Plant Physiol. (1990) 94, 1689–1695
0032-0889/90/94/1689/07/$01.00/0

Received for publication May 29, 1990
Accepted August 26, 1990

Sequence and Expression of a HSP83 from *Arabidopsis thaliana*[1]

Timothy W. Conner*[2], Peter R. LaFayette, Ronald T. Nagao, and Joe L. Key
Department of Botany, University of Georgia, Athens, Georgia 30602

ABSTRACT

A full-length cDNA encoding a heat shock protein (hsp) belonging to the 83 to 90 kilodalton hsp family of *Arabidopsis thaliana* has been isolated and sequenced. Truncated cDNA clones were isolated by nucleic acid hybridization to a truncated soybean HSP83 cDNA probe and a fragment generated from a *Drosophila* HSP83 gene. A single strand DNA vector/primer based extension procedure was employed to obtain the full-length cDNA. The level of transcripts homologous to this cDNA (AtHS83) is low in 2-week-old *Arabidopsis* plants but is rapidly enhanced by elevated temperatures. DNA sequence comparison between this cDNA and hsp83–90 sequences from human, yeast and *Drosophila* reveal amino acid identities of 63 to 69%, typical identities for interspecies comparisons between hsp83 to 90 kilodalton proteins. Genomic DNA blot analysis performed with probes derived from AtHS83 indicate the presence of a HSP83 gene family estimated to be comprised of at least three genes.

The induction of the hsp[3], a group of proteins synthesized in response to elevated temperatures, exists in almost every organism studied to date (28). The role of hsp has been under intense investigation; their role in thermotolerance or protection of cells from a variety of other stresses is the subject of recent reviews (15, 18).

The hsp83 to 90 kD gene family in eucaryotes encodes a highly conserved group of proteins ranging from approximately 80 to 108 kD. Proteins encoded by these genes have been localized to the cytoplasm, endoplasmic reticulum, and nuclei (18). Like the better characterized 70 kD hsp family, 83 to 90 kD hsp are expressed in unstressed conditions, and are differentially expressed during development. The multiple sites of localization in cells and the specific associations of 83 to 90 kD hsp with a variety of other proteins in vertebrate cells such as cellular kinases, and steroid hormone complexes are established. These studies lead to the suggestion that hsp83–90 kD proteins are involved in a general mechanism of regulation, controlling a variety of cellular functions (18). An additional function of hsp83–90 kD proteins during the heat shock response is postulated to be the modulation of eIF-2a phosphorylation and subsequent regulation of translation of cellular mRNAs (27).

Genomic and cDNA clones of hsp83 to 90 sequences have been isolated and characterized from a number of species including human (13, 26), mouse (21), *Drosophila* (4), fungi (5, 9), trypanosomes (8), and maize and cabbage (personal communication cited in reference 24); a procaryotic counterpart, the *E. coli htpG* gene has also been identified and characterized (1). While the synthesis of hsp induced by elevated temperatures has been studied in a variety of plant species (23, 30), very little characterization of hsp83–90 genes and their transcripts has been reported. In this study, we report the complete sequence of a plant hsp83 cDNA isolated from *Arabidopsis thaliana,* and examine its expression in response to heat shock.

MATERIALS AND METHODS

Arabidopsis Plants

Arabidopsis thaliana (ecotype Columbia) plants used for heat shock experiments were grown at 22°C either in a soil mixture (peat, spermus, sphagus, vermiculite) for 3 weeks in continuous light or in Petri dishes for 2 weeks in continuous light on hormone-free MS media (pH 6) containing 8 g/L agar. MS media contained 4.3 g MS salts (Gibco), 30 g sucrose, 1 mL B5 vitamins (11), and 0.63 g Mes (2-[N-Morpholino] ethanesulfonic acid) per liter.

Nucleic Acid Isolation from Plants and Cells

Total cellular RNA isolations were performed as described (3). Genomic DNA was prepared from 3-week-old plants by a CsCl/EtBr gradient procedure as described (17, 20) with minor modifications. After the CsCl/EtBr gradient and removal of the EtBr by CsCl saturated isopropanol extractions, the DNA was precipitated twice with 2 M NH_4OAc and 2 volumes of 95% ethanol. After centrifugation to collect the DNA, the DNA was resuspended in 0.1× SSC, adjusted to 2× SSC, and digested with 100 μg/mL α-amylase at 37°C for 2 h. The DNA was extracted once with phenol, extensively extracted with chloroform/octanol (24:1), precipitated again with NH_4OAc and ethanol, washed with 70% ethanol, and resuspended in TE (pH 8.0).

[1] This work was supported by Department of Energy Grant DE-FG09-86ER13602 to J. L. K. T. W. C. was supported by Public Health Service postdoctoral fellowship award GM12626 and the University of Georgia Research Foundation. The sequence data presented in this article have been submitted to the EMBL/Gen Bank Data Libraries under the accession number M36960.

[2] Present address: Monsanto Company, 700 Chesterfield Village Parkway, St. Louis, MO 63198.

[3] Abbreviations: hsp, heat shock protein; CsCl/EtBr, Cesium chloride/ethidium bromide; bp, base pair.

Construction and Isolation of cDNA Clones

An *Arabidopsis* cDNA library was constructed in the phage vector λ gt10 using poly(A)$^+$ RNA isolated from leaves of 3 week-old plants which were heat shocked at 36°C for 2 h. The library was screened by plaque hybridization to radiolabeled fragments from 83 kD heat shock sequences from a soybean cDNA (VR34) and a *Drosophila* genomic fragment using standard techniques (20). Fragments from recombinant cDNA clones to be used for sequencing, cDNA extension, and nucleic acid hybridization were subcloned into the phagemid vectors pUC119 or pUC118 (29).

