Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 4949–4953, August 1985
Biochemistry

Isolation and functional analysis of a human 70,000-dalton heat shock protein gene segment

(DNA sequence/transient expression/mammalian cells/*Xenopus* oocytes/hybrid genes)

RICHARD VOELLMY*, AFTAB AHMED*, PAUL SCHILLER*, PETER BROMLEY†, AND DURI RUNGGER‡

*Department of Biochemistry, University of Miami School of Medicine, Miami, FL 33101; †Battelle Memorial Institute, Geneva Research Center, CH-1227 Carouge, Switzerland; and ‡Department of Animal Biology, University of Geneva, CH-1224 Chene-Bougeries, Switzerland

Communicated by V. Prelog, April 8, 1985

ABSTRACT **A human 70-kDa heat shock protein (hsp70) gene segment has been isolated. The segment contains 3.15 kilobase pairs (kbp) of 5′ nontranscribed sequence, an RNA leader of 119 bp, and a protein-coding region of 741 bp. The human protein sequence shows a high degree of homology to hsp70 sequences from other species. Expression experiments in *Xenopus* oocytes and mammalian cells indicate that a region that includes only 105 bp of 5′ nontranscribed sequence contains all elements required for the efficient heat-controlled expression of the human gene. Two adjacent identical sequence elements, which are partly homologous to the *Drosophila* "heat shock consensus" sequence, are located 57 to 76 bp upstream from the capping site. Interestingly, the capping site itself is flanked by inverted repeat sequences.**

The heat shock response in *Drosophila* has been described extensively (see ref. 1 for review). The genes coding for seven major heat shock proteins (hsp) are expressed at high levels at elevated temperature (35°C–37°C) but, typically, are inactive or only weakly active at normal temperature (25°C). The heat-induced expression of these *Drosophila* genes is controlled at both the transcriptional and the translational levels. The *Drosophila* genes have been isolated and most of them have been sequenced completely. Extensive expression studies with one of these genes (coding for hsp70) have led to the definition of a promotor element, the "heat shock consensus" sequence (2, 3), which is required for the regulated expression of the gene.

Proteins with structures similar to those of the *Drosophila* hsp have been found in a variety of other eukaryotic and even in prokaryotic organisms (1). *Drosophila* hsp70-like genes have been identified in yeast (4) and *Escherichia coli* (5) and, most recently, after the material presented here had been assembled, also in *Xenopus* (6). That the *Drosophila hsp70* genes are expressed in a heat-regulated fashion in frog, mouse, and monkey cells (2, 3, 7, 8) implies that not only the structures, but even the mechanism of transcription regulation of these genes have been conserved throughout evolution.

To extend our understanding of the evolutionary relationship between heat shock response mechanisms in different organisms, we have isolated and functionally tested a human *hsp70* gene. This study also provides the basis for future experiments in which the newly isolated human *hsp* gene will be used to investigate the mechanism through which transforming genes activate cellular genes: *hsp70* genes are induced as a result of the action of the adenoviral *E1A* gene (9) and even of a c-*myc* gene (10).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Screening of a Human DNA Library. A phage λ library of human genomic DNA (11) was kindly provided by Tom Maniatis. This library was screened by plaque hybridization (12) using as probe a 2-kbp *Sal* 1 fragment from plasmid SalO (13), which includes a complete *Drosophila* hsp70 RNA-coding region. After hybridization under standard conditions, the filters were washed at 42°C in 5× NaCl/Cit (1× NaCl/Cit = 0.15 M NaCl/0.015 M Na citrate). About 5×10^5 plaques were screened. As a final test, DNAs were prepared from the positive clones (14) and were hybridized to labeled *Drosophila hsp70* gene DNA in dot blot experiments (15).

DNA Sequence Analysis. Plasmid DNAs were prepared by CsCl banding from cleared lysates. DNA sequence analysis was carried out according to Maxam and Gilbert (16) except that DNA fragments were end-labeled by DNA polymerase fragment A and (α-^{32}P)dXTP.

