The heat shock response of *Synechocystis* sp. PCC 6803 analysed by transcriptomics and proteomics

Iwane Suzuki1,*, William J. Simon2 and Antoni R. Slabas2

1 Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi Ibaraki-ken 305-8572, Japan
2 Department of Biochemical and Biomedical Sciences, University of Durham, Durham DH1 3LE, UK

Received 20 August 2006; Accepted 6 February 2006

Abstract

When cells of the cyanobacterium *Synechocystis* sp. PCC 6803 are exposed to high temperature they perceive changes in the growth conditions and regulate the expression of genes and synthesize heat-inducible proteins as a response to the heat stress. DNA microarray analysis revealed that genes for chaperonins and proteases, such as *groESL1*, *groEL2*, *htpG*, *hspA*, and *clpB1* were transiently induced after incubation of the cells at 44 °C for 20 min. Quantitative two-dimensional gel electrophoresis revealed that the levels of these chaperonins and proteases were elevated after incubation of cells at 44 °C for 60 min. These findings indicated that levels of the mRNAs and proteins of chaperonins were well correlated in the cells of *Synechocystis*. However, the level of elongation factors are mainly regulated at the protein level. These results indicated that acclimation to the heat-shock conditions might be governed by transcriptional and translational regulation in *Synechocystis*.

Key words: Chaperonin, cyanobacteria, heat stress, proteomics, *Synechocystis*, transcriptomics, translation.

Introduction

The cyanobacterium *Synechocystis* sp. PCC 6803 is a photosynthetic prokaryote which is believed to be a progenitor of higher plant chloroplasts (Bryant, 1986). This organism has a relatively simple genome containing 3264 genes on the chromosome and its genome has been completely sequenced (Kaneko et al., 1996). From a biochemical point of view it contains few subcompartments: there are two membrane systems, the thylakoid membrane and the plasma membrane, and the cytoplasm which contains soluble proteins. The availability of a DNA microarray containing 3079 of the 3264 genes of *Synechocystis*, plus the relative simplicity of its proteome, make it an attractive model to develop for studies on perception, response, and adaptation to alterations in environmental conditions. This is particularly so as gene disruption, via homologous recombination, is highly efficient in *Synecho-cystis* allowing one to generate clean knockouts which can be used to elucidate the potential importance of specific genes in cellular metabolism.

Numerous studies have been performed on the response of the genome of *Synechocystis* to environmental stress conditions using DNA microarray technology, and these have identified both histidine kinases and response regulators involved in signal perception and transduction (Hihara et al., 2001, 2003; Suzuki et al., 2001, 2004; Huang et al., 2002a; Kanesaki et al., 2002; Yamaguchi et al., 2002; Inaba et al., 2003; Marin et al., 2003; Singh et al., 2003; Paithoonrangsarid et al., 2004; Wang et al., 2004; Hubschmann et al., 2005; Shoumskaya et al., 2005). However, few proteomic studies, utilizing the recent developments in mass spectrometry, have been carried out with *Synechocystis* (Huang et al., 2002b; Kashino et al., 2002; Simon et al., 2002, Slabas et al., 2006). In the area of proteomics, developments using differential fluorescent dye labelling (2D-DIGE, Tonge et al., 2001) have allowed quantitative measurements of the changes in the entire soluble proteome using similar labelling technology to that employed in DNA microarray experiments, however, there are few cases where this has been applied to investigations in cyanobacteria.

Living organisms are frequently exposed to changes in environmental conditions. These changes can be quite...
dramatic and data from a study in the USA has shown that diurnal changes between 25–50 °C occur daily on the Salt Plains National Wildlife Refuge. (http://www.okstate.edu/artsci/lexen/LEXEn.htm). If an organism is to survive such changes then it has to be able to perceive the changes, protect the existing metabolic machinery, and adapt its metabolism to suit the new environmental conditions. Heat shock is one of the most extensively studied systems with regard to environmental responses and its study has led to the discovery of ‘heat shock proteins’, which have subsequently been shown to be chaperons and associated proteases important in protein folding and degradation and so playing a pivotal role in cellular adaptation.

In this study the heat shock response in Synechocystis has been investigated using both a transcriptomic and a proteomic approach. The results demonstrate that changes at the transcriptome level can be rapid and transient. Whilst there is a correlation between changes in the level of the major chaperonins and proteases observed using proteomics and transcriptomics following heat shock, there are areas where this correlation does not exist. More extensive kinetic studies aimed at defining levels of transcripts, stability of mRNA, translational rates, and stability/turnover of proteins will be required to understand the dynamics of cellular responses to environmental stress. Comments are made on the requirements in this area and potential experimental ways forward.

