Functional characterization of two new proteins, Ssl2245 and Sll1130, encoded by a dicistronic operon in *Synechocystis* sp. PCC 6803

Thesis submitted to the University of Hyderabad for the award of

Doctor of Philosophy

By

Pilla Sankara Krishna

(Regd. No. 07LPPH04)

Department of Plant Sciences

School of Life Sciences

University of Hyderabad

Hyderabad - 500046

India

June, 2013

Functional characterization of two new proteins, Ssl2245 and Sll1130, encoded by a dicistronic operon in *Synechocystis* sp. PCC 6803

Thesis submitted to the University of Hyderabad
for the award of

Doctor of Philosophy

By
Pilla Sankara Krishna
(Regd. No. 07LPPH04)

Department of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad - 500046
India
June, 2013

University of Hyderabad

(A Central University established in 1974 by act of parliament)

HYDERABAD – 500 046, INDIA

CERTIFICATE

This is to certify that this thesis entitled **"Functional characterization of two new proteins, Ssl2245 and Sll1130, encoded by a dicistronic operon in *Synechocystis* sp. PCC 6803"** is a record of bonafide work done by **Mr. Pilla Sankara Krishna** a research scholar for Ph.D. programme in Department of Plant Sciences, School of Life Sciences, University of Hyderabad under my guidance and supervision. The thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma

Dr. J. S. S. Prakash
(Supervisor)

Head,
Department of Plant Sciences.

Dean,
School of Life Sciences.

University of Hyderabad
(A Central University established in 1974 by act of parliament)
HYDERABAD – 500 046, INDIA

DECLARATION

I **Pilla Sankara Krishna** here by declare that this thesis entitled **"Functional characterization of two new proteins, Ssl2245 and Sll1130, encoded by a dicistronic operon in *Synechocystis* sp. PCC 6803"** submitted by me under the guidance and supervision of **Dr. Jogadhenu S. S. Prakash** is an original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Pilla Sankara Krishna
(07LPPH04)

Dr. J. S. S. Prakash
(Supervisor)

CONTENTS

Acknowledgements

This Report would not have been possible without the essential and gracious support of many individuals.

I take immense pleasure in thanking **Dr. Jogadhenu S. S. Prakash**, *Head,* ***Laboratory of microbial genes to functions*** *group, who had been a source of inspiration and role model in any given way, for his constant moral and personal support and guidance throughout my doctoral research. The leadership and critical assessment capabilities of* **Dr. Jogadhenu S. S. Prakash**, *allowed me to mobilize my collective knowledge and capacity in the best way possible and created an enabling environment to complete my research work. I thank the present and former Deans, school of Life Sciences, prof. R.P.Sharma, Prof. M Ramanadham and Prof. A. S. Raghavendra, the present and former Heads, Department of Plant Sciences Prof. A. R. Reddy and Prof. Apparao Podile for allowing me to use the facilities of the School and the Department. I sincerely acknowledge the infrastructural support provided by UGC-SAP, DST-FIST, DBT-CREBB, to the Dept. of Plant Sciences and the CSIR, Govt. of India for the research fellowship.*

I wish to express my deep sense of gratitude to Dr. S. Shivaji, who graciously extended his laboratory facilities and support during my research work.

I wish to express my deep sense of gratitude to Dr. Iwane Suzuki, who graciously extended his laboratory facilities during initial microarray experiments and northern blotting experiments.

This thesis is not only the result of my scientific works but the outcome of a long road and I am glad to acknowledge everyone who has contributed to the build-up of my scientific personality and therefore, to the completion of this thesis.

Here again I thank my supervisor **Dr. Jogadhenu S. S. Prakash** *for his love, care and affection, which made me to forget my home.*

My heartfelt thanks are also due to my colleague scholars of Microbial genes to functions group Dr. K. Sirisha, P. Bhavani devi, M. Karthik Mohan, M. Subhashini, B. Radha Rani, P.V.P. Sai Arun, Deepak Kumar Singh, Reetika Chaurasia, Linga Swamy Bhantu and Afshan Sri Kumar for their timely help and constant encouragement.

I would like now to mention "some" of Project students who have helped me, while learning B. Yesu Nadha Rao, D. SomaSekhar, Surjeet Kumar Arya, A. Venkateswara Rao, Santosh and B. Ranjith .

I wish to place on record my deep sense of thankfulness to Prof. Prasanna Mohanty for his suggestions and motivation during his visit to HCU.

I would like to thank all my dearest friends P. Raj Sheel, A. Kranthi kiran, J. Prasanna Kumar, J. Anil Kumar, P. Mallikarjun Rao, Md. Mujahid and I. Aravind who made my stay always cheerful in HCU.

Eight long years in the same place is for sure a long time and with time come friends. I thank all my friends at University of Hyderabad for making my stay a delightful one.

This dissertation work at Laboratory of Microbial genes to functions (MGF) group could never run smoothly without all the help I got from lab helpers. I wish to thank G. Ramesh, Ch. Sridhar Yadav for their assistance in the lab.

I am thankful to all my batch mates and other research scholars of the School of Life Sciences for their timely help.
I highly acknowledge the efforts of Ramesh, Mohan Rao, Sudarshanam and other non-teaching staff .

Saving the most important for the last, I would like to express my heartfelt thanks to my guide, Dr. J S S. Prakash and his family for their blessings, wishes, support and constant encouragement throughout my Ph.D tenure.

Pilla Sankara Krishna

ABBREVIATIONS

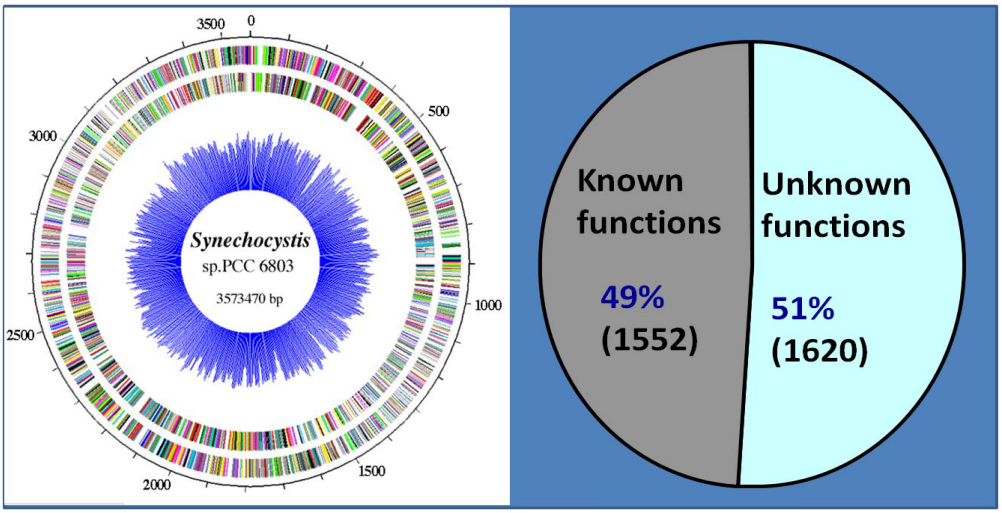
µg	:	microgram
µM	:	micromolar
°C	:	degree centigrade/degree Celsius
Abs	:	absorption
ALP	:	alkaline phosphatase
ATP	:	adenosine tri phosphate
BCIP	:	5-bromo-4-chloro-3-indolyl phosphate
Bp	:	base pair (bp)
BSA	:	bovine serum albumin
CD	:	circular dichroism
cDNA	:	complementary DNA
Chl	:	chlorophyll
C-terminal	:	carboxy terminal
Cyt C	:	cytochrome C
DEPC	:	diethylpyrocarbonate
DMSO	:	dimethyl sulfoxide
DNA	:	deoxy ribonucleic acid
dNTPs	:	deoxy nucleotide triphosphates
DTT	:	dithiothreitol
EDTA	:	ethylene diamine tetra acetic acid
gm	:	gram
h	:	hour(s)
Hepes	:	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid sodium salt, N-(2- Hydroxyethyl) piperazine-N′-(2-ethanesulfonic acid) sodium salt
IgG	:	immunoglobulin G
IPTG	:	isopropyl β-D-thiogalactoside
kb	:	kilobase pair
kDa	:	kilodalton
L	:	litre
LB	:	Luria-Bertani
M	:	molar
Mb	:	Mega base pair
mg	:	milligram
min	:	minute
ml	:	milliliter
mM	:	millimolar
NBT	:	nitroblue tetrozolium
Ni-NTA	:	nickel-nitroacetic acid agarose
nm	:	nanometers

N-terminal	:	amino terminal
OD	:	optical density
ORF	:	open reading frame
PAGE	:	polyacrylamide gel electrophoresis
PBS	:	phosphate buffered saline
PCR	:	polymerase chain reaction
PMSF	:	phenylmethylsulfonylfluoride
qRT-PCR	:	quantitative real time PCR
RNA	:	ribonucleic acid
rpm	:	revolutions per minute
RT-PCR	:	reverse transcriptase-polymerase chain reaction
SDS	:	sodium dodecyl sulphate
Sec(s)	:	seconds
SEM	:	scanning electron microscopy
TA	:	toxin-antitoxin
TE	:	Tris-EDTA
TEM	:	transmission electron microscopy
Tris	:	tris-(Hydroxymethyl) aminoethane
V	:	volts

List of figures

List of tables

INTRODUCTION



1. INTRODUCTION

Functionally unknown protein is a protein whose existence has been proven but with no clear idea about its function. Hypothetical protein is a protein whose existence has been predicted. When a bioinformatic tool used for the gene identification finds an open reading frame without an analog in the protein data base, it returns "hypothetical protein" as an annotation remark.

Functional characterization of unknown proteins

With the advantage of Next generation sequencing (NGS) technology every year several genomes have been sequenced and deposited in the public data bases. By the end of 2012, total of 2,272 bacterial genomes with more than or equivalent to 1 Mb genome size were sequenced and deposited in the NCBI database. Microbial genome sequence data has revealed that on average ~30–40% of the proteins encoded by typical bacterial genome are not functionally characterized and are annotated as either hypothetical proteins or functionally unknown proteins (Osterman and Overbeek 2003; Galperin and Koonin 2004). The well studied model systems like *E. coli*, *Lactobacillus* and *S. aureus* have about 82%, 70% and 68% of functionally annotated proteins respectively (Hanson *et al.*, 2010). Moreover functions of many annotated proteins are uncertain, as they were annotated based on sequence similarity but not based on experimental evidence. Even in a well studied organism like *E. coli* experimental evidence supporting the function of a protein is available only for 54% of the total proteins (Frishman 2007). There is a large gap in the information on protein's function and the sequence deposition. Therefore characterization of proteins

is required to fill the gap. This gap is high in case of eukaryotes *i.e* plants and animals (Hanson *et al.*, 2010).

Table 1.1: Functionally unknown and hypothetical proteins in cyanobacteria

Name of the organism	Total genes	Unknown proteins	Hypothetical proteins	% (Unknown+ hypothetical)
Acaryochloris marina MBIC11017	8,462	4	4,434	52.4
Nostoc punctiforme ATCC 29133	6,794	22	2,617	38.8
Arthrospira platensis NIES-3	6,676	4	4,084	61.2
Microcystis aeruginosa NIES-843	6,363	1,404	1,994	53.4
Anabaena sp. PCC 7120	6,223	1,756	1,769	56.6
Cyanothece sp. ATCC 51142	5,359	0	3,009	56.1
Gloeobacter violaceus PCC 7421	4,484	959	1,636	57.9
Synechocystis sp. PCC 6803	3,725	682	1,238	51.5
Synechococcus elongatus PCC 6301	2,580	1	946	36.7
Thermosynechococcus elongatus BP-1	2,528	249	837	43.0

Table 1.1. Total number of genes and their annotations for each cyanobacterial species were compiled from the cyanobacterial genome database, cyanobase (http://genome.microbedb.jp/cyanobase/). The first sequenced cyanobacterium is *Synechocystis* sp. PCC 6803 by Kaneko *et al.*, 1996. Though a number of research groups have been using this organism as model system to understand details of photosynthetic adaptation, abiotic stress responses, biochemistry, physiology and metabolism of cyanobacteria, even till to date only 50% of its genes are functionally characterized.

In case of human genome ~40% of the proteins are functionally known and higher plants with larger genome contains much more unknown proteins, Arabidopsis has more than 60% of unknown proteins. About ~50% of the proteins are functionally unknown in different cyanobacterial model systems (Table 1.1).

Importance of characterization of functionally unknown proteins

Genetic manipulation of bacteria for the production of a desired compound or metabolite has been an area of interest for many research groups in the world. Presence of unknown proteins probably indicates existence of yet to be identified metabolic pathways. Therefore, there is a need to understand the metabolism (functions of all proteins) of the cell completely prior to targeted manipulation of metabolic pathway(s) using recombinant DNA technology. Metabolic engineering will be possible once we acquire complete understanding of metabolic pathway(s) related to the selected compound. Moreover, characterization of unknown proteins will answer the biological questions like role of unknown proteins as receptors or carriers, identification of novel metabolic pathways and cellular chemical transformations, etc.

Conserved unknown proteins are the top targets for characterization

Several of bacterial unknown proteins are conserved and have orthologues in other bacterial genomes of different taxonomic groups. In some cases, these proteins' conservation is broad, and includes most or even all forms of life from bacteria, Archaea, higher plants and animals (Tatusov *et al.*, 2001; Galperin and Koonin 2004). There are many such widely distributed (henceforth 'conserved') proteins, as shown by the OrthoMCL database of orthologous protein families (Chen *et al.*, 2006b). For

instance, most bacteria share >400 orthologue families with *Arabidopsis* and human; of these, about half lack known functions. Conserved proteins are plainly ancient in origin and must have crucial functions in metabolism, transport or core cellular processes such as translation that are shared by all organisms (Hedges *et al.,* 2004). Thus, among all the families of unknown proteins, the conserved ones merit the highest priority for functional characterization because they have the greatest potential payoff in new biological knowledge (Koonin and Galperin 2002; Galperin and Koonin, 2004). It is assumed that the genes coding for hypothetical proteins that have been retained in at least several distantly related organisms may have important physiological function and could be expected to be essential for survival of the cell. Functions of a hypothetical protein can be predicted based on the presence of a structural motif, which was previously identified and characterized in other proteins. For example, hypothetical proteins with GTPase, ATPase, methyl transferase or RNA helicase motif can be predicted to have corresponding putative biochemical function. However, this does not mean that their exact biological function is known. Therefore, to elucidate an exact biological function of such proteins, direct experimentation is necessary.

Proteins expressed in response to abiotic stress

Genomic and proteomic studies revealed that genes coding for hypothetical proteins have been up-regulated under one or more abiotic stress conditions in bacteria (Ye *et al.,* 2000; Hihara *et al.,* 2001; Suzuki *et al.,* 2001; Zheng *et al.,* 2001; Kanesaki *et al.,* 2002; Suzuki *et al.,* 2006; Kanesaki *et al.,* 2007; Prakash *et al.,* 2010). Induced expressions of hypothetical genes under abiotic stress conditions indicate that execution of several unidentified short term adaptive molecular mechanisms by the

cell to cope up with the stress. Assigning a functional role to such stress regulated hypothetical proteins is a matter of utmost importance, in order to decode the role played by them under various stresses.

Strategies for characterization of functionally unknown proteins

Several strategies have been in use for functional characterization of hypothetical proteins by different research groups in the world. The most widely used methods of functional characterization are mutagenesis and phenotypic characterization, functional complementation, heterologous expression, biochemical and biophysical characterization, gene expression profiling and protein-protein interaction studies to get clues on function. Functional characterization of hypothetical proteins in a model cyanobacterium *Synechocystis*, using one or the other methods mentioned above are described in the following sections.

Mutagenesis and phenotypic characterization

The widely used experimental strategy for unraveling function of a hypothetical protein is by generating knockout mutant of the hypothetical gene by targeted mutagenesis, and subsequent characterization of the mutant by physiological, biochemical and biophysical methods. Several research articles have been published from different laboratories using above strategy and reported the function of various hypothetical proteins in *Synechocystis.* For instance, hypothetical protein, Sll0088 was identified as a regulator of photosystem I biogenesis (Yu *et al.*, 2003). The *sll0088* gene encodes a protein predicted to be a transcriptional regulator with sequence similarities to transcription factors in other prokaryotic and eukaryotic organisms including *Arabidopsis thaliana*. Mutagenesis of *sll0088* by insertion of a

kanamycin resistance cassette, physiological characterization, presence of a typical helix-turn-helix DNA-binding motif, Sll0088 was identified as a negative transcriptional regulator and reported as regulator of biogenesis of photosystem I (Yu *et al.,* 2003). Putative ATP/GTP binding protein, Sll1717 was reported to affect the redox state of plastoquinone pool by modulating the quinol oxidase activity in thylakoids (Kufryk and Vermaas 2006). By generation and characterization of *Δssr2998,* and analysis of co-elution of proteins it is associated with, Ssr2998 was reported as a regulator of photosynthetic electron transfer (Volkmer *et al.,* 2007). By deletional inactivation and subsequent DNA microarray analyses of a gene coding for putative RNA helicase protein, CrhR, it has been reported that it is regulator of the expression of heat shock genes under low temperature stress (Prakash *et al.,* 2010). The availability of large-scale bacterial and plant knockout collections, along with databases on knockout phenotypes, can quickly show whether a gene is essential or is associated with a particular phenotype (Gerdes *et al.,* 2006; Tzafrir *et al.,* 2003). Besides revealing associations directly (e.g. when auxotrophy connects a gene with a biosynthetic pathway) phenotype data, especially essentiality data, pinpoint important genes. Essentiality data for bacteria are integrated into the SEED database (http://www.theseed.org/wiki/Main_Page); plant phenome databases include RAPID, SeedGenes (http://www.seedgenes.org/) and Chloroplast2010 (http://plastid.msu.edu/). Details of cyanobacterial mutants and their corresponding phenotypes are presented in cyanobase (http://genome.microbedb.jp/cyanobase/).

Complementation or heterologous expression as an alternative strategy

A second strategy for characterization of a hypothetical protein is by functional complementation in model organisms like *E. coli*, expression, purification and, *in vitro* biochemical and biophysical analysis of the purified protein. In *Synechocystis,* the ORF, *sll1468* was identified as a gene for β-carotene hydroxylase by functional complementation of an *E. coli* mutant in which, its homologue has been mutated in a β-carotene producing *E. coli* (Masamoto *et al.,* 1998). Putative ADP-ribose pyrophosphatases were characterized by recombinant expression and purification in *E. coli* and by *in vitro* biochemical characterization (Okuda *et al.,* 2005). Together with targeted mutagenesis and, recombinant expression and biochemical characterization of the expressed protein, Sll1575 was reported as a protein required for motility of *Synechocystis* cells (Kamei *et al.,* 2001).

Protein-Protein interaction studies to get clues on functions

A third strategy for getting clues on function of a hypothetical protein is by identification of proteins, that are physically associated with the protein (unknown) in question. If the functions of interacting protein partners are already known, it is possible to predict the function of hypothetical protein in question. Kashino *et al.,* has identified five novel proteins (Sll1414, Sll1390, Sll1252, Sll1130, Sll1638) as association partners of the His-tag affinity purified PSII complex (Kashino *et al.,* 2002). Since, these proteins were identified as associating partners of PSII, it was assumed that they play important roles in structure and function of PSII. Subsequently, one of these novel uncharacterized proteins, Sll1252 was shown as a

necessary protein for energy balancing in photosynthetic electron transport in *Synechocystis* (Kashino *et al.*, 2011). Another protein, Sll1638, was found to be homologous to plant PsbQ protein (17 kDa protein). Sll1638 was reported to be located at the lumenal side of PSII and play role in water oxidation (Thornton *et al.*, 2004; Kashino *et al.*, 2006). Auxiliary functions to support the assembly and repair of PS II complexes have been shown for Sll1414 (Psb29) and TLP18.3 protein, a homologue of Sll1390 in *Arabidopsis thaliana* (Keren *et al.*, 2005; Sirpio *et al.*, 2007), as found in Psb27 (Chen *et al.*, 2006a; Nowaczyk *et al.*, 2006; Roose & Pakrasi 2008).

Gene expression profile for characterization of unknown proteins

In higher organisms regulation of spatial and temporal expression of genes is very important for the development of the organism. So in plants and animals the site (spatial) or developmental stage (temporal) in which a gene is expressed provides us vital clues about the function (de Cr´ecy-Lagard and Hanson2007). If a gene is specifically expressed in a tissue it is highly possible, that it has role only in that particular tissue's function. Functional associations of a protein can be derived from its expression profile and co-expression datasets (from microarrays, 2D electrophoresis and iTRAQ), which are now well developed for model bacteria as well as for plants and animals (e.g. Gollub *et al*., 2006; Laule *et al.*, 2006; Obayashi *et al.*, 2007). Microarray databases and tools include Microbes Online and GenExpDB for bacteria (http://genexpdb.ou.edu/main/), and ATTED and the Golm Transcriptome Database (http://csbdb.mpimp-golm.mpg.de/index.html) for *Arabidopsis* and exclusively for cyanobacteria a web server named cyanoexpress

(http://cyanoexpress.sysbiolab.eu/) has the data of gene expression profiles. From these databases its expression profile and co-expression datasets for gene in question can be derived.

Three dimensional structure, presence of common regulatory sites and a common regulatory protein, gene clustering, and comparative genomics are some of the methods that provide clues for functional characterization of a hypothetical protein. There are several other examples revealing functions of genes coding for hypothetical proteins using one or the other above mentioned strategies (Kufryk and Vermaas 2001; Kufryk and Vermaas 2006).

Cyanobacterium *Synechocystis* sp. PCC 6803 as model

Cyanobacteria are model microorganisms to study photosynthesis, carbon and nitrogen assimilation, evolution of plastids, and abiotic stress adaptation. *Synechocystis* sp. PCC 6803 (here after *Synechocystis)* is one of the extensively studied cyanobacterium as it can grow autotrophically or heterotrophically in the absence of light. The photosynthetic apparatus of it is very similar to the one found in plants. *Synechocystis* was the first photosynthetic organism for which the entire genome sequence was determined (Kaneko *et al*., 1996). In 1996, Tabata and coworkers at the Kazusa DNA Research Institute finished the genomic sequence of this organism and made the information available in very useful format on a website named CyanoBase (http://www.kazusa.or.jp/cyano/cyano.html). *Synechocystis* is a naturally transformable cyanobacterium with several copies of the genome in it (Labarre *et al.*, 1989; Griese *et al.*, 2011). As it is naturally transformable and incorporates DNA into its genome by homologous recombinaiton, targeted

inactivation of genes in its genome has been standardized. Cyanobacterial vectors for complementation of the mutants and protein expression were developed (Zinchenko *et al.*, 1999) and successfully used in *Synechocystis* for complementation (Suzuki *et al.*, 2005; Prakash *et al.*, 2010). Minimal requirements for growth, easy maintenance, natural competency, and availability of genome sequence, methods of mutagenesis and microarray chips attracted the research groups to work with *Synechocystis*. Due to availability of DNA microarray chips this model was extensively studied in terms of responses to different stress conditions at the level of gene expression (Hihara *et al.*, 2001; Suzuki *et al.*, 2001; Mikami *et al.*, 2002; Hihara *et al.*, 2003; Inaba *et al.*, 2003; Suzuki *et al.*, 2005).

Genome of *Synechocystis*

The genome of the *Synechocystis* consists of a single chromosome and several plasmids of different sizes. Till date seven native plasmids were identified and the nucleotide sequences of all seven plasmids (pSYSM:120 kb, pSYSX:106 kb, pSYSA:103 kb, pSYSG:44 kb, pCC5.2: 5.2 kb, pCA2.4: 2.4 kb, and pCB2.4:2.3 kb) have already been deposited in the database. *Synechocystis* genome is 3,956,957 bp including all native plasmids. It contains total 3725 genes involved in various cellular processes and genes for functionally uncharacterized proteins (unknown and hypothetical proteins) (http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome&cmd=Retrieve&dopt=overview &list_uids=112). *Synechocystis* genome, has 1920 functionally unknown proteins (51% of total genes, 682 unknown and 1238 hypothetical proteins), suggesting a number of unknown molecular events /chemical transformations may be occurring in the cell. Computational prediction and analysis

of the native plasmids has indicated the presence of 409 potential genes (397 in four large plasmids and 12 in three small plasmids), but little information was obtained about the functional relationship of plasmids to cell, as majority of the predicted genes (77%) are of unknown function (Kaneko *et al.,* 1996 and 2003). pSYSX, showed a high degree of sequence similarity (higher than 95% at the amino acid level) to the adjacent chromosomal genes *sll0789* (*rre34*) and *sll0790* (*hik31*), respectively. Moreover these gene sets in chromosome and plasmids are flanked by genes involved in cation transduction suggesting both these gene sets in plasmids and genome has similar function of transduction of cation signals. The occurrence of genes on plasmids is divergent, except *parA* which is the only gene common to all four large plasmids. The sequences and the order of five to seven protein-encoding genes were significantly conserved between pSYSM and pSYSX implies either an evolutional relationship between these two plasmids or the occurrence of a segmental transfer between the two plasmids. Presence of similar genes in genome and plasmid and presence of response regulators in the plasimds shows a cross talk between plasmid and genome of *Synechocystis* in regulation of gene expression. Very little is known about the regulation of these plasmid genes till date. Therefore, functional elucidation of these plasmid genes and their regulation is of evolutionary importance.

Ssl2245 and Sll1130 are two unknown proteins in *Synechocystis*

In this study, we focused on two functionally unknown proteins Ssl2245 and Sll1130 in the genome of *Synechocystis* (Figure 1.1). These two proteins Ssl2245 (88 aa) and Sll1130 (115 aa) are coded by an over lapping gene pair. The stop codon (ATGA) of the first gene (*ssl2245*) overlaps with the start codon (ATGA) of the second one (*sll1130*)

(http://genome.microbedb.jp/cyanobase/Synechocystis/genes/sll1130). This type of dicistronic overlapping gene pairs are present in a wide variety of gram positive and gram negative bacteria (Lillo and Krakauer 2007; Summer *et al.,* 2007). Sll1130 encoded by second gene of the operon was shown to be a novel protein associated with PSII complex. It was reported to be co-eluted with the other proteins of PSII complex (Kashino *et al.,* 2002). Since it is a probable member of PSII protein complex, we assumed that this protein may have important role in structure and function of PSII. We have chosen this protein for functional characterization as this protein is a novel member of PSII complex. Recently, it has been reported that mutation of *sll0822,* an AbrB like gene lead to the upregulation of the *sll1130* and *ssl2245* along with some nitrate limitation inducible genes (Ishii and Hihara 2008). DNA microarray analysis revealed that both *ssl2245* and *sll1130* genes were down regulated upon heat shock in *Synechocystis* (Suzuki *et al.,* 2005). There are no reports available that suggest probable function of Ssl2245, the protein coded by the first gene in the operon. Even the sequence comparison, did not show similarity to any known protein, including toxins and antitoxins. However, comparison of Sll1130 protein sequence with the other proteins in the database shows similarity with transcriptional modulator of MazE/toxin, MazF of other bacterial species. As Sll130 shows similarity with PemK domain, it is assumed that this protein may be involved in growth regulation and / or plasmid maintenance. On the other hand, large scale protein-protein interaction studies have shown that Sll1130 protein is physically associated with the Ssl2245 (http://ppi.kazusa.or.jp/Synechocystis/genes/show/sll1130) (Sato *et al.,* 2007). This type of interacting protein pairs coded by overlapping DNA regions are reported to be

members of toxin-antitoxin systems (TA system) (Pandey and Gerdes 2005). Based on available literature, organization of genes in this operon, presence of PemK domain in Sll1130 and its conservation in various bacterial species, we have chosen this operon for functional characterization.

Figure 1.1: *ssl2245* and *sll1130* in genome context (gene map)

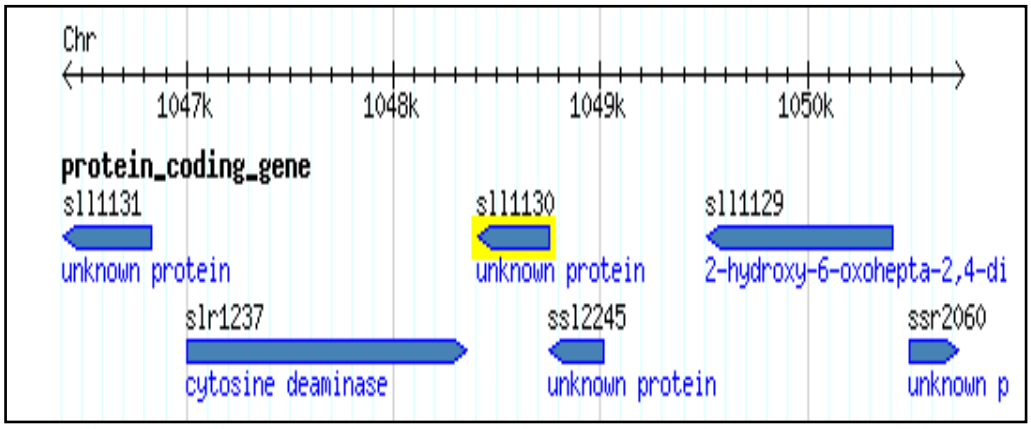


Figure 1.1. Location of the dicistronic operon, *ssl2245-sll1130* in the genome of *Synechocystis*. *Synechocystis* has a 3.6 Mbp chromosome in which *ssl2245-sll1130* are located between 1049016and 1048406. Both structural genes were annotated as unknown proteins in the genome map. These genes are oriented in the reverse orientation (in the *sll1130* annotation 's' stands for *Synechocystis,* 'l' stands for long and second 'l' stands for left orientation and in *ssl2245* 's' stands for *Synechocystis,* second 's' stands for small and 'l' stands for left orientation). Taken from cyanobase (http://genome.microbedb.jp/cyanobase/Synechocystis/genes/sll1130).

Sll1130 was co-eluted with the PSII complex

Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 revealed the presence of novel polypeptides (Kashino *et. al.,* 2002). In this research the photosystem II (PSII) complex was purified from the HT-3 strain of the widely used cyanobacterium *Synechocystis,* in which the CP47 polypeptide has been genetically engineered to

contain a polyhistidine tag at its carboxyl terminus. Out of the total 31 polypeptides identified in photosystem II of *Synechocystis*, 26 polypeptides are functionally well characterized while 5 polypeptides are not characterized. The five novel polypeptide encoding genes are *sll1414, sll1252, sll1390, sll1638,* and *sll1130*. These genes show BLAST2 homology with certain open reading frames in *Arabidopsis* and other bacterial proteins. Sll1130 protein is a conserved protein in different groups of bacteria with PemK domain. Characterization of conserved Sll1130 is important because of its association with PSII an important protein complex.

Photosystem II an important protein complex of *Synechocystis*

Photosystem II (PSII) is a membrane-protein complex located in the thylakoid membrane (Hankamer *et al.*, 1997). It uses solar energy to carry out the primary processes of photosynthesis: charge separation, quinone reduction and water oxidation to molecular oxygen. CP43 and CP47 are two core-antenna complexes in PSII. They accept excitation energy that is harvested by the light-harvesting complex and transfer the energy directly to the PSII reaction center (RC) (Bricker 1990). Photosystem II (of cyanobacteria and green plants) is composed of around 20 subunits (depending on the organism) as well as other accessory, light-harvesting proteins. Each photosystem II contains at least 99 cofactors: 35 chlorophyll a, 12 beta-carotene, two pheophytin, two plastoquinone, two heme, one bicarbonate, 20 lipid, the Mn_4CaO_5 cluster (including chloride ion), and one non heme Fe^{2+} and two putative Ca^{2+} ion per monomer (Guskov *et al*., 2009). Strong illumination results in the photoinhibition of photosynthesis with PSII as the main target. Under high light illumination, the PSII membrane suffers impairment of pigments (Telfer *et al.*, 1991; Bumann and

Oesterhelt 1995) degradation of proteins such as the D1 protein (Aro *et al.*, 1990; Shipton and Barber 1991) and loss of photosynthetic activity. Both CP43 and CP47 are found to be impaired during photoinhibition (Salter *et al.*, 1992). Photosynthetic oxygen evolution is an extremely heat-sensitive process and incubation of spinach Photosystem II (PSII) membranes at 40^{o}C for several minutes leads to its complete inactivation. When 33-kDa manganese stabilizing protein was substituted by a homologue protein, isolated either from the thermophilic cyanobacterium *Phormidium laminosum*, or from *Escherichia coli* as a recombinant thermophilic cyanobacterial protein, showed a significant increase in tolerance to heat inactivation of the oxygen-evolving activity (Pueyo *et al.*, 2002). PSII activity of a cell during high light recovery and heat treatment shows the structural integrity and repair kinetics of PSII.

***ssl2245-sll1130* mRNA is down regulated in heat stress condition**

DNA microarray analysis revealed that both *ssl2245* and *sll1130* genes were down regulated upon heat shock in *Synechocystis* (Suzuki *et al.*, 2005). Heat treated (44^{o}C) *Synechocystis* cells were applied on microarrays at different time points and the genes that are differentially expressed with the heat stress were identified. *sll1130*, *ssl2245* were down regulated with heat stress with in 20 min of incubation at (44^{o}C) and the levels remained low even after 60 min (Figure 1.2). Genes that code for desaturases were also down regulated with heat stress with in 20 min and their levels remained low even after 60 min. This gives us a clue that down regulation of *sll1130* and *ssl2245* has some role in heat acclimation. On the other hand, large scale protein-protein interaction studies have shown that Sll1130 protein is physically associated with the Ssl2245 (Sato *et al.*, 2007). Large scale protein-protein

interaction studies have identified the interacting protein partners in *Synechocystis* genome (http://ppi.kazusa.or.jp/Synechocystis/genes/show/sll1130).

