Biochemistry

 Volume 30, Number 1 January 8, 1991

Perspectives in Biochemistry

Transcriptional Regulation of the Heat-Shock Response: A Plant Perspective†

William B. Gurley
Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611
Joe L. Key*
Departments of Botany and Biochemistry, The University of Georgia, Athens, Georgia 30602
Received June 6, 1990; Revised Manuscript Received August 10, 1990

The response of organisms to high-temperature stress has been widely researched since the heat-shock (HS) response was first observed in *Drosophila* (Ritossa, 1962). The HS phenomenon is characterized by dramatic and rapid changes in both transcription and translation with the onset of heat stress [see Ashburner and Bonner (1979)]. The HS response occurs in most, if not all, organisms ranging from bacteria and lower eukaryotes to mammals and plants. Major physiological and biochemical aspects of the HS response have been covered in many reviews [e.g., Browder et al. (1989), Craig (1985), Lindquist (1986), Lindquist and Craig (1988), Nagao et al. (1986), Nagao and Key (1989), Neidhardt (1987), Neumann et al. (1989), Nover (1987), Pelham (1985), Schlesinger (1986), and Tanguay (1988)] and monograph/symposium volumes [e.g., Atkinson and Walden (1985), Morimoto et al. (1990), Pardue et al. (1989), and Schlesinger et al. (1982)]. The HS response is evolutionarily highly conserved, suggestive of an essential function or functions for survival of organisms exposed to excessively high temperatures (e.g., thermotolerance).

The HS response after an 8–12 °C shift up from the normal growing temperature is characterized by very fast induction of HS gene transcription coupled with a precipitous decline in the transcription of most other genes. The response is enhanced further by the selective translation of HS mRNAs at HS temperatures (or rapid turnover of non-HS mRNAs in bacteria and yeast for example), resulting in the selective and rapid accumulation of heat-shock proteins (hsp) at the elevated temperature. Plants generally have a more complex set of hsps than most other organisms; the complexity and abundance of families of low molecular weight hsps in plants are particularly unique (Key et al., 1981; Mansfield & Key, 1987). The response is fairly short-lived, ranging from 10–15 min in bacteria to a few hours in most higher eukaryotes. The transient nature and the rapid, specific induction of a set of genes are the hallmarks of the HS response in all organisms studied to date.

The focus of this paper is on the HS response in plants relating to the regulation of transcription of the HS genes. Since several aspects of the knowledge base are greater in other systems, this paper does provide a range of comparative analyses. Emphasis is placed on promoter analysis of HS genes, since this area of investigation has progressed at a faster pace than has the understanding of the mechanism of down-regulation (also referred to as self-regulation or autoregulation) of HS gene expression. The ability to dissect this regulatory mechanism is further complicated by the fact that many HS genes, in addition to their induction during heat stress, often exhibit low-level constitutive expression and are also responsive to developmental cues. Transduction of the HS stimulus into functional biochemistry remains the least well understood process [see Parker-Thornburg and Bonner (1987), and Tanguay (1988)], and conseuqently it is not reviewed here.

Transcriptional Regulation of Plant Heat-Shock Genes Is Similar to That of Other Higher Eukaryotes. In plants, as well as other eukaryotes, the rapid accumulation of HS-specific mRNAs is primarily due to a large increase in the transcription of HS genes (Schoffl & Key, 1982; Schoffl et al., 1987; Kimpel et al., 1990). In soybean, a new set of mRNAs representing 20% or more of the total poly(A)+ RNA is induced by 2 h of HS. These HS-specific mRNAs generally correspond to the small HS proteins (hsps) and accumulate to levels estimated at 20000 copies each per cell (Schoffl & Key, 1982). The dramatic thermoinducibility of HS genes has been attributed to the presence of cis elements designated as heat-shock consensus elements (HSEs), which are located in the 5′ flanking sequences. The occurrence of multiple HSE-like motifs within the TATA-proximal and upstream regions of plant genes strongly implied that the mechanism of tran-

† The research from the authors' laboratories was supported by DOE Contract DE-FG09-86ER13602 to J.L.K. and by NIH Grant GM39732 to W.B.G.
* To whom correspondence should be addressed.

0006-2960/91/0430-1$02.50/0

scriptional activation of HS genes was conserved between animals and higher plants (Czarnecka et al., 1985; Nagao et al., 1985; Raschke et al., 1988; Rochester et al., 1986; Schoffl et al., 1984; Takahashi & Komeda, 1989; Wu et al., 1988). The HSE was originally identified in *Drosophila hsp70* by deletion analysis and sequence comparison as a 14 base pair (bp) palindrome with the sequence 5′-CTnGAAnnTTCnAG-3′ (Mirault et al., 1982; Pelham, 1982). More recent evidence has modified the definition of the HSE to be a cluster of three or more inverted repeats of the sequence 5′-nGAAn-3′ arranged with alternating polarity (Amin et al., 1988; Xiao & Lis, 1988; Lis et al., 1990). The HSE was subsequently shown to serve as a binding site for the heat-shock transcription factor (HSF) that mediates the thermal activation of transcription (Goldenberg et al., 1988; Parker & Topol, 1984; Wu, 1984a,b).

The universality of the HS response in terms of the mechanism of transcriptional induction has been well documented by the heterologous expression of the *Drosophila hsp70* gene in amphibian, nematode, insect, and vertebrate systems [for reviews, see Nover (1987), and Tanguay (1988)]. The first direct demonstration that plants also shared this highly conserved mode of promoter activation was seen in the heat-inducible transcription of the *Drosophila hsp70* promoter in transgenic tobacco callus tissue (Spena et al., 1985) and later confirmed in regenerated tobacco plants (Spena & Schell, 1987). In these studies a 457-bp fragment from the *Drosophila hsp70* promoter was fused to the neomycin phosphotransferase II coding region. The *hsp70* fragment included 258 bp of 5′ flanking sequences containing four HSEs known to bind *Drosophila* HSF in vitro (Topol et al., 1985). The level of HS induction in tobacco ranged from 5 to 50 times the basal activity. Direct involvement of the HSE was demonstrated by Strittmatter and Chua (1987), who showed that two overlapping HSE-like sequences from the TATA-proximal promoter domain of the soybean *hsp17.3B* gene (−76 to −41; Figure 1) conferred thermoinducibility to heterologous promoters in transgenic tobacco plants. Thermal induction was obtained when the two overlapping HSEs were placed at −46 upstream of a truncated cauliflower mosaic virus (CaMV) 35S promoter or when they were fused to the pea *rbcS-3A* promoter. When the HSEs were placed in the upstream region of the *rbcS-3A* promoter at positions −49 and −410, HS induction was obtained in each case, demonstrating both the TATA-proximal and long-range activity of the soybean HSEs in plants.

This ability to act both as a TATA-proximal element and as an enhancer element was first demonstrated for animal HS genes. For example, HSEs are naturally located between −370 and −270 in the *Drosophila hsp27* gene but still retain function when placed >2 kb upstream (Riddihough & Pelham, 1986). The ability of the HSE to act at long range as an enhancer element is also seen in the *Xenopus hsp70* gene where the three most proximal HSEs are positioned between −260 and −100 (Bienz & Pelham, 1986). Removal of multiple upstream HSEs from position −298 to −181 of soybean HS gene *Gmhsp17.3B* by 5′ deletion eliminated heat inducibility (Baumann et al., 1987). Analysis of the closely related *Gmhsp17.5E* gene of soybean by internal deletions showed that this promoter could be fully inactivated by deletion of either the TATA motif, the TATA-proximal HSE, or a low-homology HSE located at −100 to −87 (Czarnecka et al., 1989).

In animal systems, a minimum of one HSE positioned in either orientation upstream of the TATA motif is generally sufficient for thermoinducibility [Amin et al., 1987; Kay et al., 1986; Wei et al., 1986; for review, see Tanguay (1988)].

