystis* sp. strain PCC6803, J. Bacteriol. 190 (2008) 1011–1019.
[16] J. Sjöholm, P. Oliveira, P. Lindblad, Transcription and regulation of the bidirectional hydrogenase in the cyanobacterium *Nostoc* sp. strain PCC 7120, Appl. Environ. Microbiol. 73 (2007) 5435–5446.
[17] E. Devine, M. Holmqvist, K. Stensjo, P. Lindblad, Diversity and transcription of proteases involved in the maturation of hydrogenase in *Nostoc punctiforme* ATCC29133 and *Nostoc* sp. strain PCC7120, BMC Microbiol. 9 (53) (2009) (doi: 10-1186/1471-2180-9-53).
[18] G. Mazón, I., S. Campoy Erill, A.R. Fernandez de Hemestrosa, P. Candau, J. Barbe, LexA-binding sequences in Gram-positive and cyanobacteria are closely related, Mol. Gen. Genomics. 271 (2004) 40–49.
[19] D. Ferreira, E. Leitao, J. Sjoholm, P. Oliveira, P. Lindblad, P. Moradas-Ferreira, P. Tamagnini, Transcription and regulation of hydrogenase(s) accessory genes, *hypFCDEAB*, in the cyanobacterium *Lyngbya majuscula* CCAP1446/4, Arch. Microbiol. 188 (2007) 609–617.
[20] L.M. Patterson-Fortin, K.R. Colvin, G.W. Owttrim, A LexA-related protein regulates redox-sensitive expression of the cyanobacterial RNA helicase, *crhR*, Nucleic Acids Res. 34 (2006) 3446–3454.
[21] L.M. Patterson-Fortin, G.W. Owttrim, A *Synechocystis* LexA-orthologue binds direct repeats in target genes, FEBS Lett. 582 (2008) 2424–2430.

[22] A. Kizawa, A. Kawahara, K. Takashima, Y. Takimura, Y. Nishiyama, Y. Hihara, The LexA transcription factor regulates fatty acid biosynthetic genes in the cyanobacterium *Synechocystis* sp. PCC 6803, Plant J. 92 (2017) 189–198.
[23] S. Li, M. Xu, Z. Su, Computational analysis of LexA regulons in cyanobacteria, BMC Genomics 11 (2010) 527.
[24] A. Kirti, A. Kumar, H. Rajaram, Differential regulation of *ssb* genes in the nitrogen-fixing cyanobacterium, *Anabaena* sp. strain PCC7120, J. Phycol. 53 (2017) 322–332.
[25] A. Kumar, A. Kirti, H. Rajaram, LexA protein of cyanobacterium *Anabaena* sp. strain PCC7120 exhibits *in vitro* pH-dependent and RecA-independent autoproteolytic activity, Int. J. Biochem. Cell Biol. 59 (2015) 84–93.
[26] P. Oliveira, P. Lindblad, Novel insights into the regulation of LexA in the cyanobacterium *Synechocystis* sp. strain PCC6803, J. Bacteriol. 193 (2011) 3804–3814.
[27] N. Blot, D. Mella-Flores, C., G. Le Corguille Six, C. Boutte, A. Peyrat, A. Monnier, M. Ratin, P. Gourvil, D.A. Campbell, L. Garczarek, Light history influences the response of the marine cyanobacterium *Synechococcus* sp. WH7803 to oxidative stress, Plant Physiol. 156 (2011) 1934–1954.
[28] A.K. Chaurasia, A. Parasnis, S.K. Apte, An integrative expression vector for strain improvement and environmental applications of the nitrogen fixing cyanobacterium, *Anabaena* sp. strain PCC7120, J. Microbiol. Methods 73 (2008) 133–141.
[29] H.S. Yoon, J.W. Golden, Heterocyst pattern formation controlled by a diffusible peptide, Science 282 (1998) 935–938.
[30] P.S. Raghavan, H. Rajaram, S.K. Apte, Nitrogen status dependent oxidative stress tolerance conferred by overexpression of MnSOD and FeSOD proteins in *Anabaena* sp. strain PCC7120, Plant Mol. Biol. 77 (2011) 407–417.
[31] H. Rajaram, S.K. Apte, Differential regulation of *groESL* operon expression in response to heat and light in *Anabaena*, Arch. Microbiol. 192 (2010) 729–738.
[32] R.W. Castenholz, Culturing Methods for Cyanobacteria, Methods Enzymol, vol. 167, Academic Press New York, London, 1988, pp. 68–93.
[33] G. Mackinney, Absorption of light by chlorophyll solutions, J. Biol. Chem. 140 (1941) 315–322.
[34] B. Panda, B. Basu, H. Rajaram, S.K. Apte, Methyl viologen responsive proteome dynamics of *Anabaena* sp. strain PCC7120, Proteomics 14 (2014) 1895–1904.
[35] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH image to ImageJ: 25 years of image analysis, Nat. Methods 9 (2012) 671–675.
[36] D. Chakravarty, M. Banerjee, S.C. Bihani, A. Ballal, A salt inducible Mn-catalase (KatB) protects cyanobacterium from oxidative stress, Plant Physiol. 170 (2016) 761–773.
[37] Y.Y. He, D.P. Hader, UV-B-induced formation of reactive oxygen species and oxidative damage of the cyanobacterium *Anabaena* sp: protective effects of ascorbic acid and N-acetyl-L-cystein, J. Photochem. Photobiol. B66 (2002) 115–124.
[38] P.S. Raghavan, H. Rajaram, S.K. Apte, Membrane targeting of MnSOD is essential for oxidative stress tolerance of nitrogen-fixing cultures of *Anabaena* sp. strain PCC7120, Plant Mol. Biol. 88 (2015) 503–514.
[39] S. Pandey, A. Kirti, A. Kumar, H. Rajaram, The SbcC and SbcD homologs of the cyanobacterium *Anabaena* sp. strain PCC7120 (Alr3988 and All4463) contribute independently to DNA repair, Funct. Integ. Genomics. (2018), https://doi.org/10.1007/s10142-018-0599-7.
[40] D.W. Mount, K.B. Low, S.J. Edmiston, Dominant mutations (*lex*) in *Escherichia coli* K-12 which affect radiation sensitivity and frequency of ultraviolet light-induced mutations, J. Bacteriol. 112 (1972) 886–893.
[41] L.-N. Liu, X.-L. Chen, Y.-Z. Zhang, B.-C. Zhou, Characterization, structure and function of linker polypeptides in phycobilisomes of cyanobacteria and red algae: an overview, Biochim. Biophys. Acta 1708 (2005) 133–142.
[42] D.A. Bryant, Phycoerythrocyanin and phycoerythrin; properties and occurrence in cyanobacteria, J. Gen. Microbiol. 128 (1982) 835–844.
[43] S.C. Singh, R.P. Sinha, D.-P. Haeder, Role of lipids and fatty acids in stress tolerance in cyanobacteria, Acta Theriol. 41 (2002) 297–308.
[44] T. Liu, J.W. Golden, D.P. Giedroc, A zinc(II)/lead(II)/cadmium(II)-inducible operon from the cyanobacterium *Anabaena* is regulated by AztR, an alpha 3N ArsR/SmtB metalloregulator, Biochemist 44 (2005) 8673–8683.
[45] I. Kolodkin-Gal, H. Engelberg-Kulka, Induction of *Escherichia coli* chromosomal *mazEF* by stressful conditions causes an irreversible loss of viability, J. Bacteriol. 188 (2006) 3420–3423.

Appendix A: *E. coli* stains harbouring recombinant plasmids

Strain	Characteristics	Source/Reference
DH5α(pET*lexA*)	Cb^r, DH5α strain harbouring pET*lexA* plasmid	This study
DH5α (pET*lexA*-CTD)	Cb^r, DH5α strain harbouring pET*lexA*-CTD plasmid	This study
DH5α (pET*lex*A-NTD)	Cb^r, DH5α strain harbouring pET*lexA*-NTD plasmid	This study
DH5α (pET*lexA*ΔGLI)	Cb^r, DH5α strain harbouring pET*lexA*ΔGLI plasmid	This study
DH5α (pET*lexA*-A84S)	Cb^r, DH5α strain harbouring pET*lexA*-A84S plasmid	This study
DH5α (pET*lexA*-G85S)	Cb^r, DH5α strain harbouring pET*lexA*-G85S plasmid	This study
DH5α (pET*lexA*-S118A)	Cb^r, DH5α strain harbouring pET*lexA*-S118A plasmid	This study
DH5α (pET*lexA*-K159R)	Cb^r, DH5α strain harbouring pET*lexA*-K159R plasmid	This study
DH5α(pET*lexA*-E96Q)	Cb^r, DH5α strain harbouring pET*lexA*-E96Q plasmid	This study
BL21(pET*lexA*)	Cm^r, Cb^r, BL21(p*lysS*) cells harbouring pET*lexA* plasmid	This study
BL21(pET*recA*)	Cm^r, Cb^r, BL21(p*lysS*) cells harbouring pET*recA* plasmid	This study
BL21(pET*lexA*-CTD)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-CTD plasmid	This study
BL21(pET*lexA*-NTD)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-NTD plasmid	This study
BL21(pET*lexA*-ΔGLI)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-ΔGLI plasmid	This study
BL21(pET*lexA*-A84S)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-A84S plasmid	This study
BL21(pET*lexA*-A85S)	Cm^r, Kan^r, BL21(p*lysS*) cells	This study

	harbouring pET*lexA*-A85S plasmid	
BL21(pET*lexA*-S118A)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-S118A plasmid	This study
BL21(pET*lexA*-K159R)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-K159R plasmid	This study
BL21(pET*lexA*-E96Q)	Cm^r, Kan^r, BL21(p*lysS*) cells harbouring pET*lexA*-E96Q plasmid	This study
DH5α(pAM-P_x)	Kan^r, DH5α strain harbouring the plasmid pAM-P_x, wherein P_x= P_{glgB}, P_{galE}, P_{cpcB}, P_{apcE}, P_{rbcL}, P_{ahpC}, P_{ndk}, P_{galE}, P_{petH} , P_{tdk}, P_{dnaK}, P_{glnA} or P_{recA}.	This study
DH5α(pAM-P_{ssb1})	Kan^r, DH5α strain harbouring the plasmid pAM-P_{ssb1}	(171)
DH5α(pAM-P_{ssb2})	Kan^r, DH5α strain harbouring the plasmid pAM-P_{ssb2}	(171)
BL21(pET16b)(pAM-P_x)	Cm^r, Cb^r, BL21(p*lysS*) cells harbouring pET16b plasmid, co-transformed with Kan^r pAM-P_x plasmid; wherein P_x= P_{glgB}, P_{galE}, P_{cpcB}, P_{apcE}, P_{rbcL}, P_{ahpC}, P_{ndk}, P_{galE}, P_{petH} , P_{tdk}, P_{dnaK}, P_{glnA}, P_{recA}, P_{ssb1} or P_{ssb2}	This study
BL21(pET*lexA*)(pAM-P_x)	Cm^r, Cb^r, BL21(p*lysS*) cells harbouring pET*lexA* plasmid, co-transformed with Kan^r pAM-P_x plasmid as defined above	This study

Appendix B

GenBank

Nostoc sp. PCC 7120 LexA (lexA) gene, complete cds

GenBank: **KF269537**.1

FASTA Graphics

Go to:

```
LOCUS       KF269537                606 bp    DNA     linear   BCT 28-JUL-2013
DEFINITION  Nostoc sp. PCC 7120 LexA (lexA) gene, complete cds.
ACCESSION   KF269537
VERSION     KF269537.1
KEYWORDS    .
SOURCE      Nostoc sp. PCC 7120 (Anabaena sp. PCC 7120)
  ORGANISM  Nostoc sp. PCC 7120
            Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   Kumar,A., Rajaram,H. and Apte,S.K.
  TITLE     LexA regulon of Anabaena 7120
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 606)
  AUTHORS   Kumar,A., Rajaram,H. and Apte,S.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2013) Molecular Biology Division, Bhabha Atomic
            Research Centre, Trombay, Mumbai, Maharashtra 400085, India
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..606
                     /organism="Nostoc sp. PCC 7120"
                     /mol_type="genomic DNA"
                     /strain="PCC 7120"
                     /db_xref="taxon:103690"
                     /PCR_primers="fwd_name: lexa fwd primer, fwd_seq:
                     gccgccatatggaacgcctaacagaa, rev_name: lexa rev primer,
                     rev_seq: gaaggatcctcacatataaccgcgcca"
     gene            1..606
                     /gene="lexA"
     CDS             1..606
                     /gene="lexA"
                     /codon_start=1
                     /transl_table=11
                     /product="LexA"
                     /protein_id="AGR66231.1"
                     /translation="MERLTEAQQELYEWLAEYIRIHQHSPSIRQMMQAMNLKSPAPIQ
                     SRLEHLRTKGYIEWTEGKARTIRVLQPIKQGVPVLGAIAAGGLIEPFTDAVEHIDFSN
                     FVLPAQTYALRVTGDSMIEDLITDGDLVFLRPVPEPDQLKNGTIVAARVDGYGNTLKR
                     FYRSGDRITLKPANPKYNPIEVAAIQVEVQGSLVGVWRGYM"
ORIGIN
        1 atggaacgcc taacagaagc gcaacaagaa ctttatgaat ggctggcaga atatatccgt
       61 attcaccagc attcgccctc aattcgtcag atgatgcagg cgatgaattt aaaatcacct
      121 gcacctatcc aaagtcgttt agaacattta cgcacgaaag gatacattga atggactgaa
      181 ggtaaagccc gaacaattcg agttttacaa cctattaagc aaggtgtgcc ggtttttggga
      241 gcgatcgccg caggtggttt aatagaacca ttcactgatg ctgtcgagca tatcgacttt
      301 tctaatttcg ttttacctgc tcaaacttat gctttgcggg taactggtga cagcatgatt
      361 gaagatttaa ttaccgatgg ggatttggta tttttgcgcc cagttcccga accagatcaa
      421 ttaaaaaatg gtactatcgt cgctgccaga gtggatggtt atggtaatac attaaaacgt
      481 ttttatcgaa gtggcgatcg catcactctt aaaccagcca accccaaata taaccccatt
      541 gaagttgcag ccatacaggt agaggtgcaa ggttctctgg ttggtgtgtg gcgcggttat
      601 atgtga
```

Appendix-C

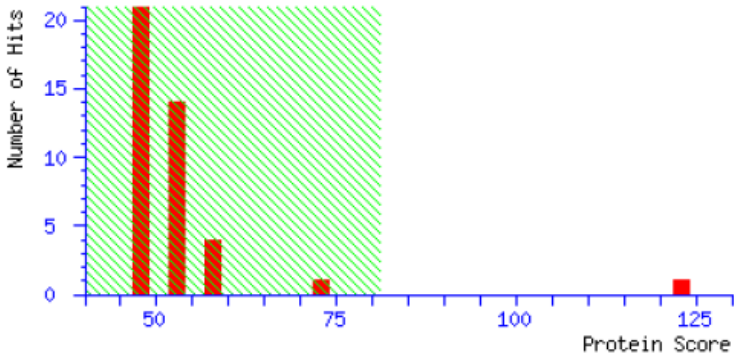
GlgB

MATRIX SCIENCE Mascot Search Results

```
User            : Arvind Kumar
Email           : parvind@barc.gov.in
Search title    : spot1
MS data file    : peaklist.xml
Database        : NCBInr 20150912 (71310511 sequences; 25955008131 residues)
Taxonomy        : Other Bacteria (7064089 sequences)
Timestamp       : 15 Sep 2015 at 09:08:35 GMT
Top Score       : 123 for gi|75704401, Glycoside hydrolase, family 13-like protein [Anabaena
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 81 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▼ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▼

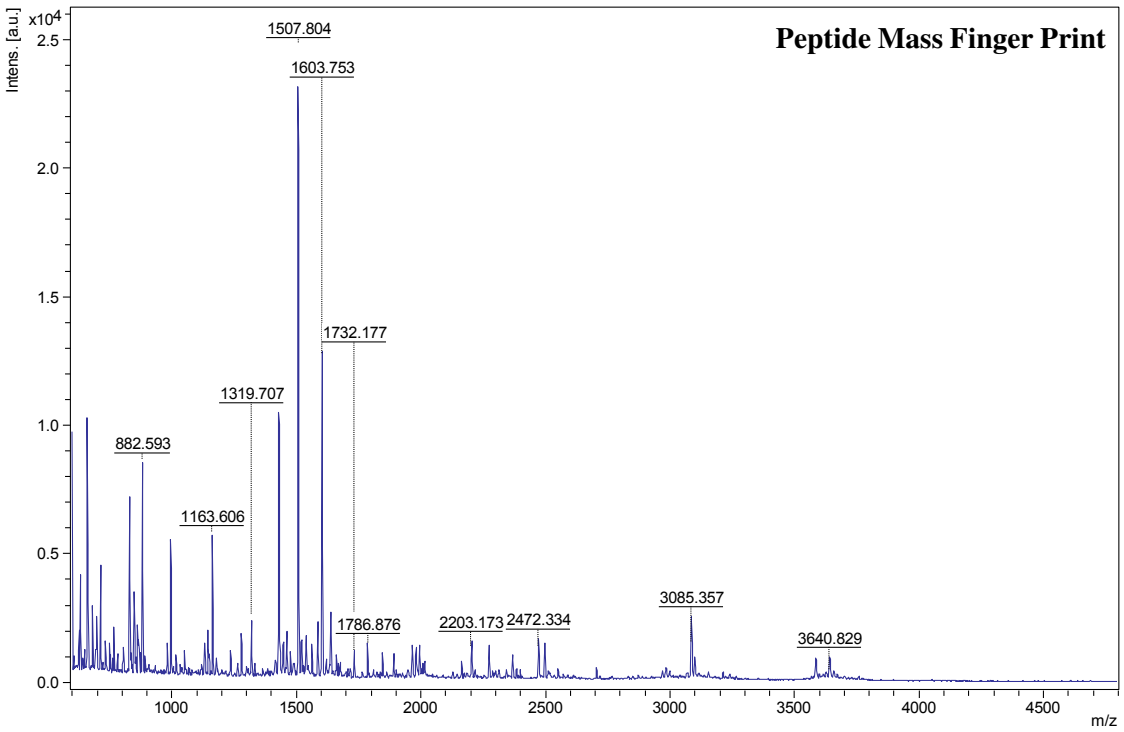
Re-Search All | Search Unmatched

1. gi|75704401 **Mass:** 64033 **Score:** 123 **Expect:** 3.5e-06 **Matches:** 16
Glycoside hydrolase, family 13-like protein [Anabaena variabilis ATCC 29413]
gi|499304274 **Mass:** 64022 **Score:** 110 **Expect:** 7.1e-05 **Matches:** 15
alpha-amylase [Nostoc sp. PCC 7120]

Protein sequence coverage: 33%

Matched peptides shown in ***bold red***.

