JOURNAL OF BACTERIOLOGY, Aug. 1993, p. 5242–5252
0021-9193/93/165242-11$02.00/0

Vol. 175, No. 16

Characterization of Twenty-Six New Heat Shock Genes of *Escherichia coli*†

SHUANG-EN CHUANG AND FREDERICK R. BLATTNER*

Laboratory of Genetics, University of Wisconsin—Madison, 445 Henry Mall, Madison, Wisconsin 53706

Received 8 February 1993/Accepted 10 June 1993

Most organisms respond to heat by substantial alteration of the pattern of gene expression. This has been particularly well studied with *Escherichia coli* although the response has by no means been completely characterized. Here we report the characterization of 26 new heat shock genes of *E. coli*, termed *hsl*, discovered by global transcription analysis with an overlapping lambda clone bank. We have measured the molecular weights of the corresponding heat shock proteins and mapped each of them to within a few kilobases on the *E. coli* genome. In vitro, 16 of them can be activated by the $E\sigma^{32}$ RNA polymerase, which specifically transcribes heat shock genes. In vivo expression kinetics of seven of eight examined new proteins were found to be similar to those of the four most studied heat shock proteins, DnaK, DnaJ, GroEL (MopA), and GroES (MopB). In the course of this work, we confirmed that the catalytic subunit of the ATP-dependent Clp protease (also known as Ti protease), ClpP, is derived from a larger precursor protein. Possible assignments of some of the *hsl* genes to known proteins are discussed.

The heat shock response is an important adaptation enhancing survival of organisms at high temperature in which a battery of genes is activated by heat. In *Escherichia coli*, these genes define the heat shock stimulon (33). The stimulated genes include those coding for chaperones which can protect cellular proteins from thermal denaturation and proteases which degrade damaged proteins in the cell. In *E. coli*, more than 20 heat shock genes are known, of which 15 have been genetically mapped and functionally studied. Others are still known only as protein spots on two-dimensional protein gels (15, 33). Some of the *htr* (high-temperature requirement) genes have been found to be heat shock genes, for example, *htrA* (26), *htrC* (34), and *htrM* (35). Most of the genes of the heat shock stimulon in *E. coli* are under the control of the transcriptional initiation factor σ^{32}, encoded by *rpoH*. The RNA polymerase holoenzyme $E\sigma^{32}$ recognizes promoters on heat shock genes which are very different in sequence from the general promoters recognized by the holoenzyme containing σ^{70}, encoded by *rpoD* (10). Another class of promoters, controlled by σ^{E}, are induced only at very high temperatures (12). These include the *htrA* promoter and the third promoter (P3) of the *rpoH* gene.

Several heat shock proteins have been studied extensively. For example, *groEL* (*mopA*), *dnaK*, and *htpG* of *E. coli* are the genes which code for the major chaperones Hsp60, Hsp70, and Hsp90. The *clpP* (22, 40), *lon* (23), and *htrC* (34) genes are all heat shock genes with roles in proteolysis. DnaK, DnaJ, and GrpE and of course RpoH are involved in the autoregulation of heat shock response (11, 14, 25, 39).

Despite this extensive information, the inventory of heat shock genes and proteins is by no means complete. To this end, we have employed a global transcription mapping technique to simultaneously detect, clone, and map genes of *E. coli* that respond to heat by increased or decreased transcription (9). This technique employs hybridization to ordered λ clones on membrane DNA dot blots. In this communication, we have extended the analysis by examining the proteins produced by the active λ clones in response to heat shock. This was done both in vivo by the UV-irradiated host protein expression system (38) and in vitro with and without the addition of $E\sigma^{32}$. This allows the determination of which genes are members of the *rpoH* heat shock regulon. Finally, we used Southern blot analysis to map these heat shock genes more precisely to the genomic restriction map. In this study, we tested only those clones which were globally detected previously (9). These do not include clones which contain *htrA*, *htrC*, *htrM*, or *rpoH*.

MATERIALS AND METHODS

Strains and materials. *E. coli 159 uvrA* (λ*ind*) was from our laboratory stock (FB no. 418). The *E. coli* miniset λ phages of Kohara et al. (21) were obtained from F. C. Neidhardt. [^{35}S]methionine (800 Ci/mmol) and [α-^{32}P]dCTP (3,000 Ci/mmol) were purchased from Amersham. Heat shock-specific RNA polymerase holoenzyme $E\sigma^{32}$ was kindly provided by C. A. Gross. Avian myeloblastosis virus reverse transcriptase was purchased from Life Sciences Inc. MOPS minimal medium (32) contains 10 mM NH_4Cl, 1.32 mM potassium phosphate, 0.53 mM $MgCl_2$, 0.28 mM K_2SO_4, 0.01 mM $FeSO_4$, 0.5 mM $CaCl_2$, 50 mM NaCl, 4 mM Tricine, 40 mM MOPS (3-[*N*-morpholino]propane-sulfonic acid [pH 7.2]), trace elements, and appropriate carbon sources. NZC medium contains 1% NZ amine (Sheffield Products), 0.5% NaCl, 0.1% Casamino Acids, and 8 mM $MgSO_4$. Primary λ phages were grown by the plate lysate method on an NZC plate, according to Maniatis et al. (29). The plate lysates were extensively dialyzed against phage dialysis buffer (0.3 M NaCl, 0.1 M Tris [pH 7.9]) before they were used for expression in the UV-irradiated host system. The titers were generally about 10^{11}/ml.

Preparation of λ clone DNAs for directing in vitro protein synthesis. To avoid using RNase and DNase to remove bacterial RNA and DNA from phage lysates, we used DEAE-cellulose (preswollen anion exchanger DE52, from Whatman Biosystems, Ltd.). DE52 was first washed three

* Corresponding author. Electronic mail address: blattner@wiscmacc.bitnet.

† Paper no. 3362 from the University of Wisconsin, Laboratory of Genetics.

times with water, three times with 150 mM NaCl–1 mM EDTA–10 mM Tris (pH 8.1), and twice with NZC medium and made to 50% slurry in NZC. Two volumes of the 50% DE52 slurry was added to each volume of phage plate lysate, incubated at room temperature for 10 min, mixed at intervals, and spun in a microcentrifuge for 2 min. The supernatant was extracted with phenol-chloroform twice and with chloroform once. Clone DNA was then precipitated with ethanol and resuspended in water. DNA concentration was estimated by comparison with λ DNA standards in an ethidium bromide-stained agarose gel after electrophoresis.

Expression of heat shock-inducible proteins in vivo. Differential expression of heat shock-inducible proteins from λ clones was measured by infecting UV-irradiated host cells, *E. coli 159 uvrA* (λ*ind*), with and without heat shock pretreatment. The host was grown in 15 ml of MOPS minimal medium with 0.4% maltose and 0.2% glycerol at 37°C to mid-log phase. Heat shock pretreatment was at 50°C for 7 min. Cells were then subjected to UV irradiation of 1,000 J m^{-2} (the bacterial culture was held in a petri dish 6 cm below the germicidal UV lamp and irradiated for 30 s with constant shaking). This destroys the host's DNA completely so that the only suitable template for transcription is the subsequently infecting phage DNA. Irradiated cells were then spun down and resuspended in 3 ml of the same medium, and 0.1 ml was dispensed to each Eppendorf tube and was infected with 10^9 dialyzed λ phage in 20 μl and incubated for 20 min. Newly synthesized proteins were labeled by addition of 25 μCi of [^{35}S]methionine. After 30 min of labeling, 0.1 ml of chasing solution (0.2% methionine in MOPS medium) was added, incubated for 5 min, and then chilled on ice for 10 min. Cells were pelleted and resuspended in 20 μl of loading buffer (20% glycerol, 2% sodium dodecyl sulfate [SDS], 5% β-mercaptoethanol, 0.1% bromophenol blue, 60 mM Tris [pH 6.8]) and incubated in a boiling water bath for 3 min before SDS-polyacrylamide gel electrophoresis (PAGE) analysis.

