The global transcriptional response of *Escherichia coli* to induced $\sigma^{32}$ protein involves $\sigma^{32}$ regulon activation

followed by inactivation and degradation of $\sigma^{32}$ *in vivo*

Kai Zhao$^1$, Mingzhu Liu$^{2,3}$, Richard R. Burgess$^1$,

$^1$McArdle Laboratory for Cancer Research, $^2$Department of Genetics, $^3$Department of Computer Science, University of Wisconsin, Madison, WI 53706

Running title: $\sigma^{32}$ regulon in *E. coli*

Keywords: sigma factor, regulon, microarray, Western blot, *E. coli*, anti-sigma factor

$^#$ Address correspondence to: Richard R. Burgess, McArdle Laboratory for Cancer Research, University of Wisconsin-Madison, 1400 University Avenue, Madison, WI 53706, Email: rburgess@wisc.edu, Phone: 608-263-2635, Fax: 608-262-2824

$\sigma^{32}$ is the first alternative sigma factor discovered in *E. coli* and can direct transcription of many genes in response to heat shock stress. To define the physiological role of $\sigma^{32}$, we have used transcription profiling experiments to identify, on a genome-wide basis, genes under the control of $\sigma^{32}$ in *E. coli* by moderate induction of a plasmid-borne *rpoH* gene under defined, steady-state growth conditions. Together with a bioinformatics approach, we successfully confirmed genes previously known to be directly under the control of $\sigma^{32}$ and also assigned many additional genes to the $\sigma^{32}$ regulon. In addition, to better understand the functional relevance of the increased amount of $\sigma^{32}$ to changes in the transcriptional level of $\sigma^{32}$-dependent genes, we measured the protein level of $\sigma^{32}$ both before and after induction by a newly developed quantitative Western blot method. We found that, at a normal constant growth temperature (37°C), the $\sigma^{32}$ protein level rapidly increased, plateaued, and then gradually decreased after induction, indicating $\sigma^{32}$ can be regulated by genes in its regulon and that the mechanisms of $\sigma^{32}$ synthesis, inactivation, and degradation are not strictly temperature dependent. The decrease in the transcriptional level of $\sigma^{32}$-dependent genes occurs earlier than the decrease in full-length $\sigma^{32}$ in wild-type strain, and the decrease in the transcriptional level of $\sigma^{32}$-dependent genes is greatly diminished in a ΔDnaK strain, suggesting that DnaK can act as an anti-sigma factor to functionally inactivate $\sigma^{32}$ and thus reduce $\sigma^{32}$-dependent transcription *in vivo*.
genes upon heat shock stress. Most known $\sigma^{32}$-dependent genes were identified either by monitoring synthesis rates of individual proteins before and after heat shock on two-dimensional gels (8) or by hybridizing cDNA (generated mRNA from heat-shocked cells) with membrane filters containing an ordered Escherichia coli genomic library (9). However, in response to temperature upshift, the induction of $\sigma^{5}$ was shown in a Western blot experiment and the induction of $\sigma^{5}$-dependent genes was confirmed by using $\sigma^{5}$-dependent promoter-lacZ fusion approach (10). Also, Taylor et al. (11) found $\sigma^{54}$-controlled genes are another group of genes that can be induced by heat shock in addition to $\sigma^{32}$ and $\sigma^{E}$ regulons (12).

Therefore, although the heat shock response is mainly mediated by $\sigma^{32}$, there are some other global gene regulators that increase and turn on genes during the heat shock response. This makes the heat shock stimulon a complicated group of different regulons.

Here we report the results of using transcription profiling experiments to identify the $\sigma^{32}$ regulon in Escherichia coli. Our basic strategy is to minimally perturb steady-state growth (Escherichia coli MG1655 growing exponentially in minimum medium at $37^\circ C$) by moderate induction of $\sigma^{32}$. We then monitor global RNA transcript abundance changes as a function of time using Affymetrix GeneChip® Escherichia coli Antisense Genome Arrays. This approach allows us to reduce the possibility of induction of genes under the control of other sigma factors that was found in the previous heat shock response studies. Meanwhile, to characterize how the transcriptional level of $\sigma^{32}$-dependent genes is regulated by the amount of $\sigma^{32}$ in vivo, we measure the protein level of $\sigma^{32}$ both before and at various time points after induction by a quantitative Western blot analysis. On the basis of this first systematic study of the $\sigma^{32}$ regulon in Escherichia coli including $\sigma^{32}$ protein level determination, we gain insight into the complex network regulated by $\sigma^{32}$ and how it contributes physiological adaptation to changes in the external environment.

**EXPERIMENTAL PROCEDURES**

*Reagents and Media* - All reagents were purchased from Sigma Chemical Company (St. Louis, MO) unless otherwise indicated. 10X MOPS minimal media was prepared as described in Neidhardt et al. (13). The media was filter sterilized through a 0.2 μm filter and stored at 4°C. The defined media for cell log-phase growth contained 1 X MOPS minimal media, 0.1% glucose, 0.66 mM K$_2$HPO$_4$.

*Bacteria strains and plasmids* - In order to controllably and quantitatively overexpress $\sigma^{32}$ in Escherichia coli, we constructed an overexpression vector derived from the pZ-vectors developed by Lutz and Bujard (14). The pZ-vector system features a tightly-regulated low-copy-number plasmid with a widely controllable regulatory range. A DNA fragment that contains the entire $\sigma^{32}$ protein coding region was cloned into the KpnI and AvaiI sites of pZA31-luc plasmid, putting the entire *rpoH* ORF under the control of the P$_{Ltet}$ promoter to produce $\sigma^{32}$ overexpression vector pTet32.

A Tet repressor (TetR) expression plasmid, pIL4, carrying the entire TetR gene served as PCR template. A 690bp-long Tet repressor gene was PCR-amplified by using primers TetRUp (5'-AACTGCAGAACCAATGCATTGGTGGTAAAATAACTCTATCAA-3'; PstI site underlined) and TetRDown (5'-TCCCCCCGGGGGATTTAAGACCCACTTTCACTTTCAATT-3'; SmaI site underlined). A kanamycin-resistance (Km) coding sequence was amplified from plasmid pACYC177 by PCR using primer pairs of KmUp (5'-GGAATTCACGTTCGTAAGCCATTTCC-3'; EcoRI site underlined) and KmDown (5'-GGGGTACCACGTTCGTAAGCCATTTCCAAGTGAGTAGG-3'; KpnI site underlined). Both resulting PCR products were cloned into pMOD-2 Transposon construction vector (Epicentre).

Because the Escherichia coli Genochip probe set is based on the sequenced Escherichia coli K-12 strain MG1655 (λ- F- ilvG- rfb 50 rph-1, prototroph) (15), we chose this bacterial strain to use in our study. For tight control of the P$_{Ltet}$ promoter, we constructed a derivative Escherichia coli strain MG1655K1, integrating Tet repressor gene and a selectable marker kanamycin-resistance gene into the MG1655 chromosome. Tn5 transposon and transposase were used following the procedure of Goryshin et al. (16).

*Growth conditions, preparation of cell lysates* - All cultures were grown in a New Brunswick...
Gyrotory water bath shaker (model G76) with vigorous aeration (225 rpm) unless otherwise indicated. For cultures of cells carrying antibiotic resistance markers, the media were supplemented with ampicillin (100 µg/ml), chloramphenicol (30 µg/ml), or kanamycin (50 µg/ml) where appropriate. For induction of σ^{32} under the control of the anhydrotetracycline (aTc)-regulated promoter, aTc was added at a final concentration of 100 ng/ml.

*E. coli* strain MG1655K1 containing a σ^{32} overexpression plasmid (pTet32) was grown overnight in MOPS minimal media at 37°C in an air shaker. 2 mL of the overnight culture was used to inoculate 100 mL of fresh MOPS minimal medium. When the culture density reached optical density of 0.2, a 1000 µL portion of culture was harvested into a pre-chilled 1.5 mL Eppendorf tube and then immediately put on ice for 1 minute. This sample served as the control for Western blot analysis. To measure changes in the σ^{32} intracellular level, cells were then harvested every five minutes after induction, immediately put on ice for 1 minute and centrifuged at 10,000g (12,000 rpm for BECKMAN Microfuge®) for 10 minutes at 4°C. The supernatant was removed and the cell pellet resuspended immediately in 40 µl lysis buffer (1 X SDS) and heated at 75°C for 5 minutes to quickly lyse the cells and prevent changes in the intracellular levels of the sigma factors being measured.