```
  1 MAKPIEFTLF APYNKGAALI ASFSDWQEIP MKKGDDGYFR TTVELEDGTY
 51 QYKFRVQTRS WFFEEDQWVD VTDPYATDID ESSGKDNSIA RIKDGEKIVD
101 TYVWQHDDKP LPADHELVIY ELHVGDFSGG EDDPYARGKY KHVIEKLDYL
151 CELGINAIEL LPVKEYPGDY SWGYNPRYFF ATESSYGSTA DLKKLVDECH
201 QRGIRIIMDG IYNHSEASSP LTQIDHDYWY HHEPRDPDNN WGPEFNYEHY
251 DENLETYPAR KFIGDTVRYW VGEYHLDGIR YDAARQIANY DFMHWIAQEA
301 KKTAGAKPFY NVAEHIPETT SITNLDGPMD GCWHDSFYHT IKAHICGDTF
351 DLENLKDVID PKRQGFLGAT NVVNYLTNHD HDHIMVELGN REIFHDEAFR
401 RAKLGTAILM TAVGVPLIWM GEEFGEYKPK QQDQSKIDWT LLGNDLNRSL
451 FDYHKGLIGL RKNNHALYTE NIDFIHENPE AKVLAYSRWN DEGSRVVVVT
501 NFSEDFLAGY HVPNFPSGGT WHEWTGDYDV EAGDDGIITD IGPYEAKVFV
551 WQ
```



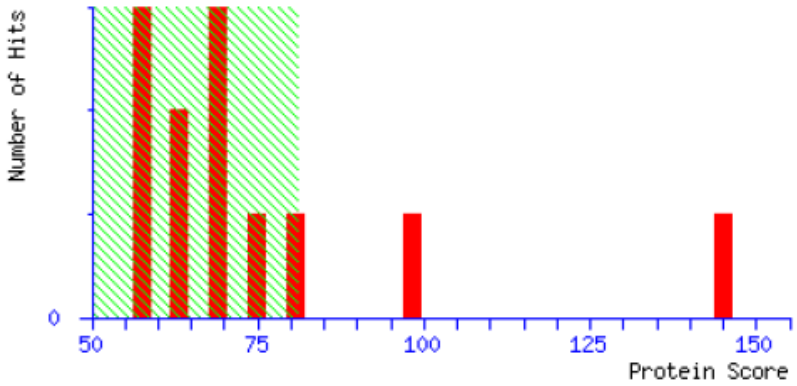
DnaK

MATRIX SCIENCE Mascot Search Results

```
User            : Arvind Kumar
Email           : parvind@barc.gov.in
Search title    : spot2
MS data file    : peaklist.xml
Database        : NCBInr 20150912 (71310511 sequences; 25955008131 residues)
Taxonomy        : Other Bacteria (7064089 sequences)
Timestamp       : 15 Sep 2015 at 09:17:51 GMT
Top Score       : 145 for gi|499305135, chaperone protein dnaK2 [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 81 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▼ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▼

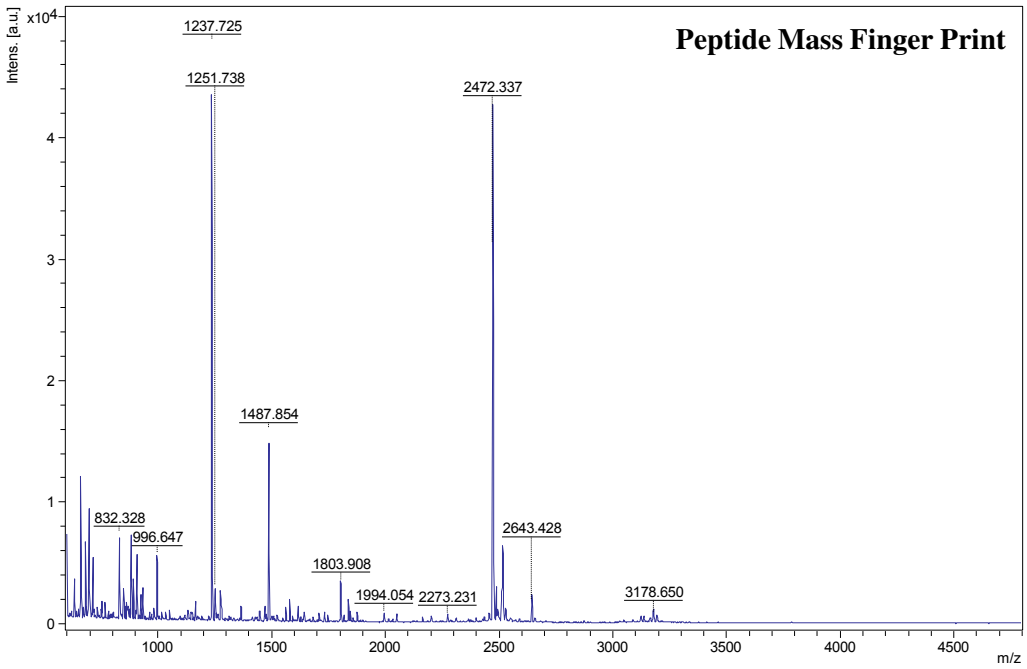
Re-Search All | Search Unmatched

1. gi|499305135 **Mass:** 68037 **Score:** 145 **Expect:** 2.2e-08 **Matches:** 18
chaperone protein dnaK2 [Nostoc sp. PCC 7120]

Protein sequence coverage: 35%

Matched peptides shown in ***bold red***.

```
  1 MAKVVGIDLG TTNSCVAVME GGKPTVIANA EGFRTTPSVV AFAKNGDTLV
 51 GQIAKRQAVM NPENTFYSVK RFIGRRYDEV TNEATEVSYK VLSSGGNVKV
101 DSPGAGKQFA PEEISAKVLR KLVEDASKYL GETVTQAVIT VPAYFNDSQR
151 QATKDAGKIA GIEVMRIINE PTAASLAYGF DKKSNETILV FDLGGGTFDV
201 SVLEVGDGVF EVLATSGDTH LGGDDFDKKI VDFLAEQFRK DEGIDLRKDK
251 QALQRLTEAA EKAKIELSSV TQAEINLPFI TATQDGPKHL DMTLTRAKFE
301 ELCADLIDRC RIPVEQALRD AKLKKENIDE VVLVGGSTRI PAVQQLVKNL
351 LGREPNQTVN PDEVVAVGAA IQAGVLGGEV TGILLLDVTP LSLGVETLGG
401 VMTKIIPRNT TIPTKKSEVF STAVDGQTNV EIHVLQGERE FANDNKSLGT
451 FRLDGIPPAP RGVPQIEVVF DIDANGILNV TAKDKGTGKE QSISITGAST
501 LDKTDIDRMV REAEQNASSD KERREKIERK NQADSLAYQA EKQLQELGDK
551 VPEADKTKVE GLVKDLREAV AKEDDEQIKK LTPELQQALF AVGSNIYQQA
601 GGGAAPGAAP QDGGTTSSDG GDDVIDADFT ETK
```



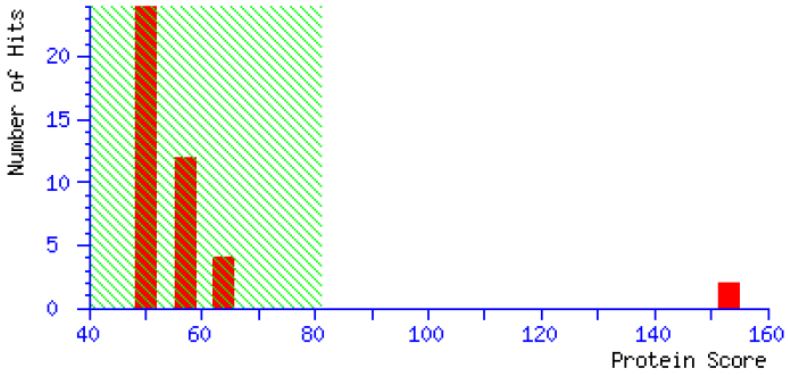
GlnA

MATRIX SCIENCE Mascot Search Results

```
User            : Arvind Kumar
Email           : parvind@barc.gov.in
Search title    : spot3
MS data file    : peaklist.xml
Database        : NCBInr 20150912 (71310511 sequences; 25955008131 residues)
Taxonomy        : Other Bacteria (7064089 sequences)
Timestamp       : 15 Sep 2015 at 10:14:23 GMT
Top Score       : 153 for gi|223805, Gln synthetase
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 81 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▼ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▼

Re-Search All | Search Unmatched

1. gi|223805 Mass: 53430 Score: 153 Expect: 3.5e-09 Matches: 16
   Gln synthetase

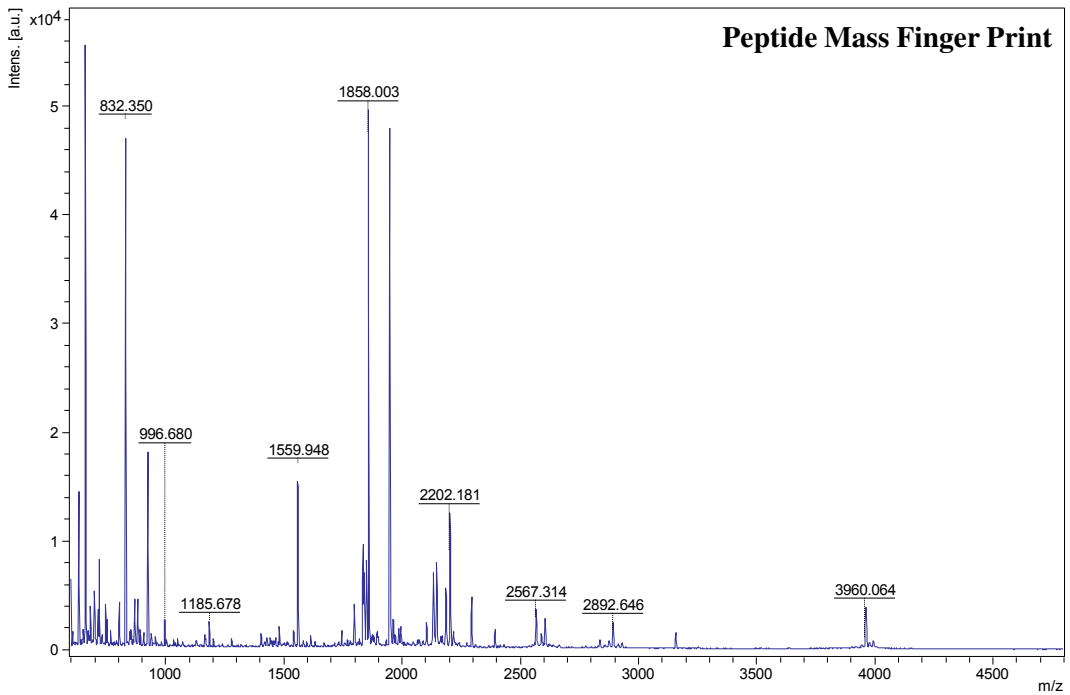
---

2. gi|38891 Mass: 53386 Score: 153 Expect: 3.5e-09 Matches: 17
   unnamed protein product [Nostoc sp. PCC 7120]
   gi|499305709 Mass: 53445 Score: 153 Expect: 3.5e-09 Matches: 17
   glutamine synthetase [Nostoc sp. PCC 7120]

Protein sequence coverage: 47%

Matched peptides shown in ***bold red***.

```
  1 MTTPQEVLKR IQDEKIELID LKFIDTVGTW QHLTLYQNQI DESSFSDGVP
 51 FDGSSIRGWK AINESDMTMV LDPNTAWIDP FMEVPTLSIV CSIKEPRTGE
101 WYNRCPRVIA QKAIDYLVST GIGDTAFFGP EAEFFIFDSA RFAQNANEGY
151 YFLDSVEGAW NSGKEGTADK PNLAYKPRFK EGYFPVSPTD SFQDIRTEML
201 LTMAKLGVPI EKHHHEVATG GQCELGFRFG KLIEAADWLM IYKYVIKNVA
251 KKYGKTVTFM PKPIFGDNGS GMHCHQSIWK DGKPLFAGDQ YAGLSEMGLY
301 YIGGLLKHAP ALLAITNPST NSYKRLVPGY EAPVNLAYSQ GNRSASIRIP
351 LSGTNPKAKR LEFRCPDATS NPYLAFAAML CAGIDGIKNK IHPGEPLDKN
401 IYELSPEELA KVPSTPGSLE LALEALENDH AFLTDTGVFT EDFIQNWIDY
451 KLANEVKQMQ LRPHPYEFSI YYDV
```



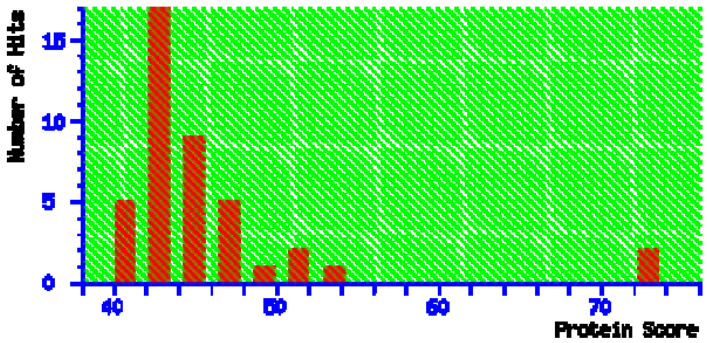
TalB

MATRIX SCIENCE Mascot Search Results

User : Bhakti Basu
Email : bhaktibasu@gmail.com
Search title :
Database : NCBInr 20131113 (34201960 sequences; 12001222213 residues)
Taxonomy : Other Bacteria (3793221 sequences)
Timestamp : 12 Dec 2013 at 10:55:01 GMT
Top Score : 73 for gi|75906716, transaldolase B [Anabaena variabilis ATCC 29413]

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 78 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary Help
Significance threshold p< 0.05 Max. number of hits 20
Preferred taxonomy: All entries

Re-Search All Search Unmatched

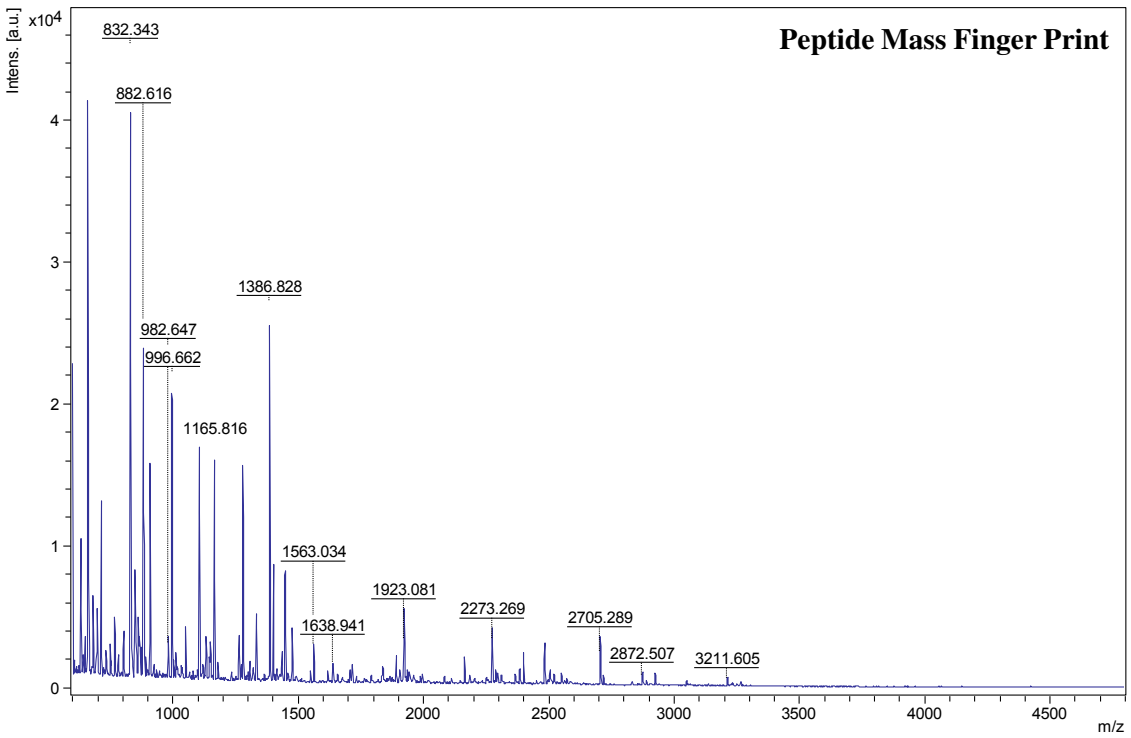
1. gi|75906716 Mass: 36256 Score: 73 Expect: 0.19 Matches: 7
transaldolase B [Anabaena variabilis ATCC 29413]
gi|1261622 Mass: 36290 Score: 42 Expect: 2.7e+02 Matches: 5
transaldolase [Anabaena variabilis]

2. gi|17230055 Mass: 36269 Score: 72 Expect: 0.22 Matches: 8
transaldolase B [Nostoc sp. PCC 7120]

Protein sequence coverage: 23%

Matched peptides shown in ***bold red***.