Expression of heat shock-inducible proteins in vitro. The in vitro-coupled transcription-translation reactions were conducted in 15 μl with 0.4 μg of each phage DNA. In parallel reactions, 0.5 μg (in molar excess over the phage DNA input) of the heat shock RNA polymerase holoenzyme $E\sigma^{32}$ was added in addition to the general $E\sigma^{70}$ present in the kit (Amersham Corp.). Proteins synthesized were labeled with 25 μCi of [^{35}S]methionine for 60 min at 37°C. Protein was precipitated by adding 15 μl of 20% trichloroacetic acid and washing twice with ethanol containing 100 mM Tris, pH 8.1. The pellet was dissolved in 40 μl of loading buffer, and protein products were resolved by SDS-PAGE. Five microliters was loaded in each lane. Protein synthesis was not observed without DNA input (data not shown).

Processing of the 24-kDa protein product of the λ clone 148. The same UV-irradiated host system procedure as described above was used, except that labeling was continued for only 2 min, followed by chasing with an excess amount of nonradioactive methionine (final concentration, 0.2%) for various lengths of time (up to 30 min) before harvesting.

Kinetics of heat shock response. The same UV-irradiated host system procedure was used, except that cells were grown at 30°C and heat shock pretreatment of the host cells was at 42°C for various lengths of time (up to 60 min) before UV irradiation.

Localization of the heat shock genes by Southern blot analysis of the heat shock clones. One microgram of each phage clone DNA was digested with various restriction endonucleases and resolved in 1% agarose gels. DNA in the gel was denatured in 0.5 M NaOH–1.5 M NaCl at room temperature for 30 min. Capillary Southern blotting was conducted in 20 mM NaOH–1 M NaCl, DNA was blotted onto the GeneScreen Plus nylon membrane (DuPont) overnight, and blots were neutralized and rinsed in 1× SSC (0.15 M NaCl, 15 mM Na citrate, pH 7.2) and air dried at 65°C for 2 h before use. [^{32}P]dCTP-labeled cDNA probes were synthesized with avian myeloblastosis virus reverse transcriptase (50 U) from 20 μg of total RNA isolated from control cells (grown in MOPS minimum medium with 0.4% glucose as the carbon source at 37°C) or from total RNA isolated from cells which were heat shocked for 7 min at 50°C. Three micrograms of random hexamers (dpN_6, Pharmacia) was used as primers. [^{32}P]cDNA was denatured, and the RNA templates were simultaneously degraded by addition of a final concentration of 50 mM EDTA and 0.25 M NaOH and incubated at 65°C for 30 min. After neutralization with HCl and Tris buffer, cDNA probe was added directly to the blot in the hybridization bag. Standard hybridization procedure (29) was followed. The same blot was used for hybridization to both control and heat shock [^{32}P]cDNA with an alkaline stripping to remove the probe in between.

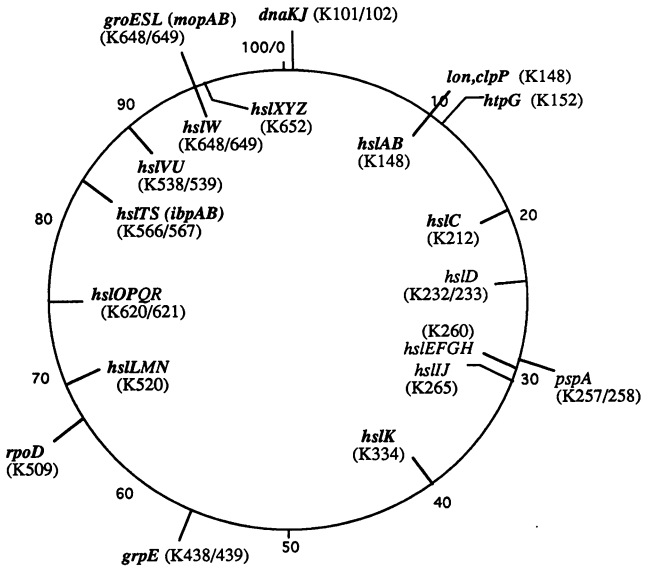

FIG. 1. Circular map of the heat shock genes. Genes outside the circle are previously known; those inside are previously uncharacterized. The new ones are each given a serial name, *hsl*, for heat shock locus. Genes activated by $E\sigma^{32}$ are in boldface type. Corresponding Kohara clones, beginning with the letter K, are shown in parentheses.

RESULTS

Detection of new heat shock genes. Previously, we have detected more than a dozen chromosomal regions which were induced by a heat shock treatment (9). These regions have been characterized in more detail as summarized in Fig. 1 as a circular map with new heat shock genes inside the circle and those previously known outside. We have termed the new genes *hsl* for heat shock locus.

A total of 10 of the 11 previously known heat shock proteins examined were detected. LysU, the heat shock-inducible lysyl-tRNA synthetase II, encoded at min 93.8 and residing in both Kohara clones 646 and 647, was not detected

by this technique (data not shown). A missing metabolic signal, as suggested by Matthews and Neidhardt (30), in the UV-irradiated host system probably was the cause. It is also possible that neither of the clones contains an intact *lysU* gene.

Expression of new heat shock proteins. Expression of heat shock proteins from λ clones was performed both in vivo in the UV-irradiated host system and in vitro in the coupled transcription-translation system. To verify the experimental systems, five of the known heat shock proteins were tested in the UV-irradiated host system as shown in the top panel of Fig. 2, from clones 101, 152, 438, and 509. Their corresponding genes are *dnaK* and *dnaJ* (coding for proteins of 69 and 39 kDa, respectively) from clone 101. Clone 102 also expressed *dnaKJ* (data not shown). The *htpG* gene codes for a 70-kDa protein (first identified and sequenced by Bardwell and Craig [4]) from clone 152. The *grpE* gene codes for a 26-kDa protein from clone 438. The sequence-calculated molecular mass is 21,668 Da (27). Clone 439 also expressed GrpE. The *rpoD* gene codes for the major transcriptional initiation factor σ^{70}, from clone 509.

The following 12 clones were examined for unknown heat shock genes both in vivo and in vitro. Clone 148 (at min 10) produced four proteins stimulated by heat shock in vivo, as well as by $E\sigma^{32}$ in vitro. The sizes of these proteins are 89, 65, 46, and 24 and 22 kDa. The 89-kDa protein matches well with the size of the Lon protease. The 65- and 46-kDa proteins were previously unknown. The stimulated 22-kDa in vivo and 24-kDa in vitro proteins are mature and precursor proteins of ClpP, respectively (see below). ClpP is the catalytic subunit of the ATP-dependent Clp protease (also known as Ti protease, composed of two subunits of ClpA and between 10 and 12 subunits of ClpP [18]) and recently proved to be the heat shock protein HtpL (two-dimensional spot F21.5 [22]). It has been shown that ClpP is a processed protein (31). We confirmed that the 24-kDa protein is a precursor for the 22-kDa protein by the pulse and chase experiment in the UV-irradiated host system shown in Fig. 3. After 5 min of chasing, about one-half of the precursor was converted into the mature 22-kDa protein.

Clone 212 at min 19.2 expressed a heat shock-inducible protein of 80 kDa which was slightly stimulated by $E\sigma^{32}$ in vitro (Fig. 2).

Clone 257 (also clone 258; data not shown) at min 29.2 expressed a 28-kDa protein which was induced by heat shock in vivo but not by $E\sigma^{32}$ in vitro. This protein matches the characteristics of the phage shock protein PspA (induced by infection of the filamentous phage f1 [5, 6, 41]) in four respects: inducibility by heat shock only at an extreme temperature of 50°C, *rpoH* independence, genetic map position, and molecular weight.