Extension of Truncated cDNA Clone

The 900 bp truncated cDNA (HS11), subcloned into the *Eco*RI site of pUC119, was extended by a primer extension method (Fig. 1). Template DNA (4 μg), prepared as described (29), was annealed to RD_{20} oligomer (NEB) and linearized by digestion with 20 units *Eco*RI at 42°C as described (7). After complete linearization, the DNA was deleted using the 3′ to 5′ exonuclease activity of T4 DNA polymerase (incubation times were adjusted to delete 100–200 nucleotides) to remove any extra linker sequences between the pUC *Eco*RI site and the beginning of the cDNA sequence. After phenol extractions and ethanol precipitation, the DNA was resuspended in hybridization buffer (40 mM Pipes-HCl [pH 6.4], 1 mM EDTA, 0.4 M NaCl, and 80% formamide) at a concentration of 0.1 μg/μL (6). Poly(A)$^+$ RNA (10 μg) was isolated from 3-week-old plants heat shocked as described above, precipitated and resuspended in hybridization buffer at a concentration of 0.5 μg/μL. Linearized template DNA (2 μg) was added to the RNA; the mixture was denatured at 65°C for 5 min, then

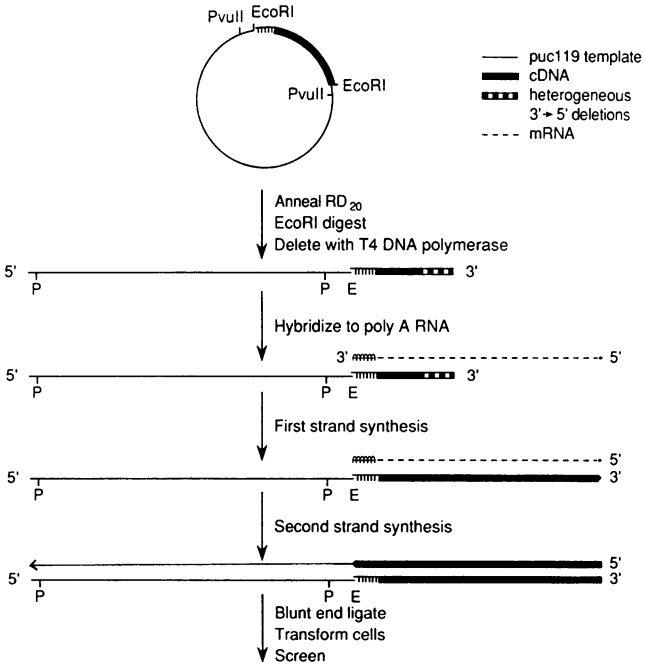

Figure 1. Strategy for extension of truncated cDNA clone. TTTTTT-poly T strand of the cDNA, P- *Pvu*II, E- *Eco*RI.

incubated at 37°C for 4 h, and ethanol precipitated. The pellet was resuspended in 25 μL TE, precipitated with ethanol containing 2% KOAc, and washed twice with 70% ethanol. The DNA/RNA hybrid was resuspended in 10 μL H_2O; first and second strand cDNA synthesis was performed as described (12). After phenol extraction and ethanol precipitation of the second strand cDNA reaction products, the double strand cDNA/plasmid was resuspended in 50 μL H_2O, adjusted to 50 mM Tris-Cl (pH 7.6), 10 mM $MgCl_2$, 5% PEG 8000 (w/v), 1 mM ATP, 1 mM DTT, and 200 units T4 DNA ligase (NEB) in 100 μl and ligated 20 h at 15°C. The ligation mixture was phenol extracted, ethanol precipitated, and resuspended in 25 μl H_2O. Hanahan-competent DH1 cells transformed with <2 μL of the phenol clarified ligation products were plated on YT agar plates containing 50 μg/mL ampicillin. *In situ* colony lifts were performed with nitrocellulose filters and hybridized to a radiolabeled 5′ probe derived from HS22 (Fig. 2) which did not overlap the 3′ cDNA sequence from HS11.

DNA Sequencing

Fragments identified above were sequenced by the strategy shown in Figure 2. Overlapping sequential deletions were generated as described (7). DNA sequencing was performed by the dideoxy chain termination method using the U.S. Biochemical sequenase sequencing method and 6% polyacrylamide sequencing gels. The complete sequence (100%) was determined from both strands for HS83 and HS22. Only the 3′ and 5′ ends were sequenced for HS11 and HS16 truncated cDNAs.

Preparation of Radiolabeled Probes

Fragments for use in DNA and RNA blot hybridizations were isolated by standard procedures and radiolabeled by the random oligonucleotide priming method (10).

Nucleic Acid Hybridization

Genomic DNA blots and nitrocellulose filter lifts for plaque and *in situ* colony hybridization were prehybridized at 60°C in 6× SSC, 20 mM sodium phosphate (pH 6.5), 5× Denhardt's solution, 0.5% SDS, 10 μg/mL denatured sheared salmon testes DNA, and 10 μg/mL yeast RNA for 6 to 8 h. Hybridization was performed in fresh hybridization solution to which denatured probe was added. RNA blots were prehybridized and hybridized at 42°C in 50% formamide, 5× SSC, 50 mM sodium phosphate (pH 6.5), 5× Denhardt's solution, 0.1% SDS, 10 μg/mL sheared salmon testis DNA, and 10 μg/mL yeast RNA. Prehybridization was performed for 4 to 6 h, after which fresh formamide hybridization solution containing denatured probe was added. Both DNA and RNA blots were hybridized for 10 to 12 h. Hybridized filters were rinsed with 2× SSC, 0.1% SDS by shaking 15 min at room temperature, and washed with three changes of 2× SSC, 0.1% SDS at 60°C for a total of 45 min, and 0.2× SSC, 0.1% SDS at 60°C for 15 min.

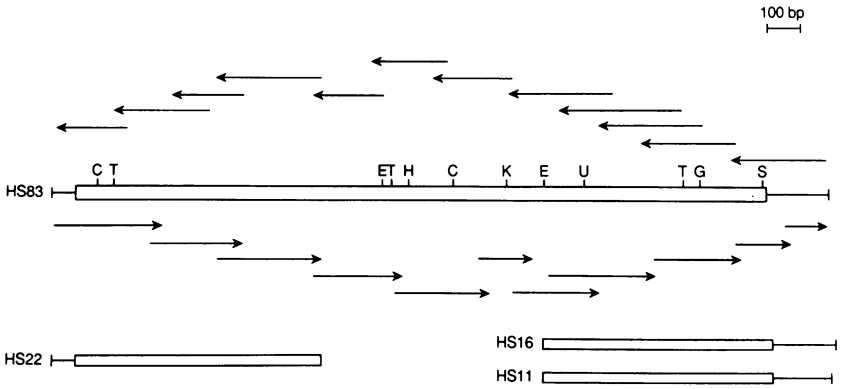

Figure 2. Restriction maps and sequencing strategy for HS11, HS16, HS22, and HS83 (AtHS83) cDNA clones. Arrows indicate the direction and extent of sequencing. Open regions denote the coding sequences flanked by 5′ and 3′ untranslated sequences (attached line). All cDNAs were cloned into the *Eco*RI site of the pUC119 polylinker. C- *Bcl*I, E- *Eco*RI, G- *Bgl*II, K- *Kpn*I, S- *Sal*I, T- *Sst*I, U- *Stu*I.

