November 2015

THE EVOLUTION OF THERMOTOLERANCE A CHARACTERIZATION OF A DIRECTIONALLY EVOLVED CYANOBACTERIUM

nathen Emil Bopp
University of Massachusetts Amherst

THE EVOLUTION OF THEMOTOLERANCE: A CHARACTERIZATION OF A DIRECTIONALLY EVOLVED CYANOBACTERIUM

A Thesis Presented

by

NATHEN BOPP

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE

SEPTEMBER 2015

Biochemistry and Molecular Biology

THE EVOLUTION OF THEMOTOLERANCE: A CHARACTERIZATION OF A DIRECTIONALLY EVOLVED CYANOBACTERIUM

A Thesis Presented

by

NATHEN BOPP

Approved as to style and content by:

_____
Elizabeth Vierling, Chair

_____
Courtney Babbitt, Member

_____
Steven Sandler, Member

_____
Danny Schnell, Member

_____
Danny Schnell, Program Director
Department of Biochemistry

ACKNOWLEDGEMENTS

I would like to acknowledge and thank Dr. Elizabeth Vierling for the tremendous amount of opportunity she has provided me with. I would also like to thank my lab members for being some of the best people I have known.

Thank you to my committee members, Dr. Danny Schnell, Dr. Courtney Babbitt, and Dr. Steven Sandler for your suggestions and assistance with my work.

My family who has supported me throughout my seemingly endless amount of time at Umass.

Finally, I would like to acknowledge the National Institutes of Health (NIH) for providing the funding for this research.

ABSTRACT

THE EVOLUTION OF THEMOTOLERANCE: A CHARACTERIZATION OF A DIRECTIONALLY EVOLVED CYANOBACTERIUM

SEPTEMBER 2015

NATHEN BOPP, B.S., UNIVERSITY OF MASSACHUSETTS AMHERST

M.S. UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Elizabeth Vierling

Chaperone proteins are essential components in the maintenance and turnover of the proteome. Many chaperones play integral functions in the folding and unfolding of cellular substrates under many conditions, including heat stress. Most chaperones can be characterized into two categories; the typical ATP dependent chaperones and the ATP independent chaperones. One ATP independent chaperone class it the Small Heat Shock Proteins (sHSPs), which as molecular life vests and are thought to protect misfolding proteins from irreversible aggregation. SHSPs have a conserved basic domain structure, containing an N-terminal arm of variable length, a conserved α-crystalline domain, and a short C-terminal extension. sHSPs are found in virtually all forms of life and are essential in some organisms for acquired thermotolerance to elevated temperatures. One such organism, the cyanobacterium *Synechocystis sp.* PCC. 6803, is an excellent model for the study and understanding of these proteins and their functions in vivo. The genome of *Synechocystis* encodes only one sHSP, Hsp16.6, and it has be shown to be essential for acquired thermotolerance. Two mutant derivatives of Hsp16.6 with single amino acid substitutions in the N-terminal arm (L9P and E25K) have loss-of-function phenotypes similar to knock out strains, but each has very different biochemical properties. The mutant L9P has an inability to interact with putative substrates during heat stress in vivo, while the mutant E25K

appears unable to release substrates. Using a directed evolution approach, suppressors have been isolated that recover the lost thermotolerance of their respective parent strains, either L9P (16 suppressors) or E25K (10 suppressors). These suppressors show thermotolerance equivalent to wild type under acute heat stress, while displaying superior chronic heat stress tolerance. This suggests that the response to acute and chronic heat stress may have different components. Additionally, deleting the L9P or E25K Hsp16.6 allele in each of the suppressor strains or

replacing it with wild type Hsp16.6 revealed that 16 of the suppressors tested so far (9 L9P and 7 E25K) no longer require Hsp16.6 for acute heat tolerance. Illumina sequencing and comparative genomics have been used to identify alterations in the genomes of the suppressor strains in order to define genetic circuits involved in thermotolerance. Our genome resequencing will also lead to a better understanding of how directed evolution impacts the genome of *Synechocystis*.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

BACKGROUND

Introduction

Perturbation of cellular processes is a continuous problem that all organisms are forced to cope with. These perturbations come in many types ranging from biotic stress, such as viral infection, to abiotic stresses like increased temperature. During the evolution of life, these stressors acted as selective forces, leading to the development of highly conserved, highly effective stress responses. Key components of these responses include chaperones, proteins responsible for protecting other proteins, and proteases, proteins responsible for the degradation of other proteins. Understanding the mechanisms of stress responses, their evolution and their potential therapeutic uses is an important area of research.

Synechocystis

Cyanobacteria are arguably the most important organisms on the planet. Thought to have been the cause of the oxygenation of early Earth, they continue to be the significant producers of gaseous O_2 today (Schopf et al, 2012). Additionally, cyanobacteria are believed to be the precursors of the eukaryotic chloroplast. One cyanobacterium, *Synechocystis sp.* PCC 6803 (hereafter referred to as *Synechocystis*), has been adopted as a model for understanding photosynthesis, the metabolism of photosynthetic organisms, and stress responses. Among the many interesting attributes of this single-celled organism that make it a model system, four important traits differentiate it from other cyanobacteria. First, *Synechocystis* was the first photosynthetic organism to have its complete, 3.95 mega base genome sequenced (Nakamura, Y.

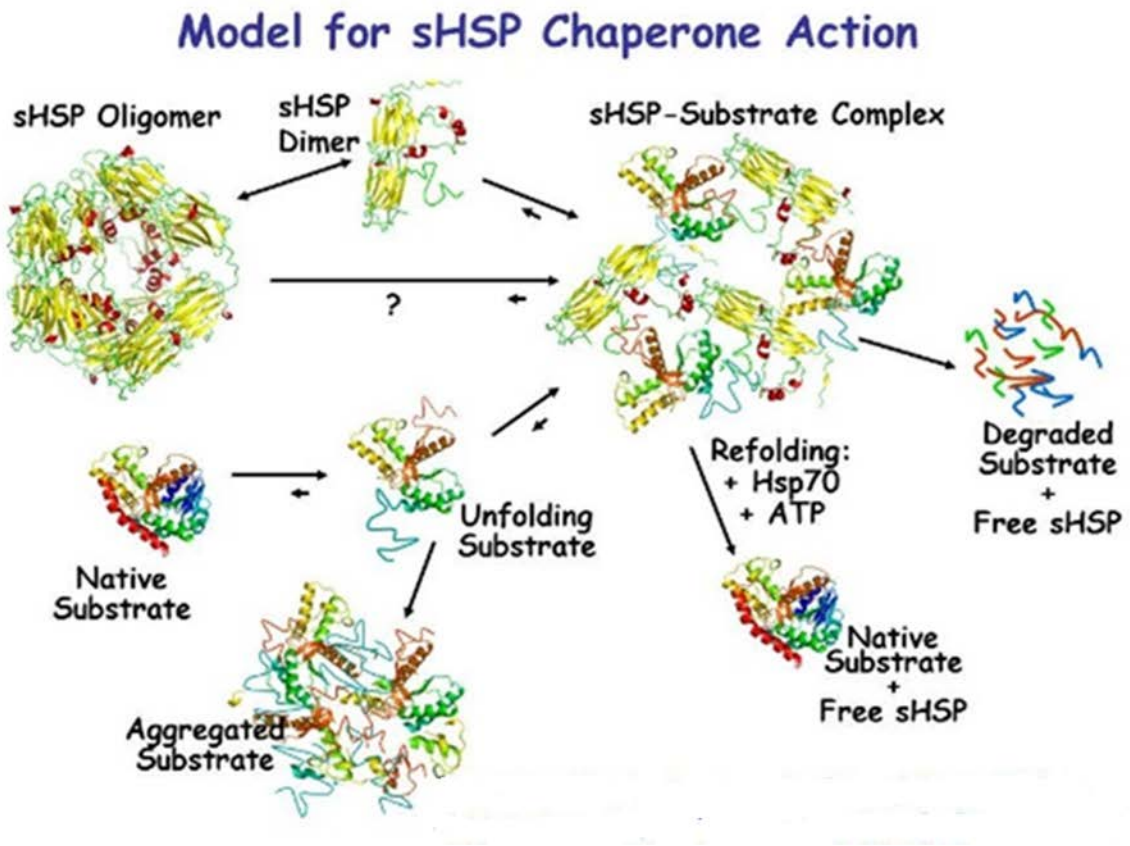
et al, 1998). Second, *Synechocystis* is a naturally competent for transformation and recombines DNA into its genome at homologous loci. Several attributes of *Synechocystis* make it an excellent model organism for the study of photosynthesis. Primarily, the rapid rate of growth, eight hours per generation, makes it a great choice for biochemical studies over slower growing plants. Another important attribute is the organism's ability to grow autotrophically, mixotrophically, and heterotrophically (with the addition of short pulses of blue light) (Anderson et al, 1991). In addition to these features, an important characteristic for the studies described in here, is the fact that the *Synechocystis* genome encodes only a single small heat shock protein (sHSP), Hsp16.6, and deletion of the Hsp16.6 gene does not impair normal growth, but renders cells heat sensitive (Lee et al 1998). Thus, *Synechocystis* has enabled analysis of the relationship of sHSP structure and function in vivo at a level of detail that has not been possible in other model systems, due to presence of multiple redundant sHSP genes and/or the absence of a readily score able phenotype (Geise et al 2005).

Small Heat Shock Proteins (sHSPs)

Small heat shock proteins (sHSPs) are a ubiquitous class of ATP- independent chaperones that are found in nearly every known organism. sHSPs range in size from 12-42 kilo Daltons and can be found in large oligomers ranging from 12 to >32 subunits. The typical architecture of a sHSP monomer is composed of three domains; a disordered N-terminal arm, a beta-sandwich α-crystallin domain, and a flexible C-terminal extension. The N-terminal domain is the most variable region of the protein, having little conservation between species. The α-crystallin domain is the most highly conserved region and adopts β-sandwich conformation composed of 7 to 8 anti-parallel β-strands (Basha et al, 2012). The C-terminus contains an I-X-I, a motif which is important for oligomerization (Basha et al, 2012). This motif is essential for formation of sHSP oligomers and when removed from sHSP Hsp16.6, only dimers form (Giese et al 2002).
Multiple have proposed various mechanisms of sHSP function. Data from previous studies suggest that the inactive state of the sHSP is the oligomer. Upon stress, these oligomers disassemble into active species, which are thought to be dimers. These dimers interact with misfolding substrates and may prevent their aggregation. Additionally, sHSPs may be essential facilitators for downstream refolding, unfolding, or degradation of some of their substrates. These mechanisms are still unknown and work is being done to address the roles that sHSPs have in all of these pathways (Figure 1).

Figure 1.

Model of sHSP chaperone action. sHSP oligomers disassociate under stress into active dimers. These dimers form sHSP-substrate complexes with denaturing protein based on unknown mechanisms of substrate recognition. sHSPs in these complexes are thought to facilitate the interaction of substrates with other molecular machines for disaggregation, refolding, or degradation (Van Montfort et al, 2001-Basha et al. 2012, Haslbeck and Vierling 2002).



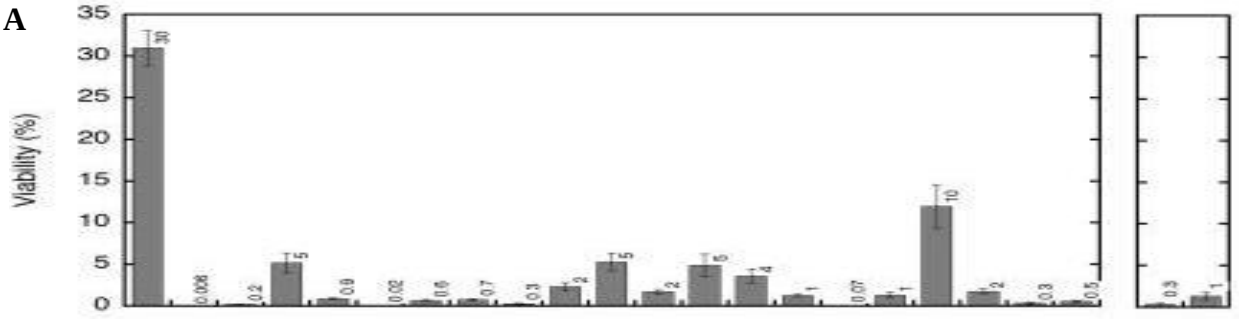
Hsp16.6 and previous work

As stated above, Hsp16.6 is the only sHSP encoded in the *Synechocystis* genome. It is found in position 460250 to 460690 of the chromosome and is annotated as sll1514 (Nakamura et al, 1998). It is 146 amino acids in length and has the same conserved architecture as other sHSPs. Mass spectrometry data suggest that this protein adopts a wide range of higher order oligomeric species ranging from a 12mer to a 28mer, with the predominant species being a 18mer (J. Benesch, Oxford, U.K., personal communication). Previous work has shown that ΔHsp16.6 *Synechocystis* have a loss of viability under stresses that wild type cells would normally survive. Under permissive laboratory conditions, 30°C under 100μmol/$m^{2/}$s light, Hsp16.6 is undetectable by western analysis. After treatment at 42°C for 2 hours, Hsp16.6 is expressed highly and accumulates to a maximum level of 0.5% of the total cellular protein (Basha et al 2004). Under these conditions, the viability of wild type cells and Δ16.6 cells is equivalent. However at 44°C for 8 hours, a >90% reduction in Δ16.6 viability can be observed, while wild type appears unaffected. Additional work set out to question the mechanics of the sHSPs in vivo by screening for point mutants of Hsp16.6 that phenocopied the heat sensitivity of the knock out strain (Figure 2). This study found several mutations that altered the ability of Hsp16.6 to oligomerize and mutations that altered the interaction of Hsp16.6 with substrates. Only mutants with altered substrate interaction and not altered oligomerization were considered for further study. Two of these mutations, E25K and L9P, have opposing biochemical phenotypes. E25K is unable to release substrate during the same amount of time that the wild type protein does. While L9P does not interact with substrates at any stage of heat stress (Figure 3) (E. Basha, personal communication). Additional in vitro data show that these mutants function well as chaperones with model substrates, in contrast to their inability to confer heat tolerance in vivo (Giese et al 2005). Other work has been performed to identify substrates Hsp16.6 so as to further understand the role of this protein in vivo.

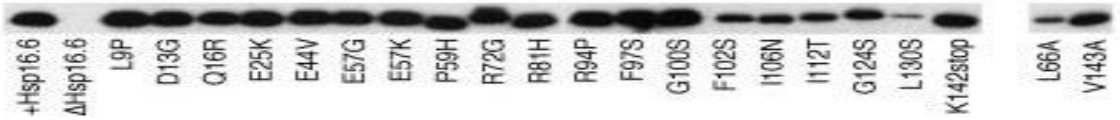
Figure 2.