Figure 1.2: Differentially expressed genes of *Synechocystis* upon heat stress

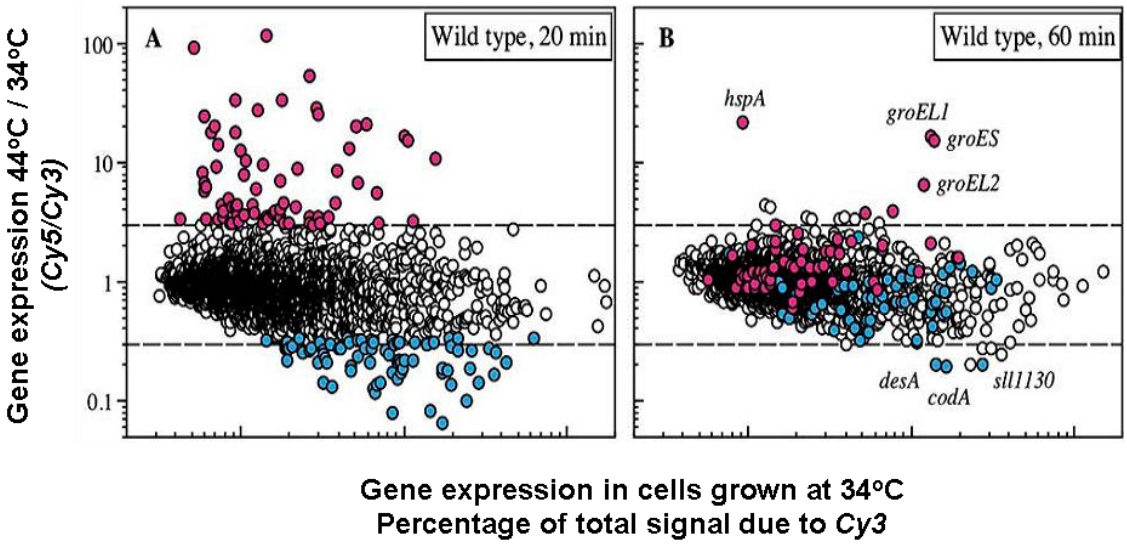


Figure 1.2. DNA microarray analysis of changes in genome-wide expression of genes in response to heat shock at 44°C in wild-type cells for 20 min (A) and 60 min (B). RNA extracted from cells that had been grown for 16 h at 34°C and then incubated at 44°C for 20 min or 60 min was labeled with Cy5 (as the sample) and RNA extracted from the respective cells that had not been exposed to heat shock (as the control) was labeled with Cy3. Genes whose levels of expression rose more than 3-fold and genes whose expression was suppressed to less than one-third of the control level in wild-type cells after incubation at 44°C for 20 min are indicated in red and blue, respectively. Upper and lower dashed lines indicate a 3-fold increase and a decrease to one-third, respectively. This figure was adapted from Suzuki *et al.*, 2005.

Heat stress response at the level of gene expression in *Synechocystis*

Temperature is one of the key environmental factors that influence the survival of microorganisms. Sudden shift of microbial cultures from optimum growth temperature to high temperature triggers a ubiquitous and homeostatic cellular

response (Webb and Sherman 1994; Feder and Hofmann, 1999). Bacteria acclimatize to heat by rapid upregulation of specific set of genes that encode heat shock proteins (HSP) (Gottesman *et al.,* 1997; Lund 2009). During heat acclimation, majority of the heat shock proteins act as molecular chaperones or proteases and facilitate refolding of heat denatured proteins, stabilization of the protein structure, solubilization of aggregated proteins and degradation of irreversibly damaged proteins (Gottesman *et al.,* 1997; Horwich *et al.,* 2007; Liberek *et al.,* 2008; Lund 2009).

In cyanobacteria, the expression of heat shock genes is regulated at the transcriptional level under heat shock and other stress conditions (Webb and Sherman 1994; Kanesaki *et al.,* 2002; Nakamoto *et al.,* 2003) and HSPs play important roles in the tolerance of cyanobacterial cells to heat, cold and oxidative stress (Nakamoto *et al.,* 2001; Hossain and Nakamoto 2002, 2003). Thus these genes are generally referred to as stress responsive genes, essentially required for thermal tolerance in *Synechocystis* (Kojima and Nakamoto 2005). There are several other genes like *isiA* and *isiB* and *frpC* induced during heat stress response (Suzuki *et al.,* 2005, 2006). Inactivation of heat responsive genes, such as *htpG, hspA, dnaK2,* (molecular chaperones) *isiA* and *isiB* (iron and heat regulated proteins), *sigB* and *sigC* (alternative sigma factors), induce thermal sensitivity in *Synechocystis* indicating their physiological importance under heat stress (Tanaka and Nakamoto 1999; Nakamoto *et al.,* 2000; Török *et al.,* 2001; Fang and Barnum 2003; Tuominen *et al.,* 2006; Kojima *et al.,* 2006). In *Synechocystis* the small heat shock protein, HspA has been reported to be an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding (Török *et al.,* 2001). *Synechocystis* becomes sensitive to high temperature on disruption of *hspA* (Lee *et al.,*

1998; Fang and Barnum 2003). Inactivation of *htpG*, another important member of heat shock genes in *Synechocystis*, led to comparatively more sensitivity to high temperature than those of *E. coli*, *B. subtilis* and *A. actinomycetemcomitans* (Bardwell and Craig 1988; Winston *et al.,* 1996; Versteeg *et al.,* 1999; Fang and Barnum 2003) indicates its significance in photosynthetic organisms. Insertional inactivation of *isiA*, which was originally reported to be upregulated under iron limiting conditions, exhibited heat sensitive phenotype in *Synechocystis* (Kojima *et al.,* 2006). Thus collectively literature suggests that accumulation of major HSPs seems to be essentially required for thermal tolerance in *Synechocystis* (Kojima and Nakamoto, 2005). However regulation of such important heat shock genes is poorly understood in *Synechocystis*.

In *Synechocystis* the two component signal transduction systems involving a sensory histidine kinase and a cognate response regulator perceive the environmental signals to the target genes, apart from regulation of genes by various transcription factors (Paithoonrangsaid *et al.,* 2004; Shoumskaya *et al.,* 2005; Los *et al.,* 2008; Kanesaki *et al.,* 2007). It has been recently reported that a sensory kinase, Hik34 is involved in the down regulation of several heat shock genes and the Δ*hik34* deletion mutant showed an increased thermal tolerance (Suzuki *et al.,* 2005). In addition to the signal transduction involving Hik34, a positive regulation by alternative sigma factors, SigB and SigC was suggested to activate the transcription of *hspA* and *groEL* genes in *Synechocystis* (Tuominen *et al.,* 2006, 2008). In various gram negative and gram positive bacterial systems, different negative control mechanisms regulate heat shock genes (Narberhaus *et al.,* 1999). It clearly shows despite of the same function of these *hsp*s they follow different mechanisms of regulation. Positive and negative

regulation of *hsp*s in different model systems was presented in a table (Table 1.2 and 1.3). For instance, HAIR/HspR controls *dnaK* operon and *clpB* in *Streptomyces albus* (Bucca *et al.,* 1995), CtsR, transcriptional repressor negatively regulates *clpC, clpP* and other heat shock genes in *Bacillus subtilis* (Derré *et al.,* 1998), heat shock genes, *htpG, hspA* and others are known to be regulated by an unknown transcriptional repressor in *Bacillus subtilis, Clostridium acetobutylicum* and *Leuconostoc oenos* (Narberhaus *et al.,* 1999) and CIRCE/HrcA system (Hecker *et al.,* 1996), act as a negative regulator of *grpE/dnaK/dnaJ* and or *groE* operons (Van-Asseldonk *et al.,* 1993; Zuber and Schumann, 1994; Yuan and Wong, 1995; Nakamoto *et al.,* 2003). Transcriptional regulators like HrcA/CIRCE were known to involve in regulation of *groES*, *groEL* chaperonins in *Synechocystis* and *Bacillus sp.* (Zuber and Schumann1994; Nakamoto *et al.,* 2003). Small RNA mediated regulation of *isiA* gene was reported. Though majority of heat shock genes are under the control of Hik-response regulator little is known about their alternative transcriptional regulation. Certain genes which are induced in multiple stress conditions like *hspA*, *htpG*, *isiA* should be having a different method of control other than Hik-Rre two-component regulation. Indeed, expression of heat shock protiens, is tightly regulated in the acclimation processes. There was no report to specify the factors involved in regulation of *frpC* and *sll1130* and *ssl2245* which are differentially expressed in heat stress treatment. Several transcriptional regulators like *hrcA*, *lexA* and *sll0082* etc. have been reported in *Synechocystis,* showing the possibility of another transcription factor involvement in regulation of *hsp*s.

Table 1.2: Positive regulation of heat shock genes in different bacteria

Regulatory Protein	organism	Regulated gene(s) or operon(s)	References
σ32 (RpoH)	*Escherichia coli*	*dnaKJ, grpE, groESL* and more	Grossman *et al.*, 198
σE (σ24 RpoE)	*Escherichia coli*	*htrA, rpoH, rpoE* operon and more	Raina *et al.*, 1995
σB (SigB)	*Bacillus subtilis*	gspA, csbA, katE and more (classII)	Hecker *et al.*, 1996
σ54 (σN, RpoN) +Pspf	*Escherichia coli*	*psp*ABCDE	Weiner *et al.*, 1991
SigB and SigC	*Synechocystis* sp. PCC 6803	*hspA, groEL*	Tuominen *et al.*, 2006 and 2008

Table 1.2. Transcriptional regulators involved in positive regulation of heat shock genes in different bacteria.

Table 1.3: Negative regulation of heat shock genes in different bacteria

Regulator	**organism**	**Regulated gene(s) or operon(s)**	**References**
CIRCE/HrcA	*Bacillus subtilis*	*groESL, dnaK* operon (class I)	Hecker *et al.,* (1996)
HAIR/HspR	*Streptomyces albus*	dnaK operon *clpB*	Bucca *et al.*,(1995)
OrfY	*Streptomyces albus*	*hsp18*	Servant and Mazodier, (1996).
ROSE	*Bradyrhizobioum japonicum*	*hsp*A *rpo*H, and more	Narberhaus *et al.,* 1998
CtsR	*Bacillus subtilis*	*clpC, clpP* and more (classIII)	Hecker *et al.,* (1996)
Unknown	*Bacillus subtilis*	*htpG, ftsH, ion* and more (classIV)	Hecker *et al.,* (1996)
Unknown	*Clostridium acetobutylicum*	*hsp18*	Narberhaus 1999
Unknown	*Leuconostoc oenos*	*hsp18*	Narberhaus 1999
Unknown	*Synechocystis* sp. PCC 6803	*hspA* and *htpG*	Narberhaus 1999
CIRCE/HrcA	*Synechocystis* sp. PCC 6803	*groES, groEL*	Nakamoto *et al.,* 2003

Table 1.3 Negative regulation of heat shock genes in different bacteria. Regulation mechanism of different heat shock proteins was presented in table.

Toxin-Anti-toxin systems (TA systems) prokaryotic growth regulators:

Like in eukaryotes and higher animals, regulation of cell growth and cell death is important in bacteria, under various stress conditions. Almost all bacteria and some fungi contain the TA systems in their genomes, and constitutively express toxins and their cognate antitoxins and *E. coli* contains at least 34 TA systems. The TA systems are not essential for cell growth but are considered to play important roles in survival under stress conditions. Usually these toxins are small proteins approximately 100 amino acid residues in length. This toxin may inhibit cell growth by targeting a key molecule in any one of several essential cellular processes, including DNA replication, mRNA stability, protein synthesis, cell-wall biosynthesis, and ATP synthesis (Pandey and Gerdes 2005; Shao *et al.,* 2011). In normally growing cells, a toxin is associated with its cognate antitoxin and forms a stable complex there by prevents the toxin from exerting its toxicity. Hence, these toxins are co-transcribed and co-translated with their cognate antitoxins from an operon called a toxin-antitoxin (TA) operon (Pandey and Gerdes 2005; Yamaguchi *et al.*, 2011). As a result, in the normally growing cells, in order to constantly inhibit the toxin function, cognate antitoxins have to be continuously synthesized (Yamaguchi and Inouye 2009). The TA operon is classified into three types, Type I, II, and III, according to the nature of the antitoxin and the composition of the TA systems. In the Type I TA system, the toxin gene expression is regulated by an antisense RNA transcribed from the same toxin region in the reverse orientation (Fozo *et al.,* 2008; Gerdes and Wagner 2007). In the Type II TA system, both toxins and antitoxins consist of proteins, and the toxin function is neutralized by forming a complex with its cognate antitoxin (Yamaguchi and Inouye 2009). The Type III TA system, in which RNA antitoxin is proposed to inhibit toxin function by

directly interacting with the toxin protein by forming a RNA-protein complex (Fineran *et al.,* 2009) has been recently identified. The RNA antitoxin is not an antisense RNA, and therefore it does not inhibit the translation of the toxin mRNA like the Type I TA system. Among the TA systems, the Type II TA system has been most extensively studied. Examples of this TA system include MazE-MazF (Aizenman *et al.,* 1996), RelE-RelB (Pedersen *et al.,* 2003), YefM-YoeB (Kamada and Hanaoka 2005; Zhang and Inouye 2009) and MqsR-MqsA (Brown *et al.,* 2009; Yamaguchi *et al.,* 2009). Almost all the Type II TA systems share a number of common features (Gerdes *et al.,* 2005; Yamaguchi and Inouye 2009; Yamguchi *et al.,* 2011). Type II TA cassettes have a characteristic organization in which the gene for the antitoxin component precedes the toxin gene; the two loci often overlap, reflecting a common auto-regulatory mechanism exerted by both components. Although most TA modules conform to this arrangement, there are examples of TA cassettes in which the gene order is reversed, where the antitoxin alone exerts the regulatory effect or where the product of a third gene is implicated (Engelberg-Kulka and Glaser 1999; Pandey and Gerdes 2005).

PemK proteins have an insert beta-sheet sub domain and C-terminal helix domain, which are characteristic to PemK super family members. PemK-like protein or PemK is a growth inhibitor in *E. coli* known to bind to the promoter region of the Pem operon, auto-regulating synthesis. This Pfam family consists of the PemK protein in addition to ChpA, ChpB and other PemK-like proteins. PemK family members involved in stable maintenance of plasmid in *E. coli* (Masuda *et al.,* 1993). Sometimes these toxin antitoxin members act as modules of bacterial stress managers (Buts *et al.,* 2005). Summary of the well studied TA systems was presented as a table

(Table 1.2) and from the literature it is clear that PemK homologues are toxins, involved in either regulating growth or plasmid maintenance by killing the plasmid free cells. Homologues of Maz, YdcE, PemK, ChpBK cleave free RNAs there by Inhibit translation in *E. coli*. MazF/PemK family protein coded by *ydcDE* operon in *B. subtilis* is an endoribonuclease and regulates the translation. Pem, plasmid maintenance system may turn into growth inhibitor upon degradation of PemI (corresponding antitoxin). Table 1.4 illustrates the well studied TA systems.

Table 1.4: Properties of Toxin-Antitoxin systems in bacteria.

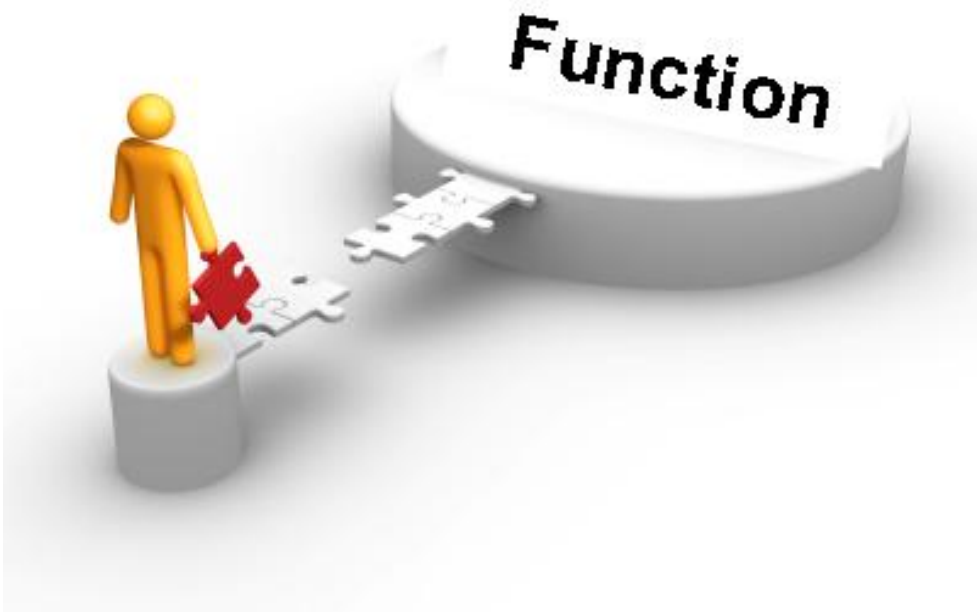
TA operon	Toxin	Antitoxin	Localization	Toxin homologies
parD(pem)	Kid (PemK)	Kis (PemI)	R1/R100 plasmid	CcdB/ChpAK/ChpBK
ccd	CcdB	CcdA	F1 Plasmid	Kid/ChpAK/ChpBK
chpA (mazEF)	ChpAK (MazF)	ChpAI (MazE)	*E. coli* chromosome	PemK (Kid)/ ChpBK/CcdB
chpB	ChpBK	ChpBI	*E. coli* chromosome	PemK (Kid)/ ChpAK/CcdB
relBE	RelE	RelB	*E. coli* chromosome	RNase T1
parDE	ParE	ParD	RK2/RP4 plasmid	RelE
hipBA	HipA	HipB	*E. coli* chromosome	CDK2/ cyclin A

Table 1.4. Summary of the well studied TA systems. Toxin homologies refer to proteins sharing a similar structure or amino acid sequence. The identity between the *parD–pem* and *chpA–mazEF* systems or their proteins is indicated in parenthesis. This table was adapted from Diago-Navarr *et al.*, 2010.

In *Synechocystis* TA systems have been very poorly experimentally characterized so far. Functionality of chromosomal encoded TA systems in cyanobacteria and autotrophic bacteria was unknown. Recently a chromosomal gene pair *ssr1114/slr0664*, named *relNE*s, of *Synechocystis* is identified as a TA system belonging to *rel* family. Ectopic expression of *Synechocysits* RelE led to the slow growth in *E. coli* due to its toxic activity (Ning *et al.,* 2011). *sll7003/ssl7004* TA system encoded on plasmid pSYSA of the *Synechocystis* is identified as a toxin with Mg^{2+} dependent RNA endonuclease activity targeting single-stranded RNA regions. Furthermore, four to seven additional TA system components are predicted on pSYSA (Kopfmann and Hess 2013), cyanobacteria are of high ecological importance and are considered promising for the production of biofuels, so characterization of these TA systems to understand the cyanobacterial responses and cell death in different conditions is very important.

Based on available literature, organization of genes in this operon, presence of PemK domain in Sll1130 and its conservation in various bacterial species, association of Sll1130 with the PSII complex and their heat dependent down regulated expression we have chosen this interacting protein pair (Ssl2245-Sll1130) with unknown function for functional characterization.

OBJECTIVES



2. OBJECTIVES

From the genome database it is clear that *Synechocystis* contains more than 50% of functionally unknown genes. *ssl2245-sll1130* is one overlapping gene pair coding for unknown proteins. Basal RNA levels of *sll1130* in the cell; at optimal conditions and report by Kashino *et al.*, (2002) shows that *sll1130* gene may be involved in the cell metabolism or photosynthesis. Sll1130 sequence similarity and its operon structure shows it can be a toxin evolved from bacterial TA systems. Present work was designed to identify the function of Sll1130 and Ssl2245 in *Synechocystis* by mutation, phenotype characterization, microarray analysis and by studying the purified protein characters.

Proposed objectives

I. Functional characterization of Sll1130

1. Targeted mutagenesis of *sll1130* in *Synechocysits* and to analyze the effect of mutation on function of PSII.
2. Over expression and purification of Sll1130 in *E. coli,* in order to check whether the protein has MazF like toxin properties.
3. DNA Microarray analysis of Δ*sll1130* to identify the role of Sll1130 in gene regulation as by sequence similarity it can be a possible transcriptional regulator.
4. Studies on expression pattern of *sll1130* and the genes that are differentially expressed in Δ*sll1130* upon heat stress.

II. Functional characterization of Ssl2245

1. Targeted mutagenesis of *ssl2245*, expression of an extra copy of *ssl2245-sll1130* in *Synechocystis* using cyanobacterial complementation vector pVZ321.
2. Phenotypic characterization of Δ*ssl2245*, Δ*sll1130*, Δ*DM (*Δ*ssl2245::*Δ*sll1130)* and *WT+*(WT with extra copy of *sll1130-ssl2245*).
3. DNA Microarray analysis of Δ*ssl2245* to identify the role of Ssl2245.

III. Role of Ssl2245 and Sll1130 interaction

1. Heterologous over expression of Ssl2245 and Sll1130, growth analysis and role of individual proteins on bacterial growth.
2. *In vitro* characterization of Sll1130 and Ssl2245 proteins.

In the present study we introduce an integrated strategy for elucidating functions of uncharacterized proteins, Sll1130 and Ssl2245. We combine targeted insertional inactivation of *sll1130,* and *ssl2245* and subsequent transcriptomic analysis of the mutants for getting clues on the function of this protein pair. Suitable physiological and biochemical experiments were performed to completely reveal the function of Sll1130 and Ssl2245. As will be presented, Sll1130 is a transcriptional regulator of certain group of genes and Ssl2245 is a genome encoded switch of genes in large native plasmids.

MATERIALS AND METHODS

SIGMA-ALDRICH
CORPORATION

GE Healthcare

3. MATERIALS AND METHODS

Escherichia coli

Escherichia coli (DH5α) was used for the maintenance of plasmid clones, cloning and DNA modification experiments while *E. coli* BL-21 (DE3) pLysS was used for the expression of His-tagged recombinant proteins. Bacterial cultures were grown/maintained in Luria Bertani (LB) broth or LB plates, respectively, with appropriate antibiotic(s) at 37 °C.

Luria and Bertani medium (L.B. MEDIUM)

Tryptone (10gm), Yeast extract (5gm), NaCl (10gm) and pH adjusted to 7-7.2. All these ingredients were added to 950ml of double distilled water (DDW). The pH was adjusted to 7.2 with 1N HCl/1N NaOH and DDW was added to make up to 1 liter. In order to prepare LB Agar, 1.5% of agar agar (w/v) was added to the LB solution. This is then sterilized by autoclaving at 15psi pressure for 15-20 minutes. In case of LB Agar the media was allowed to cool to 50°C before adding appropriate antibiotics. About 25ml of media was poured in 90mm Petri plate and allowed it to solidify. These solidified LB plates were stored at 4°C.

Preparation of Antibiotic LB medium

To give selection pressure appropriate antibiotics were added just prior to inoculation to the above mentioned LB medium. Filter sterilized Kanamycin, Spectinomycin and Chloromphenacol were added to a final concentration of 25µg/ml, 20µg/ml and 20µg/ml respectively.

Synechocystis sp. PCC 6803 GT-1:

Synechocystis, a glucose tolerant strain was originally obtained from Dr. J. G. K. Williams, (Dupont de Nemours, Wilmington, DE) and served as the wild-type. Wild-type cells were grown photoautotrophically, at 34°C in BG-11 medium (Stanier *et al.*, 1971), buffered with 20 mM HEPES-NaOH (pH 7.5), under continuous illumination at 70 µmol photons m^{-2} s^{-1} as described previously (Wada and Murata 1989). The liquid culture was bubbled with air. Δ*sll1130* culture, in which the *sll1130* gene in the genome had been disrupted by inserting kanamycin-resistance gene (Kan^R) cassette, was grown under the same conditions as described above with the exception that the culture medium contained kanamycin at 25 µg mL^{-1} in pre-cultures. Growth of the culture was monitored by measuring OD at 730 nm using a spectrophotometer (Model, UV-160A; Shimadzu Co., Japan). For heat treatment, wild-type *Synechocystis* culture was grown to mid log phase (an absorbance of ~0.7 OD at 730 nm) at 34°C and then shifted to a water bath maintained at 42°C, with continuous illumination (70 µmol photons m^{-2} s^{-1}). Cells were harvested before and after heat treatment, for isolating either total cellular proteins or total RNA. For heat shock and recovery experiment, 50 mL of wild-type and Δ*sll1130* mutant cultures were grown at 34°C to mid log phase and then incubated at 50°C for 30 min in a water bath. The heat shocked culture tubes were shifted back to 34°C and allowed to recover from the heat shock. Cells were collected before heat shock and at 48 h of recovery for viability test.

Preparation of BG-11 for *Synechocystis* sp PCC6803 culture:

BG-11 (supplemented with 20mM HEPES –NaOH, pH 7.5)

Stock 1: Citric acid-0.3g, Ferric ammonium citrate-0.3g and EDTA-0.05g weighed and adjusted to 100 ml with milliQ water, filter sterilized it separately and stored in dark and low temperature. **Stock 2:** $NaNO_3$-30g, K_2HPO_4-0.7g and $MgSO_4.7H_2O$-1.5g weighed and adjusted to 1litre with milliQ water. **Stock 3:** $CaCl_2.2H_2O$ -1.9g weighed and adjusted to 100ml with milliQ water. **Stock 4:** Na_2CO_3-2g weighed and adjust to 100ml with milliQ water. **Stock 5:** H_3BO_3-2.86g, $MnCl_2.4H_2O$-1.81g, $ZnSO_4.7H_2O$-0.222g, $Na_2Mo_4.2H_2O$-0.391g, $CuSO_4.5H_2O$-0.079g and $Co(NO_3)_2.6H_2O$-0.049g weighed and adjustd to 1 litre with milliQ water. **Stock 6:** HEPES-119.15g weighed and dissolved in 750 ml of milliQ water, adjusted pH to 7.5 with 2M NaOH and volume made up to 1 litre with milliQ water. All stocks are kept at cool temperature.

Table 3.1: Preparation of BG-11 from stocks

Stock solution	1X - 1000ml	2X - 500ml
Stock 2	50ml	50ml
Stock 3	2ml	2ml
Stock 4	1ml	1ml
Stock 5	1ml	1ml
Stock 6	40ml	40ml
Milli Q water	Up to 1000ml	Up to 500ml

Table 3.1. Preparation of BG-11 nutrient medium from the stocks

Stock 1 is added just prior to inoculation, 2ml in 1000ml of 1X BG-11solution after Autoclaving.

Preparation of BG-11 Agar plates

For preparation of agar plates 2X BG-11 and 2X agar agar (3% agar agar) were autoclaved separately and mixed in 1:1 ratio after autoclaving, poured in sterile petri plates after addition of stock-1 and100mM $Na_2S_2O_3$ (10ml for 1 lit).

Preparation of BG-11 Antibiotic medium

To give a selection pressure appropriate antibiotics were added just prior to inoculation to the above mentioned Bg-11 medium. Filter sterilized Kanamycin, Spectinomycin and Chloromphenicol were added to a final concentration of 25µg/ml, 20µg/ml and 20µg/ml respectively.

Preparation of BG-11 nutrient limited medium

To give nitrogen limitation, nitrate in the stock-II was substituted with equimolar concentration of NaCl. 20.6 gm of NaCl, 0.78 gm of K2HPO4 and 1.5 gm of $MgSO_4$ $7H_2O$ were added to water and made it to a final volume of 1 Lit and used instead of BG-11, Stock-II solution. For creating iron limitation, to the BG-11, stock-I solution ferric ammonium citrate was not added.

Preparation of BG-11 nutrient limited medium

To give nitrogen limitation, nitrate in the stock-II was substituted with equimolar concentration of NaCl. 20.6 gm of NaCl, 0.78 gm of K2HPO4 and 1.5 gm of $MgSO_4$ $7H_2O$ were added to water and made it to a final volume of 1 Lit and used instead of BG-11, Stock-II solution.

Kits, enzymes, chemicals and reagents

Molecular biology kits and enzymes were procured from Simga-Aldrich (USA), Qiagen (Germany), MBI Fermentas (Germany) and Takara bio (Japan). Precautions were taken as per manufacturers' instructions. The chemicals and reagents were of

analytical grade and obtained from Sigma-Aldrich (USA), GE health care (USA), Fermentas (Germany), Agilent technologies (USA) Himedia (India), SRL (India) and Qualigens fine chemicals (India).

Molecular biology protocols

Plasmid and genomic DNA isolation, restriction digestion, agarose gel electrophoresis, ligation, competent cell preparation and transformation were according to Sambrook *et al.,* (1989) and/or as per manufacturers' protocol.

Plasmid DNA vectors

Commercially available T-vector of fermentas /Bangalore genei was used for simple TA-cloning experiments and blue white screening. pVZ321 a cyanobacterial vector was used to express the *ssl2245-sll1130* operon with its native promoter in *Synechocystis*. pET28a(+) a high copy number prokaryotic expression vector was used for the cloning and expression of Sll1130, Ssl2245 separately and together.

Quantification of DNA and RNA:

The quality and concentration of RNA and DNA was examined by ethidium bromide-stained agarose gel electrophoresis and spectrophotometric analysis (nanodrop ND-1000). The concentration of DNA and RNA was determined using spectrophotometry by measuring the absorbance at 260 nm and 280 nm. A value of OD260 =1 corresponds to 50 µg/ml for DNA, while OD260 =1 corresponds to 40 µg/ml for RNA. A value of OD_{260}/OD_{280} between 1.8 and 2.0 and between 1.9 and 2.1 was considered as pure form of DNA and RNA, respectively.

Oligonucleotides and sequencing

All DNA oligonucleotides were synthesized by either Sigma Aldrich or MWG Biotech (Eurofins). Nucleic acid sequencing was carried out by Eurofins as per

requirement. All the expression clones used for purification were prior confirmed at sequence level.

PSII activity measurement

Photosynthetic oxygen evolution was measured in 1 mL of cell suspension (OD730 of ~1, about 5 μg mL^{-1} chlorophyll), using an oxygen electrode (Oxygraph plus, Hansatech Instruments Ltd., Norfolk, England). The cells were cultivated photoautotrophically under aeration at 70 μmol photons m^{-2} s^{-1} light. PSII activity was determined in the presence of 1.0 mM p-benzoquinone (PBQ) and 1.0 mM potassium ferricyanide ($K_3Fe[CN]_6$). Photosynthetic oxygen evolution was recorded at 400 μmol photons m^{-2} s^{-1}. Three independent cultures of wild type and *Δsll1130* cells were measured with technical triplicates.

Estimation of chlorophyll

Chlorophyll was estimated by a modified method from Arnon *et al.*, (1974). 1 ml of cells, of density 0.4 (O.D 730nm) was centrifuged at a speed of 15k rpm and supernatant was discarded to the maximum extant possible. Cell pellet was resuspended in 200 μL of water and added 800 μL of acetone (to final 80% (v/v) conc. Acetone) kept on shaking for 15min and centrifuged at high speed (15k rpm) for 15 min pellet containing phycobilisomes is deep blue. Absorbance of the supernatant was measured at 663nm taking 80% acetone as the blank. The amount of chlorophyll was calculated using the formula; [Chl μg / ml = abs at 663nm X 12.189]. Chlorophyll content was calculated from the peak heights of absorption spectra (live cells) also using the equations of Arnon *et al.*, (1974).

Viability test

Synechocystis wild-type and Δ*sll1130* cells were stained with ViaGram^{TM} Red^{+} Bacterial Gram Stain and Viability Kit according to the manufacturer's instructions (V-7023, Molecular probes, Invitrogen, Carlsbad, CA). Three µL of SYTOX stain was diluted to 60 µL with water. To the 50 µL of cell suspension 2.5 µL of the diluted SYTOX Green was added and incubated at room temperature (28^{o}C) for 15 min. Ten µL of the stained cell suspension was examined with a confocal microscope (TCSSP-2, AOBS 4 channel UV and visible; Leica, Heidelberg, Germany).

Gel permeation chromatography of Sll1130

Gel permeation chromatography of purified Sll1130 protein (100µg) was performed on a Sephacryl S-100 column (dimensions: 50 x 0.5 cm) in 50mM phosphate buffer (pH 7.0). The gel filtration molecular weight markers, bovine serum albumin (BSA) and carbonic anhydrase (CA) were dissolved at a concentration of 0.1 mg/ml in 100 µl of purified Sll1130 protein solution (Cat No: A8654 and C5024, Sigma-Aldrich) and loaded onto the chromatography column (100 µl sample on a Sephacryl S-100, GE Healthcare). About 37 fractions (each equivalent to 0.25 ml) were collected for analysis. Elutes were estimated for the presence of proteins by UV absorption at 280 nm. Equal volumes of collected fractions, were resolved on SDS-PAGE (15%) and subsequently the amount of protein present in each band was estimated densitometrically. The Coomassie blue stained gels were scanned in a Gel Logic 212 PRO (Kodak) and intensity of each protein band corresponding to molecular weight markers and Sll1130 were further calculated by Carestream Molecular Imaging software (Carestream Health, New Haven).

Denaturing and semi-denaturing SDS gels

Proteins were separated on SDS-PAGE, 12% or15% as per the requirement. The over-expressed and purified Sll1130 protein was examined under both denaturing (loading dye containing β-mercaptoethanol was mixed with the protein sample and boiled before loading in the well) and semi-denaturing conditions (Purified protein mixed with loading dye without β-mercaptoethanol and not boiled before loading in the well) to find oligomerization state of Sll1130 protein. Purity of the purified proteins was also confirmed by resolving the protein in SDS-PAGE (15%), denaturing conditions. Protein confirmation was done by MALDI-TOF-TOF analysis of gel excised protein.

DNA microarray chip

Takara chip (DNA microarray chip) DNA microarray covers 3079 (97%) of the 3165 genes on the chromosome of *Synechocystis* (99 genes for transposases are excluded from this calculation). Moreover, the entire sequences of all four of the plasmids (pSYSM, pSYSX, pSYSA and pSYSG) harbored by *Synechocystis* have been reported (Kaneko *et al*., 2003). There are 132, 110, 106 and 49 putative genes, respectively, on these plasmids, namely, 397 genes in total. However, these genes are not included in the DNA microarray from Takara Bio Co. **Agilent chip (DNA microarray chip)** Agilent's *Synechocystis* 15K oligo DNA microarray used in these studies was custom designed and constructed using Agilent noncontact inkjet technology (Agilent). The probes printed on these arrays are 60-mer oligomers that are known to yield excellent sensitivity and specificity. Candidate oligonucleotides representing 3,459 genes *in Synechocystis* were selected from the 3' end of genes. In the case of genes present on plasmids, target sequences were selected based on

Kazusa annotation (Kanesaki *et al*., 2003). The expression microarray holds probe sets for all annotated genes from the chromosome (NC_000911) as well as the seven plasmids (pSYSA: NC_005230, pSYSG: NC_005231, pSYSM: NC_005229, pSYSM:NC_005232, pCA2.4, pCB2.4, pCC5.2 available at http://genome.kazusa.or.jp/cyanobase/Synechocystis/. On average, 3 to 5 probes per transcript were present on each array.

