Ryffel, *Cell* **46**, 1053 (1986); P. M. Druege *et al.*, *Nucleic Acids Res.* **14**, 9329 (1986); I. Theulaz, unpublished results.
5. P. Walker, M. Brown-Luedi, F. Givel, W. Wahli, *Nucleic Acids Res.* **12**, 8611 (1984).
6. K. Gorski, M. Carneiro, U. Schibler, *Cell* **47**, 161 (1986).
7. C. S. Parker and J. Topol, *ibid.* **37**, 273 (1984); P. A. Weil, D. S. Luse, J. Segall, R. G. Roeder, *ibid.* **18**, 469 (1979); J. L. Manley, A. Fire, A. Cano, P. A. Sharp, M. L. Gefter, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3855 (1980); J. D. Dignam, R. M. Lebowitz, R. G. Roeder, *Nucleic Acids Res.* **11**, 1475 (1983); Y. Suzuki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9522 (1986); W. Lee, A. Haslinger, M. Karin, R. Tjian, *Nature (London)* **325**, 368 (1987); C. Lefevre *et al.*, *EMBO J.* **6**, 971 (1987).
8. L. A. Laimins, G. Khoury, C. Gorman, B. Howard, P. Gruss, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6453 (1982).
9. M. A. Hayward, M. L. Brock, D. J. Shapiro, *Nucleic Acids Res.* **10**, 8273 (1982).
10. B. Westley and J. Knowland, *Cell* **15**, 365 (1978).
11. B. ten Heggeler-Bordier *et al.*, *EMBO J.* **6**, 1715 (1987).
12. S. L. McKnight and R. Kingsbury, *Science* **217**, 316 (1982).
13. We thank U. Schibler for advice on preparing the nuclear extracts, and O. Hagenbüchle and P. Shaw for comments on the manuscript. Supported by the Swiss National Science Foundation and the Etat de Vaud.

10 September 1987; accepted 13 January 1988

Germline Transformation Used to Define Key Features of Heat-Shock Response Elements

HUA XIAO AND JOHN T. LIS

The heat-shock consensus element (HSE), CTNGAANNTTCNAG, is found in multiple copies upstream of all heat-shock genes. Here, the sequence requirements for heat-shock induction are tested by *Drosophila* germline transformation with an *hsp*70–*lacZ* gene fused to a pair of synthetic HSEs. Certain single-base substitutions in either HSE cause a dramatic reduction (fortyfold) in expression. Surprisingly, variations in sequences immediately flanking the HSEs also reduced levels of induction. One such variant that contains two perfect 14–base pair HSEs, which are correctly spaced relative to each other and the TATA box, retained only 7% of wild type–induced expression. These and additional analyses indicate that the heat-shock regulatory element includes sequences beyond the 14–base pair HSE and may be better described as a dimer of a 10–base pair sequence, NTTCNNGAAN.

THE REGULATORY REGIONS OF highly expressed eukaryotic genes usually have multiple short sequence elements that function as binding sites for one or more regulatory proteins (*1*). A typical example is the heat-shock consensus element (HSE), CTNGAANNTTCNAG, which is found in multiple copies upstream of the transcription start of heat-shock genes from a wide variety of organisms (*2*). Sites containing this 14-bp sequence have been shown to bind purified protein, heat-shock factor (*3–5*). A *Drosophila hsp*70 gene with an upstream regulatory region that contains two HSEs is expressed at high levels in response to heat shock when introduced into the *Drosophila* genome, whereas the same gene with one HSE is expressed at 1/100 the level (*6, 7*). This led to the hypothesis that two copies of the HSE are absolutely required to achieve high levels of induced expression.