Viability of mutants and Hsp16.6 model. A) Percent viability of different Hsp16.6 mutants after heat treatment at 44°C for 8 hours. (Giese et al, 2005). **B)** Western blots of Hsp16.6 variants from whole cell extracts of each sHSP mutant after 2 hours at 42°C, confirming expression of each mutant protein (Giese et al 2005). **C)** A model of a monomer of Hsp16.6 based on the structure of Hsp16.9 from wheat (PDB: 1GME). Red: N-terminal arm; the green is the α-crystalline domain, blue: C-terminal extension. In the N-terminal arm, L9P and E25K are shown as spheres.

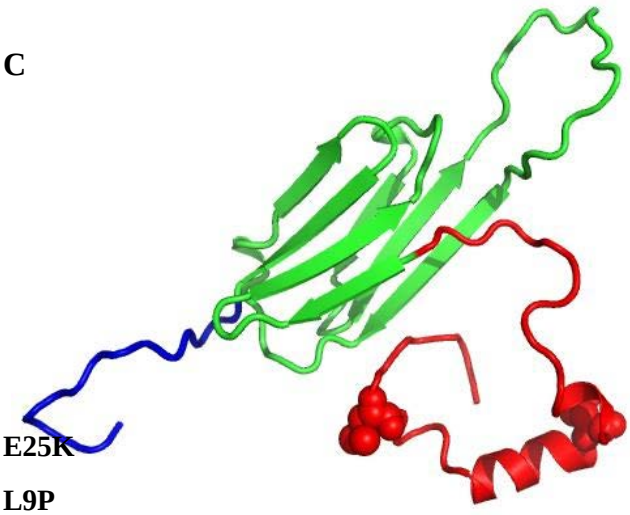
A



B



C

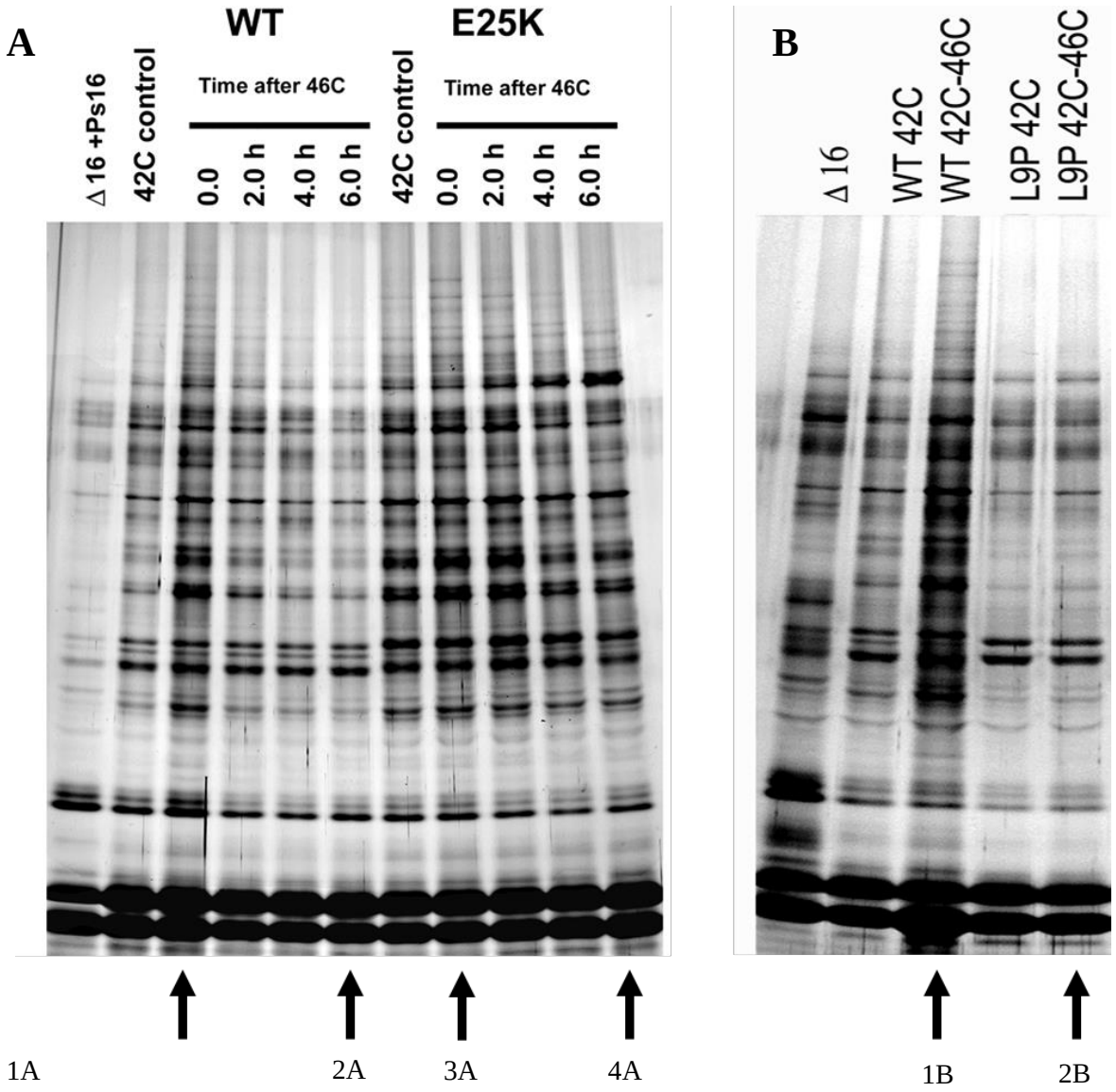


E25K

L9P

Figure 3.

L9P and E25K pull downs. A) Pull downs of strep tagged Hsp16.6 and associated proteins following heat stress.Lane 1A: wild type Hsp16.6 and associated proteins immediately after heat stress. Lane 2A: pulldown of wild type Hsp16.6 and associated proteins 6 hours after heat stress. Lane 3A: pulldown of the mutant E25K Hsp16.6 and associated proteins immediately after heat stress. Lane 4A: pulldown of the mutant E25K Hsp16.6 and associated proteins 6 hours after heat stress. (Eman Basha personal communication, unpublished) **B)** Pulldown of Hsp16.6 and associated proteins during heat stress. Lane 1B: pulldown of the wild type Hsp16.6 and associated proteins during heat stress. Lane 2B: pulldown of the mutant L9P Hsp16.6 and associated proteins during heat stress. (Eman Basha personal communication, unpublished)


A
WT
E25K
Δ 16 +Ps16
42C control
Time after 46C
0.0
2.0 h
4.0 h
6.0 h
42C control
Time after 46C
0.0
2.0 h
4.0 h
6.0 h
B
Δ 16
WT 42C
WT 42C-46C
L9P 42C
L9P 42C-46C
1A
2A
3A
4A
1B
2B

Directed evolution of thermotolerance

Directed evolution is the use of selective forces and natural variation to acquire new proteins, rRNAs, or organisms with specific enhanced desired attributes. Many groups have used directed evolution to evolve proteins to be more proficient at their function under certain conditions. Take for example the evolution of *E.coli* GroEL to better fold substrate (Wang et al 2002). Other groups have used directed evolution to evolve new abilities in target proteins, such as the evolution cytochrome P450 to have multiple catalytic activities (Behrendorff et al 2015). Directed evolution has also been performed on entire organisms in order to increase certain abilities. One ability that has been evolved in several studies is thermotolerance. Caspeta et al evolved *S. cerevisae* to survive and grow at temperatures >40°C (permissive growth of *S. cerevisae* <34°C). This group found that altered sterols and membrane composition were the partial causation of the increased thermotolerance (Caspeta et al 2014). Closer to the work presented in this thesis, is the work of Tillich et al in 2014, in which this group directly evolved *Synechocystis sp* PCC. 6803 to grow at temperatures exceeding 45°C. Cultures in this work were exposed to increasingly higher temperatures and monitored for stable growth at which point the temperature was once again increased. Once achieving growth at temperatures higher than 45°C (the max being 45.8°C) they sequenced and found several alterations in the genomes of the evolved strains. Roughly 30 mutations were found in various frequencies between the seven evolved strains. Most notably, several strains had mutations located in the gene encoding ClpC, which is a component of the ClpP protease system (Tillich et al 2014). These studies show directional evolution is a powerful tool that can be used to understand many different systems and system dynamics.

Thesis Overview

Many questions remain about the function and mechanism of Hsp16.6. In addition, there are also numerous questions involving the general stability of the proteome during times of stress, as well as the global evolution of thermotolerance. The work presented in this thesis set out to address some basic questions regarding these ideas. Is the in vitro model of sHSP function the actual mechanism in vivo? Do sHSPs protect substrates by binding and preventing their insolubility? Is the thermotolerance of *Synechocystis* absolutely dependent on Hsp16.6, or can adaptations be made to circumvent the need for the sHSP? To address these questions, I set out to find suppressors of the two Hsp16.6 mutants; L9P and E25K. Each of these mutants have decreased viability after heat stress compared to wild type with L9P showing a 1,500 fold decrease and E25K a 15,000 fold decrease. Yet, unlike cells lacking Hsp16.6 (Δ16.6), the presence of these different mutant Hsp16.6 proteins may apply different selective pressures to the proteome that could cause different cellular adaptations. These adaptations may occur in associated proteins to facilitate sHSP activity or in substrates to provide them greater stability against heat denaturation, or potentially in pathways that have no direct association with Hsp16.6. Experiments to address these questions are described in the next six chapters. Chapter 2 contains the methods and conditions used in all experiments. Chapter 3 describes the development of the suppressor screen and the cell lines obtained from it. Chapter 4 pertains to the characterization of the suppressor strains. Chapter 5 explains the bioinformatics analysis performed on the re-sequenced genomes of the suppressors, which was performed in order to determine causative mutations. Chapter 6 is a description of the biochemical characterization of the sHSP and its substrates. Chapter 7 outlines future directions for this project.

CHAPTER 2

METHODS

Strains

A glucose tolerant, non-motile parent strain of *Synechocystis sp.* PCC 6803 from which subsequent strains were derived, was obtained from Dr. Fukuzawa (Kyoto University) in the 1990s. Four parental strains were derived from this strain;

1) sll1514 $^{\text{WT-sll1514}}$; (Null ClpB1) slr1641 $^{\text{tm1(ermE)}}$

2) sll1514 $^{\text{tm1(aada)}}$; (Null ClpB1) slr1641 $^{\text{tm1(ermE)}}$

3) sll1514 $^{\text{L9P-sll1514}}$; (Null ClpB1) slr1641 $^{\text{tm1(ermE)}}$

4) sll1514 $^{\text{E25K-ssl1514;}}$; (Null ClpB1) slr1641 $^{\text{tm1(ermE)}}$

These strains are referred to as Wild type, Δ16.6, L9P, and E25K, respectively (Torok, Goloubinoff et al. 2001, Giese and Vierling 2002). All four strains are deleted for ClpB1, one of two ClpB genes in Synechocystis, which encodes an AAA+ protein disaggregase in the Hsp100 family. (Note that deletion of ClpB2 is lethal). Previous work indicated that the heat sensitivity phenotype of Δ16.6 strains was more severe in the background of the ClpB1 deletion.

***Synechocystis* growth**

Cells were grown in standard liquid BG-11 media (pH 7.8; buffered with 10mM HEPES) containing either 5mM glucose or no glucose, depending on the experiment. Cells were grown under continuous 100 μmol/m^2/s fluorescent light at 30°C. Small 5 mL cultures were rotated at 200 RPM on a New Brunswick culture wheel, while large 100mL cultures were grown in 250mL Erlenmeyer flasks in an Innova Incubator under 100 μmol/m^2/s fluorescent light while shaking at 200 RPM. All strains are preserved at -80°C in 30% glycerol/ 70% BG-11 media. Supplemental Table 1 lists of all strains currently stored at -80°C.

Directed evolution

To screen for strains that had increased heat tolerance, cells of each of the four parental strains, wild type (as a control), Δ16.6, L9P, and E25K, were first grown to an O.D. of 1.0 at 730 nm. No mutagen was employed, rather results depended on the appearance of spontaneous mutations (transitions, transversions, insertions, deletions, duplications etc.). Approximately 100,000 of cells of each strain were plated on solid, 1.5% Bacto agar (Fisher Scientific, Waltham Ma.) BG-11 media containing no glucose (one cell line per plate). Cells were allowed to acclimate to the solid media for 24 hours at 30°C under continuous 100 μmol/m^2/s fluorescent light prior to being subjected to screening conditions. Following acclimation, cells were exposed to a prohibitive temperature of 44°C for sequentially increasing amounts of time (19 exposures beginning at 5% of 8 hours and increasing in 5% time increments). Between heat stresses, cells were allowed 12 hours recovery at 30°C. Throughout the screen, light was held constant at 100 μmol/m^2/s. Following this screen, individual colonies that survived were streaked and retested to confirm the heat tolerance phenotype.

Plasmid construction and transformation

Two plasmids were used to transform *Synechocystis* in this study. Both originated from pNaive, an in house plasmid constructed by Kim Giese. PNaive contains the sequences 500 bp upstream and downstream of sll1514 that allow for homologous recombination. PHK was constructed by swapping the aaDA gene ($Spec^R$) with a Kan^R gene from phk-2r (A gift from Dr. Fukazawa at Kyoto University) and was used for knocking out the entire Hsp16.6 gene (sll1514). This was done by cutting both plasmids with the restriction enzyme Pst1, which cuts at both ends of the resistance cassettes. The Kan^r gene was gel purified using a gel purification kit (Qiagen, Redwood City Ca.). The cut pNaive plasmid was not purified, but rather treated with Antarctic Phosphatase (New England Biolabs) before being used for T4 ligation. Ligations were transformed into DH5α cells (New England Biolabs) and verified through PCR. The same protocol was performed for PNK. PNK replace the sll1514 locus with a wild type copy of the Hsp16.6 gene in addition to a Kan^R gene.

Thermotolerance assay

Synechocystis cells were grown to an O.D. of 1.0 at 730 nm in 5 mL liquid BG-11 media containing no glucose. A ten-fold dilution series was performed from 1 to 10^{-8} in fresh BG-11 media. 5 µL of each dilution was dotted onto solid BG-11 media containing 140 mM $MgSO_4$, 5mM glucose, and 1.5% w/v Bacto-Agar (Fisher Scientific). Cells were allowed to acclimate to solid media for 24 hours before being subjected to an 8 hour stress at 44°C in the dark. Following this stress, cells were allowed to recover at 30°C for 4 to 7 days under 100 µmol/m^2/s continuous fluorescent light before imaging.