**RNA isolation, cDNA synthesis, labeling and hybridization** - For preparing the total RNA for microarray experiments, 15 mL samples of cells were taken at 5, 10 and 15 minutes after induction, immediately mixed with a double volume of RNAprotect Bacterial reagent (Qiagen) and then incubated at room temperature for 10 minutes. Cells were centrifuged at 5,800g for 20 minutes and cell pellets were stored at -80°C prior to RNA extraction. Total nucleic acid was isolated using MasterPure kits (Epicentre) as described by the manufacturer. DNase I (Epicentre) was used to remove genomic DNA contamination. The quality and integrity of the isolated RNA was checked by visualizing the 23S and 16S rRNA bands on a 2% agarose gel. 10 µg of total RNA was mixed with 500 ng random hexamers and then was reverse transcribed for first strand cDNA by using the Superscript II system (Invitrogen). RNA was removed by using RNase H (Life Technologies) and RNase A (Epicentre). cDNA was purified by using Qiag Quick PCR purification kit (Qiagen) and followed by partial DNase I digestion to fragment cDNA to an average length of 50-100bp. The fragmented cDNA was 3’-end-labeled by using terminal transferase (New England Biolabs) and biotin-N6-ddATP (PerkinElmer) and was added to hybridization solution to load on Affymetrix GeneChip® *E. coli* Antisense Genome Arrays. Hybridization was carried out at 45°C for 16 h. The arrays were then washed and subsequently stained with streptavidin, biotin-bound anti-streptavidin antibody and streptavidin-phycocerythrin (Molecular Probes) to enhance the signal. Arrays were scanned at 570 nm with a 3 µm resolution using a confocal laser scanner (Hewlett-Packard). For each time point, two independent cultures were prepared and the RNA was analyzed in microarray experiments.

**Data analysis** - Image analysis was carried out by Affymetrix® Microarray Suite 5.0 software. Cell intensity files were first generated from the image data files. An absolute expression analysis then computes the detection call, detection p-value and signal (background-subtracted and adjusted for noise) for each gene. Genes were considered up-regulated relative to the 0 time point (before induction) sample if they had a 2-fold increase in signal intensity and the signal intensity in the experiment had a log_{2} value of at least 8.0 with a “present” detection call; The higher log_{2} intensity values were used to limit the analysis to those genes for which we have a high degree of confidence in their level of expression.

**RNA isolation, cDNA synthesis, labeling and hybridization** - For preparing the total RNA for microarray experiments, 15 mL samples of cells were taken at 5, 10 and 15 minutes after induction, immediately mixed with a double volume of RNAprotect Bacterial reagent (Qiagen) and then incubated at room temperature for 10 minutes. Cells were centrifuged at 5,800g for 20 minutes and cell pellets were stored at -80°C prior to RNA extraction. Total nucleic acid was isolated using MasterPure kits (Epicentre) as described by the manufacturer. DNase I (Epicentre) was used to remove genomic DNA contamination. The quality and integrity of the isolated RNA was checked by visualizing the 23S and 16S rRNA bands on a 2% agarose gel. 10 µg of total RNA was mixed with 500 ng random hexamers and then was reverse transcribed for first strand cDNA by using the Superscript II system (Invitrogen). RNA was removed by using RNase H (Life Technologies) and RNase A (Epicentre). cDNA was purified by using Qiag Quick PCR purification kit (Qiagen) and followed by partial DNase I digestion to fragment cDNA to an average length of 50-100bp. The fragmented cDNA was 3’-end-labeled by using terminal transferase (New England Biolabs) and biotin-N6-ddATP (PerkinElmer) and was added to hybridization solution to load on Affymetrix GeneChip® *E. coli* Antisense Genome Arrays. Hybridization was carried out at 45°C for 16 h. The arrays were then washed and subsequently stained with streptavidin, biotin-bound anti-streptavidin antibody and streptavidin-phycocerythrin (Molecular Probes) to enhance the signal. Arrays were scanned at 570 nm with a 3 µm resolution using a confocal laser scanner (Hewlett-Packard). For each time point, two independent cultures were prepared and the RNA was analyzed in microarray experiments.

**Data analysis** - Image analysis was carried out by Affymetrix® Microarray Suite 5.0 software. Cell intensity files were first generated from the image data files. An absolute expression analysis then computes the detection call, detection p-value and signal (background-subtracted and adjusted for noise) for each gene. Genes were considered up-regulated relative to the 0 time point (before induction) sample if they had a 2-fold increase in signal intensity and the signal intensity in the experiment had a log_{2} value of at least 8.0 with a “present” detection call; The higher log_{2} intensity values were used to limit the analysis to those genes for which we have a high degree of confidence in their level of expression.

**Purification and fluorescence labeling of proteins and MAbs** - Purified core RNA polymerase was made from *E. coli* MG1655 according to the method of Thompson et al. (17). Purified sigma factors and monoclonal antibodies (MAbs) were made as described in Anthony et al (18). Mouse MAbS used in this experiment were anti-α (4RA2), anti-β’ (NT73), anti-σ^{70} (2G10) and anti-σ^{32} (3RH3). Fluorescence dye, IC5-OSu (Dojindo), was used to label the primary antibodies according to previously described methods (19). The IC5-labeled MAbS, stored at a final concentration of 1 mg/ml, were diluted 1:2000 for use in this experiment.
Electrophoresis and immunoblot assay - Lysate samples were electrophoresed on a 4-12% NuPAGE gel. The purified core RNA polymerase as well as the purified sigma factor protein was also loaded on the same gel to serve as controls. The gel was run at a constant voltage of 125 V until the bromphenol blue loading dye had almost run off the bottom of the gel. Proteins in the gel were transferred electrophoretically to a 0.45 µm nitrocellulose membrane at 50 V for 2 hour at room temperature. The membrane was blocked with 1% (w/v) non-fat dry milk (Blotto) for 30 minutes at room temperature or overnight at 4°C. The blot was then probed in Blotto for 1 hour at room temperature with fluorescence-labeled MAb specific to the sigma factor under study and a subunit of core RNA polymerase. The blot was rinsed three times with 25 ml PBS buffer and scanned with a Typhoon FluorImager in the red fluorescence-scanning mode. Signal intensities of the bands were quantified using the ImageQuant program.

To measure soluble protein levels in vivo, the cell pellet was harvested and resuspended in buffer A as described by Anthony (18) before lysozyme was added to facilitate cell breakage. Cells were then sonicated for 90 seconds to completely break cell walls and the sample was centrifuged at 20,000 x g for 15 minutes at 4°C. Any insoluble inclusion bodies plus cell debris would be in the pellet. The supernatant containing soluble protein was then measured in Western blot assay.

Real-time PCR - Quantitative reverse transcription (RT)-PCR primers were designed using Primer Express software (Applied Biosystems) and were synthesized by the UW Biotechnology Center. Two steps of real-time quantitative RT-PCR are performed. 5 µg of the DNase-treated total RNA was reverse transcribed for first strand cDNA by using Superscript II system (Invitrogen) as mentioned above. Reactions were then performed using 1 ng cDNA and 100 nM of each primer in a 50 µl volume with SYBR Green I mixture. Controls lacking AmpliTaq Gold DNA Polymerase or template were used. Reactions were run on an ABI 7700 instrument (Applied Biosystems) using the following cycling parameters: 95°C for 10 minutes, 40 cycles of denaturation at 94°C for 15 sec and extension at 60°C for 1 minute. Relative gene expression data analysis was carried out with the standard curve method (20). Changes in expression will be calculated using the time zero sample as the reference.