Extraction of total RNA

RNA from the Synechocystis cells was isolated essentially by the method of Los *et al*., 1997. Actively growing *Synechocystis* cells (treated or untreated (or) mutant/WT) in bubbling culture (O.D-0.4-0.6) was immediately killed by addition of equal volume of ice cold 5% (w/v) phenol in ethanol, cells were harvested at 4^{o}C and cell pellet is stored in -80^{o}C. Cells were lysed by heating the cells in presence of 1% SDS and using acid phenol method. To the obtained nucleic acids DNaseI treatment was given, to remove the contamination of genomic and plasmid DNA and the obtained RNA pellet was dissolved in 30 µl of DEPC water and stored in -80^{o}C.

Expression and purification of recombinant proteins (His-tag)

E. coli BL21 DE3 competent cells were transformed with confirmed recombinant plasmids and plated on LB kanamycin (25 µg/ml) plate. Single colony was inoculated on to 5 ml of LB broth with antibiotics and incubated over night with shaking. Overnight culture was diluted 50 times with fresh LB broth with antibiotics and incubated at afore said conditions. After reaching OD_{600} 0.5, cells were induced with 400 µM IPTG to express its peptides. Cells were harvested after 4 h of induction and washed with 100 mM Tris.Cl pH 8.0 with 300 mM NaCl and resuspended in same buffer. Cells were disrupted by sonication and insoluble debris was removed by

centrifugation. Protein preparations were compared with un-induced controls to check the expression of peptides by SDS PAGE. Five ml of Ni-NTAgarose (Sigma-Aldrich) was packed in a polypropylene column and equilibrated with 25 ml of Tris-NaCl buffer with 5 mM imidazole. Sonicated supernatant was applied to the column and flow through was collected and stored at -20°C. The column was washed with 20 ml of washing buffer (Tris-NaCl buffer with 25 mM imidazole). Finally the protein was eluted with 15 ml of elution buffer (Tris-NaCl buffer with 200 mM imidazole) and fractions were collected. Purity of the protein was analyzed by SDS-PAGE, followed by Coomassie blue staining. Fractions containing pure protein were pooled and stored at -20°C after addition of glycerol to final concentration of 10% (v/v).

Generation of *sll1130* mutant, Δ*sll1130*

We generated a Δ*sll130* mutant of *Synechocystis* by insertional inactivation of the *sll1130* gene. A DNA fragment containing the *sll1130* ORF with 69-bp upstream and 197-bp downstream flanking regions was amplified by PCR with the following primers: *sll1130*-F (5' TAT GGC TGC CAC CGC CGA CAC TAT GAC 3') and *sll1130*-R (5' GTC CAG CGG ATC TAG TTA TCT TTC CAG 3'). The PCR amplified 614 bp fragment was ligated to a linear T- vector (GeNei™ INSTANT Cloning kit, Cat No: 107416, Bangalore Genie Pvt. Ltd., Bangalore, India) and the resultant plasmid pT-*sll1130* was used to inactivate the *sll1130* ORF by performing *in vitro* transposon reaction according to the protocol provided by the manufacturer (EZ::Tn5™ <KAN-2> Insertion kit, Cat No: EZ1982K, Epicentre, Madison, WI, USA). Plasmid DNA construct in which *sll1130* ORF was disrupted with the kanamycin resistance gene cassette was designated as pT*sll1130::kan*^{R}. This

construct was used to transform *Synechocystis* cells. The site of insertion of kanamycin cassette was located by sequencing the pT*sll1130::kan*^{R} construct using Kan-RP1 primer.

For transformation, wild-type *Synechocystis* cells were grown to an OD_{730} of about 0.6, adjusted to OD_{730} of 2 by centrifuging cell culture at 3,000 *g* for 5 min and re-suspending the resulted cell pellet in a small volume of fresh BG-11 medium. About 5 µg of pT*sll1130::kan*^{R} was mixed with 150 µL of cell suspension and incubated for 16 h in a sterile microfuge tube at 30°C in light at 50 µmol photons m^{-2} s^{-1} without shaking. The cells were then spread on BG11 agar plates that contained a concentration gradient of kanamycin. Single colonies, which had developed on the BG11-containing agar in the region where kanamycin was present at a low concentration, were re-plated on the BG11-containing agar supplemented with 10 µg mL^{-1} kanamycin. Thus cells that had been grown actively were subsequently re-plated on the BG11 medium, which contained increasing concentrations of kanamycin. The maximum concentration of kanamycin used was 25 µg mL^{-1}. As *Synechocystis* contains multiple identical genome copies, increasing antibiotic pressure is necessary to completely replace all wild-type copies of the *sll1130*, with the *in vitro* disrupted copy of the gene, *sll1130::kan*^{R} by homologous recombination. Genomic DNA of the *Δsll1130* mutant cells grown for several rounds in BG11 medium was prepared and extent of the replacement of wild-type copies of *sll1130* with the *sll1130::kan*^{R} was checked using primers *sll1130*-F and *sll1130*-R. The mutant thus generated was named Δ*sll1130*.

Targeted mutagenesis of *ssl2245* (Δ*ssl2245*):

We generated a Δ*ssl2245* mutant by insertional inactivation of the *ssl2245* gene. A DNA fragment containing the *ssl2245* ORF and *sll1130* ORF in downstream as flanking region was amplified using PCR with the *ssl2245*-exp-F and *sll1130*-exp-R primers (F 5' ATGTCTATCAATGCTTACAAACTAGCTACG 3' and R 5' GCGAAGCTTACC GAGTTTAAAAACATGGGG 3'). The resulting 615 bp fragment was ligated to a linear T vector (GeNei™ INSTANT Cloning kit, Cat No: 107416, Bangalore Genie Pvt. Ltd., Bangalore, India). The resultant plasmid pT-*ssl2245* was used to inactivate the *ssl2245* ORF by performing a EcoRI digestion and subsequent blunting of it using *pfu* DNA polymerase followed by ligation with omega spectinomycin cassette and transformation into *E. coli*. Selected colonies were screened for omega spectinomycin by colony PCR, confirmed by DNA sequencing. Plasmid with disrupted gene was transformed to *Synechocystis* sp. PCC 6803, as mentioned earlier, except that the selection pressure was given by using spectinomycin. Cells that had been grown actively were subsequently re-plated on the BG11 medium, which contained increasing concentrations of spectinomycin and finally maintained on BG-11 plates with 20µg/ml final concentration of spectinomycin.

Synechocystis strain with an extra copy of the *ssl2245-sll1130* (WT+)

We generated a WT+ strain (WT with an extra copy of *ssl2245-sll1130* with native promoter expressed in pVZ-321) was generated. *ssl2245-sll1130* operon and an upstream of 300 bp was PCR amplified using specific primers WT+ F: 5' CATACTTTAGCTTGTCGTCCCGGGC 3' *sll1130*-exp-R 5' GCGAAGCTTACC GAGTTTAAAAACATGGGG 3'). The resulting 900 bp fragment was ligated to

smaI digested pVZ321 vector. The resultant plasmid pVZ-*ssl224-sll1130* was transformed into WT *Synechocystis* by triparental mating (Zinchenko *et al*., 1999). Transformed colonies were selected with antibiotic selection pressure of 25 µg of chloromphenicol/ml on BG-11 agar medium. The so formed strain was named as WT+ due to presence of an extra copy of the gene pair (*ssl2245-sll1130*).

Preparation of antibodies against Sll1130

Antibodies against Sll1130 were raised in rabbits with His-tagged Sll1130 of *Synechocystis,* which had been overexpressed in *E. coli* as described before (Prakash *et al*., 2010). The *sll1130* ORF was PCR amplified with the following primers: *sll1130*-ExF (5' GCG CCA TGG AT ACA ATT TAC GAA CAA TTT G 3') and *sll1130*-ExR (5' GCG AAG CTT ACC GAG TTT AAA AAC ATG GGG 3'). The NcoI and HindIII restriction sites are underlined. The ninth nucleotide 'G' was substituted for 'A' to create NcoI site in the forward primer. The amplified ORF of *sll1130* was inserted into pET28a(+) at the NcoI and HindIII sites to generate pET-sll1130. The C-terminally His-tagged Sll1130 protein was expressed in BL21(DE3)pLysS, which had been transformed with pET-sll1130 and was purified using HIS-Select^{TM} Nickel Affinity gel (SIGMA - P6611) according to the supplier's instructions. The expression of Sll1130 protein was induced by addition of 400 µM (Final concentration) IPTG. Bacterial cells were collected by centrifugation at 10,000 *g* for 10 min and pelleted cells were disrupted with a sonic oscillator (Model, UV2070, probe MS-72, Bandelin Electronic, Berlin) operated for 10 min at 50% power, with 1 min pulse interval, in 100 mM Tris-HCl (pH 8.0) and 200 mM NaCl. Insoluble materials were removed by centrifugation at 20,000 *g* for 20 min at 4^{o}C. The supernatant was loaded onto a HIS-Select ^{TM} Nickel Affinity column. The column

was washed with 50 mM NaH_2PO_4 (pH 7.5), 200 mM NaCl and 10 mM imidazole and sequentially with the same buffer containing 40 mM imidazole. Then His tagged Sll1130 was eluted with 50 mM NaH_2PO_4 (pH 7.5), 200 mM NaCl and 200 mM imidazole. The purity of each fraction was examined by SDS gel electrophoresis. The fractions which gave a single protein band at the expected region on the gel were combined and dialysed against 50 mM NaH_2PO_4 (pH 7.5) using amicon ultra filters (Millipore-UFC800324). The resultant protein was used to generate anti-Sll1130 antibody in New Zealand white rabbits with Freund's complete adjuvant (primary) and Freund's incomplete adjuvant (booster). Serum which was collected after the second booster was diluted 3000 times in 3% BSA and used for detecting Sll1130 protein.

Gel mobility shift assays

A 283 bp DNA fragment containing the region upstream from a translation start site of *slr1788* was obtained by PCR using Cy3 labeled primers 5'-GAG AAA CTG ATC TTG AAG AAG TGG-3' and 5'- GAT TTT GGT AAT TGG ATC ATG GCG-3'. Cy-3 labeled primers were obtained from Eurofins Genomics Pvt. Ltd. Bangalore, India. 270 ng of Cy3 labeled PCR fragment was incubated with various concentrations of His-Sll1130 in 25 µL of binding buffer [10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.5 mM EDTA, 2% glycerol, 0.5 mM DTT, and 1 mM $MgCl_2$ and 4 µg of poly dI-dC**]** for 30 min at room temperature. The mixtures were subjected to electrophoresis on a native 6% polyacrylamide gel at 60 V. A specific competitor DNA was prepared by annealing the following oligonucleotides; 5' TGC CCC AAG CAA GAA TGG CGA TCG CCG TTG GGA TTC GGA GC 3' and 5' GCT CCG AAT CCC AAC GGC GAT CGC CAT TCT TGC TTG GGG CA-3'. Inverted-repeat,

which was predicted to be the binding site of Sll1130 is underlined in the oligonucleotides. Analyses of gel images were carried out using epi-luminescence and Cy3 filter in a Syngene G:box image analyzer.

Preparation of cDNA for DNA microarray analysis

Wild-type *Synechocystis* cells and Δ*sll1130* mutant cells that were grown at 70 µmol photons m^{-2} s^{-1} (50 mL) were killed instantaneously by the addition of 50 mL of ice cold 5%-phenol in ethanol (w/v) and then total RNA was extracted as described previously (Los *et al.,* 1997). The RNA was treated with DNase I (Nippon Gene, Tokyo, Japan) to remove contaminating DNA. cDNAs, labeled with fluorescent dyes (Cy3 and Cy5; Amersham Pharmacia Biotech) were prepared from 10 µg total RNA with an RNA fluorescence labeling core kit (M-MLV, version 2.0 Takara Co. Ltd., Kyoto, Japan) according to the manufacturer's instructions. For Agilent chip, c-DNA was prepared and labeled using Fair play three microarray labeling kit (Agilent, La Jolla, CA) as per the manufacturer's instructions.

DNA microarray analysis of Δ*sll1130*/WT and Δ*ssl2245*/WT

Genome-wide analysis of transcript levels was performed with DNA microarrays as described previously (Kanesaki *et al.*, 2002; Prakash *et al.*, 2010). In brief, we used a *Synechocystis* DNA microarray (CyanoCHIP, Takara Co. Ltd.) that covered 3079 of the 3168 open-reading frames (97% of all genes, with the exception of transposon-related genes) of the *Synechocystis* genome. Hybridization of labeled cDNAs to the DNA microarray was allowed to proceed at 65^{o}C for 16 h. After hybridization, the microarrays were rinsed with 2x SSC (1x SSC is 150 mM NaCl and 15 mM sodium citrate) at room temperature. They were washed with 2x SSC at 60^{o}C for 10 min and with 0.2x SSC, 0.1% SDS at 60^{o}C for 10 min and then rinsed with

distilled water at room temperature for 2 min. Moisture was removed with an air spray prior to analysis with the array scanner (GMS418; Affimetrix, Woburn, MA, USA). Each signal was quantified with the ImaGene ver. 4.0 program (BioDiscovery, Los Angeles, CA, USA). The signal from each gene on the microarray was normalized by reference to the total intensity of signals from all genes, with the exception of genes for rRNAs. Then we calculated the change in the level of the transcript of each gene relative to the total amount of mRNA.

Genome-wide analysis of transcript levels of Δ*ssl2245* was performed with DNA microarrays in brief, we used a *Synechocystis* DNA microarray (Agilent chip, Agilent, La Jolla, CA) that covered 3,459 genes including the genes of all native plasmids in *Synechocystis* genome. cDNA was prepared using fair playIII microarray labeling kit (Cat.No.252009, Agilent, La Jolla, CA). Cy3 and Cy5 dyes (Cat.No. PA23001 and PA25001, GE healthcare) were couple to the prepared cDNA as per the manufacturer's instructions. Dye coupled cDNA was purified using microspin columns. Hybridization of labeled cDNA to the DNA microarray was done as per the manufacturer recommendations. Total 45µl of hybridization mix with Cy3 and Cy5 labeled DNA, and other hybridization contents were allowed to hybridize at 65^{o}C for 16-18hr in a hybridization chamber. Washings were done using Agilent gene expression wash buffer one and two provided by Agilent as per manufacturer's instructions. After washings the chip was scanned using Agilent microarray chip scanner (G2505B, Microarray scanner, Agilent technologies) with red and green dye channel. Scanning was done using green and red PMT at 100% (XDR Hi 100%) and 70% also. Feature extraction was done using Agilent feature extraction (FE) software version 9.5.1 as per protocol mentioned in web site

(www.agilent.com/chem/feprotocols). The signal from each gene on the microarray was normalized by reference to the total intensity of signals from all genes. Then we calculated the change in the level of the transcript of each gene relative to the total amount of mRNA.

Quantitative PCR

The RNA isolated from wild-type and Δ*sll1130* cells was used for cDNA synthesis with Affinity Script cDNA synthesis kit, according to the manufacturer's protocol (Cat No: 600559, Agilent, La Jolla, CA). qRT-PCR was carried out using the Power SYBR Green master mix kit (Cat No. 4368577, Applied biosystems). Each reaction was carried out in a 25 µL volume containing 12.5 µL Power SYBR green master mix, 0.2 µM of each primer and 5 µL diluted cDNA (35 ng). All reactions were run in duplicates using a qRT-PCR instrument (Model, Mx3005P, Agilent, La Jolla, CA). The instrument was programmed at 95^{o}C for 10 min; and then 40 cycles of 30 s at 95 °C, 30 s at 60°C and 30 s at 72°C. For each reaction the melting curve was analyzed and the PCR product was run on the agarose gel in order to confirm the specificity of the RT-PCR. Expression levels were normalized using either *gap1* or 16SrRNA gene as an internal reference. Primers used for qRT-PCR are listed in a table in chapter titled primer sequences.

Western blotting analysis

Soluble proteins were extracted by mechanical disruption of *Synechocystis* cells using glass beads (106-µm diameter, Sigma). *Synechocystis* cells in 200 µL of 50 mM Tris-HCl (pH 8.0) and 1 mM EDTA buffer were mixed with 330 mg glass beads in a thick walled glass tube and disrupted by vigorous vortex mixing at maximum speed on a vortex mixer for 1 min followed by 2 min cooling on ice.

Vortex mixing and cooling on ice was repeated 10 times to ensure maximum disruption of cells. Then mechanically disrupted cells were centrifuged at 25,000 *g* for 20 min to separate soluble proteins from insoluble material. The resultant supernatant which contained soluble proteins from control and heat treated cells, were loaded on to SDS-PAGE gel on equal protein basis (30 µg soluble protein in each well). After electrophoresis, the separated proteins were blotted onto polyvinylidene fluoride membrane (Immobilon-P; Millipore) in a semidry transfer apparatus (TE77-PWR semi dry transfer unit, GE health care). Levels of Sll1130 were determined immunologically with NBT-BCIP chromogenic detection system according to the manufacturer's instructions (Sigmafast BCIP/NBT, Cat No: B5655, Sigma-Aldrich). Polyclonal antibodies raised in the rabbit against His-Sll1130 protein were used as primary antibody and alkaline phosphatase-linked antibody raised in goat against rabbit IgG was used as the secondary antibody.

Northern blotting analysis:

Total RNA was extracted from cells, and Northern blotting analysis was performed as described by Los *et al*. (1997). At various time points of treatment cells were harvested for RNA extraction and subsequent Northern blotting analysis. DNA fragments corresponding to *sll1130* and *rnpB* gene were conjugated with alkaline phosphatase (Cat. No. RPN3690, CDP star Alkphos Direct kit; Amersham Pharmacia Biotech) and the resultant conjugates were used as probes. After hybridization, the blots were soaked with the CDP-star solution (Amersham Pharmacia Biotech) and signals from hybridized mRNAs were detected with a Fujifilm, LAS-4000 luminescence image analyzer.

Identification of a common *cis*-regulatory element upstream of genes upregulated by mutation of *sll1130*

A 500 bp upstream and/or intergenic DNA region of each gene whose expression was upregulated due to mutation in *sll1130* was submitted to MEME ver. 4.3.0 (Motif extraction by multiple expectation maximization). MEME was run using the default parameters. MEME identified a common inverted-repeat (*cis*-regulatory element) upstream of majority of the genes whose expressions were upregulated by the mutation. A consensus nucleotide sequence was generated by aligning the *cis*-regulatory elements upstream of all the upregulated genes.

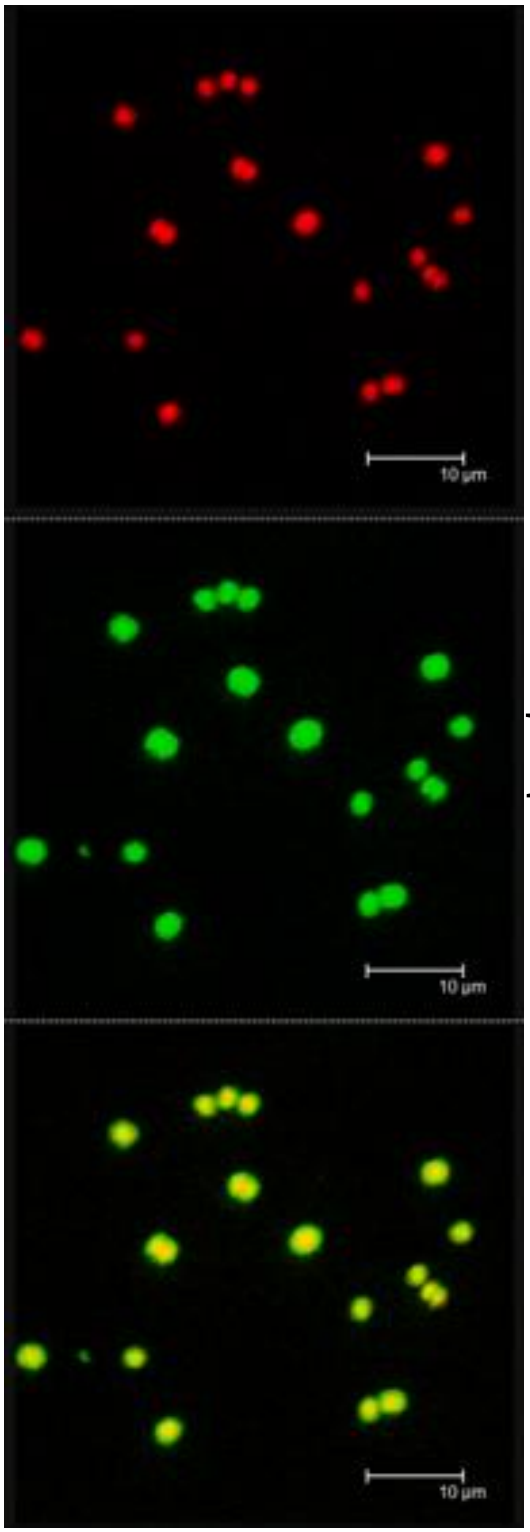
Localization of Sll1130 by Western blotting analysis:

Western blotting analysis was performed as described previously with some modifications (Prakash *et al*., 2010). Polyclonal antibodies raised in rabbit against His-Sll1130 protein were used as the primary antibody, and alkaline phosphatase-linked antibody raised in goat against rabbit IgG was used as the secondary antibody. After SDS-PAGE separation, proteins were blotted onto polyvinylidene fluoride membrane (Immobilon-P; Millipore) in a semidry transfer apparatus (TE77-PWR semi dry transfer unit, GE health care). Levels of Sll1130 were determined immunologically with NBT-BCIP chromogenic detection system according to the manufacturer's instructions (Sigmafast BCIP/NBT, Cat No: B5655, Sigma-Aldrich).

Over expression of *ssl2245* and *ssl2245- sll1130* operon:

The *ssl2245* ORF is PCR amplified with primers FP: 5’ GCG CCA TGT CTA TCA ATG CTT ACA AAC TAG CTA CG3’ RP: 5’ GCG AAG CTT TCA TAG GTG TCG GTA TGC AGA ATT ATC AGC 3’ the operon is amplified with the same forward primer and *sll1130* exp reverse primer (5’GCG AAG CTT ACC GAG TTT

AAA AAC ATG GGG3') and cloned as a recombinant fusion protein containing a N-terminal His tag in pET28a(+) (Novagen) by digesting with NdeI and HindIII to generate pET-*ssl2245*. The His-tagged Ssl2245 (HIS-Ssl2245) protein was expressed in BL21(DE3)pLysS by transforming pET-his-*ssl2245-sll1130*. Cells were induced with IPTG(0.4mM) in log phase, Induced bacterial pellet was lysed in 0.05M TRIS, 0.1 M NaCl, by ultra sonication in UV2070 sonicator and insoluble materials were removed by centrifugation at 10,000 rpm for 30min. so formed proteins were resolved on SDS-PAGE.



RESULTS

4. RESULTS

Functional characterization of Sll1130

About 50% of *Synechocystis* genes code for functionally unknown proteins. Sll1130 is one such functionally unknown protein. Sll1130 may have role in structure and function of PSII, as this protein was reported to be associated with PSII (Kashino *et al.*, 2002). Hence, we inactivated Sll1130 and attempted to assess its role in PSII structure and function.

Sll1130 is a conserved protein with an unknown function

The open reading frame, *sll1130* encodes an unknown protein in *Synechocystis* (Kaneko *et al.*, 1996). Figure 4.1 shows the alignment of top 10 BlastP hits of Sll1130 protein. The protein is well conserved and its orthologs are present in bacteria of different taxonomic phyla such as Proteobacteria, Cyanobacteria and Spirochaetes. Secondary structure prediction by GORIV suggests that Sll1130 protein has four helices spread throughout its length (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html). The first three helical regions were well conserved among all orthologs of Sll1130 (Figure 4.1). The protein has a putative conserved PemK superfamily domain. Proteins containing the PemK domain are common to many bacterial species and are reported to play roles in stress-mediated growth regulation or plasmid maintenance (Masuda *et al.*, 1993). Some protein members of the PemK superfamily are known to be toxin / antitoxin proteins that regulate cell growth under antibiotic or abiotic stress (Gerdes *et al.*, 2005).

Figure 4.1: *In silico* analysis of the Sll1130 protein

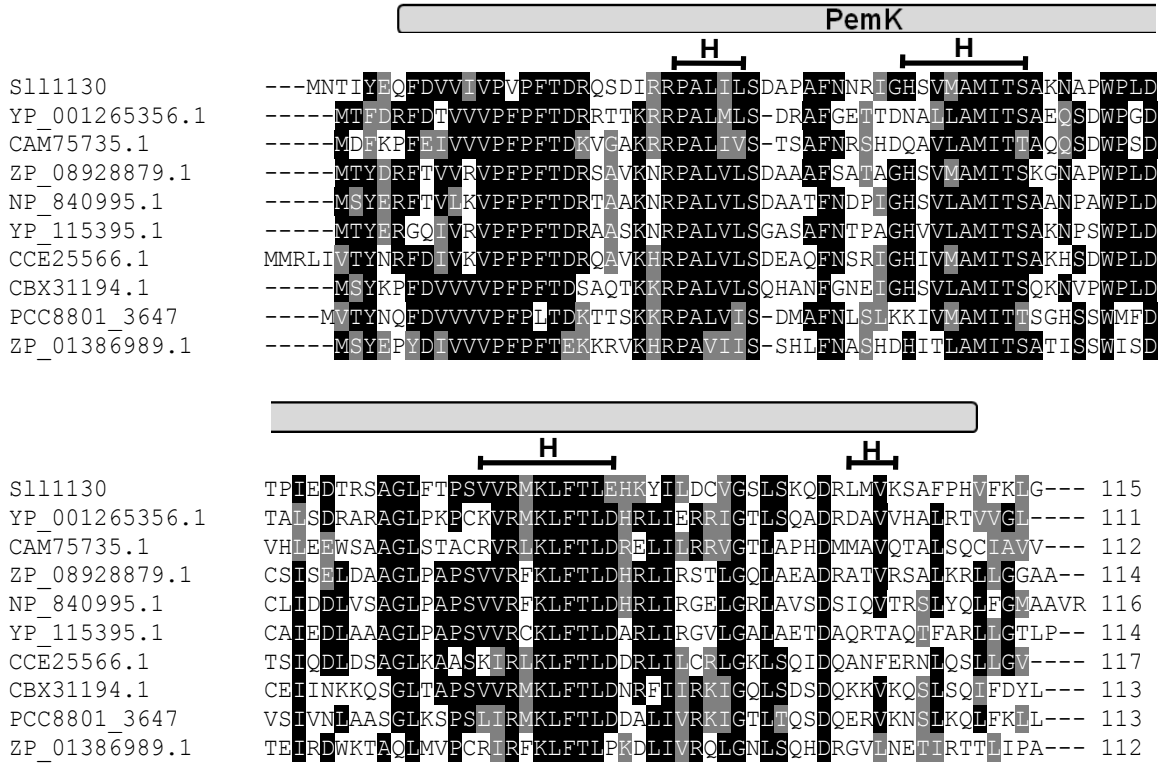


Figure 4.1. Multiple sequence alignment of the Sll1130 from *Synechocystis* sp. PCC 6803 and its homologs retrieved from the NCBI database. The sequences were aligned using the ClustalW algorithm. Identical amino acids are shown in black and conserved amino acid substitutions are in grey. The first three helices (H), starting from the N-terminal end of Sll1130 are conserved among all orthologs of Sll1130. Helix regions were predicted using the GORIV secondary structure prediction tool. PemK domain on Sll1130 protein was identified by Pfam database search and is indicated as a grey bar. Sll1130, *Synechocystis* sp. PCC 6803; YP_001265356.1, *Sphingomonas wittichii RW1*; CAM75735.1, *Magnetospirillum gryphiswaldense MSR-1*; ZP_08928879.1, *Thioalkalivibrio thiocyanoxidans ARh4*; NP_840995.1, *Nitrosomonas europaea ATCC 19718*; YP_115395.1, *Methylococcus capsulatus str. Bath*; CCE25566.1, *Methylomicrobium alcaliphilum*; CBX31194.1, *Desulfobacterium* sp.; PCC8801_3647, *Cyanothece sp. PCC 8801*; ZP_01386989.1, *Chlorobium ferrooxidans DSM 13031*.

Despite of having PemK domain, Sll1130 is not a toxin

Since the Sll1130 protein is well conserved among various bacterial species and showed similarity to the PemK superfamily proteins, it is likely that it may have an important physiological function and play role(s) in survival of the cell. In order to check whether Sll1130 is a toxic protein, it was expressed in *E. coli* cells using pET expression system. A non-toxic protein, Sll1961 of *Synechocystis* was also expressed using pET expression system in *E. coli* as a control. *E. coli* cells harboring pET28a(+) expression vector without any *Synechocystis* gene cloned into it was used as negative control. Figure 4.2 clearly indicate the synthesis and accumulation of Sll1130 and Sll1961 proteins after 4 hours of IPTG induction in *E. coli*. Sll1130 protein band can be seen at 15 kDa region and Sll1961 protein band at 40 kDa respectively (Figure 4.2 A and 4.2 B). Growth profiles of *E. coli* expressing these proteins were monitored in comparison with *E. coli* harboring only pET28a(+) vector without any *Synechocystis* ORF cloned into it (Figure 4.2 C). It is clear from Figure 4.2 C that all three *E. coli* strains exhibited similar growth profiles, suggest that Sll1130 is not a toxic protein when it is expressed in *E. coli*.

Figure 4.2: Growth analyses of *E. coli* cells expressing *Synechocystis* proteins, Sll1130 and Sll1961

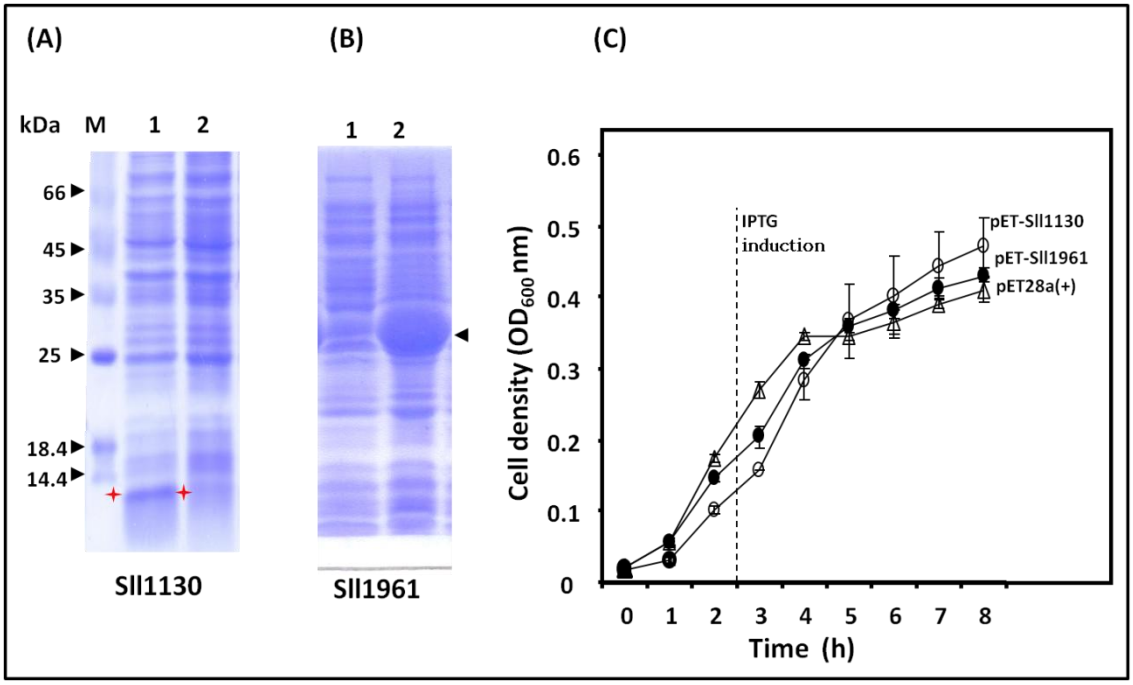


Figure 4.2. Over expression of Sll1130 and Sll1961 proteins in *E. coli* (A and B) and Growth profiles of *E. coli* expressing Sll1130, Sll1961 and *E. coli* harbouring only pET 28a(+) vector without any *Synechocystis* gene as insert (C). Total proteins of *E. coli* cells after 4 h of induction with 0.4 mM IPTG were resolved on SDS-PAGE. (A) *E. coli* expressing Sll1130; M, marker; 1, Total protein before addition of IPTG for Sll1130 induced expression; 2, Total protein after 4 hours of Sll1130 induction by IPTG; Sll1130 band is indicated with red stars. (B) *E. coli* expressing Sll1961; Total protein before addition of IPTG for Sll1961 induced expression; 2, Total protein after 4 hours of Sll1961 induction by IPTG; Sll1130, induced Sll1130 band was labeled with (red stars). (C), growth profiles of *E. coli* expressing Sll1130 (-○-); Sll1961 (-●-) and only vector (-Δ-). Vertical dotted line indicates the time at which IPTG was added to the culture.

Mutagenesis of *sll1130*

In an attempt to ascertain the function of the Sll1130 protein, *sll1130* was inactivated as shown in the schematic representation (Figure 4.3 A). In this mutant all copies of the *sll1130* gene were replaced by the disrupted copies of *sll1130* genes as confirmed by the fact that when genomic DNA of wild type cells was used as template with specific primers (*sll1130*-F and *sll1130*-R), a PCR product of 614 bp correspondingto the *sll1130* ORF and its upstream and downstream flanking regions was amplified (Figure 4.3B). In contrast, when the genomic DNA of Δ*sll1130* cells was used as template with the same set of primers, a 1814 bp DNA fragment corresponding to the wild type fragment (614 bp) plus the inserted kanamycin cassette (1200 bp) was amplified (Figure 4.3 B) confirming the insertion of kanamycin cassette into genome of Δ*sll1130* and complete segregation of the mutant copy of *sll1130*.

Δ*sll1130* mutant cells have not exhibited any phenotype at optimal conditions

Both wild type and Δ*sll1130* mutant cells of *Synechocystis* exhibited similar growth profiles at 34^{o}C despite of complete inactivation of *sll1130* gene, in the Δ*sll1130* mutant (Figure 4.4). Cell density in terms of OD at 730nm during the course of incubation of cells at optimal growth conditions were shown in Figure 4.4. As can be evidenced from Figure 4.4, there is no difference in the growth between wild type and mutant cells.