Clone 260 (min 29.7) expressed four heat shock-induced proteins of 60, 51, 41, and 39 kDa in vivo. None of them were stimulated by $E\sigma^{32}$ in vitro.

Clone 265 (min 30.6) expressed two heat shock-inducible proteins of 36 and 14 kDa in vivo. While synthesis of the 36-kDa protein was not affected by $E\sigma^{32}$, it is not clear for the 14-kDa protein, because of the strong protein synthesis of the same size in vitro.

Clone 334 at min 39.8 expressed a 49-kDa protein which was induced by heat shock in vivo and also stimulated in vitro by $E\sigma^{32}$.

Clone 520 at min 69.2 expressed three proteins of 70, 31, and 27 kDa which were stimulated by heat shock in vivo. Two of them, the proteins of 70 and 27 kDa, were also stimulated by $E\sigma^{32}$ in vitro.

Clone 538 at min 89 expressed two proteins of 49 and 21 kDa which were stimulated both in vivo and in vitro. Clone 539 also expressed the two heat shock-inducible proteins (data not shown). The two corresponding genes have been sequenced, and an $E\sigma^{32}$-specific promoter is found in front of these genes. They are transcribed counterclockwise and belong to the same operon (8).

Clones 566 and 567 (at min 83) were found previously to be the strongest heat shock-induced clones (9). Indeed, synthesis of a 15-kDa protein was extraordinarily stimulated both in vivo and in vitro from clone 566 as shown in Fig. 2. Clone 567 also expressed this heat shock-inducible protein (data not shown). We have sequenced this region. It is a bicistronic operon controlled by an $E\sigma^{32}$-specific promoter. Both genes specify a protein of 15 kDa; the two proteins show 48% identity to each other in amino acid sequences and are transcribed counterclockwise (8). While we were preparing this article, Allen et al. (1) also reported these sequences and their roles in forming the inclusion bodies in response to heterologous protein expression. The *hslS* and *hslT* genes are the same as the *ibpB* and *ibpA* genes, respectively.

Clone 620 (min 75) expressed four proteins of 33, 30, 24, and 18 kDa which were stimulated by heat shock in vivo, and two of them, the 33- and 30-kDa proteins, were also stimulated in vitro by $E\sigma^{32}$. The adjacent clone 621 expressed all of them but the 24-kDa protein (data not shown).

Clone 649 at min 94.2 contains one of the most studied heat shock operons, *groESL* (*mopBA*), which codes for proteins of 60 and 10.7 kDa, respectively (17). The 60-kDa GroEL (MopA) is an *E. coli* analog of HSP60 or chaperone 60. The apparent molecular mass of GroES (MopB) estimated by SDS-PAGE has been around 15 kDa. Both are strongly induced heat shock proteins and are induced in an *rpoH*-dependent manner. We confirmed this both in vivo and in vitro. Surprisingly, an additional new protein of 22 kDa was also found to be stimulated both in vivo and in vitro. Another point worth mentioning is that many bands trailing down the GroEL (MopA) band were also found to be stimulated upon induction. Whether they were degradation intermediates of the major heat shock proteins was unclear. Clone 648 also expressed the same three heat shock-inducible proteins (data not shown).

Clone 652 at min 94.8 expressed three proteins of 51, 45, and 37 kDa, respectively. All three were stimulated both in vivo and in vitro. Likely candidates for these proteins are HflX, HflK, and HflC, of 50-, 46-, and 38-kDa apparent molecular masses, respectively (2). HflK and HflC are subunits of a multimeric protease which specifically degrades λ phage protein CII. The function of HflX is unclear (7).

Kinetics of the heat shock response. Heat shock genes are activated to a maximal level in about 5 min after an upshift in temperature, and then after 10 to 20 min, the response fades to a new level which is only slightly higher than the level before temperature shift.

We compared the characteristics of two well-studied heat shock operons (*dnaKJ* and *groESL*) with those of two of the newly detected heat shock loci by the differential expression method in the UV-irradiated host system. The treatment with heat shock was different from that described for Fig. 2 in that the host cells were grown at 30°C and then heated to 42°C for various lengths of time as is normally done to see the transient response. As shown in Fig. 4, responses all peaked at 5 min after heat treatment for the genes in clones 101 and 648, *dnaKJ* and *groESL* (*mopBA*), and the gene for the 22-kDa protein in clone 648. The same kinetic results

