Cloning and Characterization of RNA Binding Protein All 4377 from Anabeana PCC 7120

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

Master of Science
in
Plant Biotechnology

Submitted by

Tushita Rai

Supervisor	**Co-Supervisor**
Prof. L.C. Rai	***Dr. Rajesh Kumar***

DEPARTMENT OF GENETICS AND PLANT BREEDING
INSTITUTE OF AGRICULTURAL SCIENCES
BANARAS HINDU UNIVERSITY
VARANASI – 221005

ID No. 19430PLB024 **2021** **Enrollment no. 419073**

काशी हिन्दू विश्वविद्यालय BANARAS HINDU UNIVERSITY

कृषि विज्ञान संस्थान INSTITUTE OF AGRICULTURAL SCIENCES

Prof. L.C Rai
Distinguished Professor
E-mail-lcrbhu15@gmail.com

Department of Botany
Institute of Sciences
Varanasi- 221005

Ref.No…………….

Date………..

CERTIFICATE

To,
The Registrar
Banaras Hindu University
Varanasi-221005 (India)

Through The Head,
Department of Genetics and Plant Breeding Institute
of Agricultural Sciences
Banaras Hindu University, Varanasi-221005 (India)

Dear Sir,

I have great pleasure in forwarding the thesis entitled **"Cloning and Characterization of RNA Binding Protein All 4377 from Anabaena PCC 7120"** submitted by Tushita Rai**, ID. No. 19430PLB024** in partial fulfillment of the requirements for the degree of **Master of Science** in **Plant Biotechnology** from Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi.

I certify that the entire scheme of investigation reported herein, was planned and carried out by the candidate under my guidance to the best of my knowledge and belief; the data presented in the thesis are genuine and original. No part of the work has been submitted for any degree/or distinction.

Thanking you,

Yours sincerely

Prof. B. Sinha	Dr. Ravindra Prasad	Prof. L.C. Rai
Head	Course coordinator	Supervisor

Cloning and Characterization of RNA Binding Protein ALL 4377 from *Anabeana PCC 7120*

By

TUSHITA RAI

Thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
IN
PLANT BIOTECHNOLOGY

DEPARTMENT OF GENETICS AND PLANT BREEDING
INSTITUTE OF AGRICULTURAL SCIENCES
BANARAS HINDU UNIVERSITY
VARANASI - 221 005

ID. No.19430PLB024 **2021** **Enrolment No. – 419074**

APPROVED BY ADVISORY COMMITTEE

Advisor : **Prof. L.C Rai**
Distinguished professor
Centre of Advanced studies
Department of Botany
Institute of Sciences
Banaras Hindu University
Varanasi, India

Co-advisor : **Dr. Rajesh Kumar**
Assistant professor (Plant biotechnology)
Department of Genetics and Plant breeding
Institute of Agricultural sciences
Rajiv Gandhi South Campus , BHU ,
Barakchha, Mirzapur, India

Member : **Dr. Ashok Kumar**
Assistant Professor, (Plant biotechnology)
Department of Genetics and Plant breeding,
Institute of Agricultural Sciences,
Rajiv Gandhi South Campus, BHU, Barkachha,
Mirzapur, India

Member : **Dr. Vishal Srivashtav**
Assistant Professor (Plant Biotechnology)
Department of Genetics and Plant breeding,
Institute of Agricultural Sciences,
Rajiv Gandhi South Campus, BHU,
Barkachha, Mirzapur, India

EXTERNAL EXAMINER :

Dedicated

to

my beloved

Daddy.

Acknowledgements

I would first like to thank the Almighty for blessing me with sound health and providing me with strength and patience to complete my four months dissertation successfully.

I sincerely thank **Prof. N.K Dubey,** Head, Department of Botany, Institute of Science, Banaras Hindu University, Uttar Pradesh for giving me the opportunity to work in the department.

I would like to express my deepest gratitude to my supervisor, Emeritus Professor, **Prof. L.C. Rai,** Department of Botany, Institute of Science, B.H.U. I have been extremely fortunate to have a guide, who consistently provided me with the liberty to explore my potential and for providing me such a good lab environment.

I am really grateful to my co-supervisor **Dr. Rajesh Kumar**. Assistant Professor, Department of Genetics and Plant Breeding, Banaras Hindu University for his irreplaceable suggestions, guidance and kind attitude leading my path to achieve the destination during the entire move despite of his heavy schedule of work.

I am also thankful to **Dr. Ravindra Prasad**, Coordinator, Department of plant biotechnology, Banaras Hindu University for allowing me to work on this dissertation report.

My sincere thanks to my advisory committee: **Dr. Vishal Srivashtava** (Assistant professor) and **Dr. Ashok Kumar** (Assistant Professor), for their encouragement, insightful comments and hard questions.

I am greatly thankful to my senior ***Miss Sonam Srivastava*** *for guiding me throughout the journey and for their passionate participation and devoting their precious time in the project work, without whom this endeavor of mine would not have been successfully conducted.*

I would like to give special thanks to ***Ruchi Rai, Shilpi Singh, Vigya kesari, Krishna Rai, Alka Raj ,Shradha Pandey, Pooja Pandey*** *for their support and guidance. They have helped me learning many new things and and non-teaching staff . Mr. Vinod Yadav for their valuable advice, pertinent suggestion and ever willing help which definitely helped me in my research work.*

Last but not the least, I express my profound gratitude to my parents and all my other friends for providing me with unfailing support and continuous encouragement. This accomplishment would not have been possible without them.

I just hope that I can use the piece of work I learned here for the betterment of society, as 'One can never pay in gratitude, but can pay somewhere else in life'.

Thank you

Tushita Rai

Contents

LIST OF ABBREVIATIONS

CAN	Acetonitrile
APS	Ammonium per sulfate
ATP	Adenosine 5’-triphosphate
Bp	Base Pair
Dntp	Deoxynucleotide triphosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra acetic acid
Gm	Gram
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
Kb	Kilobase
KDa	Kilodalton
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PMSF	Phenyl Methyl Sulfonyl Fluoride
RNA	Ribonucleic acid
RT-PCR	Reverse Transcription Polymerase Chain Reaction
Tris	Tris (hydroxymethyl) aminomethane

LIST OF TABLES

LIST OF FIGURES

Chapter I

GENERAL INTRODUCTION

Cyanobacteria thrive in every illuminated aquatic environment and possess the unifying property of performing plant type oxygenic photosynthesis. Phylogenetically, these cyanophytans comprises a coherent assembly of evolutionarily primitive, morphologically divergent, and ecologically significant plant-type bacteria. This oxygen -evolving conservative domain of bacteria originated 3.5 billion years ago and is believed to be responsible for the oxygenation of Earth’s atmosphere. During the course of evolution, they are believed to be the common putative progenitor of plastids in plants (Douglas 1994, Gray 1989) therefore act as an ancestor of plant chloroplasts.

Many cyanobacteria love to dwells in agricultural soils and are responsible for biological nitrogen fixation, phosphate solubilization, and mineral release along with improvement in the fertility of soil and escalation in crop productivity. The composition of lipid in plasma and thylakoid membranes of cyanobacteria is similar with the chloroplast of higher plant. Consequently, these bluish-green bacteria emerged as a powerful and impressive model for deciphering the studies of the molecular mechanisms of the responses and acclimation of plants to several types of environmental stress. From their early history, cyanobacteria help in adapting a wide variety of habitats from aquatic to terrestrial ecological niches up to the present day. These wide arrays of distribution enlighten us with their ability to resist and adapt in variety of stress, such as nutrient deprivation, luminosity along with osmotic, thermal, and oxidative stresses.

As organisms which evolved million years ago (Brock, 1973) and have effectively survived against diverse variety of environmental stresses, cyanobacteria provides outstanding opportunities for researchers to study diverse stress response (Pentii, 1892). The survival of all living organisms depends on the adaptation to unavoidable environmental stress, under extreme abiotic stress conditions, organic molecules such as lipids, proteins, and nucleic acids are prone to damage or degradation (Miranda, 2011).

Environmental stresses regulate a plethora of physiological and metabolism related activities in living organisms. Cyanobacteria respond to different physiochemical environmental changes with their genetically programmed physiological adaptations which include the inhibition or extinction of some proteins and induction or enhancement of other proteins.

Cyanobacteria do not possess any innate mechanism to govern their own fluctuating temperature but have advanced themselves with diverse strategies to adapt and combat themselves in varying environmental temperatures. However, adaption to high temperature has been analyzed in detail; we have limited knowledge regarding cold- shock response which is certainly a serious problem in the field of agriculture. In order to nullify the detrimental effects of cold injury, I want to determine different adaptive approaches involved during cold stress using famous photosynthetic bacteria as our model organism.

Conventionally low temperature induces the formation of ice crystals which eventually results in reducing the functions of several enzymes and consequently reduces physiological and biochemical activities of the cell. Hence, cyanobacteria acclimate to cold stress by varying the physiological, biochemical, and molecular activities which totally relies on the duration and intensity of temperature. This domain of bacteria adapts to cold stress by (1) Alteration in fatty-acids of cell membrane (2) aggregation of osmo-protectants (3) accumulation of stress-responsive proteins. This different type of adaptation helps to protect the cyanobacteria from the detrimental effects of cold injury.

During the course of my study, I found an RNA-binding protein gene *rbp*

A. The level of the transcript of *RBP* was drastically increased when the growth temperature was shifted from 38 to 22°C. The Rbp protein which was prepared by expression in *Escherichia coli* cells had a strong affinity with RNA isolated from *Anabaena* cells. This small RBP protein of about 100 amino acid sequence is characterized by the presence of a single RRM (RNA-recognition motif) with a short C terminal glycine-rich domain and N terminal acidic domain. RBPs play a major role in determining RNA metabolism, such as splicing, modification, maintenance of stability, and translation. Basically, the involvement of cold- induced genes enhances the efficiency of transcription and translation to remunerate for the decrease in the efficiency of these processes that would otherwise occur at the lower temperature. (Sato, 1994). The relationship between cold stress responses is neither surprising nor unique.

Although different pathways have been studied on a variety of cyanobacterial species, there is still a lacuna regarding the molecular mechanism underlying RNA Binding Protein during cold stress in *Anabaena* sp. PCC7120. Thus, there is a strong need to identify and characterize the novel RBP genes. Our study will definitely provide a new direction in cyanobacterial biological research to understand the role and uncover the hidden wealth of cold stress coping RBP genes for their future application to develop cold-resistant transgenic plants. Certainly, stresses hamper the survival and cell cycle of a living organism, however, most organisms have developed attractive mechanisms to withstand various environmental stress conditions. In response to environmental changes of various kinds, the ability to trigger and coordinate suitable adaptive mechanisms depends on the ability of these bacteria to rapidly sense the physical stimuli present and to appropriately transduce the signals perceived into gene expression and modulation processes.

In addition to the above conclusion, the underlying mechanism by which cyanobacteria and plant responds towards environmental signals and transmit these signals to cellular machinery to activate adaptive responses is of immense importance to molecular biology (Xiong et al., 2002).The ability of an organism to survive in stressful situations is usually the result of possession and selective expression of certain novel genes. Information on such genes and their regulation is an important pre-requisite for devising future strategies aimed at building stress tolerant bacteria and crop plants (Bohnert and Jensen., 1996). The gene *all 4377* of *Anabaena* sp. PCC7120 is a protein coding gene which belongs to the RRM (RNA Recognition Motif) super family in cyanobase (http://genome.Microbedb.jp/cyanobase). Anabaena PCC 7120 is a conventional tool for genetic manipulation and accepted as a model microorganism for biological and biochemical studies of nitrogen fixation pattern formation and cell differentiation as its genome is fully sequenced.

The defense mechanism via RBP protein is used by an array of living organism as a method of protection against cold stresses generated in cell either in vivo or as a result of seasonal change in environment. Thus to investigate the RRM family member, All 4377 from *Anabaena* sp. PCC7120 and their role in cold stress tolerance in *E. coli*, present study was undertaken with the following objectives:-

a) Insilco analysis of All 4377(sequence, 3-D structure and folding analysis)
b) Cloning of All 4377
c) Heterologous expression and Purification of recombinant protein All 4377 from *E. coli*
d) Transcript analysis of All4377 under cold stress
e) Ectopic expression of A*ll4377* in *E. coli* to analysis stress tolerance capacity towards different cold stresses through change in temperature.

◆

Chapter II

REVIEW OF LITERATURE

2. 1 *Cyanobacteria* (An overview)

Cyanobacteria pertain to be a peculiar member of the bacterial clade, which consists of gram-negative phototrophic bacteria. They are probably the first large successful diverse group of organisms to evolve on the earth with the ability to carry out oxygenic photosynthesis. This unique domain of bacteria is monophyletic, primitive, and exhibits cosmopolitan distribution in almost all possible illuminating environments. The above-mentioned unicellular prokaryotic bacteria are characterized by their ability to form phycobilin and phycocyanin pigment. Accumulation of these pigments leads to the bluish-green color of the cyanobacteria, and hence commonly known as blue-green algae.

Figure 1: **Image of cyanobacteria**

Oxygenic photosynthesis being the most fundamental and unifying metabolic processes that have evolved on Earth as it has paved the way for complex life forms. In addition to this, this process played a vital role in the development of biological and geological processes that led to the emergence of oxygen. Production and liberation of oxygen as a byproduct of photosynthesis aids the conversion of the early reducing atmosphere into an oxidizing one and eventually responsible for the “Great Oxygenation Event” and the “Rusting of Earth” (Bacteria, 2016). The history of Earth truly represents how cyanobacteria have triggered key evolutionary events, because of their ability to produce oxygen, endosymbiotic existence, and fixing free N_2 and CO_2.

According to many geologists and geochemists, cyanobacteria are the oldest photosynthetic organism with a long evolutionary history on earth, which originated approximately 2.6-3.5 billion years ago. As a result of abundance in their fossil, the Proterozoic era (2500-570Ma) was regarded as the “Age of Cyanobacteria” (Bekker et al., 2004; Lyons et al., 2014). Their wet biomass is composed of about 10^{15} gm. (i.e., about a thousand million tons) or global biomass of $3x10^{14}$ gm. of carbon (Garcia-Pichel et al 2003) and is significantly responsible for half of the global net primary productivity.

The existence of novel genome sequences and phylogenetic analysis offers subterranean insights into the evolution of morphology and habitat within the members of cyanophyce. Eventually, their resolution becomes apparent in resolving deep–branching relationships within cyanobacterial members (Blank & Sánchez-Baracaldo, 2010; Shih et al., 2013; Schirrmeister et al., 2015).

The morphology of cyanobacteria is highly diversified because of its existence in variety of different forms from unicellular, planktonic or benthic to filamentous, and colonial or coccoid ones. The most widespread occurring photosynthetic organism can thrive in a wide range of ecological habitats, ranging from marine, limnetic freshwater, to terrestrial environments, and often survive as pioneer species in almost all habitats such as drylands, bare rock , glaciers, and the open ocean (Castenholz, 2001; Blank & Sánchez-Baracaldo, 2010). Moreover, these blue-green algae subsist as symbiotic associations with higher plants, protists, fungi, and algae (Singh, 2014). Biochemical investigation of cyanobacterial's cell wall fatty acid composition showed a correlation with their morphology. This relativity allows us to distinguish between Anabaena and Nostoc (Holton, 1968). Additionally, these photosynthetic prokaryote can withstand a wide range of temperatures in almost all environments ranging from Antarctica (-20^{o} C) to hot springs (70^{o} C) (Psenner, Sattler ; Ward,1998).

The most unifying property of cyanobacteria is the potential to perform various types of metabolism and amazing ability to switch rapidly between different vogues of metabolic processes. Their metabolism makes them distinct and leading players in the biosphere because of their impact on the global carbon and nitrogen cycles. They are significantly responsible for half of the global net primary productivity and the ability to fix nitrogen with the help of heterocyst held to be the chief source of combined nitrogen in the marine environment.

The typical organization of cyanobacteria resembles the prokaryotic cell, characterized by the presence of an elaborate system of eukaryotic internal membranes which are directly involved in respiratory and photosynthetic electron transport. The most prevalent and universal storage products of these bacteria are polyphosphate (phosphorus storage compound,), cyanophycin or phycobillin protein pigment (nitrogen storage products), and glycogen (storage product of carbon and energy). The most general and popular method to perform water-oxidizing photosynthesis by using Z-scheme with the help of Photosystem I and Photosystem II but under oxygen deficit they prefer Photosystem I only over the combination of photosystems , mostly associated with purple bacteria. Surprisingly, photosynthesis and respiration rely on the same electron transport machinery.

Cyanobacteria constitute a fundamental component of paddy field microflora in tropical countries, play a vital role in the nitrogen economy of the soil (Singh, 1962), and improve crop productivity. The symbiotic association of Anabaena with Azolla exclusively presents in the paddy fields and increases the production by fivefold greater. It fixes about 20-30 kg nitrogen per hectare in rice fields (Karthikeyan, 2009). According to the survey carried out on 2213 soils samples of Indian paddy fields, approximately only 33% contains nitrogen-fixing cyanobacteria (Venkataraman, 1981). They are the most natural colonizing group of rice roots (Singh, 2014). In the south of 35˚ N latitude, and about 12% of the total algal flora are the nitrogen-fixing cyanobacteria in the south of 35˚ N latitude whereas, in the north, it constitutes only about 2%.

The ubiquitous nature of cyanobacteria helps them to expose to various environmental conditions. Their growth depends upon the nutrient, PH, temperature, osmolarity of the medium, and any alterations in these conditions may affect the growth (Borbély et al 1990), leading to physiological depression (Murata et al 2005) with change in gene expression pattern (Borbély teal 1990), sometimes proving fatal to the organism (Murata et al 2005).

Their unparalleled and extraordinary ability to adapt and withstand harsh environmental changes by evolving a myriad of bio-molecules and secondary metabolites with the intention of their growth and survival is a matter of appreciation.

These unique organisms inhabited with the unique ability of plant-like photosynthesis and surprisingly have higher photosynthesis and biomass production rates than plants. They have the ability to convert 3–9% of the solar energy into biomass over the land plants which can only convert ≤0.25–3% and they require less land area for cultivation than the terrestrial plant which eventually diminish the competition for other crops plants. Cyanobacteria are prokaryotic bacteria with relatively simple genetic backgrounds that facilitate manipulation. In addition to the above extractable usable products, the left-over residual biomasses of cynobacteria can be used as animal fodder or converted into organic fertilizer.

Cyanobacteria act as efficient bio-fertilizers and have various applications in the field of pharmaceuticals also. Additionally, they are distinguishable due to the production of secondary metabolites having antifungal, antibiotic, anti-carcinogenic, antiviral, anti-malarial, anticoagulant, anti-inflammatory, and antitumor properties (Venkataraman et al 1981). A photoautotrophic property makes them suitable for the production of bio-fuels (Dismukes et al 2008).

Currently, cyanobacteria are promising organisms in the biotechnological industry because they act as inexpensive microbial factories that aid in the carbon trapping and storing as well as facilitate the production of secondary metabolites and bio-fuels due to their simple nutritional requirements, their metabolic plasticity, and the powerful genetics of some model strains.

In consideration of above-mentioned inherent facts of cyanobacteria reveals as the most attractive and novel model for biotechnological studies. Recent discovery in genetic engineering technologies paved their investigation to analyze the full potential of cyanobacteria with the help of more than cyanobacterial genome sequences. Their possible and credible use of cyanobacteria in the field of agriculture , pharmaceuticals, effluent bacteria, bio-fuels, bioremediation and various secondary metabolites such as vitamins ,toxins and various enzymes.

These photoautotrophs represents as an excellent and magnificent model for unraveling the mysteric puzzle associated with different metabolical, biochemical and physiological processes which is similar to higher plants. Due to their rapid and potential utilization in different biotechnological and commercial fields of biology, there is an urgent requisite requirement to genetically modify these tiny bacteria in such a manner that they can endure and acclimatize in variegated environmental stress conditions. Adaptation to different stress depends on the precise control of genes that produce novel proteins and enzymes.

Thus, the most utmost and important objective is to study various biotic and abiotic stress responses which includes post-genomic investigations and revolves around diverse proteomic changes utilizing advanced proteomics, synthetic and structural biology related function. Recent advent and developments in this field will help in deciphering and manipulation of various physiological and metabolical pathways genes of cyanobacteria with the aim of understanding the stress-induced proteomic and genomic defenses actively.