Protein sample preparation

Immunoblotting was performed on both whole cell lysates and fractionated lysates from cells that were stressed under several different conditions, depending on the experimental design. Cells used for quantitation of specific proteins were stressed under three different regimes depending on the experiment; 2 hours at 42°C followed by 1 hour at 46°C (solubility assay 1); 2 hours at 42°C followed by a 12 hours recovery period at 30°C followed by 1 hour at 46°C (solubility assay 2); or 2 hours at 44°C (sHSP expression). To fractionate the proteome for solubility assays, following heat stress cells were pelleted at 4000xg for 10 min. Pelleted cells were resuspended in 400µL of lysis buffer (25 mM HEPES-KOH, 0.2 M NaCl, 0.5% Triton X-100, 5 mM ϵ-aminocaproic acid, 1 mM benzamidine, and 5 unit/ml potato apyrase) per 10 mL of cells. Resuspended cells were lysed with the addition of 0.1 mm glass beads and vortexing at the highest speed (6 x 30s). Between vortexing cycles, samples were cooled on ice for 1 min. 300µl of supernatant was then placed in a new tube and centrifuged for 15 min at 16,250 x g at 4°C. The supernatant was removed from the pellet and the pellet was resuspended in a volume of lysis buffer equivalent to the soluble supernatant (300µl) (Basha et al. 2004). 4x Laemmli buffer was then added to make the final sample 1x Laemmli buffer. Samples were then boiled for 15 minutes before use. Protein concentration for both sample preparations was determined using a Coomassie G-250 dot blot assay and a Li-Cor fluorescence imager (Li-Cor Lincoln NE).
Total protein samples used for sHSP expression were prepared by pelleting 2ml of cells at an O.D of 1.0 at 730nm at 4000 x g for 10 mins. Pellets were then resuspended in 1x Laemmli buffer, 100 µL per 0.01 mg of cell pellet and boiled for 15 mins. Protein concentration for both sample preparations was determined using a Coomassie G-250 dot blot assay and a Li-Cor fluorescence imager (Li-Cor Lincoln NE).

Immunoblotting

Immunoblotting was performed using four different polyclonal rabbit antibodies; anti-Hsp16.6 (1:10,000) (*Synechocystis* sll1514*),* anti-GroEL (1:10,000) (*E.coli* b4143*),* anti-FBA II (1:5000) (*Synechocystis* sll0018*),* and anti-EF-G (1:5000) (*Synechocystis* slr1461). Volume equivalent to 10 µg of total protein was loaded on a 15% SDS page gel for separation. Proteins were transferred to nitrocellulose membrane using a semidry transfer protocol. Following transfer, membranes were washed 3 times for 5 mins each in PBST. Blocking was performed in 6% milk-PBST for 1 hr before the addition of primary antiserum. Incubation in primary antiserum was performed overnight at 4°C on an orbital shaker (50 RPM). Washing was then performed 3 times with PBST for 5 mins each time before the addition of the fluorescently labeled anti-rabbit antibody (LiCor, Lincoln NE). Incubation in secondary antiserum was performed for 50 mins in darkness. Following incubation, 3 washes were performed with PBS for 5 mins each. Membranes were then imaged with an Odyssey Clx imager (LiCor, Lincoln NE).

Growth curves

Synechocystis cells were grown to an approximate O.D. of 1.0 at 730nm before being diluted 5- fold into fresh BG-11 media to an OD of 0.2 at 730 nm for analysis of growth rates. Cells were grown either fully autotrophically in BG-11 media without glucose, or mixotrophically in media containing 5 mM glucose. Permissive growth conditions were as follows; continuous 100 µmol/m^2/sec fluorescent white light with an ambient temperature of 30°C. Prohibitive growth conditions were as follows: continuous 100 µmol/m^2/sec fluorescent light with an ambient temperature of 42°C. All strains were grown in biological triplicate, starting from the same parent culture. Spectral reads were taken at 12 hours, 36 hours, and 72 hours using an Implen Nanophotometer^{tm} spectrophotometer (Implen, München, Germany) to read 500 µl of each sample at 730nm.

Fluorescence activated cell sorting (FACS)

Cells were grown to an O.D. of 1.0 at 730 nm in both 5 mL glucose containing (5 mM glucose) and 5 mL glucose free BG-11 media. 2 mL of culture were removed and placed in a sterile culture tube and treated with rifampicin to a final concentration of 30 µg/mL for 12 hours at 30°C under continuous 100 µmol/m^2/sec fluorescent light. Cells were centrifuged at 4000xg and the pellet resuspended in 70% ethanol for 10 hours at 4° C. Following fixation, cells were pelleted at 6000 x g and washed three times with FACS prep buffer (50mM NaCitrate; pH □6.0). Samples were incubated for 1 hour at 37°C following suspension in FACS preparation buffer containing 0.25 mg/ml RNAse (Chien Lab, UMass Amherst). Additional FACS preparation buffer equivalent to one volume of sample was added following incubation. Following RNA digestion, 0.3 µl of sample was added to 200 µL of FACS running buffer (10mM Tris HCl, 1mM EDTA from 0.5M stock at pH 8.5, 50mM NaCitrate unpHed, 0.01% Triton X-100, Sytox green 1uM (Thermofisher Agawam, MA) and run on a Millipore Guava System FACS analyzer. Data were analyzed using FlowJo™ (Mycyte.org)

Genomic DNA extraction and preparation

Genomic DNA was extracted from cells using a Qiagen Plant Genomic DNA extraction kit (Redwood City CA). Cells were grown in 100 mL of BG-11 media without glucose until an O.D. of ~2.0 at 730nm. Cells were pelleted at 4000xg for 10 mins at room temperature. A cell pellet of approximately 100 mg was used in the extraction (see Qiagen Protocols.com). Following extraction, the concentration of the purified DNA, which had been resuspended in 100µl of Qiagen elution buffer, was determined using a Thermofisher Nanodrop, and purity/quality was determined from analyzing 100 ng of each sample on a 0.7% agarose gel run at 70 volts. Final concentration and purity were determined using a Qubit Fluorimeter (Life Technologies, Carlsbad, CA).

Next generation sequencing

Sequence data from six samples, Wild type, Δ16.6, L9P, A6, A4, and C3, are reported here, and 24 more samples have been prepared for sequencing. The six samples were sent to the Tufts University Genomics Core (Tufts University, Boston MA). These samples were prepared using a TruSeq Nano DNA kit (Illumina, San Diego CA.). Samples were run on a MiSeq Illumina sequencer (Illumina, San Diego CA.) using version 2 (second generation), 500 cycle (total number of synthesis rounds) chemistries. Samples were multiplexed on the same MiSeq lane and sequenced.

Bioinformatics

Data obtained from the MiSeq runs was received in a FASTQ format already demultiplexed. FASTQ files were groomed using FastQ Groomer (Blankenberg et al 2012). BWA-mem was used to assemble the reads against the reference genome (Heng et al 2013). All default settings were used when mapped against GCF_000009725.1_ASM972v1_genomic from Ensemble Genomes (Kersey et al 2013). BWA-mem assembled genomes were used for variance detection, statistical analysis, and annotation. Variance detection was performed using Varscan (Koboldt et al 2012) using a frequency parameter of 33.3% and a minimum quality score of 15. Statistics pertaining to coverage and proper alignment were generated using SAM tools (Li et al 2012). Annotation was completed with snpEff (Cingolani P et al 2009), using GCF_000009725.1_ASM972v1_genomic as the reference genome and Synechocystis_sp_pcc_6803_gca_000009725_1.GCA_000009725.1.27.gff3 as a gene annotation file (Kersey et al 2013).

CHAPTER 3

SUPRESSORS OF LOSS-OF-FUNCTION HSP16.6 MUTANTS

Introduction

Previous work on Hsp16.6 uncovered several dozen point mutations that altered the function of this sHSP in vivo (Fig. 2). Further work focused on a subset of these mutants that retained the oligomeric architecture of the wild type protein and were able to protect specific model substrates from heat denaturation in vitro. Proteins with mutations at two positions in the N-terminal arm of the sHSP, L9P and E25K, displayed opposing biochemical phenotypes. L9P appears unable to interact with substrates in vivo during heat stress, while E25K does not release its associated proteins during recovery from heat stress. Viability of cells carrying these mutant sHSPs is approximately equivalent to the Δ16.6 strain (Fig. 2). In order to further characterize these mutants, an intragenic suppressor screen was performed by in vitro random mutagenesis of either mutant Hsp16.6 gene, reintroducing the mutant genes into *Synechocystis* and screening for enhanced heat tolerance. This screen produced several suppressors of L9P and no suppressors of E25K. However, one suppressor of L9P, mutation E24K, recovered this mutant's loss of function. It was later discovered that L9P and E25K also complemented one another (Basha, unpublished). These intragenic mutants were used to further characterize the biochemistry of Hsp16.6. However, this in vitro method only allowed for the production of intragenic suppressors. In order to discover extragenic suppressors, a second screen was designed. This screen relied on the principles of directed evolution and yielded both extragenic and intragenic suppressors of both L9P and E25K.

Results

A number of trials were required to develop conditions that resulted in the isolation of heat tolerant suppressors from the sHSP mutant strains. The first screen utilized a solid media approach. Cells were plated on standard BG-11 solid media containing no glucose and grown for 24 hrs at 30°C under continuous 100 µmol/m^2/s light. Every 24 hours, the temperature was increased 2°C until a final temperature of 44°C was achieved after 7 days. Following this treatment cells were found to be chlorotic (dead) and no thermotolerant strains were recovered. A second attempt with the same basic approach, but in which longer intervals between temperature increases were used, yielded similar results. Therefore, a second approach was adopted in which cells were exposed to the non-permissive temperature of 44°C for increasing amounts of time under continuous 100 µmol/m^2/s light. In between exposures to the non- permissive temperature, cells were allowed 12 hours of recovery at 30°C. Cells were first exposed to 44°C for 10% of 8 hours (48 minutes) and subsequent exposures increased by increments of 5% (24 min) until 8 hours at 44°C was achieved (Figure 4).

The first screen done in this fashion yielded three L9P suppressor strains; A6, A4, C3. The strains were detected as healthy green colonies while the remainder of the colonies on the plate were severely chlorosed. A second screen using the same method and with the assistance of programed temperature changes in the incubator, yielded a total of 24 mutants; 13 L9P and 11 E25K. Thermotolerance assays were performed on each of these strains to confirm that they had increased thermotolerance over their respective parent strain (Figure 5). The Hsp16.6 gene was amplified from genomic DNA of each strain and sequenced. All L9P suppressor strains were confirmed to only have the corresponding L9P mutation in the Hsp16.6 gene. Of the 11 E25K suppressors, 10 were found to have no additional mutations beyond the original E25K mutation, while one mutant had both the E25K mutation as well as a mutation that coincidentally altered Leu 9 to Pro. Therefore, at total of 27 extragenic suppressors were obtained.

Figure 4.

Directed evolution workflow. Time spent at 44°C increased in increments of 5% beginning at 10% of 8 hours (48 minutes). Total time spent at 44°C was 79.6 hours (3.31 days).

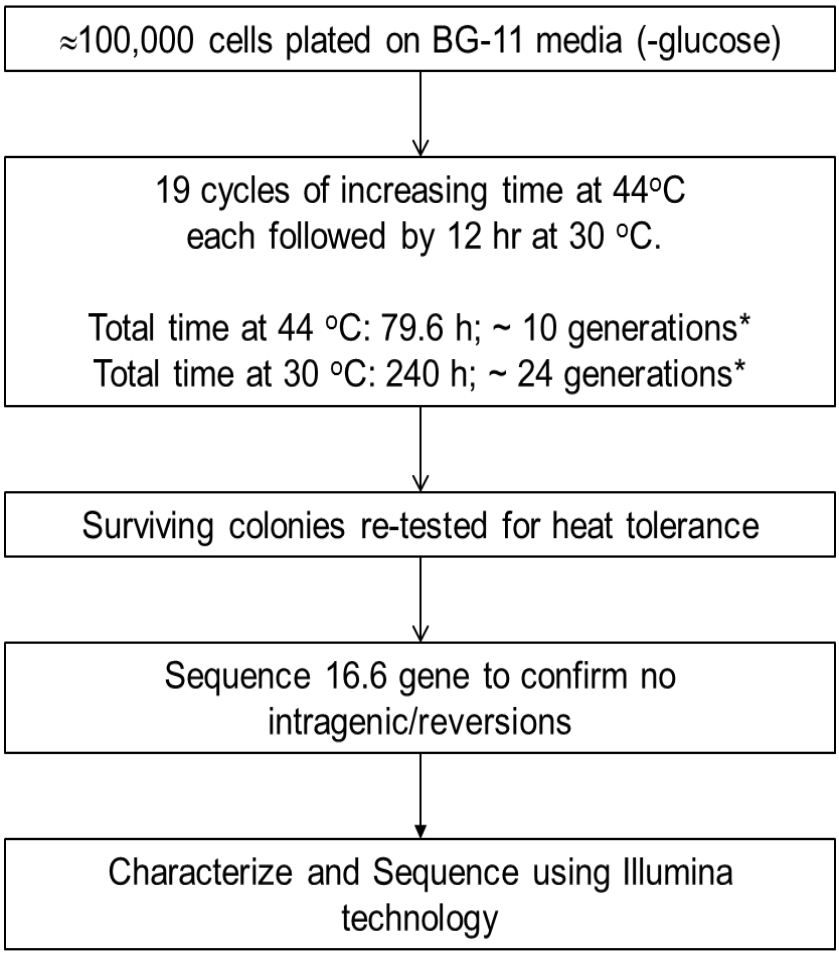
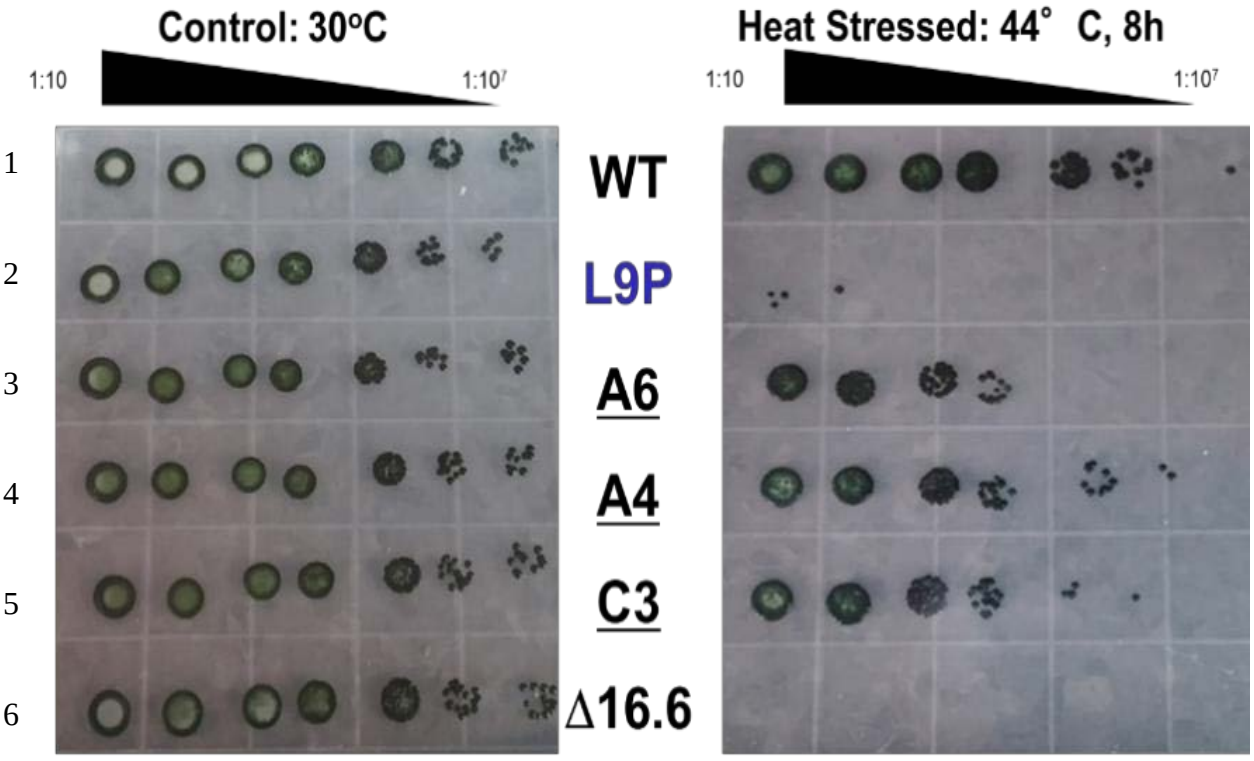


Figure 5.