Materials and methods

**Strains and culture conditions**

*Synechocystis* sp. PCC 6803, a glucose tolerant strain, was kindly provided by Dr JGK Williams (Du Pont de Nemours, Wilmington, DE) and served as the wild type. Cells of *Synechocystis* were grown at 34 °C in BG-11 medium (Stanier et al., 1971) with aeration with CO₂-enriched air (1%, v/v) under continuous illumination at 70 μmol photons m⁻² s⁻¹ (Wada and Murata, 1989).

**DNA microarray analysis**

Three independent biological replicates were performed for each of these experiments. A 50 ml aliquot of culture was mixed with an equal volume of an ice-cold mixture of 5% phenol in ethanol (w/v) to stop cellular metabolism. Cells were harvested from the mixture by centrifugation at 4000 g, for 5 min at 4 °C, and then frozen immediately and stored at −80 °C. Total RNA was extracted with hot phenol as described previously (Suzuki et al., 1996), subsequently been shown to be chaperons and associated proteases important in protein folding and degradation and so playing a pivotal role in cellular adaptation.

**Protein gel electrophoresis, labelling and imaging**

**Sample preparation**: At least four independent biological replicates were performed for each experiment. For DIGE experiments sample clean-up was performed using GE Healthcare’s, 2-D Clean-Up kit and resolubilized in lysis buffer containing 25 mM TRIS-HCl pH 9.5, to give an approximate final protein concentration of 5 mg ml⁻¹. The final pH of the samples was checked using pH indicator strips and adjusted to pH 8.5, if required by addition of 1 M NaOH.

**Cyanine dye labelling**: Protein samples were labelled using CyDye DIGE Fluor minimal dyes (GE Healthcare, Amersham, UK) according to the Ettan™ 2D-DIGE protocol. 1 mM (1 nmol μl⁻¹) stock CyDye solutions in anhydrous N,N-dimethylformamide were made, from which 0.04 mM working dye solutions (in DMP) were prepared. This working dye solution was added to the prepared protein samples at a final ratio of 50 μg protein: 400 pmol dye. Samples were vortexed and incubated on ice for 30 min in the dark. After this time labelling was stopped by adding 1 μl of 10 mM Lysine. Samples were vortexed and left on ice for a further 10 min in the dark. The final protein concentration was 2.5 μg ml⁻¹ for each cyanine-labelled sample.

**Preparative electrophoresis**: Equal 12.5 μg quantities of control (Cy3-labelled) and treated (Cy5-labelled) samples were mixed by vortexing with the same quantity (12.5 μg) of pooled internal standard (Cy2-labelled). Lysis buffer was added to each mix to give a final volume of 70 μl, to which DTT to 1% and Pharmalytes (pH 4–7) to 2% were added. These samples were subjected to 2-DE as described previously (Slabas et al., 2006). Samples were run which corresponded to a minimum of four independent biological replicates.

**Imaging**: Immediately following electrophoresis cyanine dye-labelled protein gels were imaged directly within the glass cassettes using a Typhoon variable mode imager (GE Healthcare, Amersham, UK) according to manufacturer’s recommendations at 100 nm resolution. The resultant gel images were processed and analysed using DeCyder Biological Variation Analysis (BVA) software version 5.0 (Amersham Biosciences). Only spots that satisfied a 95% confidence limit and were present on all gels analysed were considered to have altered.

**Protein spot picking, protease digestion, and mass spectrometric analyses**

**Preparative electrophoresis**: 2D gels with a high protein load (400 μg) for spot picking were visualized with MS compatible silver staining, as described by Shevchenko et al. (1996). The spots were recognized by superimposition on the fluorescent images.

**Protein spot picking and trypsin digestion**: Selected spots were excised by manual picking using One Touch Plus Spot Pickers (The Gel Company, San Francisco, California) into separate wells of 96-well microtitre plates (Genomic Solutions, Cambridge UK). Microtitre plates were then transferred to a ProGest workstation (Genomic Solutions, Cambridge UK) and the samples were destained and digested with trypsin according to the standard ProGest long trypsin digestion protocol. Following digestion, peptide extracts were lyophilized in a vacuum centrifuge and resuspended in 10 μl of 0.1% formic acid.