Electrophoretic mobility-shift assays (EMSA) - The DNA fragments (~300bp) used for gel mobility shift assays were amplified by PCR from the upstream sequence of five genes (yceP, ldhA, macB, mutM, ybbN) that were highly up-regulated in our microarray data. The DNA was labeled at the 5’ end using T4 polynucleotide kinase (Life Technologies) and [γ-32P]ATP (5,000 Ci/mmol; PerkinElmer Life Science) at 37°C for 45 minutes. The unincorporated nucleotides were removed by passing the labeling reaction mixture through a G-50 Sephadex microspin column (Amersham). Core RNA polymerase and σ32 were purified using the procedures described earlier (18). The labeled DNA fragment (1.15 nM) was incubated with different concentrations of core RNA polymerase and σ32 in a buffer containing 20 mM Tris-acetate (pH 8.0), 0.1 mM EDTA, 1 mM dithiothreitol, 50 mM NaCl, 4 mM magnesium acetate, 5% glycerol (v/v), and 200 ng of poly(dI-dC) · poly(dI-dC) (Amersham) in a total volume of 20 µl. The mixtures were left on ice for 30 minutes before being incubated at 30°C for 15 minutes. The samples were loaded directly onto a 4-10% native Tris-glycine Novex Gel (Invitrogen) and were run at 4°C in 25 mM Tris, 190 mM glycine (pH 8.3) at 200 V for 1h. The gel was fixed in a solution of 10% acetic acid and 10% methanol for 15 minutes and dried at 80°C on a Slab Dryer (BioRad). Biomax MS film (Kodak) was used for autoradiography. The gels were scanned using a PhosphorImager (Molecular Dynamics), and the intensities of the bands were determined using ImageQuant version 5.2 software.

RESULTS

The construction of E. coli strain MG1655K1 - The sequenced E. coli K-12 strain MG1655 (λ’ FlvG rfb 50 rph-1, prototroph) (15), on which E. coli Affymetrix Genechip probe design is based, was chosen for our studies. For controllable induction of individual sigma factors in vivo, we used the Ptet promoter to construct an overexpression vector (14). The Ptet promoter is
controlled by the repressor TetR. A downstream gene can be induced in the presence of inducer, anhydrotetracycline (aTc). The TetR repressor gene as well as the kanamycin resistance gene were cloned into the pMOD-2 transposon construction vector as described in Materials and Methods. To ensure stable and defined conditions for the synthesis and maintenance of the regulatory protein TetR, the gene encoding this repressor molecule was integrated into the chromosome of this sequenced E. coli strain by using the Tn5 Transposon as described by Goryshin et al. (16). Analysis of several kanamycin-resistant colonies by PCR and Southern blots showed that the transcription unit encoding TetR as well as the kanamycin resistance marker were stably integrated into the MGI655 genome (data not shown).

Quantitation of $\sigma^{32}$ after induction - A newly developed quantitative Western blot method (19) was used to monitor the intracellular level of $\sigma^{32}$ in vivo before and after induction. The $\alpha$- or $\beta'$-subunits of core RNA polymerase were also examined to serve as internal controls because their intracellular levels remain constant under various conditions (21,22). The signal intensities of the proteins were immunodetected by corresponding IC5-labeled monoclonal antibodies. Our results (Fig. 1A) show that while the signal intensities of $\alpha$- or $\beta'$-subunits are quite constant at all the time points before and after induction, $\sigma^{32}$ levels vary at different time points. Generally, the $\sigma^{32}$ protein level, which is normalized to the $\beta'$ subunit of RNA polymerase, rapidly increased 5 minutes after induction with almost 7.4 fold change and then stayed at a high level for 10 minutes (~8.2 fold ) before it gradually decreased (Fig. 1B).

To eliminate the possibility that this decrease of $\sigma^{32}$ 20 minutes after induction was due to our overexpression system, we performed the same experiment for $\sigma^{70}$ overexpression by cloning the entire $\sigma^{70}$ protein coding region under the control of the same inducible $P_{\text{List}}$ promoter. The same experiment was performed to test the intracellular level of $\sigma^{70}$ after induction. We extended our induction time to 60 minutes and found the $\sigma^{70}$ protein level kept increasing as shown in Fig. 1A & 1B. Apparently, there was no decrease in $\sigma^{70}$ protein level. From this comparison, we can conclude that this decrease in $\sigma^{32}$ protein after induction is not due to the overexpression system that we used in the experiment and there might be a feedback regulatory system in vivo that causes this decrease (see below).

The more significant increase (fold change) of $\sigma^{32}$ at 5-minute time point, compared with $\sigma^{70}$ induction level, is due to the fact that the experiment was performed at log-phase (OD$_{600}$=0.2) in minimum medium, in which $\sigma^{70}$ is the dominant sigma factor and has a much higher protein level than $\sigma^{32}$ before induction. Thus, although the two sigma factors are under the same $P_{\text{List}}$ promoter control, the fold change of $\sigma^{32}$ is much higher because of its lower initial protein level.

Known $\sigma^{32}$-dependent genes are induced after $\sigma^{32}$ overexpression - To characterize the effect of the increasing $\sigma^{32}$ protein level in vivo on gene expression, global RNA transcript abundance was monitored at 5, 10, 15 minutes after $\sigma^{32}$ induction with cells grown in log-phase (OD$_{600}$=0.2) in MOPS minimal medium at 37°C. Transcriptional profiles were obtained as described in Materials and Methods. The sample at time zero was used as the reference to identify genes whose transcript abundance had significantly changed after $\sigma^{32}$ overexpression.

DNA microarray results showed most of the well-characterized genes belonging to the $\sigma^{32}$ regulon are induced following $\sigma^{32}$ overexpression. Several known $\sigma^{32}$-dependent genes (such as rpoD, a $\sigma^{32}$-dependent gene, whose transcriptional level increased 1.9 fold after 5 minutes induction) were not included into our data set due to our cutoff level (2-fold increase). In Table 1, we show some known $\sigma^{32}$-dependent genes that were up-regulated at least two fold in 5 minutes and also list the transcriptional levels of these genes at 10 and 15 minutes. Most of those genes were initially identified in heat shock stress and can be divided into three functional groups as shown in Table 1: (1) adaptation (heat shock related, atypical); (2) proteases (degradation of proteins/peptides); (3) chaperones.

Because our Western blot assay showed that $\sigma^{32}$ protein is degraded in vivo after induction, we paid specific attention to the genes belonging to proteases and chaperones groups. We observed...
that Lon, the first ATP-dependent protease isolated from *E. coli* (23,24) that plays an important role in general protein degradation, increased 3.6 fold after 5 minutes of induction. Meanwhile, the transcriptional levels of the Clp family genes, which encode the two-component Clp protease (the catalytic subunits ClpP and ClpQ[HslV]) and the regulatory subunits ClpX, ClpY[HslU] and possibly ClpB), are significantly induced after induction. These cytoplasmic proteases can degrade a variety of proteins as well as some specific substrates *in vivo* (25-28). The transcriptional level of a membrane-bound metalloprotease FtsH (HflB), which was first implicated as a protease responsible for σ^32- degradation(29,30), increased 4.4 fold at the 5-minute time point.

A number of heat shock protein chaperones were also involved in degrading abnormal proteins (27,28). Among the induced chaperone proteins, DnaK/DnaJ/GrpE chaperone team is involved in folding of nascent chains and plays a significant role in cellular folding reactions (31-35). A similar function can be found in the GroEL and GroES chaperone team (31,36). The transcriptional level of these five genes (*dnaK, dnaJ, grpE, groEL, groES*) increased 4.8, 2.3, 10.5, 7.6 and 4.8 fold, respectively, at 5 minutes after induction. The potential role of chaperones to promote σ^32- degradation is that chaperones can compete with RNAP to bind σ^32- and then make σ^32- unstable and more easily degraded by the protease machinery (27,28).