Suppressor thermotolerance assay. Left: Control plate grown continuously at 30°C; Right: Plate was exposed to 44°C for 8 hrs. Row 1wild type control, row 2: L9P parent strain, rows 3 to 5: A6, A4 andC3 suppressors of the L9P mutation, and row 6: Δ16.6 strain. Note that plates were imaged approximately two weeks after being heat stressed. This additional time accounts for the white centers in the denser colonies that are a result of cell death due to drying on the plates.



Discussion

The first goal of this research was to determine whether or not strains carrying mutant sHSPs could spontaneously mutate to regain thermotolerance. While initial attempts to directly evolve thermotolerance were unsuccessful, the second approach succeeded in yielding a total of 28 mutants in two separate screens. The low yield of the first successful screen (3 suppressors) was most likely due to performing temperature changes manually. Because of human error, additional time was added to some of the steps (both stress steps and recovery). The second screen had a significantly higher yield of suppressors 13 L9P and 11 E25K, most likely due to the precise temperature changes conducted using the automated settings on the incubator. Note that a total of 56 colonies, 34 L9P and 22 E25K, were picked from the plates a putative suppressors in the second screen and only 24 were reconfirmed as suppressors (42.9%). It is estimated that the cells in this screen underwent a total of 34 generations, based on measures of permissive and non- permissive growth rates. This means that the approximately 100,000 cells that were plated at the start the screen had the potential of generating 1.7*10^15 daughter cells. Based on the number of plates, the total possible number of cells, and the number of surviving colonies at the end of this assay, it can be assumed that the random acquisition of mutations conferring thermotolerance is very low. Thermotolerance of all of the suppressors appears equivalent to that of Wild type under the conditions tested. We can conclude from these data that the loss-of-function of a chaperone can be recovered by extragenic mutations. The nature of these mutations may vary, as they may appear in co-chaperones, substrates, or potentially in pathways that are not directly related to the sHSP. Further investigation and characterization are required in order to determine the mechanism of suppression and how it may relate to sHSP function.

CHAPTER 4

CHARACTERIZATION OF THE SUPPRESSOR MUTANTS

Introduction

Several questions arose from the isolation of suppressors of the mutant Hsp16.6 strains. One major question was if thermotolerance in *Synechocystis* was dependent sHSP. No other organism previously shown to require a sHSP for thermotolerance has been shown to evolve to no longer need it. To test whether our mutants had evolved to no longer require Hsp16.6, a knock-out of the Hsp16.6 locus in each strain was performed. Could these evolved lines survive heat stress without a sHSP at all? Additionally, it was of interest to investigate other attributes of the suppressor strains in order to determine if all strains had evolved a similar mechanism for regaining thermotolerance. While the thermotolerance assay is a good measure of acute heat stress tolerance, would tolerance also be observed under conditions of chronic heat stress? Are the mechanisms of acute thermotolerance and chronic thermotolerance the same?

Results

The dependence of thermotolerance on the small heat shock protein

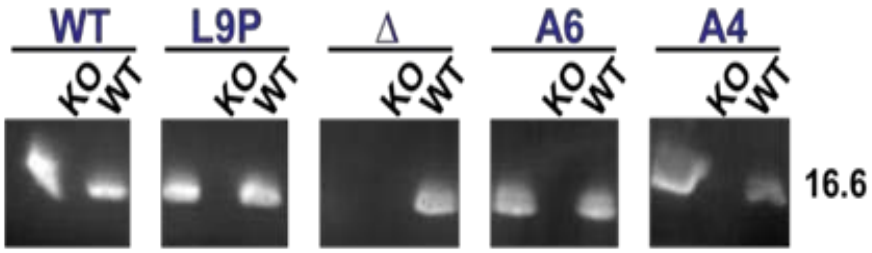
Each suppressor strain was transformed with the PNK plasmid designed to replace the parental mutant Hsp16.6 allele with a wild type sll1514 gene (Hsp16.6) linked to a Kan^R gene, and with the PHK plasmid designed to knock-out the sll1514 locus (Hsp16.6) with a Kan^R gene. Thus, the plasmids were used to create isogenic suppressor strains with and without Hsp16.6. Confirmation of either the knockout, by the PHK plasmid, or the knock-in of the wild type Hsp16.6 gene was performed by western analysis (Figure 6A), and Sanger sequencing. Following confirmation of the genotype, strains

were assayed for thermotolerance to check for the requirement of Hsp16.6 in the suppressor back ground. Wild type, Δ16.6, L9P, and E25K strains transformed with these plasmids all showed thermotolerance when transformed with PNK and a loss of thermotolerance when transformed with PHK. Of the suppressor mutants successfully transformed and confirmed, only four appeared dependent on the presence of either a mutant or wild type Hsp16.6 allele. The other 16 suppressors that were successfully transformed displayed thermotolerance in the presence and absence of Hsp16.6 (Table 1; Appendix 1).

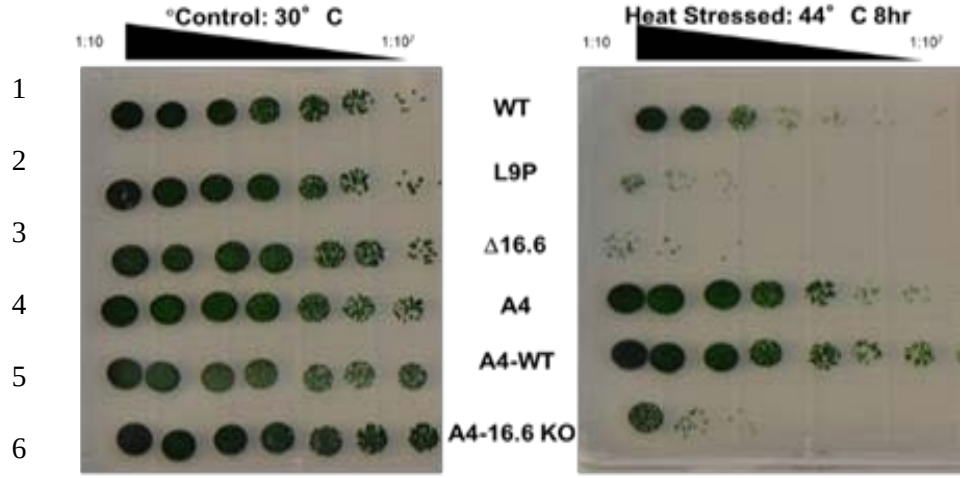
Figure 6.

Hsp16.6 dependence assay. A) Western analysis of strains transformed with the PNK and PHK plasmids. Anti-16.6 was used to detect Hsp16.6 in total cell protein following a 2 hours 42°C treatment. 10 µg of total protein was loaded per lane. **B)** Thermotolerance assay of the L9P suppressor strain A4. A4 is an example of a suppressor mutant that is dependent on Hsp16.6. Row 1 is wild type, row 2 is parental (L9P), row 3 is the Δ16.6, row 4 is the parent suppressor (A4), row 5 is the A4 transformed with PNK (wild type 16.6 gene knock in), row 6 is the A4 transformed with PHK (16.6 locus knock-out). **C)** Thermotolerance assay of the L9P suppressor strain A6. A6 is an example of a suppressor mutant that is independent of the sHSP. Row 1 is wild type, row 2 is parental (L9P), row 3 is the Δ16.6, row 4 is the parent suppressor (A6), row 5 is the A6 transformed withPNK (wild type 16.6 gene knock in), row 6 is the A6 transformed with PHK (16.6 locus knock-out).

A


WT
L9P
Δ
A6
A4
KO
WT
16.6

B


°Control: 30° C
Heat Stressed: 44° C 8hr
1:10
1:10⁷
1 WT
2 L9P
3 Δ16.6
4 A4
5 A4-WT
6 A4-16.6 KO

C

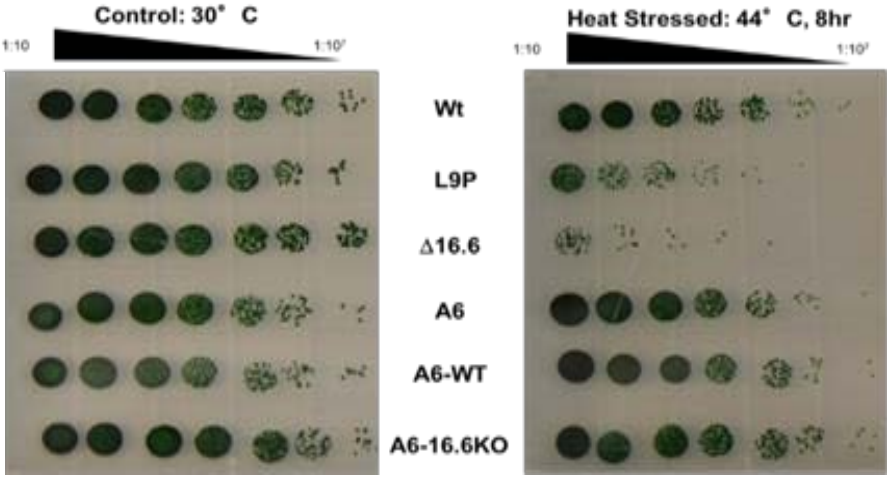

Control: 30° C
Heat Stressed: 44° C, 8hr
1:10
1:10⁷
1 Wt
2 L9P
3 Δ16.6
4 A6
5 A6-WT
6 A6-16.6KO

Table 1.

Hsp16.6 dependence assay for all tested suppressors. Results from all successfully transformed and confirmed Hsp16.6 knockouts in suppressor backgrounds. Rows marked as "Yes" denote strains that require a sHSP gene (wild type or mutant) to be present in order to be thermotolerant. Rows marked as "No" denote strains that show thermotolerance in the absence of the sHSP. Rows marked N.A. are strains that still require testing. All strains reported were tested in three separate experiments.

Strain	sHSP Dependence	Strain	sHSP Dependence
Wt	Yes	L9P-17-Syn	No
Del	Yes	L9P-19-Syn	Yes
E25K	Yes	L9P-22-Syn	No
L9P	Yes	L9P-24-Syn	No
L9P-A6-Syn	No	L9P-33-Syn	Yes
L9P-A4-Syn	Yes	E25K-2-Syn	N.A
L9P-C3-Syn	Yes	E25K-3-Syn	No
L9P-1-Syn	No	E25K-4-Syn	No
L9P-2-Syn	N.A	E25K-6-Syn	N.A.
L9P-5-Syn	No	E25K-8-Syn	No
L9P-7-Syn	N.A	E25K-9-Syn	No
L9P-11-Syn	No	E25K-11-Syn	No
L9P-12-Syn	No	E25K-17-Syn	No
L9P-14-Syn	N.A	E25K-18-Syn	N.A
L9P-15-Syn	N.A	E25K-19-Syn	No
L9P-16-Syn	No		

Suppressor mutants survive chronic heat stress better than Wild Type

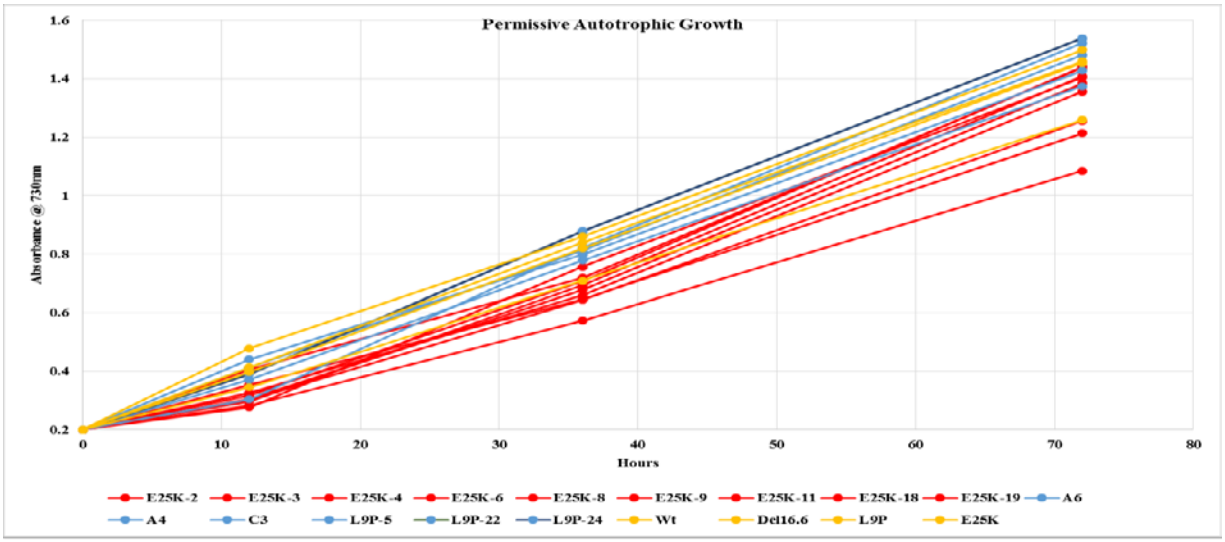
In order to further group the independently generated suppressor strains, additional assays were performed. One of these assays aimed to test the ability of the suppressors to handle chronic heat stress.

Grown under permissive temperatures, all strains show equal growth rates suggesting that any mutations conferring thermotolerance do not negatively affect the ability of these strains to grow (Figure 7A). Additionally, no strains appear to have a higher rate of growth. Several stress conditions were tested at which no cell lines grew. Continuous growth at 46°C, as well as 44°C led to the death of all strains. Continuous growth at 42°C led to the death of the four parent strains; Wild type, Δ16.6, L9P, and E25K. In contrast, suppressor strains continued to show linear growth at this temperature, albeit with 2-fold lower rates than seen at permissive temperatures (Figure 7B). Note that not all lines are shown in this figure. Further testing is required. Additionally, several other conditions were tested on a subset of strains; wild type, L9P, Δ16.6, A6, A4, C3.