**Protein identification by MALDI-TOF MS**: MALDI-TOF peptide mass fingerprinting (MALDI-PMF) was performed using an Applied Biosystems (Warrington UK) Voyager DE-STR Biospectrometry
workstation (0.5 µl of each concentrated digest was spotted onto the surface of a thin film of 1:1 4-cyan-4-hydroxycinnamic acid matrix, 1% nitrocellulose previously dried onto the surface of a MALDI target plate. Once spotted the samples were allowed to air dry before being washed with 5 µl of cold 0.1% TFA solution. For each spot accumulated spectra were acquired from 480 laser shots using system parameters optimized for the mass range 800–3500 amu. Automated peak detection, noise reduction, and peak de-isotoping was carried out on the spectra using the Applied Biosystems Data Explorer software. De-isotoped spectra were internally calibrated using the trypsin autolysis peaks at 842.5 m/z and 2211.11 m/z present in the spectra. The calibrated peak list of peptide masses for each sample was used in an unrestricted MASCOT (www.matrixscience.com) search of the NCBInr database allowing for a single mis cleavage, the presence of oxidized methionines and carbamidomethyl cysteines, and using a peptide mass accuracy of 50 ppm.

Positive protein identifications using the above criteria were accepted when the MASCOT MOWSE probability score was significantly higher than the probability cut-off for a random event, a minimum of eight peptides matched to the identified protein, and the minimum sequence coverage was greater than 18%. Further confidence in the protein assignment being correct was that the top hit reported from an unrestricted search was from Synechocystis and that, although not reported here, experience from other studies on this organism is that whenever identification using MALDI-PMF with the above criteria was verified using MS-MS sequencing, the identifications were confirmed.

Results and discussion

Experimental design considerations

These experiments were concerned with identifying changes at both the mRNA and protein levels. The doubling rate of Synechocystis is approximately 8 h, and, for experimental reasons, time points were selected which would be meaningful bearing in mind differences in the speed of responses at the protein and mRNA levels and the expense of each experiment. Previous experiments with Synechocystis have demonstrated that responses at the mRNA level occur rapidly and can be seen in the 10–60 min time frame. Accordingly, two time points were chosen for the transcriptome analyses: 20 min and 60 min. In order to select a time point for the proteome analysis, a (35S)-methionine labelling experiment (data not shown) was performed to determine the time frame during which new protein synthesis occurred following heat shock. Major differences could be seen at 30 min and the response levelled out at 120 min, accordingly a single time point of 60 min was chosen for the proteomic analyses.

Alterations in the levels of mRNA and proteins of chaperonins and proteases in the cells following heat shock indicate a strong component of transcriptional control

Following heat shock brought about by the transfer of cells from 34–44 °C samples were removed for RNA and protein isolation. Temperature measurements showed that it took 5 min to go from 34–44 °C. Twenty minute and 60 min samples had their RNA isolated for genome-wide transcriptome analysis. Levels of expression of heat-shock genes, such as groEL1, groEL2, hspA, htpG, dnaK2, and clpB1 genes, which encode chaperonins, rose more than 8-fold during the first 20 min at 44 °C, some of the transcripts such as groES, groEL2, hspA, and clpB1 were elevated over 40-fold (Table 1). The levels of mRNA of the clpC and hhoA genes which encode proteases were also elevated following heat shock, but to a lesser extent (Table 1). Expression of most of the heat-shock genes except groEL1 and clpC were decreased at the 60 min time point, compared with that at 20 min, following incubation at 44 °C. Changes in levels of the soluble proteins were examined by 2D differential gel electrophoresis (2D-DIGE; GE Healthcare Amersham). Spots of the protein were visualized by silver staining and the proteins identified by matrix-associated laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry peptide fingerprinting after in-gel digestion of the proteins with trypsin. In the 2D-DIGE analyses, typically several spots were identified as corresponding to the same protein. Levels of GroES, GroEL1, GroEL2, HspA, HtpG, DnaK2, ClpB1, ClpC, and

Table 1. Induction of levels of mRNAs and proteins of chaperonins and proteases in the cells of Synechocystis under the heat-shock conditions