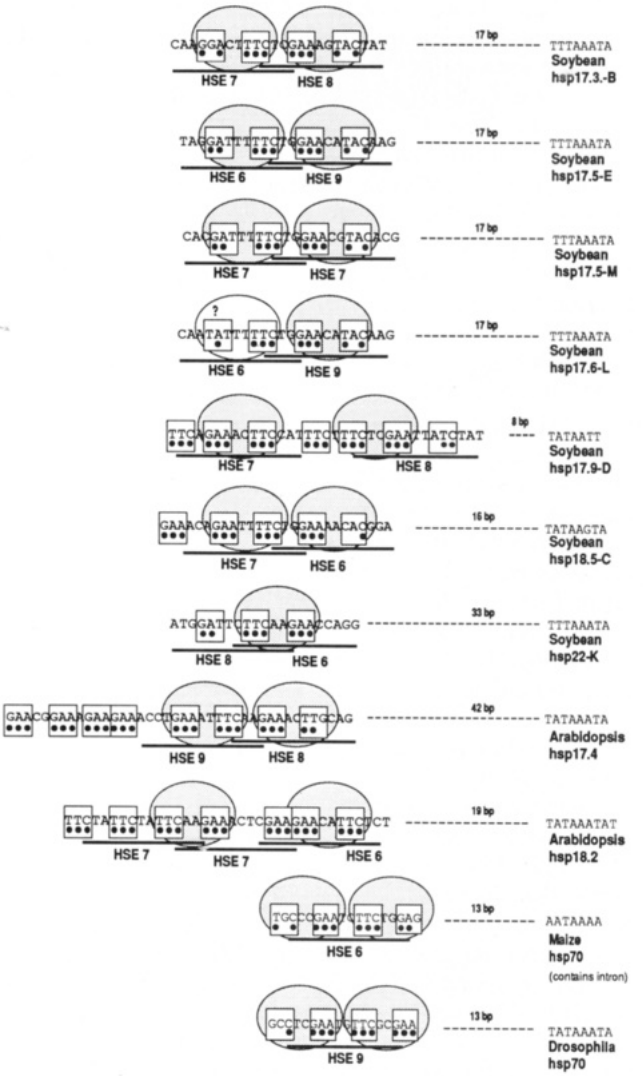

FIGURE 1: TATA-proximal HSEs of plant HS genes. Partial homology to the HSE as originally defined (Pelham, 1982) is indicated by the underline. The number of matching bases out of 10 is shown underneath the line (e.g., HSE at site 1 of the *hsp17.5E* gene is shown under the line to have a 9 out of 10 homology or HSE 9). The conserved bases of the 5-bp core repeats are boxed (GAA or TTC). Both perfect and imperfect repeats are shown with dots designating matches to the consensus. Hypothetical HSF trimer binding sites are indicated by shaded ovals. Only two contact points are shown although three may be possible in certain cases. The nonshaded oval in soybean *hsp17.6L* represents questionable HSF binding. The *Drosophila* HSE at site 1 is included for comparison. DNA sequences are taken from the following references: soybean genes *Gmhsp17.3B* (Schoffl et al., 1984), *Gmhsp17.5E* (Czarnecka et al., 1985), *Gmhsp17.5M* (Nagao et al., 1985), *Gmhsp17.6L* (Nagao et al., 1985), *Gmhsp17.9D* (Raschke et al., 1988), *Gmhsp18.5C* (Raschke et al., 1988), and *Gmhsp22K* (R. Nagao, personal communication); *Arabidopsis* genes *hsp17.4* (Takahashi & Komeda, 1989) and *hsp18.2* (Takahashi & Komeda, 1989); maize gene *hsp70* (Rochester et al., 1986); *Drosophila hsp70* (Pelham, 1982).

However, in most native promoter configurations, a single HSE confers only slight activity; full transcriptional activation generally requires at least two HSEs.

The question of how many HSEs are sufficient to obtain thermal induction has also been addressed in plants. In experiments by Schoffl et al. (1989), synthetic oligomers homologous to the TATA-proximal overlapping HSEs of *Gmhsp17.3B* (Figure 1) were placed upstream of a truncated CaMV 35S promoter. When assayed in transgenic tobacco plants, the overlapping HSE pair was able to confer HS induction, with increasing activity as the number of overlapping

HSE oligomers was increased up to five copies. In similar experiments by Wing et al. (1989), segments of various lengths from the *Drosophila hsp70* promoter were placed upstream of a minimal promoter (−78 to +8) from the T-DNA gene *iaaH* (from a tumor-inducing plasmid of *Agrobacterium tumefaciens*), which retained a CCAAT-like motif in addition to TATA and the start site for transcription. A single HSE (HSE1; −72 to −35; Figure 1) from the *hsp70* promoter was sufficient for a 5–8-fold thermal induction of the *iaaH* promoter in transformed tobacco calli. DNA fragments containing additional upstream HSEs from the *hsp70* gene resulted in activities ranging from 5 to 30 times the constitutive level. Thus, it appears that, in the proper context, a single nonoverlapping HSE is sufficient to confer HS activation of plant genes.

Significance of Distal HSEs. Far upstream HSEs (located upstream of approximately −120) seem to be of minor importance in plants. In a deletion analysis of the *Gmhsp17.5E* promoter, distal HSEs made only minor contributions to the overall activity of the promoter (Czarnecka et al., 1989). The region from −411 to −358 contains six HSEs with homologies ranging from six out of ten to eight out of ten matches with the *Drosophila* HSE. Yet, removal of these sequences by 5′ deletion resulted only in a 25% loss of activity when assayed in transformed sunflower tumors. A more substantial contribution of a distal HSE to promoter activity was obtained with the *Gmhsp17.3B* gene (Baumann et al., 1987). When the activity was assayed in transgenic tobacco plants, the removal of an HSE (nine out of ten match) located between −173 and −160 resulted in an approximately 50% loss of thermoinducible activity as compared to the activity of a 5′ deletion to position −298. The distance between the distal HSE and TATA is not as great in *Gmhsp17.3B* as that of the HSE clusters in *Gmhsp17.5E*. The greater contribution of the *Gmhsp17.3B* distal HSE is therefore consistent with the general observation that HSEs are diminished in strength with increasing distance from TATA (Amin et al., 1987; Pauli et al., 1986).

An unresolved question regarding the function of HSEs as enhancer elements in plants is whether or not there is a requirement for TATA-proximal elements, such as CCAAT, in addition to the TATA motif. This question is posed since there appears to be a difference in the requirement for additional TATA-proximal elements between *Drosophila* and vertebrate systems. In *Xenopus hsp70*, two HSEs located between positions −260 and −100 of the *hsp70* gene require a CCAAT box in the TATA-proximal domain for thermal activation of transcription when transfected into mammalian cells (Bienz & Pelham, 1986). The CCAAT box is also essential for thermal activation when HSEs from *Xenopus hsp70* are placed upstream of the β-globin promoter. This requirement for a CCAAT box and other elements in the TATA-proximal domain is also seen for the stimulation of the β-globin promoter by the SV40 enhancer (Dierks et al., 1983; Grosveld et al., 1982). In contrast, the *Drosophila hsp27* gene contains a cluster of three HSEs (centered at approximately −500) that are capable of activating transcription at long range in the absence of other elements between it and TATA when analyzed by transient assay in *Drosophila* tissue culture (Riddihough & Pelham, 1986). The *Drosophila hsp27* gene, however, is not active in mammalian cells. This question has not been directly addressed in plants. Although the issue will not be resolved without further experimentation, it may be that HSEs in plants do not require additional elements TATA-proximal for long-range activation by HSEs, since other plant enhancer-like sequences do not appear to require CCAAT or other elements in the TATA-proximal domain (Bruce et al., 1988). However, when the TATA-proximal overlapping set of two HSEs was deleted from the *Gmhsp17.5E* promoter, thermal induction was not seen even though TATA-distal clusters of HSEs were still intact (Czarnecka et al., 1989), suggesting that the total configuration of the promoter may be important in ways that are not understood.

TATA-Proximal Promoter Structure. A compilation of TATA-proximal HSEs present in the promoters of plant HS genes is shown in Figure 1. It is evident that an overlapping configuration of HSEs is often present within 20 bp upstream of TATA. Although this pattern of HSE location is common in plant as well as animal HS genes, HSE function is not limited to its assignment to this position, nor is the overlapping HSE configuration a requirement for activation in plants. For example, maize *hsp70* lacks the overlap motif, and the experiments of Wing et al. (1989) discussed above show moderate HS-inducible activity by a single nonoverlapping HSE from the *Drosophila hsp70* gene.

Recent studies in *Drosophila* have identified the core component of the HSE to be the 5-bp repeat 5′-nGAAn-3′, which is present as a head to head inverted repeat in the HSE as originally defined (Perisic et al., 1989). In view of this finding, a reappraisal of the HSE as an array of core repeats is warranted (Perisic et al., 1989; Xiao & Lis, 1988, 1989; Lis et al., 1990). Binding studies with HSF and synthetic HSEs indicate that at least two of the 5-bp core repeats are needed to form a stable complex in vitro (Perisic et al., 1989). The core repeats may be arranged either head to head or tail to tail for binding to occur. Naturally occurring (native) HSEs are frequently embedded in an array of perfect and imperfect core repeats. Since imperfect core repeats are often present adjacent to perfect repeats, it is assumed that the HSF is able to bind the imperfect core repeats if they are adjacent to perfect repeats of the opposite orientation (i.e., nGAAnnTTCn). Supporting this view is the DNase I footprinting of site 2 of *Drosophila hsp70* (Topol et al., 1985), which contains only perfect–imperfect pairs (Figure 2). Within the core, the guanine residue is highly conserved and is one of the sites of close interaction with the HSF (Perisic et al., 1989; Shuey & Parker, 1986). The HSF of *Drosophila* and yeast exists as a trimer in solution and when bound to DNA. Trimerization is thought to involve a 35 amino acid α-helical region of the HSF monomer capable of forming a coiled coil. A model of HSF binding to core repeats proposed by Perisic et al. (1989) predicts that the orientation of DNA binding sites of the trimer is in the tail to tail and head to head configuration (← → ←). Efficient binding of the HSF trimer to DNA requires that at least two of the monomer subunits be bound to HSE core repeats. Presumably synergistic interactions are involved in binding of an imperfect repeat adjacent to an intact core.

