

Transformation and Expression of a Na⁺- Dependent High Affinity Bicarbonate Transporter SbtA in *Nicotiana* Chloroplasts

THESIS

Submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

VANDANA

Under the Supervision of
Dr. Sandhya Mehrotra



BITS Pilani
Pilani | Dubai | Goa | Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

2018

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE,
PILANI**

CERTIFICATE

This is to certify that the thesis entitled “**Transformation and Expression of a Na⁺-Dependent High Affinity Bicarbonate Transporter SbtA in *Nicotiana Chloroplasts***” and submitted by **Vandana**, ID No. **2012PHXF0415P** for award of Ph.D. degree of the Institute embodies original work done by her under my supervision.

Signature of the Supervisor:

Name (Supervisor) : SANDHYA MEHROTRA, Ph.D.

Designation : Assistant Professor,

Department of Biological Sciences,

BITS Pilani, Pilani Campus (from November 2008 to
December 2017)

Present Address (BITS, Pilani, KK Birla Goa Campus)

Date

Abstract

“Enhancing plant productivity is the need of the hour”

Increasing yield potential of crop plants to meet burgeoning demand for food is challenging under the changing environmental conditions. Developing crop varieties which possess the capability of utilizing more CO₂ for photosynthesis and achieving higher food production is an urgent requirement. Furthermore, the expected increase in yields of C₃ crops due to the increasing atmospheric CO₂ concentrations is much less than anticipated due to a phenomenon known as photosynthesis acclimation. Therefore, it is high time to explore all the possibilities for improving C₃ photosynthesis.

Improving Rubisco catalysis, increasing the rate of Ribulose-1, 5-bisphosphate (RuBP) regeneration and lowering the photorespiratory fluxes are few of the strategies tried till date. Major thrust area includes the introduction of highly efficient photosynthetic components known as carbon concentrating mechanisms (CCMs), from C₄ or CAM plants or from certain lower organisms such as algae and cyanobacteria to the C₃ photosynthetic system. For incorporation of a functional CCM into C₃ plants, some distinct features such as a carboxysomal structure and inorganic carbon uptake mechanisms need to be replicated in a C₃ plant chloroplast. It has been assumed that out of all these features, transporters are capable of enhancing photosynthesis independently of others.

The present study envisages introduction of one of the simplest components of the cyanobacterial HCO₃⁻ transport system viz., SbtA to the chloroplasts of model C₃ plants. SbtA transporter has a low K_m and is therefore presumed to be more effective in reducing the CO₂ compensation point and improving the water use efficiency (WUE) under dry-air conditions more efficiently in comparison to other transporters. In order to attain the desired objective, initially *in-silico* studies were carried out to find out the distribution of inorganic carbon transporters in cyanobacteria and also to predict the mode of their evolution. As, *sbtA* gene in some lower organisms exists in an operon along with another gene known as *sbtB*, efforts were also made to study the co-evolution of both the genes.

Since the major objective of the study was expression of SbtA transporter in model plants, suitable genetic constructs were synthesized by fusing *sbtA* transporter gene from

Synechococcus elongatus PCC 7942 with transit peptide sequences (TP) from chloroplast inner envelope proteins of *Arabidopsis thaliana*. For this, several proteins localized to chloroplast inner envelope of *A. thaliana* were screened and TNaXTP (TP from Bile acid: sodium symporter) and TMDTP (TP from multidrug transporter) were shortlisted as final candidates. The fusion genes were initially cloned into intermediate vectors and then into plant expression vectors (*pRII01-AN* and *pCAMBIA1302* with *gus* and *mgfp5* reporter genes respectively). Preliminary studies were carried out by transforming *Nicotiana tabacum* with *gus-pRII01-AN* constructs and transient protein expression was confirmed by GUS reporter gene assays (histochemical and fluorometric) as well as by gene specific PCR analysis.

Transient expression of the gene was studied by transforming *Nicotiana benthamiana* by agroinfiltration using recombinant *pCAMBIA1302* constructs. Expression of the transgene in *N. benthamiana* was confirmed at DNA, RNA and protein levels and subcellular localization of the targeted protein was checked by confocal microscopy. Confocal images of the protoplasts reveal chloroplastic location of the protein in both the constructs. One of the constructs even depicted targeting to the chloroplast envelope membrane. The constructs depicting proper targeting to the chloroplasts were further used to generate stable plants by *Agrobacterium* mediated co-culture method.

Thus, the present work provides a preliminary yet valuable basis for future manipulations of chloroplast function for improving plant photosynthesis.

Declaration

I hereby declare that this Ph.D. thesis entitled “**Transformation and Expression of a Na⁺-Dependent High Affinity Bicarbonate Transporter SbtA in *Nicotiana Chloroplasts***” is my original work which was carried out for the degree of Doctor of Philosophy under the supervision of Dr. Sandhya Mehrotra at Birla Institute of Technology and Science, Pilani, Pilani campus, India. I have duly acknowledged all the sources of information used in the work presented in the thesis.

Signature:

Name: Vandana

Date:

Acknowledgements

I am eternally beholden by the unrestrained and boundless blessing of the almighty by virtue of which I have been able to complete my Ph.D efficaciously and capable to present this modest piece of work, for which I am evermore obligated.

Ardently and humbly, I appreciate the unpretentious cooperation, encouragement and cordiality offered to me by my supervisor, Dr. Sandhya Mehrotra, right from the commencement of my research work to ship-shaping of the thesis. I have been amazingly fortunate to have an advisor like her, who gave me the freedom to explore on my own, and at the same time providing me the guidance to recover when my steps faltered. As my teacher and mentor, she has taught me more than I could ever give her credit for. Her expedient recommendations, constructive criticism and assiduous guidance will be forever cherished. Thank you, dear mam for this beautiful journey and I eagerly look forward to future expeditions down the lane.

I wish to express my earnest gratitude to the members of my Doctoral Advisory Committee, Prof. Rajesh Mehrotra and Prof. Shibasish Chowdhury for their ever willing help and constant incentives, rendered at all stages of this endeavour and also for their time and effort in checking the thesis in short time period. Special thanks to Prof. Rajesh Mehrotra for his ceaseless appreciation and for providing a pleasant and homely environment during my stay in Pilani. He always had kind concern and consideration regarding my academics and research. Thank you, sir.

With interminable delight, I extend my indebtedness and deep sense of gratitude to Director Prof A.K. Sarkar, Dean, Academic & Research Division Prof. S.K. Verma BITS-Pilani, Pilani Campus, for infrastructural support.

Words are inadequate in offering my thanks to present and former heads of the Department of Biological Sciences, Prof. Prabhat Nath Jha, Prof. Rajesh Mehrotra and Prof. Jitendra Panwar for their unswerving help and providing the resources in the department which paved path for successful accomplishment of present study. I am indebted to DRC committee members Prof. Jitendra Panwar, Prof. Ashis Kumar Das, Prof. Shibasish Chowdhary, Prof. Lalita Gupta and Dr. Sandhya Mehrotra for their time to time motivation

and for drafting policies in support of students where and when necessary. I owe my deepest gratitude to all the faculty members, Department of Biological Sciences who supported me in many respects. Their insightful comments and constructive criticisms at different stages of my research were thought-provoking and helped me to focus my ideas.

I wish to express my gratefulness to my dear seniors, Dr. Purva, Dr. Gurpreet and Panchsheela Mam for their constant support and valuable assistance throughout the research work. Their unparalleled love and encouragement will go a long way throughout my life. Special thanks to Dr. Gagandeep, Dr. Zarna Pala and Dr. Bhoopati for their kind support and guidance.

Collective and individual acknowledgments are also owed to my lab mates and colleagues Nisha Jangir, Asha Jhakar, Vidushi Asati, Chetna Sangwan and Sneha Chauhan for their support and co-operation. They helped in maintaining a jovial atmosphere in the lab and boosting up my moral when I was low in spirits. Heartfelt thanks to Monika, Poonam, Sandeep, Vikram, Vikas, Shahid, Shraddha, Leena, Shubhra, Manohar, Divya Mam and Parva Sir. I will always be short of words to express my thanks to Tania, Zaiba and Pinky for the help and moral support and also for creating a family like atmosphere during the course of my stay in Pilani. I also thank Stuti Singh for helping me in carrying out experiments and providing inputs from a new perspective which helped a lot in this thesis.

It's pleasure to acknowledge the support and help provided by Mr. Naresh Kumar Saini, Mr. Mukesh, Mr. Manoj Kumar Mr. Raghuveer and Mr. Mahipal. Heartfelt thanks to Ajay ji and Subhash ji for always rendering a helping hand whenever I needed help related to any experiment or instruments.

I am thankful to Prof. G. Dean Price (ANU, Canberra, Australia) for providing us the SbtA transporter constructs and to Dr. Kundan Kumar (BITS Pilani K.K. Birla Goa campus) for providing us the plant expression vector. I feel extremely grateful to Dr. Siddharth Tiwari (NABI, Mohali) for allowing me to perform the particle bombardment experiments in his lab and Anshu Alok for assisting while performing the experiment. My sincerest thanks to Dr. Sneha Lata Singla Pareek (ICGEB, New Delhi) for permitting me to work in her lab for protoplast isolation experiment and Brijesh for helping in using the lab facilities and instruments.

Diction is not enough to express my gratitude and affection to my beloved father Mr. Anand Prakash Tomar and mother Mrs. Saroj Tomar for bringing me up in the best of ways, for rendering me the best of education, for nurturing in me the best of ideals and for helping me to see the best of times. They are the ultimate role models, whose love and guidance are with me in whatever I pursue. I would like to thank my in-laws, Mr. Jagdish Singh Deopa and Mrs. Sudha Deopa as well as my brother-in-law, Chetan Deopa for their words of encouragement and support. There is no match to express my love for my lovely sister Sneha and brother Siddharth who made my life the happiest.

I dedicate this thesis to my beloved husband Mr. Gaurav Deopa, whose role in my life is and will always remain immense. For, it was because of his faith and confidence in my abilities, that I kept working through the hurdles and failures that I had faced during this entire time. He continually inspires me with his strength of character and spirits. He is the love of my life.

My study needs special acknowledgement to the DST-INSPIRE for providing financial assistance in the form of fellowship and to BITS Library, for giving me the best of the knowledge and resources for my course of study.

Last but not the least, I thank my alma mater, BITS Pilani for providing me the opportunity to be a part of this esteemed institute and providing a congenial environment during my stay in the campus. I feel so empowering and spirited on being hailed as a BITSian. It was undoubtedly a magical experience residing here.

At the end, but till the end of my life, humble thanks to the omnipresent GOD for answering my prayers for giving me the strength to plod on despite my constitution wanting to give up and throw in the towel, thank you so much Dear Lord.

Date:

Vandana

CONTENTS

List of Tables	i-iii
List of Figures	iv-ix
Abbreviations	x- xiv
Key Words	xv
Chapter 1: Introduction and Review of Literature	1-1 to 1-39
Chapter 2: Materials and Methods	2-1 to 2-54
Chapter 3: Results and Discussions	3-1 to 3-90
Chapter 4: Conclusions and future prospects	4-1 to 4-4
Summary	S-1 to S-3
Bibliography	B-1 to B-15
Appendices	
List of publications	A-1
List of conferences and workshops attended	A-2 to A-3
Biography of supervisor	A-4
Biography of candidate	A-5
Reprint of publications	

List of Tables

Table No.	Table Title	Page No.
Table 1-1	Regulators of CCM in cyanobacteria, proteobacteria and algae, their function and their co-regulatory metabolites	1-19
Table 2-1	The list of the cyanobacterial strains used for the analysis of distribution of <i>cmpABCD</i> , <i>ndhD₃/ndhF₃/cupA</i> and <i>ndhD₄/ndhF₄/cupB</i> operons	2-1
Table 2-2	The name, phyla and the taxonomic IDs of the organisms used for the analysis of distribution of <i>sbtAB</i> operon	2-3, 2-4
Table 2-3	List of primers used in the study for the amplification of transit peptide sequences, <i>sbtA</i> gene, fusion constructs and analysis of putative transformants	2-11
Table 2-4	The working concentrations of the various PCR components used in the study	2-12
Table 2-5	The PCR conditions standardized for various amplifications	2-12
Table 2-6	The composition of ligation reaction along with recommended controls for TA cloning	2-17
Table 2-7	The details of the restriction enzymes used for cloning of various DNA fragments	2-25
Table 2-8	The composition of reaction mixture used for setting up restriction reactions in the study	2-25
Table 2-9	The composition of MS medium and the concentrations of different stock solutions used for preparing MS media	2-33
Table 2-10	Composition of phytohormones stock solutions used in the study	2-34
Table 2-11	The selected media for tissue culture experiments	2-42
Table 2-12	The observation parameters for tissue culture responses of <i>N. tabacum</i> leaf discs on different compositions of MS medium	2-43

Table 2-13	The composition of reverse transcription reaction mixture used for cDNA synthesis	2-47
Table 2-14	PCR conditions for diagnostic PCRs performed using genomic DNA and cDNA as template	2-48
Table 2-15	Composition of PAGE gel used in the study	2-50
Table 3-1	Distribution of inorganic carbon transporters (CO ₂ and HCO ₃ ⁻ uptake systems) in α- and β-cyanobacteria	3-3
Table 3-2	The gene symbol, function, nucleotide and amino acid length of <i>sbtA</i> and <i>sbtB</i> genes in the seventy eubacterial phyla used for the analysis	3-16 to 3-20
Table 3-3	The correlation values between <i>sbtA</i> , <i>sbtB</i> and <i>sbtAB</i> operon genes obtained by Mirrortree server	3-21
Table 3-4	Analysis of TM helices in SbtA transporter protein using several topology prediction tools	3-25
Table 3-5	Proteins shortlisted from list of inner envelope membrane proteins provided in the AT_CHLORO database	3-36
Table 3-6	Prediction of the length and the presence of chloroplast transit peptide (cTP) cleavage sites in the shortlisted proteins by using ChloroP 1.1 Prediction Server	3-38
Table 3-7	TargetP 1.1 Server results showing the frequency of presence of chloroplastic transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP) in the proteins along with their location inside the cell	3-39
Table 3-8	Prediction of length and number of transmembrane helices, membrane protein topology and signal peptides using TOPCONS, TMHMM and OCTOPUS servers	3-40
Table 3-9	Details of the transit peptide, their source protein and its function retrieved from TAIR database	3-41

Table 3-10	Caulogenesis percentage, number of shoots and length of shoots in <i>N. tabacum</i> leaf explant inoculated on MS medium supplemented with four different concentrations of BAP (0.5, 1.0, 1.5 and 2.0 mg/L)	3-59
Table 3-11	Caulogenesis percentage, number of shoots and length of shoots in <i>N. tabacum</i> leaf explant inoculated on MS medium supplemented with constant BAP (2.0 mg/L) and varying NAA (0.05, 0.1, 0.15 and 0.2 mg/L) concentrations	3-62
Table 3-12	Rhizogenesis percentage, number of roots per shoot and length of roots per shoot in <i>N. tabacum</i> leaf explant inoculated on MS medium supplemented with different concentrations of IBA (0.5, 1.0, 1.5 and 2.0 mg/L)	3-64

List of Figures

Figure No.	Figure Title	Page No.
Figure 1-1	The representation of a cyanobacterial CCM and its components	1-2
Figure 1-2	The topology of SbtA transporter from <i>Synechococcus elongatus</i> PCC 7942	1-4
Figure 1-3	The taxonomic profile of <i>sbtA</i> from eggnog database	1-5
Figure 1-4	The activators and repressors of CCM in cyanobacteria	1-21
Figure 1-5	The representation of CCM in algae (<i>Chlamydomonas reinhardtii</i>)	1-24
Figure 1-6	The representation of CCM in diatoms (<i>Phaeodactylum tricornutum</i>)	1-26
Figure 2-1	A schematic depicting the total number of <i>A. thaliana</i> chloroplast proteins and their sub-plastidial localization present on the AT_CHLORO database	2-6
Figure 2-2	<i>A. thaliana</i> , <i>N. benthamiana</i> and <i>N. tabacum</i> plants grown in plant tissue culture room under 16/8 hrs of photoperiod	2-8
Figure 2-3	The agarose gel electrophoresis image for the genomic DNA (gDNA) isolated from <i>A. thaliana</i> leaves	2-10
Figure 2-4	A schematic depicting cloning of <i>sbtA</i> gene into <i>pGEM®-T</i> vector	2-19
Figure 2-5	A schematic depicting cloning of transit peptide sequences (TNaXTP/TMDTP) into <i>pGEM®-T</i> vector	2-20
Figure 2-6	A schematic depicting cloning of transit peptide sequences (TNaXTP/TMDTP) and <i>sbtA</i> gene into <i>pColdI</i> vector	2-21
Figure 2-7	A schematic depicting cloning of <i>gus</i> gene from <i>pBI101</i> vector into <i>pRI101-AN</i> vector	2-22
Figure 2-8	A schematic depicting cloning of fusion genes (TNaXTP/TMDTP- <i>sbtA</i> gene) upstream of the <i>gus</i> gene into <i>pRI101-AN</i> vector	2-23

Figure 2-9	A schematic depicting cloning of fusion genes (TNaXTP/TMDTP- <i>sbtA</i>) into <i>pCAMBIA1302</i> vector	2-24
Figure 2-10	The growth stages of <i>N. tabacum</i> plantlets grown on MS basal medium	2-35
Figure 2-11	Standard curve for the relative fluorescence of 4-methylumbelliferone (4-MU) standards	2-39
Figure 3-1	Arrangement of the genes encoding the cyanobacterial Ci transporters	3-6, 3-7
Figure 3-2	Phylogenetic tree of <i>cmpABCD</i> operon (BCT1) in sixteen cyanobacterial species	3-8
Figure 3-3	Phylogenetic tree of <i>ndhI₃</i> operon in eighteen cyanobacterial species	3-9
Figure 3-4	Phylogenetic tree of <i>ndhI₄</i> operon in twenty one cyanobacterial species	3-10
Figure 3-5	The alignment of protein sequence from <i>Synechocystis</i> PCC 6803, <i>Synechococcus elongatus</i> PCC 7942, <i>Prochlorococcus marinus</i> MED4 and <i>Prochlorococcus marinus</i> str. MIT 9312	3-22
Figure 3-6	Phylogenetic tree for <i>sbtAB</i> operon conducted in MEGA6	3-23
Figure 3-7	The amino acid alignment of SbtA from <i>Synechocystis</i> sp. PCC 6803 and <i>Synechococcus elongatus</i> PCC 7942	3-26
Figure 3-8	Secondary Structure predictions for SbtA protein using PSIPRED	3-27, 3-28
Figure 3-9	HHpred results for SbtA from <i>Synechococcus elongatus</i> PCC 7942	3-29
Figure 3-10	The protein model for SbtA protein from <i>Synechococcus elongatus</i> PCC 7942, generated by threading using RaptorX software	3-30
Figure 3-11	Validation of 3-dimensional structure of SbtA using Ramachandran Plot	3-31
Figure 3-12	Validation of three dimensional structure of SbtA protein using ERRAT score	3-31

Figure 3-13	Secondary structure of SbtB from <i>Synechococcus elongatus</i> PCC 7942 predicted by Phyre ² server	3-33
Figure 3-14	The sequence logo for SbtB sequences from 20 cyanobacterial species provided in the Kazusa cyanobase	3-33
Figure 3-15	Relative adaptiveness plots for each amino acid of SbtA transporter with the genome of <i>Synechocystis sp.</i> PCC 6803	3-43
Figure 3-16	Relative adaptiveness plots for each amino acid of SbtA transporter with the genome of <i>A. thaliana</i>	3-44
Figure 3-17	Relative adaptiveness plots for each amino acid of SbtA transporter with the genome of <i>N. tabacum</i>	3-45
Figure 3-18	Vector map of <i>sbtA</i> - <i>pGEMT</i> clone	3-47
Figure 3-19	PCR products of TNaXTP, TMDTP and <i>sbtA</i> gene	3-47
Figure 3-20	Recombinant vector maps and restriction digestion analysis to confirm cloning of (A) <i>sbtA</i> in <i>pGEM®-T</i> vector (B) TNaXTP in <i>pGEM®-T</i> vector and (C) TMDTP in <i>pGEM®-T</i> vector	3-48, 3-49
Figure 3-21	Recombinant vector maps and restriction analysis of recombinant <i>pColdI</i> clones (TNaXTP/TMDTP- <i>sbtA</i> - <i>pColdI</i>)	3-50
Figure 3-22	Recombinant vector map and restriction analysis for confirmation of cloning of <i>gus</i> gene into <i>pRI101-AN</i> vector	3-52
Figure 3-23	Schematic representation of the recombinant <i>gus</i> - <i>pRI101-AN</i> vector with TNaXTP/TMDTP- <i>sbtA</i> insert engineered for plant transformation	3-53
Figure 3-24	Recombinant vector map and restriction analysis of recombinant TNaXTP- <i>sbtA</i> - <i>gus</i> - <i>pRI101-AN</i> clone	3-53
Figure 3-25	Recombinant vector map and restriction analysis of recombinant TMDTP- <i>sbtA</i> - <i>gus</i> - <i>pRI101-AN</i> clone	3-54
Figure 3-26	Fluorometric quantification of GUS activity (Mean \pm SD). Expression of MU (Methylumbelliferone) in control (wild type) and TNaXTP- <i>sbtA</i> - <i>gus</i> - <i>pRI101-AN</i> and TMDTP- <i>sbtA</i> - <i>gus</i> - <i>pRI101-AN</i> plasmid bombarded samples	3-55
Figure 3-27	Transformation of <i>A. thaliana</i> by floral dip method	3-57

Figure 3-28	Transfer of antibiotic resistant plantlets of <i>A. thaliana</i> (T ₁ generation) in soilrite mix	3-58
Figure 3-29	Transfer of antibiotic resistant plantlets of <i>A. thaliana</i> (T ₂ generation) in soilrite mix	3-58
Figure 3-30	Effect of MS media supplemented with four different concentrations of BAP on frequency of caulogenesis in <i>N. tabacum</i>	3-60
Figure 3-31	Effect of MS media supplemented with four different concentrations of BAP on number of shoots per culture in <i>N. tabacum</i>	3-60
Figure 3-32	Effect of MS media supplemented with four different concentrations of BAP on length of shoots in <i>N. tabacum</i>	3-61
Figure 3-33	Effect of BAP and NAA interaction on frequency of caulogenesis in <i>N. tabacum</i>	3-62
Figure 3-34	Effect of BAP and NAA interaction on number of shoots per culture in <i>N. tabacum</i>	3-63
Figure 3-35	Effect of BAP and NAA interaction on length of shoots in <i>N. tabacum</i>	3-63
Figure 3-36	Effect of MS basal and MS media supplemented with four different concentrations of IBA on frequency of rhizogenesis in <i>N. tabacum</i>	3-65
Figure 3-37	Effect of MS basal and MS media supplemented with four different concentrations of IBA on number of roots/plantlet in <i>N. tabacum</i>	3-65
Figure 3-38	Effect of MS basal and MS media supplemented with four different concentrations of IBA on number of roots/plantlet in <i>N. tabacum</i>	3-66
Figure 3-39	GUS histochemical staining of callus tissue regenerated from agro-infected <i>N. tabacum</i> leaves	3-67
Figure 3-40	<i>Agrobacterium</i> mediated co-culture method for regeneration of stable <i>N. tabacum</i> plants	3-68

Figure 3-41	PCR analysis of the transformed plants regenerated by Agroinfection of <i>N. tabacum</i> by recombinant <i>pRI101-AN</i> constructs (TNaXTP/TMDTP- <i>sbtA-gus-pRI101-AN</i>)	3-69
Figure 3-42	The agarose gel electrophoresis image of the amplification of the fusion genes (<i>sbtA-TNaXTP/TMDTP</i>) by using <i>pColdI</i> recombinant plasmids as template	3-70
Figure 3-43	Recombinant vector map and restriction analysis for confirmation of the cloning of fusion constructs (TNaXTP/TMDTP- <i>sbtA</i>) into <i>pCAMBIA1302</i> vector	3-71
Figure 3-44	Schematic representation of the recombinant <i>pCAMBIA1302</i> vector with TNaXTP/TMDTP- <i>sbtA</i> insert engineered for plant transformation	3-72
Figure 3-45	Agroinfiltration of <i>N. benthamiana</i> leaves with <i>Agrobacterium</i> cells harboring fusion constructs (TNaXTP/TMDTP- <i>sbtA-pCAMBIA1302</i>)	3-73
Figure 3-46	Genomic DNA isolated from wild type (WT) and Agroinfiltrated samples	3-74
Figure 3-47	The agarose gel images representing PCR products amplified from genomic DNA of agroinfiltrated leaves of <i>N. benthamiana</i>	3-75
Figure 3-48	The agarose gel images representing PCR products amplified from genomic DNA of agroinfiltrated leaves of <i>N. benthamiana</i> .	3-76
Figure 3-49	Total RNA isolated from wild type (WT) and Agroinfiltrated samples	3-77
Figure 3-50	Agarose gel electrophoresis image showing absence of PCR products in RNA isolated from <i>N. benthamiana</i> leaves infiltrated with TNaXTP/TMDTP- <i>sbtA-pCAMBIA1302</i> recombinant constructs	3-78
Figure 3-51	The agarose gel images representing PCR products amplified from cDNA of Agroinfiltrated leaves of <i>N. benthamiana</i>	3-79
Figure 3-52	Total protein isolated from wild type and Agroinfiltrated leaf samples of <i>N. benthamiana</i>	3-80

Figure 3-53	Western blot analysis of total protein isolated from WT (wild type) and Agroinfiltrated leaves of <i>N. benthamiana</i>	3-82
Figure 3-54	Confocal images of <i>N. benthamiana</i> protoplasts isolated from wild type and agroinfiltrated leaves	3-84, 3-85
Figure 3-55	Non-transformed leaf discs of <i>N. tabacum</i> cultured on shoot induction medium containing different concentrations of hygromycin	3-86
Figure 3-56	Effect of hygromycin concentration on regeneration frequency of non-transformed leaf explants grown on shoot elongation medium	3-87
Figure 3-57	GFP fluorescence in transgenic shoots generated using TNaXTP- <i>sbtA-pCAMBIA1302</i> constructs monitored using Zeiss AXIO SCOPE A1 Microscope	3-87
Figure 3-58	Growth stages of shoots directly regenerated from leaf explant on MS media containing BAP (2 mg/L) + NAA (0.2 mg/L) + Hygromycin (40 mg/L) + Cefotaxime (250 mg/L)	3-88
Figure 3-59	PCR analysis of the transformed plants regenerated by Agro-infection of <i>N. tabacum</i> by recombinant <i>pCAMBIA1302</i> constructs (TNaXTP/TMDTP- <i>sbtA-pCAMBIA1302</i>)	3-89

ABBREVIATIONS

μ L	Microliter
μ m	Micrometer
2, 4-D	2, 4-Dichlorophenoxyacetic acid
ABC	ATP Binding Cassette
AMQW	Autoclaved Milli-Q Water
ANOVA	Analysis of Variance
APS	Ammonium per sulfate
ATP	Adenosine triphosphate
BAP	6-Benzylaminopurine
BIC	Bayesian Information Criterion
BLAST	Basic Local Alignment Search Tool
BMC	Bacterial microcompartment
bp/kb	Base pair/kilo base pair
BPB	Bromophenol blue
BSA	Bovine Serum Albumin
bZIP	Basic Leucine Zipper Domain
C.D.	Critical Difference
CA	Carbonic anhydrase
CAI	Codon Adaptation Index
CAM	Crassulacean Acid Metabolism
cAMP	Cyclic adenosine monophosphate
CBB	Calvin Benson Bassham
CBB	Coomassie Brilliant Blue
CCM	Carbon Concentrating Mechanism
CCRE	CO ₂ -cAMP-Responsive <i>cis</i> -elements
CD	Conserved Domain
cDNA	Complementary DNA

Ci	Inorganic Carbon
CO ₂	Carbon Dioxide
CRD	Completely Randomized Design
CREB	cAMP response element binding
CS	Cleavage Site
cTP	Chloroplast Transit Peptide
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
DOPE	Discrete Optimized Protein Energy
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
FBPase	Fructose-1, 6-bisphosphatase
GCUA	Graphical Codon Usage Analyser
gDNA	Genomic DNA
GFP	Green Fluorescent Protein
GOE	Great oxidation event
GUS	β-glucuronidase
HCl	Hydrogen Chloride
HCO ₃ ⁻	Bicarbonate ion
HMM	Hidden Markov Model
hr/hrs	Hour/hours
HRP	Horseradish peroxidase
HTH	Helix-turn-helix
IBA	Indole-3-butyric acid
IEM	Inner Envelope of Chloroplast
kDa	Kilodalton
LB	Luria Bertani
LC-MS	Liquid chromatography–mass spectrometry

LTTR	LysR Type Transcriptional Regulators
MCL	Maximum composite likelihood
MCS	Multiple Cloning Site
MEGA	Molecular Evolutionary Genetics Analysis
MEME	Multiple EM for motif elicitation
MES-K	2-(N-morpholino) ethanesulfonic acid
mg	Milligram
mGFP5	Modified Green Fluorescent Protein 5
min	Minute/minute
mL	Millilitre
ML	Maximum Likelihood
mM	Millimolar
mRNA	Messenger RNA
MS	Murashige and Skoog
MSA	Multiple Sequence Alignment
MU	4-Methyl Umbelliferone
MUG	4-methylumbelliferyl- β -D-glucuronide
NAA	Naphthalene acetic acid
NAD-ME	NAD-dependent malic enzyme
NADP	Nicotinamide Adenine Dinucleotide Phosphate
NADP-MDH	NADP-Malate dehydrogenase
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
NFW	Nuclease Free Water
ng	Nanogram
nm	Nanometer
NN	Neural Network
$^{\circ}$ C	Degree Celsius
OD	Optical Density

PAGE	Polyacrylamide Gel Electrophoresis
PBD	Protein Data Bank
PCR	Polymerase Chain Reaction
PDS	Particle Delivery System
PEP	Phosphoenolpyruvate
PEPC	Phosphoenolpyruvate carboxylase
PG	Phosphoglycolate
PGA	Phosphoglyceric acid
PGR	Plant Growth Regulators
pH	Potenz /Power of Hydrogen
POGs	Putative Orthologous Groups
PPDB	Plant Proteome Database
PPDK	Pyruvate phosphate dikinase
PVDF	Polyvinylidene difluoride
PVPP	Polyvinylpolypyrrolidone
RC	Reliability count
RFU	Relative Fluorescent Unit
RIPE	Realizing Increased Photosynthetic Efficiency
RLP	Rubisco-like protein
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
RuBisCO	Ribulose-1, 5-Bisphosphate Carboxylase/Oxygenase
RuBP	Ribulose-1, 5-bisphosphate
S.E	Standard Error
SBPase	Sedoheptulose-1, 7-bisphosphatase

SDS	Sodium Dodecyl Sulfate
sec	Seconds
SP	Signal Peptide
SSU	Small subunits
STAS	Sulphate transporter anti-sigma factor-like domain
SUBA	The subcellular localization database for <i>Arabidopsis</i>
TAE	Tris Acetate EDTA
TAIR	The <i>Arabidopsis</i> Information Resource
Taq	Taq
TBS	Tris Buffer Saline
T-DNA	Transfer DNA
TE	Tris-EDTA
TEMED	N, N, N', N'-Tetramethylethylenediamine
T _m	Melting temperature
TM	Transmembrane
TMB	3, 3', 5, 5'-Tetramethylbenzidine
TMD	Trans-membrane domain
TMDTP	Transit peptide from a multidrug transporter
TMH	Trans-membrane helices
TNaXTP	Transit peptide from a bile acid: sodium symporter
TP	Transit Peptide
UV	Ultraviolet
v/v	Volume/volume
wt/v	Weight/volume
X-Gluc	5-Bromo-4-Chloro-3-Indolyl- β -D glucuronide
α -KG	α -Ketoglutarate

KEY WORDS

1. *Agrobacterium* mediated co-culture
2. Agroinfiltration
3. *Arabidopsis thaliana*
4. C₃ plants
5. Carbon concentrating mechanisms
6. Carbon dioxide
7. Chloroplast
8. Chloroplast Inner envelope
9. Cyanobacteria
10. Evolution
11. Floral dip
12. GUS Reporter
13. Inorganic carbon transporters
14. mGFP5 reporter
15. *Nicotiana benthamiana*
16. *Nicotiana tabacum*
17. Particle bombardment
18. Photosynthesis
19. Phytohormones
20. Plant transformation
21. Regulation
22. Rubisco
23. SbtA transporter
24. SbtB
25. Transit peptide



Dedicated

To

**My loving parents and my
beloved husband**

**For their support, patience, faith
and for always believing in me...**

CHAPTER 1

INTRODUCTION AND REVIEW OF LITERATURE

1.1. Introduction

Photosynthesis primarily revolves around the capability of the Rubisco enzyme to fix atmospheric CO₂ into energy rich molecules like glucose. Rubisco is the most abundant enzyme and accounts for about 50% of soluble leaf protein in C₃ plants (Evans and Seemann, 1989; Evans and Poorter, 2001). Unfortunately, the enzyme has a low turnover rate and cannot discriminate between its two substrates viz., CO₂ and O₂ and thus leading to wastage of nitrogen and water resources of plants through a process known as photorespiration (Long, 1991). The oxygenase activity of the enzyme has to be kept in check if enhancing plant productivity is the long term objective. In our major cereal crops photorespiratory processes lead to loss of about 25% of carbon fixed during photosynthesis (Furbank et al., 2015). Some lower organisms and higher plants (C₄ and CAM) have developed mechanisms to overcome these losses by the acquiring CCMs (Carbon concentrating mechanisms). CCMs are biophysical processes which aid in concentrating CO₂ around Rubisco so as to enhance the frequency of carboxylation reactions over oxygenation ones. While in prokaryotes, CCMs consists of carboxysomes [compartments harboring Rubisco and carbonic anhydrases (Price and Badger, 1989)] and inorganic carbon uptake systems [HCO₃⁻ and CO₂ transporters (Price et al., 2002)], in eukaryotes like algae and diatoms, pyrenoids (which bear resemblance to carboxysomes) help in concentrating CO₂. In higher plants, C₄ pathway carries out CO₂ fixation by using certain anatomical and biochemical features and CAM plants have mechanisms to counteract arid environments by keeping their stomata closed during the day hours and opening them at night (Cockburn et al., 1979). The C₃ plants are devoid of CCM of any kind as they evolved from eukaryotic algae several million years before the acquisition of CCM by microalgae and cyanobacteria (Badger and Price, 1994).

Amongst all, the cyanobacterial CCM is the most well studied (Figure 1-1). It is the most efficient single celled mechanism to accumulate 1000 folds C_i in the vicinity of Rubisco (Price et al., 2007). The genome sequences for most of the cyanobacteria are available and hence all genes encoding CCM components have been annotated specifically in the well characterized strains. The heart of a cyanobacterial CCM is a small polyhedral compartment known as the carboxysome which contains all the machinery required for CO₂

fixation including Rubisco and carboxysomal carbonic anhydrases. The carboxysomes are protein shells and have specifically charged pores (most probably positively charged) that allow negative metabolites like HCO_3^- to enter through them while restricting the entry of molecules that lack a dipole moment (CO_2 and O_2) and therefore these molecules enter the shell through diffusion. The carboxysomal shell proteins also own an air-lock mechanism that allows passage of RuBP and 3-PGA in and out of the shell respectively thus keeping a close eye on the loss of important metabolites like CO_2 from the carboxysome. Except *Atelocyanobacterium thalassa* (previously *Cyanobacterium UCYN-A*), all cyanobacterial strains possess carboxysomes. In addition to carboxysomal compartments, inorganic carbon transporters also form an indispensable component of a cyanobacterial CCM. There are basically five transporters known, three of which transport HCO_3^- , one belonging to traffic ATPase (BCT1) and two Na^+ dependent HCO_3^- transporters (SbtA and BicA). For CO_2 uptake, two transporters based on modified NADPH dehydrogenase complexes (NDH-I₃ and NDH-I₄) are present.

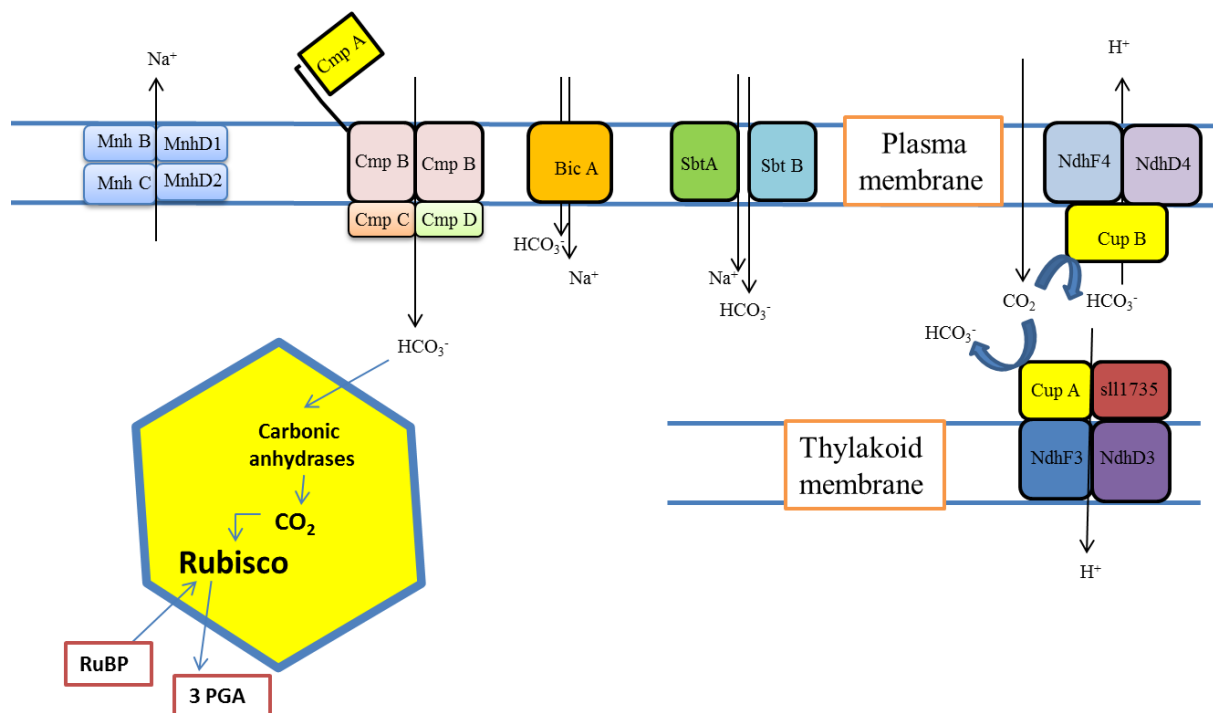


Figure 1-1: Legends on the next page

Figure 1-1: The representation of cyanobacterial CCM and its components. The cyanobacterial CCM consists of three bicarbonate transporters (BCT1, BicA and SbtA) and two CO₂ uptake systems (NdhI₃ and NdhI₄). These transporters help in generating a bicarbonate pool in the cytosol which is utilized within the carboxysome. The carboxysomes dehydrates the bicarbonate to CO₂ which is further fixed into organic carbon by the Calvin cycle.

Since, a cyanobacterial CCM is extremely effective in elevating levels of CO₂ near Rubisco in comparison to other CCMs and involves less transformation experiments, it seems to be the most promising approach of engineering CCMs to C₃ plants for increasing their photosynthetic performance. Additionally, the structure of a cyanobacterial cell is quite similar to chloroplast which makes the cyanobacterial CCMs to be a better alternative. Several attempts have been made to engineer C₄/CAM pathway enzymes as well as algal CCM components into model C₃ plants but the transgenic plants developed did not show any improvement in photosynthetic performance (Singh et al., 2014). Hence, installing components of a cyanobacterial CCM into C₃ plants seems to be the most feasible approach for increasing crop productivity. The entire course would be a complex coup of gene engineering, the first step of which would include incorporation of active bicarbonate pumps to chloroplast envelope followed by assembly of a carboxysome and carboxysomal Rubisco inside the chloroplast and thereafter removal of stromal carbonic anhydrases and endogenous Rubisco.

Amongst the five transporters, SbtA and BicA are the most appropriate candidates for plant transformation as both of them are single gene encoded and therefore their assembly and integration into plant chloroplasts would be easier. In addition to this leaf cytosol has 1-3 mM Na⁺ which is sufficient enough to create gradient across chloroplast envelope (SbtA and BicA require only 1 mM Na⁺). The uptake affinities of HCO₃⁻ for SbtA and BicA are 5-15 μM and 90-170 μM respectively which can be easily addressed by leaf cytosol which has 250 μM HCO₃⁻ (Price et al., 2011). Further, *Arabidopsis thaliana* chloroplasts have also shown existence of several Na⁺/H⁺ antiporters which could aid in creation of concentration gradient (Bassil et al., 2011). The SbtA transporter in comparison to BicA transporter has a lower K_m and is therefore presumed to be more effective in reducing the CO₂ compensation point and improving the WUE under dry-air conditions more efficiently than BicA. The BicA transporter also failed to show any activity when

expressed from the plastid genome (Pengelly et al., 2014) as it was not properly activated. Furthermore, in *Escherichia coli*, BicA failed to be active whereas SbtA was functional in this system suggesting that it could also be active in plant chloroplasts (Du et al., 2014).

The SbtA transporter, first characterized in *Synechocystis* PCC 6803 (freshwater cyanobacteria), is a high affinity bicarbonate transporter (member of sodium solute symporter family) having 10 transmembrane helices and a molecular weight of 39.7 KDa (Shibata et al., 2002) [Figure 1-2]. Being a high affinity transporter, it is induced only in low CO₂ conditions. Its activity is dependent on concentration of Na⁺ ions and along with a bicarbonate ion entering the cell two Na⁺ ions are also symported (McGinn et al., 2004).

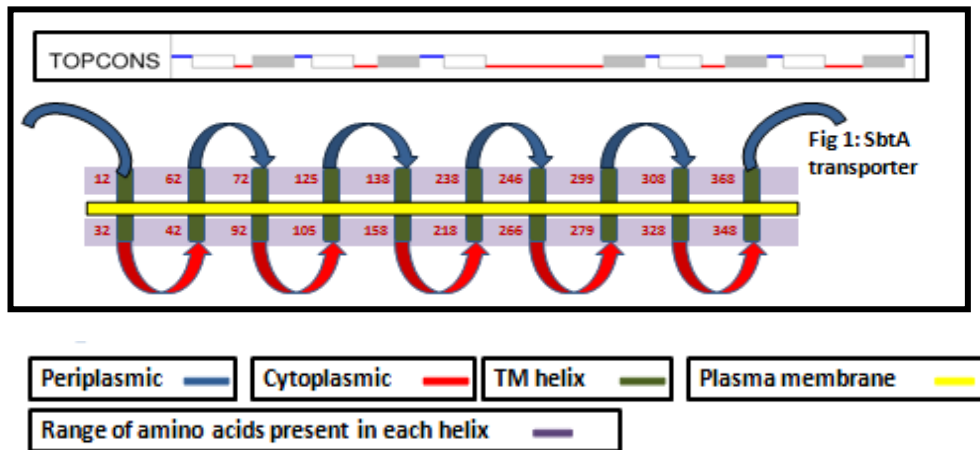


Figure 1-2: The topology of SbtA transporter from *Synechococcus elongatus* PCC 7942. The diagram has been created by the information gathered from TOPCONS software.

In *Synechocystis* 6803, the SbtA transporter requires the presence of NtpJ (a Na⁺/K⁺ pump) which is essential to establish the $\Delta\mu\text{Na}^+$ for active HCO₃⁻ accumulation. The topology studies of the transporter depict that both N- and C- termini are periplasmic. It has a K_m [HCO₃⁻] of about 2 μM and requires about 1 mM Na⁺ for half-maximal HCO₃⁻ transport activity. The 95 member family of SbtA transporter has related homologs within a large number of β -cyanobacteria and a number of α-cyanobacteria and proteobacteria (Figure 1-3).

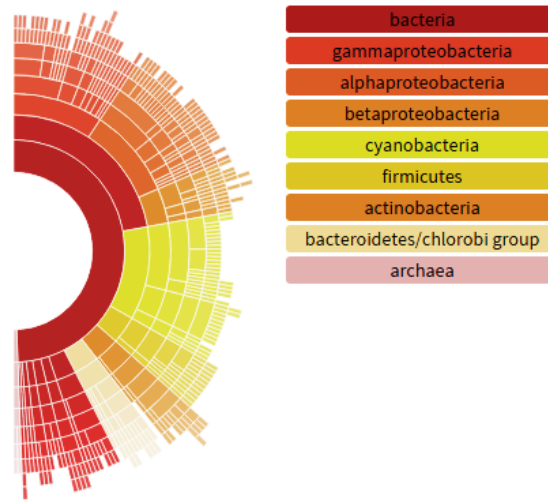


Figure 1-3: The taxonomic profile of SbtA from eggnog database (eggnogdb.embl.de/). The SbtA protein has orthologs in various eubacterial strains as well as in a few archaea.

The transfer of SbtA to C_3 chloroplast inner envelope could be done either by nuclear or plastid transformation. However when plastid transformation was attempted in case of BicA transporter, most of the protein was targeted to the thylakoid instead of the chloroplast inner envelope membrane (Pengelly et al., 2014). Therefore, it seems that nuclear transformation could be the method of choice for integrating SbtA to crop species. The primary requisite for the same would be the presence of pre-sequence of few bases in length for proper targeting as it lacks a cTP (chloroplast transit peptide) of its own (Emanuelsson et al., 1999). Choosing transit peptide sequence from the host plant itself would be a better strategy and selecting these sequences from proteins located in the chloroplast inner envelope of a model C_3 plants (eg.: *A. thaliana*) would be the best approach.

For this, several proteins located on the chloroplast inner envelope membrane were shortlisted and out of them two were finalized as source of TP elements to generate fusion constructs with SbtA transporter from *Synechococcus elongatus* PCC 7942 in plant

expression vectors. The constructs were transformed into *A. thaliana*, *Nicotiana tabacum* and *Nicotiana benthamiana* via several transformation methods like particle bombardment, floral dip, agro-infiltration and *Agrobacterium* mediated gene transfer. The positive transformants were screened and analyzed at the DNA, mRNA and protein level for the presence of transgene. The localization of the protein was confirmed by confocal microscopy analysis. The stable transgenic plants carrying *sbtA* transporter gene were raised under controlled conditions of light and humidity and the seeds were collected from them on maturity.

1.2. CCM: A tool for sequestering carbon

The CCMs evolved as a necessity to cope up with the increasing oxygen concentration in the environment due to the advent of great oxidation event. Other adaptation to the rising O₂ concentration was development of more efficient Rubiscos having better catalysis, higher affinity for CO₂ and better discriminating ability between the two substrates. A basic photosynthetic microorganism's CCM consist of inorganic carbon uptake systems, enzymes for Ci conversion and a Rubisco microcompartment. In nature, CCMs exist in diverse forms viz., carboxysomes in cyanobacteria and proteobacteria, pyrenoids in algae and diatoms, C₄ and Crassulacean acid metabolism (CAM) in higher plants. The CCMs are present in almost all the cyanobacteria, few proteobacteria, in all algae apart from a few members of the Chlorophyta and Rhodophyta as well as in some hornworts, lycopsids and ferns and in a significant number of flowering plants (Winter and Smith, 1996; Badger et al., 1998; Sage and Monson, 1998; Colman et al., 2002). The CCM mechanism has enabled these organisms to achieve higher rate of carbon assimilation per unit of Rubisco although it is associated with additional resource cost of maintenance and synthesis of mechanisms that are involved in active accumulation of carbon (Evans, 2013).

In higher organisms having C₄ and CAM pathways, CO₂ intake occurs via diffusion through stomatal pores whereas in algae and cyanobacteria, CO₂ is taken inside by the help of active CO₂ and HCO₃⁻ uptake systems, the reason being the differences in the rate of CO₂ diffusion in marine and terrestrial environments. In cyanobacteria, energization of Ci accumulation occurs through the NAD(P)H-dependent PSI cyclic electron flow (Mi et al.,

1992). However, eukaryotic microalgae rely on the pH gradient set up across the chloroplast thylakoid membrane in the presence of light (Moroney and Ynalvez, 2007).

1.3. Types of CCM

1.3.1. Cyanobacterial CCM

A cyanobacterial CCM is the most efficient single-celled CCM as it can concentrate 1000 folds of C_i in the vicinity of Rubisco. It comprises of (a) Carboxysomes which are specialized protein compartments harboring Rubisco and Carbonic anhydrases and (b) Inorganic carbon transporters which are further subdivided into CO_2 uptake systems (NdhI₃ and NdhI₄) and HCO_3^- transporters (BCT1, SbtA and BicA). An average cyanobacterial cell has 5 to 15 carboxysomes per cell and the carboxysome shell is composed of six to eight proteins (Price et al., 2007). During the evolution of CCM in cyanobacteria, the first step would have been the acquisition of a carboxysome and thereafter adoption of low and high affinity transporters (Badger and Price, 2003). Carboxysomes are anticipated to have evolved from the smaller shell proteins or from more distantly related bacterial microcompartments (BMC) having role in some other metabolic pathway (BMCs are involved in Pdu (propanediol utilization) and the Eut (ethanolamine metabolism). Cyanobacteria can possess upto five distinct transport systems for C_i uptake and the induction of multiple transporters can occur at the same time. The cyanobacterial cells are capable of achieving higher HCO_3^- concentrations inside in comparison to ambient ones by the help of transporters located at the plasma membrane which actively transport HCO_3^- and CO_2 from the environment into the cytosol. In the cytosol, the accumulated CO_2 is actively hydrated to HCO_3^- . Since there is no CA activity in the cytosol and the molecules like HCO_3^- have low permeability through plasma membrane, a high HCO_3^- concentration is maintained in the cytosol. The bicarbonate molecules can enter the carboxysomes through specifically charged pores in the shell and by the aid of CA localized to the inner side of the shell and is rapidly converted to CO_2 . This leads to a higher CO_2 concentration around Rubisco thus favoring carboxylation over the oxygenation reactions (Price et al., 2011; Espie and Kimber, 2011).

The cyanobacterial CCMs in comparison to CCMs in other organisms is well studied and looked upon as a source of potential genes to enhance photosynthetic performance of C_3 crop species (Price et al., 2007). The major reason being the structural similarities between the chloroplasts of higher plants and cyanobacteria due to common ancestry (have homologous membranes and compartments) (Raven and Allen, 2003), and hence the idea/presumption that engineering them into crop plants will require less anatomical changes. In addition, all the genes and proteins involved in the functioning of a cyanobacterial CCM have been well characterized which will help in identifying the minimum number of genes to be introduced into the C_3 system to carry out the functions with least manipulations of the host genome.

1.3.1.1. Components of a Cyanobacterial CCM

(i) Carboxysomes

The carboxysomes are proteinaceous microcompartments that harbor the Rubisco enzyme along with carboxysomal anhydrases in cyanobacteria and in some proteobacteria (Yeates et al., 2008). An average cyanobacterial cell normally possesses 5 to 15 carboxysomes (Price et al., 2011). The outer surface of the carboxysome has innumerable number of shell proteins with a central pore of 2-3 Å radius (Cai et al., 2015). The carboxysome shell permits passage of ribulose-1, 5-bisphosphate (RuBP), 3-PGA (3-Phosphoglyceric acid), Mg^{2+} and HCO_3^- in between the carboxysome and cytoplasm but offers a diffusion barrier to CO_2 efflux and O_2 influx. The carboxysomal CAs dehydrate HCO_3^- to CO_2 and this CO_2 is fixed into sugars by Rubisco. Carboxysomes have been classified into two types based on Rubisco phylogeny and on the basis of cyanobacterial species in which they are present, the α -carboxysomes are found in α -cyanobacteria, which live in oceanic environments and in some proteobacteria and the β -carboxysomes are present mainly in β -cyanobacteria which inhabit freshwater/estuarine environments (Price et al., 2007).

The carboxysomes possess an icosahedral geometry. The proteins that form this geometry belong to two domain classes: the BMC/Pfam00936 domain i.e. bacterial microcompartment proteins which form hexamers that form the flat facets of the shell and the CcmL/EutN (Pfam 03319) domain proteins that form the pentamers that occupy the

vertices of the carboxysome shell (Kerfeld et al., 2005). The carboxysome shell is composed of CcmK1, CcmK2, CcmK3, CcmK4, CcmL, CcmM, CcmO and CcmP in β -carboxysomes and CsoS1A, CsoS1B, CsoS1C, CsoS1D, CsoS1E, CsoS2, CsoS4A and CsoS4B in α -carboxysomes (Kinney et al., 2011; Rae et al., 2013). CcmM encodes a protein with an N-terminal homologous to γ -CA (Alber and Ferry, 1994) and a C-terminal region having three to four Rubisco small subunit repeats (SSU repeats) (Ludwig et al., 2000). Along with this the types of carbonic anhydrases present in the α - and β -carboxysomes are also different, α -carboxysomes possess γ -CAs (CcaA/IcfA/CcmM) whereas the β -carboxysomes have the α -CA (CsoS3).

(ii) Inorganic carbon transporters

Five types of inorganic carbon transporters are involved in acquisition of Ci in cyanobacteria. Three of them i.e., BCT1, NdhI₃ and NdhI₄ are multi-subunit transporters. BCT1 uses ATP as energy source for HCO₃⁻ transport while NdhI₃ and NdhI₄ use NADPH or reduced ferredoxin for CO₂ uptake. The other two i.e., BicA and SbtA are single gene encoded transporters which derive energy by coupling HCO₃⁻ transport to an electrochemical Na⁺ gradient (Badger and Price, 2003).

(a) NDH-I₃ uptake system

NDH-I₃ is a CO₂ uptake system based on a specialized NADPH dehydrogenase (NDH-I) complexes. It is inducible in limiting Ci conditions and possesses a higher uptake affinity for CO₂ (Shibata et al., 2001; Maeda et al., 2002). The transporter has a low flux rate with photosynthetic affinity ($K_{0.5}$) of 1-2 μ M CO₂. The subunits of this complex are encoded by *ndhF3*, *ndhD3*, *cupA* and *sll1735* genes.

(b) NDH-I₄ uptake system

Ndh-1₄ is a constitutively transcribed CO₂ uptake system encoded by *ndhF4*, *ndhD4*, *cupB* genes. The low affinity transporter has high flux rate with photosynthetic affinity ($K_{0.5}$) of about 10-15 μ M CO₂. This uptake system has been reported to be located on the thylakoid membranes where it is functional in converting the CO₂ diffusing in the cell from outside, or that arising from leakage from the carboxysomes into HCO₃⁻ (Ohkawa et al., 2002).

(c) BCT1 transporter

BCT1 is a high affinity HCO_3^- transporter located on the plasma membrane of the cyanobacteria which is induced under low CO_2 conditions (Omata et al., 1999). It is a member of ABC transporter family and is encoded by *cmpABCD* operon. Among the four polypeptides, CmpA codes for a high affinity solute-binding lipoprotein which is involved in binding to HCO_3^- , CmpB for an integral membrane permease and forms an ion transport channel, CmpC for an ATPase/solute-binding fusion protein that is involved in regulation of transport and CmpD for a cytoplasmic ATPase. CmpC possess an extra domain than CmpD and is proposed to be involved in inactivation of the BCT1 transporter in dark by a mechanism quite similar to NrtC protein of NRT1 transporter (Maeda and Omata, 1997). BCT1 transporter is mainly present in freshwater β -cyanobacterial strains but few transition strains of α -cyanobacteria have also acquired the transporter by horizontal gene transfer (Price et al., 2007)

(d) BicA transporter

BicA is a low affinity and high flux, sodium dependent HCO_3^- transporter and requires about 1 mM Na^+ for half-maximal transport activity. It is a single subunit transporter with 12 transmembrane helices which is expressed under limiting substrate conditions but is even expressed in low levels at high Ci conditions (Price and Howitt, 2010). BicA is a member of SulP/SLC26 family of transporters and its members are present in several eukaryotic and prokaryotic organisms (Shelden et al., 2010). The C-terminus of the transporter has a STAS (sulphate transporter anti-sigma factor-like) domain which has been found to be important in the regulation of the transporter on the basis of studies conducted in certain mammalian homologs of the transporter (Ko et al., 2004).

(e) SbtA transporter

SbtA is a high affinity and low flux HCO_3^- transporter with the $K_m[\text{HCO}_3^-]$ of about 2 μM . SbtA just like BicA, is a single gene encoded transporter and requires about 1 mM Na^+ for half-maximal HCO_3^- transport activity. It belongs to sodium solute symporter family and has 10 transmembrane helices in “5+5” structure/arrangement. Since it is a high affinity transporter, it is induced only in low CO_2 conditions. Its activity depends on concentration

of Na⁺ ions. Along with a bicarbonate ion entering the cell two Na⁺ ions are also taken in. SbtA in *Synechocystis* 6803 requires the presence of NtpJ (a Na⁺/K⁺ pump) which is essential to establish the $\Delta\mu_{\text{Na}^+}$ for active HCO₃⁻ accumulation. Topology studies of the transporter depict that both N and C-termini are periplasmic (Figure 1-2). SbtA has a 95 member family widely distributed in cyanobacteria, proteobacteria and also present in a few bacteria but does not have homologs in archaea and eukaryotes (Price and Howitt, 2014). On the basis of phylogeny SbtA can be of two types, one with 370–374 and the other with 324–339 amino acids. Although, both types have quite low sequence homology but both possess 10 membrane-spanning domains which are conserved. SbtA from *Synechococcus elongatus* PCC 7942 and *Cyanobium gracile* PCC 6307 require less Na⁺ ions (1.5 mM and 0.8 mM respectively) for half maximal activities in comparison to the other SbtA homologs (Du et al., 2014). In most of the cyanobacterial species *sbtA* gene is present in the operon along with another gene, *sbtB* which encodes a small periplasmic protein of 90-110 amino acids. SbtB has been speculated to be important in suppressing the activity of SbtA in dark and the SbtA transporter from several cyanobacterial species showed normal uptake of HCO₃⁻ in the absence of SbtB (Du et al., 2014).

1.3.2. Algal CCM

The presence of a CCM is not mandatory for algae, heterokont algae belonging to groups chrysophyte and synurophyte lack CCM. The green alga, *Coccomyxa* is unusual in that it has no CCM activity and CO₂ enters via diffusion. Some marine red algae also have no CCM, as they live in low light conditions where even less CO₂ levels can fulfill the needs of the cell. Pyrenoids are also not ubiquitous, for instance they are absent in some members of the genus *Chloromonas* (Morita et al., 1998; Morita et al., 1999). For the successful operation of CCM in algae sixteen proteins have been identified as crucial components. These include six transporters [reviewed by Wang et al., (2011)], six CAs [reviewed by Moroney et al., (2011)], two regulatory proteins (Fukuzawa et al., 2001) and two soluble proteins which are presumed to localize around the pyrenoid both in response to light and to low CO₂ conditions (Yamano et al., 2010).

In algae, the availability of CO₂ and HCO₃⁻ for photosynthesis is dependent on the fluctuations in the inorganic carbon concentrations and the pH in the aqueous environment

(Wang et al., 2011). At an acidic pH, CO₂ is the major form of inorganic carbon, whilst at alkaline pH; most Ci is in the form of HCO₃⁻ (Beardall et al., 1998; Gehl et al., 1987). Inorganic carbon (in the form of HCO₃⁻) from the environment is taken in by transporters located on the plasma membrane (HLA3/LCI1) and thereafter taken in to the stroma by the help of transporters on the chloroplast envelope (LCIA/CCP1/CCP2). The stromal HCO₃⁻ is transported into the thylakoid lumen by a putative transporter present on the thylakoid membrane where HCO₃⁻ is rapidly converted to CO₂ by CAH3, a constitutively expressed CA. This CO₂ then diffuses into the pyrenoid matrix from the thylakoid lumen and is fixed by Rubisco present in the pyrenoids. Carbonic anhydrases CAH1, CAH2 and CAH8 are located in the periplasmic space, CAH9 in the cytoplasm, CAH6 and CAH7 in the chloroplast stroma and CAH3 in the thylakoid lumen. There are proteins of unknown functions named as LCIB and LCIC that surround the pyrenoid when CO₂ concentration drops (Mackinder et al., 2017).

1.3.3. Proteobacterial CCM

The phylum Proteobacteria is divided into classes designated by Greek letters i.e., α , β and γ . Proteobacteria possess different forms of Rubisco as a mechanism to adapt to fluctuations in levels of CO₂ and O₂ and they therefore use different forms of Rubisco under different environmental conditions. The forms of Rubisco found in various members of the group are form II, IC, IAq, and IAc. The α -proteobacteria contain four main forms of Rubisco i.e., form II, IC, IAq and IAc while the β -proteobacteria possess form II, IC, IAq and IAc Rubisco. The Rubisco forms IAq and IAc are predominant in γ -proteobacteria along with forms IC and II. Proteobacteria lack form IB Rubisco which is present in β -cyanobacteria. It is also presumed that Form IA Rubisco is the progenitor of Form IB Rubisco (Badger and Bek, 2008).

Unlike cyanobacterial CCM, scattered information is available about how CO₂ is taken and concentrated by a proteobacterial CCM. In two of the γ -proteobacteria viz., *Thiomicrospira crunogena* and *Halothiobacillus neapolitanus*, CCM is present whereas in *Thiobacillus versutus* (an α -proteobacteria) it is reported to be absent (Dobrinski et al., 2005; Karagouni and Kelly, 1989). There also exist differences in the preference of carbon species taken in by the different species of proteobacteria; for instance; CO₂ is the predominantly

utilized form by *H. neapolitanus*, whereas *T. crunogena* can utilize either CO₂ or HCO₃⁻ for Ci uptake. It has been observed that most of the α -carboxysome-containing proteobacteria are acidophilic and utilize CO₂ for their CCM. Recent studies have identified genes for carbonic anhydrase and potential HCO₃⁻ transporters in the sequenced genomes of α -proteobacterium, *Rhodopseudomonas palustris* and the ammonia-oxidizing β -proteobacterium, *Nitrosomonas europaea*. The carboxysomal carbonic anhydrases have also been identified in a chemolithoautotroph, *H. neapolitanus* which are involved in similar functions like their cyanobacterial counterparts (Dobrinski et al., 2005).

1.3.4. Diatomic CCM

The diatoms are marine microalgae that belong to the group Chromista which includes eukaryotes that evolved by multiple endosymbiotic events and therefore they possess plastids with four membranes. Diatoms are unique in that they lack Rubisco activase, SBPase and Oxidative pentose phosphate pathway (Kroth et al., 2008; Gruber et al., 2009). The CO₂ acquisition in diatoms is determined by the ambient CO₂ concentration which influences the transcription of CO₂ responsive genes via cAMP signaling. The two β -CAs (PtCA1 and PtCA2) along with Rubisco are located inside the pyrenoid where they participate in the saturation of Rubisco with CO₂. Other classes of CA, like the α -CAs are expressed independently of CO₂ concentration (Tachibana et al., 2011) and the γ -CAs located in the mitochondria have no role in photosynthesis. Reinfelder et al., (2004) proposed the presence of a C₄ type biochemical CCM in the diatom, *Thalassiosira weissflogii*. However no such mechanism was found in other diatomic species when searched using *in-silico* methods and radio tracer techniques implied it to be species specific.

1.3.5. CCM in C₄ plants

The C₄ pathway carries out CO₂ fixation at the site of Rubisco by using certain anatomical and biochemical features which are characteristic of C₄ photosynthesis. The C₄ plants perform better than their C₃ counterparts and exhibit features like high rate of photosynthesis, increased nitrogen, light and water use efficiency and remarkably low rates of photorespiration (possess carbon concentrating mechanism that reduces photorespiration)

(Leegood, 2002). The C_4 photosynthesis is spatially coordinated between two cell types viz, mesophyll and the bundle sheath cells. The process begins in the cytosol of mesophyll cells where CO_2 is fixed into a four carbon compound (Oxaloacetate) by the help of phosphoenolpyruvate carboxylase (PEPC) enzyme. The oxaloacetate is then converted to malate by NADPH or to aspartate by transamination reaction; malate diffuses to bundle sheath cells through plasmodesmata where it forms pyruvate and CO_2 by decarboxylation and oxidation reactions by the action of malic enzymes. The released CO_2 is fixed by Rubisco enzyme through CBB cycle (Gowik et al., 2011). Pyruvate forms Phosphoenol pyruvate (PEP) by Pyruvate Pi Dikinase (PPDK enzyme) in the mesophyll chloroplasts. The decarboxylation of malate releases higher concentration of CO_2 in the bundle sheath cells of C_4 plants which helps them to achieve higher rates of photosynthesis in comparison to C_3 plants (Bräutigam et al., 2011). Furthermore, since the levels of CO_2 relative to O_2 in bundle sheath cells is higher, rates of photorespiration in C_4 plants is lower than in C_3 plants.

The C_4 plants are of three subtypes depending on the type of major decarboxylation enzyme present: (a) NADP-ME type (present in Chenopodiaceae family e.g., *Halothanmus glaucus*) (b) NAD-ME type (Chenopodiaceae family e.g., *Salsola laricina*) and (c) PEP-CK (Poaceae family e.g., *Spartina anglica*) (Wang et al., 2014). In all three classes, PEPC is localized in mesophyll cells whereas the decarboxylases and Rubisco are present in the bundle sheath cells. The enzyme PEPCK (Phosphoenolpyruvate carboxykinase) is present in the cytosol, NADP-ME in chloroplast and NAD-ME in mitochondria. PEPC, which is the first enzyme of C_4 cycle is present in all subtypes and PEPCK type can coexist with either NADP-ME or NAD-ME type.

1.3.5.1. Variations of C_4 photosynthesis

(a) C_4 mechanism in single cell

Species like *Bienertia cycloptera* and *Borszczowia aralocaspica* which belong to family Chenopodiaceae are able to carry out C_4 photosynthesis in a single chlorenchymatous cell without using Kranz anatomy (Edwards et al., 2004). The single celled C_4 photosynthesis in terrestrial chenopods is of the NADP-ME type. These plant species depict a unique type of anatomy by intracellular partitioning of enzymes and organelles into two compartments

which are similar to mesophyll and bundle sheath cells in their functions. The chloroplasts of these species are dimorphic which means the chloroplasts have two concentric sheaths of cells around each vascular bundle. The inner sheath of cells is described as the bundle sheath and the outer sheath of cells is known as the mesophyll sheath and it contains chloroplasts having grana similar to those of C₃ plants. They are, however, different in that they do not normally contain starch grains (Montes and Bradbeer, 1976).

(b) Single cell C₃-C₄ intermediates

These species are intermediate between C₃ and C₄ plants, in a way that they possess chloroplasts with bundle sheath cells (a feature of C₄ plants) but their cell wall is not as thick as that of a C₄ plant. They also possess low rates of photorespiration in comparison to C₃ plants (Monson et al., 1984). A partially developed kranz anatomy is a characteristic feature of these plants (Moore et al., 1987) and they are able to refix the photorespired CO₂ without using C₄ photosynthesis. The fixation of CO₂ occurs in the mesophyll cells via C₃ cycle which leads to the production of glycolate by the action of RuBP oxygenase enzyme. The glycolate is converted to glyoxylate and glycine by the action of glycolate oxidase and serine/glyoxylate aminotransferase respectively (Nakamura et al., 1983). The glycine is transported to bundle sheath mitochondria where it is converted to serine by the help of glycine decarboxylase/serine hydroxymethyl transferase enzyme. The process leads to the release of CO₂ which is further fixed by chloroplasts in the bundle sheath cells, thereby reducing photorespiration (Edwards et al., 2004).

(c) Single cell aquatic C₄ photosynthesis

Single celled C₄ photosynthesis is present in certain aquatic plants like *Hydrilla verticillata*. *Hydrilla* is an example of facultative C₄ species (NADP-ME type) which can switch from C₃ to C₄ photosynthesis when exposed to low CO₂ conditions. Under high concentrations of C_i (Inorganic carbon) and in cooler climates it carries out C₃ photosynthesis whereas under low C_i concentrations and in warm climates it performs C₄ photosynthesis (Xu et al., 2013). Another example is the *Eleocharis* sp. (family Cyperaceae) which performs C₃ photosynthesis when submerged and on emerging from water performs C₄ or C₃-C₄ intermediate photosynthesis (Bruhl and Wilson, 2007).

1.3.6. CCM in CAM plants

The CAM plants have developed mechanisms to counteract arid environments by keeping their stomata closed during the day hours and opening them at night (temporal regulation). During the night, CO₂ fixation occurs in the mesophyll cells by the help of a reaction catalyzed by PEPC enzyme in a reaction similar to C₄ pathway. This enzyme carries out the carboxylation of PEP to oxaloacetate and Pi. The oxaloacetate is further converted to malate which is thereafter transported and stored in vacuoles. During the day, malate enters the stroma of the chloroplast where it gets converted to pyruvate. The reaction releases a molecule of CO₂ which is then fixed into sugars via the Calvin cycle. The expression of CAM in plants is dictated by environmental and developmental factors. The factors like salinity and osmotic stress also enhance the expression of CAM metabolism. In *Kalanchoe* sp., CAM expression is limited to mature leaves and the younger ones exhibit a more C₃ like metabolism (Lüttge, 2004). In *Kalanchoe blossfeldiane* cv. Tom thumb, short day photoperiod conditions are capable of inducing flowering, thus showing that CAM expression is also regulated by environmental factors. *Mesembryanthemum crystallinum*, a halophyte exhibits CAM metabolism with increasing age (Cushman and Bohnert, 1997). Another CAM species, *Isoetes howellii* expresses CAM photosynthesis on submerged conditions but when water level regresses and leaves become exposed to environment, CAM photosynthesis may be lost and C₃ photosynthesis takes over the carbon metabolism (Keeley, 1983).

1.3.7. C₃ plants lack CCM

The development of CCM took place well after the endosymbiotic origins of land plant chloroplasts, thereby depriving the vascular plants of the CCM (Price et al., 2012). Therefore, C₃ plants do not possess any CCM at leaf or chloroplast level. A typical C₃ plant invests up to 30% of soluble protein and about 25% of leaf nitrogen into Rubisco. They have Rubisco with higher specificity but lower turnover in comparison to the C₄ and CAM plants which have CCM and hence are not capable of achieving high rates of photosynthetic CO₂ fixation. Moreover, any increase in temperature can reduce the amount of CO₂ available to the organism and thereby further reducing the photosynthetic efficiency by favoring photorespiration (Kubien and Sage, 2008). The reaction with O₂ rather than CO₂ results in

loss of some fixed carbon and ammonia molecules and also causes consumption of more ATP (reviewed by Parry et al., 2012). Other than the high rates of photorespiration, C_3 plants also offer diffusive resistance to CO_2 by stomatal pores followed by cytoplasm and then through chloroplasts which results in a drawdown or deficit in the steady-state CO_2 concentration in the chloroplast relative to that in the ambient air (von Caemmerer and Evans, 2010; Evans et al., 2009). This lower CO_2 concentration in the chloroplast further results in reduced Rubisco carboxylation and increased oxygenation, thereby enhancing the expenditure of metabolic energy to recycle carbon and nitrogen via photorespiration.

1.4. Regulation of CCM

The supply of inorganic carbon, presence of regulatory metabolites and transcription factors are some of the important factors which play a vital role in triggering the expression of genes required for working of a CCM (Yamano et al., 2008). Some of the activities of the CCM are constitutive while others are genetically inducible under low or high inorganic carbon (C_i) conditions (McGinn et al., 2003). In a cyanobacterial CCM, the genes coding for Rubisco, carboxysome proteins and low affinity C_i transporters are constitutively expressed (Kaplan and Reinhold, 1999; Price et al., 1998; Price et al., 2002) whereas the expression of high affinity C_i transporters such as high affinity CO_2 uptake system (NDH- I_3), SbtA and BCT1 (encoded by *cmpABCD*) are induced under limiting C_i conditions (Price et al., 2007). Most of the regulators identified till date to be involved in this regulation belong to the LTTR group of transcription factors (Maddocks and Oyston, 2008). Some of the examples of such regulators include CmpR (activator of *cmpABCD* and *psbAII* in cyanobacteria), CcmR (repressor of *sbt* operon, *ndhI₃* operon, *mnh* operon and *ubiX* in cyanobacteria), CbbR (positive regulator of CO_2 fixation in proteobacteria), Ycf 30 (activator of Rubisco operon during dark to light shift in red alga) and NtcB (involved in regulation of nitrate assimilation) (Maddocks and Oyston, 2008; Burnap et al., 2015; Daley et al., 2012) etc.