Our study uses germline transformation of *Drosophila melanogaster* (*8*) to examine the specific sequence requirements for heat-shock inducibility of the *hsp*70 gene. In germline transformants, an *hsp*70–*lacZ* fusion gene with regulatory sequences to position −89 shows the same range of inducibility (about 200-fold induction) as the endogenous *hsp*70 genes; however, this fusion gene and other related *hsp*70 fusion genes show relatively high basal level expression or low induced levels (3- to 20-fold induction) in transient expression assays in transfected cell cultures (*9*). The full range of inducibility in germline transformants provides a more sensitive measure of the effects of mutations on the expression of the *hsp*70 gene. *Hsp*70–*lacZ* hybrid genes with variant regulatory regions were generated, and stable transformants containing single copies of these genes were used to assay heat-induced expression of the transformed gene (*9*). The effect of line-to-line variation was minimized by obtaining an average value of 3 to 8 independent transformant lines for each variant (a total of 94 independent lines that have a single inserted *hsp*70–*lacZ* gene were examined in this study).

Two synthetic HSEs, which are perfect matches to the 14-bp HSE (*10*) conferred 27% of wild-type induction to an *hsp*70–

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853.

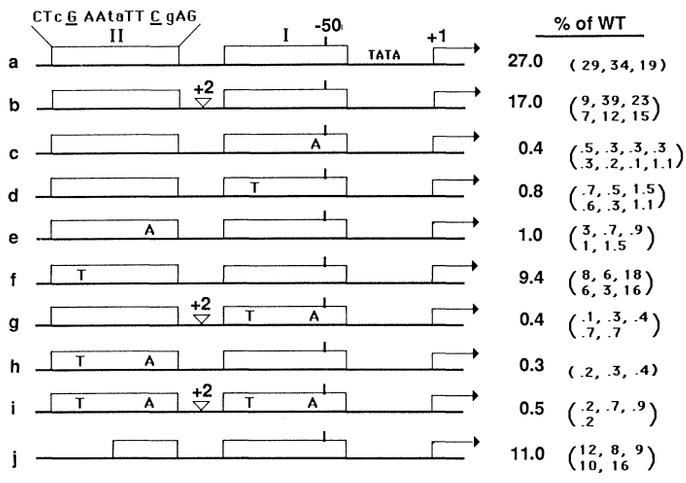

Fig. 1. Heat-shock–induced expression of *hsp*70 genes containing mutations in conserved nucleotides of the 14-bp HSE. Diagrams of *hsp*70–*lacZ* fusion genes shown in (**a** to **j**) represent the synthetic regulatory regions (*13, 14*) that were substituted for the native *hsp*70 regulatory region by fusion to position −50 (*15–17*). The transcription start of the genes is designated +1. The TATA homology between site I and the transcription start is shown in the diagrams. The two HSEs of sites I and II are shown as boxes, and the sequence of the synthetic 14-bp HSE is written above site II (see Fig. 3d for the complete sequence of the synthetic regulatory region of this construct). The underlined G and C residues were replaced by the noncomplementary transversions to T and A in the constructs indicated. The cloning to generate the constructs in this figure and in Figs. 3 and 4 was performed by standard methods (*18*). The fusion genes were subcloned into a P-element transformation vector c70T1, a Carnegie 20 derivative that contains the transcription termination sequence from *hsp*70 (*9*). Germline transformants with single inserts were backcrossed to the parental strain, Adh^{n6}, *cn*; ry^{502}. Heat-shock induction and measurements of β-galactosidase activity from adult female progeny by means of the substrate chlorophenol red–β-D-galactopyranoside were performed as described previously (*9*). Each transformant line was assayed in duplicate and the values are reported here and in experiments of Figs. 3 and 4 as the percentage of wild-type activity (% of WT), where wild type is an identical fusion gene containing native *hsp*70 regulatory sequences to position −89. In large type are the mean values of all measurements made on three to eight independent transformant lines (shown in parentheses). Since constructs g and i contain an extra 2 bp between the two mutant HSEs, construct b with this same 2-bp insert between two perfect synthetic HSEs is included for comparison.

lacZ hybrid gene (Fig. 1A). As shown in Fig. 1, c to e, a substitution for a single conserved base of either of the two HSEs reduced the heat-shock expression of the *hsp*70 gene to 0.4 to 1% of wild type (about 1/40 that of construct 1a). The supposition that this reduction is due simply to disruption of rotational symmetry is ruled out by the observation that the double-point mutations, which reestablish a perfect but different inverted repeat sequence (Fig. 1, g to i), also reduced expression to less than 1% of wild type.