**New candidate genes for σ^32- regulon** - Expression profiling of transcripts corresponding to the complete set of ORFs in *E. coli* genome revealed that the response to induced σ^32- level *in vivo* was quite broad. As a result of simple mass action, an increase in the level of σ^32 relative to the other sigma factors should lead to an increase in the expression of genes in the σ^32 regulon due to the increase of corresponding σ^32 holoenzyme. In addition to identifying known σ^32-dependent genes, our microarray data also allowed us to assign many additional new candidate genes to the σ^32 regulon. There are 129 (3.0% of genome), 116 (2.7% of genome) and 51 (1.2% of genome) genes up-regulated two fold or more, at 5 minutes, 10 minutes and 15 minutes after induction, respectively. In this paper, we show a group of genes whose transcriptional level increases more than four fold at 5 minutes and keeps at a high level (more than two fold) at 10 minutes in Table 2.

To further confirm new genes in the σ^32 regulon, we chose the top five up-regulated genes (yeceP, ldhA, macB, mutM and ybbN) in Table 2 for native gel shift assays. Interestingly, in choosing promoter sequences for gene macB, yeceP and ybbN, we found that, as shown in Fig. 2, their upstream genes in the same predicted operon (37,38) show no change or even a slight decrease in both our σ^32 over-expression study and previous heat shock microarray data (Liu et al. personal communication). Therefore, we predict that there are additional promoters which have not been discovered or annotated in the DNA sequences upstream of *macB*, yeceP and ybbN genes.

Native gel shift experiments were performed to test the binding of purified σ^32 holoenzyme to the promoter regions of these genes. The upstream sequence of *fliL* gene that contributes flagella biosynthesis function was chosen as a negative control for the gel-shift assay because transcription of this gene is regulated by σ^70 and σ^F and is not σ^32-dependent (39). In our σ^32 overexpression microarray data, the transcriptional level of this gene (*fliL*) is downregulated 2.6 fold at 5 minutes after induction.

Binding of each promoter region by σ^32- associated holoenzyme was examined at three different molar ratios (1:0, 1:2.5, 1:5) of core RNA polymerase to σ^32 protein. Electrophoretic mobility-shift assay (EMSA) results (Fig. 3C) show that the DNA fragment generated from the upstream DNA sequences of these five up-regulated genes can be shifted by σ^32 holoenzyme. Although *yeceP* is the most up-regulated among the five genes as indicated by the microarray data, its transcript abundance is lower than *ldhA* and almost the same as that of other genes. Therefore we are not surprised that its relative promoter binding preferences is not the most efficiently bound by holoenzyme as determined by EMSA. The *fliL* promoter region showed no binding, suggesting its down-regulation, mentioned above, is not due to negative control by Erσ^32 binding, but more likely to competition of sigmas for core *in vivo*. We have observed that many of σ^F-
dependent genes are down-regulated upon $\sigma^{32}$ induction (data not shown).

Results from promoter region consensus analysis using the algorithms MEME (40) and BioProspector (41) revealed $\sigma^{32}$ binding sites in the upstream regulatory sequences of these genes (Fig. 3B). Note that four of the five promoter regions contain two potential $\sigma^{32}$-holoenzyme binding sites that were predicted by computer programs. We do not know if the binding observed is due to binding at one or both of these sites. While these possible two-block DNA consensus sequences provide the primary interaction with holoenzyme, additional transcriptional activators such as FIS and CRP might be utilized to strengthen the promoter-holoenzyme interaction in vivo which is not available in our in vitro gel shift assay. In addition, deviation from the consensus sequence is common and contributes to reduce the binding strength of holoenzyme to the promoter (42).

**Gene expression patterns as a function of time after $\sigma^{32}$ induction** - One of the interesting observations in this time course microarray analysis was that the global changes in gene expression upon $\sigma^{32}$ increase in vivo were quite transient. The consistent pattern is: the $\sigma^{32}$ regulon is rapidly induced in response to $\sigma^{32}$ protein level increase with RNA levels increasing by 5 minutes and generally declining 10 minutes after induction (Fig. 4). The maximum number of up-regulated genes was found 5 minutes after induction and signal intensities of the genes which represent their transcriptional level are highest at this same time point. A slight decrease in both the number and the transcriptional level of the up-regulated genes occurs by 10 minutes. By 15 minutes after induction, a significant decrease of the number of up-regulated genes occurs (Fig. 4A) and the transcriptional levels of those genes up-regulated at 5 minutes become low or almost return to pre-induction levels (Fig. 4B).

Combined with the observed $\sigma^{32}$ protein level change measured by a quantitative Western blot assay, we found that the decrease in the transcriptional level of $\sigma^{32}$-dependent genes occurs earlier than the decrease in full-length $\sigma^{32}$ in our assay (Fig. 4C). The induced high $\sigma^{32}$ protein level at 10 and 15 minutes does not maintain the high transcriptional levels of its regulon. This indicates at least part of this $\sigma^{32}$ is not functional in vivo.

One possible reason is that more of the $\sigma^{32}$ is present in inclusion bodies at 10 and 15 minutes than at 5 minutes after induction. Therefore, although we might detect the total $\sigma^{32}$ increase in whole cell lysates at 10 and 15 minutes time points using the specific mAb, the $\sigma^{32}$ in inclusion bodies would not functionally bind to core RNA polymerase to form holoenzyme and then transcribe $\sigma^{32}$-dependent genes. To test this possibility, instead of measuring the total $\sigma^{32}$ in whole cell lysates, we performed experiments to measure soluble $\sigma^{32}$ protein levels before and after induction. Results showed the soluble $\sigma^{32}$ still remained at a high level 10 and 15 minutes after induction and the overall trend of increased soluble $\sigma^{32}$ is similar to that of increased total $\sigma^{32}$ we measured before (Fig. 4C). This result suggests that although inclusion body production is common in protein overexpression (usually the induction time is 3 or 4 hours), most overexpressed $\sigma^{32}$ observed here is soluble and does not reach the threshold of $\sigma^{32}$ aggregation in that short time period (0-20 minutes). Therefore, transcription of $\sigma^{32}$-dependent genes rapidly decreases as a result of decrease in $\sigma^{32}$ activity rather than in $\sigma^{32}$ level or solubility.

**DnaK is responsible for inactivating $\sigma^{32}$ in vivo** - Inactivation and degradation of $\sigma^{32}$ under conditions of excess $\sigma^{32}$ regulon expression as shown in our assay suggests that $\sigma^{32}$ can be feedback regulated by genes in its regulon. The possible $\sigma^{32}$-dependent gene expression that is involved in inactivation of $\sigma^{32}$ might be chaperones, particularly the DnaK-DnaJ chaperones. Physical interaction (binding) between $\sigma^{32}$ and DnaK is well documented, both in crude lysates and in a purified system (43,44). To test the function of DnaK in our $\sigma^{32}$ overexpression system, we made a dnaK in-frame deletion strain, where chromosomal position from 12193 to 14049 in the DnaK coding region has been deleted, following the description of Datsenko et al (45,46). A $\sigma^{32}$ overexpression experiment like that performed earlier in this paper has been performed in this ΔDnaK strain. Instead of using a microarray approach, we used a real-time RT-PCR assay to measure the transcriptional
level changes of two well-known $\sigma^{32}$-dependent genes (lon and grpE) as described below. Results showed that the decrease of the transcriptional level of these two $\sigma^{32}$-dependent genes was diminished and more parallel to the decrease of $\sigma^{32}$ protein level in $\Delta$DnaK strain (Fig. 5), indicating DnaK contributes to inactivation of $\sigma^{32}$ and causes a decrease of $\sigma^{32}$-dependent gene transcription in our assay.

Comparison of microarray data with RT-PCR results - For comparison with the array data, we independently determined the degree of induction of mRNA for several different genes by the quantitative real-time RT-PCR approach. Five ORFs, exhibiting high, moderate, and low expression as identified by microarray analysis, were selected for this purpose. Their quantitative values were obtained by using the comparative threshold cycle (CT) method recommended by Applied Biosystems. Gene expression levels were normalized to that of the $rpoA$ gene, because its expression was found to be invariant under different time points before and after $\sigma^{32}$ induction in our array data. The relative expression of each gene was determined in each of the two experimental RNA samples and is expressed as the fold difference in quantity of cDNA molecules present at 5, 10 and 15 minutes time points relative to that present at the zero time point. The resulting gene expression ratio was plotted against the average log ratio values obtained by microarray analysis (Fig. 6).