```
  1 MTKNLLEQLR EMTVVVADTG DIQAIEKFTP RDATTNPSLI TAAAKMPEYQ
 51 EIVDQTLLQA KKDAGAGASK GQIVSLAFDR LAVSFGLKIL QIIPGRVSTE
101 VDARLSYDTE ATITKARELI AQYKAAGIGP ERVLIKIAST WEGIKAAEIL
151 EKEGIHCNLT LLFGLHQAIA CAEAGITLIS PFVGRILDWY KKETGRDSYP
201 SAEDPGVISV TTIYNYYKKF GYTTEVMGAS FRNIGEITEL AGSDLLTISP
251 GLLGELQATI GELPRKLDPA KAATLDIEKI SIDKATFDKM HAADRMAYDK
301 LDEGIKGFTK ALEELETLLA ERLARLEVVA SH
```



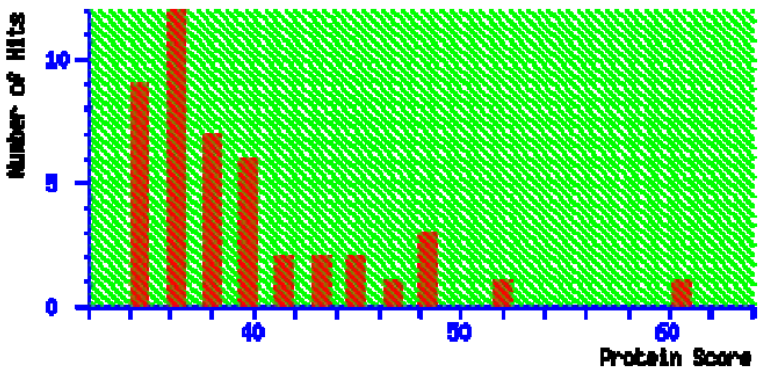
GalE

MATRIX SCIENCE Mascot Search Results

User : Bhakti Basu
Email : bhaktibasu@gmail.com
Search title :
Database : NCBInr 20131113 (34201960 sequences; 12001222213 residues)
Taxonomy : Other Bacteria (3793221 sequences)
Timestamp : 20 Nov 2013 at 11:44:08 GMT
Top Score : 61 for gi|17232205, UDP-glucose 4-epimerase [Nostoc sp. PCC 7120]

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 78 are significant (p<0.05).



Concise Protein Summary Report

Format As [Concise Protein Summary] Help

Significance threshold p< [0.05] Max. number of hits [20]

Preferred taxonomy [All entries]

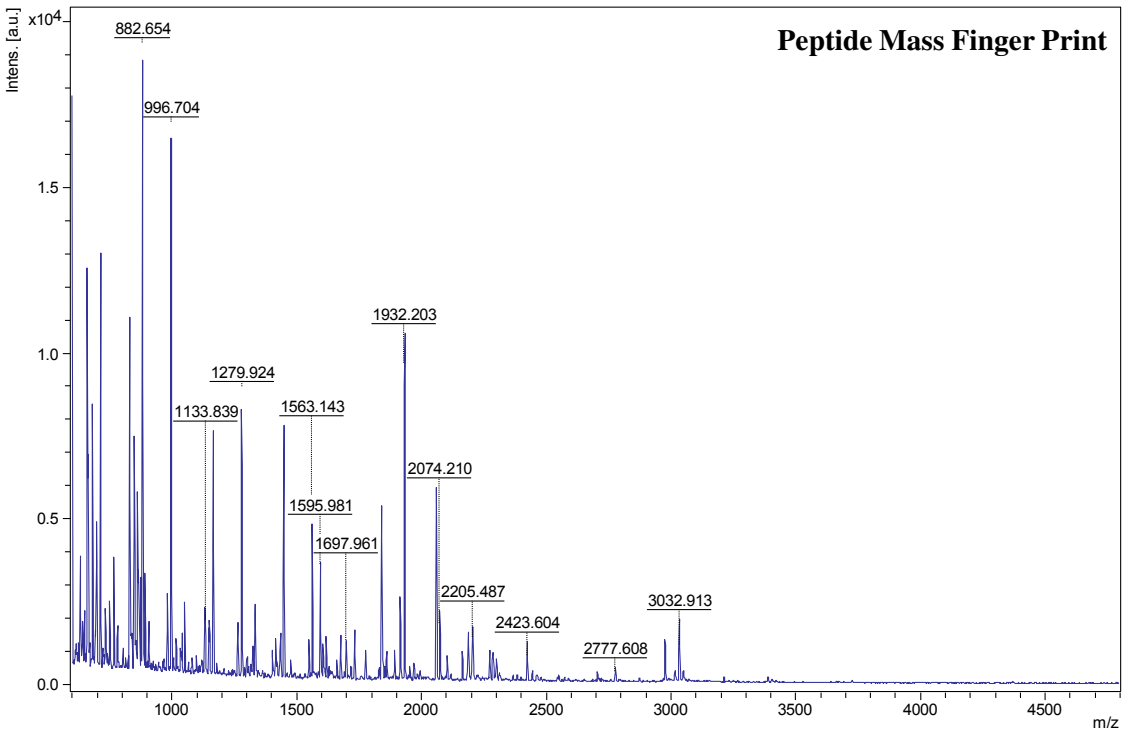
Re-Search All | Search Unmatched

1. gi|17232205 **Mass:** 36522 **Score:** 61 **Expect:** 3.3 **Matches:** 9
UDP-glucose 4-epimerase [Nostoc sp. PCC 7120]

Protein sequence coverage: 46%

Matched peptides shown in ***bold red***.

```
  1 MSPGKPSILV TGGAGYIGSH TVLALKQAGY DVVILDNLVY GHRDLVEKVL
 51 QVELVVGDTG DRPLLDELFK SRHFDAVMHF SAYAYVGESV SDPAKYYRNN
101 VLGTLTLLEA MLAASINKFV FSSTCATYGV PKTVPIPEDH PQNPINPYGA
151 TKLMVERILA DFDVAYGLKS VRFRYFNAAG ANPDGLLGED HNPETHLIPL
201 VLLTALGKRK FISIFGTDYP TPDGTCIRDY IHVNDLADAH VLGLKYLLKG
251 GDSEVFNLGN GQGFSVREVI AAGEQVTGLP ITVEECDRRP GDPPSLIGSG
301 EKARKILGWQ PQYSSIKDIV SHAWQWHQKR HQ
```



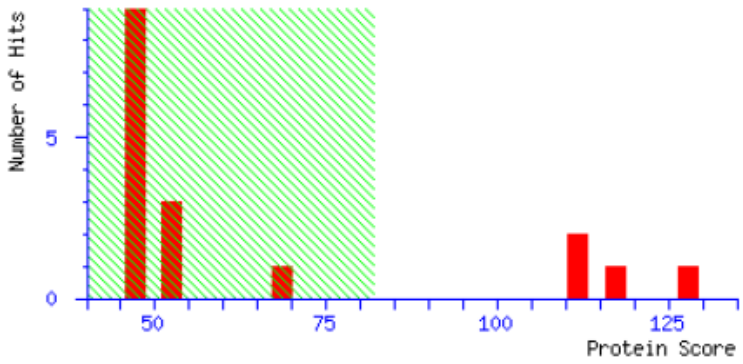
PetH-1

MATRIX SCIENCE Mascot Search Results

```
User            : arvind Kumar
Email           : parvind@barc.gov.in
Search title    : spot11
MS data file    : peaklist.xml
Database        : NCBInr 20160328 (84059059 sequences; 30803194823 residues)
Taxonomy        : Other Bacteria (8074972 sequences)
Timestamp       : 31 Mar 2016 at 10:07:44 GMT
Top Score       : 128 for gi|2498066, RecName: Full=Ferredoxin--NADP reductase; Short=FNR
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▾ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▾

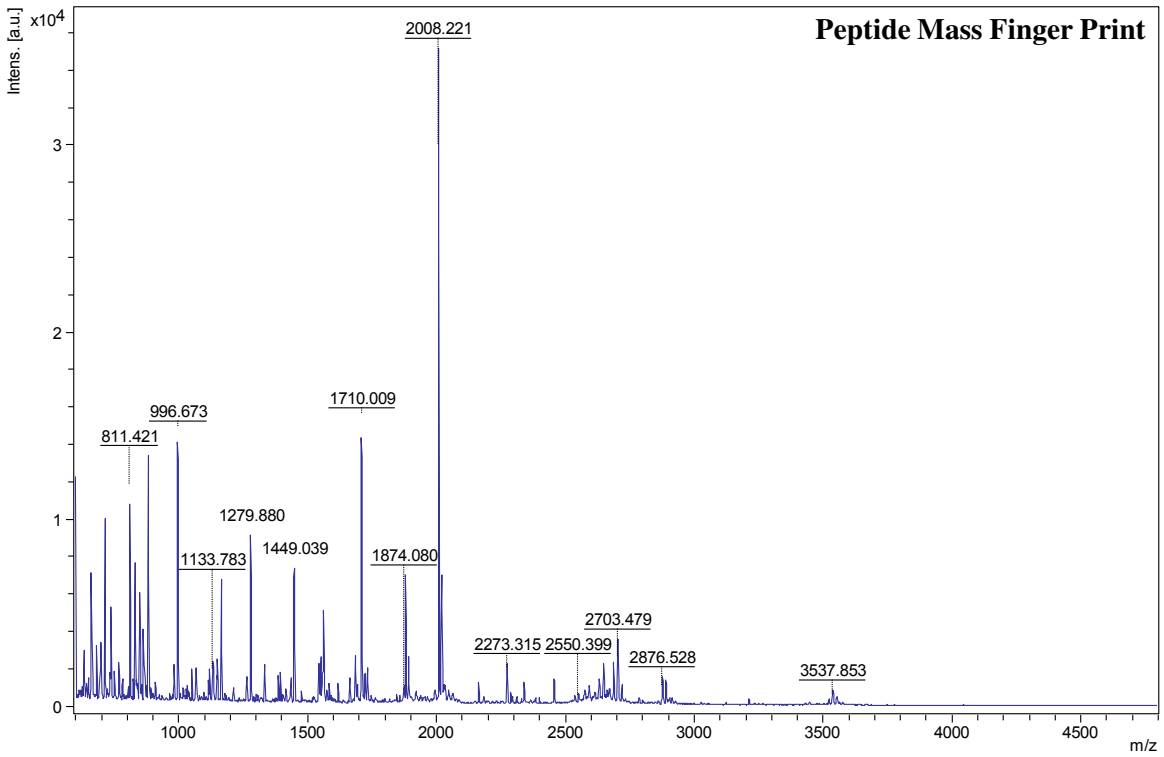
Re-Search All | Search Unmatched

4. gi|499307485 Mass: 48979 Score: 114 Expect: 3.2e-05 Matches: 18
ferredoxin--NADP(+) reductase [Nostoc sp. PCC 7120]

Protein sequence coverage: 45%

Matched peptides shown in ***bold red***.

```
  1 MSNQGAFDGA ANVESGSRVF VYEVVGMRQN EETDQTNYPI RKSGSVFIRV
 51 PYNRMNQEMQ RITRLGGKIV SIQTVSALQQ LNGRTTIATV TDASSEIAKS
101 EGNGKATPVK TDSGAKGFAK PPAEEQLKKK DNKGNTMTQA KAKHADVPVN
151 LYRPNAPFIG KVISNEPLVK EGGIGIVQHI KFDLTGGNLK YIEGQSIGII
201 PPGVDKNGKP EKLRLYSIAS TRHGDDVDDK TISLCVRQLE YKHPESGETV
251 YGVCSTYLTH IEPGSEVKIT GPVGKEMLLP DDPEANVIML ATGTGIAPMR
301 TYLWRMFKDA ERAANPEYQF KGFSWLVFGV PTTPNILYKE ELEEIQQKYP
351 DNFRLTYAIS REQKNPQGGR MYIQDRVAEH ADELWQLIKN EKTHTYICGL
401 RGMEEGIDAA LSAAAAKEGV TWSDYQKDLK KAGRWHVETY
```



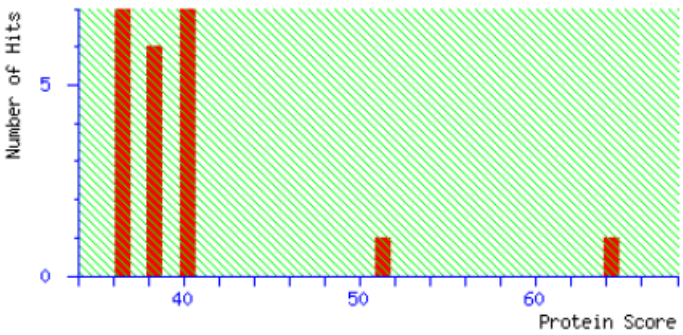
PetH-2

MATRIX SCIENCE Mascot Search Results

```
User            : Arvind Kumar
Email           : parvind@barc.gov.in
Search title    : Spot 13
MS data file    : peaklist.xml
Database        : NCBInr 20150912 (71310511 sequences; 25955008131 residues)
Taxonomy        : Other Bacteria (7064089 sequences)
Timestamp       : 16 Sep 2015 at 09:19:57 GMT
Top Score       : 64 for gi|296863364, Chain A, Ferredoxin-Nadp Reductase Mutant With Tyr 303
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 81 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▾ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▾

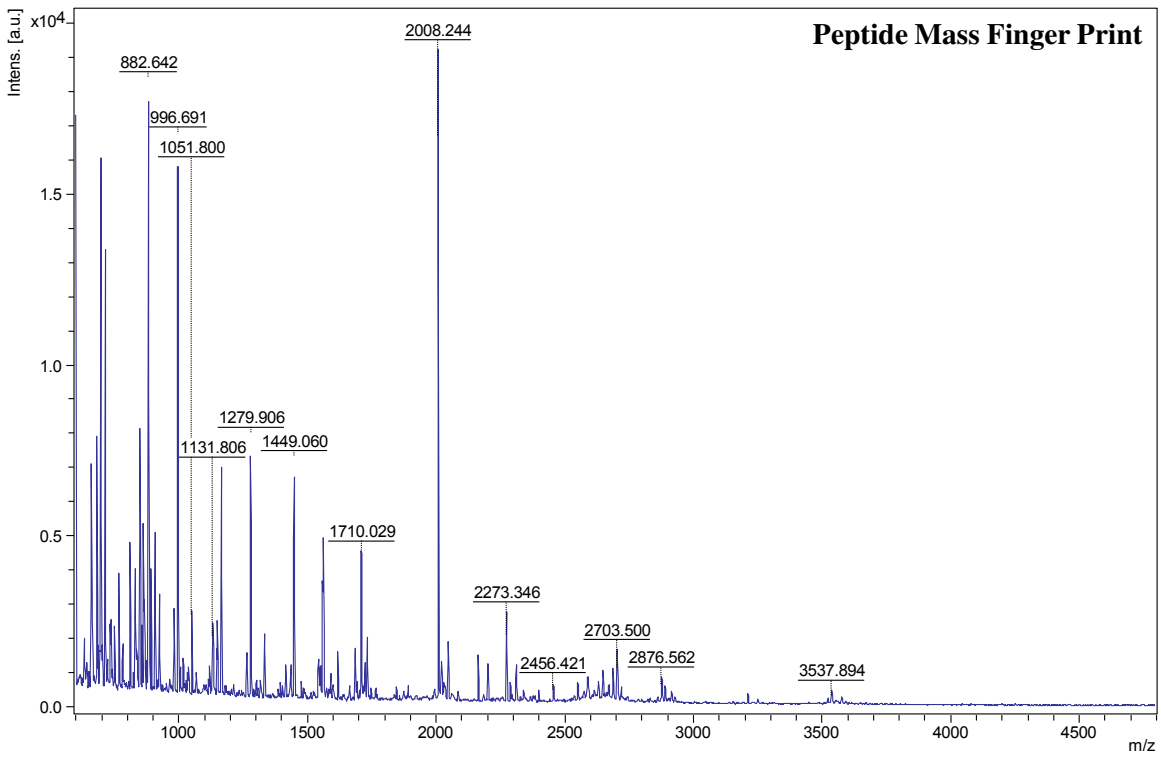
Re-Search All | Search Unmatched

```
1.    gi|296863364    Mass: 34131    Score: 64    Expect: 2.6  Matches: 9
      Chain A, Ferredoxin-Nadp Reductase Mutant With Tyr 303 Replaced By Phe (Y303f)
      gi|13096126    Mass: 34147    Score: 64    Expect: 2.6  Matches: 9
      Chain A, Anabaena Pcc7119 Ferredoxin:ferredoxin-Nadp+-Reductase Complex
      gi|558704689    Mass: 34132    Score: 64    Expect: 2.6  Matches: 9
      Chain A, Ferredoxin-nadp Reductase Mutant With Ser 80 Replaced By Ala (s80a)
      gi|39251    Mass: 34278    Score: 63    Expect: 3.6  Matches: 9
      ferredoxin--NADP(+) reductase [Anabaena variabilis]
      gi|499307485    Mass: 48979    Score: 53    Expect: 32  Matches: 9
      ferredoxin--NADP reductase [Nostoc sp. PCC 7120]
```

Protein sequence coverage: 41%

Matched peptides shown in ***bold red***.