Some other transcription factors that have been identified in regulation of CCM are members of Abr family (CyAbrB, regulator of the expression of CCM genes in *Synechocystis* sp. PCC 6803 (Yamauchi et al., 2011), zinc-finger family type transcription regulator (CIA5/CCM1 in *Chlamydomonas reinhardtii* (Fang et al., 2012) and Myb type transcription factors (LCR1 in *C. reinhardtii* (Yoshioka et al., 2004) (Table 1-1).

The proteins of LTTR family are known for inducing DNA bend in promoter regions and possess N-terminus having a helix–turn–helix (HTH) motif, which provides a means of binding to DNA (Schell, 1993). The regulatory interactions of these LTTRs are influenced by the presence of allosteric modulators like Ribulose 1, 5 Biphosphate (RuBP), Adenosine triphosphate (ATP), Fructose-1,6-bisphosphate (FBP) and Nicotinamide adenine dinucleotide phosphate (NADPH) (Daley et al., 2012). It was observed that RuBP and 2-phosphoglycolate (2PG) enhanced the binding of *cmpR* (transcriptional activator of *cmp* operon and *psbA* genes) to the operator region of *cmpA-D* operon (Nishimura et al., 2008).

1.4.1. Regulation of CCM in Cyanobacteria

The cyanobacterial CCM is capable of concentrating 1000 fold of C_i in the cytosol of cyanobacterial cells in comparison to extracellular concentrations (Rae et al., 2013). CCMs in cyanobacteria as mentioned before include carboxysomes (absent in *Cyanobacterium* UCYN-A (Zehr et al., 2008) and inorganic carbon transporters which includes both bicarbonate (BCT1, BicA and SbtA) and CO_2 uptake (NdhI₃ and NdhI₄) systems (Badger and Price, 2003).

The cyanobacterial CCMs tend to minimize the loss of CO_2 from the cell by employing several approaches. First, the carboxysome shell is composed of many shell proteins (formed of CcmK homologs viz., CcmK2 and CcmK4 proteins in β -cyanobacteria) which have different pore sizes and biochemical properties that allow entry of only specific or selective moieties like HCO_3^- , RuBP⁻ etc., (Kinney et al., 2011). For instance, pores with positive charges might be involved in selective transport of negatively charged ions like HCO_3^- , whereas shell proteins with larger uncharged pores may play a role in transport of big molecules and co-factors. Second, the absence of carbonic anhydrase (CA) activity in the cytosol helps in increasing the HCO_3^- bulk in the cytosol by preventing the conversion of HCO_3^- to CO_2 (Price, 2011). Also, the NdhI₃ uptake system located in the thylakoid membrane participates actively in recycling the leaked CO_2 , if any, back to the cytosolic pool (Price et al., 2007).

Table 1-1: Regulators of CCM in cyanobacteria, proteobacteria and algae, their function and their co-regulatory metabolites. The domains have been identified using KEGG database (http://www.kegg.jp/ssdb-bin/ssdb_motif).

S. No.	Regulator	Class	Domain	Function	Co-regulatory metabolites	Organisms
1	CcmR	LysR family	Bacterial regulatory helix-turn-helix protein, lysR family, LysR substrate binding domain, PBP superfamily domain	Repressor of <i>sbt</i> operon, <i>ndhI₃</i> operon, <i>mnh</i> operon and <i>ubiX</i>	α -ketoglutarate and NADP ⁺	All Cyanobacteria except <i>Synechococcus</i> PCC 7942
2	CmpR	LysR family	MarR family, Bacterial regulatory helix-turn-helix protein, lysR family, Helix-turn-helix domain, Bacterial regulatory protein, arsR family, LysR substrate binding domain and PBP superfamily domain	Activator of <i>cmpABCD</i> operon (BCT1 transporter)	RuBP and 2PG	Cyanobacteria
3	CyAbrB2 (SII0822)	Arb family	Arb like motif	Regulation of high affinity transporters like SbtA	-	Cyanobacteria
4	Ycf30 (SII0998)	LysR family ORF	Bacterial regulatory helix-turn-helix protein, lysR family, Tubulin/FtsZ family, GTPase domain, ABC transporter, phosphonate, periplasmic substrate-binding protein, LysR substrate binding domain and PBP superfamily domain	Orthologous to CbbR, a <i>cbb</i> operon transcriptional regulator of the lysR family of proteobacteria	RuBP, 3PG and NADPH	Cyanobacteria
5	CbbR	LysR family	Bacterial regulatory helix-turn-helix protein, lysR family, LysR substrate binding domain and PBP superfamily domain	Regulation of <i>cbb</i> operon of several chemo and photoautotrophic bacteria	RuBP, ATP, Fructose 1,6 bisphosphate (FBP) and NADPH	Proteobacteria
6	CIA5/CCM1	Zinc finger	Znf-C2H2/integrase-DNA-binding domain	Regulation of CA's (CAH1, CAH3, CAH4/5) Ci uptake proteins HLA3, CCP1/2, LCI1, NAR1.2, LCIB and LCIC	-	Algae
7	LCR1	Myb (myeloblastosis) family	Myb-type HTH DNA-binding domain	Regulates expression of CAH1, LCIL, LCI6	-	Algae

CcmR (also called NdhR), is a negative regulator (repressor) of SbtA and NdhI₃ transporters. CcmR in *Synechocystis* controls the expression of gene clusters of *sbtA/sbtB* (*sbtAB* operon), *ndhF3/ndhD3/cupA/sll1735* (*ndhI₃* operon), multisubunit Na⁺/H⁺ antiporters (*mnh* operon) and ubiX (encoding protein involved in quinone biosynthesis) (Figure 1-4) (Woodger et al., 2007). The binding of CcmR to cognate repressor control sequences of promoter regions of Ci uptake system genes is assisted by molecules like NADP⁺ and α-ketoglutarate, which are therefore referred to as co-repressors (Daley et al., 2012).

During high Ci concentrations, RuBP is carboxylated to 3-Phosphoglyceric acid (3-PG) which in turn leads to formation of Glyceraldehyde-3-phosphate (G3P). A part of G3P is utilized in synthesis of sugars and the rest in regeneration of RuBP. The carbon skeletons formed further contribute to formation of α-ketoglutarate. Further, the NADPH synthesized in this process is utilized readily due to high rates of carbon fixation leading to a low NADPH/NADP⁺ ratio (Daley et al., 2012). Hence, the C₃ cycle keeps replenishing the levels of co-repressors namely α-ketoglutarate and NADP⁺, giving an advantage to CcmR to stay bound to the repressor complex, leading to repression in transcription of high affinity Ci uptake systems (Daley et al., 2012). But when the Ci concentration drops, the oxygenation of RuBP leads to accumulation of 2-PG and depletion of 3-PG, subsequently leading to low G3P, sugars and finally α-ketoglutarate and RuBP. The lack of the co-repressors causes CcmR to dissociate from the repressor complex leading to activation of high affinity Ci uptake genes (Woodger et al., 2007).

Another regulator, CmpR (cytoplasmic membrane protein regulator), a member of LysR family stimulates the transcription of BCT1 transporter (encoded by *cmpABCD* operon) an ATP-dependent bicarbonate uptake system in cyanobacteria (Badger and Price, 2003) (Figure 1-4). It has also been reported to act as an enhancer of transcription of photosystem II genes (*psbAII* and *psbAIII*) in low CO₂ and high light conditions (Takahashi et al., 2004). The studies report RuBP and 2-PG to be the main modulators of CmpR and that their presence leads to efficient binding of CmpR to the operator region of the *cmp* operon in *Rhodobacter capsulatus* (Dubbs et al., 2004). Being a high affinity ABC type bicarbonate transporter system, BCT1 is required when Ci concentrations are low (Du et al., 2014).

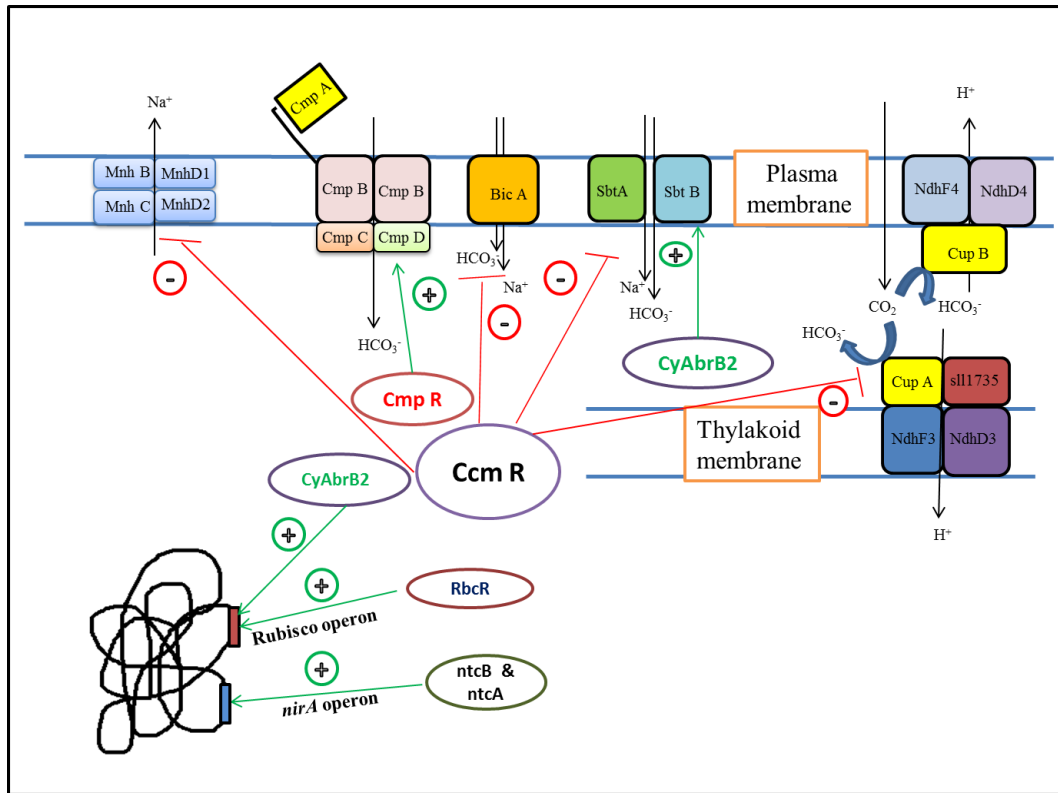


Figure 1-4: The activators and repressors of CCM in cyanobacteria: CcmR acts as a negative regulator of *sbt* operon, *ndhI₃* operon, *mnh* operon and *bicA* in cyanobacteria whereas CmpR (activator of *cmpABCD*) and CyAbrB2 (positive regulator of Rubisco and *sbtAB* operon) function as enhancers. RbcR is a positive regulator of Rubisco operon and *ntcA* and *ntcB* act as positive inducers of *nirA* operon.

When CO_2 concentrations are limiting, the oxygenation reactions of the Rubisco are favored over the carboxylation ones, leading to production of 2-PG (by oxygenation of RuBP) and reduced synthesis of G3P (Calvin cycle) and α -ketoglutarate (Krebs cycle). This leads to accumulation of RuBP (substrate of Rubisco enzyme for carboxylation reactions). Both 2-PG and RuBP act as co-inducers for CmpR (Burnap et al., 2015), leading to activation of BCT1 transporter. It should be noted that the RuBP showed effect as a co-activator only at a very high concentration in comparison to 2-PG (effective even at one-hundredth concentration of RuBP) in cyanobacteria.

1.4.2. Regulation of CCM in Proteobacteria

The proteobacteria are gram negative chemoheterotrophic bacteria which are categorized into α , β and γ subgroups. The forms of Rubisco present in α and β - proteobacteria are II, IAc, IAq, IBc and IC while in γ -proteobacteria, form II, IAc and IAq are more predominant (Badger and Bek, 2008). It is interesting to note that more than one form of Rubisco can exist in proteobacteria, depending upon the need of the organism. This in turn generates the need for a specialized mechanism of regulation for the expression of a particular Rubisco form under a particular set of conditions (Badger and Bek, 2008). The calvin cycle enzymes in proteobacteria are encoded by three unlinked transcriptional units that are the *cbb* operon, the *gap-pgk* operon and the *tpi* gene (Meijer et al., 1996). Amongst the three, Rubisco and phosphoribulokinase, the key enzymes of the Calvin cycle, are encoded by the *cbb* operon, which is under the control of CbbR (van Keulen et al., 2000). The CBB cycle regulator (CbbR) activates the *cbb* operon of several proteobacteria (Kusian and Bowien, 1997) during metabolic shift from chemoheterotrophic to photoautotrophic growth conditions.

In *R. palustris*, the genes for CO₂ assimilation pathway exist in *cbb_I* and *cbb_{II}* gene clusters (Dangel and Tabita, 2015). The *cbb_I* gene cluster consists of (a) *cbbLS* (encodes form I Rubisco), (b) *cbbRRS* (encodes for CbbRRS system) and (c) *cbbR* (encodes for CbbR) genes which are under the control of CbbR, a LysR type transcriptional regulator. The *cbbLS* operon is regulated by a redox regulated two component system (*cbbRRS*) which has (a) A hybrid sensor kinase (*cbbSR*) and (b) Two response regulators (*cbbRR1* and *cbbRR2*) (Romagnoli and Tabita, 2007). The metabolic shift from chemoheterotrophic to photoautotrophic conditions signals the phosphorylation of the response regulators by the sensor kinases. The response regulator *cbbR1* binds to CbbR forming a CbbR1-CbbR complex. This binding is further stabilized by CbbR2 (Joshi et al., 2012) in the presence of the co-inducers (RuBP, ATP, FBP and NADPH) (Joshi et al., 2011).

The presence of co-inducers like ATP and FBP change the conformation of DNA by inducing changes in the DNA helix and a concentration of about 500 μ M of each of them was reported to exhibit maximum binding of CbbR to the regulatory regions (Joshi et al., 2011). Another inducer, NADPH also enhanced CbbR-DNA binding to several folds but the response of CbbR to NADPH was found to be organism specific. For example, in

Xanthomonas flavus, NADPH had an evident effect in reducing the DNA bend and in relieving the torsion created by *cbbR* binding whereas in *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*, NADPH did not influence CbbR binding (Joshi et al., 2012). In *Ralstonia eutropha*, Phosphoenolpyruvate (PEP) reduced DNA binding by CbbR. Methanol, formate and hydrogen enhanced the transcription of *cbb* operon in *X. flavus* whereas presence of carbon source reduced the transcription (Van Keulen et al., 2003).

1.4.3. Regulation of CCM in Algae

Carbon concentrating mechanisms in algae basically relies on the interplay of pH across the thylakoid membrane and the dominance of C_i species existing at a particular pH value (Yamano et al., 2015). Algae possess several isoforms of CAs which carry out the inter conversions of CO_2 and HCO_3^- and also certain special compartments known as pyrenoids that harbor Rubisco, thus mimicking the carboxysome shell of cyanobacteria (Figure 1-5) (Moroney et al., 2001). When algal cells possessing a functional CCM were visualized, they were found to possess a starch sheath around the pyrenoids (Wang et al., 2015). These starch granules display rearrangements upon fluctuations in CO_2 concentrations. They show a homogenous distribution in stroma at high CO_2 concentrations but at low CO_2 concentrations form a compact shell to sequester more CO_2 at the site of Rubisco (Moroney and Ynalvez, 2007).

Since the pH gradient is set up by light driven electron transport chain, presence of light is the primary requisite for efficient working of a CCM in algae (Moroney and Ynalvez, 2007). Wang et al., (2015) proposed 14 genes to be crucial for efficient working of an algal CCM which includes five C_i transporters, four CAs, two pyrenoid proteins, a putative methyl transferase and two nuclear transcriptional regulators. Atkinson et al., (2016) successfully transferred eight CCM components to appropriate location in leaves of higher plants. Brueggeman et al., (2012) reported the discovery of an extensive system of head to head (HTH) gene pairs in CCM genes in their comparative transcriptome analysis. According to them, 1845 CCM genes had HTH conformation out of which 212 gene pairs were upregulated together and in the other 1116 pairs only one of the genes was upregulated. The reason behind this bidirectional regulation was the short distance between the start codons of HTH gene pairs and high median correlation coefficient of 0.674 in

comparison to 0.522 in all other gene pairs. Fang et al., (2012) investigated the entire transcriptome of *C. reinhardtii* to understand the induction of genes on exposure to low CO₂ conditions. They reported that the expression of Ci transporters, several CAs and LCR1 (low CO₂ induced gene 1) essentially requires the expression of CIA5, wherein CIA5 expression is independent of CO₂ levels. LCR1 is involved in regulation of CAH1, LCI1 and LCI6. Yamano et al., (2015) showed that HLA3 and LCIA are localized to plasma membrane and chloroplast respectively and the absence of LCIA sends retrograde signals from the chloroplast to the nucleus to reduce mRNA expression of HLA3. These studies highlight that LCIA plays a key role in HLA3 stability and the duo behaves in co-operative fashion to drive HCO₃⁻ transport in *C. reinhardtii*.

One of the recent discoveries in algal CCM is of Essential Pyrenoid Component 1 (EPYC1), which is presumed to bind to Rubisco to form the pyrenoid matrix (Mackinder et al., 2016). EPYC1 possesses four identical repeats of 60 amino acids each which contain Rubisco binding sites within them. Each EPYC1 can bind to four Rubisco holoenzymes and each Rubisco can bind to eight EPYC1 molecules. This makes possible for a single Rubisco molecule to connect to twelve Rubiscos as a whole (Mackinder et al., 2016). The binding of this protein to Rubisco is still not clear but postulated to be regulated by its phosphorylation at low CO₂ conditions. Other mechanisms of regulation of EPYC1 might be the methylation of cysteine residues and N-terminal methylations (Taylor et al., 2001).

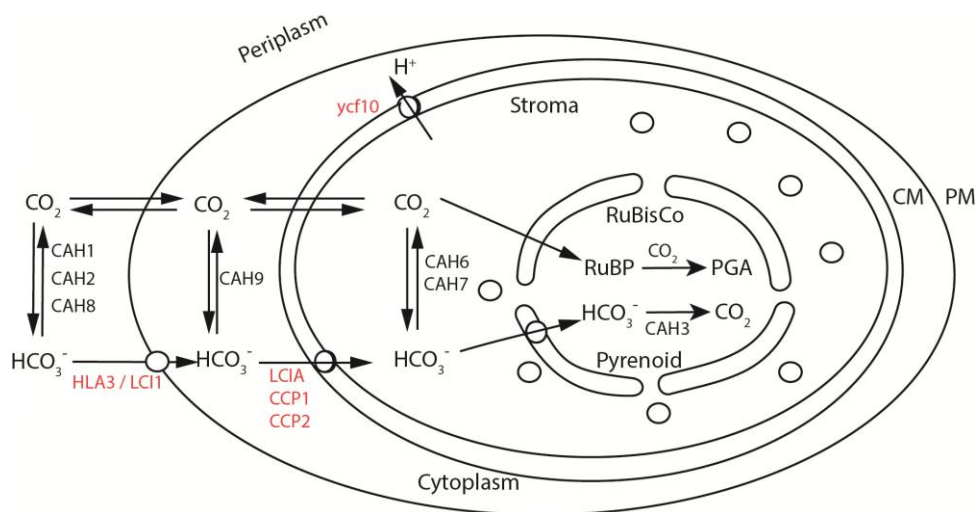


Figure 1-5: Legends on the next page

Figure 1-5: Representation of CCM in algae (*Chlamydomonas reinhardtii*): PM stands for plasma membrane and CM for chloroplast membrane. CAH1, CAH2, CAH3, CAH6, CAH7, CAH8 and CAH9 are various isoforms of carbonic anhydrases. Transporters located at the plasma membrane (HLA3/LCI1) and at the chloroplast envelope (LCIA/CCP1/CCP2) help in C_i acquisition. Ycf10 is involved in extrusion of protons from the stroma into the cytoplasm and thus plays role in CO_2 uptake by creating low pH near the inter-membrane space of chloroplast envelope. LCIB (depicted by 'O' in the figure) flips between two locations i.e., stromal and peri-pyrenoid locations, both in response to light and to CO_2 levels (Yamano et al., 2010).

Most of the CCM genes in *C. reinhardtii* as evident from above are inducible under limiting CO_2 conditions (Spalding and Jeffrey, 1989). The exact mechanism of regulation of these is still unclear but two important regulators have been found out to have a major role in regulation of CCM related genes. One of them is CIA5/CCM1 which through mutant studies has been found to be involved in acclimation to limiting CO_2 conditions by regulating the expression of several putative C_i transporters and inducible CAs (Fukuzawa et al., 2001). The other is LCR1 which has been implicated to regulate the expression of CAH1, LCI1 and LCI6 (Yoshioka et al., 2004).

1.4.4. Regulation of CCM in Diatoms

Diatoms are the principal carbon-fixing organisms in the marine environment and can fix upto 20% of global CO_2 (Raven and Falkowski, 1999). The presence of four plastid envelope membranes (diatom plastids originated from secondary endosymbiosis with a red alga) and association of chloroplasts with the nucleus linked by an ER network are unique features of diatoms (Kroth et al., 2008). Diatoms have type-I Rubisco (Badger et al., 1998). Rubisco along with CAs are localized to the pyrenoid which is a lens-shaped body residing in the central part of the stroma (Tachibana et al., 2011). Carbon concentrating mechanisms in diatoms consist of several plasma membrane located pumps belonging to SLC4 family and different classes of CAs located in between plastidial layers (Figure 1-6). Rubisco activase and sedoheptulose-1, 7-bisphosphatase (SBPase) are absent in diatoms (Kroth, 2015).

The regulation of CCM in diatoms occurs via cAMP mediated cell signaling cascade involving interaction between bZIP proteins and CCRE's (CO_2 /cAMP response elements)

(Tanaka et al., 2016). The cAMP mediated cell signaling is evident in pyrenoid located CA i.e., PtCA1 (*Phaeodactylum tricornutum* carbonic anhydrase 1) which is a low CO₂ inducible protein and is repressed under high CO₂ conditions. The core regulatory region of the PtCA1 promoter has 3 CCREs (CO₂/cAMP response elements) which are site of binding of CREB (cAMP-responsive-element-binding proteins) (Ohno et al., 2012). CREBs are a type of bZIP proteins (PtbZIP11 of *P. tricornutum*) which bind to CCREs to form a repressor complex on the PtCA1 promoter, thus repressing expression of PtCA1 under high CO₂ conditions (Matsuda and Kroth, 2014). Another pyrenoidal CA, PtCA2 basically shows only slight response to low CO₂ conditions and does not display complete repression under high CO₂ conditions (Harada et al., 2005). It has 5 CREs in the region between -630 and -330 upstream of transcription start site.

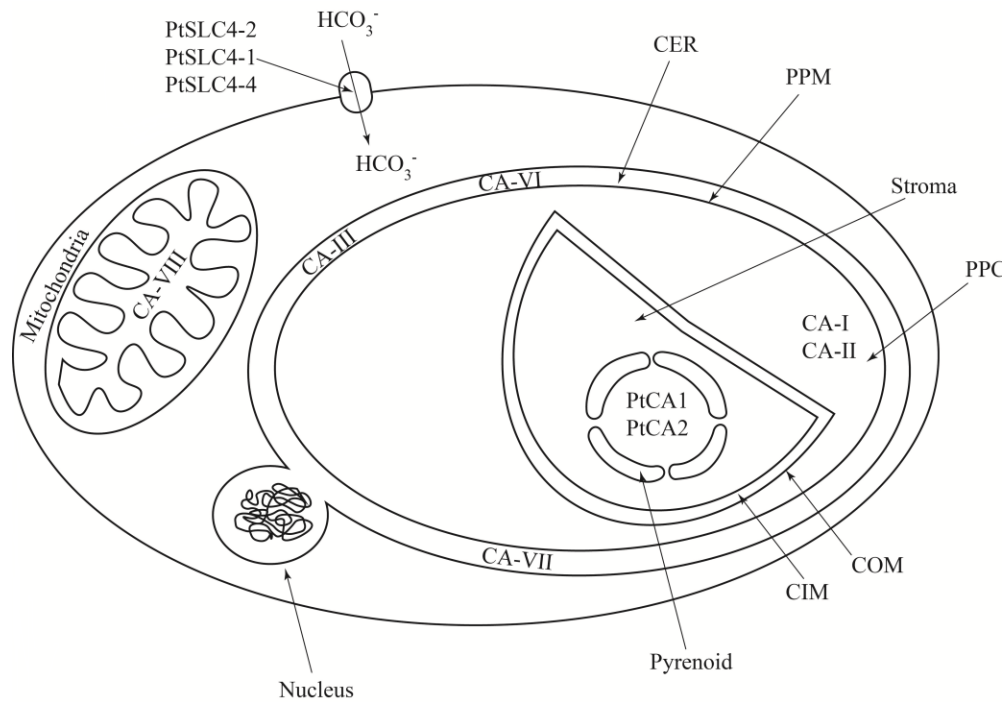


Figure 1-6: Legends on the next page

Figure 1-6: Representation of CCM in diatoms (*Phaeodactylum tricornutum*). CA-I, CA-II, CA-III, CA-VI, CA-VII, CA-VIII, PtCA1 and PtCA2 are isoforms of carbonic anhydrases. CER- Chloroplast endoplasmic reticulum, PPM -Periplastidic membrane, PPC- Periplastidial compartment, COM - Chloroplast outer membrane, CIM - Chloroplast inner membrane. PtSLC4-2, PtSLC4-1 and PtSLC4-4 are the only known plasma membrane bicarbonate transporters in diatoms which work under low CO₂ conditions. In *P. tricornutum*, CA-I and II are localized in the periplastidial compartment, CA-III, VI, VII in the chloroplast endoplasmic reticulum and CA-VIII in the mitochondria. PtCA1 and PtCA2 are the β -CAs localized to the pyrenoid and are involved in increasing the flow of inorganic carbon near the Rubisco, under CO₂ limitation. Diatoms do not contain genes homologous to the other algal transporters or cyanobacterial bicarbonate transporters.

Out of the ten putative bicarbonate transporter genes identified till date in *P. tricornutum*, seven show sufficiently high similarity to human SLC4 proteins (Nakajima et al., 2013). Among the seven, PtSLC4-2, PtSLC4-1 and PtSLC4-4 are responsible for the majority of DIC influx in diatoms (Matsuda and Kroth, 2014). The PtSLC4-2 transporter is located on plasma membrane and plays role both in inorganic carbon uptake (Figure 1-6) as well as in photosynthesis (Matsuda and Kroth, 2014). The other two transporters viz., PtSLC4-1 and PtSLC4-4 are low CO₂ inducible putative inorganic carbon transporters which are repressed in high CO₂ (5% CO₂) (Nakajima et al., 2013).

1.5. Approaches for enhancing the photosynthetic efficiency of crops (C₃ plants)

For increasing the photosynthetic potential of C₃ crop species, several targets have been identified. These include improving kinetic properties of Rubisco (Whitney et al., 2011; Parry et al., 2012), installation of the C₄ pathway into C₃ crops (Gowik and Westhoff, 2011), enhanced activity of sedoheptulose biphosphatase (Raines, 2011), rapid relaxation from photoprotection (Murchie and Niyogi, 2011) and improved canopy architecture. Ort et al., (2011) proposed that improving the light penetration into the leaves and reducing the photosystem antenna size can also lead to increments in photosynthetic yields. Chen and Blankenship (2011) proposed that cyanobacterial chlorophyll 'd' and 'f' if transferred into higher plants, would help in capturing additional light by extending the waveband of sunlight available for photosynthesis. Improving stomatal conductance (Condon et al., 1990; Rebetzke et al., 2013) as well as mesophyll conductance have also been attempted to increase CO₂ uptake from atmosphere. Modifications in Rubisco activase, a catalytic

chaperone involved in modulating the activity of Rubisco has been also suggested by Salvucci and Crafts-Brandner (2004) and Parry et al., (2010). Efforts have also been made for reducing photorespiratory losses by installation of alternative salvage pathways by Peterhansel and Maurino (2011) in *A. thaliana* which have resulted in enhanced growth and biomass. In addition, strategies like introduction of C₄-physiology into C₃ cells (Covshoff and Hibberd, 2012), engineering novel catabolic by-pass pathways (Kebeish et al., 2007), enhancing Calvin cycle RuBP regeneration by increasing sedoheptulose-1,7-bisphosphatase (SBPase) activity (Rosenthal et al., 2011) and increasing the thermotolerance of Rubisco activase to sustain Rubisco activity under moderately elevated temperatures (Kurek et al., 2007; Kumar et al., 2009) have also been attempted.

After many years of fundamental discovery, enough was known about carbon-concentrating mechanisms in cyanobacteria that Price et al., (2008) proposed that bicarbonate pumps should be engineered into the chloroplast envelope of C₃ plants. To realize the full potential of these bicarbonate pumps, they could be engineered along with carboxysomes (Price et al., 2012; Zarzycki et al., 2012) or pyrenoids (Meyer and Griffiths, 2013) which would further elevate the levels of CO₂ around Rubisco.

1.5.1. Modification of Rubisco

The Rubisco enzyme is an inefficient catalyst as it not only has a low turnover number, but also catalyzes two competing reactions: carboxylation and oxygenation (Portis and Parry, 2007). Due to the oxygenation activity, the carbon flow is directed towards the photorespiratory pathway leading to loss of about 25% and 30% of the carbon fixed. The reduction in oxygenase activity of Rubisco can lead to enhanced carbon assimilation thereby leading to higher photosynthetic efficiency (Long et al., 2006). Environmental changes through geological time have shaped the evolution and diversification of Rubisco by providing selective pressures that favor changes in the structure of Rubisco, resulting in improved performance (Tcherkez et al., 2006; Christin et al., 2008).

Rubisco is found in almost all autotrophic organisms, ranging from diverse prokaryotes, including photosynthetic and chemolithoautotrophic bacteria and archaea, to eukaryotic algae and higher plants. The phylogenetic analysis of Rubisco indicates that there

are three classes of bonafide Rubisco proteins, forms I, II and III, which catalyse the same reactions. In addition to these, there exists another form of Rubisco, form IV, which does not catalyse RuBP carboxylation or oxygenation. The form IV is actually a homologue of Rubisco and is called the Rubisco-like protein (RLP) (Ashida et al., 2003). Both Rubisco and RLP appear to have evolved from an ancestor protein in a methanogenic archaeon, and a comprehensive study indicates that the different forms (I, II, III and IV) contain various subgroups, with individual sequences derived from representatives of all three kingdoms of life (Tabita et al., 2008). Form I is widespread among higher plants, eukaryotic algae, cyanobacteria and proteobacteria. There appear to be four subclasses of form I Rubisco, termed as IA, IB, IC and ID, found in different organisms (Tabita, 1999). Form I Rubisco has also been classified into the green and red forms. Form I 'green' which has IA, IB forms of Rubisco is present in higher plants and cyanobacteria while the 'red' one which has IC and ID forms of Rubisco is present in rhodophytes and diatoms (non-green algae). The form II Rubisco is found in proteobacteria, chemoautotrophs and dinoflagellates (Whitney et al., 2011) and Form III is present in archaea (Whitney et al., 2011).

Since, the cyanobacterial Rubisco is phylogenetically more similar to land plant Rubisco (Badger et al., 2002), replacing the endogenous land plant Rubisco with a cyanobacterial enzyme that assembles into a microcompartment can provide an advantage in high CO₂ environment. Replacement of C₃ plant Rubisco with those from different species by deleting the endogenous *rbcL* could be readily accomplished in those species in which chloroplast transformation is possible. Whitney et al., (2011) and Hanson et al., (2012) have been successful in generating transgenic tobacco plants in which tobacco *rbcL* gene was replaced with a *rbcL* gene from another species but unfortunately the plants had impaired photosynthesis.

Other strategies of Rubisco modifications tried include targeting photorespiratory activity of Rubisco (Whitney et al., 2011) and reengineering Rubisco protein using site-directed mutagenesis (Parry et al., 2007). Rubiscos present in eukaryotic phytoplanktons of the red plastid lineage possess superior kinetic properties than the green-type Rubiscos found in higher plants and have the potential to increase the yield of crops like rice and wheat by as much as 30% (Loganathan et al., 2016). For this, several mutants of *C.*

reinhardtii and cyanobacterial Rubisco which depict better catalytic properties have been selected. Genkov and Spreitzer (2009) have created hybrid Rubisco by using large subunit from *Chlamydomonas* and small subunit from higher plants. Sharwood et al., (2008) observed that replacement of large subunit of Rubisco (RbcL) in tobacco by that of sunflower produced hybrid Rubisco comprising of sunflower large and tobacco small subunits. In spite of so many efforts of Rubisco engineering, no significant photosynthetic improvement has been reported till date. One of the basic reasons is the incompatibility of foreign Rubisco with host chaperones. Since most of the strategies attempted till now have focused around modification/engineering of large subunit, engineering of both the plastid (*rbcL*) and nuclear (*rbcS*) genomes needs attention.

1.5.2. Photorespiratory bypass

The photorespiratory pathway has long been seen as a target for crop improvement because of the high energetic cost that it imposes on plant metabolism by releasing assimilated CO₂ in mitochondria (Long et al., 2006). Mutagenized plants that have naturally reduced rates of photorespiration are being searched in several C₃ species (Peterhansel et al., 2010). Other than looking for natural variants, engineering photorespiratory bypasses or altering photorespiratory fluxes are also anticipated to improve plant yield (Peterhansel et al., 2013). Kebeish et al., (2007) transformed *A. thaliana* with five chloroplast-targeted bacterial genes of glycolate catabolic pathway of *E. coli* which coded for enzymes like glyoxylate carboligase, glycolate dehydrogenase and tartronic semialdehyde reductase. These plants had reduced flux of photorespiratory metabolites through peroxisomes and mitochondria as glycolate was directly converted to glycerate. The transgenic plants generated had faster growth rate and biomass and improved photosynthesis. Other bypass proposed by Maier et al., (2012) comprised of introduction of the complete glycolate catabolic cycle in chloroplasts of *A. thaliana* which could oxidize 2PG to CO₂. The approach comprised of engineering glycolate oxidase (GO), malate synthase (MS), and catalase (CAT) to facilitate metabolization of H₂O₂ and glycolate in the chloroplasts.

Carvalho et al., (2011) transformed tobacco plants with *gcl* and *hyi* genes encoding for glyoxylate carboligase and hydroxypyruvate isomerase respectively from *E. coli* and targeted them to the peroxisomes. The transgenic tobacco plants generated produced

bacterial glyoxylate carboligase but not hydroxypyruvate isomerase, leading to higher concentrations of amino acids like glutamine and asparagine and a decrease in level of soluble sugars in the leaves. As a part of the RIPE project, Kebeish bypass, Maier bypass and Carvalho bypass have been successfully engineered in plants and have given promising results from preliminary experiments. Betti et al., (2016) reviewed that the theoretical improvements by establishment of photorespiratory bypass are approximately 15% less in comparison to other engineering approaches for enhancing photosynthesis and the negative effects associated with reducing/modifying photorespiration cannot be overlooked.

1.5.3. Improving Non-photochemical quenching (NPQ)

Non-photochemical quenching is a mechanism by which photosynthetic organisms protect themselves from high intensity of light. When light intensity is too high, the photosynthetic antenna complexes could get overexcited which may cause oxidative damage to the photosynthetic pigments of the thylakoid membrane, so NPQ harmlessly dissipates excess excitation energy by converting it into heat (Müller et al., 2001). Higher light intensity causes protonation of the PsbS subunit of PS II as well as elimination of an epoxide group from violaxanthin (a type of xanthophyll, helps in transfer of photon to chlorophyll 'a') causing the formation of antheraxanthin and zeaxanthin. Both of these molecules are involved in dissipation of the energy in form of heat thereby protecting photosystem II from photodamage. It has been proposed that increase in the relaxation rate of NPQ could be one of the very promising strategies for improving crop photosynthetic efficiency (Murchie and Niyogi, 2011). This approach has also been adopted as a part of RIPE project. Kromdijk et al., (2016) increased amounts of PsbS (photosystem II subunit) as well as speeded up the interconversion of violaxanthin and zeaxanthin in tobacco. The tobacco plants with this scheme had 15% greater plant biomass production in natural field conditions.

1.5.4. Ribulose-1, 5-bisphosphate (RuBP) regeneration

It is believed that if RuBP regeneration is somehow optimized by re-engineering multiple enzymes involved in the C_3 cycle, it can lead to improvement in photosynthesis and growth. Enzymes like SBP/FBPases play an important role in the regulation of carbon flow and in regeneration of RuBP (ribulose-1,5-bisphosphate). Several studies have proven/indicated

that minor reductions in either SBPase or FBPase levels in the C₃ cycle had negative impact on photosynthesis signifying the importance of these enzymes in regulating carbon assimilation rates (Raines et al., 1999; Raines and Paul, 2006). Lefebvre et al., (2005) observed that the transgenic tobacco plants having higher expression of cyanobacterial SBPase/FBPase or SBPase alone showed improvements in photosynthetic carbon fixation. Simkin et al., (2015) generated transgenic tobacco plants with different combinations of fructose-1,6-bisphosphate aldolase, sedoheptulose-1,7-bisphosphatase and *ictB* (a cyanobacterial putative-inorganic carbon transporter B) and observed that increasing the levels of all the three proteins enhanced the rate of photosynthetic carbon assimilation, biomass yield and leaf area individually and to a greater extent when used in combination.

Quite contrary to the results in tobacco, increase in the activity of SBPase in rice had no impact on photosynthesis. The transgenic plants with increased SBPase showed higher photosynthesis rates only when they were subjected to salt stress or heat conditions (Feng et al., 2007a; Feng et al., 2007b). The experiments highlight the fact that manipulation of SBPase for increasing photosynthesis is dependent not only on the plant species involved but also on the growth conditions. Henkes et al., (2001) proposed that decrease in the activity of plastid transketolase significantly reduced photosynthesis and phenylpropanoid metabolism in tobacco. Thus, increasing the activity of transketolase in plants can increase the carbon assimilation and phenylpropanoid metabolism in C₃ plants.

1.5.5. Transfer of C₄ pathway components to C₃ plants

Engineering of C₄ photosynthesis in a C₃ plant requires an understanding of leaf architecture, tissue differentiation and enzymes involved in C₄ cycle. The reason being the compartmentalization in a C₄ leaf, as in a C₄ plant the process of photosynthesis is distributed between the mesophyll and the bundle sheath cells whereas in a C₃ plant, the entire process of photosynthesis occurs in mesophyll cells. Since, the C₄ photosynthesis is more efficient than C₃ and has lower rate of photorespiration, several research consortia are interested in introducing C₄ pathway into C₃ plants for improving crop productivity (von Caemmerer et al., 2012). The discovery of single-celled C₄ carbon-concentrating mechanism in aquatic plant like *Hydrilla verticillata* and species like *Bienertia cycloptera* and *Borszczowia aralocaspica* belonging to Chenopodiaceae family (Sage, 2002) have provided

new incentives for efforts to convert C_3 crops into C_4 plants. Moreover, engineering C_3 plants with a single-celled C_4 cycle would not need an introduction of Kranz anatomy unlike a two celled C_4 mechanism.

Rice (*Oryza sativa*) is the most sought after species for a possible C_3 to C_4 conversion (Ueno et al., 2006) because of its anatomical features which are quite similar to C_4 grasses (the thickness of mesophyll cells, ratio of mesophyll to bundle sheath cells and the arrangement of mesophyll cells around the vascular bundles resemble the kranz anatomy. *Brachypodium distachyon*, relative of major cereal crops like wheat and barley may also serve as an alternate to rice due to its small size, a short life cycle, self-fertility and its small and well annotated genome (Brkljacic et al., 2011).

Attempts have been made to overexpress key enzymes of C_4 pathway like PEPC, PPDK, NADP-MDH and NADP-ME in C_3 plants. Since these enzymes play housekeeping function in C_3 plants, their expression if not restricted to photosynthetic tissue may lead to several pleiotropic effects (Miyao et al., 2011). It has been now well established that incorporation of C_4 mechanism in C_3 plants is perhaps one of the most challenging genetic engineering approach ever undertaken as all the steps involved in C_4 biochemistry are still unidentified (Schuler et al., 2016). A major breakthrough in this area is the discovery of grain specific C_4 photosynthesis in the cross and tube cells of the pericarp of the developing wheat grains (Rangan et al., 2016). Given that there is already a C_4 pathway in the wheat grain during developmental stages, there are great chances of extending the pathway to the entire plant easily.

1.5.6. Engineering CAM to C_3 plants

The CAM pathway which evolved as an adaptation to deal with arid conditions is characterized by night time fixation of CO_2 . CAM increases the CO_2 concentration in the vicinity of Rubisco by ten folds and also suppresses photorespiration. Engineering CAM into C_3 plants in order to improve the resource-use efficiency of C_3 photosynthesis requires a better understanding of the regulatory mechanisms responsible for the spatial and temporal events that occur in CAM plants. In comparison to C_4 metabolism which occurs in the mesophyll and bundle sheath cells, CAM is an attractive alternative as it involves only the

mesophyll cells. CAM plants perform CO₂ uptake during night and therefore have less evapotranspiration rates in comparison to C₄ and C₃ plants (Borland et al., 2009). Kebeish et al., (2012) expressed PEPC of *Solanum tuberosum* in *A. thaliana* under a constitutive as well as a dark induced promoter. The transgenic plants showed marked increase in stomatal conductance, dark respiration rates and transpiration in comparison to wild-type plants.

DePaoli et al., (2014) have enlisted the modules that need to be transferred to C₃ plant for establishing an efficient CAM pathway. These include (i) Carboxylation module responsible for CO₂ fixation and accumulation of malic acid in vacuole at night, (ii) Module for decarboxylation for releasing CO₂ from malate, (iii) A module for controlling night time opening and day time closing of guard cells of stomata and (iv) A module for anatomical requirements for CAM which comprise of the leaf anatomy as well as leaf and stem succulence in order to facilitate CO₂ diffusion within these tissues. Borland et al., (2014) identified *Agave* and *Opuntia* as candidate CAM species that could be used as source for transferring CAM traits to potential C₃ hosts like *Arabidopsis* and *Populus*.

1.5.7. Introduction of algal CCM components to C₃ plants

An algal CCM consists of several inorganic carbon transporters located at the plasma membrane and chloroplast envelope along with a structure called pyrenoid which is analogous to cyanobacterial carboxysome. One of the greatest advantages of an algal CCM is its eukaryotic nature which enables its components to localize to the appropriate locations in higher plants as cTPs from *Chlamydomonas* could be easily recognized by plants (von Heijne, 1991). Engineering algal CCM components to C₃ plants has been reported by Atkinson et al., (2016) where eight CCM components of *Chlamydomonas* were expressed both transiently (in tobacco) and stably (in *Arabidopsis*) and were successfully transferred into appropriate locations in leaves of higher plants. Although the experiments lead to stable expression of transporters like HLA3 or LCIA but failed to achieve any enhancements in growth or photosynthetic efficiency in the transgenic plants. This highlights the possibility of requisite/requirement of an additional transporter activity for their function (Yamano et al., 2015) quite similar to studies conducted on BicA and SbtA (Pengelly et al., 2014).

Other aspects like the assembly of Rubisco to form a pyrenoid-like structure will also be necessary for a fully functional algal CCM to be active in plants (Mackinder et al., 2016). Mackinder et al., (2016) proposed that incorporation of α -helical sequences of RbcS of *Chlamydomonas* into the native SSU of angiosperms, along with proteins like EPYC1 could lead to the formation of a pyrenoid like structure. Atkinson et al., (2017) re-engineered *Arabidopsis* Rubisco by generating a hybrid Rubisco that could aggregate pyrenoid like structure by co-expressing SSU (Rubisco small subunit) α -helices from *Chlamydomonas* with RbcL of *Arabidopsis*. The *Arabidopsis rbcS* mutants stably expressed hybrid Rubisco which was active and had catalytic properties similar to those of native Rubisco.

1.5.8. Incorporation of cyanobacterial CCM to C₃ plants

Transferring cyanobacterial CCM into crop plants has emerged as one of the most promising approaches for engineering CCM to C₃ species. The cyanobacterial CCM is perhaps the most effective of any photosynthetic organism as it is capable of achieving a large increase in the CO₂ concentration around Rubisco and can elevate the CO₂ concentration (1000 folds) in the Rubisco-containing carboxysome structures (Badger and Price, 2003). Along with this the incorporation of a cyanobacterial CCM will require least anatomical changes to the leaf tissue.

Trajectories for engineering a cyanobacterial CCM into C₃ chloroplasts will include:

Step 1: The transfer of bicarbonate pumps to the inner chloroplast envelope of C₃ plants.

Step 2: Assembly of a cyanobacterial carboxysome in the stroma of chloroplasts.

Step 3: Elimination of Carbonic anhydrases from the stroma.

Step 4: Establishment of a functional CO₂ uptake system (For e.g., Ndh-I_{3/4}) in the thylakoid membrane.

Transfer of all the CCM components simultaneously to crop plants may seem to be quite challenging and some of the proteins if transferred without the full apparatus may be deleterious. Along with this, the accurate folding, exact localization, proper assembly and functioning of the proteins are few aspects that need consideration (Price et al., 2012).

Recently, Bonacci et al., (2012) established a functional carboxysome of *H. neapolitanus* in *E. coli*. Lin et al., (2014) attempted engineering carboxysomes into the chloroplast of higher plants chloroplast by agroinfiltration technique where they were able to successfully express fluorescently labeled β -carboxysomal proteins (CcmK2, CcmM, CcmL, CcmO and CcmN) in *N. benthamiana*. The proteins formed structures resembling the empty microcompartments/carboxysomes in the chloroplasts as was evident by confocal and electron microscopy analysis. Cai et al., (2015) in a recent study demonstrated that a mixture of shell proteins from α - and β -carboxysomes can form chimeric microcompartments.

There have been no reports on incorporation of cyanobacterial multi-subunit transporters like BCT1, NdhI₃ and NdhI₄ into crop plants to date. However, few attempts have been made to transfer single subunit transporters like SbtA and BicA to chloroplasts of higher plants. Price et al., (2011) through theoretical modelling approaches proposed that about 5-15% improvement in photosynthetic CO₂ fixation could be achieved by installation of BicA and/or SbtA transporters into the chloroplast inner envelope of C₃ plants.

For the transporters to be functional in plant chloroplasts, leaf cytosol should have sufficient amount of bicarbonate. A typical C₃ leaf cytosol has been reported to have about 250 μ M of HCO₃⁻ which is sufficient enough to drive HCO₃⁻ uptake by BicA and SbtA which have uptake affinities of approximately 90-170 μ M and 5-15 μ M respectively (Evans and von Caemmerer, 1996). Since these transporters are sodium symporters, along with HCO₃⁻ ions, sufficient concentration of Na⁺ ions is also needed for generation of gradient. SbtA and BicA require ~1mM Na⁺ for their activity and it has been anticipated that 1-3 mM Na⁺ is already present in the leaf cytosol (Karley et al., 2000). In addition to it, numerous Na⁺ coupled transporters and Na⁺ /H⁺ antiporters homologous to cyanobacterial forms have been identified in the *Arabidopsis* chloroplast envelope (Rolland et al., 2003). Furthermore, ATP requirement for net CO₂ assimilation by these transporters is less (0.25 ATP for BicA and 0.5 ATP for SbtA) which will make the CO₂ fixation more energy efficient (Price et al., 2011). SbtA transporter because of its lower *K_m* has been presumed to be more effective at reducing the compensation point than the BicA transporter (Price et al., 2011).

Transformation of C₃ plants with the transporter gene could be done by transforming either the nuclear or plastid genome. Generation of fusion constructs with the gene of

interest and the transit peptide is required only while transforming the nuclear genome. The transformation of plastid genome does not require the fusion of chloroplast targeting signals to the proteins of interest (Rolland et al., 2016). Pengelly et al., (2014) attempted to transfer BicA transporter to tobacco via plastid transformation but the transformants had impaired photosynthesis as the transporter was possibly not correctly activated in the heterologous system. When BicA and SbtA were expressed in *E.coli*, BicA did not show any detectable uptake of HCO_3^- whereas six forms of SbtA were functional in this system suggesting that it could also be active in plant chloroplasts (Du et al., 2014).

Recently, Rolland et al., (2016) and Uehara et al., (2016) stated that along with TP sequence, specific amino acid sequences from the N-terminus of certain mature protein candidates may also play a key role in the correct targeting of the protein. However none of the groups have reported any improvements in plant physiology.

1.6. Gaps in existing research

Due to the inherent defect/s in Rubisco enzyme i.e., its low catalytic turnover rate and fickle specificity for its substrate, C_3 photosynthesis cannot operate efficiently with present environmental conditions (Raines, 2011). Therefore, it is quite crucial to manipulate C_3 pathway to increase crop productivity so that the challenge of food security could be addressed. Several targets like modifications in the enzyme Rubisco or establishment of a C_4 pathway into C_3 plants have been identified in an attempt to improve photosynthesis but enhanced photosynthesis still seems to be a far-fetched dream.

Introduction of components of the highly efficient cyanobacterial carbon concentrating mechanisms to C_3 plants has been identified as one of the most reliable methods for improving crop production. However, stacking of all the components together, their probable deleterious effects in heterologous system and complex regulatory processes involved in their functioning are still unexplored and hence limit our understanding. Further, several modelling studies have suggested that the introduction of cyanobacterial bicarbonate transporters like SbtA and BicA individually or together can lead to improvements in C_3 photosynthesis by 5%-15% (Price et al., 2011). Even though these transporters are encoded by single genes each, targeting foreign membrane proteins to chloroplast inner envelope is

quite challenging (Rolland et al., 2016). In the present study, we have attempted to install SbtA transporter from *Synechococcus elongatus* PCC 7942 to C₃ model plants by transforming the nuclear genome. The studies could serve as stepping stone for our long term objective of enhancing agricultural productivity of important crop plants.

1.7. Objectives of the Proposed Research

In the light of available information and in order to address the gaps in existing research, following objectives were undertaken:

1. The study of distribution and arrangement of inorganic carbon transporters in various cyanobacterial species.
 - 40 cyanobacterial phyla were screened to identify genes coding for cyanobacterial Ci transporters and an attempt was made to decipher the trend in evolution of the transporters with respect to transporters present in *Gloeobacter violaceus* PCC 7421, an early-diverging cyanobacterium.
2. Analysis of *sbtA* gene and transit peptide elements.
 - Analysis of SbtA transporter and identification of conserved residues.
 - Prediction of tertiary structure of SbtA by homology modelling.
 - Screening of *A. thaliana* proteome to find out candidate proteins which could serve as a source of transit peptide sequence/s for targeting SbtA transporter to chloroplast inner envelope.
3. Synthesis of fusion construct/s of *sbtA* transporter gene with selected transit peptide element/s and its transformation into C₃ host plant/s followed by *gus* reporter gene assay.
 - Preparation of recombinant constructs (TP-*sbtA*) in appropriate plant expression vector (*pRII01-AN*) having *gus* reporter gene.
 - Optimization of the concentrations of plant growth regulators (PGRs) for *in-vitro* shoot and root induction in *N. tabacum*.

-
- Transformation of chimeric constructs in *N. tabacum* and *A. thaliana* by particle bombardment, *Agrobacterium* mediated co-culture and floral dip methods followed by analysis of transformants by GUS assay.
4. Preparation of chimeric constructs with fluorescent tags for transient expression and localization studies.
- Transformation of *N. benthamiana* by recombinant constructs having *mgfp5* reporter gene by *Agrobacterium* mediated agroinfiltration method followed by the analysis of the presence and expression of transgene at DNA, mRNA and protein level.
 - Study of the sub-cellular localization patterns of GFP tagged fluorescent fusion proteins using confocal microscopy.
5. Generation of stable plants carrying *sbtA* by *Agrobacterium* mediated co-culture method.
- *Agrobacterium* mediated co-culture of *N. tabacum* leaf discs and analysis of transformants by gene specific PCR analysis.
 - Maintenance of transgenic plants under controlled conditions of light and humidity and collection of seeds.

CHAPTER 2

MATERIALS AND METHODS

The chapter describes the general protocols and methodologies followed in the present research. The chapter has been divided into two major parts viz. the *in-silico* analysis and the *in-vivo/in-vitro* analysis (not necessarily in the exact order of their performance). Most of the basic molecular biology protocols have been adapted from the standard book “Molecular Cloning: A Laboratory Manual” (Sambrook et al., 1989; Sambrook and Russell, 2001).

2.1. In-silico Analysis

2.1.1. Distribution and arrangement of inorganic carbon transporters in various cyanobacterial species

2.1.1.1. Retrieval of gene sequences coding for *cmpABCD*, *ndhI₃* and *ndhI₄* operons

In order to predict the evolutionary trends of these genes, the nucleotide sequences coding for *cmpABCD*, *ndhD₃/ndhF₃/cupA* and *ndhD₄/ndhF₄/cupB* operons were searched in forty cyanobacterial species available at the Kazusa cyanobase (<http://genome.microbedb.jp/cyanobase>) and NCBI (<http://www.ncbi.nlm.nih.gov/>) using *Gloeobacter violaceus* PCC 7421 as query. The homologs were determined based on BLAST statistics with the E-value threshold of $\leq 10^{-20}$, the identity of ≥ 30 and coverage percentage of ≥ 60 (Table 2-1).

Table 2-1: The list of the cyanobacterial strains used for the analysis of distribution of *cmpABCD*, *ndhD₃/ndhF₃/cupA* and *ndhD₄/ndhF₄/cupB* operons

S.No.	Name of cyanobacterial strain	S.No.	Name of cyanobacterial strain
1.	<i>Gloeobacter violaceus</i> PCC 7421	21.	<i>Synechococcus</i> 9902
2.	<i>Thermosynechococcus elongates bp-1</i>	22.	<i>Synechococcus</i> sp.WH 8102
3.	<i>Synechocystis</i> 680	23.	<i>Synechococcus</i> sp. WH 7803
4.	<i>Nostoc punctiforme</i> ATCC 29133	24.	<i>Synechococcus</i> sp.PCC 9605
5.	<i>Anabaena</i> 7120	25.	<i>Synechococcus</i> sp. CC 9311
6.	<i>Anabaena variabilis</i> 29413	26.	<i>Synechococcus</i> sp. BL107
7.	<i>Synechococcus</i> 6301	27.	<i>Prochlorococcus</i> 9313
8.	<i>Synechococcus</i> 7942	28.	<i>Prochlorococcus</i> SS120
9.	<i>Synechococcus</i> sp. PCC 7002	29.	<i>Prochlorococcus</i> 9312
10.	<i>Arthrospira platensis</i> NIES-39	30.	<i>Prochlorococcus</i> MED4
11.	<i>Trichodesmium erythraeum</i> IMS101	31.	<i>Prochlorococcus marinus</i> str. NATL1A
12.	<i>Cyanothece</i> sp. PCC 7424	32.	<i>Prochlorococcus marinus</i> str. NATL2A
13.	<i>Cyanothece</i> sp. PCC 7425	33.	<i>Prochlorococcus marinus</i> str. AS9601
14.	<i>Acaryochloris marina</i> MBIC 11017	34.	<i>Prochlorococcus marinus</i> str. MIT 9515
15.	<i>Cyanothece</i> sp. PCC 8801	35.	<i>Prochlorococcus marinus</i> str. MIT 9303
16.	<i>Cyanothece</i> sp. ATCC 51142	36.	<i>Prochlorococcus marinus</i> str. MIT 9301
17.	<i>Synechococcus</i> sp. RCC 307	37.	<i>Prochlorococcus marinus</i> str. MIT 9215
18.	<i>Microcystis aeruginosa</i> NIES-843	38.	<i>Prochlorococcus marinus</i> str. MIT 9211
19.	<i>Synechococcus</i> sp. JA-3-3Ab	39.	<i>Prochlorococcus marinus subsp.marinus</i> str. CCMP1375
20.	<i>Synechococcus</i> sp. JA-2-3B'a (2-13)	40.	<i>Prochlorococcus marinus subsp.marinus</i> str. CCMP1986

2.1.1.2. Phylogenetic analysis of *cmpABCD*, *ndhI₃* and *ndhI₄* operons

The sequences of these genes were subjected to multiple sequence alignment using BioEdit v 7.2.5 (Hall, 1999). The phylogenetic trees of the operons were constructed by the Maximum likelihood method using MEGA software version 5.1 (Tamura et al., 2011) using

1000 bootstrap replicates. The distribution of various transporters among the cyanobacterial species under study was correlated with the evolution of the cyanobacteria which was in turn inferred from 16s rRNA analysis performed by Nelissen et al., 1995.

2.1.2. Analysis of nucleotide and protein sequence of SbtA transporter

2.1.2.1. Phylogenetic analysis of the *sbtAB* operon

In order to carry out the phylogenetic analysis, the nucleotide sequences of seventy homologs of *sbtA* and *sbtB* genes under study were retrieved from NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Kazusa cyanobase (<http://genome.microbedb.jp/>) using an E-value threshold of $\leq 10^{-6}$. The name, phyla and the taxonomic IDs of the organisms used for the analysis of distribution of *sbtAB* operon have been mentioned in Table 2-2. The sequences of individual genes (*sbtA* and *sbtB*) and operon (*sbtAB*) from the phyla were aligned by MSA using ClustalW. The best model on basis of BIC score was then used to carry out phylogenetic analysis (maximum likelihood method based on best fit substitution model) with 1000 replicates of bootstrapping using MEGA 6.0 (Tamura et al., 2013). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining (NJ) and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The phylogenetic trees (in newick format) were submitted to the MirrorTree server (<http://csbg.cnb.csic.es/mtserver/index.php>) to study the co-evolution as well as the matching clades between the two trees (Ochoa and Pazos, 2010).

2.1.2.2. Secondary structure prediction

Three dimensional crystal structure of SbtA protein is not available in PDB (Protein Data Bank). In the absence of 3-D structure, prediction of secondary structure is a useful intermediate step which can facilitate the process of 3-D structure prediction and also help in identifying conserved domains. Secondary structures of SbtA and SbtB proteins were generated using PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) and Phyre² (<http://www.sbg.bio.ic.ac.uk/phyre2/>) servers to find out the total number of α -helices and β -sheets present so as to depict the conformation of amino acid residues that are seen repeatedly in the proteins.

Table 2-2: The name, phyla and the taxonomic IDs of the organisms used for the analysis of distribution of *sbtAB* operon

S. No.	Organism	Name of Phylum	Tax. ID	S. No.	Organism	Name of Phylum	Tax. ID
1.	<i>Synechocystis</i> sp. PCC 6803	Cyanobacteria	1117	36.	<i>Synechocystis</i> sp. PCC 6714	Cyanobacteria	1117
2.	<i>Synechococcus elongatus</i> PCC 7942	Cyanobacteria	1117	37.	<i>Synechococcus</i> sp. UTEX 2973	Cyanobacteria	1117
3.	<i>Acaryochloris marina</i> MBIC 11017	Cyanobacteria	1117	38.	<i>Calothrix</i> sp. 336/3	Cyanobacteria	1117
4.	<i>Cyanothece</i> sp. ATCC 51142	Cyanobacteria	1117	39.	<i>Calothrix</i> sp. PCC 7507	Cyanobacteria	1117
5.	<i>Anabaena variabilis</i> ATCC 29413	Cyanobacteria	1117	40.	<i>Chloroflexus aggregans</i> DSM 9485	Chloroflexi	200795
6.	<i>Arthrospira platensis</i> NIES-39	Cyanobacteria	1117	41.	<i>Pararhodospirillum photometricum</i> DSM 122	Alpha Proteobacteria	28211
7.	<i>Cyanothece</i> sp. PCC 8801	Cyanobacteria	1117	42.	<i>Deinococcus deserti</i> VC D115	Deinococcus-thermus	1297
8.	<i>Cyanothece</i> sp. PCC 8802	Cyanobacteria	1117	43.	<i>Comamonadaceae bacterium</i> A1	Beta Proteobacteria	28216
9.	<i>Cyanothece</i> sp. PCC 7425	Cyanobacteria	1117	44.	<i>Comamonadaceae bacterium</i> B1	Beta Proteobacteria	28216
10.	<i>Geitlerinema</i> sp. PCC 7407	Cyanobacteria	1117	45.	<i>Roseibacterium elongatum</i> DSM 19469	Alpha Proteobacteria	28211
11.	<i>Halothece</i> sp. PCC 7418	Cyanobacteria	1117	46.	<i>Salinispira pacifica</i> strain L21-RPul-D2	Spirochaetes	136
12.	<i>Calothrix</i> sp. PCC 6303	Cyanobacteria	1117	47.	<i>Pseudomonas mendocina</i> ymp	Gamma Proteobacteria	1236
13.	<i>Chroococciopsis thermalis</i> PCC 7203	Cyanobacteria	1117	48.	<i>Pseudomonas pseudoalcaligenes</i> CECT 5344	Gamma Proteobacteria	1236
14.	<i>Cyanobium gracile</i> PCC6307	Cyanobacteria	1117	49.	<i>Pseudomonas alcaligenes</i> str. NEB 585	Gamma Proteobacteria	1236
15.	<i>Dactylococcopsis salina</i> PCC 8305	Cyanobacteria	1117	50.	<i>Moorea producens</i> PAL-8-15-08-1	Cyanobacteria	1117
16.	<i>Microcoleus</i> sp. PCC 7113	Cyanobacteria	1117	51.	<i>Thermus parvatiensis</i> str. RL	Deinococcus-thermus	1297
17.	<i>Nostoc</i> sp. PCC 7107	Cyanobacteria	1117	52.	<i>Rugosibacter aromaticivorans</i>	Beta Proteobacteria	28216
18.	<i>Nostoc</i> sp. PCC 7120	Cyanobacteria	1117	53.	<i>Rhodothermus marinus</i> SG0.5JP17-172	Deinococcus-thermus	1297
19.	<i>Nostoc</i> sp. PCC 7524	Cyanobacteria	1117	54.	<i>Zhongshania aliphaticivorans</i> str. SM2	Gamma Proteobacteria	1236
20.	<i>Prochlorococcus marinus</i> str. AS9601	Cyanobacteria	1117	55.	<i>Meiothermus ruber</i> DSM 1279	Deinococcus-thermus	1297
21.	<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> CCMP 1986(MED 4)	Cyanobacteria	1117	56.	<i>Pseudarthrobacter phenanthrenivorans</i> sphe3	Actinobacteria	201174
22.	<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> CCMP 1375	Cyanobacteria	1117	57.	<i>Nostoc punctiforme</i> PCC 73102	Cyanobacteria	1117

23.	<i>Prochlorococcus marinus</i> str. MIT 9215	Cyanobacteria	1117	58.	<i>Phycisphaera mikurensis</i> NBRC102666 DNA	Planctomycetes	203682
24.	<i>Prochlorococcus marinus</i> str. MIT 9301	Cyanobacteria	1117	59.	<i>Pseudohongiella spirulinae</i> str.KCTC 32221	Gamma Proteobacteria	1236
25.	<i>Prochlorococcus marinus</i> str. MIT 9303	Cyanobacteria	1117	60.	<i>Thermus thermophilus</i> SG0.5JP17-16	Deinococcus-thermus	1297
26.	<i>Prochlorococcus marinus</i> str. MIT 9312	Cyanobacteria	1117	61.	<i>Candidatus nitrosopumilus</i> sp. AR2	Archaeobacteria	2157
27.	<i>Prochlorococcus marinus</i> str. MIT 9313	Cyanobacteria	1117	62.	<i>Azoarcus</i> sp.BH72	Beta Proteobacteria	28216
28.	<i>Prochlorococcus marinus</i> str. MIT 9515	Cyanobacteria	1117	63.	<i>Thioalkalivibrio nitratireducens</i> DSM 14787	Gamma Proteobacteria	1236
29.	<i>Prochlorococcus marinus</i> str.NATL1A	Cyanobacteria	1117	64.	<i>Melioribacter roseus</i> P3M	Bacteroidetes/Chlorobi	68336
30.	<i>Prochlorococcus marinus</i> str.NATL2A	Cyanobacteria	1117	65.	<i>Pseudomonas stutzeri</i> CCUG 29243	Gamma Proteobacteria	1236
31.	<i>Pseudanabaena</i> sp. PCC 7367	Cyanobacteria	1117	66.	<i>Pseudomonas balearica</i> DSM 6083	Gamma Proteobacteria	1236
32.	<i>Rivularia</i> sp. PCC 7116	Cyanobacteria	1117	67.	<i>Oscillatoria nigroviridis</i> PCC 7112	Cyanobacteria	1117
33.	<i>Staniera cyanosphaera</i> PCC7437	Cyanobacteria	1117	68.	<i>Leptolyngbya</i> sp. PCC 7376	Cyanobacteria	1117
34.	<i>Synechococcus elongatus</i> PCC 6301	Cyanobacteria	1117	69.	<i>Fischerella</i> sp.JSC-11	Cyanobacteria	1117
35.	<i>Synechococcus</i> sp. PCC 7002	Cyanobacteria	1117	70.	<i>Chloroflexus</i> sp.Y-400-fl	Chloroflexi	200795

2.1.2.3. Prediction of 3-D structure of SbtA

(a) Comparative modelling using MODELLER 9v11

To predict the 3-D model of the protein, template model was selected using HHpred server (Söding et al., 2005). For this, the amino acid sequences of SbtA transporter was submitted to HHpred online server. Based on the sequence similarity and query coverage of the target by the template and resolution of template structure, the best template was selected. This selected crystal structure was used as a template for protein structure prediction using MODELLER 9v11 (Šali and Blundell, 1993). Discrete Optimized Protein Energy (DOPE) score was calculated for each structure and one with minimum DOPE score was selected. To evaluate the reliability of three dimensional structures of proteins, structure validation was accomplished by using different online available servers, where quality evaluation of the models for the environment profile was done with PROCHECK (Laskowski et al., 1993) and ERRAT (Colovos and Yeates, 1993). Complete assessment and evaluation of the generated models were performed by Ramachandran plot analysis using RAMPAGE (Lovell et al., 2003).

(b) Protein Threading using RaptorX structure prediction tool

A three dimensional model of SbtA protein was also built using the RaptorX structure prediction tool (Källberg et al., 2012). RaptorX, distinguishes itself from other servers as it delivers high-quality structural models for many targets with only remote templates by exploiting structure information rather than sequence information. The model was further evaluated for quality and atomic content using different online servers like PROCHECK (Laskowski et al., 1993), VERIFY3D (Eisenberg et al., 1997), ERRAT (Colovos and Yeates, 1993) and RAMPAGE (Lovell et al., 2003). The obtained model was further refined by energy minimization using online server YASARA (Krieger et al., 2009).

2.1.3. Analysis of transit peptide elements

2.1.3.1. Shortlisting of inner chloroplast envelope proteins

For selection of transit peptide (cTP) for SbtA transporter, several proteins localized to inner chloroplast envelope of *A. thaliana* were selected from the AT_CHLORO database (Ferro et al., 2010). The AT_CHLORO database maintains information for all the proteins that have been identified from *A. thaliana* chloroplast. The database provides information about 1856 proteins that have been categorized as per their sub-plastidial localization i.e., envelope, stroma, thylakoid and grana/stroma-lamellae proteins. Accessory information which includes the accession numbers, sub-plastidial localization, analytical coordinates of all peptides, curated and predicted localization and functions of all the proteins have also been enlisted in the database. It is also linked to websites like TAIR, AtProteome, PPDB, Aramemnon, POGs, SUBA.

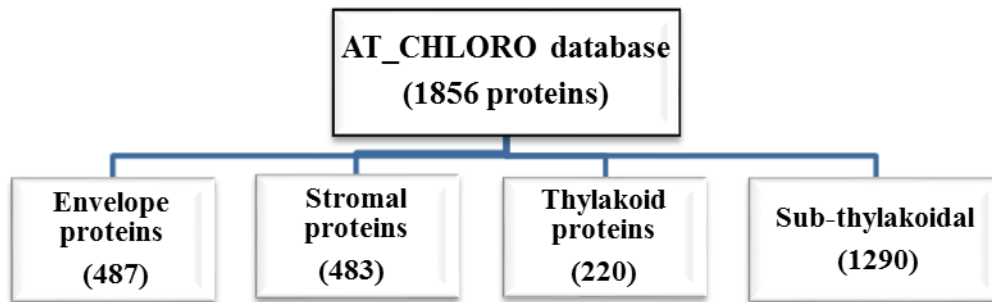


Figure 2-1: A schematic depicting the total number of *A. thaliana* chloroplast proteins and their sub-plastidial localization present on the AT_CHLORO database

2.1.3.2. Screening for the presence of transit peptide

The sub-cellular location of the selected proteins was predicted using TargetP 1.1 Server (Emanuelsson et al., 2000) and the presence of chloroplast transit peptides (cTP) in protein sequences and the location of potential cTP cleavage sites were found out using ChloroP 1.1 Server (Emanuelsson et al., 1999). ChloroP is a neural network (NN) based program which provides information about the presence or absence of cTP, the cTP length and the corresponding cleavage site (CS). The program provides a NN output and CS score which is calculated from an automatic algorithm called MEME. TargetP is also an NN based program

and it predicts mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP) in addition to cTP. The output generated by the program includes the predicted presequence length (TPlen) and reliability class (RC) values which indicate the reliability of prediction.

2.1.3.3. Prediction of trans-membrane helices and orientation of N- and C- terminus

Prediction of trans-membrane protein topology was done using TMHMM server v. 2.0 (Krogh et al., 2001), OCTOPUS (Viklund and Elofsson, 2008) and TOPCONS server (Bernsel et al., 2009). Although both TOPCONS and TMHMM are based on Hidden Markov Model (HMM), TOPCONS server was used to determine the number of trans-membrane helices as well as the orientation of the N- and C-terminus of proteins as it provides cumulative results of five different softwares viz., SCAMPI-single, SCAMPI-multi, OCTOPUS, PRO-TMHMM and PRODIV-TMHMM.

2.1.4. Codon usage and rare codon analysis

Codon usage analysis of the *sbtA* gene was carried out to find the differences and preferences in codon usage with respect to the genome of plants (*Arabidopsis* and *Nicotiana*) that were used for expression of the gene. Graphical Codon Usage Analyzer (GCUA) tool (Fuhrmann et al., 2004) was used to determine relative adaptiveness (http://gcua.schoedl.de/seqoverall_v2.html). The GCUA tool displays the codon usage results in a graphical manner. The low frequency codons were determined by the GenScript Codon Analysis Tool (<https://www.genscript.com/tools/codon-analysis>). Another web server, CAIcal (<http://genomes.urv.es/CAIcal>) (Puigbò et al., 2008) was used to determine the codon adaptation values.

2.2. In-vitro/In-vivo analysis

2.2.1. Optimization of growth conditions of *A. thaliana*

A. thaliana (ecotype Columbia) seeds used in the study were procured from Lehle seeds, USA. The seeds were germinated in two different sets for the experiments. For DNA isolation experiments, the seeds were germinated in ½ MS media. For this, the seeds were surface sterilized with 0.1% mercuric chloride (HgCl₂) solution for 1 min followed by

washing with autoclaved Milli-Q water for 4-5 times in order to remove any residual HgCl_2 traces which might be toxic to the seeds. After washing, the seeds were inoculated in flasks which were first incubated in dark for 2 days and then maintained in plant growth chamber under controlled growth conditions of 16/8 hrs of photoperiod at 22°C temperature and 75% humidity.

For, floral dip and *Agrobacterium* mediated co-culture experiments, the seeds were grown in soilrite mix procured from Keltech Energies Limited, Bangalore. For this experiment, about 10-12 seeds were inoculated in each plastic pot filled with soilrite mix. The pots were watered, covered with aluminium foil (in order to maintain proper humidity) and kept in dark at 4°C for two days for stratification. After stratification, the plants were shifted to plant growth chamber under controlled growth conditions of 16/8 hrs of photoperiod at 22°C temperature and 75% humidity.

Note: Cold stratification of the seeds is the process of exposing seeds to low temperature conditions in order to break their dormancy.

2.2.2. Optimization of growth conditions of *N. tabacum* and *N. benthamiana*

N. tabacum and *N. benthamiana* plants were grown in soil containing a mixture of vermiculite, clay soil and sandy soil in the ratios of 1:1:3. The soil was sterilized thrice by autoclaving before preparation of the pots. The plants were maintained at 25-28°C temperature in a plant tissue culture room under 16/8 hrs of photoperiod.