Not all of the four related single-point mutations are functionally equivalent. A point mutation in either half of site I (Fig. 1, c and d) or the TATA-proximal half of site II (Fig. 1e) reduced heat shock–induced expression to barely detectable levels, while a point mutation in the distal half of site II caused only moderate reduction in expression (9.4%) (Fig. 1f). These results suggest that the proximal portion of site II in conjunction with a complete HSE in site I may provide a moderately active regulatory region. This was confirmed by constructing a promoter that consists of one and a half HSEs, one at the position of site I and a half in the proximal region of site II. The level of induced expression of this construct is 11% of wild type (Fig. 1j) and is similar to the point mutation of construct 1f. Thus, a half HSE in the proximal region of site II can complement a complete HSE in site I to render the *hsp*70 gene partially inducible.

The ability of a half HSE to complement a full HSE prompted a search for the half-HSE sequences in the previously defined functional regulatory regions for heat-shock genes. Heat-shock gene regulatory regions contain multiple copies of both halves of the previously defined 14-bp HSE, and surprisingly most of these do not lie side by side to generate a complete 14-bp HSE, but more often, they overlap by four bases to produce a 10-bp sequence, NTTCNNGAAN (Fig. 2). Overlapping HSEs have been previously noted in regulatory regions of heat-shock genes in a variety of organisms (*10*). A search of the regulatory regions of *Drosophila* heat-shock genes for this 10-bp sequence revealed a striking coincidence of this new consensus sequence with regions defined as critical for induction of these genes (Fig. 2). For example, sites I and II of the *hsp*70 gene, identified originally by their match to the 14-bp HSE consensus (*6*), can instead be viewed as each being composed of a tandem dimer of the 10-bp unit sequence. This enlarged 20-element is centered over the original 14-bp HSE and contains three additional base pairs at each end.

Some of the sequences immediately flanking the 14-bp HSEs were altered during the replacement of the native regulatory region with the synthetic regulatory region (Figs. 1a and 3d), and the resulting gene has a reduced level of expression (27% of wild type). This synthetic regulatory region retains the native sequence between the TATA box and the synthetic HSE of site I, thereby preserving a good match to the 10-bp consensus in the TATA-proximal region of site I (Figs. 2 and 3d). As a simpler test of the importance of these flanking bases, a construct was made in which the native TATA-proximal gTTCgcGAAa sequence (the conserved bases of the 10-bp unit are shown in capital letters) was changed to gTTCgAGtcg (it should be noted that the substitution of the central c with A creates a better match to the 14-bp HSE consensus). These substitutions (underlined) reduced the heat-induced expression of the gene to 33% of wild type (Fig. 3b). Additional constructs with one or more substitutions in the sequence