```
  1 TQAKAKHADV PVNLYRPNAP FIGKVISNEP LVKEGGIGIV QHIKFDLTGG
 51 NLKYIEGQSI GIIPPGVDKN GKPEKLRLYS IASTRHGDDV DDKTISLCVR
101 QLEYKHPESG ETVYGVCSTY LTHIEPGSEV KITGPVGKEM LLPDDPEANV
151 IMLATGTGIA PMRTYLWRMF KDAERAANPE YQFKGFSWLV FGVPTTPNIL
201 YKEELEEIQQ KYPDNFRLTY AISREQKNPQ GGRMYIQDRV AEHADELWQL
251 IKNQKTHTYI CGLRGMEEGI DAALSAAAAK EGVTWSDYQK DLKKAGRWHV
301 ETF
```



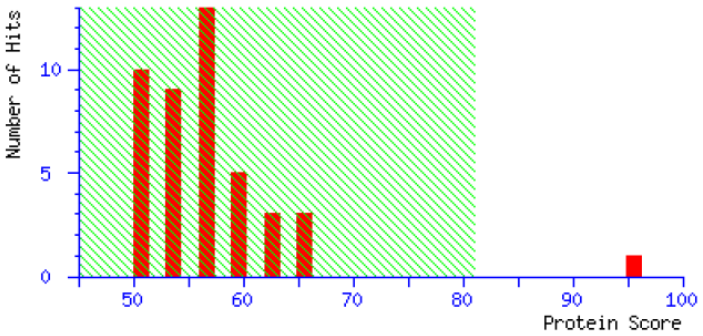
CpcC-1

MATRIX SCIENCE Mascot Search Results

```
User            : Arvind Kumar
Email           : parvind@barc.gov.in
Search title    : Spot 16
MS data file    : peaklist.xml
Database        : NCBInr 20150912 (71310511 sequences; 25955008131 residues)
Taxonomy        : Other Bacteria (7064089 sequences)
Timestamp       : 16 Sep 2015 at 09:49:17 GMT
Top Score       : 96 for gi|499303931, photosystem I reaction center subunit XII [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 81 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary ▼ Help

Significance threshold p< 0.05 Max. number of hits AUTO

Preferred taxonomy: All entries ▼

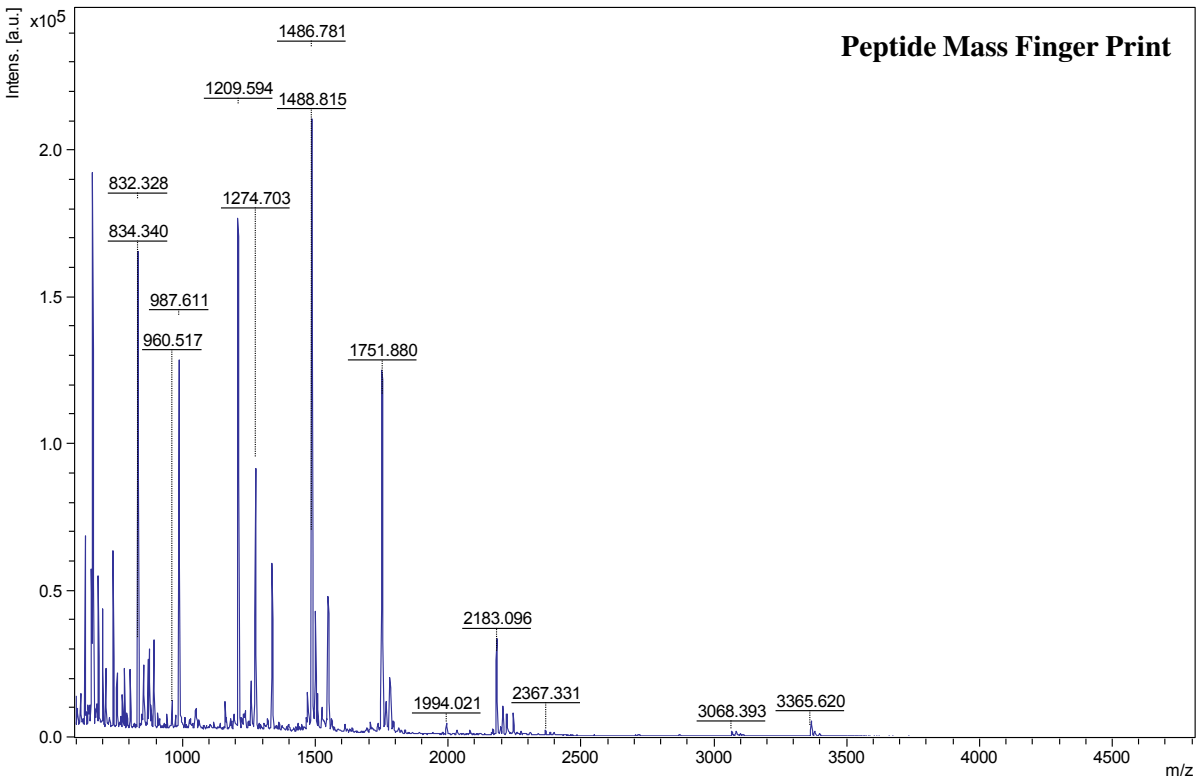
Re-Search All | Search Unmatched

1. gi|499303931 Mass: 32190 Score: 96 Expect: 0.002 Matches: 10
photosystem I reaction center subunit XII [Nostoc sp. PCC 7120]

Protein sequence coverage: 40%

Matched peptides shown in ***bold red***.

```
  1 MAITTAASRL GTEPFSDAPK VELRPKASRE EVESVIRAVY RHVLGNDYIL
 51 ASERLVSAES LLRDGNLTVR EFVRSVAKSE LYKKKFFYNS FQTRLIELNY
101 KHLLGRAPYD ESEVVYHLDL YQNKGYDAEI DSYIDSWEYQ SNFGDNVVPY
151 YRGFETQVGQ KTAGFNRIFR LYRGYANSDR AQVEGTKSRL ARELASNKAS
201 TIVGPSGTND SWGFRASADV APKKNLGNAV GEGDRVYRLE VTGIRSPGYP
251 SVRRSSTVFI VPYERLSDKI QQVHKQGGKI VSVTSA
```



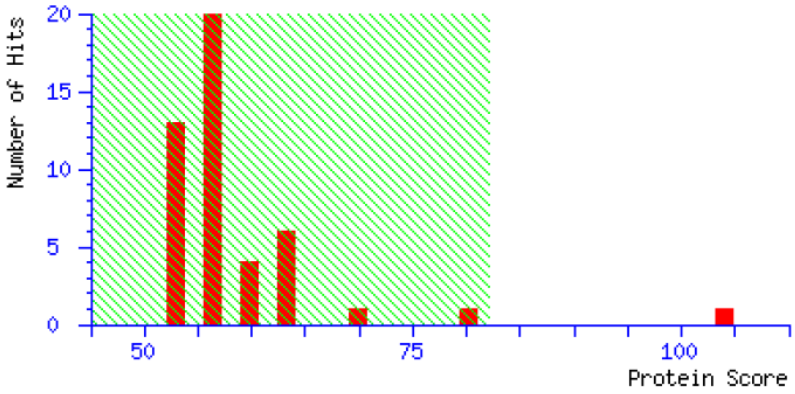
CpcC-2

MATRIX SCIENCE Mascot Search Results

```
User            : arvind Kumar
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 28 Mar 2016 at 10:52:44 GMT
Top Score       : 104 for gi|38896, unnamed protein product [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary ▾ | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries ▾

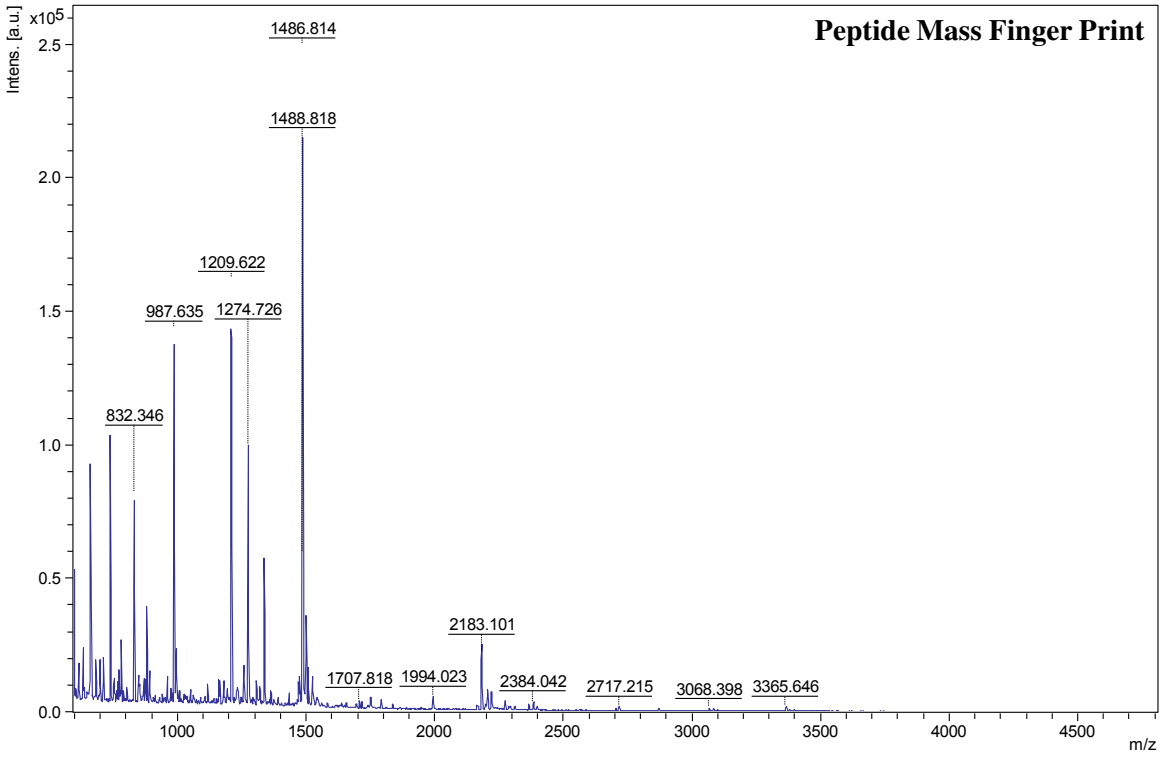
Re-Search All | Search Unmatched

1. gi|38896 Mass: 32133 Score: 104 Expect: 0.00032 Matches: 11
   unnamed protein product [Nostoc sp. PCC 7120]
   gi|499303931 Mass: 32190 Score: 104 Expect: 0.00032 Matches: 11
   photosystem I reaction center subunit XII [Nostoc sp. PCC 7120]

Protein sequence coverage: 44%

Matched peptides shown in ***bold red***.

```
  1 MAITTAASRL GTEPFSDAPK VELRPKASRE EVESVIRAVY RHVLGNDYIL
 51 ASERLVSAES LLRDGNLTVR EFVRSVAKSE LYKKKFFYNS FQTRLIELNY
101 KHLLGRAPYD ESEVVYHLDL YQNKGYDAEI DSYIDSWEYQ SNFGDNVVPY
151 YRGFETQVGQ KTAGFNRIFR LYRGYANSDR AQVEGTKSRL ARELASNKAS
201 TIVGPSGTND SWGFRASADV APKKNLGNAV GEGDRVYRLE VTGIRSPGYP
251 SVRRSSTVFI VPYERLSDKI QQVHKQGGKI VSVTSA
```



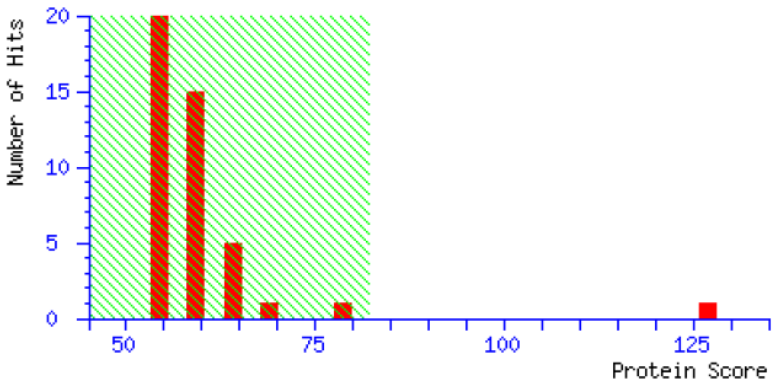
PrxA

MATRIX SCIENCE Mascot Search Results

```
User            : spot 21
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 28 Mar 2016 at 17:15:51 GMT
Top Score       : 127 for gi|499307997, peroxiredoxin [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary ▾ | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries ▾

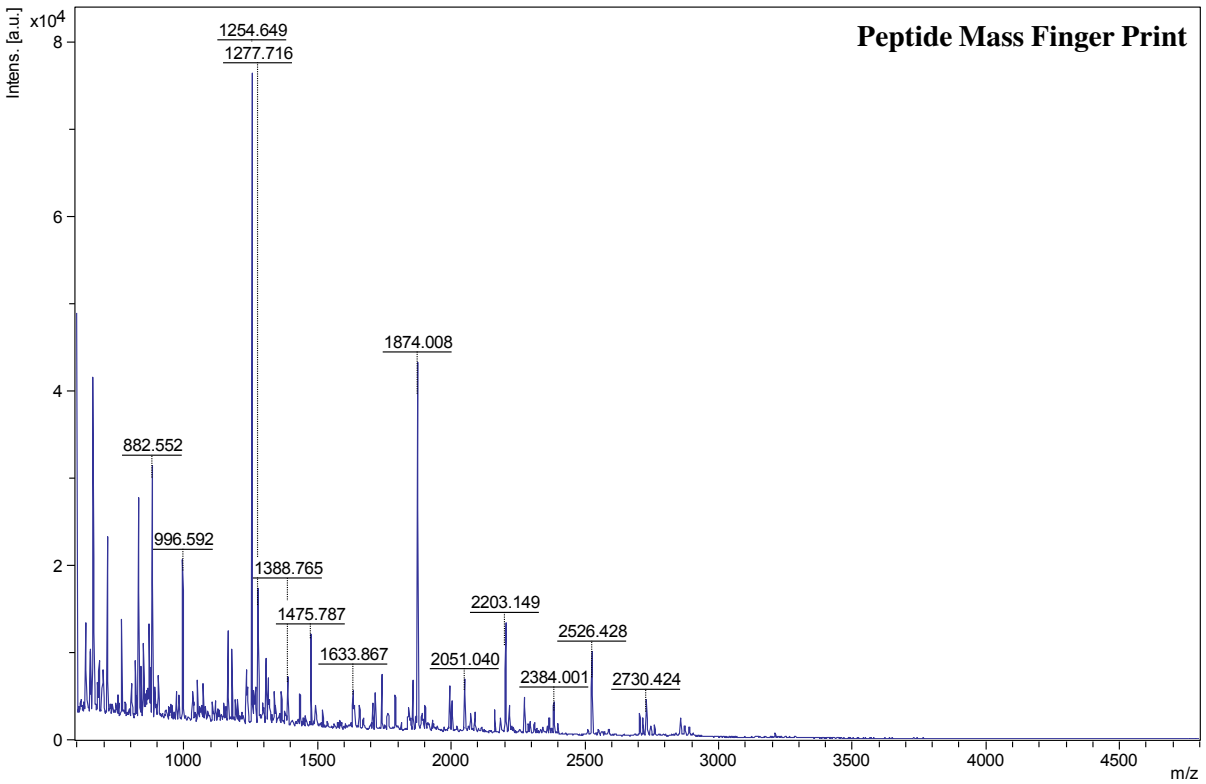
Re-Search All | Search Unmatched

1. gi|499307997 Mass: 22731 Score: 127 Expect: 1.6e-06 Matches: 10
peroxiredoxin [Nostoc sp. PCC 7120]

Protein sequence coverage: 59%

Matched peptides shown in ***bold red***.

```
  1 MSITYGTQES LRVGQQAPDF TATAVVDQEF KTIKLSDYRG KYVVLFFYPL
 51 DFTFVCPTEI TAFSDRYEEF KKLNTEILGV SVDSEFSHLA WIQTDRKSGG
101 VGDLNYPLVS DIKKEVSDAY NVLDPAAGIA LRGLFIIDKD GIIQHATINN
151 LAFGRSVDET LRTLQAIQYV QSHPDEVCPA GWQPGEKTMT PDPVKSKVYF
201 AAV
```



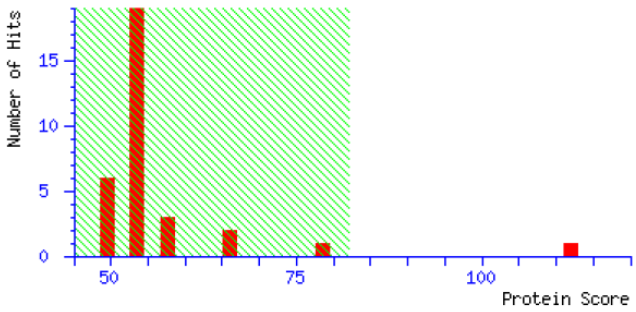
CpcG1-1

MATRIX SCIENCE Mascot Search Results

```
User            : spot 22
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 28 Mar 2016 at 17:20:26 GMT
Top Score       : 112 for gi|499303935, phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary ▾ | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries ▾

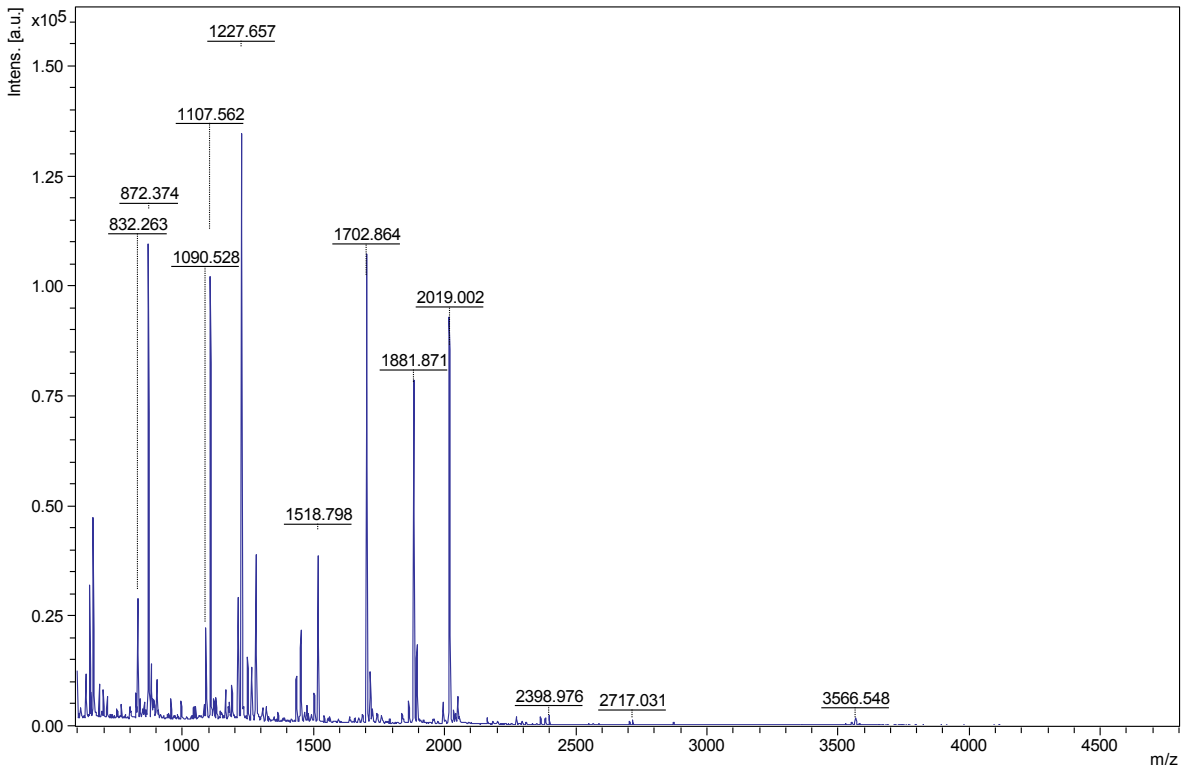
Re-Search All | Search Unmatched

1. gi|499303935 **Mass:** 31915 **Score:** 112 **Expect:** 5.1e-05 **Matches:** 11
phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]

Protein sequence coverage: 41%

Matched peptides shown in ***bold red***.