2.2 Anabeana PCC 7120 (*model organism*)

Anaebaena PCC 7120 is a filamentous cyanobacterium. It contains heterocyst, which helps in fixing nitrogen molecules by providing anaerobic condition in the oxygen containing environment with the help of the dinitrogenase enzyme. It provides nitrogen in the form of glutamine and receives carbohydrate formed as a result of photosynthesis in exchange of that. Generally, in presence of nitrogen, most of the cells are vegetative in nature but during nitrogen limiting condition, about 10% of the cells undergo terminal differentiation to form heterocyst (Wei et al 1994). It is used as an important model organism to study the genetics, physiology and mechanism of nitrogen fixation.

Various gene manipulating techniques are available for this along with a conjugation system (Wolket al 1994). The chromosomal size of *Anabaena* PCC7120 is approximately 6.42 Mb (Bancroft etal 1989) consisting a chromosome and six plasmids (α, β, γ, δ, ε, δ). Since, *Anabaena* PCC 7120, has fully sequenced genome and closely resembles higher plants photosynthesis, survives under a wide array of abiotic stresses, therefore they are selected as a model organism (Rai et al 2013).

	LENGTH	AVERAGE GC CONTENT
Chromosome	6,413,771	41.3
Pcc7120α	408,101	40.5
Pcc7120β	186,614	40.2
Pcc7120γ	101,965	41.0
Pcc7120δ	55,414	41.6
Pcc7120ε	40,340	40.9
Pcc7120ζ	5,584	44.2

Table 1: Size and average GC content of each replicon of *Anabaena* sp. PCC 7120

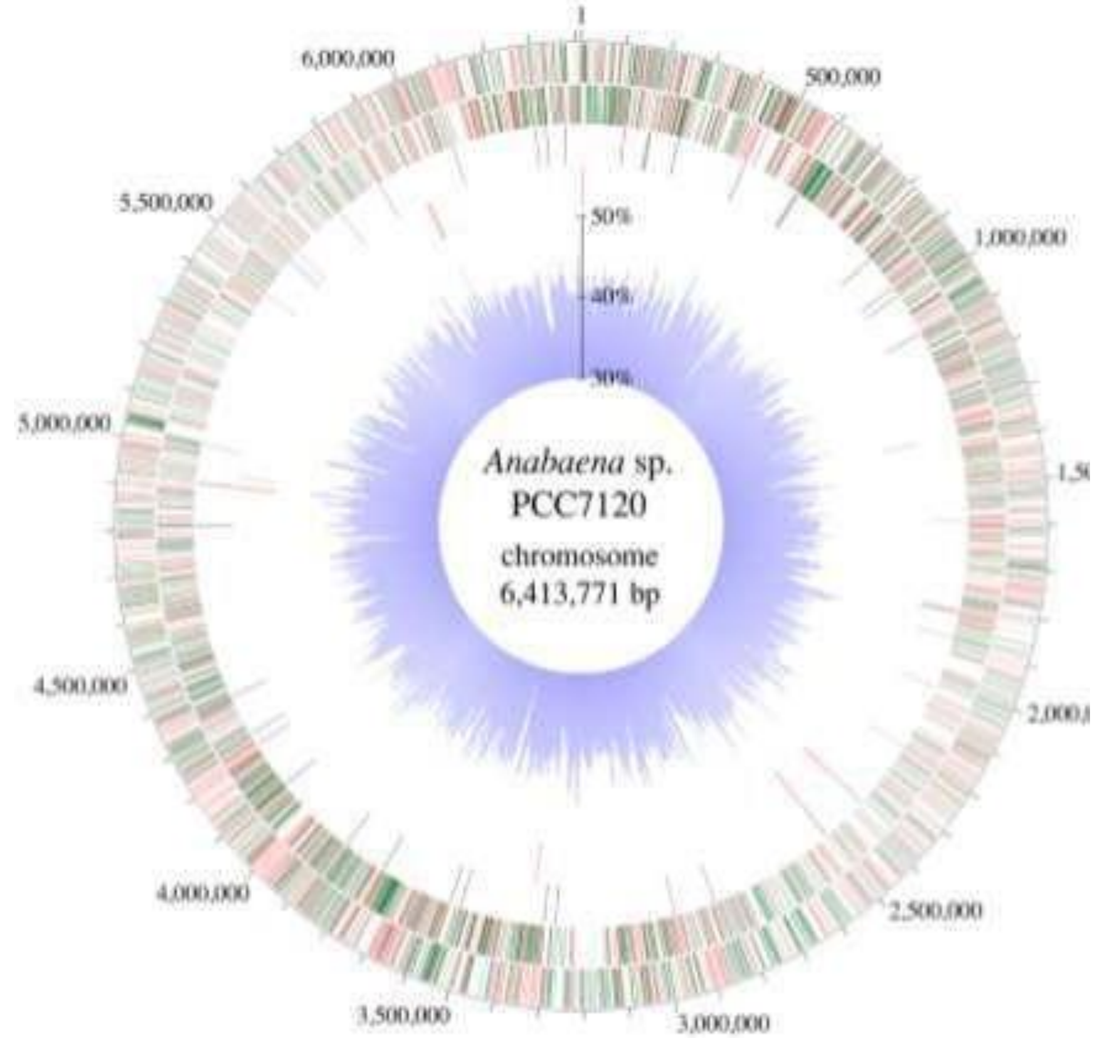


Figure 2: Circular representation of the chromosome of *Anabaena* sp. PCC 7120.

2.3 Cold stress in *Anabaena PCC 7120*

Stress is nothing but actually modification of different eco-factors that leads to a response in terms of adaptation against biological organisms. Extensive investigation of responses towards cold stress is of utmost importance because of their ability in determining the behavior of organisms in terms of attack and counter-attack in different environmental modifications. Variety of stresses influences different biological and chemical structures inside the organism which consists of ultraviolet radiation, pH, or salinity, or thermal stress, among which cold stress plays a vital role in biological studies of bacterial organism.

Response towards the cold stress depends on the prevailing temperature and more specifically whether it is greater than or less than 0°C. Extremely low temperature (Below 0°C) is associated with the freezing of free water of the cell which eventually led to the death of cells (Haines, 1938). This is also associated with the degradation of cell membrane and to the nucleic acid of the cell (El-Kest (1992). Long exposure of organism to low temperatures (below 0°C) results in an active response of bacteria by the rapid accumulation of specific cold shock proteins, leading to a transient metabolic adaptation in low temperatures. The adaptation of cyanobacteria to cold stress is mainly through biochemical modifications, physiological changes and metabolical induction.

Modification in environmental parameters leads to an active and rapid response in biological organisms. Different responses of cyanobacteria towards variety of stress have been examined because of their utmost importance in determining the ability of organisms to respond to or counter attack against the prevailing environmental modifications. Stresses that affect biological structures of organism may be thermal o non-thermal, such as ultraviolet radiation, pH, or salinity, or thermal. Response against cold stress in bacteria takes place in two phases, a continuous acclimation phase and a transient shock response. Cold temperature of bacteria approves their growth with the immediate defense response by the synthesis of cold shock proteins (CSPs).

Existence of cyanobacteria at low temperatures requires unique mechanisms to combat with the present thermodynamic limitation. These thermodynamic limitation involves depletion in the quantity of available water which eventually leads the formation of ice crystal , decrease rate of enzyme catalysis along with membrane fluidity, and stabilization of molecular structures and eventually lowers the growth of organism (Cavicchioli et al. 2000).

Membrane lipids stiffen by undergoing a transition from liquid crystal to gel at low temperatures. This stiffening in lipid membrane makes the embedded proteins unable to transport the nutrients within the membrane effectively and eventually led to starvation in the cell. (Nedwell 1999; Pomeroy and Wiebe 2001). Researchers found that the bacteria will inhibit their proper functioning when 50–90% of their membrane lipids converted from liquid to gel phase (Jackson and Cronan 1978; Melchior 1982). The fluidity associated with the cell membrane can be achieved by increasing the amount of unsaturated lipids, reducing the acyl chain length and branch-chained lipids, and by the synthesis of compatible solutes in the organism (Russell and Fukunaga 1990).

In low temperatures, protein structures become less flexible, one of the major reasons for their reduced activity. When enzymes function, they use the energy of the environment to drive conformational motions. At low temperature, movement of enzyme which eventually ceases the catalytic function of enzyme (Fields 2001).In order to function properly, cold adapted enzymes reduces their weak stabilizing interactions (ion pairs, hydrogen bonds, hydrophobic and inter-subunit interactions) and increase solvent interactions with a polar or interior residues and finally reduces proline and arginine content, and/or clustering of glycine residues (Feller et al. 1996; Russell 2000).Cold temperatures generally hinder the crucial metabolical process of transcription, translation, and DNA replication in order to maintain secondary structure of nucleic acids. The organism which is adapted to cold alleviates the above-mentioned process by the use of specialized helicases and chaperones (Jiang et al. 1997; Chamot and Owttrim 2000; Phadtare et al. 2002). Cold-adapted microorganisms generally amplify the presence of dihydrouridine content of t-RNA and reduce the production of the GC content present in their 16S r-RNA which is the prerequisite requirement involved in the preservation and maintenance of conformational flexibility of functional RNA (transfer and ribosomal RNA) at low temperatures (Dalluge et al. 1997; Noon et al. 2003; Khachane et al. 2005).

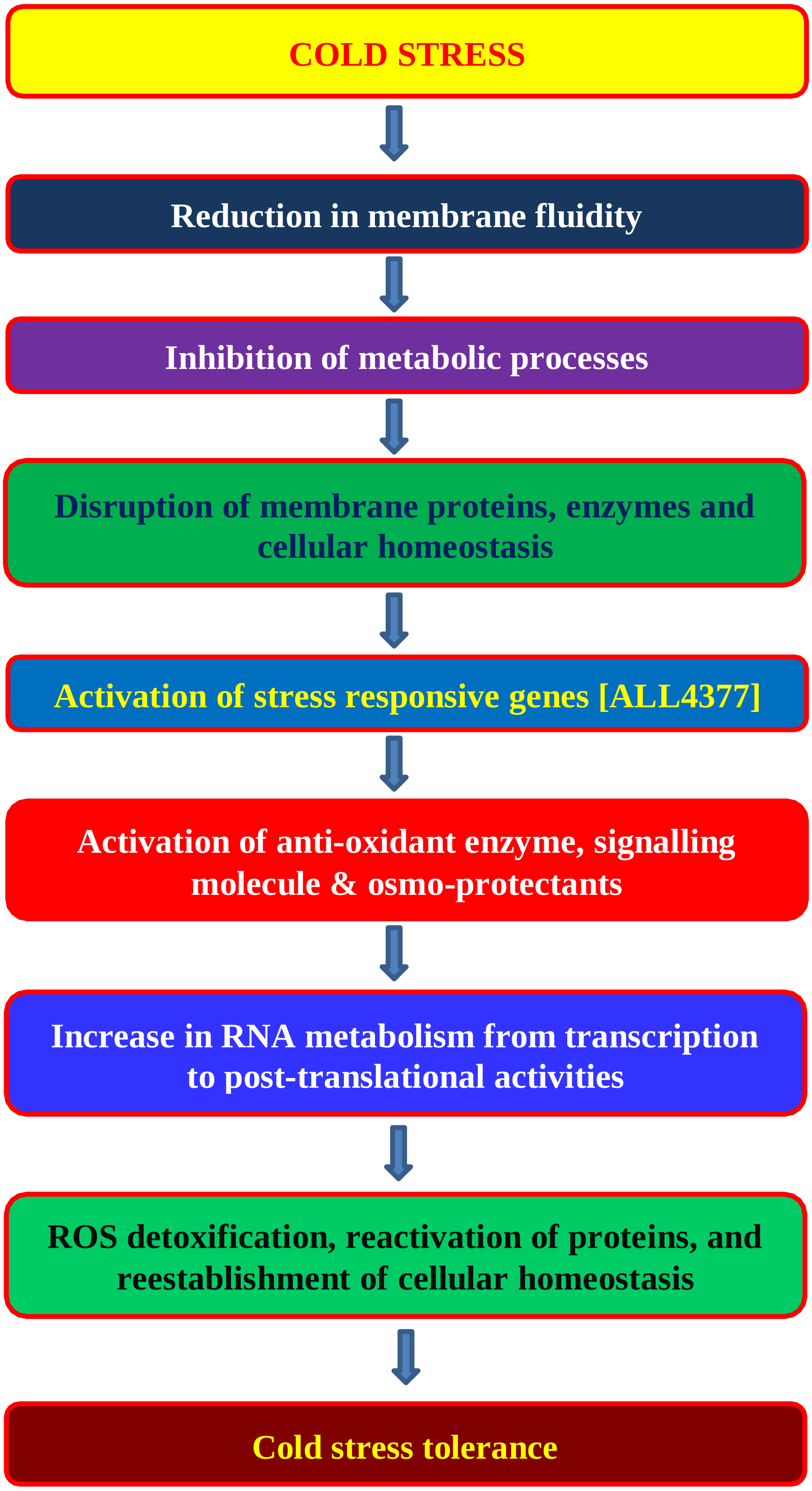


Figure 3 : **Flow chart showing cold stress tolerance in cyanobacterial**

organism

2.4 RNA Binding Protein (ALL 4377)

In the years following the discovery of RNA, it was considered only as a supporting agent during the transfer of genetic and cellular information from DNA to proteins. However, recent studies on new cellular forms of RNAs and their uniquely diverse functions in biotic and abiotic stress have shaken the ground of traditional molecular cell biology and sparked a keen intense interest in RNA biology. RNA molecules cannot exist independently, which causes the stable assembly of RNA molecules in conjunction with an array of proteins to form ribonucleoprotein complexes (RNPs). The numerous proteins which form stable complexes with RNA and have been named as RNA Binding Protein. RBPs are indispensable partner not in their metabolism and regulation but also in RNA localization, splicing surveillance, translation and in their degradation and localization biology. The increased production of RNAs results in the elevated generation of RBPs. The biogenesis of RNPs and their functional roles in every condition must be orchestrated with noble fidelity.

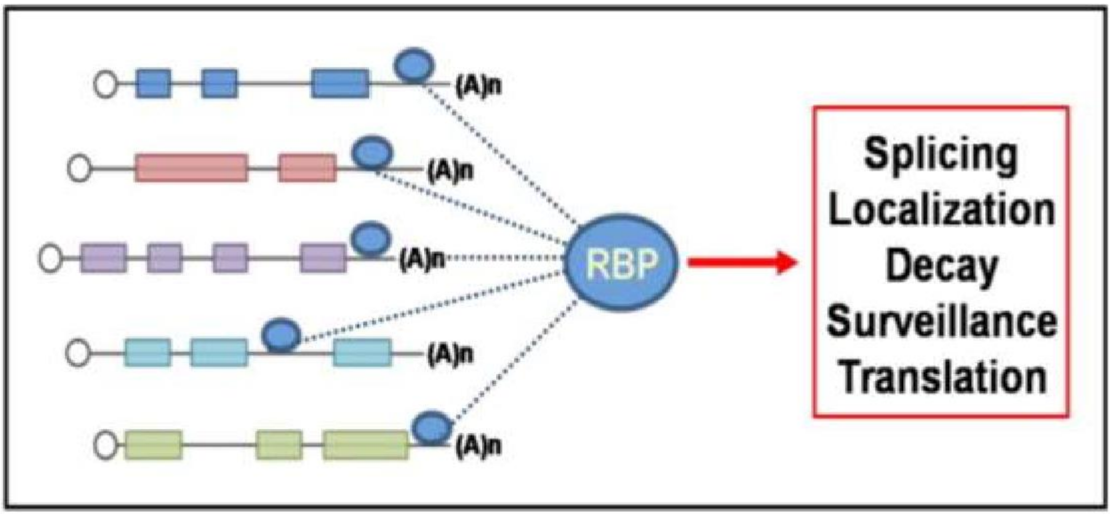


Fig. 4: Post-transcriptional RNA regulon.

The RNA-binding proteins (RBPs) are a large class of more than 2,000 proteins intimately involved with transcripts in many different RNA-driven ways. RNA-binding proteins use a variety of structures and mechanisms for binding and regulating RNA. Many studies have found that RNA-binding proteins (RBPs) play a crucial role in regulating transcripts during their complete life cycle (Lorkovic, 2012). There are several ways in which RBPs interact with RNA, from simple, single-protein interactions to multi-protein interactions such as spliceosomes. The existence of different RNA-binding domains in the particular protein is the chief reason of RNA binding. Variety of such domains is very common in a single RBP, and their arrangement within the domain helps in coordinating and enhancing binding of RNA (Clery and Allain, 2012; Lunde et al., 2007).

A variety of RNA-binding proteins are involved in RNA metabolism, including splicing, modification, stability, and translation (Nagai et al. 1995). RNA-recognition motif (RRM), sometimes referred to as the consensus-sequence type RNA-binding domain, is one of the best-characterized RNA-binding domains. The presence of one or more RRMs was observed in more than 100 rbp proteins from a variety of prokaryotic and eukaryotic organisms like plants, animals, and microorganisms (Birney et al. 1993).

This study involves the characterization of RNA-binding protein (rbp G) from the cyanobacterium Anabaena PCC 7120, and expression of this gene is known to be regulated at low temperature (Sato 1994). Subsequent study of this gene revealed that it belongs to a multi-gene family designated as the RRM_SF family (Sato 1995). A closely related cyanobacterium appears to have homologous rbp genes, Anabaena variabilis M3 (Mulligan et al. 1994), and orthologous rbp genes exist in distantly related cyanobacteria, namely, Chlorogloeopsis sp. PCC6912 (Mulligan et al. 1994), Synechococcus sp. PCC6301 (Sugita and Sugiura 1994) and Synechococcus sp. PCC7942 (Dolganov and Grossman 1993, Belbin and Mulligan 1995).

All of these cyanobacterial Rbp proteins have been isolated from Anabaena (Sato 1995) or Synechococcus (Sugita and Sugiura, 1994) and exhibit a close association for poly (U) and poly (G) rather than for poly (A) and poly(C), when compared with the RNA-binding proteins isolated from chloroplasts (Ye and Sugiura 1992) and plant nucleoli (Ludevid et al. 1992). Despite of their immense presence of RBPs in eukaryotic organisms and cyanobacteria, the presence of identical RNA-binding proteins in other clade of prokaryotic organism is in vain.

Higher plants chloroplasts contain two types of RBP, a carboxy-terminal glycine-rich domain and an amino-terminal acidic domain (Sato, 1995). The glycine-rich domain of RBPs of the cyanobacterium is divided into two subgroups: those with a glycine-rich carboxy-terminal domain and without a carboxy-terminal domain. Cold-regulated rbp genes express a glycine-rich carboxy-terminal domain. At 38^0C, the transcripts of cold-inducible RBP genes are barely detectable however they become easily available within 15-30 min after cells are transferred from 38^0C to 22^oC. The level of corresponding proteins also increased dramatically after this temperature change.

A lac Z reporter gene fused to several modified promoter regions (Sato, 1998) localized putative cold-responsive cis-acting elements in the 5”-

untranslated region of the rbp gene. The 150-bp region of DNA, is present between the initiation site of transcription and a ribosome-binding site, which plays an important role in cold-induced transcription of the rbpA1 gene. Deletions within this region result in constitutive transcription at both 38^0C and 22^oC (Sato and Nakamura, 1998).

A clear physiological role for cyanobacterial RNA-binding proteins is yet to be determined; however, the cold regulation through the expression of these proteins may provide insight into their cellular and metabolic function. Most of these proteins contain one or more copies of the RRM (about 80 residues long), along with multiple auxiliary domains, such as an arginine-rich domain, a glycine-rich domain and acidic domain. Researchers have deciphered the three-dimensional structure of the RRM domain of the U1A protein and the SXL protein. The results revealed that the beta-sheet includes four strands, which provide an RNA-binding surface, and two alpha-helices that sustain the beta-sheet from the back. RNA is also supposed to interact with a protruding loop (loop 3).The site where RNA binds consist of two central strand which is an octapeptide called the RNP-1 motif together with a hexapeptide called the „RNP-2 motif, the combination of these two RNP motifs has been identified as the RNA-binding sites. The very common interaction between hydrophobic side chain and RNA bases takes place within these motif usually with the help of stacking.