RESULTS

Isolation of AtHS83 cDNA

The λgt10 cDNA library constructed from poly(A)$^+$ RNA isolated from leaves of heat shocked plants was screened with VR34, a truncated soybean cDNA which yields a 83 kD hsp in hybrid selection/*in vitro*, translation experiments (not shown), and a 2 kb fragment from D301.1, a *Drosophila* hsp83 genomic clone (14). Two recombinant cDNAs, HS11 and HS16 (containing 920 and 980 bp inserts, respectively), were isolated from plaque hybridizations with VR34 (Fig. 2). A third cDNA, HS22 (1.0 kb), was isolated from plaque screens with the 2.0 kb *Eco*RI, *Pst*I radiolabeled fragment from D301.1. Radiolabeled inserts from all of these recombinant phage hybridized to a 2.4 kb transcript whose levels increased during heat shock conditions. These clones were characterized by partial nucleotide sequencing and oriented relative to the human 90 kD hsp cDNA sequence (26). HS11 and HS16 possessed identical 5′ termini but differed at their polyadenylation sites. HS22 contained sequences homologous to the amino terminal sequences of human and yeast hsp83–90 sequences, but does not overlap with HS11 and HS16 sequences.

We performed a cDNA extension procedure to isolate a full length contiguous hsp83 cDNA (see "Materials and Methods"). Screening of HS11 extended transformant colonies by hybridization with an 800 bp 5′ probe derived from HS22 yielded several hundred extended cDNAs. Two of the extended cDNAs chosen contained cDNA inserts of 2.4 kb.

Sequence and Characterization of the AtHS83 cDNA

The complete sequence of one of the extended cDNAs, HS83, was determined. The partial restriction map and sequencing strategy of HS83, hereafter referred to as AtHS83, is shown in Figure 2. This cDNA contains 2391 bp with an open reading frame of 2115 bp (Fig. 3). The cDNA contains two ATG codons at positions 70 and 85 bp downstream of the 5′ terminus of the cDNA. These two putative initiation codons are in the same translational reading frame and would encode proteins with predicted molecular masses of 81 kD and 80.5 kD.

The predicted pIs of the respective translation products obtained from sequence analysis would be 5.73 and 5.76; these correspond to the pI of the 82 kD protein translated *in vitro* from hybrid selection of homologous transcripts (Fig. 4). Comparison of the predicted amino acid sequence between AtHS83 and human, *Drosophila*, and yeast 83 to 90 kD hsp reveals sequence identities of 68.2%, 63.1%, and 66.9%, respectively (Fig. 5). As with nearly all other hsp/hsc 83 to 90 kD sequences examined to date in eucaryotes, this plant 83 kD hsp contains the MEEVD carboxy-terminal amino acid sequence (5).

AtHS83 is Encoded by a Multigene Family

The size of the *Arabidopsis* hsp83 gene family was estimated by hybridization of *Arabidopsis* genomic DNA blots to probes spanning the entire cDNA, probes containing 5′ or 3′ portions of the AtHS83 cDNA, or a probe containing the last three codons and the 3′ untranslated sequence (Fig. 6). The probe spanning the entire cDNA hybridized to one high molecular mass fragment from *Bam*HI digested genomic DNA and several genomic fragments from *Eco*RI, and *Hin*dIII digestions (Fig. 6A). The 5′ probe hybridized with one to four fragments (Fig. 6B). A nonoverlapping 3′ probe detected one to three fragments (Fig. 6C). The lesser number of bands that hybridized in each case was from the *Bam*HI digested DNA. *Bam*HI genomic fragments that hybridized were >20 kb, and may contain multiple AtHS83 homologous genes that comigrate during electrophoresis or are clustered on this fragment. Excluding the *Bam*HI genomic DNA blot results for the reasons mentioned above, the 5′ and 3′ nonoverlapping probes detected only one comigrating band of 2.6 kb (Fig. 6, B, C, lane 3). The detection of multiple fragments by small, nonoverlapping probes is indicative of a multigene family. These results indicate that the *Arabidopsis* genome contains a hsp83 family of at least three genes.