```
  1 MSIPLLEYAP SSQNQRVEGY EVPNEDTPTI YRLAAAIDDA DVDAIIWAGY
 51 RQIFSEHLII KSNRQSFLES QLRNRAINVR DFIRGLGKSE VYRTQVADLN
101 SNYRLVDITL KRFLGRAAYN QDEEIAWSIV IGSQGLHGFI DALLDSDEYR
151 ENFGDDIVPY QRRRYKDRPF NLVNPRYNAY WRDRQTLNAL GGRSFYSART
201 SGTLTKDDIR RAIPANFMAL AGKILTPERN YQRTIASVTS QIKDIKIPDT
251 SREVTTPEVT VKPVAVALPY RYIPGNKTT
```



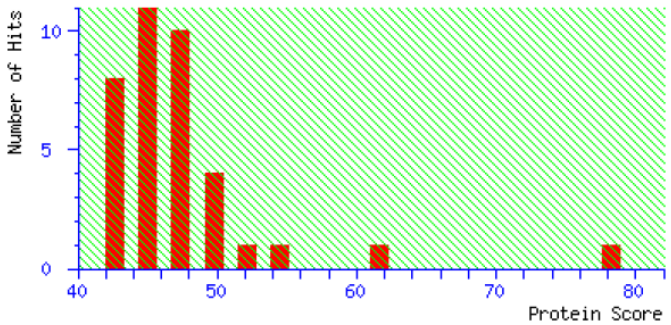
CpcG1-2

MATRIX SCIENCE Mascot Search Results

```
User            : spot 23
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 28 Mar 2016 at 17:27:08 GMT
Top Score       : 78 for gi|499303935, phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As [Concise Protein Summary ▾] Help

Significance threshold p< [0.05] Max. number of hits [AUTO]

Preferred taxonomy [All entries ▾]

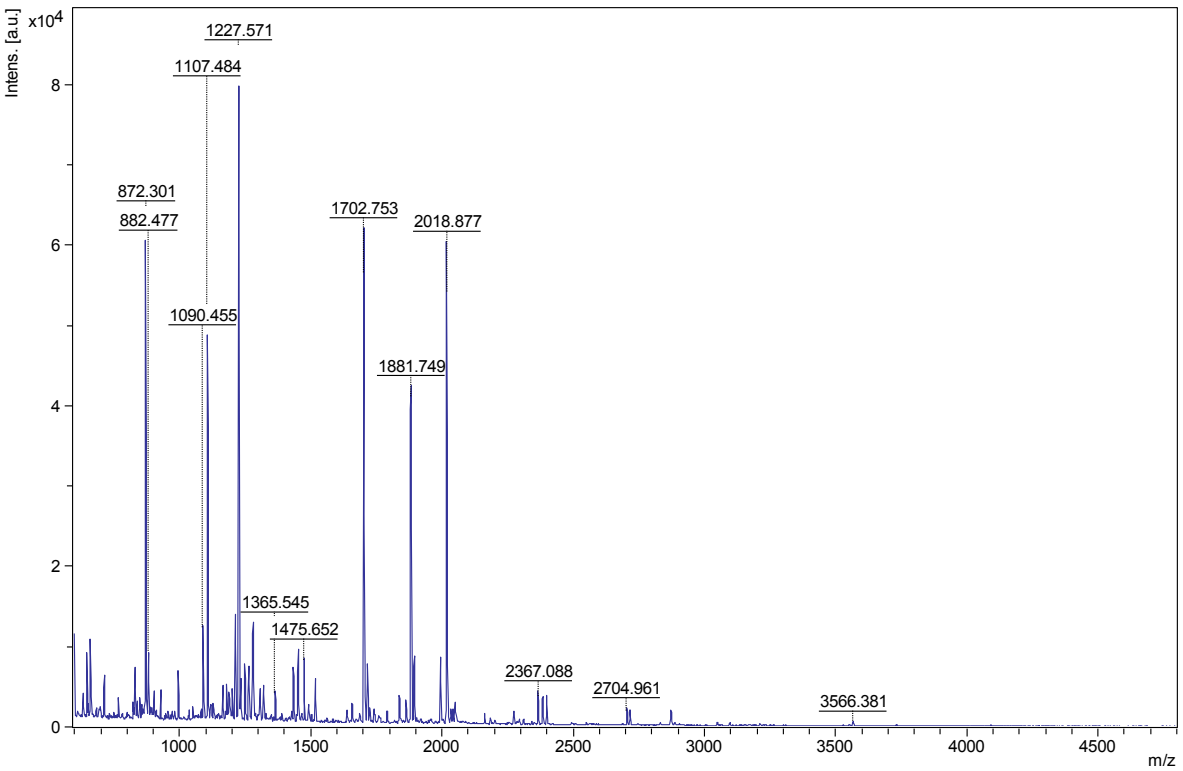
[Re-Search All] [Search Unmatched]

```
1.    gi|499303935    Mass: 31915    Score: 78    Expect: 0.12  Matches: 8
      phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]
```

Protein sequence coverage: 33%

Matched peptides shown in ***bold red***.

```
  1 MSIPLLEYAP SSQNQRVEGY EVPNEDTPTI YRLAAAIDDA DVDAIIWAGY
 51 RQIFSEHLII KSNRQSFLES QLRNRAINVR DFIRGLGKSE VYRTQVADLN
101 SNYRLVDITL KRFLGRAAYN QDEEIAWSIV IGSQGLHGFI DALLDSDEYR
151 ENFGDDIVPY QRRRYKDRPF NLVNPRYNAY WRDRQTLNAL GGRSFYSART
201 SGTLTKDDIR RAIPANFMAL AGKILTPERN YQRTIASVTS QIKDIKIPDT
251 SREVTTPEVT VKPVAVALPY RYIPGNKTT
```



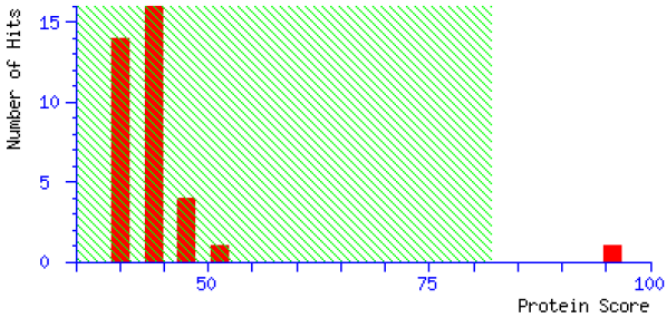
CpcG1-3

MATRIX SCIENCE Mascot Search Results

```
User            : spot 24
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 28 Mar 2016 at 17:31:14 GMT
Top Score       : 96 for gi|499303935, phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As [Concise Protein Summary ▾] Help

Significance threshold p< [0.05] Max. number of hits [AUTO]

Preferred taxonomy [All entries ▾]

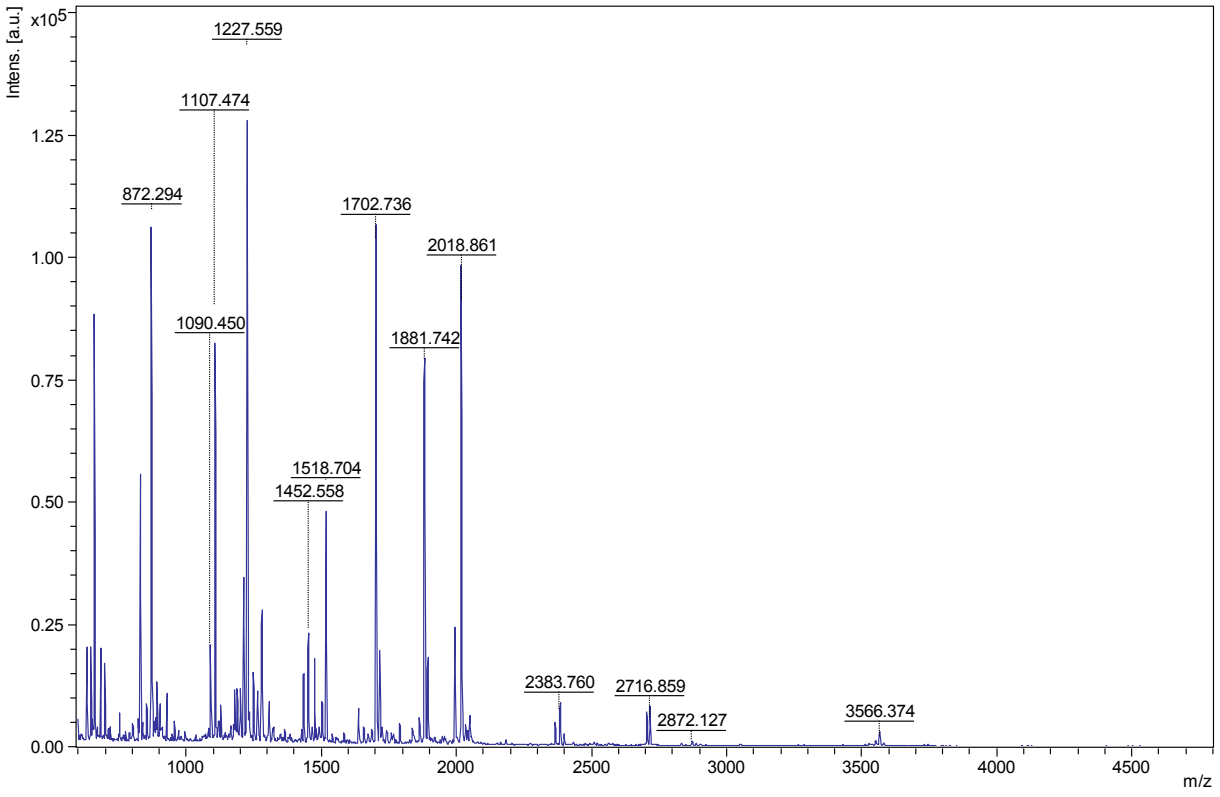
[Re-Search All] [Search Unmatched]

1. gi|499303935 Mass: 31915 Score: 96 Expect: 0.0022 Matches: 9
phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]

Protein sequence coverage: 37%

Matched peptides shown in ***bold red***.

```
  1 MSIPLLEYAP SSQNQRVEGY EVPNEDTPTI YRLAAAIDDA DVDAIIWAGY
 51 RQIFSEHLII KSNRQSFLES QLRNRAINVR DFIRGLGKSE VYRTQVADLN
101 SNYRLVDITL KRFLGRAAYN QDEEIAWSIV IGSQGLHGFI DALLDSDEYR
151 ENFGDDIVPY QRRRYKDRPF NLVNPRYNAY WRDRQTLNAL GGRSFYSART
201 SGTLTKDDIR RAIPANFMAL AGKILTPERN YQRTIASVTS QIKDIKIPDT
251 SREVTTPEVT VKPVAVALPY RYIPGNKTT
```



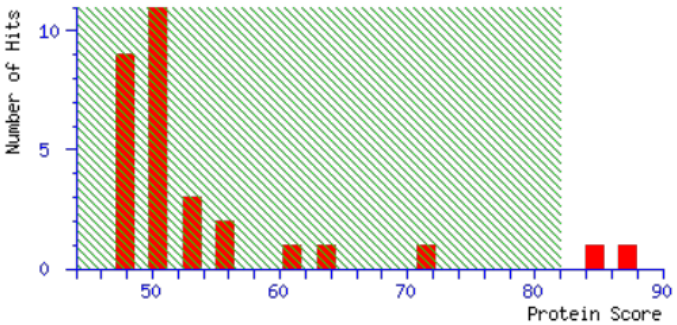
Ndk

MATRIX SCIENCE Mascot Search Results

```
User            : spot 26
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 06:28:58 GMT
Top Score       : 87 for gi|835293708, hypothetical protein [Chryseobacterium sp. YR561]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries

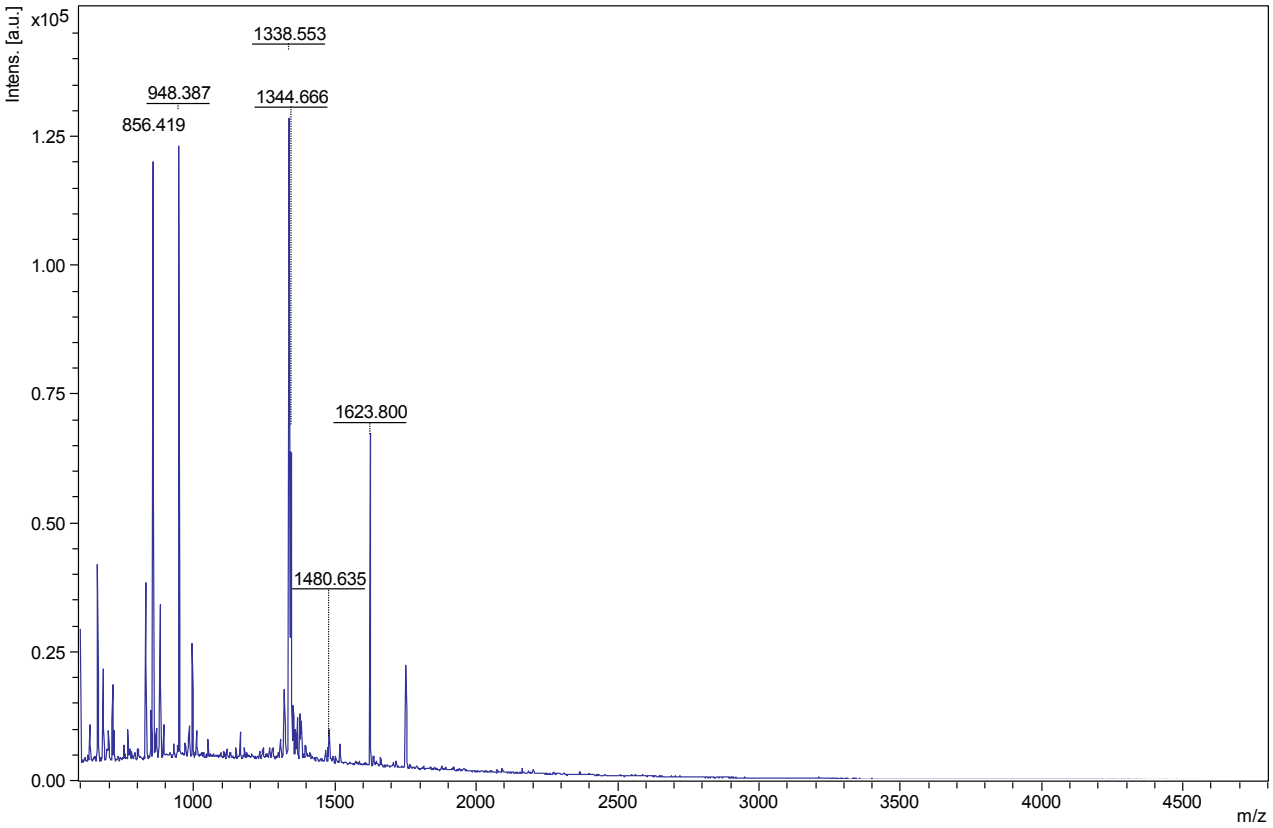
Re-Search All | Search Unmatched

2. gi|670494352 Mass: 26183 Score: 85 Expect: 0.023 Matches: 7
dimethylmenaquinone methyltransferase [Arenibacter algicola]
gi|499639412 Mass: 16594 Score: 70 Expect: 0.81 Matches: 5
nucleoside-diphosphate kinase [Nostoc sp. PCC 7120]

Protein sequence coverage: 36%

Matched peptides shown in ***bold red***.

```
  1 MERTFLAIKP DGVQRGLVGE IIRRFETKGF TLVGLKFLQV SKELAEQHYG
 51 VHRERPFFPS LVEFITSGPV VAMVWEGDGV IASARKIIGA TNPLTAEPGT
101 IRGDFGINIG RNLIHGSDAP ETAQKEVSLW FTDAELVNWQ PHLTPWLHE
```



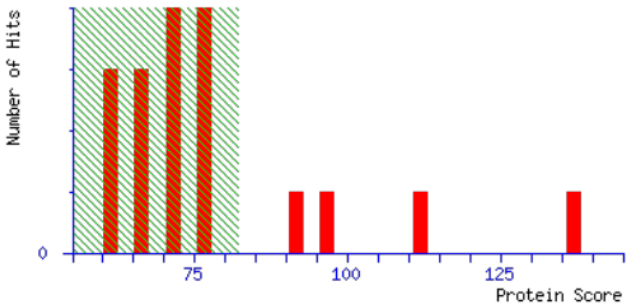
RbcL-1

MATRIX SCIENCE Mascot Search Results

```
User            : spot 32
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 07:22:34 GMT
Top Score       : 137 for gi|504925593, ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy All entries

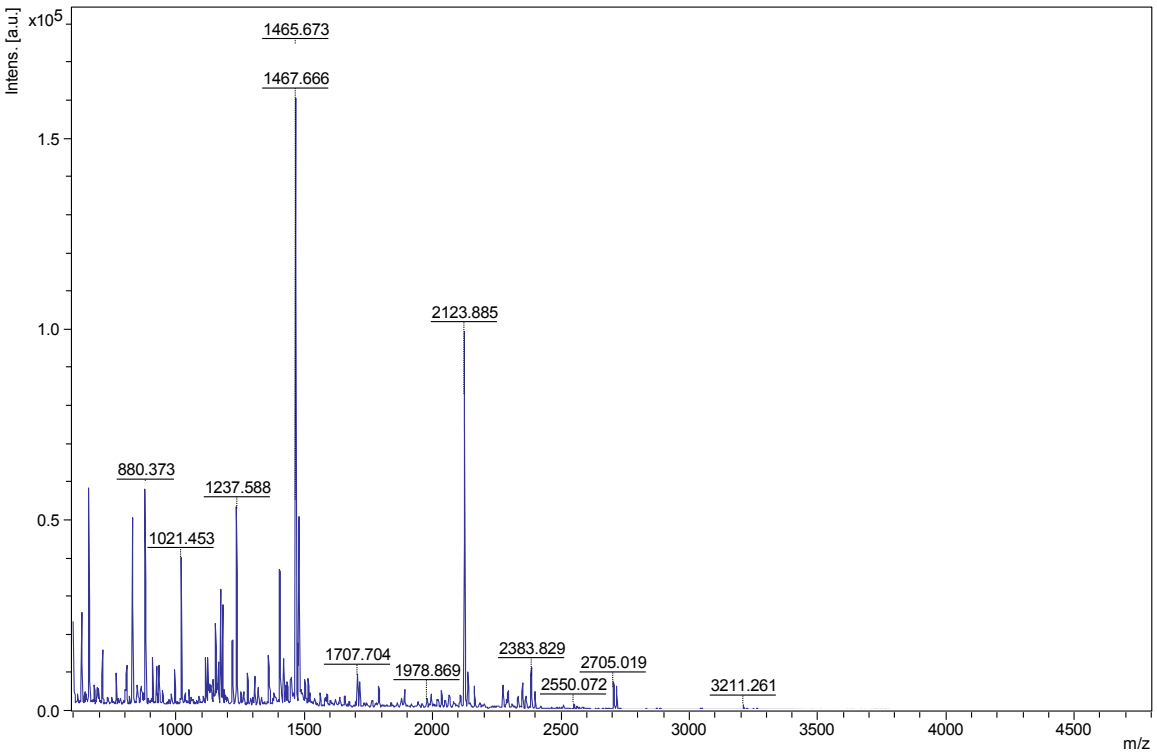
Re-Search All | Search Unmatched

```
1.    gi|504925593    Mass: 53409    Score: 137    Expect: 1.6e-07  Matches: 21
      ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
      gi|499304918    Mass: 53411    Score: 124    Expect: 3.2e-06  Matches: 20
      ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7120]
```

Protein sequence coverage: 35%

Matched peptides shown in ***bold red***.