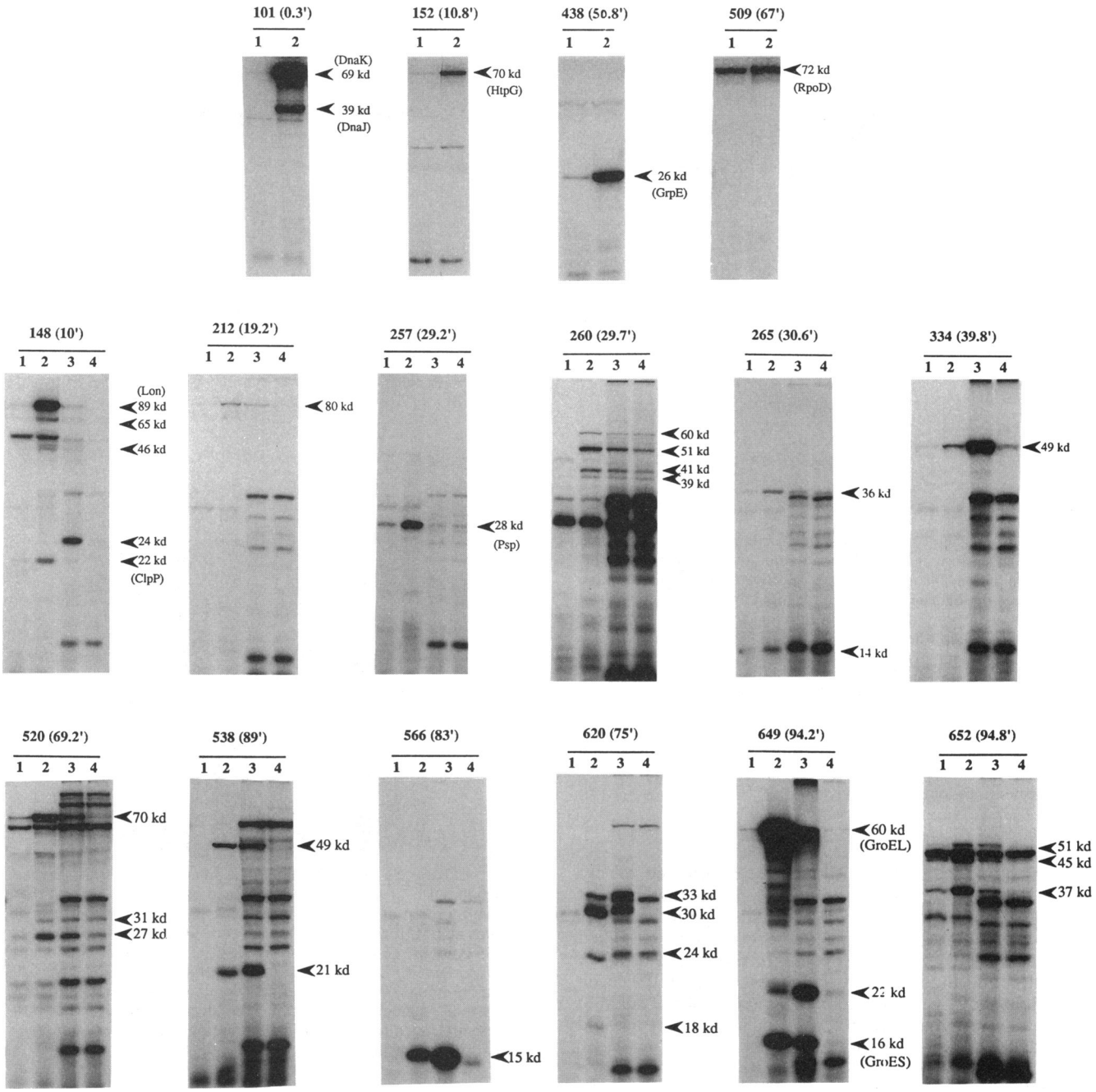

FIG. 2. Detection of heat shock proteins. Differential expression was conducted in vivo in the UV-irradiated host system and in vitro in the coupled transcription-translation system (see Materials and Methods). Numbers on top of each gel are the clone names of the miniset in the Kohara library. Numbers in parentheses are the map position of the midpoint of each clone (in minutes). Lanes 1 and 2 show expression in vivo without and with a heat shock pretreatment of the host cells, respectively. Lanes 3 and 4 are expression in vitro with and without addition of $E\sigma^{32}$, respectively. Clones 101, 152, 438, and 509 (top panel, containing the known heat shock genes *dnaKJ*, *htpG*, *grpE*, and *rpoD*, respectively) were analyzed only in vivo. Names of known protein products are given in parentheses, accompanying their corresponding molecular sizes.

were observed for seven of the eight newly discovered proteins in clones 260 ($E\sigma^{32}$ independent) and 620 ($E\sigma^{32}$ stimulated). Synthesis of the 60-kDa protein from clone 260 decreased only slightly during the heat treatment, in contrast to other proteins tested.

Physical mapping of the heat shock genes. In the genomic library, each clone contains a bacterial DNA insert of between 9 and 21 kb in length. To map the heat shock genes, Southern blot analysis was performed, taking advantage of the availability of the restriction map for each clone. DNA fragments from restriction endonuclease digestion were resolved in agarose gels and blotted onto the GeneScreen Plus

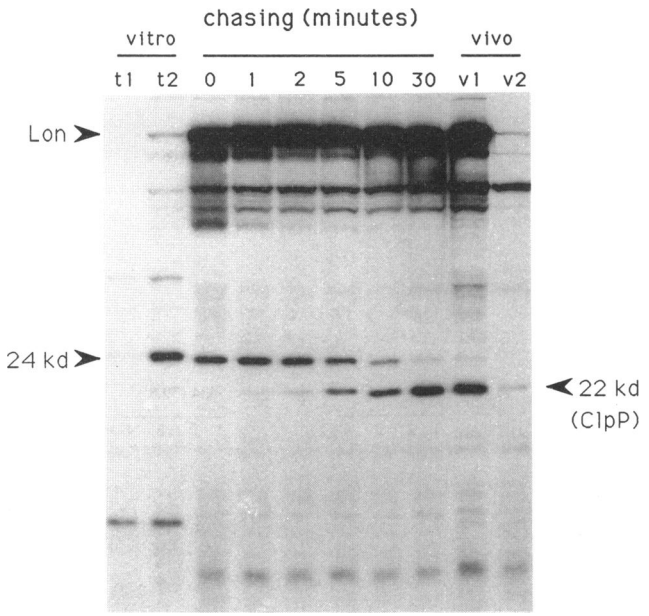

FIG. 3. Processing of the heat shock protein ClpP. Heat shock-activated expression in the UV-irradiated host system was performed with clone 148 as in Fig. 2, except that [^{35}S]methionine labeling was only for 2 min, and then the labeling was chased with an excess amount of nonradioactive methionine for various lengths of time as indicated in the middle six lanes. Lanes t1 and t2 are differential expression in vitro without and with the addition of Eσ^{32}, respectively. Lanes v1 and v2 are in vivo differential expression in the UV-irradiated host system with and without a heat shock pretreatment of the host cells, respectively. In 5 min, about one-half of the 24-kDa protein was processed into 22-kDa protein, which is consistent with the size of the mature ClpP protein. Almost all ClpP synthesized in vitro is the 24-kDa precursor, whereas the 22-kDa mature protein predominates in vivo.

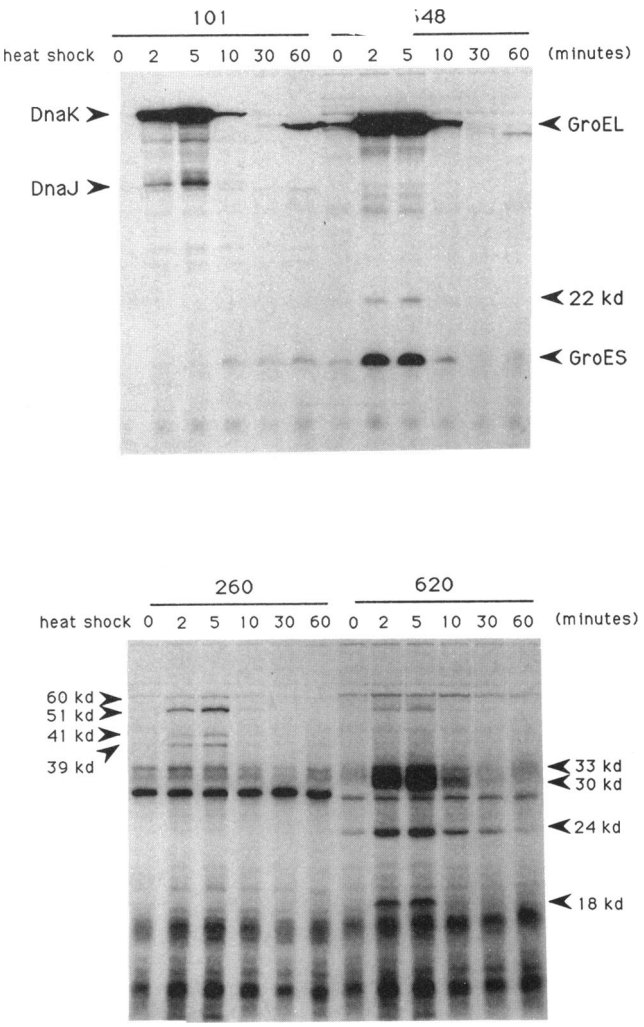

FIG. 4. Kinetics of heat shock response. Expression in the UV-irradiated host system was performed as in Fig. 2, except that cells were grown at 30°C and heat shock pretreatment of the host cells was at 42°C for various lengths of time as indicated. Clones 101 and 648 contain the known heat shock genes, the *rpoH*-dependent *dnaKJ* and *groESL* (*mopBA*), respectively. Clones 260 and 620 contain previously uncharacterized heat shock-inducible genes which are Eσ^{32} independent and dependent, respectively. The sixth and seventh lanes from the left of the upper panel were contaminated by overflow from the eighth lane during loading. Numbers below each clone are the duration of heat treatment in minutes at 42°C.

nylon membranes (DuPont). Blots were hybridized to ^{32}P-labeled total cDNA probe of control and to that of heat shock. By comparison of the hybridization patterns, DNA fragments containing the heat shock gene(s) were identified (Fig. 5). Figure 6 graphically presents the localized DNA fragment(s) and those previously known genes within each clone. Heat shock genes of clones 212, 260, and 520 failed to show enhanced hybridization to particular bands and could not be mapped by this method.