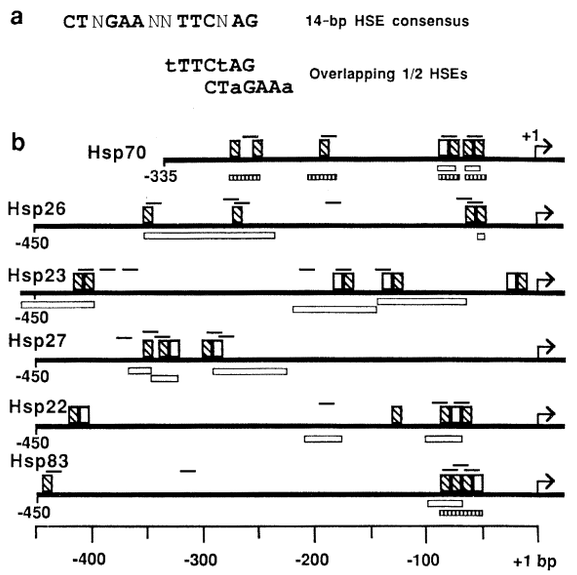

Fig. 2. Distribution of the 10-bp heat-shock consensus sequence and known heat-shock regulatory regions. (**a**) The 14-bp HSE consensus sequence derived by Pelham (*19*) is an inverted repeat. Computer searches of regulatory regions of heat-shock genes with the sequence of one arm of the inverted repeat reveals frequent 4-bp overlapping matches shown (overlapping half HSEs). (**b**) The 10-bp consensus sequence (NTTCNNGAAN) derived from the overlap of two half HSEs was used to search *Drosophila* heat-shock genes. The search was conducted with a program (*20*) that allows weights to be assigned to each nucleotide. The weights assigned (11, 41, 36, 40, 15, 15, 40, 36, 41, and 11 for the bases TTTCTAGAAA) represent the frequency of the preferred nucleotide in the corresponding positions of the 14-bp HSEs that have been found in regulatory regions of heat-shock genes [summarized in (*9*)]. The best 10-bp elements (diagonally hatched boxes) possess a numerical sum of matches >200 and possess correctly positioned C and G nucleotides. These matches are often found in tandem arrays, and less stringent matches to this consensus (open boxes) are shown when they immediately flank these best matches. The protein coding regions of *hsp*70, *hsp*26, *hsp*27, *hsp*23, and *hsp*22 gene contain only three "best" matches in a total of 4626 bp. Matches to the 14-bp HSE at seven or more of the ten specified positions are indicated by the solid lines above each map. Regions containing upstream regulatory elements identified by deletion analyses are designated by the open bars under the maps. The precise limits shown above are from the work of many laboratories: *hsp*70 (*6, 7, 9, 11*); *hsp*26 (*21, 22*); *hsp*23 (*23, 24*); *hsp*27 (*25*), further upstream elements have also been noted (*26*); *hsp*22 (*27, 28*); *hsp*83 (H. Xiao and J. T. Lis, unpublished results). Regions of *hsp*70 that bind heat-shock transcription factor (*3, 4*) or of *hsp*83 that bind heat-shock activator protein (*5*) are designated by the dashed bars under the maps. The drawing is to scale with +1 designating the transcription start site.

Fig. 3. Effects of base substitutions in the sequence immediately flanking the 14-bp HSEs. The sequence relevant to this study covers the region from positions −45 to −89. Matches to the HSE consensus are shown in black boxes in sites I and II, and those nucleotides that differ from the native sequence are underlined. The 10-bp consensus is shown below the TATA-proximal 10-bp sequence. The heat shock–induced β-galactosidase levels listed as % of WT were determined as in Fig. 1 from the raw data given in parentheses. (**a**) Native *hsp*70 regulatory sequence to position −89 (*7*). (**b**) An Xho I–Bam HI linker in place of the native sequence between site I and the TATA box. This construction was made by recombining pXNI and pMTI (*29*). (**c**) The synthetic 14-bp HSE in place of native site I and was generated by ligating native site II to pw1 (*13*). (**d**) The same construct shown in Fig. 1a. (**e**) The synthetic 14-bp HSE [from pHSE1 (*13*)] in place of native site II. (**f**) Two synthetic 14-bp HSEs separated by 9 bp and cloned into the Xho I site of pMTI (*29*).

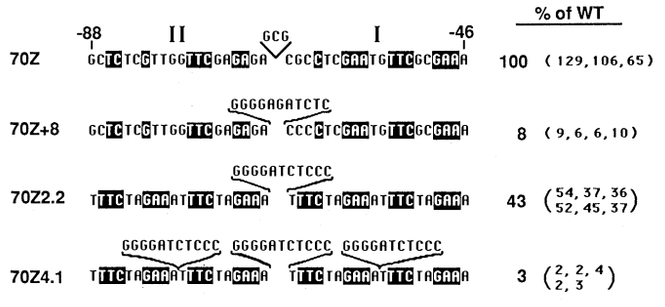

Fig. 4. The ability of a synthetic 10-bp sequence TTTCTAGAAA to specify heat-shock induction. Nucleotides that match the 10-bp consensus are in black boxes. The construct 70Z is the same as shown in Fig. 3a. The regulatory region of 70Z+8 contains an 8-bp insertion between native sites I and II, generated by inserting a 12-bp Bgl II linker into the Bss HII site, which has had its 4-bp sticky ends removed. Construct 70Z2.2 contains two copies of a tandem dimer of the 10-bp sequence, while 70Z4.1 contains four separated copies of the 10-bp sequence (*29*).