The process of gene expression involved during RNA metabolism originates with transcription and further proceeds with post-transcriptional RNA metabolism, which involves RNA splicing, translation, 3‟ end-formation, RNA import and export along with their degradation ,decay ,and surveillance. Traditionally, gene expression has been considered separate and independent post-transcriptional steps. However, current views on gene expression emphasize the interconnectivity and coordination of these steps (Mainitis, 2002). Functional multiplicity of RBPs would provide a mechanism for connecting the post-translational steps.

According to the RNA regulon hypothesis (Keene, 2007), post-transcriptional regulation might coordinate multiple RNAs within the cell. RBPs play as a chief regulator of the RNA regulon, which gives them the ability to modulate an array of target RNAs. RNA-binding proteins bind to consensus RNA elements in target RNA molecules, probably coordinating functionally related RNA molecules. Identification of RBP binding RNA elements can be investigated experimentally by various methods such as Systemic Evolution of Ligands by EXponential enrichment (SELEX) technique and cross-linking immunoprecipitation (CLIP) method. This technology should enable us to analyze the complex RNA-RBP network at the cellular level using global analyses.

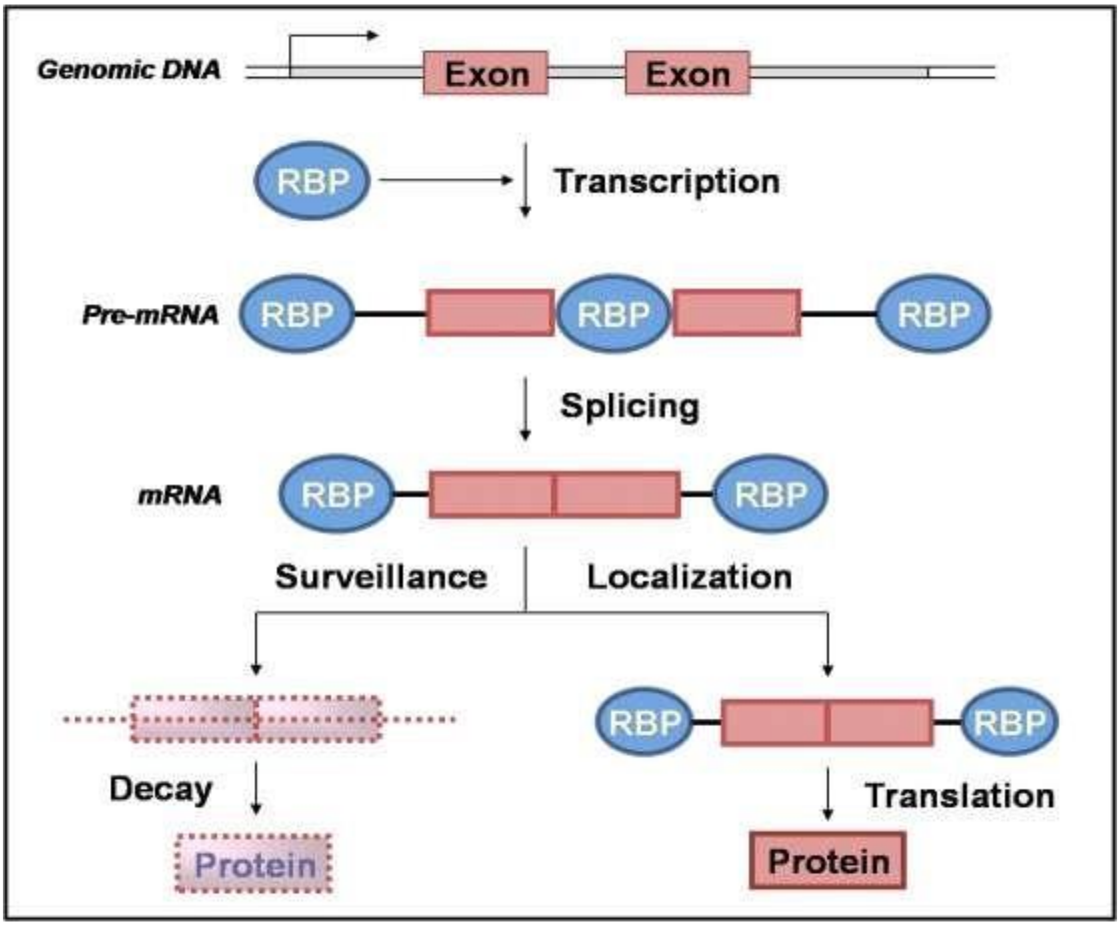


Fig. 5: <u>Multifunctionality of RBP in RNA metabolism.</u>

The chemical interactions between RNA nucleotides and RBP residues give RBPs their ability to bind RNA targets. As the molecules approach this resolution, the distinction between protein and RNA disappears; as the same intermolecular forces shape the structures of both the molecules and also stitch the two molecules together. It has been shown that these interactions involve dynamic rearrangements of RNA and protein (Hainzl et al., 2005; Leulliot and Varani, 2001) Most proteins can interact with RNA by addressing any of their main chains and side chains.

A consistent study in RNA-Protein side chain interaction evidently demonstrates the presence of 71.5% of hydrogen bonds and 76% of Vanderwal interactions with the particular nucleic acid. These predominating interactions depicts the presence of popular polar amino acids Serine and Asparagine along with positively charged amino acids Lysine and Arginine, which interacts with RNA via strong ionic hydrogen bonds (salt bridges) (Gupta and Gribskov, 2011; Han and Nepal, 2007; Hoffman et al., 2004; Perez-Cano and Fernandez-Recio, 2010; Treger and Westhof, 2001). The presence of similar amino acid is also observed in Vanderwal interactions as present in hydrogen bonds (Ellis et al., 2007; Han and Nepal, 2007; Jones et al., 2001; Treger and Westhof, 2001).Apart from hydrogen and vanderwal interactions, the occurrence of few hydrophobic interactions (50% of interaction) also have been observed in protein-RNA structures. (Hu et al., 2018).

2.5 Future perspective of RNA Binding Protein

The study of cold-shock response has recently attracted attention because of its numerous commercial and health implications. There is evidence that cyanobacterial RNA binding proteins contribute to cold shock resistance and further promise to prevent industrial disasters. As refrigeration is the most commonly used storage method for food, it is vitally important to understand how food borne pathogens react to cold temperatures. Better cryotolerance can be seen in cells that experience cold shock before freezing. The direct freezing of food-spoilage bacteria makes them susceptible to damage caused by cold environments (Willimisky et al, 1992)

Low temperature is one of the concerns within the agricultural industry since it reduces the efficiency of bio-fertilizers. To address the above problem, RBP mediated rhizobial bio-fertilizers are used to cope up with cold stress. A decrease in temperature reduces grain yields and quality. For this, plants use RBP by using the gene transfer method from bacteria to plants resulting in enhances chilling resistance. RBPs must play a role in the inefficient expression of proteins or proteolysis by expressing cold-inducible expression systems for the expression of these proteins. The catalytic efficiency of the enzymes was enhanced with the use of cold-adapted enzymes.

Chapter III

MATERIALS AND METHODS

3. INSILICO ANALYSIS OF ALL 4377

3.1.1 Physicochemical properties of All 4144

We predicted the physical, chemical, and functional properties of protein with the help of the Protparam tool (http://web.expasy.org/protparam) which is a part of Swiss-Prot or TrEMBL. This tool provides data which is used for deep-analysis for elucidation of the protein by predicting its physicochemical properties that include molecular weight, theoretical isoelectric point (pI), amino acid composition, atomic composition, instability index, half-life of proteins. The output value of Protparam is employed to compute other metrics such as aliphatic index, GRAVY, and PEPIB in the database of CyanoPhyChe (http://genome.kazusa.or.jp/cyanobase). These two databases contain information about the physical and chemical properties of protein All 4144.

3.1.2 Conserve domain search analysis

The hierarchy of homologous sequences was constructed by analyzing conserved domains of All 4377 based upon the sequence similarity search with close orthologous family members. For this purpose, a tool – InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan) which combines different protein signature recognition methods (Zdobnov and Apweiler, 2001) and National Center for Biotechnology Information (NCBI)-conserved domain database which consists of a collection of well-annotated multiple sequence alignment models for ancient domains and full-length proteins and HMMER (https://www.ebi.ac.uk/Tools/hmmer) were used. By analyzing conserved domains in proteins, it is possible to reflect various aspects of evolutionary history and perform phylogenetic and structural analysis.

3.1.3 Sequence alignment, phylogenetic analysis, and motif elucidation

The full- length primary protein sequences of RNA Binding protein from cyanobacterial sources were retrieved from the database of Cyanobase (http://genome.microbedb.jp/cyanobase) .To find out the degree to which conserved regions of the All 4377 matches with consensus regions of the proteins from other cyanobacterial species, Basic local alignment search tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used. Multiple sequence alignment for individual profiles was performed using Clustal W. The Clustal W alignment file was imported into the Box

Shade sequence alignment editor. Identical and similar amino acids were shaded or colored. Phylogenetic analysis of protein sequences was generated using the alignment obtained with Clustal W as an output and input for creating phylogenetic tree using MEGA 7 software. The discovered motifs were further used to search their protein family using Pfam at the DDBJ MOTIF server (http://www.genome.jp/tools/motif/).

3.1.4 Structural analysis of All 4377 protein

A 2-D model of protein was constructed using protein structure homology model building program PSIPRED (http://www.bioinf.cs.ucl.ac.uk/psipred) and PREDATOR (https://npsa-prabi.ibcp.fr/cgi-bin) in order to predict the function of All 4377. All the protein sequences of RNA binding protein were used in prediction of secondary structure. PREDATOR acknowledge the input protein sequence in the form of a FASTA formatted file and then determine the secondary structure by using profiles present in the STRIDE database of PREDATOR.

A 3-D model of cyanobacterial RNA Binding protein was designed using the protein structure homology model building program SWISS-MODEL with energy minimization parameters. protein was framed using I-TASSER and swiss model for structural analysis and verification

3.1.5 Homology modeling and verification

The modeled 3D structures were estimated and confirmed with the I-TASEER(https://zhanglab.dcmb.med.umich.edu/I-TASSER) and RAMPAGE program (Laskowski *et al.* 1996). PyMOL software is used to calculate the RMSD value and superimposition between the built structure and its template by using the built 3-D structure obtained from homology modeling.

3.1.6. Functional analysis and protein-protein interaction study

In order to predict the interaction of RNA Binding protein of Anabeana PCC 7120 with other closely related proteins STRING v10.0 (http://string-db.org/) server was used. All 4377 gene sequence was selected as query sequence, and functional protein association network was generated. In addition to it ,the selected query sequence was examined in order to determine the protein famiy. For this, motif finder server (http://www.genome.jp/tools/motif/) was used. Active site was predicted from COFACTOR (http://zhanglab.ccmb.med.umich.edu/COFACTOR/).

3.2 GROWTH AND MAINTENANCE OF MODEL ORGANISM

3.2.1 Cyanobacteria (*Anabaena* sp. PCC7120) its growth conditions and maintenance

Anabeana PCC 7120 was grown photoautotrophically in BG-11 Medium (Rippika et al., 1979) buffered with 10 mM HEPES-NaOH, pH maintained at 7.5 at 24±2°C under day light fluorescent tubes radiating 72 μmol photon m^{-2} s^{-1} PAR (photosynthetically active radiation) light intensity having a definite photoperiod of 14:10 h. Composition of BG-11 medium is given below .

1ml of each MgSO4.7H2O, CaCl2.2H2O, Na2CO3 and micronutrient were taken from respective 100X stock solution to make 1 liter BG-11 media and pH of media was adjusted to approximately 7.5g/L. HEPES was added to the final medium as a buffer to reduce PH from 8.0 to 7.5. K2HPO4.3H2O, citric acid combined with ferric ammonium citrate and EDTA were autoclaved separately and finally added to the cold sterilized culture medium.

The culture were shaken 2-3 times daily for aerated condition. Incubated in an air conditioned illuminated by 40W fluorescent tubes for 14 hour daily. Light intensity varied between 72-75 micromol photons m^{-1} s^{-1}.

Macronutrients	Conc.(g L^{-1})	Micronutrients	Conc. (mg L^{-1})
$K_2HPO_4.3H_2O$	0.040	H_3BO_3	2.76
$MgSO_4.7H_2O$	0.075	$MnCl_2.4H_2O$	1.81
$CaCl_2.2H_2O$	0.036	$ZnSO_4.7H_2O$	0.222
Citric Acid combined with Ferric Ammonium Citrate	0.006 0.006	Na $MoO_4.5H_2O$	0.390
EDTA*	0.001	$CuSO_4.5H_2O$	0.079
Na_2CO_3	0.020	$Co(NO_3)_2.6H_2O$	0.0494

TABLE 2: Detailed composition of modified BG-11 stock solution and 1X media

3.2.2 Maintenance of *Anabaena* stock cultures

Maintaining organism stock is crucial part in creating cell culture of an organism. 1LCulture media was prepared in an autoclaved haffkine of 3-4L capacity. The culture medium was sterilized with non-absorbent cotton plug and autoclave at 15lb inch $^{-2}$ for 15 minutes. Culture medium was preserved in air conditioned culture room illuminated by 40 W florescent tubes for 14 h daily. The culture received 72 µM photons m-2s-1 light at 24 ± 2˚C and was hand shaken 2-3 times daily. All experiments were done using exponentially growing cultures.

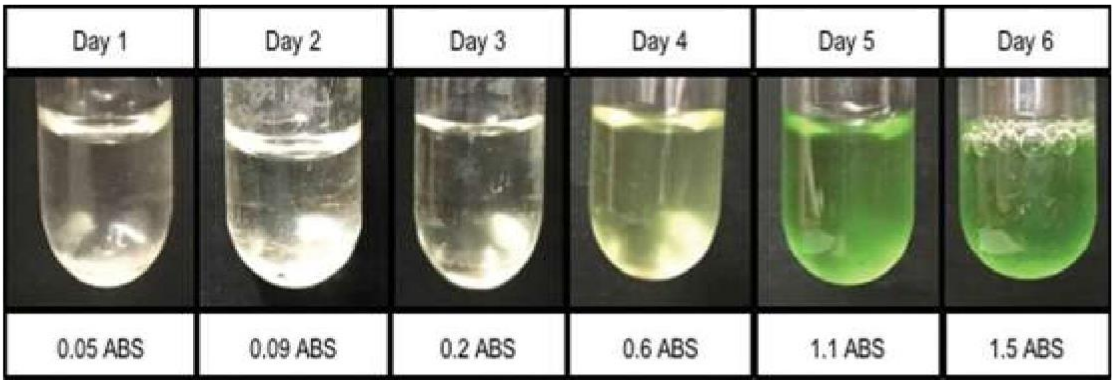


Figure 6: Cyanobacterial culture and their absorbance

3.2.3 Bacterial strains and plasmid

Gene cloning and expression of our target protein were achieved using E. coli strain DH5α and BL21 (DE3) (Novagen) as hosts, respectively. The Plasmid Pet-21a (Novagen) acts as a vector for the biological process cited earlier. E. coli cultures were prepared with the help of Luria – Bertani medium containing 2gm LB powder in 100 ml of distilled water and autoclaved immediately to destroy contaminants.

The E. coli culture media were maintained on Luria–Bertani (LB) plates at 37°C containing autoclaved 1.5% (w/v) agar media [2gm LB + 1.5GM Agar + 100ml distilled water] and stored as 10% (v/v) glycerol stocks at −80°C. The cells harboring recombinant plasmids have been grown and maintained with LB medium supplemented with 100 g/ml ampicillin (Sambrook and Russell, 2001).

3.2.4 Luria and Bertani medium

Luria-Bertani (LB) broth is the most popular and broadly used medium extensively used for bacterial growth for the growth of bacteria (Bertani 1951). The unique composition of the LB medium is given in Table 2. pH of the medium was upholded approximately at 7.5. In order to prepare culture, LB media was prepared by adding 2gm of LB powder in 100Ml of demonized water and eventually followed by sterilization by autoclaving at 15psi, from 121-124^0C for 15 minutes. For solid medium, 1.5% agar was added in same solution and sterilized before using.

Components	**Concentration (%)**
Tryptone	1.0%
Yeast extract	0.5%
Sodium chloride	1.0%

TABLE -3: Components of Luria-Bertani medium

3.3 ISOLATION AND QUANTIFICATION OF GENOMIC DNA AND PLASMID

3.3.1 Genomic DNA Isolation

DNA, deoxyribonucleic acid, is the Bio-molecule present in every cell and responsible for life of the cell, therefore commonly known as, molecule of life. DNA extraction from cyanobacterial cell is a methodology which involves isolation and purification by using different physical and chemical methods from a given amount of sample which eventually separates DNA from nuclear membrane, cell membranes, proteins, and other cellular components. The utilization of DNA isolation process involves efficient extraction having refined quality and quantity of DNA which is completely pure and having no amount of contaminants like RNA and protein. Extraction of DNA from the cell involves breakdown of cell membrane and nuclear membrane and solubilization of DNA in chemical environment followed by removal of macromolecules, lipids, RNA and proteins with the help of different chemical and enzymatic method. Total genomic DNA was extracted according to the protocol of Srivastava *et al.,* (2007):

BUFFER /CHEMICALS	***COMPONENTS***	***PURPOSE***
1. WASH BUFFER	*a) 50mMTris HCl [pH 8.0]*	*Maintain Ph of buffer .Interact with LPS and destablize membrane*
	b) 50mM NaCl	*Precipitation by neutralizing - charge on DNA.*
	c) 50mM EDTA	*Destablize membrane by eliminating Ca+2 and Mg+2.*
2. TE BUFFER	*a)50mM Tris-HCl [Ph-8.0]*	*Maintain Ph of cell solution containing fragmented DNA ,Protein*
	b) 50mM EDTA	*Protect fragmented DNA from nucleases.*

3. EXTRACTION BUFFER	*a) 50mM Tris-HCl [pH 8.0]*	*Used to maintain stable Ph*
	b) 20mM EDTA	*Chelating agent*
	c) 2.5 M NaCl	*Neutralize charge on the sugar phosphate backbone of DNA*
	d) 3% CTAB	*Disrupts membrane and precipitate protein*
	e) 1% sarkosyl	*It is detergent and disrupts membranes*
	f) 2- mercaptoethanol	*Reducing agent denature Rnases*
4. PROTEINASE K	*2 mg ml-1*	*Digest proteins and degrade nucleases.*
5. CHLOROFORM ISOAMYL ALCOHAL	*24:1*	*Separates protein from nucleic acid*
6. 70% ETHANOL		*Rinse excess salt that come along with extraction buffer from the pellet*
7. SODIUM ACETATE	*3M [Ph 5.2]*	*Na+ neutralize negative charge on the nucleic acid makes molecule less soluble in water.*
8. TE SUSPENSION BUFFER	*a) 10mM Tris-HCl [Ph-8.0]*	*Alkalinity of TE buffer prevents acid hydrolysis which disrupt the stability of DNA stored in water*
	b) 1mM EDTA	*Chelate Mg+2 in solution to protect DNA from DNAase activity*

TABLE 4: <u>Buffer/reagents involved during DNA isolation and their uses</u>

Protocol:

1. The 50 ml of 0.5 O.D. cyanobacterial culture was centrifuged at 5000 rpm for 10 min and Pellet was washed twice in 2 mL wash buffer

suspended in the same buffer and homogenized using sterile glass beads on a vortex for 5 min.

2. The pellet was further washed with the same buffer, suspended in 2 mL TE buffer and sonicated for 10 min in ultrasonic processor (Heat systems, USA) under pre-chilled condition.

3. 2 mg ml-1 of proteinase K were added to the sonicated samples and incubated at 37°C for one hour in water bath shaker.

4. 2 ml of pre-warmed extraction buffer were added to the samples and kept at 65°C for one hour in a water bath shaker.

5. After cooling to room temperature, chloroform: isoamyl alcohol (24:1) solution was added in equal volume. The centrifuge tube containing the above solution was subjected to gentle rolling till the two distinct layers disappeared and a homogenous suspension was formed.

6. This suspension was centrifuged at 9,168 g for 15 min; the upper aqueous transparent layer was taken.

7. To this 5 µl of 3 M sodium acetate (pH 5.2) and double volume of pre-chilled ethanol were added and left over night at –20°C (Vestfrost deep freezer, Blue Star, India) for DNA precipitation.