Construction of Human *hsp70–E. coli* β-Galactosidase Hybrid Genes. To construct plasmid 671SX, an *Xho* I linker sequence was introduced into the unique *Bgl* II site of the previously described plasmid 671 (17), which contains, in between the *Xma*III and *Bam*HI sites of pSVOd (18), a 3-kbp *E. coli* β-galactosidase fragment from pMC1871 (19) and downstream from it, 2.3 kbp of eukaryotic 3′ gene flanking sequence. The construction of p173 is described in Fig. 3*a*. Sticky ends of restriction fragments were filled in by DNA polymerase fragment A. To construct p173P, a 5-kbp *Pst* I fragment from p173 was isolated that includes 600 bp of 5′ nontranscribed and 500 bp of transcribed human *hsp70* gene sequence, the β-galactosidase-coding region and 3′ flanking sequences. This fragment was inserted into the *Pst* I site of pUC8 (20). Plasmid 173S was prepared by digesting p173 with *Sph* I. Ends were made blunt by T4 DNA polymerase. After digestion with *Sma* I, the material was religated and used for transformation of *E. coli* MC1061. Plasmid 173X was made by introducing an *Xho* I linker into the *Sph* I site of p173. To construct p173OR, an *Xho* I/*Cla* I fragment from p173X containing 2.95 kbp of human hsp70-coding and upstream sequences and the first one-third of the β-galactosidase gene was inserted in between the *Xho* I and *Cla* I sites of p622a (17).

Microinjection and Transient Expression Experiments. Injections into sets of 10–20 *Xenopus* oocytes (3 ng of DNA in 10 nl per oocyte) were carried out as described (8, 21). To measure the hybrid gene-directed synthesis of β-galactosidase, the oocytes were lysed by pipetting in 700 μl of Z buffer (22). The lysates were centrifuged in the cold at 12,000 × *g* for 10 min, and 400 μl of the clear supernatants was recovered; 100 μl of *o*-nitrophenyl-β-D-galactopyranoside (4 mg/ml) was added, and the samples were incubated for 1.5 hr at 37°C. Activities were estimated as described (22). Human

Abbreviations: kbp, kilobase pair(s); hsp, heat shock protein(s).

Wish cells (ATCC:CCL25) grown in basal Eagle's medium (BME)/Earle's salts and 10% fetal calf serum were transfected with p173 DNA (4 μg per 3.5-cm dish) using $CaCl_2$ (23). They were then either heat-treated for 3 hr at 43°C or kept at 37°C. After an additional 16-hr incubation at 37°C, the cells were fixed, and a peroxidase assay was carried out with rabbit antibodies directed against *E. coli* β-galactosidase and peroxidase-conjugated goat anti-rabbit antibodies.

RNA and S1 Nuclease Analysis. Total RNA from heat-treated (90 min at 36°C) or untreated (kept at 21°C) oocytes was isolated according to Probst *et al.* (24). Human cytoplasmic RNA was prepared from heat-treated (3 hr at 43°C) or control Wish cells by Dounce homogenization in a buffer containing 0.5% Nonidet P-40, followed by sedimentation through CsCl gradients (25). S1 nuclease analysis was carried out by the procedure of Weaver and Weissmann (26) except that DNA fragments were not heated in pure formamide but were resuspended and denatured directly in the hybridization buffer. Excess RNA was eliminated after S1 nuclease treatment by incubation of the samples in 0.1 M NaOH at 45°C for 45 min. Electrophoresis was on 6% acrylamide/urea gels.

RESULTS AND DISCUSSION

Isolation and Initial Analysis of a Human *hsp70* Gene Segment. Since corresponding hsp from different organisms have similar structures, the DNA sequences encoding these hsp could be expected to cross-hybridize, at least under conditions of relaxed stringency. We therefore decided to attempt the isolation of human heat shock genes by using *Drosophila hsp* genes as hybridization probes. Eighteen phages with inserts hybridizing to *Drosophila hsp70* gene sequences were isolated. One of these phages was analyzed further. It contains an insert fragment of ≈13 kbp, which could be divided into three 4- to 4.5-kbp fragments by *Eco*RI digestion. These fragments were subcloned into pUC8. Only one of the subclones, p17, was found to hybridize to *Drosophila* hsp70 RNA-coding sequences. A restriction map of the human DNA insert of p17 was established (Fig. 1*D*). Southern blot experiments indicated that the human sequences that hybridize to (the 5′ half of) the *Drosophila hsp70* gene are located to the right of the *Sac* I site in the p17 map.