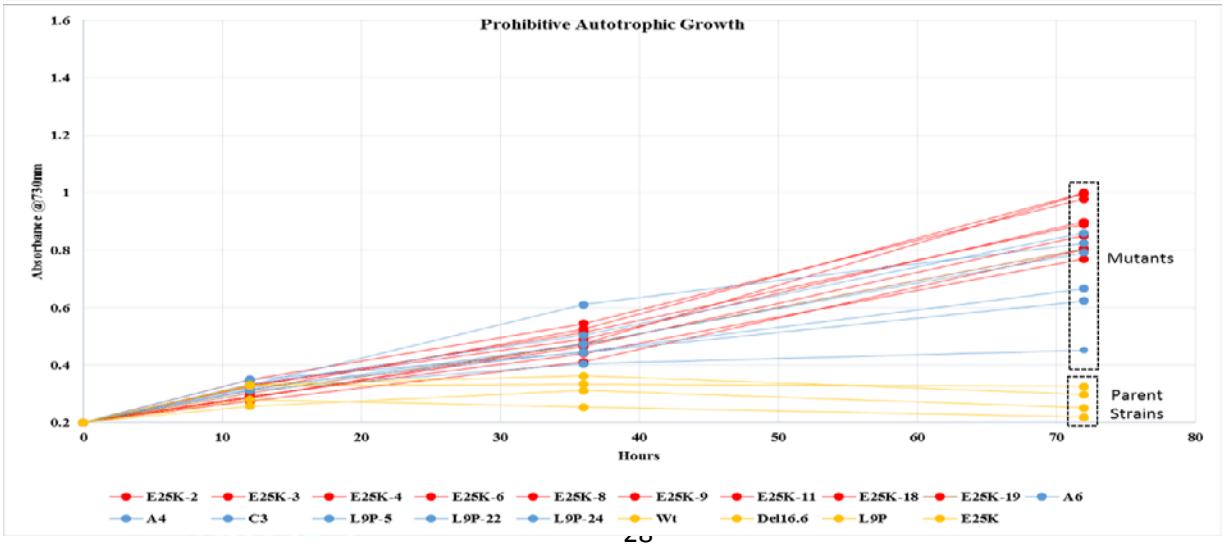
Figure 7.

Permissive and stress autotrophic growth curves. A) Autotrophic growth curves of suppressor lines from both screens 1 and 2 as well as wild type, Δ16.6. L9P, and E25K. These curves were generated from cells growing at 30°C under continuous 100µmol/m^2/s light in 5 mL Bg-11 cultures containing no glucose. **B)** Autotrophic growth curves of suppressor line from both screens 1 and 2 as well as wild type, Δ16.6. L9P, and E25K. These curves were generated from cells growing at 42°C under continuous 100µmol/m^2/s light in 5 mL Bg-11 cultures containing no glucose.

A



B



Suppressor and parental strains have the same ploidy, which doubles during mixotrophic growth

In order to determine the DNA content of the *Synechocystis* strain, FACS analysis with dsDNA intercalating dye was used. Cells grown under autotrophic conditions had fluorescence emissions comparable to 4-6N, based off of the known *Caulobacter crescentus* standard. *C. crescentus* has a recorded N=2 prior to division and N=1 post division, with experimental FACS data suggesting emissions to be 711 and 355, respectively. Using the genome ratio 0.982 (*Synechocystis -3.95Mb/ C. crescentus-4.02Mb)* multiplied by *C. crescentus* N=1 emission of 355, we can estimate the *Synechocystis* genome to have an N=1 emission value of 348.81. Under mixotrophic conditions, ploidy doubled in all cell lines to an extrapolated range of 10-13N (Table 2). Figure 5 is an example of the spectra obtained from a wild type sample grown under autotrophic and mixotrophic conditions. Additional experiments, not shown, confirm that the addition of glucose alone is not the cause of the emissions increase. Two hours prior to cell fixation, glucose was added to each cell culture and allowed to incubate in a shaking incubator under 100 μmol/m^2/s light. No emission shift was detected suggesting that mixotrophically grown cells have higher ploidy than autotrophically grown cells.

Figure 8.

FACS spectra. A) *Caulobacter crescentus* B) Autotrophically grown wild type *Synechocystis* Emission= 1711 *C)* Mixotrophically grown wild type *Synechocystis* Emission=3342. *C. crescentus* has an N=2 in this sample, equivalent to an emission of 711. Using relative ratios based on the respective genome sizes of each organism, N values for *Synechocystis* strains can be extrapolated.

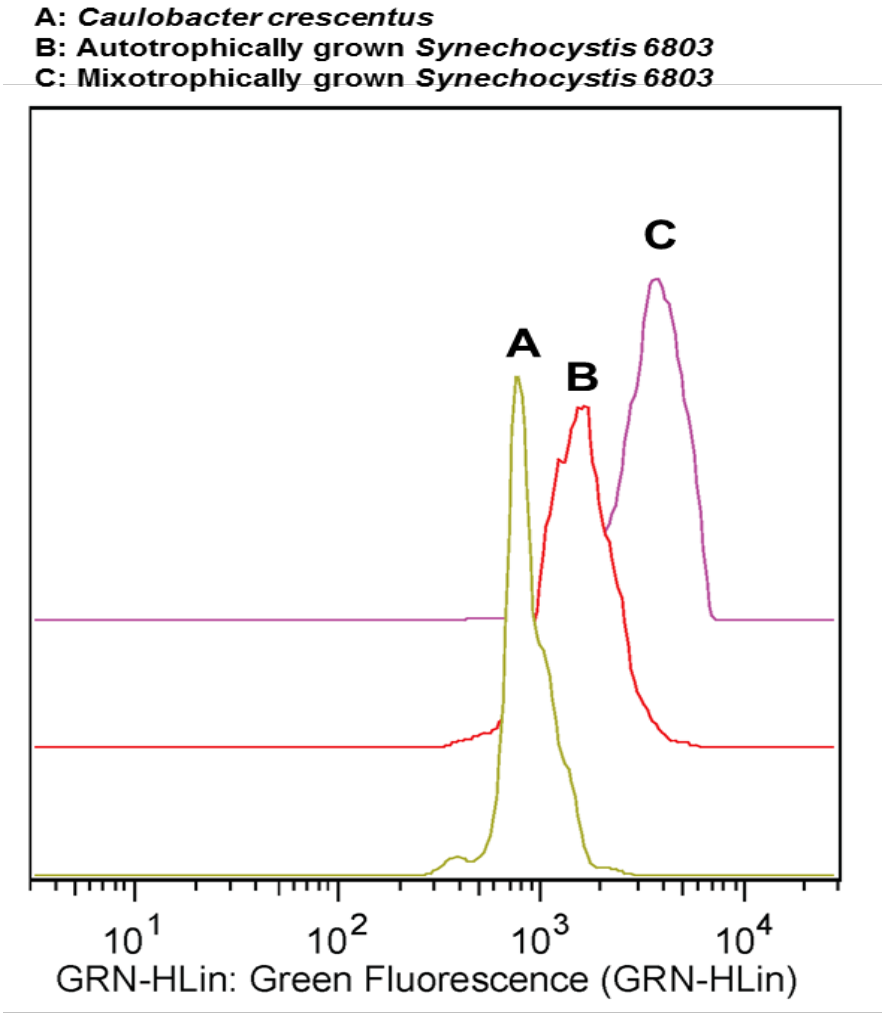


Table 2.

FACS suppressor values. Values for each strain under either autotrophic and mixotrophic conditions. Ploidy values were extrapolated based on the genome size of *C.crescentus*.

	Autotrophic		Mixotrophic	
Strain	Calculated Ploidy	Emission	Calculated ploidy	Emission
Wld type	4.91	1711.00	9.58	3342.00
Δ16.6	5.27	1839.00	10.33	3603.00
L9P	5.62	1959.00	10.72	3738.00
E25K	4.80	1674.00	11.22	3914.00
L9P-A6	4.99	1741.00	11.14	3886.00
L9P-A4	5.47	1907.00	10.80	3768.00
L9P-C3	4.91	1713.00	10.74	3747.00
L9P-1	5.63	1963.00	9.22	3217.00
L9P-2	4.95	1728.00	10.78	3761.00
L9P-5	4.60	1604.00	10.93	3811.00
L9P-7	4.50	1568.00	10.30	3592.00
L9P-11	4.93	1721.00	9.84	3433.00
L9P-12	5.41	1887.00	10.79	3763.00
L9P-14	4.44	1547.00	10.52	3669.00
L9P-15	4.82	1680.00	9.13	3185.00
L9P-16	5.44	1897.00	11.03	3849.00
L9P-17	4.70	1640.00	9.26	3230.00
L9P-19	5.22	1821.00	10.33	3604.00
L9P-22	4.44	1547.00	10.59	3693.00
L9P-24	4.68	1632.00	11.07	3862.00
L9P-33	4.55	1587.00	9.10	3175.00
E25K-2	4.94	1723.00	11.43	3987.00
E25K-3	4.72	1647.00	9.62	3357.00
E25K-4	5.32	1855.00	10.89	3798.00
E25K-6	4.94	1724.00	9.69	3380.00
E25K-8	5.38	1878.00	10.80	3767.00
E25K-9	4.60	1603.00	10.48	3657.00
E25K-11	5.57	1944.00	9.42	3285.00
E25K-17	5.40	1885.00	9.91	3457.00
E25K-18	4.60	1603.00	10.05	3504.00
E25K-19	5.66	1974.00	9.22	3217.00
	Average= 5.01		Average=10.28	

Discussion

From the analyzed strains, it appears that the predominant pathway of evolution towards thermotolerance resulting in strains becoming sHSP independent. Due to the unknown nature of the causative mutations, it is difficult to speculate on the mechanisms involved in bypassing the requirement for sHSPs. One possibility is the enhancement of the other components the protein quality control network. A second explanation for these data is that a mutation or mutations occurred in an essential target protein of the sHSP that provides that protein with increased thermal stability. If this is the case, it would suggest a substrate hierarchy; that is, certain substrates require protection more than others in order for cells to survive. In regard to the four mutants displaying sHSP dependence, these data suggest a possible interactor/co-chaperone that is helping to directly recover/enhance sHSP function.

Regardless of whether or not they are sHSP dependent, all mutant strains appear to display similar growth phenotypes. Under optimal temperatures, all strains grew at the same rate as their respective parent strains and wild type. Under heat stress conditions, phenotypic differences were observed. Note that previous data show that suppressor mutants and wild type show no substantial loss of viability under acute stress. Under chronic stress (continuous 42°C) however, all parent strains show diminished growth, including the wild type strains. In comparison, most mutants, with the exception of L9P-22, show continued linear growth that is approximately 50% lower than their growth at permissive temperatures. Further investigation of L9P-22 would be of interest due to its ability to handle acute stress but not chronic stress.

The last characterization performed was a ploidy analysis using FACS. Data suggest that there is no substantial difference between the ploidy of each suppressor strain or the parental strain. One interesting observation was that cells grown in the presence of glucose (mixotrophic growth) had double the amount of DNA that autotrophically grown cells had. Additional data (Appendix Figure 4) show that the doubling time of *Synechocystis* is almost tripled in the presence of glucose. This observation, in line with the increased growth rates, suggest a differential proteome is being expressed in the presence of glucose and possibly other exogenous carbon sources. This differential proteome may contain targets of the sHSP that were not detected under the original substrate screen performed on mixotrophically grown cells.

CHAPTER 5

COMPARATIVE GENOMICS AS A TOOL TO UNDERSTANDING THERMOTOLERANCE

Introduction

Inevitably, the goal of this project is to understand the evolution of thermotolerance and how sHSPs are required or not in the resulting thermotolerance mechanism. Unlike genetic screens in other organisms, like *Arabidopsis thaliana*, one cannot easily map the general location of mutations with systematic crosses. Luckily, due to the cost of next generation sequencing, the availability of inexpensive bioinformatics tools, and the extensive genomic work previously done on *Synechocystis*, I approached identifying potential causative mutations through a large scale comparative genomics study.

Results

Currently, the complete sequencing of six strains has performed and another 24 strains were prepared for sequencing. These sequenced strains are; Wild type, Δ16.6, L9P, A6, A4, and C3. High quality sequencing results were obtained and analyzed using a series of bioinformatics tools. Average read length between each strain was 248 base. Mapping performed by BWA-mem (Heng et al 2013) yielded >90% properly paired reads and over 99% total genome coverage (see Figure 9B). Average depth varied between the chromosome and mega plasmids within each sample but was comparable between strains. Average chromosomal coverage for all stains was in the range of 240-280x and coverage for all mega plasmids was in the range of 40-75x (see Figure 9A). Alignments were then analyzed using Varscan (Li et al 2009) to identify single nucleotide,

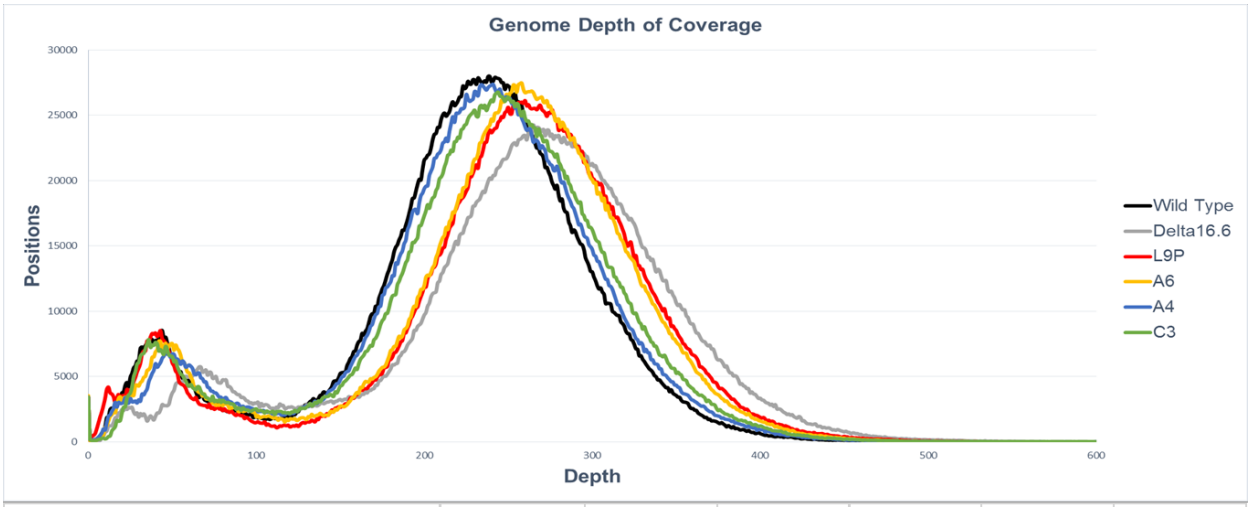
multi nucleotide, and insertion/deletion polymorphisms. A variation frequency of 33.3% (# Differences called/Reference called) was used to search for variants. This parameter allowed for the discovery of over 100 variants in each strain. Integrated Genome Browser (Nicol et al 2009) was used to look for structural changes apparent in the mapping. As a positive control for structural changes, a cross comparison at the sll1514 (Hsp16.6) locus was performed (Figure 10). Using Varscan, mutations were called in each strain with each strain having the following; wild type has 54 variants, Δ16.6 has 208 variants, L9P has 125 variants, A6 has 148 variants, A4 has 155 variants, and C3 has 150 variants. These single nucleotide variants were found in both coding and non-coding regions. Only unique variants (variants found in variants but not control samples) are reported in table 3.