8. This was followed by centrifugation at 20,000 g for 15 min at 4°C. The

DNA pellet was washed in 70% ethanol and resuspended in TE suspension buffer.

3.3.2 Plasmid isolation (alkaline lysis method)

In molecular biology, Plasmid is an extra chromosomal small circular deoxyribonucleic acid (DNA), which duplicates independently from chromosomal DNA. In molecular biology alkaline lysis is method involves in separation of plasmid DNA from other cell component.

Bacteria with the plasmid of interest are cultured and then allowed to undergo lysis with an alkaline lysis buffer containing a detergent sodium-dodecyl sulphate (SDS) and a strong base sodium hydroxide. The major function of detergent is to break the phospholipid bilayer plasma membrane and the alkali involves denatures the protein which was used to maintain the structure and integrity if cell membrane.

Protocol:

1. The E. coli cells having pET-28a vector were incubated at 37^0C for overnight.

2. 100 microlitre of overnight grown *E. coli* having pET-28a culture was added to 5 ml of LB medium with 5 microlitre ampicillin.

3. Incubation was done for two hours at 37^0C and 250 rpm in a orbital shaking incubator.

4. The culture was centrifuged at 10,000 rpm for 10 minute.

5. The supernatant was discarded and pellet was re-suspended in 100 ml of solution first (50 Mm glucose, 25 Mm tris-HCl, 10 mM EDTA) mixed well by pipetting, and store at 0^0C.

6. 200 microlitre of solution 2 (0.2 N NaOH, 1%w/v SDS) was added slowly in the tube mixed by gently inserting the tube till the solution become clear and then tube was again store in ice for 5 minutes.

7. 150 microlitre of solution 3 (5 M potassium acetate-60 ml, glacial acetic acid -11.5 ml, water-28.5 ml) was added and mixed rapidly and stored in ice for 30-40 min.

8. Tube was again centrifuged at 10000 rpm for 15 min.

9. Then supernatant was collected in other tube and 200 microlitre tris-saturated phenol was added and mixed gently by inverting tube for 5 to 10 times.

10. The sample was again centrifuge at 10000 rpm for 10 min. at 4 ^{0}C.

11. Upper layer was collected and 1 ml of chilled absolute ethanol was added to it.

12. The tubes were stored at -20oC for 1.5 hour then the sample was centrifuged at 10000 rpm for 10 min. at 4^0.

13. A milky white pellet was obtained and washed with 250 microlitre of 70 % ethanol by inverted tube gently 3-4 times and again centrifuged at 10000 rpm for 10 min.

14. Supernatant was discarded and pellet was air dried completely then pellet was re-suspended in 30 microlitre of TE-Buffer containing 3 microlitre of RNase.

15. Isolated pET-21a vector plasmid was loaded (2µl plasmid mixed with 3µl loading dye) in 0.8 % Agarose gel.

3.3.4 Expression vector

To characterize the function of RNA binding protein, we introduce our target gene into the specific host cell with the help of a vector. Along with this, this delivery vector hijacks the host cell's protein synthesis machinery to produce only RBP protein, expressed by the gene all4377.

A Plasmid has to modify genetically with regulatory sequences, like enhancer and promoter regions, that eventually help in the transcription of the gene carried by the vector.

In context with the above statement, Studier and his colleagues designed the pET vector originally. pET vector is Plasmid for Expression by T7 RNA polymerase vector and belongs to the family of expression vectors that uses phage T7 promoters to regulate the synthesis of the cloned gene product. This system is the most dominant and widely used system which helps in cloning and expressing the recombinant proteins in E. coli.

The normal function of pET involves no protein expression as Lac1 represses transcription by blocking T7 RNA polymerase while active expression of the protein alters pET function because of the binding of IPTG to LAC repressor and express T7 RNA polymerase for transcription. The T7 RNA polymerase helps in identification of the T7promoter on the vector and express the recombinant gene into the Pet vector. At the time of induction, only our target gene will be expressed by the highly active polymerase.

Plasmids containing pET vectors allow us to clone our target genes under the control of highly effective bacteriophage T7 transcription and (optionally) translation signals; the source of T7 RNA polymerase in the host cell induces the expression. T7 RNA polymerase is highly selective and dynamic to such an extent that it can convert almost all of the cell’s resources to target gene expression upon full induction. pET expression vectors developed from the pBR322 plasmid and can be engineered to utilize the advantages of T7 bacteriophage that promote high expression in terms of transcription and translation. RNA polymerase encoded by bacteriophages is highly precise to promoter sequences in the T7 phage genome but is rarely present in other genomes besides the T7 phage genome. In turn, this helps the T7 promoter to hide from the host cell RNA polymerase. Thus target genes are transcriptionally silent in the uninduced state, which is the utmost important feature for efficient expression of a toxic gene in the cell.

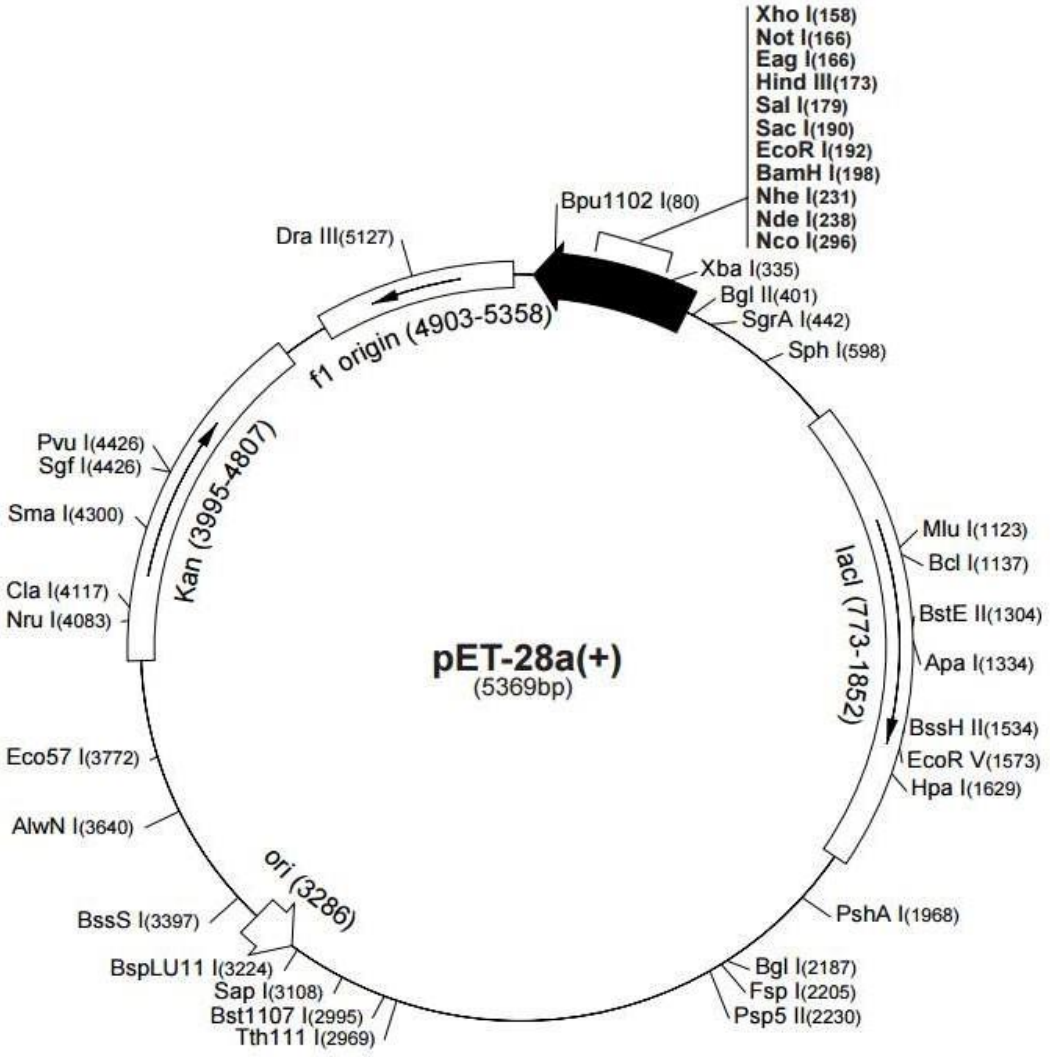


Figure-7: <u>The pET -28a (+) vector expression system</u>

3.3.5 DNA quantification and electrophoresis

According to the book Sambrook and Russell, the concentration of DNA can be determined by UV absorbance which involves measuring the absorbance/transmission of light through a liquid to determine the concentration of substances in the liquid. Prepared DNA concentrations are calculated by measuring the maximum absorbance at 260 nm. The ratio of A260/ A280 ratio measured in the NanoVue Plus spectrophotometer was used to check the purity of DNA (GE Healthcare Life Sciences). Genomic DNA was electrophoresed at 100 mA in 0.8% agarose gel containing 1 μg/ml ethidium bromide using TAE electrophoresis buffer (40 mM Tris-acetate and 1 mM EDTA, pH 8.0) and photographed in the G: BOX gel documentation system (Syngene, USA).

3.4 CLONING AND EXPRESSION ANALYSIS

3.4.1. PCR amplification of *Anabaena* sp. PCC7120 ORF *All4377*

The open reading frame *All 4377*, encoding a hypothetical protein was amplified by polymerase Chain reaction using genomic DNA as the template. The PCR primer pairs used were All 4377(forward primer) and All 4377(reverse primer), the Nde1 and Xho1 recognition sites are underlined. The PCR was done in a reaction mixture of 25 μl for 30 cycles at 94 °C for 1.30 min, 61.7 °C for 1 min, and 72 °C for 2 min using standard PCR conditions (100 ng of DNA, 2.5 μl of 10× PCR buffer with 15 mM MgCl2, 200 μM dNTPs, 10 pmol of each primer and 0.2 U Taq DNA polymerase in an Icycler (Bio-Rad, USA).

3.4.2 Extraction and purification of PCR product from agarose gel

The PCR product was applied to 1% agarose gel, the DNA band was excised from the gel. The amplified product was gel purified using a QIA quick gel extraction kit (Qiagen).

3.4.3 Double Digestion of purified *All 4377* PCR product and pET-28a vector

Purified PCR product of All 4377 and pET-28a expression vector were double digested with Nde1 (NEB) and Xho1 (NEB) restriction enzyme by incubating at 37°C for overnight. The restriction endonuclease enzyme Nde1 and Xho1 are 6 base pair cutter. Digestion reaction component include the purified PCR product, buffer (Cutsmart buffer), NdeI, XhoI enzyme and water.

3.4.4 Purification of digested product

Digested PCR product and vector was run on 1% agarose gel and purification was done using a QIA quick gel extraction kit (Qiagen).

3.4.5 Ligation (Construction of the recombinant plasmid)

The final step in the construction of a recombinant plasmid is inserting the gene of interest All 4377 into a compatibly digested pET28-a vector backbone. This is accomplished by covalent connecting the sugar backbone of the two DNA fragments.

- **PROCEDURE OF LIGATION**

Before setting up the ligation reaction it is important to determine the amount of insert and vector. The volume of vector DNA and insert DNA used in the ligation will vary depending upon their concentration. For ligation a molar ratio of 3:1(Insert: Vector) was taken and ligation was performed in the following way. Two reaction mixtures were used during ligation. One reaction mixture containing only the pET vector (no PCR product) used as control and the other containing *al4377* (PCR product) and digested pET, both the mixture having volume of 10 μl. Mixture was incubated overnight at 16^0C. Reaction mixture included PCR product (insert), Vector, Ligase buffer, T4 DNA Ligase and DD water making the final volume of 10 μl. The purified PCR product after digested with Nde1 and Xho1 (NEB), and the resultant DNA fragment was ligated with the pET-28a (Novagen) expression vector digested with the same restriction enzymes. Vector pET21a incorporate hexa histidine (6x His) tag to aid purification. After ligation the recombinant plasmids and empty vector were introduced into *E. coli* strain BL28 (DE3) for expression studies.

3.4.6 Preparation of competent cell for transformation

Single colony of *E. coli* (DH5α and/or BL21), was inoculated in 5 ml Luria broth (LB) and allowed to grow overnight at 37^0 C shaker. Overnight grown culture was diluted in 50 ml LB (1:100 dilutions) and allowed to grow till O.D. comes 0.4 (2-3 hr). Then the culture was chilled on ice for1 h and centrifuged at 4000 rpm for 5 min at 4oC. Pellet was suspended in 10 ml of ice cold CaCl2 and kept on ice for 30 min. Cells were harvested by centrifugation at 5000 rpm for 4 min. at 4°Cand pellet was resuspended in 3 ml of 20% glycerol in 0.1M CaCl2. Aliquotes of 100 µL were frozen in liquid nitrogen and then stored at -80^0C till use. The required solutions and centrifuge or microfuge tubes were used and kept in the cold during the whole process.

3.4.7 Transformation of ligated product in *E. coli* BL-21(DE3) cells

Heat shock method of transformation was used to transform bacterial cells. Approximately 50 ng of ligated mixture was used for transformation of competent *E. coli* cells (BL21 cells). Ligated mixture was added in 100 µl of competent cells, kept on ice for 30 min, then a brief heat shock (90 sec at 42°C) was given to the cell and tubes were replaced on ice for 10 min. After that 1 mL of LB medium was added and the cells were allowed to grow at 37°C for 90-120 min in shaking incubator. Finally, cells were harvested by centrifugation at 10,000 rpm for 1 min, resuspended in 100 µl of LB and spread on LB agar plate containing ampicillin. Plates were incubated at 37°C overnight and then positive clones were identified using colony PCR and double digestion. From well grown template plates, master plates was made by streaking method and inoculated in 5 ml of LB tubes for plasmid isolation (recombinant). Culture was allowed to grow in tubes as well as on master plate in incubator at 37^{o}C for overnight. The master plate was stored at 4°C for further use.

3.4.8 Selection of positive clones by colony PCR and further confirmation by double digestion

For the confirmation of positive clones two methods were employed; colony PCR and double digestion. (i) Colony PCR was performed by taking cells from master plates and then drawn up and expelled the PCR mix (containing 2.5 µl of 10x Taq buffer with 15 mM MgCl2, 200 µM dNTPs, 10 pmol of each primer, and 0.2 U Taq DNA polymerase) in the PCR tube a few times with the innoculated tip. Once all samples were ready polymerase chain reaction was performed for the targeted gene. (ii)

For double digestion plasmids from transformed clones were isolated and then digested with Nde1 and Xho1 restriction endonuclease for 12 h at 37˚C. The digested product was analyzed on 1% agarose gel.

3.5 EXTRACTION, EXPRESSION & ANALYSIS of RECOMBINANT PROTEIN

3.5.1 Heterologous expression analysis of recombinant protein All 4377

BL21 (DE3) cells transformed with recombinant plasmid pET-28a-All4377 were grown overnight in Luria-Bertani (LB) medium containing 100 µg/ml ampicillin. Overnight culture was diluted 1:100 in fresh LB broth (5x 100ml culture) containing 100µg/ml ampicillin and then incubated at 37°C for 3–4 h (or till OD600 reached ~0.6). Once the desired optical density reached, the cultures were cooled at RT and induced with a final concentration of .1mM IPTG by incubating at 16°C for 16/17 h and shaking at 200 rpm.

For isolation of protein, harvested bacteria were washed twice with extraction buffer (50mM Tris–HCl, 10mM MgCl2, 20mM KCl pH7.5) and centrifuged at 12,000×g.

The pellet was resuspended in 2 ml extraction buffer and subjected to grinding under liquid

nitrogen for protein extraction. The extract was centrifuged at 10,000 rpm for 60 min. Supernatant was mixed with×1 SDS loading dye, and 12% SDS-PAGE was carried out. Solubility of protein was also checked by loading pellet and the supernatant fraction in different lanes of SDS-PAGE.

3.5.2 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis, also known as SDS-PAGE, is a discontinuous electrophoretic system developed to measure and analyzes proteins quantitatively. The most commonly used method involves separating proteins based on their size. It is specifically designed for monitoring protein purification because the procedure relies on the separation of proteins according to their molecular weight.

The combined use of sodium dodecyl sulfate (SDS) and polyacrylamide gel allows eliminating the influence of structure, charge, and proteins are separated solely based on differences in their molecular weight. SDS is an anionic detergent that binds quantitatively to protein, giving them linearity and uniform charge so that they can be separated solely based on molecular weight and size. According to Sambrook and Russell (2001), 1-dimensional SDS PAGE took place in 12% polyacrylamide gel.

Buffers and reagents:

BUFFERS / REAGENTS	**COMPONENETS**	**PURPOSE**
1. STACKING & RESOLVING GEL	30% Acrylamide	*Helps in the formation of gels via linear polymerization*
	1.5M Tris[pH -8.8]	*Acts as a lysis buffer*
	10% SDS	*Anionic detergent. Unfolds and denatures proteins, coating proteins in negative charge*
	10s%APS	*APS catalyzes the polymerization of acrylamide and bis-acrylamide*
	TEMED	*Catalyst for gel polymerization*
	Water	*Provide medium for solution*

2. TRIS GLYCINE BUFFER	25mM Tris	*Acts as a buffer*
	250mM Glycine	*Acts as a a charge factor.*
	0.1% SDS	*Make the protein linear*

3. 2X Gel Loading Buffer	100mM Tris-Cl	*acts as a buffer*
	4% SDS	*Denature protein and provide uniform charge and aids the size wise separation*
	0.2% Bromophenol Blue	*Serves as indicator & migration dye ,facilitate in loading of samples into wells*
	20% (v/v) Glycerol	*increase the density of sample so that it will fall into the bottom of the well and helps in forming the layer of the sample*
	200mM 2-mercaptoethanol	*Reducing agent*

4. STAINING SOLUTION	2.5gm Coomassie stain	*Used to stain the proteins*
	400mL Water	*Acts as a medium*
	500mL Methanol	*denature the protein & provide*
	100mL Glacial acetic acid	*an acidic environment enhancing the interactions with dye*

5. DESTAINING SOLUTION	400mL Water	*medium for methanol & acetic acid*
	500mL Methanol	*Prevent the gel from swelling , remove unbound dye & reduce the background*
	100mL Glacial acetic acid	*facilitate the precipitation & fix the protein in the gel*

TABLE 4: buffer/ reagents involved during Sds-page and their uses

Protocol:

1. To prepare the gel, first, the glass plates were cleaned and dried, and then the base of the plate was sealed with agarose.

2. When agarose becomes solidified poured 12% resolving gel (for a total volume of 20 mL, the following were mixed in the given order: 4.6 mL H2O, 10.0 mL 30% acrylamide, 5.0 mL 1.5 M Tris pH 8.8, 200 µL 10% SDS, 200 µL 10% APS and 8.0 µL TEMED) between the two glass plates and simultaneously it was over layered with iso-propanol.

3. After 30- 45 minutes when gel was polymerized, iso-propanol was removed by washing several times with distilled water and the surface dried as much as possible.

4. The stacking gel was prepared (for a total volume of 8 mL the following were mixed in the given order: 5.5 mL H2O, 1.3 mL 30% acrylamide, 1.0 mL 1.0 M Tris pH 6.8, 80 μL 10%SDS, 80 μL 10% APS and 6.0 μL TEMED).

5. Comb was fitted and the gel was kept in a vertical position at room temperature for polymerization.

6. After polymerization was complete the comb was removed carefully and the wells were washed with distilled water to remove any unpolymerized acrylamide.

7. The gel was now transferred into the SDS-PAGE assembly. Tris glycine electrophoresis buffer was added to the top and bottom reservoirs.

8. Samples were prepared (20 μL protein sample containing 15ng of protein+ 20 μL 2X gel loading buffer), heated for 2-3 minutes and then loaded in the wells.

9. The electrophoretic apparatus was now attached to an electric power supply. The gel was run at a constant current of 15 mA and after the dye front reached the resolving gel the current was increased to 25 mA.

10. The dye front was allowed to reach the bottom of the gel and power turned off.

11. The gel was removed and stained overnight in staining solution and destained in destaining solution.

3.5.3 Extraction and Purification of recombinant protein

The transformation occurred by the recombination event of BL21 (DE3) cells with recombinant plasmid pET-28a-ALL 4377 were grown on Luria-Bertani (LB) medium containing 100 g/mL ampicillin and induced with a final concentration of 1mM IPTG by incubating at 16°C for 16/17 h (shaking at 200 rpm).

