Published Online: 9 May 2019

Proc Indian Natn Sci Acad **85** No. 4 December 2019 pp. 791-802
© *Printed in India*.
DOI: 10.16943/ptinsa/2019/49592

Review Article

Voyaging Around ClpB/Hsp100 Proteins and Plant Heat Tolerance

RATNESH CHANDRA MISHRA* and ANIL GROVER
Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi 110 021, India

(Received on 04 September 2018; Revised on 18 January 2019; Accepted on 27 February 2019)

Temperature is one of the key physical parameters that fine tunes plant growth and development. However, above the optimal range, it can negatively affect the physiology of plants. Supraoptimal temperature brings incongruity in cellular proteostasis resulting in the build-up of insoluble toxic protein aggregates. To prevent protein misfolding and aggregation, cells deploy different strategies including synthesis of heat shock proteins (Hsp) belonging to different families, like small Hsps (sHsps), Hsp40, Hsp60, Hsp70. Once these aggregates are formed, their dissolution and recovery of the functional proteins occurs by the action of Caseinolytic Protease B (ClpB)/Hsp100, which are evolutionarily conserved in bacteria, fungi and plants. ClpB function during heat stress (HS) is important and appears indispensable, as mutant bacteria, yeast as well as plants lacking ClpB protein fail to survive HS. Genetic expression of ClpB proteins is modulated both by high temperature as well as developmental cues. Plant contains three isoforms of ClpB/Hsp100, one each localized to cytoplasm (ClpB-C), chloroplast (ClpB-P) and mitochondria (ClpB-M), against one in bacteria and two in yeast. Among these, ClpB-C protein in particular governs the thermotolerance response in plants. This review introduces plant ClpB proteins, summarizes the knowledge gained hitherto in ClpB biology, critically analyzes the recent findings and brings forth the areas requiring thrust in the upcoming research on ClpBs.

Keywords: Aggregates; ClpB/Hsp100; Plant; Protein; Renaturation; Thermotolerance

Introduction

Plants, which are not as privileged as animals to move and escape the undesirable environmental conditions, have evolved and developed enormous sensitivity towards monitoring and responding to their environment. Consequently, besides identifying and adjusting with subtle changes, plants have also developed sophisticated mechanisms to sustain environmental extremes. One of the key physical parameters that fine tunes the growth and development of plants is the ambient temperature. On one hand, diurnal temperature variation entrains the plants' biological clock, on the other hand, seasonal variations decide the timing of reproduction (Kumar and Wigge, 2010). While the subtle changes in temperature provide important cues for plant development, high temperature represents a significant stress for cellular proteostasis. Heat stress (HS) affects both the vegetative as well as the reproductive processes in plants, with the latter being more susceptible (Bita and Gerats, 2013). As reproductive processes involve flowering and fruit/seed setting, any adversity during these events are directly translated into the crop productivity losses. In the wake of the predicted increase in earth's ambient temperature in the future due to global warming and the growing food demand of the 21st century, there is a need to develop genetically modified crop plants with high thermotolerance.

Research in the past two decades has provided important insights into the plant heat shock response mechanisms. It has emerged that Caseinolytic Protease B/heat shock protein 100 (ClpB/Hsp100) is pivotal in modulating plants' thermotolerance trait (Mishra and Grover, 2016). Mutant *Arabidopsis*, maize and rice plants with defect in the ClpB/Hsp100 function/expression show extreme sensitivity to HS (Hong and Vierling, 2000; Nieto-Sotelo *et al.*, 2002; Lin *et al.*, 2014). Further, genetic expression of ClpB/Hsp100 genes from different plants have been shown

*Author for Correspondence: E-mail: ratnesh.id@gmail.com

to complement the heat sensitivity of the yeast Hsp100 mutant "*Schsp104* (Lee *et al.*, 1994; Schirmer *et al.*, 1994; Singh *et al.*, 2010). Furthermore, overexpression of ClpB/Hsp100 in rice and *Arabidopsis* resulted in the enhancement of overall thermotolerance (Queitsch *et al.*, 2000; Katiyar-Agarwal *et al.*, 2003). In the light of the above arguments, ClpB/Hsp100 represents an important candidate gene to be targeted for crop improvement amidst the uncertainties of the future climate. This review discusses an understanding in ClpB biology, with emphasis on its function, genetic regulation and strategies to genetically employ ClpB proteins for thermotolerance enhancement. The first few sections provide the readers with a general account of ClpB proteins, which is necessary to realize the applicability of ClpB biology in crop development discussed in later sections of the review.

Proteostasis, Heat Stress and Importance of ClpB/Hsp100

Nearly all the vital activities in a cell are performed by properly folded, functionally active proteins. Proteostasis, as it comprises biogenesis, correct folding, unfolding, trafficking and turnover of functional proteins, is thus of the utmost importance for the destined cell functioning. Any anomaly in this process, either spontaneous or under cellular stress, may deviate the entire proteostasis scheme and prove fatal. Every living cell is thus equipped with molecular chaperones in the form of heat shock proteins (Hsp) that besides assisting the proper protein folding also ensure correct refolding, immediately after a protein gets misfolded. Molecular chaperone such as GroEL/GroES (Hsp60/Hsp10) and DnaK/DnaJ/GrpE [Hsp70/Hsp40/nucleotide exchange factors (NEF)] are involved in identifying and refolding the misfolded proteins back into their active conformations (Mishra and Grover, 2016). However, when there is an unprecedented increase in the levels of misfolded proteins, like during HS, it goes beyond the capacity of aforementioned chaperones to bring them all back to their native states. This leads to an overall increase in the levels of misfolded proteins and as these species carry exposed hydrophobic residues, they tend to club together forming toxic protein aggregates. To clear up these toxic bodies, a cell can either target them to proteolytic degradation or somehow recover the functional proteins from these intricate proteinaceous masses. The targeting of aggregated proteins is accomplished with the help of ClpA, C, D, M, N, X and Y proteins. To recover functional proteins from the aggregates, plants cell are endowed with ClpB/Hsp100 proteins (Mishra and Grover, 2016). During HS, where most of the vital functional proteins are misfolded, their recovery gets decisive over proteolytic degradation for proper cell functioning and survival. Thus, the presence/function of ClpB/Hsp100 is indispensable and critical for plants under HS. A simplified overview of the proteostasis steps that involves chaperone proteins, including ClpB/Hsp100 during HS, is depicted in Fig. 1. The model highlights the point of intervention and importance of ClpB/Hsp100 under HS conditions.