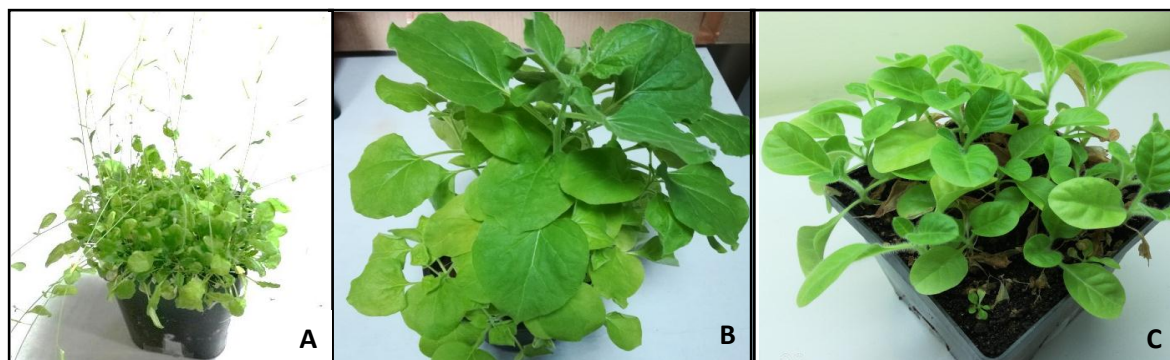


Figure 2-2: *A. thaliana* (A), *N. benthamiana* (B) and *N. tabacum* (C) plants grown in plant tissue culture room under 16/8 hrs of photoperiod. *A. thaliana* plants were grown in soilrite whereas *N. benthamiana* and *N. tabacum* plants were grown in soil mixture.

2.2.3. Plant genomic DNA isolation

Genomic DNA was isolated from *A. thaliana*, *N. tabacum* and *N. benthamiana* leaves using DNeasy Plant Mini Kit (Qiagen).

Procedure

1. 100 mg of leaves were collected from 8-9 weeks old plants and grounded to a fine powder in liquid nitrogen (liq. N₂) using a clean and pre-chilled mortar and pestle. The tissue powder was instantly transferred to a microcentrifuge tube before it thawed.
2. 400 µL of buffer AP1 and 4 µL of RNase A solution (stock concentration of 100 mg/mL) was added to the powdered sample and mixed properly by invert mixing. The samples were then incubated at 65°C in a water bath for 10-12 min for proper lysis of cells.
3. After incubation, 130 µL of buffer P3 was added to the lysate and was incubated on ice for 5 min. The mixture was then centrifuged at 14,000 rpm for 5 min at room temperature.
4. After centrifugation, the supernatant was added into the QIA shredder spin column placed in a 2 mL collection tube and centrifuged at 14,000 rpm for 2 min. The flow through was cautiously collected in a 2 mL microcentrifuge tube (without disturbing the pellet, if present).
5. To the flow-through collected, 1.5 volumes of buffer AW1 was added and mixed by pipetting.
6. 650 µL of the mixture was added into the DNeasy Mini Spin Column placed in a 2 mL collection tube and centrifuged at 8,000 rpm for 1 min. If the total volume of the mixture was more than 650 µL, this step was repeated using the same column and flow through was discarded.
7. The spin column was thereafter placed in a fresh collection tube and 500 µL of buffer AW2 was added to it. The column was centrifuged at 8,000 rpm for 1 min. After centrifugation, flow through was discarded.
8. Buffer AW2 (500 µL) was again added to the column and centrifuged at 14,000 rpm for 2 min.
9. The DNeasy Mini Spin Column was carefully removed from the collection tube and transferred to a 1.5 mL microcentrifuge tube and 30 µL nuclease free water (NFW) was

added to it. After incubating for 5-10 min at RT, the eluted DNA was collected by centrifugation at 8,000 rpm for 1 min.

10. The eluted genomic DNA was visualized on 0.8% agarose gel (Figure 2-3). The remaining samples were stored at -20°C .

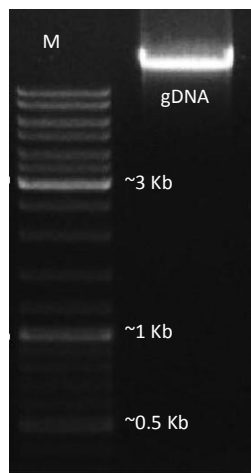


Figure 2-3: The agarose gel electrophoresis image for the genomic DNA (gDNA) isolated from *A. thaliana* leaves; M: Gene ruler SM0331 (Thermoscientific)

2.2.4. Primer designing and PCR amplification

Primers were designed manually by incorporating appropriate restriction enzyme sites in both forward and reverse primer. The primer sequences were analysed using DNAMAN (Lynnon Biosoft), Oligoanalyzer 3.1-Integrated DNA Technologies (<http://eu.idtdna.com/calc/analyzer>) to rule out the possibilities of primer dimer formation and hairpin structure formation. The probabilities of non-specific binding of the primers was determined by BLAST search at NCBI (www.ncbi.nlm.nih.gov/BLAST/BASTn). For determination of melting temperature (T_m) and % GC content of the primers, Promega BioMath calculator (<http://www.Promega.com/apps/biomath/index.html?calc=tm>) was used. Primer sequences used in the study have been enlisted in Table 2-3.

Table 2-3: List of primers used in the study for the amplification of transit peptide sequences, *sbtA* gene, fusion constructs and analysis of putative transformants

S. No.	Primer name	Primer sequence (5' to 3')	Primer length (bases)	T _m (°C)
1.	TNaXTP- <i>NdeI</i> -F	ACTCATATGGCGATAGCTAGTA	22	61.0
2.	TNaXTP- <i>KpnI</i> -F	TTGGGTACCTCTAATGGATCTG	22	62.0
3.	TMDTP- <i>NdeI</i> -F	ATACATATGGCGATTCTACA	21	57.0
4.	TMDTP- <i>KpnI</i> -R	TTCGGTACCTTTGGTGTTT	19	60.0
5.	<i>sbtA</i> - <i>KpnI</i> -F (full length)	AGAGGTACCCGTGGATTTCTTG	22	65.0
6.	<i>sbtA</i> - <i>BamHI</i> -R (full length)	TAAGGATCCGCCGCCAATGG	20	68.0
7.	TNaXTP- <i>sbtA</i> - <i>BglII</i> -F	GAAGGTAGATCTATGGCGATAGCTAGTA	28	63.7
8.	TNaXTP- <i>sbtA</i> - <i>SpeI</i> -R	AAGCTTGAATTCAGTAGTGCCGCCAATG	28	65.1
9.	TMDTP- <i>sbtA</i> - <i>BglII</i> -F	ATATCGAAGGTAGATCTATGGCGATTTC	28	62.2
10.	TMDTP- <i>sbtA</i> - <i>SpeI</i> -R	AAGCTTGAATTCAGTAGTGCCGCCAATG	28	65.1
11.	<i>mgfp5</i> -F (full length)	GGTAGATGGATCCAGTAAAGGAGAAG	26	56.0
12.	<i>mgfp5</i> -R (full length)	GGTGGTGGAGCTCTTTGTATAGTTC	25	58.0
13.	<i>sbtA</i> -RTF (internal)	TCATCGGGTCAAAATCTGGC	20	57.3
14.	<i>sbtA</i> -RTR (internal)	ACGCAGCTCACTAATTCTCG	20	57.3
15.	<i>mgfp5</i> - RTF (internal)	CAGGATCGAGCTTAAGGGAATC	22	55.0
16.	<i>mgfp5</i> - RTR (internal)	GGTAATGGTTGTCTGGTAAAAGG	23	54.0

The primers were procured from Eurofins Analytical Services India Pvt. Ltd in lyophilized form, and were diluted in the 1/10 TE buffer [10 mM Tris-HCl (pH-8) and 1 mM EDTA (pH-8)]. For preparing the final working concentration of 10 pM/ μ L, the primers were diluted with NFW. The PCR components used and their working concentration have been provided in Table 2-4. All PCR reactions were performed in Veriti™ Thermal Cycler machine from Applied Biosystems using Taq DNA polymerase (Bangalore Genei).

Table 2-4: The working concentrations of the various PCR components used in the study

S.No.	Component	Working concentration
1.	Taq Buffer (BufferA) with MgCl ₂	1.5 mM
2.	Primers (Forward and Reverse)	10-30 pM
3.	dNTPs	200 μMoles (50 μMoles of each dNTP)
4.	Taq DNA Polymerase (3 Units/uL)	1-2 units
5.	DNA template	~10 ng
6.	Water (Nuclease Free)	As required (Volume make up)

For amplification of TNaXTP and TMDTP, *A. thaliana* genomic DNA was used as template whereas *sbtA* gene was amplified from *pGEMT-sbtA* construct gifted by Dr. Dean Price, ANU, Australia. For *pCAMBIA1302* cloning, fusion constructs were amplified from *pColdI*-(TNaXTP/TMDTP)-*sbtA* plasmids. The PCR conditions standardized for various genes have been provided in Table 2-5. The PCR products obtained were analyzed by agarose gel electrophoresis. The percentage of gel varied depending on the expected size of the amplicon.

Table 2-5: The PCR conditions standardized for various amplifications

S. No.	Amplicon names	Initial denaturation	Denaturation	Annealing	Extension	Final extension	No of cycles
1.	TNaXTP	95°C, 3 min	95°C, 1 min	50°C, 30 sec	72°C, 1:30 min	72°C, 3 min	35
2.	TMDTP	95°C, 3 min	95°C, 1 min	48°C, 30 sec	72°C, 1:30 min	72°C, 3 min	35
3.	<i>sbtA</i>	95°C, 3 min	95°C, 1 min	56°C, 1 min	72°C, 1:30 min	72°C, 3 min	35
4.	TNaXTP- <i>sbtA</i>	95°C, 3 min	95°C, 1 min	63°C, 1 min	72°C, 1:30 min	72°C, 3 min	35
5.	TMDTP- <i>sbtA</i>	95°C, 3 min	95°C, 1 min	56°C, 1 min	72°C, 1:30 min	72°C, 3 min	35

2.2.5. Agarose gel electrophoresis

Requirements

- Agarose
- 1X Tris-Acetate-EDTA (TAE) buffer: 40 mM Tris Base, 20 mM Glacial Acetic acid and 1 mM EDTA (pH-8.0).
- 6X gel loading dye: 1mM EDTA (pH-8.0), 50% (v/v) Glycerol, 0.4% (w/v) Bromophenol Blue and 0.4% (w/v) Xylene Cyanol FF.
- TE Buffer (pH-8.0): 100mM Tris HCl (pH-8.0) and 10mM EDTA (pH-8.0).
- EtBr (Stock concentration - 10mg/mL): 1µg/mL in TE buffer (pH-8.0).

The required amount of agarose was suspended in 1X TAE and was dissolved by heating in a microwave oven. The melted agarose gel was cooled down to approximately 50°C. EtBr was added to the solution and then it was poured into appropriate gel casting trays containing comb. The Agarose gel was allowed to set for about 30 min and after that the combs were removed. The gel was then submerged into a 1X TAE running buffer. DNA samples were mixed with 1X loading dye and loaded into the wells and then the gel was run at 80 volts/cm. DNA was visualized in a gel documentation system (Molecular Imager® Gel Doc™ XR-Bio-Rad). Quantification of DNA samples was done using the Quantity One software from Bio-Rad® using three standard bands.

2.2.6. PCR amplicon purification by Gel Elution Protocol (using QIAquick® gel extraction kit, Qiagen)

1. The PCR products were run on Agarose gel in freshly prepared 1X TAE buffer and the band of desired DNA was excised by the help of a clean scalpel.

Note: Exposure to UV light for long durations may lead to mutations by forming pyrimidine dimers in DNA. Therefore, UV exposure should be kept minimal.

2. The excised piece of gel containing the desired band was weighed in pre-weighed microcentrifuge tube and 3 volumes of buffer QG was added to the gel (300 µL QG buffer was added to 100 mg of gel).
3. The samples were incubated at 50°C for 10-15 min until the agarose piece melted properly to give a clear solution.

4. Equal volume of isopropanol was added to the sample and mixed properly for DNA precipitation. The samples were kept at room temperature for 5 min.
5. 750 μL of sample was added to the QIA quick spin column placed in a 2 mL collection tube and centrifuged at 13000 rpm for 1 min at room temperature.
6. After centrifugation, the flow through was discarded and the column was placed back in the same collection tube. The previous step was repeated with remaining sample.
7. A volume of 500 μL of QG buffer was added to QIA quick spin column and centrifuged at 13000 rpm for 1min at room temperature. Flow through was discarded.
8. To the QIA quick spin column, 750 μL of PE buffer was added and was incubated at room temperature for 5 min.
9. The column was centrifuged at 13000 rpm for 1 min; flow through was discarded and an empty spin was given at 13000 rpm for 1min to remove any remaining PE buffer in the column.
10. The spin column was transferred to fresh 1.5 mL microcentrifuge tube and NFW (10-20 μL , depending upon the initial DNA concentration) was added to centre of the spin column. The tube was then incubated at room temperature for 5-10 min.
11. After incubation, tube was centrifuged at 13,000 rpm for 2 min at room temperature to collect the eluted DNA. The sample was checked on agarose gel, quantified and stored at -20°C .

2.2.7. Plasmid DNA isolation by Alkaline lysis method (Birnboim and Doly, 1979)

Requirements

- Solution I: 50 mM Glucose, 25mM Tris Cl (pH-8.0), 10mM EDTA (pH-8.0)
- Solution II: 0.2N NaOH, 1% SDS (should be prepared fresh each time)
- Solution III: 3M Sodium acetate (pH-5.2; chilled)
- Tris saturated phenol (Merck), Chloroform and Isoamyl alcohol (25:24:1 v/v)
- Isopropanol, 70% Ethanol (Merck) and NFW

Procedure

1. A single colony from the agar plate containing bacterial culture was inoculated in 5mL LB broth containing appropriate antibiotic/s and allowed to grow for 12-14 hrs at 37°C, with shaking at 200 rpm.
2. To collect the cells, the overnight grown culture was centrifuged at 5000 rpm for 5 min. Supernatant was discarded and tubes were kept on tissue paper in an inverted position to remove any remnants of supernatant.
3. The bacterial pellet was suspended in 100 µL of ice cold Solution I and the suspension was then vortexed to completely dissolve the pellet.
4. To this, freshly prepared Solution II (200 µL) was added and incubated at room temperature for 5 min with intermittent mixing.
5. To the above mix, 150 µL of chilled Solution III was added, gently mixed and incubated on ice for 15 min. The sample was then spun at 10,000 rpm for 10 min at 4°C and the supernatant was transferred to a fresh tube.
6. The clear transparent supernatant was transferred into a fresh microcentrifuge tube. DNase free RNaseA (20 µg/mL) was added and the sample was incubated at 37°C for 1 hour.
Note: RNase solution is heated at 100°C before use, to inactivate any DNase (if present).
7. Equal volume of phenol: chloroform: isoamyl alcohol (25:24:1 v/v) was then added to the supernatant to remove all the proteins. The solutions were mixed by inverting the tubes to get a homogenous mixture. The sample was then centrifuged at 10,000 rpm for 10 min.
8. The upper aqueous phase was carefully removed and transferred to fresh 1.5 mL microcentrifuge tube. DNA was precipitated by adding equal volume of chilled absolute isopropanol and the sample was incubated at -20°C for 15-30 min.
9. The precipitated DNA was pelleted by centrifuging the tubes at 10,000 rpm for 30 min at 4°C. The supernatant was discarded and the pellet was washed with chilled 70% ethanol.
10. The pellet was air dried and suspended in an appropriate volume of NFW (20-30 µL). The integrity of the plasmid DNA was checked by running on 1% agarose gel and the remaining sample was stored at -20°C.

Note: For *pGEM®-T* and *pColdI* vectors, ampicillin (100 µg/mL) while for *pRI101-AN*, *pBI101* and *pCAMBIA-1302* vectors, kanamycin (50 µg/mL) was used while seeding cultures for plasmid isolation.

2.2.8. Cloning of PCR products and generation of fusion constructs in multiple vectors

In the present study, several vector systems and inserts were used to prepare optimal constructs that could be expressed successfully in a plant system (Figures 2-4-to 2-9). The preliminary step included the amplification of the transit peptide sequences and *sbtA* gene from *A. thaliana* genomic DNA and *pGEMT-sbtA* construct respectively. This was followed by cloning of the PCR products (TNaXTP, TMDTP and *sbtA*) into *pGEM®-T* vector so as to incorporate appropriate restriction enzyme sites. The individual genes were then excised from *pGEM®-T* clones and cloned in intermediate cloning vector i.e., *pColdI* to generate fusion constructs (TNaXTP/TMDTP-*sbtA*). The fused segment from *pColdI* clones were amplified and subcloned in plant expression vectors (*pRI101-AN* and *pCAMBIA1302*). The clones generated in plant expression vectors were used for plant transformation studies.

2.2.8.1. Cloning of PCR products in *pGEM®-T* vector

The *pGEM®-T* vector (Promega) is a linearized vector with a single 3'-terminal thymidine at both ends. The vector possesses T overhangs which readily ligate with A-overhangs of insert generated by polymerases lacking 3' to 5' exonuclease activity (Taq DNA polymerase gives 3'A overhangs as it lacks 3' to 5' exonuclease activity). Additionally, *pGEM®-T* is a high copy number plasmid having T7 and SP6 RNA polymerase promoters on either side of the MCS region, within the α -peptide coding region of β -galactosidase. Therefore, positive recombinants can be differentiated from negative ones by blue/white screening. The ligation reaction with the *pGEM®-T* cloning includes:

- *pGEM®-T* vector (100 ng)
- 2X ligation buffer (working concentration - 1X)
- PCR product with A overhangs (concentration depending upon the molar ratio between vector and insert)
- T4 DNA ligase (3 Weiss U/µL).

The formula for calculation of required concentration of insert depending upon the molar ratio is as follows:

$$\text{ng of insert} = \frac{\text{ng of vector} \times \text{Kbp size of insert} \times \text{insert: vector molar ratio}}{\text{Kbp size of vector}^*}$$

*The size of *pGEM®-T* vector is 3000 bp.

To set up the ligation reaction standard molar ratio of 3:1 (insert: vector) was used.

Procedure

1. All the components of the *pGEM®-T* kit were thawed on ice before setting up the reaction. The 2X rapid Ligation Buffer was vortexed vigorously, before use.
2. Added all the components sequentially as mentioned in table below into a 0.5 mL microcentrifuge tube. The contents were then mixed by pipetting.

Table 2-6: The composition of ligation reaction along with recommended controls for TA cloning

S.No.	Reaction component	Test Reaction	Positive control	Background control
1.	2X rapid Ligation Buffer, T4 DNA Ligase	5 µL	5 µL	5 µL
2.	<i>pGEM®-T</i> Vector (50ng)	1 µL	1 µL	1 µL
3.	PCR product	x µL*	-	-
4.	Control Insert DNA	-	2 µL	-
5.	T4 DNA Ligase (3Weiss U/µL)	1 µL	1 µL	1 µL
6.	Nuclease free water to final volume	10 µL	10 µL	10 µL

*Molar Ratio of insert: vector varied for different inserts

3. A positive control with T-vector and control DNA (with A overhangs) was set up to check the efficiency of the enzyme and vector. Along with this a background control was also set up to check self ligation of the vector (Table 2-6).
4. After mixing, the tubes were incubated at 4°C for overnight.

5. After 12-13 hrs, the entire ligation mixture was transformed into high-efficiency competent cells.
6. The transformed cells were screened by plating on IPTG+X-gal plates. Plasmid was isolated from white colored colonies and digestion of the plasmid was performed by appropriate enzymes (for which the sites are present in the primers) to check the presence of insert.

Note: The major concern associated with TA cloning is that the gene has 50% chance of getting cloned in the reverse direction as directional cloning is not possible with TA vectors. Hence, in order to get the orientation of the insert in right frame, different enzymes sites were introduced in the forward and reverse primers.

*For setting up ligations in *pColdI*, *pRI101-AN* and *pCAMBIA1302* vectors, standard molar ratio of 3:1 (insert: vector) was used. In all the reactions, 100 ng of the vector was taken and the insert concentration was calculated using the formula mentioned above. The gene sequences and vectors were ligated together using T4 DNA ligase enzyme (Thermo Fisher Scientific/Promega) and the ligation reactions were performed overnight at 4°C or 16°C (on the basis of manufacturer's instructions).

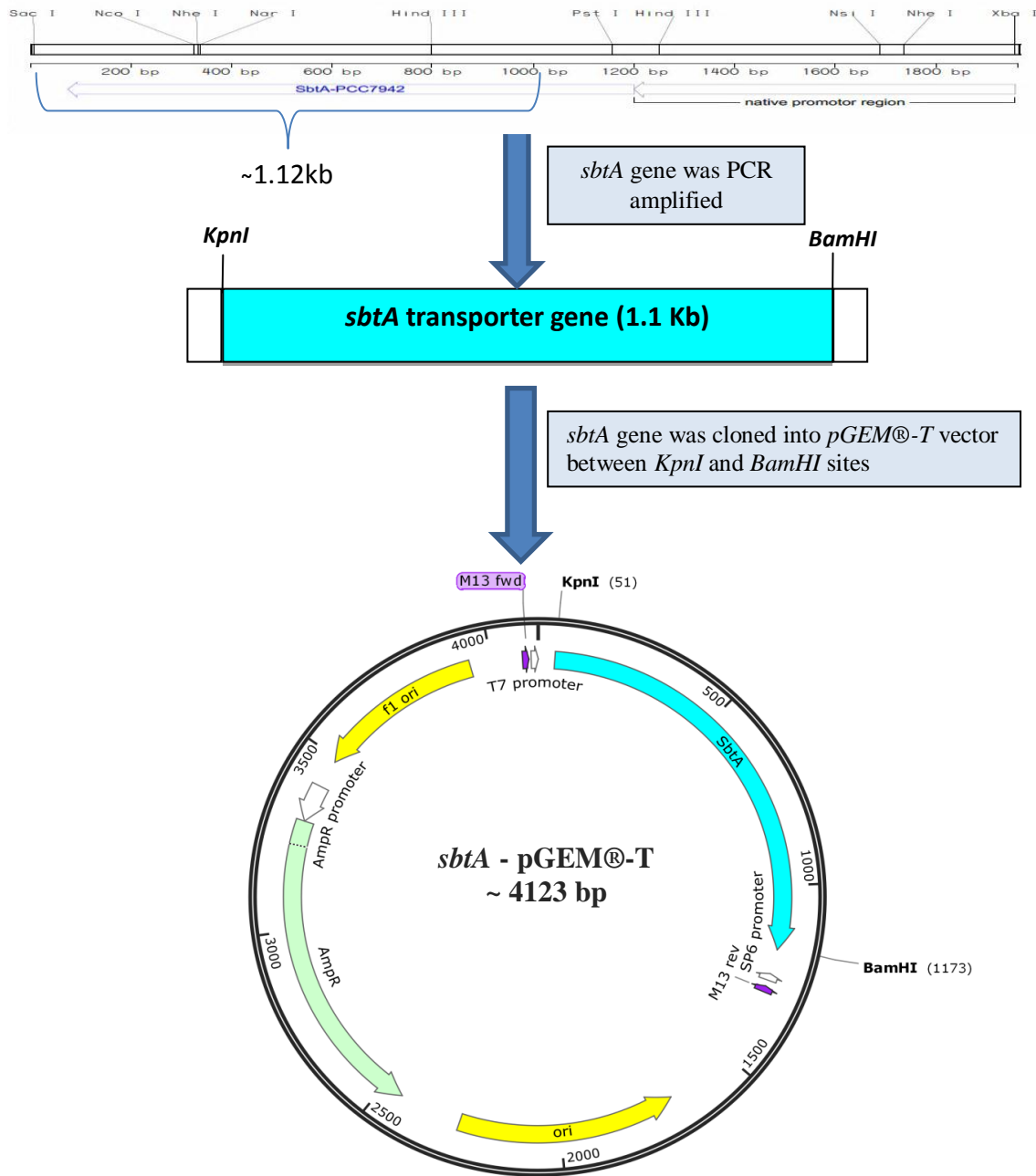


Figure 2-4: A schematic depicting cloning of *sbtA* gene into *pGEM*[®]-*T* vector. The vector map has been generated by using SnapGene[®] 4.0.5 software.

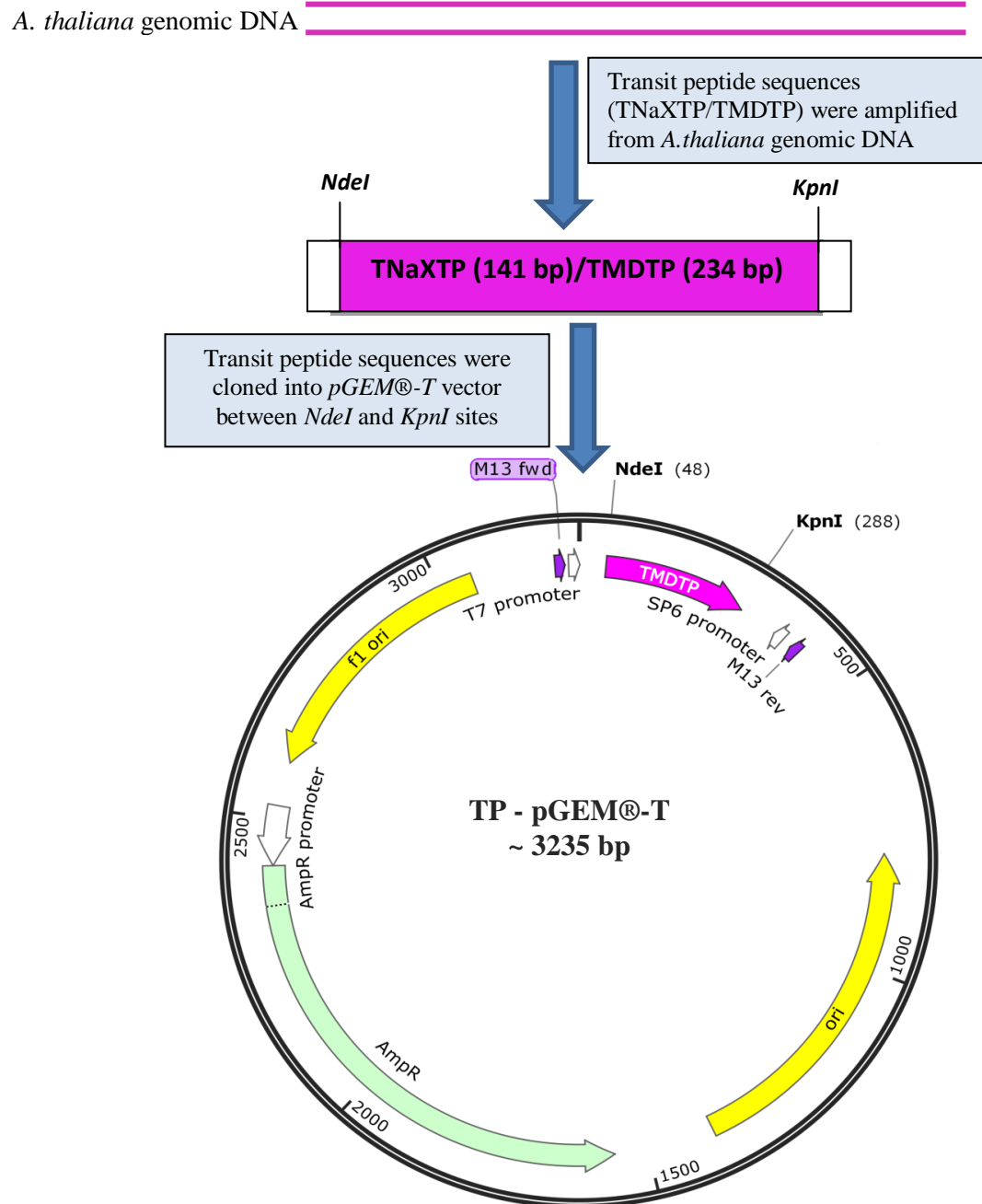


Figure 2-5: A schematic depicting cloning of transit peptide sequences (TNaXTP/TMDTP) into *pGEM*[®]-*T* vector. The vector map has been generated by SnapGene[®] 4.0.5 software using TMDTP as a reference transit peptide.

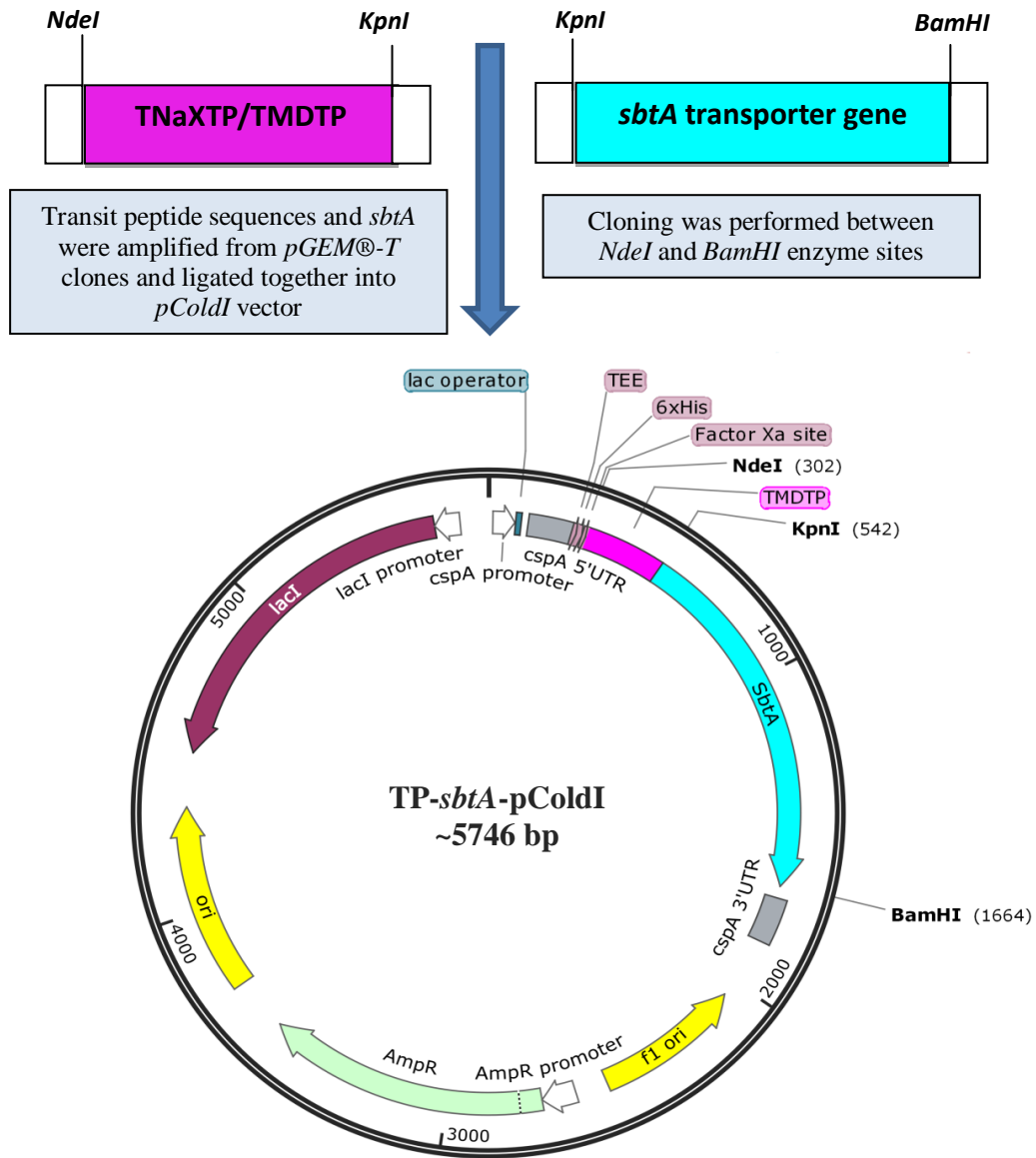


Figure 2-6: A schematic depicting cloning of transit peptide sequences (TNaXTP/TMDTP) and *sbtA* gene into *pColdI* vector to generate fusion genes. The vector map has been generated by SnapGene® 4.0.5 software using TMDTP as a reference transit peptide.

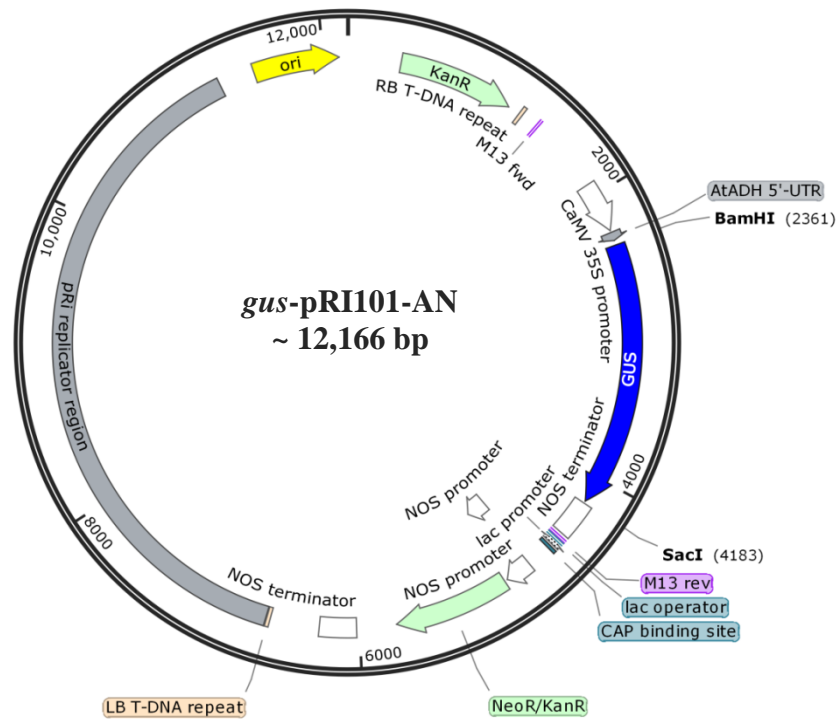
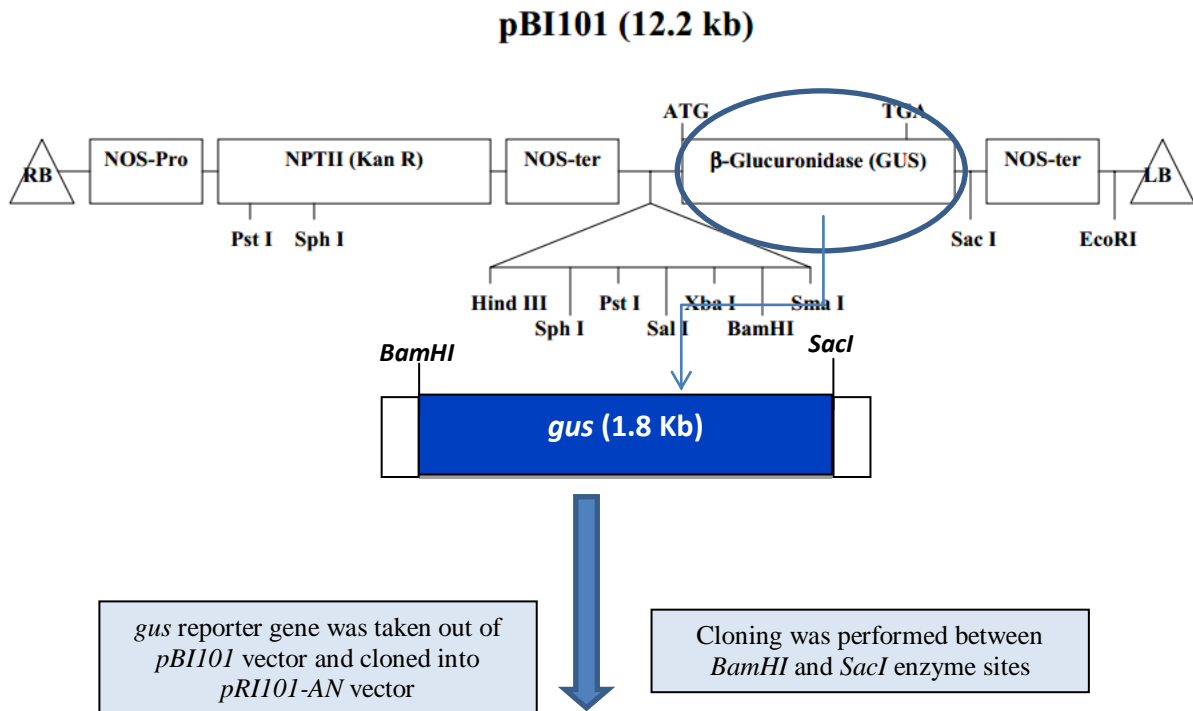


Figure 2-7: A schematic depicting cloning of *gus* gene from *pBI101* vector into *pRI101-AN* vector. The vector map has been generated by using SnapGene® 4.0.5 software.

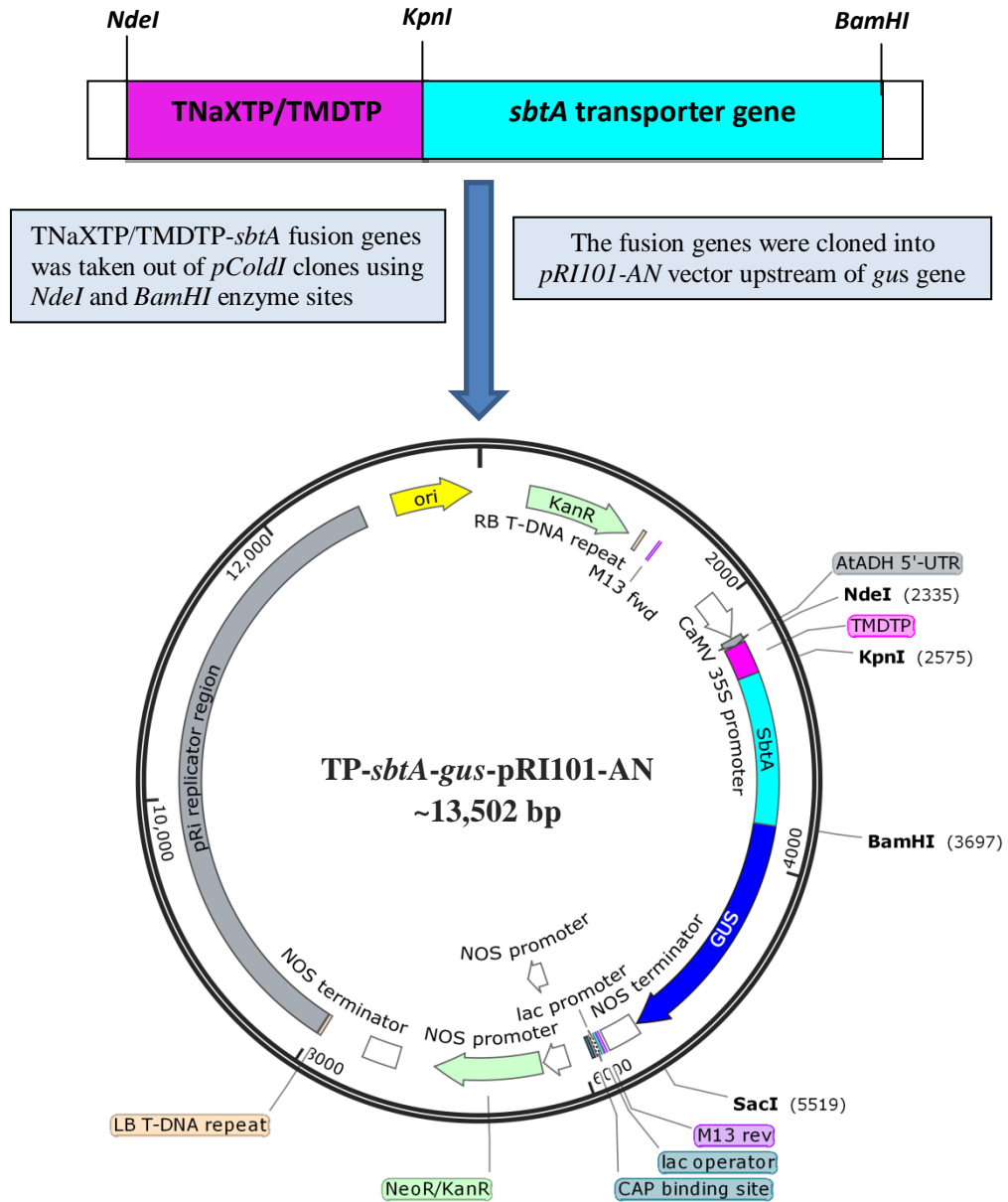


Figure 2-8: A schematic depicting cloning of fusion genes (TNaXTP/TMDTP-*sbtA* gene) upstream of the *gus* gene into *pRI101-AN* vector. The vector map has been generated by SnapGene® 4.0.5 software using TMDTP as a reference transit peptide.

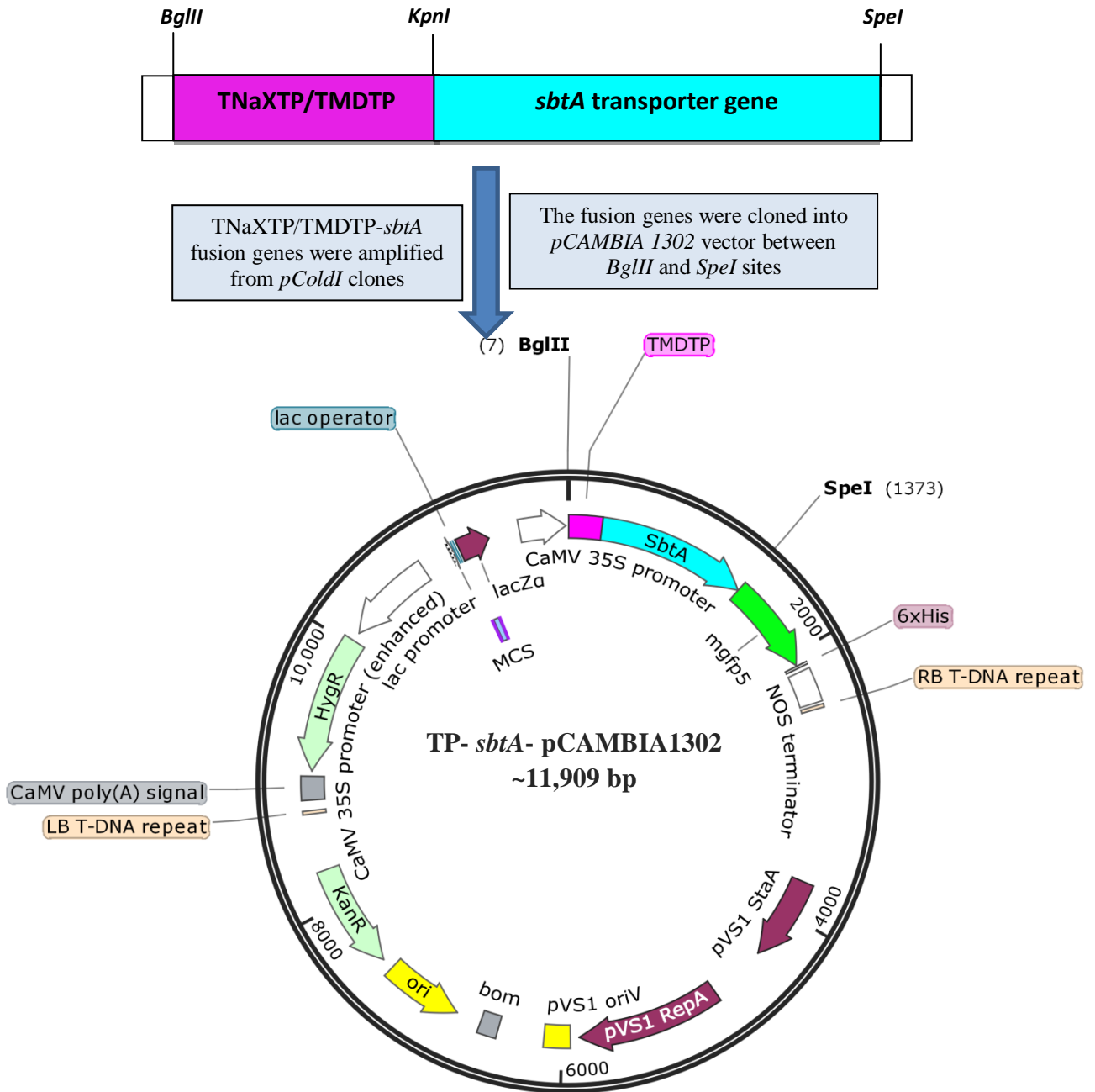


Figure 2-9: A schematic depicting cloning of fusion genes (TNaXTP/TMDTP-*sbtA* gene) into *pCambia1302* vector. The vector map has been generated by SnapGene® 4.0.5 software using TMDTP as a reference transit peptide.

2.2.9. Digestion of inserts (PCR amplicons) and vectors using appropriate restriction enzymes

For making the choice of the restriction enzyme site to be incorporated in the insert to be cloned, *in-silico* analysis of the DNA of interest was done using NEB cutter (Vincze et al., 2003). The enzyme sites that had single occurrence in the vector and absence in the DNA insert were chosen and incorporated into the primers used for PCR. The amplicons and the vector of interest were then digested with chosen enzymes to generate compatible ends. The details of the enzymes used in the study are mentioned in Table 2-7.

Table 2-7: The details of the restriction enzymes used for cloning of various DNA fragments

S. No.	Name of RE	Recognition sequence of RE
1.	<i>NdeI</i>	CA ^V TATG GTAT _Δ AC
2.	<i>KpnI</i>	GGTAC ^V C C _Δ CATGG
3.	<i>BamHI</i>	G ^V GATCC CCTAG _Δ G
4.	<i>SacI</i>	GAGCT ^V C C _Δ TCGAG
5.	<i>SpeI</i>	A ^V CTAGT TGATC _Δ A
6.	<i>BglIII</i>	A ^V GATCT TCTAG _Δ A

For setting up the digestion reaction all the components given in Table 2-8 were mixed and incubated at 37°C in a circulating waterbath for 1 hr to 12 hrs (overnight) depending on the manufacturer's instructions.

Table 2-8: The composition of reaction mixture used for setting up restriction reactions in the study

S.No.	Component	Amount
1.	DNA	1 μg
2.	Buffer (CutSmart®) (10X)	1X
3.	Enzyme 1	1Unit
4.	Enzyme 2	1Unit
5.	Nuclease Free Water	Volume make-up

Note: Since all the enzymes used in the study were HF enzymes all of them had compatibility with CutSmart® Buffer. Hence, CutSmart® Buffer was used for all digestions.

2.2.10. Preparation of chemically competent cells of *E. coli* DH5 α

Requirements

- 100 mM CaCl₂ (50 mL) - Autoclaved and pre-cooled to 4°C
- 60% glycerol (Autoclaved - pre-cooled to 4°C)
- 5 mL LB broth for 1° inoculation - Autoclaved
- 50 mL LB broth for 2° inoculation - Autoclaved
- Oakridges and Microcentrifuge tubes (1.5 and 2 mL) - Autoclaved and stored at 4°C
- Autoclaved tips (1 mL and 200 μ L)
- Autoclaved cut tips (1 mL and 200 μ L)
- LB Agar plates

Procedure

1. *E. coli* DH5 α cells were streaked on a LB Agar plate and grown overnight at 37°C.
2. For 1° inoculation, single colony of *E. coli* DH5 α cells was picked and inoculated into 5 mL LB broth. The culture was allowed to grow overnight (14-16 hrs) at 37°C with shaking at 150 rpm.
3. The following day, 1% inoculum from 1° culture was taken, i.e. 500 μ L for 50 mL medium, and secondary culture was inoculated and incubated at 37°C for 2-3 hrs, until the O.D. at 600 nm reached 0.3-0.4.
4. After this, the culture was chilled on ice for 10 min and subsequently centrifuged at 6000 rpm, 4°C for 8 min. Supernatant was discarded, and any traces of media left in the tube was drained by keeping the tubes inverted on a paper towel.
5. The pellet was slowly resuspended in ice cold 100mM CaCl₂ (30 mL for 50 mL culture) using cut tips with repeated pipetting. It was made sure not to keep tubes out of ice for more than 2-3 consecutive min.

6. After gently resuspending, the cells were again centrifuged at 8000 rpm, 4°C for 8 min. The supernatant was discarded and the pellet was resuspended in 100mM CaCl₂ (2 mL CaCl₂ for 50 mL culture).
7. To the suspended cells, 0.9 mL of 60% glycerol was added in order to make the final concentration to 15%. The cells were gently mixed with glycerol by tapping and then distributed as 300 µL aliquots in labelled and pre-chilled microcentrifuge tubes.
8. The tubes were immediately frozen (in liq.N₂) and stored at -80°C.

2.2.11. Transformation of recombinant DNA into *E. coli* DH5α competent cells

1. The competent cells of *E. coli* (DH5α) were thawed on ice (for approximately 20-30 min). The thawed cells were then distributed into 3 different microcentrifuge tubes (100 µL each).
2. To one of the part (having 100 µL of competent cells), ligation mixture was added and mixed by gentle tapping. This was followed by incubation on ice for 30 min. The remaining two were used as positive and negative controls for transformation.
3. A heat shock was given at 42°C (water bath) for 90s and the tubes were incubated on ice for 2-3 min. Immediately after heat shock, 800 µL of LB broth was added to the cells to help them to recover after heat shock.
4. The cells were thereafter incubated at 37°C for 45 min in an orbital shaker with slow speed. After 45 min, the cells were centrifuged at 3000 rpm, for 5 min at RT.
5. From each of the tubes, 750 µL of LB media was removed from the supernatant without disturbing the pellet and the cells were suspended in the remaining 150 µL supernatant.
6. Tubes containing the ligation mixture and the negative control were spread on LB agar plate with antibiotic/s (depending on resistance in plasmid transformed) while for positive control only Luria Agar plate was used.
7. The plates were incubated overnight (12-14 hrs) at 37°C. The bacterial colonies were checked for recombinant plasmid by isolation of plasmid and restriction analysis.

Note: In a successful experiment, one would observe few to many number of colonies on the plate containing the ligation mixture. There should be no colonies on negative control plate and there should be a bacterial lawn on the positive control plate.

2.2.12. Screening of putative transformants

The colonies obtained were screened for the presence of gene by performing colony PCR, restriction analysis and DNA sequencing. For the positive colonies, master plate was prepared which was further used to streak a fresh plate to obtain single colonies.

2.2.12.1. Colony PCR

Procedure

1. A small portion of the bacterial colony was picked up and resuspended in 30 μ L of NFW in a 1.5 mL microcentrifuge tube using a sterile toothpick or pipette tip.
2. The sample was then lysed at 100°C for 4 min and thereafter incubated on ice for 4 min.
3. After incubation, the sample was centrifuged at 12,000 rpm for 4 min at RT.
4. The supernatant obtained after centrifugation was used as PCR template. PCR reaction was set up using gene specific primers and about 3-5 μ L of the supernatant was used as template.
5. The amplicons were analysed on agarose gel for the presence of the desired gene.

2.2.12.2. Restriction analysis

The presence of gene in the positive clones was also confirmed by digesting the recombinant plasmids with specific restriction enzymes and the digested samples were compared with control samples to confirm the expected band size (depending on the size of the gene). Samples that gave bands of expected sizes were sent for DNA sequencing. The sequencing results were also cross verified by NCBI blast wherein the gene sequences were aligned with the sequences obtained after performing DNA sequencing.

2.2.13. Preparation of glycerol stocks

For preparing glycerol stocks, a single colony from the positive clones was inoculated in 5 mL of LB broth medium containing appropriate antibiotic/s and was allowed to grow overnight at 37°C with shaking at 200 rpm (temperature requirements differ for different

cells, for *E. coli* DH5 α , the temperature required is 37°C). A volume of 800 μ L of this bacterial culture was mixed with 200 μ L of 100% sterile glycerol in cryovials (labelled) and immediately transferred to a deep freezer at -80°C for long term storage.

2.2.14. Preparation of chemically competent cells of *Agrobacterium tumefaciens* strain GV3101

1. A single colony of *Agrobacterium tumefaciens* strain GV3101 was inoculated in 5 mL of LB broth medium (1 $^{\circ}$ culture) containing rifampicin (50 μ g/mL) and gentamicin (30 μ g/mL). The cells were grown at 28°C for 36-48 hrs at 220 rpm in an orbital shaker.

Note: Since *Agrobacterium tumefaciens* strain GV3101 harbors Ti plasmid pMP90, which codes for gentamicin resistant gene and its chromosomal background contains rifampicin resistance gene, therefore both these antibiotics were added into the culture medium while preparing competent cells.

2. For preparing 2 $^{\circ}$ culture, 1% of primary culture was inoculated in 50 mL of LB broth with both the antibiotics and under the same conditions.
3. The cells were grown till O.D. at 600 nm wavelength reached 0.3-0.4 (8-10 hrs) and then were pelleted by centrifugation at 4500 rpm for 10 min at 4°C.
4. After centrifugation, the supernatant was discarded and the pellet was resuspended in 20 mM of CaCl₂ solution (5 mL).
5. The cells were then centrifuged at 4500 rpm at 4°C for 5 min and the pellet was again resuspended in 1mL CaCl₂ solution (20 mM).
6. The cells were distributed as aliquots of 200 μ L into sterile microcentrifuge tubes. Tubes were frozen in liquid N₂ and immediately stored in a deep freezer at -80°C.

2.2.15. Transformation of recombinant DNA into chemically competent *Agrobacterium* cells

1. The competent cells (aliquot of 200 μ L) were taken out from the deep freezer (-80°C) and thawed on ice for 20-30 min.

2. After thawing, the cells were distributed into three microcentrifuge tubes. To one of the microcentrifuge tubes, recombinant DNA was added and mixed by pipetting while the remaining two were used as positive and negative controls.
3. The tubes were then incubated on ice for 15-20 min followed by freezing in liquid N₂ for 5 min and then giving a heat shock at 37°C for 5 min. Just after the heat shock, the tubes were incubated on ice for 5 min.
4. After incubation, 800 µL of LB broth was added to the tubes and were incubated at 28°C for 3-4 hrs in an orbital shaker with slow speed.
5. The bacterial cells were collected by centrifugation at 3000 rpm for 5 min at RT. After removing 750 µL of supernatant, the remaining supernatant was mixed properly and plated on the LB Agar plate containing proper antibiotics.
6. The plates were incubated at 28°C for approximately 48 hrs. The bacterial colonies obtained were checked for recombinant plasmid by colony PCR analysis followed by isolation of plasmid and restriction analysis.

2.2.16. Preparation of electrocompetent cells of *Agrobacterium tumefaciens* strain GV3101

1. 1° culture: A single colony of *Agrobacterium tumefaciens* strain GV3101 was inoculated in 5 mL of LB broth containing rifampicin (50 µg/mL) and gentamicin (30 µg/mL). The cells were grown at 28°C for 36-48 hrs at 220 rpm in an orbital shaker.
2. 2° culture: 1% of primary culture was inoculated in 50 mL of LB broth with both the antibiotics and under the same conditions. The cells were grown till O.D. at 600 nm wavelength reached 0.3-0.4 (8-10 hrs) and then were pelleted by centrifugation at 4500 rpm for 10 min at 4°C.
3. The bacterial pellet was resuspended in 10 mL of 1 mM of HEPES buffer (pH-7). The cells were centrifuged at 4500 rpm at 4°C for 5 min and the pellet obtained was again resuspended in 1 mM of HEPES buffer containing 10% glycerol.

Note: The HEPES buffer was filter sterilized through 0.22 µm syringe filter before use.

4. This step was repeated twice and the cells were first resuspended in 2 mL and then in 200 μ L of buffer (1 mM of HEPES buffer + 10% glycerol).
5. Cells were aliquoted in a volume of 100 μ L, transferred to sterile microcentrifuge tubes and immediately stored in a deep freezer at -80°C .

2.2.17. Transformation of electrocompetent *Agrobacterium* cells with recombinant plasmids

1. The competent cells were taken out from the deep freezer (-80°C) and thawed on ice for 20-30 min. 100 μ L of cells were equally distributed into three microcentrifuge tubes.
2. To one of the the microcentrifuge tubes recombinant plasmid DNA was added and mixed properly while the remaining two were used as positive and negative controls.
3. After adding DNA, the cells were again incubated on ice for 15-20 min. Just after incubation the cells were transferred to three chilled electrocuvettes (0.1 cm) which were still kept on the ice.
4. For performing electroporation, MicropulserTM (Electroporation apparatus- BioRad) was used and for electroporation of *Agrobacterium* cells, a preset default program “Agr” was used, in which five pulses (each for 5 sec) at 2.20 kV were given to the sample.

Note: All electroporations were performed on ice.

5. Soon after delivering the pulses, 800 μ L LB broth medium was added into the cuvettes and after proper mixing the entire contents were transferred to 1.5 mL microcentrifuge tubes. The tubes were incubated at 28°C for 3-4 hrs in an orbital shaker at slow speed.
6. The bacterial cells were then collected by centrifugation and were spread on LB agar plates with the appropriate antibiotics and grown at 28°C for about 48 hrs.
7. The colonies were screened by colony PCR and glycerol stocks were prepared from positive colonies which were stored in a deep freezer at -80°C .

2.2.18. Preparation of MS (Murashige and Skoog) medium

The Murashige and Skoog medium (Murashige and Skoog, 1962) was used as the basal medium for all the plant transformation experiments i.e., callusing and plant regeneration. MS basal medium was augmented with different concentrations and combinations of different phytohormones to generate a range of media.

A. Preparation of stock solutions for MS media

The stock solutions for major salts, minor salts, iron, vitamins and amino acids were prepared separately by dissolving the required measured amount of constituents in AMQW (Table 2-9). However iron salts were dissolved in distilled water that was boiled for 30 min to make it oxygen free.

Similarly, separate stock solutions for different phytohormones like 2,4-D, NAA and IBA and BAP were also prepared. For preparing stock solutions of auxins, the required amount of constituents were first dissolved in 2-3 mL of ethanol and then AMQW water was slowly added to make up the requisite volume. Similarly for preparing stocks of cytokinins the constituents were dissolved in 5-10 mL of N/10 NaOH and then AMQW water was used to make up the required volume (Table 2-10). The stock solutions were stored at 4°C and brought to room temperature before use.

Table 2-9: The composition of MS medium and the concentrations of different stock solutions used for preparing MS media

S. No.	Constituent	Stocks	Amount (mg/L)	Amount in stock soln. (mg)	Dissolved in AMWQ	Conc.	Volume used/L
1.	NH ₄ NO ₃	Major elements	1650	16500	500 mL	20X	50 mL
2.	KNO ₃		1900	19000			
3.	CaCl ₂ .2H ₂ O		440	4400			
4.	MgSO ₄ .7H ₂ O		370	3700			
5.	KH ₂ PO ₄		170	1700			
6.	H ₃ BO ₃	Minor elements	6.20	620	500 mL	200X	5 mL
7.	MnSO ₄ .H ₂ O		22.3	2230			
8.	ZnSO ₄ .7H ₂ O		8.6	860			
9.	KI		0.83	83			
10.	Na ₂ MoO ₄		0.25	25			
11.	CuSO ₄ .5H ₂ O		0.025	2.5			
12.	CoCl ₂ .6H ₂ O	0.025	2.5				
13.	Na ₂ EDTA	Iron sources	37.3	373	250 ml	40X	25ml
14.	FeSO ₄ .7H ₂ O		27.8	278			
15.	Glycine	Amino acids & vitamins	2.0	200	250 mL	400X	2.5 mL
16.	Myoinositol		100	10,000			
17.	Nicotinic acid		0.5	50			
18.	Pyridoxine hydrochloride		0.5	50			
19.	Thiamine hydrochloride		0.1	10			
20.	Sucrose		30,000				
21.	Agar		8000				

Table 2-10: The composition of phytohormones stock solutions used in the study

S. No.	Phytohormones	Weight (mg)	Dissolved in	Volume make up by AMQW	Strength
1.	NAA	25	3 mL ethanol	250 mL	1mg =10 mL
2.	IBA	25	3 mL ethanol	250 mL	1mg =10 mL
3.	BAP	25	10 mL N/10 NaOH solution	250 mL	1mg =10 mL
4.	2,4-D	25	10 mL N/10 NaOH solution	250mL	1mg =10mL

B. Preparation of MS medium

For preparation of 1L medium, all the components of MS basal and required phytohormones were added from their respective stock solutions to the measuring flask. Thereafter 30g sucrose was added and dissolved by adding distilled water. In order to make final volume to 1L required volume of AMQW was added. The pH of the medium was adjusted to 5.8 using dilute NaOH or HCl. In a conical flask, 8g of agar was taken, mixed with the prepared media and was heated for 10-15 min for homogenizing agar. When the medium turned transparent, 20 mL of it was poured into culture tubes of dimensions 25×150 mm. The culture tubes were then plugged with nonabsorbent cotton and autoclaved. After sterilization the culture tubes and bottles were taken out and cooled to room temperature. The media was allowed to set and used for inoculation on the next day.

Note: For preparing media for *Agrobacterium* mediated co-culture experiment, antibiotics (kanamycin, hygromycin and cefotaxime) and acetosyringone were added to the media after autoclaving and just before pouring inside the laminar hood.

2.2.19. Transformation of *N. tabacum* by particle bombardment/ biolistics method

Particle bombardment is a physical and direct gene transfer method which allows direct entry of recombinant plasmid DNA into plant cells. The method could be used for the genetic transformation of diverse range of organisms. It is based on the simple principle that DNA-coated microscopic particles (microcarriers) if accelerated at high speed by helium gas within a vacuum travels at such a velocity so as to penetrate the target cells. *N. tabacum* leaves were transformed by recombinant DNA using the Biolistic® PDS-1000/He Particle Delivery System (Bio-Rad). The main purpose of the experiment was to study the transient expression level of *gus* reporter gene in the transformed *N. tabacum* leaves.

2.2.19.1. Preparation of plant material for transformation by particle bombardment

1. The *N. tabacum* seeds (Lehle seeds, USA) were surface sterilized by 0.1% HgCl₂ solution, washed thrice with autoclaved Milli-Q water and then inoculated on the MS basal medium in flasks.
2. The flasks were maintained at 25-28°C temperature in a plant tissue culture room under 16/8 hrs photoperiod.
3. When the leaves attained appropriate size (1.5 to 4 cm²), they were excised from the plants and dissected into 1.5 cm X 1.5 cm sized explants after removing the midrib.
4. The leaf discs were inoculated on MS basal medium and incubated for 12 hrs at 22°C. These leaf discs were further used to perform particle bombardment experiment.



Figure 2-10: Growth stages of *N. tabacum* plantlets grown on MS basal medium: (A) 15-20 days old (B) 20-30 days old and (C) 6-7 weeks old plantlets

2.2.19.2. Isolation and purification of recombinant plasmids using Plasmid Mini Kit (Qiagen)

Note: For particle bombardment, *pRII01-AN-TNaXTP/TMDTP-sbtA-gus* constructs were used.

1. A single bacterial colony (obtained after streaking from glycerol stock) was inoculated in 5 mL of LB broth with kanamycin and was allowed to grow at 37°C for 12-14 hrs at 200 rpm in an orbital shaker.
2. The culture was pelleted by centrifugation at 6000g for 15 min at 4°C and the pellet was resuspended in 0.3 mL of buffer P1 (containing RNase A) by vortexing.
3. To the suspension obtained in step 2, a volume of 0.3 mL of buffer P2 was added and mixed properly by inverting the tubes 4-6 times. The tubes were then incubated at RT for 5 min.
4. After incubation, 0.3 mL of chilled buffer P3 was added and the tubes were again kept on ice for 5 min after mixing the contents properly.
5. The samples were then centrifuged at ~18,000 x g at 4°C for 10 min and the supernatant was collected in 2 mL microcentrifuge tube. Meanwhile, a Qiagen-tip 20 column was equilibrated with 1 mL buffer QBT and the column was allowed to get empty by gravity flow.
6. When the column was completely empty, the supernatant collected in step 5 was transferred to it and allowed to pass through the column resin by gravity flow. The Qiagen-tip 20 column was then washed twice with 2 mL buffer QC.
7. The DNA bound to the column was eluted in 0.8 mL buffer QF in a 1.5 mL microcentrifuge tube.
8. To the eluate, a volume of 0.56 mL of isopropanol was added for DNA precipitation. This was followed by centrifugation at $\geq 15,000$ x g at 4°C for 30 min (to pellet the DNA) and the supernatant was carefully decanted.
9. The pellet was washed with 1 mL of 70% ethanol and centrifuged at 15,000g at 4°C for 10 min.
10. The supernatant was decanted carefully without disturbing the pellet. The pellet was air dried by inverting the tubes on a tissue paper for 20-30 min at RT and dissolved in an appropriate volume of NFW (20-30 μ L).

11. The concentration of DNA was determined by taking the absorbance readings at 260 nm wavelength in a U.V. spectrophotometer.

2.2.19.3. Preparation of microcarriers (gold particles coated with recombinant DNA)

Note: This experiment was performed in triplicates for both the constructs (*pRI101-AN-TNaXTP/TMDTP-sbtA-gus* constructs) hence, a total of six bombardments were performed.

A. Baking of gold particles

1. Weighed 3 mg of gold particles (size: 1 μm) into 1.5 mL sterile microcentrifuge tube and baked overnight at 192°C in a hot air oven.
2. The gold particles were washed with 70% ethanol (1 mL) by vortexing for 5 min, they were then allowed to settle down by incubation at RT for 15 min.
3. The particles were then given a short spin for 5 sec and supernatant was discarded.
4. The residual ethanol was then removed by washing the particles with 1 mL of autoclaved Milli-Q water and vortexing for 1 min and then allowing them to settle at RT.
5. The above step was repeated three times and the pellet obtained in the last step was dissolved in 50 μL glycerol (50%) and mixed properly by vortexing for 5 min.

Note: Vortexing was done before every step as gold particles have the tendency to form clumps.

6. The contents were divided into two tubes and to both the tubes, 5 μL plasmid DNA (0.5 $\mu\text{g}/\mu\text{L}$), 25 μL sterile calcium chloride (2.5 M stock solution) and 10 μL filter sterilized spermidine (0.1 M stock) were added with continuous stirring at a low speed.
7. After addition of all the components, vortexing was done for 3 min and the sample was incubated at RT for 5 min. A short spin was given for 5 sec and supernatant was discarded.
8. To the pellet, 70 μL ethanol (70%) was added (slowly without disturbing) and a short spin was given for 2 sec after which the supernatant was removed by pipette.
9. Again, 70 μL ethanol (100%) was added followed by a short spin for 2 sec and the supernatant was pipetted out.
10. Finally, the pellet was dissolved in 24 μL ethanol (100%) by gentle tapping and continuous vortexing for 3 sec.

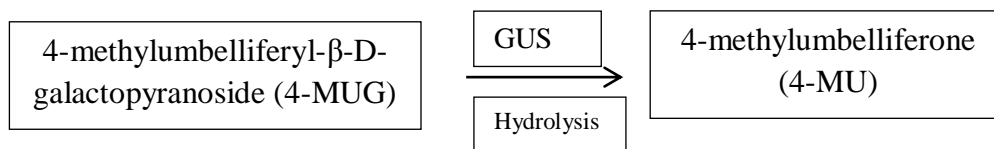
11. From the above suspension (gold particles coated with DNA/microcarriers), 8 μL was spread onto the macrocarrier.

2.2.19.4. Bombardment of macrocarriers on *N. tabacum* leaves

The particle bombardment apparatus was assembled, the rupture disk, the stopping screen and the microcarrier assembly were loaded. DNA coated macrocarriers were fixed on the macrocarrier holder and the distance between the plates (containing leaf discs on MS media) and stopping screen was adjusted. The vacuum was created by pressing the VAC button and holding it until it reached the desired value. Then the FIRE button was pressed until the rupture disk automatically bursted. The bombardment was performed at 1100 lbs/cm³ pressure using 1100 psi rupture disk (Bio-Rad) with the Biolistic® PDS-1000/He Particle Delivery System (Bio-Rad). Once the bombardment was finished, the explants were incubated onto the same medium (MS basal medium) for 48 hrs under 16/8 hrs photoperiod at 25°C in a plant growth chamber. The bombarded leaf discs were then analyzed for the expression of *gus* gene by GUS fluorometric assay by keeping the non- bombarded leaves as negative control for the experiment.

2.2.20. GUS fluorometric assay (Jefferson et al., 1987)

The assay is based on the principle that GUS enzyme acts on its substrate i.e., MUG to form a fluorogenic reaction product called 4-MU (4-Methyl umbelliferone).



Procedure

1. The bombarded leaf discs (10-15 mg) were removed from the media after 48 hrs and were crushed in liquid N₂ using mortar-pestle.
2. The powder was resuspended in 150 μL of extraction buffer [50 mM sodium phosphate buffer (pH-7), 10 mM DTT, 1 mM EDTA (pH-8), 0.1% SDS and 0.1% Triton-X) and the entire contents were transferred to microcentrifuge tubes.
3. The samples were then centrifuged at 13,000 rpm for 20 min at 4°C (to pellet the debris) and the supernatant was collected in a fresh microcentrifuge tube.

- To 90 μL of the supernatant, 10 μL of assay buffer [50 mM sodium phosphate buffer (pH-7), 10 mM DTT, 1 mM EDTA (pH 8), 0.1% SDS, 0.1% Triton-X, 1 mM 4-Methylumbelliferyl- β -D-glucuronide (MUG)] was added and incubated at 55°C for 20 min.
- After 20 min, 20 μL of methanol was added to inhibit intrinsic GUS activity and the further reaction was carried out at 37°C for 2 hrs. The incubation reaction was stopped by addition of 900 μL of 0.2 M sodium carbonate solution and proper mixing of the contents.
- To detect the relative fluorescence of 4-methylumbelliferone (MU), 200 μL of the sample was analysed using spectrofluorometer (Thermoscientific-Fluoroskan Ascent).
- For comparison, all the samples were used in triplicates, extraction buffer was taken as blank and non-transformed sample was taken as control. Readings were taken for the test sample (transformed), MU, blank and negative control. The reading of Blank sample i.e., extraction buffer was subtracted from all the readings.
- MU standard curve was prepared and by its help, the value of 1 Relative Fluorescent Unit (RFU) was calculated and a formula was derived to calculate specific GUS activity.

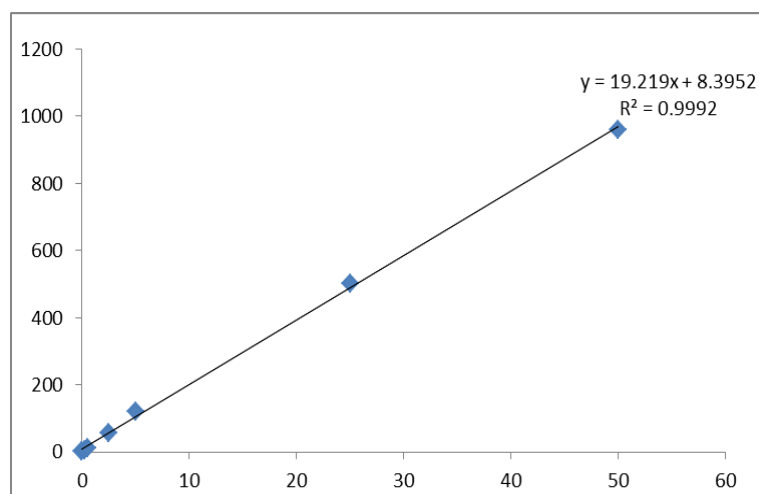


Figure 2-11: Standard curve for the relative fluorescence of 4-methylumbelliferone (MU) standards. The standards were prepared from concentrations ranging from 0-50 μM .

The formula used to calculate specific GUS activity:

$$\text{Number of nanomoles of MU/mg of protein/minute} = 0.07 \text{ X/Y}$$

Where, X = RFU value of sample

Y = Total protein concentration of sample

9. Total soluble protein concentration was determined in the leaf extract supernatant by Bradford assay (Bradford, 1976). A volume of 100 μL of sample was mixed with 100 μL of Bradford reagent and similarly Bovine Serum Albumin (BSA) standard samples were prepared with variable concentrations.
10. The samples were added to a 96 well microtitre plate and incubated in dark at 37°C for 30 min at low shaking speed. The absorbance values were measured at 595 nm by using an ELISA plate reader (STAT FAX 2100).
11. Using the BSA standard curve, total protein concentration was determined for control and the test samples. The values obtained were then used to calculate specific GUS activity.
12. After estimating the total protein concentration and concentration of MU in the test as well as in control samples, Specific GUS activity was calculated and plotted in a graph.