After the completion of the incubation period, recombinant cells were harvested by centrifugation and resuspended in lysis buffer containing 20mM sodium phosphate (pH 7.4), 0.5M NaCl, and 40mM imidazole, and disruption of cells by grinding with of 10µM PMSF (Sigma) in the presence of liquid nitrogen. After centrifugation at 4°C and 20,000g for 30 minutes, the pellet includes harvested cell debris, and the supernatant consists of recombinant His-tagged fusion protein.

The supernatant containing the recombinant His-tagged fusion protein was purified using the Nickel beads column (Quaigen) by following the manufacturer’s instructions. The purified protein was analyzed and estimated by using the Bradford assay. The purity of the protein was examined by analyzing the collected fractions on 12% SDS-PAGE and stained with Coomassie Brilliant Blue R-250. The relative protein concentration was determined using 280 nm absorbance measurements.

3.5.4 Protein estimation by Bradford reagent

Protein concentration was estimated spectrophotometrically using the Bradford method (1976). For this purpose, Bradford reagent was prepared by adding 2 ml of .2% Coomassie Brilliant blue G-250 and 4 ml of orthophosphoric acid in 14 ml of millipore water. Then 100 µl of purified protein was mixed with 1 ml of Bradford reagent which resulted in development of blue colour and absorbance was recorded at 595 nm.

3.6 ISOLATION AND QUANTIFICATION OF RNA

3.6.1 RNA isolation and primer design

Ribonucleic acid (**RNA**) is an important biological macromolecule and ubiquitously present in all living cell.The major function of this nucleic acid is synthesis of required protein, carrying genetic information from the DNA which itself contain information required for developing and sustaining life of the living organism.

The extracted pellets of cyanobaceria were freezed in liquid nitrogen and RNA was extracted by using the RNA Sure mini kit (Nucleo-pore) following the manufacturer’s instructions. RNA purification and quantification were performed with the help of NanoVue Plus spectrophotometer GE Healthcare Life Sciences. The gene related to primer usage were designed using PRIMER3 software.

3.6.2 cDNA synthesis

The formation of DNA from RNA template, via reverse transcription, results in complementary DNA (cDNA). The cDNA can be prepared using iScript cDNA synthesis kit of Bio-Rad which involves reverse transcription of one microgram of total RNA in a 20 µl reaction mixture. Reactions were performed in a total volume of 20 µl including 4 µL 5 x iScript reaction mixtures, 1 µL iScript reverse transcriptase, 1 µg freshly prepared RNA and nuclease free water. The reaction conditions were 22˚C for 5 min, 42˚C for 90 min, and 85˚C for 5 min.

3.6.3 Quantitative real time PCR

Total RNA was extracted from 50 mL culture (OD750nm 0.6) of *Anabaena* PCC 7120 before and after 1 day of heat (48°C for 30 min.), cold (16^0C for 30 min), and desiccation (30°C for10h) treatment using the RNASure mini kit (Nucleo-pore). One microgram of total RNA was reverse transcribed in a 20µl reaction mixture using the iScript cDNA synthesis kit (Bio-Rad).

For the transcript analysis, following gene specific primer sets were designed using primer3 software for reference rnpB gene. 15ng of cDNA extracted from each sample was used for quantitative real-time PCR (qRT-PCR).

Reactions were performed in a total volume of 20µl including 10 pmol of forward and reverse primers and 1x Sso fast evagreen qPCR super mix (Bio-Rad). Samples were amplified with the following program: 95°C for 2 min, 40 cycles at 95°C for 5 sec and 60°C for 30 sec, followed by 95°C for 5 sec. CFX-96 (Bio-Rad) was used for PCR and for detection of fluorescence change. Transcript levels were normalized to 16S transcript and calculated relative to 0 h using the 2−ΔΔCt method. The comparative ΔCt method was used to evaluate the relative quantities of each amplified product in the samples. The threshold cycle (Ct) was automatically determined for each reaction by the system set with default parameters. The specificity of the PCR was determined by melting curve analysis of the amplified products.

◆

Chapter IV

RESULTS AND DISCUSSION

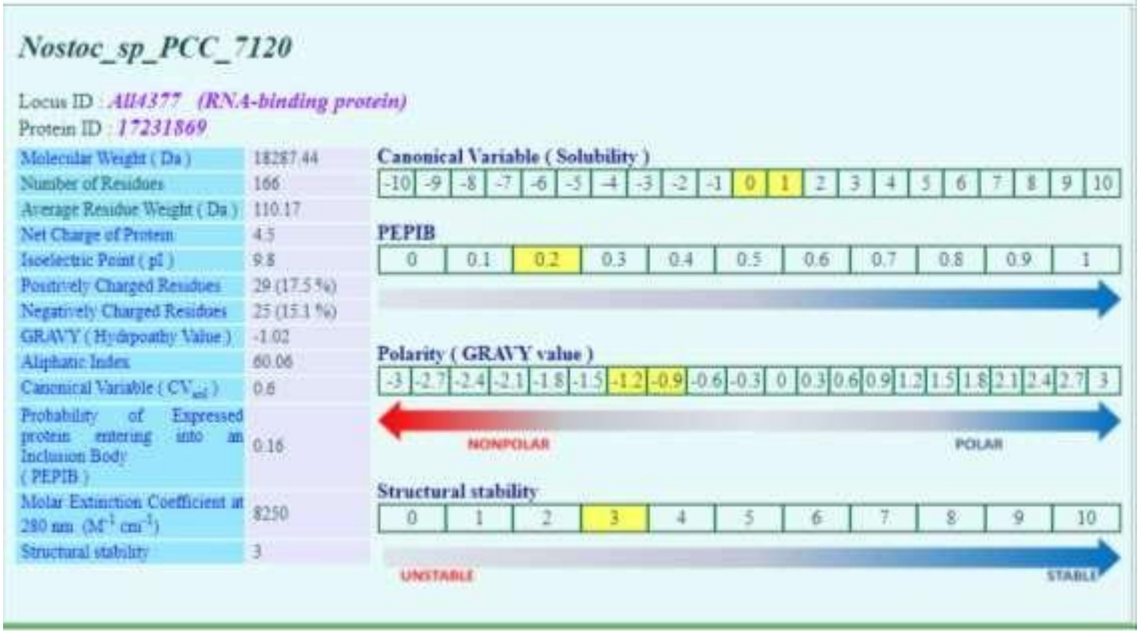
4.1 IN-SILICO ANALYSIS OF RNA BINDING PROTEIN ALL 4377

4.1.1 Physicochemical properties of Alr8073

The detailed analysis of the physical and chemical properties of All 4377 deciphered that this RNA binding protein has a total of 166 amino acids in its sequence. The molecular weight of the protein All 4377 was 18287.44 daltons. The theoretical iso-electric point was found to be 6.09. The maximum number of amino acids present in the sequence was found to be that of Lysine (K) (10.2%) and Threonine (T) (10.2 %). The least number of amino acids present in the sequence was Histidine (H) (0.6%) and Tryptophan (W) (0.6%). The total number of positively charged residue (Arginine +lysine) is 29 and the total number of negatively charged residue (Aspartic acid +Glutamic Acid) is 25.

The GRAVY value was calculated to be -1.024, which indicates non-polar nature of protein along with hydrophilic and soluble nature of target protein. The predicted aliphatic index was found to be 60.06 which indicate the least stability of the protein-expressing towards higher temperatures (Ikai et al., 1980). The probability of expressed protein entering into inclusion bodies (PEPIB) was 0.2, which clearly reveals the probability of our protein getting expressed into the soluble fraction (the supernatant) is high as compared to that of the protein which is entering into the inclusion bodies.

Studies show that there is a positive correlation between the structural stability and aliphatic amino acid content of the protein. The structural stability of our protein was calculated 3 which makes it highly unstable. However, the instability index of the protein was computed as 35.52 which provide an estimate of the high stability of All 4377 in the test tube. The other details of the Physico-chemical properties of the protein and its orthologs are presented in the snapshot below



Nostoc_sp_PCC_7120

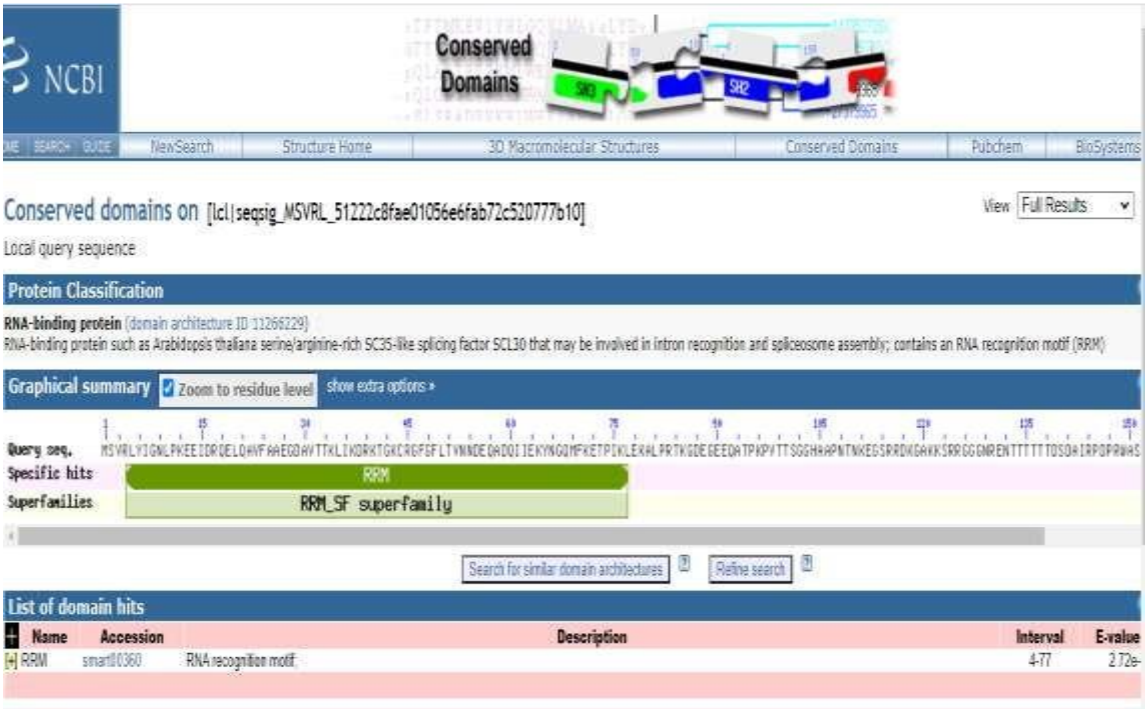
Locus ID : *All4377 (RNA-binding protein)*
Protein ID : *17231869*

Molecular Weight (Da)	18287.44
Number of Residues	166
Average Residue Weight (Da)	110.17
Net Charge of Protein	4.5
Isoelectric Point (pI)	9.8
Positively Charged Residues	29 (17.5 %)
Negatively Charged Residues	25 (15.1 %)
GRAVY (Hydropathy Value)	-1.02
Aliphatic Index	60.06
Canonical Variable (CV_{sol})	0.6
Probability of Expressed protein entering into an Inclusion Body (PEPIB)	0.16
Molar Extinction Coefficient at 280 nm (M^{-1} cm^{-1})	8250
Structural stability	3

Figure 8 - A snapshot showing physic-chemical properties of All 4377 from CyanoPhyChe database

4.1.2 Conserved domain search

Domains are the evolutionarily conserved units of proteins, which are commonly used for the classification of protein sequences and predicts their function. Therefore, different procedures are pre-requisite in order to select appropriate domain annotations involved for the protein. NCBI-conserved domains from the Conserved Domain Database (CDD) utilize ancestral information in the form of phylogenetic tree and structural information to construct hierarchies of homologous domain models so that they can easily reveal their evolutionary histories.In conformity with the NCBI's Conserved Domain Database (CDD), the gene All 4377 belongs to the member of RRM_SF Super families and pertains RNA Recognition Motif (RRM) Domain in the region starting from amino acid 4 to amino acid 78. Several members of this family have been implicated in various functions, one of which was characterized during my lab work and involves increasing gene expression and translation of gene all 4377 in order to provide resistance to cold stress in cells.



Conserved domains on [lcl|seqsig_MSVRL_51222c8fae01056e6fab72c520777b10]

Local query sequence

Protein Classification

RNA-binding protein (domain architecture ID 11266229)

RNA-binding protein such as Arabidopsis thaliana serine/arginine-rich SC35-like splicing factor SCL30 that may be involved in intron recognition and spliceosome assembly; contains an RNA recognition motif (RRM)

Graphical summary

List of domain hits

Name	Accession	Description	Interval	E-value
[+] RRM	smart00360	RNA recognition motif	4-77	2.72e-

Figure 9: **<u>Showing result of NCBI conserved domain analysis</u>**.

RRM domain

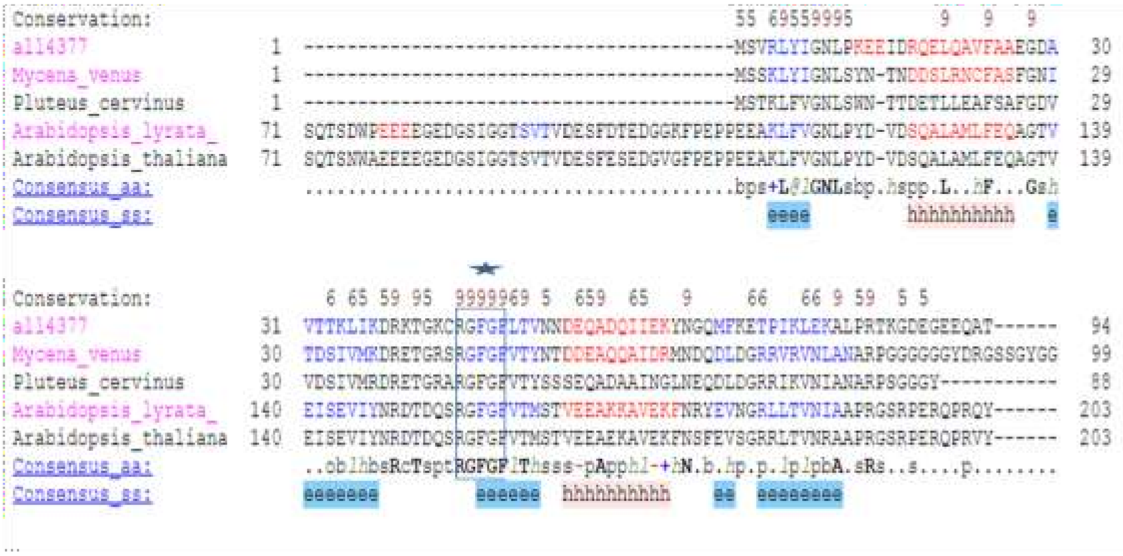
his is a SMART **RRM** domain (full annotation).

Position:	4 to 78
E-value:	3.91113522839262e-17 (HMMER2)

Figure 10 : <u>Conserved Domain analysis of ALL 4377 gene by SMART</u>

4.1.3 Multiple sequence alignment

The All 4377 protein sequences retrieved from CyanoBase was taken as input to find out the similar protein sequences across different species involved in RNA Binding Activity using BLASTP. Multiple sequence alignment of All 4377 revealed conserved Arginine, Glycine and phenylalanine amino acid residues that are common to RRM domain of other organisms. It has been suggested that these residues tends to play an important role in catalytic site of enzyme and through evolutionary time they remain conserved among members of different species. The active binding residues showed in the box and marked by blue star.All4377 from *Anabaena sp.* PCC 7120; *Arabidopsis lyrata and thaliana*, *Mycena venus* and *Pluteus cervinus.* The presence of RGFGF Amino acid are highly conserved sequence, which might play important roles as RNA recognition motif (RRM) in *Anabaena sp.*PCC7120.



```
Conservation:                                                       55 69559995       9   9   9
all4377               1   ----------------------------------------MSVRLYIGNLPKEEIDRQELQAVFAAEGDA   30
Mycena_venus          1   ----------------------------------------MSSKLYIGNLSYN-TNDDSLRNCFASFGNI   29
Pluteus_cervinus      1   ----------------------------------------MSTKLFVGNLSWN-TTDETLLEAFSAFGDV   29
Arabidopsis_lyrata_  71   SQTSDWPEEEEGEDGSIGGTSVTVDESFDTEDGGKFPEPPEEAKLFVGNLPYD-VDSQALAMLFEQAGTV  139
Arabidopsis_thaliana 71   SQTSNWAEEEEGEDGSIGGTSVTVDESFESEDGVGFPEPPEEAKLFVGNLPYD-VDSQALAMLFEQAGTV  139
Consensus_aa:             ........................................bps+L@lGNLsbp.hspp.L..hF...Gsh
Consensus_ss:                                                         eeee          hhhhhhhhh   e

                                         ★
Conservation:                 6 65 59 95 9999969 5  659  65   9    66   66 9 59  5 5
all4377              31   VTTKLIKDRKTGKCRGFGFLTVNNDEQADQIIEKYNGQMFKETPIKLEKALPRTKGDEGEEQAT------   94
Mycena_venus         30   TDSIVMKDRETGRSRGFGFVTYNTDDEAQQAIDRMNDQDLDGRRVRVNLANARPGGGGGGYDRGSSGYGG   99
Pluteus_cervinus     30   VDSIVMRDRETGRARGFGFVTYSSSEQADAAINGLNEQDLDGRRIKVNIANARPSGGGY-----------   88
Arabidopsis_lyrata_ 140   EISEVIYNRDTDQSRGFGFVTMSTVEEAKKAVEKFNRYEVNGRLLTVNIAAPRGSRPERQPRQY------  203
Arabidopsis_thaliana 140  EISEVIYNRDTDQSRGFGFVTMSTVEEAEKAVEKFNSFEVSGRRLTVNRAAPRGSRPERQPRVY------  203
Consensus_aa:             ..oblhbsRcTsptRGFGFlThsss-pApphl-+hN.b.hp.p.lplpbA.sRs..s....p........
Consensus_ss:             eeeeeee        eeeeee  hhhhhhhhhh    ee  eeeeeeee
...
```

Figure 10: **<u>Multiple sequence alignment of all 4377 with consensus sequence of RBP protein of other organisms</u>.**

4.1.4 Phylogenetic tree Analysis

Phylogenetic analyses made by using RRM Domain of different organisms revealed the close relationships of our protein *All 4377* of *Anabaena* PCC7120 with RNA Binding protein of plant system like *Mycena* , *Pluteus* and *Arabidopsis.*

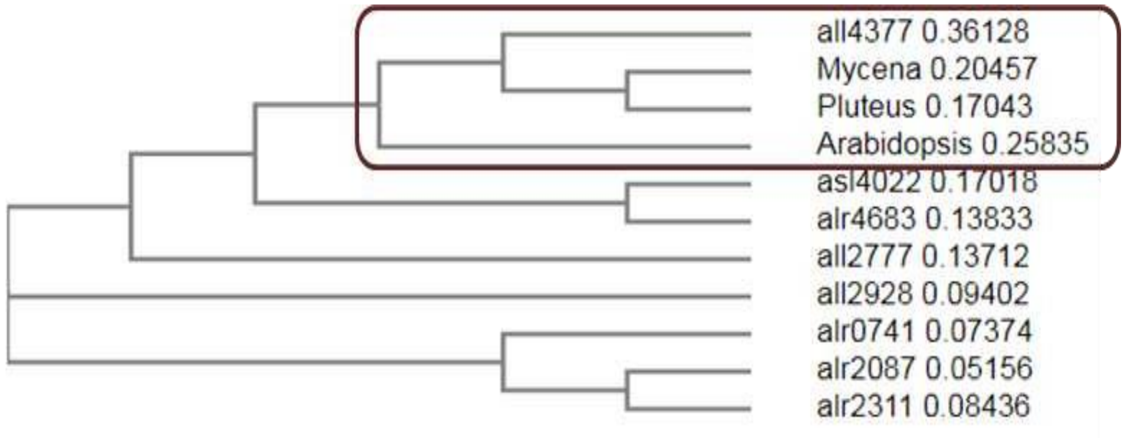


Figure 11: ***<u>Phylogenetic tree illustrating the relatedness of All4377 to putative homologues</u>***