A high level of concordance ($r = 0.985$) was observed between microarray and RT-PCR data. The overall trends (i.e., high or low expression) seen in the data derived from microarrays were consistent with those derived from the real-time PCR analysis. This validation study by real-time RT-PCR indicates that our microarray approach produces accurate fold change differences with sufficient sensitivity to identify differentially regulated transcripts.

Computer prediction of $\sigma^{32}$-related promoter elements - A computer program was used to examine upstream DNA sequence of upregulated genes in our microarray data to look for regulatory sequence motifs. As prokaryotic promoter motifs often occur in two blocks with a gap of variable length, BioProspector (41), a C program which is capable of modeling motifs with two blocks and uses a Gibbs sampling strategy, was used to find the -10 and -35 consensus regions for $\sigma^{32}$ binding. Upstream sequences (300 bases from the first genes in transcription units that contain 2 fold up-regulated genes in our microarray data) were extracted as input sequences. A number of overall highest scoring motifs as position-specific probability matrices were reported. According to the reported highest scoring motif and its site locations on the input sequence, a graphical display of the results was generated using SEQUENCE LOGO (47) (Fig. 7). The resulting consensus is represented as a ggcTTGa-(N)_{12-20}eCCCCAT, where lower case indicates a less highly conserved site. Higher sequence conservation is observed in the -10 region. This consensus agrees well with previously reported $\sigma^{32}$ consensus which was aligned to maximize alignment (CTTGAA-(N)_{13-17}CCCCCATNT) in the -35 and -10 regions of several published $\sigma^{32}$ promoters (3,27,28,48).

Comparison of the $\sigma^{32}$ regulon with heat shock stimulons - The heat-shock response is a cellular protective and homeostatic response to cope with stress-induced damage to proteins. Many heat shock proteins (HSPs) play major roles in protein folding, assembly, transport, repair and turnover under stress and non-stress conditions. Activation of $\sigma^{32}$ in response to heat shock is well documented. In E. coli, induction of HSPs occurs primarily by increase of the master regulator $\sigma^{32}$, which specifically directs RNA polymerase to transcribe from heat shock promoters (27,28). Therefore, the $\sigma^{32}$ regulon is often equated with the heat shock regulon. To extend our studies, we compared our $\sigma^{32}$ overexpression data with the expression profiles from exponentially growing cultures that was subject to 10 minutes heat shock by a shift in growth temperature from 30°C to 50°C. The cells responded with over 300 genes that are up-regulated more than two fold in heat shock experiments (Liu et al. unpublished) to cope with heat-induced damage in bacteria. We are not surprised that more genes have been turned on in the heat shock response. Compared to the minimum perturbation by moderate induction $\sigma^{32}$ in the steady-state growth cells in our assay, heat shock stress is a much stronger stress response and will induce other sigma factors (10-12,27,48,49) to turn on multiple regulons to meet the complex
cellular requirement that protect cells against cytoplasmic or periplasmic protein damage. Therefore, although the heat shock stimulen, to some extent, overlaps with σ^{32} regulon, there is a significant difference between the two sets of genes (Fig. 8).

**DISCUSSION**

The power and utility of microarray approaches to define stimulons and regulons have been applied in different species of bacteria (50-56). In this study, we have constructed a system for controllable induction of individual sigma factors to study genes under σ^{32} control by using microarrays. Because σ^{32} is a major sigma factor in the heat shock response, previous studies on the σ^{32} regulon were mostly focused on a group of heat shock proteins (HSPs). Several global transcription analyses in response to growth temperature variation have been carried out in various bacterial species (50,53,56). Since no experiments have been performed so far with *E. coli* to systematically understand the genes that are directly under control of σ^{32}, for purposes of discussion here, we define the σ^{32} regulon as those genes up-regulated after the induction of σ^{32}.

Through studies on *E. coli* and many other organisms, it has become clear that a major means of gene regulation occurs at the level of transcription initiation. Transcription initiation is regulated by sigma factors and the first step of the transcription initiation pathway is the binding of a sigma factor to core RNA polymerase to form a holoenzyme complex which then binds to a specific set of promoters and initiates transcription. The binding of the different sigma factors to core RNA polymerase ultimately results in the expression of a set of genes or a regulon. As a result of simple mass action, an increase in the level of σ^{32} relative to the other sigma factors would result in an increase in the level of the corresponding σ^{32} holoenzyme. This would lead to an increase in the expression of the σ^{32} regulon.

Compared with heat shock experiments in which several other sigma factors are also induced *in vivo*, our system has the advantage of reducing the upregulation of genes controlled by other sigma factors and allowing specific study the σ^{32} regulon. Using our specific monoclonal antibodies for different sigma factors, we tested σ^{70}, σ^{8}, σ^{14}, σ^{54} protein level changes as a function of time in our σ^{32} overexpression experiments and found there are no significant changes of these sigma factors in our assays (data not shown). We also examined the expression of several well-known genes that are under direct control of other six sigma factors. The transcriptional level of these genes showed either no change or decreased in our σ^{32} overexpression microarray data. Taken together, these results make us confident that the group of upregulated genes in our experiment is predominantly due to the increased σ^{32} *in vivo*.

Another advantage of our approach for studying the σ^{32} regulon is its relatively higher induced σ^{32} protein level. Under the control of strong P_{Ltet} promoter, the induced σ^{32} protein level in our assay is higher than the σ^{32} level caused by the temperature upshift in previous heat shock experiments. The higher level of σ^{32} in our assay will turn on a group of genes which have a weak σ^{32}-dependent promoter to detectable levels. When compared with the sample at time zero as reference, the significant changes of genes in this group will be detected in our data but missed in previous heat shock studies due to no or low transcriptional level of those genes. Therefore, we think our approach provides a “purer” and more complete set (if not all) of genes in the σ^{32} regulon.

We also have utilized a new method of quantitative Western blotting (19) to measure the intracellular protein levels of sigma factors. Measuring both the protein level of σ^{32} and the transcriptional levels of σ^{32}-controlled genes as a function of time provides valuable information for exploring the complex network regulated by σ^{32} and gives us an insight into how the σ^{32} protein level regulates the transcriptional level of σ^{32}-controlled genes *in vivo*. Our results showed that although the σ^{32} protein level remained at a high level 10 and 15 minutes after induction, the number of induced genes and the transcriptional level of σ^{32}-controlled genes were significantly down-regulated at these same time points (Fig. 4), that is, the decrease in the transcriptional level of σ^{32}-dependent genes occurs earlier than the decrease in full-length σ^{32}.
In heat shock response, the induced synthesis of $\sigma^{32}$ usually takes place in 5 minutes and then declines to a new steady-state level, two- to three-fold higher than the pre-shift level (57). Meanwhile, a group of heat shock proteins decreases as well after their initial increase (57). The decrease of the $\sigma^{32}$ protein and the heat shock proteins virtually take place at the same time. Therefore, the amount of $\sigma^{32}$ in the cell was believed to be one of the key regulatory elements in the heat shock response (57).

The activity control of $\sigma^{32}$ that is involved in regulation of a limited number of HSPs was found in cold shock experiments (58) and in mutants lacking FstH function (59). Instead of measuring the synthesis of several heat-shock proteins at the translational level as carried out in these early studies, we measured the global transcriptional level change which better represents the level of active $\sigma^{32}$ in vivo by microarray assays. The transcriptional levels of several genes have been confirmed by real-time PCR. Results clearly showed that, whereas the increased amount of $\sigma^{32}$ largely accounts for initial induction of $\sigma^{32}$ regulon, the constitutive production of $\sigma^{32}$ from our overexpression system does not maintain the activation of $\sigma^{32}$ regulon in later time points. This indicates that decrease in $\sigma^{32}$ activity rather than in $\sigma^{32}$ level or solubility is responsible for the rapid shutdown of the transcription of $\sigma^{32}$-dependent genes in our assays. In addition, we measured the soluble $\sigma^{32}$ level to eliminate the explanation that decrease in $\sigma^{32}$ activity was due to insolubilities.