immediately flanking the 14-bp consensus in site I or site II also greatly reduced the induction of the gene (Fig. 3, c to e). Construct f (Fig. 3) retained only 7% of wild type–induced expression, although it contains two synthetic HSEs with perfect matches to the 14-bp consensus. However, sequences flanking the 14-bp HSEs contain substitutions that no longer match the dimer of the 10-bp unit consensus (Fig. 3). In all of these constructs, the wild-type spacing between the TATA box and the HSEs and between the two HSEs is preserved. These results suggest that a complete regulatory site of the *hsp*70 gene is a 20-bp sequence composed of two tandem 10-bp units.

Recently, Amin *et al.* (*11*) have generated a variety of mutations in the *Drosophila hsp*70 regulatory region and assayed their induction by transfection into cultured cells. In agreement with our findings, they conclude that the 14-bp HSE is insufficient to specify heat-shock induction. Support for the role of sequences flanking the 14-bp HSE was also obtained in a recent mutational analysis of the regulatory regions of a yeast *hsp*70 gene (*12*). We note that bases flanking these HSEs also match the ends of the tandem 10-bp unit.

To directly test the role of the 10-bp unit in heat-shock regulation, the regulatory region of the *hsp*70 gene was replaced with two different combinations of a synthetic 10-bp unit. The capacity of these synthetic promoters to drive heat shock–induced expression of the *Drosophila hsp*70 gene is shown in Fig. 4. The 70Z2.2 construct contains two copies of a tandem dimer of the 10-bp unit, and the two resulting 20-bp sequences are separated by 11 bp, equivalent to a turn of the DNA helix. The heat shock–induced level of expression of this construct is 43% of wild type. This is sixfold higher than the construct containing two perfect 14-bp HSEs (Fig. 3f). That the induced expression of 70Z2.2 is lower than that of the wild-type 70Z may be a consequence of the increased spacing between site I and site II in 70Z2.2. Adding 8 bp between the native sites I and II of 70Z to increase the spacing of these sites to equal that between the two synthetic elements of 70Z2.2 reduced the induction to 8% wild type (Fig. 4, 70Z+8). Thus 70Z2.2 is expressed at a fivefold higher level than an *hsp*70 gene which has two comparably spaced native HSE elements (70Z+8).

To test whether the 10-bp unit can act when not in a direct tandem array, we generated 70Z4.1. This construct has four 10-bp units, which are separated from one another by 11 bp. The level of its heat shock–induced expression is only 3% that of wild type (Fig. 4; also compare 70Z2.2 to 70Z4.1). Thus, the four 10-bp sequences in this separated configuration work less well than when arranged in tandem dimers.

The DNA sequences within the *hsp*70 regulatory region that bind heat-shock transcription factor have recently been examined at high resolution (*4*). Alkylation interference mapping has shown that the three to four bases that extend beyond each side of the 14-bp HSE consensus are important for binding heat-shock factor. This binding study is consistent with the data presented here that the heat-shock sites I and II of *hsp*70 are each composed of a tandem dimer of the 10-bp unit.

Is the 10-bp unit the basic building block from which heat-shock regulatory elements are built? The 10-bp consensus unit is itself a dyad of a 5-bp inverted repeat; therefore, each of the 5-bp repeats may be a site of recognition for a subunit or domain of heat-shock factor. Tandem arrays of these sequences may not need to be obligatorily arranged in whole multiples of the 10-bp consensus unit. Indeed, the most upstream binding site for heat-shock transcription factor (the region −262 to −248 of *hsp*70, Fig. 2) has a best fit to one and a half 10-bp sequences (aGAAtaTTCtaGAAt) (*3*).