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene name</th>
<th>Product Description</th>
<th>Transcriptomic</th>
<th>Proteomic</th>
</tr>
</thead>
<tbody>
<tr>
<td>sr2075</td>
<td>groES</td>
<td>10 kDa Chaperonin, Groes Protein</td>
<td>44.75 15.46 3.24</td>
<td>60 min</td>
</tr>
<tr>
<td>sr2076</td>
<td>groEL1</td>
<td>60 kDa Chaperonin 1</td>
<td>8.70 16.45 1.37–7.17 10 spots</td>
<td>60 min</td>
</tr>
<tr>
<td>sl0416</td>
<td>groEL2</td>
<td>60 kDa Chaperonin 2</td>
<td>42.14 6.46 4.78–4.88 4 spots</td>
<td>60 min</td>
</tr>
<tr>
<td>sl0154</td>
<td>hspA</td>
<td>16.6 kDa Small Heat Shock Protein</td>
<td>68.09 21.65 19.92</td>
<td>60 min</td>
</tr>
<tr>
<td>sl0430</td>
<td>htpG</td>
<td>Heat Shock Protein 90</td>
<td>34.61 3.72 2.56–2.76 4 spots</td>
<td>60 min</td>
</tr>
<tr>
<td>sl0170</td>
<td>dnaK2</td>
<td>Heat Shock Protein 70</td>
<td>22.30 2.06 2.36–2.64 6 spots</td>
<td>60 min</td>
</tr>
<tr>
<td>str1641</td>
<td>clpB</td>
<td>ClpB protein</td>
<td>92.46 1.89 3.31–5.77 5 spots</td>
<td>60 min</td>
</tr>
<tr>
<td>sl0080</td>
<td>clpC</td>
<td>ATP-dependent Clp protease ATPase subunit</td>
<td>1.83 2.49 1.46</td>
<td>60 min</td>
</tr>
<tr>
<td>sl1679</td>
<td>hhoA</td>
<td>Periplasmic protease HhoA</td>
<td>1.86 1.66 1.28</td>
<td>60 min</td>
</tr>
</tbody>
</table>
HhoA increased following incubation of the cells at 44 °C for 60 min (Table 1). Although the ratios of induction of the proteins were not the same as the ratios of induction of the expression of gene which was determined by DNA micro-array, there was a distinct trend toward higher amounts of the proteins corresponding to those genes transcripts which were elevated and remained high in this set. These results clearly indicate that accumulation of these heat-shock proteins was dependent on the induction of expression of the genes under the heat-shock conditions. It may well be that, because these genes are involved with protein folding and the removal of protein aggregates, there is a strong requirement for an elevated level of them to aid both removal of unwanted proteins and to help in the correct folding of proteins which are newly synthesized to adapt metabolism to the new conditions.

**Alteration in the level of components in protein translation caused by heat shock seems to be regulated most at the post-translational level**

The heat-inducible accumulation of chaperonin proteins appear to be correlated with the increase in the levels of transcripts. However, an increase of all the heat-inducible proteins was not correlated with increases in the mRNA levels. For example, the level of mRNA of the tsf gene coding for an elongation factor EF-Ts was transiently increased up to 2.97-fold after incubation at 44 °C for 20 min and returned almost to the initial level after incubation for 60 min in the heat conditions. The level of the EF-Ts protein was also increased to 1.31-fold (Table 2). This profile is very similar to those of chaperonins as shown in Table 1. It is not clear why the expression of the tsf gene is induced under the heat-shock conditions. Interestingly other proteins, which are related to the process of the elongation of translation such as EF-G1, EF-G2, and EF-Tu, also increase 1.32-, 1.68-, and 1.43-fold, respectively, during the heat treatment. However, levels of the mRNA of these genes were not increased (Table 2), suggesting that levels of the elongation factors were regulated at the translational level, except possibly for EF-Ts. Although the mechanism of regulation of translation of the elongation factors is not clear, the elongation of translation might be severely inhibited under the heat conditions and perhaps the retardation of elongation activity might induce the synthesis of the elongation factors via translation itself. In prokaryotes, however, most of the proteins are thought to be regulated at the transcriptional level (Malhotra et al., 1996). These results clearly indicated that translational regulation might also play an important role in the cellular responses to environmental stresses. It is clear that a deeper analysis of the proteomic experiment will be possible using iTRAQ technology (Ross et al., 2004). This will allow proteins to be taken into account which are unsuitable for analysis using 2D gel-based technology, due to isoelectric point and molecular weight considerations, and which cannot to be resolved using current gel-based technologies. Such experiments are currently being conducted. To reach definitive conclusions on regulation at the level of translation, further experiments will be required taking into account the rate of transcription and protein turnover of the level of individual proteins.