Figure 4.3: Mutation of *sll1130* and PCR based confirmation of mutation

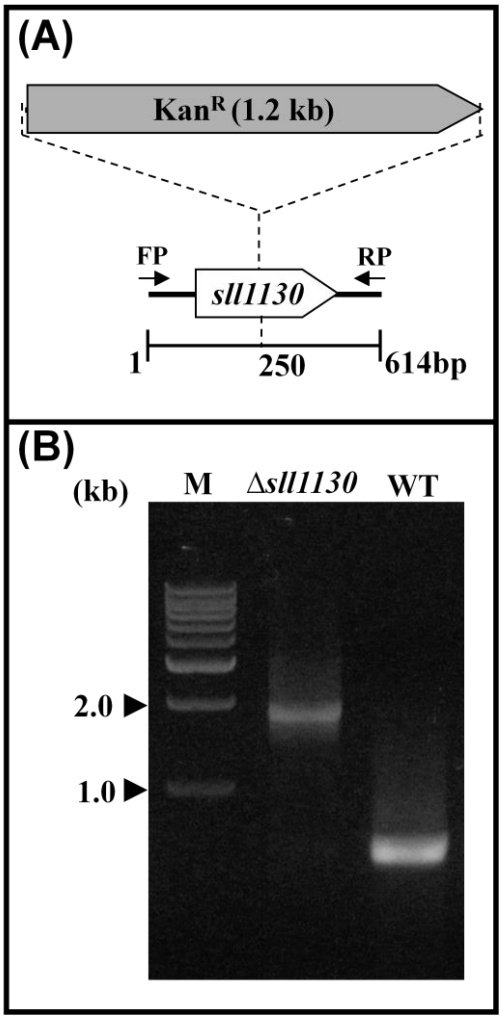


Figure 4.3. Strategy for disruption of the *sll1130* gene in the genome of *Synechocystis* sp. PCC 6803. The wild type copy of the *sll1130* gene was completely replaced by the mutated copy of the gene in Δ*sll1130* mutant cells. A 614 bp DNA fragment from the *sll1130* gene was insertionally inactivated with a kanamycin resistance cassette (1200bp). (A) Schematic representation of the genotype of the Δ*sll1130* mutant. The *sll1130* gene and the kanamycin-resistance gene (*KanR*) cassette are shown in the open and filled arrows, respectively. Thick arrows indicate *sll1130-F* (FP) and *sll1130-R* (RP), the two primers that were used for PCR amplification of the wild type copy of the *sll1130* gene and that of the *KanR* cassette. (B) Genomic PCR analysis with the primers indicated in (A). M, 1-kb DNA ladder; Δ*sll1130*, PCR product with Δ*sll1130* DNA as template; WT, PCR product with wild type genomic DNA as template.

Figure 4.4: Growth profiles of Δ*sll1130* and WT

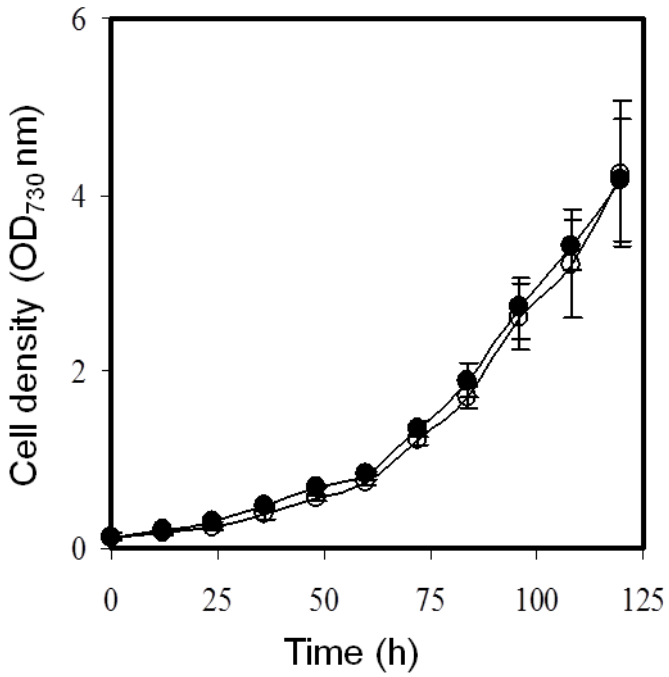


Figure 4.4. Growth profiles of wild type (-●-) and Δ*sll1130* mutant cells (-○-) of *Synechocystis* cells at optimal growth temperature (34°C). Similar results were obtained in three independent experiments, and the data are represented as means ± S.D.

Sll1130 is not functionally associated with PSII

Kashino *et al*., reported that Sll1130 is a novel protein associated with the PSII complex (Kashino *et al*., 2002). Therefore, it is expected that the targeted inactivation of Sll1130 may lead to an impairment of PSII structure and function. Hence, PSII-activity was measured in both wild type and Δ*sll1130* mutant cells. PSII activity was measured using an artificial electron accepter p-benzoquinone at 25°C using a Clarke-

type electrode. The PSII activity in the wild type and mutant cells were 474±37 µmol O_2 evolved/mg Chl/h and 447±39 µmol O_2 evolved/mg Chl/h respectively under 34^oC, 70 μmol of photons/m^2/s, suggests lack of *sll1130* has no influence on PSII activity (Figure 4.5 A).

The cells were subjected to heat shock at 42^oC for 24 h, as PSII is reported to be sensitive to heat treatment (Pueyo *et al*., 2002). Hence, comparison of heat sensitivity between wild type and mutant would suggest the role of Sll1130 in PSII structure and function. The activities of intact cells before and after heat treatment (24 h) were measured in BG-11 medium supplemented with 1mM benzoquinone and 1mM Potassium ferricyanide ($K_3Fe(CN)_6$). The PSII activities in heat treated wild type and mutant cells were 349±45 µmol O_2 evolved/mg Chl/h and 327±35 µmol O_2 evolved/mg Chl/h respectively (Figure 4.5 A), suggests, targeted inactivation of *sll1130* has no influence on PSII heat sensitivity (Figure 4.5 A). Certain auxiliary proteins of PSII were shown to play role in repair of photodamaged-PSII (Nixon *et al.*, 2010). Since, Sll1130 was also reported to be a PSII associated protein; it is assumed that it may be also involved in repair of photo-damaged PSII. Photosystem II was damaged by exposing wild type and mutant cells to high light (1200 μmol of photons/m^2/s) for 30 min, followed by recovery at 70 μmol of photons/m^2/s light. High light treatment for a period of 30 min resulted in loss of 80% of original activity in both wild type and mutant cells (Figure 4.5 B). PSII activity was recovered to 80% of the original by 3 h incubation of high light treated cells at optimal growth conditions. (Figure 4.5 B). PSII repair and recovery was complete after 4 h, resumed its total initial activity. Rate of PSII damage with high light and recovery were almost same in wild type and Δ*sll1130* (Figure 4.5 B). Despite of complete mutation of

sll1130 in *Synechocystis* neither PSII activity nor photo damaged PSII repair was affected. This clearly shows that Sll1130 has no role in PSII structure, function and assembly.

Figure 4.5: Role of Sll1130 in function and repair of PSII

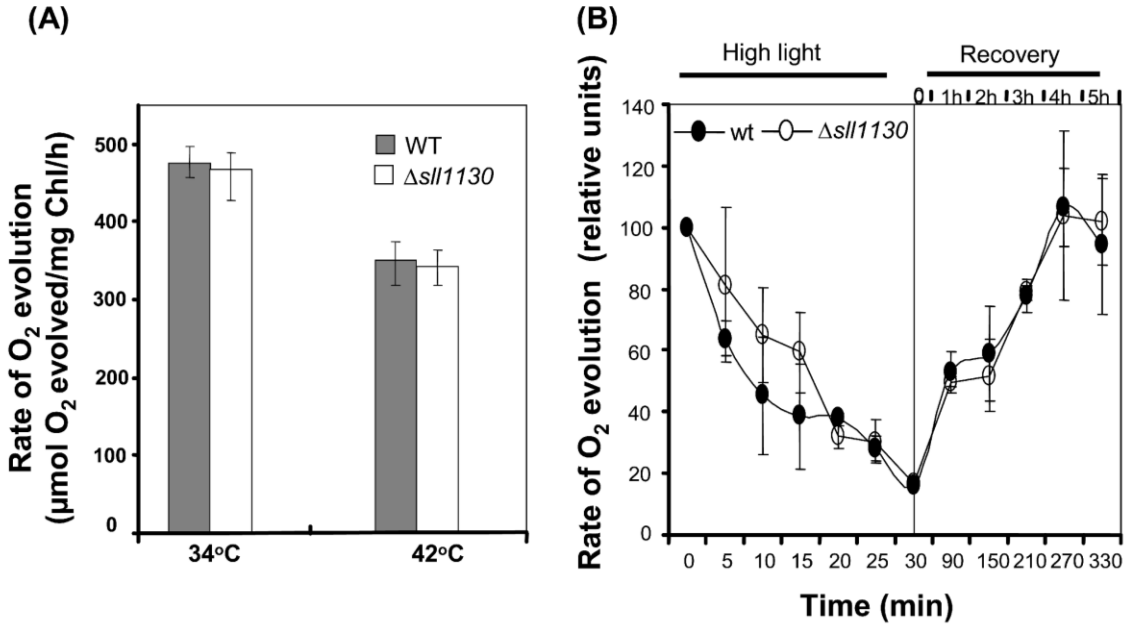


Figure 4.5. Measurement of PSII activity and repair in wild type and mutant. (A) PSII activity in terms of oxygen evolution was measured at 34°C and 42°C, ■ WT and □ Δ*sll1130*. (B) PSII repair after photo damage was represented -●- WT and -○- Δ*sll1130*.

Sll1130 is a constitutively expressing soluble protein

As the PSII functional studies indicated that Sll1130 may not have any role in PSII structure and function, we wanted to verify and localize the protein by western blotting analysis. Hydropathy profile of Sll1130 by web based tool indicated that Sll1130 is a soluble protein (http://bp.nuap.nagoya-u.ac.jp/sosui/cgi-bin/adv_sosui.cgi). Western blotting analysis was carried out using anti-Sll1130

polyclonal antibody raised against purified His-Sll1130 in Rabbit. Figure 4.6, shows the immuno-detection of Sll1130 in the soluble and insoluble membrane fractions of wild type and Δ*sll1130* mutant cells. The anti-Sll1130 antibody detected Sll1130 in the soluble fraction of wild type cells as well as the purified His-tagged Sll1130 protein. Sll1130 protein was not detected in the insoluble/membrane fraction of the cells (Figure 4.6). Thus, it is clear that the Sll1130 is a soluble protein and was absent in the insoluble membrane fraction (Figure 4.6). This also confirms that the Sll1130 is not associated with photosystem II, as was reported earlier by Kashino *et al.*, (2002). Sll1130 protein was detected neither in the soluble nor insoluble fractions of Δ*sll1130*, suggesting absence of Sll1130 functional protein in the mutant and its complete inactivation (Figure 4.6). Detection of Sll1130 in cultures grown in optimal conditions confirms that it is a constitutively expressing protein.

Figure 4.6: Localization of Sll1130 by western blotting experiment

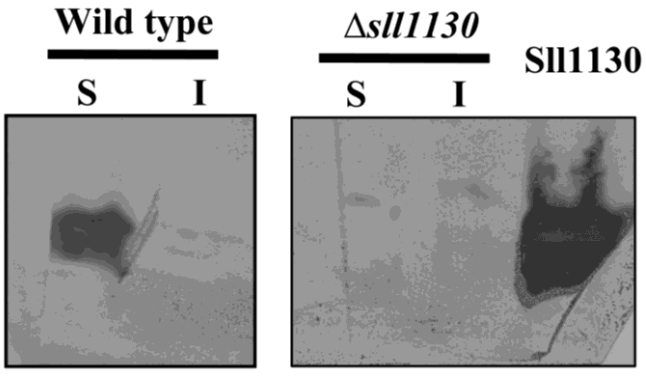


Figure 4.6. Immuno detection of Sll1130 protein in the soluble (S) and insoluble membrane (I) fractions extracted from wild type and Δ*sll1130 Synechocystis* cells; Sll1130, purified His-Sll1130 was detected by anti-Sll1130 antibody.

Δ*sll1130* exhibited increased thermo tolerance

As described previously, both wild-type and Δ*sll1130* mutant cells of *Synechocystis* exhibited similar growth profiles at 34◦C (Figure 4.4). But when exponentially growing cells of wild-type and Δ*sll1130* mutant strains were heat-treated at 50◦C for 30 min and then allowed to recover at 34◦C, it was observed that *sll1130* mutant cells were relatively more tolerant to heat than wild-type cells (Figure 4.7). We measured cell survival after a high temperature treatment by taking small aliquots of cell suspension and counting the viable and nonviable cells as described in materials and methods. Before heat shock, both wild type and Δ*sll1130* cells appeared as bright red fluorescent cells, indicating that the cells were equally viable at the optimal growth temperature, 34^{o}C (Figure 4.7 A and B). Following heat shock at 50^{o}C for 30 min, cells were allowed to recover at 34^{o}C for 48 h and viability was again measured. While only 30% of the total wild type cells were viable following this heat shock (cells emitting red fluorescence over green fluorescence) (Figure 4.7 C and E), 70% of the Δ*sll1130* cells remained viable (Figure 4.7 D and E). This result suggests that disruption of *sll1130* leads to increased thermo-tolerance to *Synechocystis* cells.

Figure 4.7: Viability staining of the WT and Δ*sll1130* cells after heat shock & recovery

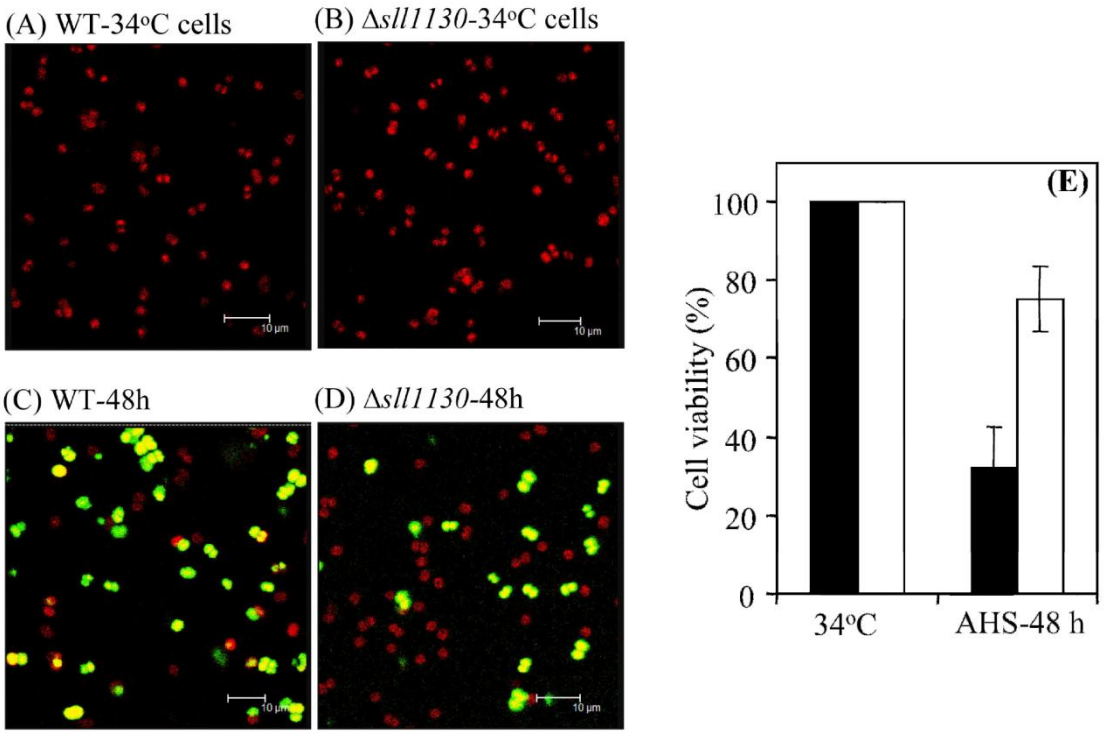


Figure 4.7. Effect of heat shock on the survival of *Synechocystis* wild type and Δ*sll1130* mutant cells. Wild type and Δ*sll1130* cells were grown at 34°C for 16 h (A, WT-34°C cells; B, Δ*sll1130*-34°C cells). Wild type and Δ*sll1130* cells were grown at 34°C for 16 h, subjected to heat shock at 50°C for 30 min and then shifted back to 34°C for 48 h (C, WT-48h; D, Δ*sll1130*-48h). The white bar indicates 10 µm. The cells were stained with SYTOX Green dye and the percentage of viable cells (red) and dead cells (green) in the wild type and Δ*sll1130* mutant before and after heat shock were determined (E) by counting under 40X magnification using a confocal microscope. In (E), 34°C, wild type or Δ*sll1130* cells grown at 34°C for 16 h; AHS-48 h, 34°C-cells were incubated at 50°C for 30 min and then shifted back to 34°C for 48h. ■, Wild type cells; □, Δ*sll1130* mutant cells. Data represents the mean and standard deviation of 3 independent experiments.

Identification of target genes of Sll1130

It is likely that Sll1130, which we predicted as a DNA binding protein with conserved helices, may serve as a transcriptional regulator in *Synechocystis.* We therefore sought to identify the target genes of Sll1130. We performed DNA microarray analysis of wild type and Δ*sll1130* mutant cells grown under optimal growth conditions (70 μmol photons/m^2/S, 34^{o}C). Table 4.1 shows the list of genes whose expression levels were significantly affected by the disruption of *sll1130*. Genes whose mean induction (Δ*sll1130* / wild type) was greater than 2.0 were considered to be upregulated, whereas those with an induction less than 0.5 are listed as down-regulated (Table 4.1). It is notable that following inactivation of *sll1130* the expression of certain heat-responsive genes, such as *hspA*, *htpG*, *isiA* and *isiB*, were upregulated along with several other genes that encode hypothetical proteins (Table 4.1). Four genes, *ssl2245*, *sll1130*, *cysS* (cysteinyl t-RNA synthetase) and *codA* (cytosine deaminase) were downregulated due to mutation in *sll1130.* Both the *ssl2245* gene, which is located upstream to the *sll1130* ORF, and *codA* gene, which is located downstream of the *sll1130* ORF (www.kazusa.or.jp/cyanobase/Synechocystis/genes/sll1130), were strongly down-regulated due to insertional inactivation of *sll1130.* These observations suggest that there is a common regulatory mechanism which controls the expression of these two genes. Observed down-regulation of *cysS*, *codA* and *sll1130* genes is consistent with the DNA microarray expression data generated upon heat shock in *Synechocystis* (Suzuki *et al.*, 2005).

Table 4.1: Effect of *sll1130* mutation on genome-wide expression of genes.

Genes upregulated due to mutation in *sll1130*			
ORF No.	**Gene**	**Product**	**Δ*sll1130*/wt**
slr1788		hypothetical protein	8.3±1.0
slr1789		hypothetical protein	6.3±0.6
sll1009	*frpC*	iron-regulated protein	4.4±0.8
slr1920		hypothetical protein	4.2±0.3
sll0910		hypothetical protein	3.6±0.4
slr1594	*rre5*	two-component response regulator (PatA subfamily)	3.6±0.4
sll1722		hypothetical protein	3.5±0.4
sll0249		hypothetical protein	3.4±0.4
slr0006		hypothetical protein	3.2±0.2
ssr2194		hypothetical protein	3.0±0.2
slr1931	*pilA8*	type 4 pilin-like protein	2.9±0.0
slr1963		water-soluble carotenoid protein	2.9±0.0
slr0294		hypothetical protein	2.8±0.1
sll1514	*hspA, hsp17*	16.6 kDa small heat shock protein, molecular chaperone	2.8±0.3
slr0476		hypothetical protein	2.7±0.0
sll0247	*isiA*	iron-stress induced chlorophyll-binding protein	2.7±0.1
slr0869		hypothetical protein	2.7±1.2
slr0581		hypothetical protein	2.7±0.1
sll0447		hypothetical protein	2.6±0.4
ssl2501		hypothetical protein	2.5±0.1
sll1660		hypothetical protein	2.5±0.2
slr1930	*pilA7*	type 4 pilin-like protein	2.2±0.2
sll0248	*isiB*	Flavodoxin	2.2±0.9
sll1780	*ISY203b*	transposase gene of IS4 family insertion sequence	2.2±0.0

slr0870		hypothetical protein	2.1±0.0
sll0430	*htpG*	HtpG, heat shock protein 90, molecular chaperone	2.1±0.0
sll1549		salt-enhanced periplasmic protein	2.1±0.2

Genes down-regulated due to mutation in *sll1130*

ORF No.	Gene	Product	Δ*sll1130*/wt
ssl2245		hypothetical protein	0.3±0.15
sll1130		hypothetical protein	0.2±0.04
slr0958	*cysS*	cysteinyl-tRNA synthetase	0.1±0.04
slr1237	*codA*	cytosine deaminase	0.1±0.01

Table 4.1. Wild type and Δ*sll1130* cells were grown at 34^{o}C for 16 h. Each value indicates the ratio of levels of mRNA from Δ*sll1130* mutant cells to wild type cells. The values shown are the means ± range of two independent experiments. The numbering of the ORF corresponds to that of Kaneko *et al*. (Kaneko *et al*., 1996). The complete list of gene expression in wild-type and Δ*sll1130* mutant can be accessed at http://www.genome.jp/kegg-bin/gethtext?htext=Exp_DB&hier=1. The list of genes that are upregulated in the Δ*sll1130* mutant ≥2folds and the list of genes that are down regulated in the Δ*sll1130* were presented (Krishna *et al.*, 2013).

qRT-PCR analysis confirms DNA microarray expression changes

Differences in the expression of five genes, namely *slr1788, isiA, frpC, hspA* and *ssl2245,* observed in the *sll1130* mutant compared to the wild type were further confirmed by qRT-PCR analysis (Figure 4.8 A and Table 4.1). These results suggest that Sll1130 acts as a transcriptional negative regulator of *hspA, htpG, isiA, isiB, frpC* and several other hypothetical genes, and therefore these genes may be upregulated by de-repression in Δ*sll1130* cells. Genes coding for HspA and HtpG are known to play important roles in heat acclimation and the survival of *Synechocystis* cells at high temperature (Tanaka and Nakamoto 1999; Nakamoto *et al.,* 2000; Fang and Barnum 2003; Lee *et al.,* 1998; Nakamoto *et al.,* 2001). IsiA and IsiB have been reported to be upregulated under heat and iron-limiting conditions and play important roles in protection from heat stress conditions (Kojima *et al.,* 2006). Heat-induced expression of *hspA* gene was studied both in wild type cells and Δ*sll1130* mutant by qRT-PCR to ascertain reasons for the heat tolerant phenotype (Figure 4.8 B). The temporal expression of *hspA* gene for a period of 180 minutes indicated an initial increase in expression by 10 min and a gradual decline by 180 min in both the wild type and Δ*sll1130* mutant cells. However, in the mutant cells *hspA* expression was significantly greater than in the wild type cells all through the heat treatment period (Figure 4.8 B). Relatively higher levels of *hspA* mRNA, in Δ*sll1130* mutant cells over wild type cells explain the heat tolerant phenotype of Δ*sll1130* mutant. Thus, collectively our results together with the available literature suggest that the enhanced thermotolerance of the Δ*sll1130* strain could be due to elevated transcript and corresponding protein levels of these important heat stress inducible genes in *Synechocystis* (Table 4.1, Figures 4.7 and 4.8). We further analysed the *frpC* levels

during heat stress in wild type and Δ*sll1130* mutant by qRT-PCR (Figure 4.8 C). The temporal expression of *frpC* gene for a period of 180 minutes indicated an initial increase in expression by 10 min and maintained at steady state in Δ*sll1130* mutant cells, a 2 fold upregulation was observed in wild type after 30 min of heat stress and mRNA levels further decreased and reached the study state levels in wild type, but in Δ*sll1130 frpC* levels were high and they increased further upon heat stress and remained high throughout the period tested (Figure 4.8 C).

A common *cis*-regulatory element was detected upstream of genes upregulated in Δ*sll1130*

If the Sll1130 is a negative regulator of transcription of all the genes that were upregulated by its inactivation, then a common *cis*-regulatory element that is recognized by the Sll1130 protein would be expected upstream of these genes. As predicted, we found a common inverted-repeat GGCGATCGCC sequence upstream of most of the genes whose expression was upregulated due to mutation in *sll1130* (Figure 4.9 A and B). This common inverted-repeat may be the target binding site of Sll1130 and probably acts as a *cis*-regulatory element. We predicted this inverted-repeat computationally using MEME software as described in Materials and Methods. In some of these upregulated genes the identified DNA element GGCGATCGCC is located at two places in the upstream region, similar to that of LacI repressor site (Reznikoff *et al.*, 1974) (Figure 4.9 C).

Figure 4.8: Gene expression profiles in *Synechocystis* by qRT-PCR

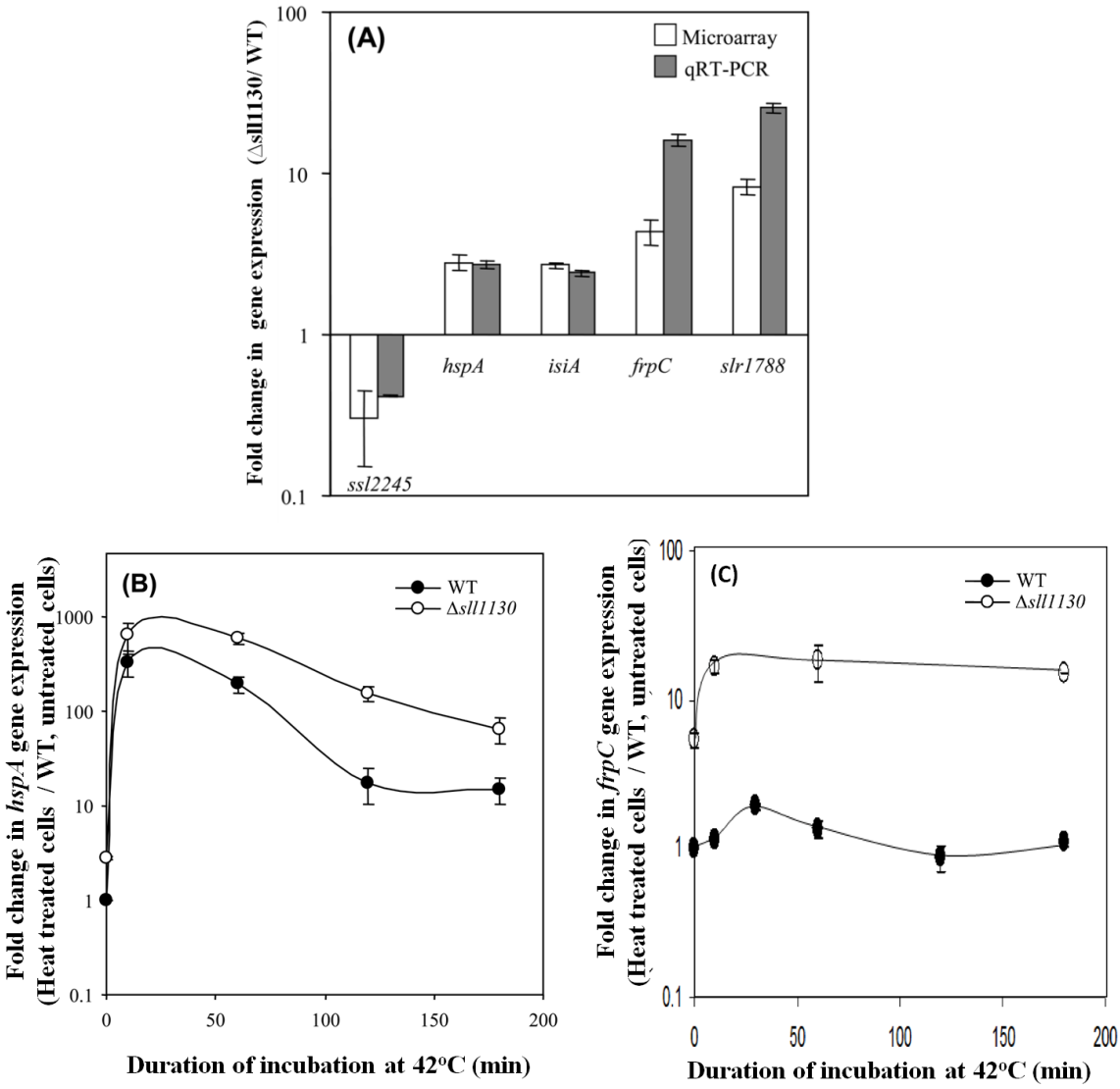


Figure 4.8. (A) Confirmation of DNA microarray results by qRT-PCR. *slr1788, frpC, isiA,* and *hspA* showed a higher level of expression in the Δ*sll1130* mutant than wild-type as measured by microarray (white bars) and by qRT-PCR (black bars). Similar results were obtained in two independent experiments; data are represented as means ± SD. (B) *hspA*-mRNA levels during the course of heat treatment in the wild type and Δ*sll1130* cells. (C) *frpC*-mRNA levels during the course of heat treatment in the wild type and Δ*sll1130* cells. Both wild type (-●-) and Δ*sll1130* (-○-) cells were grown at 34°C for 16 h and then subjected to heat stress (42°C) for 10, 60, 120 and 180 min. The fold change of *hspA* and *frpC*-mRNA levels in heat-treated wild type and Δ*sll1130* cells was expressed relative to its levels in WT-34°C cells (means ± SD of three independent experiments).

Figure 4.9: *In silico* analysis of the upstreams of genes upregulated in Δ*sll1130*

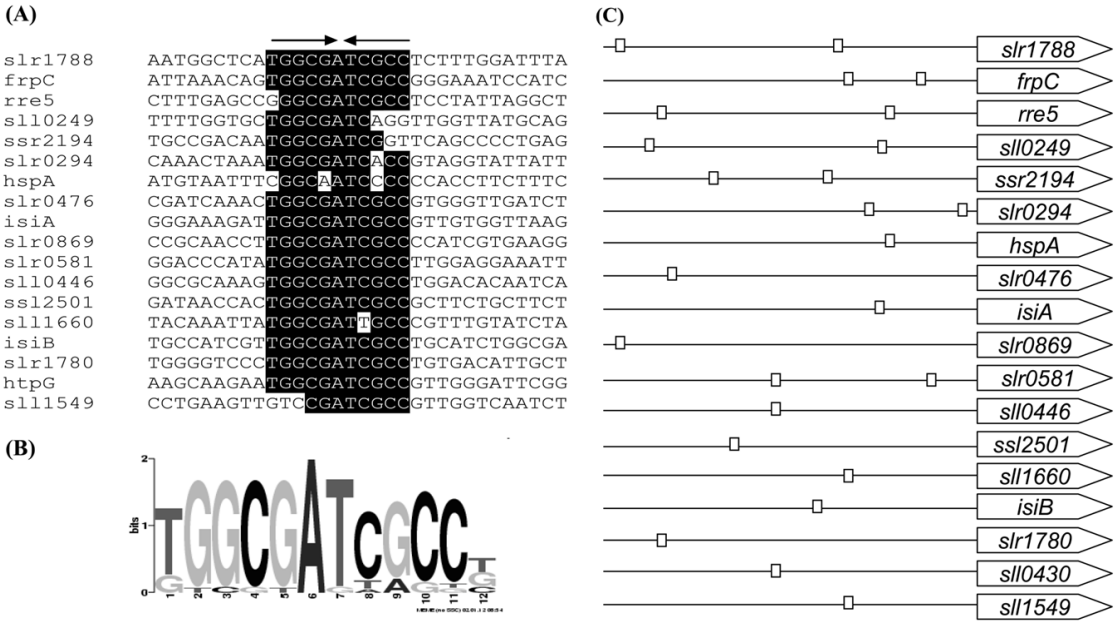


Figure 4.9. Presence of a conserved inverted-repeat in the upstream region of genes that are upregulated due to mutation in *sll1130*. Upstream DNA region of genes that are induced due to mutation in *sll1130* were submitted to MEME (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) for prediction of the conserved DNA element. (A) Shows the alignment of an inverted-repeat in the upstream region of genes upregulated due to mutation in *sll1130*. (B) Logo representation of the conserved regulatory element. (C) Location of the inverted-repeat in the upstream region of genes upregulated due to mutation in *sll1130*. Open arrows in the schematic representation are coding regions of upregulated genes. Straight line, length of upstream DNA region; rectangles, inverted-repeat.

Specific binding of Sll1130 to the *cis*-regulatory element located upstream of *slr1788*

The binding of the *cis*-regulatory element located upstream of *slr1788* with Sll1130 protein was studied using gel mobility shift assay. As shown in Figure 4.10, purified His-Sll1130 protein retarded the electrophoretic mobility of the upstream DNA fragment of *slr1788*, and this shift was observed to be concentration dependent (Figure 4.10). We confirmed that the observed retardation was due to binding of Sll1130 protein with the upstream DNA fragment of *slr1788*, by adding 100-fold excess of double-stranded oligonucleotides (40 bp) covering the computationally predicted *cis*-regulatory element, which completely eliminated the retarded fragment (Figure 4.10). Taken together with the presence of the same *cis*-regulatory element upstream of the other genes upregulated by inactivation of *sll1130*, these data clearly implicate Sll1130 as a negative regulator of these genes.

Expression of the *sll1130* gene is down-regulated upon upward shift in temperature

DNA microarray analysis has revealed that expression of the *Sll1130* gene is down-regulated upon shift of wild type *Synechocystis* cells to high temperature (Suzuki *et al.*, 2005). We observed down regulation in the *sll1130* expression in *Synechocystis* cells grown at 34^{o}C and then shifted to 42^{o}C. Significant levels of *sll1130* mRNA were detected before exposure to 42^{o}C, indicating that the *sll1130* gene is constitutively expressed (Figure 4.11 A). Within 10 min of an upward shift in temperature to 42^{o}C the level of *sll1130* transcript decreased by about 80% and remained at a similar level up to seven hours after the shift (Figure 4.11A and C). Western blot analysis with anti-Sll1130 antibody also indicated that the levels of

Sll1130 protein decreased significantly when *Synechocystis* cells grown at 34^{o}C for 16 h were shifted to 42^{o}C (Figure 4.11 B and C). The anti-Sll1130 antibody detected Sll1130 at a molecular mass of approximately 15 kDa (Figure 4.11 B). The level of Sll1130 protein decreased to 60% within 30 min of incubation at 42^{o}C and reached a minimum of 50% of the original level within 180 min, suggesting that *Synechocystis* cells down regulate the expression of the *sll1130* gene and reduced its protein upon heat treatment (Figure 4.11 C). This is in agreement with our previous DNA microarray study in which there was down-regulation of *sll1130* upon a 20 min heat shock at 42^{o}C (Suzuki *et al.*, 2005). These observations indicated that *Synechocystis* cells down-regulated the expression of the *sll1130* gene and corresponding Sll1130 protein upon upward shift in temperature.