Additionally the same frequency parameter was used to identify small indels. Wild type has 10 variants, Δ16.6 has 15 variants, L9P has 13 variants, A6 has 13 variants, A4 has 17 variants, and C3 has 19 variants. Annotation of these variants was performed using snpEff (Cingolani et al 2009). Cross comparison of genomes yielded 22 total chromosomal variants (A6=6, A4=9, C3=12) for future investigation.

Figure 9.

NGS coverage histograms. A) Global genome coverage of all six *Synechocystis* strains (plotted with Excel): Black; Wild Type Grey; Δ16.6.6 Red; L9P Gold; A6 Blue; A4 Orange; C3 B) Key statistics regarding the sequencing and mapping of Illumina reads of all strains generated with SAM tools. Row 1: Total coverage of the chromosome Row 2: Total coverage of the all of the mega plasmids Row 3: Average depth of coverage across the entire chromosome of each strain. Row 4: Average depth of coverage across all of the mega plasmids. Row 5: Percent of the mapped reads that were properly paired.

A



B

	Wild Type	Deletion	L9P	A6	A4	C3
1) Chromosomal Coverage	>99.9%	>99.9%	>99.9%	>99.9%	>99.9%	>99.9%
2) Mega Plasmid Coverage	>99.9%	>99.9%	>99.9%	>99.9%	>99.9%	>99.9%
3) Avg. Chromosomal Depth	240.5x	277.5x	267.6x	264.2x	247.2x	251.0x
4) Avg. Mega plasmid Depth	46.6x	71.25x	43.8x	43.78x	54.1x	55.31x
5) Properly Paired	96.66%	96.84%	97.33%	92.47%	97.21%	97.24%

Figure 10.

Integrated genome browser structural identification. A) An image captured from Integrated Genome Browser (Nicol et al, 2009) shows the 6 strains in the following order (top to bottom); Δ16.6, Wild type, L9P, A6, A4, and C3. It shows the deletion of the sll1514 locus (Hsp16.6) in the Δ16.6 strain but none of the other lines.

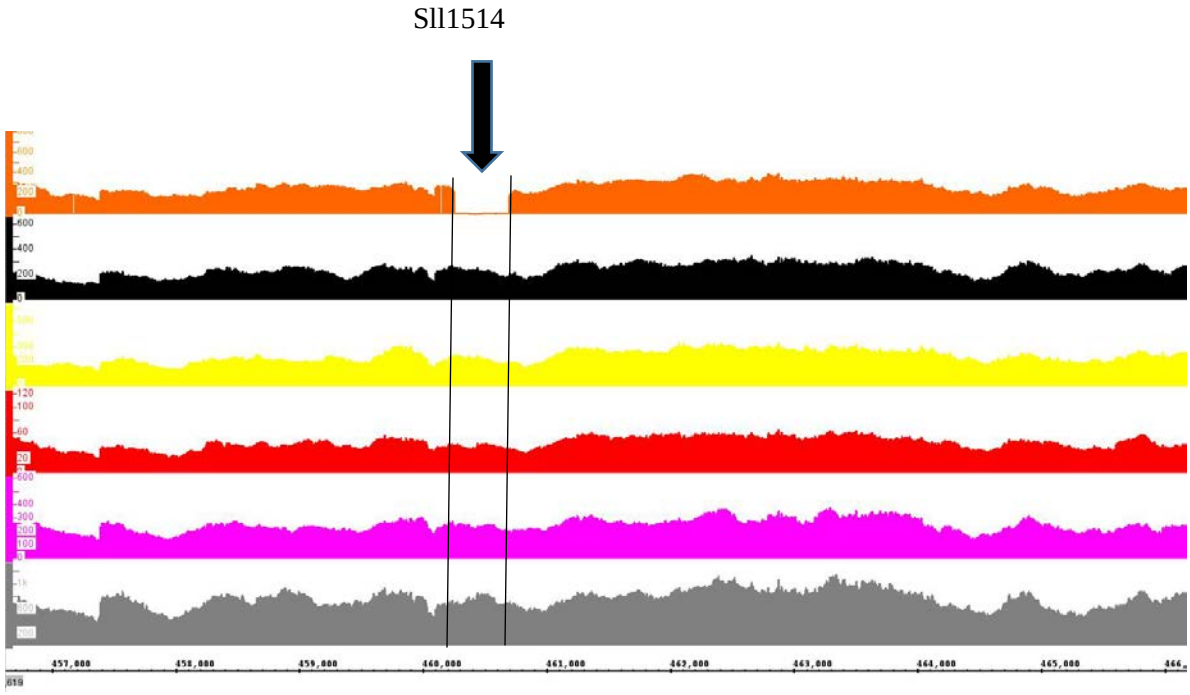


Table 3

Summary of genome variants Table 3 lists the unique chromosomal variants of the three sequenced suppressor strains; A6, A4, C3. Variants were cross compared with one another as well as L9P, wild type, and Δ16.6 variants. Contained in this table is the position of the mutation, the gene name, the gene product, the depth of reads at that position, the frequency of the alternate allele, the type of alteration, and which strain it appears in.

Position	Gene	Gene product	Depth	Frequency	Alteration	A6	A4	C3
2462141	Sll0024	Unknown	176, 174,176	95.16%, 96.34%,	Ins. FS	x	x	x
2082516	slr0442	Unknown	69	30.43%	Synonomous		x	
318397	slr1037	rre10 CheY	212	21.87%	Leu138Phe	x		
460049	slr1597	parA	144	20.83%	Ile284Phe			x
378018	slr1074	Unknown	141	20.57%	Leu113Phe	x		
2953250	sll0590	Unknown	166	16.87%	Pro336Ser			x
1013073	sll1633	ftsZ	152	16.45%	Leu422Pro			x
2782673	slr0907	Unknown	134	16.42%	Val424Gly		x	
3312087	slr1462	Hydrogenase	124	16.13%	Thr142Asn			x
1428593	sll1950	unknown	202	14.85%	Asp113Ala			x
1781251	slr1971	hypothetical	151	14.29%	Thr15Pro			x
2475440	Slr0006	unknown	148	13.91%	Synonomous		x	
1036516	Slr1229	sulfate permease	176	12.65%	Leu308Phe	x	x	x
2353411	slr0364	hypothetical	207, 221	11.59%,11.59%	Thr697Pro		x	x
1952032	slr1670	Unknown	182	10.99%	Asn807Thr		x	
3213072	sll0542	acetyl-coenzyme A synthetase	207	10.63%	Ser159Pro	x		
1586875	sll1425	roline-tRNA ligase	218	10.09%	Val49Gly	x		
2922002	slr0511	putative transposase	253	8.30%	Gly115stop			x
1803789	sll1867	coproporphyrinogen III oxidase	280	7.45%	Del. FS			x
2361148	slr0366	unknown	496	5.04%	Asn231Thr		x	
2362043	slr0366	Unknown	409	2.38%	Synonomous			x
2534398	slr0230	transposase	397	2.04%	Ins. FS		x	

Discussion

Determining the cause of the enhanced thermotolerance of the suppressor strains is essential to unraveling the roles of the sHSP in vivo. Sequencing of the Hsp16.6 gene showed that the causative mutation was not intragenic. In order to determine genomic alterations, Illumina sequencing of the entire genome was performed. These data were than analyzed using several bioinformatics tools.

General statistics pertaining to the mapping and assembly of the sequence data showed high mapping averages, with properly mapped reads being in the >90% range. Coverage of the genome in the strains was extremely high with the coverage of the chromosome roughly 4.92 times higher than the average coverage of the mega plasmids. This may result from either the DNA purification technique or the ploidy of the cells. The spin column method of DNA purification may have a preference for binding larger pieces of DNA. If this is the case, we would expect the chromosome, which is between 30x-70x larger in size than any individual mega plasmid, to be preferentially enriched. The second possibility is that per cell, the number of chromosomes out numbers the number of mega plasmids and that is the reason for the 6-fold higher coverage. FACS analysis suggests a ploidy average of five genomes. Because the FACS data only show total genomic material, it is not known if the chromosomes and mega plasmids are in the same copy number. Fitting the sequencing data to the FACS data, if each cell has a 4.92:1 ratio of chromosomes to plasmids (based on coverage), the chromosome fluorescence (4.92n*3.65Mbp) in addition to the mega plasmids (1n*0.3Mbp) would add up to 18.258Mbp while FACS data (5.01n*3.95Mbp) equaling 19.79Mbp. These values appear to coincide enough to allow for further bioinformatic analysis. However further confirmation is required.

Using the genome browser, I was able to see these two known structural deviations; the deletion of the

Hsp16.6 gene in the Δ16.6 strain and the ClpB1 gene in all strains. No other large variations were detected, however more bioinformatics must be performed to confirm this. Though no large alterations were apparent, I did investigate further the possibility of small alterations, single nucleotide variations, and multi-nucleotide variations.

However, due to the ploidy of the *Synechocystis* genome, it has been difficult to confirm genomic alterations in the suppressor strains compared to the parental strains. Because of the polyploid nature of these organisms, several assumptions must be made in order to identify of the causative mutations. The first assumption is that the mutation or mutations conferring thermotolerance in these suppressor mutants must be present on at least one of the chromosomes in each cell. If this assumption is false and only a subset of cells have this mutation, I suggest there would be disparities in the times of the growth curves. Strains with a mixed population (thermotolerant and not thermotolerant) would have slower starts than populations with a distributed allele as half of the population would die during the chronic stress while the other would replicate. This would appear as lag phase in which the O.D. would not change for several hours before beginning to increase. However, because the curves of all suppressors all grow at similar initial rates and no lag was observed, it is unlikely that the mutation are not distributed throughout the population. A second assumption is that the causative mutations are not under negative selection. This assumption is based off the observation that continuously grown cultures do not appear to revert back to non-thermotolerant states. If the causative mutation were found to be deleterious, it would be segregated out of a population that is not under increased thermal selection and the population would become heat sensitive, which was not observed.

Additional statistical analysis in the form of reversion plating (strains are plated at a known number of colony forming units and following heat stress, the number of colonies that lost thermotolerance are

compared with cells grown at permissive temperatures. Based on these assumptions, polymorphism analysis was performed. Because the bioinformatics data suggests a ploidy between 4 and6 in the autotrophically grown strains (which are thermotolerant) and FACS data suggests N=5, a ploidy of 6 was chosen in our analysis. N=6 was chosen in order to ensure all chromosomes in a theoretical cell are covered, whereas lower ploidy assumptions would begin aggregating variants and potentially lead to missed identities. A minimal read depth of 60x was set in order to have a theoretical representation of each possible chromosome in a theoretical cell at a 10x coverage.

Cross comparison of the suppressor strains with their parent strain, L9P, as well as the wild type data yielded several interesting observations as well as positive results. In all suppressors as well as the L9P data, the missense mutation in sll1514, Hsp16.6, was detected (100% in all strains). This positive result suggests that we are in fact detecting variants properly. Another positive result is the detection of a 7 base pair known insertion (occurs when transformed with pNaive) that is present down stream of sll1514, Hsp16.6 (between sll1514 and the aaDa gene). With two positive results, we can be relatively confident in the precision of our variant detection. For the continuation of this analysis, only chromosomal alterations will be focused on. As shown in figure 9, the coverage of the plasmids is much lower than that of the chromosomes. Because of the high minimum coverage that was used to analyze the chromosomal variants, a significant number of alterations in the lower coverage plasmids may have been ignored due coverage limits. Each plasmid should be analyzed individually based of their coverages in order to attain data.

Ignoring the introduction of variants from the plasmid analysis, and cross comparing variants of each suppressor line with both L9P and wild type variants, we see a much lower pool of potential targets. A6 appears to contain 6 variants from the controls, A4 contains 9, and C3 contains 12 (Table 3). One

potential mutation is found solely in A6. A missense mutation in slr1037, which encodes for a two-component receptor previously characterized to function in chemotaxis, is found in high frequency in this strain. This strain of *Synechocystis* is not motile so the function of this protein is not known in this strain. It may have a second role in stress relay. This mutation is a candidate for further investigation. In C3, two unique mutations to this strain show promise. The first is found in slr1597 and causes a missense mutation in its gene product, ParA. ParA has previously been shown to be involved in chromosomal partitioning. This mutation may allow ParA to interact more efficiently with the altered sHSP. The frequency of this mutation is similar to the CheY mutation in A6, at 20.83%. The second mutation in this line that compliments the ParA mutation is a mutation in sll1633, which encodes FtsZ. This protein is responsible for septum formation during replication. Two proteins involved in DNA partitioning and replication occurring at a high frequency in the same cell line seems like a likely first set of candidates to investigate. Several unknown mutations exist in A4 that may be the cause of its re-acquired thermotolerance. Without knowledge of the gene products, it is difficult to speculate on potential causative agents. The only mutation found in all three lines is located in sll0024, which has an unknown product. Though investigation of this candidate is important, previous data suggest that A6 is independent of the sHSP while A4 and C3 are dependent. This suggests that the cause of thermotolerance in these two groups is different. Additionally, confirmation of these variations is under way through traditional Sanger sequencing. DNA from 24 additional mutant strains is in the queue to be sequenced and these data will be used to narrow the field of candidates genes assuming the same mutations have been selected for in multiple strains. No mutations found in the mutants discovered by Tillich et al were found in the mutants of this screen (Tillich et al 2014)

CHAPTER 6

SMALL HEAT SHOCK PROTEIN: SUBSTRATE INTERACTIONS

Introduction

Considerable work has been performed to understand the roles and mechanisms of sHSPs in vitro. However, it has been much more difficult to study their function in vivo due to a lack of substrates and measurable activity. Previous work by Eman Basha involved the pulldown and identification of native *Synechocystis* proteins that interacted with Hsp16.6 following heat stress. Two of these identified proteins, Fructose Bisphosphate Aldolase II (FBA) and Elongation Factor G (EFG), were selected for further study. Yichen Zhang's work showed in vitro interactions at high temperatures and Hsp16.6 forming soluble complexes with FBA but not EF-G (Yichen Zhang, Ms Thesis). FBA was found forming complexes with Hsp16.6 at 45°C. EF-G was destabilized at temperatures as low as 40°C but only found to form soluble self-aggregates. This might suggest that FBAs solubility during heat stress is Hsp16.6 dependent while EF-Gs is not (Yichen Zhang, Ms Thesis). To examine the behavior of these two substrates in vivo, a western blotting analysis was performed. Soluble/insoluble fractionation of cells following heat stress should yield varied results in the wild type strain versus the Δ16.6 if the role of the sHSP is to keep substrates soluble. However, no variation in substrate solubility was detected in the presence or absence of Hsp16.6 suggesting other potential mechanisms of Hsp16.6 that are not detectable with this experiment