**Global changes of genes and proteins following heat shock and the functional classes into which they fall**

Changes identified at the protein and mRNA levels following heat shock and their distribution amongst functional categories in cellular metabolism have been studied. In this study, changes in 90 heat-inducible proteins and 113 heat-inducible genes were identified (Table 3). When viewed in this way there seems to be major differences in the pattern of alteration of cellular metabolism detected by proteomic and transcriptomic methods. It has to be borne in mind that the proteomic analysis was conducted on soluble proteins and approximately 50% of the proteins of *Synechocystis* are membrane bound, so a strict comparison is not valid. The chaperonins and associated proteins are

**Table 2. Induction of levels of mRNAs and proteins of components of translation in the cells of Synechocystis under the heat-shock conditions**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene name</th>
<th>Product</th>
<th>Transcriptomic</th>
<th>Proteomic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 min</td>
<td>60 min</td>
</tr>
<tr>
<td>sul0744</td>
<td><em>infB</em></td>
<td>Translation initiation factor IF-2</td>
<td>1.44</td>
<td>0.70</td>
</tr>
<tr>
<td>sc33441</td>
<td><em>infA</em></td>
<td>Translation initiation factor IF-1</td>
<td>0.58</td>
<td>0.44</td>
</tr>
<tr>
<td>slr0974</td>
<td><em>infC</em></td>
<td>Translation initiation factor IF-3</td>
<td>1.03</td>
<td>0.98</td>
</tr>
<tr>
<td>slr1261</td>
<td><em>tsf</em></td>
<td>Elongation factor EF-Ts</td>
<td>2.97</td>
<td>0.84</td>
</tr>
<tr>
<td>slr1463</td>
<td><em>fusA</em></td>
<td>Elongation factor EF-G1</td>
<td>0.97</td>
<td>0.72</td>
</tr>
<tr>
<td>slr1105</td>
<td><em>fus</em></td>
<td>Elongation factor EF-G</td>
<td>0.80</td>
<td>0.64</td>
</tr>
<tr>
<td>slr11098</td>
<td><em>fusB</em></td>
<td>Elongation factor EF-G2</td>
<td>0.71</td>
<td>1.05</td>
</tr>
<tr>
<td>slr11099</td>
<td><em>infA</em></td>
<td>Elongation factor EF-Tu</td>
<td>0.61</td>
<td>0.84</td>
</tr>
</tbody>
</table>
in the class ‘cellular processes’. The most notable differences between the two data sets are the strong changes seen at the protein level for components involved in amino acid biosynthesis, energy metabolism, and translation which were not seen as affected to the same extent at the mRNA level. If new proteins are to be synthesized in response to stress then clearly amino acids, mRNA translation, and energy supply will be of importance. In the transcriptome data set there is a much larger proportion of genes in the ‘hypothetical and unknown’ category, which probably represents the higher sensitivity of DNA array technology and coverage of the entire genome. There are clearly a number of genes for which function has still to be assigned.

**Perspectives for future experimentation**

Understanding the full extent of cellular changes and the interaction of components is one of the challenges of systems biology. In this study there has been a snapshot of the major changes at the transcriptome level following heat shock. It is fortunate for *Synechocystis* that there is a whole-genome DNA array. The coverage of the proteome has not been complete as only the soluble protein fraction has been studied using 2D-DIGE and a single time point. To obtain greater coverage of the proteome, iTRAQ technology will be used in the future (Ross et al., 2004) which is a gel-free system, based on peptide tagging and identification. This should overcome any difficulties with sampling proteins whose *M*<sub>r</sub>, pI, and hydrophobicity fall outside of the range of proteins which can currently be analysed using gel-based technologies. At a recent HUPO (Human Proteome Organization) meeting in Munich, methodology was described which allowed an investigation into the rates of both transcription and translation of mRNAs and protein stability in yeast (O’Shea et al., 2005). To understand fully how heat stress responses are controlled, it may be that such experimentation has to be performed in *Synechocystis*. An experiment can be envisaged where mRNA stability can be followed by treating cells at specific times with transcriptional inhibitors, such as rifampicin, and using DNA array technology to follow the level of transcripts at a certain point in time. In a similar way the use of translational inhibitors, like chloramphenicol, could be used to track protein turnover. Perhaps one of the more interesting lines of investigation would be to perform knockouts on the genes which have the highest changes and see if they adversely affect the thermal tolerance of the organism. Knockouts for *hspA* already exist and they are both heat-sensitive and have a loss of acquired heat tolerance (Lee et al., 2000). HhoA has also been implicated in thermal tolerance in *Synechocystis* (Sokolenko et al., 2002).

**References**

photosynthetic apparatus. Ottawa, Canada: Department of Fisheries and Oceans, 423–500.