AtHS83 Transcript Levels are Enhanced during Heat Shock

Heat inducibility of hsp83 transcripts was examined by incubating 2-week-old *Arabidopsis* plants in incubation buffer or in air incubators at control and elevated temperatures. RNA isolated from 2 week-old *Arabidopsis* plants incubated for 2 h in incubation buffer at 26, 32, and 36°C was hybridized with the complete cDNA (Fig. 7A, lanes 1–3). The probe detected a 2400 bp heat-inducible transcript with its highest level detected at 36°C. Enhanced levels of transcripts hybrid-

```
                                                                  Met Ala Asp Val    4
TCGAAGTTCCAAATTTTCTCTTAGCATTCTCTTTCGTTTCTCGTTTGCGTTGAATCAAAGTTCGTTGCG   ATG GCG GAT GTT   81

Gln Met Ala Asp Ala Glu Thr Phe Ala Phe Gln Ala Glu Ile Asn Gln Leu Leu Ser Leu Ile Ile   26
CAG ATG GCT GAT GCA GAA ACT TTT GCT TTC CAA GCT GAG ATT AAC CAG CTT CTT AGC TTG ATC ATC  147

Asn Thr Phe Tyr Ser Asn Lys Glu Ile Phe Leu Arg Glu Leu Ile Ser Asn Ser Ser Asp Ala Leu   48
AAC ACG TTC TAC AGC AAC AAA GAA ATC TTC CTC CGT GAG CTC ATC AGT AAC TCT TCT GAT GCT CTT  213

Asp Lys Ile Arg Phe Glu Ser Leu Thr Asp Lys Ser Lys Leu Asp Gly Gln Pro Glu Leu Phe Ile   70
GAC AAG ATC CGA TTT GAG AGC TTA ACG GAT AAG AGC AAG CTC GAT GGA CAG CCT GAA CTC TTC ATT  279

Arg Leu Val Pro Asp Lys Ala Asn Lys Thr Leu Ser Ile Ile Asp Ser Gly Ile Gly Met Thr Lys   92
AGA TTG GTT CCT GAC AAG GCT AAT AAG ACG CTC TCA ATT ATT GAC AGT GGT ATT GGC ATG ACC AAA  345

Ala Asp Leu Val Asn Asn Leu Gly Thr Ile Ala Arg Ser Gly Thr Lys Glu Phe Met Glu Ala Leu  114
GCA GAT TTG GTG AAC AAC TTG GGA ACC ATT GCG AGG TCT GGA ACA AAA GAG TTT ATG GAG GCG CTT  411

Gln Ala Gly Ala Asp Val Ser Met Ile Gly Gln Phe Gly Val Gly Phe Tyr Ser Ala Tyr Leu Val  136
CAA GCT GGA GCT GAT GTA AGC ATG ATA GGA CAA TTT GGT GTT GGT TTC TAC TCT GCT TAT CTT GTT  477

Ala Glu Lys Val Val Val Thr Thr Lys His Asn Asp Asp Glu Gln Tyr Val Trp Glu Ser Gln Ala  158
GCA GAG AAG GTT GTT GTC ACT ACA AAG CAC AAT GAT GAT GAA CAA TAC GTT TGG GAG TCT CAA GCT  543

Gly Gly Ser Phe Thr Val Thr Arg Asp Val Asp Gly Glu Pro Leu Gly Arg Gly Thr Lys Ile Ser  180
GGT GGT TCC TTC ACT GTC ACT AGG GAT GTG GAT GGG GAA CCA CTT GGT AGA GGA ACT AAG ATC AGC  609

Leu Phe Leu Lys Asp Asp Gln Leu Glu Tyr Leu Glu Glu Arg Arg Leu Lys Asp Leu Val Lys Lys  202
CTC TTC CTT AAG GAC GAT CAG CTT GAA TAC TTG GAG GAG AGG AGA CTC AAA GAC TTG GTG AAG AAG  675

His Ser Glu Phe Ile Ser Tyr Pro Ile Tyr Leu Trp Thr Glu Lys Thr Thr Glu Lys Glu Ile Ser  224
CAC TCT GAG TTC ATC AGT TAC CCT ATC TAC CTT TGG ACC GAG AAA ACC ACC GAG AAG GAG ATC AGT  741

Asp Asp Glu Asp Glu Asp Glu Pro Lys Lys Glu Asn Glu Gly Glu Val Glu Glu Val Asp Glu Glu  246
GAC GAT GAG GAT GAA GAT GAA CCA AAG AAA GAA AAC GAA GGT GAG GTT GAA GAA GTT GAT GAG GAG  807

Lys Glu Gys Asp Gly Lys Lys Lys Lys Lys Ile Lys Glu Val Ser His Glu Trp Glu Leu Ile Asn  268
AAG GAG AAA GAT GGT AAA AAG AAG AAG AAA ATC AAG GAA GTC TCT CAC GAG TGG GAA CTC ATC AAC  873

Lys Gln Lys Pro Ile Trp Leu Arg Lys Pro Glu Glu Ile Thr Lys Glu Glu Ser Ala Ala Phe Tyr  290
AAG CAG AAA CCG ATC TGG TTG AGG AAG CCA GAA GAG ATC ACT AAG GAA GAG TCT GCT GCT TTC TAC  939

Lys Ser Leu Thr Asn Asp Trp Glu Asp His Leu Ala Val Lys His Phe Ser Val Glu Gly Gln Leu  312
AAG AGC TTG ACC AAT GAC TGG GAA GAT CAC TTA GCC GTG AAA CAC TTC TCA GTG GAG GGT CAG CTA 1005

Glu Phe Lys Ala Ile Leu Phe Val Pro Lys Arg Ala Pro Phe Asp Leu Phe Asp Thr Arg Lys Lys  334
GAA TTC AAG GCC ATT CTC TTT GTA CCA AAG AGA GCT CCG TTT GAT CTC TTT GAC ACG AGG AAG AAG 1071

Leu Asn Asn Ile Lys Leu Tyr Val Arg Arg Val Phe Ile Met Asp Asn Cys Glu Glu Leu Ile Pro  356
TTG AAC AAC ATC AAG CTT TAT GTC AGG AGG GTG TTC ATT ATG GAC AAC TGT GAA GAG CTA ATC CCA 1137

Glu Tyr Leu Ser Phe Val Lys Gly Val Val Asp Ser Asp Asp Leu Pro Leu Asn Ile Ser Arg Glu  378
GAG TAC CTC AGC TTT GTG AAA GGT GTT GTT GAC TCT GAT GAC TTG CCA CTC AAC ATC TCT CGT GAG 1203

Thr Leu Gln Gln Asn Lys Ile Leu Lys Val Ile Arg Lys Asn Leu Val Lys Lys Cys Ile Glu Met  400
ACG CTT CAA CAG AAC AAG ATC CTT AAG GTG ATC AGG AAG AAT CTA GTG AAG AAG TGC ATT GAG ATG 1269

Phe Asn Glu Ile Ala Glu Asn Lys Glu Asp Tyr Thr Lys Phe Tyr Glu Ala Phe Ser Lys Asn Leu  422
TTC AAC GAG ATT GCT GAG AAC AAA GAG GAC TAC ACC AAA TTC TAT GAG GCT TTC TCC AAG AAT CTC 1335

Lys Leu Glu Ile His Glu Asp Ser Gln Asn Arg Gly Lys Ile Als Asp Leu Leu Arg Tyr His Ser  444
AAA TTG GGT ATC CAT GAA GAC AGT CAG AAC AGG GGA AAG ATT GCT GAT CTT CTA CGG TAC CAC TCC 1401

Thr Lys Ser Gly Asp Glu Met Thr Ser Phe Lys Asp Tyr Val Thr Arg Met Lys Glu Gly Gln Lys  466
ACA AAG AGT GGT GAT GAA ATG ACG AGC TTC AAA GAT TAC GTC ACA AGG ATG AAG GAA GGT CAA AAG 1467

Asp Ile Phe Tyr Ile Thr Gly Glu Ser Lys Lys Ala Val Glu Asn Ser Pro Phe Leu Glu Arg Leu  488
GAC ATT TTC TAC ATC ACT GGT GAA AGC AAA AAG GCG GTG GAG AAT TCT CCC TTC TTG GAG AGG CTG 1533
                                                    HS16 |-->
Lys Lys Arg Gly Tyr Glu Val Leu Tyr Met Val Asp Ala Ile Asp Glu Tyr Ala Val Gly Gln Leu  510
AAG AAG AGA GGA TAC GAG GTA CTT TAC ATG GTG GAT GCG ATT GAC GAA TAC GCT GTT GGA CAA TTG 1599

Lys Glu Tyr Asp Gly Lys Lys Leu Val Ser Ala Thr Lys Glu Gly Leu Lys Leu Glu Asp Glu Thr  532
AAG GAG TAT GAC GGT AAG AAA CTT GTT TCT GCG ACT AAA GAA GGC CTC AAA CTT GAA GAT GAG ACC 1665

Glu Glu Glu Lys Lys Lys Arg Glu Glu Lys Lys Lys Ser Phe Glu Asn Leu Cys Lys Thr Ile Lys  554
GAA GAA GAG AAG AAA AAG AGG GAA GAG AAG AAG AAG TCC TTC GAG AAT CTG TGC AAG ACG ATT AAG 1731

Glu Ile Leu Gly Asp Lys Val Glu Lys Val Val Val Ser Asp Arg Ile Val Asp Ser Pro Sys Cys  576
GAA ATT CTC GGG GAC AAG GTT GAG AAG GTT GTG GTC TCA GAC AGG ATT GTG GAC TCT CCC TGC TGT 1797

Leu Val Thr Gly Glu Tyr Gly Trp Thr Ala Asn Met Glu Arg Ile Met Lys Ala Gln Ala Leu Arg  598
CTA GTA ACT GGT GAA TAT GGA TGG ACT GCA AAT ATG GAG AGG ATT ATG AAG GCA CAG GCG TTG AGA 1863

Asp Ser Ser Met Ser Gly Tyr Met Ser Ser Lys Lys Thr Met Glu Ile Asn Pro Asp Asn Gly Ile  620
GAT AGC AGC ATG AGT GGT TAC ATG TCG AGC AAG AAA ACA ATG GAG ATC AAC CCC GAC AAC GGT ATA 1929

Met Glu Glu Leu Arg Lys Arg Ala Glu Ala Asp Lys Asn Asp Lys Ser Val Lys Asp Leu Val Met  642
ATG GAG GAG CTC AGG AAG AGA GCT GAA GCA GAC AAG AAT GAC AAG TCT GTT AAA GAT CTT GTC ATG 1995

Leu Leu Tyr Glu Thr Ala Leu Leu Thr Ser Gly Phe Ser Leu Asp Glu Pro Asn Thr Phe Ala Ala  664
TTG CTG TAT GAG ACA GCT TTG TTG ACG TCT GGA TTC AGT CTT GAT GAA CCG AAC ACT TTT GCT GCT 2061

Arg Ile His Arg Met Leu Lys Leu Gly Leu Ser Ile Asp Glu Asp Glu Asn Val Glu Glu Asp Gly  686
AGG ATT CAC AGG ATG TTG AAG TTG GGT CTG AGT ATT GAT GAG GAT GAG AAC GTT GAG GAA GAT GGT 2127

Asp Met Pro Glu Leu Glu Glu Asp Ala Ala Glu Glu Ser Lys Met Glu Glu Val Asp  *          705
GAT ATG CCT GAG TTG GAG GAG GAC GCT GCT GAA GAG AGC AAG ATG GAG GAA GTC GAC TAA GAGATGA 2194

AGAAATTGCTCTTATGGTTCTGAAAACTTCTAATATGTCGAGTTGTTCTGAGTTTTAAGATTTTCCAAAATGTCTGTCTTTTTTTTT 2281

TATCTTTAGAGTTACTTGAACATTGTGACTACTTCTAGGGTTGGGTTTGTGTCAGGTCTGTTATATCGTGTGGTGGGTCTGTCTAAT 2368

ACTGATTCGAGTTTTTGTTATTCAAAAAA   End of HS83                                              2391
                              GCTAAGGAACTTTTCTTGTTTCTGAACAAAAAA   End of HS16
```