ClpB Proteins: The Underpinnings

Historically, Clp ATPase protein was first discovered as a protease system from bacteria that is capable of hydrolyzing casein *in vitro* (Hwang *et al.*, 1987; Katayama-Fujimura *et al.*, 1987). It was later found to be a two-component complex comprising a ClpA regulatory ATPase unit and a proteolytic component ClpP (Schelin *et al.*, 2002). The Clp family is comprised of several members, like ClpA, B, C, D, M, N, X and Y. These members are further sub categorized into class I and class II sub-families depending upon their structural features. The Clp family itself belongs to the large superfamily of AAA^+ (ATPase associated with diverse cellular activities) proteins. The detailed overview of classification, distribution and molecular structure of the Clp proteins is described elsewhere (Mishra and Grover, 2016). Briefly, Clp proteins are widely distributed, with its one or the other member present across the spectrum of life. ClpB protein is specifically present in bacteria, fungi and plants. However, ClpB proteins are mostly lacking in animals. Possibly, the ability to move away from the HS may have caused the loss of ClpB protein in animals during evolution. ClpB proteins were discovered after the discovery of ClpA in *E. coli* (Gottesman *et al.*, 1990). It was initially believed that ClpB forms are the homolog of ClpA proteins with similar functions. However, it was later discovered that ClpB cannot substitute ClpA in ClpP mediated casein degradation (Zolkiewski, 2006). Subsequent work revealed that while other Clp members facilitate the clearance of protein aggregates through proteolytic degradation, ClpB performs in a non-degradative manner and

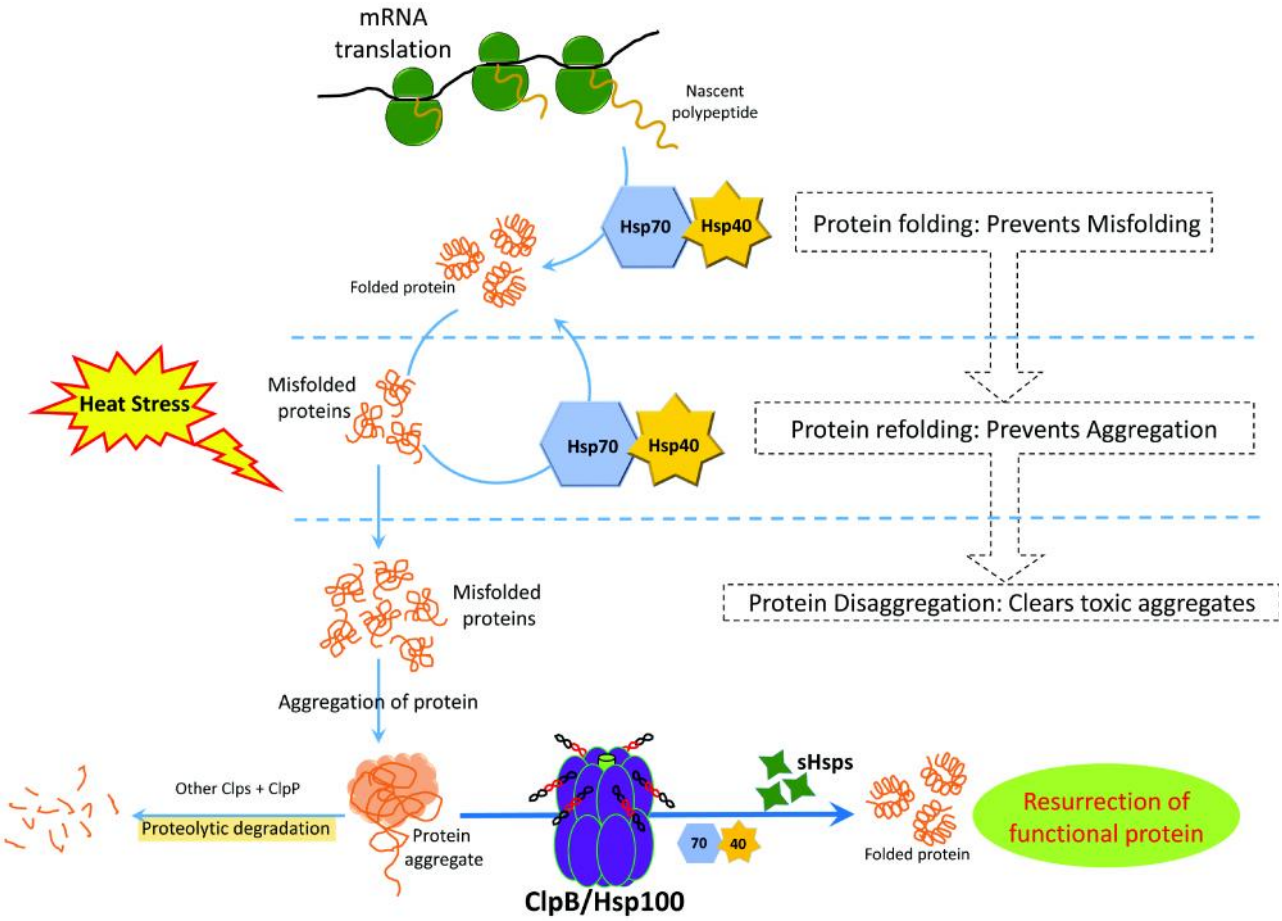

Fig. 1: Proteostasis and the importance of ClpB/Hsp100. The current model represents a simplified scheme of cellular proteostasis highlighting the indispensable role of ClpB during heat stress. Proteostasis maintains a balance between the amount of proteins being formed and targeted for degradation in a cell. Hence, it ensures an optimal level of functional proteins at a given time. Harmony in this process is of utmost importance for the correct physiological state of the cell. Being so critical, it is guided by different families of chaperone proteins/Hsps at different steps. As soon as an mRNA is translated and a nascent polypeptide comes into existence, Hsp70/Hsp40 (DnaK/DnaJ) escorts the folding of the polypeptide until a correct functional three-dimensional structure is attained. This checks misfolding of the proteins. Still, if something goes wrong and a protein is accidentally misfolded, Hsp70/Hsp40 recognizes such species and fold them back to their native functional form. This prevents their aggregation and thus formation of toxic protein aggregates. HS exaggerates protein-misfolding and leads to formation of toxic protein aggregates. Protein aggregates, in general, is cleared either by proteolytic degradation or resolublization of entangled proteins. However, under HS conditions, where most of the functional proteins form aggregates, ClpB functions by disentangling individual proteins from these aggregates. In concert with Hsp70/Hsp40 system, ClpB liberates nascent polypeptides that are believed to undergo protein folding under guidance of sHsps. Through the resurrection activity, ClpB helps cell to survive HS conditions

renatures back the native functional proteins from aggregates. Mechanistically, while other Clp proteins associate with the proteolytic ClpP component and function as the regulatory unit of the two-component proteolytic system, ClpB lacks the motif (IGF/L) for ClpP association and thus solely functions as a renaturase (Mishra and Grover, 2016). Following its discovery in bacteria, ClpB was identified in yeast (Sanchez and Lindquist, 1990) where it is commonly called as Hsp104. In addition to Hsp104, yeast cell also contains a mitochondrial isoform of ClpB, which is known as Hsp78 (Krzewska *et al.*, 2001). It was only in early 1990s when the reports stating the presence of Hsps ranging a molecular weight of 100-110 kDa in plants came (Singla and Grover, 1993; Schirmer *et al.*, 1994). In 1993, the identification of a

heat-inducible 100 kDa rice protein, sharing the immunological properties with yeast Hsp104 (*Schsp104*), marked the beginning of plant ClpB research (Singla and Grover, 1993).