2.2.21. Transformation of *A. thaliana* by floral dip method (Clough and Bent, 1998; Davis et al., 2009; Zhang et al., 2006)

A. thaliana ecotype Col-0 plants were grown (as mentioned in section 2.2.1) in soilrite mixture (Keltech energies ltd.) in plastic pots at 22°C temperature and a photoperiod of 16/8 hr and 75% humidity for about 4-6 weeks. When plants started developing primary bolts main inflorescence shoots were trimmed off as soon as they had bolted to encourage more secondary shoot formation. *A. tumefaciens* strain GV3101 harboring recombinant *pRII01-AN* plasmids (TNaXTP-*sbtA-gus-pRII01-AN* and TMDTP-*sbtA-gus-pRII01-AN*) was grown in LB broth medium containing appropriate antibiotics at 28°C for about 48 hrs which was then used to inoculate secondary culture. The secondary culture was grown overnight until an optimal O.D. of ~1.5 was achieved at 600 nm wavelength. The cells were centrifuged at 4500 rpm for 10 min at RT and were resuspended in equal volume of 5% (w/v) freshly prepared sucrose solution containing 0.02% (v/v) Silwet L-77 (Lehle seeds, USA). All the

aerial parts (inflorescence) of the plants were submerged into above mentioned infiltration medium for 10-20 seconds. Transformed plant pots were completely covered in a plastic bag or aluminium foil to maintain the appropriate humidity and kept under dark for 12 hrs at 22°C. After 12 hrs, the plastic cover was removed and plants were kept under normal growth conditions. For increasing the transformation efficiency, floral dip procedure was repeated after 3-5 days using the same plants. The plants were grown normally until they set seeds, thereafter seeds were collected, surface sterilized and germinated on kanamycin containing MS medium for selection of transformed plantlets. When resistant plantlets (T₁ generation) achieved 2-4 leaf stage (this took around 2-3 weeks), they were transferred to soilrite mix in plastic pots. T₁ plants were grown for 6-8 weeks until the siliques started drying. Thereafter T₂ seeds were collected and again screened on ½ MS medium containing kanamycin and resistant plants were transferred to soilrite mix in plastic pots.

2.2.22. Transformation of *N. tabacum* by *Agrobacterium* mediated co-culture method (Gallois and Marinho, 1995; Krügel et al., 2002)

Agrobacterium mediated co-culture technique consists of incubating plant tissues with an *Agrobacterium* suspension for 20-30 min followed by a co-culture period of 2 to 3 days and then transfer to regeneration medium until the desired response is observed. For obtaining promising results in short duration it is crucial to optimize the concentration of growth regulators to be added in the co-culture and regeneration medium used for plant transformation.

2.2.22.1. Optimization of the concentrations of plant growth regulators (PGRs) for *in-vitro* shoot and root induction in *N. tabacum*

The non-transformed *N. tabacum* leaf explants were cultured on MS media containing several combinations of PGRs and the media showing the best tissue culture responses was selected for inoculation of transformed leaves. The shoots were regenerated on MS media with BAP alone as well as in combination with NAA. For rooting, MS media with four different concentrations of IBA was used (Table 2-11).

Table 2-11: The selected media for tissue culture experiments. The media (M_1 to M_8) were used to study the effect of BAP (with or without NAA) on shoot regeneration frequency of *N. tabacum* leaf explants, while M_0 and M_9 to M_{12} media were tested for rooting efficiency.

Media name	Composition	Media name	Composition
M_0	MS Basal	M_7	MS + 2.0 mg/L BAP + 0.15 mg/L NAA
M_1	MS + 0.5 mg/L BAP	M_8	MS + 2.0 mg/L BAP + 0.2 mg/L NAA
M_2	MS + 1.0 mg/L BAP	M_9	MS + 0.5 mg/L IBA
M_3	MS + 1.5 mg/L BAP	M_{10}	MS + 1.0 mg/L IBA
M_4	MS + 2.0 mg/L BAP	M_{11}	MS + 1.5 mg/L IBA
M_5	MS + 2.0 mg/L BAP + 0.05 mg/L NAA	M_{12}	MS + 2.0 mg/L IBA
M_6	MS + 2.0 mg/L BAP + 0.1 mg/L NAA		

Procedure

N. tabacum cv. xanthi plants were grown in a mixture of 1:0.5:3 ratios of vermiculite, clay and sandy soil at 25-28°C under 16/8hr photoperiod (as mentioned in section 2.2.2). The leaves were taken from 2-3 months old plant and washed with Tween-20 for 15-20 min followed by treatment with 0.1% Mercuric chloride solution for 1 min under sterile conditions inside the laminar hood. It was followed by rinsing with distilled water 4-5 times and left immersed in distilled water until inoculation. Leaf pieces of 2-3 mm² were excised using a scalpel from surface sterilized leaves in petriplates inside the laminar air flow. Thereafter the prepared explants were inoculated with the help of forceps on to the MS medium (M_1 - M_8) over the flame of a spirit lamp. The test tubes were then plugged and kept on test tube stands. The inoculated culture tubes were transferred to the culture room for further growth under 16/8 hrs photoperiod at 22°C. The culture tubes were observed at regular time intervals. Significant responses were photographed and data was recorded. Differentiated shoots were subcultured into fresh media for further shoot differentiation and elongation. Shoots were also subcultured on different media for root formation (M_9 - M_{12}).

The observations were made for the frequency of differentiation of shoots and roots. Observations were also made for the number of differentiated shoots and roots per culture. The growth of existing shoots or differentiated shoots and differentiated roots were observed as per the table provided (Table 2-12).

Table 2-12: The observation parameters for tissue culture responses of *N. tabacum* leaf discs on different compositions of MS medium

Explanation	Measurement of growth	
	Shoot (Height in cm)	Root (length in cm)
No growth	-	-
Low growth	Less than 0.5	Less than 0.5
Medium growth	0.5 – 1.0	0.5 – 1.0
Good growth	1.0 – 2.0	1.0 – 2.0
Excellent growth	More than 2.0	More than 2.0

The experiments were set up in a completely randomized design (CRD) with a minimum of 10 culture tubes per treatment and four replications. Data was analyzed by one way analysis of variance (ANOVA) and the significances of differences among means was carried out using Duncan's multiple range test at $P = 0.05$. The data was analyzed statistically by using SPSS 16.0 (SPSS Inc., Chicago, USA). The results were expressed as a mean \pm SE of four replication.

The media in which the leaf explant depicted highest frequency for caulogenesis and rhizogenesis were used as regeneration and root induction medium for *Agrobacterium* mediated co-culture experiment.

B. *Agrobacterium* mediated co-culture of *N. tabacum* leaf discs**Procedure**

Agroinfection- Incisions were made on the all four sides of explants/leaves using a sterile surgical blade and co-cultured with *Agrobacterium* strain GV3101 (harbouring recombinant *pRII01-AN* and *pCAMBIA 1302* clones) containing acetosyringone (20 µM) for 10-20 min at 28°C, so that agroinfection could take place. (*Agrobacterium* was grown in LB broth medium containing appropriate antibiotics at 28°C for about 48 hrs which was then used to inoculate secondary culture. The secondary culture was grown till an OD₆₀₀ of 0.6-0.8 was attained).

Co-cultivation- After Agroinfection, explants were blotted dry on sterilized filter paper, then transferred to co-cultivation medium (MS salts + phytohormone (PGR*) + Acetosyringone] and kept for two days in dark conditions at 22°C.

Regeneration- After two days of co-cultivation, explants were washed with MS liquid medium containing cefotaxime (500 µg/mL) to kill the excess *Agrobacterium* and then dried on sterile filter paper. After washing explants were transferred to regeneration medium [MS medium+ PGR* + Selection antibiotic** + cefotaxime (250 µg/mL)] for further growth under 16/8 hrs photoperiod at 22°C. Sub-culturing of callus and multiple shoots was performed every 15 days. After 30–45 days of selection on the medium containing antibiotics, putative transgenic shoots were transferred onto shoot and root regeneration media without antibiotics.

*PGRs used in the case of callus induction was 2,4-D (4 mg/L) while for multiple shoot regeneration BAP (2 mg/L) and NAA (0.2 mg/L) was used.

**Selection antibiotic varied according to the vector used. For *pRII01-AN*, Kanamycin (50 mg/L, Himedia) was used as selective agent whereas for *pCAMBIA1302*, Hygromycin (40 mg/L, Himedia) was used.

The callus regenerated on 2,4-D (4 mg/L) medium was analyzed by GUS histochemical assay while from the regenerated shoots, genomic DNA was isolated and gene specific PCR was performed for confirming the presence of transgene.

Root induction, Acclimatization and pot transfer- The regenerated shoots which showed the presence of transgene were transferred to the root induction media for root formation. The regenerated plantlets with well-developed roots were selected for pot transfer. The

plantlets were made free from adhered agar by rinsing them several times in tap water and by cautiously removing the agar using a soft brush without injuring the roots. The rooted plantlets were transferred to mixture of soilrite, vermiculite, clay soil and sandy soil in the ratios of 1:1:1:3 in plastic pots. The pots were progressively acclimatized for less humidity.

2.2.23. GUS histochemical assay (Vitha et al., 2012)

For performing GUS histochemical assay a small callus tissue was excised from the callus regenerated from the leaves transformed by *Agrobacterium* mediated co-culture method as well as from the untransformed callus (negative control) and was incubated in GUS assay staining solution at 37°C for 1-24 hrs.

Components of GUS assay staining solution:

1. 20 mM Phosphate buffer (made of 0.1 M NaH₂PO₄ and 0.1 M Na₂HPO₄) - pH 7.0.
2. 2M X-gluc substrate solution
3. 1 mM K₃Fe(CN)₆
4. 0.1% Triton-X
5. 10 mM EDTA (pH- 8)

On incubation with GUS assay staining solution, X-Gluc i.e., 5-bromo-4-chloro-3-indolyl β-D glucuronide (a component of staining solution) is oxidized by GUS enzyme to form a blue colored precipitate. The images of the stained callus were captured by using the compound microscope from Olympus.

2.2.24. Transformation of *N. benthamiana* by *Agrobacterium* mediated Agroinfiltration method

N. benthamiana plants were grown in a mixture of 1:0.5:3 ratios of vermiculite, clay and sandy soil and plants were maintained at 22-25°C temperature in a plant tissue culture room under 16/8 hrs photoperiod (as mentioned in section 2.2.2). The *Agrobacterium* cells (harbouring recombinant *pCAMBIA1302* constructs) were grown in LB broth medium containing appropriate antibiotics at 28°C for about 48 hrs which was then used to inoculate secondary culture (having 20 μM Acetosyringone). The secondary culture was grown until an optimal OD₆₀₀ of 0.4 was achieved. The cells were harvested by centrifugation at 4500 rpm for 15 minutes and the pellet obtained was resuspended in the equal volume of resuspension solution [10 mM MgCl₂ and 10 mM MES-K (pH 5.6) and 100 μM

acetosyringone]. This solution was kept at RT for 2-3 hrs and used to infiltrate the leaves of 6-8 weeks grown *N. benthamiana* plants. In order to facilitate the entry of *Agrobacterium* cells into the plant, a needleless 2 mL sterile syringe was pressed on the abaxial surface of the leaves. Post- infiltration, the plants were kept under normal growth conditions away from direct light and after 4 days of incubation, the infiltrated leaves were excised and used for the experiments.

2.2.25. Total RNA isolation from the Agroinfiltrated leaves using RNeasy Plant Mini Kit (Qiagen)

Total RNA was also isolated from agroinfiltrated leaves 4 days post-infiltration using RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. All the solutions were prepared in 0.1% DEPC treated water, plasticware and glassware required for RNA work were made RNase-free by 0.1% DEPC treatment and then autoclaved to remove traces of DEPC. RNaseZAP (Sigma) solution was used to clean working area, pipettes and gloves which were used throughout the isolation procedure. The running buffer i.e. 1X TAE was freshly prepared in RNase-free water and electrophoresis unit was cleaned with 0.5% SDS then rinsed with DEPC water and finally wiped with 70% ethanol to remove RNase contamination.

Procedure

1. Weighed 100 mg of agroinfiltrated plant leaves and grounded them to a fine powder using pre-chilled liquid N₂ by the help of mortar and pestle. The powdered sample was quickly transferred into a microcentrifuge tube before it thawed.
2. To the sample, 450 µL of buffer RLT was added and it was then vortexed vigorously to lyse the sample.

Note: 1% v/v β-mercaptoethanol was added to RLT buffer before use.

3. The lysed sample was added into the QIAshredder spin column placed in a 2 mL collection tube and centrifuged at 14,000 rpm for 2 min. The flow through was collected in a 2 mL microcentrifuge tube.
4. To the collected flow through, about 0.5 volume of 100% ethanol (Merck) was added. The mixture was added into the RNeasy Mini spin column, placed in a 2 mL collection tube and centrifuged at 10,000 rpm for 15 sec. The flow through was discarded.

5. A volume of 700 μL of buffer RW1 was added to the above column and it was incubated at RT for 5 min and then centrifuged at 10,000 rpm for 15 sec. The flow through was discarded.
6. A volume of 500 μL of buffer RPE was added to the above column and centrifuged at 10,000 rpm for 15 sec. The flow through was discarded. Again, 500 μL of buffer RPE was added to the above column and centrifuged at 10,000 rpm for 2 min. To dry the membrane, the empty column was again centrifuged for 1 min.
7. To elute out the RNA from column, 30 μL RNase-free water was added to the column which was placed in 1.5 mL microcentrifuge tube and incubated for 5-10 min at RT and then centrifuged at 10,000 rpm for 1 min.
8. The eluted RNA sample was analyzed on 1% agarose gel and the remaining sample was stored at -80°C .

2.2.26. cDNA synthesis using QuantiTect[®] Reverse Transcription kit (Qiagen)

For cDNA synthesis, RNA samples were initially treated with gDNA wipeout buffer provided with the kit for removing genomic DNA contamination (if any). For this, RNA samples were taken out from -80°C deep freezer and thawed on ice. After thawing about 1-2 μg of template RNA was mixed with 2 μL of gDNA wipeout buffer (7X) and RNase-free water (variable) and was incubated at 42°C for 2-5 min in a PCR machine (Veriti[™] Thermal Cycler, Applied Biosystems). The mixture (~14 μL) was then immediately transferred to ice and used for synthesis of cDNA by reverse transcription reaction.

Table 2-13: The composition of reverse transcription reaction mixture used for cDNA synthesis

S.No.	Component	Required Volume
1.	Template RNA	14 μL
2.	Quantiscript reverse transcriptase	1 μL
3.	Quantiscript RT buffer (5X)	4 μL
4.	RT primer mix	1 μL
5.	Total	20 μL

The reaction mixture mentioned above was incubated at 42°C for 15-30 min for cDNA synthesis and thereafter at 95°C for 3 min in order to inactivate the reverse transcriptase enzyme. The cDNA synthesized was stored at -20°C.

2.2.27. Diagnostic PCRs of genomic DNA and cDNA of Agroinfiltrated leaves

The genomic DNA from agroinfiltrated leaves was used to amplify *sbtA* and *mgfp5* using gene specific primers for full-length amplicon while cDNA was used to perform PCR using gene specific primers for partial amplicon. Primers used along with their details have been mentioned in Table 2-3 in section 2.2.4.

Table 2-14: PCR conditions for diagnostic PCRs performed using genomic DNA and cDNA as template

S. No.	Template used	Primers used	PCR conditions						Expected size (bp)
			Initial denaturation	Denaturation	Annealing	Extension	Final extension	No. of cycles	
1.	Genomic DNA from TNaXTP- <i>sbtA</i> - <i>mgfp5</i> and TMDTP- <i>sbtA</i> - <i>mgfp5</i> infiltrated leaves of <i>N. Benthamiana</i>	<i>sbtA</i> gene specific primers	95°C, 3 min	95°C, 1 min	56°C, 1:30 min	72°C, 1 min	72°C, 3 min	35	~ 1.1 kb
2.	Genomic DNA from TNaXTP- <i>sbtA</i> - <i>mgfp5</i> and TMDTP- <i>sbtA</i> - <i>mgfp5</i> infiltrated leaves of <i>N. benthamiana</i>	<i>mgfp5</i> gene specific primers	95°C, 3 min	95°C, 1 min	56°C, 1:30 min	72°C, 1 min	72°C, 3 min	35	~ 711 bp
3.	cDNA from TNaXTP- <i>sbtA</i> - <i>mgfp5</i> and TMDTP- <i>sbtA</i> - <i>mgfp5</i> infiltrated leaves of <i>N. benthamiana</i>	<i>sbtA</i> gene specific primers (partial amplicon)	95°C, 3 min	95°C, 30 sec	56°C, 30 sec	72°C, 1 min	72°C, 3 min	35	~ 205 bp
4.	cDNA from TNaXTP- <i>sbtA</i> - <i>mgfp5</i> and TMDTP- <i>sbtA</i> - <i>mgfp5</i> infiltrated leaves of <i>N. benthamiana</i>	<i>mgfp5</i> gene specific primers (partial amplicon)	95°C, 3 min	95°C, 30 sec	54°C, 30 sec	72°C, 1 min	72°C, 3 min	35	~239 bp

2.2.28. Total protein isolation using G-Biosciences kit

Total protein was isolated from *N. benthamiana* leaves four days post-infiltration using Total Protein Extraction kit, TPE™ (G-Biosciences).

Procedure

1. Weighed 100 mg of agroinfiltrated leaves and grounded them to a fine powder using liquid N₂ in pre-chilled mortar-pestle.
2. To the powdered sample, 250 µL of TPE Buffer-I containing 5 µL protease inhibitor cocktail (Sigma) was added and after proper mixing, the lysate was transferred into a microcentrifuge tube.

Note: 5% PVPP and 2 mM DTT were also added to TPE Buffer-I before use to enhance the yield of proteins.

3. A volume of 40 µL of warm TPE Buffer-II was added to the tube and it was vortexed for 30 sec followed by incubation at 100°C for 30 sec.
4. The samples were vortexed and heated repeatedly until a clear solution was observed. At last the tubes were incubated at 100°C for additional 10 min.
5. The tubes were centrifuged at 15,000 g for 10 min at 4°C, the supernatant (total protein extract) was collected in a fresh tube and the pellet containing the debris was discarded. The protein samples were stored at -20°C.

2.2.29. Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE is a technique of separating proteins based on the differences in their molecular weight. Since the proteins covered by SDS are negatively charged therefore when they are placed in an electric field, they migrate towards the anode (positively charged electrode) and get separated by a molecular sieving effect based on size. The polyacrylamide gel is made up of two gels viz., the stacking and the resolving gel. The stacking gel (5%) is a lower polyacrylamide concentration gel that is placed on top of the more concentrated resolving gel (12%). The function of the stacking gel is to concentrate/pack all the proteins in one band, so that they will start migrating in the resolving gel all at the same time. The resolving gel separates the proteins in the sample based on their molecular weight.

Table 2-15: Composition of SDS-PAGE gel used in the study

S.No.	Component	Volume in 12% Resolving gel (5 mL)	Volume in 5% Stacking gel (2.5 mL)
1.	30% Acrylamide solution	2 MI	415 μ L
2.	1.5 M Tris (pH 8.8)	1.25 mL	-
3.	0.5 M Tris (pH 6.8)	-	315 μ L
4.	10% SDS	50 μ L	25 μ L
5.	10% Ammonium per sulfate (APS)	50 μ L	25 μ L
6.	N, N, N', N'-Tetramethyl-ethylenediamine (TEMED)	5 μ L	2.5 μ L
7.	Water	1.65 mL	1.7 mL

Note: 30% Acrylamide solution was prepared by dissolving 29.2g Acrylamide and 0.8g N,N'-Methylene Bis-Acrylamide in 100 mL AMQW.

Procedure

Protein samples (10 μ L, 50-100 μ g in concentration) were mixed with equal volume of 2X SDS gel loading buffer/ Sample lysis buffer [100 mM Tris-Cl (pH-6.8), 20% glycerol, 4% SDS, 10% β -mercaptoethanol and 0.1% bromophenol blue] and incubated at 100°C for 10 min. Samples were loaded onto the gel placed in vertical electrophoresis apparatus (Bio-Rad) and electrophoretic run was carried out at 80-100 volts/cm for 2-3 hrs in 1X Tris-glycine buffer (0.125 mM Tris base, 0.96 M Glycine and 0.5% SDS).

After the run, the gel was carefully removed and was stained with Coomassie Brilliant Blue (CBB) R-250 staining solution (0.1% CBB R-250 and Methanol: Water: Glacial acetic acid in the ratio of 5:4:1) for about 3 hrs. After staining, the gel was destained overnight in the destaining solution containing 30% methanol and 10% acetic acid. The gel was visualized by the gel documentation system (Bio-Rad) under white light.

2.2.30. Western Blot Analysis

Western Blot analysis helps in identifying specific protein from a complex mixture of proteins extracted from cells. The proteins are first separated on the basis of size by running them on a gel, then transferred to a solid support (membrane) and thereafter the target protein is marked using a specific primary and secondary antibody. If the protein is present, the antibodies will bind to it and this could be checked by the addition of an enzyme conjugated secondary antibody (Horse-radish Peroxidase or Alkaline Phosphatase) specific against the primary antibody. To confirm the expression of the SbtA fusion proteins, the protein samples were run on SDS-PAGE and thereafter transferred to polyvinylidene fluoride (PVDF) membrane (0.45 μm) by wet transfer method using western transfer equipment (Biorad).

Requirements

1. Transfer buffer: 45 mM Tris base, 39 mM Glycine and 20% Methanol
2. Tris buffered saline (TBS) (pH 7.5): 50 mM Tris base and 50 mM NaCl
3. TBST [TBS (Tris Buffer Saline) + Tween-20]/ Washing Buffer: 50 mM Tris base and 50 mM NaCl; pH 7.5 + 0.05% Tween-20.
4. Blocking buffer: 5% skimmed milk powder in 1X TBST
5. Poncaeu-S solution: 0.1% w/v Poncaeu-S in 5% v/v acetic acid
6. Primary antibody solution: anti-GFP primary antibody raised in rabbit (GFP tag polyclonal antibody; Thermo Fisher Scientific; A-11122 with 1:2000 dilution in blocking buffer or anti-Tic110 primary antibody raised in rabbit (polyclonal antibody; Agrisera; AS08293 with 1:1000 dilution in blocking buffer).
7. Secondary antibody solution: Horseradish peroxidase (HRP) conjugated anti-rabbit secondary antibody raised in goat (Goat anti-rabbit IgG - HRP; Merck; 621140380011730 with 1:5000 dilution.
8. Substrate solution: 3, 3', 5, 5'-Tetramethylbenzidine (TMB) Liquid Substrate System for Membranes (Himedia).

Procedure

1. After running the SDS-PAGE gel, the gel area consisting of the desired bands was cut and immersed in chilled transfer buffer for 15 min. The sponges and filter pads were also soaked in the pre-cooled transfer buffer. The membrane was first activated by incubating in 100% methanol for 10 min and then incubated in the transfer buffer.

Note: A pre-run and unstained SDS polyacrylamide gel was used for the transfer of protein on the PVDF membrane.

2. The transfer cassette containing equilibrated sponges, filter pads, gel and membrane was assembled in such a manner that there were no air bubbles and the gel and the membrane were in tight contact. Assembly was set with the gel at the negative end and the membrane at the positive end.
3. For electrotransfer, the cassette was placed in tank containing transfer buffer, cooling pack was placed to avoid generation of excess heat and the transfer was done at 50 volts/cm for 2.5 hrs inside the cold room.
4. After the transfer process, the membrane was first stained with Ponceau-S to check the transfer efficiency (pink colored protein bands are visible after staining) and then the Ponceau-S stain was removed by washing with autoclaved Milli-Q water. The membrane was thereafter incubated with blocking buffer for 4-5 hrs at RT with slow shaking.
5. After 5 hrs, the membrane was rinsed with TBS for 7 min (two times) and then once with TBST for 5 min.
6. After properly washing the blocking buffer, the membrane was incubated with primary antibody solution for 12-15 hrs in cold room at 25 rpm. This was followed by washing as mentioned in step 5.
7. Once washing was over, the membrane was incubated with HRP conjugated secondary antibody solution for 2 hrs at RT with shaking at 25 rpm. The membrane was again washed as mentioned in step 5 then incubated with substrate solution incubated for 15-20 min in dark. The reaction was stopped by adding distilled water and bands obtained (blue in color) on the membrane were analyzed by gel documentation system.

2.2.31. Protoplast isolation from agroinfiltrated *N. benthamiana* leaves (Jabnourne et al., 1993; Tahami and Chamani, 2018)

Requirements

1. Washing solution-I (WS-I): 0.5 Mannitol, 4 mM MES-K (pH-5.5) and 20 mM KCl
2. Enzyme solution: 1.5% cellulase R-10, 0.4% macerozyme R-10, 0.4 M mannitol, 20 mM KCl, 20 mM MES (pH-5.5), 0.1% BSA and 10 mM CaCl₂
3. Washing solution-II (WS-II): 154 mM NaCl, 125 mM CaCl₂, 5 mM KCl and 2 mM MES (pH-5.7)
4. MMG solution: 0.4 M Mannitol, 15 mM MgCl₂ and 4 mM MES (pH-5.7)

Procedure

1. The Agroinfiltrated leaves were cut into 0.5-1 cm pieces using a sterile scalpel blade and transferred into plastic petriplate containing 3 mL of WS-I.
2. The petriplates containing leaf discs were gently mixed by shaking with hand. The WS-I solution was slowly removed by pipette and freshly prepared enzyme solution was added to it.
3. The mixture was incubated in a shaker at 25-28°C for 4-5 hrs with low speed shaking at 50-70 rpm in dark.
4. After the incubation, the mixture was transferred into a falcon tube and an equal volume of WS-II was added to it. After gentle mixing, the complete mixture was passed through nylon mesh filter of 80 µM (to remove leaf debris).
5. The filtrate was centrifuged at 50g for 10 min and the pellet containing protoplasts was washed 2-3 times with WS-II and finally dissolved in 1 mL of MMG solution.

2.2.32. Confocal microscopy analysis

For confocal imaging, protoplasts samples (2-3 µL) were mounted with an equal volume of Antifade Mounting Medium vectashield (Vector laboratories). The images were acquired by LEICA TCS SP5 Confocal laser scanning microscope equipped with a 63X oil immersion objective. The wavelength of 488 nm was used for excitation of both GFP and chlorophyll. Emission of GFP was recorded at 499-550 nm and chlorophyll autofluorescence was recorded at 630-735 nm.

2.2.33. Regeneration of stable transgenic plants of *N. tabacum* by *Agrobacterium* mediated co-culture method

The leaf discs of *N. tabacum* were co-cultured with *Agrobacterium* cells harboring recombinant *pCAMBIA1302* constructs (TNaXTP/TMDTP-*sbtA-pCAMBIA1302*). The procedure has been mentioned in detail in sec. 2.2.22. The multiple shoots obtained on shoot induction media {MS + BAP (2.0 mg/L) + NAA (0.2 mg/L) + hygromycin (40 mg/L) + cefotaxime (250 mg/L)} were visualized by fluorescence microscopy and the presence of gene was tested by gene specific PCR analysis. After 45 days of culture on shoot induction media, the transgenic shoots were then cultured on antibiotic free shoot induction medium {MS + BAP (2.0 mg/L) + NAA (0.2 mg/L)} and later on to the root induction medium (MS Basal). The plantlets with well developed roots were transferred to plastic pots (described in detail in sec. 2.2.22) for acclimatization/hardening. After a month the plants were transferred to the larger sized pots where they were grown till maturity and the seeds were harvested.

CHAPTER 3

RESULTS AND DISCUSSIONS

3.1. Distribution and arrangement of inorganic carbon transporters in cyanobacteria

Cyanobacteria or blue green algae are one of the most diverse groups of organisms which colonize various niches including freshwater (rivers, ponds and lakes), polar caps, hot springs, alkaline, estuarine, open as well as saline oceans. During the ancient environmental adversities these organisms along with certain algal species developed active systems known as carbon concentrating mechanisms (CCM) to concentrate CO_2 at the site of Rubisco activity which in turn led to a positive impact on photosynthetic ability of these organisms (von Caemmerer and Evans, 2010). This strategy was also adopted by the land plants and they developed certain anatomical features (Kranz anatomy) to increase the carboxylation activity of Rubisco thus reducing its photorespiratory losses (Raven et al., 2008).

CCMs in cyanobacteria and proteobacteria as a whole include (a) Rubisco and carbonic anhydrases enclosing micro-compartments known as carboxysomes and (b) C_i transporters, which regulate the $\text{CO}_2/\text{HCO}_3^-$ influx and efflux at the site of Rubisco activity and hence lead to a marked increase in the CO_2 concentration in the vicinity of Rubisco.

Cyanobacteria are dependent on active accumulation of C_i to achieve a satisfactory rate of CO_2 fixation and growth (Badger et al., 2002). The efficacy of any CCM relies on the ability to minimize the loss of CO_2 from the CO_2 elevation zone or the dissolved C_i accumulation zone (Price et al., 2007). It can be accomplished in four ways. First, the accumulation of bicarbonate instead of CO_2 reduces the chances of C_i leakage, since the former is less permeable through the plasma membrane as compared to the latter. Secondly, the absence of CA (carbonic anhydrase) activity in the cytosol minimizes leakage due to wasteful conversion to CO_2 and subsequent diffusion back to the external medium. Thirdly, the carboxysome protein shell is proposed to have specifically charged pores which allow entry of only specific polar moieties like HCO_3^- , RuBP etc., while retarding the entry/exit of non-polar molecules like CO_2 and O_2 which helps to reduce photorespiration and increase efficient CO_2 fixation. Finally, thylakoid located CO_2 pumps (e.g., NdhI_3 in *Synechococcus* sp.) play a key role in recycling CO_2 that leaks from the carboxysome back into the HCO_3^- pool/cytosol.

The type of Ci transporters present in cyanobacteria is proposed to be governed by the habitat in which they dwell (Badger et al., 2005). Basically, they are of five different types which can be broadly categorized into two types viz., (a) Bicarbonate transporters (Substrate - HCO_3^-) and (b) CO_2 uptake complexes (Substrate - CO_2). The bicarbonate transporters include BCT1, BicA and SbtA. BCT1 is an inducible, high affinity bicarbonate transporter, encoded by *cmpABCD* operon and comes under traffic ATPase family. BicA and SbtA are sodium dependent and probably act as $\text{Na}^+/\text{HCO}_3^-$ symporters (Price and Howitt, 2010). The CO_2 uptake systems include the NdhI₃ (*ndhF₃/ndhD₃/cupA*) and NdhI₄ (*ndhF₄/ndhD₄/cupB*) complexes, which are based on specialized NADPH dehydrogenase complex.

Higher affinity Ci transporters are switched on only when the Ci concentration in the environment becomes limiting. There is less variety in terms of transporters existing in marine species like *Prochlorococcus* sp. which possess a limited range of HCO_3^- transporters and no CO_2 uptake systems but at the same time certain marine species exhibit enormous variations in Ci transporter complement. This fluctuation is observed due to varying Ci levels and temperature in different habitats. The freshwater species also show a range of Ci transporters. The species which occupy lake environments and peak in their abundance during summer contain the most complete complement of transporters, being correlated with the most extreme environmental fluctuations in Ci levels, temperature, oxygen and nutrients. Species with reduced sets of transporters are correlated with growth in symbiotic environments, thermal hotspots and calcareous rocks, where environmental fluctuations are expected to be less extreme.

BicA, a sodium dependent bicarbonate transporter is present in almost all cyanobacterial species but is absent in *Gloeobacter* sp. This cyanobacterial species is slow growing in nature and is found in an environment which is rich in both CO_2 and HCO_3^- resources which makes it a suitable candidate to possess low affinity rather than high affinity bicarbonate transporters (Price et al., 2011). BCT1 and SbtA are abundant in β -cyanobacteria but lack presence in α -cyanobacteria. Among the CO_2 uptake systems, NdhI₄ complex is present in all cyanobacterial species whereas NdhI₃ is thought to exist only in certain fresh water species (Table 3-1).

Table 3-1: Distribution of inorganic carbon transporters (CO₂ and HCO₃⁻ uptake systems) in α - and β -cyanobacteria

Strains	CO ₂ uptake Systems		HCO ₃ ⁻ uptake systems		
	NdhI ₃	Ndh I ₄	BCT1	SbtA	BicA
β-cyanobacteria					
<i>Gloeobacter violaceus</i> PCC 7421	+	+	+	-	-
<i>Thermosynechococcus elongates bp-1</i>	+	+	+	-	-
<i>Synechocystis</i> 6803	+	+	+	+	+
<i>Nostoc punctiforme</i> ATCC 29133	+	+	+	-	+
<i>Anabaena</i> 7120	+	+	+	+	+
<i>Anabaena variabilis</i> 29413	+	+	+	+	+
<i>Synechococcus</i> 6301	+	+	+	+	+
<i>Synechococcus</i> 7942	+	+	+	+	+
<i>Synechococcus</i> sp. PCC 7002	+	+	-	+	+
<i>Arthrospira platensis</i> NIES-39	+	+	+	+	+
<i>Trichodesmium erythraeum</i> IMS101	+/?	+/?	-	-	-
<i>Cyanothece</i> sp. PCC 7424	+	+	+	+/?	-
<i>Cyanothece</i> sp. PCC 7425	+	+	+	+/?	-
<i>Acaryochloris marina</i> MBIC 11017	+	+	-	+	-
<i>Cyanothece</i> sp. PCC 8801	+	+	+	+	-
<i>Cyanothece</i> sp. ATCC 51142	+	+/?	+	+	+
<i>Synechococcus</i> sp. RCC 307	-	+	+/?	-	-
<i>Microcystis aeruginosa</i> NIES-843	+	+/?	+	+/?	+/?
<i>Synechococcus</i> sp. JA-3-3Ab	+	+	+	-	-
<i>Synechococcus</i> sp. JA-2-3B'a (2-13)	+	+	+	-	-
α- cyanobacteria					
<i>Synechococcus</i> 9902	-	+	-	-	-
<i>Synechococcus</i> sp.WH 8102	-	+	-	-	-
<i>Synechococcus</i> sp. WH 7803	-	-	-	-	-
<i>Synechococcus</i> sp.PCC 9605	-	+	-	-	-
<i>Synechococcus</i> sp. CC 9311	-	+	-	-	-
<i>Synechococcus</i> sp. BL107	-	+/?	-	-	-
<i>Prochlorococcus</i> 9313	-	-	-	+	-
<i>Prochlorococcus</i> SS120	-	-	-	+	-
<i>Prochlorococcus</i> 9312	-	-	-	+	+/?
<i>Prochlorococcus</i> MED4	-	-	-	+	-
<i>Prochlorococcus marinus</i> str. NATL1A	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. NATL2A	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. AS9601	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. MIT 9515	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. MIT 9303	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. MIT 9301	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. MIT 9215	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> str. MIT 9211	-	-	-	-	+/?
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str.CCMP1375	-	-	-	+	+/?
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str.CCMP 1986	-	-	-	+	+/?

The transporter complexes for both CO₂ and HCO₃⁻ uptake are located on the plasma membrane of the cyanobacteria. The only exception to it is the NdhI₃ uptake system for CO₂ which is present on the thylakoids. In *Gloeobacter*, this complex is situated at the plasma membrane itself as it lacks thylakoids. The fresh water species (abundantly the β -cyanobacteria) possess a wide array of transporters whereas the α -cyanobacteria have few uptake systems. Among the α -cyanobacterial HCO₃⁻ uptake systems, BCT1 and SbtA are absent whereas BicA is widely distributed. The thylakoid located high affinity transporter

(NdhI₃) is a less abundant CO₂ transport system in α -cyanobacteria as compared to its low affinity counterpart (NdhI₄).

There have been several speculations regarding the first evolutionary steps towards developing a cyanobacterial CCM (Badger et al., 2002). The initial step for the same, during the stages of CO₂ decline would have been the evolution of a carboxysome structure for Rubisco in order to accumulate CO₂ near Rubisco and thereby increase the efficiency of CO₂ fixation (Parry et al., 2003). This structure is essential for concentration of CO₂ and hence Ci transporters are ineffectual without it. A carboxysomal carbonic anhydrase would probably also have been acquired at this stage as the rate of a non-catalyzed chemical conversion of HCO₃⁻ to CO₂ would have been too slow to support photosynthetic CO₂ supply. As CO₂ limitation became more severe, the CCM would probably have been improved by the development of a diverse array of both CO₂ and HCO₃⁻ uptake systems of varying affinities in order to acquire Ci actively from the surrounding environments. It was during this time the cyanobacterial species acquired Ci transporters for actively pumping Ci at the site of CO₂ fixation.

The evolutionary studies based on 16s rRNA suggest that *Gloeobacter* and *Thermosynechococcus* sp. are amongst the most primitive cyanobacteria and possibly the ancestors of all cyanobacteria (Nelissen et al., 1995). *G. violaceus* diverged very early during the cyanobacterial radiation, in an ancient lineage preceding the cyanobacterial chloroplast ancestors. It possesses a unique molecular structure of photosystems I and II and an unusual morphology of its phycobilisomes (PBS) which enable it to harvest light and transfer energy in a manner which is different from other photosynthetic organisms (Mimuro et al., 2010). *Gloeobacter* carries a single HCO₃⁻ transporter (BCT1) and two CO₂ uptake systems. Therefore, it can be presumed that a relational study of transporters present in this organism can pave the path for determination of evolutionary pattern of Ci transporters in cyanobacteria. So, in the present study the Ci transport systems present in *Gloeobacter* sp. were analyzed in relation to other cyanobacteria and efforts were made to comprehend the arrangement of the genes constituting these transport systems and influence of habitat and the great oxidation event (GOE) on their evolution.

3.1.1. Gene organization of *cmpABCD*, *ndhI₃* and *ndhI₄* operons in cyanobacteria

The cyanobacterial genes encoding for the C_i transportation are present not only as single gene encoded units (e.g., BicA and SbtA) but also exist as a part of transport machinery i.e., in form of operons (e.g., *cmpABCD*, *ndhI₃* and *ndhI₄*). BCT1 is encoded by *cmpABCD* operon in which *cmpB* forms transmembrane helices and *cmpA*, *C* and *D* aid in bicarbonate entry and exit. The CO₂ uptake systems exist as NdhI₃ and NdhI₄ complexes which are present as *ndhD₃-ndhF₃-cupA* and *ndhD₄-ndhF₄-cupB* operons respectively. The arrangement of genes encoding the BCT1 transporter viz., *cmp ABCD* operon are depicted in (Figure 3-1a) and those of *NdhI₃* and *NdhI₄* are represented by (Figure 3-1b) and (Figure 3-1c) respectively.

The cyanobacterial evolution according to 16S rRNA, has been elucidated by Rudi et al., (1997). In accordance with this classification, *Gloeobacter violaceus* PCC 7421 is the most primitive cyanobacteria and hence the transporters present in *G. violaceus* can also be considered as the most primitive ones. The arrangement of the genes in the operon was found to be same in all the cyanobacterial strains studied, which implies that possibly all the genes for the transporter evolved as a unit or co-evolved in a related pattern.

Cyanobacteria show diversity in terms of distribution of the inorganic carbon transporters. Among the five most widely known transporters, SbtA and BicA were less widely distributed in comparison to BCT1, NdhI₃ and NdhI₄ transporters. *Prochlorococcus* sp. which is an α -cyanobacteria was devoid of both CO₂ transporters and harbored only one bicarbonate transporter, i.e., either BicA or SbtA. *Prochlorococcus* MED4, *Prochlorococcus* 9313 and *Prochlorococcus* SS120 showed presence of SbtA, a high affinity bicarbonate transporter whereas other *Prochlorococcus* sp. showed the presence of homogs of BicA, a low affinity bicarbonate transporter. The β -cyanobacteria had a mixed array of CO₂ as well as bicarbonate transport systems in which mutisubunit transporters were more predominant in comparison to the bicarbonate ones.

A

Serial number	Name of the organism	Arrangement of <i>cmpABCD</i> operon(BCT1 transporter)
1.	<i>Gloeobacter violaceus</i> PCC 7421	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
2.	<i>Thermosynechococcus elongatus</i> bp-1	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
3.	<i>Synechococcus</i> 7942	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
4.	<i>Synechocystis</i> sp. PCC 6803	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
5.	<i>Synechococcus</i> 6301	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
6.	<i>Anabaena</i> 7120	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
7.	<i>Anabaena variabilis</i>	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
8.	<i>Nostoc punctiforme</i> PCC 73102	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
9.	<i>Arthrospira platensis</i> NIES-39	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
10.	<i>Microcystis aeruginosa</i> NIES-843	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
11.	<i>Cyanothece</i> sp. PCC 7424	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
12.	<i>Cyanothece</i> sp. ATCC 51142	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
13.	<i>Cyanothece</i> sp. PCC 8801	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
14.	<i>Cyanothece</i> sp. PCC 7425	3' → [purple arrow] ← [green arrow] ← [red arrow] ← [blue arrow] → 5'
15.	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'
16.	<i>Synechococcus</i> sp. JA-3-3Ab	5' → [blue arrow] → [red arrow] → [green arrow] → [purple arrow] → 3'

cmpA
 cmpB
 cmpC
 cmpD

B

Serial number	Name of the organism	Arrangement of <i>ndhI₃</i> operon
1.	<i>Gloeobacter violaceus</i> PCC 7421	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
2.	<i>Thermosynechococcus elongatus</i> bp-1	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
3.	<i>Synechococcus</i> 7942	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
4.	<i>Synechocystis</i> sp. PCC 6803	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'
5.	<i>Synechococcus</i> 6301	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'
6.	<i>Anabaena</i> 7120	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
7.	<i>Anabaena variabilis</i>	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'
8.	<i>Nostoc punctiforme</i> PCC 73102	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
9.	<i>Synechococcus</i> sp. PCC 7002	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
10.	<i>Arthrospira platensis</i> NIES-39	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
11.	<i>Microcystis aeruginosa</i> NIES-843	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
12.	<i>Cyanothece</i> sp. PCC 7424	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
13.	<i>Cyanothece</i> sp. ATCC 51142	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
14.	<i>Acaryochloris marina</i> MBIC 11017	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'
15.	<i>Cyanothece</i> sp. PCC 8801	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
16.	<i>Cyanothece</i> sp. PCC 7425	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'
17.	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	5' → [red arrow] → [blue arrow] → [green arrow] → 3'
18.	<i>Synechococcus</i> sp. JA-3-3Ab	3' → [green arrow] ← [blue arrow] ← [red arrow] → 5'

ndh D₃
 ndh F₃
 cup A

C

Serial number	Name of the organism	Arrangement of <i>ndhI₄</i> operon
1.	<i>Gloebacter violaceus</i> PCC 7421	3' ← green ← blue ← red → 5'
2.	<i>Thermosynechococcus elongatus</i> bp-1	5' → red → blue → green → 3'
3.	<i>Synechococcus</i> 7942	3' ← green ← blue ← red → 5'
4.	<i>Synechocystis</i> sp. PCC 6803	3' ← green ← blue ← red → 5'
5.	<i>Synechococcus</i> 6301	5' → blue → red → green → 3'
6.	<i>Anabaena</i> 7120	5' → red → blue → green → 3'
7.	<i>Anabaena variabilis</i>	5' → red → blue → green → 3'
8.	<i>Nostoc punctiforme</i> PCC 73102	3' ← green ← blue ← red → 5'
9.	<i>Synechococcus</i> sp. PCC 7002	5' → red → blue → green → 3'
10.	<i>Arthrospira platensis</i> NIES-39	5' → red → blue → green → 3'
11.	<i>Synechococcus</i> sp. WH 8102	5' → red → blue → green → 3'
12.	<i>Synechococcus</i> sp. CC 9605	5' → red → blue → green → 3'
13.	<i>Synechococcus</i> sp. CC 9311	3' ← green ← blue ← red → 5'
14.	<i>Cyanothece</i> sp. PCC 7424	5' → red → blue → green → 3'
15.	<i>Synechococcus</i> sp. RCC 307	5' → red → blue → green → 3'
16.	<i>Synechococcus</i> sp. CC 9902	3' ← green ← blue ← red → 5'
17.	<i>Acaryochloris marina</i> MBIC 11017	3' ← green ← blue ← red → 5'
18.	<i>Cyanothece</i> sp. PCC 8801	3' ← green ← blue ← red → 5'
19.	<i>Cyanothece</i> sp. PCC 7425	5' → red → blue → green → 3'
20.	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	3' ← green ← blue ← red → 5'
21.	<i>Synechococcus</i> sp. JA-3-3Ab	5' → red → blue → green → 3'

Figure 3-1: Arrangement of the genes encoding the cyanobacterial Ci transporters. (A) Schematic representation of the arrangement of genes in BCT1 transporter (*cmpABCD*). (B) Arrangement of the genes encoding the *NdhI₃* uptake system. (C) Arrangement of the genes encoding the *NdhI₄* uptake system. The *ndhI₄* operon was found to be present in most of the cyanobacterial phyla.

3.1.2. Phylogenetic analysis of *cmpABCD*, *ndhI₃* and *ndhI₄* operons

The BCT1 transporter showed its presence in sixteen out of forty species studied. The phylogenetic tree depicted two major clades. The lower clade mostly consisted of late diverging cyanobacteria like *Cyanothece* sp., *Arthrospira platensis* and *Nostoc punctiforme* while the upper clade comprised mostly of the early diverging ones (Figure 3-2). The only exception was *Anabaena* 7120 which although being a late diverging was found in the clade of *Gloeobacter violaceus* PCC 7421, an early diverging bacteria. When the protein sequences of CmpA, B, C and D in *Synechococcus elongatus* PCC 7942 were further studied they depicted conserved domains which were characteristics of the ABC type transporters. CmpA has the TauA domain and belongs to periplasmic binding protein type 2 superfamily. This domain showed about 48-50% similarity to the ABC-type nitrate/sulfonate transport system of several archaea, proteobacteria and plants. The CmpB protein showed 50-60% similarity to the bicarbonate/nitrate ABC transporter permease of various proteobacteria and few archaea and plants. The CmpC and CmpD have the ntrCD domain which is 55-60% similar to nitrate ABC transporter ATP-binding protein of proteobacteria and archaea.

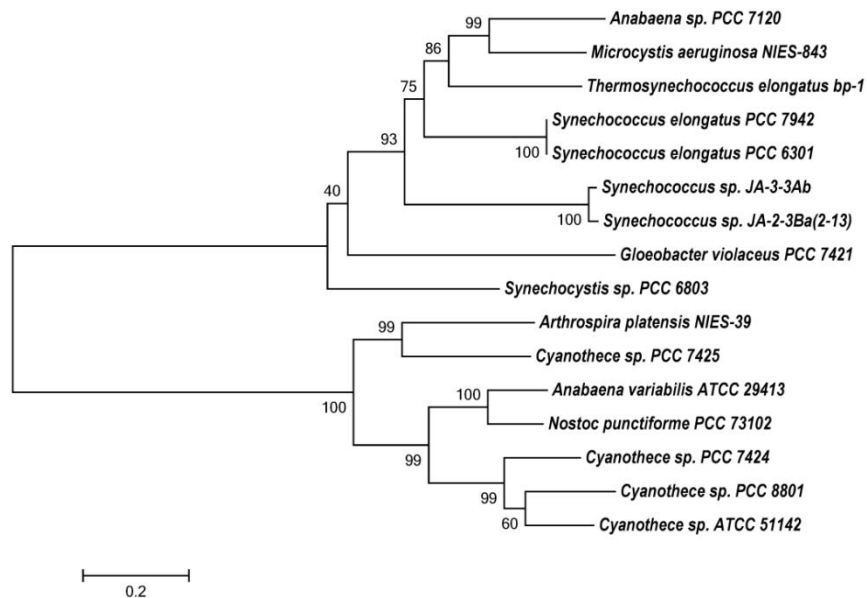


Figure 3-2: Phylogenetic tree of *cmpABCD* operon (BCT1) in sixteen cyanobacterial species. The trees are constructed using maximum likelihood method. Bootstrapping was done for 1000 iterations. The scale bars indicate the number of nucleotide substitutions per site.

The *ndhI₃* operon was present in eighteen species out of the forty investigated (Figure 3-3). There were two major clades, the lower one primarily consisting of late diverging cyanobacteria except *Synechococcus elongatus* PCC 6301 and *Synechococcus* sp. JA-3-3Ab (both are early diverging). The upper clade was further diverged into two subclades having early and late diverging cyanobacteria respectively.

The low affinity CO₂ uptake system viz., NdhI₄ was the most widely distributed among the three transporters studied and was found in twenty one species. The phylogenetic tree of this transporter revealed two major clades. The upper one mainly consisted of late diverging cyanobacteria while the lower one was a combination of both. *Synechococcus* sp. JA-3-3Ab, *Thermosynechococcus elongatus* bp-1 and *Synechococcus elongatus* PCC 6301 which are closely related to *G. violaceus* (an early diverging cyanobacteria) were present in the upper clade which included mostly the late diverging cyanobacteria (Figure 3-4).

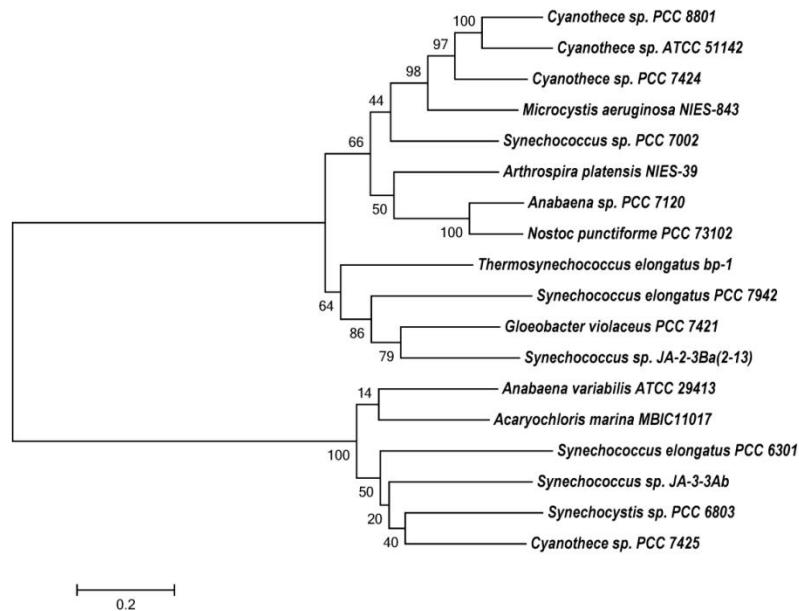


Figure 3-3: Phylogenetic tree of *ndhI₃* operon in eighteen cyanobacterial species. The trees are constructed using maximum likelihood method. Bootstrapping was done for 1000 iterations. The scale bars indicate the number of nucleotide substitutions per site.

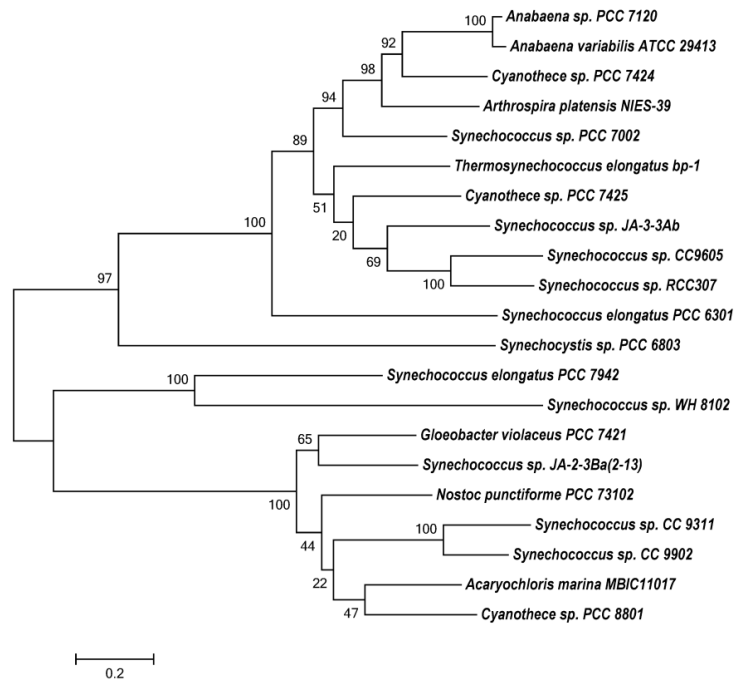


Figure 3-4: Phylogenetic tree of *ndhI4* operon in twenty one cyanobacterial species. The trees are constructed using maximum likelihood method. Bootstrapping was done for 1000 iterations. The scale bars indicate the number of nucleotide substitutions per site.

It should be noted that BCT1 of *Cyanothece* sp. PCC 8801, *Cyanothece* sp. PCC 7424 and *Cyanothece* sp. ATCC 51142 was far-related from *G. violaceus* which was located in separate clade. As far as *NdhI3* is concerned, *Cyanothece* sp. PCC 8801 was more closely related to *G. violaceus* PCC 7421 and were located in the same clade. For *NdhI4*, all the *Cyanothece* sp. under study seemed to show diversity in their nucleotide sequences as they appeared to be far related to each other.

Another important observation was that *G. violaceus*, *Synechococcus* sp. JA-3-3Ab, *Synechococcus* sp. JA-2-3B'a(2-13) and *Synechococcus elongatus* PCC 6301 were closely related to each other for BCT1 gene while in case of *NdhI3* and *NdhI4*, *Synechococcus* sp. JA-3-3Ab was far related from the other three. For the high affinity transporters (BCT1 and *NdhI3*), *G. violaceus* and *Anabaena* sp. PCC 7120 appeared to be closely related and were a part of the same major clade. In the low affinity transporter (*NdhI4*), the early and late

diverging cyanobacteria were far related with the exception of *Cyanothece* sp. PCC 8801 (a late diverging cyanobacteria).

In the CO₂ uptake systems, CO₂ hydration proteins cupA (chpY) and cupB (chpX) which belong to chpXY superfamily were found to be present only in cyanobacteria whereas NdhD₃/NdhD₄ and NdhF₃/NdhF₄ both had homologs in archaea, proteobacteria as well as plants (Maeda et al., 2002).

The analysis revealed that the low affinity CO₂ uptake system (NdhI₄) was abundant in most of the cyanobacterial species which reflects that this transporter was possibly present in all the cyanobacteria prior to the GOE as the environment was rich in CO₂. The other two transporters (BCT1 and NdhI₃) perhaps evolved on the basis of organisms need in accordance with the changing environment by a functional uptake of certain pre-existing transporters.

3.1.3. Diversity in suite of Ci transporters in cyanobacteria is governed by habitation

Cyanobacteria show a diverse collection of transporters depending on Ci availability and the environment in which they dwell. The cyanobacterial strains which inhabit the marine environments (α -cyanobacteria) possess less varied complement of bicarbonate transporters in comparison to the ones living in fresh water i.e. β -cyanobacteria (Badger et al., 2005). BCT1, a high affinity bicarbonate transporter is absent in marine forms as the marine environment already possesses a substantially high content of bicarbonate which deters the need of a transporter with high energy investment (BCT1 requires ATP energization) and favours employment of less energy requiring sodium symporters like BicA and SbtA. Therefore, out of the forty species studied BCT1 marked its presence only in thirteen which were majorly the β -cyanobacteria. BicA and SbtA have been reported to be present in almost all cyanobacterial strains but were found to be absent in *G. violaceus* which infers that these transporters were present before the branching of α - and β -cyanobacterial clades and were possibly lost from *G. violaceus*. Both CO₂ uptake systems viz., NdhI₃ and NdhI₄ are present in *Gloeobacter* sp. which suggests that these were acquired during early stages of evolution but the high affinity NdhI₃ uptake system was gradually lost from some marine species.

NdhI₄ persisted and is still found in many cyanobacterial species inhabiting a wide regime of temperature and Ci fluctuations (Paerl, 2012).

3.1.4. Impact of Great Oxidation Event on evolution of Ci transporters

The evolution of the genes related to photosynthesis greatly depends on environmental conditions. Earlier the environment was CO₂ rich but after the great oxidation event (GOE), there was a drift from anoxygenic environment to an oxygenic one. The origin of cyanobacteria and the evolution of multicellularity are estimated before or at the beginning of the GOE. Relatively soon after GOE, three major cyanobacterial clades originated which included the unicellular cyanobacteria, terminally differentiated taxa and marine phytoplankton. The phylogeny showed that the cyanobacteria which came into existence before the great oxidation event included the β -cyanobacteria which comprised of *Gloeobacter*, *Thermosynechococcus* and certain *Synechococcus* species, while the cyanobacteria which evolved after the GOE included species like *Nostoc* and *Anabaena*. Since primitive cyanobacteria like *Gloeobacter* had evolved before the GOE it is likely that the CO₂ low affinity transporters like NdhI₄ existed before GOE when the environment was CO₂ rich. But when the environment suffered dramatic change from a reducing to non-reducing one, the high affinity transporters like NdhI₃ must have evolved by acquisition of the function of CO₂ transport probably by the modification of the already existing NAD(P)H dehydrogenase systems which were earlier involved in respiratory activities. Our hypothesis is in consonance with the findings of Howitt et al., (1999) who have revealed that the type II NADH dehydrogenases in the cyanobacterium *Synechocystis* PCC 6803 are involved in regulation rather than respiration. One of the cyanobacterial bicarbonate transporters, BicA is also reported to be member of the SulP superfamily (Saier Jr et al., 1999), which includes the plant SulP sulphate transporter family and the mammalian SLC26 family of anion transporters as well as numerous members in a wide range of prokaryotes (Price and Howitt, 2010). BCT1, an ABC type transporter shows a strong similarity to the nitrate transporter (NRT1) located on the plasma membrane and its subunits NrtA and NrtC exhibits 46.5% and 30% identity to CmpA and CmpC respectively (Omata, 1991).

According to various reports it has been found that the two types of CO₂ uptake systems may have been acquired before the branching of α -cyanobacteria as α -cyanobacteria

like *Prochlorochoccus* which dwell in marine habitats do not possess any of the CO₂ uptake systems and only possess HCO₃⁻ uptake systems (Badger and Price, 2003) as HCO₃⁻ is the predominant form of Ci present in oceans. *Synechococcus* 9902 and *Synechococcus* 8102 possess only the low affinity CO₂ transporter (NdhI₄). After further branching of α-cyanobacteria, the low affinity CO₂ transporters might have been lost. The high affinity CO₂ uptake system (NdhI₃) is generally absent in marine strains but some marine strains may have acquired these systems by horizontal gene transfer. β-cyanobacteria from freshwater face the greatest immoderations in Ci accessibility, influenced by conditions of temperature and pH and therefore these species possess the largest number of different types of transporters.

BCT1 being a high affinity and high energy demanding transporter would not have been in great demand before GOE. The advent of GOE raised the demand of these high affinity transporters due to scarcity of CO₂ in the vicinity of RuBisCO. Therefore, it can be anticipated that the cyanobacteria which evolved after GOE might have acquired this transporter through modification of the existing transporters like NRT1. NdhI₄, being a low affinity and low energy investment transporter would have been more prevalent in early environment when CO₂ concentration was probably high.

The phylogenetic analysis of these transporters in different phyla of cyanobacteria belonging to both α- and β-cyanobacterial classes suggested that the early environmental conditions have played a major role in shaping the course of evolution of bicarbonate transporters. Depending upon the habitat the cyanobacteria enriched themselves with the array of transporters. As evolution is never a directed one, variations which were found beneficial were incorporated. Since the organization of the genes encoding the transporters exhibited a random order in different cyanobacterial clades, these could also be proposed to have been acquired by horizontal gene transfer comprising of single gene lateral transfer event from an ancestral archaeal, proteobacterial or even a cyanobacterial strain. There is also the possibility that both Rubisco and carboxysome genes could be inherited by a single lateral gene transfer event as in many cyanobacteria they are found ordered on contiguous regions of the chromosome (Badger et al., 2005). The knowledge of the evolution of these

genes can be a way to predict the best candidate cyanobacteria which can be used to contribute genes for development of efficient carbon concentrating mechanism even in plant species which lack such mechanisms.

3.2. Analysis of *sbtA* gene

SbtA transporter is a single subunit transporter that was originally identified in *Synechocystis* sp. strain PCC 6803 (Shibata et al., 2002). In *Synechocystis* sp. PCC 6803, the transporter consists of 1125 nucleotides and 373 amino acid residues. As far as the topology is concerned, the transporter has been predicted to have 10 transmembrane helices (TMHs) with both N- and C-terminus facing periplasm. SbtA is a high affinity transporter and is induced under low C_i conditions (Zhang et al., 2004). Till date, SbtA proteins from *Synechocystis* (Shibata et al., 2002) and *Synechococcus* PCC 7002 (Price et al., 2004) are the only forms of the SbtA protein that have been well characterized and studied.

Among the β -cyanobacteria, SbtA transporter was found to be absent in *G. violaceus*, *Thermosynechococcus elongatus* bp-1 and *Trichodesmium erythraeum* which are few of the early evolving cyanobacteria which projects the fact that the cyanobacteria might have acquired this transporter after the branching of α -cyanobacteria.

3.2.1. Phylogenetic analysis of *sbtAB* operon

The *sbtA* gene in cyanobacteria exists in an operon with another gene *sbtB* which codes for SbtB protein that possibly functions as SbtA regulator. Several studies have postulated that the sequences flanking the *sbtA* and *sbtB* genes in α -cyanobacterial strains like *Synechococcus* WH 5701 and *Cyanobium* PCC 7001 have significantly lower than average 'G+C' contents (Rae et al., 2011). These lower 'G+C' content have been found to be associated with insertions of transposons (Bolhuis et al., 2010), which indicate that these sequences might have been acquired by horizontal gene transfer in these transitions strains which most probably was a transposon-mediated one.

Since, no method/tool is available for detection of horizontal gene transfer in organisms; the relatedness between organisms under study was first studied by constructing phylogenetic tree of individual genes (*sbtA* and *sbtB*) and operon (*sbtAB*). The phylogenetic

trees were then subjected to analysis of co-evolution by Mirrortree server. The DNA sequences coding for *sbtA*, *sbtB* along with the flanking sequences (150 bp on either side) were retrieved from seventy organisms belonging to various phyla of Eubacteria namely, Proteobacteria, Cyanobacteria, Chloroflexi, Planctomycetes, Deinococcus-thermus, Bacteroidetes/Chlorobi, Actinobacteria and Spirochaetes using Kazusa cyanobase (<http://genome.microbedb.jp/cyanobase/>) and NCBI (<https://www.ncbi.nlm.nih.gov/>). The homologs for *sbtA* and *sbtB* gene were searched using *sbtA* and *sbtB* sequence from *Synechococcus elongatus* 7942 as query. Since SbtB is a member of nitrogen regulatory protein P_{II} family it may have several copies in the genome, the one situated closest to SbtA was used for the study. In order to avoid ambiguity in the results, only those organisms which had complete set of *sbtAB* operon in them were used for the study of *sbtA*, *sbtB* and *sbtAB* operon (Table 3-2).

The phylogenetic trees of individual genes (*sbtA* and *sbtB*) as well as *sbtAB* operon were constructed using the Maximum likelihood method. The trees were exported in newick format and were analyzed by the help of the Mirrortree web server. The server helps in studying the co-evolution of two genes/protein families and also investigates their functional relationships and possible interactions in a taxonomic context (Ochoa and Pazos, 2010). The output generated gives the correlation values between the genes/protein families, higher values of correlation (close to 1) infer to the co-evolution of the genes/proteins. Since the correlation values were more than 0.8 for *sbtA*, *sbtB* and *sbtAB* interactions, the genes are likely to have evolved together (Table 3-3).

High correlation values indicate that the transfer of the genes of the operon might have occurred as a single unit during horizontal gene transfer event. In order to find out the most probable organisms who would have been the source of the genes in far-related organisms, the gene sequences of *sbtA* and *sbtB* along with intermediary and flanking sequences were retrieved and a phylogenetic tree was constructed (Figure 3-3).

Table 3-2: The gene symbol, function, nucleotide and amino acid length of *sbtA* and *sbtB* genes in the seventy eubacterial phyla used for the analysis

S. No.	Organism	SbtA				SbtB			
		Gene symbol	Gene name	Nucleotide length	Amino acids	Gene symbol	Gene name	Nucleotide length	Amino acids
1	<i>Synechocystis</i> sp. PCC 6803	slr1512	sodium-dependent bicarbonate transporter	1125	347	slr1513	Periplasmic protein, function unknown	333	110
2	<i>Synechococcus elongatus</i> PCC 7942	Syn pcc 7942_1475	sodium-dependent bicarbonate transporter	1122	373	Synpcc7942_1476	conserved hypothetical protein	315	104
3	<i>Acaryochloris marina</i> MBIC 11017	AM1_4164	DUF897 domain membrane protein	1095	364	AM1_4165	conserved hypothetical protein	333	110
4	<i>Cyanothece</i> sp. ATCC 51142	cce_2939	sodium-dependent bicarbonate transporter	1128	375	cce_2937	Hypothetical protein	318	105
5	<i>Anabaena variabilis</i> ATCC 29413	Ava_3027	Protein of unknown function DUF897	1110	369	Ava_3028	Plasma membrane protein of unknown function	333	110
6	<i>Arthrospira platensis</i> NIES-39	NIES39_E03120	sodium-dependent bicarbonate transporter	1137	378	NIES39_E03140	Hypothetical protein	318	105
7	<i>Cyanothece</i> sp. PCC 8801	PCC8801_1249	Protein of unknown function DUF897	1101	366	PCC8801_1248	Hypothetical protein	333	110
8	<i>Cyanothece</i> sp. PCC 8802	PCC8802_1280	Protein of unknown function DUF897	1101	366	PCC8802_1279	Conserved hypothetical protein	333	110
9	<i>Cyanothece</i> sp. PCC 7425	Cyan7425_5063	Protein of unknown function DUF897	1134	377	Cyan7425_5062	Conserved hypothetical protein	333	110
10	<i>Geitlerinema</i> sp. PCC 7407	GEI7407_1945	Protein of unknown function DUF897	1131	376	GEI7407_1946	Hypothetical protein	315	104
11	<i>Halothece</i> sp. PCC 7418	PCC7418_0056	Protein of unknown function DUF897	1158	385	PCC7418_0057	Hypothetical protein	315	104
12	<i>Calothrix</i> sp. PCC 6303	Cal6303_1190	Protein of unknown function DUF897	1122	373	Cal6303_1189	Plasma membrane protein of unknown function	333	110
13	<i>Chroococcidiopsis thermalis</i> PCC 7203	Chro_1748	Protein of unknown function DUF897	1131	376	Chro_1749	Hypothetical protein	333	110
14	<i>Cyanobium gracile</i> PCC6307	Cyagr_2515	Putative permease	1008	335	Cyagr_2514	Hypothetical protein	315	104
15	<i>Dactylococcopsis salina</i> PCC 8305	Dasca_3500	Putative permease	1140	379	Dasca_3499	Hypothetical protein	315	104

16	<i>Microcoleus</i> sp. PCC 7113	Mic7113_1307	Putative permease	1134	377	Mic7113_1308	Hypothetical protein	333	110
17	<i>Nostoc</i> sp. PCC 7107	Nos7107_3597	Protein of unknown function DUF897	1110	369	Nos7107_3598	Plasma membrane protein of unknown function	333	110
18	<i>Nostoc</i> sp. PCC 7120	all2134	Hypothetical protein	1113	370	all2133	Hypothetical protein	333	110
19	<i>Nostoc</i> sp. PCC 7524	Nos7524_2740	Putative permease	1134	377	Nos7524_2739	Hypothetical protein	333	110
20	<i>Prochlorococcus marinus</i> str. AS9601	A9601_02311	Putative sodium dependent bicarbonate transporter	999	332	A9601_02301	Conserved hypothetical protein	282	93
21	<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> CCMP 1986(<i>Prochlorococcus marinus</i> MED4)	PMM0213	Putative sodium dependent bicarbonate transporter, SbtA	999	332	PMM0212	Conserved hypothetical protein	282	93
22	<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> CCMP 1375	Pro_0241	Predicted permease	996	313	Pro_0240	Uncharacterized protein	282	93
23	<i>Prochlorococcus marinus</i> str. MIT 9215	P9215_02321	Putative sodium dependent bicarbonate transporter	999	332	P9215_02311	Conserved hypothetical protein	282	93
24	<i>Prochlorococcus marinus</i> str. MIT 9301	P9301_02331	Putative sodium dependent bicarbonate transporter	999	332	P9301_02321	Conserved hypothetical protein	288	95
25	<i>Prochlorococcus marinus</i> str. MIT 9303	P9303_07971	Putative sodium dependent bicarbonate transporter	1026	341	P9303_07981	Conserved hypothetical protein	282	93
26	<i>Prochlorococcus marinus</i> str. MIT 9312	PMT9312_0215	Putative sodium dependent bicarbonate transporter	999	332	PMT9312_0214	Nitrogen regulatory proteinP-II(GlnB,GlnK)	282	93
27	<i>Prochlorococcus marinus</i> str. MIT 9313	PMT_1213	Putative sodium dependent bicarbonate transporter	1020	339	PMT_1212	Conserved hypothetical protein	282	93
28	<i>Prochlorococcus marinus</i> str. MIT 9515	P9515_02421	Putative sodium dependent bicarbonate transporter	999	332	P9515_02411	Conserved hypothetical protein	282	93
29	<i>Prochlorococcus marinus</i> str. NATL1A	NATL1_02901	Putative sodium dependent bicarbonate transporter	993	330	NATL1_02891	Conserved hypothetical protein	282	93
30	<i>Prochlorococcus marinus</i> str. NATL2A	PMN2A_1580	Permease	993	330	PMN2A_1579	Conserved hypothetical protein	282	93
31	<i>Pseudanabaena</i> sp. PCC 7367	Pse7367_3848	Protein of unknown function DUF897	1008	335	Pse7367_3849	Hypothetical protein	318	105

32	<i>Rivularia</i> sp. PCC 7116	Riv7116_3232	Putative permease	1185	394	Riv7116_3233	Hypothetical protein	321	105
33	<i>Staniera cyanosphaera</i> PCC7437	Sta7437_3040	Protein of unknown function DUF897	1131	376	Sta7437_3039	Hypothetical protein	333	110
34	<i>Synechococcus elongatus</i> PCC 6301	Syc2461_c	Sodium dependent bicarbonate transporter, SbtA	1122	373	Syc2462_d	Hypothetical protein	315	104
35	<i>Synechococcus</i> sp. PCC 7002	SYNPCC7002_A0470	Sodium dependent bicarbonate transporter, SbtA	1113	370	SYNPCC7002_A0472	Conserved hypothetical protein	315	104
36	<i>Synechocystis</i> sp. PCC 6714	D082_29370	Putative sodium dependent bicarbonate transporter, SbtA	1122	373	D082_29380	Hypothetical protein	333	110
37	<i>Synechococcus</i> sp. UTEX 2973	M744_08365	Sodium dependent bicarbonate transporter	1122	373	M744_08370	Membrane protein	315	104
38	<i>Calothrix</i> sp. 336/3	IJ00_04285	Sodium dependent bicarbonate transporter	1113	370	IJ00_04290	Membrane protein	333	110
39	<i>Calothrix</i> sp. PCC 7507	Cal7507_4199	Protein of unknown function DUF897	972	323	Cal7507_4198	Hypothetical protein	308	102
40	<i>Chloroflexus aggregans</i> DSM 9485	Cagg_1405	Sodium dependent bicarbonate transport family permease	974	324	Cagg_1404	Nitrogen regulatory protein P-II	314	104
41	<i>Pararhodospirillum photometricum</i> DSM 122	RSPPHO_02673	Membrane protein containing DUF897 domain	1217	405	RSPPHO_02674	Putative uncharacterized	314	104
42	<i>Deinococcus deserti</i> VC D115	Deide_17120	Sodium dependent bicarbonate transport family permease	1100	366	Deide_17130	Hypothetical protein	302	100
43	<i>Comamonadaceae bacterium</i> A1	SRAA_0072	Sodium dependent bicarbonate transport family permease	1007	335	SRAA_0073	Hypothetical protein	314	104
44	<i>Comamonadaceae bacterium</i> B1	SMCB_0106	Sodium dependent bicarbonate transport family permease	1004	334	SMCB_0107	Hypothetical protein	314	104
45	<i>Roseibacterium elongatum</i> DSM 19469	roselon_00805	Sodium dependent bicarbonate transport family permease	1054	350	roselon_00806	Membrane protein	317	105

46	<i>Salinispira pacifica</i>	L21SP2_1540	Sodium dependent bicarbonate transport family permease	1007	335	L21SP2_1541	Hypothetical protein	317	105
47	<i>Pseudomonas mendocina</i> ymp	Pmen_4477	Sodium dependent bicarbonate transport family permease	941	313	Pmen_4476	Hypothetical protein	302	100
48	<i>Pseudomonas pseudoalcaligenes</i> CECT 5344	BN5_4307	Sodium dependent bicarbonate transport family permease	941	313	BN5_4306	Hypothetical protein	302	100
49	<i>Pseudomonas alcaligenes</i> str. NEB 585	AOT30_RS01430	Sodium dependent bicarbonate transport family permease	935	311	AOT30_RS01425	Hypothetical protein	302	100
50	<i>Moorea producens</i> PAL-8-15-08-1	BJP34_33730	Sodium dependent bicarbonate transport family permease	1124	374	BJP34_33735	Hypothetical protein	314	104
51	<i>Thermus parvatiensis</i> str.RL	AV541_08120	Sodium dependent bicarbonate transport family permease	956	318	AV541_08115	Hypothetical protein	302	100
52	<i>Rugosibacter aromaticivorans</i>	PG1C_09855	Permease	941	313	PG1C_09850	Nitrogen regulatory protein P-II	302	100
53	<i>Rhodothermus marinus</i> SG0	Rhom172_2682	Protein of unknown function DUF897	959	319	Rhom172_2681	Nitrogen regulatory protein P-II	302	100
54	<i>Zhongshania aliphaticivorans</i> str.SM2	AZF00_RS08520	Sodium dependent bicarbonate transport family permease	944	314	AZF00_RS08515	Hypothetical protein	302	100
55	<i>Meiothermus ruber</i> DSM 1279	Mrub_1132	Sodium dependent bicarbonate transport family permease	968	322	Mrub_1131	Hypothetical protein	302	100
56	<i>Pseudarthrobacter phenanthrenivorans</i> sphe3	Asphe3_35730	Sodium dependent bicarbonate transport family permease	944	314	Asphe3_35720	Hypothetical protein	302	100
57	<i>Nostoc punctiforme</i> PCC 73102	Npun_F1009	Sodium dependent bicarbonate transport family permease	968	322	Npun_F1010	Nitrogen regulatory protein P-II	311	103
58	<i>Phycisphaera mikurensis</i>	PSMK_06670	Sodium dependent	1031	343	PSMK_06660	Hypothetical protein	317	105

	NBRC102666 DNA		bicarbonate transport family permease						
59	<i>Pseudohongiella spirulinae</i> str.KCTC 32221	PS2015_ 2263	Permease	938	312	PS2015_2262	Hypothetical protein	317	105
60	<i>Thermus thermophilus</i> SG0.5JP17-16	Ththe16_ 0514	Sodium dependent bicarbonate transport family permease	956	318	Ththe16_0515	Hypothetical protein	302	100
61	<i>Candidatus nitrosopumilus</i> sp. AR2	NSED_ 02425	Sodium dependent bicarbonate transport family permease	1079	359	NSED_02430	Hypothetical protein	302	100
62	<i>Azoarcus</i> sp.BH72	Azo1481	Predicted permease	938	312	Azo1480	Nitrogen regulatory protein P-II	311	103
63	<i>Thioalkalivibrio nitratireducens</i> DSM 14787	TVNIR_ 2980	Putative sodium dependent bicarbonate transporte	1007	335	TVNIR_2979	Nitrogen regulatory protein P-II	320	106
64	<i>Melioribacter roseus</i> P3M	MROS_ 2614	Sodium dependent bicarbonate transport family permease	944	314	MROS_2613	Nitrogen regulatory protein P-II	299	99
65	<i>Pseudomonas stutzeri</i>	A458_ 20345	Permease	941	313	A458_20340	Hypothetical protein	302	100
66	<i>Pseudomonas balearica</i> DSM 6083	CL52_ 19775	Sodium dependent bicarbonate transport family permease	941	313	CL52_19770	Nitrogen regulatory protein P-II	302	100
67	<i>Oscillatoria nigro-viridis</i> PCC 7112	Osc7112_ 0025	Sodium dependent bicarbonate transport family permease	971	323	Osc7112_0024	Nitrogen regulatory protein P-II	308	102
68	<i>Leptolyngbya</i> sp. PCC 7376	Lepto7376 _1337	Sodium dependent bicarbonate transport family permease	970	322	Lepto7376_133 6	Nitrogen regulatory protein P-II	311	103
69	<i>Fischerella</i> sp.JSC-11	FJSC11DR AFT_0014	Sodium dependent bicarbonate transport family permease	968	322	FJSC11DRAFT _0013	Nitrogen regulatory protein P-II	311	103
70	<i>Chloroflexus</i> sp.Y-400-fl	Chy400_ 1931	Sodium dependent bicarbonate transport family permease	974	324	Chy400_1932	Nitrogen regulatory protein P-II	317	105

Table 3-3: The correlation values between *sbtA*, *sbtB* and *sbtAB* operon genes obtained by Mirrortree server (csbg.cnb.csic.es/mtserver/)

S.No.	Genes/operon	Correlation values (obtained by Mirrortree server)
1.	<i>sbtA</i> and <i>sbtB</i>	0.810
2.	<i>sbtA</i> and <i>sbtAB</i> operon	0.985
3.	<i>sbtB</i> and <i>sbtAB</i> operon	0.890

On analysis, it was found that the β -cyanobacteria having *sbtA* with high certainty of transporter function (70-80% identity to *sbtA*-PCC 7942) formed a separate clade at the bottom. The *Prochlorococcus* strains (23% identity), having a distant homologue of *sbtA* forms a different clade from the rest of the cyanobacteria. The reason might be attributed to the fact that the SbtA protein in *Prochlorococcus* strains is of 324-339 amino acids whereas in β -cyanobacteria it is of 370-374 amino acids. There also occurs a variation in the length of the loop present between 5th and 6th helix in *Prochlorococcus* and other cyanobacterial strains (the loop is shorter in *Prochlorococcus* due to the deletion which might have occurred during the event of horizontal gene transfer) (Figure 3-5).