The DnaK-DnaJ-GrpE chaperone team is involved in $\sigma^{32}$ degradation in vivo (60,61), as mutations in each of the corresponding genes decrease the rate of $\sigma^{32}$ degradation (58). Tomoyasu, et al found a small increase (1.5-fold) in the DnaK-DnaJ level reduced the level and activity of $\sigma^{32}$ and caused faster shutdown of heat shock response (59,62), whereas a small decrease in the chaperone level caused inverse effects. The loss of the DnaK function leads to markedly impaired downregulation of the transcriptional level $\sigma^{32}$-dependent genes in our assay. This indicates that DnaK is a factor (or at least one of multiple factors) that is involved in inactivating $\sigma^{32}$ in vivo. However, the particular role played by DnaK or other factor(s) in promoting inactivation is not clear. A possible mechanism is that DnaK and core RNAP appear to compete with each other in binding to $\sigma^{32}$ at specific region(s) (28,63,64). The DnaK binding leads to $\sigma^{32}$ inactivation, whereas the RNAP binding stabilizes $\sigma^{32}$. From our in vitro studies that compare all seven purified E. coli sigma subunits binding affinities to the core RNA polymerase, we found $\sigma^{32}$ has highest binding affinity to core RNA polymerase among seven known sigma factors in E. coli (Bergendahl et al. unpublished data). Although the present study clearly reveals that the increased level of the $\sigma^{32}$ regulon expression exerts a negative feedback regulation on the intracellular protein level and activity of $\sigma^{32}$ and that DnaK is a factor involved in this process, more work needs to be done. We will explore whether DnaK or other factor(s) can only associate with free $\sigma^{32}$ and then prevent its formation of functional holoenzyme with core or whether they can bind to and remove $\sigma^{32}$ from a tightly bound RNAP-$\sigma^{32}$ complex by using a luminescence resonance energy transfer (LRET)-based homogeneous binding assay (65).

REFERENCES

1. Burgess, R. R., Travers, A. A., Dunn, J. J., and Bautz, E. K. (1969) *Nature* **221**, 43-46
2. Burgess, R. R., and Travers, A. A. (1970) *Fed Proc* **29**, 1164-1169
3. Wosten, M. M. (1998) *FEMS Microbiol Rev** **22**, 127-150
4. Helmann, J. D., and Chamberlin, M. J. (1988) *Annu Rev Biochem* **57**, 839-872
5. Haldenwang, W. G. (1995) *Microbiol Rev* **59**, 1-30
6. Sauer, U., Santangelo, J. D., Treuner, A., Buchholz, M., and Durre, P. (1995) *FEMS Microbiol Rev* **17**, 331-340
7. Gomez, J. E., Chen, J. M., and Bishai, W. R. (1997) *Tuber Lung Dis* **78**, 175-183
8. Lemaux, P. G., Herendeen, S. L., Bloch, P. L., and Neidhardt, F. C. (1978) *Cell* **13**, 427-434
9. Chuang, S. E., Daniels, D. L., and Blattner, F. R. (1993) *J Bacteriol* **175**, 2026-2036
10. Muffler, A., Barth, M., Marschall, C., and Hengge-Aronis, R. (1997) *J Bacteriol* **179**, 445-452
11. Kuczynska-Wisnik, D., Laskowska, E., and Taylor, A. (2001) *Biochem Biophys Res Commun* **284**, 57-64
12. Ericsson, J. W., and Gross, C. A. (1989) *Genes Dev* **3**, 1462-1471
13. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) *J Bacteriol* **119**, 736-747
14. Lutz, R., and Bujard, H. (1997) *Nucleic Acids Res* **25**, 1203-1210
15. Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277**, 1453-1474
16. Goryshin, I. Y., and Reznikoff, W. S. (1998) *J Biol Chem* **273**, 7367-7374
17. Thompson, N. E., Hager, D. A., and Burgess, R. R. (1992) *Biochemistry* **31**, 7003-7008
18. Anthony, L. C., Foley, K. M., Thompson, N. E., and Burgess, R. R. (2003) *Methods Enzymol* **370**, 181-192
19. Bergendahl, V., Glaser, B. T., and Burgess, R. R. (2003) *J Immunol Methods* **277**, 117-125
20. Winer, J., Jung, C. K., Shackel, L., and Williams, P. M. (1999) *Anal Biochem* **270**, 41-49
21. Ishihama, A. (2000) *Annu Rev Microbiol* **54**, 499-518
22. Charette, M. F., Henderson, G. W., and Markovitz, A. (1981) *Proc Natl Acad Sci U S A* **78**, 4728-4732
23. Gross, C. A. (1996) *Function and regulation of heat shock proteins. Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd ed (Neidhardt, F. C., Curtiss III, R., Ingraham, J.L., et al, Ed.), ASM Press, Washington, D.C.
24. Takashi Yura, M. K., and Miyo T. Morita. (2000) *The Heat Shock Response: Regulation and Function. Bacterial Stress Responses* (G. Storz, R. H.-A., Ed.). Chapter 1 vols., ASM Press, Washington, D.C.
25. Schroder, H., Low, K. B., Fenton, W. A., Hirshfield, I. N., and Furtak, K. (1993) *Cell* **74**, 909-917
26. Bockhorst, J., Qiu, Y., Glasner, J., Liu, M., Blattner, F., and Craven, M. (2003) *Bioinformatics* **19 Suppl 1**, i34-43
27. Salgado, H., Gama-Castro, S., Martinez-Antonio, A., Diaz-Peredo, E., Sanchez-Solano, F., Peralta-Gil, M., Garcia-Alonso, D., Jimenez-Jacinto, V., Santos-Zavaleta, A., Bonavides-Martinez, C., and Collado-Vides, J. (2004) *Nucleic Acids Res* **32 Database issue**, D303-306
28. Liu, X., and Matsumura, P. (1996) *Mol Microbiol* **21**, 613-620
29. Bailey, T. L., and Gribskov, M. (1998) *J Comput Biol* **5**, 211-221
41. Liu, X., Brutlag, D. L., and Liu, J. S. (2001) *Pac Symp Biocomput*, 127-138
42. Hawley, D. K., and McClure, W. R. (1983) *Nucleic Acids Res* **11**, 2237-2255
43. Gamer, J., Bujard, H., and Bukau, B. (1992) *Cell* **69**, 833-842
44. Liberek, K., Galitski, T. P., Zylicz, M., and Georgopoulos, C. (1992) *Proc Natl Acad Sci U S A* **89**, 3516-3520
45. Datsenko, K. A., and Wanner, B. L. (2000) *Proc Natl Acad Sci U S A* **97**, 6640-6645
46. Kang, Y., Durfee, T., Glasner, J. D., Qiu, Y., Frisch, D., Winterberg, K. M., and Blattner, F. R. (2004) *J Bacteriol* **186**, 4921-4930
47. Schneider, T. D., and Stephens, R. M. (1990) *Nucleic Acids Res* **18**, 6097-6100
48. Yura, T., Nagai, H., and Mori, H. (1993) *Annu Rev Microbiol* **47**, 321-350
49. Wang, Q. P., and Kaguni, J. M. (1989) *J Bacteriol* **171**, 4248-4253
50. Richmond, C. S., Glasner, J. D., Mau, R., Jin, H., and Blattner, F. R. (1999) *Proc Natl Acad Sci U S A* **97**, 6640-6645
51. Cao, M., Kobel, P. A., Morshedi, M. M., Wu, M. F., Paddon, C., and Helmann, J. D. (2002) *J Mol Biol* **316**, 443-457
52. Baichoo, N., Wang, T., Ye, R., and Helmann, J. D. (2002) *Mol Microbiol* **45**, 1613-1629
53. Helmann, J. D., Wu, M. F., Kobel, P. A., Gamo, F. J., Wilson, M., Morshedi, M. M., Navre, M., and Paddon, C. (2001) *J Bacteriol* **183**, 7318-7328
54. Britton, R. A., Eichenberger, P., Gonzalez-Pastor, J. E., Fawcett, P., Monson, R., Losick, R., and Grossman, A. D. (2002) *J Bacteriol* **184**, 4881-4890
55. Molle, V., Fujita, M., Jensen, S. T., Eichenberger, P., Gonzalez-Pastor, J. E., Liu, J. S., and Losick, R. (2003) *Mol Microbiol* **50**, 1683-1701
56. Stintzi, A. (2003) *J Bacteriol* **185**, 2009-2016
57.Straus, D. B., Walter, W. A., and Gross, C. A. (1987) *Nature* **329**, 348-351
58. Straus, D. B., Walter, W. A., and Gross, C. A. (1989) *Genes Dev* **3**, 2003-2010
59. Tatsuta, T., Tomoyasu, T., Bukau, B., Kitagawa, M., Mori, H., Karata, K., and Ogura, T. (1998) *Mol Microbiol* **30**, 583-593
60. Keller, J. A., and Simon, L. D. (1988) *Mol Microbiol* **2**, 31-41
61. Straus, D. B., Walter, W. A., and Gross, C. A. (1988) *Genes Dev* **2**, 1851-1858
62. Tomoyasu, T., Ogura, T., Tatsuta, T., and Bukau, B. (1998) *Mol Microbiol* **30**, 567-581
63. Nagai, H., Yuzawa, H., Kanemori, M., and Yura, T. (1994) *Proc Natl Acad Sci U S A* **91**, 10280-10284
64. McCarty, J. S., Rudiger, S., Schonfeld, H. J., Schneider-Mergener, J., Nakahigashi, K., Yura, T., and Bukau, B. (1996) *J Mol Biol* **256**, 829-837
65. Bergendahl, V., and Burgess, R. R. (2003) *Methods Enzymol* **370**, 192-205