```
  1 MSYAQTKTQA KSGYKAGVQD YRLTYYTPDY TPKDTDLLAA FRVTPQPGVP
 51 FEEAAAAVAA ESSTGTWTTV WTDLLTDLDR YKGRCYDIEP VPGEDNQFIA
101 YIAYPLDLFE EGSITNVLTS IVGNVFGFKA LRALRLEDLR FPVAYIKTFQ
151 GPPHGIQVER DKLNKYGRPL LGCTIKPKLG LSAKNYGRAV YECLRGGLDF
201 TKDDENINSA PFQRWRDRFL FVADAITKAQ AETGEIKGHY LNVTAPTCEE
251 MLKRAEYAKE LKQPIIMHDY LTAGFTANTT LARWCRDNGV LLHIHRAMHA
301 VIDRQKNHGI HFRVLAKALR LSGGDHIHTG TVVGKLEGER GITMGFVDLL
351 RENYVEQDKS RGIYFTQDWA SLPGVMAVAS GGIHIWHMPA LVEIFGDDSV
401 LQFGGGTLGH PWGNAPGATA NRVALEACVQ ARNEGRNLAR EGNDIIREAA
451 KWSPELAVAC ELWKEIKFEF EAMDTV
```



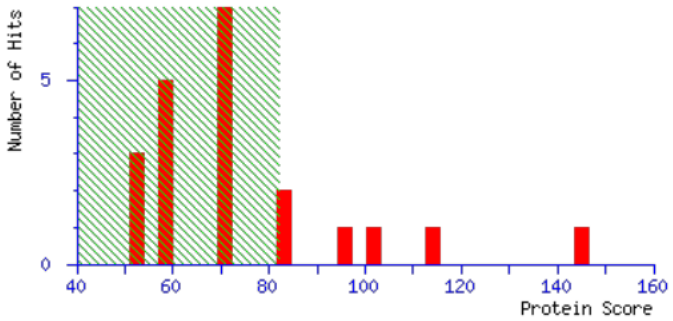
RbcL-2

MATRIX SCIENCE Mascot Search Results

```
User            : spot 33
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 07:38:52 GMT
Top Score       : 145 for gi|504925593, ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries

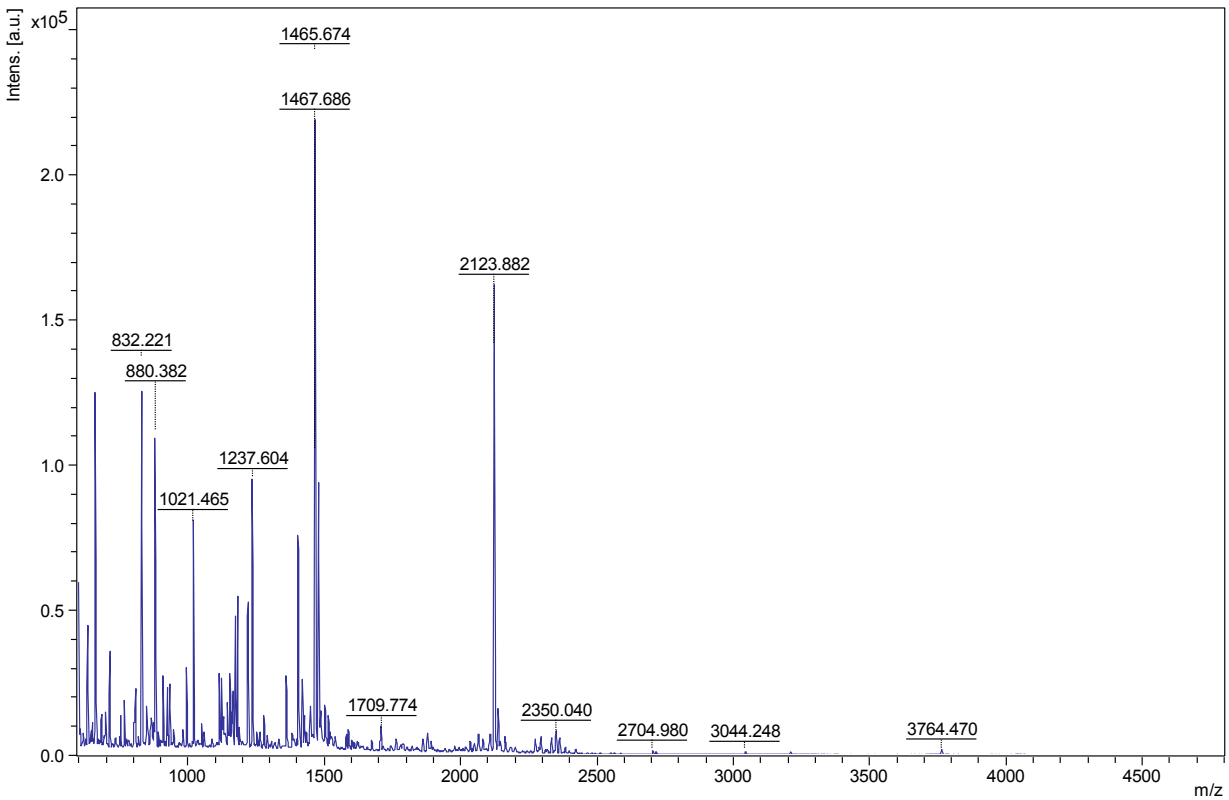
Re-Search All | Search Unmatched

1. gi|504925593 Mass: 53409 Score: 145 Expect: 2.5e-08 Matches: 21
   ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
   gi|499304918 Mass: 53411 Score: 129 Expect: 1e-06 Matches: 20
   ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7120]

Protein sequence coverage: 34%

Matched peptides shown in ***bold red***.

```
  1 MSYAQTKTQT KSGYKAGVQD YRLTYYTPDY TPKDTDILAA FRVTPQPGVP
 51 FEEAAAAVAA ESSTGTWTTV WTDLLTDLDR YKGRCYDIEP VPGEDNQFIA
101 YIAYPLDLFE EGSITNVLTS IVGNVFGFKA LRALRLEDIR FPVAYIKTFQ
151 GPPHGIQVER DKLNKYGRPL LGCTIKPKLG LSAKNYGRAV YECLRGGLDF
201 TKDDENINSA PFQRWRDRFL FVADAITKAQ AETGEIKGHY LNVTAPTCEE
251 MLKRAEYAKE LKQPIIMHDY LTAGFTANTT LARWCRDNGV LLHIHRAMHA
301 VIDRQKNHGI HFRVLAKALR LSGGDHIHTG TVVGKLEGER GITMGFVDLL
351 RENYVEQDKS RGIYFTQDWA SLPGVMAVAS GGIHVWHMPA LVEIFGDDSV
401 LQFGGGTLGH PWGNAPGATA NRVALEACVQ ARNEGRNLAR EGNDVIREAA
451 KWSPELAVAC ELWKEIKFEF EAMDTV
```



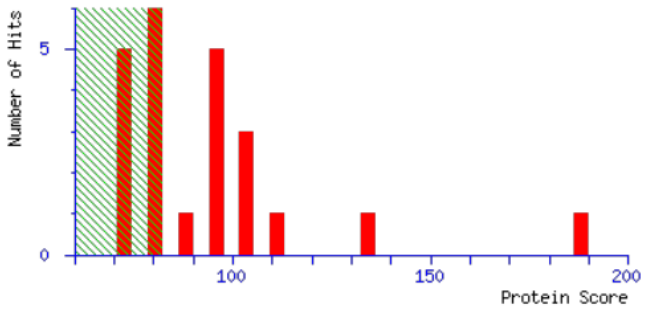
RbcL-3

MATRIX SCIENCE Mascot Search Results

```
User            : spot 34
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 09:18:04 GMT
Top Score       : 188 for gi|504925593, ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries

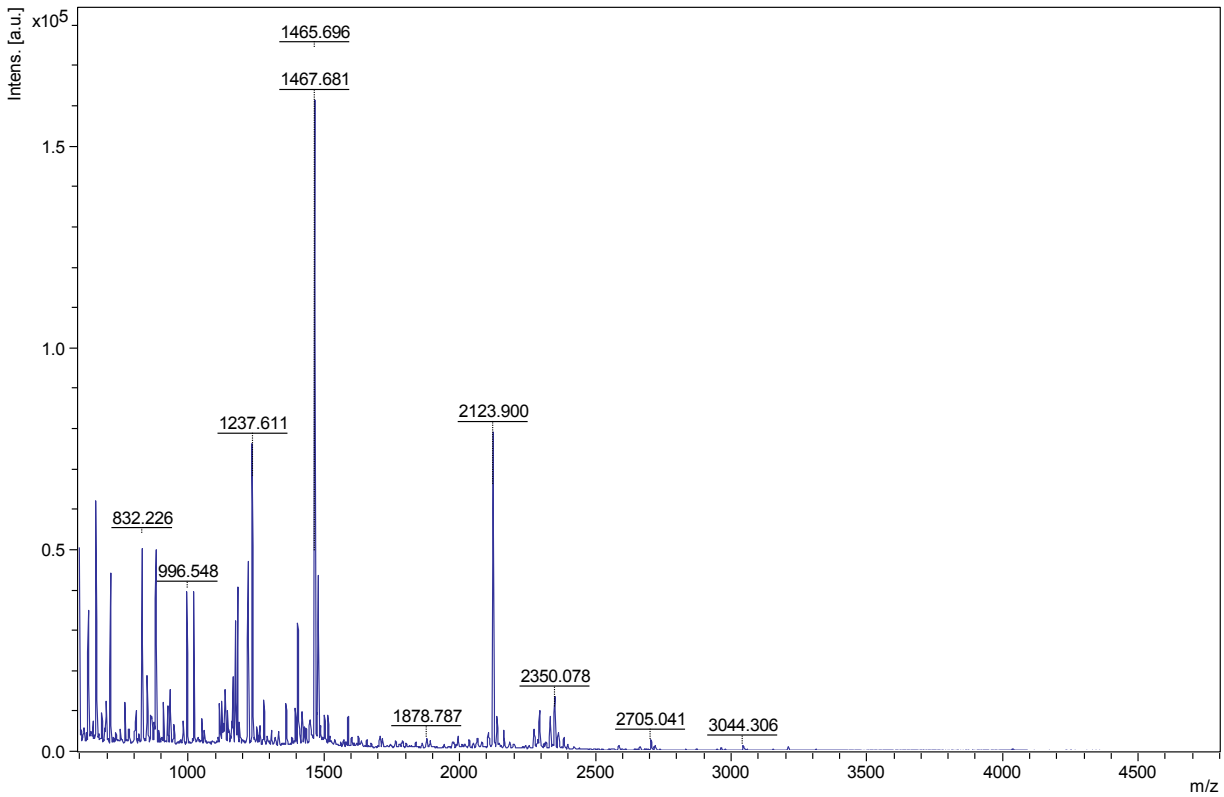
Re-Search All | Search Unmatched

```
1.    gi|504925593    Mass: 53409    Score: 188    Expect: 1.3e-12  Matches: 23
      ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7107]
      gi|499304918    Mass: 53411    Score: 174    Expect: 3.2e-11  Matches: 22
      ribulose bisphosphate carboxylase large chain [Nostoc sp. PCC 7120]
```

Protein sequence coverage: 36%

Matched peptides shown in ***bold red***.

```
  1 MSYAQTKTQT KSGYKAGVQD YRLTYYTPDY TPKDTDILAA FRVTPQPGVP
 51 FEEAAAAVAA ESSTGTWTTV WTDLLTDLDR YKGRCYDIEP VPGEDNQFIA
101 YIAYPLDLFE EGSITNVLTS IVGNVFGFKA LRALRLEDIR FPVAYIKTFQ
151 GPPHGIQVER DKLNKYGRPL LGCTIKPKLG LSAKNYGRAV YECLRGGLDF
201 TKDDENINSA PFQRWRDRFL FVADAITKAQ AETGEIKGHY LNVTAPTCEE
251 MLKRAEYAKE LKQPIIMHDY LTAGFTANTT LARWCRDNGV LLHIHRAMHA
301 VIDRQKNHGI HFRVLAKALR LSGGDHIHTG TVVGKLEGER GITMGFVDLL
351 RENYVEQDKS RGIYFTQDWA SLPGVMAVAS GGIHVWHMPA LVEIFGDDSV
401 LQFGGGTLGH PWGNAPGATA NRVALEACVQ ARNEGRNLAR EGNDVIREAA
451 KWSPELAVAC ELWKEIKFEF EAMDTV
```



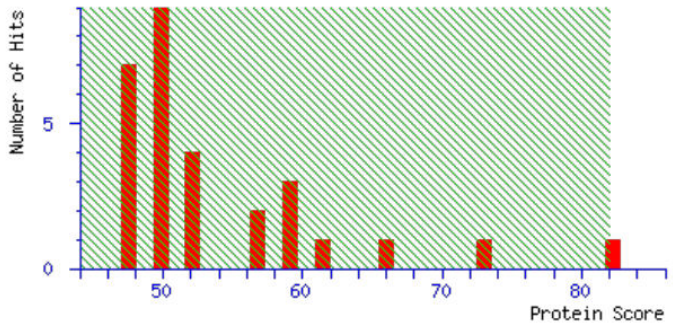
PetH

MATRIX SCIENCE Mascot Search Results

```
User            : spot 36
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 10:12:16 GMT
Top Score       : 82 for gi|2498066, RecName: Full=Ferredoxin--NADP reductase; Short=FNR
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries

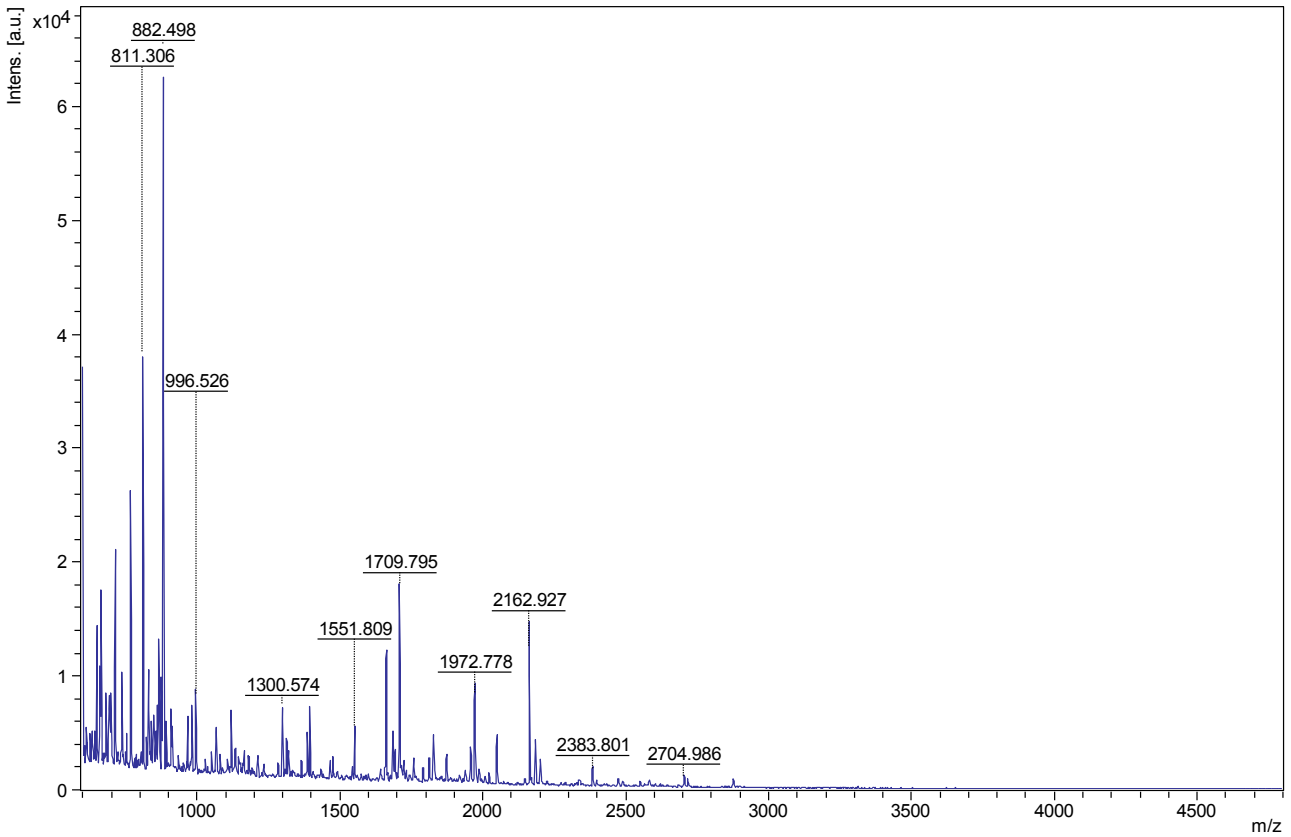
Re-Search All | Search Unmatched

1. gi|2498066 Mass: 48967 Score: 82 Expect: 0.049 Matches: 12
RecName: Full=Ferredoxin--NADP reductase; Short=FNR
gi|499307485 Mass: 48979 Score: 82 Expect: 0.052 Matches: 12
ferredoxin--NADP(+) reductase [Nostoc sp. PCC 7120]

Protein sequence coverage: 28%

Matched peptides shown in ***bold red***.

```
  1 MSNQGAFDGA ANVESGSRVF VYEVVGMRQN EETDQTNYPI RKSGSVFIRV
 51 PYNRMNQEMQ RITRLGGKIV SIQTVSALQQ LNGRTTIATV TDASSEIAKS
101 EGNGKATPVK TDSGAKGFAK PPAEEQLKKK DNKGNTMTQA KAKHADVPVN
151 LYRPNAPFIG KVISNEPLVK EGGIGIVQHI KFDLTGGNLK YIEGQSIGII
201 PPGVDKNGKP EKLRLYSIAS TRHGDDVDDK TISLCVRQLE YKHPESGETV
251 YGVCSTYLTH IEPGSEVKIT GPVGKEMLLP DDPEANVIML ATGTGIAPMR
301 TYLWRMFKDA ERAANPEYQF KGFSWLVFGV PTTPNILYKE ELEEIQQKYP
351 DNFRLTYAIS REQKNPQGGR MYIQDRVAEH ADELWQLIKN EKTHTYICGL
401 RGMEEGIDAA LSAAAAKEGV TWSDYQKDLK KAGRWHVETY
```