These observations suggest that the transporter might have first evolved in the β -cyanobacterial strains and then could have been transferred to the α -cyanobacterial transition strains (which moved from oligotrophic marine environments to freshwater habitat) through horizontal gene transfer. The evolution of cyanobacteria has been influenced a lot by the horizontal gene transfer (Zhaxybayeva, 2009).

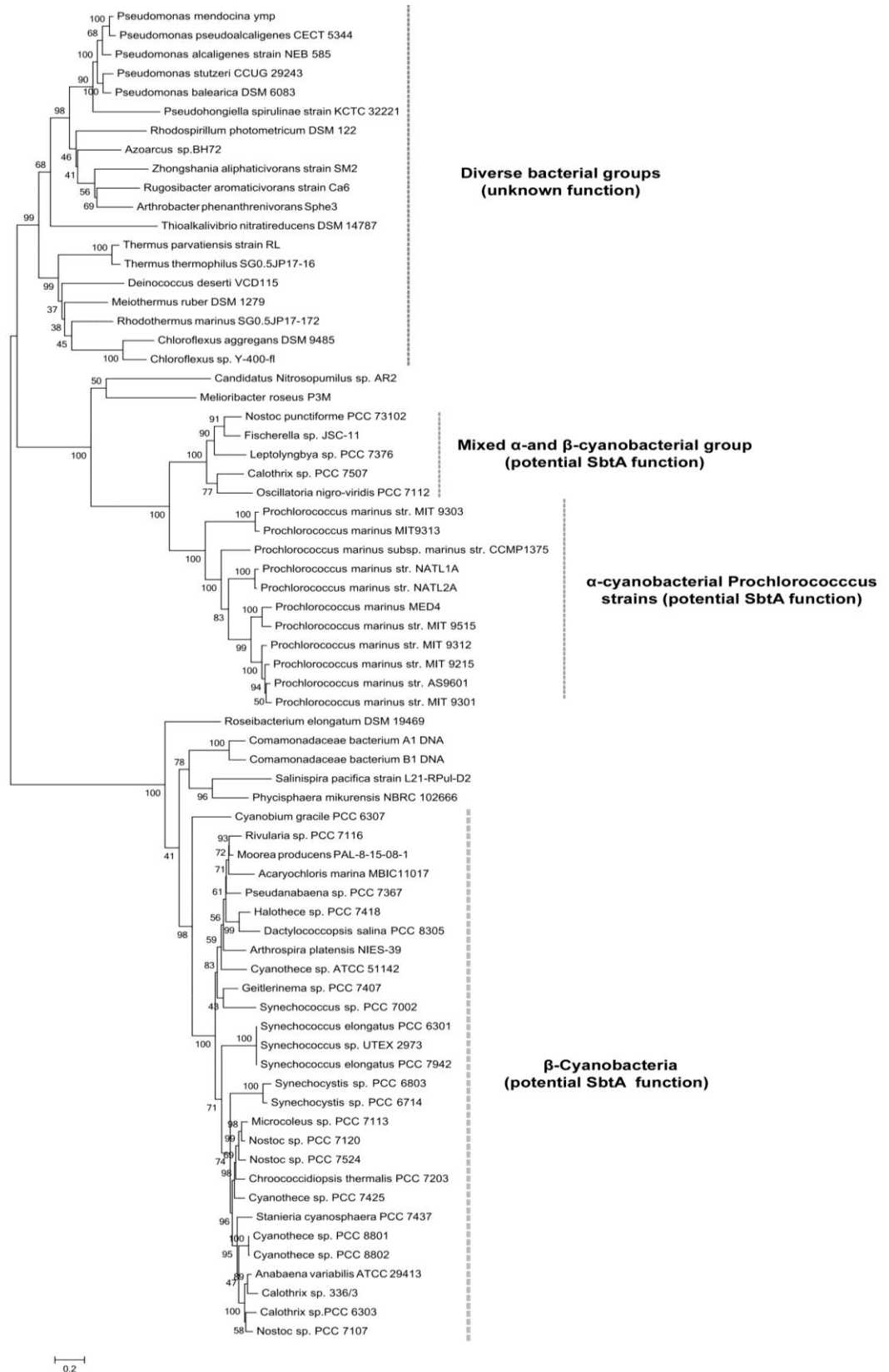


Figure 3-6: Legends on the next page

Figure 3-6: Phylogenetic tree for *sbtAB* operon conducted in MEGA6. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000). The tree with the highest log likelihood (-9412.4050) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites {5 categories (+G, parameter = 2.2953)}. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 70 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 222 positions in the final dataset. Bootstrap support is annotated on the corresponding branch.

3.2.2. Analysis of nucleotide and protein sequence of SbtA transporter

The SbtA transporter from *Synechococcus elongatus* PCC 7942 has been identified as one of the most suitable candidates for expression in crops because of its lower requirements for Na^+ ions, which is only 1.5 mM for half maximal uptake of HCO_3^- . The sequence of *sbtA* transporter gene from *Synechococcus elongatus* PCC 7942 was retrieved from Kazusa cyanobase by homology search using *sbtA* gene sequence from *Synechocystis* sp. PCC 6803 (slr1512) as query (84% identity).

Since SbtA is a transmembrane transporter protein; analysis of the number of transmembrane helices present in the protein was also carried out using several web servers/online available tools. The results obtained after analysis through different tools showed discrepancy in the number of helices as the number of helices predicted ranged from as few as 7 to 10 (Table 3-4). Price et al., (2011) had determined the membrane topology of the SbtA protein experimentally using *phoA/lacZ* fusion reporter method and confirmed the presence of 10 transmembrane helices in the proteins.

Table 3-4: Analysis of TM helices in SbtA transporter using several topology prediction tools

S. No.	Web server/tool used for prediction of TM helices	No. of TM helices
1.	TMHMM	7
2.	DAS-TM filter server	9
3.	Phobius	10
4.	OCTOPUS	10
5.	TTMOD	9
6.	TOPCONS	10

The activity of the SbtA transporter is regulated at multiple levels. At the transcriptional level, its activity is modulated by CcmR, a negative regulator. The conditions of low C_i favor the unbinding of CcmR from repressor leading to the activation of the transcription of the *sbtA* gene (Woodger et al., 2007). It has been observed that inorganic carbon transporters in cyanobacteria are activated on transition from darkness to illuminated conditions (post-translational regulation). This activation/ inactivation has been proposed to occur either by a redox regulatory signal (Kaplan et al., 1987) or by phosphorylation/de-phosphorylation mechanism which acts specifically on the serine or threonine residues present in the regulatory loops of transporter. Therefore, the amino acid sequence of SbtA from *Synechococcus elongatus* PCC 7942 was aligned with its homologue from *Synechocystis* sp. PCC 6803 and the conserved residues in the protein were marked. Along with this the potential residues involved in dark inactivation of protein were also identified.

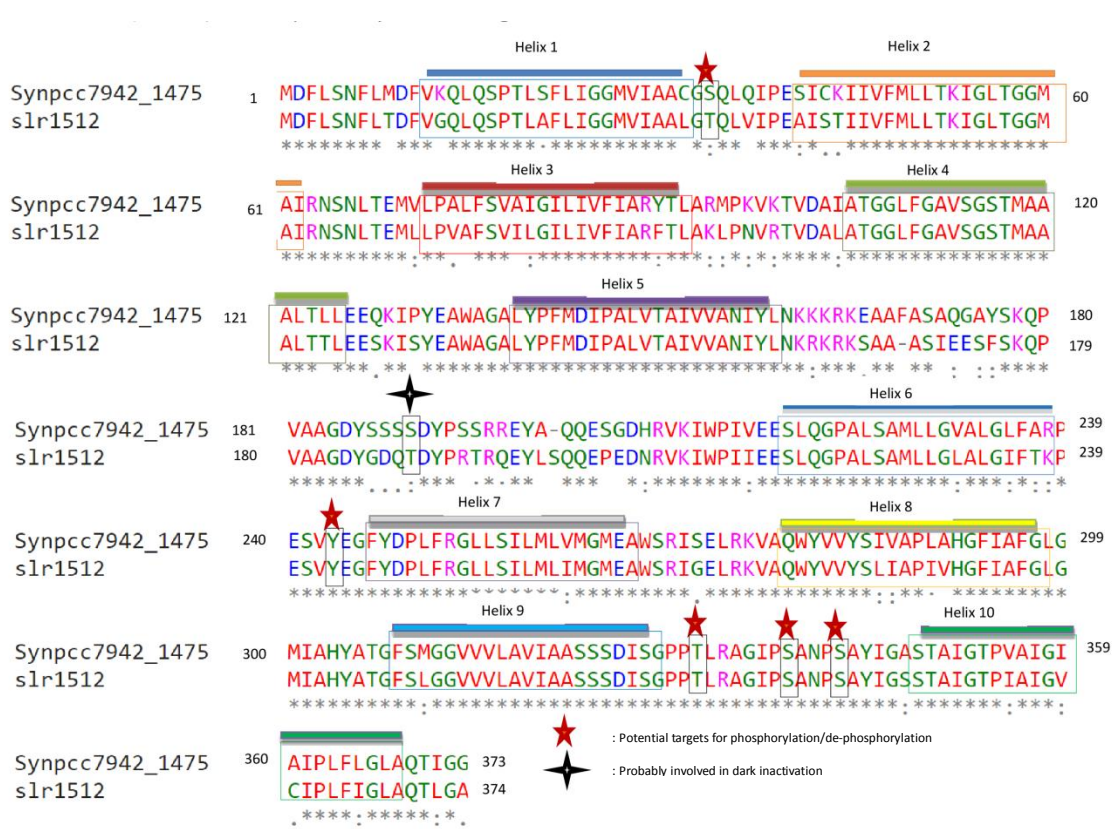


Figure 3-7: The amino acid alignment of SbtA from *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942. The sequences were aligned using EMBOSS Needle pairwise sequence alignment tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/). The residues identified as potential kinase targets and the ones involved in dark inactivation (marked in red and black colored stars) were identified manually on the basis of studies conducted by Price et al., (2011).

The loop between 9th and 10th helix had the most conserved residues and hence T³³¹, S³³⁸, S³⁴² were identified as potential kinase targets. These residues have also been reported to be conserved in the 95-member SbtA family by Price et al., (2011). Along with this a conserved residue was identified in the loop between helix 1 and 2 (S/T³⁴) and in between helix 6 and 7 (Y²⁴⁴) (Figure 3-7). These residues could be of particular importance in the goal of understanding how SbtA transporter is activated in cyanobacteria as well as in discovering ways to adapt active transporter to higher plants.

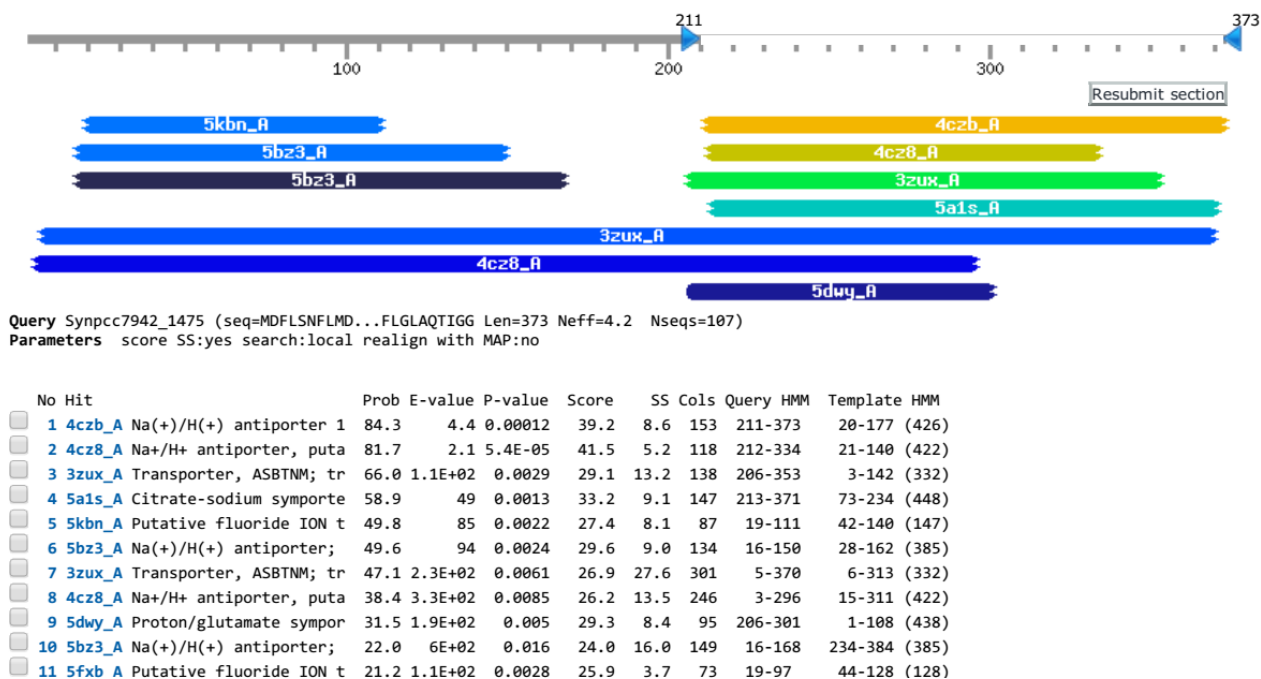


Figure 3-9: HHpred results for SbtA from *Synechococcus elongatus* PCC 7942. The templates with high homology are colored ‘yellow’ while the ones having remote homology are ‘blue’ in color. The column ‘Prob’ gives the probability that the hit is homologous to the query.

Among the hits obtained, 4czb_A had maximum probability (84.3) to SbtA from *Synechococcus elongatus* PCC 7942 but had less query cover. Therefore the protein showing maximum query cover (3zux_A) was initially used as template and the model was built using the program MODELLER 9v11 (Šali and Blundell, 1993). It is worth noting that no Na⁺/HCO₃⁻ symporter structure has been submitted to PDB till date. The best model with minimum DOPE score was selected and further evaluated for quality and atomic content using different online servers like PROCHECK (Laskowski et al., 1993), VERIFY3D (Eisenberg et al., 1997), ERRAT (Colovos and Yeates, 1993), WHAT IF (Vriend, 1990) and RAMPAGE (Lovell et al., 2003).

The model generated on assessment with protein structure verification algorithm had poor quality. Further, all the hits obtained by HHpred server (Figure 3-9) were sequentially used to predict the structure but the low sequence identities and low query cover combined to yield a relatively poor quality alignments and models for the protein. Finally, a more

sensitive threading tool viz., RaptorX structure prediction tool (Källberg et al., 2012) was used. RaptorX is an integer programming based protein threading software which distinguishes itself from other servers as it delivers high-quality structural models for many targets with only remote templates. It uses probabilistic graphical models in combination to statistical inference to model proteins from both single template and multiple templates by protein threading. A three dimensional model of SbtA was built using the RaptorX (Figure 3-10). The crystal structure of sodium proton antiporter, napA (PDB ID: 5BZ3) with resolution of 2.3 Å was used as a template for SbtA protein structure prediction.

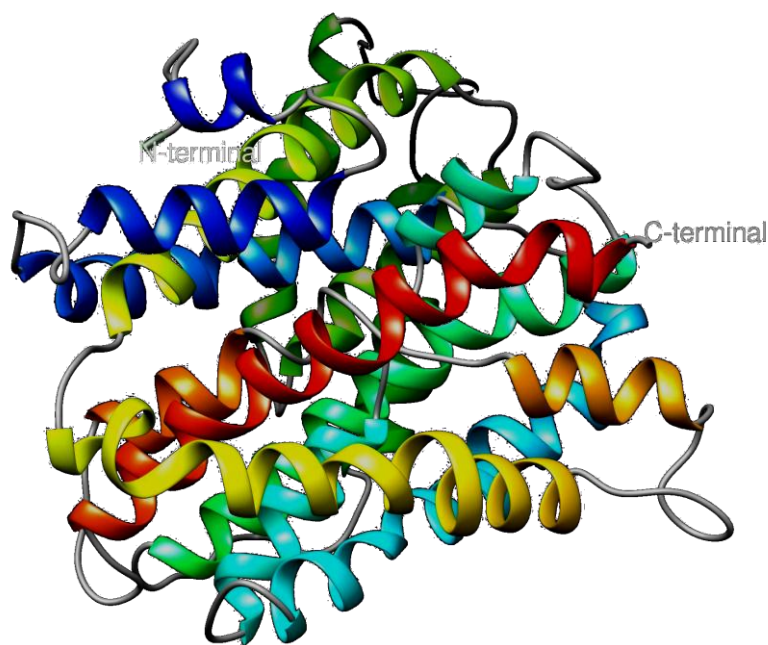


Figure 3-10: The protein model for SbtA protein from *Synechococcus elongatus* PCC 7942, generated by threading using RaptorX software (Källberg et al., 2012)

The model was further evaluated for quality and atomic content using different online servers like PROCHECK, ERRAT, VERIFY3D and RAMPAGE. The obtained model was further refined by energy minimization using online server YASARA (Krieger et al., 2009).

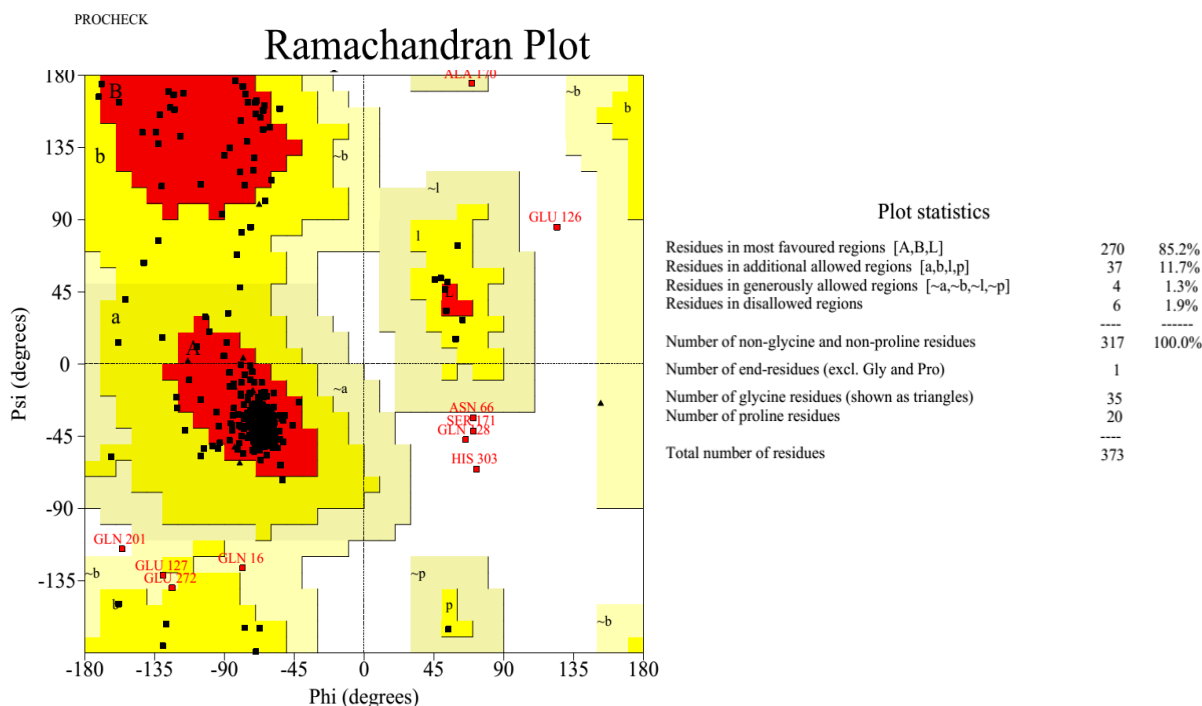


Figure 3-11: Validation of 3-dimensional structure of SbtA using Ramachandran Plot

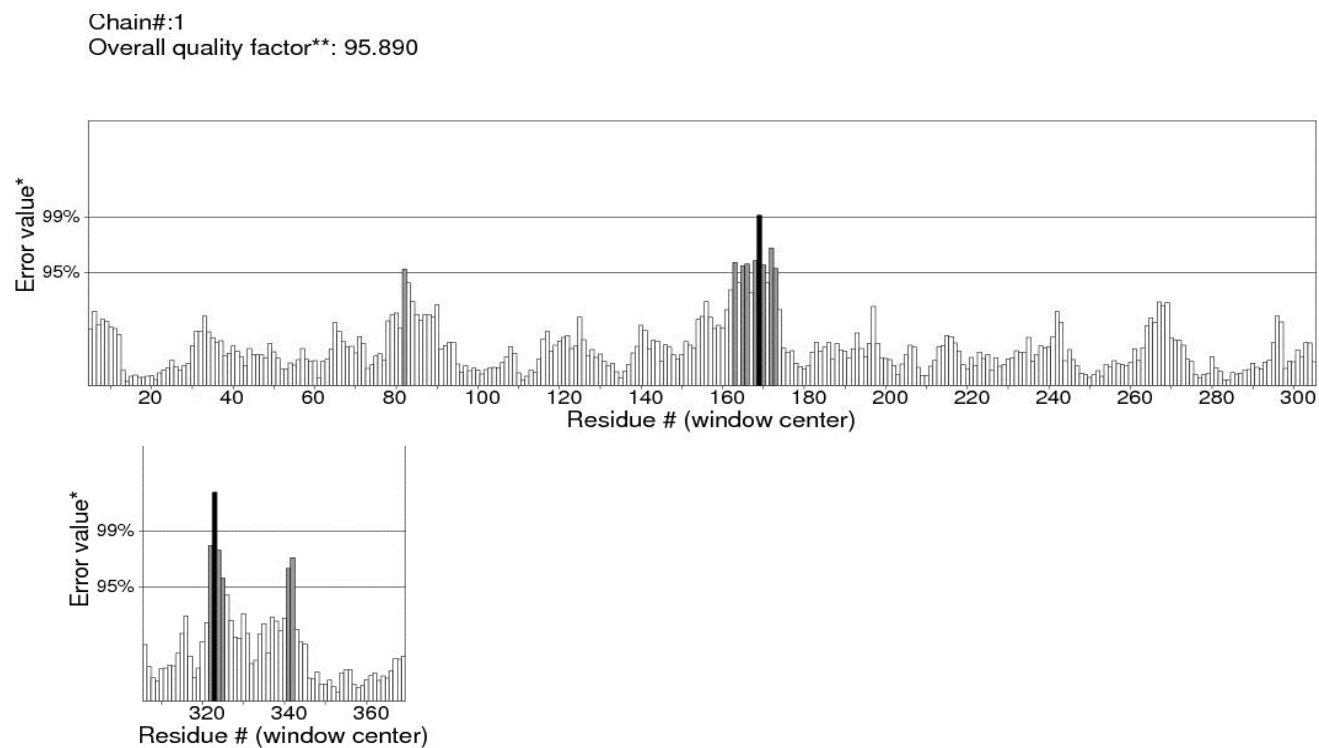


Figure 3-12: Validation of three dimensional structure of SbtA protein using ERRAT score

The structure showed a G-score 'BETTER' when evaluated by PROCHECK while an average profile was obtained with VERIFY3D. The main chain conformations for 85.2% of amino acid residues were within the favored or allowed regions in the Ramachandran plot analysis (Figure 3-11). The ERRAT score for the protein obtained was 95.890, which indicated that the molecular geometry of the model is of good quality (Figure 3-12).

For SbtA protein to be functionally active in higher plant chloroplasts, an understanding of the regulation of SbtA is essential. It has been speculated that SbtB is involved in regulation of SbtA by inhibiting its activity in dark by post-translational regulation. The expression of SbtB protein (a peptide of 110 amino acids) also occurs during low Ci conditions.

Recently, Du et al., (2014) observed that SbtB inhibits the activity of SbtA by directly interacting with it. The mechanism involved and the regulations which mediate the regulatory activity of SbtB are still unknown. The SbtA-SbtB interaction has been assumed to be quite similar to AmtB-GlnK interaction (Du et al., 2014). The AmtB-GlnK is involved in controlling ammonium influx in response to the intracellular nitrogen status in bacteria, fungi and plants. The regulation of AmtB by GlnK occurs by modification (uridylylation) of a key residue in GlnK (Y⁵¹) protein. Once uridylylated, GlnK cannot bind to AmtB and when deuridylylated, it can easily bind to AmtB (Conroy et al., 2007).



Figure 3-13: Secondary structure of SbtB from *Synechococcus elongatus* PCC 7942 predicted by Phyre² server. The green helices represent α -helices, blue arrows indicate β -strands and faint lines indicate coil. The 'SS confidence' line indicates the confidence in the prediction, with red being high confidence and blue low confidence. The 'Disorder' line predicts the disordered regions in the protein and such regions are indicated by question marks (?). The 'Disorder confidence' line indicates the confidence in the prediction of disordered regions.

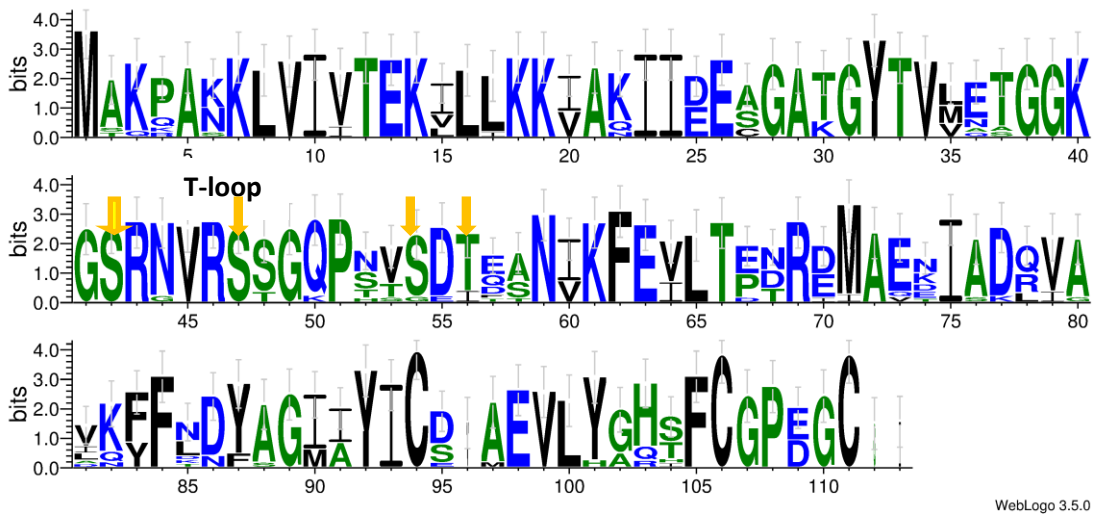


Figure 3-14: The sequence logo for SbtB sequences from 20 cyanobacterial species provided in the Kazusa cyanobase. The areas marked by yellow arrow show the conserved residues in the T-loop of the sequence.

The secondary structure of SbtB protein reflects that the protein has 4 stranded β -sheets and 2 helices (Figure 3-13). The structure of SbtB differs from other cyanobacterial P_{II} proteins [16.5% sequence identity and 31.9% similarity, (Forchhammer and Lüddecke, 2016)] as it lacks the adenylate-binding motifs which is a characteristic of P_{II} proteins. The T-loop (18 residues long) between the 2nd and 3rd β -sheet (similar to T-loop in GlnK) has conserved S/serine and T/threonine residues which could act as potential sites for phosphorylation/dephosphorylation (Figure 3-14). These residues could be phosphorylated/dephosphorylated in the presence of light/or by presence of certain metabolites whose concentration increases in presence in light thus rendering SbtB incapable of interacting with SbtA. The inefficiency of SbtB to repress SbtA activity keeps the transporter functionally active. An exactly opposite phenomenon might occur in dark where SbtB binds to SbtA and inactivates it, thereby preventing the uptake of bicarbonate in dark.

3.3. Screening of transit peptides (TNaXTP and TMDTP) from *A. thaliana* proteome

The primary requisite for installing SbtA to the inner envelope of chloroplast is the presence of signal sequences upstream of the gene which is essential for directing the protein into chloroplasts (Price et al., 2011). Transit peptides are responsible for the transport of a protein encoded by a nuclear genome to a particular organelle via secretory pathway. It is a 5-30 amino acid long stretch having a core of hydrophobic amino acids and most of them possess a short positively charged stretch of amino acids in the beginning (Bhushan et al., 2006). At the end there are specific sequences which are recognized and cleaved by signal peptidases either during or after completion of translocation to generate a free signal peptide and a mature protein. Since SbtA transporter is located on the plasma membrane in cyanobacteria and the structure of inner envelope membrane (IEM) is analogous to cyanobacterial plasma membrane, transit peptides were chosen from proteins located in the chloroplast inner envelope of *A. thaliana*.

The inner envelope membrane houses many proteins; mostly transporters (Linka and Weber, 2010), whereas the outer envelope membrane contains only about 30 different proteins (in *A. thaliana*), which include proteins of the outer membrane import machinery, lipid biosynthetic enzymes, few solute carriers and proteins involved in organelle division (Inoue, 2011). Most of the IEM proteins in vascular plants are almost exclusively encoded

by the nucleus (Ferro et al., 2002), except for the chloroplast genome (plastome) encoded *cemA* gene product (Wicke et al., 2011). These proteins possess a cTP (chloroplast transit peptide) or N-terminal peptide sequences of ~90-115 amino acids in front of their first transmembrane domain which is presumed to direct them to the desired location (Bhushan et al., 2006). The pre-sequences of these proteins thus can be used to target heterologous proteins of novel functions to inner envelope membrane by nuclear transformation. It is also important to know about the in-membrane topology of the transporter before making the choice of the cTP. For instance, *SbtA* has both N- and C-termini towards the periplasm. The cTP from a transporter with the same orientation in the membrane might increase the likelihood of similar alignment when introduced into the host plant.

An extensive *in-silico* analysis of *A. thaliana* envelope proteins was done from AT_CHLORO database in order to select most suitable protein candidates that could be used as the source of transit peptide (TP) elements for *SbtA* transporter of cyanobacteria for designing fusion gene constructs. The AT_CHLORO database stores the information about 1856 *A. thaliana* proteins and compiles results from LC-MS/MS analyses of highly purified sub-fractions of the four major chloroplast sub-compartments viz., envelope, stroma, thylakoids (Ferro et al., 2010) and sub-thylakoidal (Grana, stroma-lamellae). Among 1856 proteins, 487 proteins were found to be located on the envelope out of which 169 proteins were located exclusively on the inner chloroplast envelope and were therefore used for screening of most suitable protein candidates. Shortlisting of relevant/suitable proteins was done on the basis of the following criteria:

The protein must:

- Be localized at the inner envelope membrane of the C₃ plant chloroplast
- Be nuclear encoded and chloroplast targeted
- Contain transmembrane (TM) domains
- Play a role in the transport of ions or other molecules

Further, based on the criteria given above, 40 proteins were shortlisted, 38 of which were located at chloroplast inner membrane whereas 2 were localized to chloroplast envelope and chloroplast thylakoid (Table 3-5).

Table 3-5: Proteins shortlisted from list of inner envelope membrane proteins provided in the AT_CHLORO database having transporter like function and which form an integral component of chloroplast inner envelope

S.No.	Accession	Length (AA)	Location	Description and Function	S.No.	Accession	Length (AA)	Location	Description and Function
1	AT1G06950	1016	Ch/E/IM	Translocon at inner envelope membrane of chloroplast (Tic)	21	AT4G25450	714	Ch/E/IM	ATNAP8 (<i>Arabidopsis thaliana</i> non-intrinsic ABC protein 8)
2	AT1G32080	512	Ch/E/IM	Membrane putative protein/ transporter	22	AT4G25650	536	Ch/E/IM	ACD1-LIKE; electron carrier (Protein targeting Tic)
3	AT1G65410	345	Ch/E/IM	ATNAP11 (Transporter ABC lipids)	23	AT4G28210	348	Ch/E/IM	EMB1923 (transporter aa)
4	AT1G70610	700	Ch/E/IM	ATTAP1 (Transporter ABC)	24	AT4G32400	392	Ch/E/IM	SHS1 (sodium hypersensitive 1)
5	AT1G78180	418	Ch/E/IM	Transporter ATP/ADP	25	AT4G33350	268	Ch/E/IM	Chloroplast inner membrane import protein Tic22, putative
6	AT1G78560	401	Ch/E/IM	Bile acid:sodium Symporter family protein (Transporter Na/X)	26	AT4G33460	271	Ch/E/IM	ATNAP13 (Transporter ABC)
7	AT1G80300	624	Ch/E/IM	ATNTT1: ATP:ADP antiporter	27	AT5G01500	415	Ch/E/IM	Mitochondrial substrate carrier family protein
8	AT2G02590	324	Ch/E/IM	similar to unnamed protein product (GB:CAO41025.1), Transporter multidrug	28	AT5G03910	634	Ch/E/IM	ATATH12 (<i>Arabidopsis</i> ABC Transporter homolog 12)
9	AT2G15290	296	Ch/E/IM	ATTIC21/CIA5/PIC1/TIC 21 (chloroplast import apparatus)	29	AT5G10490	673	Ch/E/IM	MSL2 (MSCS- LIKE 2)
10	AT2G24820	539	Ch/E/IM	TIC55	30	AT5G14100	278	Ch/E/IM	ATNAP14 (Non-intrinsic ABC protein 14)
11	AT2G29650	512	Ch/E/IM	Inorganic phosphate transporter, putative	31	AT5G16150	546	Ch/E/IM	GLT1/PGLCT (glucose transporter 1)
12	AT2G38550	335	Ch/E/IM	similar to unknown protein (TAIR:AT3G57280.1)	32	AT5G17520	415	Ch/E/IM	RCP1 (ROOT CAP 1), Transporter maltose
13	AT2G42770	232	Ch/E/IM	Peroxisomal membrane 22 kDa family protein	33	AT5G17630	417	Ch/E/IM	glucose-6-phosphate/phosphate transporter
14	AT2G47840	208	Ch/E/IM	TIC20 protein	34	AT5G22830	459	Ch/E/IM	GMN10 (<i>Arabidopsis thaliana</i> Mg transporter 10)
15	AT3G13070	661	Ch/E/IM	CBS domain-containing protein (transporter Mg& Co)	35	AT5G33320	408	Ch/E/IM	Transporter Pi/PEP
16	AT3G46980	533	Ch/E/IM	transporter- related(transporter Pi)	36	AT5G42130	412	Ch/E/IM	Mitochondrial substrate carrier family protein
17	AT3G51870	381	Ch/E/IM	Mitochondrial substrate carrier family protein (Transporter MCF Ca/X)	37	AT5G43745	817	Ch/E/IM	phosphotransferase-related transporter ion channel
18	AT3G56160	436	Ch/E/IM	Bile acid:sodium Symporter(Transporter Na/X)	38	AT5G59250	558	Ch/E/IM	Sugar transporter family protein (transporter D- xylose-H ⁺)
19	AT4G00370	541	Ch/E/IM	ANTR2 (anion transporter 2); organic anion transmembrane transporter	39	AT2G26900	409	Ch/E & Ch/Th	Bile acid: sodium symporter family protein(transporter Na/X)
20	AT4G24460	431	Ch/E/IM	similar to unknown protein (TAIR:AT5G19380.1)	40	AT5G12860	557	Ch/E & Ch/Th	DIT1 (dicarboxylate transporter 1)

3.3.1. Transit peptide prediction

For transit peptide prediction in the shortlisted proteins, FASTA sequences of all the forty proteins were retrieved from TAIR database (<https://www.arabidopsis.org/>) and subjected to transit peptide prediction using ChloroP 1.1 and TargetP 1.1 servers (Table 3-6 and Table 3-7). On the basis of results obtained by ChloroP 1.1 server, two proteins were excluded (Proteins with S. No. 16 and 19) because they lacked the presence of chloroplastic transit peptide (cTP). Further, four proteins (S. No. 13, 16, 24, 36) were eliminated on the basis of TargetP 1.1 server results because of uncertainty in their localization and one more protein (S. No. 21) because of its mitochondrial location. Thus, a total of 6 proteins were excluded from the 40 proteins after analysis of the proteins by transit peptide prediction tools.

3.3.2. Membrane topology analysis

All the 40 proteins were also subjected to topology prediction for determining the orientation of N- and C-termini using TOPCONS, TMHMM and OCTOPUS topology prediction softwares (Table 3-8). As per the results obtained by the topology prediction softwares, eleven proteins were excluded either because of no TM helix detection or because of uncertainty in the topology of their N- or C- terminus.

Four proteins were found to be most suitable candidates based on the number of transmembrane helices, function as well as the location of the N- and C-termini. The proteins were AT1G78560 (putative transporter Na/X), AT2G26900 (putative transporter Na/X), AT3G56160 (putative transporter Na/X) and AT2G02590 (putative transporter multidrug). Although, all the four proteins equally justified all the criteria required to be chosen as the source of TP element, only two of the proteins were finally used to generate fusion constructs with *sbtA* which were AT3G56160.1 (TNaXTP) and AT2G02590.1 (TMDTP) each of them having 141 and 234 nucleotides respectively (Table 3-9).

Table 3-6: Prediction of the length and the presence of chloroplast transit peptide (cTP) cleavage sites in the shortlisted proteins by using ChloroP 1.1 Prediction Server. Score: indicates the certainty for the presence of cTP in a given protein sequence, cTP: Chloroplastic transit peptide, CS score: cleavage site score and Y: chloroplast located (Yes).

S.No	Accession No.	Protein Length	Score	cTP	CS-score	cTP-length
1	AT1G06950	1016	0.569	Y	8.643	50
2	AT1G32080	512	0.557	Y	1.671	13
3	AT1G65410	345	0.564	Y	0.119	46
4	AT1G70610	659	0.555	Y	6.905	59
5	AT1G78180	418	0.535	Y	2.947	82
6	AT1G78560	401	0.576	Y	3.671	70
7	AT1G80300	624	0.502	Y	2.919	21
8	AT2G02590	324	0.579	Y	0.531	78
9	AT2G15290	296	0.57	Y	5.092	15
10	AT2G24820	539	0.552	Y	5.286	48
11	AT2G29650	512	0.578	Y	7.259	59
12	AT2G38550	335	0.54	Y	7.985	72
13	AT2G42770	232	0.508	Y	1.862	46
14	AT2G47840	208	0.588	Y	5.579	49
15	AT3G13070	661	0.518	Y	0.328	71
16	AT3G46980	533	0.482	-	1.231	52
17	AT3G51870	381	0.51	Y	0.612	26
18	AT3G56160	436	0.554	Y	2.79	47
19	AT4G00370	541	0.493	-	2.954	28
20	AT4G24460	431	0.567	Y	5.592	79
21	AT4G25450	714	0.549	Y	3.091	24
22	AT4G25650	536	0.582	Y	4.858	55
23	AT4G28210	348	0.587	Y	6.679	72
24	AT4G32400	392	0.538	Y	1.431	88
25	AT4G33350	268	0.545	Y	2.398	59
26	AT4G33460	271	0.556	Y	2.791	48
27	AT5G01500	415	0.511	Y	0.943	61
28	AT5G03910	634	0.561	Y	6.255	51
29	AT5G10490	673	0.502	Y	2.415	75
30	AT5G14100	278	0.531	Y	-1.634	49
31	AT5G16150	546	0.566	Y	10.898	80
32	AT5G17520	415	0.568	Y	7.208	47
33	AT5G17630	417	0.54	Y	2.049	55
34	AT5G22830	459	0.567	Y	3.074	62
35	AT5G33320	408	0.567	Y	3.342	30
36	AT5G42130	412	0.502	Y	0.578	92
37	AT5G43745	817	0.562	Y	0.967	55
38	AT5G59250	558	0.551	Y	1.475	31
39	AT2G26900	409	0.543	Y	9.061	74
40	AT5G12860	557	0.587	Y	7.601	69

Table 3-7: TargetP 1.1 server results showing the frequency of presence of chloroplastic transit peptide (cTP), mitochondrial targeting peptide (mTP) or secretory pathway signal peptide (SP) in the proteins along with their location inside the cell. 'C' stands for Chloroplast, 'M' for Mitochondria and 'RC' for reliability class values.

S.No	Accession No.	Protein Length	cTP	mTP	SP	Other	Location	RC
1	AT1G06950	1016	0.941	0.033	0.005	0.078	C	1
2	AT1G32080	512	0.93	0.051	0.018	0.058	C	1
3	AT1G65410	345	0.886	0.106	0.055	0.05	C	2
4	AT1G70610	659	0.874	0.451	0.001	0.035	C	3
5	AT1G78180	418	0.724	0.038	0.018	0.574	C	5
6	AT1G78560	401	0.986	0.092	0.005	0.009	C	1
7	AT1G80300	624	0.542	0.412	0.005	0.165	C	5
8	AT2G02590	324	0.989	0.011	0.025	0.077	C	1
9	AT2G15290	296	0.85	0.181	0.003	0.027	C	2
10	AT2G24820	539	0.897	0.054	0.038	0.082	C	1
11	AT2G29650	512	0.986	0.086	0.002	0.025	C	1
12	AT2G38550	335	0.942	0.076	0.003	0.074	C	1
13	AT2G42770	232	0.158	0.274	0.027	0.782	-	3
14	AT2G47840	208	0.909	0.31	0.004	0.026	C	3
15	AT3G13070	661	0.794	0.08	0.003	0.278	C	3
16	AT3G46980	533	0.281	0.115	0.02	0.405	-	5
17	AT3G51870	381	0.513	0.113	0.009	0.493	C	5
18	AT3G56160	436	0.953	0.059	0.014	0.034	C	1
19	AT4G00370	541	0.684	0.21	0.003	0.152	C	3
20	AT4G24460	431	0.743	0.042	0.002	0.305	C	3
21	AT4G25450	714	0.539	0.697	0.002	0.018	M	5
22	AT4G25650	536	0.981	0.124	0.008	0.004	C	1
23	AT4G28210	348	0.972	0.07	0.009	0.005	C	1
24	AT4G32400	392	0.276	0.134	0.015	0.667	-	4
25	AT4G33350	268	0.679	0.07	0.004	0.083	C	3
26	AT4G33460	271	0.754	0.367	0.003	0.027	C	4
27	AT5G01500	415	0.764	0.042	0.01	0.316	C	3
28	AT5G03910	634	0.734	0.272	0.002	0.035	C	3
29	AT5G10490	673	0.862	0.342	0.002	0.015	C	3
30	AT5G14100	278	0.891	0.036	0.043	0.216	C	2
31	AT5G16150	546	0.674	0.5	0.005	0.029	C	5
32	AT5G17520	415	0.94	0.046	0.006	0.083	C	1
33	AT5G17630	417	0.901	0.206	0.003	0.02	C	2
34	AT5G22830	459	0.888	0.038	0.017	0.151	C	2
35	AT5G33320	408	0.975	0.08	0.002	0.005	C	1
36	AT5G42130	412	0.525	0.033	0.013	0.592	-	5
37	AT5G43745	817	0.439	0.283	0.031	0.012	C	5
38	AT5G59250	558	0.907	0.102	0.043	0.035	C	1
39	AT2G26900	409	0.732	0.201	0.002	0.053	C	3
40	AT5G12860	557	0.937	0.057	0.002	0.013	C	1

Table 3-8: Prediction of length and number of transmembrane helices, membrane protein topology and signal peptides using TOPCONS, TMHMM and OCTOPUS servers

S. No.	Gene Acc. No.	Length	No. of TM helix			N-termini			C-termini		
			TOP	THM	OCT	TOP	THM	OCT	TOP	THM	OCT
1	AT5G33320	408	10	6	9	IN	IN	IN	IN	IN	OUT
2	AT5G12860	557	14	13	14	IN	OUT	IN	IN	IN	OUT
3	AT1G32080	512	11	12	12	OUT	IN	IN	IN	IN	IN
4	AT1G80300	624	10	9	11	IN	IN	IN	IN	OUT	OUT
5	AT4G24460	431	10	10	9	IN	IN	OUT	IN	IN	IN
6	AT5G16150	546	12	10	12	IN	IN	IN	IN	IN	IN
7	AT5G59250	558	12	11	12	IN	IN	IN	IN	OUT	IN
8	AT1G78560	401	9	9	9	IN	OUT	IN	OUT	IN	OUT
9	AT3G46980	533	12	12	12	IN	IN	IN	OUT	IN	IN
10	AT4G00370	541	12	10	12	IN	IN	IN	IN	IN	IN
11	AT2G26900	409	9	9	9	IN	IN	IN	OUT	OUT	OUT
12	AT3G56160	436	8	7	8	IN	OUT	OUT	IN	IN	OUT
13	AT2G02590	324	7	6	6	OUT	OUT	OUT	IN	OUT	OUT
14	AT2G29650	512	12	10	10	IN	IN	IN	IN	IN	IN
15	AT5G17520	415	9	9	5	IN	OUT	OUT	OUT	IN	IN
16	AT5G17630	417	8	7	10	IN	OUT	IN	IN	IN	IN
17	AT3G13070	661	5	4	5	IN	OUT	IN	OUT	OUT	OUT
18	AT5G10490	673	5	5	4	IN	IN	OUT	OUT	OUT	IN
19	AT4G28210	348	9	3	7	IN	OUT	IN	OUT	IN	OUT
20	AT4G25450	714	6	5	6	IN	IN	IN	IN	OUT	IN
21	AT5G03910	634	6	2	5	IN	OUT	IN	IN	OUT	IN
22	AT1G70610	659	4	2	4	IN	OUT	OUT	IN	OUT	IN
23	AT1G78180	418	2	0	6	OUT	-	OUT	OUT	-	OUT
24	AT2G15290	296	4	3	4	IN	IN	IN	IN	OUT	IN
25	AT2G38550	335	4	3	3	IN	IN	IN	IN	OUT	OUT
26	AT5G42130	412	0	3	6	-	OUT	OUT	-	IN	OUT
27	AT4G25650	536	2	2	2	IN	OUT	IN	IN	OUT	IN
28	AT5G43745	817	3	2	2	IN	OUT	IN	OUT	OUT	IN
29	AT2G24820	539	2	2	2	IN	OUT	IN	IN	OUT	IN
30	AT2G42770	232	3	2	3	IN	IN	IN	OUT	IN	OUT
31	AT2G47840	208	4	4	4	IN	IN	IN	IN	IN	IN
32	AT3G51870	381	0	2	6	-	IN	OUT	-	IN	OUT
33	AT4G32400	392	0	0	6	-	-	OUT	-	-	OUT
34	AT4G33350	268	0	0	0	-	-	-	-	-	-
35	AT5G01500	415	0	0	6	-	-	OUT	-	-	OUT
36	AT1G65410	345	0	0	0	-	-	-	-	-	-
37	AT1G06950	1016	3	0	2	IN	-	IN	OUT	-	IN
38	AT4G33460	271	0	0	0	-	-	-	-	-	-
39	AT5G14100	278	0	0	0	-	-	-	-	-	-
40	AT5G22830	409	0	5	0	-	IN	-	-	OUT	-

Table 3-9: Details of the transit peptide, their source protein and the function of the protein in the inner envelope membrane based on the information retrieved from TAIR database

S. No.	Name of Transit Peptide	Gen. Acc. No. of TP source protein	No. of TM helices	Length (nt)	Length (aa)	Sequence of TP	Function of the protein
1.	TNaXTP	AT3G56160	8	141	47	MAIASTLASTQNPFCLRQ PPSPGNRSVVFRRCDPCG RRWISRSIR	Bile acid: Sodium symporter
2.	TMDTP	AT2G02590	7	234	78	MAISTLLSISSFTLSSSSSTK THLLTSIPASRVYSYPISPK WKQIRFLQSQSSSLFYPLR RNFTRFCSSPDGFLRNTK	Transporter Multidrug [similar to unnamed protein product (GB:CAO41025.1)]

3.4. Codon usage analysis of *sbtA* transporter gene

For SbtA transporter to be functionally active in plant system, it is essential to determine the similarity/preferences in the codon usage between *sbtA* and the genome of the model plants. The determination of codon usage plays a critical role when recombinant proteins are expressed in different organisms as it can help in deciding whether the sequence needs to be optimized for heterologous gene expression or not. Hence, codon usage of the *sbtA* gene was analyzed with respect to codon preferences in cyanobacteria (*Synechocystis* sp. PCC 6803) and plants (*Arabidopsis thaliana* and *Nicotiana tabacum*) using the GCUA (Fuhrmann et al., 2004), GenScript Codon Analysis Tool and the CAIcal software (Puigbò et al., 2008). The details of the softwares have been provided in sec. 2.1.4 in chapter 2.

The GCUA software calculates the relative adaptiveness of the codon usage between the gene and genome of the host organism and gives the mean difference values. The codon usage in *sbtA* was determined with respect to the genomes of three organisms viz., *Synechocystis* sp. PCC 6803, *A. thaliana* and *N. tabacum*. The *Synechocystis* sp. PCC 6803 genome was used as a reference cyanobacterial genome because the genome of *Synechococcus elongatus* PCC 7942 was not available in the software used for analysis. The analysis of relative adaptiveness of codon usage in *sbtA* gene depicted a mean difference of approximately 24.34% in cyanobacteria (Figure 3-15) and 33.13% and 35.91% in *A. thaliana* and *N. tabacum* respectively (Figures 3-16 and 3-17). The results reflect that out of

the 100% value, approximately 65% codons of *sbtA* gene find good preference in the plant genome.

The percentage of low frequency codons was also analyzed using the GenScript Codon Analysis Tool, which gave a value of 3% and 8% in *A. thaliana* and *N. tabacum* respectively which is quite less than the acceptable percentage of low frequency (30%). The values more than 30% can reduce the efficiency of translation or even disengage the translational machinery.

Further, the similarity in the codon usage between the gene and genome of the host plants was calculated using the CAIcal software. The values of CAI range from 0 to 1 where a CAI of 1 is considered ideal. The lower the value, the higher the chance that the desired gene will be expressed poorly. The CAI values obtained for all the three organisms were 0.7, 0.66 and 0.63 in *Synechocystis* sp. PCC 6803, *A. thaliana* and *N. tabacum* respectively. The values were more towards maximum limit which indicated that there are good chances of expressing this cyanobacterial gene in the plants. The analysis depicted significant CAI values and low frequency of rare codons which offer good chances of the expression of the protein in the plant system.

SbtA transporter (red):
 sequence derived from *Synechococcus elongatus* PCC 7942

Codontable (black):
*Synechocystis*_PCC6803

Mean difference: 25.34 %

Ordinate (y-axis): relative adaptiveness

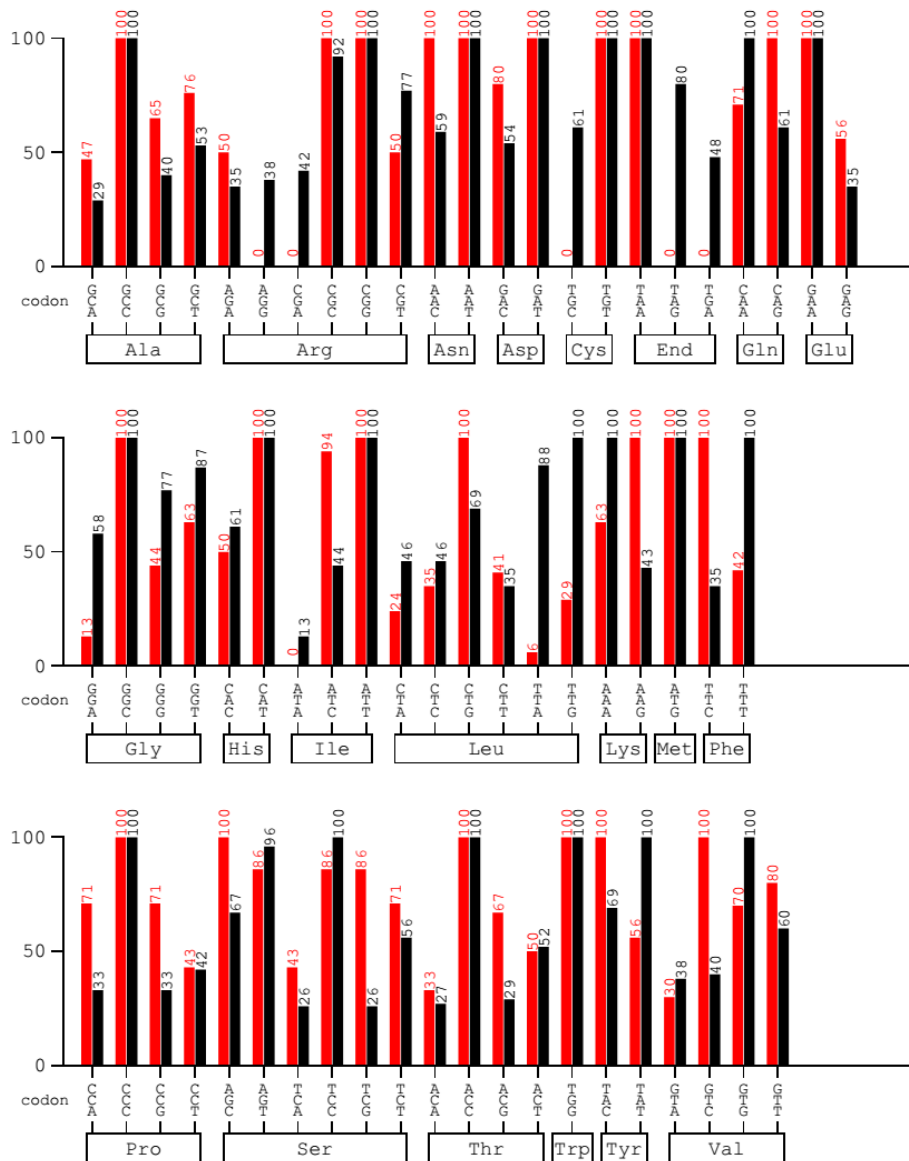


Figure 3-15: Relative adaptiveness plots for each amino acid of SbtA transporter (denoted by red color) with genome of *Synechocystis* sp. PCC 6803 (denoted by black color)

SbtA transporter (red):
 sequence derived from *Synechococcus elongatus* PCC 7942

Codontable (black):
Arabidopsis thaliana

Mean difference: 33.13 %

Ordinate (y-axis): relative adaptiveness

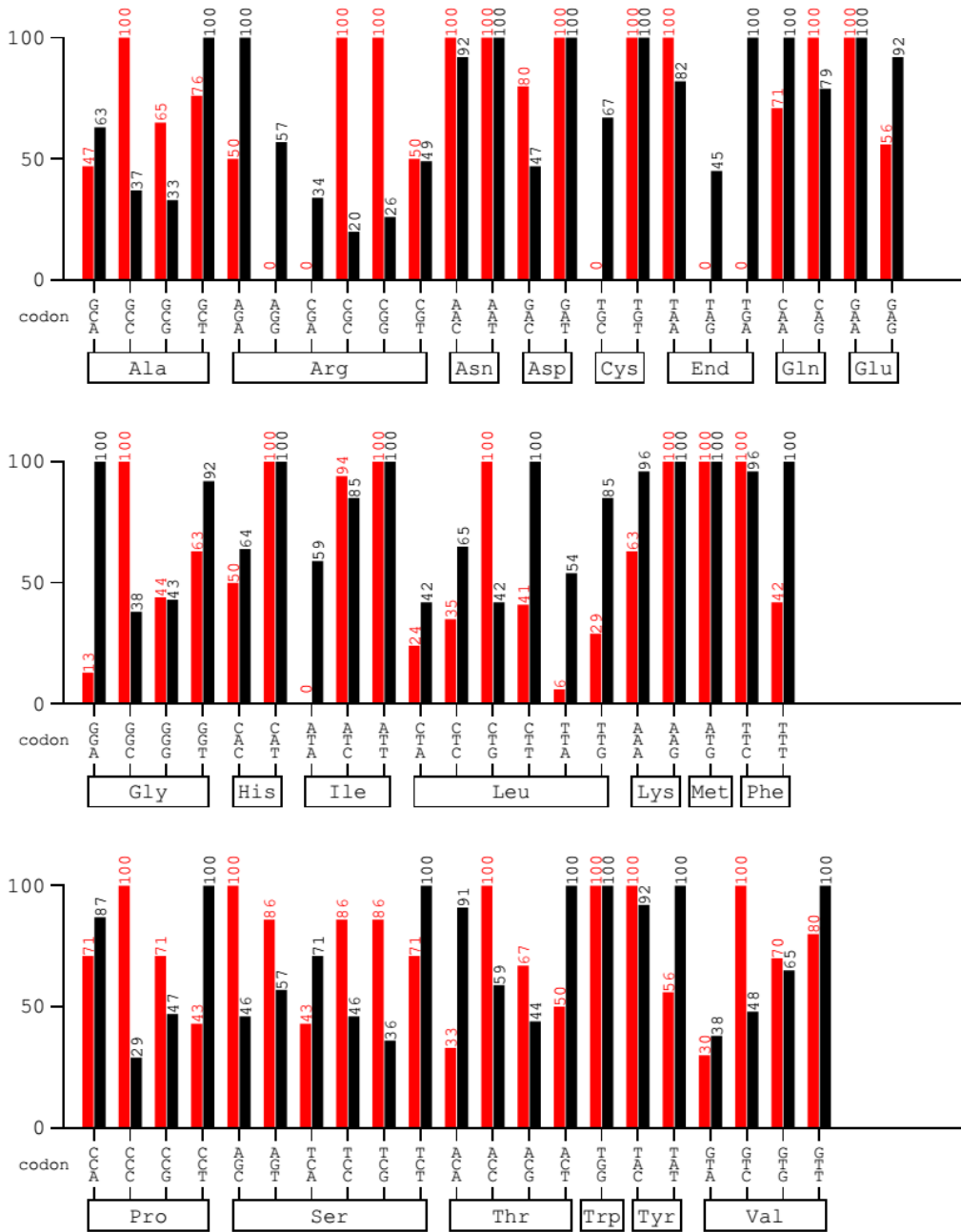


Figure 3-16: Relative adaptiveness plots for each amino acid of SbtA transporter (denoted by red color) with genome of *A. thaliana* (denoted by black color)

SbtA transporter (red):
 sequence derived from *Synechococcus elongatus* PCC 7942

Codontable (black):
Nicotiana tabacum

Mean difference: 35.91 % Ordinate (y-axis): relative adaptiveness

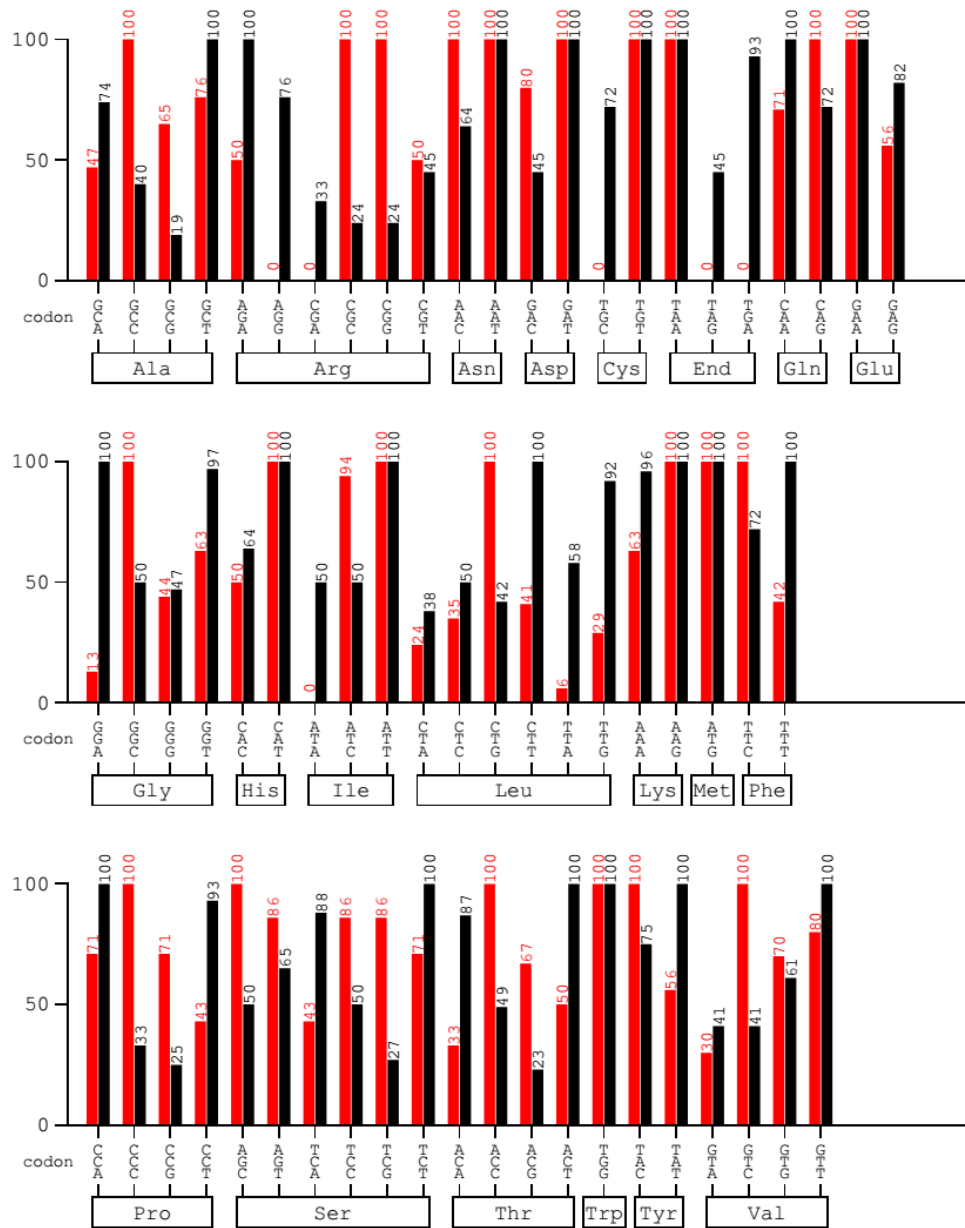


Figure 3-17: Relative adaptiveness plots for each amino acid of SbtA transporter (denoted by red color) with genome of *N. tabacum* (denoted by black color)

3.5. Synthesis of fusion constructs and their transformation into C₃ plants

For proper targeting of SbtA transporter to the chloroplast envelope of the plant system, it should possess a transit peptide sequence at its N-terminal region. Since SbtA does not naturally have any such sequence therefore it was fused with the TP sequences of plant origin. Suitable genetic constructs of *sbtA* transporter gene were prepared by using appropriate chloroplast targeting sequences (transit peptides) from N-terminus of chloroplast inner envelope proteins of *A. thaliana* named as TNaXTP and TMDTP respectively. The fusion genes (TNaXTP/TMDTP-*sbtA*) were initially cloned in intermediate cloning vector (*pColdI*) and thereafter were finally cloned into plant expression vectors viz. *pRII01-AN* and *pCAMBIA1302* having *gus* and *mgfp5* reporter genes respectively. Several trials were made to transform the model plants (*A. thaliana*/ *N. tabacum*/ *N. benthamiana*) by recombinant constructs. The preliminary studies were carried out by transforming *N. tabacum* leaves using recombinant *pRII01-AN* constructs (TNaXTP/TMDTP-*sbtA-gus-pRII01-AN*) in which the presence of gene was confirmed by PCR analysis of the recombinant DNA and the transient protein expression was confirmed by GUS reporter gene assay. Further, *N. benthamiana* leaves were transiently transformed using recombinant *pCAMBIA1302* constructs (TNaXTP/TMDTP-*sbtA-pCAMBIA1302*) and the incorporation of the fusion genes was confirmed at DNA and mRNA levels by performing diagnostic PCRs using gene specific primers. The expression was confirmed at the protein level by western blotting using anti-GFP primary antibody and horseradish peroxidase (HRP) conjugated secondary antibody. The localization of the protein was detected by confocal microscopy of isolated protoplasts. The recombinant *pCAMBIA1302* constructs were also used to generate transgenic plants by *Agrobacterium* mediated transformation. The plants were screened for the presence of gene and the positive transformants were raised to collect seeds.

3.5.1. PCR amplification of *sbtA* and transit peptide sequences

The amplification of transit peptides viz., TNaXTP and TMDTP was carried out by using *A. thaliana* genomic DNA as template and the *sbtA* gene (from *Synechococcus* sp. PCC 7942) was amplified from *sbtA-pGEMT* clone (gift from Price GD, ANU, Australia, Figure 3-18).

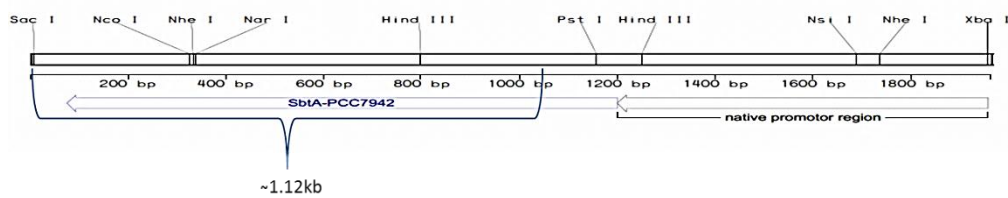


Figure 3-18: Vector map of *sbtA-pGEMT* clone (received as a kind gift from Prof. G. Dean Price, ANU, Canberra)

The *sbtA* gene was extracted by PCR due to absence of restriction enzyme site for the excision of the gene from the clone. The genomic DNA from *A. thaliana* was isolated using DNeasy Plant Mini Kit (Qiagen), detailed procedure has been mentioned in sec 2.2.3 in chapter 2. The details of primer sequences and the PCR conditions used have been provided in details in sec 2.2.4 in chapter 2. The expected size of the PCR products obtained from these amplifications were 141 bp, 234 bp and 1.1 kb for TNaXTP, TMDTP and *sbtA* respectively (Figure 3-19). Appropriate restriction sites were incorporated in forward and reverse primers for both transit peptides and *sbtA* to facilitate further cloning. In the case of *sbtA* amplicon, *KpnI* enzyme site was introduced by the forward primer and *BamHI* enzyme site by the reverse primer while for both transit peptides, *NdeI* enzyme site was incorporated by the forward primer and *KpnI* enzyme site by the reverse primer.