Figure 3. Nucleotide sequences of the AtHS83 cDNA and predicted amino acid sequences. The nucleotide sequence was determined by the strategy shown in Figure 2. The 5′ untranslated region is numbered beginning with the 5′ end of the cDNA as 1. Numbers at the end of each line denote the last amino acid residue number (top) and nucleotide number (bottom) on that line. The arrow at position 1509 denotes the 5′ end of HS16; the HS16 3′ terminus and extension is shown near the end of the sequence.

Figure 4. Hybrid selection and *in vitro* translation of transcripts homologous to AtHS83 from heat shocked poly(A)$^+$ RNA. AtHS83 cDNA was spotted onto nitrocellulose, and hybridized to poly(A)$^+$ RNA; the filters were washed and the selected RNA was eluted (20). The RNA was translated in a wheat germ cell free translation system (S23) (22) and analyzed by two-dimensional PAGE (25). Dashes on the left margin indicate the migration of protein standard markers (200, 97.4, 68, and 43 kD, top to bottom). Numbers indicate the pI range of the gel.

izing to a 3′ probe derived from the last three codons and the 3′ untranslated region of AtHS83 were also observed with similar kinetics to that detected by the complete cDNA probe (Fig. 7B, lanes 1–3). This 3′ gene probe (an *Eco*RI, *Sal*I 220 bp fragment) detected only one fragment on a genomic DNA blot which implies that only transcripts derived from one gene were detected (Fig. 6D). To determine whether the heat-inducible expression of the transcript was similar under different experimental heat shock conditions, 2-week-old plants were incubated in air incubators for 2 h at temperatures of 24, 28, 32, 36, and 40°C (Fig. 7; lanes 4–8). The complete probe and the 3′ probe both detected transcripts at 32°C that were enhanced relative to the 24 and 28°C incubated plants. The level of transcripts detected from plants heat shocked in the air incubator was highest at 40°C. The enhanced levels of AtHS83 homologous transcripts obtained from plants heat shocked in incubation buffer and those heat shocked by elevating the air temperature indicated that the accumulation of AtHS83 homologous transcripts was heat-inducible.

The accumulation of AtHS83 transcripts was rapid, being detected within 15 min of heat stress (Fig. 8). The transcript increased to a maximum level at 2 h and declined after 2 to 4 h. The accumulation of homologous transcripts was similar with both the complete probe and the 3′ probe.