Figure 4.10: Electrophoretic mobility shift assay with purified Sll1130 and *slr1788* upstream

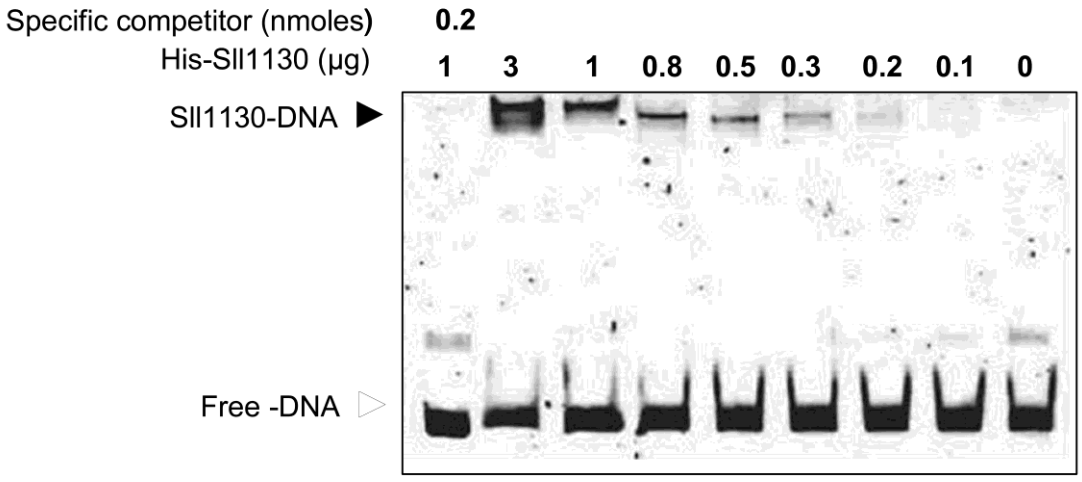


Figure 4.10. Binding of Sll1130 to the region upstream of *slr1788*. Gel mobility shift assay was performed with different concentrations of His-Sll1130 protein (0 to 3 µg) and the region upstream of *slr1788*. A 283 bp Cy3-labelled region of DNA upstream of *slr1788* starting from -6 to -289 bp with respect to the translation start site and a 41 bp specific competitor, 5' TGC CCC AAG CAA GAA TGG CGA TCG CCG TTG GGA TTC GGA GC 3' (0.2 nmoles) that also includes the predicted GGCGATCGCC *cis*-regulatory binding element were used for binding assays. Open arrow indicates free Cy3 labeled DNA and closed arrow indicates His-Sll1130 DNA complex.

Figure 4.11: Expression profile of *sll1130* upon heat stress in *Synechocystis*

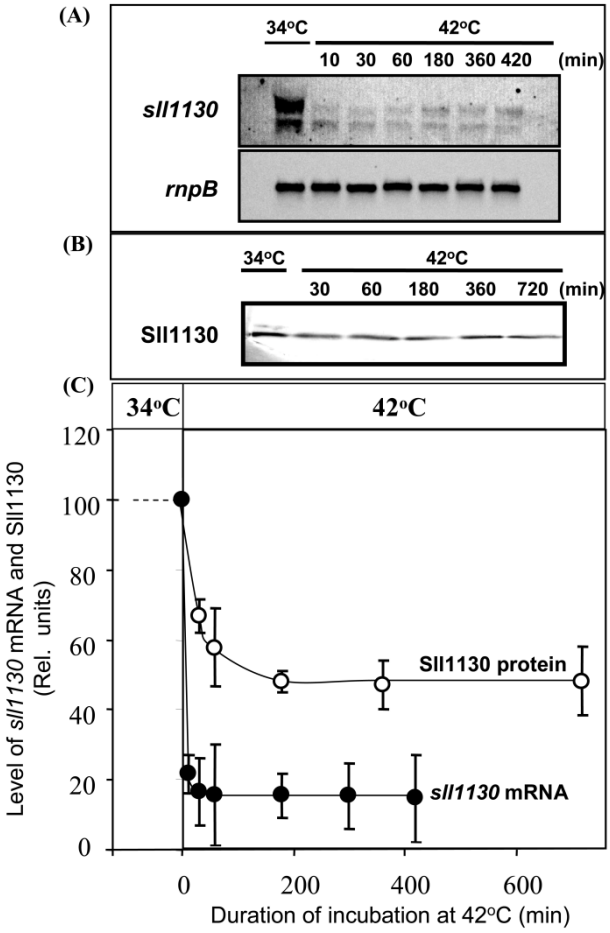


Figure 4.11. Northern and Western blot analyses of changes in the expression of the *sll1130* gene upon increase in growth temperature in wild type cells of *Synechocystis*. (A) Northern blotting of *sll1130* mRNA. Total RNA was extracted from wild type cells that were grown at 34°C for 16 h or were grown at 34°C for 16 h and then incubated at 42°C for 10, 30, 60, 180, 360 and 420 min. Aliquots (20 µg) of the extracted RNA were electrophoresed on 1.2% agarose gels that contained 1.4 M formaldehyde. (B) Western blotting of Sll1130. Soluble proteins were extracted from wild type cells that were grown at 34°C for 16 h or were grown at 34°C for 16 h and then incubated at 42°C for 30, 60, 180, 360 and 720 min. Samples equivalent to 30 µg proteins were loaded in each well of a polyacrylamide gel (15%) that contained 0.1% SDS. Sll1130 antibody produced in rabbit was used to detect Sll1130 on the blot. (C) Quantitative expression of *sll1130* mRNA (●) and Sll1130 (o) upon high temperature treatment of wild type cells (mean ± SD of two independent experiments).

Functional characterization of Ssl2245

Sll1130 exhibits similarity with MazF (PemK), though Ssl2245 does not exhibit any similarity with already reported antitoxins, it is located just upstream to Sll1130 with 4nt overlapping, a characteristic feature of most toxin-antitoxin system encoding operons. Upstream genes of Sll1130 orthologues encode AbrB like proteins or MazE antitoxin proteins. it is probably because this gene pair is evolved from TA systems. Present work has been designed for functional characterization of Ssl2245 by mutagenesis, phenotype characterization and transcriptome analysis of Δ*ssl2245*.

ssl2245 and *sll1130* are in the same operon

Open reading frames *ssl2245* and *sll1130* has a four nucleotide overlap in their coding region. To validate whether they are in same operon *i.e* under the control of the same promoter reverse transcribed mRNA with gene specific primers covering *ssl2245* and *sll1130*. Total RNA was isolated from the wild type cells and cDNA was prepared using random hexamers. PCR amplification using the primers of *ssl2245* forward primer (*ssl2245*-exp-FP) and *sll1130* reverse primer (*sll1130*-exp-RP) has given a product of ~ 600 bp which corresponds to the size of entire operon. This result confirms that *ssl2245* and *sll1130* are co-transcribed and are in a dicistronic operon (Figure 4.12 A). Additionally we demonstrated that both Ssl2245 and Sll1130 proteins are translated from the same transcript by cloning the entire operon into pET28a(+) expression vector. Ssl2245 over expressed along with Sll1130 protein (*ssl2245-sll1130* operon) in pET28a(+) in such a way that the *ssl2245* and *sll1130* ORFs are in downstream to T7 promoter of the vector. Upon induction with IPTG resulted synthesis of 6x-His tagged Ssl2245 protein and Sll1130 (Figure 4.12 B). Next we cloned these two structural genes, in which *ssl2245* had been insertionally

inactivated by Ω-spectinomycin cassette into pET28a(+). Upon induction with IPTG we did not observe the expression of both Ssl2245 and Sll1130 (Figure 4.12 C). This result further confirms that *ssl2245* and *sll1130* are in the same operon under the control of same promoter and are transcribed as a single transcript. Our results also suggest that the insertional inactivation of first gene, *ssl2245* of the operon blocks the synthesis of transcription of the second gene, *sll1130.*

Figure 4.12: *ssl2245* and *sll1130* are in same operon

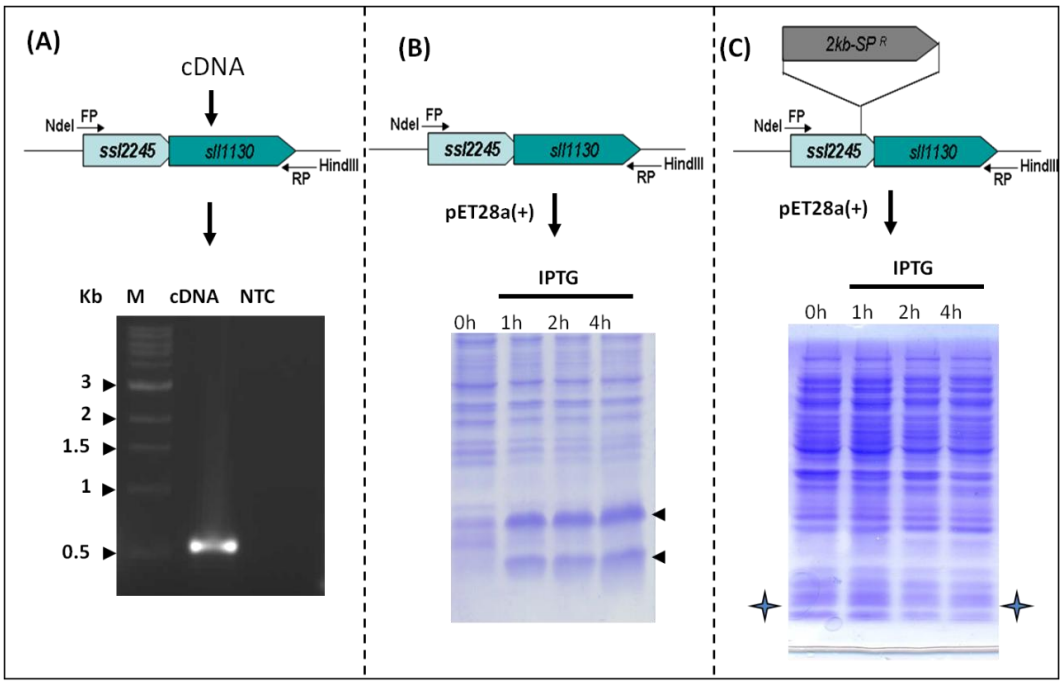


Figure 4.12. *ssl2245* and *sll1130* are in same operon. (A) Total RNA was isolated and cDNA was prepared, using cDNA as the template DNA and with specific primers (*ssl2245* exp FP and *sll1130* exp RP) PCR reaction was performed and PCR product was separated on 1.2% agarose gel, M, marker; cDNA, cDNA as template; NTC, no template DNA. (B) *ssl2245-sll1130* are PCR amplified and cloned into pET28a(+) and transformed to *E. coli* BL21 strain. Transformed *E. coli* was induced with 0.4mM IPTG (4 h) and total cellular proteins were resolved on SDS-PAGE (15%). Two protein bands appeared with IPTG induction. (C) *ssl2245-sll1130* with Ω-spectinomycin casette insertion in the *ssl2245* ORF was PCR amplified and cloned into pET28a(+) and transformed to *E. coli* BL21 strain. Transformed *E. coli* was induced with 0.4mM IPTG (4 h) and total cellular proteins were resolved on SDS-PAGE (12%). No difference in protein profile was observed with IPTG induction.

Ssl2245 is an unknown protein

The open reading frame, *ssl2245* encodes an unknown protein in *Synechocystis* (Kaneko *et al.*, 1996). Figure 4.13 shows the alignment of top 8 BlastP hits of Ssl2245 protein. The protein is well conserved with other unknown proteins in its N-terminal region but the C-terminal half is not well conserved. Secondary structure prediction of the Ssl2245 protein by GORIV suggests that the protein contains two helices in its C-terminal half (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_ automat.pl?page =npsa_gor4.html) (Figure 4.13). The protein has no putative domains predicted (http://smart.embl-heidelberg.de/smart/show_motifs.pl) except that a 17 a.a length low complexity region between 59 a.a - 75 a.a. The orthologues of Sll1130 have a conserved gene located in their upstreams with overlapping base pair region. Some of these genes located upstream to Sll1130 orthologues were annotated as MazE, an antitoxin, and some others as AbrB like protein. In case of *Nitrosomonas europea*, *sll1130* ortholog, upstream overlapping gene was not present (Figure 4.14).

Figure 4.13: Sequence wise alignment of Ssl2245 BLAST-P results

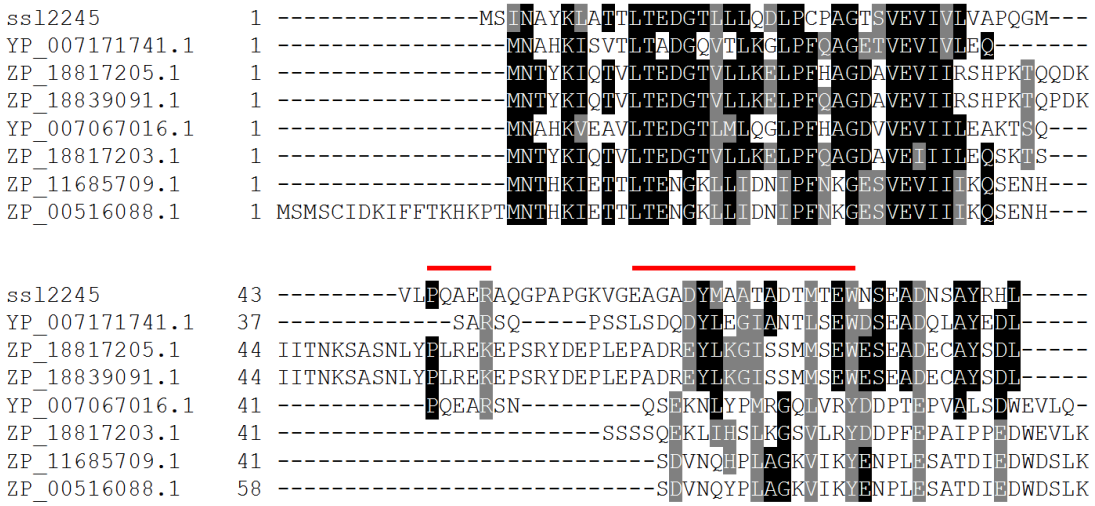


Figure 4.13. Multiple sequence alignment of the Ssl2245 *Synechocystis* sp. PCC 6803 and its homologues retrieved from the NCBI database. The sequences were aligned using the ClustalW algorithm. Identical amino acids are shown in black and conserved amino acid substitutions are in grey. Helix regions were predicted using the GORIV secondary structure prediction tool (depicted with a red line in figure). Ssl2245, *Synechocystis* sp. PCC 6803; YP_007171741.1, hypothetical protein Dacsa_1720 (*Dactylococcopsis salina* PCC 8305); ZP_18817205.1, conserved hypothetical protein (*Microcystis aeruginosa* PCC 9432); ZP_18839091.1 conserved hypothetical protein (*Microcystis aeruginosa* PCC 9808); YP_007067016.1, hypothetical protein Cal7507_3793 (*Calothrix* sp. PCC 7507); ZP_11685709.1, hypothetical protein CWATWH0003_2523 (*Crocosphaera watsonii* WH 0003); ZP_18817203.1, conserved hypothetical protein (*Microcystis aeruginosa* PCC 9432); ZP_00516088.1, hypothetical protein CwatDRAFT_3921 (*Crocosphaera watsonii* WH 8501).

Figure 4.14: Schematic representation of Sll1130 orthologues and their upstream genes

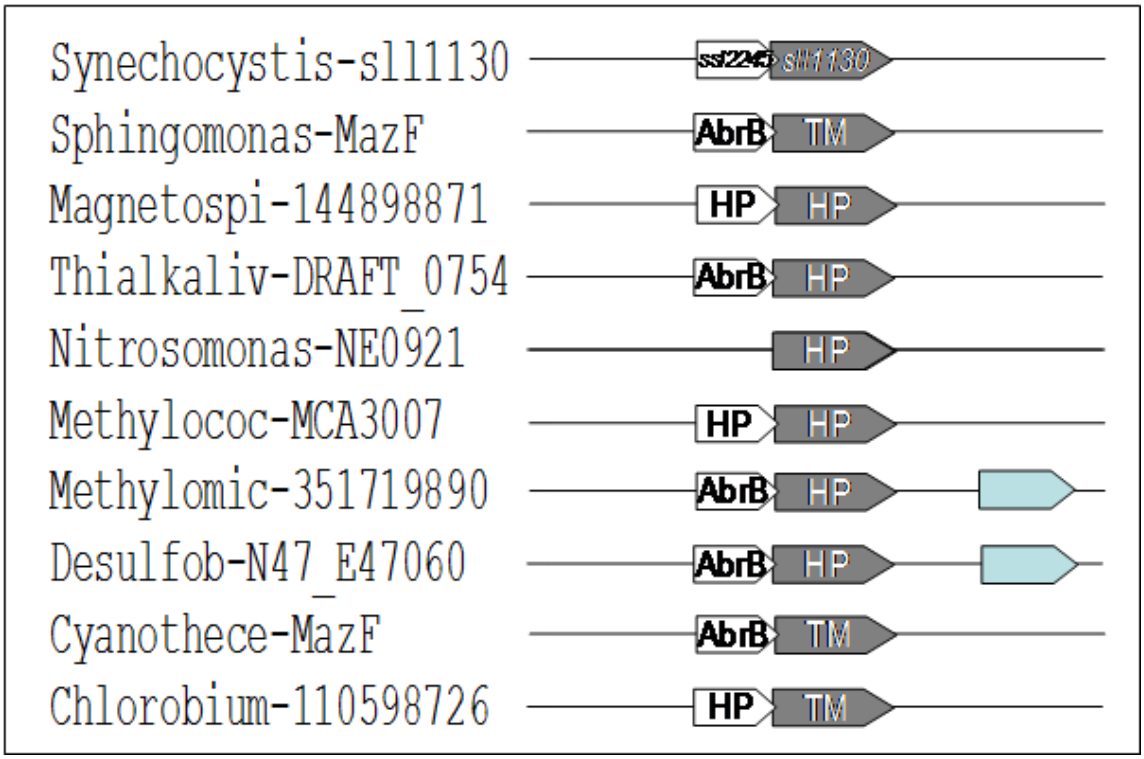


Figure 4.14: Schematic representation of ORFs of Sll1130 orthologues and their upstreams. Orthologues of Sll1130 were taken and their upstream gene products (protein) were down loaded. Except *Nitrosomonas europea* all other bacteria has an upstream gene with overlapping region. Block arrows indicate ORF; AbrB, AbrB like protein coding ORF; HP, hypothetical protein coding ORF. Ssl2245 and Sll1130 ORFs are represented with their respective numbers.

Mutagenesis of *ssl2245* and generation of double mutant Δ*ssl2245*:: Δ*sll1130*

In an attempt to ascertain the function of the Ssl2245 protein, *ssl2245* was inactivated as shown in the schematic representation (Figure 4.15 A). In this mutant all copies of the *ssl2245* gene were not replaced by the disrupted copies of *ssl2245* even after several generations of antibiotic selection pressure (spectinomycin 20µg/ml) as confirmed by the fact that when genomic DNA of wild type cells was used as template with specific primers (*ssl2245*-F and s*sl2245*-R), a PCR product of 260 bp corresponding to the *ssl2245* ORF was amplified (Figure 4.15 B). In contrast, when the genomic DNA of Δ*ssl2245* cells was used as template with the same set of primers, a 2260 bp DNA fragment corresponding to the wild type fragment (260 bp) plus the inserted Ω-spectinomycin gene cassette (2000 bp) was amplified along with 260 bp product (Figure 4.15 B lane labeled Δ*2245*). A TA plasmid (pTZ57R/T, Cat.No.K1213, Fermentas) with the disrupted *ssl2245* gene was transformed to fully segregated mutant of *sll1130* and allowed to grow on spectinomycin gradient BG-11 plate with kanamycin 25µg/ml and selected colonies were analyzed for the disruption of both the genes by PCR analysis, we labeled this double mutant as Δ*DM*. The so generated double mutant was complete in case of *sll1130* mutation but very few copies of *ssl2245* were left in the genome (Figure 4.15 B lane labeled Δ*DM*).

Figure 4.15: Schematic presentation of mutagenesis and segregation analysis of *ssl2245*

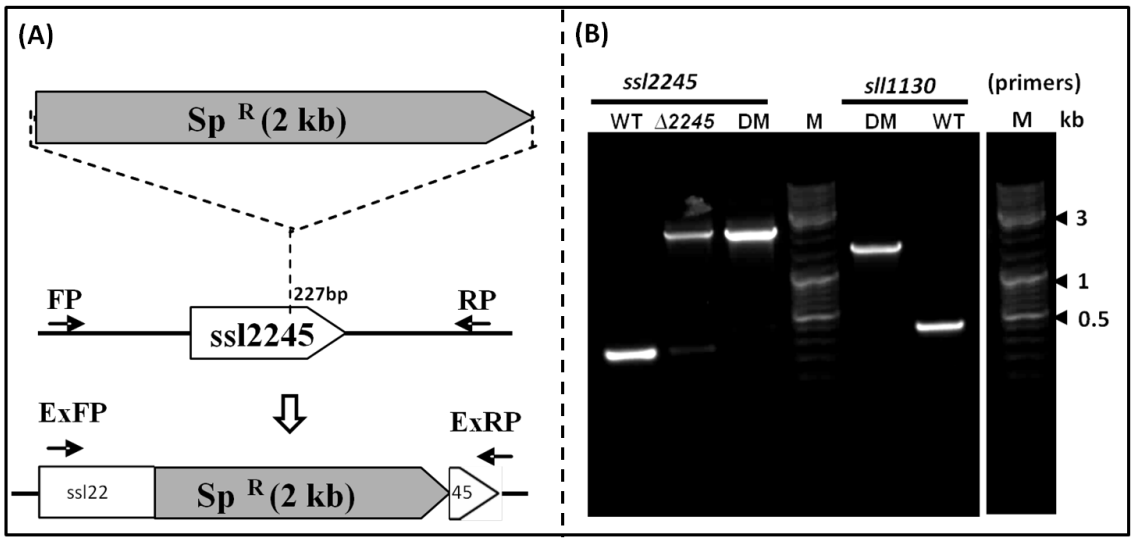


Figure 4.15. Strategy for disruption of the *ssl2245* gene in the genome of *Synechocystis* sp. PCC 6803. The wild type copy of the *ssl2245* gene was not completely replaced by the mutated copy of the gene in Δ*ssl2245* mutant cells. A 614 bp DNA fragment from the *ssl2245* gene was insertionally inactivated with a Ω-spectinomycin resistance cassette (2000bp). (A) Schematic representation of the genotype of the Δ*ssl2245* mutant. The *ssl2245* gene and the Ω-spectinomycin - resistance gene (SP^R) cassette are shown in the open and filled arrows, respectively. Thick arrows indicate *ssl2245-F* (FP) and *ssl2245-R* (RP), the two primers that were used for PCR amplification of the wild type copy of the *ssl2245* gene and that of the SP^R cassette. (B) Genomic PCR analysis with the primers designed for expression of Ssl2245 and Sll1130. M, 1-kb DNA ladder; Δ*2245*, PCR product with Δ*ssl2245* DNA as template; WT, PCR product with wild type genomic DNA as template; DM, double mutant (Δ*ssl2245*:: Δ*sll1130*) genomic DNA as template; the set of primers used for amplification whether ssl2245 primers or sll1130 primers was mentioned on top of the gel and as ssl2245 (primers of Ssl2245 expression), *sll1130* (primers of Sll1130 expression).

Inactivation of neither *ssl2245* nor entire operon (Δ*ssl2245::sll1130*) did not affect growth but expression of an extra copy of *ssl2245-sll1130* operon exhibited the slow growth.

Synechocystis expressing an extra copy of *ssl2245-sll1130* operon exhibited slow growth. To identify the functional role of Ssl2245 individually or in combination with Sll1130, growth of wild type, Δ*ssl2245*, Δ*sll1130*, Δ*DM* and WT+ was analyzed at optimal growth conditions. All strains exhibited similar growth profiles and no change in phenotypes were observed (Figure 4.16). This results shows just like TA systems, inactivation of these genes, there by absence of the corresponding proteins has no effect on cell growth under optimal conditions (Tsilibaris *et al.*, 2007). Inactivation of *sll1130* caused down regulation of the first gene, *ssl2245* and mutation of the first gene *ssl2245* resulted in inhibition of *sll1130* expression (Table 4.1 and 4.2 respectively). In order to see the effect of Ssl2245 and Sll1130 when they are expressed at higher concentration than the normal, an extra copy of the *ssl2245-sll1130* with its native promoter was expressed in *Synechocystis* using a cyanobacterial expression vector pVZ321. DNA region covering *ssl2245-sll1130* operon with its native promoter (upstream 300 bp) was PCR amplified and cloned into pVZ321 vector as mentioned in materials and methods. pVZ321 vector with *ssl2245-sll1130* was transformed to *Synechocystis* by triparental mating explained by Zinchenko *et al*., (1999) and screened under chloromphenicol antibiotic selection pressure. So formed strain was named as WT+ and it has exhibited slow growth phenotype when compared to the all other mutant strains (Δ*ssl2245*, Δ*sl1130* and ΔDM) and wild type. Growth of the cultures at 34^{o}C and 70 μmol photons/m^{2}/s was measured as the cell density in terms of OD at 730nm (Figure 4.16).

Figure 4.16: Growth profiles of different mutant strains

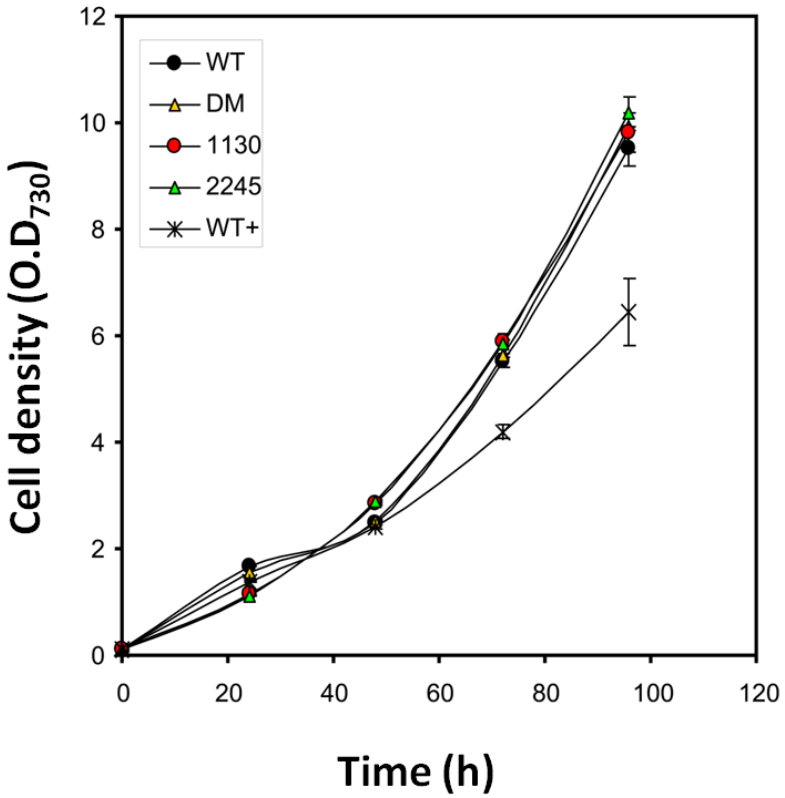


Figure 4.16. Growth profiles of wild type, Δ*sll1130*, Δ*ssl2245*, DM and WT+. WT (-●-), wild type; DM (-▲-), double mutant (**Δ*ssl2245*::*sl1130*);** 1130 (-●-)**,** Δ*sll1130*; 2245 (-▲-), Δ*ssl2245* and WT+ (-X-), wild type with an extra copy of the *ssl2245* and *sll1130* in pVZ-321.

Identification of genes with altered expression due to inactivation of *ssl2245*

Wild type and Δ*ssl2245* cells were grown photoautotrophically at 34◦C in BG-11 medium buffered with 20 mM Hepes/NaOH (pH 7.5) under continuous illumination at 70 μmol of photons/m^2/s. Total RNA was isolated and differentially expressed genes in Δ*ssl2245* when compared to wild type were identified using microarray (custom made 15KAgilent DNA microarray chip). Microarray analysis of the Δ*ssl2245* mutant in comparison with wild type *Synechocystis* cells, under optimal conditions has revealed upregulation of several genes. In Δ*ssl2245*, 63 genes have

shown ≥ 3 fold expression in comparison to WT. Out of the top 30 upregulated genes, 22 are plasmid associated genes. We observed that the mutation in Δ*ssl2245* also resulted in repression of the downstream gene, *sll1130* (Table 4.2). As a result, the genes repressed by Sll1130 under optimal conditions, such as *slr1788*, *slr1789* and *sll1009* were also upregulated in Δ*ssl2245* (Table 4.2). Repression of *sll1130* expression is may be due to unavailability of promoter to *sll1130* due to insertional inactivation of *ssl2245*. Upregulated genes in Δ*ssl2245* with Δ*ssl2245*/WT ≥ 3.00 were presented as upregulated genes and Δ*ssl2245*/WT ≤ 0.3 are represented as down regulated genes (Table 4.2).

Table 4.2: Effect of *ssl2245* mutation on genome-wide expression of genes.

Genes upregulated due to mutation in *ssl2245*			
GENE	**function**	**Δ*ssl2245*/WT ± SD**	**Plasmid**
slr0179	hypothetical protein	23.51 ± 8.55	
sll7064	unknown protein	20.43 ± 2.13	pSYSA
sll7090	unknown protein	20.01 ± 0.96	pSYSA
slr7023	hypothetical protein	18.78 ± 1.63	pSYSA
sll7086	unknown protein	17.63 ± 1.22	pSYSA
sll7089	unknown protein	17.24 ± 0.71	pSYSA
slr1788	unknown protein	17.12 ± 5.33	
sll7087	unknown protein	16.74 ± 0.36	pSYSA
slr1764	similar to tellurium resistance protein TerE	15.95 ± 6.03	
sll7065	unknown protein	15.79 ± 2.12	pSYSA
ssr6027	unknown protein	15.78 ± 3.03	pSYSX
ssr6024	unknown protein	14.98 ± 3.13	pSYSX
slr1789	unknown protein	12.60 ± 1.76	
ssr6026	unknown protein	10.76 ± 1.84	pSYSX

slr0105	hypothetical protein	10.75 ± 4.23	
slr7091	hypothetical protein	10.33 ± 0.05	pSYSA
sll7067	unknown protein	9.55 ± 0.86	pSYSA
sll7066	unknown protein	8.40 ± 0.40	pSYSA
slr6025	probable antirestriction protein	8.18 ± 3.80	pSYSA
sll7085	unknown protein	8.17 ± 0.43	pSYSA
sll7063	unknown protein	8.16 ± 1.97	pSYSA
sll1009	unknown protein	8.10 ± 1.10	
sll7062	unknown protein	7.66 ± 0.49	pSYSA
sll7075	unknown protein	6.88 ± 1.34	pSYSA
sll5097	hypothetical protein	6.45 ± 0.96	pSYSM
ssr6030	unknown protein	6.14 ± 0.49	pSYSX
slr0869	hypothetical protein	6.01 ± 0.37	
sll0327	unknown protein	5.67 ± 1.06	
slr6028	unknown protein	5.40 ± 1.84	pSYSX
ssr6089	unknown protein	5.21 ± 0.36	pSYSX
sll7069	hypothetical protein	5.16 ± 1.80	pSYSA
slr0106	unknown protein	4.84 ± 0.92	
sll1660	hypothetical protein	4.72 ± 1.35	
slr1920	unknown protein	4.59 ± 0.91	
sll0444	unknown protein	4.23 ± 0.44	
slr7097	hypothetical protein	4.19 ± 0.55	
slr2017	type 4 pilin-like protein, essential for motility	4.12 ± 0.65	
sll7070	unknown protein	4.11 ± 0.47	pSYSA
slr2016	type 4 pilin-like protein, essential for motility	3.95 ± 0.74	
slr7092	hypothetical protein	3.90 ± 0.13	pSYSA
slr6029	hypothetical protein	3.87 ± 0.82	pSYSX
sll1239	unknown protein	3.70 ± 1.56	
sll0445	unknown protein	3.70 ± 0.40	

slr7025	hypothetical protein	3.63 ± 2.15	pSYSA
ssr6046	hypothetical protein	3.52 ± 1.00	pSYSX
sll0041	accepting chemotaxis protein	3.44 ± 1.25	
slr1441	hypothetical protein	3.36 ± 2.72	
slr5111	unknown protein	3.36 ± 0.71	pSYSM
slr7024	hypothetical protein	3.35 ± 0.85	pSYSA
slr0820	probable glycosyltransferase	3.35 ± 0.39	
sll1240	unknown protein	3.34 ± 1.12	
sll1366	putative SNF2 helicase	3.30 ± 2.45	
sll0443	unknown protein	3.29 ± 0.24	
sll0909	unknown protein	3.26 ± 1.38	
slr0871	unknown protein	3.25 ± 0.97	
ssl8024	unknown protein	3.23 ± 0.60	pSYSG
sll1599	manganese transport system ATP-binding protein MntA	3.19 ± 2.29	
slr6012	unknown protein	3.15 ± 1.38	pSYSX
slr6107	hypothetical protein	3.08 ± 2.18	pSYSX
slr7096	hypothetical protein	3.07 ± 1.25	pSYSA
slr0870	hypothetical protein	3.07 ± 0.95	
sll1578	phycocyanin alpha subunit	3.00 ± 0.38	

Genes downregulated due to mutation in *ssl2245*

GENE	**function**	**Δ*ssl2245*/WT ± SD**	**Plasmid**
slr1928	type 4 pilin-like protein	0.29 ± 0.04	
slr1704	hypothetical protein	0.28 ± 0.04	
ssl6023	unknown protein	0.28 ± 0.09	pSYSX
sll1723	probable glycosyltransferase	0.27 ± 0.01	
sll0857	unknown protein	0.25 ± 0.01	
sll0858	hypothetical protein	0.24 ± 0.02	
sll1722	hypothetical protein	0.20 ± 0.02	
sll0856	RNA polymerase ECF-type (group 3) sigma-E factor	0.18 ± 0.01	

slr0909	unknown protein	0.15 ± 0.05	
ssl2245	unknown protein	0.14 ± 0.05	
sll1396	unknown protein	0.14 ± 0.05	
sll6010	unknown protein	0.06 ± 0.00	pSYSX
sll1130	unknown protein	0.02 ± 0.01	

Table 4.2. Wild type and Δ*ssl2245* cells were grown at 34^{o}C for 16 h. Each value indicates the ratio of levels of mRNA from Δs*s2245* mutant cells to wild type cells. The values shown are the means ± sd of two independent experiments. The numbering of the ORF corresponds to that of Kaneko *et al*., (Kaneko *et al*., 2003 and Kaneko *et al*., 1996).