ClpB stands out from other members of Clp family not just on how it functions, rather it is also unique in its characteristic HS induced expression. It is because of its HS induced expression and the molecular weight, which is in the range of 100 kDa, plant ClpBs are referred to as Hsp100 proteins (Singh and Grover, 2010; Kim *et al.*, 2012; Mishra and Grover, 2016). The subsequent account of this review mainly focuses on plant ClpB/Hsp100 proteins, with occasional references of bacterial and yeast ClpB forms wherever necessary.

ClpB Distribution and Homologs

As mentioned earlier, ClpB proteins are present in bacteria, fungi and plants. The prokaryotes and yeast contain one and two (cytoplasmic *Schsp104* and mitochondrial *Schsp78*) ClpB homologs, respectively. Plants are strikingly endowed with cytoplasmic (ClpB-C), mitochondrial (ClpB-M) and plastidial (ClpB-P) forms of ClpB proteins. Mishra and Grover (2016) presented a detailed phylogenetic analysis of different ClpB forms from a plethora of species, including archaebacteria, bacteria, cyanobacteria, slime molds, budding fungi yeast and lower to higher flowering plants. Presence of three ClpB isoforms in plants is a general feature; more than 30 species of flowering plants (including both dicots and monocots) were analyzed and noted to contain separate cytoplasmic and organeller ClpB forms. Their analysis also revealed that plant organeller ClpBs (ClpB-M and ClpB-P) are more closely related, whereas the ClpB-C has evolved independently. Further, organeller ClpBs are more closely related to the bacterial and cynobacterial ClpB forms, whereas ClpB-C lineage comprised ClpB proteins from eukaryotic members like *Dictyostelium* (protists), algae, bryophytes and pteridophytes. Also, *Schsp104* was noted to be remarkably diverse from bacterial as well as plant ClpB forms. The relatedness of organeller ClpBs with bacterial and cyanobacterial ClpBs corroborates the theory of endosymbiotic origin of organelles in eukaryotic cells. According to this theory, chloroplast and mitochondria represent formerly free-living bacteria that were taken inside by another cell as endosymbionts. Interestingly, the closeness of organellar ClpB lineage is more with the cyanobacterial lineage as compared to any other bacterial species. This observation, on one hand, reflects that the plastidial homologue possibly originated from a cyanobacterial endosymbiont, and on the other hand indicates a possible origin of mitochondrial homologue from chloroplastic form rather than from a separate mitochondrial endosymbiont. Apparently, the unicellular green algae appears to be the most appropriate candidate to host the later event. *In silico* analysis entailing green alga *Chlamydomonas* strongly supports this idea. Search for the *Chlamydomonas* proteins on the Phytozome database (Goodstein *et al.*, 2012) using the keyword "ClpB" fetched three sequences – two with well-defined signature domains of ClpB/Hsp100 protein (with 925 aa and 1012 aa) and a third 1411 amino acid long protein. The 925 and 1012 aa sequences shared high homology with *Arabidopsis* and rice ClpB-C and ClpB-P forms, respectively. Protein with 1411 aa showed high similarity with *Arabidopsis* ClpB-M. Additionally, 3 more ClpB related sequences of 1065, 1085 and 1053 aa were revealed. The proteins with 1065 and 1085 aa were noted to be homologous to rice ClpB-M (Schroda, 2004). The foregoing account advocates that the evolution of organellar ClpB forms took place in the green alga host. As obvious from the trend in their evolutions, the expression and function of organeller ClpBs diverged from that of ClpB-C form (discussed later in this review).

Expression Attributes of ClpB Proteins

One of the characteristic and noteworthy features of ClpB proteins is its conserved genetic regulation. ClpB protein is characteristically heat-inducible throughout the species where it has been studied. Moreover, it also shares its characteristic pattern of developmental regulation. Broadly, in vegetative structures, the expression of ClpB is strictly stress inducible (primarily HS), however, transition from vegetative to reproductive stages induces constitutive expression of ClpB proteins.

Developmental Regulation

In yeast, expression of Hsp104 is beyond detection in exponentially growing cells, but upon transition to stationary growth stage, Hsp104 expresses

constitutively (Sanchez and Lindquist, 1990). Similarly, plant ClpB-C genes show constitutive expression in reproductive structures. *Arabidopsis* express ClpB-C protein constitutively in pollen (Mishra and Grover, 2014). ClpB-C was also detected in tassel, pre-anthesis anther, silk tissues and pollen of maize plants (Young *et al.*, 2001). Rice ClpB-C promoter was found to express in anthers, ovary and style together with embryonic half of the seeds (Singh *et al.*, 2012). In fact, high constitutive expression in seeds is a hallmark of ClpB-C's genetic regulation in plants: many plant species, like *Brassica juncea*, *Zea mays*, *Triticum aestivum* and *Oryza sativa* expressed high levels of ClpB-C protein in the seeds (Singla *et al.*, 1998). More specifically, it is the mature embryos in the seeds that accumulated high levels of ClpB-C protein (Mishra and Grover, 2016). Unlike in rice, the expression of ClpB-C in *Arabidopsis* seeds increases further during the first 24 h of germination (Mishra and Grover, 2014; Mishra *et al.*, 2016), suggesting a possible involvement of the ClpB-C function in the germination of *Arabidopsis* seeds. The function of heat induced ClpB-C as a protein chaperone is obvious, however, the relevance of plant development-related ClpB-C accumulation remains elusive.