DISCUSSION

We have presented results of studies on the expression and sequence analysis of a HSP83 sequence in a higher plant. The cDNAs analyzed in this study were constructed from poly(A)$^+$

```
At  M---ADVQMADAETFAFQAEINQLLSLIINTFYSNKEIFLRELISNSSDALDKIRFESLTDKSKLDGQPELFIRLVPDKANKT  80
H    PEEVHHGEEEV         A  M                        A        Y     P    SGK  K DII NPQER  83
D       -----EE          A  M                        A        Y     P    SGK  Y K I N TAQ  75
S        ------AS        T  M      V                 A        YK  S PKQ ETE D    IT KPEQ V  74

At  LSIIDSGIGMTKADLVNNLGTIARSGTKEFMEALQAGADVSMIGQFGVGFYSAYLVAEKVVVTTKHNDDEQYVWESQAGGSF 162
H    TLV T          I       K    A           I                    IR        A   S      165
D    T  T                   K    A           I                D   T   S N       S      157
S    E R             I       K    A     S                   LF   DR Q IS S     I   N    156

At  TVTRDVDGEPLGRGTKISLFLKDDQLEYLEERRLKDLVKK-HSEFISYPIYLWTEKTTEKEI--SDDEDEDEP--KKENEGE 239
H     RA -H    I M   VI H  E  T       V EV     Q  G   T YL  ER          A E KGE E EDKD 243
D     RA -NS         V YI E  TD    S I EI N    Q  G   K LV  ER    V     AD  KKE-GDEKK  234
S      L EVN RI     ILR            K I E- I R    VD   Q VVT EV   VPIPEE KK  EK  D EKK- 235

At  VE--------EVD---E--EK-EKDGKKKKKIKEVSHEWELINKQKPIWLRKPEEITKEESAAFYKSLTNDWEDHLAVKHFS 307
H   D  E KPKIED GSDE DDSGKD -K  T     KYIDQ EL  T     T N DD  Q  YGE                    322
D   M TD EPKIED G ED DAD KD  A  T     KYT D EL  T     T N DD SQ  YGE                    314
S   D DDKKPKLE    EE      K---P T  -V  EVQ I EL  T  L T N SD  Q  YN     IS     P Y      310

At  VEGQLEFKAILFVPKRAPFDLFDTRKKLNNIKLYVRRVFIMDNCEELIPEYLSFVKGVVDSDDLPLNISRETLQQNKILKVI 389
H          R L  I R          ENK  K                 S D        N IR       E          M   S      404
D          R L  I R T        ENK  R                           D    N M      E          M     V    396
S          R    I             ESK  K                    T EA D     W         E     L    M       M  392

At  RKNLVKKCIEMFNEIAENKEDYTKFYEAFSKNLKLGIHEDSQNRGKIADLLRYHSTKSGDEMTSFKDYVTRMKEGQKDIFYI 471
H       I     L L S  L   D   N K                          T  RRLSE      TSQ        LSE  S     T  S Y   486
D             TM LIE LT D   N K   DQ          V     N  A L   F  F TSA     D FC LA   S    DN   HVYF  478
S       I     L  A       DS QFE   S      I    V     T   AAL K     N      V   L   LT      P H   N Y   474

At  TGESKKAVENSPFLERLKKRGYEVLYMVDAIDEYAVGQLKEYDGKKLVSATKEGLKL-E-DETEEEKKKRE-EKKKSFENLC 550
H         EQ A   A V  VR    F  V  TEP     C Q     F    S    V       E P     E--     M    S AK        564
D         DQ S   A V  V A    F  V  TEP     VIQH     K   Q    V       E P     S--       D AK   S    556
S          L     K     DA  AK F   FLT P      FT     FE   T   DI   -DFE E T   ---   AE   K I -EY P T  551

At  KTIKEILGDKVEKVVVSDRIVDSPCCLVTGEYGWTANMERIMKAQALRDSSMSGYMSSKKTMEINPDNGIMEELRKRAE-AD 631
H     LM     DK       TI N L S     I   ST                      N TM    MA   HL      HP V T   QK       645
D     LM S   DN           N L       I   SQF  S                  TATM    AG   QL      HP V T   QK D     637
S     AL       Q         YKLL A AAIR  QF  S                        S          F  S KSP IK  K   VDEGG  633

At  KNDKSVKDLVMLLYETALLTSGFSLDEPNTFAARIHRMLKLGLSIDEDE-NVEEDGDM-----PELEEDAAEESKMEEVD*  705
H         A     V  F     S      ED Q HSN   Y   I     G       VAA  PNAAVPDEI P - G EDA R      *  724
D         A     I  F  S  S      S QVH S   Y   I     G       PMTT DAQSAGDA S V  TEDA H      *  718
S   AQ  T     TK                 TS  S   N LIS   N          -- TETAPEASTA-APV EVPADTE      *  709
```

Figure 5. AtHS83 amino acid sequence comparisons. The deduced amino acid sequence of AtHS83 is listed on the top line. Only amino acid differences are listed for the predicted protein sequences of human 90 (26), *Drosphila* hsp83 (4), and yeast HSP90 (9) represented as H, D, S, respectively. Dashes are inserted to maximize alignment. Asterisks indicate the stop codon of each protein. Numbers on the right margin indicate the position of the last amino acid residue on that line.

RNA isolated from heat shocked plants; all three truncated cDNAs (HS11, HS16, and HS22) were colinear with the full-length extended cDNA. The independent isolation of three cDNAs corresponding to the extended cDNA implies that they represent the most abundant hsp83–90 kD transcript in heat shocked plants. Genomic DNA blot hybridization experiments indicate that AtHS83 corresponds to one gene of a small multiple gene family.

Like nearly all eucaryotic 83 to 90 kD hsp/hsc genes examined to date, the carboxy-terminal amino acid sequence MEEVD is conserved in an otherwise divergent region of the protein (Fig. 5). The significance of these five amino acids is not known, and the presence of the same four carboxy-terminal amino acids (EEVD) of most hsp70 proteins (5) indicates a potentially important sequence, whose removal may alter the protein's specificity, stability, or compartmentation. Like similar interspecies comparisons between mammalian, yeast, *Drosophila,* and trypanosome hsp83–90 sequences, the amino acid identities between this plant hsp83 sequence and the amino acid sequences of hsp83 from human, *Drosophila,* and yeast are highest in the amino-terminal 25% of the amino acid coding sequence. However, unlike the carboxy-terminus, the amino terminal end of these genes are often divergent. Among vertebrates, the hsp83–90 family belongs to two classes based on their amino terminal sequence domain; some genes encode proteins with a glutamine-rich segment interrupting an acidic stretch of amino acids. The functional importance of this basic domain is unknown and has not been reported for nonvertebrate hsp83–90 genes including the plant cDNA presented in this study.

The AtHS83 cDNA possesses two methionine codons (amino acids 1 and 6) in the same translational reading frame (Fig. 3). Using the consensus translation initiation consensus sequence surrounding the AUG for plant sequences, (AACAAUGGC) determined by base frequencies of plant genes (19), the first methionine codon (TGCGAUGGC) would be the preferred site relative to methionine 6 (TCAGAUGGC) for initiation of translation. It possesses a purine (G) at the −3 position which is next preferred to an optimal A base at that position.