Gene expression profile of *frpC* in Δ*ssl2245,* Δ*sll1130* and WT+

To get clues on function of Ssl2245, one gene that is upregulated in Δ*ssl2245* and Δ*sll1130* was selected (*frpC*) and its gene expression profile was tested in Δ*ssl2245,* Δ*sll1130* and WT+. Total RNA was isolated from the wild type, Δ*ssl2245,* Δ*sll1130* and WT+ cells and cDNA was prepared using random hexamers. qRT PCR was performed with specific primers of *frpC*. Gene expression profile of *frpC* in Δ*ssl2245,* Δ*sll1130* and WT+ is expressed in relative to WT (Figure 4.17). Fold change in gene expression of *frpC* in Δ*ssl2245* and Δ*sll1130* was high as observed in microarrays (Table 4.1 and 4.2), where as in WT+ reduced levels of *frpC* was observed. This suggests that that expression of an extra copy of *ssl2245* and *sll1130* will alter the expression profile of *frpC* in opposite to that of the mutation.

Figure 4.17: Gene expression profile of *frpC* in Δ*ssl2245*, Δ*sl1130* and WT+

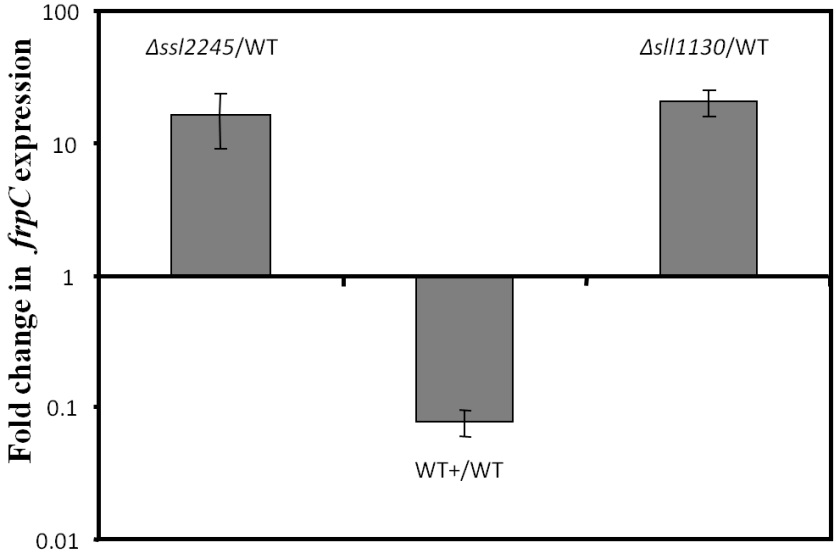


Figure 4.17. Gene expression profiles of *frpC* in Δ*ssl2245*, Δ*sll1130* and WT+. Δ*sll1130*/WT, fold change in gene expression of *frpC* in Δ*sll1130* relative to WT; Δ*ssl2245*/WT fold change in gene expression of *frpC* in Δ*ssl2245* relative to WT; WT+/WT, fold change in gene expression of *frpC* in WT+ relative to WT. Similar results were obtained in two independent experiments; data are represented as means ± SD.

Role of Ssl2245-Sll1130 interaction

Sato *et al.*, 2007, reported that Ssl2245 and Sll1130 interact with each other. Presence of overlapping region between *ssl2245* and *sll1130* ORFs shows a possible way of post transcriptional regulation. Present work has been designed for biochemical characterization of Sll1130, to validate the interaction of Ssl2245 with Sll1130 and to see the effect of these proteins on *E. coli* when they are over expressed in different combinations.

Over expression and purification of Ssl2245 and Sll1130

To study the biochemical properties of purified proteins, Ssl2245 and Sll1130 were over expressed using pET28a(+) in *E. coli* BL-21 (DE3) pLysS as mentioned in

materials and methods. Protein expression was induced with 0.4 mM IPTG and the expressed protein was purified using Ni-NTA column by His-affinity chromatography. Ssl2245 protein elutes exhibited ~3-4 kDa larger size (16 kDa) than expected (12 kDa) on SDS-PAGE (15%) (Figure 4.18 A). MALDI TOF-TOF analysis of purified protein bands resolved on SDS-PAGE confirmed the purified protein as His-Ssl2245. (Figure 4.19). Whereas, purified His-Sll1130 protein resolved at the expected size of 14.7 kDa on 15% SDS-PAGE gel (Figure 4.18 B).

Figure 4.18: Over expression and purification of Ssl2245 and Sll1130

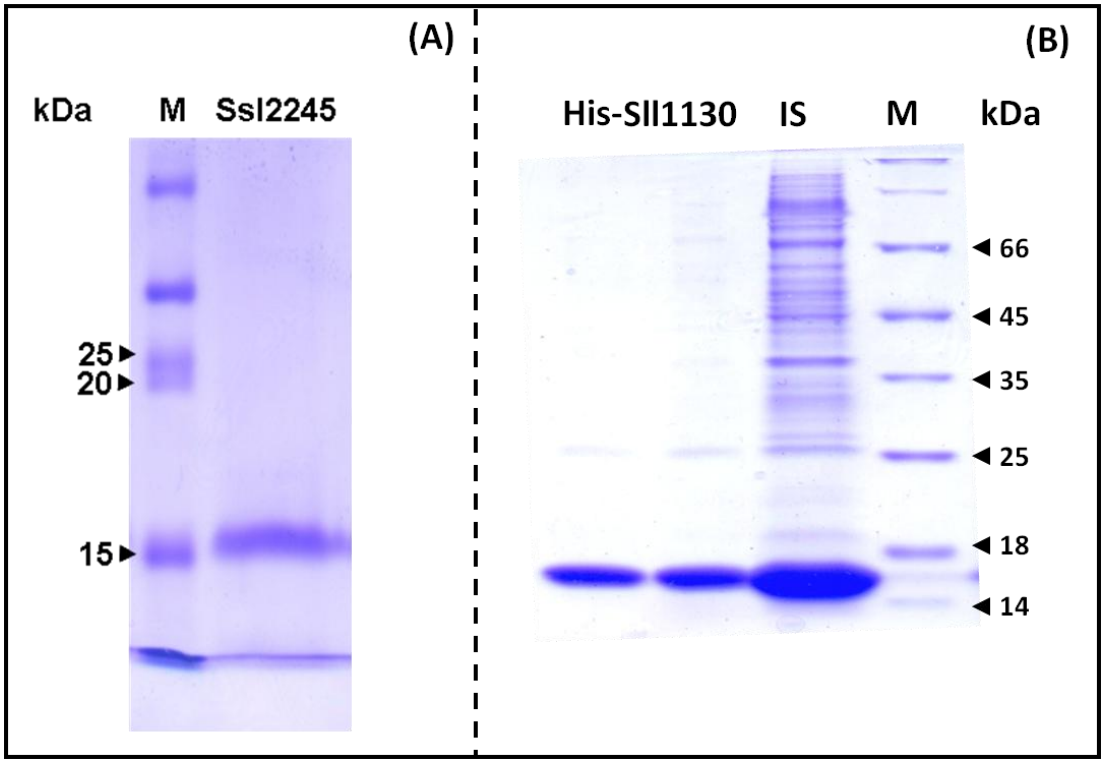
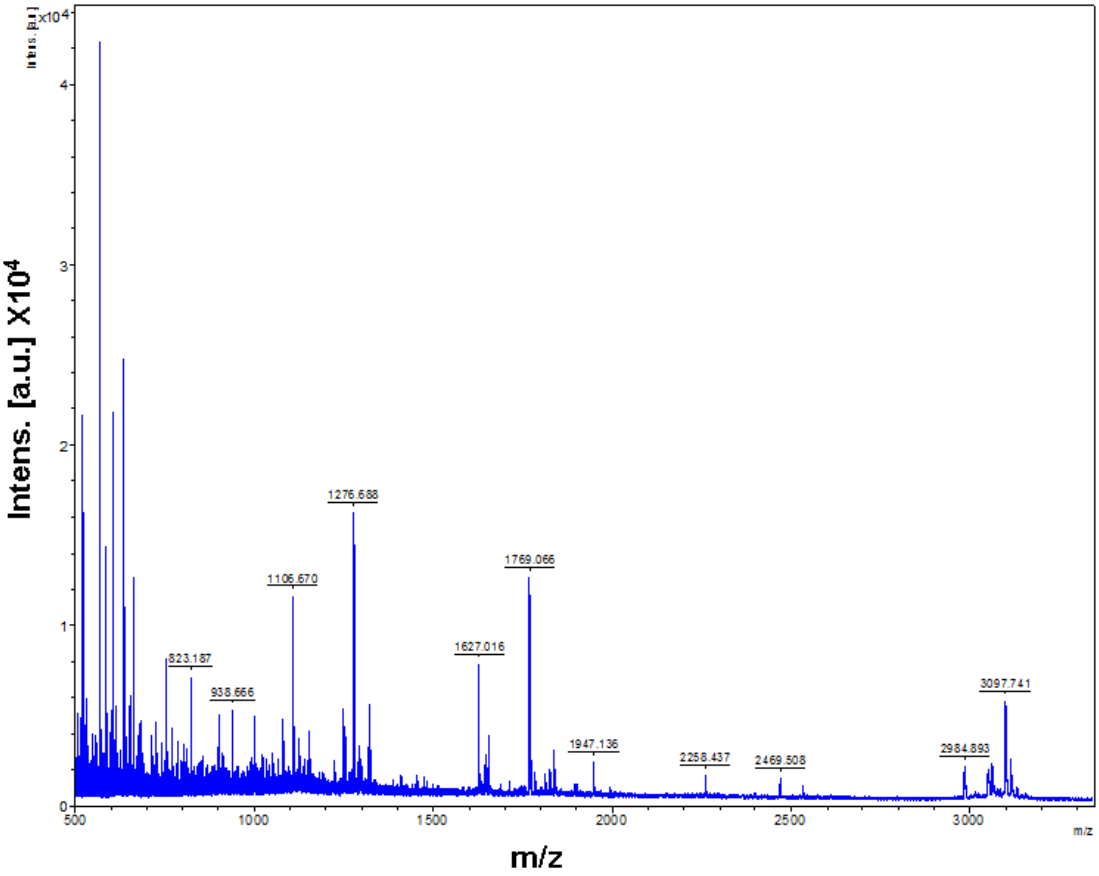


Figure 4.18. Over expression and purification of Ssl2245 and Sll1130. Sll1130 (C-terminal His tag) and Ssl2245 (N-terminal His tag) are over expressed in *E. coli* and purified using Ni-NTA column. (A), Purified Ssl2245 was resolved on SDS-PAGE (15%), M, marker; Ssl2245, purified His-Ssl2245. (B) Purified Sll1130 was resolved on SDS-PAGE (15%), M, marker; IS, Induced (IPTG) soluble fraction; His-Sll1130, purified His-Sll1130.

Figure 4.19: MALDI confirmation of the purified Ssl2245 protein



Matched peptides shown in **bold red**.

>Ssl2245

MSINAYKLATTLTEDGTLLLQDLPCPAGTSVEVIVLVAPQGMVLPQAER**AQGPAPGK**
VGEAGADYMAATADTMTEWNSEADNSAYRHL

Figure 4.19. MALDI-TOF/TOF validation of the purified Ssl2245 protein. Purified protein was given for trypsin digestion followed by MALDI-TOF/TOF analysis. Identified peptides based m/z values were matching with the Ssl2245 sequence.

Sll1130 protein is a tetramer

We determined the oligomeric state of Sll1130 by fractionating the purified native Sll1130 protein along with known appropriate markers, such as BSA (66 kDa) and carbonic anhydrase (29 kDa) on a gel exclusion column. Protein quantification by densitometric analysis of elutes indicated that BSA and carbonic anhydrase eluted with maximum intensity in fraction numbers 15 and 19 respectively (Figure 4.20 A). We observed that the native Sll1130 protein eluted with maximum intensity in fraction number 15 along with BSA. The estimated molecular mass of native Sll1130 was four times higher than that calculated based on its amino acid sequence, suggesting Sll1130 in its native form exists as a tetramer. In addition, on native-PAGE, Sll1130 protein appeared as a single protein band above the 43 kDa marker protein (Ovalbumin) (Figure 4.20 B). However, the pure His-Sll1130 protein resolved as a single protein band of 14.6 kDa on SDS-PAGE, implying that the molecular weight of its monomeric form is 12.9 kDa after subtracting 1.7 kDa accounted for by the His-tag (Figure 4.20 C). Interestingly, under semi-denaturing conditions(in the absence of β-mercaptoethanol), in addition to the protein band that was expected at 14.6 kDa, another protein band at around 30 kDa was observed. This suggests the involvement of disulphide bridge between two monomers in dimer formation, facilitated by a cysteine residue that can promote the formation of a disulfide bridge. Incubation of purified Sll1130 with SDS resulted in dissociation of tetramers into monomers and dimers. But it was observed that the monomer constituted a major fraction suggesting that the tetrameric form consists of a dimer associated with two monomers by non-covalent interactions. The monomer and

Dimer protein bands observed on semi-denaturing gel were excised and confirmed them as Sll1130 by MALDI-TOF TOF (Figure 4.21).

Figure 4.20: Oligomeric state of Sll1130

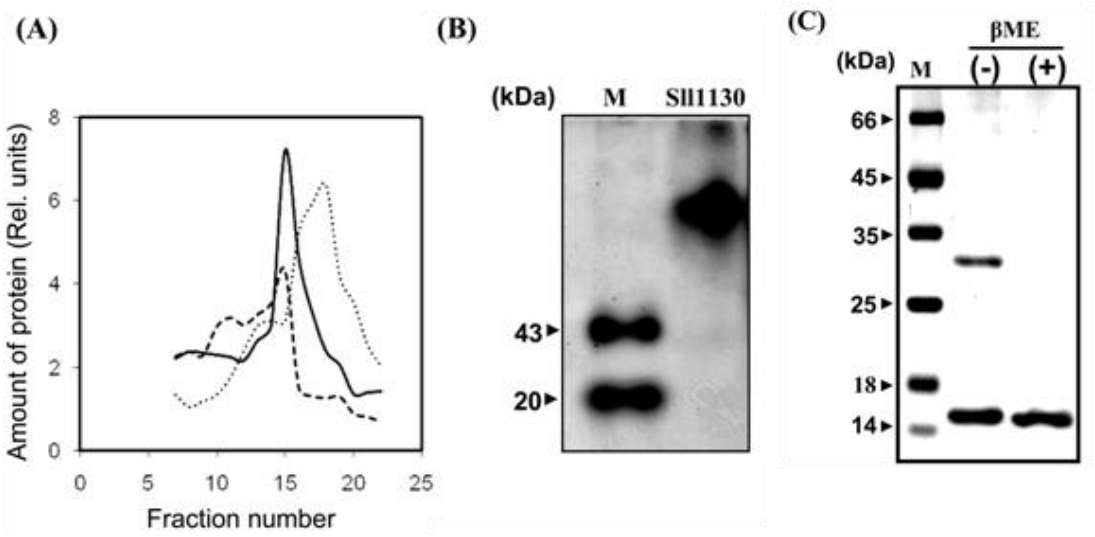
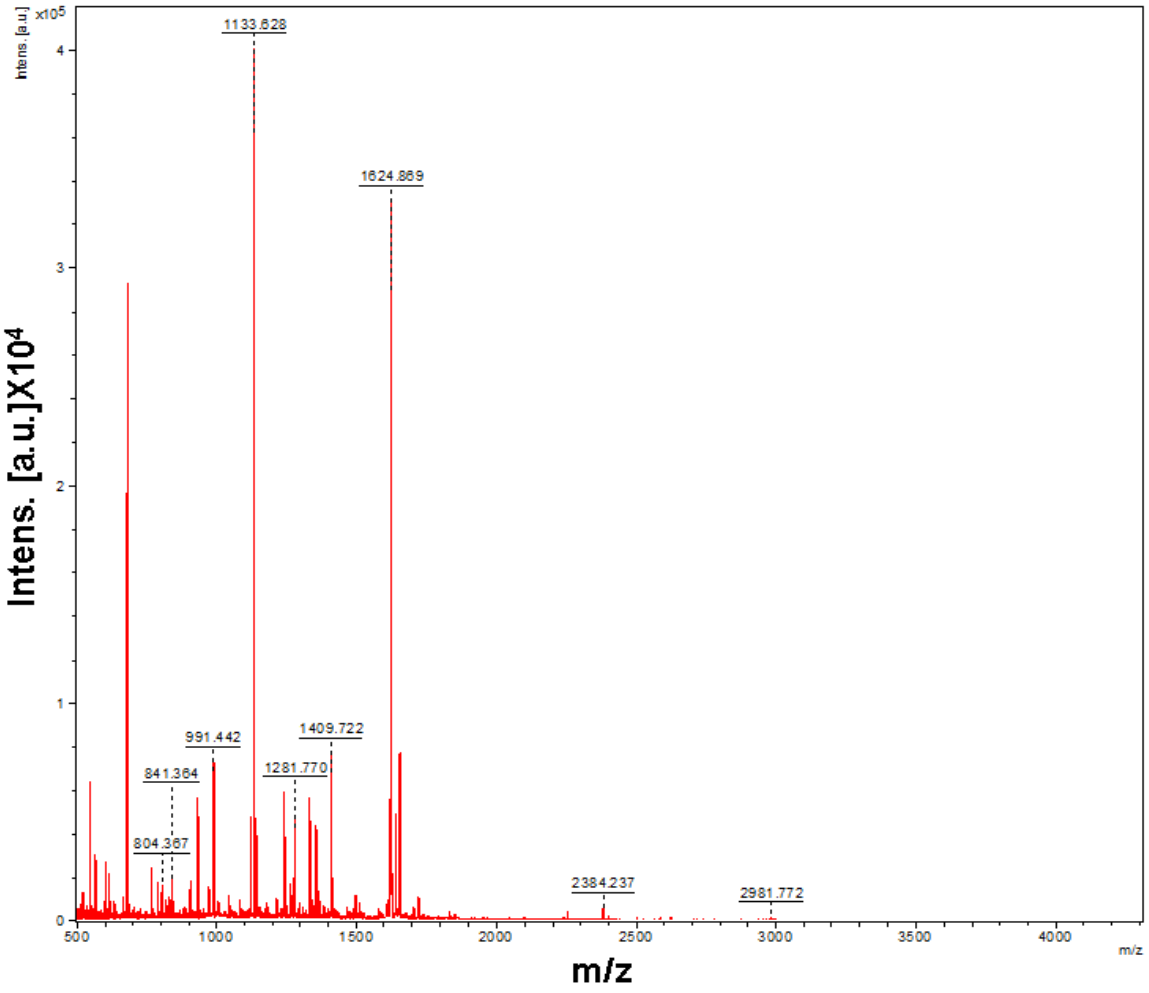


Figure 4.20. Oligomeric state of over expressed and purified His-Sll1130 protein (A) Gel exclusion chromatographic analysis of His-Sll1130 protein on Sephacryl S-100 column. His-Sll1130 protein mixed with gel filtration markers was loaded on to the column. Carbonic anhydrase, (dotted line); BSA (dashed line); Sll1130, (solid line). (B) Native non-denaturing PAGE separation of purified His-Sll1130. M, native protein marker contained ovalbumin (43 kDa) and trypsin soyabean inhibitor (20.1 kDa); Sll1130, purified His-Sll1130 protein. (C) SDS-PAGE (15%) separation of purified His-Sll1130 protein. βME (+), in the presence of β-mercaptoethanol; βME (-), in the absence of β-mercaptoethanol; M, protein marker 14-66 kDa

Figure 4.21: MALDI confirmation of Sll1130 dimer



Matched peptides shown in **bold red**.

>Sll1130

MNTIYEQFDVVIVPVPFTDRQSDIR**RPALILSDAPAFNNR**IGHSVMAMITSAK**NAPW**
PLDTPIEDTRSAGLFTPSVVRMKLFTLEHKYILDCVGSLSKQDRLMVKSAFPHVFKL
G

Figure 4.21. MALDI-TOF/TOF validation of the Sll1130 dimer. Purified His-Sll1130 protein was resolved on SDS-PAGE (15%) in semi-denaturing conditions (in absence of β-mercaptoethanol) and the bands corresponding to 30kDa were given for trypsin digestion followed by MALDI-TOF/TOF analysis. Identified peptides based on m/z values were matching with the Sll1130 sequence.

Sll1130 tetramer includes only one di-sulphide bridge and non covalent associations between two homodimers

The experimental results of gel exclusion chromatography, semi-denaturing PAGE and native PAGE analysis has confirmed that the protein exists as a tetramer. However, the protein has only one cystine in it and is expected to form a single disulphide bridge with other monomer. However, we observe a tetramer formation. Therefore, it is proposed that the two dimmers may be involved in formation of a tetramer by non-covalent interactions. A model has been proposed to explain the arrangement of the Sll1130 molecules in formation of tetramer (Figure 4.22).

Figure 4.22: Model proposed to explain the Sll1130 tetramer formation

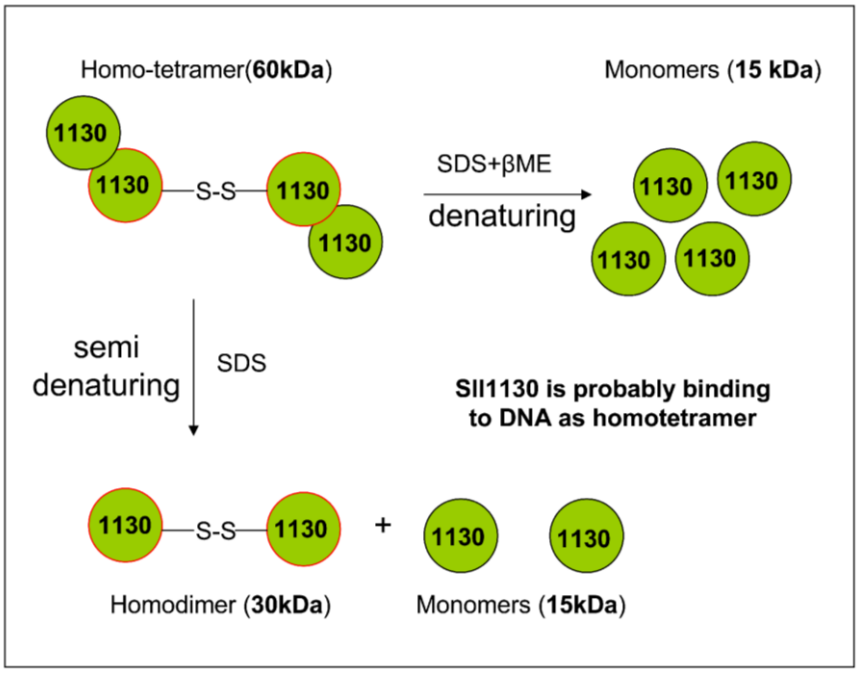


Figure 4.22. A model explaining the oligomerization of the Sll1130; purified His-Sll1130 (represented as 1130) exists in tetrameric state in native conditions. Formation of tetramer includes only one cysteine di-sulphide bond. Non-covalently associated Sll1130 dimers (two) are linked to each other by a single cysteine disulphide bridge.

Over expression of Ssl2245 in *E. coli* lead to slow growth

Position wise the upstream genes of Sll1130 orthologues are MazE or AbrB like protein coding genes, it is likely that Ssl2245 may have an important physiological function and play role(s) in survival of the cell by inhibiting activity of its counter toxin. However Ssl2245 from *Synechocystis* exhibited toxic property rather than antitoxin activity (Figure 4.23). Growth profiles of *E. coli* expressing Ssl2245 in pET28a(+), Ssl2245-Sll1130 together in pET28a(+), *ssl2245* mRNA in pBAD-24 were compared. *E. coli* strain expressing Ssl2245 protein has exhibited slow growth prior to induction, probably due to leaky expression of Ssl2245. Induction of Ssl2245 expression with IPTG resulted in a slow growth phenotype of host *E. coli* strain (Figure 4.23). When Ssl2245 and Sll1130 are expressed together did not exhibit any slow growth phenotype (Figure 4.23). However, *ssl2245* cloned into pBAD vector, out of reading frame lead to transcription of only *ssl2245* mRNA, but not the Ssl2245 protein, but a different protein (because *ssl2245* ORF was out of frame due to insertion of a single nucleotide after the ATG start codon of vector), did not cause growth retardation of *E. coli* host strain upon induction with arabinose. From these results it is clear that Ssl2245 accumulation in the cell inhibits the growth of *E. coli*, but not in presence of Sll1130 showing that the growth inhibitory effect of Ssl2245 was nullified by Sll1130 protein (Figure 4.23). The toxic affect of Ssl2245 was due to the protein in the cell but not due to RNA or siRNA mediated inhibition of growth as presence of *ssl2245* mRNA did not exhibit any slow growth (Figure 4.23). Viability assays using Sytox green fluorescent dye has confirmed that the action of Ssl2245 is by growth inhibition but not by killing the cells.

Figure 4.23: Effect of Ssl2245 expression on growth of *E. coli*

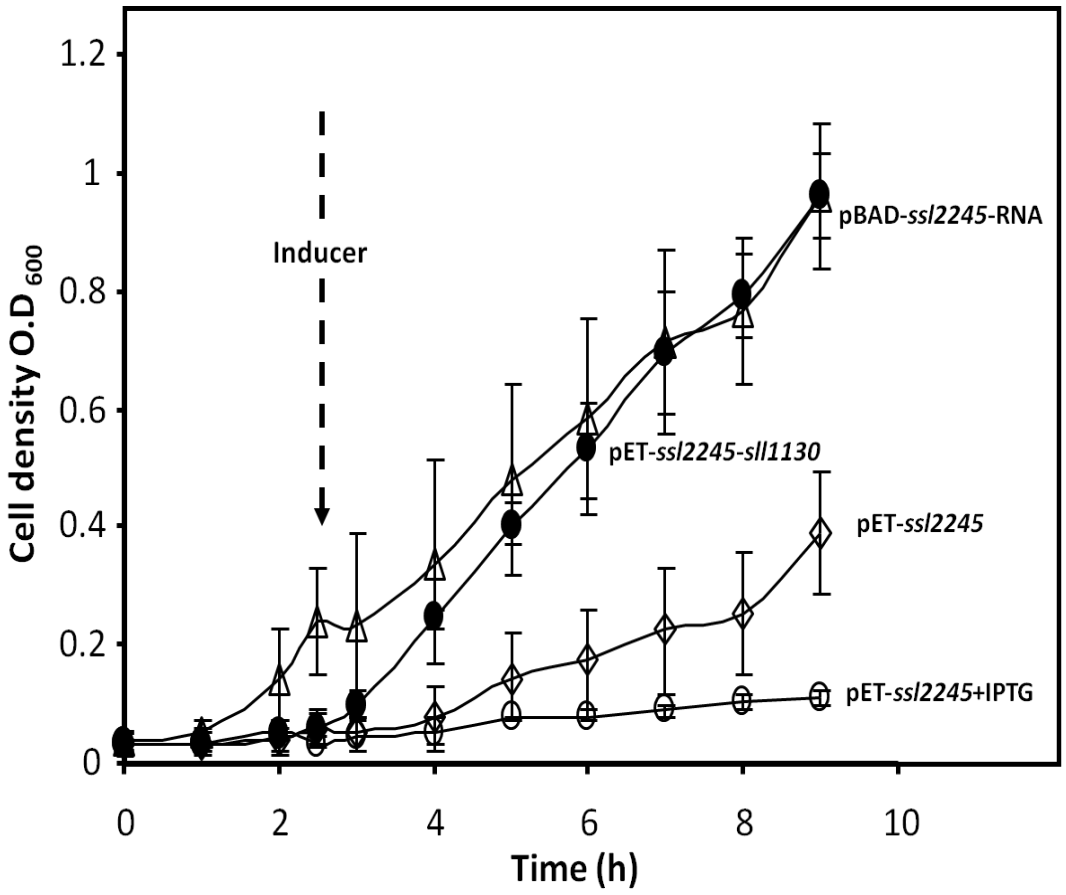


Figure 4.23. Growth analysis of the *E. coli* strains expressing Ssl2245 and Sll1130 in different combinations. *E. coli* cells were grown for two and half hours and induced with the specific inducer and the growth was monitored in side arm flasks in colorimeter with the filter having absorbance maximum at 600nm. pET-ssl2245, is *E. coli* BL21 with pET vector harboring *ssl2245* ORF (His-Ssl2245) but not induced ; pET-ssl2245+IPTG, is *E. coli* BL21 with pET vector harbouring *ssl2245* ORF (His-Ssl2245) and induced with 0.4mM IPTG; pET-ssl2245-sll1130, , is *E. coli* BL21 with pET vector harboring *ssl2245-sll1130* ORFs (His-Ssl2245 and Sll1130) and induced with 0.4 mM IPTG. pBAD-24-*ssl2245*-RNA, is *E. coli* BL21 with pBAD-24 vector harboring *ssl2245* ORF (ORF is out of frame so no protein will be produced) and induced with arabinose.

Ssl2245 is probably a stress mediated growth regulator:

Literature indicated that the Ssl2245 and Sll1130 are upregulated in Δ*sll*0082 along with several nitrate limitation inducible genes (Ishii and Hihara 2008). If nitrate limitation induces the expression of *ssl2245-sll1130* operon, it will create a situation where there is only Ssl2245 protein in the *Synechocystis* cells but not Sll1130. To see whether presence of Ssl2245 alone has growth inhibitory action on *Synechocystis* or not, *Synechocystis* WT, Δ*sll1130,* Δs*sl2245,* Δ*DM* and WT+ cells were grown under nitrate limiting conditions. Nitrate limitation was given by replacing the $NaNO_3$ in the BG-11 medium with equimolar concentration of NaCl. Under nitrate limiting growth conditions, Δ*sll1130* cells showed slow growth than the other strains (Figure 4.24). This confirms that the presence of Ssl2245 alone in the cell might have decreased the cell division and growth.

Ssl2245 and Sll1130 are physically associated with each other

Sll1130 is counteracting the growth inhibitory effect of Ssl2245, may be by interacting with Ssl2245 (Figure 4.23). To validate the physical association of Ssl2245 and Sll1130, we over expressed these two proteins in *E. coli* using pET28a (+) vector system. We have designed and made a DNA construct, in which *ssl2245-sll1130* operon was ligated to pET28a(+) in frame as described in materials and methods. Upon induction with IPTG, Ssl2245 with N-terminal His-tag is expected to be synethesized and the Sll1130 is synthesized without any tag. Ni^{+2}-NTA column elutes were resolved on on SDS-PAGE, identified two protein bands, one equivalent to 12 kDa and second one of 18 kDa(Figure 4.25). MALDI-TOF/TOF analysis of the resolved protein bands indicated that 12 kDa protein band corresponds to the Sll1130 and larger 18 kDa protein band to Ssl2245. However, the results clearly demonstrated

that Sll1130 is physically associated with His-Ssl2245, as they were co-eluted in Ni^{+2}-NTA agarose column (Figure 4.25).

Figure 4.24: Growth of all mutants in nitrate limited condition

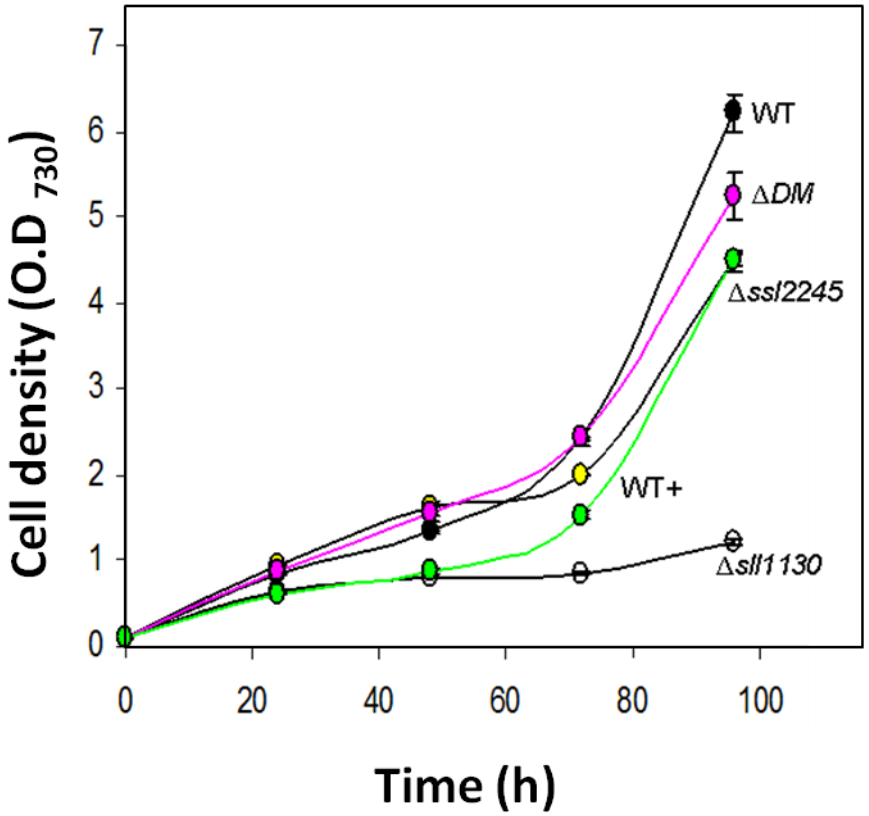


Figure 4.24. Growth analysis of different strains of *Synechocystis* in nitrate limited condition. Different strains of *Synechocystis* cells were grown for overnight in BG-11 medium, cells were harvested and resuspended in nitrate free BG-11, growth was monitored for 4 days. Growth of the WT, Δ*sll1130*, Δ*ssl2245*, DM and WT+ were allowed to grow at 34^{o}C and the growth was monitored (OD at 730nm).