Stress-regulation

The expression of ClpB/Hsp100 in vegetative tissues is characteristically heat-inducible. In bacterial cells, ClpB is induced after an exposure of high temperature (Squires *et al.*, 1991). Similarly, expression of *Schsp104* in yeast is also HS inducible (Sanchez and Lindquist, 1990). Notably, the expression of ClpB-C protein is characteristically heat-inducible in the vegetative tissues of diverse plants studied so far, like wheat, rice, maize, tomato, soybean and *Arabidopsis* (Mishra and Grover, 2016). This evidently highlights the requirement and importance of this protein in plants during HS. Besides HS, ClpB-C also shows slight induction after few other stresses that are believed to compromise the integrity of functional proteins. For instance, expression of ClpB-C was also induced by heavy metal stress (Mishra and Grover, 2014), like in yeast, where arsenite stress reportedly induced *Schsp104* expression (Sanchez *et al.*, 1992). Further, expression of *Schsp104* was also induced by ethanol – a dehydrating agent (Sanchez *et al.*, 1992). Corroboratively, expression of ClpB-C in wheat was found to be induced by dehydration stress as well as ABA (Campbell *et al.*, 2001). Further, oxidative stress also up-regulates ClpB-C expression (Singh *et al.*, 2010; Singh *et al.*, 2012). Another study found induction of ClpB-C in response to oilseed rape mosaic virus infection in *Arabidopsis* (Carr *et al.*, 2006). It is important to be highlighted here that despite its expression under diverse cues, most of the studies have addressed HS-induced expression of ClpB-C in plants. Of course, instances of expression other than during HS are occasional and limited, but we believe that referring to such reported case studies is necessary to understand complete ClpB-C biology.

Modulation of ClpB Expression

Studies suggest that heat-inducible expression of ClpB/Hsp100 is governed by interaction between the heat responsive cis-elements in the promoter of ClpB genes and specific transacting factors. In case of *E. coli* the heat responsive ClpB promoter is modulated by a heat shock sigma factor (Kitagawa *et al.*, 1991). In yeast, the heat induced expression of *Schsp104* is regulated by the interaction between the only heat shock transcription factor (Hsf) – Hsf1 – and the heat shock element (HSE) present in the promoter of *Schsp104* (Singh and Grover, 2010). *In silico* studies combined with genetic analyses suggest the presence of consensus HSEs in the plants' Hsps promoters that enable them to respond HS by inducing gene expression. The importance of HSE for heat dependent transcriptional regulation in plants has been validated by (1) promoter deletions and (2) the capacity of synthetic HSE sequence integrated in a truncated CaMV35S promoter to stimulate heat inducible reporter gene expression in transgenic tobacco (Schoffl *et al.*, 1989). Maize and *Arabidopsis* ClpB-C promoters contain five and six HSEs, respectively (Nieto-Sotelo *et al.*, 1999; Hong and Vierling, 2001). However, rice ClpB-C promoter harbors only one functional HSE (Singh *et al.*, 2012). Our group unveiled the importance of this single HSE in rice ClpB-C promoter through site-directed mutagenesis, which altered the expression attributes of the promoter (Singh *et al.*, 2012). Further, we revealed that among a plethora of Hsfs in rice, the ones that specifically regulate the heat-induced expression of ClpB-C are *OsHsfA2c* and *OsHsfA6a* (Singh *et al.*, 2012; Lavania *et al.*, 2018; Singh *et al.*, 2018). Although the research on *Arabidopsis* and maize ClpB-C is much advanced, the identity of the

Hsf regulating the heat-induction of ClpB-C in these species needs further investigation.

Studies entailing the mechanism of genetic regulation of rice ClpB-C gene have thrown several important observations. We noted that heat-induced expression of rice ClpB-C is not just transcriptionally regulated rather it involves post-transcriptional regulation as well. To envisage the necessity of such a regulatory mechanism, it is important to first note that during HS, where most of the cellular translational machinery is inactivated, ClpB-C is still able to maintains high level of protein expression. In principle, the phosphorylation and thus function of mRNA 7-methyl guanosine cap binding factor – eIF4E, which initiates translation – is compromised during HS (Echevarria-Zomeno *et al.*, 2013). This indicated the mediation of a cap-independent mode of ClpB-C translation during HS. Viruses have long been known to employ **i**nternal **r**ibosome **e**ntry **s**ites (IRES) present in the 5'UTRs of their mRNAs for cap-independent translation; through which, they overcome the translational inhibition posed by the stressed host cell. We revealed that the 5'UTR of ClpB-C mRNA contains possibly an IRES, which enables docking of ribosomes in cap-independent manner and thus advantageously allows uninterrupted translation during HS (Mishra *et al.*, 2016). We also showed that presence of the 5'UTR is an added advantage for the transcript to form protein under persistent HS; absence of 5'UTR doesn't affect the translation of the protein under optimal temperature (Mishra *et al.*, 2016). Involvement of 5'UTR-mediated cap-independent translation of ClpB-C has also been evidenced in maize (Dinkova *et al.*, 2005). In Arabidopsis, however, the mechanism through which ClpB-C maintains high expression levels during HS still needs investigation. Our own analysis entailing the deletion of few promising bases of *Arabidopsis* ClpB-C 5'UTR, close to translation start codon, which may function as an IRES, annulled the possibility of IRES mediation (Mishra and Grover, 2014). *Arabidopsis* ClpB-C 5'UTR harbors an intron, which we believe may have a role in the heat regulation of this gene. This hypothesis needs to be tested in the upcoming ClpB-C research.

While the genetic regulation of ClpB-C under HS has been an area of major focus since the beginning, the question of how the developmental expression of ClpB-C is regulated has gained less attention. There exist a major gap in our knowledge regarding the cis-elements and trans-acting factors involved in ClpB-C's expression in reproductive tissues. In *Arabidopsis*, HsfA9 has been implicated in induction of ClpB-C in seeds (Kotak *et al.*, 2007). As an extension to our study on rice ClpB-C 5'UTR, we reported that the deletion of 5'UTR doesn't affect the constitutive expression in seeds. Therefore, the presence of 5'UTR in ClpB-C transcripts expressed in seeds appears to be accessory just to ensure effective germination under HS conditions (Kotak *et al.*, 2007).

Unlike ClpB-C, which is strictly stress inducible in vegetative tissues, organellar ClpBs also have constitutive basal expressions (Mishra and Grover, 2016). This hints towards a possible housekeeping role of organellar ClpBs in plants.