FOOTNOTES

We thank Dr. Larry Anthony for advice on using the P_Ltet promoter system, Katherine Foley for monoclonal antibodies to sigma factors and subunits of RNA polymerase, Dr. Nancy Thompson for critical reading of the manuscript, the Dr. Frederick Blattner group for technical advice in microarray experiments and assistance in data analysis. We also thank Sandra Splinter-BonDurant, Wayne Davis, John Leucke at the Gene Expression Center for excellent technical assistance. The research was supported by a grant from NIGMS (GM28575) to R.R.B. In compliance with the University of Wisconsin-Madison’s Conflict of Interest Committee Policy, RRB acknowledges financial interest in NeoClone, LLC., (Madison, WI) which markets proteins and MAbs in this study.

FIGURE LEGENDS

**Fig. 1.** $\sigma^{32}$ and $\sigma^{70}$ protein level changes before and after $\sigma^{32}$ induction as measured by quantitative western blot analysis. The $\alpha$ and $\beta'$ subunits of core RNA polymerase served as internal controls. Panel A, the $\alpha$
and β' subunits of core RNA polymerase are relatively constant across time points. σ^{70} keeps increasing after induction and gradually reaches a maximum expression level after 30 minutes, while the maximum detected σ^{32} protein level appears around 15 minute after induction and gradually decreases. Panel B shows the fold changes of two sigma factors after normalized to signal values of the core RNA polymerase subunits. Signal intensities are determined using ImageQuant version 5.2 software. Error bars represent standard deviation in three different experiments.

**Fig. 2.** Evidence indicating potential new transcriptional start sites in three previously predicted operons. Transcription profiling in both σ^{32} overexpression and heat shock assays show that the transcription level of second genes (macB, yceP, ybbN) in the respective operons, has been highly induced while the transcription level of first genes (macA, dinl and ybbO, respectively) in the same operons exhibits no change or a slight decrease. This indicates that new transcription start sites maybe exist just upstream of macB, yceP and ybbN. X-axis indicated different experiments. Solid arrows represent previously predicted transcription start sites, dashed arrows represent potential new transcription start sites.

**Fig. 3.** Electrophoretic mobility-shift assays to test the binding of σ^{32} holoenzyme to DNA fragments carrying putative promoter elements. Panel A, SDS-PAGE gel show purified core RNA polymerase and σ^{32} as well as MultiMark MW Standard. Panel B, potential σ^{32} consensus binding sites of each gene are predicted and aligned. Panel C: native gel shift assays are performed at three different molar ratios (1:0, 1:2.5 and 1:5) of core enzyme to σ^{32} protein with each DNA sample. The upstream sequence of fliL served as a negative control.

**Fig. 4.** Temporal changes of the number of up-regulated genes and their transcript abundance across time points. Panel A, The table shows the numbers of up-regulated genes (≥2 fold) in different function groups. Panel B, transcriptional abundance changes of the genes in relevant function groups. The Y-axis is log2 intensity that represents the transcriptional level and the X-axis is different time-points (0, 5, 10, 15 minutes) after induction. Panel C, Dynamic changes of total σ^{32} protein level, soluble σ^{32} protein level and the number of up-regulated genes after induction. Comparative analysis showed that transcription of σ^{32}-dependent genes does not strictly parallel σ^{32} level.

**Fig. 5.** Temporal changes of the σ^{32} protein and transcript abundance of two σ^{32}-dependent genes in wild type and ΔDnaK strain across time points. The induced σ^{32} protein level is slightly higher in ΔDnaK strain than in wild type and the temporal decrease of the transcriptional level of lon and grpE is reduced in ΔDnaK strain.

**Fig. 6.** Comparison of gene expression levels measured by microarray and RT-PCR approaches. Gene expression level measured by RT-PCR from σ^{32} overexpression samples at different time points is plotted against corresponding microarray data value.

**Fig. 7.** Determination of the σ^{32} consensus binding site. σ^{32} related two-block promoter element is aligned using Bioprospector (41) from the upstream sequence of upregulated genes in σ^{32} overexpression data and displayed using SEQUENCE LOGO (47). The height of each column reflects the non-random bias of particular residues at that position, the size of each residue letter reflecting its frequency at that position.

**Fig. 8.** Comparison of the number of upregulated genes in heat shock experiment and in σ^{32} overexpression experiment. The Venn diagram shows the σ^{32} regulon is overlapping but not identical to the heat-shock stimulon. Boxed values represent fold increase of the four sigmas at 5 or 10 minutes after σ^{32} induction or 10 minutes after heat shock. The total number of regulated genes in each experiment is in parentheses.
| bnumber | Gene | Product | Function                                    | 5 min  | 10 min | 15 min |
|---------|------|---------|---------------------------------------------|--------|--------|--------|
| b3686   | ibpB | small heat shock protein                     | Adaptations, atypical conditions            | 20.5*  | 10.1   | 6.2    |
| b3687   | ibpA | small heat shock protein                     | Adaptations, atypical conditions            | 12.9   | 9.5    | 5.1    |
| b1379   | hslJ | heat shock protein hslJ                      | Adaptations, atypical conditions            | 9.7    | 7.3    | 3.9    |
| b1829   | htpX | heat shock protein, integral membrane protein| Adaptations, atypical conditions            | 5.2    | 2.6    | 1.4    |
| b3989   | htrC | heat shock protein htrC                      | Adaptations, atypical conditions            | 2.1    | 2.6    | 2.7    |
| b2592   | clpB | ATP-dependent protease, part of multi-chaperone system with DnaK, DnaJ, and GrpE | Degradation of proteins, peptides           | 16.6   | 11.3   | 6.6    |
| b3931   | hslU | ATPase component of the HslUV protease, also functions as molecular chaperone | Degradation of proteins, peptides           | 10.2   | 7.1    | 4.6    |
| b3932   | hslV | peptidase component of the HslUV protease    | Degradation of proteins, peptides           | 6.9    | 3.9    | 3.1    |
| b4175   | hflC | with HflK, part of modulator for protease specific for FtsH phage lambda cII repressor | Degradation of proteins, peptides           | 5.4    | 4.0    | 3.1    |
| b3178   | hflB | ATP-dependent zinc-metallo protease          | Degradation of proteins, peptides           | 4.4    | 3.2    | 2.1    |
| b0438   | clpX | ATPase, chaperone subunit of serine protease | Degradation of proteins, peptides           | 2.8    | 2.8    | 1.1    |
| b1779   | gapA | glyceraldehyde-3-phosphate dehydrogenase A  | Energy metabolism (carbon)                  | 2.1    | 2.0    | 1.8    |
| b3179   | ftsJ | 23 S rRNA methyltransferase                  | Cell division                               | 4.5    | 2.9    | 2.9    |
| b0473   | htpG | chaperone Hsp90, heat shock protein C 62.5   | Chaperones                                  | 31.8   | 18.9   | 20.4   |
| b4143   | mopA | chaperone Hsp60 with peptide-dependent ATPase activity, affects cell division | Chaperones                                  | 7.6    | 6.4    | 4.1    |
| b4142   | mopB | chaperone Hsp10, affects cell division       | Chaperones                                  | 4.8    | 4.9    | 3.5    |
| b0014   | dnaK | chaperone Hsp70 in DNA biosynthesis/cell division | Chaperones                                  | 4.8    | 4.3    | 3.1    |
| b0015   | dnaJ | heat shock protein (Hsp40), co-chaperone with DnaK | Chaperones                                  | 2.3    | 2.3    | 1.5    |
| b0439   | lon | DNA-binding ATP-dependent protease La; heat shock K-protein | Global regulatory functions                | 3.6    | 2.2    | 1.6    |
| b2614   | grpE | Hsp 24 nucleotide exchange factor            | Phage-related functions                      | 10.5   | 8.3    | 5.7    |
| b3401   | yflI | heat shock protein 33, redox regulated chaperone | Unknown                                     | 5.7    | 3.9    | 2.5    |
| b3913   | b3913 | periplasmic repressor of cpx regulon by interaction with CpxA | Unknown                                     | 2.2    | 1.3    | 1.7    |