Results

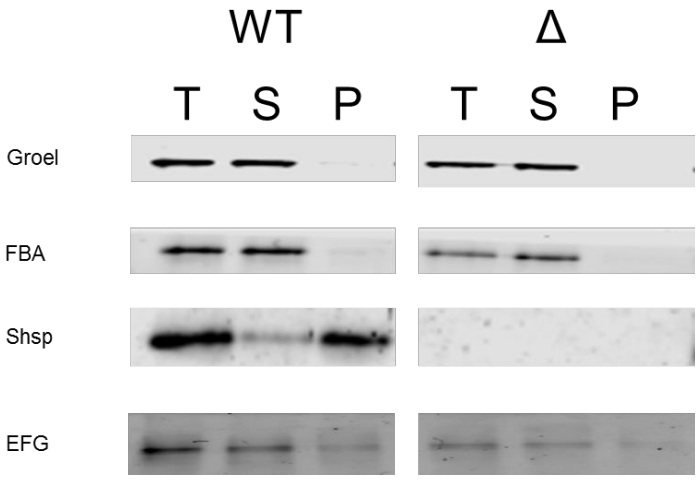
Varied heat stresses were used in an attempt to see differences in the solubility of FBA, EFG and Hsp16.6 in both wild type and Δ16.6 cells. Figure 11a shows a western analysis of the total protein (T), soluble fractions (S) and insoluble fractions (P) of wild type samples and Δ16.6 samples under the same heat stress conditions used by Eman Basha: 42°C for 2 hours followed by 12 hours at 30°C followed by 1 hour at 46°C. GroEL was used as a loading control, as previous data (Nathen Bopp, unpublished) showed little variability in its levels between these two strains under both permissive and heat stress conditions. FBA levels appear relatively constant under both acute heat stress and permissive conditions between the two strains, as do the levels of EF-G. As expected, no Hsp16.6 was detected in the Δ16.6 strain, while significant levels were detected in the wild type strain. Minimal FBA was detected in the insoluble fraction while insoluble EF-G was found in similar concentrations in both strains. Note that almost all Hsp16.6 is found in the insoluble fraction following heat stress. The conditions tested led to a high loss of viability in the Δ16.6 strain in comparison to the wild type strain, shown in Figure 11B.

Figure 11.

Western analysis of sHSP substrates. A) Western analysis of the fractionated lysates of wild type and Δ16.6. Cells were heat stressed for 2 hours at 42°C followed by 12 hours at 30°C and then heat stressed at 46°C for 1 hour prior to fractionation. Total protein and equal proportions of soluble and insoluble cell fractions were separated on 15% SDS PAGE, and used for Western analysis. T: Total protein; S: Soluble protein; P: Pellet, insoluble protein.

B) Viability of wild type and Δ16.6 cells following heat stress, 2 hours at 42°C followed by 12 hours at 30°C and then heat stressed at 46°C for 1 hour. Following stress, cells were plated in a ten-fold dilution series and plated on solid BG-11 agar and allowed to recover and grow under 100μmol/m^2/sec fluorescent light at 30°C for 5 days. Corresponding unstressed cells were not examined this experiment, though previous data show Δ16.6 has no viability issues in comparison to wild type when grown at permissive temperatures (Figure 5).

A



B

Discussion

In vitro data suggests a temperature dependent interaction between Hsp16.6 and its substrates. In order to test this in vivo, several different heat stress conditions were used to stress both wild type and Δ16.6 strains. Under conditions optimized by Eman Basha (Basha et al 2004), no solubility changes in FBA or EFG were detected. More extreme stresses such as 2 hours at 42°C immediately followed by 2 hours at 46°C did not affect the solubility of FBA or EF-G. Heat stress longer than two hours at 46°C dramatically decreases the viability of wild type but did not appear to affect the solubility of either substrate. These data suggests that, in the case of FBA and EF-G, the sHSP does not alter these proteins solubility during heat stress. However, these data does not completely disprove an interaction. These substrates may naturally stay soluble in vivo, even under heat stress, and the sHSP may help protect these substrates from misfolding or losing activity. Additional experiments may be performed to try and see alterations in the behavior or activity of these substrates in the absence of the sHSP in vivo. Native electrophoresis and western analysis have been attempted to examine altered oligomeric structures of FBA in vivo but with no success (not shown). Additional substrates can be probed in future solubility experiments and FBA activity in vivo before, during, and after heat stress could be attempted. Conversely, there may not be a biologically relevant interaction between Hsp16.6 and FBA or EFG and the interactions observed could be artificial. Further investigation is required.

CHAPTER 7

FUTURE DIRECTIONS

The immediate future of this work is the continuation of the bioinformatics analysis and the sequencing of the additional mutants. With more genomes sequenced, comparative analysis has the potential to yield more common factors that may lead to the causes of the increased thermotolerance. Additionally, further western analysis on these strains to could be performed in order to determine the levels of other chaperones under heat stress conditions, as they could be the cause of the enhanced thermotolerance.

The ground work for multiple avenues of further experimentation have been established by this project. In addition to questions regarding the suppressor mutants, fundamental questions about the behavior of the proteome under heat stress have been raised. If the sHSPs do not alter the solubility of its associated proteins, what functions is it performing in the cell? To address this question, a mass spectrometry approach has been devised. Because *Synechocystis* cannot fix atmospheric nitrogen, it receives all of its bioavailable nitrogen from the addition of ammonium nitrate to its growth media. Using N^{15} ammonium nitrate allows us to grow heavy cells which could be used to perform comparative proteomics between strains. Using the mass differences of the heavy and light cells, alterations in the soluble and insoluble proteomes of each strain could then be compared before and after heat stress. A second avenue that this project has opened up is the investigation of other chaperones and their roles in the evolution of thermotolerance. It might be possible that partial loss-of-function mutants in other chaperones could be used in a directed evolution screen to further understand the complexity of stress tolerance. Potential chaperones that could be targeted include DnaK, GroeL, and ClpB.

APPENDIX

SUPPLEMENTAL DATA AND FIGURES

Figure A.1

Western analysis of all successfully transformed suppressors with the PHK plasmid. Cells were treated for 2 hours at 42°C before being lysed in 1x SDS running buffer. Note that samples in which Hsp16.6 was detected under permissive temperatures were also in media containing antibiotics. *Accidental mislabeling of L9P-2 parent strain as L9P-2-PHK.

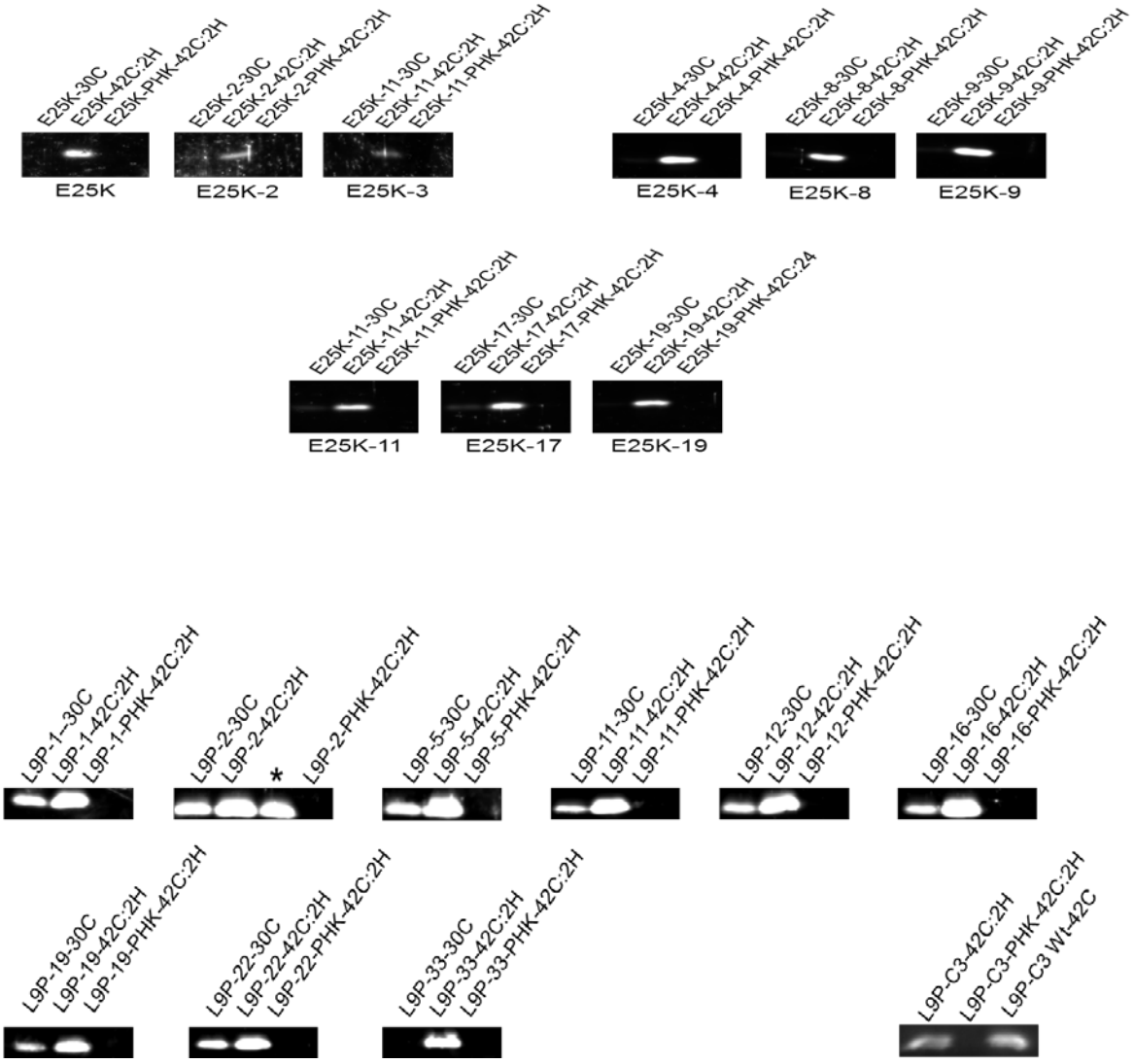


Figure A.2

Wild type cells grown in 5 mL liquid Bg-11 media under continuous 100µmol/m2/s fluorescent light while shaking at 200 RPM. The blue line is the mixotrophically grown Wild Type cells, achieving an O.D at 730nm of >3.0 in about 24 hours. At 24 hours, the autotrophically grown cells were roughly at an O.D of 0.6 at 730nm.

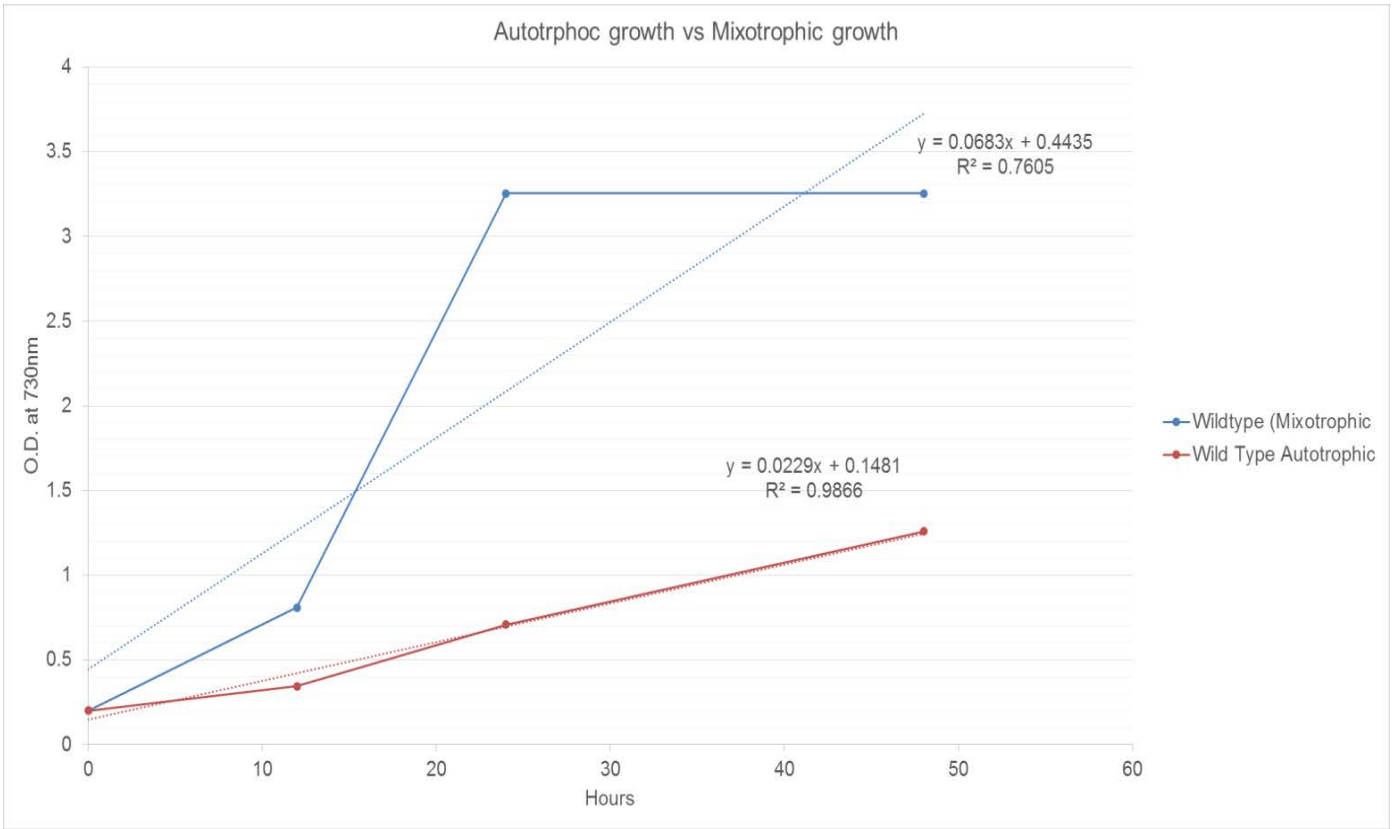


Table A.1

Additional stress conditions tested on wild type, Δ16.6, L9P, A6, A4, and C3 to determine if strains had differential responses. These conditions were derived from previous publications pertaining to other forms of abiotic stress (Blasi et al 2012). Note that none of these additional stresses elicited significantly differential responses between strains. Light stress was performed on Bg-11 solid agar plates. Plates were allowed to recover under normal light levels, 100 µmol/m^2/sec, for 5 days before imaging. Growth in NaCl, sorbitol, cobalt chloride, nickel chloride, and cadmium chloride was performed in liquid culture. 5 mL BG-11 cultures were grown, beginning at an O.D. of 0.2 at 730nm. Values listed are OD 730nm after 72 hours of growth at 30°C under continuous 100µmol/m^2/sec light.