In view of the models of Perisic et al. (1989), Xiao and Lis (1988, 1989), and Lis et al. (1990), the HSE and overlapping HSEs are best considered as tandem arrays of the 5-bp core repeat. In the case of the high-affinity HSE (Topol et al., 1985), which is TATA-proximal in the *Drosophila hsp70* promoter, the array is comprised of three perfect cores and one imperfect core (Figure 1). The two overlapping HSEs located TATA-proximal in five of the sequenced soybean HS genes are comprised of two perfect repeats formed by the overlap of two "flawed" HSEs (Figure 1). In this configuration, each of the perfect repeats is flanked by an imperfect repeat. This type of array has the potential of binding two

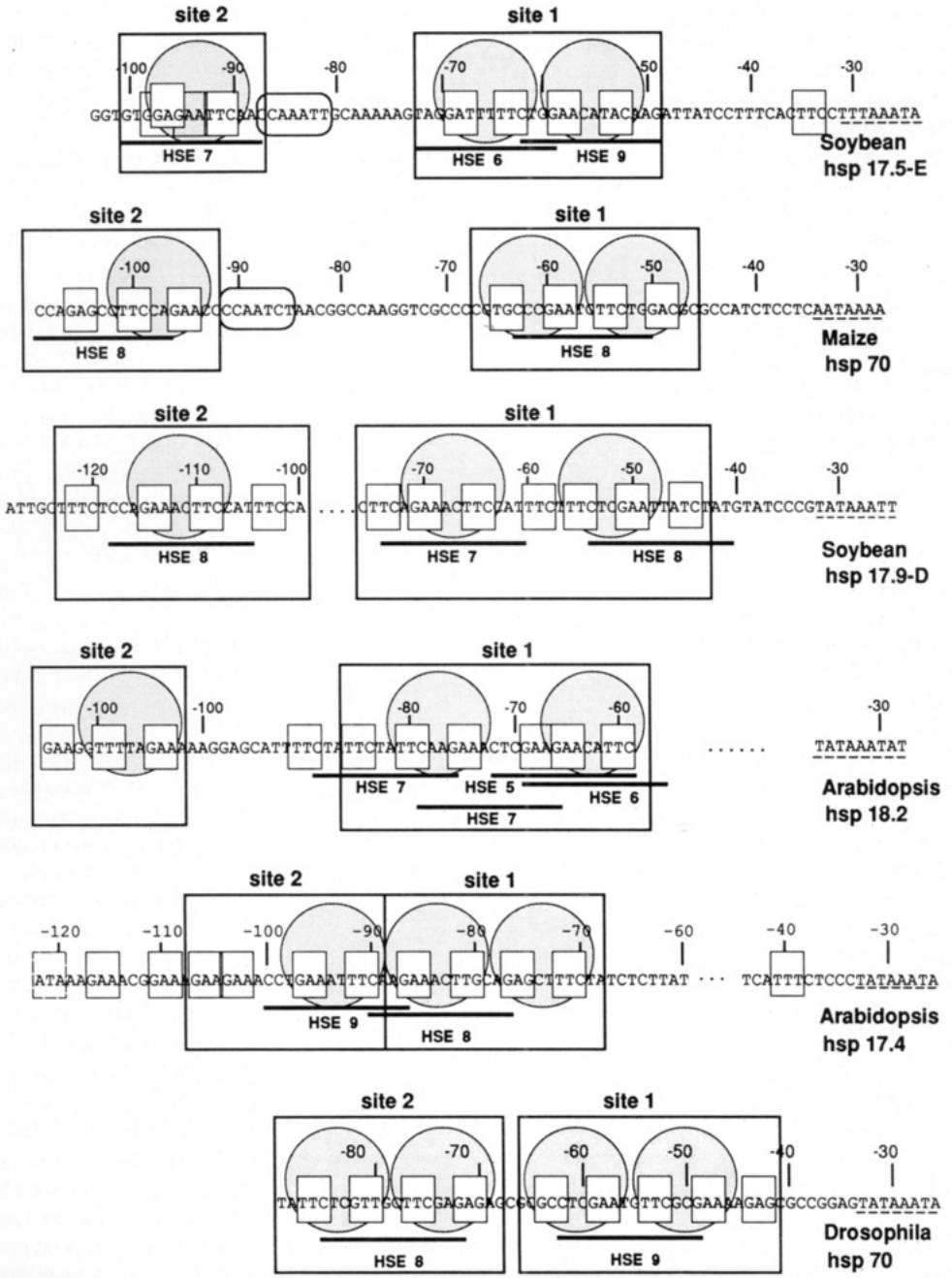

FIGURE 2: Two-site model for TATA-proximal domain of plant HS promoters. Clusters of hypothetical HSF binding sites (shaded circles) are grouped into two sites on the basis of the analogy with *Drosophila hsp70* and the mutation results for soybean *Gmhsp17.5E* (Czarnecka et al., 1989) indicating the presence of two TATA-proximal sites of HS regulation. CCAAT-like sequence motifs are noted by rounded rectangles. The legend and sources of DNA sequence information are the same as in Figure 1.

trimers of the HSF, one bound to each perfect–imperfect pair. Two HSF trimers are also predicted to be bound to the high-affinity HSE at site 1 of the *Drosophila hsp70* promoter; the difference is that, in the case of *hsp70*, the TATA-distal pair of core repeats are both perfect.

Reinterpretation of the HSE sequence (5′-CTnGAAn-nTTCnAG-3′) based on the occurrence of core repeats suggests that the configuration of two central perfect core repeats flanked by two imperfect repeats is often utilized in native HS promoters. The flanking imperfect repeats are represented in the HSE as single bases, with the highly conserved cytosine and guanine residues at either end.

The pattern of core repeats comprising native HSEs in plant (and animal) genes suggests that the optimum array of core repeats is able to bind two trimers of the HSF. Implicit in this model of the HSE is the assumption that flanking imperfect repeats are able to participate in HSF binding. Figure 1 shows the theoretical pattern of HSF trimer binding to TATA-proximal core repeat arrays. A two-trimer model for the TATA-proximal HSE can be accommodated in every case except for the soybean *hsp17.6L* and *hsp22K* genes. The exceptions imply that a single trimer also retains a degree of activity depending on its context within the promoter. It also seems likely that the two HSF trimers do not necessarily need to be juxtaposed as seen in the predicted HSF binding for soybean *hsp17.9D* and *Arabidopsis hsp18.2* genes. It seems noteworthy that in these two examples where the predicted HSF binding sites are separated, each site is comprised of two perfect core repeats, raising the possibility that cooperativity between HSF trimers may be necessary for binding to perfect–imperfect repeats.

Support for the two-trimer model of binding can be indi-

rectly inferred from the experiments of Wing et al. (1989). In this study, HS induction was obtained when the *Drosophila hsp70* HSE at site 1 was placed upstream of a CCAAT-like sequence at −78 in the truncated *iaaH* promoter. However, no inducible transcription was observed when the bases from −72 to −58 (relative to *hsp70*) were deleted from the HSE leaving three core repeats (Figure 2). This change from an array of five core repeats (three perfect, two imperfect) to three repeats is predicted to result in the loss of one HSF trimer binding site with the remaining HSF site comprised of two perfect and one imperfect repeat. Even though three perfect repeats have been shown to bind a single HSF trimer with high affinity (Perisic et al., 1989), they failed to confer heat inducibility to the chimeric *iaaH* promoter.

Two-Site Configuration of the TATA-Proximal Domain. Deletion analysis of the soybean *hsp17.5E* gene indicated that HS induction was conferred by two sites containing HSEs within the TATA-proximal domain (Czarnecka et al., 1989). This two-site model of the TATA-proximal promoter structure is analogous to sites 1 and 2 present in the *Drosophila hsp70* promoter as defined by Topol et al. (1985). In the *Drosophila* gene, site 1 is TATA-proximal and is comprised of a single HSE with relatively high affinity for purified HSF in vitro. Site 2 contains a HSE that shows 12.5-fold less affinity for HSF and is occupied after site 1 has been filled. The cooperativity involved in the binding of HSF to site 2 is postulated to form the basis of a sharply delineated switch mechanism for promoter activation in response to HS. A comparison of the HSE homologies at sites 1 and 2 of the soybean *hsp17.5E* gene promoter (Figure 2) suggests a similar arrangement of high- and low-affinity sites for HSF binding. Consistent with this interpretation are observations (Czarnecka et al., 1990) that a synthetic oligomer homologous to the overlapping HSEs at site 1 binds nuclear proteins from soybean plumules with much higher affinity than an oligomer homologous to the HSE at site 2.

In spite of the striking similarity in the pattern of HSEs present in the TATA-proximal domains of *Drosophila hsp70* and soybean *Gmhsp17.5E*, there are also notable differences. For example, the four sites of HS regulation in the *Drosophila hsp70* promoter are comprised entirely of HSEs (Topol et al., 1985). In *Gmhsp17.5E*, only the TATA-proximal site 1 is a pure array of core repeats, whereas site 2 has potential for more complex interactions. This potential for multiple regulation at site 2 is based on the presence of a low-homology HSE (7 out of 10 match) and an overlapping GT motif (5′-GGTGTGGAGAATTC-3′) with an 11 out of 14 base match with the SV40 enhancer core (Laimins et al., 1982). Similar GT motifs (box II) have also been described in the *rbcS-3A* light-regulated promoter, which binds a nuclear protein that acts to repress transcription in the dark (Kuhlemeier et al., 1987). Another GT motif found in environmentally regulated plant promoters is present in *Arabidopsis Adh*, where it is associated with the region responsible for anaerobic regulation (McKendree et al., 1990); this motif appears to be involved in close protein contacts in vivo (Ferl & Laughner, 1989). Site 2 is flanked downstream by a sequence with CCAAT box homology (5′-CAACCAAATTG-3′). Although the function of the CCAATT-like sequence is untested for *Gmhsp17.5E*, the CCAAT factor has been postulated to form an anchor for long-range interactions between HSF bound to upstream HSEs and the TATA-proximal domain in vertebrate HS genes (Bienz & Pelham, 1986).

The high degree of similarity in the organization of the TATA-proximal domains of *Drosophila hsp70* and soybean *Gmhsp17.5E* suggests that a two-site structure of the TATA-proximal promoter may be common among heat-inducible genes. A comparison of the pattern of HSEs found in the TATA-proximal domains of several plant HS genes is presented in Figure 2. The clusters of core repeats have been grouped into hypothetical sites 1 and 2 to illustrate the common occurrence of this type of arrangement. Although the TATA-proximal site (site 1) usually has, as a minimum, the number of core repeats (both perfect and imperfect) required to bind two HSF trimers, site 2 sometimes has three or less with the proper spacing and orientation, suggesting that the binding of a single HSF trimer is sufficient in this context. Perhaps interactions between proteins bound at sites 1 and 2, or between adjacent heterologous elements, stabilize HSF binding to the low-affinity site 2, eliminating the requirement for a second HSF trimer in the formation of an active module. This principle of promiscuous cooperativity may have been involved in the partial restoration of HS induction by the substitution of an AT-rich element for the HSE at site 2 of soybean *Gmhsp17.5E* (Czarnecka et al., 1989).