This approach was validated by identifying all the previously known and sequenced heat shock genes that we examined here (compare Fig. 5 and 6): the *dnaKJ* operon in the expected 3.8-kb *Bgl*I fragment of clone 101; *lon* in the expected 1.9- and 2.4-kb *Eco*RV-*Hin*dIII fragments of clone 148 (the 1.9-kb fragment also contains *clpX* and part of *clpP*); *htpG* in the expected 2.7-kb *Eco*RI-*Eco*RV fragment of clone 152; the *psp* operon in the expected 4.8-kb *Eco*RI fragment of clone 257; *grpE* in the expected 3.9-kb *Eco*RI-*Eco*RV fragment of clone 439; *rpoD* in the expected 3.2-kb *Pvu*II fragment of clone 509; and the *groE* (*mop*) operon in the expected 2- and 2.7-kb fragments of clone 649 (see below). In clone 101, the 5-kb fragment was probably from a partial digestion of *Bgl*I (the 3.6-kb fragment combined with an adjacent 1.5-kb fragment, Fig. 5 and 6). For the *htpG*-containing clone 152, although there are two *Eco*RI-*Eco*RV fragments of a similar size of 2.7 kb, as shown in Fig. 6, the downstream fragment has no capacity for encoding a protein of 70 kDa. For the *grpE*-containing clone 439, instead of two adjacent fragments of 2.2 and 1.7 kb being identified, as would be predicted from Fig. 6, a single 4.1-kb fragment was observed. Apparently, the *Eco*RV site in the middle was not present in this clone. For the *groE* (*mop*)-containing clone 649, Fig. 6 predicts two 2.6-kb fragments, in contrast to the observed 1.9- and 2.8-kb fragments. But our observation was consistent with the original data of Kohara et al. (21).

The remaining 10 clones, potentially containing new heat shock genes, will be discussed below. For clone 212, we failed to identify any DNA fragment which might contain the 80-kDa-protein-encoding gene.

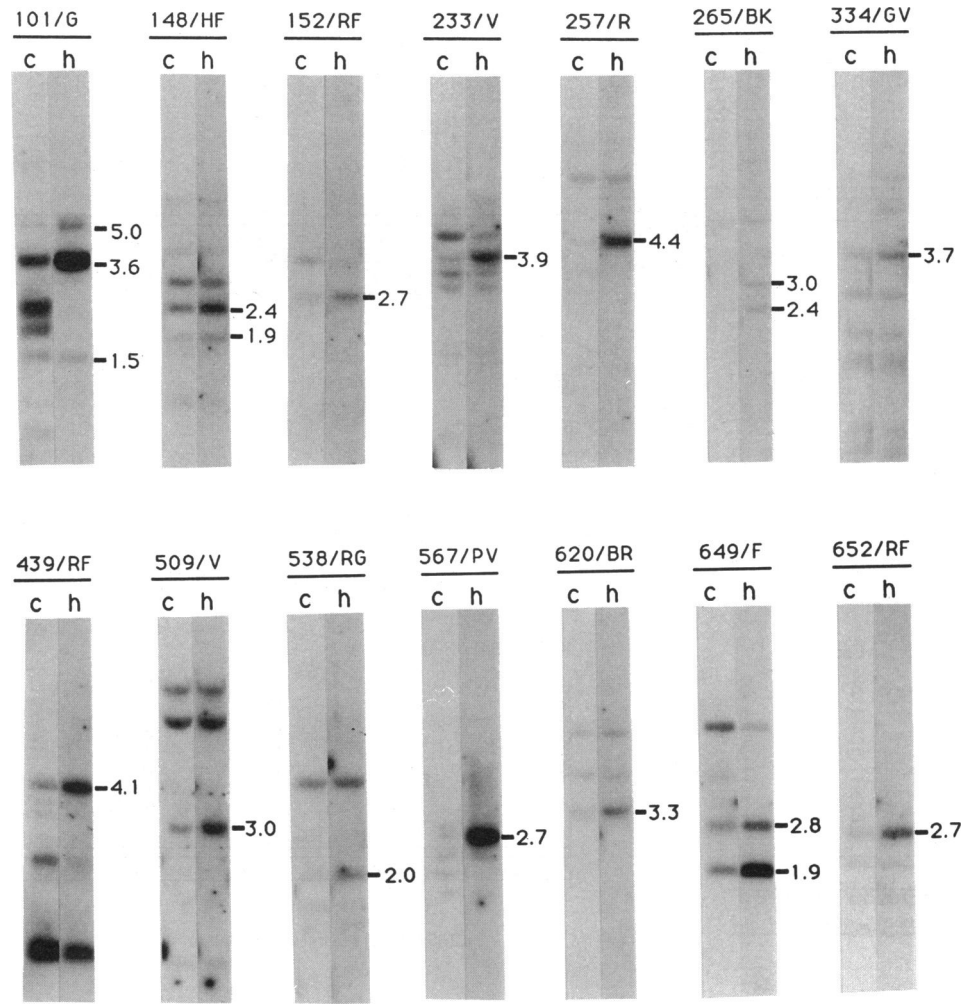

FIG. 5. Localization of the heat shock genes by Southern blot analysis. DNA of each clone was digested with various restriction enzymes as indicated (B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI; F, *Eco*RV; G, *Bgl*I; K, *Kpn*I; P, *Pst*I; and V, *Pvu*II). The transferred Southern blot was hybridized to control total cDNA probe (c) and heat shock total cDNA probe (h). The same blot was used for hybridization to both probes with an alkaline stripping procedure in between. Sizes of the marked DNA bands are in kilobase pairs.

Although clone 233 failed to express heat shock-inducible protein in the UV-irradiated host system (data not shown), transcription was observed to be stimulated upon heat shock in this region (9). Indeed, a 3.9-kb *Pvu*II fragment was identified in the Southern analysis (Fig. 5), which was mapped in Fig. 6. The *htrB* gene has already been shown not to be under the heat shock regulation (19, 20).

The four σ^{32}-independent heat shock-inducible genes in clone 260 were not localized by this analysis. They could be anywhere in the 18-kb blank region counterclockwise of *fnr* as shown in Fig. 6.

Clone 265 lies in the genetically vacant region at 30.6 min, around the defective prophage rac. No genes have been physically mapped in the identified 5.4-kb *Bam*HI-*Kpn*I fragment so far.

For clone 334, a 3.7-kb band was identified (Fig. 5). The closest fragment expected from Fig. 6 would be the 3-kb *Pvu*II fragment as shown. Both the *pabB* and *sdaA* genes code for a matching-size protein of 50 kDa, but they reside in a predicted 5-kb *Bgl*I fragment clockwise of the one indicated.

For clone 520, no DNA fragment was identified. But, as shown in Fig. 6, no sequenced genes in this clone match in size any of the three proteins detected in Fig. 2. Therefore, the new heat shock genes must be between *argG* and *greA*. The *hflB* locus (coding for or regulating a protease that acts on cII protein of phage λ), which is 85% cotransducible with the *argG* gene (3), might be a good candidate.

For clone 538, a 2-kb band was identified in the *Eco*RI-*Bgl*I double digestion; this band was either the *Eco*RI-*Bgl*I fragment or the following *Bgl*I fragment as indicated in Fig. 6. According to the sequence, genes for the 19- and 49-kDa proteins are in the *Bgl*I fragment (data not shown).

For clone 567, a 2.7-kb band was identified in the *Pst*I-*Pvu*II double digestion. Four fragments would be predicted as indicated in Fig. 6. According to the sequence, the two new heat shock genes, both specifying a protein of about 16 kDa, are in the rightmost *Pst*I-*Pvu*II fragment (data not shown).