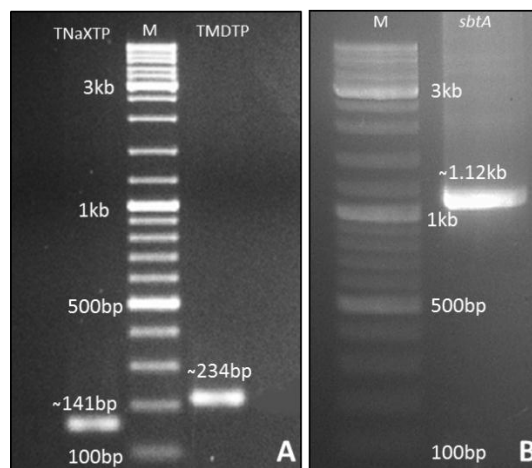


Figure 3-19: PCR products (A) Lane1: TNaXTP, Lane 2: Gene ruler SM0331 (Thermo- scientific), Lane 3: TMDTP. The products were run on a 1.5% agarose gel. (B) Lane 1: Gene ruler SM0331, Lane 2: *sbtA*. The product was run on a 0.9% agarose gel.

3.5.2. Cloning of PCR products in *pGEM®-T* vector

Since it was difficult to amplify transit peptide sequences from *A. thaliana* genome because of smaller size of the products which lead to non-specific amplifications, the PCR products were cloned in a TA vector (*pGEM®-T*). The detailed procedure has been mentioned in sec 2.2.8.1 in chapter 2 and in Figure 2-4. The TP-*pGEM®-T* and *sbtA-pGEM®-T* clones were used for the extraction of genes for further clonings. The cloning of TNaXTP, TMDTP and *sbtA* into *pGEM®-T* vector was confirmed by restriction analysis (Figure 3-20). The restriction analysis confirmed the presence of genes (TNaXTP: ~141 bp), (TMDTP: ~234 bp) and *sbtA*: ~1122 bp).

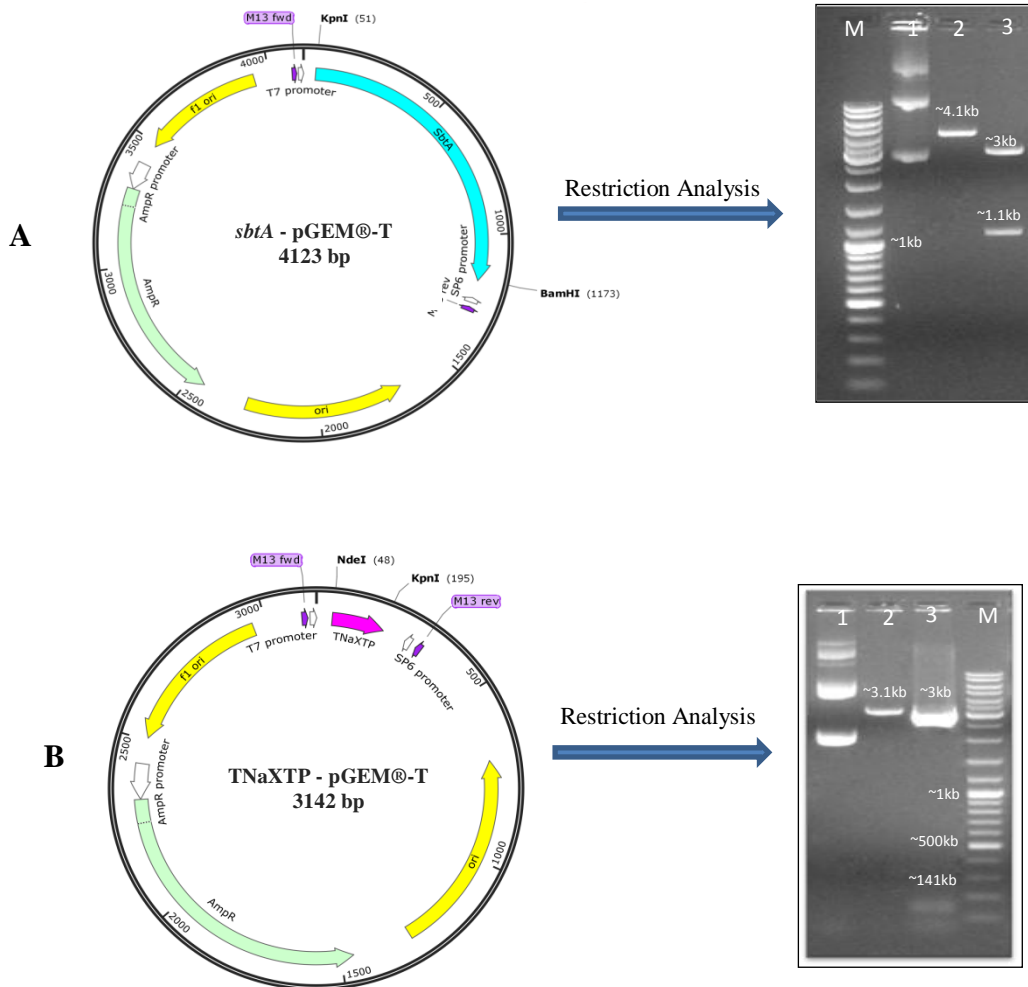


Figure 3-20: Legend on next page

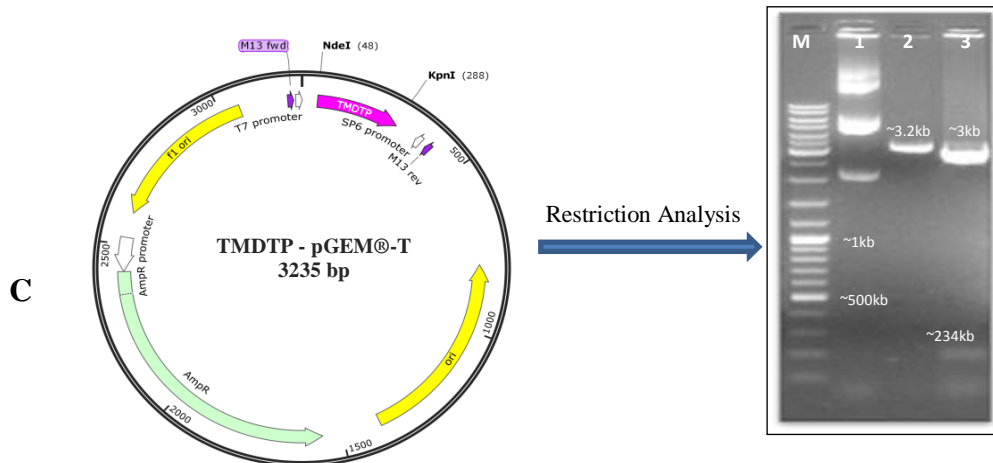


Figure 3-20: Recombinant vector maps and restriction digestion analysis to confirm cloning of (A) *sbtA* in *pGEM®-T* vector: Lane 1: Gene ruler SM0331 (Thermoscientific), Lane 2: Undigested *sbtA-pGEM®-T* recombinant plasmid, Lane 3: *sbtA-pGEM®-T* recombinant plasmid digested with *BamHI* and Lane 4: *sbtA-pGEM®-T* recombinant plasmid double digested with *BamHI* and *KpnI* enzymes. (B) *TNaXTP* in *pGEM®-T* vector: Lane 1: Undigested *TNaXTP-pGEM®-T* recombinant plasmid, Lane 2: *TNaXTP-pGEM®-T* recombinant plasmid digested with *NdeI* restriction enzyme, Lane 3: *TNaXTP-pGEM®-T* recombinant plasmid double digested with *NdeI* and *KpnI* enzymes and Lane 4: Gene ruler SM0331 (Thermoscientific) (C) **TMDTP** in *pGEM®-T* vector: Lane1: Gene ruler SM0331 (Thermoscientific), Lane 2: Undigested *TMDTP-pGEM®-T* recombinant plasmid, Lane 3: *TMDTP-pGEM®-T* recombinant plasmid digested with *NdeI* restriction enzyme and Lane 4: *TMDTP-pGEM®-T* recombinant plasmid digested with *NdeI* and *KpnI* enzymes.

3.5.3. Cloning of TP and *sbtA* gene in *pColdI* vector

In order to achieve proper targeting of *sbtA* gene into the desired location in the plant system, generation of fusion construct of *sbtA* with the TP sequence of plant origin is the prerequisite. Therefore, recombinant fusion constructs of *sbtA* were synthesized with TP sequences from two different chloroplast inner envelope located proteins viz., *TNaXTP* and *TMDTP*. For this, *TNaXTP-pGEM®-T*, *TMDTP-pGEM®-T* and *sbtA-pGEM®-T* recombinant clones were double digested with *NdeI* and *KpnI* (for taking out *TNaXTP* and *TMDTP*) and *KpnI* and *BamHI* (for taking out *sbtA* gene) enzymes. The sticky end products of *TNaXTP* and *TMDTP* obtained were purified and individually fused with *sbtA* gene by cloning them together into appropriate sites (*NdeI* and *BamHI*) of *pColdI* vector using

standard ligation procedure mentioned in sec 2.2.8.1 in chapter 2. The cloning was confirmed by restriction analysis (Figure 3-21) and DNA sequencing.

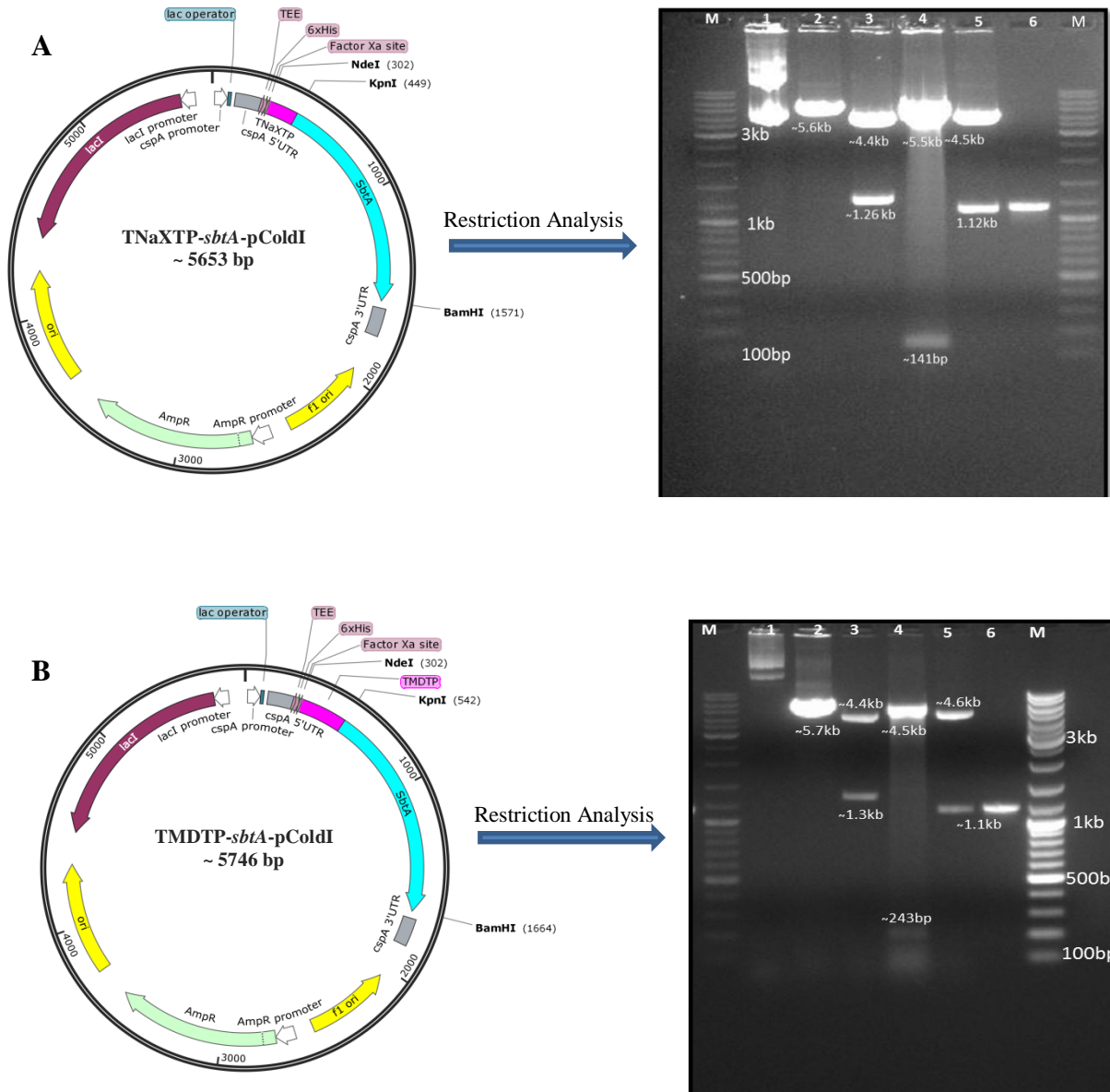


Figure 3-21: Legends on the next page

Figure 3-21: Recombinant vector maps and restriction analysis of recombinant *pColdI* clones (TNaXTP/TMDTP-*sbtA-pColdI*). **(A)** M- Gene ruler SM0331 (Thermoscientific), Lane 1: TNaXTP-*sbtA-pColdI* plasmid uncut, Lane 2: TNaXTP-*sbtA-pColdI* plasmid digested with *BamHI*, Lane 3: TNaXTP-*sbtA-pColdI* plasmid double digested with *NdeI* and *BamHI*, Lane 4: TNaXTP-*sbtA-pColdI* plasmid double digested with *NdeI* and *KpnI*, Lane 5: TNaXTP-*sbtA-pColdI* plasmid double digested with *KpnI* and *BamHI* and Lane 6: Gel eluted PCR product of *sbtA*. **(B)** M- Gene ruler SM0331 (Thermoscientific), Lane 1: TMDTP-*sbtA-pColdI* plasmid uncut, Lane 2: TMDTP-*sbtA-pColdI* plasmid digested with *BamHI*, Lane 3: TMDTP-*sbtA-pColdI* plasmid double digested with *NdeI* and *BamHI*, Lane 4: TMDTP-*sbtA-pColdI* plasmid double digested with *NdeI* and *KpnI*, Lane 5: TMDTP-*sbtA-pColdI* plasmid double digested with *KpnI* and *BamHI* and Lane 6: Gel eluted PCR product of *sbtA*.

3.5.4. Cloning of *gus* gene into *pRI101-AN* vector

The preliminary step of all plant transformation experiments is the cloning of desired genes into a plant expression vector. Therefore, the fusion genes (TNaXTP/TMDTP-*sbtA*) were extracted from *pColdI* clones and cloned into *pRI101-AN* vector. The *pRI101-AN* vector does not contain any reporter gene and therefore before cloning of the fusion genes a recombinant vector with *gus* reporter gene was synthesized. The *gus* gene was obtained from *pBI101* vector by double digestion with *BamHI* and *SacI* restriction enzymes and cloned between the same enzymes in the MCS region of *pRI101-AN* vector. The cloning was confirmed by restriction analysis of the recombinant clone (Figure 3-22) with *BamHI* and *SacI* enzymes and a band corresponding to *gus* gene (1.8 kb) was obtained.

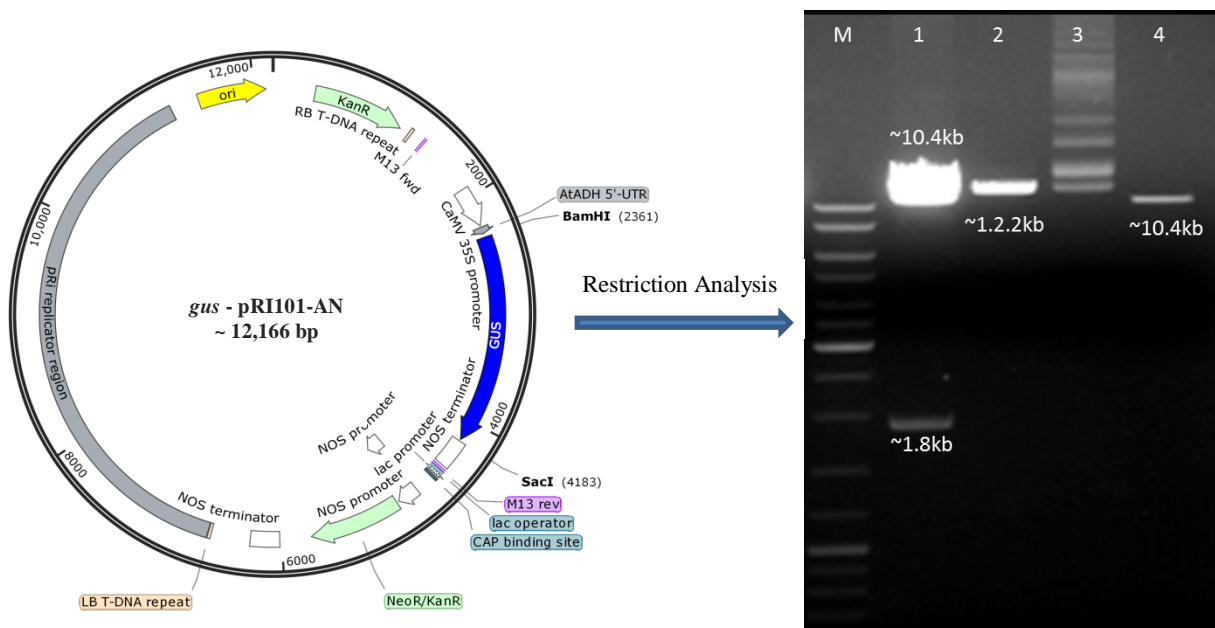


Figure 3-22: Recombinant vector map and restriction analysis for confirmation of cloning of *gus* gene into *pRI101-AN* vector. M: Gene ruler SM0331 (Thermoscientific). Lane 1: Recombinant *gus-pRI101-AN* plasmid double digested with *BamHI* and *SacI* restriction enzymes, Lane 2: Recombinant *gus-pRI101-AN* plasmid digested with *BamHI* enzyme, Lane 3: Uncut *pRI101-AN* plasmid, Lane 4: Empty *pRI101-AN* plasmid double digested with *BamHI* and *SacI* enzymes.

3.5.5. Cloning of fusion genes (TNaXTP/TMDTP-*sbtA*) into recombinant *gus-pRI101-AN* vector

The recombinant *pColdI* clones were digested with *NdeI* and *BamHI* enzymes and the extracted fusion genes (TNaXTP/TMDTP-*sbtA*) were cloned upstream of *gus* gene in *gus-pRI101-AN* vector using the same restriction enzymes (Figure 3-23). The cloning was confirmed by restriction analysis of both clones (Figures 3-24 and 3-25).

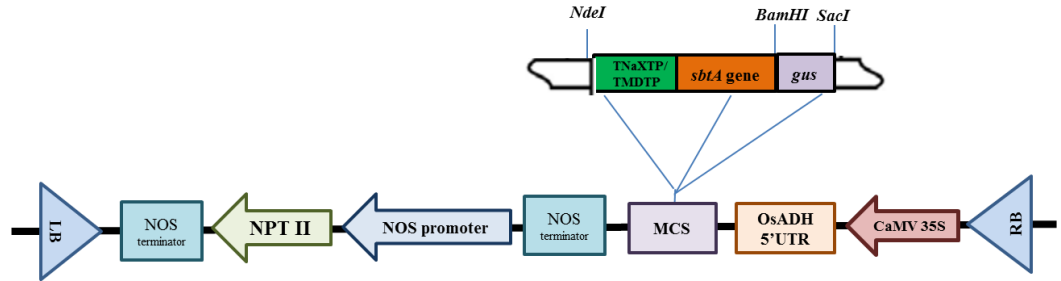


Figure 3-23: Schematic representation of the recombinant *gus-pRI101-AN* vector with TNaXTP/TMDTP-*sbta* insert engineered for plant transformation

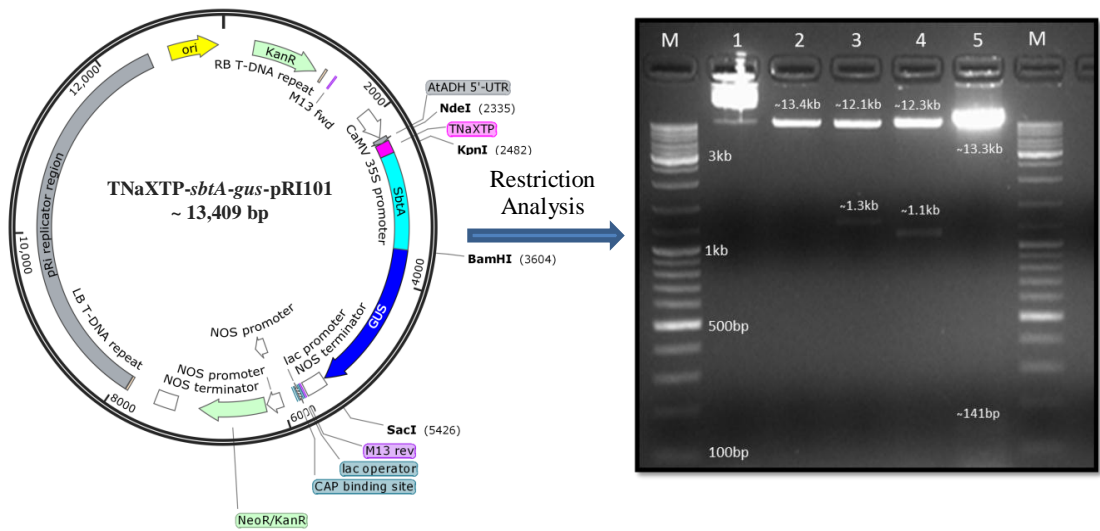


Figure 3-24: Recombinant vector map and restriction analysis of recombinant (TNaXTP-*sbta-gus-pRI101-AN*) clone. M-Gene ruler SM0331 (Thermoscientific), Lane 1: TNaXTP-*sbta-gus-pRI101-AN* plasmid uncut, Lane 2: TNaXTP-*sbta-gus-pRI101-AN* plasmid digested with *BamHI* enzyme, Lane 3: TNaXTP-*sbta-gus-pRI101-AN* plasmid double digested with *NdeI* and *BamHI* restriction enzymes, Lane 4: TNaXTP-*sbta-gus-pRI101-AN* plasmid double digested with *KpnI* and *BamHI* enzymes and Lane 5: TNaXTP-*sbta-gus-pRI101-AN* plasmid double digested with *NdeI* and *KpnI* enzymes.

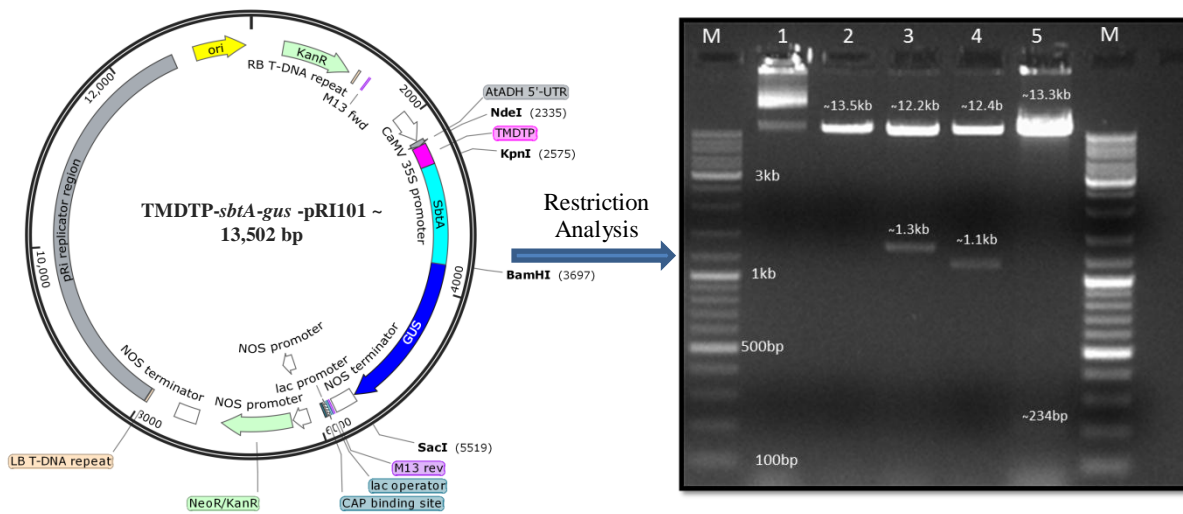


Figure 3-25: Recombinant vector map and restriction analysis of recombinant (TMDTP-*sbtA-gus-pRI101-AN*) clone. M- Gene ruler SM0331(Thermoscientific), Lane 1: TMDTP-*sbtA-gus-pRI101-AN* plasmid uncut, Lane 2: TMDTP-*sbtA-gus-pRI101-AN* plasmid digested with *BamHI* enzyme, Lane 3: TMDTP-*sbtA-gus-pRI101-AN* plasmid double digested with *NdeI* and *BamHI* enzymes, Lane 4: TMDTP-*sbtA-gus-pRI101-AN* plasmid double digested with *KpnI* and *BamHI* restriction enzymes and Lane 5: TMDTP-*sbtA-gus-pRI101-AN* plasmid double digested with *NdeI* and *KpnI* enzymes.

3.6. Plant transformation using recombinant *pRI101-AN* constructs

To study the effect of the *sbtA* gene on plant system, plant transformation studies were carried out by transforming *A. thaliana* and *N. tabacum* by both recombinant *pRI101-AN* constructs (TNaXTP/TMDTP-*sbtA-gus-pRI101-AN*). Since SbtA is itself a large hydrophobic protein and due to the large size of the binary vector (*gus-pRI101-AN*, 12,166 bp) used for transformation, several methods of plant transformation viz., particle bombardment, floral dip and *Agrobacterium* mediated co-culture method were attempted on multiple host system (*Arabidopsis* and *Nicotiana*) and their responses were analyzed. *Nicotiana* being easier to transform and having large infiltrable leaves was used to study both transient and stable expression of the gene while *A. thaliana* being best suited for vacuum infiltration (McIntosh et al., 2004) was used for floral dip experiments (Clough and Bent, 1998; Bechtold and Pelletier, 1998).

3.6.1. Transformation of *N. tabacum* by particle bombardment method

Both recombinant *pRI101-AN* plasmids (TNaXTP/TMDTP-*sbtA-gus-pRI101-AN*) were bombarded on *N. tabacum* leaves to study transient gene expression. The crude protein extract was extracted from the bombarded leaves after 48 hrs and GUS fluorometric assay was performed. The detailed procedure has been mentioned in sec 2.2.19 in chapter 2. In GUS fluorometric assay, the hydrolyzing action of GUS enzyme on the substrate 4-MUG yields a fluorescent product known as 4-MU. Using the relative fluorescence values of 4-MU and total protein concentration (present in the crude extract), the specific GUS activity in the samples was calculated. Specific GUS activity was found to be the highest for TNaXTP-*sbtA* construct in comparison to TMDTP-*sbtA* which was 0.19 and 0.04 nM-4MU/mg/min respectively. Since there was higher expression level of GUS in both the constructs (TNaXTP/TMDTP-*sbtA*) in comparison to the control sample (protein extracted from wild type leaves which have not undergone bombardment) it was confirmed that constructs were transformed successfully into the plant system (Figure 3-26).

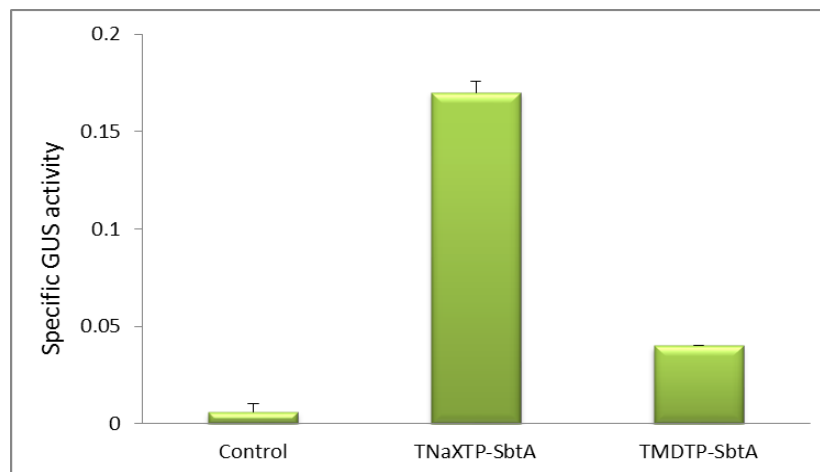


Figure 3-26: Fluorometric quantification of GUS activity (Mean \pm SD). Expression of MU (Methylumbelliferone) in control (wild type) and TNaXTP-*sbtA-gus-pRI101-AN* and TMDTP-*sbtA-gus-pRI101-AN* plasmid bombarded samples. The specific GUS activity is represented in nM-4MU/mg/min

Although particle bombardment technique is widely used for production of transient transformants in many animals and plants (Birch, 1997), the transformation efficiency of this method is not high (Kawamoto et al., 2016). Kikkert et al., (2005), obtained five transient transformants after bombardment by approximately 8000 particles per dish with a particle gun. There are high chances of gold particles getting washed away during fixation or dehydration (Kawamoto et al., 2016). Other than these, some levels of endogenous GUS activities in untransformed tissues (which may give false positives) has also been reported by some research groups (Hu et al., 1990). Wenzler et al., (1989) detected low levels of endogenous GUS activities in the leaves and roots of the potato and tobacco plants with the highly sensitive quantitative fluorometric assays. So as to avoid the chances of obtaining false positive results, intrinsic GUS activity was inhibited by adding methanol during protein extraction and the GUS activity values obtained in case of control sample were deducted from the samples while calculating specific GUS activity.

3.6.2. Transformation of *A. thaliana* by *Agrobacterium* mediated floral dip method

Floral dip is one of the most commonly used/followed methods of plant transformation because of its simplicity, efficiency and low cost (Mara et al., 2010). Other than this it does not require extensive tissue culture and regeneration unlike the *Agrobacterium* mediated co-culture method (Clough and Bent, 1998). The flowering shoots of 2-2.5 months old *A. thaliana* plants were transformed by *pRI101-AN* recombinant constructs (TNaXTP/TMDTP-*sbtA-gus-pRI 101-AN*) (Figure 3-27). The detailed procedure has been mentioned in sec 2.2.21 in chapter 2.

The transformed plants were returned to normal growth and were allowed to set (T_1) seeds. The seeds were screened on $\frac{1}{2}$ MS + Kan (50 mg/L) media (Figure 3-27). The seeds which were transformed with the recombinant constructs germinated on the media containing kanamycin while the untransformed ones could not survive. When the plants reached four leaflet stage they were removed cautiously from the media and transferred to soilrite mix in plastic pots. The plants were raised in growth chamber at 75% humidity and 16/8 hrs photoperiod. The transgenic plantlets when compared to the wild type were slower in growth and their acclimatization frequency was also lower. They successfully reached the flowering stage but the number of flowers was few and the siliques were thin and carried

only a few seeds in them. The seeds were harvested from the siliques and were again grown on $\frac{1}{2}$ MS + Kan (50 mg/L) media; the plantlets were transferred to pots and let to set seeds (T_2). The seeds were collected and stored.

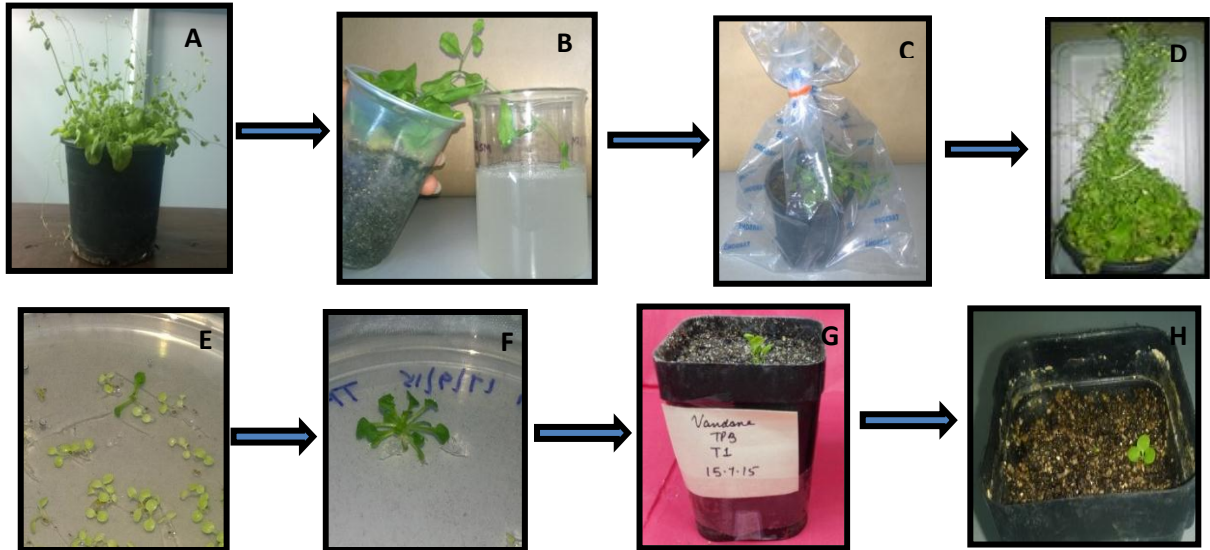


Figure 3-27: Transformation of *A. thaliana* by floral dip method. (A) *A. thaliana* plant (2-3 months old), (B) Dipping of *A. thaliana* flowers into *Agrobacterium* suspension harboring the recombinant TNaXTP/TMDTP-*sbtA-gus-pRI101-AN* plasmids (C) Plant covered in plastic bag, after floral dip (D) Plant showing silique formation after 3 weeks of floral dip (E and F): Screening of T_0 seeds on $\frac{1}{2}$ MS media containing kanamycin (50 mg/L) (G) Regeneration of T_1 plants and (H) Regeneration of T_2 plants.

When the constructs were transformed using floral dip method, T_1 and T_2 generation were obtained successfully. The plants regenerated but had less flowers, fewer and thin siliques, less number of seeds and gradually began to die off on seed set (Figures 3-28 and 3-29). Although no clear cause of the observation is known, the reason might be the integration of T-DNA into gene sequences responsible for silique development and seed formation. There have been reports where mutations at insertion sites in the plant genome resulted in misexpression of important endogenous plant traits, in part because transgenes inserted into or near functional gene sequences (Gheysen et al., 1987; Latham et al., 2006). Forsbach et al., (2003) and Chen et al., (2003) have stated that approximately 27%–63% of T-DNA insertions disrupt known gene sequences in *A. thaliana* and *Oryza sativa*. Integration of

multiple copies of the transgene has been reported to lead not only to the suppression of transgene expression, but also to the co-suppression of the homologous endogenous gene (Kohli et al., 2010)

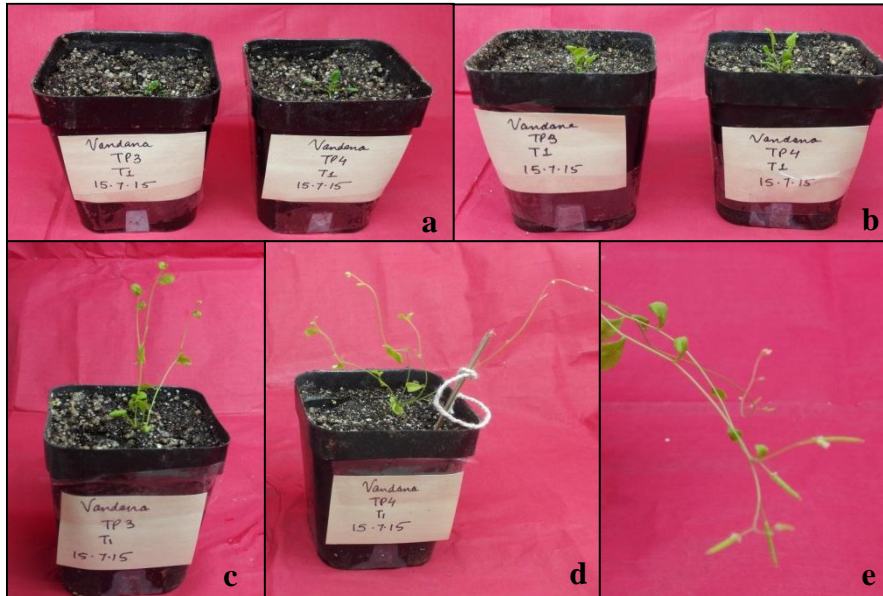


Figure 3-28: Transfer of antibiotic resistant plantlets (T_1 generation) on soilrite mix. Plants after 5 (a), 15 (b), 30 (c and d) and 45 (e) days of pot transfer. Silique formation (e) started after 45 days of pot transfer.



Figure 3-29: Transfer of antibiotic resistant plantlets (T_2 generation) on soilrite mix. Plants after 30 (a and b) and 45 (c) days of pot transfer

3.6.3. Optimization of the concentrations of plant growth regulators (PGRs) for *in-vitro* shoot and root induction in *N. tabacum*

Since the plants generated by floral dip method had retarded growth and less number of siliques and seeds, *Agrobacterium* mediated co-culture method of plant transformation was tried to transform and express the cyanobacterial gene in plant system. Before starting with actual agroinfection experiments, several combinations of phytohormones were tested on the non-transformed leaf discs of *N. tabacum* and the ones giving the best responses were finally used as a component of the regeneration medium.

(A) Effect of BAP concentrations on shoot regeneration in *N. tabacum*

Four different concentrations of BAP viz. 0.5, 1.0, 1.5 and 2.0 mg/L namely M₁, M₂, M₃, and M₄ respectively were used and supplemented with MS basal medium. The percentage of caulogenesis, number of shoots along with their length responses was observed at progressive stages of tissue culture. Maximum caulogenesis was recorded on M₄ (89.16±0.27) medium, which was found to be significantly superior to M₁ (66.31±3.81), M₂ (78.19±1.18) and M₃ (81.94±3.25) medium where, M₁, M₂ and M₃ were found to be at par with each other. The highest number of multiple shoots were obtained on M₄ (5.24±0.08) which was found to be significantly superior to M₁ (2.41±0.34), M₂ (4.25±0.25) and M₃ (4.25±0.34) medium. The maximum length of shoots were also obtained on M₄ (4.17±0.12) medium followed by M₃ (3.70±0.21), M₂ (3.48±0.10) and M₁ (2.73±0.07) medium.

Table 3-10: Caulogenesis percentage, number of shoots and length of shoots in *N. tabacum* leaf explant inoculated on MS medium supplemented with four different concentrations of BAP (0.5, 1.0, 1.5 and 2.0 mg/L)

Treatments (BAP)	Caulogenesis (% ± SE)	Number of shoots ± SE	Length of shoots (cm ± SE)
M ₁ (0.5 mg/L)	66.31±3.81 ^a	2.41±0.34 ^a	2.73±0.07 ^a
M ₂ (1.0 mg/L)	78.19±1.18 ^b	4.25±0.25 ^b	3.48±0.10 ^b
M ₃ (1.5 mg/L)	81.94±3.25 ^{ab}	4.25±0.34 ^b	3.70±0.21 ^b
M ₄ (2.0 mg/L)	89.16±0.27 ^c	5.24±0.08 ^c	4.17±0.12 ^c
C.D. at 5%	8.04	0.86	0.43
SE (m)	2.58	0.27	0.13

Value followed by the same letter in column is not significantly different using Duncan's new multiple range test at 5%.

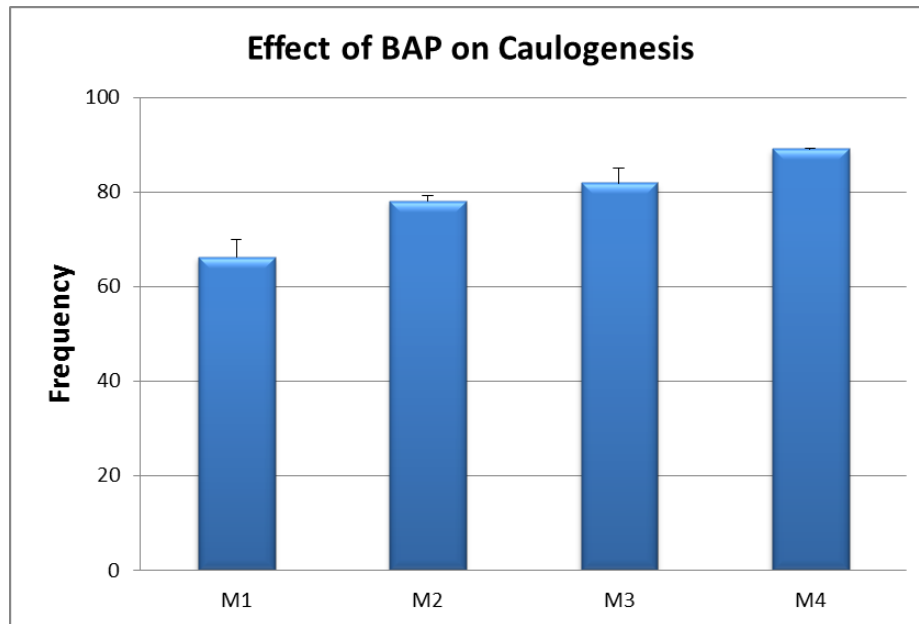


Figure 3-30: Effect of MS media supplemented with four different concentrations of BAP on frequency of caulogenesis in *N. tabacum*. The M₁, M₂, M₃ and M₄ media had 0.5, 1, 1.5 and 2 mg/L concentration of BAP respectively.

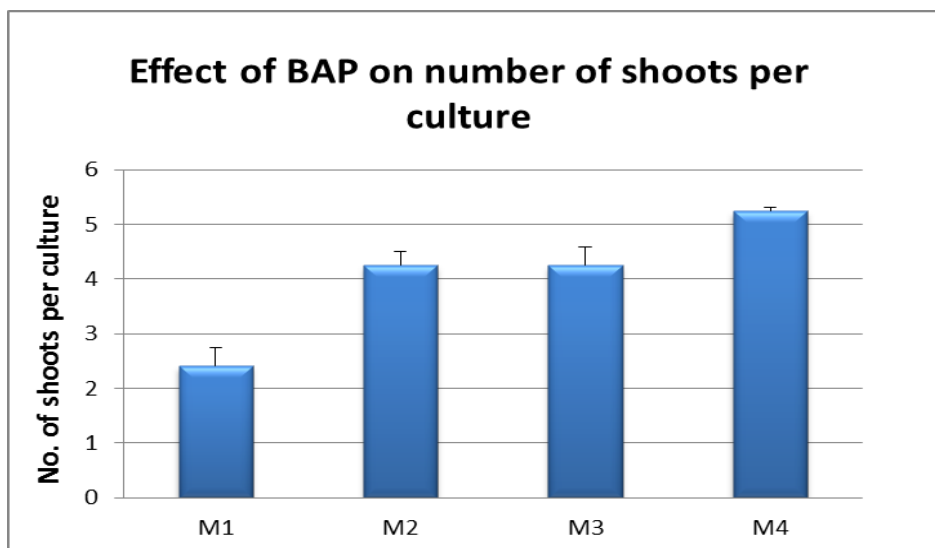


Figure 3-31: Effect of MS media supplemented with four different concentrations of BAP on number of shoots per culture in *N. tabacum*. The M₁, M₂, M₃ and M₄ media had 0.5, 1, 1.5 and 2 mg/L concentration of BAP respectively.

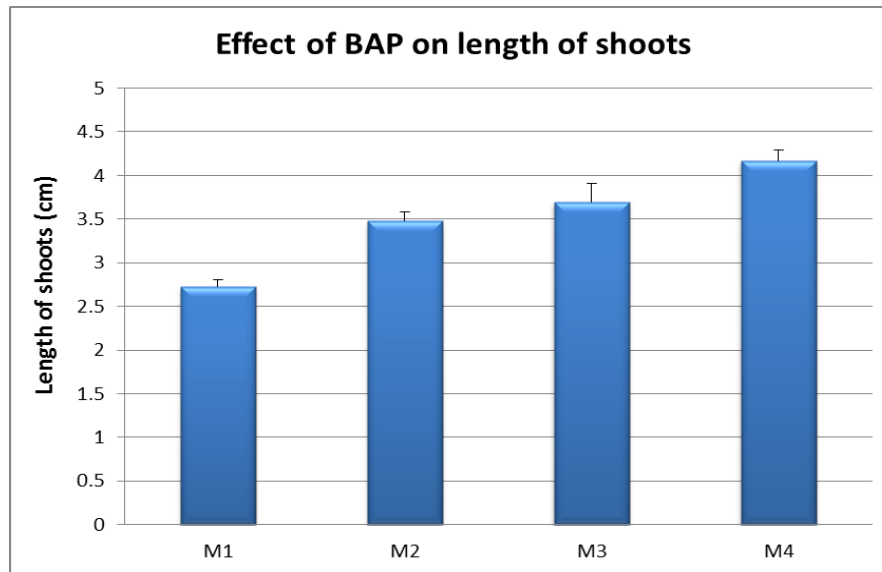


Figure 3-32: Effect of MS media supplemented with four different concentrations of BAP on length of shoots in *N. tabacum*. The M₁, M₂, M₃ and M₄ media had 0.5, 1, 1.5 and 2 mg/L concentration of BAP respectively.

(B) Effect of BAP and NAA interaction on shoot regeneration in *N. tabacum*

For the BAP constant (2 mg/L) and NAA varying (0.05, 0.1, 0.15 and 0.2 mg/L) media, the maximum caulogenesis was obtained on M₈ (89.44±0.32) medium, followed by M₇ (85.41±2.50), M₆ (78.47±3.47) and the lowest on M₅ (65.62±3.12) medium. The highest number of shoots were obtained on M₈ (14.08±0.15) medium followed by M₇ (10.83±0.55) and the lowest on M₅ (6.58±0.34) medium. The M₈ medium depicted the maximum length of shoots (4.82±0.10) followed by M₇ (4.43±0.02), M₆ (4.00±0.08) and the least was recorded on M₅ (3.70±0.04) medium.

Table 3-11: Caulogenesis percentage, number of shoots and length of shoots in *N. tabacum* leaf explant inoculated on MS medium supplemented with constant BAP (2.0 mg/L) and varying NAA (0.05, 0.1, 0.15 and 0.2 mg/L) concentrations

Treatments (BAP + NAA)	Caulogenesis (% \pm SE)	Number of Shoots \pm SE	Length of Shoots (cm \pm SE)
M ₅ [BAP(2 mg/L+NAA(0.05 mg/L)]	65.62 \pm 3.12 ^a	6.58 \pm 0.34 ^a	3.70 \pm 0.04 ^a
M ₆ [BAP(2 mg/L+NAA(0.1 mg/L)]	78.47 \pm 3.47 ^b	9.75 \pm 0.65 ^b	4.00 \pm 0.08 ^b
M ₇ [BAP(2 mg/L+NAA(0.15 mg/L)]	85.41 \pm 2.50 ^{bc}	10.83 \pm 0.55 ^b	4.43 \pm 0.02 ^c
M ₈ [BAP(2 mg/L+NAA(0.2 mg/L)]	89.44 \pm 0.32 ^c	14.08 \pm 0.15 ^c	4.82 \pm 0.10 ^d
C.D. at 5%	8.31	1.46	0.22
SE(m)	2.67	0.46	0.07

Value followed by the same letter in column is not significantly different using Duncan's new multiple range test at 5%.

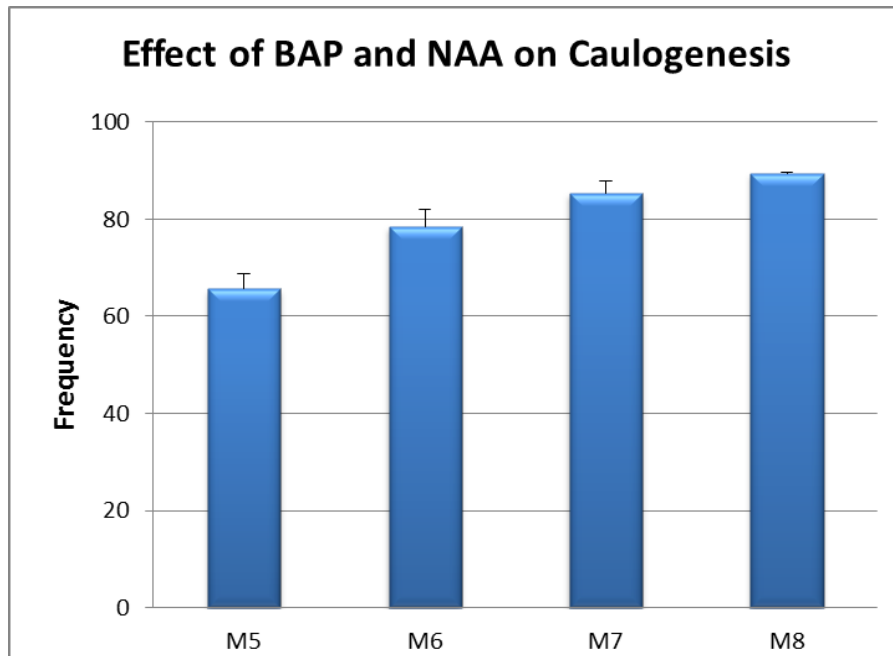


Figure 3-33: Effect of BAP and NAA interaction on frequency of caulogenesis in *N. tabacum*. The M₅, M₆, M₇ and M₈ media had constant BAP (2.0 mg/L) and varying NAA (0.05, 0.1, 0.15 and 0.2 mg/L) concentrations.

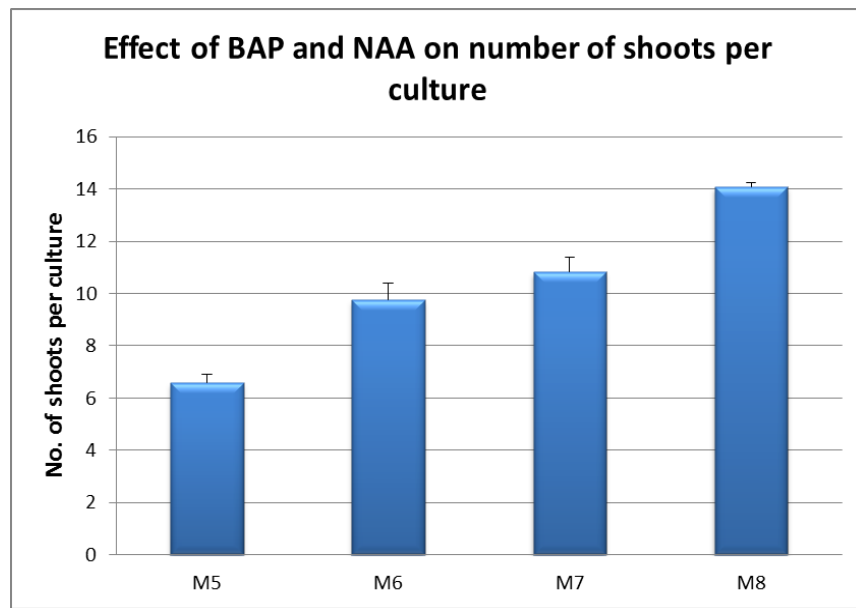


Figure 3-34: Effect of BAP and NAA interaction on number of shoots per culture in *N. tabacum*. The M₅, M₆, M₇ and M₈ media had constant BAP (2.0 mg/L) and varying NAA (0.05, 0.1, 0.15 and 0.2 mg/L) concentrations.

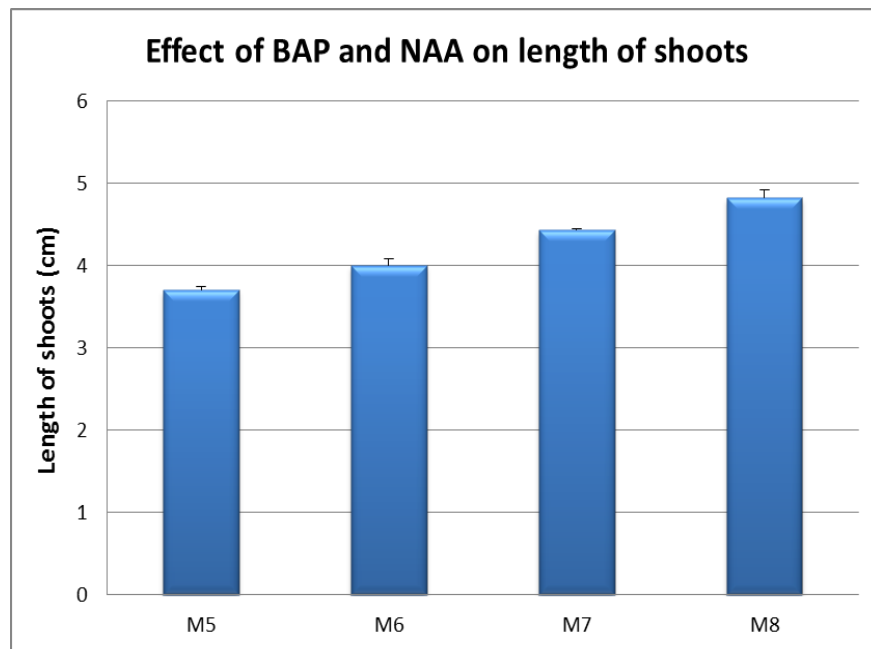


Figure 3-35: Effect of BAP and NAA interaction on length of shoots in *N. tabacum*. The M₅, M₆, M₇ and M₈ media had constant BAP (2.0 mg/L) and varying NAA (0.05, 0.1, 0.15 and 0.2 mg/L) concentrations.

(C) Effect of IBA on root induction in *N. tabacum*

Four different concentrations of IBA viz. 0.5, 1.0, 1.5 and 2.0 mg/L supplemented with MS basal medium were used for rooting of plantlets developed in multiple shoot media. The maximum rhizogenesis was observed on MS basal medium and on M₁₀ medium i.e. 97.50±2.50 which was found to be significantly superior to M₉ (85.00±2.88) and M₁₂ (87.50±2.50), but at par to M₁₁ (92.50±2.50) medium.

The maximum number of roots was obtained on MS basal (8.53±0.37) medium which was found to be significantly superior to other treatments of IBA. The least number of roots was recorded on M₉ (3.50±0.16) medium. Maximum length of roots was also recorded on MS basal medium (6.20±0.17) which was at par to M₁₀ (5.96±0.12) but significantly superior to all other treatments.

Table 3-12: Rhizogenesis percentage, number of roots per shoot and length of roots per shoot in *N. tabacum* leaf explant inoculated on MS medium supplemented with different concentrations of IBA (0.5, 1.0, 1.5 and 2.0 mg/L).

Treatments (IBA)	Rhizogenesis (% ± SE)	Number of Roots ± SE	Length of Roots ± cm
M ₉ (0.5 mg/L)	85.00±2.88 ^a	3.50±0.16 ^a	4.44±0.33 ^a
M ₁₀ (1.0 mg/L)	97.50±2.50 ^b	7.25±0.16 ^c	5.96±0.12 ^b
M ₁₁ (1.5 mg/L)	92.50±2.50 ^{ab}	4.58±0.34 ^b	4.96±0.17 ^a
M ₁₂ (2.0 mg/L)	87.50±2.50 ^a	3.58±0.43 ^a	4.74±0.15 ^a
MS Basal	97.50±2.50 ^b	8.53±0.37 ^d	6.20± 0.17 ^b
C.D. at 5%	2.58	0.31	0.20
SE (m)	7.85	0.96	0.62

Value followed by the same letter in column is not significantly different using Duncan's new multiple range test at 5%.

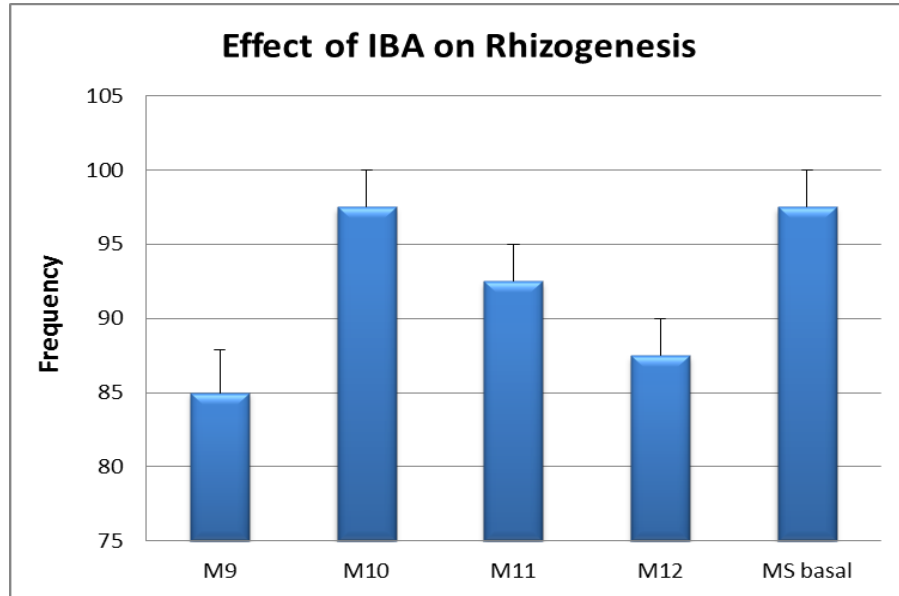


Figure 3-36: Effect of MS basal and MS media supplemented with four different concentrations of IBA on frequency of rhizogenesis in *N. tabacum*. The M₉, M₁₀, M₁₁ and M₁₂ media had 0.5, 1, 1.5 and 2 mg/L concentration of IBA respectively.

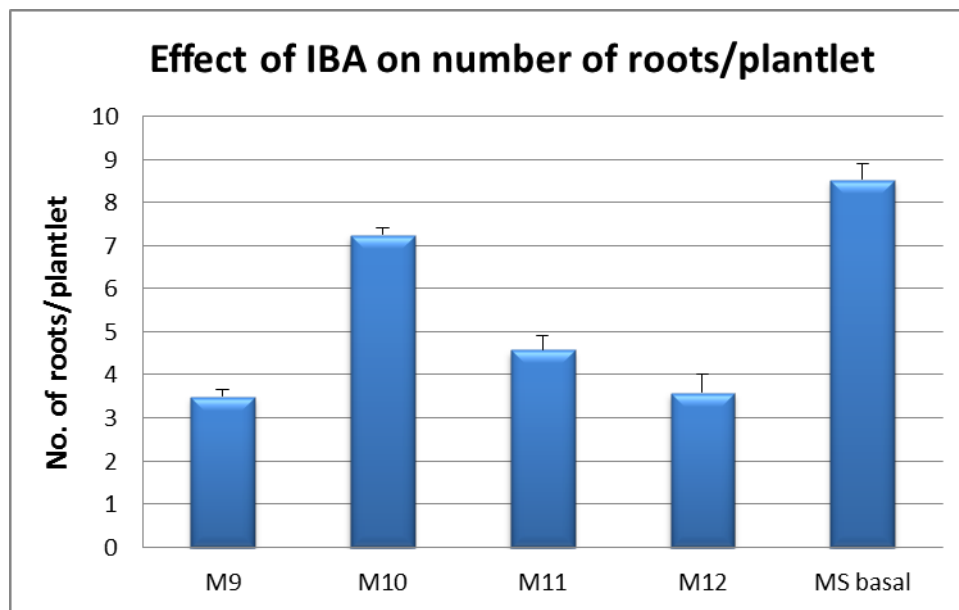


Figure 3-37: Effect of MS basal and MS media supplemented with four different concentrations of IBA on number of roots/plantlet in *N. tabacum*. The M₉, M₁₀, M₁₁ and M₁₂ media had 0.5, 1, 1.5 and 2 mg/L concentration of IBA respectively.

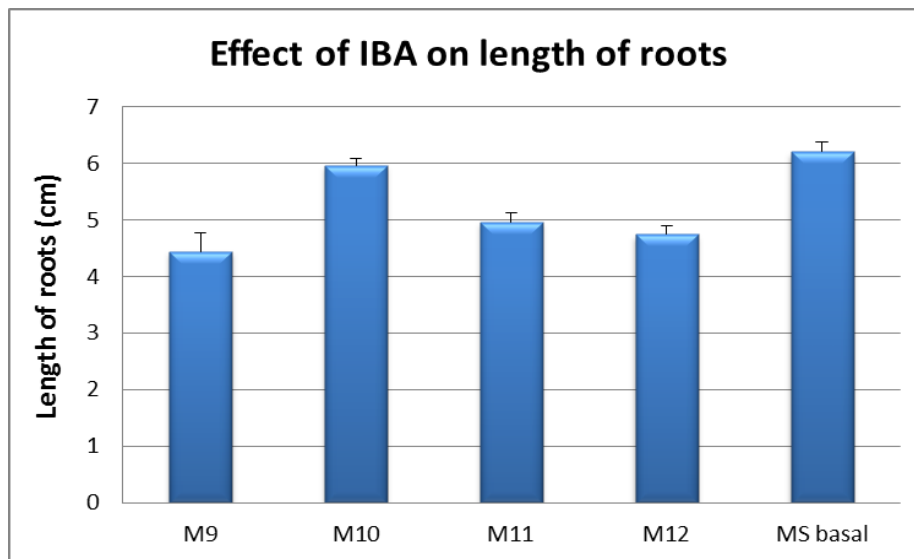


Figure 3-38: Effect of MS basal and MS media supplemented with four different concentrations of IBA on length of roots/plantlet in *N. tabacum*. The M₉, M₁₀, M₁₁ and M₁₂ media had 0.5, 1, 1.5 and 2 mg/L concentration of IBA respectively.

As per the tissue culture responses of the non-transformed leaf explants of *N. tabacum*, M₈ medium [BAP (2 mg/L + NAA (0.2 mg/L)] showed the best response for shoot induction while for root induction, MS basal outperformed the other media combinations. Thus, while performing transformation of *N. tabacum* leaf discs by *Agrobacterium* mediated co-culture method, these media combinations were used for shoot regeneration and root induction.

3.6.4. Transformation of *N. tabacum* using recombinant *pRII01-AN* constructs by *Agrobacterium* mediated co-culture method

In *Agrobacterium* mediated gene transfer experiment, leaf discs of *N. tabacum* were co-cultured with *Agrobacterium* harboring recombinant *pRII01-AN* (TNaXTP/TMDTP-*sbtA-gus-pRII01-AN*) constructs. The details have been provided in sec 2.2.22 in chapter 2. The transformed leaf discs were regenerated into shoots and also into callus. For, callogenesis, leaf explant was inoculated on MS media + 2,4-D (4 mg/L) + kanamycin (50 mg/L) + cefotaxime (250 mg/L) while for direct shoot regeneration, the leaf explant was cultured on shoot induction medium which contained MS+ BAP (2.0 mg/L) + NAA (0.2 mg/L) + kanamycin (50 mg/L) + cefotaxime (250 mg/L). The regenerated callus along with the

control callus (callus generated from non-transformed wild type plant leaves) was used to study the transient *gus* gene expression by GUS histochemical assay. Lopez et al., (2015) used the same method for induction of callus and for analysis of GUS reporter gene assay in tomato (*Solanum lycopersicum* L.) leaves transformed with *Agrobacterium tumefaciens*. The presence of blue colour in the transgenic callus tissue in comparison to the control callus clearly indicated the successful transformation of *sbtA* into *N. tabacum* (Figure 3-39). The stained callus tissues and the control callus were analyzed under a compound microscope and images were captured.

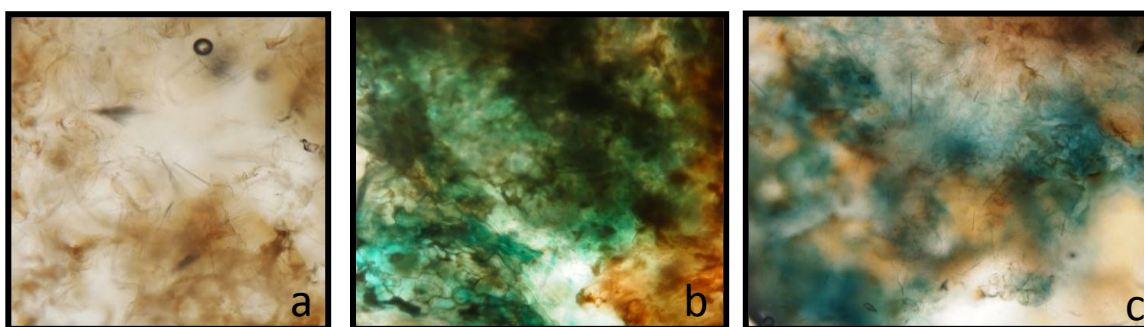


Figure 3-39: GUS histochemical staining of (a): Non-transformed callus (control), (b): Callus tissue regenerated from agro-infected *N. tabacum* leaves using recombinant TNaXTP-*sbtA-gus-pRI101-AN* construct and (c): Callus tissue regenerated from agro-infected *N. tabacum* leaves using recombinant TMDTP-*sbtA-gus-pRI101-AN* construct.

The multiple shoots that were regenerated on shoot induction (MS+ BAP (2.0 mg/L) + NAA (0.2 mg/L) + kanamycin (50 mg/L) + cefotaxime (250 mg/L) were maintained on the same medium for about 45 days (Figure 3-40). After 30–45 days of selection on the medium containing antibiotics, putative transgenic shoots were transferred onto shoot and root regeneration media without antibiotics. The genomic DNA was isolated from the regenerated shoots to check the presence of transgene. Out of the several plantlets screened by gene specific PCR, the ones that tested positive were further transferred on the rooting medium (½ MS medium without any phytohormone) (Figure 3-41). After development of roots the plantlets were acclimatized into plastic pots carrying soilrite mix.

The acclimatization frequency of the transgenic plantlets in the *in-vivo* conditions was quite low and they did not survive for long in the natural conditions. The *in-vitro*

cultured plants have been reported to possess non-functional stomata, weak root system and poorly developed cuticle and therefore show a high mortality rate on transfer to *ex-vitro* conditions (Mathur et al., 2008). Pospóšilová et al., (1999) observed decreased stomatal density and changes in the size and morphology of stomata on both sides of newly formed leaves on acclimatization of *N. tabacum* plantlets to *ex-vitro* conditions.

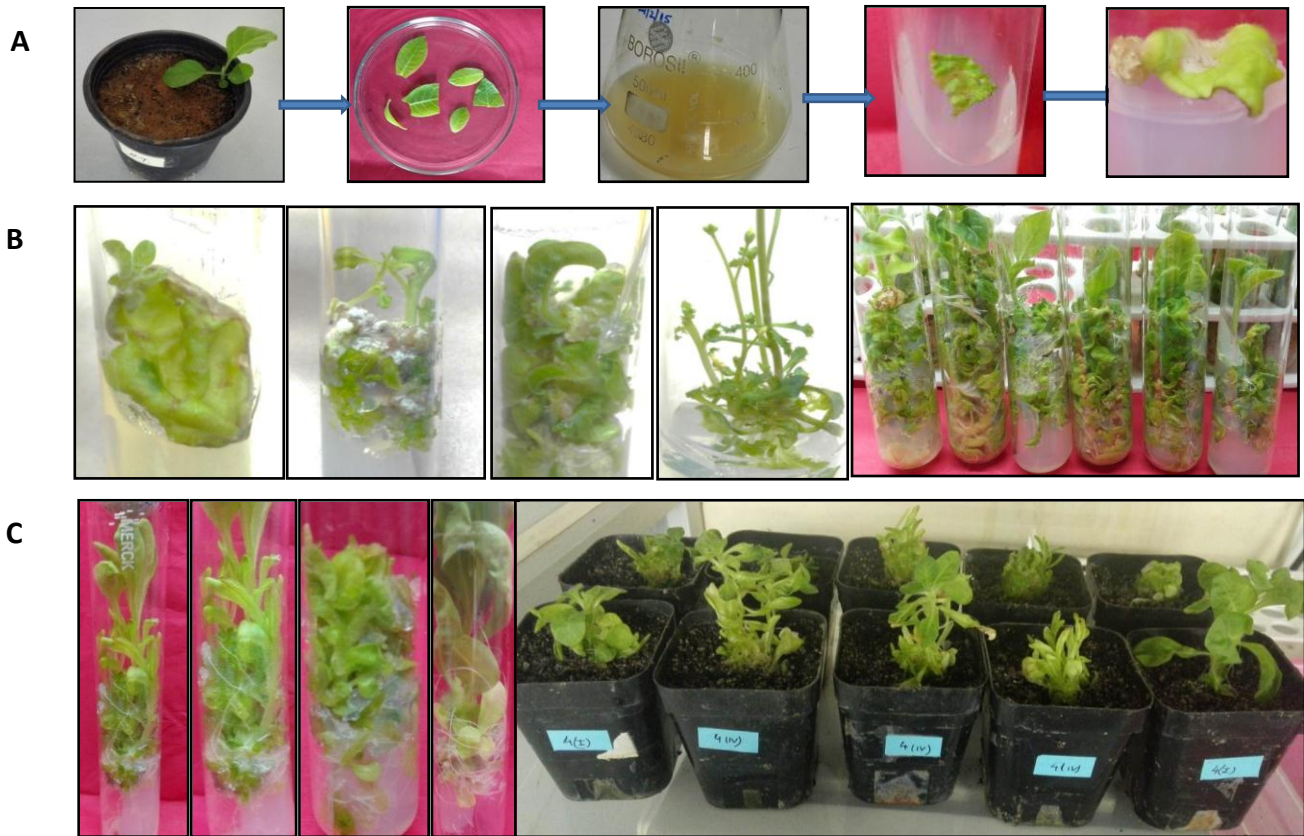


Figure 3-40: *Agrobacterium* mediated co-culture method for regeneration of stable *N. tabacum* plants. (A) Leaf discs of *N. tabacum* transformed with *Agrobacterium* cells harboring *pRI101-AN* constructs (TNaXTP/TMDTP-*sbtA-gus-pRI101-AN*). (B) Growth stages of shoots directly regenerated on MS media containing BAP (2.0 mg/L), NAA (0.2 mg/L), Kan (50 mg/L) and Cefotaxime (250 mg/L). (C) Induction of roots on rooting media ($\frac{1}{2}$ MS) followed by pot transfer in soilrite mix.

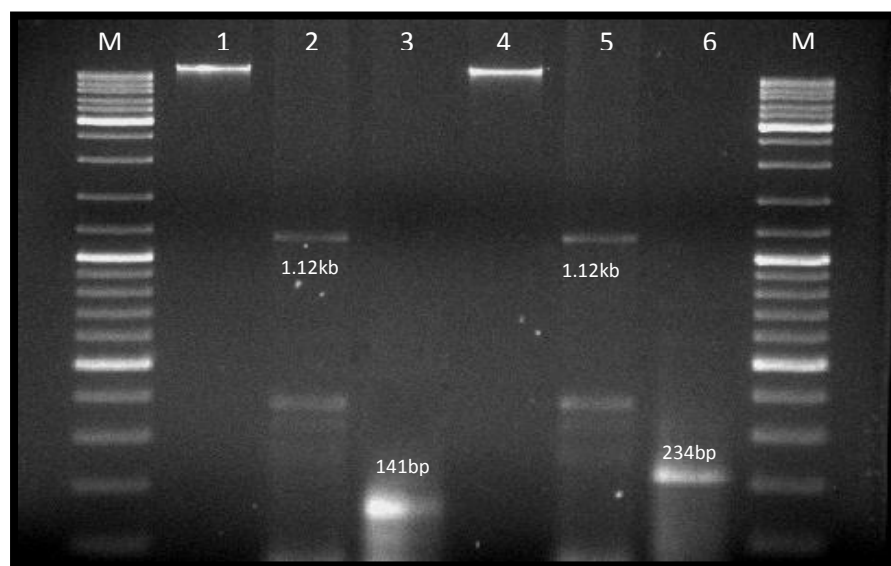


Figure 3-41: PCR analysis of the transformed plants regenerated by agroinfection of *N. tabacum* by recombinant *pRII01-AN* constructs (TNaXTP/TMDTP-*sbtA-gus-pRII01-AN*). The amplification was performed using gene specific primers for *sbtA* and TP sequences. Lane 1: gDNA of transgenic plant regenerated after transformation with TNaXTP-*sbtA-gus-pRII01-AN* construct. Lane 2: gDNA of transgenic plant with TNaXTP-*sbtA-gus* construct amplified by using *sbtA* specific primers, Lane 3: gDNA of transgenic plant with TNaXTP-*sbtA-gus* construct amplified by using TNaXTP specific primers, Lane 4: gDNA of transgenic plant regenerated after transformation with TMDTP-*sbtA-gus-pRII01-AN* construct, Lane 5: gDNA of transgenic plant with TMDTP-*sbtA-gus* construct amplified by using *sbtA* specific primers and Lane 6: gDNA of transgenic plant with TMDTP-*sbtA-gus* construct amplified by using TMDTP specific primers. M: Gene ruler SM0331 (Thermoscientific).

In the plant transformation studies conducted using recombinant *pRII01-AN* constructs, positive results were obtained in transformation of *N. tabacum* whereas an acute recalcitrance was observed in case of *A. thaliana*. Despite intensive research on plant transformation it is still poorly understood why some plant species can be transformed easily, while others are recalcitrant to *Agrobacterium* mediated transformation. One of the reasons attributed to such observation is the incompatibility of the host species with the transformation protocol which may trigger host defense responses leading to low transformation success (Hwang et al., 2015). Therefore, the success of transformation varies from species to species hence, some species are easier to transform in comparison to others. Additionally, the nature (hydrophobic) and size (large) of SbtA, makes its transformation and expression difficult.

After the confirmation of the gene integration by PCR analysis of the transformed *N. tabacum* plantlets (generated by *Agrobacterium* mediated co-culture method), the interest was to know whether the gene is being successfully transcribed and translated in the heterologous system. Since, it was not possible to study the exact localization of the protein by GUS constructs, the fused segments were cloned into a GFP containing vector i.e., *pCAMBIA1302*.

3.7. Cloning of fusion genes (TNaXTP/TMDTP-*sbtA*) into *pCAMBIA1302* vector

The fusion genes (TNaXTP/TMDTP-*sbtA*) were amplified using recombinant *pColdI* plasmid DNA as template with *BglIII* and *SpeI* restriction enzyme sites in forward and reverse primer respectively. The details of primer sequences and PCR conditions used have been mentioned in detail in sec 2.2.4 in chapter 2. The PCR products corresponding to a size of approximately 1.26 kb and 1.35 kb were obtained respectively for TNaXTP-*sbtA* and TMDTP-*sbtA* (Figure 3-42).

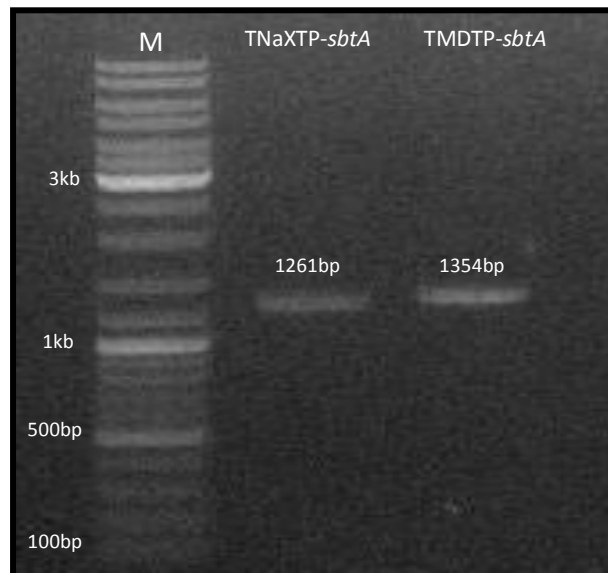


Figure 3-42: The agarose gel electrophoresis image of the amplification of the fusion genes (*sbtA*-TNaXTP/TMDTP) by using *pColdI* recombinant plasmids as template. Lane 1: M- Gene ruler SM0331 (Thermoscientific), Lane 2: PCR product of TNaXTP- *sbtA* (~1261 bp) and Lane 3: PCR product of TMDTP-*sbtA* (~1354 bp).

Figure 3-43: Recombinant vector map and restriction analysis for confirmation of the cloning of fusion constructs (TNaXTP/TMDTP-*sbtA*) into *pCAMBIA1302* vector. **(A)** M: Gene ruler SM0331 (Thermoscientific), Lane1: Recombinant uncut TNaXTP-*sbtA*-*pCAMBIA1302* plasmid, Lane 2: Recombinant TNaXTP-*sbtA*-*pCAMBIA1302* plasmid digested with *Bgl*III, Lane 3: Recombinant TNaXTP-*sbtA*-*pCAMBIA1302* plasmid digested with *Kpn*I (~938 bp). **(B)** M: Gene ruler SM0331 (Thermoscientific), Lane 1: Recombinant uncut TMDTP-*sbtA*-*pCAMBIA1302* plasmid, Lane 2: Recombinant TMDTP-*sbtA*-*pCAMBIA1302* plasmid digested with *Bgl*III, Lane 3: Recombinant TMDTP-*sbtA*-*pCAMBIA1302* plasmid digested with *Kpn*I (~1031bp)

3.7.1. Transformation of recombinant *pCAMBIA1302* constructs into *N. benthamiana* by *Agrobacterium* mediated agroinfiltration method

N. benthamiana is lately the most widely used experimental host in plant biology for production of recombinant proteins at high levels as it has a short life cycle, carries relatively large, easily infiltrable leaves and the leaf does not show necrosis upon infiltration with most *Agrobacterium* strains (D'Aoust et al., 2008). Four to five weeks old *N. benthamiana* plants were chosen for agroinfiltration. At this stage the plant has at least five fully developed true leaves and no visible flower buds. The filling of apoplastic spaces by the *Agrobacterium* solution is also clearly visible in form of dark green patches on the abaxial surface of the leaf (Figure 3-45).

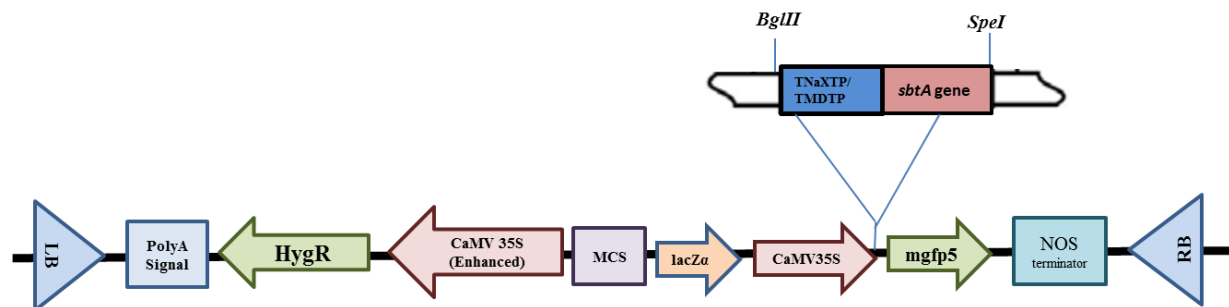


Figure 3-44: Schematic representation of the recombinant *pCAMBIA1302* vector with TNaXTP/TMDTP-*sbtA* insert engineered for plant transformation



Figure 3-45: Agrobacterium infiltration of *N. benthamiana* leaves with *Agrobacterium* cells harboring recombinant constructs (TNaXTP/TMDTP-*sbtA-pCAMBIA1302*). (A) 2.5 to 3 months old *N. benthamiana* plant. (B) Abaxial (lower) surface of the leaf before infiltration. (C) Infiltration of *Agrobacterium* suspension inside the air spaces of the mesophyll cells of the leaf using needleless syringe. (D) Leaf after infiltration, the infiltrated areas can be clearly seen as wet patches

Both the recombinant *pCAMBIA1302* constructs (TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302*) along with the empty *pCAMBIA1302* vector were transformed into *N. benthamiana* by agrobacterium infiltration method (Figure 3-45). The detailed procedure followed for agrobacterium infiltration has been mentioned in sec 2.2.24 in chapter 2. After infiltration, the plants were stored away from the direct light till the plants showed maximum transgene expression (2-4 days).

Although there are reports that the accumulation of the recombinant protein by agrobacterium infiltration can be achieved two days post infiltration, no fluorescence was observed in protoplasts harvested after 2 days of infiltration. The optimal accumulation of protein was observed 4–8 days post infiltration (dpi). The transformed leaves were used for extracting genomic DNA, total RNA, total protein and protoplast samples four days after infiltration. For all the experiments, samples were collected from the leaves agrobacterium infiltrated with (a) TNaXTP-*sbtA-pCAMBIA1302* (b) TMDTP-*sbtA-pCAMBIA1302* (c) Empty *pCAMBIA1302* vector and (d) Negative control (non-transformed wild type leaves).

3.7.2. Analysis of transgene expression at DNA level in the agroinfiltrated leaves of *N. benthamiana*

Genomic DNA was isolated from the transformed leaves using DNeasy Plant Mini Kit (sec 2.2.3 in chapter 2) for all the four samples four days after infiltration (Figure 3-46).

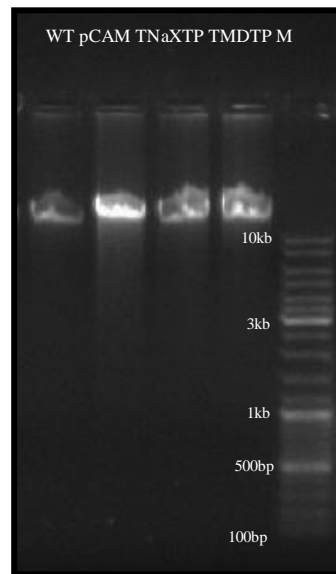


Figure 3-46: Genomic DNA isolated from wild type (WT) and agroinfiltrated samples {pCAM (*pCAMBIA 1302* empty vector control), TNaXTP (TNaXTP-*sbtA-pCAMBIA1302*) and TMDTP (TMDTP-*sbtA-pCAMBIA1302*)}, M: Gene ruler SM0331 (Thermoscientific)

The presence of transgene was detected at DNA level, by performing diagnostic PCRs by using TNaXTP/TMDTP, *sbtA* gene and *mgfp5* gene specific primers. The detailed protocol for PCR conditions used as well as the sequences of primers have been mentioned in detail in sec 2.2.4 in chapter 2. The reaction conditions used for amplification have been provided in sec 2.2.27 in chapter 2, Table 2-14.

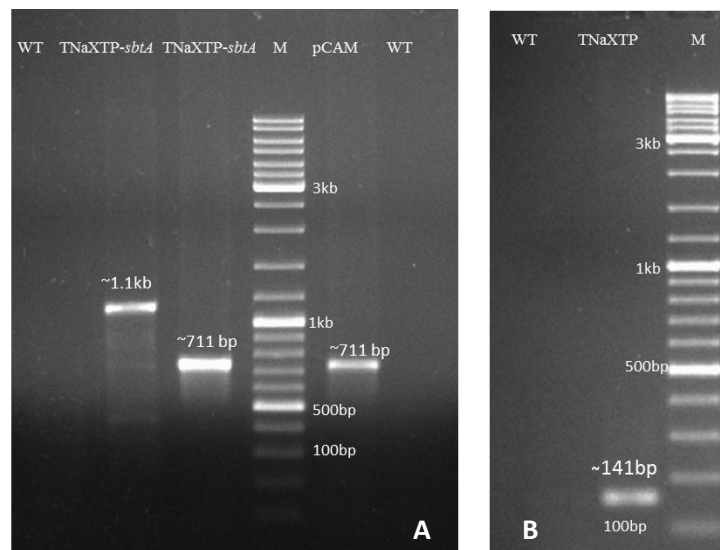


Figure 3-47: The agarose gel images representing PCR products amplified from genomic DNA (A and B) of agroinfiltrated leaves of *N. benthamiana*. **(A)** Amplification using *sbtA* specific and *mgfp5* specific primers. Lane 1: wild type control (PCR using *sbtA* specific primers), Lane 2: Amplification of *sbtA* from TNaXTP-*sbtA*-*pCAMBIA1302* recombinant gDNA, Lane 3: Amplification of *mgfp5* from TNaXTP-*sbtA*-*pCAMBIA1302* recombinant gDNA, Lane 4: M- Gene Ruler SM0331 (Thermoscientific), Lane 5: Amplification of *mgfp5* from *pCAMBIA1302* recombinant gDNA and Lane 6: wild type control (PCR using *mgfp5* gene specific primers). **(B)** Amplification using TNaXTP specific primers. Lane1: wild type control (PCR using TNaXTP specific primers), Lane 2: Amplification of TNaXTP from TNaXTP-*sbtA*-*pCAMBIA1302* recombinant gDNA and Lane 3: M- Gene Ruler SM0331 (Thermoscientific)

The gene specific PCR resulted into full length amplicons of expected sizes for TNaXTP (141 bp), *sbtA* (1.12 kb) and *mgfp5* (711 bp), thereby confirming the successful integration of TNaXTP, *sbtA* and *mgfp5* gene into the genome of *N. benthamiana* (Figure 3-47).

Similarly, the sequence specific PCR analysis of genomic DNA isolated from *N. benthamiana* leaves infiltrated with TMDTP-*sbtA*-*pCAMBIA1302* construct gave full length amplicons of expected sizes for TMDTP (234 bp), *sbtA* (1.12 kb) and *mgfp5* (711 bp). These results validate successful integration of TMDTP, *sbtA* and *mgfp5* gene into *N. benthamiana* genome (Figure 3-48).

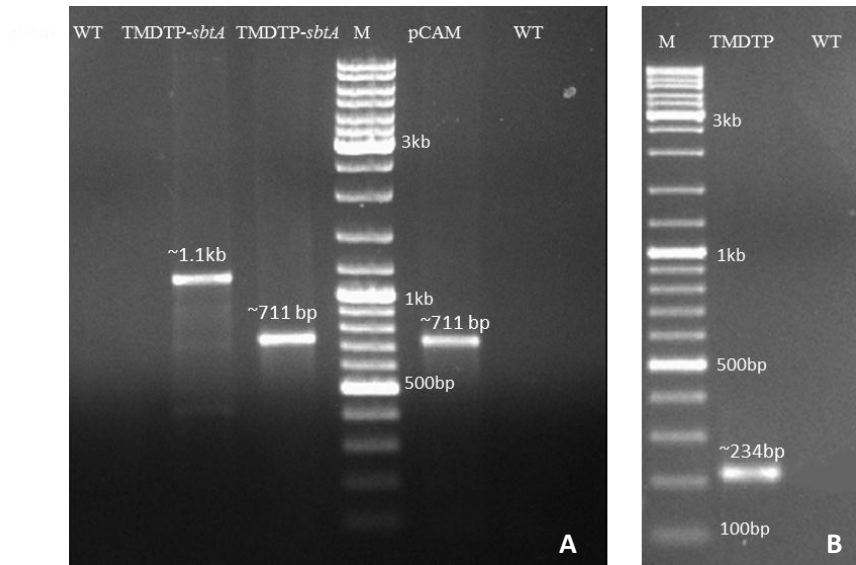


Figure 3-48: The agarose gel images representing PCR products amplified from genomic DNA (A and B) of agroinfiltrated leaves of *N. benthamiana*. (A) Amplification using *sbtA* specific and *mgfp5* specific primers. Lane 1: wild type control (PCR using *sbtA* specific primers), Lane 2: Amplification of *sbtA* from TMDTP-*sbtA*-pCMBIA1302 recombinant gDNA, Lane 3: Amplification of *mgfp5* from TMDTP-*sbtA*-pCMBIA1302 recombinant gDNA, Lane 4: M- Gene Ruler SM0331 (Thermoscientific), Lane 5: Amplification of *mgfp5* from pCMBIA1302 recombinant gDNA and Lane 6: wild type control (PCR using *mgfp5* gene specific primers). (B) Amplification using TMDTP specific primers. Lane1: M- Gene Ruler SM0331 (Thermoscientific), Lane 2: Amplification of TMDTP from TMDTP-*sbtA*-pCMBIA1302 recombinant gDNA, Lane 3: wild type control (PCR using TMDTP specific primers)

3.7.3. Analysis of transgene expression at mRNA level in the agroinfiltrated leaves of *N. benthamiana*

Total RNA was extracted from all the four samples using RNeasy Plant Mini Kit (sec 2.2.25 in chapter 2, Figure 3-49). The isolated RNA was used for cDNA synthesis which was further analyzed for the transgene expression.

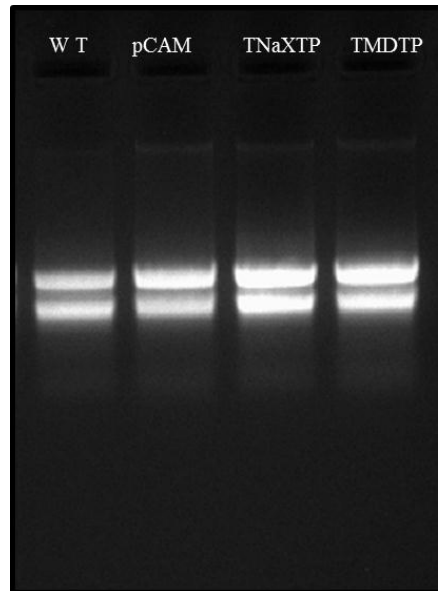


Figure 3-49: Total RNA isolated from wild type (WT) and agroinfiltrated samples {pCAM (*pCAMBIA1302* empty vector control), TNaXTP (*TNaXTP-sbtA-pCAMBIA1302*) and TMDTP (*TMDTP-sbtA-pCAMBIA1302*)}. RNA was isolated from agroinfiltrated leaves four days after infiltration.

Any genomic DNA contamination in RNA samples was removed by using gDNA wipeout buffer (a component of QuantiTect® Reverse Transcription kit). A negative control PCR was also performed using gDNA wiped out RNA samples to nullify the possibility of genomic DNA contamination in RNA samples by using gene specific primers. The absence of bands in the samples ascertains that gDNA wipeout was successful in eliminating genomic DNA contamination present, if any (Figure 3-50).

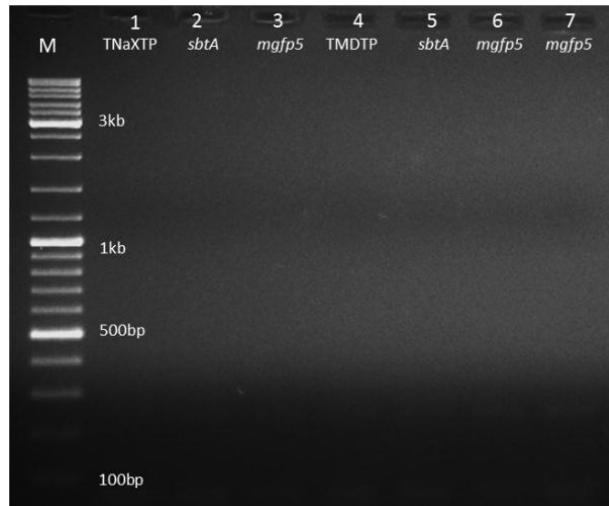


Figure 3-50: Agarose gel electrophoresis image showing absence of PCR products in RNA isolated from *N. benthamiana* leaves infiltrated with TNaXTP-*sbtA*-*pCAMBIA1302* (Lanes 1-3) and TMDTP-*sbtA*-*pCAMBIA1302* (Lanes 4-6) using TNaXTP, TMDTP, *sbtA* and *mgfp5* primers and to check the possibilities of DNA contamination in the RNA samples. Last lane (Lane 7) is amplification of *mgfp5* from (*pCAMBIA1302*) empty vector RNA.

The RNA samples were used to synthesize cDNA using the QuantiTect® Reverse Transcription kit. The detailed protocol has been mentioned in sec 2.2.26 in chapter 2. The cDNA samples were analyzed for the presence of mRNA transcript of TP sequences (TNaXTP/TMDTP), *sbtA* and *mgfp5*. For TP sequences, full length primers were used (TP sequences being smaller in size) while for *sbtA* and *mgfp5*, gene specific primers for the internal region of the sequence were used.

A mixture of oligo(dT) and random primers was used for performing RT-PCRs. Since Oligo(dT) primers are most suitable for full length cDNA synthesis only, adding random hexamers to them can improve cDNA yield by increasing binding at multiple sites on the same template. This can help in detection of even the low expressing genes. The reaction conditions used for performing diagnostic PCRs have been described in sec 2.2.27 in chapter 2.

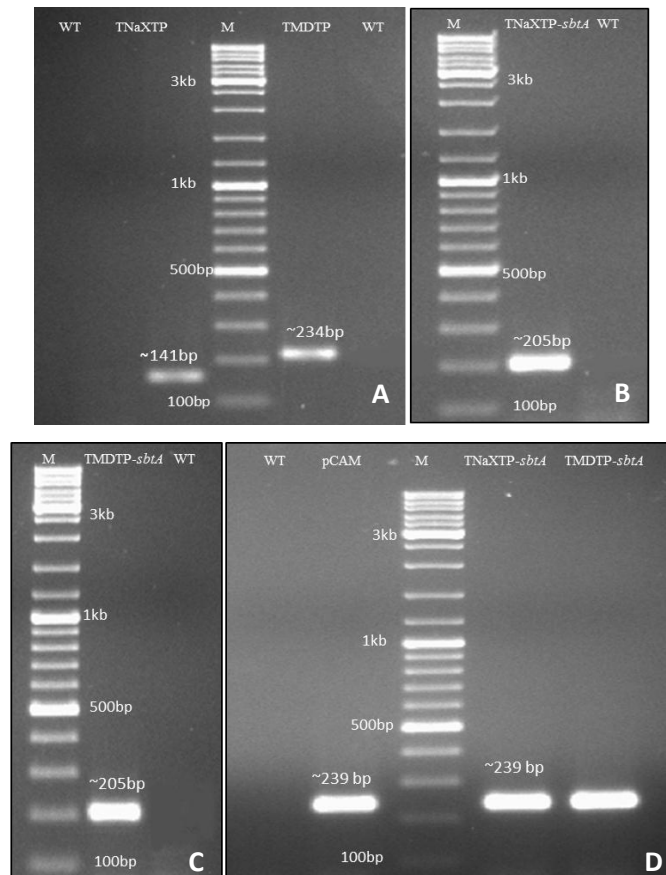


Figure 3-51: The agarose gel images representing PCR products amplified from cDNA (A, B, C and D) of agroinfiltrated leaves of *N. benthamiana*. (A) **Amplification using TNaXTP and TMDTP specific primers.** Lane 1: wild type control (PCR using TNaXTP specific primers), Lane 2: Amplification of TNaXTP from recombinant TNaXTP-*sbtA-pCAMBIA1302* cDNA, Lane 3: M-Gene ruler, SM0331 (Thermoscientific), Lane 4: Amplification of TMDTP from recombinant TMDTP-*sbtA-pCAMBIA1302* cDNA and Lane 5: wild type control (PCR using TMDTP specific primers). (B) and (C) **Amplification of partial *sbtA* gene.** (B) Lane 1: M- Gene ruler, SM0331 (Thermoscientific), Lane 2: Amplification of partial *sbtA* gene from TNaXTP-*sbtA-pCAMBIA1302* recombinant cDNA and Lane 3: wild type control. (C) Lane 1: M-Gene ruler, SM0331 (Thermoscientific), Lane 2: Amplification of partial *sbtA* gene from TMDTP-*sbtA-pCAMBIA1302* recombinant cDNA and Lane 3: wild type control. (D) **Amplification of partial *mgfp5* gene.** Lane 1: wild type control, Lane 2: Amplification of partial *mgfp5* gene from empty vector control (pCAM) recombinant cDNA, Lane 3: M-Gene ruler, SM0331 (Thermoscientific), Lane 4: Amplification of partial *mgfp5* gene from TNaXTP-*sbtA-pCAMBIA1302* recombinant cDNA and Lane 5: Amplification of partial *mgfp5* gene from TMDTP-*sbtA-pCAMBIA1302* recombinant cDNA

The results obtained in diagnostic PCRs indicate the expression of TP sequences (TNaXTP/TMDTP), *sbtA* gene and *mgfp5* reporter gene at DNA and mRNA level in the transiently transformed leaves of *N. benthamiana*.

3.7.4. Analysis of transgene expression at protein level in the agroinfiltrated leaves of *N. benthamiana*

The total protein was extracted from the agroinfiltrated leaves using Total Protein Extraction (TPE™) kit (Figure 3-52). The detailed protocol has been mentioned in sec 2.2.28 in chapter 2. About 10 µL of protein samples (approximately 50-100 µg) of the extracted proteins were run on a 12% SDS-PAGE gel.

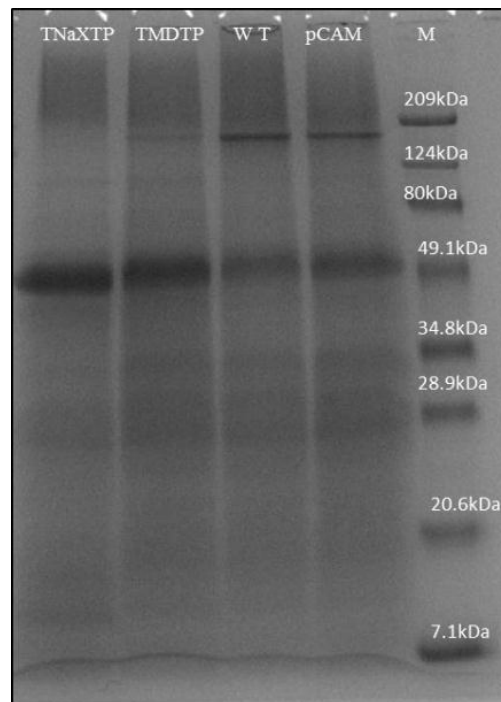


Figure 3-52: Total protein isolated from wild type and agroinfiltrated leaf samples of *N. benthamiana*. Samples were run on a 12% SDS-PAGE. Lane 1: Protein isolated from leaves agroinfiltrated with TNaXTP-*sbtA*-*pCAMBIA1302* construct, Lane 2: Protein isolated from leaves agroinfiltrated with TMDTP-*sbtA*-*pCAMBIA1302* construct, Lane 3: Protein isolated from leaves of wild type, Lane 4: Protein isolated from leaves agroinfiltrated with pCAM (*pCAMBIA1302*) empty vector control and Lane 5: Prestained SDS-PAGE Standards, broad range (Biorad).

The band of fusion protein viz., SbtA-Mgfp5 (~65.7 kDa) was not visible on the gel because of the low expression levels of the protein. Therefore, in order to check whether the desired GFP tagged proteins were synthesized or not, western blotting was performed using antibody against the GFP tag. As an internal control for chloroplast membrane proteins, Tic110 was chosen as it is the translocon located at the inner envelope of chloroplasts. All the four protein samples were checked for the presence of Tic110 and mGFP5 protein using antibodies against Tic110 and Mgfp5.

The Tic110 protein (internal control) had constitutive expression in the wild type (~100 - 110 kDa) and in all the three agroinfiltrated leaf samples of *N. benthamiana*. The presence of Tic110 protein in the total protein extract confirms that the isolate contains almost all the transmembrane proteins of the plant cell including the ones located on the inner envelope which would include the SbtA (a transmembrane transporter). In case of empty vector control (pCAM), a band at ~26 kDa of mGFP5 protein was observed. In leaves transformed with TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302* constructs, bands of ~ 65-70 kDa [39.7 kDa (SbtA) + 26 kDa (Mgfp5)] were observed in both cases. These bands corresponded to GFP fused transporter protein (SbtA and Mgfp5) for both the constructs. This band was absent in WT (Figure 3-53). The TP sequence length is not taken into account here because these should be cleaved off by plant peptidases after protein targeting into the chloroplast.

The results of western blotting confirms the presence of near full length SbtA protein in the protein isolate which ascertains that the SbtA-GFP fusion constructs have successfully integrated and expressed in *N. benthamiana*.

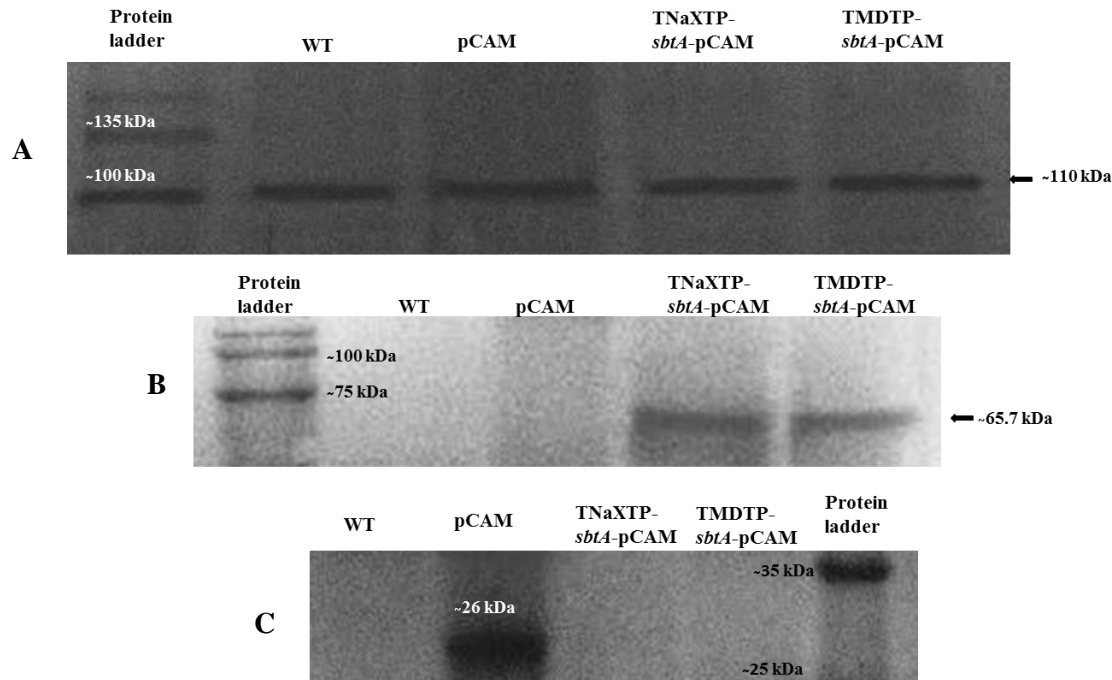


Figure 3-53: Western blot analysis of total protein isolated from WT (wild type leaves which were not agroinfiltrated) and agroinfiltrated *N. benthamiana* leaves with pCAM (*pCAMBIA1302* empty vector control), TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302* constructs immunoblotted with antibodies against Tic110 (A) and Mgfp5 (B and C). (A) Expression of Tic110 (band at ~110 kDa) in all the protein samples on immunoblotting with anti-Tic110 antibody. (B) Expression of the fusion protein (Mgfp5+SbtA, band at ~65.7 kDa) in both TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302* constructs and (C) Expression of Mgfp5 in pCAM (band at ~26 kDa) on immunoblotting with antibody against Mgfp5. Protein ladder used in images is BLUeye prestained protein ladder (BR Biochem).

3.7.5. Analysis of the localization of the targeted transporter protein by confocal microscopy

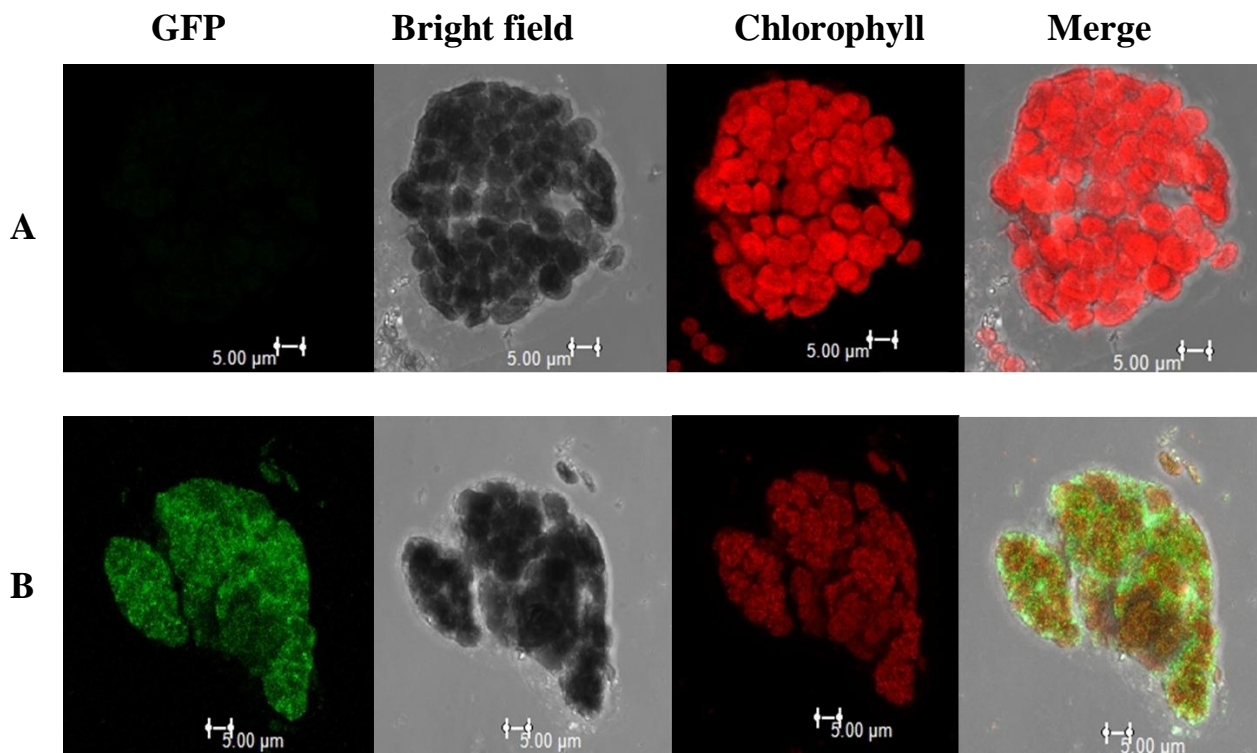
For studying the sub-cellular localization of the transporter-reporter fusion product (SbtA-Mgfp5), protoplasts were isolated from all the four samples 4 days post infiltration. The procedure of protoplast isolation has been mentioned in sec 2.2.31 in chapter 2. All the four samples viz., (a) Protoplasts prepared from wild type plant leaves (negative control), (b) Protoplasts prepared from leaves infiltrated with TNaXTP-*sbtA-pCAMBIA1302* construct, (c) Protoplasts prepared from leaves infiltrated with TMDTP-*sbtA-pCAMBIA1302* construct

and (d) Protoplasts prepared from leaves infiltrated with empty *pCAMBIA1302* vector were analyzed by confocal laser scanning microscopy.

The subcellular fluorescent localization patterns in transformed protoplasts were compared to wild type and constructs designed for expression in the cytosol (pCAM) as control (Figure 3-54). The autofluorescence of chlorophyll served as chloroplast marker. In the wild type sample only chlorophyll autofluorescence was observed whereas the empty vector (*pCAMBIA1302*) was located in the cytosol because of the absence of any targeting sequence and was indicated by non-overlapping GFP fluorescence. In TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302* constructs, the merged images showed overlap between chlorophyll autofluorescence and GFP fluorescence which assures their chloroplastic locations.

Although both constructs showed chloroplastic expression, the exact location in the organelle (chloroplast sub-compartment) remains unsettled. In case of TNaXTP-*sbtA-pCAMBIA1302* construct, fluorescent signal seems to be associated almost exclusively with the envelope of the chloroplasts although it is quite difficult to discriminate clearly between inner and outer envelope membranes. It has been speculated that TMDs (Transmembrane domains) determine the exact location of the protein (Froehlich and Keegstra, 2011). Since our protein is an integral component of plasma membrane in cyanobacteria, its chloroplastic location signifies integration into either of the two membranes, viz., the inner envelope or the thylakoid membrane, as both are present in continuum in the chloroplast (Rosado-Alberio et al., 1968). Recently, Rolland et al., (2016) and Uehara et al., (2016) were able to target SbtA completely to chloroplast inner envelope by using a longer N-terminal sequence consisting of cleavable chloroplast transit peptide (cTP) and a membrane protein leader (MPL) of about 92–115 amino acids. Though the addition of MPL along with the cTP may help in proper integration of the protein, there are high chances that it could also inhibit the activity of the native bicarbonate transporter (Uehara et al., 2016). If the latter happens to be the case; the sequences additional to cTP have to be removed from the transporter by using a protease.

Recently, Singhal and Fernandez (2017) have proposed the possible mechanism of targeting of multi-transmembrane domain proteins to specific chloroplast compartments. They predicted the localization of two nuclear encoded proteins, SCY1 and SCY2 which form the component of chloroplast SEC translocase in *A. thaliana*. Both these proteins in spite of having conserved structure with 10 transmembrane domains were targeted to different compartments in chloroplast. SCY1 was targeted to thylakoid membrane and SCY2 to inner envelope membrane. Rolland et al., (2017) postulated that targeting of non-chloroplastic proteins to chloroplast inner-envelope membrane could be done either by addition of extra TMDs of inner envelope membrane proteins like those of SCY2 (TMD3/TMD4 individually or both together) as proposed by Singhal and Fernandez (2017) or by engineering the TMDs of the cargo protein so that there is no thylakoid targeting sequence in the protein which would decrease the chances of mistargeting of the protein.



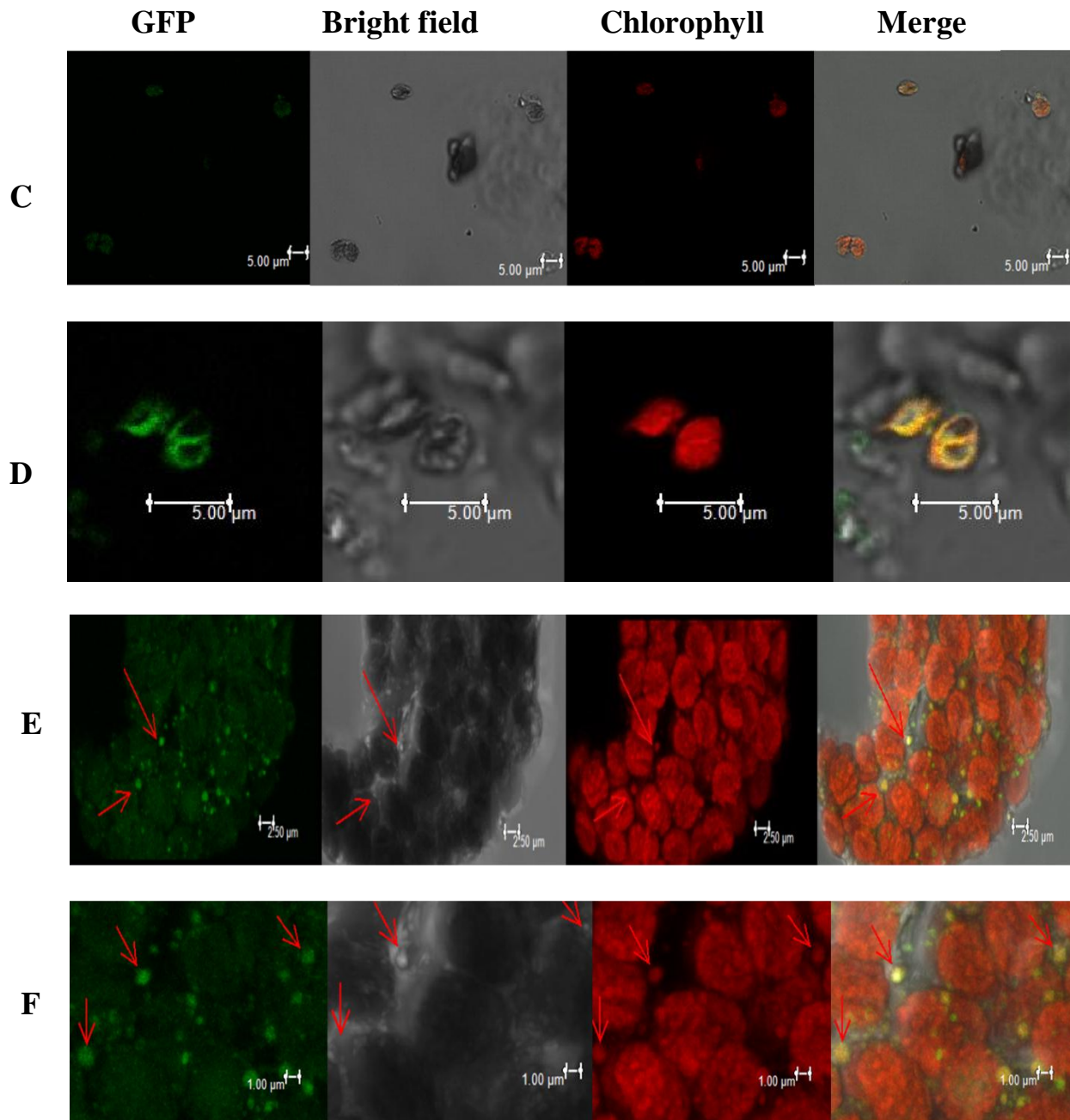


Figure 3-54: Confocal images of *N. benthamiana* protoplasts (A) Wild type, protoplasts isolated from a non-agroinfiltrated leaf, (B) Protoplast from a leaf agroinfiltrated with *pCambia1302* empty vector carrying *mgfp5* reporter gene, (C) and (D) Protoplasts agroinfiltrated with fusion construct TNaXTP-*sbtA-pCambia1302* showing transient expression of GFP, (E) and (F) Protoplasts agroinfiltrated with fusion construct TMDTP-*sbtA-pCambia1302* showing transient expression of GFP. Chlorophyll autofluorescence is shown in red. Arrows denote the localization of GFP in chloroplasts. Scale bars are specified at the lower panel of each image.

3.8. Transformation of *N. tabacum* using recombinant *pCAMBIA1302* constructs by *Agrobacterium* mediated co-culture method

After confirmation of the localization of the fusion protein (SbtA-Mgfp5) in the chloroplasts of *N. benthamiana*, the constructs were used to regenerate stable plants by *Agrobacterium* mediated co-culture method. The *Agrobacterium* cells harboring the recombinant *pCAMBIA1302* constructs (TNaXTP-*sbtA-pCAMBIA1302* and TMDTP-*sbtA-pCAMBIA1302*) were co-cultured with the leaf explants of *N. tabacum* and transferred to shoot induction medium containing MS + BAP (2.0 mg/L) + NAA (0.2 mg/L) + hygromycin (40 mg/L) + cefotaxime (250 mg/L) (Figure 3-58). Different concentrations of hygromycin were tested for efficient recovery of transgenic plants (Figures 3-55 and 3-56). 40 mg/L hygromycin was found to be the optimum concentration for shoot regeneration. In this study, hygromycin at 30 mg/L could already significantly reduce the regeneration frequency of regenerated shoots compared to the condition when no hygromycin was added in the medium. Although about 2–3% regeneration could occur when the hygromycin concentration reached 40 mg/L, the regenerants could only survive for a very short time, while no regeneration occurred when hygromycin reached 50 mg/L. Since 40 mg/L hygromycin proved to be lethal to the wild type *N. tabacum* (prevented complete shoot regeneration of non-transgenic explants) therefore this concentration was used in the subsequent transformation experiments. (Figures 3-55 and 3-56).

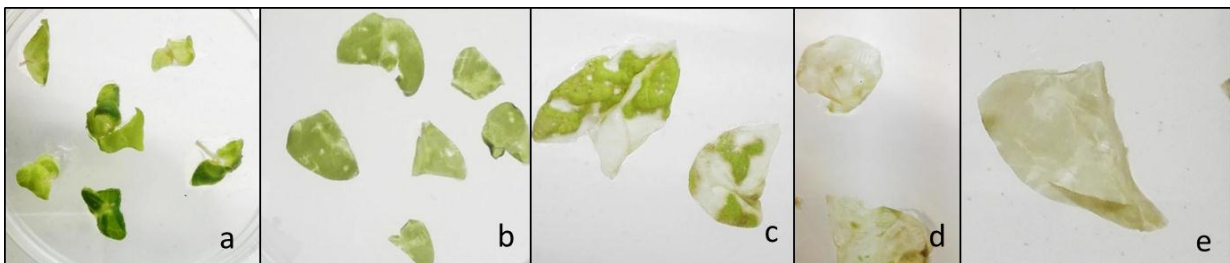


Figure 3-55: Non-transformed leaf discs of *N. tabacum* cultured on shoot induction medium containing (a) No hygromycin, (b) 20 mg/L hygromycin, (c) 30 mg/L hygromycin, (d) 40 mg/L hygromycin and (e) 50 mg/L hygromycin

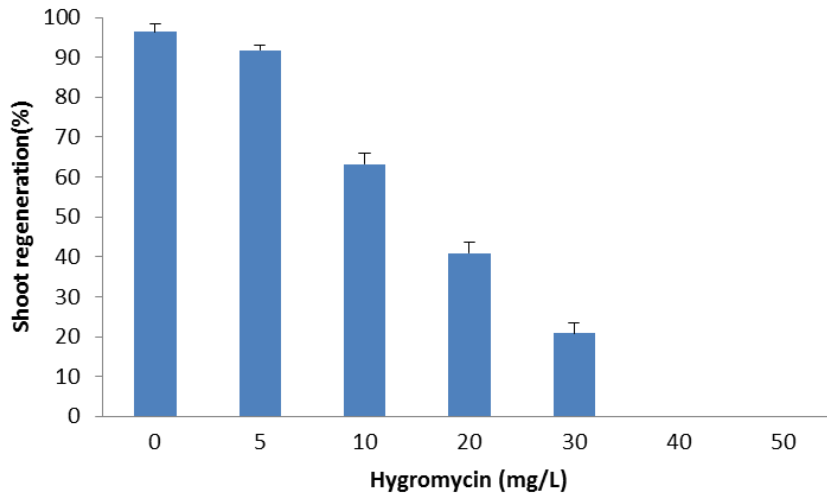


Figure 3-56: Effect of hygromycin concentration on regeneration frequency of non-transformed DEC explants grown on shoot elongation medium. Data present percent frequency in four replicates \pm SD.

The transformed leaf discs successfully proliferated into multiple shoots on the shoot induction medium (MS + BAP (2 mg/L) + NAA (0.2 mg/L) + cefotaxime (250 mg/L) and hygromycin (40 mg/L) (Figure 3-58) and after every 2 weeks, the regenerated shoots were sub-cultured onto fresh MS medium with the same composition. After 30-45 days of selection on the hygromycin medium, putative transgenic shoots were transferred onto shoot and root regeneration media without hygromycin and cefotaxime.

Meanwhile, the GFP fluorescence in transgenic shoots was monitored by microscopy analysis using the Zeiss AXIO SCOPE A1 microscope (Figure 3-57).

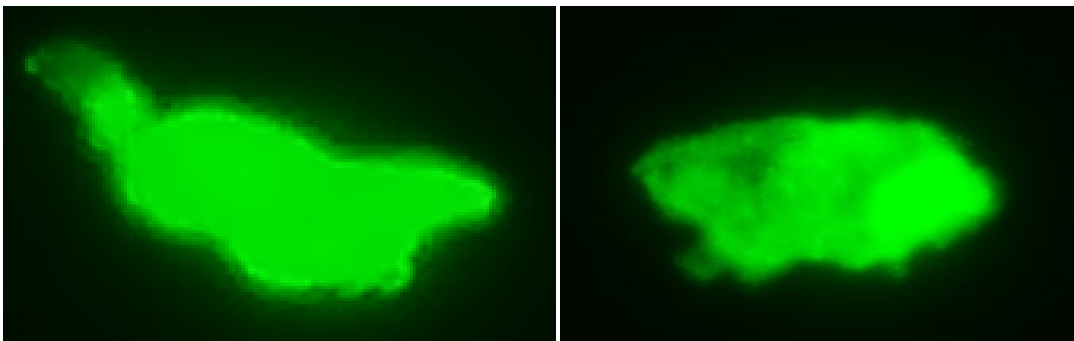


Figure 3-57: GFP fluorescence in transgenic shoots generated using *TNaXTP-sbtA-pCAMBIA1302* constructs monitored using Zeiss AXIO SCOPE A1 Microscope



Figure 3-58: Growth stages of shoots (a-d) directly regenerated from leaf explant on MS media containing BAP (2 mg/L) + NAA (0.2 mg/L) + Hygromycin (40 mg/L) + Cefotaxime (250 mg/L). Shoots transferred to antibiotic free shooting and rooting medium (e-h). Rooted plantlets growing in potting mixture (I, j and k).

The genomic DNA was also isolated from the shoots and analyzed for the presence of transgene by gene specific PCR (Figure 3-59). The plantlets were further transferred to the rooting media (MS Basal) when enough number of shoots was regenerated.

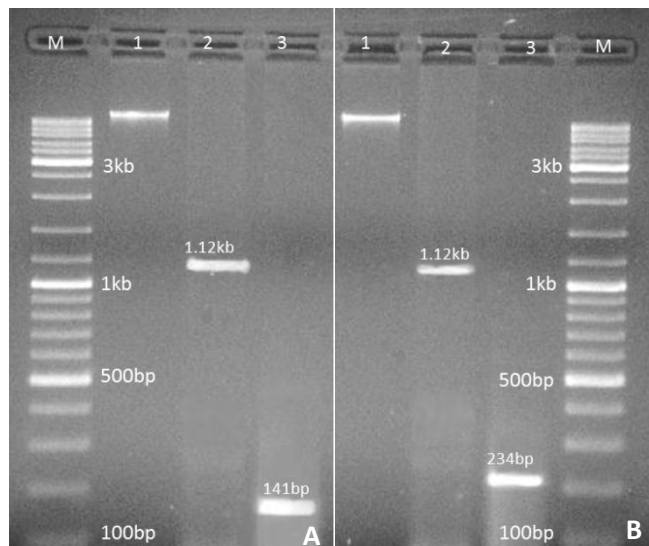


Figure 3-59: PCR analysis of the transformed plants regenerated by agro-infection of *N. tabacum* by recombinant *pCAMBIA1302* constructs (TNaXTP/TMDTP-*sbtA-pCAMBIA1302*). The amplification was performed using gene specific primers for *sbtA* and TP sequences. **(A)** Lane 1: gDNA of transgenic plant regenerated after transformation with TNaXTP-*sbtA-pCAMBIA1302* construct. Lane 2: gDNA of transgenic plant with TNaXTP-*sbtA-pCAMBIA1302* construct amplified by using *sbtA* specific primers, Lane 3: gDNA of transgenic plant with TNaXTP-*sbtA-pCAMBIA1302* construct amplified by using TNaXTP specific primers. **(B)** Lane 1: gDNA of transgenic plant regenerated after transformation with TMDTP-*sbtA-pCAMBIA1302* construct, Lane 2: gDNA of transgenic plant with TMDTP-*sbtA-pCAMBIA1302* construct amplified by using *sbtA* specific primers and Lane 3: gDNA of transgenic plant with TMDTP-*sbtA-pCAMBIA1302* construct amplified by using TMDTP specific primers. M: Gene ruler SM0331 (Thermoscientific).

The plants with well-developed roots were shifted into plastic pots containing a mixture of soilrite, clay soil and sandy soil in the ratios of 1:1:3 for hardening/acclimatization (Figure 3-58). Since the tissue cultured plantlets are grown under low intensity of artificial light and high humidity they are delicate and are not adequately “hardened off” which may make them readily lose water when exposed to ambient

conditions. It is also important to remove the traces of the agar adhered on the roots by gentle washing before transferring them to potting mixture because the agar contains sucrose and other nutrients that can serve as a medium for growth of disease-causing organisms (Perez and Hooks, 2008). The pots were covered with polythene bags to maintain humidity for initial 15-20 days and were gradually acclimatized by making small incisions on the side of the bags to facilitate aeration. The size of the cuts were slowly increased at an interval of 10 days and once the plantlets had properly regained their photosynthetic efficiency and transpiration mechanism they were transferred to large sized pots. The plants were maintained at 25-28°C temperature in a plant tissue culture room under 16/8 hrs of photoperiod. The plants were watered on regular basis until seed set and the seeds were collected and stored.

The transgenic seeds could be further raised along with the wild type seeds and analyzed for improvements in photosynthetic performance/growth rate/s. If successful, the experiments could also be replicated in C₃ crop plants like wheat and rice.

CHAPTER 4

CONCLUSIONS AND FUTURE PROSPECTS

The conclusions drawn from the present work can be summarized as mentioned below:

4.1. Distribution and arrangement of inorganic carbon transporters in cyanobacteria

- The array of transporters present in cyanobacteria varies on the basis of Ci availability and the habitat in which they dwell. The cyanobacteria living in the marine environments (α -cyanobacteria) possess less varied complement of bicarbonate transporters in comparison to the ones living in fresh water i.e., β -cyanobacteria.
- BCT1, a high affinity bicarbonate transporter was found to be absent in marine forms as the marine environment already possesses a substantially high content of bicarbonate which discourages the need of a transporter with high energy investment (BCT1 requires ATP energization) and encourages employment of less energy requiring sodium symporters like BicA and SbtA. The BCT1 transporter was present in thirteen species (belonging to β -cyanobacteria) out of the forty species studied.
- *G. violaceus*, an early diverging cyanobacterium, lacked SbtA and BicA transporters, while possessed the NdhI₃ and NdhI₄ (CO₂ uptake systems) complexes. This suggests that the CO₂ uptake systems were present during early years of evolution.
- The low CO₂ affinity transporter viz., NdhI₄ was present in both late and early diverging cyanobacterial species inhabiting a wide regime of temperature and Ci fluctuations. The high affinity CO₂ uptake system (NdhI₃) was absent in marine strains while some marine strains might have acquired these systems by horizontal gene transfer or by function uptake/modification of certain existing transporters (NAD(P)H dehydrogenase systems).

4.2. Analysis of *sbtA* gene and transit peptide elements

- The distribution of *sbtA* and *sbtB* genes was analyzed in several organisms belonging to the phyla eubacteria. No homologs of the transporter were detected in early evolving cyanobacteria like *G. violaceus*, *Thermosynechococcus elongatus* bp-1 and *Trichodesmium erythraeum*.

- Given the fact that horizontal gene transfer (HGT) has attributed a lot to the evolution of cyanobacteria, SbtA transporter also seems to have been transferred to far-related species by HGT along with SbtB which is reflected from the high correlation values obtained from Mirrortree server.
- For precise targeting of SbtA transporter to inner chloroplast envelope of the model C₃ plants, two inner chloroplast envelope located transporter proteins were shortlisted as the source of transit peptide sequences. The transit peptide (TP) from AT3G56160, a bile acid: sodium symporter was named as TNaXTP while the TP from AT2G02590, a multidrug transporter was named as TMDTP.
- A codon usage analysis was performed to ensure functional expression of SbtA in heterologous hosts (*A. thaliana* and *N. tabacum*). The analysis of relative adaptiveness of codon usage in *sbtA* gene depicted a mean difference of approximately 24.34% in cyanobacteria and 33 to 35% in the plants used in the study which implies that approximately 65% codons of *sbtA* gene find good preference in the plant genome. The percentage of low frequency codons was also found to be less than 30% (3% and 8% in *A. thaliana* and *N. tabacum* respectively).

4.3. Synthesis of fusion constructs and their transformation to C₃ plants followed by reporter gene assays

- The transit peptide sequences were amplified from *A. thaliana* genome and fused upstream of *sbtA* gene in *pColdI* vector to generate fusion constructs. The recombinant fragment (TNaXTP/TMDTP-*sbtA*) was further subcloned in plant expression vector viz., *pRII01-AN* containing *gus* reporter gene.

4.3.1. Plant transformation using *pRII01-AN* recombinant constructs

- Several plant transformation experiments were conducted to study transient and stable expression of the transporter gene. The transformation of *N. tabacum* by particle bombardment depicted a significantly higher expression level of GUS protein when analyzed by GUS fluorometric assay in both the constructs in comparison to the control sample which confirmed that the TNaXTP/TMDTP-*sbtA* fusion genes were successfully transformed into the plant system.

- Attempts were made to regenerate stable plants of *A. thaliana* by floral dip method but the transgenic plants had poor growth with less number of siliques and seeds and therefore could not survive for long.
- The transformation of *N. tabacum* leaf discs by the recombinant constructs was performed by *Agrobacterium* mediated co-culture method and the transient expression of GUS was checked in callus tissue by GUS histochemical assay. The presence of blue colour only in the transformed tissue commends the successful transformation of *N. tabacum* by the fusion genes.
- The same method (*Agrobacterium* mediated co-culture method) was also used for regeneration of stable plants of *N. tabacum*. The multiple shoots obtained on regeneration media after agroinfection with *Agrobacterium* harboring the recombinant constructs gave positive results on gene specific PCR analysis which assures the integration of fusion genes in the genome of *N. tabacum*. The plantlets which tested positive were transferred to pots after rooting but they were not capable of acclimatization and hence gradually died.

4.3.2. Plant transformation using *pCAMBIA1302* recombinant constructs

- The *pCAMBIA1302* vector has *mgfp5* reporter gene and hence the fusion genes were cloned into it for studying the sub-cellular localization patterns of the protein. The cassette with TNaXTP/TMDTP-*sbtA* fusion genes in *pCAMBIA1302* vector was transformed into *N. benthamiana* leaves by agroinfiltration method. The presence of the transgene was confirmed by gene specific PCR analysis and by RT-PCR. The expression at the protein level was confirmed by western blotting using anti-GFP primary antibody and horseradish peroxidase (HRP) conjugated secondary antibody and a band of ~65.7 kDa corresponding to SbtA-Mgfp5 fusion protein was obtained.
- The localization of the recombinant (SbtA-Mgfp5) protein in the protoplasts isolated from agroinfiltrated leaves of *N. benthamiana* was detected by confocal microscopy analysis. The expression for both the fusion constructs in the chloroplasts suggested that the transit peptide sequences (viz., TNaXTP and TMDTP) were capable of taking the cargo protein to the chloroplasts.

- The constructs were further used to generate stable transgenic plants by *Agrobacterium* mediated co-culture method. The plantlets stably expressing the gene were transferred to potting mixture and grown till seed set. The transgenic seeds could be used to raise transgenic plants which can be compared to the wild type for enhanced growth rate/s or biomass yield.

Future Prospects

While in the present study we have been successful in generating stable C₃ plants harboring the cyanobacterial transporter gene, a lot of scope for extending the work of this thesis still remains. A few directions in which this work may find future scope are:

- The interaction of SbtA-SbtB complexes could be further investigated to explore the underlying mechanism of activation/inactivation and post-translational regulation of cyanobacterial transporters.
- The seeds collected from the transgenic plants can be further raised and the T₁ generation could be analyzed for enhanced growth and photosynthetic efficiency.
- If successful in model C₃ plants, the study can also be replicated in agronomically important cereal crops like wheat and rice.
- The future prospects could also include the installation of other transporters from cyanobacteria along with SbtA or even the transfer of entire CCM assembly (Carboxysomes and the transporters) into the plant system, although the establishment of so many multi-subunit components and their activation is a challenging task.

SUMMARY

Most of our C₃ crop plants like rice, wheat and barley could not achieve their entire photosynthetic potential because of the inefficiencies of key enzyme, Ribulose 1, 5 biphosphate carboxylase/oxygenase or Rubisco which is involved in the very first step of CO₂ fixation. Rubisco is a bifunctional enzyme that also binds to molecular oxygen resulting in an investment of metabolic resources of the plant through a process called photorespiration. The problems of this enzyme have been alleviated in some lower organisms like cyanobacteria and algae that have developed certain strategies known as carbon concentrating mechanisms. Among the CCMs present in nature, the cyanobacterial CCM is considered to be the most efficient single celled CCM as it is capable of accumulating thousand folds of C_i near the enzyme Rubisco. A cyanobacterial CCM comprises of (i) a compartment called carboxysome which contains the enzymes like Rubisco and Carbonic anhydrases and (ii) Inorganic carbon transporters. According to modelling studies, the transporters specifically BicA and SbtA are the only components of the cyanobacterial CCM which can increase the photosynthetic efficiency independently of the other components. In the present work, attempts were made to integrate the cyanobacterial SbtA transporter to the chloroplasts of C₃ model plants. These studies can serve as a stepping stone for the ongoing research on enhancing crop productivity.

Firstly, the distribution and phylogeny of C_i transporters in cyanobacteria was studied with respect to *Gloeobacter violaceus* PCC 7421 (an early diverging cyanobacterial strain). *G. violaceus*, lacked SbtA and BicA transporters, while possessed the NdhI₃ and NdhI₄ (CO₂ uptake systems) complexes suggesting that the CO₂ uptake systems were present during early years of evolution. The complement of transporters present in cyanobacteria was found to differ on the basis of C_i availability and the habitat in which they lived. The cyanobacterial strains living in the marine environments (α -cyanobacteria) possessed less varied complement of bicarbonate transporters in comparison to the ones living in fresh water i.e, β -cyanobacteria.

Further, the distribution of *sbtA* and *sbtB* genes was studied in seventy eubacterial strains. The analysis revealed that SbtA is more of a recently acquired bicarbonate transporter as no homologs of the transporter were detected in *G. violaceus*, *Thermosynechococcus elongatus* bp-1 and *Trichodesmium erythraeum*. The co-evolution

studies of *sbtA* and *sbtB* disclosed that both the genes co-evolved together and the far-related organisms acquired this gene (*sbtAB*) cluster together during a single horizontal gene transfer event.

Furthermore, an extensive *in-silico* analysis of *A. thaliana* envelope proteins was done in order to select most suitable protein candidates that can be used as the source of transit peptides for targeting the SbtA transporter of cyanobacteria into chloroplast inner envelope of model C₃ plants. The information of all inner chloroplast envelope membrane proteins was retrieved from AT_CHLORO database and the short listing of most relevant protein candidates was done by using various softwares. Finally, two transit peptide (TPs) were finalized, the TP from AT3G56160, a bile acid: sodium symporter was named as TNaXTP while the TP from AT2G02590, a multidrug transporter was named as TMDTP. A codon usage analysis was performed to ensure functional expression of SbtA in plant system. The analysis of relative adaptiveness of codon usage in *sbtA* gene depicted that approximately 65% codons of *sbtA* gene found good preference in the plant genome.

The TNaXTP and TMDTP transit peptides were successfully amplified by PCR and then fused/cloned upstream of the *sbtA* gene in *pColdI* vector. The fusion genes were further sub-cloned into *gus* reporter gene containing plant expression vector viz., *pRII0-AN*. The recombinant *pRII01-AN* constructs (TNaXTP/TMDTP-*sbtA-gus-pRII01-AN*) were transformed into *N. tabacum* by particle bombardment which was followed by transient expression analysis by GUS fluorometric assay. The assay showed a significantly higher expression level of GUS protein in both the constructs in comparison to the control sample which confirmed that the fusion genes were transformed successfully into the plant system. Higher values of specific GUS activity were obtained in TNaXTP construct in comparison to TMDTP construct. The reason for the same might be attributed to the cTP used, as the TNaXTP (cTP from AT3G56160, BASS4) is a targeting sequence for a sodium bile acid symporter located at the chloroplast inner envelope membrane and the structure of SbtA resembles the structure of 3ZUX (in PDB), which is a bacterial homologue of the bile acid sodium symporter ASBT of *Neisseria meningitides*. Thus, it may be hypothesized that the similarity in structure and function of the cargo protein attached to the cTP facilitates the easy targeting of the protein.

Thereafter, efforts were made to regenerate stable plants by transforming *N. tabacum* leaf discs with *Agrobacterium* cells harboring recombinant *pRI101-AN* constructs using the *Agrobacterium* mediated co-culture method. The gene specific PCR analysis confirmed the integration of the gene in the regenerated shoots.

To study the sub-cellular localization of the fusion protein, the fusion genes were cloned into *mgfp5* reporter gene containing plant expression vector *pCAMBIA1302*. The recombinant *pCAMBIA1302* constructs were transformed into *N. benthamiana* by *Agrobacterium* mediated agroinfiltration method. The integration and expression of transgenes were confirmed at DNA and mRNA level by gene specific PCR analysis. The protein expression was confirmed by western blotting using anti-GFP antibodies which revealed the presence of full length GFP fusions in total protein extracts. The confocal microscopy analysis showed chloroplastic expression of transporter gene (*sbtA*) for both the constructs in transiently transformed protoplasts of *N. benthamiana*. In one of the constructs the fluorescence appeared to be associated with the membranes of the chloroplast. The constructs were further used for stable plant regeneration by *Agrobacterium* mediated co-culture method. The plantlets which tested positive with gene specific PCR analysis were raised and seeds were collected from them on maturity.

In the present work, the SbtA transporter from *Synechococcus elongatus* PCC 7942 was introduced and expressed in the chloroplasts of C₃ plant by fusing it with chloroplast transit peptides (cTPs) from two inner chloroplast envelope membrane proteins of *A. thaliana*. One of the constructs even depicted targeting to chloroplast envelope membrane which assures the success of our chimeric constructs. The study is the first report of targeting cyanobacterial bicarbonate transporter to the chloroplast of C₃ model plants using a minimal chimeric construct consisting of just the cTP.

BIBLIOGRAPHY

1. Alber, B.E. and Ferry, J.G. 1994. A carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *Proceedings of the National Academy of Sciences*, 91, 6909-6913.
2. Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E. and Bechner, M. 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*, 306, 79-86.
3. Ashida, H., Saito, Y., Kojima, C., Kobayashi, K., Ogasawara, N. and Yokota, A. 2003. A functional link between RuBisCO-like protein of *Bacillus* and photosynthetic RuBisCO. *Science*, 302, 286-290.
4. Atkinson, N., Feike, D., Mackinder, L., Meyer, M.T., Griffiths, H., Jonikas, M.C., Smith, A.M. and McCormick, A.J. 2016. Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant biotechnology journal*, 14, 1302-1315.
5. Atkinson, N., Leitão, N., Orr, D.J., Meyer, M.T., Carmo-Silva, E., Griffiths, H., Smith, A.M. and McCormick, A.J. 2017. Rubisco small subunits from the unicellular green alga *Chlamydomonas* complement Rubisco-deficient mutants of *Arabidopsis*. *New Phytologist*, 214, 655-667.
6. Badger, M.R., Andrews, T.J., Whitney, S., Ludwig, M., Yellowlees, D.C., Leggat, W. and Price, G.D. 1998. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO₂-concentrating mechanisms in algae. *Canadian Journal of Botany*, 76, 1052-1071.
7. Badger, M.R. and Bek, E.J. 2008. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle. *Journal of experimental botany*, 59, 1525-1541.
8. Badger, M.R., Hanson, D. and Price, G.D. 2002. Evolution and diversity of CO₂ concentrating mechanisms in cyanobacteria. *Functional Plant Biology*, 29, 161-173.
9. Badger, M.R. and Price, G.D. 1994. The role of carbonic anhydrase in photosynthesis. *Annual review of plant biology*, 45, 369-392.
10. Badger, M.R. and Price, G.D. 2003. CO₂ concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *Journal of experimental botany*, 54, 609-622.
11. Badger, M.R., Price, G.D., Long, B.M. and Woodger, F.J. 2005. The environmental plasticity and ecological genomics of the cyanobacterial CO₂ concentrating mechanism. *Journal of experimental botany*, 57, 249-265.
12. Bassil, E., Tajima, H., Liang, Y.-C., Ohto, M.-A., Ushijima, K., Nakano, R., Esumi, T., Coku, A., Belmonte, M. and Blumwald, E. 2011. The *Arabidopsis* Na⁺/H⁺ antiporters NHX1 and NHX2 control vacuolar pH and K⁺ homeostasis to regulate growth, flower development, and reproduction. *The Plant Cell*, 23, 3482-3497.
13. Beardall, J. 1989. Photosynthesis and photorespiration in marine phytoplankton. *Aquatic Botany*, 34, 105-130.
14. Beardall, J., Johnston, A. and Raven, J. 1998. Environmental regulation of CO₂-concentrating mechanisms in microalgae. *Canadian journal of botany*, 76, 1010-1017.
15. Bechtold, N. and Pelletier, G. 1998. In planta *Agrobacterium* mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Arabidopsis protocols*. Springer.
16. Bernsel, A., Viklund, H., Hennerdal, A. and Elofsson, A. 2009. TOPCONS: consensus prediction of membrane protein topology. *Nucleic acids research*, 37, W465-W468.
17. Betti, M., Bauwe, H., Busch, F.A., Fernie, A.R., Keech, O., Levey, M., Ort, D.R., Parry, M.A., Sage, R. and Timm, S. 2016. Manipulating photorespiration to increase plant productivity: recent advances and perspectives for crop improvement. *Journal of experimental botany*, 67, 2977-2988.

18. Bhushan, S., Kuhn, C., Berglund, A.-K., Roth, C. and Glaser, E. 2006. The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. *FEBS letters*, 580, 3966-3972.
19. Birch, R.G. 1997. Plant transformation: problems and strategies for practical application. *Annual review of plant biology*, 48, 297-326.
20. Birnboim, H. and Doly, J. 1979. Nucleic Acid Res., 7, 1513–1523. *CrossRef| PubMed| Web of Science*.
21. Bolhuis, H., Severin, I., Confurius-Guns, V., Wollenzien, U.I. and Stal, L.J. 2010. Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium *Microcoleus chthonoplastes*. *The ISME journal*, 4, 121.
22. Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A. and Savage, D.F. 2012. Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences*, 109, 478-483.
23. Borland, A.M., Griffiths, H., Hartwell, J. and Smith, J.a.C. 2009. Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *Journal of Experimental Botany*, 60, 2879-2896.
24. Borland, A.M., Hartwell, J., Weston, D.J., Schlauch, K.A., Tschaplinski, T.J., Tuskan, G.A., Yang, X. and Cushman, J.C. 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science*, 19, 327-338.
25. Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, 72, 248-254.
26. Bräutigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K.L., Carr, K.M., Gowik, U., Maß, J. and Lercher, M.J. 2011. An mRNA blueprint for C₄ photosynthesis derived from comparative transcriptomics of closely related C₃ and C₄ species. *Plant Physiology*, 155, 142-156.
27. Brkljacic, J., Grotewold, E., Scholl, R., Mockler, T., Garvin, D.F., Vain, P., Brutnell, T., Sibout, R., Bevan, M. and Budak, H. 2011. *Brachypodium* as a model for the grasses: today and the future. *Plant Physiology*, 157, 3-13.
28. Brueggeman, A.J., Gangadharaiyah, D.S., Cserhati, M.F., Casero, D., Weeks, D.P. and Ladunga, I. 2012. Activation of the carbon concentrating mechanism by CO₂ deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii*. *The Plant Cell*, 24, 1860-1875.
29. Bruhl, J.J. and Wilson, K.L. 2007. Towards a comprehensive survey of C₃ and C₄ photosynthetic pathways in Cyperaceae. *Aliso: A Journal of Systematic and Evolutionary Botany*, 23, 99-148.
30. Burnap, R.L., Hagemann, M. and Kaplan, A. 2015. Regulation of CO₂ concentrating mechanism in cyanobacteria. *Life*, 5, 348-371.
31. Cai, F., Dou, Z., Bernstein, S.L., Leverenz, R., Williams, E.B., Heinhorst, S., Shively, J., Cannon, G.C. and Kerfeld, C.A. 2015. Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life*, 5, 1141-1171.
32. Chen, M. and Blankenship, R.E. 2011. Expanding the solar spectrum used by photosynthesis. *Trends in plant science*, 16, 427-431.
33. Chen, P.-Y., Wang, C.-K., Soong, S.-C. and To, K.-Y. 2003. Complete sequence of the binary vector pBI121 and its application in cloning T-DNA insertion from transgenic plants. *Molecular breeding*, 11, 287-293.
34. Christin, P.A., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F. and Besnard, G. 2008. Evolutionary switch and genetic convergence on *rbcL* following the evolution of C₄ photosynthesis. *Molecular Biology and Evolution*, 25, 2361-2368.

35. Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The plant journal*, 16, 735-743.
36. Cockburn, W., Ting, I.P. and Sternberg, L.O. 1979. Relationships between stomatal behavior and internal carbon dioxide concentration in Crassulacean acid metabolism plants. *Plant Physiology*, 63, 1029-1032.
37. Colman, B., Huertas, I.E., Bhatti, S. and Dason, J.S. 2002. The diversity of inorganic carbon acquisition mechanisms in eukaryotic microalgae. *Functional Plant Biology*, 29, 261-270.
38. Colovos, C. and Yeates, T.O. 1993. Verification of protein structures: patterns of nonbonded atomic interactions. *Protein science*, 2, 1511-1519.
39. Condon, A., Farquhar, G. and Richards, R. 1990. Genotypic variation in carbon isotope discrimination and transpiration efficiency in wheat. Leaf gas exchange and whole plant studies. *Functional Plant Biology*, 17, 9-22.
40. Conroy, M.J., Durand, A., Lupo, D., Li, X.D., Bullough, P.A., Winkler, F.K. and Merrick, M. 2007. The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. *Proceedings of the National Academy of Sciences*, 104, 1213-1218.
41. Covshoff, S. and Hibberd, J.M. 2012. Integrating C₄ photosynthesis into C₃ crops to increase yield potential. *Current Opinion in Biotechnology*, 23, 209-214.
42. Cushman, J.C. and Bohnert, H.J. 1997. Molecular genetics of crassulacean acid metabolism. *Plant physiology*, 113, 667-676.
43. D'aoust, M.A., Lavoie, P.O., Couture, M.M.J., Trépanier, S., Guay, J.M., Dargis, M., Mongrand, S., Landry, N., Ward, B.J. and Vézina, L.P. 2008. Influenza virus-like particles produced by transient expression in *Nicotiana benthamiana* induce a protective immune response against a lethal viral challenge in mice. *Plant Biotechnology Journal*, 6, 930-940.
44. Daley, S.M., Kappell, A.D., Carrick, M.J. and Burnap, R.L. 2012. Regulation of the cyanobacterial CO₂-concentrating mechanism involves internal sensing of NADP⁺ and α -ketoglutarate levels by transcription factor CcmR. *PLoS One*, 7, e41286.
45. Dangel, A.W. and Tabita, F.R. 2015. CbbR, the master regulator for microbial carbon dioxide fixation. *Journal of bacteriology*, 197, 3488-3498.
46. Davis, A.M., Hall, A., Millar, A.J., Darrah, C. and Davis, S.J. 2009. Protocol: Streamlined sub-protocols for floral-dip transformation and selection of transformants in *Arabidopsis thaliana*. *Plant Methods*, 5, 3.
47. De Fc Carvalho, J., Madgwick, P.J., Powers, S.J., Keys, A.J., Lea, P.J. and Parry, M.A. 2011. An engineered pathway for glyoxylate metabolism in tobacco plants aimed to avoid the release of ammonia in photorespiration. *BMC biotechnology*, 11, 111.
48. Depaoli, H.C., Borland, A.M., Tuskan, G.A., Cushman, J.C. and Yang, X. 2014. Synthetic biology as it relates to CAM photosynthesis: challenges and opportunities. *Journal of experimental botany*, 65, 3381-3393.
49. Dobrinski, K.P., Longo, D.L. and Scott, K.M. 2005. The carbon-concentrating mechanism of the hydrothermal vent chemolithoautotroph *Thiomicrospira crunogena*. *Journal of bacteriology*, 187, 5761-5766.
50. Du, J., Förster, B., Rourke, L., Howitt, S.M. and Price, G.D. 2014. Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. *PLoS One*, 9, e115905.
51. Dubbs, P., Dubbs, J.M. and Tabita, F.R. 2004. Effector-mediated interaction of CbbRI and CbbRII regulators with target sequences in *Rhodobacter capsulatus*. *Journal of bacteriology*, 186, 8026-8035.
52. Edwards, G.E., Franceschi, V.R. and Voznesenskaya, E.V. 2004. Single-cell C₄ photosynthesis versus the dual-cell (Kranz) paradigm. *Annu. Rev. Plant Biol.*, 55, 173-196.
53. Eisenberg, D., Lüthy, R. and Bowie, J.U. 1997. [20] VERIFY3D: Assessment of protein models with three-dimensional profiles. *Methods in enzymology*. Elsevier.

54. Emanuelsson, O., Nielsen, H., Brunak, S. and Von Heijne, G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of molecular biology*, 300, 1005-1016.
55. Emanuelsson, O., Nielsen, H. and Von Heijne, G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science*, 8, 978-984.
56. Espie, G.S. and Kimber, M.S. 2011. Carboxysomes: cyanobacterial RubisCO comes in small packages. *Photosynthesis research*, 109, 7-20.
57. Evans, J. and Poorter, H. 2001. Photosynthetic acclimation of plants to growth irradiance: the relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain. *Plant, Cell and Environment*, 24, 755-767.
58. Evans, J.R. 1998. Carbon dioxide diffusion inside C₃ leaves. *Photosynthesis: Mechanisms and Effects*. Springer.
59. Evans, J.R. 2013. Improving photosynthesis. *Plant physiology*, 162, 1780-1793.
60. Evans, J.R., Kaldenhoff, R., Genty, B. and Terashima, I. 2009. Resistances along the CO₂ diffusion pathway inside leaves. *Journal of Experimental Botany*, 60, 2235-2248.
61. Evans, J.R. and Seemann, J.R. 1989. The allocation of protein nitrogen in the photosynthetic apparatus: costs, consequences, and control. *Photosynthesis*, 183-205.
62. Evans, J.R. and Von Caemmerer, S. 1996. Carbon Dioxide Diffusion inside Leaves. *Plant Physiology*, 110, 339-346.
63. Feng, L., Han, Y., Liu, G., An, B., Yang, J., Yang, G., Li, Y. and Zhu, Y. 2007a. Overexpression of sedoheptulose-1, 7-bisphosphatase enhances photosynthesis and growth under salt stress in transgenic rice plants. *Functional Plant Biology*, 34, 822-834.
64. Feng, L., Wang, K., Li, Y., Tan, Y., Kong, J., Li, H., Li, Y. and Zhu, Y. 2007b. Overexpression of SBPase enhances photosynthesis against high temperature stress in transgenic rice plants. *Plant cell reports*, 26, 1635-1646.
65. Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S. and Kieffer-Jaquinod, S. 2010. AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Molecular and Cellular Proteomics*, 9, 1063-1084.
66. Ferro, M., Salvi, D., Rivière-Rolland, H., Verinat, T., Seigneurin-Berny, D., Grunwald, D., Garin, J., Joyard, J. and Rolland, N. 2002. Integral membrane proteins of the chloroplast envelope: identification and subcellular localization of new transporters. *Proceedings of the National Academy of Sciences*, 99, 11487-11492.
67. Forchhammer, K. and Lüddecke, J. 2016. Sensory properties of the PII signalling protein family. *The FEBS journal*, 283, 425-437.
68. Forsbach, A., Schubert, D., Lechtenberg, B., Gils, M. and Schmidt, R. 2003. A comprehensive characterization of single-copy T-DNA insertions in the *Arabidopsis thaliana* genome. *Plant molecular biology*, 52, 161-176.
69. Froehlich, J.E. and Keegstra, K. 2011. The role of the transmembrane domain in determining the targeting of membrane proteins to either the inner envelope or thylakoid membrane. *The Plant Journal*, 68, 844-856.
70. Fuhrmann, M., Hausherr, A., Ferbitz, L., Schödl, T., Heitzer, M. and Hegemann, P. 2004. Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant molecular biology*, 55, 869-881.
71. Fukuzawa, H., Miura, K., Ishizaki, K., Kucho, K.-I., Saito, T., Kohinata, T. and Ohyama, K. 2001. Ccm1, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing CO₂ availability. *Proceedings of the National Academy of Sciences*, 98, 5347-5352.

72. Furbank, R.T., Quick, W.P. and Sirault, X.R. 2015. Improving photosynthesis and yield potential in cereal crops by targeted genetic manipulation: prospects, progress and challenges. *Field Crops Research*, 182, 19-29.
73. Gallois, P. and Marinho, P. 1995. Leaf disk transformation using *Agrobacterium tumefaciens*-expression of heterologous genes in tobacco. *Plant gene transfer and expression protocols*. Springer.
74. Gehl, K.A., Cook, C.M. and Colman, B. 1987. The effect of external pH on the apparent CO₂ affinity of *Chlorella saccharophila*. *Journal of experimental botany*, 38, 1203-1210.
75. Genkov, T. and Spreitzer, R.J. 2009. Highly conserved small subunit residues influence rubisco large subunit catalysis. *Journal of Biological Chemistry*, 284, 30105-30112.
76. Gheysen, G., Van Montagu, M. and Zambryski, P. 1987. Integration of *Agrobacterium tumefaciens* transfer DNA (T-DNA) involves rearrangements of target plant DNA sequences. *Proceedings of the National Academy of Sciences*, 84, 6169-6173.
77. Gowik, U., Bräutigam, A., Weber, K.L., Weber, A.P. and Westhoff, P. 2011. Evolution of C₄ photosynthesis in the genus *Flaveria*: how many and which genes does it take to make C₄? *The Plant Cell*, 23, 2087-2105.
78. Gowik, U. and Westhoff, P. 2011. The path from C₃ to C₄ photosynthesis. *Plant Physiology*, 155, 56-63.
79. Gruber, A., Weber, T., Bártulos, C.R., Vugrinec, S. and Kroth, P.G. 2009. Intracellular distribution of the reductive and oxidative pentose phosphate pathways in two diatoms. *Journal of basic microbiology*, 49, 58-72.
80. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series, 1999. [London]: Information Retrieval Ltd., c1979-c2000., 95-98.
81. Hanson, M.R., Gray, B.N. and Ahner, B.A. 2012. Chloroplast transformation for engineering of photosynthesis. *Journal of experimental botany*, 64, 731-742.
82. Harada, H., Nakatsuma, D., Ishida, M. and Matsuda, Y. 2005. Regulation of the expression of intracellular β -carbonic anhydrase in response to CO₂ and light in the marine diatom *Phaeodactylum tricorutum*. *Plant physiology*, 139, 1041-1050.
83. Henkes, S., Sonnewald, U., Badur, R., Flachmann, R. and Stitt, M. 2001. A small decrease of plastid transketolase activity in antisense tobacco transformants has dramatic effects on photosynthesis and phenylpropanoid metabolism. *The Plant Cell*, 13, 535-551.
84. Howitt, C.A., Udall, P.K. and Vermaas, W.F. 1999. Type 2 NADH Dehydrogenases in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803 Are Involved in Regulation Rather Than Respiration. *Journal of bacteriology*, 181, 3994-4003.
85. Hu, C.-Y., Chee, P.P., Chesney, R.H., Zhou, J.H., Miller, P.D. and O'brien, W.T. 1990. Intrinsic GUS-like activities in seed plants. *Plant cell reports*, 9, 1-5.
86. Hwang, H.-H., Gelvin, S.B. and Lai, E.-M. 2015. "Agrobacterium biology and its application to transgenic plant production". *Frontiers in plant science*, 6, 265.
87. Inoue, K. 2011. Emerging roles of the chloroplast outer envelope membrane. *Trends in plant science*, 16, 550-557.
88. Jabnourne, M., Secco, D., Lecampion, C., Robaglia, C., Shu, Q. and Poirier, Y. 1993. An Efficient Procedure for Protoplast Isolation from Mesophyll Cells and Nuclear Fractionation in Rice. *The Plant Cell*.
89. Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO journal*, 6, 3901-3907.
90. Joshi, G.S., Bobst, C.E. and Tabita, F.R. 2011. Unravelling the regulatory twist—regulation of CO₂ fixation in *Rhodospseudomonas palustris* CGA010 mediated by atypical response regulator (s). *Molecular microbiology*, 80, 756-771.

91. Joshi, G.S., Zianni, M., Bobst, C.E. and Tabita, F.R. 2012. Further unraveling the regulatory twist by elucidating metabolic coinducer-mediated CbbR-cbbI promoter interactions in *Rhodospseudomonas palustris* CGA010. *Journal of bacteriology*, 194, 1350-1360.
92. Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H. and Xu, J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nature protocols*, 7, 1511.
93. Kaplan, A. and Reinhold, L. 1999. CO₂ concentrating mechanisms in photosynthetic microorganisms. *Annual review of plant biology*, 50, 539-570.
94. Kaplan, A., Zenvirth, D., Marcus, Y., Omata, T. and Ogawa, T. 1987. Energization and activation of inorganic carbon uptake by light in cyanobacteria. *Plant Physiology*, 84, 210-213.
95. Karagouni, A.D. and Kelly, D.P. 1989. Carbon dioxide fixation by *Thiobacillus versutus*: apparent absence of a CO₂-concentrating mechanism in organisms grown under carbon-limitation in the chemostat. *FEMS microbiology letters*, 58, 179-182.
96. Karley, A.J., Leigh, R.A. and Sanders, D. 2000. Where do all the ions go? The cellular basis of differential ion accumulation in leaf cells. *Trends in plant science*, 5, 465-470.
97. Kawamoto, H., Suzuki, R., Ugaki, M. and Kawano, S. 2016. Location of Gold Particles and Puncture of Tobacco Leaf Epidermis by Particle Bombardment. *Cytologia*, 81, 455-458.
98. Kebeish, R., Niessen, M., Oksaksin, M., Blume, C. and Peterhaensel, C. 2012. Constitutive and dark-induced expression of *Solanum tuberosum* phosphoenolpyruvate carboxylase enhances stomatal opening and photosynthetic performance of *Arabidopsis thaliana*. *Biotechnology and bioengineering*, 109, 536-544.
99. Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.-J., Rosenkranz, R., Stähler, N., Schönfeld, B., Kreuzaler, F. and Peterhänsel, C. 2007. Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nature biotechnology*, 25, 593.
100. Keeley, J.E. 1983. Crassulacean acid metabolism in the seasonally submerged aquatic *Isoetes howellii*. *Oecologia*, 58, 57-62.
101. Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M. and Yeates, T.O. 2005. Protein structures forming the shell of primitive bacterial organelles. *Science*, 309, 936-938.
102. Kikkert, J.R., Vidal, J.R. and Reisch, B.I. 2005. Stable transformation of plant cells by particle bombardment/biolistics. *Transgenic plants: methods and protocols*. Springer.
103. Kinney, J.N., Axen, S.D. and Kerfeld, C.A. 2011. Comparative analysis of carboxysome shell proteins. *Photosynthesis research*, 109, 21-32.
104. Ko, S.B., Zeng, W., Dorwart, M.R., Luo, X., Kim, K.H., Millen, L., Goto, H., Naruse, S., Soyombo, A. and Thomas, P.J. 2004. Gating of CFTR by the STAS domain of SLC26 transporters. *Nature cell biology*, 6, 343.
105. Kohli, A., Miro, B. and Twyman, R.M. 2010. Transgene integration, expression and stability in plants: strategies for improvements. *Transgenic crop plants*. Springer.
106. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. and Karplus, K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins: Structure, Function, and Bioinformatics*, 77, 114-122.
107. Krogh, A., Larsson, B., Von Heijne, G. and Sonnhammer, E.L. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes I. *Journal of molecular biology*, 305, 567-580.
108. Kromdijk, J., Glowacka, K., Leonelli, L., Gabilly, S.T., Iwai, M., Niyogi, K.K. and Long, S.P. 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science*, 354, 857-861.
109. Kroth, P.G. 2015. The biodiversity of carbon assimilation. *Journal of plant physiology*, 172, 76-81.

110. Kroth, P.G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M.S., Stanley, M.S., Kaplan, A., Caron, L. and Weber, T. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS one*, 3, e1426.
111. Krügel, T., Lim, M., Gase, K., Halitschke, R. and Baldwin, I.T. 2002. *Agrobacterium*-mediated transformation of *Nicotiana attenuata*, a model ecological expression system. *Chemoecology*, 12, 177-183.
112. Kubien, D.S. and Sage, R.F. 2008. The temperature response of photosynthesis in tobacco with reduced amounts of Rubisco. *Plant, cell and environment*, 31, 407-418.
113. Kumar, A., Li, C. and Portis, A.R. 2009. *Arabidopsis thaliana* expressing a thermostable chimeric Rubisco activase exhibits enhanced growth and higher rates of photosynthesis at moderately high temperatures. *Photosynthesis research*, 100, 143-153.
114. Kurek, I., Chang, T.K., Bertain, S.M., Madrigal, A., Liu, L., Lassner, M.W. and Zhu, G. 2007. Enhanced thermostability of *Arabidopsis* Rubisco activase improves photosynthesis and growth rates under moderate heat stress. *The Plant Cell*, 19, 3230-3241.
115. Kusian, B. and Bowien, B. 1997. Organization and regulation of *cbb* CO₂ assimilation genes in autotrophic bacteria. *FEMS microbiology reviews*, 21, 135-155.
116. Laskowski, R.A., Macarthur, M.W., Moss, D.S. and Thornton, J.M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, 26, 283-291.
117. Latham, J.R., Wilson, A.K. and Steinbrecher, R.A. 2006. The mutational consequences of plant transformation. *BioMed Research International*, 2006.
118. Leegood, R.C. 2002. C₄ photosynthesis: principles of CO₂ concentration and prospects for its introduction into C₃ plants. *Journal of experimental botany*, 53, 581-590.
119. Lefebvre, S., Lawson, T., Fryer, M., Zakhleniuk, O.V., Lloyd, J.C. and Raines, C.A. 2005. Increased sedoheptulose-1, 7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. *Plant Physiology*, 138, 451-460.
120. Lin, M.T., Occhialini, A., Andralojc, P.J., Devonshire, J., Hines, K.M., Parry, M.A. and Hanson, M.R. 2014. β -Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. *The Plant Journal*, 79, 1-12.
121. Linka, N. and Weber, A.P. 2010. Intracellular metabolite transporters in plants. *Molecular plant*, 3, 21-53.
122. Loganathan, N., Tsai, Y.-C.C. and Mueller-Cajar, O. 2016. Characterization of the heterooligomeric red-type rubisco activase from red algae. *Proceedings of the National Academy of Sciences*, 113, 14019-14024.
123. Long, S. 1991. Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO₂ concentrations: has its importance been underestimated? *Plant, Cell and Environment*, 14, 729-739.
124. Long, S.P., Zhu, X.G., Naidu, S.L. and Ort, D.R. 2006. Can improvement in photosynthesis increase crop yields? *Plant, Cell and Environment*, 29, 315-330.
125. Lopez, E., Proano, K., Jadan, M. and Mihai, R. 2015. Callus tissue induction and analysis of GUS reporter gene expression in tomato (*Solanum lycopersicum* L.) transformed with *Agrobacterium tumefaciens*. *Romanian Biotechnological Letters*, 20, 10205.
126. Lovell, S.C., Davis, I.W., Arendall, W.B., De Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S. and Richardson, D.C. 2003. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins: Structure, Function, and Bioinformatics*, 50, 437-450.
127. Ludwig, M., Sültemeyer, D. and Price, G.D. 2000. Isolation of ccmKLMN genes from the marine cyanobacterium, *Synechococcus* sp. PCC7002 (Cyanophyceae), and evidence that CcmM is essential for carboxysome assembly. *Journal of Phycology*, 36, 1109-1119.

128. Lüttge, U. 2004. Ecophysiology of crassulacean acid metabolism (CAM). *Annals of botany*, 93, 629-652.
129. Mackinder, L.C., Chen, C., Leib, R.D., Patena, W., Blum, S.R., Rodman, M., Ramundo, S., Adams, C.M. and Jonikas, M.C. 2017. A Spatial Interactome Reveals the Protein Organization of the Algal CO₂-Concentrating Mechanism. *Cell*, 171, 133-147. e14.
130. Mackinder, L.C., Meyer, M.T., Mettler-Altmann, T., Chen, V.K., Mitchell, M.C., Caspari, O., Rosenzweig, E.S.F., Pallesen, L., Reeves, G. and Itakura, A. 2016. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences*, 113, 5958-5963.
131. Maddocks, S.E. and Oyston, P.C. 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology*, 154, 3609-3623.
132. Maeda, S.-I. and Omata, T. 1997. Substrate-binding lipoprotein of the cyanobacterium *Synechococcus* sp. strain PCC 7942 involved in the transport of nitrate and nitrite. *Journal of Biological Chemistry*, 272, 3036-3041.
133. Maeda, S.I., Badger, M.R. and Price, G.D. 2002. Novel gene products associated with NdhD3/D4-containing NDH-1 complexes are involved in photosynthetic CO₂ hydration in the cyanobacterium, *Synechococcus* sp. PCC7942. *Molecular microbiology*, 43, 425-435.
134. Maier, A., Fahnenstich, H., Von Caemmerer, S., Engqvist, M.K., Weber, A.P., Flügge, U.-I. and Maurino, V.G. 2012. Transgenic introduction of a glycolate oxidative cycle into *A. thaliana* chloroplasts leads to growth improvement. *Frontiers in plant science*, 3, 38.
135. Mara, C., Grigorova, B. and Liu, Z. 2010. Floral-dip transformation of *Arabidopsis thaliana* to examine pTSO2:: β -glucuronidase reporter gene expression. *Journal of visualized experiments: JoVE*.
136. Mathur, A., Mathur, A.K., Verma, P., Yadav, S., Gupta, M.L. and Darokar, M.P. 2008. Biological hardening and genetic fidelity testing of micro-cloned progeny of *Chlorophytum borivillianum* Sant. et Fernand. *African Journal of Biotechnology*, 7.
137. Matsuda, Y. and Kroth, P.G. 2014. Carbon fixation in diatoms. *The structural basis of biological energy generation*. Springer.
138. Mcginn, P.J., Price, G.D. and Badger, M. 2004. High light enhances the expression of low-CO₂-inducible transcripts involved in the CO₂-concentrating mechanism in *Synechocystis* sp. PCC6803. *Plant, Cell and Environment*, 27, 615-626.
139. Mcginn, P.J., Price, G.D., Maleszka, R. and Badger, M.R. 2003. Inorganic carbon limitation and light control the expression of transcripts related to the CO₂-concentrating mechanism in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Physiology*, 132, 218-229.
140. Mcintosh, K., Hulm, J., Young, L. and Bonham-Smith, P. 2004. A rapid *Agrobacterium*-mediated *Arabidopsis thaliana* transient assay system. *Plant molecular biology reporter*, 22, 53-61.
141. Meijer, W.G., Van Den Bergh, E. and Smith, L.M. 1996. Induction of the *gap-pgk* operon encoding glyceraldehyde-3-phosphate dehydrogenase and 3-phosphoglycerate kinase of *Xanthobacter flavus* requires the LysR-type transcriptional activator CbbR. *Journal of bacteriology*, 178, 881-887.
142. Meyer, M. and Griffiths, H. 2013. Origins and diversity of eukaryotic CO₂-concentrating mechanisms: lessons for the future. *Journal of experimental botany*, 64, 769-786.
143. Mi, H., Endo, T., Schreiber, U., Ogawa, T. and Asada, K. 1992. Electron donation from cyclic and respiratory flows to the photosynthetic intersystem chain is mediated by pyridine nucleotide dehydrogenase in the cyanobacterium *Synechocystis* PCC 6803. *Plant and cell physiology*, 33, 1233-1237.
144. Mimuro, M., Yokono, M. and Akimoto, S. 2010. Variations in photosystem I properties in the primordial cyanobacterium *Gloeobacter violaceus* PCC 7421. *Photochemistry and photobiology*, 86, 62-69.

145. Miyao, M., Masumoto, C., Miyazawa, S.-I. and Fukayama, H. 2011. Lessons from engineering a single-cell C₄ photosynthetic pathway into rice. *Journal of experimental botany*, 62, 3021-3029.
146. Monson, R.K., Edwards, G.E. and Ku, M.S. 1984. C₃-C₄ intermediate photosynthesis in plants. *Bioscience*, 34, 563-574.
147. Montes, G. and Bradbeer, J. 1976. An association of chloroplasts and mitochondria in *Zea mays* and *Hyptis suaveolens*. *Plant Science Letters*, 6, 35-41.
148. Moore, B.D., Franceschi, V.R., Cheng, S.-H., Wu, J. and Ku, M.S. 1987. Photosynthetic characteristics of the C₃-C₄ intermediate *Parthenium hysterophorus*. *Plant physiology*, 85, 978-983.
149. Morita, E., Abe, T., Tsuzuki, M., Fujiwara, S., Sato, N., Hirata, A., Sonoike, K. and Nozaki, H. 1998. Presence of the CO₂-concentrating mechanism in some species of the pyrenoid-less free-living algal genus *Chloromonas* (Volvocales, Chlorophyta). *Planta*, 204, 269-276.
150. Morita, E., Abe, T., Tsuzuki, M., Fujiwara, S., Sato, N., Hirata, A., Sonoike, K. and Nozaki, H. 1999. Role of pyrenoids in the CO₂-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). *Planta*, 208, 365-372.
151. Moroney, J., Bartlett, S. and Samuelsson, G. 2001. Carbonic anhydrases in plants and algae. *Plant, Cell and Environment*, 24, 141-153.
152. Moroney, J.V., Ma, Y., Frey, W.D., Fusilier, K.A., Pham, T.T., Simms, T.A., Dimario, R.J., Yang, J. and Mukherjee, B. 2011. The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. *Photosynthesis Research*, 109, 133-149.
153. Moroney, J.V. and Ynalvez, R.A. 2007. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryotic cell*, 6, 1251-1259.
154. Müller, P., Li, X.-P. and Niyogi, K.K. 2001. Non-photochemical quenching. A response to excess light energy. *Plant physiology*, 125, 1558-1566.
155. Murashige, T. and Skoog, F. 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum*, 15, 473-497.
156. Murchie, E.H. and Niyogi, K.K. 2011. Manipulation of photoprotection to improve plant photosynthesis. *Plant physiology*, 155, 86-92.
157. Nakajima, K., Tanaka, A. and Matsuda, Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proceedings of the National Academy of Sciences*, 110, 1767-1772.
158. Nakamura, Y., Salvucci, M. and Bowes, G. Glucose-6-phosphate activation and malate inhibition of *Hydrilla* PEP carboxylase. PLANT PHYSIOLOGY, 1983. AMER SOC PLANT BIOLOGISTS 15501 MONONA DRIVE, ROCKVILLE, MD 20855 USA, 42-42.
159. Nei, M. and Kumar, S. 2000. *Molecular evolution and phylogenetics*, Oxford university press.
160. Nelissen, B., Van De Peer, Y., Wilmotte, A. and De Wachter, R. 1995. An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. *Molecular Biology and Evolution*, 12, 1166-1173.
161. Nishimura, T., Takahashi, Y., Yamaguchi, O., Suzuki, H., Maeda, S.I. and Omata, T. 2008. Mechanism of low CO₂-induced activation of the *cmp* bicarbonate transporter operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus* strain PCC 7942. *Molecular microbiology*, 68, 98-109.
162. Ochoa, D. and Pazos, F. 2010. Studying the co-evolution of protein families with the Mirrortree web server. *Bioinformatics*, 26, 1370-1371.
163. Ohkawa, H., Sonoda, M., Hagino, N., Shibata, M., Pakrasi, H.B. and Ogawa, T. 2002. Functionally distinct NAD(P)H dehydrogenases and their membrane localization in *Synechocystis* sp. PCC6803. *Functional plant biology*, 29, 195-200.

164. Ohno, N., Inoue, T., Yamashiki, R., Nakajima, K. and Kitahara, Y. CO₂-cAMP-Responsive cis-Elements Targeted by a Transcription Factor with CREB/ATF-Like Basic Zipper Domain in the Marine Diatom.
165. Omata, T. 1991. Cloning and characterization of the *nrtA* gene that encodes a 45-kDa protein involved in nitrate transport in the cyanobacterium *Synechococcus* PCC 7942. *Plant and cell physiology*, 32, 151-157.
166. Omata, T., Price, G.D., Badger, M.R., Okamura, M., Gohta, S. and Ogawa, T. 1999. Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proceedings of the National Academy of Sciences*, 96, 13571-13576.
167. Ort, D.R., Zhu, X. and Melis, A. 2011. Optimizing antenna size to maximize photosynthetic efficiency. *Plant physiology*, 155, 79-85.
168. Paerl, H.W. 2012. Marine plankton. *Ecology of Cyanobacteria II*. Springer.
169. Parry, M., Andralojc, P., Mitchell, R.A., Madgwick, P. and Keys, A. 2003. Manipulation of Rubisco: the amount, activity, function and regulation. *Journal of experimental botany*, 54, 1321-1333.
170. Parry, M., Madgwick, P., Carvalho, J. and Andralojc, P. 2007. Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. *The Journal of Agricultural Science*, 145, 31.
171. Parry, M.A., Andralojc, P.J., Scales, J.C., Salvucci, M.E., Carmo-Silva, A.E., Alonso, H. and Whitney, S.M. 2012. Rubisco activity and regulation as targets for crop improvement. *Journal of Experimental Botany*, 64, 717-730.
172. Parry, M.A., Reynolds, M., Salvucci, M.E., Raines, C., Andralojc, P.J., Zhu, X.-G., Price, G.D., Condon, A.G. and Furbank, R.T. 2010. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of experimental botany*, 62, 453-467.
173. Pengelly, J., Förster, B., Von Caemmerer, S., Badger, M., Price, G.D. and Whitney, S. 2014. Transplastomic integration of a cyanobacterial bicarbonate transporter into tobacco chloroplasts. *Journal of experimental botany*, 65, 3071-3080.
174. Perez, E.A. and Hooks, C. 2008. Preparing tissue-cultured banana plantlets for field planting.
175. Peterhansel, C., Horst, I., Niessen, M., Blume, C., Kebeish, R., Kürkcüoglu, S. and Kreuzaler, F. 2010. Photorespiration. *The Arabidopsis Book*, e0130.
176. Peterhansel, C., Krause, K., Braun, H.P., Espie, G., Fernie, A., Hanson, D., Keech, O., Maurino, V., Mielewczik, M. and Sage, R. 2013. Engineering photorespiration: current state and future possibilities. *Plant Biology*, 15, 754-758.
177. Peterhansel, C. and Maurino, V.G. 2011. Photorespiration redesigned. *Plant physiology*, 155, 49-55.
178. Portis, A.R. and Parry, M.A. 2007. Discoveries in Rubisco (Ribulose 1, 5-bisphosphate carboxylase/oxygenase): a historical perspective. *Photosynthesis Research*, 94, 121-143.
179. Pospíšilová, J., Tichá, I., Kadleček, P., Haisel, D. and Plzánková, Š. 1999. Acclimatization of micropropagated plants to *ex-vitro* conditions. *Biologia Plantarum*, 42, 481-497.
180. Price, G. and Badger, M. 1989. Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO₂-requiring phenotype: evidence for a central role for carboxysomes in the CO₂ concentrating mechanism. *Plant physiology*, 91, 505-513.
181. Price, G.D. 2011. Inorganic carbon transporters of the cyanobacterial CO₂ concentrating mechanism. *Photosynthesis Research*, 109, 47-57.
182. Price, G.D., Badger, M.R. and Von Caemmerer, S. 2011a. The prospect of using cyanobacterial bicarbonate transporters to improve leaf photosynthesis in C₃ crop plants. *Plant Physiology*, 155, 20-26.
183. Price, G.D., Badger, M.R., Woodger, F.J. and Long, B.M. 2007. Advances in understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): functional components, Ci

- transporters, diversity, genetic regulation and prospects for engineering into plants. *Journal of experimental botany*, 59, 1441-1461.
184. Price, G.D. and Howitt, S.M. 2010. The cyanobacterial bicarbonate transporter BicA: its physiological role and the implications of structural similarities with human SLC26 transporters. *Biochemistry and Cell Biology*, 89, 178-188.
185. Price, G.D. and Howitt, S.M. 2014. Topology mapping to characterize cyanobacterial bicarbonate transporters: BicA (SulP/SLC26 family) and SbtA. *Molecular membrane biology*, 31, 177-182.
186. Price, G.D., Maeda, S.-I., Omata, T. and Badger, M.R. 2002. Modes of active inorganic carbon uptake in the cyanobacterium, *Synechococcus* sp. PCC7942. *Functional Plant Biology*, 29, 131-149.
187. Price, G.D., Pengelly, J.J., Forster, B., Du, J., Whitney, S.M., Von Caemmerer, S., Badger, M.R., Howitt, S.M. and Evans, J.R. 2012. The cyanobacterial CCM as a source of genes for improving photosynthetic CO₂ fixation in crop species. *Journal of experimental botany*, 64, 753-768.
188. Price, G.D., Shelden, M.C. and Howitt, S.M. 2011b. Membrane topology of the cyanobacterial bicarbonate transporter, SbtA, and identification of potential regulatory loops. *Molecular Membrane Biology*, 28, 265-275.
189. Price, G.D., Sültemeyer, D., Klughammer, B., Ludwig, M. and Badger, M.R. 1998. The functioning of the CO₂ concentrating mechanism in several cyanobacterial strains: a review of general physiological characteristics, genes, proteins, and recent advances. *Canadian Journal of Botany*, 76, 973-1002.
190. Price, G.D., Woodger, F.J., Badger, M.R., Howitt, S.M. and Tucker, L. 2004. Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. *Proceedings of the National Academy of Sciences*, 101, 18228-18233.
191. Puigbò, P., Bravo, I.G. and Garcia-Vallve, S. 2008. CAIcal: a combined set of tools to assess codon usage adaptation. *Biology direct*, 3, 38.
192. Rae, B.D., Förster, B., Badger, M.R. and Price, G.D. 2011. The CO₂-concentrating mechanism of *Synechococcus* WH5701 is composed of native and horizontally-acquired components. *Photosynthesis research*, 109, 59-72.
193. Rae, B.D., Long, B.M., Badger, M.R. and Price, G.D. 2013. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO₂ fixation in cyanobacteria and some proteobacteria. *Microbiology and Molecular Biology Reviews*, 77, 357-379.
194. Raines, C. and Paul, M. 2006. Products of leaf primary carbon metabolism modulate the developmental programme determining plant morphology. *Journal of Experimental Botany*, 57, 1857-1862.
195. Raines, C.A. 2011. Increasing photosynthetic carbon assimilation in C₃ plants to improve crop yield: current and future strategies. *Plant physiology*, 155, 36-42.
196. Raines, C.A., Lloyd, J.C. and Dyer, T.A. 1999. New insights into the structure and function of sedoheptulose-1, 7-bisphosphatase; an important but neglected Calvin cycle enzyme. *Journal of Experimental Botany*, 50, 1-8.
197. Rangan, P., Furtado, A. and Henry, R.J. 2016. New evidence for grain specific C₄ photosynthesis in wheat. *Scientific reports*, 6, 31721.
198. Raven, J. and Falkowski, P. 1999. Oceanic sinks for atmospheric CO₂. *Plant, Cell and Environment*, 22, 741-755.
199. Raven, J.A. and Allen, J.F. 2003. Genomics and chloroplast evolution: what did cyanobacteria do for plants? *Genome biology*, 4, 209.
200. Raven, J.A., Cockell, C.S. and De La Rocha, C.L. 2008. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 2641-2650.

201. Rebetzke, G.J., Rattey, A.R., Farquhar, G.D., Richards, R.A. and Condon, A.T.G. 2013. Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. *Functional Plant Biology*, 40, 14-33.
202. Reinfelder, J.R., Milligan, A.J. and Morel, F.M. 2004. The role of the C₄ pathway in carbon accumulation and fixation in a marine diatom. *Plant Physiology*, 135, 2106-2111.
203. Rolland, N., Ferro, M., Seigneurin-Berny, D., Garin, J., Douce, R. and Joyard, J. 2003. Proteomics of chloroplast envelope membranes. *Photosynthesis research*, 78, 205-230.
204. Rolland, V., Badger, M.R. and Price, G.D. 2016. Redirecting the cyanobacterial bicarbonate transporters BicA and SbtA to the chloroplast envelope: soluble and membrane cargos need different chloroplast targeting signals in plants. *Frontiers in plant science*, 7, 185.
205. Rolland, V., Rae, B.D. and Long, B.M. 2017. Setting sub-organellar sights: accurate targeting of multi-transmembrane-domain proteins to specific chloroplast membranes. *Journal of experimental botany*, 68, 5013-5016.
206. Romagnoli, S. and Tabita, F.R. 2007. Phosphotransfer reactions of the CbbRRS three-protein two-component system from *Rhodospseudomonas palustris* CGA010 appear to be controlled by an internal molecular switch on the sensor kinase. *Journal of bacteriology*, 189, 325-335.
207. Rosado-Alberio, J., Weier, T.E. and Stocking, C. 1968. Continuity of the chloroplast membrane systems in *Zea mays* L. *Plant physiology*, 43, 1325-1331.
208. Rosenthal, D.M., Locke, A.M., Khozaei, M., Raines, C.A., Long, S.P. and Ort, D.R. 2011. Over-expressing the C₃ photosynthesis cycle enzyme sedoheptulose-1-7 bisphosphatase improves photosynthetic carbon gain and yield under fully open air CO₂ fumigation (FACE). *BMC Plant Biology*, 11, 123.
209. Rudi, K., Skulberg, O.M., Larsen, F. and Jakobsen, K.S. 1997. Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. *Applied and Environmental Microbiology*, 63, 2593-2599.
210. Sage, R.F. 2002. C₄ photosynthesis in terrestrial plants does not require Kranz anatomy. *Trends in Plant Science*, 7, 283-285.
211. Sage, R.F. and Monson, R.K. 1998. *C₄ plant biology*, Academic Press.
212. Saier Jr, M.H., Eng, B.H., Fard, S., Garg, J., Haggerty, D.A., Hutchinson, W.J., Jack, D.L., Lai, E.C., Liu, H.J. and Nusinew, D.P. 1999. Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1422, 1-56.
213. Šali, A. and Blundell, T.L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*, 234, 779-815.
214. Salvucci, M.E. and Crafts-Brandner, S.J. 2004. Relationship between the heat tolerance of photosynthesis and the thermal stability of Rubisco activase in plants from contrasting thermal environments. *Plant Physiology*, 134, 1460-1470.
215. Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular cloning: a laboratory manual*, Cold spring harbor laboratory press.
216. Sambrook, J. and Russell, D.W. 2001. *Molecular cloning: a laboratory manual 3rd edition*. Coldspring-Harbour Laboratory Press, UK.
217. Schell, M.A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annual Reviews in Microbiology*, 47, 597-626.
218. Schuler, M.L., Mantegazza, O. and Weber, A.P. 2016. Engineering C₄ photosynthesis into C₃ chassis in the synthetic biology age. *The Plant Journal*, 87, 51-65.
219. Sharwood, R.E., Von Caemmerer, S., Maliga, P. and Whitney, S.M. 2008. The catalytic properties of hybrid Rubisco comprising tobacco small and sunflower large subunits mirror the kinetically equivalent source Rubiscos and can support tobacco growth. *Plant Physiology*, 146, 83-96.

220. Shelden, M.C., Howitt, S.M. and Price, G.D. 2010. Membrane topology of the cyanobacterial bicarbonate transporter, BicA, a member of the SulP (SLC26A) family. *Molecular membrane biology*, 27, 12-22.
221. Shibata, M., Katoh, H., Sonoda, M., Ohkawa, H., Shimoyama, M., Fukuzawa, H., Kaplan, A. and Ogawa, T. 2002. Genes essential to sodium-dependent bicarbonate transport in cyanobacteria function and phylogenetic analysis. *Journal of Biological Chemistry*, 277, 18658-18664.
222. Shibata, M., Ohkawa, H., Kaneko, T., Fukuzawa, H., Tabata, S., Kaplan, A. and Ogawa, T. 2001. Distinct constitutive and low-CO₂-induced CO₂ uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proceedings of the National Academy of Sciences*, 98, 11789-11794.
223. Simkin, A.J., Mcausland, L., Headland, L.R., Lawson, T. and Raines, C.A. 2015. Multigene manipulation of photosynthetic carbon assimilation increases CO₂ fixation and biomass yield in tobacco. *Journal of Experimental Botany*, 66, 4075-4090.
224. Singh, J., Pandey, P., James, D., Chandrasekhar, K., Achary, V., Kaul, T., Tripathy, B.C. and Reddy, M.K. 2014. Enhancing C₃ photosynthesis: an outlook on feasible interventions for crop improvement. *Plant biotechnology journal*, 12, 1217-1230.
225. Singhal, R. and Fernandez, D.E. 2017. Sorting of SEC translocase SCY components to different membranes in chloroplasts. *Journal of experimental botany*, 68, 5029-5043.
226. Söding, J., Biegert, A. and Lupas, A.N. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research*, 33, W244-W248.
227. Spalding, M.H. and Jeffrey, M. 1989. Membrane-associated polypeptides induced in *Chlamydomonas* by limiting CO₂ concentrations. *Plant physiology*, 89, 133-137.
228. Tabita, F.R. 1999. Microbial ribulose 1, 5-bisphosphate carboxylase/oxygenase: a different perspective. *Photosynthesis Research*, 60, 1-28.
229. Tabita, F.R., Hanson, T.E., Satagopan, S., Witte, B.H. and Kreel, N.E. 2008. Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 2629-2640.
230. Tachibana, M., Allen, A.E., Kikutani, S., Endo, Y., Bowler, C. and Matsuda, Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricorutum* and *Thalassiosira pseudonana*. *Photosynthesis Research*, 109, 205-221.
231. Tahami, S. and Chamani, E. 2018. Efficient Protocol for Protoplast Isolation and Plant Regeneration of *Fritillaria imperialis* L.
232. Takahashi, S., Nakamura, T., Sakamizu, M., Woesik, R.V. and Yamasaki, H. 2004. Repair machinery of symbiotic photosynthesis as the primary target of heat stress for reef-building corals. *Plant and Cell Physiology*, 45, 251-255.
233. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, 28, 2731-2739.
234. Tamura, K., Stecher, G., Peterson, D., Filipinski, A. and Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30, 2725-2729.
235. Tanaka, A., Ohno, N., Nakajima, K. and Matsuda, Y. 2016. Light and CO₂/cAMP signal cross talk on the promoter elements of chloroplastic β -carbonic anhydrase genes in the marine diatom *Phaeodactylum tricorutum*. *Plant physiology*, 170, 1105-1116.
236. Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J. and Andersson, I. 2001. First crystal structure of Rubisco from a green alga, *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*, 276, 48159-48164.
237. Tcherkez, G.G., Farquhar, G.D. and Andrews, T.J. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences*, 103, 7246-7251.

238. Uehara, S., Adachi, F., Ito-Inaba, Y. and Inaba, T. 2016. Specific and efficient targeting of cyanobacterial bicarbonate transporters to the inner envelope membrane of chloroplasts in *Arabidopsis*. *Frontiers in plant science*, 7, 16.
239. Ueno, O., Kawano, Y., Wakayama, M. and Takeda, T. 2006. Leaf vascular systems in C₃ and C₄ grasses: a two-dimensional analysis. *Annals of Botany*, 97, 611-621.
240. Van Keulen, G., Dijkhuizen, L. and Meijer, W.G. 2000. Effects of the Calvin Cycle on Nicotinamide Adenine Dinucleotide Concentrations and Redox Balances of *Xanthobacter flavus*. *Journal of bacteriology*, 182, 4637-4639.
241. Van Keulen, G., Ridder, A.N., Dijkhuizen, L. and Meijer, W.G. 2003. Analysis of DNA binding and transcriptional activation by the LysR-type transcriptional regulator CbbR of *Xanthobacter flavus*. *Journal of bacteriology*, 185, 1245-1252.
242. Viklund, H. and Elofsson, A. 2008. OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics*, 24, 1662-1668.
243. Vincze, T., Posfai, J. and Roberts, R.J. 2003. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic acids research*, 31, 3688-3691.
244. Vitha, S., Bene, K., Phillips, J. and Gartland, K. 2012. Histochemical localization of β -glucuronidase (GUS) reporter activity in plant tissues. *Microscopy and Imaging Center, Texas A and M University*.
245. Von Caemmerer, S. and Evans, J.R. 2010. Enhancing C₃ photosynthesis. *Plant Physiology*, 154, 589-592.
246. Von Caemmerer, S., Quick, W.P. and Furbank, R.T. 2012. The development of C₄ rice: current progress and future challenges. *science*, 336, 1671-1672.
247. Von Heijne, G. 1991. Signals for protein import into organelles. *Plant Molecular Biology 2*. Springer.
248. Vriend, G. 1990. WHAT IF: a molecular modeling and drug design program. *Journal of molecular graphics*, 8, 52-56.
249. Wang, Y., Bräutigam, A., Weber, A.P. and Zhu, X.G. 2014. Three distinct biochemical subtypes of C₄ photosynthesis? A modelling analysis. *Journal of experimental botany*, 65, 3567-3578.
250. Wang, Y., Duanmu, D. and Spalding, M.H. 2011. Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: inorganic carbon transport and CO₂ recapture. *Photosynthesis research*, 109, 115-122.
251. Wang, Y., Stessman, D.J. and Spalding, M.H. 2015. The CO₂ concentrating mechanism and photosynthetic carbon assimilation in limiting CO₂: how *Chlamydomonas* works against the gradient. *The Plant Journal*, 82, 429-448.
252. Wenzler, H.C., Mignery, G.A., Fisher, L.M. and Park, W.D. 1989. Analysis of a chimeric class-I patatin-GUS gene in transgenic potato plants: high-level expression in tubers and sucrose-inducible expression in cultured leaf and stem explants. *Plant Molecular Biology*, 12, 41-50.
253. Whitney, S.M., Houtz, R.L. and Alonso, H. 2011. Advancing our understanding and capacity to engineer nature's CO₂-sequestering enzyme, Rubisco. *Plant Physiology*, 155, 27-35.
254. Wicke, S., Schneeweiss, G.M., Müller, K.F. and Quandt, D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant molecular biology*, 76, 273-297.
255. Winter, K. and Smith, J. 1996. Crassulacean acid metabolism: current status and perspectives. *Crassulacean acid metabolism*. Springer.
256. Woodger, F.J., Bryant, D.A. and Price, G.D. 2007. Transcriptional regulation of the CO₂-concentrating mechanism in a euryhaline, coastal marine cyanobacterium, *Synechococcus* sp. strain PCC 7002: role of NdhR/CcmR. *Journal of Bacteriology*, 189, 3335-3347.

257. Xu, J., Zhang, X., Ye, N., Zheng, Z., Mou, S., Dong, M., Xu, D. and Miao, J. 2013. Activities of principal photosynthetic enzymes in green macroalga *Ulva linza*: functional implication of C₄ pathway in CO₂ assimilation. *Science China Life Sciences*, 56, 571-580.
258. Yamano, T., Miura, K. and Fukuzawa, H. 2008. Expression analysis of genes associated with the induction of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant physiology*, 147, 340-354.
259. Yamano, T., Sato, E., Iguchi, H., Fukuda, Y. and Fukuzawa, H. 2015. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, 112, 7315-7320.
260. Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S.-I., Takahashi, Y. and Fukuzawa, H. 2010. Light and low-CO₂-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant and Cell Physiology*, 51, 1453-1468.
261. Yamauchi, Y., Kaniya, Y., Kaneko, Y. and Hihara, Y. 2011. Physiological roles of the cyAbrB transcriptional regulator pair Sll0822 and Sll0359 in *Synechocystis* sp. strain PCC 6803. *Journal of bacteriology*, 193, 3702-3709.
262. Yeates, T.O., Kerfeld, C.A., Heinhorst, S., Cannon, G.C. and Shively, J.M. 2008. Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nature Reviews Microbiology*, 6, 681.
263. Yoshioka, S., Taniguchi, F., Miura, K., Inoue, T., Yamano, T. and Fukuzawa, H. 2004. The novel Myb transcription factor LCR1 regulates the CO₂-responsive gene Cah1, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. *The Plant Cell*, 16, 1466-1477.
264. Zarzycki, J., Axen, S.D., Kinney, J.N. and Kerfeld, C.A. 2012. Cyanobacterial-based approaches to improving photosynthesis in plants. *Journal of experimental botany*, 64, 787-798.
265. Zehr, J.P., Bench, S.R., Carter, B.J., Hewson, I., Niazi, F., Shi, T., Tripp, H.J. and Affourtit, J.P. 2008. Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem II. *science*, 322, 1110-1112.
266. Zhang, P., Battchikova, N., Jansen, T., Appel, J., Ogawa, T. and Aro, E.-M. 2004. Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/Sll1735 in *Synechocystis* sp PCC 6803. *The Plant Cell*, 16, 3326-3340.
267. Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W. and Chua, N.H. 2006. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nature protocols*, 1, 641.
268. Zhaxybayeva, O. 2009. Detection and quantitative assessment of horizontal gene transfer. *Horizontal Gene Transfer*. Springer.

APPENDICES

List of publications

Published articles:

1. **Vandana Tomar**, Gurpreet Kaur Sidhu, Panchsheela Nogia, Rajesh Mehrotra and Sandhya Mehrotra (2017) “Regulatory components of carbon concentrating mechanisms in aquatic unicellular photosynthetic organisms”. *Plant Cell Reports*, 1-18.
2. Panchsheela Nogia, **Vandana Tomar**, Gurpreet Kaur Sidhu, Rajesh Mehrotra and Sandhya Mehrotra (2017) “Relative efficiencies of transient transformation of *Arabidopsis thaliana* and *Nicotiana tabacum* upon heterologous expression of membrane proteins”. *Indian Journal of Biotechnology*, 16: 195-198.
3. Panchsheela Nogia, **Vandana Tomar**, Gurpreet Kaur Sidhu, Rajesh Mehrotra and Sandhya Mehrotra (2016) “Elucidation of an array of nitrate transporter paralogs in *Arabidopsis thaliana* genome”. *Journal of Plant Physiology and Pathology*, 4:3.
4. **Vandana Tomar**, Gurpreet Kaur Sidhu, Panchsheela Nogia, Rajesh Mehrotra and Sandhya Mehrotra (2016) “Role of habitat and Great Oxidation Event on the occurrence of three multi-subunit inorganic carbon uptake systems in cyanobacteria”. *Journal of Genetics*, 95:109-1.

Articles in pipeline:

5. **Vandana Tomar**, Panchsheela Nogia, Gurpreet Kaur Sidhu, Rajesh Mehrotra and Sandhya Mehrotra (2018) “Transformation and Expression of a Na⁺- dependent high affinity bicarbonate transporter SbtA in *Nicotiana* chloroplasts using a minimal chimeric construct”.

List of conference presentations

1. **Tomar, V.**, Nogia, P., Mehrotra, R. and Mehrotra, S. “An insight into the regulatory components of carbon concentrating mechanisms in aquatic unicellular photosynthetic organisms”. In: EMBO Conference on Micro and Metabolic Regulators in Plants, February 01-04, 2017, Thiruvanthapuram, Kerala, India.
2. **Tomar, V.**, Nogia, P., Mehrotra, R. and Mehrotra, S. “Strategies for installing SbtA- a cyanobacterial sodium dependent bicarbonate transporter to chloroplasts of model C₃ plants to enhance CO₂ fixation”. In National Conference of Plant Physiology: Challenges in Crop Physiology Research: From Molecular to Whole Plant (NCP 2016), December 08-10, 2016, University of Agricultural Studies (UAS), GKVK, Bengaluru, Karnataka, India.

3. Nogia, P., **Tomar, V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra, S. “Going the cyanobacterial way to increase higher plant productivity”. In: BITS Conference on Gene and Genome regulation (BCGGR 2016), February 18-20th, 2016, BITS Pilani, Pilani Campus, Pilani, India.
4. **Tomar, V.**, Mehrotra, R. and Mehrotra, S. “Occurrence and arrangement of three multisubunit inorganic carbon transporters in cyanobacteria”. In: The 3rd International Plant Physiology Congress Challenges and Strategies in Plant Biology Research (IPP Congress), Dec 11-14, 2015, JNU, New Delhi, India.
5. **Tomar, V.**, Mehrotra, R., and Mehrotra, S. “Approaches for transplastomic integration of SbtA- a high affinity sodium bicarbonate transporter to C₃ plants for enhanced CO₂ fixation”. FAF-116. International Conference on Emerging Trends in Biotechnology, ICETB 2014. November 6-9, 2014, School of environmental sciences, Jawaharlal Nehru University and Biotech Research Society of India, JNU, New Delhi, India.

Other conference presentations

6. Nogia, P., **Tomar V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra, S. “Transient expression and localization of a cyanobacterial bicarbonate transporter BicA into chloroplast of *Nicotiana benthamiana*”. In: 8th International Conference Photosynthesis and Hydrogen Energy Research for Sustainability (PRS- 2017), October 30-November 3, 2017, School of Life Sciences, University of Hyderabad, Hyderabad, India.
7. Nogia, P., **Tomar, V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra, S. “Targeting chloroplast bicarbonate transporters BicA and SbtA to chloroplast of *Nicotiana* for enhanced CO₂ fixation”. In: Global Conference on Plant Science and Molecular Biology (GPMB-2017), September 11-13, 2017, Valencia, Spain.
8. Nogia, P., **Tomar, V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra, S. “Cyanobacterial CCM genes for improving photosynthetic yield and water use efficiency during heat stress in C₃ plants” An invited lecture in International Workshop to Develop Climate Resilient Cereals, November 03-05, 2016, Punjab Agricultural University, Ludhiana, Punjab, India.
9. Nogia, P., **Tomar, V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra, S. “Going the cyanobacterial way to increase higher plant productivity”. In: BITS Conference on Gene and Genome regulation (BCGGR 2016), February 18-20, 2016, BITS Pilani, Pilani Campus, Pilani, India.
10. Nogia, P., **Tomar, V.**, Sidhu, G. K., Mehrotra, R. and Mehrotra S. “Extensive duplication and subfunctionalization of a nitrate transporter gene in *Arabidopsis*

thaliana: Multiple factors involved in paralog retention”. In Plant Genome Evolution: A Current Opinion Conference, September 06-08, 2015, Amsterdam, Netherlands.

Workshop attended

1. Two days hands-on workshop on “Flow Cytometry” , January 18-19, 2017, organized by Department of Biological Sciences, BITS Pilani, Pilani Campus, India.

Biography of Supervisor - Dr. Sandhya Mehrotra

Dr. Sandhya Mehrotra is currently working as an Assistant Professor in the Department of Biological Sciences, Birla Institute of Technology and Science, KK Birla Goa Campus, Goa, India. She joined the department in January, 2018. Prior to this, she was working as a faculty member in the Department of Biological Sciences, BITS, Pilani, Pilani Campus since November, 2008.



She obtained her Ph.D. degree from Centre for Plant Molecular Biology, National Botanical Research Institute, Lucknow in the year 2003. The dissertation involved studying expression and regulation of genes encoded by chloroplast genome and involved in photosynthesis. Though she analyzed several genes, her main interest was the primary carbon fixing enzyme RuBisCO and gene encoding its large subunit. Thereafter she did a short post-doctoral work in the area of developing anti-osteoporosis agents from Indian medicinal plants from Central Drug Research, Lucknow. In 2005, she moved for a post-doctoral fellowship to Nara Institute of Science and Technology, Nara, Japan, where she studied patterns of evolution of Calvin cycle genes.

During her stay in BITS Pilani, Pilani Campus she had been the principal investigator in projects funded by DST/SERB and UGC as well as co-investigator in Aditya Birla Group of companies funded project. She had also been involved in project related to lipase catalyzed biofuel production from microorganisms and stress responsive designer promoters in plants. Presently, she is a principal investigator in SERB major project on installment of cyanobacterial bicarbonate transporters in C_3 plants to increase crop productivity. As a BITS faculty, she is involved in teaching several courses to first degree and higher degree students. She has supervised many graduate, post-graduate and PhD students.

At the departmental level she has served/been serving as a member of Departmental committee on academics, departmental research committee, library committee, placement faculty, member of student faculty council and the institute level standing committee on student discipline. Her research endeavors at BITS, Pilani and before have resulted into publication of more than 30 research articles in international and national journals of repute and more than 35 presentations in both national and international conferences. She is also a reviewer of several peer-reviewed journals. She has been awarded the Indo-Australia science and technology early career visiting fellowship in year 2012 and has also been selected for Indian National Science Academy's bilateral exchange program in the year 2017.


Biography of Ms. Vandana



Ms. Vandana completed her B.Sc. Biotechnology (with Chemistry, Botany and Zoology) degree from Dolphin (PG) Institute of Bio-Medical and Natural Sciences affiliated to H.N.B. Garhwal Central University, Srinagar (Garhwal), Uttarakhand in 2008. In 2009, she qualified Combined Biotech Entrance Examination (CEEB) conducted by Jawaharlal Nehru University for admission to M.Sc. (Agri.) Biotechnology. In partial fulfillment of Master's degree in Agricultural Biotechnology from Dr. Rajendra Prasad Central Agricultural University, Pusa, India, she successfully defended her thesis entitled "Effect of Media and Genotype on Tissue Culture of Strawberry (*Fragaria x ananassa* Duch.)". During her M.Sc. tenure she was granted DBT fellowship and was also awarded with Gold medal for her academic achievements.

In January, 2013 she joined Department of Biological Sciences, Birla Institute of Technology and Science (BITS) Pilani as a Ph.D. Scholar. During her Ph.D. tenure, she availed DST Inspire Fellowship as a JRF for the period of two years (2013-2015) and as SRF for the period of three years (2015-2018). She has also qualified ARS-NET exam in Agricultural biotechnology in 2013 and GATE in the subject of Biotechnology in 2014. Her doctoral research work envisages the possibility of incorporating cyanobacterial bicarbonate transporter, SbtA into model C₃ plants and studying its expression and localization. She has published research articles in journals of international repute and has presented her research work in national and international conferences. During her Ph.D. tenure, she has also worked as teaching assistant at BITS Pilani for various courses such as General biology, Instrumental Methods of Analysis, Biology laboratory and Plant biotechnology.

Regulatory components of carbon concentrating mechanisms in aquatic unicellular photosynthetic organisms

Vandana Tomar¹ · Gurpreet Kaur Sidhu¹ · Panchsheela Nogia¹ · Rajesh Mehrotra¹ · Sandhya Mehrotra¹ 

Received: 14 June 2017 / Accepted: 31 July 2017
© Springer-Verlag GmbH Germany 2017

Abstract This review provides an insight into the regulation of the carbon concentrating mechanisms (CCMs) in lower organisms like cyanobacteria, proteobacteria, and algae. CCMs evolved as a mechanism to concentrate CO₂ at the site of primary carboxylating enzyme Ribulose-1, 5-bisphosphate carboxylase oxygenase (Rubisco), so that the enzyme could overcome its affinity towards O₂ which leads to wasteful processes like photorespiration. A diverse set of CCMs exist in nature, i.e., carboxysomes in cyanobacteria and proteobacteria; pyrenoids in algae and diatoms, the C₄ system, and Crassulacean acid metabolism in higher plants. Prime regulators of CCM in most of the photosynthetic autotrophs belong to the LysR family of transcriptional regulators, which regulate the activity of the components of CCM depending upon the ambient CO₂ concentrations. Major targets of these regulators are carbonic anhydrase and inorganic carbon uptake systems (CO₂ and HCO₃⁻ transporters) whose activities are modulated either at transcriptional level or by changes in the levels of their co-regulatory metabolites. The article provides information on the localization of the CCM components as well as their function and participation in the development of an efficient CCM. Signal transduction cascades leading to activation/inactivation of inducible CCM components on

perception of low/high CO₂ stimuli have also been brought into picture. A detailed study of the regulatory components can aid in identifying the unraveled aspects of these mechanisms and hence provide information on key molecules that need to be explored to further provide a clear understanding of the mechanism under study.

Keywords Carbon concentrating mechanisms · Cyanobacteria · CcmR · CmpR · CbbR · Proteobacteria · Algae · Diatoms

Introduction

Carbon dioxide concentrating mechanisms (CCMs) form an integral component of carbon acquisition in photosynthetic organisms (Badger and Price 2003). The major driving force for the evolution of CCM was the evolutionary pressures created at the dawn of the great oxidation event (GOE) about 2.5 billion years ago (Tomar et al. 2016). Due to the altered environmental conditions, organisms had to either make Rubisco more efficient (Tcherkez et al. 2006; Rae et al. 2013) or had to develop mechanisms to elevate the level of CO₂ near the active site of Rubisco so as to avoid wasteful processes like photorespiration and fix more CO₂ into sugars. The values of specificity factor ($\Omega/(S_{c/o})/\tau$) which is the ratio of catalytic efficiency of carboxylation (V_c/K_c) to that of oxygenation (V_o/K_o) reveal that in organisms such as higher plants, which are devoid of carbon concentrating mechanisms, Rubisco with higher specificity for CO₂ and low carboxylation rate has been selected, whereas organisms such as cyanobacteria and algae, which are capable of concentrating carbon, have a Rubisco with lower specificity but high turnover rate (Andersson 2008). These variations suggest

Communicated by Neal Stewart.

Electronic supplementary material The online version of this article (doi:10.1007/s00299-017-2191-3) contains supplementary material, which is available to authorized users.

✉ Sandhya Mehrotra
sandhyamehrotrabits@gmail.com

¹ Department of Biological Sciences, Birla Institute of Technology and Science, Pilani 333031, Rajasthan, India

Relative efficiencies of transient transformation of *Arabidopsis thaliana* and *Nicotiana tabacum* upon heterologous expression of membrane proteins

Panchsheela Nogia, Vandana Tomar, Gurpreet Kaur Sidhu, Rajesh Mehrotra and Sandhya Mehrotra*

Department of Biological Sciences, Birla Institute of Technology and Science, Pilani 333031, India

A cyanobacterial gene encoding a protein with multiple transmembrane helices, when transformed into *Arabidopsis thaliana* and *Nicotiana tabacum*, demonstrated several differences in the relative responses of both the host species, particularly in efficiencies of regeneration of transformed tissue from leaf explants. The gene of interest (GOI) encoded an integral membrane protein containing 12 transmembrane helices and functioned as a sodium dependent inorganic ion transporter in a native cyanobacterium. *Agrobacterium tumefaciens*-mediated transformation was carried out by coculture method. The leaves of *N. tabacum* successfully regenerated into transformed callus tissue and gave positive results upon GUS assay; while the *A. thaliana* leaves did not give any transformed callus and, a few days post agro-infection, all the explants displayed tissue necrosis. The results obtained point towards *A. thaliana*'s recalcitrance to *Agrobacterium*-mediated transformation by coculture method and the difficulty of expressing hydrophobic membrane proteins in heterologous systems.

Keywords: *Agrobacterium tumefaciens*, *Arabidopsis thaliana*, hydrophobic membrane protein, *Nicotiana tabacum*, transformation efficiency

Introduction

Transient expression studies provide a base for analyzing success rate of transgene expression in a heterologous host system. *Arabidopsis thaliana* and a few *Nicotiana* species (*N. tabacum*, *N. benthamiana* & *N. attenuata*) are widely used by researchers as model host plants to study protein expression following various transformation methods. *A. thaliana*, in terms of its competency with different transformation methods, has been well studied using a variety of explants and its performance has been investigated in various transient transformation experiments, such as, young seedlings transformed by vacuum infiltration¹, leaves transformed by agroinfiltration² and coculture³ method. However, nowadays, these methods are rarely used due to the poor demonstration of *A. thaliana* plants in transient expression studies. Instead, *A. thaliana* is now most widely used to develop stable plants by floral dip method⁴, which has proved to be a highly successful and efficient transformation method. Several *Nicotiana* species, on the other hand, are relatively easy to transform by *Agrobacterium*-mediated coculture, agroinfiltration and particle bombardment methods, and are found to be effective and resilient for both transient as well as stable protein expression experiments⁵.

The biochemical nature of proteins also influences their expression in heterologous systems. Of various categories, the membrane proteins being of hydrophobic nature offer the greatest challenge. The membrane proteins play various important roles in the living cells and function as transporters, receptors, enzymes *etc.* In most organisms, they constitute about 20-30% to the total protein. High-level expression of membrane proteins remains a challenge to this day because of their low natural abundance and tendency to form aggregates⁶. These features also make them difficult targets for structural characterization. For optimized expression of hydrophobic proteins, various host systems are tested but it is still a major challenge. Possible reasons behind the lack of success can be that these proteins may have some lethal effects on the host or have undergone alterations at the molecular level upon expression. In addition, the host machinery may or may not allow proper folding and activation of the protein.

Researchers have been using genes from different organisms and have tried their expression in various bacterial, plant and other eukaryotic heterologous host systems but the success rate is always variable with respect to both the type and complexity of host and protein⁷. Specifically, in case of heterologous systems, we need to optimize all the parameters related to host growth and transformation according to the nature of the protein. Apart from the discussed aspects so far,

*Author for correspondence
sandhyamehrotrabits@gmail.com



Elucidation of an Array of Nitrate Transporter Paralogs in *Arabidopsis Thaliana* Genome

Panchsheela Nogia, Vandana Tomar, Gurpreet Kaur Sidhu, Rajesh Mehrotra and Sandhya Mehrotra*

Abstract

Plants derive gene paralogs after both small-scale duplication (SSD) and whole genome duplication (WGD). Plants are known to retain duplicated genes and enhance their gene repertoire which intensively contributes to the evolution of their genomes. The general view is that such paralogs are both selected and preserved in evolution because they express variable levels of proteins in different temporal and spatial patterns. In *Arabidopsis thaliana*, paralogs involved in signalling and transcriptional regulation are very often retained after gene and genome duplication events. We analysed duplication events of a primitive nitrate transporter gene NTR across the genome of *Arabidopsis thaliana*. Majority of the paralogs have acquired the function of both high, low and also dual affinity nitrate transporters, some being regulated by nitrate levels themselves and involved in a variety of plant functions such as embryonic development, repression of lateral root initiation, stomatal opening, xylem and phloem transport and wound response etc. For the present analysis, 19 nitrate transporter protein sequences were retrieved from TAIR database and analysed by a web service tool PINDA as well as manually which gave analogous results. Several transporters were found to be acting as paralogs as close sequence homology was observed amongst them. The results point towards the idea that sub-functionalization of nitrate transporter proteins in higher plants happened as a result of a series of gene duplication events.

Keywords

Gene paralogs; Transcriptional regulation; *Arabidopsis thaliana*; Embryonic development

Introduction

Plants derive gene duplicates often referred to as paralogs which result from small scale duplication (SSD) and whole genome duplication (WGD) events and lead to changes at gene level or whole genome level respectively. It has been found that SSDs are responsible for major evolutionary changes in comparison to WGD [1]. Gene duplicates can be obtained from retro-position, unequal crossing over or chromosomal/genome duplication. Unequal crossing over leads to tandem gene duplication while in retro-position mRNA is being converted into cDNA which is then inserted into the genome. Chromosomal or genome duplication mainly results from the process

called as non-disjunction and this phenomenon is more common in plants than in animals [2]. Gene and genome duplications provide a source of genetic material for mutation, drift, and selection to act upon, making new evolutionary opportunities possible. Genome duplication hence has been credited as a dominant factor in the evolution of complexity and diversity [2].

The evolutionary fate of duplicated genes is difficult to predict because mechanism behind this is not well understood [3] although it can either be stably maintained or lost in the population. For duplicated genes four outcomes [4] are possible viz. (1) pseudogenization- one copy of gene will become nonfunctional [5] (2) sub-functionalization- function partitioning in both the gene copies will occur [6] but parental expression pattern will be maintained, (3) conservation of gene function- gene product will be increased by keeping both the gene copies active and (4) neo-functionalization- one gene copy will gain a new biological function [7]. The effect of duplication scale on functional evolution has not yet been explored, probably due to the lack of global knowledge of protein function and different times of duplication events.

Several studies have shown that gene duplicates are frequently preserved by sub-functionalization and one such example is the nitrate transporters of *Arabidopsis thaliana* which have retained their function but diversified to some extent by performing low and high affinity transport of nitrate ions. Another probable reason for the same would be various degenerative mutations in the duplicated paralogs which lead to sub-functionalization event [8]. It has often been found that duplicated gene pairs whose functions are being preserved by sub-functionalization may show tissue specific expression [8] of parent gene which would be termed as divergence of gene function in terms of spatial expression. Furthermore, there can be temporal regulation of duplicated genes as they express at different time points of various cellular stages viz. globular stage, cotyledon stage, embryo stage etc [9].

Nitrate is a crucial element for plant growth and development because it acts as an important macronutrient. After its uptake by the root cells of plants, nitrate undergoes processes of its assimilation. There is a huge variation in the concentration of nitrate in soil system and therefore two types of nitrate transporters have developed viz. the low affinity transporters (when nitrate concentration is above 1 mM) and high affinity transporters (when nitrate concentration is between 1 μ M and 1 mM) [10]. Genes of these transporters are generally termed as NRT1/NPF and NRT2 which encodes inducible and constitutively expressed low affinity transporters and inducible high affinity transporters respectively [11,12].

Here, in this article we will try to elucidate the role of gene duplication in the evolution of high and low affinity nitrate transporters and describe the duplicates identified herein.

Material and Methods

Two methods were followed for detection of nitrate transporter duplicates in *Arabidopsis* genome viz. analysis by manual method and by online tool PINDA.

*Corresponding author: Sandhya Mehrotra, Department of Biological Sciences, Birla Institute of Technology and Science, Pilani, Rajasthan-333031, India, Tel: +91-9772975612; +910159615676; Fax: +91-01596-244183; E-mail: sandhya@pilani.bits-pilani.ac.in; sandhyamehrotrabits@gmail.com

Received: June 04, 2015 Accepted: July 15, 2016 Published: July 22, 2016

Role of habitat and great oxidation event on the occurrence of three multisubunit inorganic carbon-uptake systems in cyanobacteria

VANDANA TOMAR, GURPREET KAUR SIDHU, PANCHSHEELA NOGIA, RAJESH MEHROTRA
and SANDHYA MEHROTRA* 

Department of Biological Sciences, Birla Institute of Technology and Science, Pilani 333 031, India

Abstract

The oxygenase reaction catalyzed by RuBisCO became an issue only after the evolution of the oxygenic photosynthesis in cyanobacteria. Several strategies were developed by autotrophic organisms as an evolutionary response to increase oxygen levels to help RuBisCO maximize its net carboxylation rate. One of the crucial advancements in this context was the development of more efficient inorganic carbon transporters which could help in increasing the influx of inorganic carbon (Ci) at the site of CO₂ fixation. We conducted a survey to find out the genes coding for cyanobacterial Ci transporters in 40 cyanobacterial phyla with respect to transporters present in *Gloeobacter violaceus* PCC 7421, an early-diverging cyanobacterium. An attempt was also made to correlate the prevalence of the kind of transporter present in the species with its habitat. Basically, two types of cyanobacterial inorganic carbon transporters exist, i.e. bicarbonate transporters and CO₂-uptake systems. The transporters also show variation in context to their structure as some exist as single subunit proteins (BicA and SbtA), while others exist as multisubunit proteins (namely BCT1, Ndh₃ and Ndh₄). The phylogeny and distribution of the former have been extensively studied and the present analysis provides an insight into the latter ones. The *in silico* analysis of the genes under study revealed that their distribution was greatly influenced by the habitat and major environmental changes such as the great oxidation event (GOE) in the course of their evolution.

[Tomar V., Sidhu G. K., Nogia P., Mehrotra R. and Mehrotra S. 2016 Role of habitat and great oxidation event on the occurrence of three multisubunit inorganic carbon-uptake systems in cyanobacteria. *J. Genet.* **95**, 109–118]

Introduction

Cyanobacteria or blue green algae are one of the most diverse groups of organisms which colonize various niches including freshwater (rivers, ponds and lakes), polar caps, hot springs, alkaline, estuarine, open as well as saline oceans. During the ancient environmental adversities, these organisms along with certain algal species developed active systems known as carbon-concentrating mechanisms (CCM) to concentrate CO₂ at the site of RuBisCO activity which in turn led to a positive impact on photosynthetic ability of these organisms (Caemmerer and Evans 2010). This strategy was also adopted by the land plants and they developed certain anatomical features (Kranz anatomy) to increase the carboxylation activity of RuBisCO, thus reducing the resultant photorespiratory losses (Raven *et al.* 2008).

The CCM present in cyanobacteria are highly efficient and can accumulate CO₂ around RuBisCO by a factor of

1000-fold above ambient levels (Price *et al.* 2011). CCMs in cyanobacteria and proteobacteria as a whole include (i) RuBisCO and carbonic anhydrases (CA) enclosed in micro-compartments known as carboxysomes and (ii) Ci transporters, which regulate the CO₂/HCO₃⁻ influx and efflux at the site of RuBisCO activity and hence lead to a marked elevation in the CO₂ concentration in the vicinity of RuBisCO.

Cyanobacteria are dependent on active accumulation of Ci to achieve a satisfactory rate of CO₂ fixation and growth (Badger *et al.* 2002). The efficacy of any CCM relies on the ability to minimize the loss of CO₂ from the CO₂ elevation zone or the dissolved Ci accumulation zone (Price *et al.* 2007). It can be accomplished in four ways. First, the accumulation of bicarbonate instead of CO₂ reduces the chances of Ci leakage, since the former is less permeable through the plasma membrane as compared to the latter. Secondly, the absence of CA activity in the cytosol minimizes leakage due to wasteful conversion to CO₂ and subsequent diffusion back to the external medium. Thirdly, the carboxysome protein shell is proposed to have specifically

*For correspondence. E-mail: sandhyamehrotrabits@gmail.com.

Keywords. inorganic carbon transporters; *Gloeobacter violaceus* PCC 7421; great oxidation event; cyanobacteria; evolution; carbon-concentrating mechanism.