The Human Sequence Is Homologous to Other hsp70-Coding Sequences. The cross-hybridizing human DNA segment and flanking regions were analyzed by direct sequencing. The human DNA sequence (Fig. 1*A*) was computer-aligned with *Drosophila hsp70*, the *Drosophila hsp70* cognate, a yeast *hsp70*, and the *E. coli dnaK* gene sequences. The derived human protein sequence is 76.5% homologous to *Drosophila* hsp70 (13, 27), 79.5% to hsc70, the presumed product of the *Drosophila hsp70* cognate gene (28), 74.5% to yeast hsp70 (YG100 gene; ref. 4), and 55.5% to the dnaK protein (5). Long blocks of amino acids that are identical in all the above-mentioned eukaryotic hsp70-like proteins are found between human amino acids 10–20, 34–42, 70–78, 155–164, 196–210, 224–231, and 233–240. The first ATG in the aligned *Drosophila* hsp70 RNA-coding sequence appears 15 bp downstream from the first ATG in the transcribed human sequence. Because of the high degree of homology between the human and other hsp70 sequences, and because the next ATG is located 261 bp further downstream, we conclude that the first ATG in the transcribed human sequence is the start of translation codon.

Mapping of Human Heat Shock RNA. A 1.75-kbp *Bgl* II fragment from the right half of the p17 insert, which contains part of the hsp-coding region as well as 5′ flanking sequences, was used as hybridization probe in the experiment shown in Fig. 2*a*. A fragment 450–500 bases long is protected by RNA from heat-treated but not from untreated human cells in culture. This places the capping site to a position 80–130 bp to the left of the *Hin*dIII site in the p17 segment.

To map the transcription start site more precisely, a 450-bp *Bam*HI/HindIII fragment from p17 was used as probe in the experiments shown in Fig. 2 *b* and *c*. A fragment of 113 ± 3 bases is protected by human heat shock RNA (mean value from four independent experiments). This places the capping site to the position indicated as 1 in Fig. 1*A*. In addition to the 113-base fragment protected specifically by heat shock RNA, variable amounts of longer fragments (mainly of 195–200 bases) are present in all samples including those that only contain probe fragment but no RNA. Their intensities appear to correlate with the amounts of protection of full-length probe fragments (see also Fig. 2*d*). Thus, these longer fragments do not result from specific protection by heat shock RNA. Additional S1 nuclease experiments were also carried out using the upstream p17 *Bgl* II fragment (1.6 kbp, left side of p17 insert) as probe. No specific protection was observed (data not shown).

Inspection of the nucleotide sequence revealed the presence of inverted repeat sequences centered at the start of transcription site of the human *hsp* gene (Fig. 1*A*, dashed boxes). Since similar elements could not be found in corresponding locations in several other genes, it appears that these extensive repeat sequences are unique features of the p17 heat shock gene. Given their location and the fact that heat shock genes are transcribed at high rates at elevated temperatures, it is conceivable that these elements might be involved in facilitating transcription initiation.

Heat-Regulated Expression of the Cloned Human Gene. To find out whether the cloned human gene segment contains functional control sequences for heat shock gene transcription and translation, expression experiments in *Xenopus* oocytes were carried out. Since *Drosophila hsp70* genes are transcribed correctly and in a heat-induced fashion in *Xenopus*, it was expected that a functional human gene would also be recognized and expressed in these cells. Total RNA isolated from heat-treated and untreated oocytes injected with p17 DNA, and from uninjected heat-treated oocytes was analyzed by S1 nuclease mapping using the 450-bp *Bam*HI/*Hin*dIII p17 fragment as probe (Fig. 2*d*). The probe does not hybridize to RNA from uninjected oocytes. A strong transcription signal is detected in heat-treated but not in control oocytes containing the human gene. The length of the specifically protected fragment is similar to that produced by human heat shock RNA (Fig. 2*d*, lanes 1, 3, and 5). Thus, the human gene in p17 contains a functional heat shock gene promotor that is recognized in *Xenopus* oocytes.

To confirm these findings and to show that the p17 gene fragment contains a functional RNA leader region and that the translational reading frame assigned to the human gene by sequence comparison is correct, a *hsp* hybrid gene, 173, was constructed by inserting a truncated *E. coli* β-galactosidase gene into the protein-coding region of the human gene in p17. From the sequence data it was predicted (Fig. 3*c*) that an in-frame fusion of the human hsp- and the β-galactosidase-coding sequences would be accomplished by the construction procedure described in Fig. 3*a*. The human *hsp–β*-galactosidase hybrid gene in p173 was injected into *Xenopus* oocytes. The oocytes were then either heat-treated at 36°C for 90 min or kept at the normal temperature of 21°C for the same time. To allow for efficient translation of the hybrid gene products, the oocytes were incubated further at 21°C overnight prior to the preparation of extracts. Substantial β-galactosidase levels were measured in heat-treated, but not in untreated, oocytes containing the p173 hybrid gene or in uninjected oocytes (Table 1). Experiments with the transcription blocker α-amanitin demonstrated that the hybrid gene is transcribed by RNA polymerase II, that its expression is

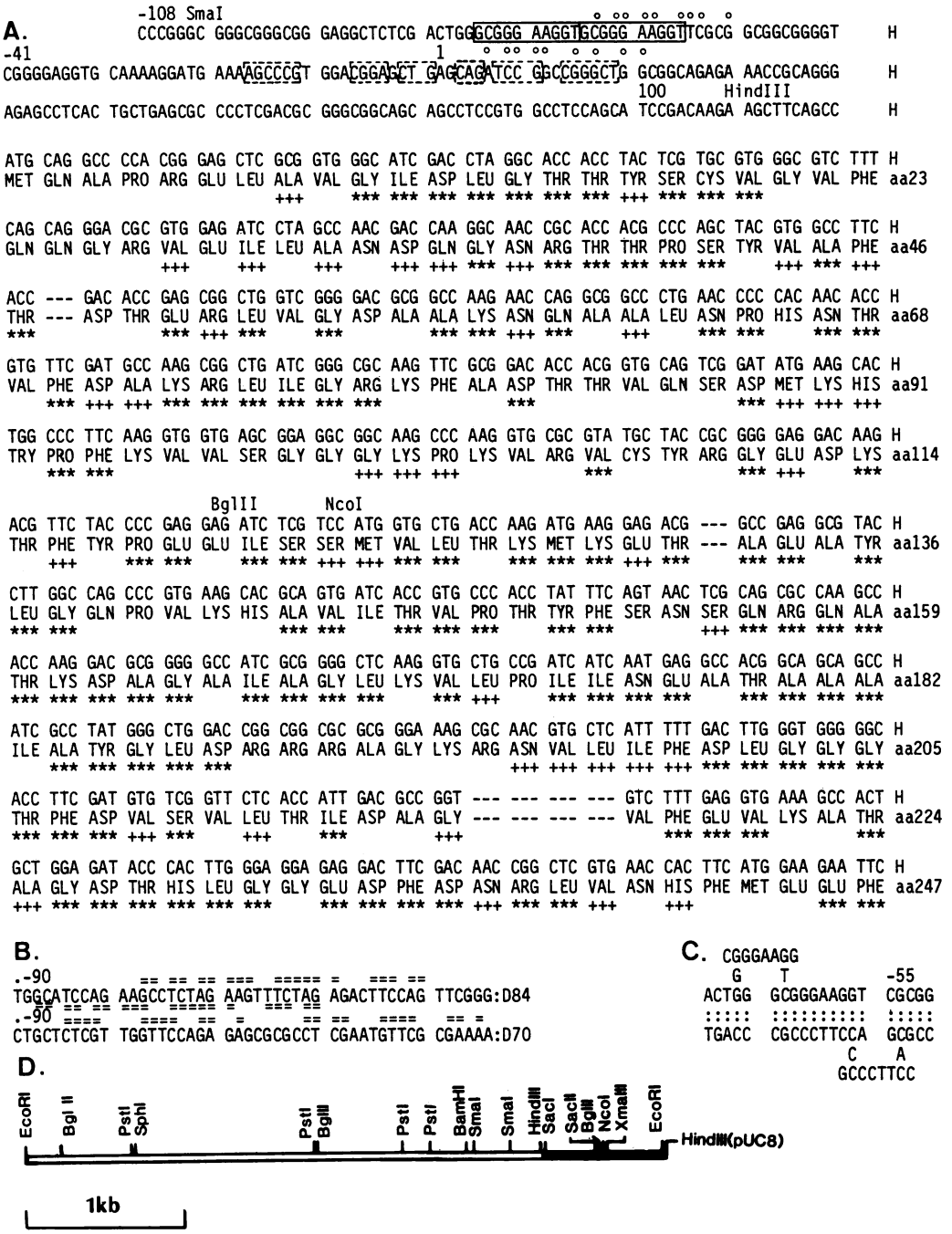

FIG. 1. (*A*) Nucleotide sequence of the p17 human *hsp70* gene segment (H). In the human nontranslated sequence as well as in the *Drosophila* D84 and D70 sequences (see *B* and *C*), 1 refers to the start of transcription site. Direct repeats in the human nontranslated sequence are shown by solid boxes and inverted repeats are shown by dashed boxes. Nucleotides matching the *Drosophila* "heat shock consensus" sequence are marked by ∘. Amino acids (aa) identical in the human hsp70, the *Drosophila* hsp and hsc70, the yeast YG100 hsp70, and the dnaK protein sequences are indicated by *, and additional amino acids common to all above hsps except the dnaK protein are indicated by +. Extra codons in other hsp sequences are indicated by ---. (*B*) Promotor sequences of the *Drosophila hsp84* (D84) and *hsp70* (D70) genes. Interrupted repeats are indicated by =. (*C*) Model structure of the human *hsp* gene promotor. (*D*) Map of p17. The filled area shows hsp70-coding sequences.

controlled at the transcriptional level, and that transcription of the hybrid gene is initiated during, and continues for some time after, heat treatment. The amounts of β-galactosidase made in heat-treated oocytes containing the human hybrid gene are comparable to those synthesized by similar *Drosophila hsp70* hybrid genes (29).

To demonstrate that the human *hsp* gene is also active and heat-regulated in mammalian cells, a derivative of p173 containing a simian virus 40 origin of replication, p173OR, was constructed. The hybrid gene was introduced into COS 1 (African Green Monkey kidney) cells that allow amplification, and it was found (see Table 1) to direct the synthesis of β-galactosidase in heat-treated, but not in untreated, cells at a level comparable to that produced by an analogous *Drosophila hsp70* hybrid gene (PR4-9; ref. 17). In a second experiment, p173 was introduced into human Wish cells by transfection. Because of the low transfection efficiency of these cells, a single cell immunoassay was used to demonstrate the heat-induced activity of the hybrid gene (Fig. 4).

Location of the Human *hsp* Gene Promotor. Mutant derivatives of p173 were prepared that include only 600 (173P) and 105 bp (173S), respectively, of 5′ nontranscribed sequence of the human gene (see Figs. 1 and 3*b*). The regulated expression of these mutant hybrid genes was examined in *Xenopus*

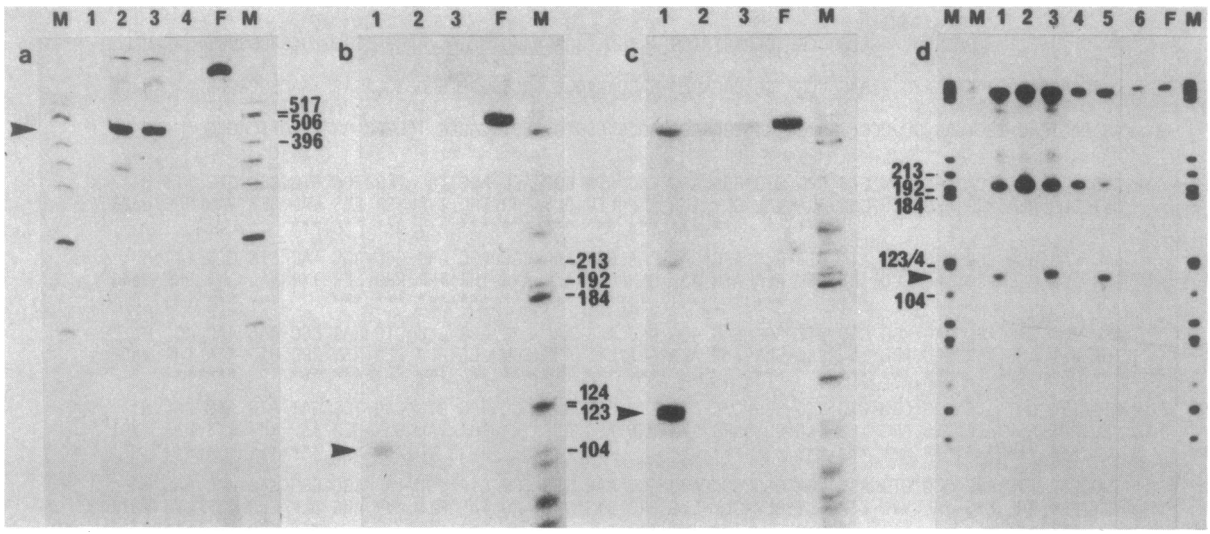

FIG. 2. Nuclease S1 analysis of *hsp* gene transcripts from human cells and *Xenopus* oocytes injected with p17 DNA. (*a*) Labeled 1.75-kb*p* *Bgl* II p17 probe fragment was hybridized to 50 μg of RNA from untreated (lanes 1 and 4) and heat-treated (lanes 2 and 3) human Wish cells. F, probe fragment; M, labeled *Hin*dIII/*Hin*fI fragments of pBR322. (*b* and *c*) The 450-bp *Bam*HI/*Hin*dIII p17 probe fragment was hybridized to 50 μg of RNA from heat-treated (lane 1) and untreated (lane 2) Wish cells, or incubated without RNA (lane 3). F, probe fragment; M, labeled *Hae* III fragments of pBR322. (*d*) The same probe as in *b* and *c* was hybridized to 300 μg of RNA from heat-treated (lanes 1 and 3) or untreated (lanes 2 and 4) *Xenopus* oocytes containing p17, or from heat-treated uninjected oocytes (lane 6), and to 50 μg of RNA from heat-treated Wish cells (lane 5). F and M are the same as in *b* and *c*.

oocytes (Table 1). The results of these experiments strongly suggest that the human *hsp* gene regulation signals are contained within a 105-bp DNA segment located immediately upstream from the capping site of the gene. The human promotor segment has many of the features that typically are associated with constitutively expressed "housekeeping" genes (30). It is G+C-rich (72%), contains several CCGCCC/GGGCGG elements (at positions −78, −97, −101, and −105), lacks a CCAAT element at the characteristic location, and does not include a typical TATA element (TGCAAAA at −33). In addition, however, it contains, between positions −52 and −75, two sequence elements that match the *Drosophila* "heat shock consensus" sequence in 9 out of 14 positions. Interestingly, these elements are almost

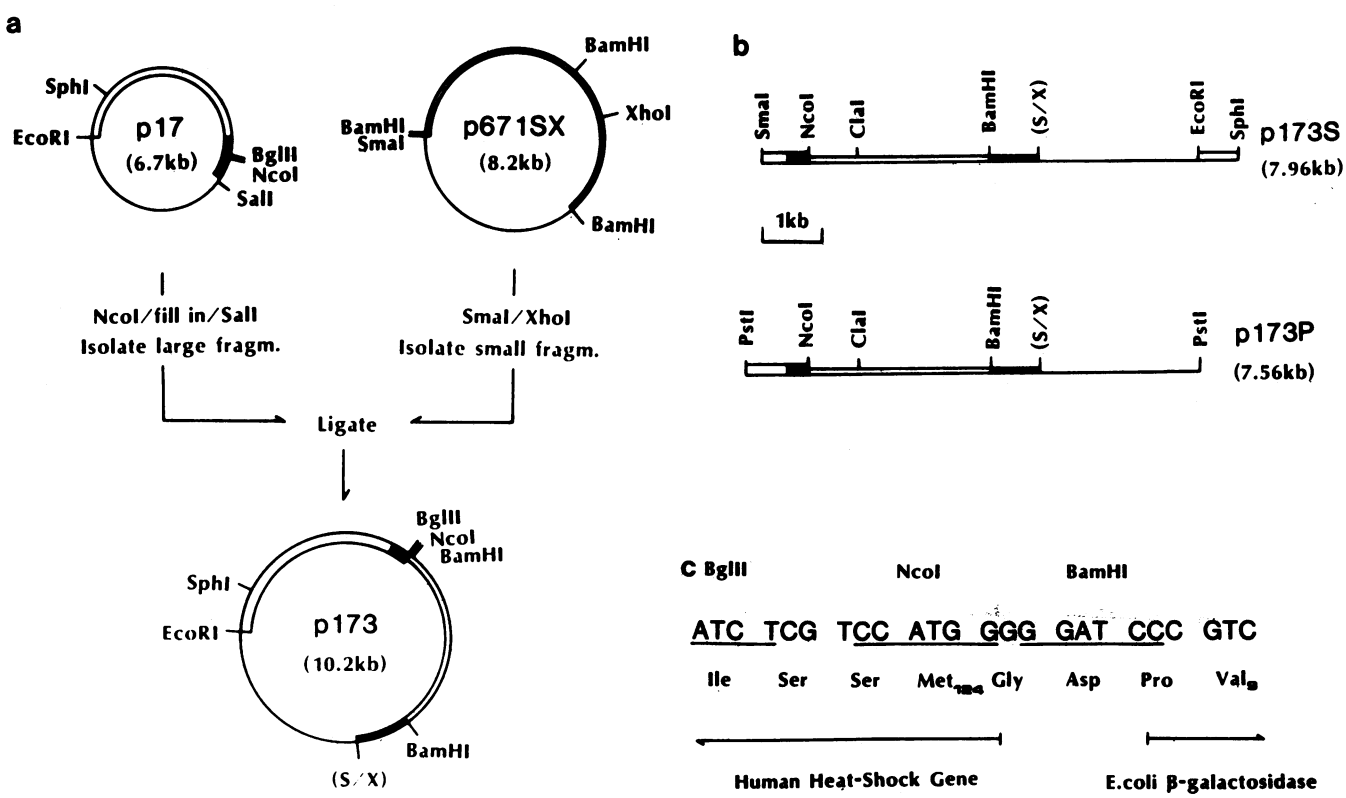

FIG. 3. Construction of human *hsp*–β-galactosidase hybrid genes. β-Galactosidase-coding regions are shown by thin double lines and 5′ nontranscribed and nontranslated human *hsp* gene sequences are shown by thick double lines. The 3′ flanking sequences are shown by thin filled areas and human hsp-coding sequences are shown by thick filled areas. Hybrid genes are read clockwise in p173 and from left to right in p173P and p173S. (S/X), *Sal* I/*Xho* I linkage.

Table 1. Expression of *hsp*–β-galactosidase hybrid genes

Xenopus oocytes			COS 1 cells		
DNAs	Treatment	β-Galactosidase activity	DNAs	Treatment	β-Galactosidase activity
None	HS	0.1	None	HS	0.0
	CO	0.1		CO	0.0
173	HS	2.0	1730R	HS	0.8
	CO	0.2		CO	0.0
	α-Amanitin		PR4-9	HS	1.1
	before HS	0.1		CO	0.0
	α-Amanitin				
	after HS	1.1			
522	HS	2.6			
	CO	0.1			
173P	HS	4.7			
	CO	0.1			
173S	HS	3.5			
	CO	0.1			

Transfections and transient assays with COS 1 cells were carried out as described (17). HS, heat shock; CO, control.

entirely included within adjacent direct repeat sequences (−57 to −76). That such direct repeats may be important features of *hsp* gene promotors is suggested by the fact that the promotors of several other *hsp* genes such as the *Drosophila hsp84* (31) and *hsp70* genes contain similar elements (see D84 and D70 in Fig. 1*B*). Significantly, promotor mutants of the *Drosophila hsp70* gene that lack the upstream repeat are not expressed efficiently in *Drosophila* cells (17, 32). These repeats may represent multiple recognition sites for a specific transcription factor such as the one described by Parker and Topol (33). Protein–DNA recognition may be aided by the transient formation of loops via a slippage mechanism (see model structure in Fig. 1*C*).

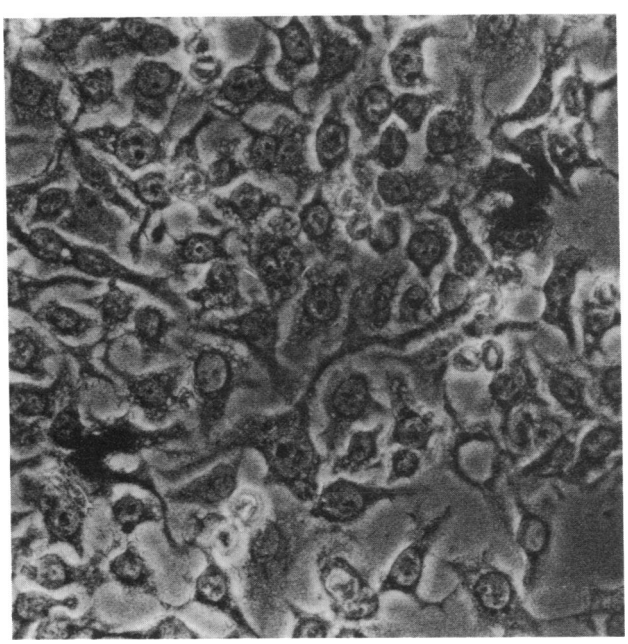

FIG. 4. β-Galactosidase activity in heat-treated human Wish cells transfected with p173 DNA. Positive cells (1%; none in untreated cultures) are darkly stained by the peroxidase color reaction, which results from first addition of anti-*E. coli* β-galactosidase antibodies followed by peroxidase-conjugated second antibodies.

We thank J. Poitry, L. Muster, K. Nemeth, H. Klapper, J. Brochot, R. Mestril, and J. Ananthan for their help with some of the experiments; A. Portianucha for the photographs; and T. Maniatis and W. Scott for gifts of materials. This work was supported by grants from Battelle Memorial Institute (Columbus, Ohio) and Prutec Ltd. (London) to R.V. and P.B., and from the Swiss National Foundation (3.244.-0.82) to D.R.

1. Schlesinger, M., Ashburner, M. & Tissières, A. (1982) *Heat Shock from Bacteria to Man* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
2. Pelham, H. R. B. (1982) *Cell* **30,** 517–528.
3. Mirault, M.-E., Southgate, R. & Delwart, E. (1982) *EMBO J.* **1,** 1279–1285.
4. Ingolia, T. D., Slater, M. R. & Craig, E. A. (1982) *Mol. Cell. Biol.* **2,** 1388–1398.
5. Bardwell, J. C. & Craig, E. A. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 848–852.
6. Bienz, M. (1984) *EMBO J.* **3,** 2477–2483.
7. Corces, V., Pellicer, A., Axel, R. & Meselson, M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 7038–7042.
8. Voellmy, R. & Rungger, D. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1776–1780.
9. Nevins, J. (1982) *Cell* **29,** 913–919.
10. Kingston, R. E., Baldwin, A. S. & Sharp, P. A. (1984) *Nature (London)* **312,** 280–282.
11. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15,** 1157–1174.
12. Benton, W. D. & Davis, R. W. (1977) *Science* **196,** 180–182.
13. Karch, F., Török, I. & Tissières, A. (1981) *J. Mol. Biol.* **148,** 219–230.
14. Maniatis, T., Fritsch, E. F. & Sambrook, S. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
16. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
17. Amin, J., Mestril, R., Lawson, R., Klapper, H. & Voellmy, R. (1985) *Mol. Cell. Biol.* **5,** 197–203.
18. Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. (1981) *Cell* **27,** 279–288.
19. Casadaban, M. J., Martinez-Arias, A., Shapira, S. K. & Chou, J. (1983) *Methods Enzymol.* **100,** 293–308.
20. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
21. Rungger, D. & Türler, H. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 6073–6077.
22. Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
23. Graham, F. L. & van der Eb, A. (1973) *Virology* **52,** 456–467.
24. Probst, E., Kressmann, A. & Birnstiel, M. L. (1979) *J. Mol. Biol.* **135,** 709–732.
25. Giorgi, C., Blumberg, B. & Kolakofski, D. (1983) *Cell* **35,** 829–836.
26. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **6,** 1175–1192.
27. Ingolia, T. D., Craig, E. A. & McCarthy, B. J. (1980) *Cell* **21,** 669–679.
28. Ingolia, T. D. & Craig, E. A. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 525–529.
29. Lawson, R., Mestril, R., Schiller, P. & Voellmy, R. (1984) *Mol. Gen. Genet.* **198,** 116–124.
30. Reynolds, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Luskey, K. L. (1984) *Cell* **38,** 275–285.
31. Holmgren, R., Corces, V., Morimoto, R., Blackman, R. & Meselson, M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3775–3778.
32. Dudler, R. & Travers, A. (1984) *Cell* **38,** 391–398.
33. Parker, C. S. & Topol, J. (1984) *Cell* **37,** 273–283.