Functional Attributes of ClpB/Hsp100 Protein

As discussed above, the bonafide function that ClpB/Hsp100 performs is renaturation of protein from aggregates. The disaggregation activity of ClpB protein was first reported in yeast. It was microscopically observed that HS leads to formation of dense stress granules (protein aggregates) that clears up during recovery phase in WT, but not in *Schsp104* mutant (Parsell *et al.*, 1994). Expression of rice ClpB-C in the *Schsp104* also resulted in clearing up of the stress granules (Agarwal *et al.*, 2003), suggesting a similar function of plant ClpB-C proteins. Although a more direct evidence of ClpB-C chaperone activity *in planta* is still lacking, the foregoing indicates a protein chaperone role of ClpB-C. An extensive overview of modular ClpB organization and the contribution of each domain in the overall chaperone activity has been provided by Mishra and Grover (2016). Precisely, ClpB monomer oligomerizes into a two-tiered hexameric ring forming an axial channel, guided by a central pore. After recognition and binding, the aggregated substrate passes through this channel in the form of unfolded polypeptide threads. Several reports suggest functional association of ClpB disaggregase with Hsp70 system for effective dissolution of protein aggregates (Mogk *et al.*, 1999). With Hsp70, ClpB forms a bi-chaperone system, which liberates unfolded, chaperone recognizable intermediate peptides. These non-native

peptides then undergo the process of protein folding aided by sHsps (Watanabe *et al.*, 2000). Hence, the aggregated proteins undergo unfolding and are finally folded back to their native tertiary/quaternary functional state. Genetic analysis in Arabidopsis revealed cooperation of sHsps with ClpB-C chaperone system (Lee *et al.*, 2005). The pressing question now is whether the recognition of substrates by ClpB-C machinery is generalized or more specific. Further, whether ClpB recognizes its substrates on its own or requires assistance of an additional accessory/adapter protein(s), like in case of other Clp proteins (Nishimura *et al.*, 2013). As discussed before, ClpB mutants show a very severe phenotype upon HS. This indirectly indicates that ClpB machinery can solubilize an array of vital substrate proteins. However, till date there is only limited research in this area. A decade ago, interaction of ClpB with anaerobic reductase and TrfA (replication initiation protein) was reported (Pontis *et al.*, 1991; Konieczny and Liberek, 2002). Very recently, employing protein-protein-interaction-based pull-down strategy, 60 potential protein substrates were identified for ClpB chaperone in *Leptospira interrogans* (Krajewska *et al.*, 2018). Most of these proteins were found to be related to major metabolic pathways, like glycolysis, TCA cycle, gluconeogenesis, fatty acid and amino acid metabolism. Proteins related to cell signaling, membrane biogenesis, protein synthesis and transcriptional process were also identified in the pool (Krajewska *et al.*, 2018). Interestingly, this study brings forth many important aspects regarding ClpB's substrates; firstly, as expected, the inventory of ClpB's substrates contains a large number of proteins, secondly, protein belonging to key metabolic pathways, crucial for proper cell functioning, constitute ClpB substrates, and lastly, ClpB substrates also comprise protein involved in signal transduction pathways, like signaling proteins and proteins involved in translation and transcription. Physical interaction of ClpB with proteins related to processes like transcription and translation hints towards a possible regulatory role of ClpB in these vital cellular processes. It is the high time to conduct high-throughput studies in plants to identify the potential substrates for plant ClpB-C machinery.

Besides the classical function of protein renaturase, ClpB also performs few other important activities. For instance, it has been noted to act as a translational activator. In one study, ClpB-C was found to interact with TMV 5' leader (Ω), forming a two-component translation modulatory complex that enhances translation up to 50-fold (Wells *et al.*, 1998). In another study, ClpB-C was noted to bind with internal light regulatory element of ferredoxin mRNA and enhance the translation (Ling *et al.*, 2000). The idea of ClpB-C's involvement in translation modulation was unfortunately limited to these scattered and scarce evidences for a long time. Interestingly, however, studies in maize show that the translation regulatory activity of ClpB-C is interlinked with its developmental regulation. Mass spectrometric studies in maize seeds revealed that ClpB-C protein constitutes one of the components of the cap-binding complex (Lazaro-Mixteco *et al.*, 2012). Through its interaction with the translational machinery, ClpB-C may affect translation of the proteins during germination. ClpB-C's presence in cap-binding complex was also hypothesized to function as a switch by which ClpB-C may activate translation of specific mRNAs during early germination (Lazaro-Mixteco *et al.*, 2012). This work thus added another dimension in the functional aspect of ClpB-C protein. From the foregoing account, two important attributes of ClpB-C can be made; first, ClpB-C physically interact with mRNAs, and second, ClpB-C expresses highly and constitutively in maturing embryos. In the light of these attributes combined with the higher number of different mRNA transcript that maturing embryos generally accumulates, we propose that ClpB-C may also perform RNA chaperone activity as well. Through later activity, ClpB-C may possibly sequester and protects the important mRNAs in mature seeds till the time of germination. Upon advent of optimal conditions, ClpB-C may then liberate the sequestered mRNAs as well as facilitate the translation of specific ones, which are crucial for the process of germination. Researchers are invited to test this hypothesis, which, of course, requires investigation and establishment of ClpB-C's RNA chaperone activity at the first place.

There is now some evidence to suggest that ClpB-C can also regulate transcriptional machinery. In this context, nuclear localization of ClpB protein is reported. Nuclear localization of *Schsp104*, rice and maize ClpB-C has been shown previously (Nieto-Sotelo *et al.*, 2002; Tkach and Glover, 2008; Singh *et al.*, 2012). Further, our group has reported interaction of ClpB-C with different Hsfs in rice, which we believe have implication in its feedback regulation

(Singh *et al.*, 2012). Nuclear shuttling of ClpB-C and its interaction with Hsfs undoubtedly indicates its involvement in transcriptional modulation. Most importantly, in yeast, it is the HS that specifically triggers the shuttling of Hsp104 from cytoplasm to nucleus (Tkach and Glover, 2008), suggesting that the transcriptional modulation by ClpB is possibly HS-induced. In plants, however, a plausible regulation of transcription by ClpB-C appears to be more general than HS-induced: Maize embryo, where ClpB-C is developmentally/constitutively expressed, contains most of its expressed ClpB-C in the nucleus (Nieto-Sotelo *et al.*, 2002). Also, a study depicting HS-induced shuttling of ClpB-C from cytoplasm to nucleus in plants is still lacking. Another important question that needs to be highlighted here is whether the transcriptional regulation by ClpB-C is just limited to its own feedback regulation? This demands more, high-throughput protein-protein interaction studies in future to investigate whether ClpB-C interacts with other transcription factors proteins also. Finally, the above discussion reveals three different facets of ClpB-C function: (1) protein renaturation, (2) translational regulation and (3) transcriptional regulation.

While ClpB-C is indispensable for survival under HS, organellar ClpBs appear to be less significant. Hitherto, there is only a single study where aberration in ClpB-P compromised acquired thermotolerance (Yang *et al.*, 2006). Mutation of ClpB-P and ClpB-M in *Arabidopsis* have no remarkable effect on the thermotolerance ability of the mutant plants (Lee *et al.*, 2006). ClpB-P mutant, however, display several developmental phenotypes; mutant seed germinates, develop few pairs of chlorotic leaves and eventually dies (thus seedling lethal). Most importantly, chloroplast development in ClpB-P mutant is anomalous (Lee *et al.*, 2006; Myouga *et al.*, 2006; Zybailov *et al.*, 2009). This goes hand in hand with the constitutive expression of ClpB-P in vegetative tissues and suggests that ClpB-P essentially is a housekeeping gene. Housekeeping role of ClpB-P is also ascertained by the study in *Arabidopsis*, where an attempt to overexpress ClpB-P for enhancing thermotolerance resulted in inhibition of chloroplast development (Myouga *et al.*, 2006). In accordance with the classical protein chaperone activity of ClpBs, it is highly plausible that organellar ClpBs are also involved in protein transport system across the organellar membranes. Nevertheless, a definitive evidence is still lacking and needs investigation in future. Overall, the knowledge gained so far in plant ClpB biology is more concentrated on ClpB-C with scarce information on organellar ClpBs. There is indeed a need to put more thrust on organellar ClpB research in future.