For clone 620, a 3.3-kb band was identified in the *Bam*HI-*Eco*RI double digestion (Fig. 5), which was mapped in the *Bam*HI-*Bam*HI fragment indicated in Fig. 6. For clone 652,

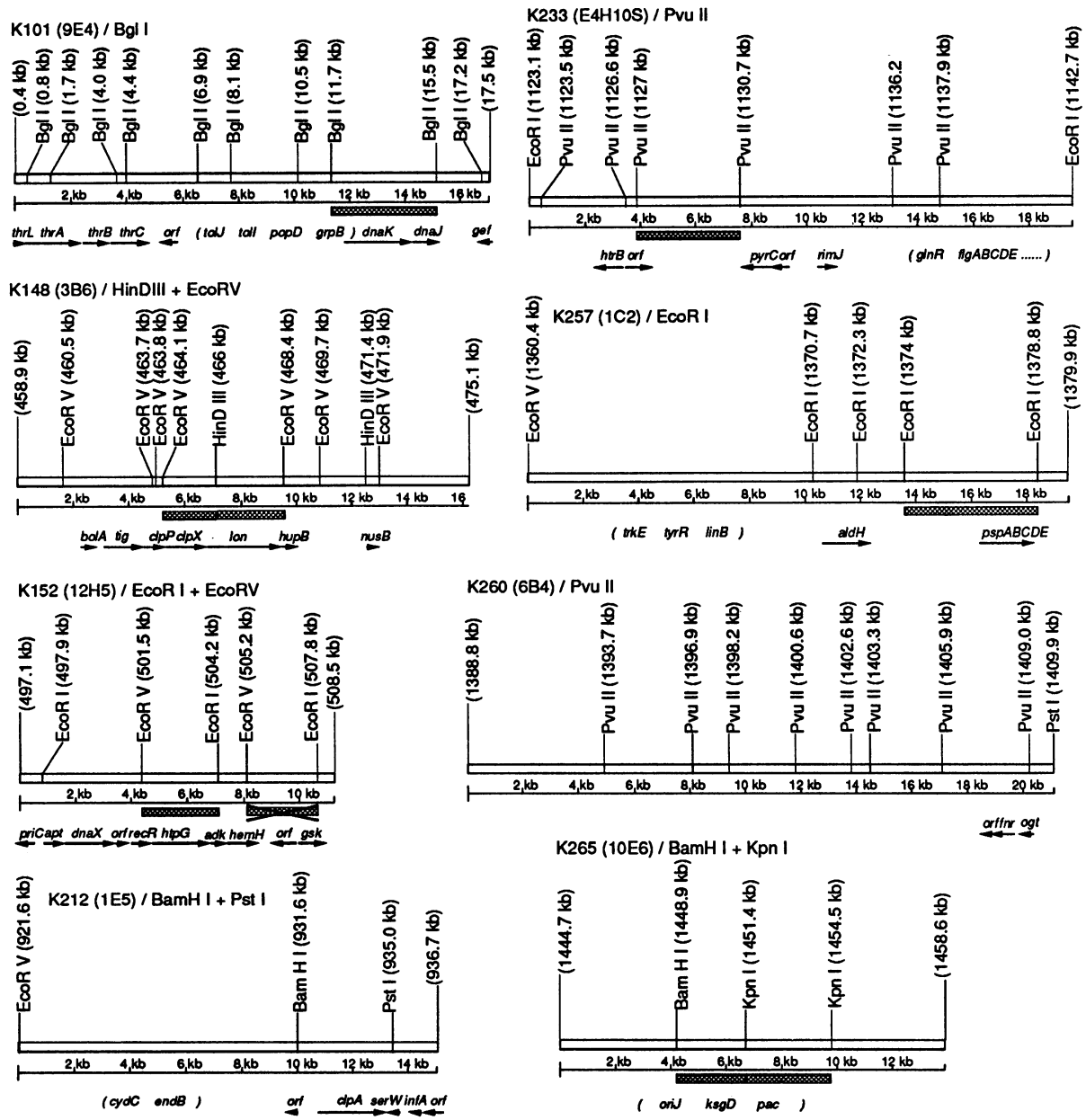

FIG. 6. Restriction maps of the heat shock clones. The consolidated eight-enzyme restriction maps with kilobase coordinates (Rudd et al., version 5 [36, 37]) of the 17 heat shock clones are presented in scale. The original clone names are in parentheses following the miniset series number. The restriction enzymes used for each clone in Fig. 5 are also indicated. The restriction fragments showing a positive increase in hybridization following heat shock are indicated by cross-hatched bars below the maps. The cross-hatched bars that are crossed out (clones K152, K538, and K567) indicate the restriction fragments that are ruled out (see text). The *Eco*RV site at 2,761.9 kb (*, clone K439) was apparently absent as shown by Southern blot analysis in Fig. 5. The 4,400.5-kb *Eco*RV site (*) of clone K649 was located about 0.6 kb closer to the *groE* (*mop*) operon in the original map of Kohara et al. (21), which was more consistent with our Southern blot analysis in Fig. 5. Sequenced genes are shown with their map positions and transcriptional directions. Other genetic markers not precisely located are shown in parentheses.

a 2.7-kb band was identified in the *Eco*RI-*Eco*RV double digestion. This mapped the responding genes in a region containing the *hflXKC* operon as shown in Fig. 6.

Summary of the heat shock genes. Table 1 summarizes the results, including the genetic map position in minutes, the Kohara clones, the apparent molecular mass of each detected protein, whether protein synthesis was stimulated by heat shock in vivo, whether protein synthesis was stimulated by the heat shock-specific RNA polymerase $E\sigma^{32}$ in vitro, assignments of the proteins, and the physical map of the heat shock-inducible genes.

DISCUSSION

We have characterized 36 heat shock-inducible genes produced from 17 chromosomal regions. The total number of