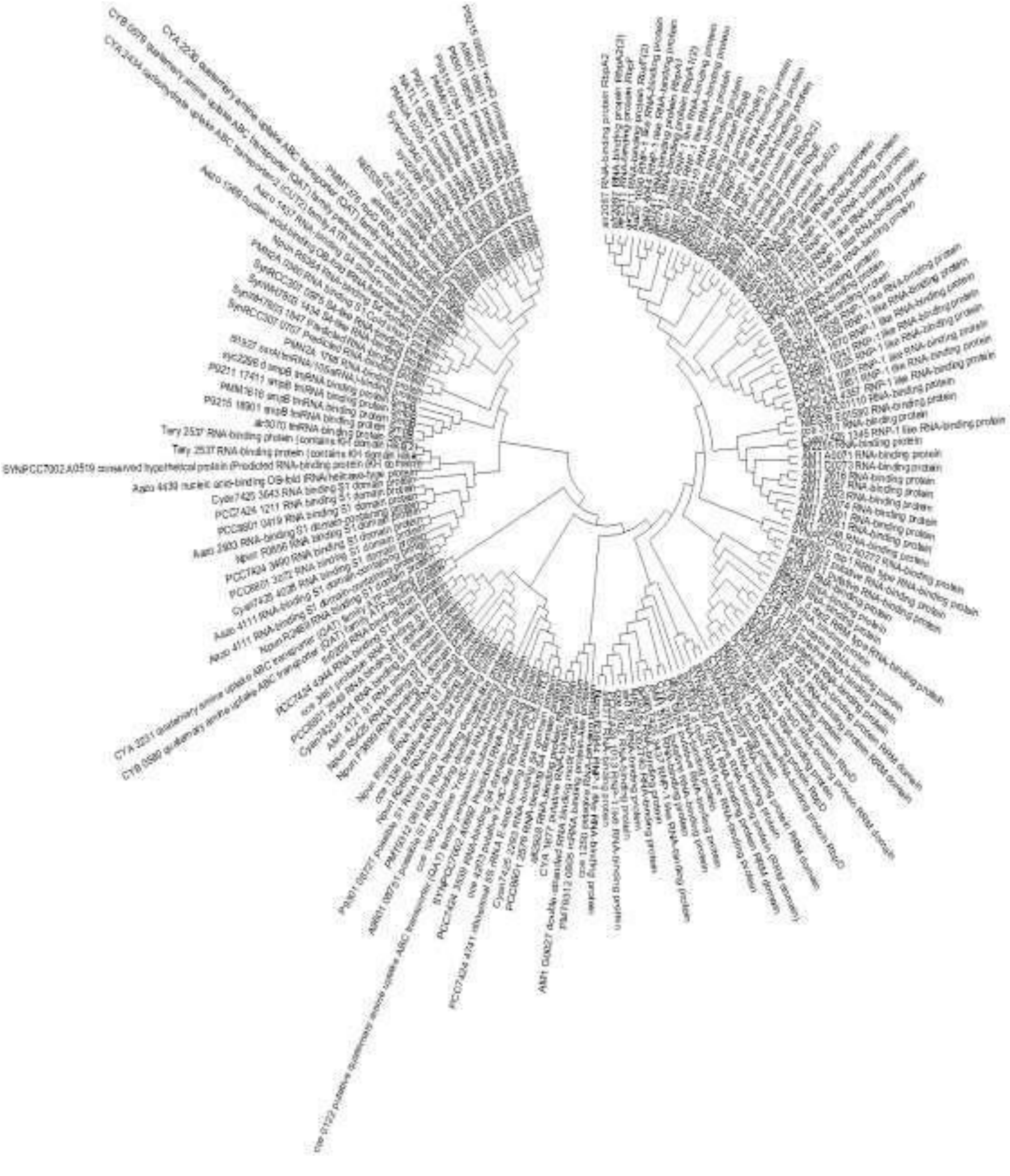


Figure 12: **<u>Phylogenetic analysis of All 4377 with consensus sequences of RNA Binding proteins of different organism</u>.**

4.1.5 Structure analysis of All 4377

Proteins play an important role in almost all living organism processes and act as the foundation of each and every cellular process. In order to predict the correct function of protein, analysis of protein structure is essential. Protein exist in three different forms or structure which includes primary structure, secondary structure and tertiary structure. Primary structure of All 4377 consists of a linear sequence of 166 amino acids which are connected in a linear sequence fashion having carboxyl group of one amino acid connected via a peptide bond with the amino group of the next amino acid.

Protein secondary structures basically the local configuration of protein’s amino acid in a polypeptide backbone. Secondary structure of protein exists in two different forms - α-helix (H) and β-strand (E), and one form present as irregular secondary structure type which is the coil region (C).Secondary structure analysis of the protein was performed as present in methods and material. Secondary structure of protein revealed the distribution of amino acid, it was observed that the distribution of amino acids in the coils is about 70.5%, in comparison to helices and Sheets which is about 6.7% and 22.8% respectively.

Protein tertiary structure or the 3-D model of the protein was visualized with the help of ITASSER which is a hierarchical protein structure modeling approach based on the secondary-structure enhanced Profile-Profile threading Alignment (PPA) and the iterative implementation of the Threading Assembly Refinement (TASSER) [http://zhang.bioinformatics.ku.edu/I-TASSER] program. The 3D model was refined using I-TASSER mod refiner and Swiss-PdbViewer (Guex and Peitsch, 1997) was used for energy minimization.

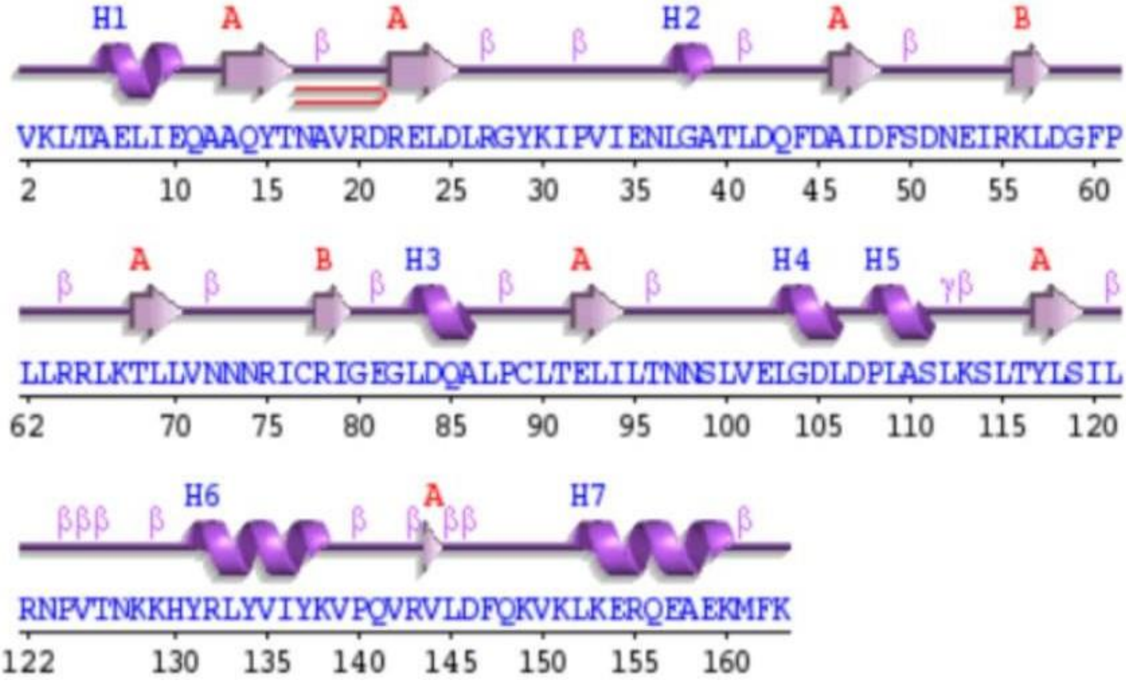


FIGURE 13. **Secondary structure prediction of all 4377 showing coil, helices and sheets.**

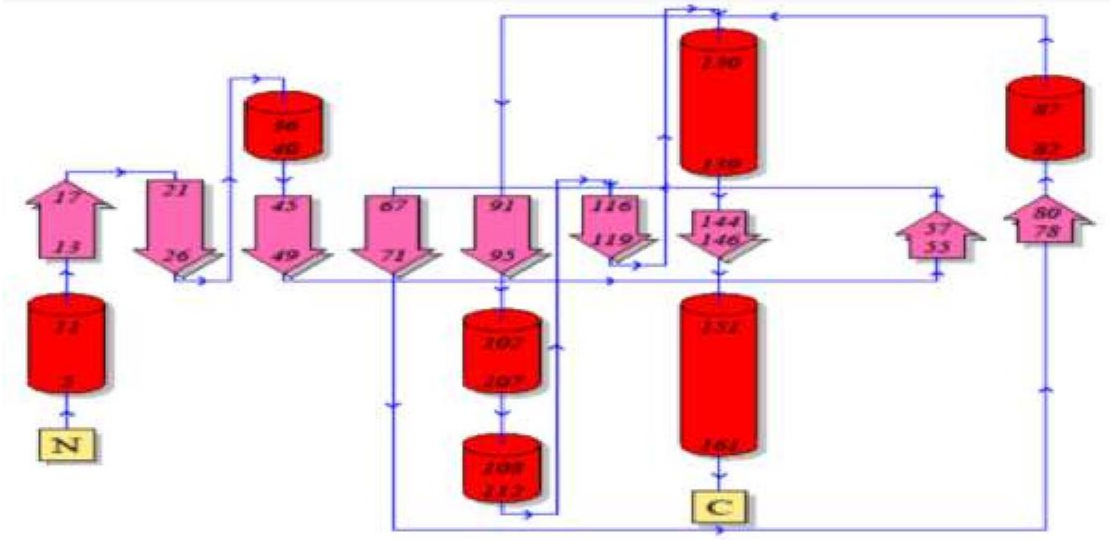
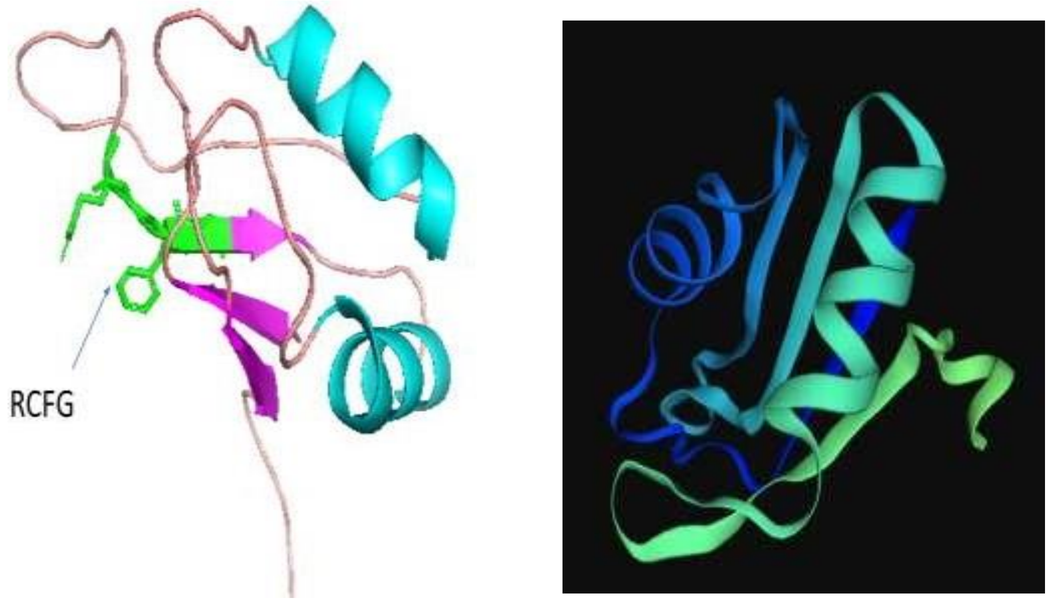


Figure 14. Topology structure of all 4377.



[A] [B]

FIGURE 15 – Predicted Tertiary structure of ALL4377 from [A] ITASEER and [B] SWISS model.

4.1.6 Structure validity and verification

the structural validation and quality assessment of this model were done using RAMPAGE. The model prepared after homology modeling with the help of swiss modeler which is then followed by the energy minimization using Modrefiner related to several structural validation and quality assessment check using RAMPAGE, RMSD and Z-score. RAMPAGE output was analyzed in the form of Ramachandran plot showed that 89.2% residues of these proteins were in the most favored regions, only 10.23% residues were in allowed regions and 0% amino acid are present in disallow region **(Fig: 7b)**. The quality of the model was determined by superimposition of the target and the template structures and finally by measuring the root mean squared deviation (RMSD). All these features attested the quality of proteins.

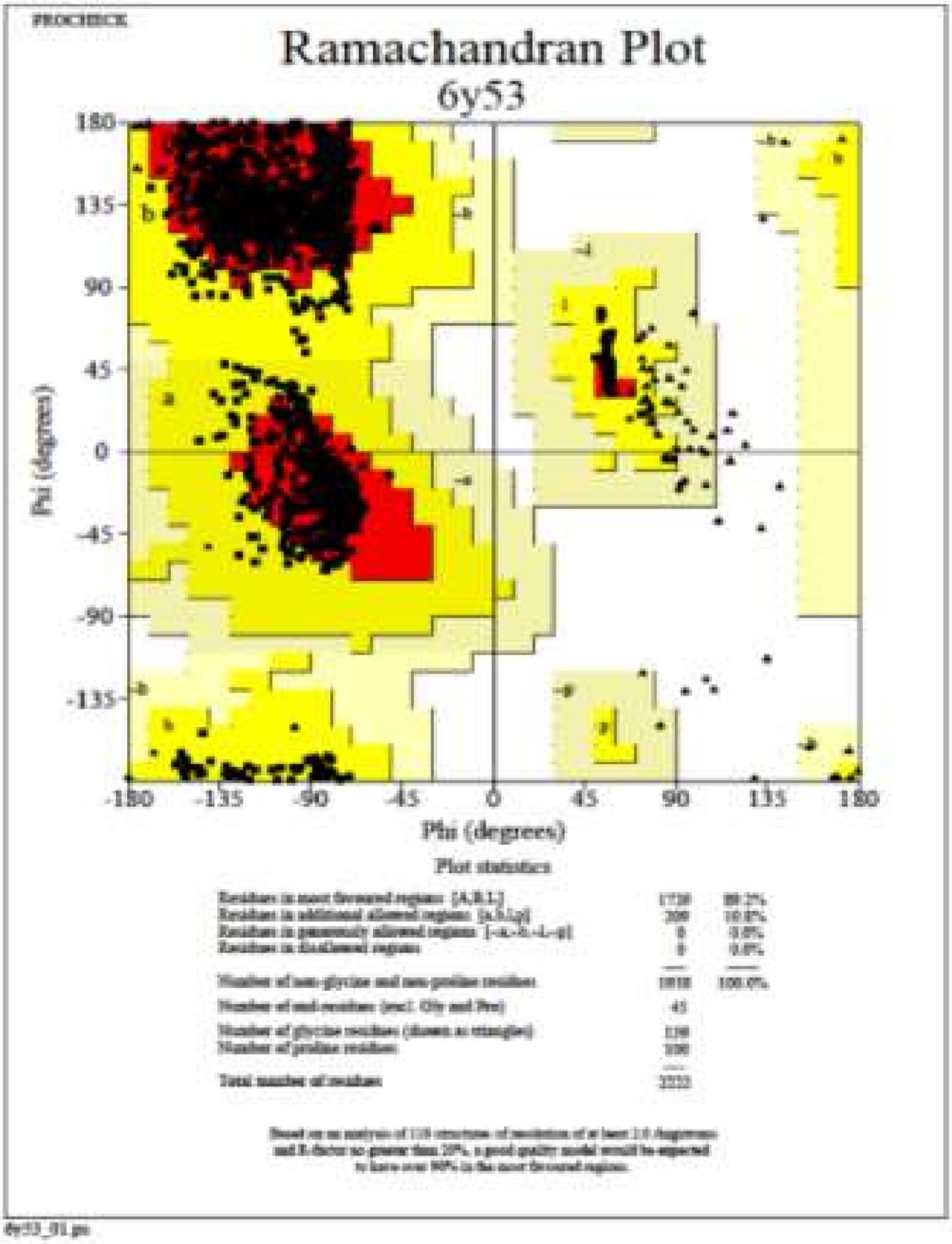


Figure 16 - Ramachandran plot an analysis of All 4377

4.6. Functional analysis

Functional analysis revealed 4 compatible and potential interacting partners of RNA binding protein in the protein interaction network as determined by STRING analysis. The query protein All 4377 (Q8YP25), appeared to contain nucleic acid binding activity. The closest interacting partner of the protein having the shortest node among all was found rbpF (RNA Binding activity), a unit while the most distant interacting protein was found to be gpml (transition metal ion binding)which helps in catalyzing the inter-conversion of 2-phosphoglycerate and 3-phosphoglycerate and gltX (Glutamate t-RNA ligase) Catalyzes the attachment of glutamate to tRNA(Glu) in a two-step reaction: glutamate is first activated by ATP to form Glu-AMP which is transferred to the acceptor end of tRNA(Glu) and belongs to the class-I amino acyl-tRNA synthetase family.

The active site amino acid residues of All 4377, was determined with the help of COFACTOR, and were found to be Histidine (H), Glycine (G), and Phenylalanine (F). It is also important to note that there is no direct method present for the detection of the active site of protein. *In silico* functional analysis reveal as the functional activities of protein which involves identifying conserved domains, studying protein-protein interactions, and detecting structural motifs of protein.

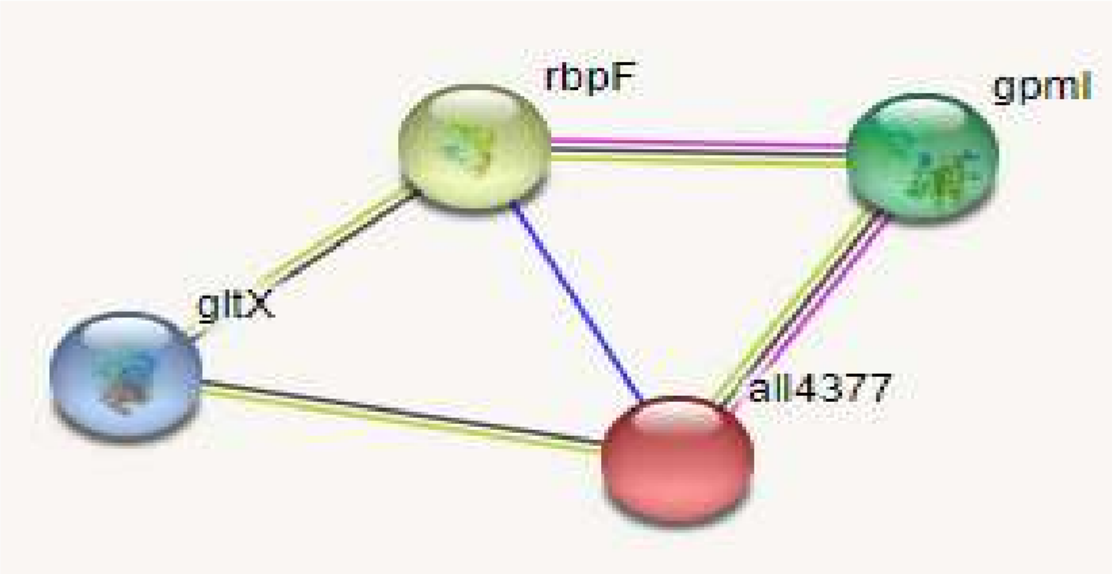
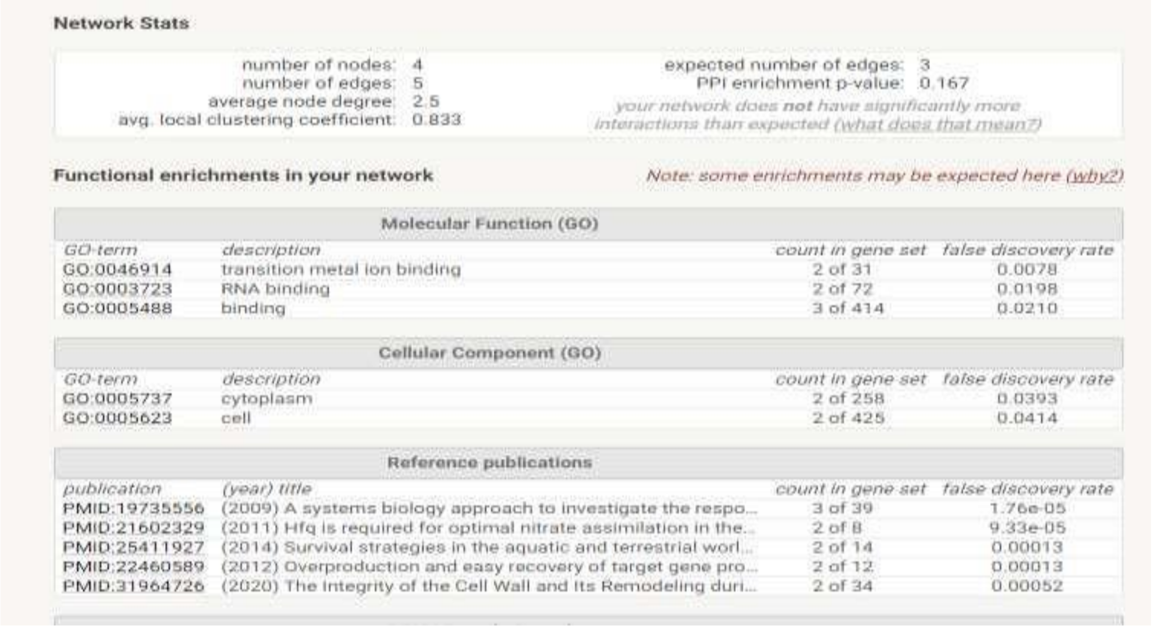


Figure 17 – Snapshot of protein-protein interaction of RNA binding protein from STRING database



Network Stats

number of nodes: 4
number of edges: 5
average node degree: 2.5
avg. local clustering coefficient: 0.833
expected number of edges: 3
PPI enrichment p-value: 0.167
*your network does **not** have significantly more interactions than expected (what does that mean?)*

Functional enrichments in your network

Note: some enrichments may be expected here (why?)

Molecular Function (GO)

GO-term	*description*	*count in gene set*	*false discovery rate*
GO:0046914	transition metal ion binding	2 of 31	0.0078
GO:0003723	RNA binding	2 of 72	0.0198
GO:0005488	binding	3 of 414	0.0210

Cellular Component (GO)

GO-term	*description*	*count in gene set*	*false discovery rate*
GO:0005737	cytoplasm	2 of 258	0.0393
GO:0005623	cell	2 of 425	0.0414

Reference publications

publication	*(year) title*	*count in gene set*	*false discovery rate*
PMID:19735556	(2009) A systems biology approach to investigate the respo...	3 of 39	1.76e-05
PMID:21602329	(2011) Hfq is required for optimal nitrate assimilation in the...	2 of 8	9.33e-05
PMID:25411927	(2014) Survival strategies in the aquatic and terrestrial worl...	2 of 14	0.00013
PMID:22460589	(2012) Overproduction and easy recovery of target gene pro...	2 of 12	0.00013
PMID:31964726	(2020) The integrity of the Cell Wall and its Remodeling duri...	2 of 34	0.00052

Figure 18 – Snapshot showing associated protein function in protein – protein interaction.

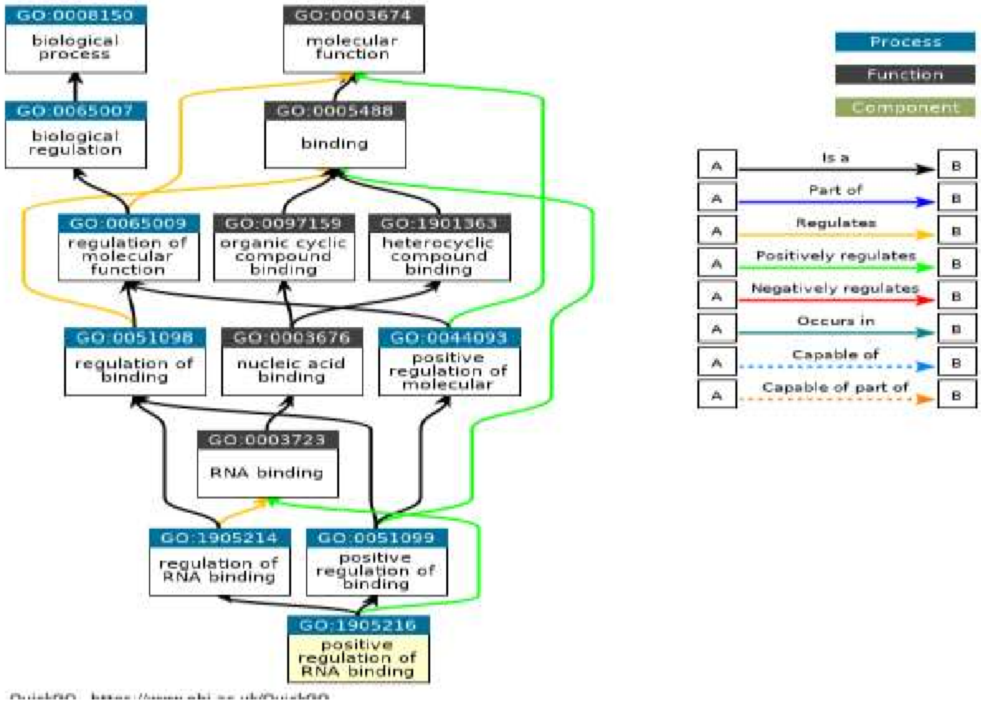


Figure 19: Gene ontology of All 4377 showing different molecular function

4.2 MOLECULAR CLONING AND OVER-EXPRESSION ANALYSIS OF ALL 4377

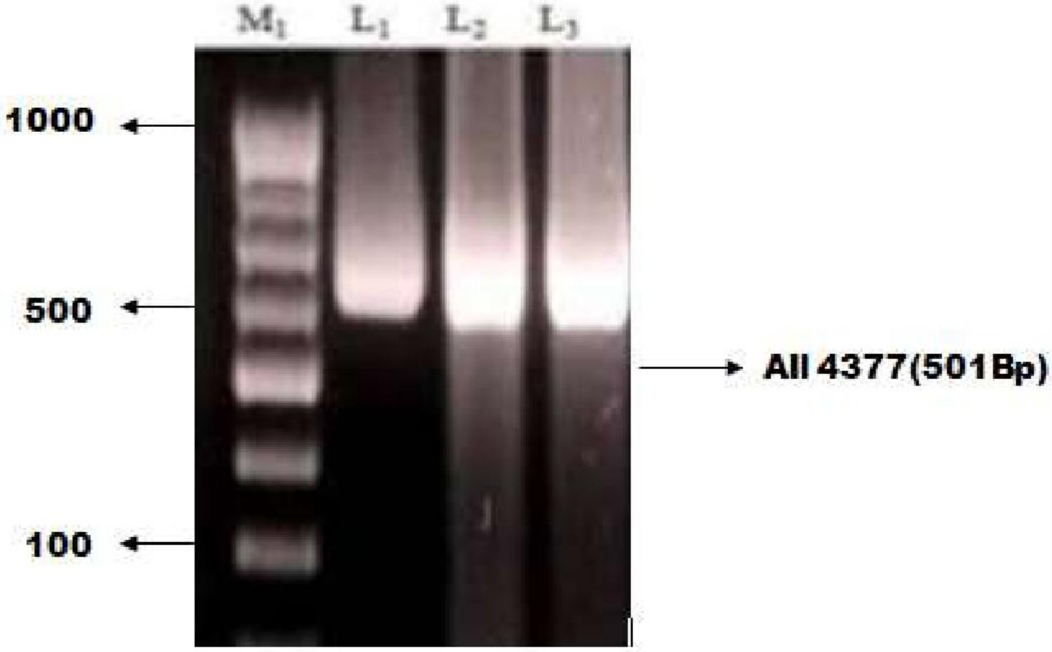
4.2.1 Isolation of genomic DNA from *Anabaena* sp. PCC7120

Genomic DNA from cyanobacterium *Anabaena* sp. PCC7120 was isolated and analyzed on 0.8% agarose gel and its concentration was 0.8µg/ml at 260 nm. The purity of DNA was checked by measuring A260/A280 ratio in a Nanoview Plus spectrophotometer (GE Healthcare Life Sciences) and it turned out to be 1.86.

FIGURE 20 : Agarose gel electrophoresis (0.8%) showing genomic DNA (g-DNA) of *Anabaena* sp. PCC7120

4.2.2 PCR amplification of *ALL 4377* gene

The PCR-amplified product of *All 4377* was approx. 518 bps when compared with 100 bps DNA ladder (NEB), which matched with the theoretical length of the *All 4377*(501 bps). Amplified PCR product was extracted, purified and digested with restriction endonuclease (Nde1 and Xho1) to clone into pET-28a vector.



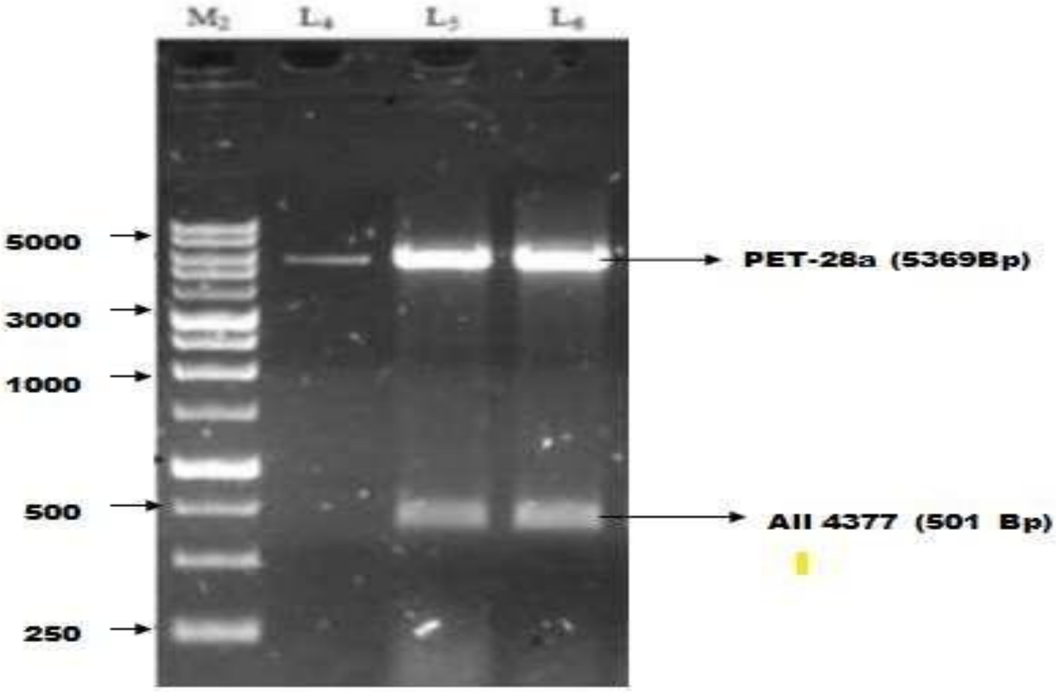


FIGURE 21: Colony PCR amplification from transformed colonies. Lane M1 showing 100 bp DNA ladder. Lanes L1-L6: amplified colony PCR product of *alr*8073 using gene specific primer

4.2.3 Cloning of *All 4377* into pET-28a vector

For cloning of *alr8073* gene into pET-28a vector, digested PCR product of *All 4377* and pET-28a vector were ligated together and transformed into *E. coli* (DH5α) cells. Transformed cells were spread on LB plate containing ampicillin (100µg/ml) and then colony PCR was performed using All 4377 gene specific primer in order to screen positive clones from the cells grown on LB plate.

4.2.4 Confirmation of positive clones by Double digestion

Transformed cells that have shown positive result of colony PCR were further confirmed with double digestion of the isolated recombinant vector with the help of Nde1 and Xho1 restriction endonuclease enzymes. Double digestion of recombinant vector produced two bands, one of 5.3 kb corresponding to molecular weight of double digested pET-28a vector and another at 501 bps. Corresponding to PCR product of All 4377. After that positive recombinant vectors were transformed in to *E. coli* (Bl-21) cells for overexpression analysis.

4.2.5 Over expression analysis of All 4377 using SDS-PAGE

Positive BL21 recombinant cells were grown overnight at 37 ºC in LB media carrying ampicillin (100µg/ml). This overnight grown culture was spiked in to LB media for secondary growth and further cells were induced with 0.1mM IPTG for 15-16 h. An overnight exposure of *E. coli* cells resulted in an increased expression on SDS-PAGE (12%) as compared with the uninduced sample.

The RNA binding protein All4377 is composed of 166 amino acids with molecular weight of approx 18.2 kDa. The expressed protein band on SDS- PAGE was in same range as of its theoretical molecular weight when compared with protein molecular weight. Induced cells were used for purification of All 4377 using Ni-NTA His-Trap column. Purified protein showed only a single band when run on SDS-PAGE (12%). The molecular weight of the purified protein was also in the same range as that of theoretical molecular weight of protein.

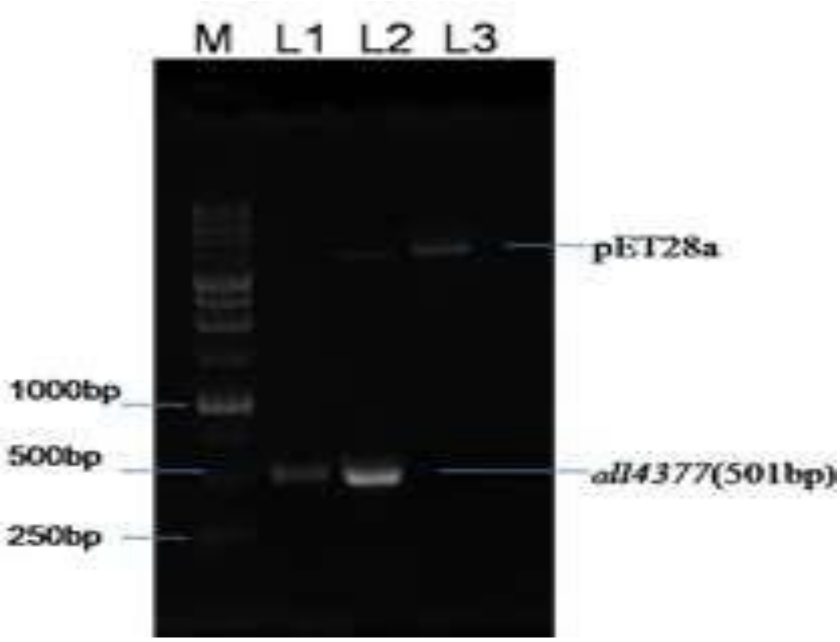


Fig. 22 : ***<u>Molecular cloning of All4377 in E. coli.</u>*** **<u>Agarose gel-separation showing double digestion of *all4377 inserted in pET-28a*. M: DNA ladder</u>**

(A) L1:PCR product of all4377

(B) L2: double digested recombinant clone containing 501bp fragment & 5.3kb pET-28 vector,

(C) L3: empty vector pET28a.

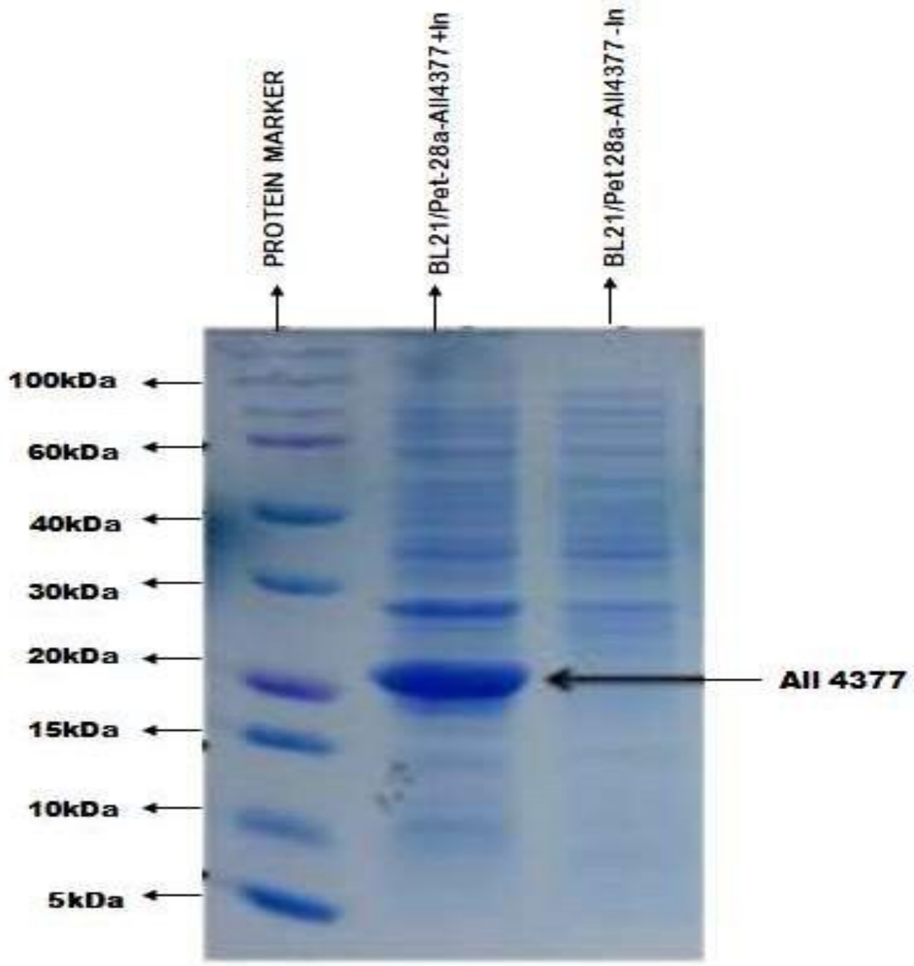


FIGURE 23 : <u>SDS-PAGE (12%) analysis showing recombinant protein expression in transformed *E.coli*</u>

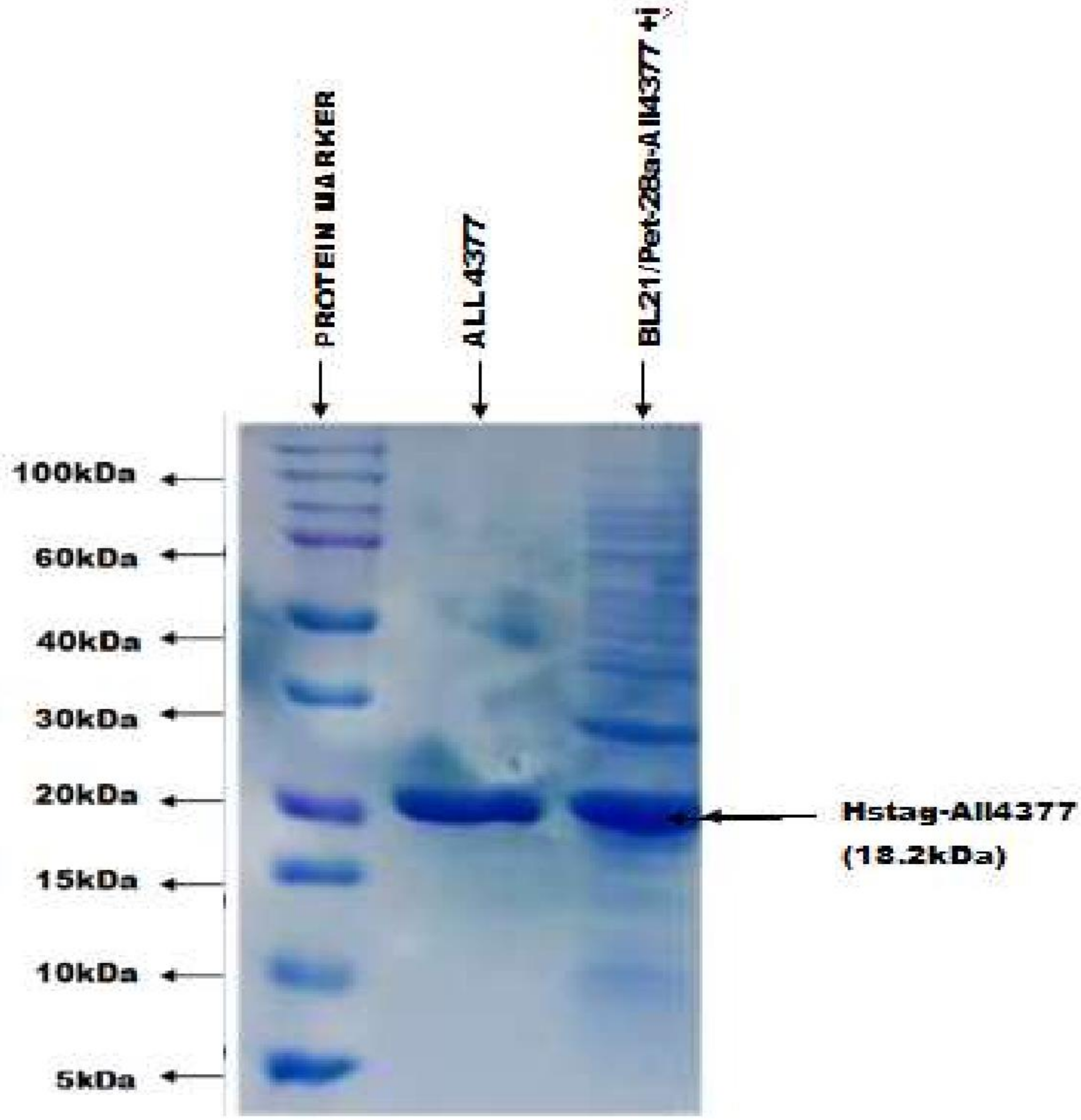


FIGURE 24 : **<u>SDS-PAGE (12%) analysis of purified recombinant protein that has been extracted from transformed *E. coli*, and purified using Ni- NTA column</u>**.