The requirement for a minimum of two arrays of core repeats (sites 1 and 2), and the ability of other elements to substitute, may reflect a general characteristic inherent in the design of the TATA-proximal promoter domain of eukaryotes. One interpretation of the two-site model is to equate the "sites" with promoter (or enhancer) modules as defined by Dynan (1989). According to this view a promoter module is comprised of either one or two subelements analogous to enhansons. In the case of HS genes, the enhanson would be equivalent to a HSF trimer binding site. Two enhansons would be required for the TATA-proximal module (site 1), while one would be sufficient for the upstream module (site 2). In other types of genes a minimum of two modules is usually required for activity as demonstrated for SV40 enhancer modules by Herr and Clarke (1986) and for HSV *tk* promoter modules (McKnight, 1982). This general requirement for two modules may be reflected in the organization of HSEs into sites 1 and 2 within the TATA-proximal region of native HS promoters.

Contribution of AT-Rich Elements to Activity of the HS Promoter. Six genes encoding low molecular weight hsps of soybean (*Gmhsp17.3B*, *-17.5E*, *-17.5M*, *-17.6L*, *-17.9D*, and *-18.5C*) often have blocks of AT-rich sequences located within the 5′ flanking region. These sequences are usually comprised of 14, or more, contiguous A-T bp positioned upstream of the TATA-proximal domain. The potential regulatory role of an AT-rich block was first indicated by Baumann et al. (1987) by deletion analysis of the *Gmhsp17.3B* gene in transformed tobacco plants. An AT-rich block including 14 adenine residues is located between positions −274 and −254. A 5′ deletion from −298 to −242 removed this sequence and resulted in a substantial loss of transcriptional activity.

Experiments using the *Gmhsp17.5E* sequence suggest that AT-rich sequences found upstream of a variety of inducible plant genes (Bustos et al., 1989; Datta & Cashmore, 1989; Jofuku et al., 1987; Jordano et al., 1989), including HS genes, may function as enhancer-like elements. Several AT-rich blocks located between −120 and −153, and far upstream from −830 to −970, of the *Gmhsp17.5E* promoter have been shown to bind nuclear proteins isolated from soybean plumules in vitro by DNase I footprint and gel mobility retardation assays (Czarnecka et al., 1990). Nuclear proteins also form stable complexes in vitro with synthetic oligomers homologous to AT blocks present in the *Gmhsp17.5E* and *Gmhsp17.3B* promoters (Czarnecka et al., 1990; Czarnecka et al., unpublished). These protein–DNA complexes are specifically competed with for

FIGURE 3: Plant HS cognate genes. Shown are the TATA-proximal regions of three genes encoding HS proteins that are either expressed under conditions other than HS, constitutively expressed, or not expressed at detectable levels. All protein-coding regions are interrupted by a single intron. CCAAT-like motifs are designated by rounded rectangles. The legend is the same as in Figure 1. DNA sequences are taken from the following references: soybean *Gmhsp26A* (Czarnecka et al., 1988) and *Arabidopsis hsp70-1* and *hsp70-2* (Wu et al., 1988).

binding by synthetic oligomers homologous to AT-rich sequences present in several other inducible genes of soybean including β-conglycinin, leghemoglobin, and lectin (Czarnecka, Ingersoll, and Gurley, unpublished). Competition for in vitro binding by a wide variety of AT-rich sequences indicates that a common protein, or a family of proteins, exists in plants that is capable of interacting with 5′ flanking regions of a variety of genes in addition to those regulated by HS. The best evidence that these AT-rich blocks may act as enhancer-like elements if the finding that multimers of a 33-bp synthetic oligomer homologous to an AT-rich block upstream of the *Gmhsp17.5E* promoter (AT composite; repeated several times between −830 and −970) are able to stimulate heterlogous expression of a truncated maize *Adh1* promoter when placed at position −140 (Czarnecka, Ingersoll, and Gurley, unpublished). It is now clear that certain AT-rich sequences play an active role in transcriptional activation of one group of HS promoters in soybean.

Similar AT-rich elements have been found upstream of the β-phaseolin gene of French bean (Bustos et al., 1989). A 55-bp synthetic oligomer homologous to this element was placed upstream of a minimal CaMV 35S promoter and shown to stimulate transcription of a β-glucuronidase (GUS) reporter gene in transgenic tobacco plants. Expression was strongest in roots of tobacco seedlings and in the cotyledons and the root–hypocotyl axis of mid-maturation to late-maturation embryos. If this element is the same as the HS AT element, its role may be, by analogy, to confer a developmental bias on HS gene expression in the embryo and seedling stages of soybean development.

Although blocks of AT-rich sequences are common to several HS genes encoding low molecular weight hsps of soybean, their distribution among plant HS genes is not universal. For example, *hsp18.2* contains a single cluster of 14 thymine residues between −583 and −568, but *Arabidopsis hsp17.4* contains no obvious AT-rich elements (Takahashi & Komeda, 1989). Likewise, maize *hsp70* also seems to lack AT-rich elements (Rochester et al., 1986). Thus, the prevalence of AT-rich elements in the sequenced HS genes of soybean may reflect the abundance of a specialized class of the low molecular weight hsp mRNAs isolated from the original cDNA library prepared from seedling RNA (Schoffl & Key, 1982) that are translated into the most abundant hsps of soybean.

Activation of HSF. In *Drosophila* and humans, both DNA binding and transcriptional activation of the HSF are enhanced by HS (Goldenberg et al., 1988; Kingston et al., 1987; Mosser et al., 1988; Sorger et al., 1987; Zimarino & Wu, 1987). This stands in contrast to activation in yeast where binding of HSF to the HSE occurs under both HS and non-HS conditions (Jakobsen & Pelham, 1988; Sorger et al., 1987) with the previously bound HSF then becoming competent to mediate transcriptional activation after HS. Although the exact mechanism of activation of the HSF is not known, there is a strong correlation between activation of transcription and phosphorylation of the HSF in both yeast and human cells (Larson et al., 1988; Sorger et al., 1987; Sorger & Pelham, 1988).

As yet there is no experimental evidence regarding the role of phosphorylation in the direct activation of plant HSF. However, Krishnan and Pueppke (1987) proposed that the phosphorylation of cellular proteins may provide a mechanism for signal transduction for sensing the stress associated with elevated temperatures in plants. This hypothesis is based on the rapidity (<10 min) of protein phosphorylation in soybean in response to heat stress. The rapid kinetics of phosphorylation correlates with the appearance of HS mRNA within 3–5 min of HS (Nagao et al., 1986; Schoffl & Key, 1982). Recent experiments with soybean nuclear extracts indicate that HSE-binding activity increases from 2- to 3-fold in response to HS (Czarnecka et al., 1990), which corresponds well with the 2–7-fold increase reported for HSF from HeLa cells (Goldenberg et al., 1988; Kingston et al., 1987). This observation indicates that higher plants are more closely related to animals than to yeast with regard to the mechanism of HS-induced activation of transcription.

Promoters of Plant Cognate Genes. There are numerous examples of genes encoding hsps that are expressed constitutively and under non-HS conditions in *Drosophila*. These genes form the broad class referred to earlier as HS cognates. Most of the cognate genes in animals have one intron and are very weakly induced, if at all, by HS. In plants, the distinction between cognate and HS genes is not as sharp in that several hsp-encoding genes contain an intron, and both are exposed under non-HS conditions, often in a tissue-specific manner (Duck et al., 1989). For example, the soybean *Gmhsp26A* and the *Arabidopsis hsp70-1* genes both contain an intron, and both are induced to a moderate extent (5–20-fold) by temperature stress (Czarnecka et al., 1984; Edelman et al., 1988; Wu et al., 1988). The tomato *hsp70* gene is expressed under non-HS conditions in root tip meristems, stems, leaves, the vascular system of the ovary, and the inner integument of developing seeds (Duck et al., 1989). The heat inducibility of this specific tomato *hsp70* gene has not been determined. A notable case of heat inducibility in HS genes containing an intron is found in two maize *hsp70* genes (Rochester et al., 1986; Shah et al., 1985) where expression is increased 40–60-fold. This high level of HS induction of maize *hsp70* warranted its inclusion in the listing of HS promoters shown in Figure 1.

The 5′ flanking sequences of three other plant cognate genes are presented in Figure 3. The HSE-like sequences are designated along with the arrays of core repeats. In each of these genes a HSE-like sequence is present. In the case of soybean *Gmhsp26A* (Czarnecka et al., 1988), two core repeats are present but overlap such that the GAA/TTC regions are adjacent instead of separated by the optimal 2 bp. In terms of the model of HS promoter structure proposed above, this gene should show very poor HS inducibility. Yet, transcriptional activity is enhanced 10–20-fold in response to elevated temperatures (Czarnecka et al., 1984; Edelman et al., 1988). Even though this gene is activated by HS, a variety of other stresses induce activity, including treatment with abscisic acid, 2,4-dichlorophenoxyacetic acid, poly(ethylene glycol), polyamines, sodium arsenite, and heavy metals (Czarnecka et al., 1988; Edelman et al., 1988). Apparently, the unique configuration of promoter elements within the 5′ flanking region of *Gmhsp26A* creates a thermoinducible promoter. However, since the modular structure of this promoter has not been determined, it is premature to draw conclusions regarding the mechanism of induction or the role of low-homology HSEs.