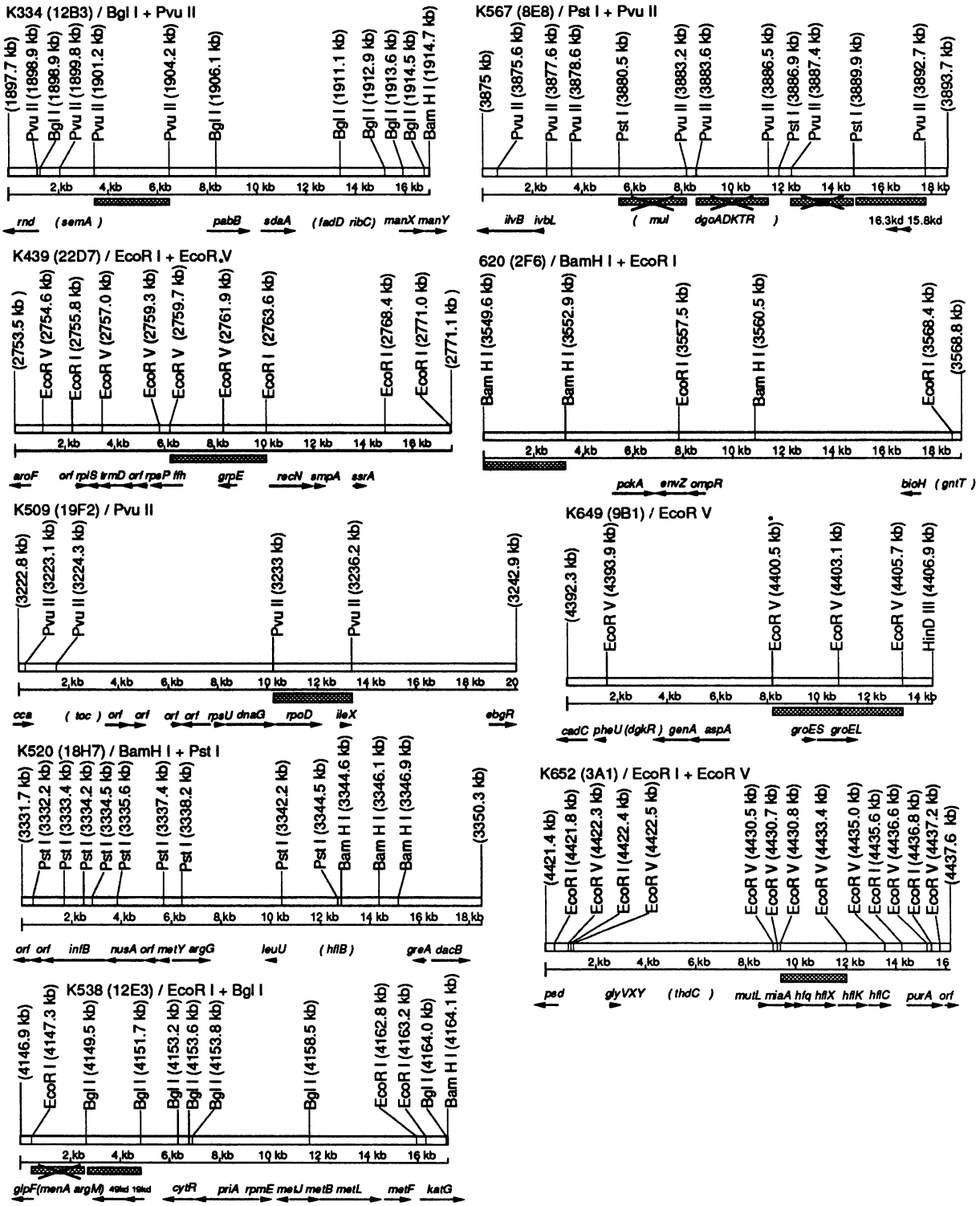

FIG. 6—*Continued.*

heat shock-inducible genes should be even higher since our previous data set covered only 92% of the genome owing to problems with the Kohara set. In addition, some genes may be masked by other genes on the same clone as previously discussed (9). The data suggest that more than two-thirds of them are σ^{32} dependent.

In early two-dimensional protein gel studies, Neidhardt and VanBogelen (33) identified 17 proteins (termed Htp)

TABLE 1. Heat shock-inducible proteins encoded in Kohara λ clones

Map position (min)	Kohara clone(s)	Molecular mass (kDa)[a]	UV[b]	RpoH[c]	Protein(s)[d]	Physical map (kb coordinates)[e]
0.3	101, 102	69	+	¶	DnaK	11.7–15.5
		39	+	¶	DnaJ	
10.0	148	89	+	+	Lon	464.1–468.4
		65	+	+	(HslA)	
		46	+	+	(HslB)	
		22 (24)	+	+	ClpP	
10.8	152	70	+	¶	HtpG	501.5–504.2
19.2	212	80	+	+	(HslC)	921.6–936.7
23.7	232, 233	?	−	ND	(HslD)	1127.0–1130.7
29.2	257, 258	28	+	−	PspA	1374.0–1378.8
29.7	260	60	+	−	(HslE)	1388.8–1409.9
		51	+	−	(HslF)	
		41	+	−	(HslG)	
		39	+	−	(HslH)	
30.6	265	36	+	−	(HslI-HtpH)	1448.9–1454.5
		14	+	?[f]	(HslJ)	
39.8	334	49	+	+	(HslK)	1901.2–1904.2
56.8	438, 439	26	+	¶	GrpE	2759.7–2763.6
67.0	509	72	+	¶	RpoD	3233.0–3236.2
69.2	520	70	+	+	(HslL)	3331.7–3350.3
		31	+	−	(HslM)	
		27	+	+	(HslN)	
75.0	620, 621	33	+	+	(HslO)	3549.6–3552.9
		30	+	+	(HslP)	
		24	+	−	(HslQ)	
		18	+	−	(HslR)	
83.0	566, 567	16.3	+	+	(HslS-IbpB-HtpE)	3889.9–3892.7
		15.8	+	+	(HslT-IbpA-HtpN)	
89.0	538, 539	49	+	+	(HslU-HtpI)	4149.5–4151.7
		21	+	+	(HslV-HtpO)	
94.2	648, 649	60	+	+	GroEL	4400.5–4405.7
		22	+	+	(HslW)	
		16	+	+	GroES	
94.8	652	51	+	+	(HslX)	4430.8–4433.4
		45	+	+	(HslY)	
		37	+	+	(HslZ)	

[a] ClpP synthesized from clone 148 was 22 kDa (in vivo) and 24 kDa (in vitro). The question mark indicates that the protein product for clone 232-233 has not been identified.
[b] Detection by the UV-irradiated host system.
[c] Stimulation by RpoH in vitro. Some proteins (¶) are previously known to be RpoH stimulated and are not repeated here. ND, not determined.
[d] Previously uncharacterized proteins are each given the three-letter name Hsl (heat shock locus) in parentheses.
[e] Rudd's physical kilobase coordinates were adopted (version 5 [36, 37]).
[f] ?, unknown.

whose synthesis was stimulated by heat shock. Using protein samples we supplied, Robert Clark of Neidhardt's group has determined that HtpH is HslI (30.6 min), HtpE is HslS-IbpB (83 min), HtpN is HslT-IbpA (83 min), and HtpO is HslV (89 min). Furthermore, HtpI, whose gene was previously mapped (15), is HslU (89 min) (data not shown). Thus, only one of the 17 Htp proteins, HtpK, at present remains uncharacterized.

Upon heat shock treatment, expression of most genes other than the heat shock ones is decreased. In contrast to the extensive studies on the mechanism of heat shock activation, little attention has been paid to the heat shock repression. As shown previously (9), some chromosomal regions were more repressed than others. Clone 626 was found to be the most repressed by heat shock at 50°C while clone 628 was most repressed at 42°C (9). It is worth noting that the S10 operon promoter (located within clone 628) is down regulated by ppGpp (13) and ppGpp levels are increased during heat shock (28). Also, the stringent response can induce the expression of the heat shock genes (16), and interestingly, depletion of Era protein reduces the expression of the heat shock genes and blocks thermal induction of ppGpp pools (24). It would be interesting to study the heat shock-repressed genes further, especially with respect to the role played by *relA*.

It has been well known that results of in vitro study do not always reflect the situation in vivo. For example, among the σ^{32}-inducible heat shock genes, some were more drastically activated in vivo while some were more stimulated in vitro. While GroES (MopB) synthesis was about equally stimulated in both conditions from clone 649 (Fig. 2), GroEL (MopA) synthesis was much more stimulated in vivo, and synthesis of the uncharacterized 22-kDa protein showed higher stimulation in vitro. Synthesis of Lon from clone 148 was also stimulated much more in vivo whereas the 49-kDa protein of clone 334 was synthesized much more in vitro. Apparently, other factors are involved in fine-tuning the response to a heat shock.

We have demonstrated here that a global approach can be quite powerful compared with the analysis of genes one at a time in studying the regulation of complex systems. Differential expression of heat shock genes by the UV-irradiated

host system was an efficient and sensitive way of identifying their protein products, as was differential Southern hybridization analysis for physically localizing the gene(s) on the genome. These new procedures will be even more productive when the genome sequence is finished and when small DNA fragments are available for every gene and open reading frame. The methods we have developed should also be applicable for the study of other globally regulated systems not limited to heat shock. Preliminary results indicate that the differential expression technique also works in the study of bacterial response to phosphate starvation. These procedures are likely applicable to other organisms as well.

ACKNOWLEDGMENTS

We are very grateful to R. Clark for his help in identifying the correspondence of some Hsl proteins to the Htp set. We thank C. A. Gross for the gift of RNA polymerase Eσ^{32}, Y. Kohara and F. C. Neidhardt for the gift of the phage collection miniset, and K. Rudd for the use of his consolidated *E. coli* sequence data base. We also want to thank the reviewer for comments on the ppGpp connection with the heat shock repression.

This work is supported by the NIH grant no. GM35682 and partly by a scholarship to S.-E.C. from the Ministry of Education, Republic of China (Taiwan).