• DISCUSSION

The present study finding concluded that majority of the cyanobacterial RBP genes are regulated by low temperature which suggests that our RNA binding proteins belongs to a new class of stress-proteins. Analogous with the heat-shock protein which act as molecular chaperons for proteins, these RBP protein might function as RNA chaperons to maintain correct higher-order structures of RNA.

These novel RBP belongs to a class of newly characterized class of RNA binding proteins. Their abundance presence can be easily detected in both filamentous and unicellular species of cyanobacteria. Apart from cyanobacteria they can be observed in other prokaryotic and eukaryotic species including plant and humans as well. The present study illuminates the characterization of the novel protein All4377 from the nitrogen-fixing cyanobacteria Anabaena PCC 7120 with the help of dry lab which includes bioinformatics along with wet lab approaches.

The bioinformatics analysis of the ALL 4377 protein reveled that it belongs to the member of RRM_SF Super families and pertains RNA Recognition Motif (RRM) Domain in the region starting from amino acid 4 to amino acid 78. Several members of this family have been implicated in various functions, one of which was characterized during my lab work and involves increasing gene expression and translation of gene all 4377 in order to provide resistance to cold stress in cells. The present protein is a short peptide protein having 166 amino acid and have a molecular weight of 18.78 Kda. Multiple sequence alignment of All 4377 showed the presence of conserved Arginine, Glycine and phenylalanine amino acid residues that are common to RRM domain of other organisms. It has been suggested that these residues tends to play an important role in catalytic site of enzyme Further functional validation of above *Insilco* prediction ,ALL 4377 was cloned in to pET-28a vector, expressed and purified from *E. coli* BL21 cells.

The present study clearly shows that both the transcripts and the protein products of RBP genes are in proper regulation by low temperature. At the time of induction, RBP transcripts was accumulated rapidly in very beginning during the first 10 min. after the temperature was shifted. However, some discrepancies between the increase of transcripts and the accumulation of proteins were also noted. These discrepancies might indicate a post-transcriptional regulation of these genes, such as translational regulation. This point should be studied in more detail in future.

Chaperon proteins possess hydrophobic patches which engaged in quality control of mis-folded client proteins and thereby control the cellular conformational ensemble of the proteome. Hence this analysis suggests that ALL4377 might be work as chaperon. Up-regulation of ALL4377 genes has been observed in *Anabaena* sp. PCC7120 under cold stresses.

RBP pathway consist of its expression which includes RNA metabolism originates with transcription and further proceeds with post-transcriptional RNA metabolism, which involves RNA splicing, translation, 3’ end-formation, RNA import and export along with their degradation ,decay ,and surveillance. Traditionally, gene expression has been considered separate and independent post-transcriptional steps. This RBP pathway keeps gene expression in check which regulates the cellular temperature homeostasis. This resulted in the survival of cyanobacteria under cold stress conditions.

In addition to this growth of transformed *E. coli* cells with pET-28a–*all 4377* was found to be better than cells transformed with pET-28a (empty vector) under cold stresses. It can be easily observed that cells transformed with pET-28a–all4377 can tolerate up to 15^{0}C whereas cells having empty vector failed to grow in such a low temperature reflecting the ability of the gene all 4377 playing important role in providing survival advantage under lower temperature to *E. coli* BL21 cells. These results vividly demonstrated that enhanced expression of ALL4377 gene from *Anabaena* sp. PCC 7120 and enhanced growth of recombinant cells under various environmental clues suggests its role in cold stress management directly or indirectly.

◆

Chapter V

CONCLUSION

This review significantly showed the high diversity of RRM–RNA recognition, binding and underlines the importance of structural data which helps in characterization of RRM proteins. This small peptide domain has the unique ability to recognize and attaches with RNA having weak or strong affinity and specificity, to interact with RNA of different length and to contact several partners at the same time.

In silico and biochemical analysis revealed that All 4377 from *Anabaena* sp. PCC7120 belongs to RRM superfamily. Biochemical pathways in the cell often involve a highly sophisticated network of intermolecular interactions. The specificity of RRMs to interact with either nucleic acid or with other proteins is best explanation for RRM protein abundance and achieving a central function in the cell. Data received after structural analysis validates the unifying ability of RRMs to interact with each and every element and support in the establishment of the structure which includes beta-strands, loops, alpha-helices which perfectly explains the extreme conservation of this motif during evolution. Moreover, the short size of the RRM domain assist its incorporation into protein genes that helps in the evolution of protein towards a new and more sophisticated function. All4377 has conserve active site residues (ARG-GLY-PHY) which is multitasking.

A unique ability of this proteins is that they can switch from one function to another depending upon environmental conditions need to be further attestation. Moreover transcript and growth analysis under low temperature stress confirm its role in protecting cell from cold shock. So, we have deciphered the role of All4377 as unique type RNA binding protein reported for the first time in cyanobacterium. The present study, therefore, concluded that ALL 4377 is a novel RNA binding protein which also functions as cold shock protein, hence recommended ALL 4377 over expressing *Anabaena* sp. PCC7120 may be an excellent bio-fertilizer in paddy fields suffering from cold stresses.

* * *

REFERENCES

Adam, S. A., Nakagawa, T. E. R. R. Y., Swanson, M. S., Woodruff, T. K., & Dreyfuss, G. I. D. E. O. N. (1986). mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Molecular and cellular biology*, *6*(8), 2932-2943.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool (1990). *Journal of Molecular Biology; 215*: 403–410.

Apte, S. K. (2001). Coping with salinity/water stress: cyanobacteria show the way. *Proceedings-Indian National Science Academy Part B*, *67*(5), 285-309.

Arun, P. P. S., Bakku, R. K., Subhashini, M., Singh, P., Prabhu, N. P., Suzuki, I., & Prakash, J. S. (2012). CyanoPhyChe: a database for physico-chemical properties, structure and biochemical pathway information of cyanobacterial proteins. *PloS One*, *7*(11), e49425.

Arun, P. P. S., Yarlagadda, V., Laxmi, G. V., & Salla, S. (2020). Computational analysis of the hypothetical protein P9303_05031 from marine Cyanobacterium Prochlorococcus marinus MIT 9303. *Journal of integrative bioinformatics*, *17*(1). P.20180087

Atlas, R. M. (1998). Microbial ecology: fundamentals and applications. *Pearson Education India*. p. 1–704.

Bandziulis, R. J., Swanson, M. S., & Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators. *Genes and Developement*, *3*(4), 431-437.

Barria, C., Malecki, M., & Arraiano, C. M. (2013). Bacterial adaptation to cold. *Microbiology*, *159*(Pt_12), 2437-2443.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., ... & Sonnhammer, E. L. (2002). The Pfam protein families database. *Nucleic acids research*, *30*(1), 276-280.

Birney, E., Kumar, S., & Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic acids research*, *21*(25), 5803-5816.

Birney, E., Kumar, S., & Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic*

acids research, *21*(25), 5803-5816.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., ... & Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science*, *277*(5331), 1453-1462.

Boyer J. S. (1982). Plant productivity and environment. *Science*, *218*(4571), 443-448.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science*, *273*(5278), 1058-1073.

Cavicchioli, R. (2016). On the concept of a psychrophile. *The International Society for Microbial Ecology Journal*, *10*(4), 793-795.

Chamot, D., & Owttrim, G. W. (2000). Regulation of cold shock-induced RNA helicase gene expression in the cyanobacterium Anabaena sp. strain PCC 7120. *Journal of bacteriology*, *182*(5), 1251-1256.

Chattopadhyay, M. K., Raghu, G., Sharma, Y. V. R. K., Biju, A. R., Rajasekharan, M. V., & Shivaji, S. (2011). Increase in oxidative stress at low temperature in an Antarctic bacterium. *Current microbiology*, *62*(2), 544-546.

Chrismas, N. A., Anesio, A. M., & Sánchez-Baracaldo, P. (2015). Multiple adaptations to polar and alpine environments within cyanobacteria: a phylogenomic and Bayesian approach. *Frontiers in microbiology*, *6*, 1070.

Cléry, A., Blatter, M., & Allain, F. H. (2008). RNA recognition motifs: boring? Not quite. *Current opinion in structural biology*, *18*(3), 290-298.

De Maayer, P., Anderson, D., Cary, C., & Cowan, D. A. (2014). Some like it cold: understanding the survival strategies of psychrophiles. *EMBO reports*, *15*(5), 508-517.

Dobretsov, N., Kolchanov, N., Rozanov, A., & Zavarzin, G. (Eds.). (2008). *Biosphere origin and evolution* (pp. 1-427). Springer US. Pp.25–42.

Dreyfuss, G., Swanson, M. S., & Piñol-Roma, S. (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends in biochemical sciences*, *13*(3), 86-91.

El-Kest, S. E., & Marth, E. H. (1992). Freezing of Listeria monocytogenes and other microorganisms: a review. *Journal of food protection*, *55*(8), 639-648.

24.Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D.,

& Venter, J. C. (1995). The minimal gene complement of Mycoplasma genitalium. *Science*, *270*(5235), 397-404.

25. Fraser, C. M., Norris, S. J., Weinstock, G. M., White, O., Sutton, G. G., Dodson, R., & Venter, J. C. (1998). Complete genome sequence of Treponema pallidum, the syphilis spirochete. *Science*, *281*(5375), 375-388.
26. Garcia-Pichel, F., Belnap, J., Neuer, S., & Schanz, F. (2003). Estimates of global cyanobacterial biomass and its distribution. *Algological Studies*, *109*(1), 213.
27. Gerber, A. P., Herschlag, D., Brown, P. O., & Eddy, S. (2004). Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS biology*, *2*(3), e79.
28. Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, *582*(14), 1977-1986.
29. Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, *582*(14), 1977-1986.
30. Gomez-Porras, J., Lewinski, M., Riaño-Pachón, D. M., & Staiger, D. (2011). Molecular evolution of RRM-containing proteins and glycine-rich RNA-binding proteins in plants. *Nature Precedings*, 1-1.
31. Haines, R. B. (1938). The effect of freezing on bacteria. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *124*(837), 451-463.
32. Huang, Y., & Steitz, J. A. (2005). SRprises along a messenger’s journey. *Molecular cell*, *17*(5), 613-615.
33. Jones, P. G., Cashel, M., Glaser, G., & Neidhardt, F. C. (1992). Function of a relaxed-like state following temperature downshifts in Escherichia coli. *Journal of Bacteriology*, *174*(12), 3903-3914.
34. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., ... & Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA research*, *3*(3), 109-136.
35. Keene, J. D. (2007). RNA regulons: coordination of post-transcriptional events. *Nature Reviews Genetics*, *8*(7), 533-543.
36. Keene, J. D. (2007). RNA regulons: coordination of post-transcriptional events. *Nature Reviews Genetics*, *8*(7), 533-543.

37. Kenan, D. J., Query, C. C., & Keene, J. D. (1991). RNA recognition: towards identifying determinants of specificity. *Trends in biochemical sciences*, *16*, 214-220.
38. Kim, M. Y., Hur, J., & Jeong, S. J. (2009). Emerging roles of RNA and RNA-binding protein network in cancer cells. *BMB reports*, *42*(3), 125-130.
39. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G. O., Azevedo, V., & Yoshikawa, H. (1997). The complete genome sequence of the gram-positive bacterium Bacillus subtilis. *Nature*, *390*(6657), 249-256.
40. Lee, A. L. Kanaar, R., Rio, D. C., & Wemmer, D. E. (1994). Resonance assignments and solution structure of the second RNA-binding domain of sex-lethal determined by multidimensional heteronuclear magnetic resonance. *Biochemistry*, *33*(46), 13775-13786.
41. Los, D. A., Zorina, A., Sinetova, M., Kryazhov, S., Mironov, K., & Zinchenko, V. V. (2010). Stress sensors and signal transducers in cyanobacteria. *Sensors*, *10*(3), 2386-2415.
42. Lukong, K. E., Chang, K. W., Khandjian, E. W., & Richard, S. (2008). RNA-binding proteins in human genetic disease. *Trends in Genetics*, *24*(8), 416-425.
43. Lukong, K. E., Chang, K. W., Khandjian, E. W., & Richard, S. (2008). RNA-binding proteins in human genetic disease. *Trends in Genetics*, *24*(8), 416-425.
44. Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nature reviews Molecular cell biology*, *8*(6), 479-490.
45. Maccario, L., Sanguino, L., Vogel, T. M., & Larose, C. (2015). Snow and ice ecosystems: not so extreme. *Research in microbiology*, *166*(10), 782-795.
46. Maniatis, T., & Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature*, *416*(6880), 499-506.
47. Maris, C., Dominguez, C., & Allain, F. H. T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS journal*, *272*(9), 2118-2131.
48. Maruyama, K., Sato, N., & Ohta, N. (1999). Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins. *Nucleic acids research*, *27*(9), 2029-2036.
49. Maruyama, K., Sato, N., & Ohta, N. (1999). Conservation of structure and cold-regulation of RNA-binding proteins in cyanobacteria: probable convergent evolution with eukaryotic glycine-rich RNA-binding proteins. *Nucleic acids research*, *27*(9), 2029-2036.
50. Moore, M. J., & Proudfoot, N. J. (2009). Pre-mRNA processing reaches back totranscription

and ahead to translation. *Cell*, *136*(4), 688-700.

51. Morita, R. Y. (1975). Psychrophilic bacteria. *Bacteriological reviews*, *39*(2), 144-167.
52. Mulligan, M. E., Jackman, D. M., & Murphy, S. T. (1994). Heterocyst-forming filamentous cyanobacteria encode proteins that resemble eukaryotic RNA-binding proteins of the RNP family. *Journal of molecular biology*, *235*(3), 1162-1170.
53. Orphanides, G., & Reinberg, D. (2002). A unified theory of gene expression. *Cell*, *108*(4), 439-451.
54. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, *372*(6505), 432-438.
55. Phadtare, S. (2004). Recent developments in bacterial cold-shock response. *Current issues in molecular biology*, *6*(2), 125-136.
56. Phadtare, S., & Inouye, M. (2004). Genome-wide transcriptional analysis of the cold shock response in wild-type and cold-sensitive, quadruple-csp-deletion strains of Escherichia coli. *Journal of bacteriology*, *186*(20), 7007-7014.
57. Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends in genetics*, *16*(6), 276-277.
58. Rodrigues, D. F., & Tiedje, J. M. (2008). Coping with our cold planet. *Applied and environmental microbiology*, *74*(6), 1677-1686.
59. Rozanov, A. Y., & Astafieva, M. M. (2009). The evolution of the early precambrian geobiological systems. *Paleontological Journal*, *43*(8), 911-927.
60. Sanford, J. R., Coutinho, P., Hackett, J. A., Wang, X., Ranahan, W., & Caceres, J. F. (2008). Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. *PloS one*, *3*(10), e3369.
61. Sanford, J. R., Ellis, J. D., Cazalla, D., & Cáceres, J. F. (2005). Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. *Proceedings of the National Academy of Sciences*, *102*(42), 15042- 15047.
62. Sanford, J. R., Gray, N. K., Beckmann, K., & Cáceres, J. F. (2004). A novel role for shuttling SR proteins in mRNA translation. *Genes & development*, *18*(7), 755-768.
63. Sato, N. (1994). A cold-regulated cyanobacterial gene cluster encodes RNA-binding protein and ribosomal protein S21. *Plant molecular biology*, *24*(5), 819-823.
64. Sato, N. (1994). A cold-regulated cyanobacterial gene cluster encodes RNA-binding protein

and ribosomal protein S21. *Plant molecular biology*, *24*(5), 819-823.

65. Sato, N. (1995). A family of cold-regulated RNA-binding protein genes in the cyanobacterium Anabaena variabilis M3. *Nucleic Acids Research*, *23*(12), 2161-2167.
66. Sato, N., & Maruyama, K. (1997). Differential regulation by low temperature of the gene for an RNA-binding protein, rbpA3, in the cyanobacterium Anabaena variabilis strain M3. *Plant and cell physiology*, *38*(1), 81-86.
67. Sato, N., & Maruyama, K. (1997). Differential regulation by low temperature of the gene for an RNA-binding protein, rbpA3, in the cyanobacterium Anabaena variabilis strain M3. *Plant and Cell Physiology*, *38*(1), 81-86.
68. Sato, N., & Wada, A. (1996). Disruption analysis of the gene for a cold-regulated RNA-Binding protein, rbpA1, in Anabaena: Cold-induced initiation of the heterocyst differentiation pathway. *Plant and cell physiology*, *37*(8), 1150-1160.
69. Sato, N., Maruyama, K., Nishiyama, Y., & Murata, N. (1997). Identification of a cold-regulated RNA-binding protein from the marine cyanobacterium Synechococcus sp. PCC7002. *Journal of Plant Research*, *110*(4), 405-410.
70. Sinetova, M. A., & Los, D. A. (2016). New insights in cyanobacterial cold stress responses: Genes, sensors, and molecular triggers. *Biochimica et Biophysica Acta General Subjects*, *1860*(11), 2391-2403.
71. Smith, C. A., Suzuki, J. Y., & Bauer, C. E. (1996). Cloning and characterization of the chlorophyll biosynthesis gene chlM from Synechocystis PCC 6803 by complementation of a bacteriochlorophyll biosynthesis mutant of Rhodobacter capsulatus. *Plant molecular biology*, *30*(6), 1307-1314.
72. Sugita, M., & Sugiura, M. (1994). The existence of eukaryotic ribonucleoprotein consensus sequence-type RNA-binding proteins in a prokaryote, Synechococcus 6301. *Nucleic acid research*, *22*(1), 25-31.
73. Swanson, M. S., Nakagawa, T. Y., LeVan, K., & Dreyfuss, G. I. D. E. O. N. (1987). Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Molecular and cellular biology*, *7*(5), 1731-1739.
74. Thomas, J., & Apte, S. K. (1984). Sodium requirement and metabolism in nitrogen-fixing cyanobacteria. *Journal of Biosciences*, *6*(5), 771-794.
75. Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D.,

& Venter, J. C. (1997). The complete genome sequence of the gastric pathogen Helicobacter pylori. *Nature*, *388*(6642), 539-547.

76. Tribelli, P. M., & López, N. I. (2018). Reporting key features in cold-adapted bacteria. *Life*, *8*(1), 8.

77. Willimsky, G., Bang, H., Fischer, G., & Marahiel, M. A. (1992). Characterization of cspB, a Bacillus subtilis inducible cold shock gene affecting cell viability at low temperatures. *Journal of Bacteriology*, *174*(20), 6326-6335.

* * *