This identification of a 10-bp heat-shock consensus unit also provides a new perspective from which to view the interaction of heat-shock transcription factor with the *hsp*83 regulatory region. This regulatory region was identified by the presence of three overlapping 14-bp consensus sequences (*7*). However, this region could be considered to have four tandem 10-bp sequences (Fig. 2) as has been independently suggested by Nover (*2*). The two internal 10-bp repeats exhibit the best match to the 10-bp consensus and might provide the tightest binding. Occupancy of all four sites should lead to a distinct repeated pattern of protection with heat-shock transcription factor on the same side of the DNA helix for each repeat.

To understand the molecular contacts required for activation of heat-shock genes, one must clearly delineate the identity of the basic DNA sequence units and the polypeptides with which they interact. The above studies necessitate a revised view of the basic sequence units from which heat-shock regulatory regions are built. However, the complete view awaits further analyses of regulatory regions in which the number and arrangement of basic sequence units are systematically varied and examined in terms of both their ability to program heat-induced expression and to bind heat-shock transcription factor.

REFERENCES AND NOTES

1. T. Maniatis, S. Goodbourn, J. A. Fischer, *Science* **236**, 1237 (1987).
2. L. Nover, *Enzyme Microb. Technol.* **9**, 130 (1987).
3. J. Topol, *et al.*, *Cell* **42**, 527 (1985).
4. D. J. Shuey and C. S. Parker, *J. Biol. Chem.* **261**, 7934 (1986).
5. V. Zimarino and C. Wu, *Nature (London)* **327**, 727 (1987).
6. R. Dudler and A. A. Travers, *Cell* **38**, 391 (1984).
7. J. A. Simon, C. A. Sutton, R. B. Lobell, R. L. Glaser, J. T. Lis, *ibid.* **40**, 805 (1985).
8. G. M. Rubin and A. C. Spradling, *Science* **218**, 348 (1982).
9. J. A. Simon and J. T. Lis, *Nucleic Acids Res.* **7**, 2971 (1987).
10. H. R. B. Pelham, *Trends Genet.* **1**, 31 (1985).
11. J. Amin, R. Mestril, P. Schiller, M. Dreano, R. Voellmy, *Mol. Cell. Biol.* **7**, 1055 (1987).
12. M. R. Slater and E. A. Craig, *ibid.*, p. 1906.
13. Two synthetic 14mers, one containing a perfect match to the heat-shock consensus, CTcGAAtaTTCgAG, and the other a double-point mutation, CTcTAAtaTTAgAG, were cloned between the Xho I and Sal I sites in plasmid pΔZX (*15*) to generate pHSE1 and pHSE2, respectively. The two synthetic sequences were then digested at the central Ssp I site and the two arms of the consensus were recombined to generate pHSE3 and pHSE4 containing the single-point mutants, CTcTAAtaTTCgAG and CTcGAAtaTTAgAG. To keep the native spacing between the synthetic consensus and the *hsp*70 TATA box, a tandem dimer of the synthetic 14mer was digested at its center with restriction enzyme Sst I, and the 4-bp sticky ends were removed with T4 DNA polymerase to generate a sequence, CTc(G/T)AAtaTT(C/A)g. This 12-bp sequence was fused to the blunt end of the Nru I site at −50 of *hsp*70 gene (CTcGAAtg*TTCg*|*cGa*, the Nru I site is in italics) to generate plasmid pw1 and the respective mutants pw2, pw3, and pw4, containing the sequence CTc(G/T)AAtaTT(C/A)g|cGa. Thus, the native spacing and sequence identity are preserved between the TATA box and the synthetic HSEs. To generate the regulatory regions containing two synthetic HSEs shown in Fig. 1, different combinations

of the synthetic HSEs were made by cloning the Xho I–Sal I fragments that contain the synthetic HSEs of pHSE1-4 into the Xho I site at the 5′ end of the HSEs of pw1-4. In these combinations, the spacing between two synthetic HSEs is 9 bp, the same as that between the two native copies in the regulatory region of the *hsp70* gene. The sequence of the synthetic regulatory regions of all the constructs was determined by DNA sequence analysis by the chain-termination method of Sanger *et al.* (*14*). Additional details of plasmid constructions are available on request.

14. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
15. An Xho I–Mbo I fragment from 5′ region of *hsp70* (Xho I at −194 and Mbo I at +260 of 132E3) (*16*) and a Bam HI fragment containing a 3-kb *lacZ* coding sequence from pMC1871 (*17*) were isolated and cloned into Xho I–Bam HI digested pΔZX [a pUC13X (*7*) derivative in which a Nar I–Eco RI fragment containing the entire *lacZ* sequence was removed]. Ligation of the Mbo I and Bam HI sticky ends put the two reading frames in phase. This hybrid gene contains only the first seven amino acids from *hsp70* fused to the eighth amino acid of *lacZ* and appears to express severalfold more β-galactosidase in transformant flies than the previously published *hsp70-lacZ* hybrid gene (*7*). All the constructs used in this study are derivatives of this new fusion gene.
16. S. Artavanis-Tsakonas, P. Schedl, M.-E. Mirault, L. Moran, J. Lis, *Cell* **17**, 9 (1979).
17. M. J. Casadaban, A. Martinez-Arias, S. K. Shapira, J. Chou, *Methods Enzymol.* **100**, 293 (1983).
18. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
19. H. R. B. Pelham, *Cell* **30**, 517 (1982).
20. Program written by and provided by R. Schleif (Brandeis University).
21. R. S. Cohen and M. Meselson, *Cell* **43**, 737 (1985).
22. D. Pauli, A. Spierer, A. Tissières, *EMBO J.* **5**, 755 (1986).
23. R. Mestril, P. Schiller, J. Amin, H. Klapper, A. Jayakumar, R. Voellmy, *ibid.*, p. 1667.
24. R. Mestril, D. Runger, P. Schiller, R. Voellmy, *ibid.* **4**, 2971 (1985).
25. G. Riddihough and H. R. B. Pelham, *ibid.* **5**, 1653 (1986).
26. E. Hoffman and V. Corces, *Mol. Cell. Biol.* **6**, 663 (1986).
27. R. Klemenz and W. J. Gehring, *ibid.*, p. 2011.
28. A. Amy, R. Southgate, A. Tissières, *J. Mol. Biol.* **182**, 469 (1985).
29. (i) To construct plasmid pMTI, we isolated an Msp I–Pst I fragment containing sequences from positions −38 to +89 of *hsp70* from pXTI (a pUC13X (*7*) derivative containing the Xho I–Pst I fragment from −89 to +89 of *hsp70* between the Xho I and Pst I sites), and the Msp I sticky ends were filled in using the Klenow fragment of DNA polymerase I. An Xho I–Bam HI adapter sequence, CCTCGAGGGATC, was added to the filled-in Msp I site, creating a Bam HI site at position −39. This fragment was then cloned into Xho I and Pst I digested pUC13X. (ii) To construct pXNI, we first digested pUC13X with restriction enzyme Xba I, and the sticky ends were removed with mung bean nuclease, and then digested with Xho I. The Xho I–Nru I fragment containing sequences from −89 to −50 of *hsp70* was isolated from pXTI and cloned into the Xba I– and Xho I–digested pUC13X. This construction recreated the 14-bp heat-shock consensus of site I with a Sal I site (in italics) immediately downstream of the consensus, CTcGAAtgTTCgAG*tcgac*. (iii) To construct plasmid pXBSB, we cloned a Bam HI–Pst I fragment from pMTI (−39 to +89 of *hsp70*) into pUC13 to generate plasmid pSB. pSB was then digested with restriction enzymes Bgl I and Sma I, and the vector fragment was purified and ligated to a purified Bgl I–Xho I nonvector fragment from pΔZX (Xho I end was filled in with the Klenow fragment), and excess Bgl II linker, GGGAGATCTCCC. The Bgl II linker was inserted between the Sma I and Xho I sites. The sequence TTTCTAGAAA was then cloned as dimer or monomer into the Sma I site of pXBSB. The Bgl II–Bam HI fragment containing the dimer or monomer was polymerized with T4 DNA ligase to generate the regulatory regions for 70Z2.2 and 70Z4.1, respectively.
30. We thank J. Werner for embryo microinjections, R. Schleif for the use of his program, and B. Glaser, O. Perisic, C. Sutton, and M. Wolfner for comments on the manuscript. Supported by NIH grant GM25232.