REFERENCES

1. **Allen, S. P., J. O. Polazzi, J. K. Gierse, and A. M. Easton.** 1992. Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. J. Bacteriol. **174:**6938–6947.
2. **Banuett, F., and I. Herskowitz.** 1987. Identification of polypeptides encoded by an *Escherichia coli* locus (*hflA*) that governs the lysis-lysogeny decision of bacteriophage λ. J. Bacteriol. **169:**4076–4085.
3. **Banuett, F., M. A. Hoyt, L. McFarlane, H. Echols, and I. Herskowitz.** 1986. *hfl*B, a new *Escherichia coli* locus regulating lysogeny and the level of bacteriophage lambda cII protein. J. Mol. Biol. **187:**213–224.
4. **Bardwell, J. C. A., and E. A. Craig.** 1987. Eukaryotic Mr 83,000 heat shock protein has a homologue in *E. coli*. Proc. Natl. Acad. Sci. USA **84:**5177–5181.
5. **Brissette, J. L., M. Russel, L. Weiner, and P. Model.** 1990. Phage shock protein, a stress protein of *E. coli*. Proc. Natl. Acad. Sci. USA **87:**862–866.
6. **Brissette, J. L., L. Weiner, T. L. Ripmaster, and P. Model.** 1991. Characterization and sequence of the *Escherichia coli* stress-induced *psp* operon. J. Mol. Biol. **220:**35–48.
7. **Cheng, H. H., P. J. Muhlrad, M. A. Hoyt, and H. Echols.** 1988. Cleavage of the cII protein of phage λ by purified HflA protease: control of the switch between lysis and lysogeny. Proc. Natl. Acad. Sci. USA **85:**7882–7886.
8. **Chuang, S. E., V. Burland, G. Plunkett III, D. L. Daniels, and F. R. Blattner.** Sequence analysis of four new heat-shock genes constituting the *hslTS* (*ibpAB*) and *hslVU* operons in *Escherichia coli*. Gene, in press.
9. **Chuang, S. E., D. L. Daniels, and F. R. Blattner.** 1993. Global regulation of gene expression in *Escherichia coli*. J. Bacteriol. **175:**2026–2036.
10. **Cowing, D. W., J. C. A. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross.** 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. Proc. Natl. Acad. Sci. USA **82:**2679–2683.
11. **Craig, E. A., and C. A. Gross.** 1991. Is hsp70 the cellular thermometer? Trends Biochem. Sci. **16(4):**135–140.
12. **Erickson, J., and C. A. Gross.** 1989. Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternate σ factor involved in high-temperature gene expression. Genes Dev. **3:**1462–1471.
13. **Freedman, L. P., J. M. Zengel, and L. Lindahl.** 1985. Genetic dissection of stringent control and nutritional shift-up response of the *Escherichia coli* S10 ribosomal protein operon. J. Mol. Biol. **185:**701–712.
14. **Gamer, J., H. Bujard, and B. Bukau.** 1992. Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor σ32. Cell **69:**833–842.
15. **Georgopoulos, C., E. Ang, A. Maddock, S. Raina, B. Lipinska, and M. Zylicz.** 1990. Heat shock response of *Escherichia coli*, p. 405–419. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
16. **Grossman, A. D., W. E. Taylor, Z. F. Burton, R. R. Burgess, and C. A. Gross.** 1985. Stringent response in *Escherichia coli* induces expression of heat shock proteins. J. Mol. Biol. **186:**357–365.
17. **Hendrix, R., and L. Tsui.** 1978. Role of the host in virus assembly: cloning of the *E. coli groE* gene and identification of its protein product. Proc. Natl. Acad. Sci. USA **75:**136–139.
18. **Hwang, B. J., K. M. Woo, A. L. Goldberg, and C. H. Chung.** 1988. Protease Ti, a new ATP-dependent protease in *Escherichia coli*, contains protein-activated ATPase and proteolytic functions in distinct subunits. J. Biol. Chem. **263:**8727–8734.
19. **Karow, M., O. Fayet, A. Cegielska, T. Ziegelhoffer, and C. Georgopoulos.** 1991. Isolation and characterization of the *Escherichia coli htrB* gene, whose product is essential for bacterial viability above 33°C in rich media. J. Bacteriol. **173:**741–750.
20. **Karow, M., and C. Georgopoulos.** 1991. Sequencing, mutational analysis, and transcriptional regulation of the *Escherichia coli htrB* gene. Mol. Microbiol. **5:**2285–2292.
21. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
22. **Kroh, H. E., and L. D. Simon.** 1990. The ClpP component of Clp protease is the σ^{32}-dependent heat shock protein F21.5. J. Bacteriol. **172:**6026–6034.
23. **Kroh, H. E., and L. D. Simon.** 1991. Increased ATP-dependent proteolytic activity in Lon-deficient *Escherichia coli* strains lacking the DnaK protein. J. Bacteriol. **173:**2691–2695.
24. **Lerner, C. G., and M. Inouye.** 1991. Pleiotropic changes resulting from depletion of Era, an essential GTP-binding protein in *Escherichia coli*. Mol. Microbiol. **5:**951–957.
25. **Liberek, K., T. P. Galitski, M. Zylicz, and C. Georgopoulos.** 1992. The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the σ^{32} transcription factor. Proc. Natl. Acad. Sci. USA **89:**3516.
26. **Lipinska, B., O. Fayet, L. Baird, and C. Georgopoulos.** 1989. Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J. Bacteriol. **171:**1574–1584.
27. **Lipinska, B., J. King, D. Ang, and C. Georgopoulos.** 1988. Sequence analysis and transcription regulation of *E. coli grpE* gene, encoding a heat shock protein. Nucleic Acids Res. **16:**7545–7562.
28. **Mackow, E. R., and F. N. Chang.** 1983. Correlation between RNA synthesis and ppGpp content in *Escherichia coli* during temperature shifts. Mol. Gen. Genet. **192:**5–9.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Matthews, R. G., and F. C. Neidhardt.** 1988. Abnormal induction of heat shock proteins in an *Escherichia coli* mutant deficient in adenosylmethionine synthetase activity. J. Bacteriol. **170:**1582–1588.
31. **Maurizi, M. R., W. P. Clark, Y. Katayama, S. Rudidoff, J. Pumphrey, B. Bowers, and S. Gottesman.** 1990. Sequence and structure of ClpP, the proteolytic component of the ATP-dependent Clp protease of *Escherichia coli*. J. Biol. Chem. **265:**12536–12545.
32. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119:**736–747.
33. **Neidhardt, F. C., and R. A. VanBogelen.** 1987. Heat shock response, p. 1334–1345. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger

(ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.

34. **Raina, S., and C. Georgopoulos.** 1990. A new *Escherichia coli* heat shock gene, *htrC*, whose product is essential for viability only at high temperatures. J. Bacteriol. **172:**3417–3426.
35. **Raina, S., and C. Georgopoulos.** 1991. The *htrM* gene, whose product is essential for *Escherichia coli* viability only at elevated temperatures, is identical to the *rfaD* gene. Nucleic Acids Res. **19:**3811–3819.
36. **Rudd, K.** 1992. Alignment of *E. coli* DNA sequences to a revised, integrated genomic restriction map, p. 2.3–2.43. *In* J. Miller (ed.), A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
37. **Rudd, K. E., W. Miller, C. Werner, J. Ostell, C. Tolstoshev, and S. G. Satterfield.** 1991. Mapping sequenced *E. coli* genes by computer: software, strategies and examples. Nucleic Acids Res. **19:**637–647.
38. **Stroker, N. G., J. M. Pratt, and I. B. Holland.** 1984. In vivo gene expression systems in prokaryotes, p. 153–177. *In* B. D. Hames and S. J. Higgins (ed.), Transcription and translation, a practical approach. IRL Press, Oxford.
39. **Tilly, K., N. McKittrick, M. Zylicz, and C. Georgopoulos.** 1983. The *dnaK* protein modulates the heat-shock response of *E. coli*. Cell **34:**641–646.
40. **Tobias, J. W., T. E. Shrader, G. Rocap, and A. Varshavsky.** 1991. The N-end rule in bacteria. Science **254:**1374–1377.
41. **Weiner, L., J. L. Brissette, and P. Model.** 1991. Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on σ^{54} and modulated by positive and negative feedback mechanisms. Genes Dev. **5:**1912–1923.