From the above discussion it is amply clear that plant ClpB-C stands out as one of the most critical genes that governs the high temperature tolerance in plants.

Global Warming and The Need to Develop Heat Tolerant Transgenics

Global warming is no more a conceived notion. Its alarming effect on earth's climatic patterns and the high and low weather extremes is now more obvious than ever before. One of the consequences of global warming is the increasing earth's ambient temperature and so is the frequency of plant HS episodes. Heat stress has already been noted to reduce the yield of the major food crops worldwide (Lavania *et al.*, 2015). Estimates suggest that for every 1^{o}C increase in night time growing temperature there is a 10% decrease in the rice productivity (Peng *et al.*, 2004). Alarmingly, the climate change models anticipate a further increase in global mean surface temperature by 1.5-4^{o}C towards the end of the 21st century (IPCC, 2013). This poses a serious threat to crop productivity and the attempts driven towards future food security. Alongside, this necessitates the genetic enhancement of crop plants to make them more tolerant to HS. Research in the past few decades have come up with several potential candidate genes that can be employed in development of crop plants against HS (Lavania *et al.*, 2015). ClpB-C, however, has emerged as one of the most important genes owing to its 'protein resurrection' property. The ability of ClpB-C to disentangle/revive functional protein from toxic aggregates makes it unique among all the molecular chaperones present in a cell. Indeed, there is no redundancy in ClpB-C's function; plant ClpB-C mutants are severely defective in their ability to survive HS, and it is only the complementation with the native functional gene that reverses this phenotype. Furthermore, attempts to overexpress ClpB-C gene genetically in a plant showed promising results in terms of thermotolerance enhancement (Queitsch *et al.*, 2000; Katiyar-Agarwal *et al.*, 2003). Molecular analysis revealed an enhanced level of ClpB-C protein

in these transgenic plants. This highlights that under the extreme HS conditions the performance of a plant is highly dependent on the levels of the functional ClpB-C it can express: higher the levels, better will be the thermotolerance. Besides, the thermotolerance trait also depends on how long the functional ClpB-C is maintained within the cell, as it ensures recovery of all the functional protein from aggregates over a period of stress. Evidently, rice Hsa32 (HS associated 32 kDa protein) mutant that fails to maintain the expressed ClpB-C levels for longer, like in wild type plants, shows compromised long-term acquired thermotolerance (Lin *et al.*, 2014). Importantly, Hsa32 mutant also show reduced ClpB-C levels in seeds, which negatively affected the seed germination during HS (Lin *et al.*, 2014). This highlights that alteration in seed ClpB-C levels also affect the basal thermotolerance of the seeds. Indeed, the complete absence of ClpB-C in *hot1-3 Arabidopsis* mutant resulted in loss of germination during HS and seed death (Hong and Vierling, 2001). Given this negative effect of reduced seed ClpB-C levels on seed basal thermotolerance, an increased level should result in enhanced seed basal thermotolerance. The key point that emerges from the above discussion is that overexpression of ClpB-C is useful for enhancing the thermotolerance. Now the important question is which expression system should be adopted – constitutive or inducible? In most cases, a beneficial gene is usually overexpressed constitutively to achieve the desired trait. Overexpression of ClpB-C, however, appears a bit tricky. Previous attempt to overexpress ClpB-C in *Arabidopsis* resulted in co-suppression (Queitsch *et al.*, 2000). Out of a number of transgenic plants generated only a few expressed higher ClpB-C levels, whereas the rest behaved like mutant plants (Queitsch *et al.*, 2000). Considering from nature's point of view, this is very much conceivable. ClpB-C is an energy-driven molecular pump that requires tremendous amount of ATP as energy source for its function. Its constitutive presence/function under the condition where it is not required may prove energetically costly. Further, as it can divert proteolytic degradation to protein renaturation, its constitutive presence in vegetative tissue may prove devastating to the selective protein turn over, which is critical to maintain harmony in various vital metabolic pathways. Also, constitutive expression may not work well in seeds, as seed specific expression is stronger than general constitutive expression through *CaMV35S* promoter. Considering the above argument, inducible expression appears to be better over constitutive expression. Choosing a native ClpB-C promoter stretch to express the transgene will take care of both heat-inducible expression as well as developmental expression in seeds and other HS sensitive reproductive tissues. Inducible expression can be achieved either by transforming plants with an assembly of inducible promoter of a desired length upstream to the ClpB-C cDNA or simply transforming the plant using a genomic fragment containing the native promoter and gene along with the 5' and 3' UTRs, if present. As genetic regulation of ClpB-C also involves post-transcriptional regulation (discussed before), which involves UTRs and perhaps introns as well, it may be a better approach to use genomic fragment instead of assembling promoter and gene individually. This may ensure higher level of expression during HS and developmentally under optimal temperature. It can be argued that overexpression of ClpB-C in developmental tissues may have undesired consequences. The possibility of this, however, is very low, as there already exist a remarkable variation of developmental ClpB-C levels among different ecotypes of a species in nature; higher levels reflect more genetic fitness (Tonsor *et al.*, 2008). Finally, if desired, inducible overexpression of ClpB-C can be combined with overexpression of Hsa32 gene/homolog through multiple gene stacking approach. Through this the enhanced ClpB-C levels can be maintained even for longer durations in the transgenic plants and can help in coping prolonged HS.

Conclusion

Genetic studies have already proven the importance of ClpB-C in plant thermotolerance trait. Fundamental research over the past three decades has produced ample information regarding ClpB-C function and genetic regulation in plants. It is the high time to start utilizing the knowledge gained so far and develop strategies for effective genetic utilization of ClpB-C machinery for enhancing crop thermotolerance. This will not just accelerate/substantiate the ClpB-C research further, but eventually also help in meeting the copious future food demand amidst the uncertainties of global agricultural climate.

Acknowledgements

R M thanks CSIR for the award of fellowship grant. AG thanks Department of Science and Technology (DST), Government of India (GOI) (SERB/SB/SO/PS/37/2013) and CARI, Department of Biotechnology; (GOI) for the financial support.

References

Agarwal M, Sahi C, Katiyar-Agarwal S, Agarwal S, Young T, Gallie D R, Sharma V M, Ganesan K and Grover A (2003) Molecular characterization of rice hsp101: complementation of yeast hsp104 mutation by disaggregation of protein granules and differential expression in indica and japonica rice types *Plant Molecular Biology* **51** 543-553

Bita C. E. and Gerats T (2013) Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops *Frontiers in Plant Science* **4** 273

Campbell J L, Klueva N Y, Zheng H G, Nieto-Sotelo J, Ho T D and Nguyen H T (2001) Cloning of new members of heat shock protein HSP101 gene family in wheat (*Triticum aestivum* (L.) Moench) inducible by heat, dehydration, and ABA(1). *Biochimica et Biophysica Acta* **1517** 270-277

Carr T, Wang Y, Huang Z, Yeakley J M, Fan J B and Whitham S A (2006) Tobamovirus infection is independent of HSP101 mRNA induction and protein expression *Virus Research* **121** 33-41

Dinkova T D, Zepeda H, Martinez-Salas E, Martinez L M, Nieto-Sotelo J and de Jimenez E S (2005) Cap-independent translation of maize Hsp101 *The Plant Journal : For Cell and Molecular Biology* **41** 722-731

Echevarria-Zomeno S, Yanguez E, Fernandez-Bautista N, Castro-Sanz A B, Ferrando A and Castellano M M (2013) Regulation of Translation Initiation under Biotic and Abiotic Stresses *Int J Mol Sci* **14** 4670-4683

Goodstein D M, Shu S, Howson R, Neupane R, Hayes R D, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, and Rokhsar D S (2012) Phytozome: a comparative platform for green plant genomics *Nucleic Acids Res* **40** D1178-1186

Gottesman S, Clark W P and Maurizi M R (1990) The ATP-dependent Clp protease of Escherichia coli. Sequence of clpA and identification of a Clp-specific substrate *The Journal of Biological Chemistry* **265** 7886-7893

Hong S W and Vierling E (2000) Mutants of Arabidopsis thaliana defective in the acquisition of tolerance to high temperature stress *Proceedings of the National Academy of Sciences of the United States of America* **97** 4392-4397

Hong S W and Vierling E (2001) Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress *The Plant journal : for cell and molecular biology* **27** 25-35.

Hwang B J, Park W J, Chung C H and Goldberg A L (1987) Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La *Proceedings of the National Academy of Sciences of the United States of America* **84** 5550-5554

Katayama-Fujimura Y, Gottesman S and Maurizi M R (1987) A multiple-component, ATP-dependent protease from Escherichia coli *The Journal of Biological Chemistry* **262** 4477-4485

Katiyar-Agarwal S, Agarwal M and Grover A (2003) Heat-tolerant basmati rice engineered by over-expression of hsp101 *Plant Molecular Biology* **51** 677-686

Kim M, Lee U, Small I, des Francs-Small C C and Vierling E (2012) Mutations in an Arabidopsis mitochondrial transcription termination factor-related protein enhance thermotolerance in the absence of the major molecular chaperone HSP101 *The Plant Cell* **24** 3349-3365

Kitagawa M, Wada C, Yoshioka S and Yura T (1991) Expression of ClpB, an analog of the ATP-dependent protease regulatory subunit in Escherichia coli, is controlled by a heat shock sigma factor (sigma 32) *Journal of Bacteriology* **173** 4247-4253

Konieczny I and Liberek K (2002) Cooperative action of Escherichia coli ClpB protein and DnaK chaperone in the activation of a replication initiation protein *The Journal of Biological Chemistry* **277** 18483-18488

Kotak S, Vierling E, Baumlein H and von Koskull-Doring P (2007) A novel transcriptional cascade regulating expression of heat stress proteins during seed development of Arabidopsis *The Plant Cell* **19** 182-195

Krajewska J, Arent Z, Zolkiewski M and Kedzierska-Mieszkowska S (2018) Isolation and Identification of Putative Protein Substrates of the AAA+ Molecular Chaperone ClpB from the Pathogenic Spirochaete Leptospira interrogans *Int J Mol Sci* **19** 1234

Krzewska J, Langer T and Liberek K (2001) Mitochondrial Hsp78, a member of the Clp/Hsp100 family in Saccharomyces cerevisiae, cooperates with Hsp70 in protein refolding *FEBS Letters* **489** 92-96

Kumar S V and Wigge P A (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis *Cell* **140** 136-147

Lavania D, Dhingra A and Grover A (2018) Analysis of transactivation potential of rice (*Oryza sativa* L.) heat shock factors *Planta* **247** 1267-1276

Lavania D, Dhingra A, Siddiqui M H, Al-Whaibi M H and Grover A (2015) Current status of the production of high temperature tolerant transgenic crops for cultivation in warmer climates *Plant Physiol Biochem* **86** 100-108

Lazaro-Mixteco P E, Nieto-Sotelo J, Swatek K N, Houston N L, Mendoza-Hernandez G, Thelen J J and Dinkova T D (2012) The Absence of Heat Shock Protein HSP101 Affects the Proteome of Mature and Germinating Maize Embryos *Journal of Proteome Research* **11** 3246-3258

Lee U, Wie C, Escobar M, Williams B, Hong S W and Vierling E (2005) Genetic analysis reveals domain interactions of Arabidopsis Hsp100/ClpB and cooperation with the small heat shock protein chaperone system *The Plant Cell* **17** 559-571

Lee U, Rioflorido I, Hong S W, Larkindale J, Waters E R and Vierling E (2006) The Arabidopsis ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development *The Plant Journal : For Cell and Molecular Biology* **49** 115-127

Lee Y R, Nagao R T and Key J L (1994) A soybean 101-kD heat shock protein complements a yeast HSP104 deletion mutant in acquiring thermotolerance *The Plant Cell* **6** 1889-1897

Lin M Y, Chai K H, Ko S S, Kuang L Y, Lur H S and Charng Y Y (2014) A Positive Feedback Loop between HEAT SHOCK PROTEIN101 and HEAT STRESS-ASSOCIATED 32-KD PROTEIN Modulates Long-Term Acquired Thermotolerance Illustrating Diverse Heat Stress Responses in Rice Varieties *Plant Physiology* **164** 2045-2053

Ling J, Wells D R, Tanguay R L, Dickey L F, Thompson W F and Gallie D R (2000) Heat shock protein HSP101 binds to the Fed-1 internal light regulator y element and mediates its high translational activity. *The Plant Cell* **12** 1213-1227

Mishra R C and Grover A (2014) Intergenic sequence between Arabidopsis caseinolytic protease B-cytoplasmic/heat shock protein100 and choline kinase genes functions as a heat-inducible bidirectional promoter *Plant Physiology* **166** 1646-1658

Mishra R C and Grover A (2016) ClpB/Hsp100 proteins and heat stress tolerance in plants *Crit Rev Biotechnol* **36** 862-874

Mishra R C, Richa Singh A, Tiwari L D and Grover A (2016) Characterization of 5'UTR of rice ClpB-C/Hsp100 gene: evidence of its involvement in post-transcriptional regulation *Cell Stress & Chaperones* **21** 271-283

Mogk A, Tomoyasu T, Goloubinoff P, Rudiger S, Roder D, Langen H and Bukau B (1999) Identification of thermolabile Escherichia coli proteins: prevention and reversion of aggregation by DnaK and ClpB *The EMBO Journal* **18** 6934-6949

Myouga F, Motohashi R, Kuromori T, Nagata N and Shinozaki K (2006) An Arabidopsis chloroplast-targeted Hsp101 homologue, APG6, has an essential role in chloroplast development as well as heat-stress response *The Plant Journal : For Cell and Molecular Biology* **48** 249-260

Nieto-Sotelo J, Kannan K B, Martinez L M and Segal C (1999) Characterization of a maize heat-shock protein 101 gene, HSP101, encoding a ClpB/Hsp100 protein homologue *Gene* **230** 187-195

Nieto-Sotelo J, Martinez L M, Ponce G, Cassab G I, Alagon A, Meeley R B, Ribaut J M and Yang R (2002) Maize HSP101 plays important roles in both induced and basal thermotolerance and primary root growth *The Plant Cell* **14** 1621-1633

Nishimura K, Asakura Y, Friso G, Kim J, Oh S H, Rutschow H, Ponnala L and van Wijk K J (2013) ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in Arabidopsis *The Plant Cell* **25** 2276-2301

Parsell D A, Kowal A S, Singer M A and Lindquist S (1994) Protein disaggregation mediated by heat-shock protein Hsp104 *Nature* **372** 475-478

Peng S, Huang J, Sheehy J E, Laza R C, Visperas R M, Zhong X, Centeno G S, Khush G S and Cassman K G (2004) Rice yields decline with higher night temperature from global warming *Proceedings of the National Academy of Sciences of the United States of America* **101** 9971-9975

Pontis E, Sun X Y, Jornvall H, Krook M and Reichard P (1991) ClpB proteins copurify with the anaerobic Escherichia coli reductase *Biochemical and Biophysical Research Communications* **180** 1222-1226

Queitsch C, Hong S W, Vierling E and Lindquist S (2000) Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis *The Plant Cell* **12** 479-492

Sanchez Y and Lindquist S L (1990) HSP104 required for induced thermotolerance *Science* **248** 1112-1115

Sanchez Y, Taulien J, Borkovich K A and Lindquist S (1992) Hsp104 is required for tolerance to many forms of stress *The EMBO Journal* **11** 2357-2364

Schelin J, Lindmark F and Clarke A K (2002) The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium Synechococcus *Microbiology* **148** 2255-2265

Schirmer E C, Lindquist S and Vierling E (1994) An Arabidopsis heat shock protein complements a thermotolerance defect in yeast *The Plant Cell* **6** 1899-1909

Schoffl F, Rieping M, Baumann G, Bevan M and Angermuller S (1989) The function of plant heat shock promoter elements in the regulated expression of chimaeric genes in transgenic tobacco *Mol Gen Genet* **217** 246-253

Schroda M (2004) The Chlamydomonas genome reveals its secrets: chaperone genes and the potential roles of their gene products in the chloroplast *Photosynthesis Research* **82** 221-240

Singh A and Grover A (2010) Plant Hsp100/ClpB-like proteins: poorly-analyzed cousins of yeast ClpB machine *Plant Molecular Biology* **74** 395-404

Singh A, Singh U, Mittal D and Grover A (2010) Genome-wide analysis of rice ClpB/HSP100, ClpC and ClpD genes *BMC Genomics* **11** 95

Singh A, Mittal D, Lavania D, Agarwal M, Mishra R C and Grover A (2012) OsHsfA2c and OsHsfB4b are involved in the transcriptional regulation of cytoplasmic OsClpB (Hsp100) gene in rice (*Oryza sativa* L.) *Cell Stress and Chaperones* **17** 243-254

Singh G, Sarkar N K and Grover A (2018) Mapping of domains of heat stress transcription factor OsHsfA6a responsible for its transactivation activity *Plant Science : An International Journal of Experimental Plant Biology* **274** 80-90

Singla S L and Grover A (1993) Antibodies raised against yeast HSP 104 cross-react with a heat- and abscisic acid-regulated polypeptide in rice *Plant Molecular Biology* **22** 1177-1180

Singla S L, Pareek A, Kush A K and Grover A (1998) Distribution patterns of 104 kDa stress-associated protein in rice *Plant Molecular Biology* **37** 911-919

Squires C L, Pedersen S, Ross B M and Squires C (1991) ClpB is the Escherichia coli heat shock protein F84.1 *Journal of Bacteriology* **173** 4254-4262

Tkach J M and Glover J R (2008) Nucleocytoplasmic trafficking of the molecular chaperone Hsp104 in unstressed and heat-shocked cells *Traffic* **9** 39-56

Tonsor S J, Scott C, Boumaza I, Liss T R, Brodsky J L and Vierling E (2008) Heat shock protein 101 effects in A. thaliana: genetic variation, fitness and pleiotropy in controlled temperature conditions *Molecular Ecology* **17** 1614-1626

Watanabe Y H, Motohashi K, Taguchi H and Yoshida M (2000) Heat-inactivated proteins managed by DnaKJ-GrpE-ClpB chaperones are released as a chaperonin-recognizable non-native form *The Journal of Biological Chemistry* **275** 12388-12392

Wells D R, Tanguay R L, Le H and Gallie D R (1998) HSP101 functions as a specific translational regulatory protein whose activity is regulated by nutrient status *Genes Dev* **12** 3236-3251

Yang J Y, Sun Y, Sun A Q, Yi S Y, Qin J, Li M H and Liu J (2006) The involvement of chloroplast HSP100/ClpB in the acquired thermotolerance in tomato *Plant Molecular Biology* **62** 385-395

Young T E, Ling J, Geisler-Lee C J, Tanguay R L, Caldwell C and Gallie D R (2001) Developmental and thermal regulation of the maize heat shock protein, HSP101 *Plant Physiology* **127** 777-791

Zolkiewski M (2006) A camel passes through the eye of a needle: protein unfolding activity of Clp ATPases *Mol Microbiol* **61** 1094-1100

Zybailov B, Friso G, Kim J, Rudella A, Rodriguez V R, Asakura Y, Sun Q and van Wijk K J (2009) Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism *Molecular and Cellular Proteomics: MCP* **8** 1789-1810.