Figure 4.25: Purification of Ssl2245 using Ni-NTA from *ssl2245-sll1130* expressing *E. coli*

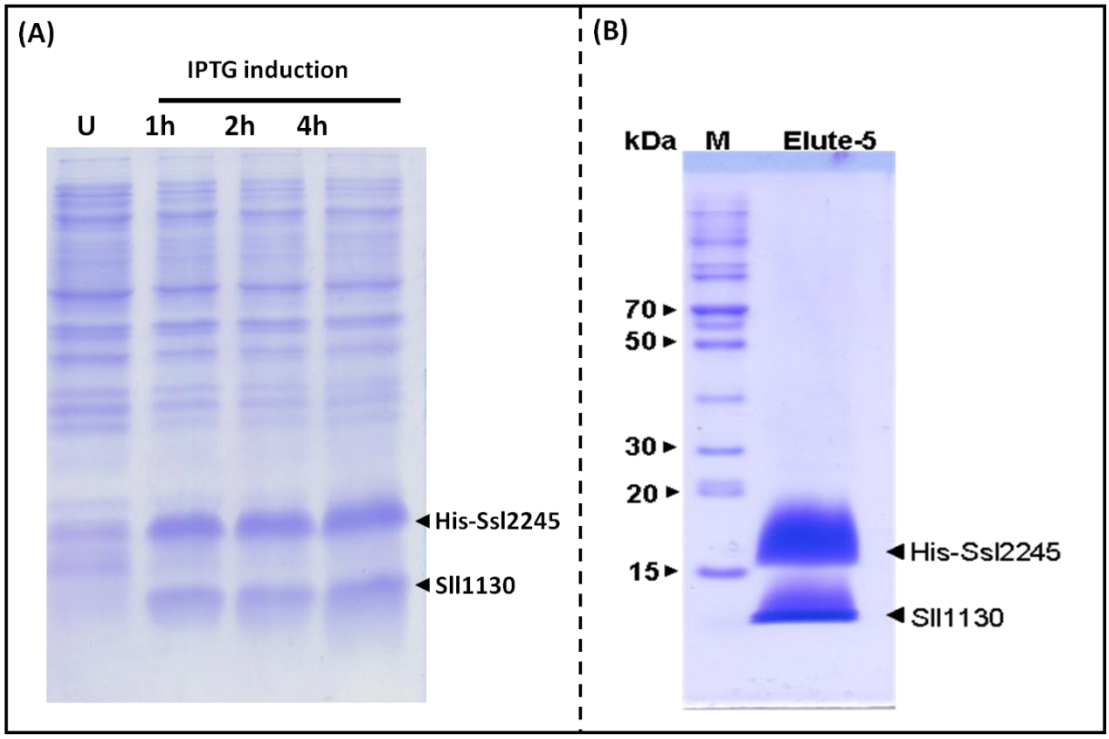
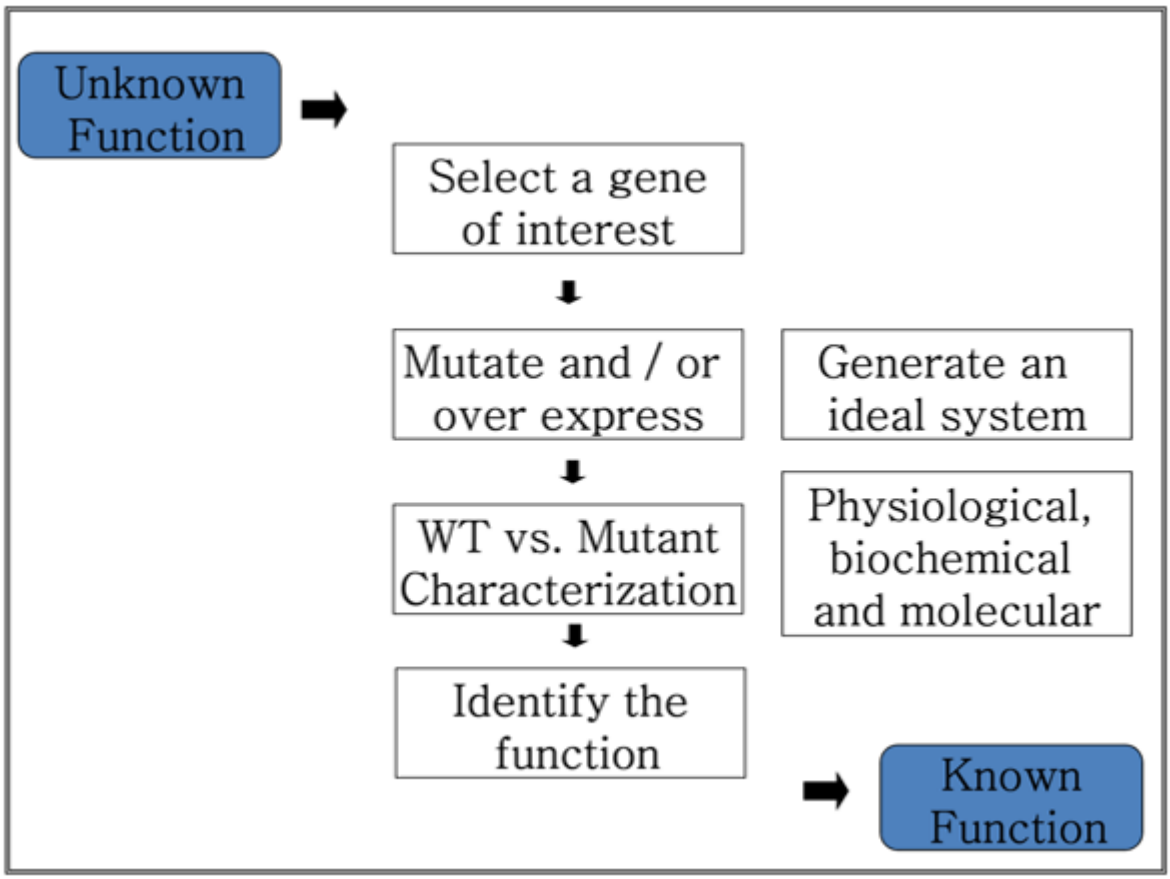


Figure 4.25. Purification of Ssl2245 using Ni-NTA from *ssl2245-sll1130* expressing *E. coli*. Protein samples are resolved in deaturing conditions in SDS-PAGE(15%). (A), Protein profile of IPTG induced *E. coli* expressing genes *ssl2245-sll1130* in pET28 a(+); (B) Ni-NTA purified protein elute-5.

Discussion


Unknown Function
Select a gene of interest
Mutate and / or over express
Generate an ideal system
WT vs. Mutant Characterization
Physiological, biochemical and molecular
Identify the function
Known Function

5. DISCUSSION

Functional characterization of Ssl2245 and Sll1130 in *Synechocystis*

The large-scale sequencing of genomes has revealed that 30 – 40% of the proteins encoded by typical bacterial genomes have functionally unknown proteins (Galperin and Koonin 2004). Well studied cyanobacterial model systems on average have more than 50% proteins uncharacterized (Table1.1). Even in a well studied organism such as *Synechocystis* only 50% of the proteins are functionally annotated. Functional annotation by homology-based approach may suite only when one of the orthologues is functionally characterized and had experimentally verified its function (Tian and Skolnick 2003). Sll1130 and Ssl2245 proteins encoded by a dicistronic operon in *Synechcoystis* chromosome are unknown proteins. Sll1130 is a conserved hypothetical protein shows similarity with PemK family proteins. Differential expression of this protein pair with heat stress shows its possible role in heat stress tolerance (Figure 1.2 and Figure 4.11).

Sll1130 is a constitutively expressing soluble protein but not PSII associated protein

Based on report of Kashino *et al*., 2002 we started verifying the role of Sll1130 in PSII structure and function. Complete mutation of *sll1130* in *Synechocystis* has affected neither growth nor PSII function (Figure 4.4 & 4.5). Further, repair of photo-damaged PSII is similar in wild type and Δ*sll1130* (Figure 4.5 B). Moreover, Sll1130 orthologues are present in bacteria of different taxonomic groups, irrespective of their mode of nutrition, clearly suggests that Sll1130 is not a

protein with role in PSII structure and function, as bacteria with no PSII complex has Sll1130 orthologues (Figure 4.1). However, previous reports by Kashino *et al.*, 2002 indicated that Sll1130 is associated with the PSII, but no one has reported the presence of Ssl2245 in association with PSII which is an interacting partner of the Sll1130 (Sato *et al.*, 2007 and Figure 4.25). Being the member of PSII complex, a membrane protein complex, Sll1130 was expected to be found in membrane fraction of *Synechocystis* cells. Absence of hydrophobic regions as predicted *in silico* and detection of Sll1130 in the soluble fraction of *Synechocystis* but not in insoluble fraction by anti-Sll1130 antibody (Figure 4.6), further confirms that the Sll1130 protein is not associated with PSII complex but is a constitutively expressing soluble protein. With these observations we report that Sll1130 is a constitutively expressing soluble protein with no role in PSII structure, function and repair.

Sll1130 is not a toxin protein

Most bacteria have Toxin-antitoxin (TA) systems. Usually toxins are small proteins approximately 100 amino acid residues in length and may inhibit cell growth by targeting a key molecule in any one of several essential cellular processes, including DNA replication, mRNA stability, protein synthesis, cell-wall biosynthesis, and ATP synthesis (Pandey and Gerdes 2005; Shao *et al.*, 2011). PemK super family proteins in *E. coli* are the growth inhibitors or plasmid maintenance proteins (Amitai *et al*., 2004). Over expression of such toxic proteins will lead to death or inhibition of growth. Sll1130 orthologues are present in bacteria of different taxons with more than 60% sequence similarity at the amino acid level and has PemK domain (Figure 4.1). Over expression of Sll1130 in *E. coli* led to the accumulation of protein in the

cell but not inhibited its growth (Figure 4.2), suggesting that it is not toxic to the cell. With these observations we conclude that despite of similarity with PemK super family proteins Sll1130 is not toxic upon over expression in *E. coli* and may not be a toxin.

Sll1130 negatively regulates heat-responsive genes in *Synechocystis* sp. PCC 6803

The major heat shock proteins that have been so far well characterized in terms of their function are GroES, GroEL, DnaJ, DnaK, HspA, HtpG and several proteases. Understanding the regulation of the heat shock genes is essential to unraveling the molecular basis of heat acclimation. In bacteria, negative regulation of several heat shock genes, such as *dnak, dnaJ, groES, groEL, grp, clpB, clpC* and *clpP*, has been previously reported (Bucca *et al.*, 1995; Derré *et al.*, 1998; Zuber and Schumann 1994; Nakamoto *et al.*, 2003). However, to the best of our knowledge, a repressor involved in negative regulation of *hspA* and *htpG* has not been reported so far (Table 1.3). This is the first report on a conserved hypothetical protein, Sll1130, as a novel heat repressible transcription factor that negatively regulates the expression of heat responsive genes such as *htpG, hspA, isiA, isiB* and several hypothetical genes. The Δ*sll1130* mutant exhibited enhanced thermotolerance compared to the wild type *Synechocystis* cells due to de-repression of *hspA, htpG, isiA* and *isiB* and accumulation of the corresponding proteins prior to heat treatment (Table 4.1; Figures 4.7 and 4.8). In support of the heat tolerant phenotype of the Δ*sll1130* mutant, over expression of *hspA, htpG* and *isiA* led to enhanced thermotolerance while mutation of these genes led to heat sensitive phenotypes in *Synechocystis* (Nakamoto *et al.*, 2000, Fang and Barnum 2003; Kojima *et al.*, 2006; Kojima and Nakamoto 2005). Gene

expression changes observed by DNA microarray and qRT-PCR analyses and the presence of a common *cis*-regulatory element in the upstream region of these genes confirmed that Sll1130 negatively regulates all these genes (Table 4.1; Figure 4.9 and 4.10). It seems likely that Sll1130 binds to its target DNA element as a tetramer, since the observed retardation of the DNA fragment was so severe that the DNA-protein complex appeared just below the wells on the polyacrylamide gel (Figure 4.10). An immediate down-regulation of *sll1130* transcript and protein levels upon upward shift in temperature suggests that the heat induced decrease in Sll1130 leads to de-repression of heat responsive genes upon heat shock (Figure 4.11 A, B and C). Collectively, our data suggests that Sll1130 binds to a *cis*-regulatory element (an inverted-repeat) located in the upstream region of major heat responsive genes *i.e., hspA, htpG, isiA, isiB* and several other genes coding for hypothetical proteins, and negatively regulate their expression (Figure 5.1A). This is the first report of a novel heat responsive transcription factor, Sll1130 that negatively regulates the expression of *htpG, hspA* and *isiAB* genes. Upon heat shock, the Sll1130 levels are down-regulated by an unknown signal transduction mechanism, leading to de-repression of heat responsive genes and, as a consequence, *Synechocystis* acclimatizes to the upward shift in temperature (Figure 5.1 B). The absence of the Sll1130 protein in the Δ*sll1130* mutant cells resulted in elevated levels of heat responsive proteins that conferred increased thermotolerance.

Figure 5.1: Proposed model explaining the function of Sll1130

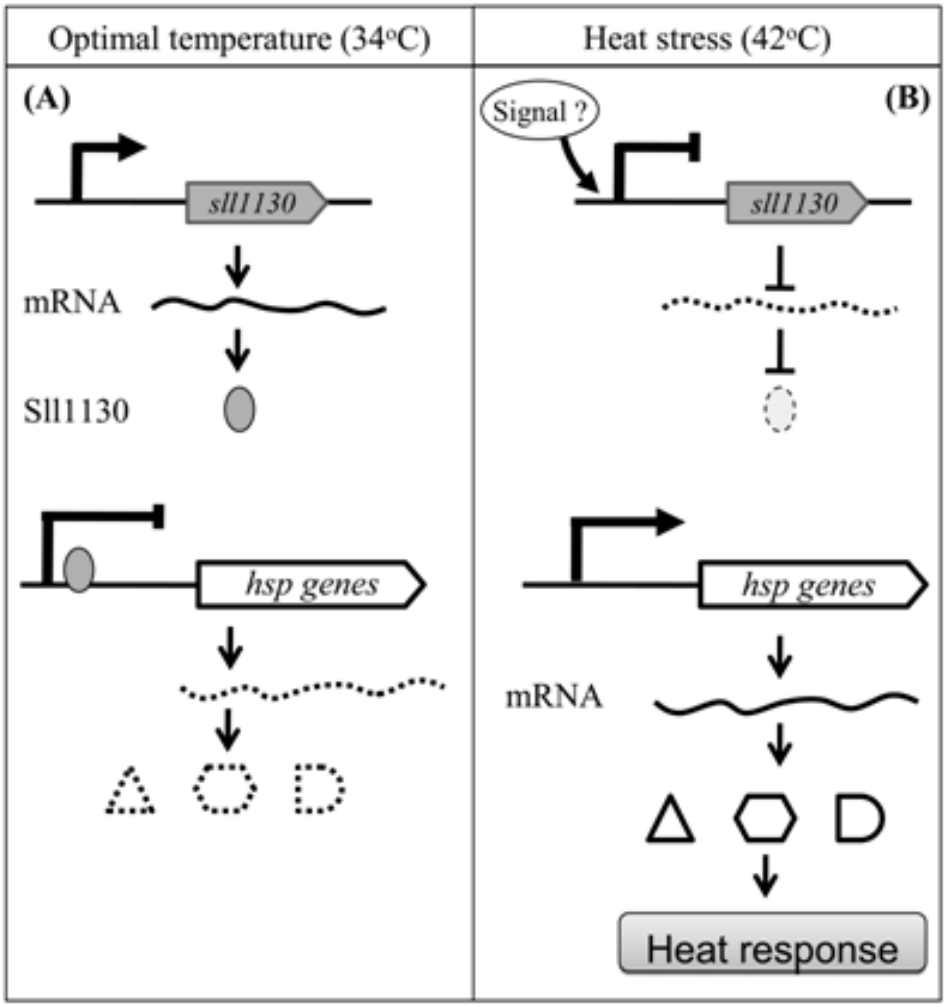


Figure 5.1. Proposed model for the regulation of gene expression by Sll1130. (A) *sll1130* is a constitutively expressed gene. At 34°C, Sll1130 negatively regulates the expression of *hspA, htpG, frpC, isiA, isiB* and other genes by binding to the *cis* regulatory element located in the upstream region. (B) Upon heat shock (42°C) *sll1130* is downregulated by an unknown mechanism of signal transduction. Reduced levels of Sll1130 de-repress *hspA, htpG, frpC, isiA, isiB* and other genes, whose protein products then take part in the process of heat acclimation.

Ssl2245 is a chromosomal encoded switch for plasmid genes in Synechocystis

Ssl2245 annotated as an unknown protein in *Synechocystis* is a conserved hypothetical protein (Figure 1.1). Sequence wise Ssl2245 shows no significant similarity with any known protein (Figure 4.13). There are no annotated proteins with significant similarity with Ssl2245 (Figure 4.14). In well characterized TA systems, the first gene is a MazE. MazE proteins are the antitoxins that regulate the toxicity of the MazF (PemK protein, with endoribonuclease activity). MazE-MazF protein pair forms a TA system in *E. coli*. Binding of MazE to MazF regulates the toxicity and prevents the cell death. Expression of MazF in absence of MazE leads to cell death (Amitai *et al.,* 2004). MazE and AbrB like proteins are phylogenetically related and are involved in stress mediated regulation of cell growth or cell death (Christensen *et al.,* 2003). AbrB-like proteins are transcriptional regulators and are reported in a wide variety of gram-positive, gram-negative and archaeal species (Vaughn *et al.,* 2000; Coles *et al.,* 2005). In most cases, the N-terminal DNA-binding domains of these proteins are conserved, whereas the C-terminal domains are significantly divergent in sequence and in size. To date, the most widely studied member of the AbrB family has been the *Bacillus subtilis* AbrB protein. AbrB, a homotetramer of 10.5-kD subunits (Vaughn *et al.,* 2000; Benson *et al.,* 2002; Bobay *et al.,* 2004) is a pivotal regulator of transition state gene expression upon entry into stationary phase and sporulation (Phillips and Strauch 2002). No DNA consensus sequence recognized by AbrB was found upstream of the target genes, and it is believed that AbrB-DNA binding interactions entail the recognition of specific stretches of three-dimensional

DNA helix configurations (Bobay *et al.,* 2004). The C-terminal domain of AbrB is required for the formation of the functional tetramer (Yao and Strauch 2005). Cyanobacterial AbrB proteins (cyAbrBs) are unique in that they have an AbrB-like DNA-binding domain in the C terminal region. Multiple copies of genes encoding cyAbrBs are found in every cyanobacterial genomic sequence now available, whereas AbrB-like regulators having DNA-binding domains at their C termini are not conserved in other bacterial species. Recently, Oliveira and indblad (2008) reported the isolation of the Sll0359 protein, one of the cyAbrBs in *Synechocystis* clearly demonstrating that the cyAbrBs are functional in cyanobacterial cells. However, their precise functions and physiological significance are still unclear. Till date there are five cyAbrBs in *Synechocystis* (http://genome.microbedb.jp/cyanobase/Synechocystis/genes/sll1130). Sll0359, Sll0822 are the chromosomal encoded AbrB like proteins. Ssr7040 (pSYSA), Ssl8028 (pSYSG) and Ssl7048 (pSYSA) are the plasmid encoded AbrB like proteins. These genes were reported as growth regulators. They may function like antitoxins or regulate the gene expression. Interestingly the *ssr7040* is upstream gene of *slr7041* a PemK like, growth regulatory protein coding gene with a 10 nt overlapping region, shows much more similarity to *ssl2245* and *sll1130* operon. Mutagenesis and further physiological characterization of the Δ *ssl2245* has shown that the growth profile of Δ *ssl2245* is similar to WT in optimal conditions (Figure 4.16). In spite of giving antibiotic selection pressure for several generations the mutation of *ssl2245* is not complete (Figure 4.15). Transcriptome analysis of the Δ*ssl2245* has shown that the mutation has affected the expression of genes in large native plasmids especially pSYSA and pSYSX (Table 4.2). Due to mutation of *ssl2245*, *sll1130* transcript was

not produced in the cell (Table 4.2), possibly due to unavailability of the promoter for transcription of *sll1130*. Mutation of *sll1130* also led to down regulation of *ssl2245* expression, *i.e* both Δ*ssl2245*and Δ*sll1130* during optimal conditions act like a double mutant (Table 4.1 and 4.2). The genes that are upregualted in the Δ*sll1130* (*slr1788*, *sll1009* etc.) are upregulated in Δ*ssl2245* also (Table 4.2). Over expression of Ssl2245 has decreased the rate of division in *E. coli.* Reduction in growth of *E. coli* is due to Ssl2245 protein in the cell but not the mRNA (Figure 4.23). However, presence of both the proteins Ssl2245 & Sll1130 has not affected the growth and division of the *E. coli* showing that the growth regulatory activity of the Ssl2245 was nullified by the presence of Sll1130. Expression of an extra copy of the Ssl2245 in the WT *Synechcocystis* using cyanobacterial expression vector led to the slow growth and nitrogen limitation conditions have resulted in decreased growth in case of the *sll1130* mutant but not *ssl2245* mutant. With all these results we conclude that Ssl2245 has the properties like AbrB like protein and regulates the expression of plasmid genes and is involved in stress mediated growth regulation.

Ssl2245 and Sll1130 are probably evolved from Toxin-Anti-toxin systems

These two proteins Ssl2245 (88 aa) and Sll1130 (115 aa) are coded by an over lapping gene pair. The stop codon (ATGA) of the first gene (*ssl2245*) overlaps with the start codon (ATGA) of the second ORF (*sll1130*). (http://genome.microbedb.jp/cyanobase/Synechocystis/genes/sll1130). Ssl2245 and Sll1130 are two interacting partners (Sato *et al.*, 2007 and Figure 4.25). These genes are under the control of same promoter and mRNA is dicistronic coding for both the proteins (Figure 4.12). Ssl2245 and Sll1130 are antagonistic to each other in function

like TA systems (Figure 4.23). The antitoxins sequester the toxins by direct protein–protein interaction but are degraded by cellular proteases (Lon or Clp). Due to lack of antitoxin replenishment, plasmid-free cells experience activation of the toxins, which in turn, reduce or prevent the growth of these cells. Curiously, plasmid and chromosome-encoded TA loci stabilize plasmids equally well (Gotfredsen & Gerdes 1998). The biological function of chromosome-encoded TA loci is less well understood. Nutritional stress with emphasis on amino acid starvation leads to the induction of TA systems in bacteria (Christensen *et al.,* 2001), but presence of homologues of these TA systems in cyanobacterial chromosomes was been an enigma. These toxins are co-transcribed and co-translated with their cognate antitoxins from an operon called a toxin-antitoxin (TA) operon (Pandey and Gerdes 2005; Yamaguchi *et al.,* 2011). As a result, in the normally growing cells, in order to constantly inhibit the toxin function, cognate antitoxins have to be continuously synthesized (Yamaguchi and Inouye 2009). Hence chromosomal encoded stress responsive TA systems exists as pairs with overlapping region under the control of same promoter. Well studied TA systems are the type –II, TA systems Almost all the Type II TA systems share a number of common features (Gerdes *et al.,* 2005; Yamaguchi and Inouye 2009; Yamguchi *et al.,* 2011). Type II TA cassettes have a characteristic organization in which the gene for the antitoxin component precedes the toxin gene; the two loci often overlap, reflecting a common autoregulatory mechanism exerted by both components. Although most TA modules conform to this arrangement, there are examples of TA cassettes in which the gene order is reversed, where the antitoxin alone exerts the regulatory effect or where the product of a third gene is implicated (Engelberg-Kulka and Glaser 1999). Ssl2245 and Sll1130 follow

the same principle that both the genes are under the control of same promoter and have a over lapping region and second gene (*sll1130*) has PemK domain like toxin and upstream gene (*ssl2245*) is position wise an AbrB like protein coding gene. This gene pair is transcribed together and post transcriptional regulation is possible due to the overlapping region of 4nt. As both the ORFs produce functional proteins it is clear that these two may be type II TA systems but not either type-I or type-III. Sometimes these toxin antitoxin members act as modules of bacterial stress managers (Buts *et al.,* 2005). Down regulation of *ssl2245* and *sll1130* in heat stress and co-expression with nitrate limited inducible genes in Δ*sll0822* confers the role of this pair in stress mediated growth regulation. With the available literature, sequence wise similarity, operon structure and overlapping region, their antagonistic activity and DNA binding properties and the characterization of the corresponding mutants we conclude that this protein pair is probably evolved from TA systems and is involved in the regulation of the cell growth. A model was proposed to explain the function of this interacting protein partners based on the above mentioned experimental results (Figure 5.2). Differential expression of this operon with abioitic stress confirms its role in stress tolerance and stress mediated growth regulation. It is clear that the relative levels of Ssl2245 and Sll1130 play role in cell growth and division. Common promoter, presence of overlapping region of nucleotides shows possibility of post transcriptional regulation of gene expression.

Figure 5.2: Model proposed to explain the role of Ssl2245 and Sll1130 interaction

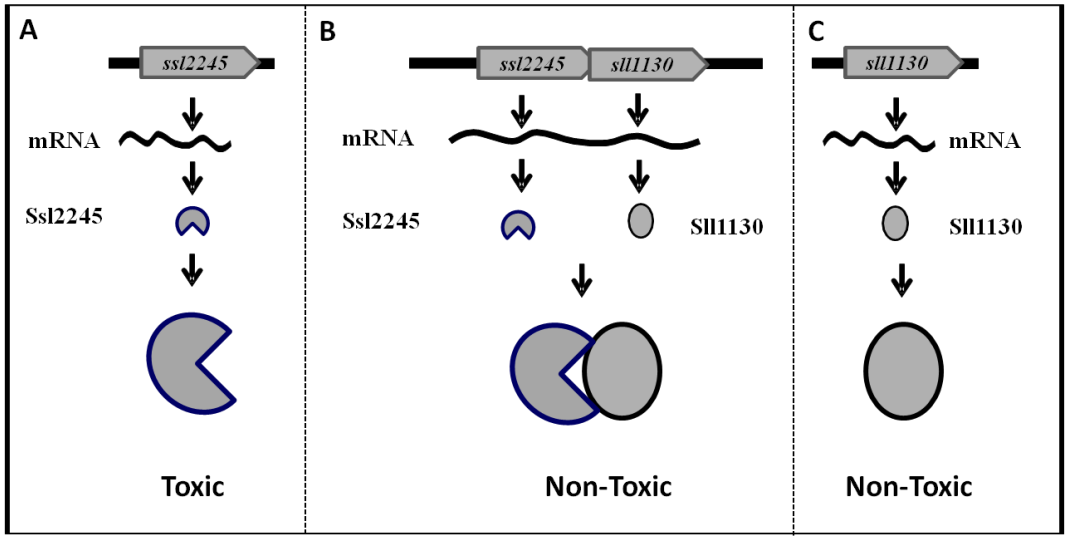
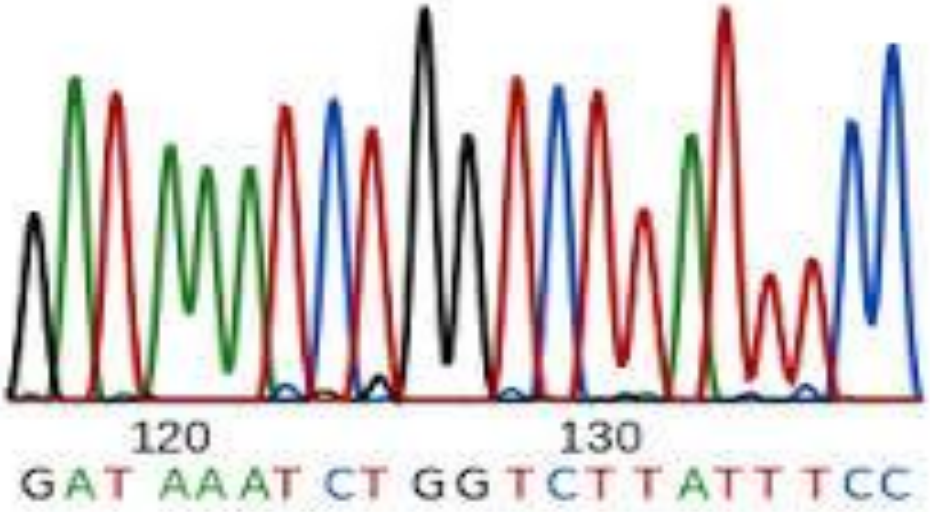


Figure 5.2. Proposed model to explain the role of Ssl2245-Sll1130 association. (A) *ssl2245* expression alone has decreased cell growth possibly due to its toxic effect. (B) growth inhibitory action of Ssl2245 was neutralized by Sll1130 by its association. (C) Sll1130 accumulation has no affect on the growth of cells. Relative concentrations of Ssl2245 and Sll1130 inside the cell might play crucial role in abiotic stress mediated growth regulation.

CONCLUSIONS

Ssl2245 and Sll1130 are two soluble proteins coded by a dicistronic operon, expressing constitutively. Sll1130 is a conserved protein with 4 predicted helical regions and DNA binding property. Despite of its sequence similarity with MazF toxin, Sll1130 accumulation is not toxic to the cell. The mRNA levels of these genes decreases in the cell with the heat stress, showing these are heat stress repressible. Mutagenesis and microarrays has shown absence of Sll1130 probably acts like a switch to these upregulated genes. Sll1130 binds to an inverted repeat in the upstream of genes that are upregulated in the mutant and regulates their transcription in optimal conditions. Sll1130 exists as a tetramer with a single disulphide bridge involvement. Ssl2245 is similar to AbrB like proteins based on operon structure. Mutagenesis and microarray has shown absence of Ssl2245 upregulates genes in the plasmids, probably acts like a genome encoded switch to plasmid genes. Expression of an extra copy of *ssl2245-sll1130* operon in *Synechocystis* led to the slow growth in optimal conditions. Over expression of Ssl2245 is toxic to the *E. coli* cells and the toxic activity of the Ssl2245 is nullified by Sll1130. Ssl2245 and Sll1130 are physically associated with each other. Similarity to the TA system at the level of operon structure, PemK domain, overlapping region, transcriptional regulator activity and differential expression in response to the abiotic stress gives us clues that this gene pair is probably evolved from TA systems and plays role in abiotic stress mediated growth regulation.

Primer-sequences



6. PRIMER SEQUENCES

Primers used for expression of *sll1130* and *ssl2245*

Name	Primer sequence 5'-3'	R.E. sites
sll1130-exp-F	GCGCCATGGATACAATTTACGAACAATTTG	NcoI
sll1130-exp-R	GCGAAGCTTACCGAGTTTAAAAACATGGGG	HindIII
ssl2245-exp-F	CGGCATATGTCTATCAATGCTTACAAACTAGC TACG	NdeI
ssl2245-exp-R	GGCAAGCTTTCA TAG GTG TCG GTA TGC AGA ATT ATC AGC	HindIII
WT+ F	TTACTCGAGCATACTTTAGCTTGTCGTCCCGG GC	- -
WT+ R	CCGATCGATGCT AAT CAT ATT TGC GGC CAG CCA GTG G	- -

Primers used for amplification of *sll1130* for probe preparation in northerns

Name	Primer sequence 5'-3'	R.E. sites
sll1130-Nor-F	CATTGTCCCAGTTCCCTTCACGGATAGG	- -
sll1130-Nor-R	GCA CTC TTC ACC ATC AAT CGG TCT TGC	- -

Primers used for Omega spectinomycin cassette amplification

Name	Primer sequence 5'-3'	R.E sites
Spec F	AAACTTTTTAAATCCTTAATTATTTGCCCACTAAAC	DraI
Spec R	ATCAAAGTTTAAAACTCCCCCAGGGTCTTAGTT C	DraI

Primers used for amplification of *slr1788* upstream and the specific inhibitor used

Name	Primer sequence 5'-3'	Modification
slr1788-GS-F	GAGAAACTGATCTTGAAGAAGTGG	5' Cy3 labeled
sr1788-GS-R	GATTTTGGTAATTGGATCATGGCG	5' Cy3 labeled
Inhibitor-F	TGCCCCAAGCAAGAATGGCGATCGCCGTT GGGATTCGGAGC	- -
Inhibitor-R	GCTCCGAATCCCAACGGCGATCGCCATTC TTGCTTGGGGCA	- -

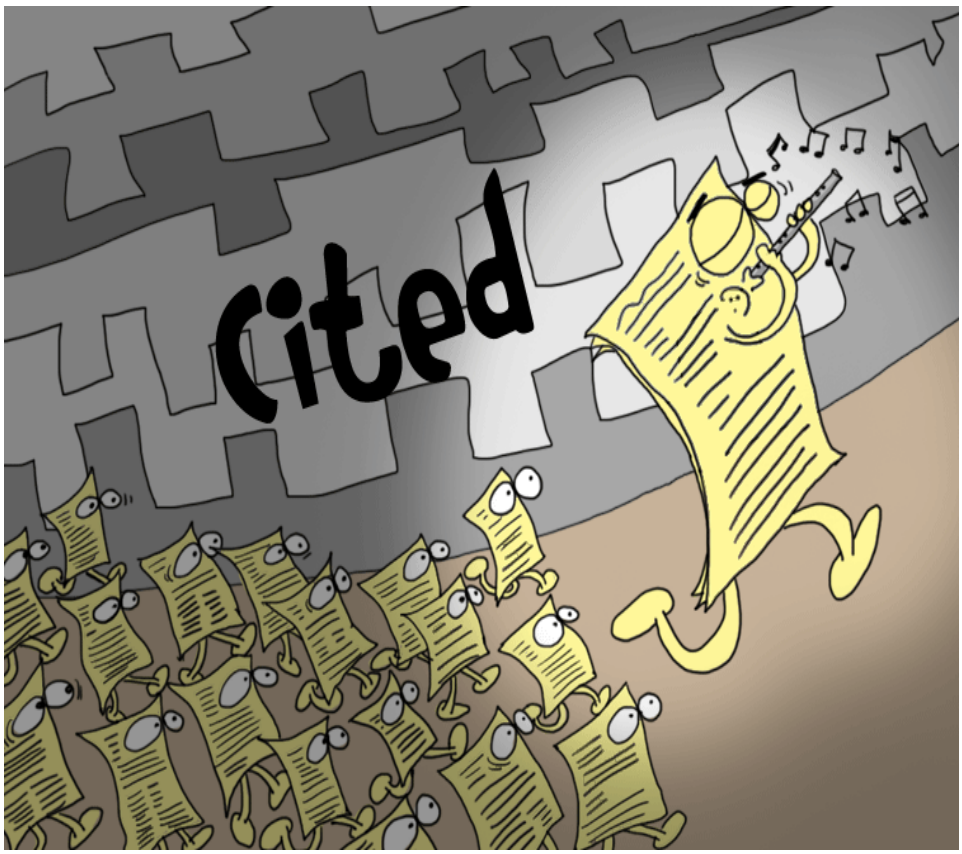
Primers used for PCR amplification of *sll1130* during mutagenesis

Name	Primer sequence 5'-3'	R.E. sites
sll1130-F	TATGGCTGCCACCGCCGACACTATGAC	- -
sll1130-R	GTCCAGCGGATCTAGTTATCTTTCCAG	- -

Primers used for qRT-PCR

Gene	Primer pair (5' -3')
slr1788	F: TTCGACTCCGTTACATCCTG R: CGTTTAAGCAGATCGTCGGT
sll1009 (frpC)	F:ATGTTGGCGATGATACCGTC R: CATTGCCGTCGCCACCATA
sll1514 (hspA)	F:GAAACTGAAGAAGCCTATGTG R: CGGTGCTATGGGTATCCTG
sll0247 (isiA)	F:CTTTGTCATTGGTAGTATTCAC R: CCAGGAACAGGAGATGATGA
ssl2245	F:CCGTAGAAGTTATTGTGTTAGT R: GGTGTCGGTATGCAGAATTAT
slr0884 (gap1)	F:ACCATCTCAACTACAACGCC R: GGTCATCAATCCTTCCACAAT
16s rRNA(*slr0808*)	F:TCGCCTACCGTTGGAGCC R: GGGCACAATGGCTTCAACAA

References



7. REFERENCES

1. Aizenman, E., Engelberg-Kulka, H. and Glaser, G. (1996) An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3_,5_-bispyrophosphate: a model for programmed bacterial cell death. Proc. Natl. Acad. Sci. USA. 93, 6059–63.
2. Amitai, S., Yassin, Y. and Engelberg-Kulka, H. (2004) MazF-mediated cell death in Escherichia coli: a point of no return. J.Bacteriol. 186, 8295-300.
3. Arnon, D., I., McSWwain, B., D., Tsujimoto, H., Y. and Wada, K. (1974) Photochemical activity and components of membrane preparations from blue-green algae, Biochim. Biophys. Acta. 357, 231–245.
4. Aro, E.V., Hundal, T., Carlberg, I. and Andersson, B. (1990) In vitro studies on light-induced inhibition of photosystem II and D 1-protein degradation at low temperatures, Biochim. Biophys. Acta. 1019, 269-275.
5. Bardwell, J.C.A. and Graig, E.A. (1988) Ancient heat shock gene is dispensable. J. Bacteriol. 170, 2977-2983.
6. Benson, L., M., Vaughn, J., L., Strauch, M., A., Bobay, B., G., Thompson, R., Naylor, S. and Cavanagh, J. (2002) Macromolecular assembly of the transition stateregulator AbrB in its unbound and complexed states probed by microelectrospray ionization mass spectrometry. Anal Biochem. 306, 222–227.
7. Bobay, B. G., Benson, L., Naylor, S., Feeney, B., Clark, A.C., Goshe, M.B., Strauch, M. A., Thompson, R. and Cavanagh, J. (2004) Evaluation of the DNA binding tendencies of the transition state regulator AbrB. Biochemistry. 43, 16106–16118.
8. Bricker, T.M. (1990) The structure and function of Cpa-1 and Cpa-2 in photosystem II, Photosynthesis Res. 24, 1-13.
9. Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J., M., Davenport, A., *et al*., 2009. Three dimensional structure of the MqsR:MqsA complex: a novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. PLoS Pathog. 5:e1000706.
10. Bucca, G., Ferina, G., Puglia, A., M. and Smith, C., P. (1995) The dnaK operon of *Streptomyces coelicolor* encodes a novel heat-shock protein which binds to the promoter region of the operon. Mol. Microbiol. 17, 663-674.
11. Bumann, D. and Oesterhelt, D. (1995) Destruction of a single chlorophyll is correlated with the photoinhibition of photosystem II with a transiently inactive donor side, Proc. Natl. Acad. Sci.USA. 92, 12195- 12199.
12. Buts, L., Lah, J., Dao-Thi, M.H., Wyns, L. and Loris, R. (2005) Toxin–antitoxin modules as bacterial metabolic stress managers TRENDS in Biochemical Sciences. 30, 672-679.
13. Chen, H., Zhang, D., Guo, J., Wu, H., Jin, M., Lu, Q., Lu, C., and Zhang, L. (2006a) A Psb27 homologue in *Arabidopsis thaliana* is required for efficient repair of photodamaged photosystem II. Plant Mol. Biol. 61, 567–575.
14. Chen, F., Mackey, A., J., Stoeckert, Jr, C., J. and Roos, D., S. (2006b) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res. 34, 363–368.