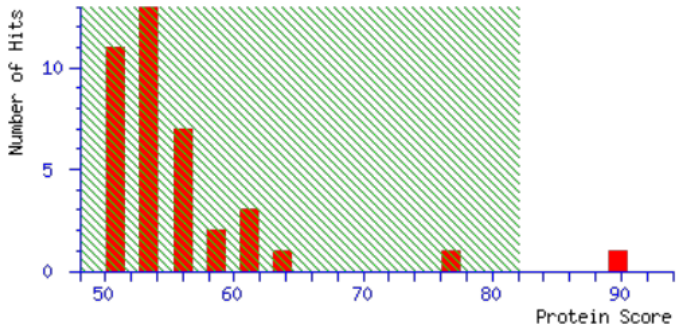
CpcC-3

MATRIX SCIENCE Mascot Search Results

```
User            : spot 37
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 11:38:03 GMT
Top Score       : 90 for gi|38896, unnamed protein product [Nostoc sp. PCC 7120]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits AUTO

Preferred taxonomy | All entries

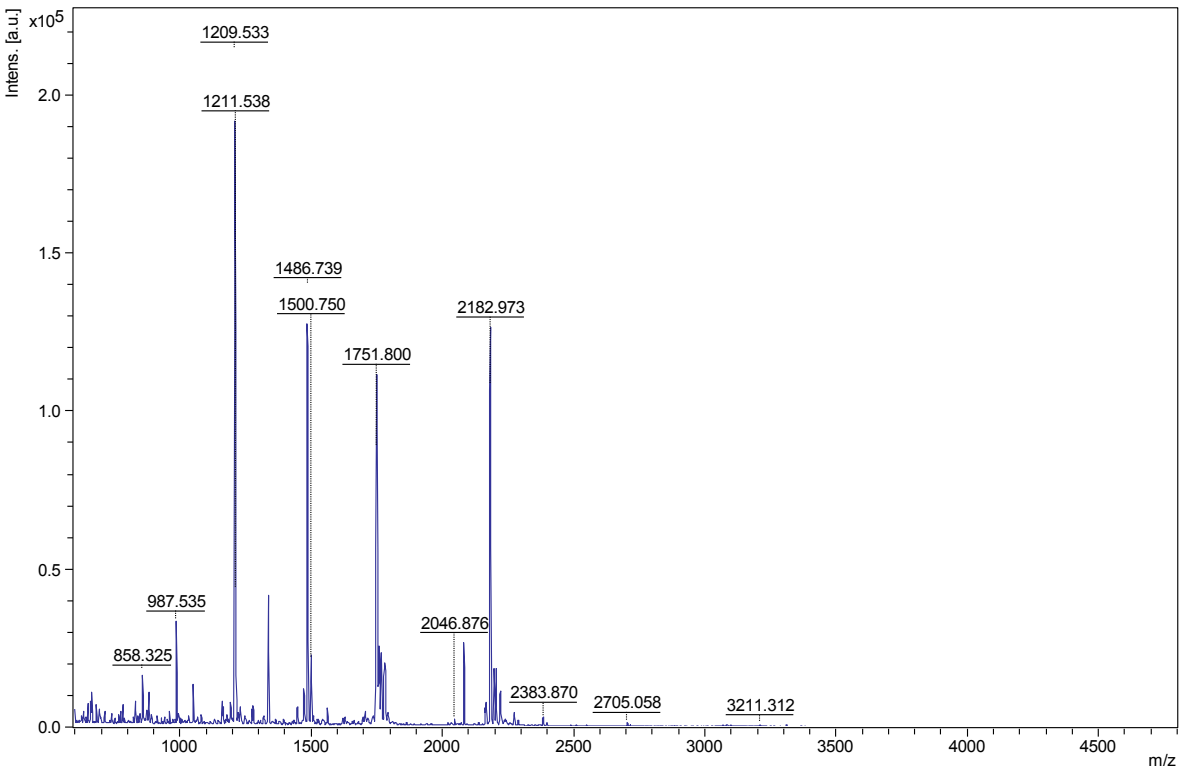
Re-Search All | Search Unmatched

1. gi|38896 Mass: 32133 Score: 90 Expect: 0.0086 Matches: 11
unnamed protein product [Nostoc sp. PCC 7120]
gi|499303931 Mass: 32190 Score: 89 Expect: 0.0097 Matches: 11
photosystem I reaction center subunit XII [Nostoc sp. PCC 7120]

Protein sequence coverage: 37%

Matched peptides shown in ***bold red***.

```
  1 MAITTAASRL GTEPFSDAPK VELRPKASRE EVESVIRAVY RHVLGNDYIL
 51 ASERLVSAES LLRDGNLTVR EFVRSVAKSE LYKKKFFYNS FQTRLIELNY
101 KHLLGRAPYD ESEVVYHLDL YQNKGYDAEI DSYIDSWEYQ SNFGDNVVPY
151 YRGFETQVGQ KTAGFNRIFR LYRGYANSDR AQVEGTKSRL ARELASNKAS
201 TIVGPSGTND SWGFRASADV APKKNLGNAV GEGDRVYRLE VTGIRSPGYP
251 SVRRSSTVFI VPYERLSDKI QQVHKQGGKI VSVTSA
```



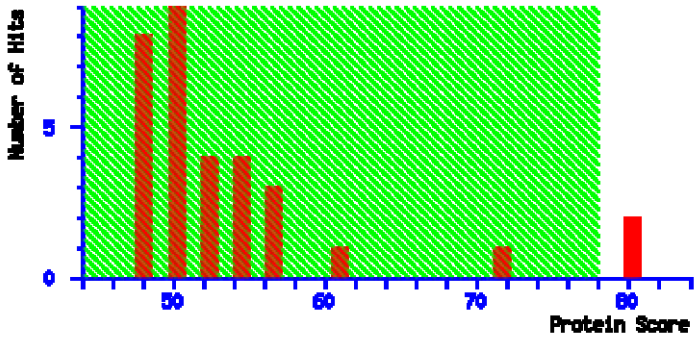
AhpC

MATRIX SCIENCE Mascot Search Results

User : Bhakti Basu
Email : bhaktibasu@gmail.com
Search title :
Database : NCBInr 20131113 (34201960 sequences; 12001222213 residues)
Taxonomy : Other Bacteria (3793221 sequences)
Timestamp : 12 Dec 2013 at 11:45:39 GMT
Top Score : 80 for gi|317135045, peroxiredoxin [Nostoc flagelliforme NX-09]

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 78 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary — Help

Significance threshold p< 0.05 — Max. number of hits 20

Preferred taxonomy: All entries

Re-Search All | Search Unmatched

1. gi|317135045 **Mass:** 23856 **Score:** 80 **Expect:** 0.036 **Matches:** 8
peroxiredoxin [Nostoc flagelliforme NX-09]
gi|427727311 **Mass:** 23879 **Score:** 51 **Expect:** 27 **Matches:** 6
peroxiredoxin [Nostoc sp. PCC 7524]

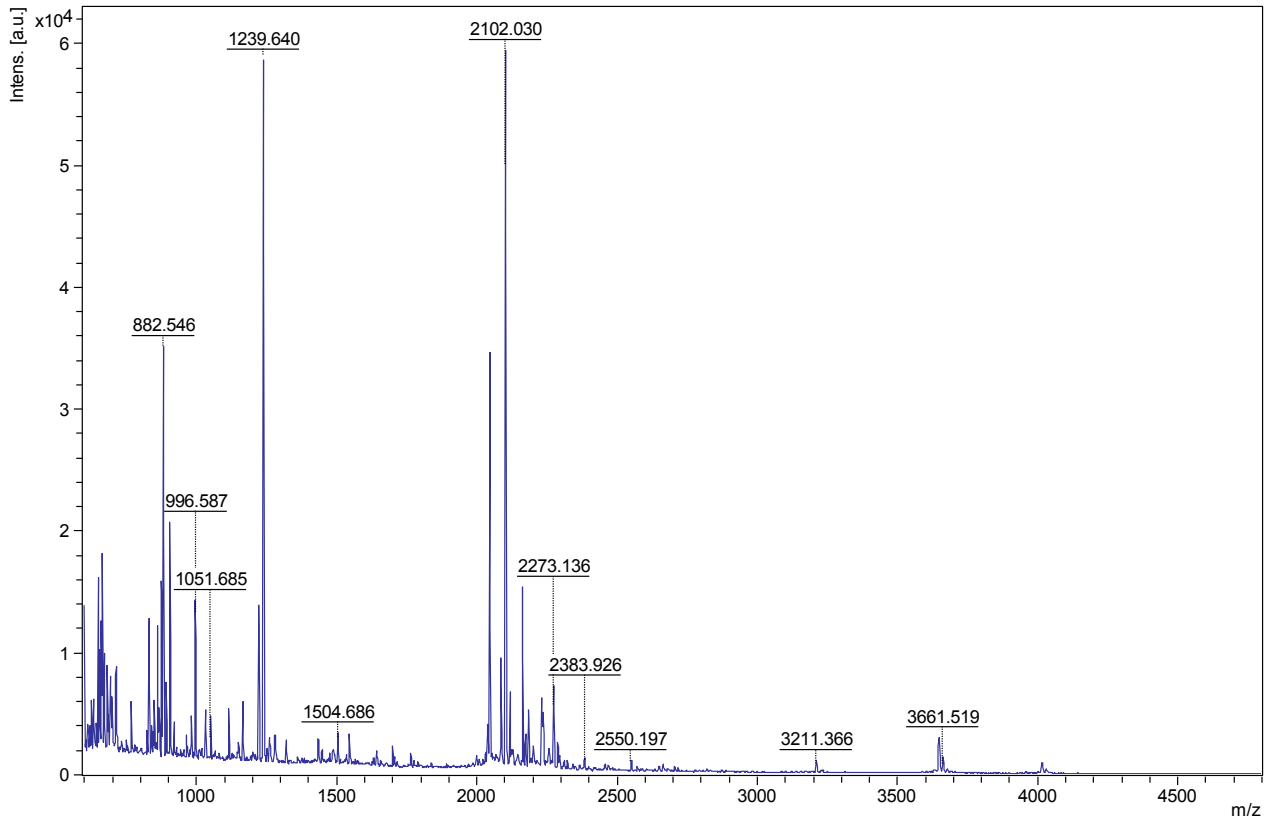
---

2. gi|17231896 **Mass:** 23740 **Score:** 79 **Expect:** 0.045 **Matches:** 8
AhpC/TSA family protein [Nostoc sp. PCC 7120]

Protein Sequence Coverage: 39 %

Matched peptide shown in ***bold red***

```
  1 MSLRLGDTVP  NFTQASTHGD  IDFYQWAGDS  WVVLFSHPAD  FTPVCTTELG
 51 TVAKLKPEFD  KRNVKAIALS  VDDVESHKGW  VGDIEETQST  TLNYPILADA
101 DRKVSDLYDM  IHPNANAAVT  VRSVFVIDPN  KKLRLTFTYP  PSTGRNFDEL
151 LRVIDSLQLT  DNYSVATPAD  WKDGEDVVIV  PSLKDPEVLK  EKFPKGYEEI
201 KPYLRITPQP  NK
```



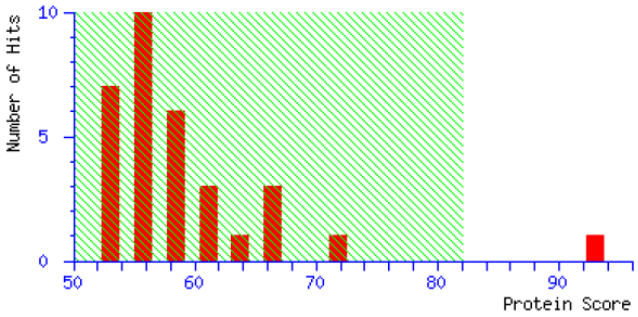
CpcG1-4

MATRIX SCIENCE Mascot Search Results

```
User            : spot 42
Email           : parvind@barc.gov.in
Search title    :
MS data file    : peaklist.xml
Database        : NCBInr 20160321 (83784708 sequences; 30681960259 residues)
Taxonomy        : Other Bacteria (8060302 sequences)
Timestamp       : 29 Mar 2016 at 16:40:09 GMT
Top Score       : 93 for gi|499303935, phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 82 are significant (p<0.05).



Concise Protein Summary Report

Format As [Concise Protein Summary ▾] Help

Significance threshold p< [0.05] Max. number of hits [AUTO]

Preferred taxonomy [All entries ▾]

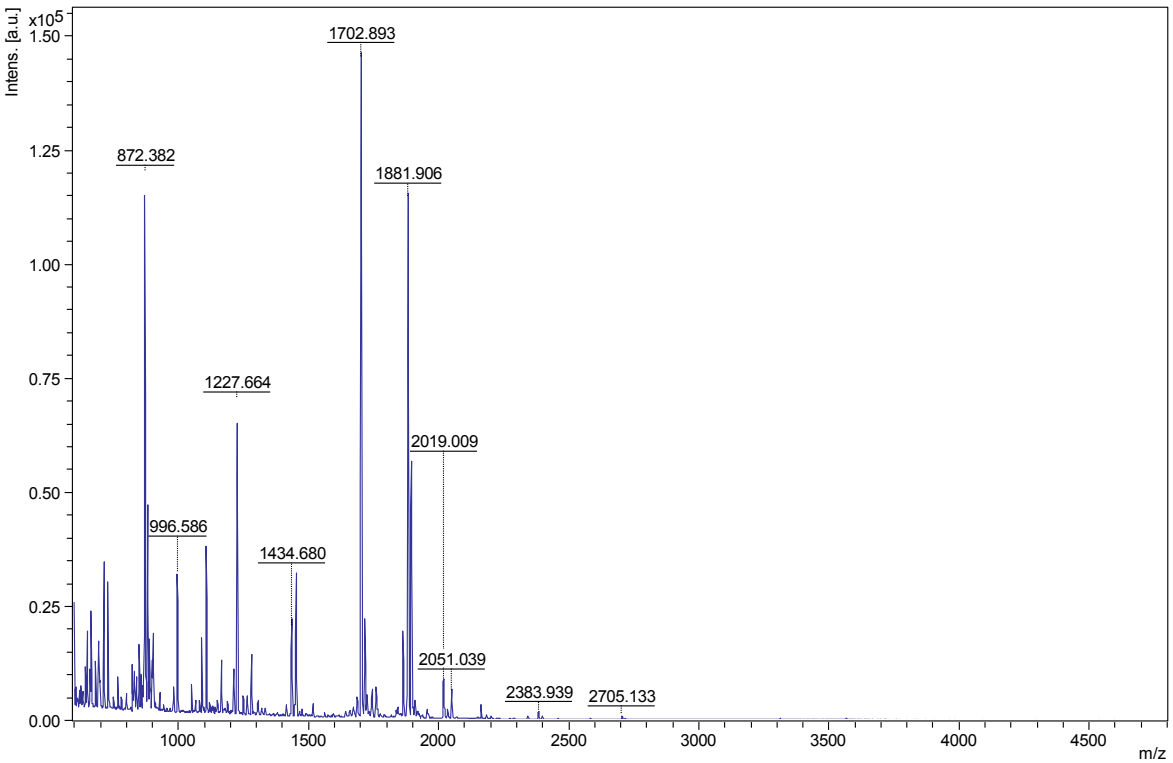
[Re-Search All] [Search Unmatched]

1. gi|499303935 Mass: 31915 Score: 93 Expect: 0.004 Matches: 9
phycobilisome rod-core linker polypeptide CpcG1 [Nostoc sp. PCC 7120]