	Wild type	Δ16.6	L9P	A6	A4	C3
Light: 4 hours @ 654µmol/m^2/sec	No Effect	No	No	No	No	No Effect
Light: 6 hours @ 654µmol/m^2/sec	No Effect	No	No	No	No	No Effect
Light: 8 hours @ 654µmol/m^2/sec	No Effect	No	No	No	No	No Effect
Control Conditions	1.9218	2.0928	2.3435	2.0483	2.17125	2.26175
350 mM NaCl	2.144	1.919	2.196	2.223	2.056	2.09
700 mM NaCl	2.073	2.098	2.2	2.117	1.8	2.063
225 mM Sorbitol	1.6	1.681	1.675	1.55	1.631	1.521
450 mM Sorbitol	1.367	1.41	1.386	1.348	1.321	1.38
8 µmol $CoCl_2$	1.413	1.488	1.637	1.723	1.7085	1.689
27 µmol $NiCl_2$	1.84	1.978	2.035	1.889	1.935	2.036
8 µmol $CdCl_2$	0.45	0.486	0.451	0.453	0.415	0.474

Table A.2

Table of all plasmids, strains, and primers used in this thesis. Column one: Strain/Plasmid Name, Column 2: genes/genotype, column 3: Function/descriptor, and column 4: Antibiotic resistance

	Genes	Function	Resistance
Pnaive-Wild type 16.6	Sll1514	Knock in WT 16.6	Spec, Amp
Pnaive-L9P-16.6	L9P Mutant sll1514	Knock in mutant L9P 16.6	Spec, Amp
Pnaive E25K 16.6	E25K Mutant sll1514	Knock in mutant E25K 16.6	Spec, Amp
Pnaive Δ16.6	Replaces sll1514 with spec gene	knock out sll1514	Spec, Amp
PHK	Replaces sll1514 with Kan gene	knocks out sll1514	Kan, Amp
PNK	Sll1514	Knocks in WT 16.6	Kan, Amp
Strains			
IDENTIFICATION	PARENT STRAIN	Function	
Wild type			Spec, Eryth
Δ16.6	Wild type	Δ16.6- Loss of thermotolerance	Spec, Eryth
L9P	Wild type	Phenocopies deletion strain	Spec, Eryth
E25K	Wild type	Phenocopies deletion strain	Spec, Eryth
Screen 1			
A6	L9P	Suppressor of L9P phenotype	Spec, Eryth
A4	L9P	Suppressor of L9P phenotype	Spec, Eryth
C3	L9P	Suppressor of L9P phenotype	Spec, Eryth
Screen 2			
S-E25K-0001	E25K		
S-E25K-0002	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0003	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0004	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0005	E25K	Not a suppressor	Spec, Eryth
S-E25K-0006	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0007	E25K	Not a suppressor	Spec, Eryth
S-E25K-0008	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0009	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0010	E25K	Not a suppressor	Spec, Eryth
S-E25K-0011	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0012	E25K	Not a suppressor	Spec, Eryth
S-E25K-0013	E25K	Not a suppressor	Spec, Eryth
S-E25K-0014	E25K	Not a suppressor	Spec, Eryth
S-E25K-0015	E25K	Not a suppressor	Spec, Eryth
S-E25K-0016	E25K	Not a suppressor	Spec, Eryth
S-E25K-0017	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0018	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0019	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0020	E25K	Not a suppressor	Spec, Eryth
S-E25K-0021	E25K	Not a suppressor	Spec, Eryth
S-E25K-0022	E25K	Suppressor of E25K phenotype	Spec, Eryth
S-E25K-0023	E25K	Not a suppressor	Spec, Eryth

S-L9P-0001	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0002	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0003	L9P	Not a suppressor	Spec, Eryth
S-L9P-0004	L9P	Not a suppressor	Spec, Eryth
S-L9P-0005	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0006	L9P	Not a suppressor	Spec, Eryth
S-L9P-0007	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0008	L9P	Not a suppressor	Spec, Eryth
S-L9P-0009	L9P	Not a suppressor	Spec, Eryth
S-L9P-0010	L9P	Not a suppressor	Spec, Eryth
S-L9P-0011	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0012	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0013	L9P	Suppressor of L9P (accidentaly lost)	Spec, Eryth
S-L9P-0014	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0015	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0016	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0017	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0018	L9P	Not a suppressor	Spec, Eryth
S-L9P-0019	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0020	L9P	Not a suppressor	Spec, Eryth
S-L9P-0021	L9P	Not a suppressor	Spec, Eryth
S-L9P-0022	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0023	L9P	Not a suppressor	Spec, Eryth
S-L9P-0024	L9P	Not a suppressor	Spec, Eryth
S-L9P-0025	L9P	Not a suppressor	Spec, Eryth
S-L9P-0026	L9P	Not a suppressor	Spec, Eryth
S-L9P-0027	L9P	Not a suppressor	Spec, Eryth
S-L9P-0028	L9P	Not a suppressor	Spec, Eryth
S-L9P-0029	L9P	Not a suppressor	Spec, Eryth
S-L9P-0030	L9P	Not a suppressor	Spec, Eryth
S-L9P-0031	L9P	Not a suppressor	Spec, Eryth
S-L9P-0032	L9P	Suppressor of L9P	Spec, Eryth
S-L9P-0033	L9P	Suppressor of L9P	Spec, Eryth
L9P-001-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-002-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-005-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-007-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0011-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0012-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0016-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0017-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0019-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0022-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-0024-PNK	L9P	Wild type 16.6 in suppressor background	Kan, Eryth
L9P-001-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-002-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-005-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-007-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0011-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0012-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0016-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0017-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0019-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0022-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth
L9P-0024-PHK	L9P	Sll1514 knockout in suppressor background	Kan, Eryth

E25K-002-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-003-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-004-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-006-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-008-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-009-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-011-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-017-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-018-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-019-PNK	E25K	Wild type 16.6 in suppressor background	Kan, Eryth
E25K-002-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-003-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-004-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-006-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-008-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-009-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-011-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-017-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-018-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
E25K-019-PHK	E25K	Sll1514 knockout in suppressor background	Kan, Eryth
PRIMERS			
SLL1514 FOR	5'GAGCTCGGCAAAAATTAATTATCGA	TM=56°C	
SLL1514 REV	5'CCCACCTTCTTTCCTGAAAACC	TM=56°C	

REFERENCES

Anderson, S L; McIntosh, L (1991). “Light-activated heterotrophic growth of the cyanobacterium Synechocystis sp. strain PCC 6803: a blue-light-requiring process.” J. Bacteriol. vol. 173 (9) 2761-2767

Basha, E., G. J. Lee, L. A. Breci, A. C. Hausrath, N. R. Buan, K. C. Giese and E. Vierling (2004). "The identity of proteins associated with a small heat shock protein during heat stress in vivo indicates that these chaperones protect a wide range of cellular functions." J Biol Chem 279(9): 7566-7575.

Basha, E., H. O'Neill and E. Vierling (2012). "Small heat shock proteins and alpha- crystallins: dynamic proteins with flexible functions." Trends Biochem Sci 37(3): 106-117.

Behrendorff, James B Y H; Huang, Weiliang; Gillam, Elizabeth M J (2015) “Directed evolution of cytochrome P450 enzymes for biocatalysis: exploiting the catalytic versatility of enzymes with relaxed substrate specificity”. The Biochemical Journal vol. 467 (1) p. 1-15

Blasi, B., Peca, L., Vass, I., & Kós, P. B. (2012). Characterization of stress responses of heavy metal and metalloid inducible promoters in synechocystis PCC6803. Journal of Microbiology and Biotechnology, 22(2), 166–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22370344

Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A; Galaxy Team. Manipulation of FASTQ data with Galaxy. Bioinformatics. 2010 Jul 15;26(14):1783-5

Caspeta, Luis; Chen, Yun; Ghiaci, Payam; Feizi, Amir; Buskov, Steen et al. (2014) "Altered sterol composition renders yeast thermotolerant" Science vol. 346 (6205) p. 75-78

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. Fly (Austin). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.", 2012 Apr-Jun;6(2):80-92. PMID: 22728672

Gidalevitz, T., E. A. Kikis and R. I. Morimoto (2010). "A cellular perspective on conformational disease: the role of genetic background and proteostasis networks." Curr Opin Struct Biol 20(1): 23-32

Giese, K. C. and E. Vierling (2002). "Changes in oligomerization are essential for the chaperone activity of a small heat shock protein in vivo and in vitro." J Biol Chem 277(48): 46310-46318

Giese, K. C. and E. Vierling (2004). "Mutants in a small heat shock protein that affect the oligomeric state. Analysis and allele-specific suppression." J Biol Chem 279(31): 32674-32683.

Giese, Kim C.; Basha, Eman; Catague, Belmund Y.; Vierling, Elizabeth (2005). Evidence for an essential function of the N terminus of a small heat shock protein in vivo, independent of in vitro chaperone activity PNAS vol. 102 (52) p. 18896-18901

Griese, Marco; Lange, Christian; Soppa, Jörg (2011) “Ploidy in cyanobacteria.” FEMS microbiology letters vol. 323 (2) p. 124-31

Haslbeck, M. (2002). "sHsps and their role in the chaperone network." Cell Mol LifeSci 59(10): 1649-1657

Herzenberg, Leonard A.; Parks, David; Sahaf, Bita; Perez, Omar; Roederer, Marioet al. The History and Future of the Fluorescence Activated Cell Sorter and Flow Cytometry: A View from Stanford (2002) Clin. Chem. vol. 48 (10) p. 1819-1827

Heng Li (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima,

M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda and S. Tabata (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 (supplement)." DNA Res 3(3): 185-209.

Koboldt, D., Zhang, Q., Larson, D., Shen, D., McLellan, M., Lin, L., Miller, C., Mardis, E., Ding, L., and Wilson, R. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing Genome Research DOI: 10.1101/gr.129684.111

Nicol, John W.; Helt, Gregg A.; Blanchard, Steven G., Jr.; Raja, Archana; Loraine, Ann E. (2009) “The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets” Bioinformatics vol. 25 (20) p. 2730-2731

Lee, S., H. A. Owen, D. J. Prochaska and S. R. Barnum (2000). "HSP16.6 is involved in the development of thermotolerance and thylakoid stability in the unicellular cyanobacterium, Synechocystis sp. PCC 6803." Curr Microbiol 40(4): 283-287.

Lee, S., D. J. Prochaska, F. Fang and S. R. Barnum (1998). "A 16.6-kilodalton protein in the Cyanobacterium synechocystis sp. PCC 6803 plays a role in the heat shock response." Curr Microbiol 37(6): 403-407.

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. In Bioinformatics, 26 (5), pp. 589–595. doi:10.1093/bioinformatics/btp698

Li, H. and Handsaker, B. and Wysoker, A. and Fennell, T. and Ruan, J. and Homer, N. and Marth, G. and Abecasis, G. and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. In Bioinformatics, 25 (16), pp. 2078–2079. doi:10.1093/bioinformatics/btp352

Li, H. (2011). Improving SNP discovery by base alignment quality. In Bioinformatics, 27 (8), pp. 1157–1158. doi:10.1093/bioinformatics/btr076

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. In Bioinformatics, 27 (21), pp. 2987–2993. doi:10.1093/bioinformatics/btr509

Nakamura, Y., T. Kaneko, M. Hirosawa, N. Miyajima and S. Tabata (1998). "CyanoBase, a www database containing the complete nucleotide sequence of the genome of Synechocystis sp. strain PCC6803." Nucleic Acids Res 26(1): 63-67.

Paul Julian Kersey, James E. Allen, Mikkel Christensen, Paul Davis, Lee J.Falin, Christoph Grabmueller, Daniel Seth Toney Hughes, Jay Humphrey, Arnaud Kerhornou, Julia Khobova, Nicholas Langridge, Mark D. McDowall, Uma Maheswari, Gareth Maslen, Michael Nuhn, Chuang Kee Ong, Michael Paulini, Helder Pedro, Iliana Toneva, Mary Ann Tuli, Brandon Walts, Gareth Williams, Derek Wilson, Ken Youens-Clark, Marcela K. Monaco, Joshua Stein, Xuehong Wei, Doreen Ware, Daniel M. Bolser, Kevin Lee Howe, Eugene Kulesha, Daniel Lawson and

Daniel Michael Staines Ensembl Genomes 2013 “Scaling up access to genome-wide data” *Nucleic acids research* 2014, 42 (D1): D546-D552

Shen R, Fan JB, Campbell D, Chang W, Chen J, Doucet D, Yeakley J, Bibikova M, Wickham Garcia E, McBride C, Steemers F, Garcia F, Kermani BG, Gunderson K, Oliphant A. High-throughput SNP genotyping on universal bead arrays. Source Illumina Inc., 9885 Towne Centre Drive, San Diego, CA 92121, USA. rshen@illumina.com

Schopf, J. W. (2012) "The fossil record of cyanobacteria. In Ecology of Cyanobacteria II (pp. 15-36). Springer Netherlands" In: Brian A. Whitton (Ed) Ecology of Cyanobacteria II: Their Diversity in Space and Time, Springer Science & Business Media. ISBN 9789400738553.

Simon, W. J., J. J. Hall, I. Suzuki, N. Murata and A. R. Slabas (2002). "Proteomic study of the soluble proteins from the unicellular cyanobacterium Synechocystis sp. PCC6803 using automated matrix-assisted laser desorption/ionization-time of flight peptide mass fingerprinting." Proteomics 2(12): 1735-1742.

Slabas, A. R., I. Suzuki, N. Murata, W. J. Simon and J. J. Hall (2006). "Proteomic analysis of the heat shock response in Synechocystis PCC6803 and a thermally tolerant knockout strain lacking the histidine kinase 34 gene." Proteomics 6(3): 845-864.

Storchova, Zuzana; Breneman, Amanda; Cande, Jessica; Dunn, Joshua; Burbank, Kendra et al. Genome-wide genetic analysis of polyploidy in yeast Nature (2006)

Tillich, Ulrich; Wolter, Nick; Franke, Philipp; Duhring, Ulf; Frohme, Marcus (2014) “Screening and genetic characterization of thermo-tolerant Synechocystis sp. PCC6803 strains created by adaptive evolution”.BMC Biotechnology vol. 14 (1) p. 66

Torok, Z., P. Goloubinoff, I. Horvath, N. M. Tsvetkova, A. Glatz, G. Balogh, V. Varvasovszki, D. A. Los, E. Vierling, J. H. Crowe and L. Vigh (2001). "Synechocystis HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding." Proc Natl Acad Sci U S A 98(6): 3098-3103

van Montfort, R., E. Basha, K.L. Friedrich, C. Slingsby, E. Vierling . Structure and assembly of a eukaryotic small heat shock protein. Nature Struct. Biol. 8:1025-1030 (2001)

Wang, Jue D; Herman, Christophe; Tipton, Kimberly A; Gross, Carol A; Weissman, Jonathan S (2002). “Directed evolution of substrate-optimized GroEL/S chaperonins.” Cell vol. 111 (7) p. 1027-39