15. Christensen, S., K., Pedersen, K., Hansen, F., G., and Gerdes, K. (2003) Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J. Mol. Biol. 332, 809–819.
16. Coles, M., Djuranovic, S., So¨ding, J., Frickey, T., Koretke, K., Truffault, V., Martin, J. and Lupas, A. N (2005) AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi betabarrels. Structure. 13, 919–928.
17. de Cr´ecy-Lagard, V. (2007) Identification of genes encoding tRNA modification enzymes by comparative genomics. Methods Enzymol. 425, 153–183.
18. Derré, I., Rapoport, G. and Msadek, T. (1998) CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in Gram-positive bacteria. Mol. Microbiol. 31, 117-131.
19. Diago-Navarro1, E., Hernandez-Arriaga1, A. M., Juan Lo´ pez-Villarejo1, Ana J. Mun˜ oz-Go´ mez1, Monique B. Kamphuis, Rolf Boelens, Marc Lemonnier and Ramo´n Dı´az-Orejas (2010) parD toxin–antitoxin system of plasmid R1 – basic contributions, biotechnological applications andrelationships with closely-related toxin–antitoxin systems. FEBS Journal. 277, 3097–3117.
20. Engelberg-Kulka, H. and Glaser, G. (1999) Addiction modules and programmed cell death and antideath in bacterial cultures. Annu. Rev. Microbiol. 53, 43–70.
21. Fang, F. and Barnum, S., R, (2003) The heat shock gene, *htpG*, and thermotolerance in the cyanobacterium, *Synechocystis* sp. PCC 6803. Curr. Microbiol. 47, 341-6.
22. Feder, M., E. and Hofmann, G., E. (1999) Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annu Rev Physiol 61: 243-282.
23. Fineran, P., C., Blower, T., R., Foulds, I., J, Humphreys, D., P., Lilley, K., S., Salmond, G., P. (2009) The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proc. Natl. Acad. Sci. USA. 106:894–99.
24. Fozo, E.M., Kawano, M., Fontaine, F., Kaya, Y., Mendieta, K.S., et al., (2008) Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol. Microbiol.* 70:1076–93.
25. Frishman, D. (2007) Protein annotation at genomic scale: the current status. Chem. Rev. 107, 3448–3466.
26. Galperin, M., Y. and Koonin, E., V. (2004) 'Conserved hypothetical' proteins: prioritization of targets for experimental study. Nucleic Acids Res. 32, 5452–5463.
27. Gerdes, K., Christensen, S., K. and Løbner-Olesen, A. (2005) Prokaryotic toxin-antitoxin stress response loci. Nature Microbiol. Rev. 3, 371-382.
28. Gerdes, S., Edwards, R., Kubal, M., Fonstein, M., Stevens, R. and Osterman, A. (2006) Essential genes on metabolic maps. Curr. Opin. Biotechnol. 17, 448–456.

29. Gerdes, K., Wagner, E., G. (2007) RNA antitoxins. *Curr. Opin. Microbiol.* 10:117–24.
30. Gollub, J., Ball, C., A. and Sherlock, G. (2006) The Stanford Microarray Database: a user's guide. Methods Mol. Biol. 338, 191–208.
31. Gotfredsen, M., Gerdes, K. (1998) The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. Mol Microbiol 29:1065–1076.
32. Gottesman, S., Wickner, S. and Maurizi, M., R. (1997) Protein quality control: triage by chaperones and proteases. Genes and Develop. 11, 815-823.
33. Griese, M., Lange, C. and Soppa, J. (2011) Ploidy in cyanobacteria FEMS Microbiol Lett. 323, 124–131.
34. Grossman, A., D., Zhou, Y., N., Gross, C., Heilig, J., Christie, G., E. and Calendar R. (1985) Mutations in the rpoH (htpR) gene of Escherichia coli K-12 phenotypically suppress a temperature-sensitive mutant defective in the sigma 70 subunit of RNA polymerase. J Bacteriol. 161, 939–943.
35. Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A. and Saenger, W. (2009). "Cyanobacterial photosystem II at 2.9 A^o resolution and the role of quinones, lipids, channels and chloride". Nat. Struct. Mol. Biol. 16, 334–42.
36. Hankamer, B., Barber, J. and Boekema, E.J. (1997) Structure and membrane organization of photosystem II in green plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 48, 641-671.
37. Hanson, A., D., Priibat, A., Waller, J., C. and Lagard V., D. (2010) 'Unknown' proteins and 'orphan' enzymes: the missing half of the engineering parts list – and how to find it. Biochem. J. 425, 1–11.
38. Hecker, M., Schumann, W. and Völker, U. (1996) Heat-shock and general stress response in *Bacillus subtilis*. Mol Microbiol. 19, 417-28.
39. Hedges, S., B., Blair, J., E., Venturi, M., L. and Shoe, J., L. (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol. Biol. 4, 2.
40. Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. (2001). DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. Plant Cell. 13, 793–806.
41. Hihara, Y., Sonoike, K., Kanehisa, M., and Ikeuchi, M. (2003). DNA microarray analysis of redox-responsive genes in the genome of the cyanobacterium *Synechocystis* sp. strain PCC 6803. J. Bacteriol. 185, 1719–1725.
42. Horwich, A.L., Fenton, W.A., Chapman, E. and Farr, G.W. (2007) Two families of chaperonin: Physiology and mechanism. Annu. Rev. Cell Dev. Biol. 23,115–145.
43. Hossain, M., M. and Nakamoto H (2002) HtpG plays a role in cold acclimation in cyanobacteria. Curr, Microbiol. 44, 291–296.
44. Hossain, M., M. and Nakamoto H (2003) Role for the cyanobacterial HtpG in protection from oxidative stress. Curr Microbiol. 46, 70–76.
45. Inaba, M., Suzuki, I., Szalontai, B., Kanesaki, Y., Los, D.A., Hayashi, H., and Murata, N. (2003). Gene engineered rigidification of membrane lipids enhances the cold inducibility of gene expression in *Synechocystis*. J. Biol. Chem. *278*, 12191–12198.

46. Ishii, A. and Hihara, Y. (2008) An AbrB-Like Transcriptional Regulator, Sll0822, Is Essential for the Activation of Nitrogen-Regulated Genes in Synechocystis sp. PCC 6803. Plant Physiology. 148, 660–670.
47. Jensen, R., B., Grohmann, E., Schwab, H., Diaz, R. and Gerdes, K. (1995) Comparison of *ccd* of F, *parDE* of RP4, and *parD* of R1 using a novel conditional replication control system of plasmid R1. Mol. Microbiol.17, 211–220.
48. Kamada, K., and Hanaoka, F. (2005) Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. Mol. Cell. 19, 497–509.
49. Kamei, A., Yuasa, T., Orikawa, K., Geng, X., X. and Ikeuchi, M. (2001) A eukaryotic-type protein kinase, SpkA, is required for normal motility of the unicellular Cyanobacterium *Synechocystis* sp. strain PCC 6803. J. Bacteriol.183, 1505-1510.
50. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., et al. and Tabata, S. (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. 3, 109–136.
51. Kaneko, T., Nakamura, Y., Sasamoto, S., Watanabe, A., Kohara, M., Matsumoto, M., Shimpo, S., Yamada, M. and Tabata, S. (2003) Structural Analysis of Four Large Plasmids Harboring in a Unicellular Cyanobacterium, *Synechocystis* sp. PCC 6803. DNA Research. 10, 221–228.
52. Kanesaki, Y., Suzuki, I., Allakhverdiev, S., I., Mikami, K. and Murata, N. (2002). Salt stress and hyperosmotic stress regulate the expression of different sets of genes in Synechocystis sp. PCC 6803. Biochem Biophys Res Commun. 290, 339–348.
53. Kanesaki, Y., Yamamoto, H., Paithoonrangsarid, K., Shoumskaya, M., Suzuki, I., Hayashi, H. and Murata, N. (2007) Histidine kinases play important roles in the perception and signal transduction of hydrogen peroxide in the cyanobacterium, *Synechocystis* sp. PCC6803. Plant J. 49, 313-324.
54. Kasari, V., Kurg, K., Margus, T., Tenson, T., Kaldalu, N. (2010) The *Escherichia coli mqsR* and *ygiT* genes encode a new toxin-antitoxin pair. J. Bacteriol. 192, 2908–19.
55. Kashino, Y., Lauber, W., M., Carroll, J., A., Wang, Q., Whitmarsh, J., Satoh, K. and Pakrasi, H., B. (2002) Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. Biochem. 41, 8004-8012.
56. Kashino, Y., Inoue-Kashino, N., Roose, J., L., and Pakrasi, H., B. (2006) Absence of the PsbQ protein results in destabilization of the PsbV protein and decreased oxygen evolution activity in cyanobacterial photosystem II. J. Biol. Chem. 281, 20834–20841.
57. Kashino, N., I., Kashino, Y., Orii, H., Satoh, K., Terashima, I. and Pakrasi H., B. (2011) S4 Protein Sll1252 is necessary for energy balancing in photosynthetic electron transport in *Synechocystis* sp. PCC 6803. Biochem. 50, 329–339.

58. Keren, N., Ohkawa, H., Welsh, E., A., Liberton, M., and Pakrasi, H., B. (2005) Psb29, a conserved 22-kD protein, functions in the biogenesis of Photosystem II complexes inSynechocystisandArabidopsis. Plant Cell. 17, 2768–2781.
59. Kojima, K. and Nakamoto, H. (2005) Constitutive expression of small heat shock protein in an *htpG* disruptant of the cyanobacterium *Synechococcus* sp. PCC 7942. Curr. Microbiol. 50, 272-276.
60. Kojima, K., Maenaka, T., S., Kikuchi, T. and Nakamoto, H. (2006) Roles of the cyanobacterial isiABC operon in protection from oxidative and heat stresses. Physiol. Plant. 128, 507-519.
61. Koonin, E., V. and Galperin, M., Y. (2002) Sequence – Evolution – Function: Computational Approaches in Comparative Genomics, Kluwer, Dordrecht.
62. Kopfmann, S. and Hess, W., R. (2013) Toxin-antitoxin systems on the large defense plasmid pSYSA of Synechocystis sp. PCC 6803. J Biol Chem. 288, 7399-409.
63. Krishna, P., S., RadhaRani, B., Mohan, M., K., Suzuki, I., Shivaji, S., Prakash, J., S., S. (2013) A novel transcriptional regulator, Sll1130 negatively regulates heat responsive genes in *Synechocystis* sp. PCC6803. Biochem. J. 449, 751-760.
64. Kufryk, G., I. and Vermaas, W., F., J. (2001) A novel protein involved in the functional assembly of the oxygen-evolving complex of photosystem II in *Synechocystis* sp. PCC 6803. Biochemistry 40:9247-9255.
65. Kufryk, G., I. and Vermaas, W., F., J. (2006) Sll1717 Affects the Redox State of the Plastoquinone Pool by Modulating Quinol Oxidase Activity in Thylakoids. J. Bacteriol. 188, 1286–1294.
66. Labarre J, Chauvat F & Thuriaux P (1989) Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803. J Bacteriol. 171, 3449–3457.
67. Laule, O., Hirsch-Hoffmann, M., Hruz, T., Gruissem, W. and Zimmermann, P. (2006) Web-based analysis of the mouse transcriptome using Genevestigator. BMC Bioinformatics 7, 311.
68. Lee, S., Prochaska, D., J., Fang, F. and Barnum S., R. (1998) A 16.6-Kilodalton protein in the cyanobacterium Synechocystis sp. PCC 6803 plays a role in the heat shock response. Curr. Microbiol. 37, 403–407.
69. Liberek. K., Lewandowska, A. and Tkiewicz, S., Z. (2008) Chaperones in control of protein disaggregation. EMBO J. 27: 328 – 335.
70. Lillo, F., Krakauer, D., C., (2007) A statistical analysis of the three-fold evolution of genomic compression through frame overlaps in prokaryotes. Biology direct. 2, 22.
71. Los, D., A., Ray, M., K. and Murata, N. (1997) Differences in the control of the temperature-dependent expression of four genes for desaturases in Synechocystis sp. PCC 6803. Mol. Microbiol. 25, 1167–1175.
72. Los, D., A., Suzuki, I., Zinchenko, V., V. and Murata, N. (2008) Stress responses in Synechocystis: regulated genes and regulatory systems. In The Cyanobacteria: Molecular Biology, Genomics and Evolution (Herrero, A., and Flores, E., eds.), pp. 117-157, Horizon Scientific Press, Norfolk, UK.

73. Lund, P., A. (2009) Multiple chaperonins in bacteria – why so many? FEMS Microbiol. Rev. 33, 785–800.
74. Masamoto, K., Misawa, N., Kaneko, T., Kikuno, R. and Toh, H. (1998) β-Carotene hydroxylase gene from the cyanobacterium *Synechocystis* sp.PCC6803. Plant Cell Physiol. 39, 560–564.
75. Masuda, Y., Miyakawa, K., Nishimura, Y. and Ohtsubo, E. (1993) chpA and chpB, Escherichia coli chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. J. Bacteriol. 175, 6850-6856.
76. Mikami, K., Kanesaki, Y., Suzuki, I., and Murata, N. (2002). The histidine kinase Hik33 perceives osmotic stress and cold stress in *Synechocystis* sp. PCC 6803. Mol. Microbiol. 46, 905–915.
77. Nakamoto, H., Suzuki, N. and Roy, S., K. (2000) Constitutive expression of a small heat-shock protein confers cellular thermotolerance and thermal protection to the photosynthetic apparatus in cyanobacteria. FEBS Lett. 483, 169-174.
78. Nakamoto, H., Tanaka, N. and Ishikawa, N. (2001) A novel heat-shock protein plays an important role in thermal stress management in cyanobacteria. J. Biol. Chem. 276, 25088–25095.
79. Nakamoto, H., Suzuki, M. and Kojima, K. (2003) Targeted inactivation of the *hrcA* repressor gene in cyanobacteria. FEBS Lett. 549, 57–62.
80. Narberhaus, F. (1999) Negative regulation of bacterial heat shock genes. Mol. Microbiol. 31, 1-8.
81. Narberhaus, F., Kowarik, M., Beck, C., and Hennecke, H. (1998) Promoter selectivity of the *Bradyrhizobium japonicum* RpoH transcription factors in vivo and in vitro. J Bacteriol. 180, 2395–2401.
82. Ning, D., Ye, S., Liu, B. and Chang, J. (2011)The proteolytic activation of the *relNEs* (*ssr1114/slr0664*) toxin-antitoxin system by both proteases Lons and ClpP2s/Xs of *Synechocystis* sp. PCC 6803. Curr Microbiol. 63, 496-502.
83. Nowaczyk, M., M., Hebeler, R., Schlodder, E., Meyer, H., E., Warscheid, B., and Rogner, M. (2006) Psb27, a cyanobacterial lipoprotein, is involved in the repair cycle of photosystem II. Plant Cell. 18, 3121–3131.
84. Obayashi, T., Kinoshita, K., Nakai, K., Shibaoka, M., Hayashi, S., Saeki, M., Shibata, D., Saito, K. and Ohta, H. (2007) ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. Nucleic Acids Res. 35, 863–869.
85. Okuda, K., Hayashi, H. and Nishiyama, Y. (2005) Systematic characterization of the ADP-ribose pyrophosphatase family in the Cyanobacterium *Synechocystis* sp. strain PCC 6803. J Bacteriol. 187, 4984-4991.
86. 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.
87. Osterman, A. and Overbeek, R. (2003) Missing genes in metabolic pathways: a comparative genomics approach. Curr. Opin. Chem. Biol. 7, 238–251.
88. Paithoonrangsarid, K., Shoumskaya, M., A., Kanesaki, Y., Satoh, S., Tabata, S., Los, D., A., Zinchenko, V., V., Hayashi, H., Tanticharoen, M., Suzuki, I. and Murata, N. (2004) Five histidine kinases perceive osmotic stress and

regulate distinct sets of genes in *Synechocystis.* J. Biol. Chem. 297, 53078–53086.

89. Pandey, D., P. and Gerdes, K., (2005) Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33, 966-976.
90. Pedersen, K., Zavialov, A., V., Pavlov, M., Y., Elf, J., Gerdes, K. and Ehrenberg, M. (2003) The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell. 112, 131–40.
91. Phillips, Z., E. and Strauch, M., A (2002) Bacillus subtilis sporulation and stationary phase gene expression. Cell Mol Life Sci. 59, 392–402.
92. Prakash, J., S., S., Krishna, P., S., Sirisha, K., Kanesaki, Y., Suzuki, I., Shivaji, S. and Murata, N. (2010) An RNA helicase, CrhR, regulates the lowtemperature-inducible expression of heat-shock genes groES, groEL1 and groEL2 in Synechocystis sp. PCC 6803. Microbiology. 156, 442–451.
93. Pueyo, J., J., Alfonso, M., Andrés, C. and Picorel, R. (2002) Increased tolerance to thermal inactivation of oxygen evolution in spinach Photosystem II membranes by substitution of the extrinsic 33-kDa protein by its homologue from a thermophilic cyanobacterium. Biochim Biophys Acta. 1554, 29-35.
94. Raina,S., Missiakas, D. and Georgopoulos, C. (1995) The rpoE gene encoding the σE (σ24) heat shock σ factor of Escherichia coli. EMBO J. 14, 1043-1055.
95. Reznikoff, W., S., Winter, R., B. and Hurley, C., K. (1974) The location of the repressor binding sites in the lac operon. Proc. Natl. Acad. Sci. USA. 71, 2314-2318.
96. Roose, J., L., and Pakrasi, H., B. (2008) The Psb27 protein facilitates manganese cluster assembly in photosystem II. J. Biol. Chem. 283, 4044–4050.
97. Salter, A., H., Virgin, I., Hagman, A. and Andersson, B. (1992) On the molecular mechanism of light-induced D1 protein degradation in photosystem II core particles. Biochemistry. 31, 3990-3998.
98. Sambrook, J., Fritsch, E., F. and Maniatis, T. (1989) Molecular Cloning: a laboratory manual. 2nd ed. NY, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press.
99. Sato, S., Shimoda, Y., Muraki, A., Kohara, M., Nakamura, Y. and Tabata, S. (2007) A Large-scale Protein–protein Interaction Analysis in Synechocystis sp. PCC6803. DNA Research. 14, 207–216.
100. Servant, P., and Mazodier, P. (1996) Heat induction of hsp18 gene expression in *Streptomyces albus* G: Transcriptional and posttranscriptional regulation. J Bacteriol. 178, 7031–7036.
101. Shao, Y., Harrison, E., M., Bi, D., Tai, C., He, X., *et al.*, (2011) TADB: a web-based resource for Type 2 toxinantitoxin loci in bacteria and archaea. Nucleic Acids Res. 39, 606–11.
102. Shipton, C., A. and Barber, J. (1991) Photoinduced degradation of the D1 polypeptide in isolated reaction centers of photosystem II: evidence for an autoproteolytic process triggered by the oxidizing side of the photosystem. Proc. Natl. Acad. Sci. USA. 88, 6691-6695.
103. Shoumskaya, M., A., Paithoonrangsarid, K., Kanesaki, Y., Los, D., A., Zinchenko, V., V., Tanticharoen, M., Suzuki, I. and Murata, N. (2005) Identical Hik-Rre systems are involved in perception and transduction of salt

signals and hyperosmotic signals but regulate the expression of individual genes to different extents in Synechocystis. J. Biol. Chem. 280, 21531-21538.

104. Sirpio, S., Allahverdiyeva, Y., Suorsa, M., Paakkarinen, V., Vainonen, J., Battchikova, N., and Aro, E., M. (2007) TLP18.3, a novel thylakoid lumen protein regulating photosystem II repair cycle.Biochem. J. 406, 415–425.
105. Stanier, R.Y., Kunisawa, R., Mandel, M. and Cohen-Bazire, G. (1971) Purification and properties of unicellular blue-green algae (order Chroococcales). Bacteriol. Rev. 35, 171–205.
106. Summer, E., J., Berry, J., Tran, T., A., Niu, L., Struck, D.K., Young, R., (2007) Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. J. Mol. Bio. 373, 1098-1112.
107. Suzuki, I., Kanesaki, Y., Mikami, K., Kanehisa, M., and Murata, N., (2001) Low temperature-regulated genes under control of the low temperature sensor Hik33 in *Synechocystis*, Mol. Microbiol. 40, 235-244.
108. Suzuki, I., Kanesaki, Y., Hayashi, H., Hall, J., J., Simon, W., J., Slabas, A., R. and Murata, N. (2005) The histidine kinase Hik34 is involved in thermotolerance by regulating the expression of heat shock genes in *Synechocystis*. Plant Physiol. 138, 1409–1421.
109. Suzuki, I., Simon, W., J. and Slabas, A., R. (2006) The heat shock response of Synechocystis sp. PCC 6803 analyzed by transcriptomics and proteomics. J. Exp. Botany. 57, 1573–1578.
110. Tanaka, N. and Nakamoto, H. (1999) HtpG is essential for the thermal stress management in cyanobacteria. FEBS Lett. 458, 117-23
111. Tatusov, R., L., Natale, D., A., Garkavtsev, I., V., Tatusova, T., A., Shankavaram, U., T., Rao, B., S., Kiryutin, B., Galperin, M., Y., Fedorova, N., D. and Koonin, E., V. (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29, 22–28.
112. Telfer, A., De Las Rivas, J. and Barber, J. ((1991) Carotene within the isolated photosystem II reaction center: photooxidation and irreversible bleaching of this chromophore by oxidized P680, Biochim. Biophys. Acta. 1060, 106-114.
113. Thornton, L., E., Ohkawa, H., Roose, J., L., Kashino, Y., Keren, N., and Pakrasi, H., B. (2004) Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem II activity in the cyanobacterium, *Synechocystis* 6803. Plant Cell. 16, 2164–2175.
114. Tian, W. and Skolnick, J. (2003) How well is enzyme function conserved as a function of pairwise sequence identity? J. Mol. Biol. 333, 863–882.
115. Török, Z., Goloubinoff, P., Horváth, I., Tsvetkova, N., M., Glatz, A., Balogh, G., Varvasovszki, V., Los, D., A., Vierling, E., Crowe, J., H. and Vigh, L. (2001) *Synechocystis* HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding. Proc. Natl. Aca. Sci. USA. 98, 3098–3103.
116. Tsilibaris, V., Maenhaut-Michel, G., Mine, N. and Van Melderen, L. (2007) What is the benefit to *Escherichia coli* of having multiple toxin-antitoxin systems in its genome? J. Bacteriol. 189, 6101-6108.

117. Tuominen, I., Pollari, M., Tyystjärvi, E. and Tyystjärvi, T. (2006) The SigB σ factor mediates high-temperature responses in the cyanobacterium *Synechocystis* sp. PCC6803. FEBS Lett. 580, 319–323.
118. Tuominen, I., Pollari, M., Wobeser, E. A. V., Tyystjärvi, E., Ibelings, B.W., Matthijs, H.C. and Tyystjärvi, T. (2008) Sigma factor SigC is required for heat acclimation of the cyanobacterium *Synechocystis* sp. strain PCC 6803. FEBS Lett. 582, 346-350.
119. Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., Patton, D. and Meinke, D. (2003) The Arabidopsis SeedGenes Project. Nucleic Acids Res. 31, 90–93.
120. Van Asseldonk, M., Simons, A., Visser, H., de Vos, W.M. and Simons, G. (1993) Cloning, nucleotide sequence, and regulatory analysis of the *Lactococcus lactis dnaJ* gene. J. Bacteriol. 175**,** 1637–1644.
121. Vaughn, J. ,L., Feher, V., Naylor, S., Strauch, M., A. and Cavanagh, J. (2000) Novel DNA binding domain and genetic regulation model of Bacillus subtilis transition state regulator abrB. Nat Struct Biol. 7, 1139–1146.
122. Versteeg, S., Mogk, A. and Schumann, W. (1999) The *Bacillus subtilis* htpG gene is not involved in thermal stress management. Mol. Gen. Genet. 261, 582-588.
123. Volkmer, T., Schneider, D., Bernát, G., Kirchhoff, H., Wenk, S.O and Rögner, M. (2007) Ssr2998 of *Synechocystis* sp. PCC 6803 is involved in regulation of cyanobacterial electron transport and associated with the Cytochrome b6/fComplex. J. Bio. Che. 282, 3730 –3737.
124. Wada, H. and Murata, N. (1989) *Synechocystis* PCC 6803 mutants defective in desaturation of fatty acids. Plant Cell Physiol. 30, 971–978.
125. Webb, R., Sherman, L., A. (1994) The cyanobacterial heat-shock response and the molecular chaperones. In DA Bryant eds, The molecular biology of cyanobacteria, Kluwer, Dordrecht pp 751–767.
126. Weiner, L., J., Brissette, L. and Model, P. (1991) Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms. Genes Dev. 5, 1912 –1923.
127. Winston, J., L., Toth, S., I., Roe, B., A. and Dyer, D.,W. (1996) Cloning and characterization of the Actinobacillus actinomycetemcomitans gene encoding a heat-shock protein 90 homologue. Gene. 179, 199-204.
128. Yamaguchi, Y. and Inouye, M., (2009) mRNA interferases, sequence-specific endoribonucleases from the toxin antitoxin systems. Prog. Mol. Biol. Transl. Sci. 85, 467–500.
129. Yamaguchi, Y., Park , J., H. and Inouye, M. (2009) MqsR, a crucial regulator for quorum sensing and biofilm formation, is a GCU-specific mRNA interferase in Escherichia coli. J. Biol. Chem. 284, 28746–53.
130. Yamaguchi, Y., Park, J., H. and Inouye, M. (2011) Toxin-Antitoxin Systems in Bacteria and Archaea. Annu. Rev. Genet. 45, 61–79.
131. Yao, F. and Strauch, M., A. (2005) Independent and interchangeable multimerization domains of the AbrB, Abh, and SpoVT global regulatory proteins. J Bacteriol. 187, 6354–6362.

132. Ye, R. W., Tao, W., Bedzuk, L., Young, T., Chen, M. and Li, L., (2000) Global gene expression profiles of Bacillus subtilis grown under anaerobic conditions. J. Bacteriol. 182, 4458–4465.
133. Yu, J., Shen, G., Wang, T., Bryant, D., A., Golbeck, J., H. and McIntosh, L. (2003) Suppressor mutations in the study of photosystem I biogenesis: sll0088 is a previously unidentified gene involved in reaction center accumulation in Synechocystis sp. strain PCC 6803. J Bacteriol. 185, 3878-87.
134. Yuan, G. and Wong, S., L. (1995) Regulation of groE expression in Bacillus subtilis: the involvement of the sigma A-like promoter and the roles of the inverted repeat sequence (CIRCE). J. Bacteriol. 177, 5427–5433.
135. Zhang, Y. and Inouye, M. (2009) The inhibitory mechanism of protein synthesis by YoeB, an Escherichia coli toxin. J. Biol. Chem. 284, 6627–38.
136. Zheng, M., Wang, X., Templeton, L., J., Smulski, D., R., LARossa, R., A. and Storz, G. (2001) DNA Microarray-mediated transcriptional profiling of the Escherichia coli response to hydrogen peroxide. J. Bacteriol. 183, 4562–4570.
137. Zinchenko, V., V., Piven, I., V., Melnik, V., A. and Shestakov, S., V. (1999). Vectors for the complementation analysis of cyanobacterial mutants.Russ J Genet. 35, 228–232.
138. Zuber, U. and Schumann, W. (1994) CIRCE, a novel heat shock element involved in regulation of heat shock operon dnaK of Bacillus subtilis. J. Bacteriol. 176, 1359–1363.

List of Publications

1. A novel transcriptional regulator, Sll1130 negatively regulates heat responsive genes in *Synechocystis* sp. PCC6803, **Krishna PS**, RadhaRani B, Mohan MK, Suzuki I, Shivaji S, Prakash JSS. **Biochem. J. (2013)** 449 (751-760).

2. RNA helicase, CrhR is indispensable for the energy redistribution and the regulation of photosystem stoichiometry at low temperature in *Synechocystis* sp. PCC6803, Sireesha K, Radharani B, **Krishna PS**, Sreedhar N, Subramanyam R, Mohanty P and Prakash JSS. **BBA, Bioenergetics (2012)** 1817, 1525–1536.

3. Sensing and molecular responses to low temperature in cyanobacteria, Prakash JSS, **Krishna PS**, Shivaji S. **(2012)** (chapter No-7) edited by Srivastava, Rai and Neilan, **Taylor and Francis** group; London.

4. An RNA helicase, CrhR, regulates the low-temperature-inducible expression of heat-shock genes *groES*, *groEL1* and*groEL2* in *Synechocystis* sp. PCC 6803, Prakash JSS, **Krishna PS**, Sirisha K, Kanesaki Y, Suzuki I, Shivaji S and Murata N. **Microbiology (2010)** 156, 442–451.

List of Poster presentations

1. A genomic approach to understand the functional diversity of RNA helicases. **Bioconvene-2007**, 19-22 Dec 2007. University of Hyderabad.

2. Delineating the mechanisms of salt and high pH conditions at molecular level in halo-alkali tolerant bacteria. **International symposium in phycological research-2010**, 25-27 Feb 2010. Banaras Hindu University.

3. Hypothetical protein, *Sll1130* is a heat responsive negative regulator of a dicistronic operon in the cyanobacterium *Synechocystis* sp. PCC6803. **ICPSPGE -2011**, 17-19 Feb 2011. Sambalpur University.

4. Hypothetical proteins to unknown cellular mechanisms: with emphasis on Sll1130, a novel heat responsive transcription factor of heat shock genes. **DFG-Germany 2012**, 22nd Feb 2012. DFG-Germany and University of Hyderabad.

5. Hypothetical protein, *Ssl2245* in *Synechocystis* sp. PCC6803 is a stress responsive growth regulator. **Indo-German symposium on Systems Biology (2012),** 27-29 Nov 2012. University of Hyderabad.

"Anyone who has never made a mistake has never tried anything new"

- Albert Einstein

"We learn from failure,

not from success"

"Success does not consist in never making mistakes but in never making the same one a second time"

— George Bernard Shaw

“I will not say I failed 1000 times, I will say that I discovered there are 1000 ways that can cause failure”

— Thomas Edison