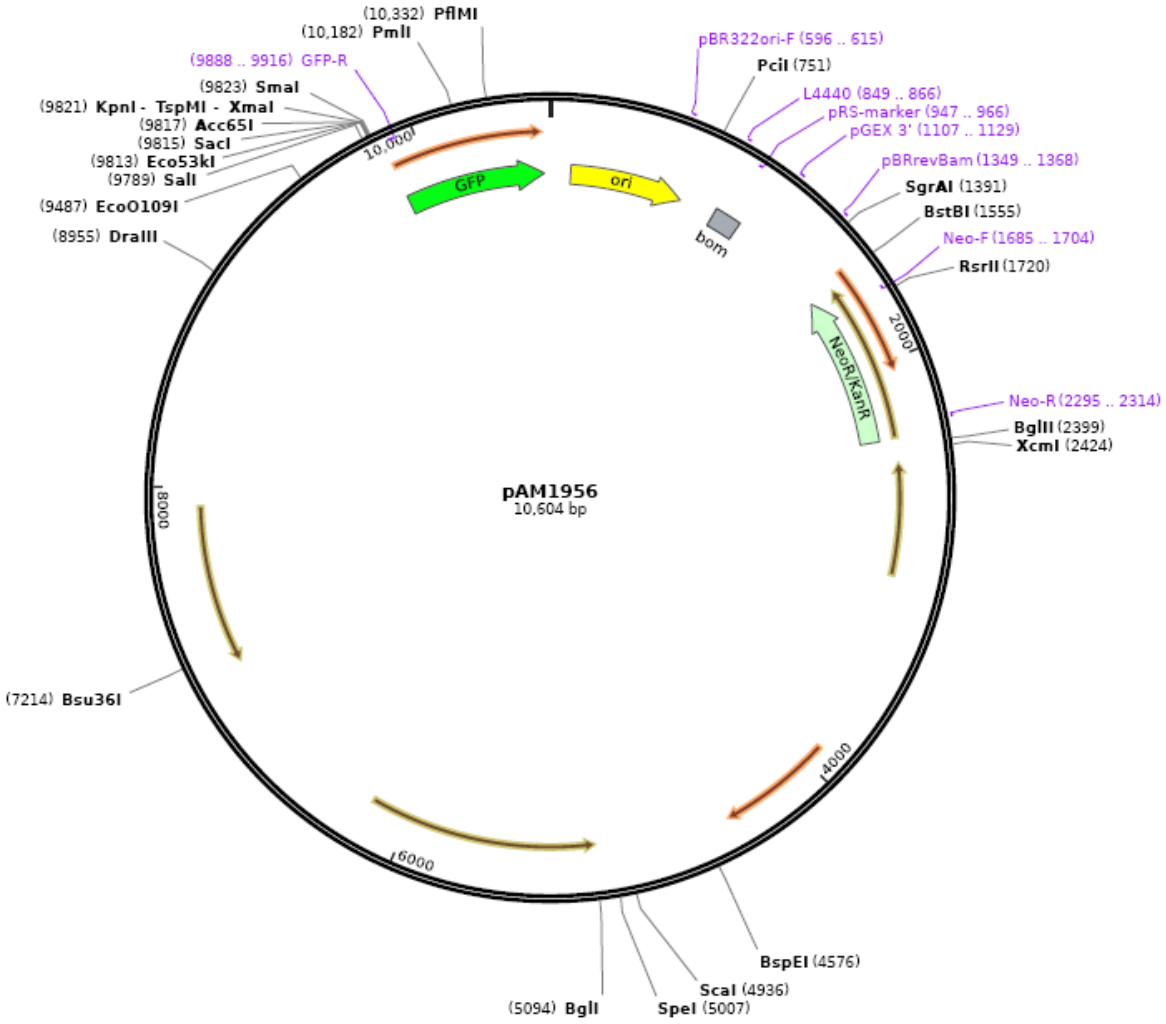
Protein sequence coverage: 35%

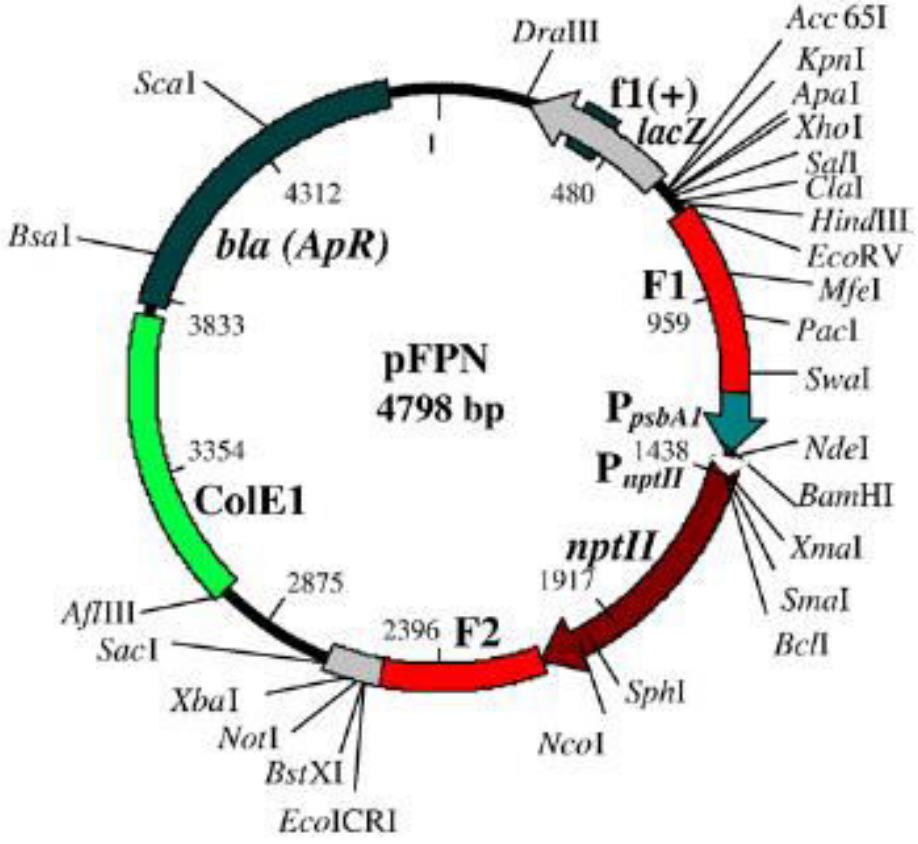
Matched peptides shown in ***bold red***.

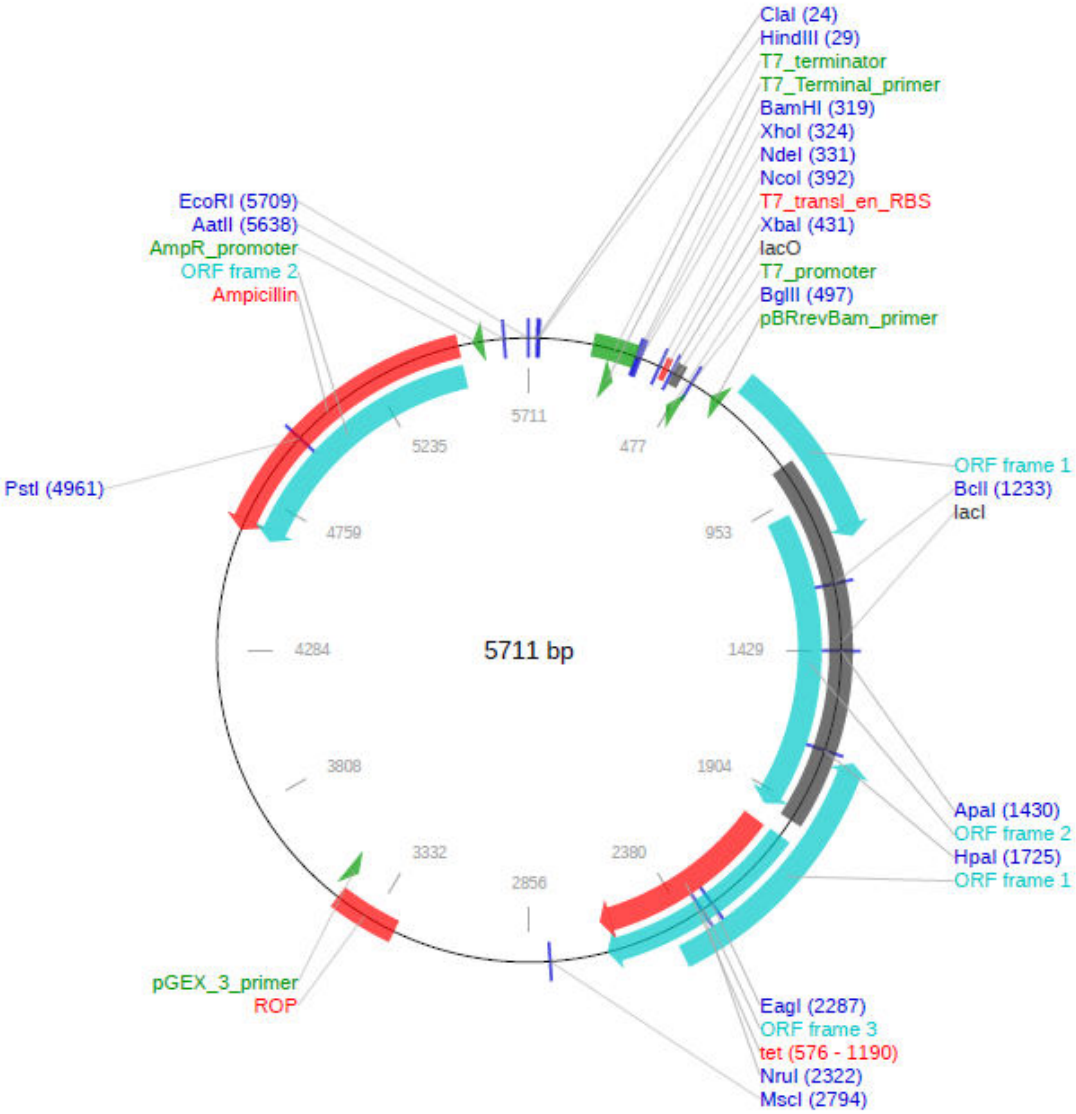
```
  1 MSIPLLEYAP SSQNQRVEGY EVPNEDTPTI YRLAAAIDDA DVDAIIWAGY
 51 RQIFSEHLII KSNRQSFLES QLRNRAINVR DFIRGLGKSE VYRTQVADLN
101 SNYRLVDITL KRFLGRAAYN QDEEIAWSIV IGSQGLHGFI DALLDSDEYR
151 ENFGDDIVPY QRRRYKDRPF NLVNPRYNAY WRDRQTLNAL GGRSFYSART
201 SGTLTKDDIR RAIPANFMAL AGKILTPERN YQRTIASVTS QIKDIKIPDT
251 SREVTTPEVT VKPVAVALPY RYIPGNKTT
```

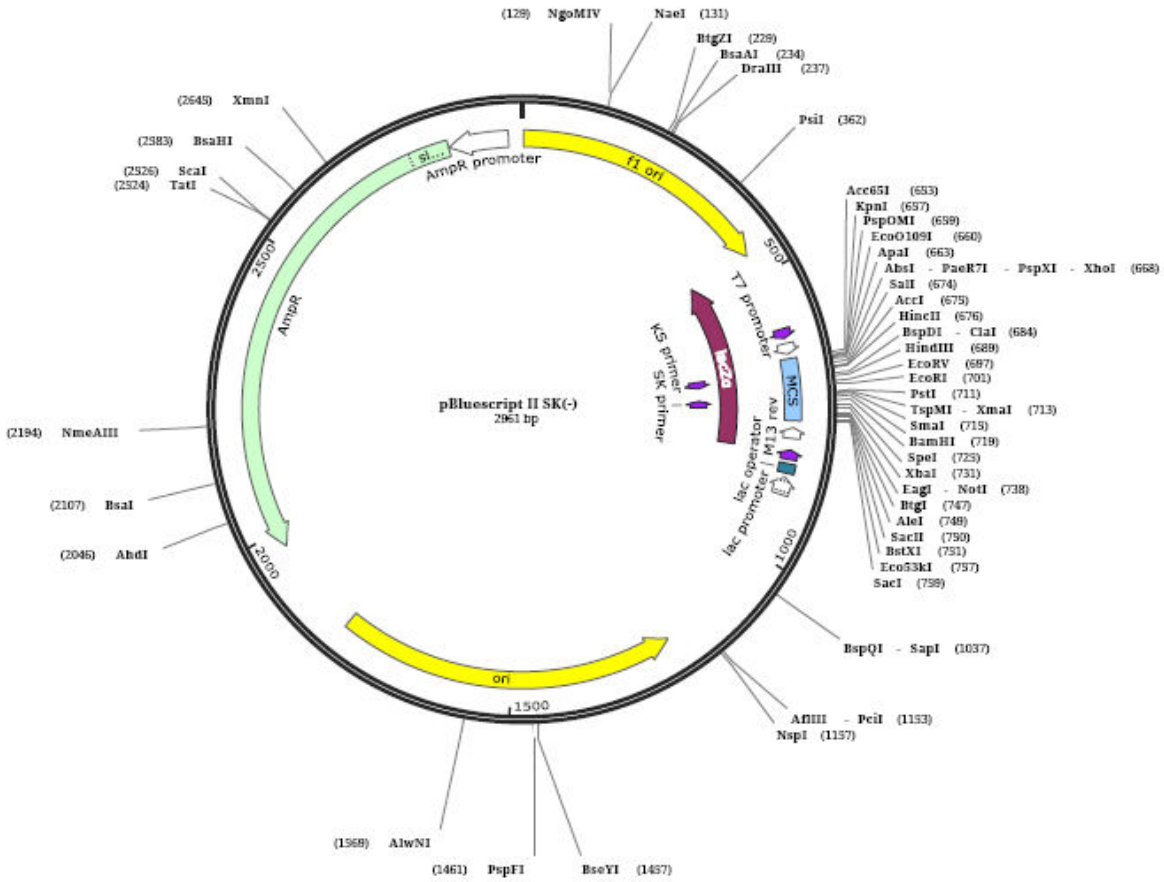


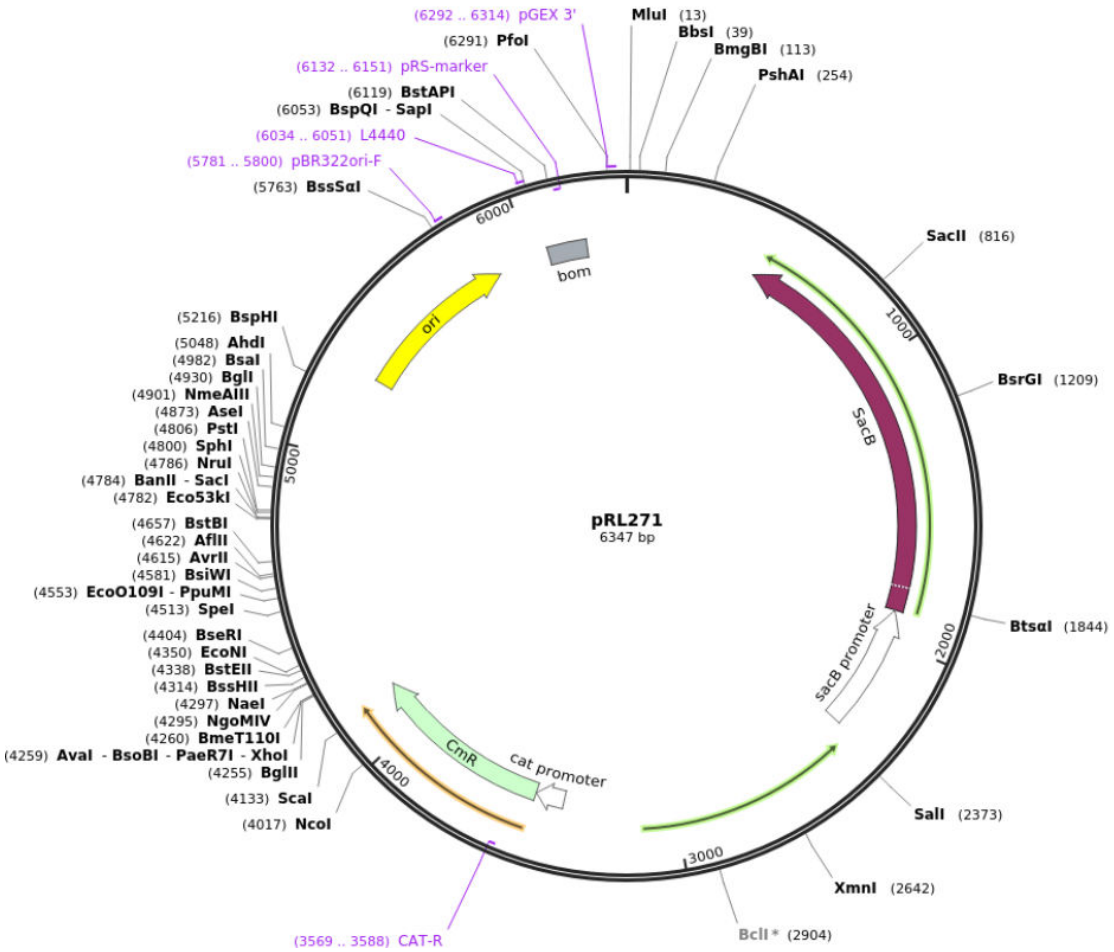
Vectors Used




pFPN
4798 bp
DraIII
Acc 65I
KpnI
ApaI
XhoI
SalI
ClaI
HindIII
EcoRV
MfeI
PacI
SwaI
NdeI
BamHI
XmaI
SmaI
BclI
SphI
NcoI
EcoICRI
BstXI
NotI
XbaI
SacI
AflIII
ScaI
BsaI
f1(+)
lacZ
F1
F2
bla (ApR)
ColE1
nptII
P psbA1
P nptII
1
480
959
1438
1917
2396
2875
3354
3833
4312


ClaI (24)
HindIII (29)
T7_terminator
T7_Terminal_primer
BamHI (319)
XhoI (324)
NdeI (331)
NcoI (392)
T7_transl_en_RBS
XbaI (431)
lacO
T7_promoter
BglII (497)
pBRrevBam_primer
EcoRI (5709)
AatII (5638)
AmpR_promoter
ORF frame 2
Ampicillin
PstI (4961)
ORF frame 1
BclI (1233)
lacI
ApaI (1430)
ORF frame 2
HpaI (1725)
ORF frame 1
EagI (2287)
ORF frame 3
tet (576 - 1190)
NruI (2322)
MscI (2794)
pGEX_3_primer
ROP
5711 bp
5711
477
953
1429
1904
2380
2856
3332
3808
4284
4759
5235


pBluescript II SK(-)
2961 bp
(129) NgoMIV
NaeI (131)
BtgZI (229)
BsaAI (234)
DraIII (237)
PsiI (362)
Acc65I (653)
KpnI (657)
PspOMI (659)
EcoO109I (660)
ApaI (663)
AbsI - PaeR7I - PspXI - XhoI (668)
SalI (674)
AccI (675)
HincII (676)
BspDI - ClaI (684)
HindIII (689)
EcoRV (697)
EcoRI (701)
PstI (711)
TspMI - XmaI (713)
SmaI (715)
BamHI (719)
SpeI (723)
XbaI (731)
EagI - NotI (738)
BtgI (747)
AleI (748)
SacII (750)
BstXI (751)
Eco53kI (757)
SacI (759)
BspQI - SapI (1037)
AflIII - PciI (1153)
NspI (1157)
BseYI (1457)
(1461) PspFI
(1568) AlwNI
(2046) AhdI
(2107) BsaI
(2194) NmeAIII
(2524) TatI
(2526) ScaI
(2583) BsaHI
(2645) XmnI
f1 ori
T7 promoter
MCS
lacZ
lac operator
lac promoter
M13 rev
KS primer
SK primer
AmpR promoter
AmpR
ori
500
1000
1500
2000
2500


pRL271
6347 bp
(6292 .. 6314) pGEX 3'
(6291) PfoI
(6132 .. 6151) pRS-marker
(6119) BstAPI
(6053) BspQI - SapI
(6034 .. 6051) L4440
(5781 .. 5800) pBR322ori-F
(5763) BssSαI
MluI (13)
BbsI (39)
BmgBI (113)
PshAI (254)
SacII (816)
BsrGI (1209)
BtsαI (1844)
SalI (2373)
XmnI (2642)
BclI* (2904)
(3569 .. 3588) CAT-R
(4017) NcoI
(4133) ScaI
(4255) BglII
(4259) AvaI - BsoBI - PaeR7I - XhoI
(4295) NgoMIV
(4297) NaeI
(4314) BssHII
(4338) BstEII
(4350) EcoNI
(4404) BseRI
(4553) EcoO109I - PpuMI
(4513) SpeI
(4581) BsiWI
(4615) AvrII
(4622) AflII
(4657) BstBI
(4782) Eco53kI
(4784) BanII - SacI
(4786) NruI
(4800) SphI
(4806) PstI
(4873) AseI
(4901) NmeAIII
(4930) BglI
(4982) BsaI
(5048) AhdI
(5216) BspHI
1000
2000
3000
4000
5000
6000
ori
bom
SacB
SacB promoter
CmR
cat promoter