Subtle differences in expression of AtHS83 homologous transcript(s) were apparent in our studies at control temperatures and during heat shock between plants incubated in incubation buffer and air. Plants incubated in buffer have higher detectable levels of AtHS83 transcripts at control temperatures than control plants in air incubators (Fig. 7). These differences may reflect an additional stress that leaves in liquid medium may be subjected to during incubation. Optimal heat shock is obtained in *Arabidopsis* plants at 36 to 37°C (30; M Mansfield, R Nagao, J Key, unpublished data). In heat shock experiments performed in air incubators, the level of AtHS83

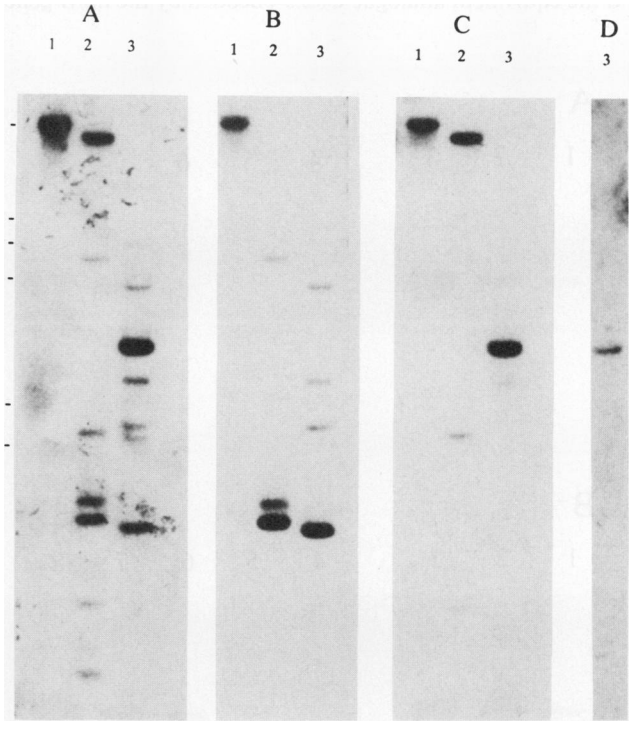

Figure 6. Restriction analysis of *A. thaliana* genomic DNA. *Arabidopsis* DNA (4.0 μg) digested with *Bam*H1 (lane 1), *Eco*RI (lane 2), or *Hin*dIII (lane 3), was electrophoresed on a 1% agarose gel in TAE and transferred to Zeta probe nylon membrane (Bio-Rad). The DNA was hybridized with a radiolabeled (A) 2200 bp *Sst*I, *Bam*HI fragment from HS83, (B) 820 bp *Sst*I, *Eco*RI fragment from HS83, (C) 920 bp *Eco*RI fragment from HS11, and (D) 220 bp *Sal*I, *Eco*RI fragment from HS11. Dashes on the left margin indicate the migration of *Hin*dIII fragments of phage λ DNA in kb (top to bottom: 23.13, 9.42, 6.56, 4.36, 2.32, 2.03, 0.56).

transcripts equivalent to the optimal heat shock in incubation buffer was detected at approximately 40°C. In incubation buffer, the temperature of the leaves is near equilibrium with the ambient temperature. Leaves heat stressed at elevated air temperatures have the ability to transpire thereby reducing their temperature several degrees relative to the ambient temperature (16).

We have characterized the cDNA from only one member of the hsp83–90 gene family in *Arabidopsis*. The presence of other hsp83–90 genes were indicated from our genomic DNA hybridization results. These genes may be expressed and induced by heat shock at lower levels. This would imply that these genes are equivalent to the heat shock cognate genes or may represent pseudogenes. One cognate gene, HSC82, corresponding to the hsp83–90 gene family has recently been characterized in yeast (5). While the amino acid sequence of yeast HSC82 is 97% identical to the yeast HSP82, its constitutive expression is higher than its hsp counterpart. However, HSC82 expression increases only moderately compared to HSP82 during heat shock.

The isolation and characterization of an AtHS83 cDNA in *Arabidopsis* provides the initial tools to examine the function of AtHS83 and to determine whether the gene product is essential in tissues or organs of a multicellular organism. Studies to probe the requirement of the 82 kD hsp of yeast and the equivalent analogue C62.5 encoded by the *htpG* gene

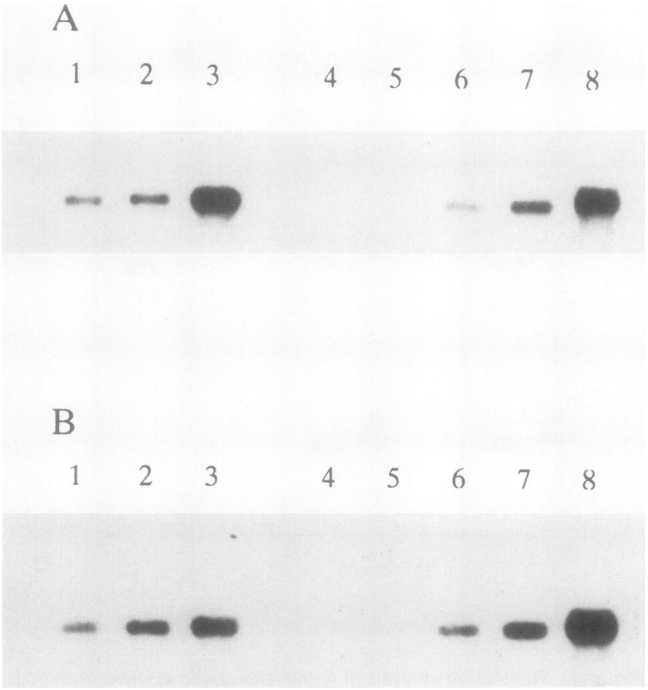

Figure 7. Induction of AtHS83 homologous transcripts in *Arabidopsis* plants. RNA was isolated from 2-week-old plants which were incubated in incubation buffer for 2 h at 26, 32, 36°C (lanes 1–3) or in air incubators for 2 h at 24, 28, 32, 36, and 40°C (lanes 4–8). Total RNA (4 μg) was electrophoresed in a 1.5% agarose formaldehyde gel in Mops running buffer (20), blotted to nitrocellulose and hybridized with (A) a 2200 bp *Sst*I, *Bam*HI radiolabeled fragment from HS83, or (B) a 220 bp *Sal*I, *Eco*RI HS11 radiolabeled fragment.

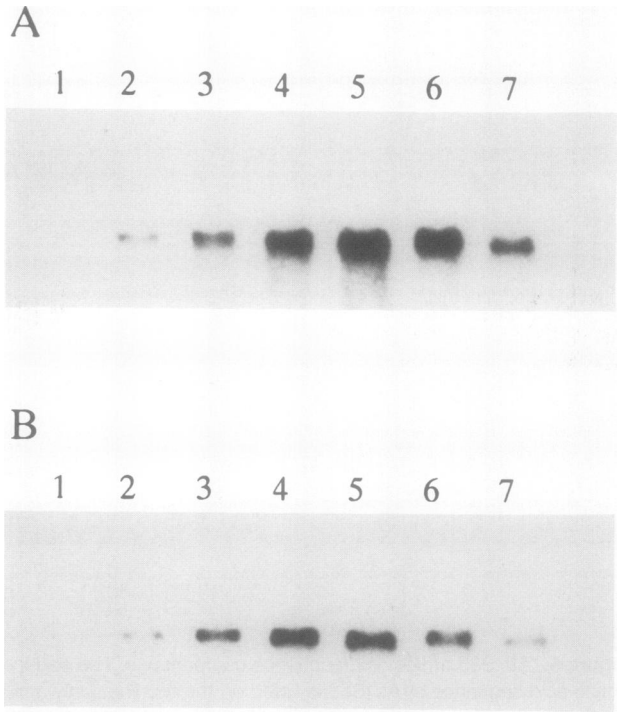

Figure 8. Accumulation of AtHS83 transcripts in *Arabidopsis* plants. RNA was isolated from 2-week-old plants and incubated in incubation buffer at 36°C for 0 min, 15 min, 30 min, 1 h, 2 h, 4 h, 8 h (lanes 1–7, respectively). RNA was electrophoresed and hybridized to (A) a 2200 bp *Sst*I, *Bam*HI radiolabeled fragment from HS83, or (B) a 220 bp *Sal*I, *Eco*RI HS11 radiolabeled fragment as described in Figure 7.