1 September 1987; accepted 5 January 1988

Multipotent Precursors Can Give Rise to All Major Cell Types of the Frog Retina

RICHARD WETTS* AND SCOTT E. FRASER

A prospective lineage analysis was performed to determine the variety of cell types that could be formed by individual precursor cells of the developing frog retina. Fluorescent dextran was iontophoretically injected into single cells of the embryonic optic vesicle. After further development of the embryo, labeled descendants were observed in all three layers of the larval retina. Furthermore, different clones were composed of various combinations of all major cell types, including the glial Müller cells. Hence, single optic vesicle cells have the potential to form any type of retinal cell, suggesting that the interactions that specify the differentiation pathway of retinal cells must occur late in development.

ONE OF THE CENTRAL QUESTIONS of developmental biology is how different cell types are created in the correct numbers and positions. One approach to this question is to examine the lineages of the cells that make up the embryo. The cell lineage reveals what cell types share common ancestors and when different lineages diverge. In some embryos, such as that of the nematode *Caenorhabditis elegans*, the complete cell lineage was elucidated by direct observation of the embryo with Nomarski optics (*1*). These data were required for subsequent analyses of both the cell interactions and the molecular mechanisms that might subserve them (*2*). Such an approach has not been possible in the vertebrate nervous system, because the large numbers and undifferentiated state of the precursor cells make it impossible to directly identify and follow single cells and their descendants. Thus, it remains possible that the wide variety of neural cell types arises either from a collection of prespecified precursors, each restricted to only one cell type, or from multipotent precursors, each able to give rise to the full range of cell types.

We have examined cell lineages in the neural retina of the frog by injecting single cells of the optic vesicle with lysinated rhodamine dextran (LRD). Fluorescent dextrans (*3*) serve well as lineage markers because their size and charge prevent their escape from the injected cell or from its descendants through either cell junctions or cell membranes. Microinjection of fluorescent dextrans into blastomeres, which are relatively large, making it easy to inject dye into them, has permitted studies of early embryonic lineages in a variety of invertebrates (*4*) and vertebrates (*5*). We have refined the techniques for iontophoretically injecting fluorescent dextrans into small and fragile cells, such as retinal precursor cells. The neural retina has been extensively studied (*6*) and offers several advantages for cell lineage studies. (i) It has relatively few cell types: the photoreceptors in the outer nuclear layer (ONL); the horizontal, bipolar, amacrine, and glial Müller cells in the inner nuclear layer (INL); and the ganglion cells in the ganglion layer (GL). The laminar organization to some extent facilitates the identification of cell types. (ii) The precursor cells in the lateral portion of the optic vesicle are accessible for microinjection. (iii) In the frog, the cells quickly develop into a functional retina. Neuron birthdates begin soon after the optic vesicle stage, at stage 29 (*7*), and the lamination begins around stage 39 (3 days after fertilization).

At the end of each microinjection (*8*), direct visual observation of a single fluorescently labeled cell confirmed a successful dye fill. In some cases this visual confirmation was ambiguous because of suboptimal viewing conditions (distortion by the overlying tissue or by the meniscus of the solution). To independently verify that the technique consistently yields single dye-filled cells, six animals with several injections each were fixed immediately (within 5 minutes) after injection. Histological sections (Fig. 1A) revealed that 16 of 17 injections resulted in single labeled cells. In the remaining case,

Department of Physiology and Biophysics and the Developmental Biology Center, University of California, Irvine, CA 92717.

*To whom correspondence should be addressed.