In *Arabidopsis* there are at least 12 hsps related to hsp70 (Wu et al., 1988). The majority of these are expressed under both HS and non-HS conditions in leaves, and at least three of these genes have introns. Expression of one cloned member of this family, *hsp70-1*, is increased 4–5-fold in response to HS. As seen in Figure 3, a low-homology HSE is positioned in the TATA-proximal domain that has a core array of a single perfect repeat flanked by sequences that match only 1 bp out of the core consensus. This HSE is predicted to accommodate only a single HSF trimer. The other *hsp70* gene, *hsp70-2*, contains an overlapping HSE in the TATA-proximal domain, but this gene is apparently not expressed. It has been proposed that the HSE cluster is positioned too close to TATA since the shortest distance reported between a functional HSE and TATA is 15 bp (Holmgren et al., 1981). Another potential problem may be the occurrence of an additional TATA-like motif immediately upstream of the overlapping HSE.

In overview, the HS cognates of plants exhibit a wide range of heat inducibility and tissue-specific expression. There do not seem to be clearly demarcated categories of heat inducibility versus constitutive or tissue-specific expression. The only common, but not unique, feature among cognates seems to be the presence of an intron within the protein coding region. The modular composition and organization of the constitutive and weakly heat-inducible promoters is largely unknown, but it is reasonable to expect the participation of promoter elements other than the HSE core repeat.

Evolution of HS Genes. The presence of introns is often considered a primitive trait in genes (Doolittle, 1987). If this is true, typical HS genes may have lost their introns in the optimization of expression under conditions of temperature stress. This type of selective pressure would not have been as great for those copies that were used in non-HS situations. Under one of several possible scenarios, the original HS genes would have contained a typical HS promoter with reliance on the HSF for heat activation. Those that were solely expressed during HS would lose the intron through evolution but retain HSF-mediated promoter activation. Those HS genes that became adapted for specialized roles in stress tolerance, or development, retained the intron but gradually lost the HSF-mediated mode of promoter activation. The low-homology HSEs in cognates may simply represent relics that no longer serve as the primary determinants of promoter activation. The differences between animals and plants in the patterns of expression of HS genes containing introns could reflect a differential in efficiencies of mRNA splicing between plants and animals. If plants were able to process mRNAs under conditions of elevated temperatures more efficiently than animals, the selection pressure toward intron removal would have been less intense. Hence, the characterized maize *hsp70* genes may be viewed as primitive since they contain an intron yet exhibit strong HS induction.

Although there is little evidence regarding the effect of temperature stress on the processing of plant RNA, splicing of primary transcripts in *Drosophila* and mouse cells is known to be inhibited by HS (Kay et al., 1987; Lindquist, 1986; Yost & Lindquist, 1986; Yost et al., 1990). A possible origin of this proposed difference between plants and animals might lie in the fact that, in general, plants have sought the sun and are nonmotile; this may have forced plants to better adapt to high temperatures, whereas animals are motile and can seek relief from temperature extremes in the shade, of course, of plants.

Self-Regulation of Heat-Shock Gene Expression. As noted above, the expression of HS genes is transient, ranging from minutes in *Escherichia coli* to a few hours in higher eukaryotes. This transient expression has been referred to as self-regulation, feedback inhibition, or autoregulation. While there is no a priori reason to expect common mechanisms to relate to the self-regulation of HS gene expression in different groups of organisms, the high degree of conservation of the basic HS response and of some hsps is suggestive that such might be the case.

A number of parameters of HS gene expression in plants have been experimentally determined to aid in elucidating the underlying mechanism(s) of regulation. In soybean seedlings, there is rapid activation of transcription of HS genes, evidenced by detection of HS mRNAs on Northern blots within 3–5 min after HS (Schoffl & Key, 1982). Nuclear runoff experiments (Schoffl et al., 1987; Kimpel et al., 1990) indicate that the accumulation of high levels of HS mRNAs is due primarily to enhanced transcription of the HS genes. The maximum level of HS mRNAs is achieved after 1–2 h of HS followed by a gradual decline in HS mRNA levels. Despite this decline, most of these HS mRNAs remain detectable on Northern blots after 12 h of continuous HS. Transcription of these genes is not detectable after 4 h of continuous HS. Thus, the continued presence of HS mRNAs at HS temperatures reflects the fact that HS mRNAs are quite stable at high temperatures. However, if seedlings are returned to the normal growth temperature after a 2- or 3-h HS, the levels of HS mRNAs decline rapidly, reaching undetectable levels within about 4 h (Nagao et al., 1986; Key et al., 1987). A similar differential stability of HS mRNA at normal versus HS temperatures occurs in *Drosophila* cells [e.g., Petersen and Lindquist (1988)]. Even though the relative transcription rates of HS genes appear to be the primary factors in determining HS mRNA levels in eukaryotes, other factors such as differential stability of the mRNAs undoubtedly contribute to the actual level of HS mRNAs and hsps. While this observation can be generalized for all HS mRNAs, the kinetics of the response vary somewhat among the different families of HS mRNAs (Kimpel et al., 1990; Petersen & Lindquist, 1988).

Plant seedlings undergo a new round of HS mRNA synthesis and accumulation when they are subjected to a second (or third) HS (e.g., 40–45 °C for 1–2 h) following a 3-h recovery at the normal growing temperature. Thermotolerance is conferred by an initial permissive HS (e.g., 40 °C, 2 h; Lin et al., 1984) since an initial HS at 45 °C for 2 h is lethal. The actual level of the HS mRNAs usually does not reach that

achieved during the initial HS period, but substantial levels do accumulate (70% during a second HS and somewhat lower levels during a third round of HS). As in the initial HS, transcriptional activation of HS genes occurs during each subsequent cycle; however, the measured level of transcription is significantly lower than during the initial HS, and it persists for a shorter time (Kimpel et al., 1990). Thus, the parameters that influence levels of HS mRNAs must vary with each successive cycle of HS. These experiments also demonstrate that HS gene transcription is readily reactivated after cessation of detectable transcription following an initial HS. A series of experiments were done with soybean seedlings by using a gradual increase in temperature, 3 °C per hour, rather than an abrupt shift to induce the HS state. The level of HS mRNAs gradually increased to significant levels up to 40 °C; accumulation continued at a slower rate up to 45 °C (Kimpel & Key, 1985; Key et al., 1985). The synthesis of hsps continued throughout the gradual rise in temperatures. Clearly, the HS response is not merely a response to abrupt shifts in temperatures. The ability of soybean seedlings (and undoubtedly all organisms that undergo a HS response) to regulate the level and magnitude of the response as a function of the type and duration of the heat treatment reflects the complexity of the regulatory controls operative in the HS response.

Results from experiments in the Lindquist laboratory using *Drosophila* cells (DiDomenico et al., 1982) showed a strong correlation between the accumulation of hsp70, the major hsp of *Drosophila*, and the self-regulation of hsp synthesis. The inclusion of an amino acid analogue (azetidine or canavanine) in the incubation medium resulted in the synthesis of nonfunctional hsps based on the inability of the analogue-containing hsps to undergo differential selective localization during HS. Analogue treatment also impaired the normal self-regulation of HS mRNA synthesis with HS mRNAs continuing to accumulate for several hours. The inclusion of cycloheximide (CH) in the incubation medium prior to HS to inhibit protein synthesis also extended the period of HS mRNA accumulation. Similarly, limiting the rate of synthesis of hsps by adding actinomycin D to reduce the level of HS mRNAs also resulted in a proportional delay in repression of hsp synthesis. In contrast, addition of CH 1 h after HS did not cause a significant change in the repression of HS mRNA synthesis. Taken together, these results indicate that a threshold level of hsp(s) was critical for the self-regulation of HS mRNA and hsp levels. In these experiments, the HS was accomplished at a temperature that resulted primarily in hsp70 synthesis after the cells were returned to 25 °C for all experimental analyses. The authors proposed that hsp70 was directly involved in its own synthesis at both the transcriptional and posttranscriptional levels. A similar conclusion was reached from studies of hsp70 synthesis in yeast (Craig et al., 1989), namely that transcriptional expression of the *SSA1* gene (a *hsp70* gene) is responsive to the level of its encoded protein.

Amino acid analogues induce the synthesis and accumulation of most HS mRNAs at the normal non-HS temperature in soybean seedlings (Lee, Lin, Nagao, and Key, unpublished data). However, some HS mRNAs are not induced by analogue treatment at normal growth temperatures. Only limited synthesis of hsps, primarily high molecular weight forms, occurs at normal growth temperatures in response to the analogues, which is similar to the observations in *Drosophila*. Analogue-induced accumulation of HS mRNAs is slower than in response to HS, but accumulation continues for at least 12 h, in contrast to the transient response to HS. A HS treatment in the presence of amino acid analogues induces a normal HS response in that a full complement of HS mRNAs and hsps is synthesized; however, accumulation continues for a longer time period than when HS is accomplished in the absence of the analogues. CH at a concentration that inhibits protein synthesis by 60–70% also extends the time by a factor of at least 2 during which HS mRNAs accumulate. These observations with the plant system generally parallel those of the Lindquist laboratory using *Drosophila*. In soybean seedlings, in contrast to *Drosophila* cells, the most abundant hsps range from 15 to 18 kDa; this group includes about 20 different but highly related hsps; the hsp70 family consists of at least 10 different proteins based on hybrid select–translation analyses, but these proteins do not accumulate to levels approaching that of the 15–18-kDa hsps. There is currently no experimental basis to attach special significance to hsp70 in the self-regulation of the HS response in plants, although hsp70 protein(s) is (are) not exlcuded from such a role. Another question about the possible role(s) of hsp(s) in the self-regulation response relates to the stability of hsps. In soybean seedlings, for example, a majority of the hsps have a half-life of many hours, possibly even days, based on pulse–chase experiments (and persistence of thermotolerance after HS). If the level of a hsp70 or another induced hsp determines self-regulation, the half-life of that hsp (or of its activity) necessarily would be much shorter than the majority in order to account for the observations during the repeat cycles of HS discussed above. It is also possible that sequestering or localization of previously synthesized hsps during the recovery from HS (i.e., incubation at 30 °C) at sites not directly involved in autoregulation may occur, resulting in a requirement for new hsp synthesis for each round of repression. This might be especially true if the site of transcriptional autoregulation was near the ribosome, perhaps involved in the folding of newly synthesized proteins. Alternatively, the sequestration inherent in the migration of some hsps to the nucleus during HS and their return to the cytosol during recovery from HS [e.g., Velazquez and Lindquist (1984) and Key et al. (1982)] may facilitate, or at least in part account for, repeated HS-induced transcription in the presence of accumulated hsps within the cell.

The strongest argument in favor of hsp involvement in self-regulation comes from studies in *E. coli*. The synthesis of hsps is transient, increasing within 1 min of the temperature shift and decreasing after 15–20 min. The activation of HS gene transcription results from a rapid increase in the level of σ^{32} (the RNA polymerase σ factor involved in HS gene transcription). The transient nature of the response is the result, at least in part, of the accumulation of a particular hsp, the dnaK protein, during HS [see Neidhardt et al. (1984), Craig (1985); and Neidhardt (1987)]. The dnaK protein is a negative regulator of the HS response. [The dnaJ and grpE hsps are involved in the negative regulation of hsp synthesis in a way similar to that of dnaK; see Gross et al. (1990) and Georgopoulos et al. (1990)]. *E. coli* strains carrying a *ts* mutation (*dnaK756*) in the dnaK protein do not turn off the HS response for several hours; consequently, these strains overproduce hsps. Revertants of this mutation regain the ability to regulate transiently the HS response. Strains of *E. coli* carrying the *dnaK* gene on an overproducing plasmid underproduce hsps during HS. Since the dnaK protein is itself a hsp, the HS response in *E. coli* is a self-regulated or autoregulated response. It is noteworthy that the dnaK protein has 45–50% amino acid identity with eukaryote hsp70 proteins. As with hsp70-related proteins in eukaryotes, the dnaK protein

is required for normal growth in *E. coli*. The dnaK protein may function via its interaction directly or indirectly with σ^{32}. A σ^{32} antibody cross-reacts with the dnaK protein, and the dnaK protein is found in association with purified preparations of RNA polymerase ($E\sigma^{70}$ in unstressed cells and $E\sigma^{32}$ in HS cells; Skelly et al., 1988). The *dnaK756* mutation that increases the rate and duration of synthesis of hsps results in a simultaneous increase in the half-life of σ^{32} by 3–5-fold (Tilly et al., 1989). The level of σ^{32} is rate-limiting to HS gene transcription under normal growth temperatures and increases rapidly upon HS. The positive control of HS gene expression thus relates to σ^{32} concentration, which is altered rapidly upon HS by both an increase in its rate of synthesis and a transient decrease in its rate of degradation (Grossman et al., 1987; Erickson et al., 1987). The cellular concentrations of σ^{32} and dnaK and their interaction thus provide for a multilevel control of the HS response in *E. coli* (Tilly et al., 1989).

Unlike the σ^{32} "transcription factor" in bacteria, the HSF is present constitutively at sufficient levels to elicit a full HS response in eukaryotes and requires only "activation" at the onset of HS. Thus, the induction of HS gene transcription at elevated temperatures between prokaryotes and eukaryotes is mechanistically quite different. Accordingly, there seems to be no particular reason to believe that the transient expression, or self-regulated expression, of HS genes will be controlled by the same mechanism in different groups of organisms. The conserved amino acid homology between eukaryote hsp70 proteins and the *E. coli* dnaK protein, in combination with the observations on their possible roles in self-regulation of the HS response, provides a basis for assuming that some commonality of mechanism may exist, however.

Since even the specific role of the dnaK (or dnaJ/grpE) protein in autoregulating the HS response in *E. coli* is not known, several possibilities for self-regulation of the HS response in eukaryotes should be considered. For example, the HSF may revert transiently to an inactive state; this could be accomplished by a reversal of the activation mechanism, namely changes in conformation and/or level of phosphorylation. The reversion of HSF to its pre-HS inactive state is an especially attractive idea in experimental situations where a brief severe HS is given (DiDomenico et al., 1982; Nagao et al., 1986; Key et al., 1987), followed by analyses of HS mRNA and hsp synthesis at normal growth temperatures. Such a model for self-regulation may not be as attractive when HS gene transcription shuts off during continuous HS; under these conditions some hsp such as hsp70, which is known to interact with a number of cellular proteins, could interact directly or indirectly with HSF to cause loss of HSE binding activity. HSF binding to the HSE is known to decrease with extended HS in animal cells (Mosser et al., 1988). Alternatively, hsp(s) might interact with other components of the transcription complex to inhibit further HS gene transcription. However, the possible involvement of hsp(s) in the deactivation of HSF is purely speculative. This hypothetical protein involved in downregulation would necessarily have a short half-life of activity or localization at the active site, particularly in systems such as plants where a HS response is achieved during successive cycles of HS and recovery as noted above.

Any number of variants on these themes could result in the transient expression of HS genes based on the number of different cis elements and trans-acting factors involved in the expression of HS genes (discussed above in detail). It is possible also that negative regulatory factors such as those described by Park and Craig (1989), which bind in an adjacent/overlapping position to a functional HSE in the *SSA1*

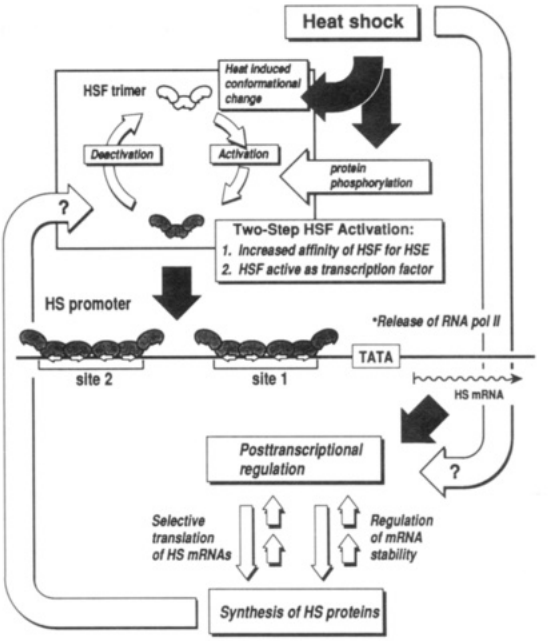

FIGURE 4: Model for regulation of HS gene expression. This model represents an integration of current views regarding the mechanism of transcriptional regulation of thermally inducible genes in higher eukaryotes. Implicit is the assumption that the thermal sensor of HS is multicomponent with elevated temperatures affecting processes involved at several regulatory points including activation of HSF, selective translation of HS mRNAs, and stability of HS mRNA. Downregulation by hsps has been postulated to counteract the stimulatory effects of HS at both the transcriptional (DiDomenico et al., 1982; Craig et al., 1989) and posttranscriptional levels (Yost et al., 1990). Black arrows connect key steps in hsp induction.

hsp70 gene of yeast, might accumulate and play a role in transient expression. Some protein corresponding to the hypothetical factor "x", which is postulated to play a role in the arrest of transcription on a *Drosophila hsp70* gene prior to HS (Rougvie & Lis, 1988), might also undergo cyclic modification or change in concentration and contribute to self-regulation of the HS response.

Concluding Remarks. A hypothetical model outlining the major regulatory control points in higher eukaryotes is summarized in Figure 4. The effects of HS are sensed in activation of the HSF for control of transcriptional expression and in posttranscriptional processes. In human cells, activation is believed to be a two-step process involving an increase in HSF binding to the HSE and an additional modification of the HSF to render it transcriptionally active. Support for this two-step activation procss is derived from the demonstration that HS-induced HSE-binding activity of HSF is separated from transcriptional activation in murine erythroleukemia cells (Hensold et al., 1990). It has been proposed (Larson et al., 1988; Hensold et al., 1990; Mosser et al., 1990; Zimarino et al., 1990b) that activation of HSE binding is due to a heat-induced change in the conformation of the HSF that is supported by the activation of HSE binding in vitro by heat (Larson et al., 1988), low pH, nonionic detergent (Mosser et al., 1990), and the lack of a requirement for new protein synthesis for HSF activation at high temperatures in *Drosophila* (Zimarino & Wu, 1987). The requirement for protein synthesis for activation of HSE-binding activity at intermediate temperatures in *Drosophila* complicates this simple model, suggesting that either a step in the activation pathway is labile at intermediate induction temperatures or newly synthesized HSF has a lower temperature threshold for activation than preexisting pools of HSF (Zimarino et al., 1990a). The in-

duction of the HSF to become transcriptionally active is closely correlated with HSF phosphorylation, which results in a reduction in electrophoretic mobility of HSF–DNA complexes in yeast and human cells (Sorger & Pelham, 1988; Larson et al., 1988). In murine erythroleukemia cells, HSE-binding activity is induced by HS, but the HSF is not transcriptionally active (Hensold et al., 1990). The underphosphorylation of the HSF in these cells appears to only partially account for differences in electrophoretic mobility between it and active HSF (Hensold et al., 1990), which raises the possibility that phosphorylation may not be the only posttranslational modification to be considered in transcriptional activation. Very little is known regarding the details of HSF-mediated activation of transcription. RNA polymerase II has been shown to be arrested in elongation after transcription of 25 bases of RNA on a *Drosophila hsp70* gene (Rougvie & Lis, 1988). Since the HSE core repeat is staggered every 5 bp, the activated HSF trimer must simultaneously interact with opposite helical faces of the DNA. It is plausible that the binding of the HSF trimers to sites 1 and 2 may alter the conformation of the preformed transcription complex comprised of TFIID and associated factors, facilitating the release of RNA polymerase for elongation.

In summary, accumulated evidence indicates that the molecular aspects of HS regulation in plants is very similar to that described for *Drosophila* and human cells. This conservation in mechanisms of regulation is evident in the organization of the HS promoter, the activation of HS mRNA accumulation, and differential stability of HS mRNAs. Distinguishing features of the plant HS response include the complexity and abundance of the low molecular weight hsps and a less stringent prohibition of introns in HS genes. With the use of transgenic expression in regenerated plants and the recent development of in vitro transcription systems from wheat germ, the path has been cleared for plant molecular biology to make a contribution to our understanding of the HS response in higher eukaryotes. It is clear that the nature of the regulation of HS gene expression is highly complex; it is equally clear that much remains to be learned in order to understand this complexity at a biochemical level.

ACKNOWLEDGMENTS

We thank Drs. Janice Kimpel, Ronald T. Nagao, and Ewa Czarnecka for comments on the manuscript during its preparation. We appreciate the technical assistance of Mary Lynn Ward in manuscript preparation.

REFERENCES

Amin, J., Mestril, R., Schiller, P., Dreano, M,. & Voellmy, R. (1987) *Mol. Cell. Biol. 7*, 1055–1062.

Amin, J., Ananthan, J., & Voellmy, R. (1988) *Mol. Cell. Biol. 8*, 3761–3769.

Ashburner, M., & Bonner, J. J. (1979) *Cell 17*, 241–254.

Atkinson, B. G., & Walden, D. B. (1985) *Changes in Eukaryotic Gene Expression in Response to Environmental Stress*, Academic Press, New York.

Baumann, G., Raschke, E., Bevan, M., & Schoffl, F. (1987) *EMBO J. 6*, 1161–1166.

Bienz, M., & Pelham, H. R. B. (1986) *Cell 45*, 753–760.

Browder, L. W., Pollock, M., Nickells, R. W., Heikkila, J. J., & Winning, R. S. (1989) in *Developmental Biology: A Comprehensive Synthesis, Vol. 6, Genomic Adaptability in Somatic Cell Specialization* (DiBerardino, M. A., & Etkin, L. D., Eds.) pp 97–147, Plenum Press, New York.

Bruce, W. B., Bandyopadhyay, R., & Gurley, W. B. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4310–4314.

Bustos, M. M., Guiltinan, M. J., Jordano, J., Begum, D., Kalkan, F. A., & Hall, T. C. (1989) *Plant Cell 1*, 839–853.

Craig, E. A. (1985) *CRC Crit. Rev. Biochem. 18*, 239–280.

Craig, E., Boorstein, W., Park, H.-O., Stone, D., & Nicolet, C. (1989) in *Stress-Induced Proteins* (Fox, C. F., Ed.) UCLA Symposia on Molecular and Cellular Biology, New Series, pp 51–61, Alan R. Liss, Inc., New York.

Czarnecka, E., Edelman, L., Schoffl, F., & Key, J. L. (1984) *Plant Mol. Biol. 3*, 45–58.

Czarnecka, E., Gurley, W. B., Nagao, R. T., Mosquera, L., & Key, J. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3726–3730.

Czarnecka, E., Nagao, R. T., Key, J. L., & Gurley, W. B. (1988) *Mol. Cell. Biol. 8*, 1113–1122.

Czarnecka, E., Key, J. L., & Gurley, W. B. (1989) *Mol. Cell. Biol. 9*, 3457–3463.

Czarnecka, E., Fox, P. C., & Gurley, W. B. (1990) *Plant Physiol. 94*, 935–943.

Datta, N., & Cashmore, A. R. (1989) *Plant Cell 1*, 1069–1077.

DiDomenico, B. J., Bugaisky, G. E., & Lindquist, S. (1982) *Cell 31*, 593–603.

Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J., & Weissmann, C. (1983) *Cell 32*, 695–706.

Doolittle, W. F. (1987) in *Cold Spring Harbor Symposium on Evolution of Catalytic Function*, pp 907–913, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Duck, N., McCormick, S., & Winter, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3674–3678.

Dynan, W. S. (1989) *Cell 58*, 1–4.

Edelman, L., Czarnecka, E., & Key, J. L. (1988) *Plant Physiol. 86*, 1048–1056.

Erickson, J. W., Vaughn, V., Walter, W. A., Neidhardt, F. C., & Gross, C. A. (1987) *Genes Dev. 1*, 419–432.

Ferl, R. J., & Laughner, B. H. (1989) *Plant Mol. Biol. 12*, 357–366.

Georgopoulos, C., Ang, D., Liberek, K., & Zylicz, M. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissieres, A., & Georgopoulos, C., Eds.) pp 191–221, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Goldenberg, C. J., Luo, Y., Fenna, M., Baler, R., Weinmann, R., & Voellmy, R. (1988) *J. Biol. Chem. 263*, 19734–19739.

Gross, C. A., Straus, D. B., Erickson, J. W., & Yura, T. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissieres, A., & Georgopoulos, C., Eds.) pp 167–189, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Grossman, A. D., Straus, D. B., Walter, W. A., & Gross, C. A. (1987) *Genes Dev. 1*, 179–184.

Grosveld, G. C., de Boer, E., Shewmaker, C. K., & Flavell, R. A. (1982) *Nature 295*, 120–126.

Hensold, J. O., Hunt, C. R., Calderwood, S. K., Housman, D. E., & Kingston, R. E. (1990) *Mol. Cell. Biol. 10*, 1600–1608.

Herr, W., & Clarke, J. (1986) *Cell 45*, 461–470.

Holmgren, R., Corces, V., Morimoto, R., Blackman, R., & Meselson, M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3775–3778.

Jakobsen, B. K., & Pelham, H. R. B. (1988) *Mol. Cell. Biol. 8*, 5040–5042.

Jofuku, K. D., Okamuro, J. K., & Goldberg, R. B. (1987) *Nature 328*, 734–737.

Jordano, J., Almoquera, C., & Thomas, T. L. (1989) *Plant Cell 1*, 855–866.

Kay, R. J., Boissy, R. J., Russnak, R. H., & Candido, E. P. M. (1986) *Mol. Cell. Biol. 6*, 3134–3143.
Kay, R. J., Fussnak, R. H., Jones, D., Mathias, C., & Candido, E. P. (1987) *Nucleic Acids Res. 15*, 3723–3741.
Key, J. L., Lin, C.-Y., & Chen, Y. M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3526–3530.
Key, J. L., Lin, C.-Y., Ceglarz, E., & Schoffl, F. (1982) in *Heat Shock: From Bacteria to Man* (Schlesinger, M. J., Ashburner, M., & Tissieres, A., Eds.) pp 329–336, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Key, J. L., Kimpel, J., Vierling, E., Lin, C.-Y., Nagao, R. T., Czarnecka, E., & Schoffl, F. (1985) in *Changes in Eukaryotic Gene Expression in Response to Environmental Stress* (Atkinson, B. G., & Walden, D. B., Eds.) pp 327–348, Academic Press, New York.
Key, J. L., Kimpel, J., & Nagao, R. T. (1987) in *Plant Gene Systems and Their Biology* (Key, J. L., & McIntosh, L., Eds.) pp 87–97, Alan R. Liss, Inc., New York.
Kimpel, J. A., & Key, J. L. (1985) *Plant Physiol. 79*, 672–678.
Kimpel, J. A., Nagao, R. T., Goekjian, V., & Key, J. L. (1990) *Plant Physiol.* (in press).
Kingston, R. E., Schuetz, T. J., & Larin, Z. (1987) *Mol. Cell. Biol. 7*, 1530–1534.
Krishnan, H. B., & Pueppke, S. G. (1987) *Biochem. Biophys. Res. Commun. 148*, 762–767.
Kuhlemeier, C., Fluhr, R., Green, P. J., & Chua, N.-H. (1987) *Genes Dev. 1*, 247–255.
Laimins, L. A., Khoury, G., Gorman, C., Howard, B., & Gruss, P. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6453–6457.
Larson, J. S., Schuetz, T. J., & Kingston, R. E. (1988) *Nature 335*, 372–375.
Lindquist, S. (1986) *Annu. Rev. Biochem. 55*, 1151–1191.
Lindquist, S., & Craig, E. A. (1988) *Annu. Rev. Genet. 22*, 631–677.
Lis, J. R., Xaio, H., & Perisic, O. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissieres, A., & Georgopoulos, C., Eds.) pp 411–428, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Mansfield, M. A., & Key, J. L. (1987) *Plant Physiol. 84*, 1007–1017.
McKendree, W. L., Paul, A.-L., DeLisle, A. J., & Ferl, R. J. (1990) *Plant Cell 2*, 207–214.
McKnight, S. (1982) *Cell 31*, 355–365.
Mirault, M. E., Southgate, R., & Delwart, E. (1982) *EMBO J. 1*, 1279–1285.
Morimoto, R. I., Tissieres, A., & Georgopoulos, C. (1990) *Stress Proteins in Biology and Medicine*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Mosser, D. D., Theodorakis, N. G., & Morimoto, R. I. (1988) *Mol. Cell. Biol. 8*, 4736–4744.
Mosser, D. D., Kotzbauer, P. T., Sarge, K. D., & Morimoto, R. I. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3748–3752.
Nagao, R. T., & Key, J. L. (1989) in *Cell Culture and Somatic Cell Genetics of Plants* (Schell, J., & Vasil, I., Eds.) Vol. 6, pp 297–328, Academic Press, New York.
Nagao, R. T., Czarnecka, E., Gurley, W. B., Schoffl, F., & Key, J. L. (1985) *Mol. Cell. Biol. 5*, 3417–3428.
Nagao, R. T., Kimpel, J. A., Vierling, E., & Key, J. L. (1986) in *Oxford Surveys of Plant Molecular & Cell Biology* (Miflin, B. J., Ed.) Vol. 3, pp 384–438, Oxford University Press, London.
Neidhardt, F. C. (1987) *Genes Dev. 1*, 109–110.
Neidhardt, F. C., VanBogelen, R. A., & Vaughn, V. (1984) *Annu. Rev. Genet. 18*, 295–329.
Neumann, D., Nover, L., Parthier, B., Rieger, R., Scharf, K.-D., Wollgiehn, R., & Nieden, U. Z. (1989) *Biol. Zentralbl. 108*, 1–156.
Nover, L. (1987) *Enzyme Microb. Technol. 9*, 130–144.
Pardue, M. L., Feramisco, J. R., & Lindquist, S. (1989) in *Stress-Induced Proteins* (Fox, C. F., Ed.) UCLA Symposia on Molecular and Cellular Biology, New Series, Alan R. Liss, Inc., New York.
Park, H. O., & Craig, E. A. (1989) *Mol. Cell. Biol. 9*, 2025–2033.
Parker, C. S., & Topol, J. (1984) *Cell 37*, 273–283.
Parker-Thornburg, J., & Bonner, J. J. (1987) *Cell 51*, 763–772.
Pauli, D., Spierer, A., & Tissieres, A. (1986) *EMBO J. 5*, 755–761.
Pelham, H. R. B. (1982) *Cell 30*, 517–528.
Pelham, H. (1985) *Trends Genet. 1*, 31–35.
Perisic, O., Xiao, H., & Lis, J. T. (1989) *Cell 59*, 797–806.
Petersen, R., & Lindquist, S. (1988) *Gene 72*, 161–168.
Raschke, E., Baumann, G., & Schoffl, F. (1988) *J. Mol. Biol. 199*, 549–558.
Riddihough, G., & Pelham, H. R. B. (1986) *EMBO J. 5*, 1653–1658.
Ritossa, F. (1962) *Experientia 18*, 571–573.
Rochester, D. E., Winter, J. A., & Shah, D. M. (1986) *EMBO J. 5*, 451–458.
Rougvie, A. E., & Lis, J. T. (1988) *Cell 54*, 795–804.
Schlesinger, M. J. (1986) *J. Cell Biol. 103*, 321–325.
Schlesinger, M. J., Ashburner, M., & Tissieres, A. (1982) *Heat Shock: From Bacteria to Man* (Schlesinger, M. J., Ashburner, M., & Tissieres, A., Eds.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schoffl, F., & Key, J. L. (1982) *J. Mol. Appl. Genet. 1*, 301–314.
Schoffl, F., Raschke, E., & Nagao, R. T. (1984) *EMBO J. 3*, 2491–2498.
Schoffl, F., Rossol, I., & Angermuller, S. (1987) *Plant, Cell Environ. 10*, 113–119.
Schoffl, F., Rieping, M., Baumann, G., Bevan, M., & Angermuller, S. (1989) *MGG, Mol. Gen. Genet. 217*, 246–253.
Shah, D. M., Rochester, D. E., Krivi, G., Hironaka, C. M., Mozer, T. J., Fraley, R. T., & Tiemeier, D. C. (1985) in *Cellular and Molecular Biology of Plant Stress* (Key, J. L., & Kosuge, T., Eds.) UCLA Symposia on Molecular and Cellular Biology, New Series, pp 181–200, Alan R. Liss, Inc., New York.
Shuey, D. J., & Parker, C. S. (1986) *J. Biol. Chem. 261*, 7934–7940.
Skelly, S., Fu, C. F., Dalie, B., Redfield, B., Coleman, T., Brot, N., & Weissbach, H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5497–5501.
Sorger, P. K., & Pelham, H. R. B. (1988) *Cell 54*, 855–864.
Sorger, P. K., Lewis, M. J., & Pelham, H. R. (1987) *Nature 329*, 81–84.
Spena, A., & Schell, J. (1987) *MGG, Mol. Gen. Genet. 206*, 436–440.
Spena, A., Hain, R., Ziervogel, U., Saedler, H., & Schell, J. (1985) *EMBO J. 4*, 2739–2743.
Strittmatter, G., & Chua, N.-H. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8986–8990.
Takahaski, T., & Komeda, Y. (1989) *MGG, Mol. Gen. Genet. 219*, 365–372.
Tanguay, R. M. (1988) *Biochem. Cell Biol. 66*, 584–593.
Tilly, K., Spence, J., & Georgopoulos, C. (1989) *J. Bacteriol. 171*, 1585–1589.

Topol, J., Rudon, D. M., & Parker, C. S. (1985) *Cell 42*, 527-537.
Velazquez, J. M., & Lindquist, S. (1984) *Cell 36*, 655-662.
Wei, R., Wilkinson, H., Pfeifer, K., Schneider, C., Young, R., & Guarente, L. (1986) *Nucleic Acids Res. 14*, 8183-8188.
Wing, D., Koncz, C., & Schell, J. (1989) *MGG, Mol. Gen. Genet. 219*, 9-16.
Wu, C. (1984a) *Nature 311*, 81-84.
Wu, C. (1984b) *Nature 309*, 229-234.
Wu, C. H., Caspar, T., Browse, J., Lindquist, S., & Somerville, C. (1988) *Plant Physiol. 88*, 731-740.
Wu, C., Zimariano, V., Tsai, C., Walker, G., & Wilson, S. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissieres, A., & Georgopoulos, C., Eds.) pp 429-442, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Xiao, H., & Lis, J. T. (1988) *Science 239*, 1139-1142.
Xiao, H., & Lis, J. T. (1989) *Mol. Cell. Biol. 9*, 1746-1753.
Yost, H. J., & Lindquist, S. (1986) *Cell 45*, 185-193.
Yost, H. J., Petersen, R. B., & Lindquist, S. (1990) in *Stress Proteins in Biology and Medicine* (Morimoto, R. I., Tissieres, A., & Georgopoulos, C., Eds.) pp 379-409, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Zimarino, V., & Wu, C. (1987) *Nature 327*, 727-730.
Zimarino, V., Tsai, C., & Wu, C. (1990a) *Mol. Cell. Biol. 10*, 752-759.
Zimarino, V., Wilson, S., & Wu, C. (1990b) *Science 249*, 546-549.

Accelerated Publications

Four-Dimensional $^{13}C/^{13}C$-Edited Nuclear Overhauser Enhancement Spectroscopy of a Protein in Solution: Application to Interleukin 1β[†]

G. Marius Clore,* Lewis E. Kay, Ad Bax,* and Angela M. Gronenborn*

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received September 18, 1990; Revised Manuscript Received October 24, 1990

ABSTRACT: A four-dimensional $^{13}C/^{13}C$-edited NOESY experiment is described which dramatically improves the resolution of protein NMR spectra and enables the straightforward assignment of nuclear Overhauser effects involving aliphatic and/or aromatic protons in larger proteins. The experiment is demonstrated for uniformly (>95%) ^{13}C-labeled interleukin 1β, a protein of 153 residues and 17.4 kDa, which plays a key role in the immune response. NOEs between aliphatic and/or aromatic protons are first spread out into a third dimension by the ^{13}C chemical shift of the carbon atom attached to the originating proton and subsequently into a fourth dimension by the ^{13}C chemical shift of the carbon atom attached to the destination proton. Thus, each NOE cross peak is labeled by four chemical shifts. By this means, ambiguities in the assignment of NOEs that arise from chemical shift overlap and degeneracy are completely removed. Further, NOEs between protons with the same chemical shifts can readily be detected providing their attached carbon atoms have different ^{13}C chemical shifts. The design of the pulse sequence requires special care to minimize the level of artifacts arising from undesired coherence transfer pathways, and in particular those associated with "diagonal" peaks which correspond to magnetization that has not been transferred from one proton to another. The 4D $^{13}C/^{13}C$-edited NOESY experiment is characterized by high sensitivity as the through-bond transfer steps involve the large $^1J_{CH}$ (130 Hz) couplings, and it is possible to obtain high-quality spectra on 1-2 mM samples of ^{13}C-labeled protein in as little as 3 days. This experiment should open up the application of protein structure determination by NMR to a large number of medium-sized proteins (150-300 residues) of biological interest.

The key to protein structure determination by NMR[1] spectroscopy lies in the identification of as many nuclear Overhauser effects (NOEs) as possible, which yield geometric information in the form of approximate interproton distance restraints (Wüthrich, 1986; Clore & Gronenborn, 1989). Over the last few years it has been convincingly demonstrated that solution structures comparable to 2-2.5 Å resolution X-ray structures can be obtained by NMR for proteins up to about 100 residues (Kline et al., 1988; Kraulis et al., 1989; Clore et al., 1990a; Qian et al., 1990; Omichinski et al., 1990; Dyson et al., 1990). This accuracy and precision depend critically on the number of NOEs that can be assigned (Clore & Gronenborn, 1991). As proteins get larger than about 120 residues (~13 kDa), resonance overlap and degeneracy are so extensive that it frequently becomes impossible to assign a large portion of the NOEs with any degree of confidence

[†] This work was supported by the AIDS-directed Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C., A.M.G., and A.B.). L.E.K. was supported by a postdoctoral Centennial Fellowship from the Medical Research Council of Canada.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; HMQC, heteronuclear multiple quantum coherence; IL-1β, interleukin 1β.