of *E. Coli* yield different conclusions. Mutants with lesions in *htpG* in *E. Coli* (2) and either one of the two loci in yeast (5) exhibit similar phenotypes; these mutants show an extreme growth disadvantage at heat shock temperatures and very little growth difference at normal temperatures in both organisms. However, the requirement for one functional gene in yeast is absolute for growth at normal temperatures. While the precise functions of 83 to 90 kD hsp are being unraveled, a pseudogenetic study allowing overexpression and inhibiting expression (antisense inhibition experiments) in a transgenic system will allow assessment of the role of AtHS83 in translation and throughout development.

ACKNOWLEDGMENTS

We wish to thank Dr. James Roberts and Dr. Elizabeth Vierling for providing us with the soybean cDNA probe VR34, Dr. Nora Plesofsky-Vig for providing us a *Neurospora* hsp83 cDNA, and Dr. Matthew Meselson for providing us with the D301.1 *Drosophila* HSP83 genomic clone.

LITERATURE CITED

1. **Bardwell J, Craig E** (1987) Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichial coli*. Proc Natl Acad Sci USA **81:** 5177–5781
2. **Bardwell J, Craig E** (1988) Ancient heat shock gene is dispensable. J Bacteriol **7:** 2977–2983
3. **Baulcombe D, Key J** (1980) Polyadenylated RNA sequences

which are reduced in concentration following auxin treatment of soybean hypocotyls. J Biol Chem **255:** 8907–8911

4. **Blackman R, Meselson M** (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila* hsp82 gene. J Mol Biol **188:** 499–515
5. **Borkovich K, Farrelly F, Finkelstein D, Taulien J, Lindquist S** (1989) hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. Mol Cell Biol **9:** 3919–3930
6. **Chirala S, Kasturi R, Pazirandeh M, Stolow D, Huang W-Y, Wakil S** (1989) A novel cDNA extension procedure. J Biol Chem **264:** 3750–3757
7. **Dale R, McClure B, Houchins J** (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18 S rDNA. Plasmid **13:** 31–40
8. **Dragon E, Sias S, Kato E, Gabe J** (1987) The genome of *Trypanosoma cruzi* contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. Mol Cell Biol **7:** 1271–1275
9. **Farrelly F, Finkelstein D** (1984) Complete sequence of the heat shock-inducible HSP90 gene of *Saccharomyces cerevisiae.* J Biol Chem **9:** 5745–5751
10. **Feinberg A, Vogelstein B** (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem **132:** 6–13
11. **Gamborg O, Miller R, Ojima K** (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res **50:** 151–158
12. **Gubler U, Hoffman B** (1983) A simple and very efficient method for generating cDNA libraries. Gene **25:** 263–269
13. **Hickey E, Brandon S, Smale G, Lloyd D, Weber L** (1989) Sequence and regulation of a gene encoding a human 89-kilodalton heat shock protein. Mol Cell Biol **9:** 2615–2626
14. **Holmgren R, Corces V, Morimoto R, Blackman R, Meselson M** (1981) Sequence homologies in the 5′ regions of four *Drosophila* heat shock genes. Proc Natl Acad Sci USA **78:** 3775–3778
15. **Key J, Nagao R, Czarnecka E, Gurley W** (1987) Heat stress: expression and structure of heat shock protein genes. *In* D von Wettstein, N-H Chua, eds, Plant Molecular Biology. Plenum, New York, pp 385–397
16. **Kimple J, Key J** (1985) Presence of heat shock mRNAs in field grown soybeans. Plant Physiol **79:** 672–678
17. **Kroner P** (1987) Gene expression during developmental and auxin-induced growth transitions in soybean seedlings. PhD dissertation, University of Georgia, Athens
18. **Lindquist S, Craig E** (1988) The heat shock proteins. Ann Rev Genet **22:** 631–677
19. **Lutcke H, Chow K, Mickel F, Moss K, Kern H, Scheele G** (1987) Selection of AUG initiation codons differs in plants and animals. EMBO J **6:** 43–48
20. **Maniatis T, Fritsch E, Sambrook J** (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. **Moore S, Kozak C, Robinson E, Ullrich S, Appella E** (1987) Cloning and nucleotide sequence of the murine hsp84 cDNA and chromosome assignment of related sequence. Gene **56:** 29–40
22. **Morch M, Drugeon G, Zagorski W, Haenni A** (1986) The synthesis of high-molecular-weight proteins in the wheat germ translation system. Methods Enzymol **118:** 154–164
23. **Nagao R, Key J** (1989) Heat shock protein genes of plants. Cell Cult Somatic Cell Genet Plants **6:** 297–328
24. **Neumann D, Nover L, Parthier B, Rieoer R, Scharf K, Wollgiehn R, Nieden U** (1989) Heat shock and other stress response systems of plants. Biol Zbl **108:** 1–156
25. **O'Farrell P** (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem **250:** 4007–4021
26. **Rebbe N, Ware J, Bertina R, Modrich P, Stafford D** (1987) Nucleotide sequence of a cDNA for a member of the human 90-kDa heat-shock protein family. Gene **53:** 235–245
27. **Rose W, Wettenhall R, Kudlicki W, Kramer G, Hardesty B** (1987) The 90-kilodalton peptide of the heme-regulated eIF-2a kinase has sequence similarity with the 90-kilodalton heat shock protein. Biochemistry **26:** 6583–6587
28. **Schlesinger M** (1986) Heat shock proteins: the search for functions. J Cell Biol **103:** 321–325
29. **Vieira J, Messing J** (1987) Production of single-stranded plasmid DNA. Methods Enzymol **153:** 3–11
30. **Wu C, Caspar T, Browse J, Lindquist S, Somerville C** (1988) Characterization of an HSP70 cognate gene family in *Arabidopsis.* Plant Physiol **88:** 731–740