*a* It is possible that one gene has several different gene names.

*b* Numbers indicate fold change relative to pre-σ<sup>32</sup> induction.
Table 2: New candidate genes for the $\sigma^{32}$ regulon

| bnumber | Gene | Product                                                                 | 5 min | 10 min | 15 min |
|---------|------|-------------------------------------------------------------------------|-------|--------|--------|
| b1060   | yceP | conserved hypothetical protein                                          | 17.9* | 8.8    | 4.3    |
| b1380   | ldhA | fermentative D-lactate dehydrogenase, NAD-dependent                    | 12.6  | 8.6    | 4.8    |
| b0879   | macB | macrolide transport protein                                             | 11.6  | 5.0    | 3.6    |
| b3635   | mutM | formamidopyrimidine DNA glycosylase                                     | 10.1  | 3.9    | 2.7    |
| b0492   | ybbN | putative thioredoxin protein                                            | 9.6   | 4.0    | 3.6    |
| b1321   | yjX  | conserved hypothetical protein                                          | 9.3   | 4.2    | 3.1    |
| b3400   | yFH  | heat shock protein 15, DNA/RNA-binding                                 | 8.6   | 5.9    | 3.7    |
| b3498   | pFC  | oligopeptidase A                                                       | 8.3   | 4.0    | 2.5    |
| b1322   | yjF  | putative membrane protein                                               | 7.5   | 5.0    | 2.1    |
| b4140   | b4140| suppress F exclusion of bacteriophage T7                                | 7.4   | 5.2    | 2.6    |
| b0631   | ybeD | conserved hypothetical protein                                          | 7.1   | 4.9    | 2.4    |
| b1664   | b1664| putative enzyme                                                         | 6.2   | 4.5    | 1.9    |
| b0966   | yccV | conserved hypothetical protein                                          | 5.7   | 3.3    | 2.3    |
| b3594   | yibA | putative lyase                                                          | 5.7   | 8.1    | 2.1    |
| b1114   | mfd  | transcription-repair ATP-dependent coupling factor                     | 5.3   | 2.8    | 2.2    |
| b0210   | yafE | putative methyltransferase                                              | 5.1   | 3.0    | 2.0    |
| b1280   | ycm  | conserved hypothetical protein                                          | 4.9   | 3.0    | 1.8    |
| b0660   | ybeZ | putative phosphate starvation-inducible protein, ATP-binding           | 4.7   | 3.2    | 2.5    |
| b1341   | b1341| conserved protein                                                       | 4.6   | 2.4    | 1.3    |
| b3497   | yhiQ | putative methyltransferase                                              | 4.5   | 3.1    | 1.9    |
| b3279   | yrdA | putative acyl transferase, ferricyochelin-binding                       | 4.4   | 2.3    | 1.3    |
| b0058   | yabO | 23S rRNA pseudouridylate 746 synthase                                  | 4.2   | 3.1    | 1.6    |
| b4714   | hflK | with HflC, part of modulator for protease specific for FtsH cII repressor| 4.2   | 2.7    | 1.6    |
| b1279   | yciS | conserved hypothetical protein                                          | 4.1   | 2.3    | 1.5    |
| b4208   | cycA | D-alanine/D-serine/glycine transport protein (APC family)              | 4.1   | 2.0    | 1.4    |
| b1274   | topA | DNA topoisomerase type I, omega protein                                 | 4.1   | 2.7    | 1.4    |

* It is possible that one gene has several different gene names.

Numbers indicate fold change relative to pre- $\sigma^{32}$ induction.
Figure 1:

A. 

B. 

Time after induction (minute):

- 0 5 10 15 20 40

$\beta^*$

$\sigma^{32}$

Time after induction (minute):

- 0 5 10 15 20 25 30 40 60

$\sigma^{70}$

$\alpha$

Fold change

- 0 1 2 3 4 5 6 7 8 9 10

Time after induction (minute):

- 0 5 10 15 20 40

Sigma32

Sigma70
Figure 2:

Predicted operon name: 184

- **Orientation (+)**: macA → macB
  - (918431-919570)
  - (919573-921516)

Predicted operon name: 1514

- **Orientation (-)**: yceP → dinl
  - (1119924-1120170)
  - (1129865-1130710)

Predicted operon name: 1624

- **Orientation (-)**: ybbN → ybbO
  - (516648-517339)
  - (517564-518373)
Figure 3:

A. 

B. 

macB1  ccgctttgaaagttcctcc---------cgaaccccatggtt
macB2  agggctttgaaacgggcagttggttgagcccccaaccag
yceP1   aatccttaaaacgagaaattatatctttctcccccttataa
yceP2   cccctttaaata----------ccgctaatagctttactta
ldhA    agttgttttaatat--------ttcaatatcgccatagtct
mutM1   gttttttttttata----------gtttcccccatattga
mutM2   tggacgtttaaaattcgatcccagtttccgcacgt
ybbN1   accagtcgaatccggg--catcgccgccggttacgct
ybbN2   cggaccttttgcgttcaaacggccagcccttctgta
Consensus  ttgaa  ccccat

C.

(native gel 1)  (native gel 2)
Figure 4:

A.

|                          | 5 min | 10 min | 15 min |
|--------------------------|-------|--------|--------|
| **Number of up-regulated genes** |       |        |        |
| Chaperones               | 9     | 9      | 7      |
| Protease                 | 10    | 7      | 5      |
| Other heat shock related genes | 7     | 7      | 5      |
| DNA-replication, repair, modification | 5     | 3      | 2      |
| Energy & intermediary metabolism | 9     | 11     | 4      |
| Not classified            | 32    | 32     | 11     |
| Unknown                  | 35    | 31     | 9      |
| Others                   | 22    | 16     | 8      |

B.

- Chaperones
- Proteases
- Other heat shock related genes

C.

- sig63 protein level changes (fold)
- number of upregulated genes

Discussion and analysis of the data presented in Figure 4.
Figure 5:
Figure 6:

| Gene      | Time (min) | Log ratio | R²     |
|-----------|------------|-----------|--------|
| rpoA      | 5, 10, 15  |           |        |
| rpoH      | 5, 10, 15  |           |        |
| grpE      | 5, 10, 15  |           |        |
| Lon       | 5, 10, 15  |           |        |
| DnaK      | 5, 10, 15  |           |        |

The regression line is given by:

\[ y = 0.984x + 0.236 \]

\[ R^2 = 0.972 \]
Figure 7: sigma32 consensus binding site
Figure 8:
The global transcriptional response of Escherichia coli to induced sigma protein involves sigma regulon activation followed by inactivation and degradation of sigma in vivo

Kai Zhao, Mingzhu Liu and Richard R. Burgess

*J. Biol. Chem.* published online March 9, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M500393200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts