during this process. Application of a large imaging force (applied vertical force >100 nN) to a selected region of a living cell produced a hole in the cell without jeopardizing cell viability (Fig. 4C). This "nanosurgery" experiment was readily reproducible. The hole may result from fusion of the two bilayers under pressure from the AFM tip and cannot be taken as direct evidence of a single bilayer penetration mechanism. Nonetheless, these data show that a living cell can withstand severe perturbation of the plasma membrane by the scanning tip, a quality that may prove very useful in future studies of living cells by the AFM.

Although the AFM is considered a surface probe, it is capable of imaging subsurface features such as F-actin within living cells. The AFM can collect 3-D data and is capable of higher resolution imaging than conventional optical microscopes on many surfaces. If this resolving capability can be realized with living cells, the AFM should prove a powerful tool for the investigation of dynamic interactions between cell surface and subsurface structures, a critical step in many signal transduction pathways.

REFERENCES AND NOTES

1. S. J. Smith, *Science* **242**, 708 (1988).
2. G. Binnig, C. F. Quate, C. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
3. S. A. C. Gould *et al.*, *J. Vac. Sci. Technol. A* **8**, 369 (1990); E. Henderson, *Nucleic Acids Res.* **20**, 445 (1992); J. H. Hoh, R. Lal, S. A. John, J.-P. Revel, M. F. Arnsdorf, *Science* **253**, 1405 (1991); B. Drake *et al.*, *ibid.* **243**, 1586 (1989); H. J. Butt *et al.*, *J. Struct. Biol.* **105**, 54 (1990); C. Bustamante *et al.*, *Biochemistry* **31**, 22 (1991); H. J. Butt, K. H. Downing, P. K. Hansma, *Biophys. J.* **58**, 1473 (1990); V. Parpura, personal communications.
4. R. D. Edstrom *et al.*, *FASEB J.* **4**, 3144 (1990); O. Marti *et al.*, *J. Microsc.* **152**, 803 (1988); C. A. J. Putman *et al.*, *Ultramicroscopy*, in press; T. Tomie *et al.*, *Science* **252**, 691 (1991); J. Vesenka *et al.*, *Ultramicroscopy*, in press; J. Vesenka *et al.*, *SPIE Proc.* **1639**, 38 (1992); A. L. Weisenhorn *et al.*, *Scanning Electron Microsc.* **4**, 511 (1990); A. L. Weisenhorn *et al.*, *Biophys. J.* **58**, 1251 (1990); A. L. Weisenhorn *et al.*, *Langmuir* **7**, 8 (1991); D. Keller *et al.*, *SPIE Proc.* **1639**, 91 (1992).
5. D. S. Sakaguchi *et al.*, *Dev. Biol.* **134**, 158 (1989).
6. J. H. Hoh and P. K. Hansma, *Trends Cell Biol.* **2**, 208 (1992).
7. C. A. J. Putnam *et al.*, *Rev. Sci. Instrum.* **63**, 1914 (1992).
8. M. Radmacher, R. M. Zimmermann, H. E. Gaub, in *The Structure and Conformation of Amphiphilic Membranes*, R. Lipowski, D. Richter, K. Kremer, Eds. (Springer Proceedings in Physics, Springer-Verlag, Berlin, 1992), vol. 66, pp. 24–29.
9. Cultures were rinsed with phosphate-buffered saline (PBS) and fixed in 4% paraformaldehyde in a 0.1 M PO_4 buffer at 22°C for 1 to 2 hours. The cultures were rinsed in PBS, blocked for 30 min in PBS plus 5% goat serum, and incubated in primary antibodies overnight at 4°C. They were then rinsed in PBS and incubated with fluorescently conjugated secondary antibodies for 90 min, rinsed again, and mounted on glass microscope slides. Primary antibodies were diluted in PBS that contained 0.3% Triton X-100 (Sigma). The antibody to β-tubulin (E7, from M. Klymkowsky) was diluted 1:20. Fluorescent secondary antibody (fluorescein isothiocyanate–conjugated goat antibody to mouse; Fisher Scientific, Springfield, NJ) was diluted 1:150 in PBS. Cultures were incubated with rhodamine-phalloidin (1:30) in 0.3% Triton X-100 in PBS for 30 min after the incubations in the secondary antibodies in the dark at room temperature. Cultures were then rinsed three times in PBS and mounted on glass microscope slides with Gelmount (Fisher). Fluorescence was detected with an Odyssey confocal laser scanning microscope (Noran Instruments, Middleton, WI). For the cytochalasin B sensitivity experiment, cytochalasin B (Sigma, dissolved at a final concentration of 20 mM in dimethyl sulfoxide) was diluted to 10 μM with growth medium. During scanning, the cytochalasin B was added from the side with a fine glass needle to the AFM fluid cell (Digital Instruments) in which the O-ring seal had been omitted. Approximately 10 μl of cytochalasin B solution was added to the 200-μl imaging chamber to give a final concentration of ~0.5 μM. Addition of the cytochalasin B took approximately 30 s; the image shown (Fig. 3B, right panel) was acquired immediately after cytochalasin B addition for a total time course of approximately 4.5 min. The ports in the fluid cell were not used to avoid detrimental hydrodynamic effects on image acquisition and cell adhesion.

10. We thank M. Klymkowsky for antibody to β-tubulin, V. Parpura for developing and sharing AFM imaging methods for living cells, and K. Radke for expert technical assistance. The authors are members of the Signal Transduction Training Group at Iowa State University supported by NSF. Supported by a grant from the Iowa State University Biotechnology Council (D.S.S.) and by NIH and NSF (E.H. and P.G.H.). P.G.H. is an Alfred P. Sloan Fellow.

23 April 1992; accepted 8 July 1992

Multiple Zinc Finger Forms Resulting from Developmentally Regulated Alternative Splicing of a Transcription Factor Gene

Tien Hsu, Joseph A. Gogos, Susan A. Kirsh,* Fotis C. Kafatos†

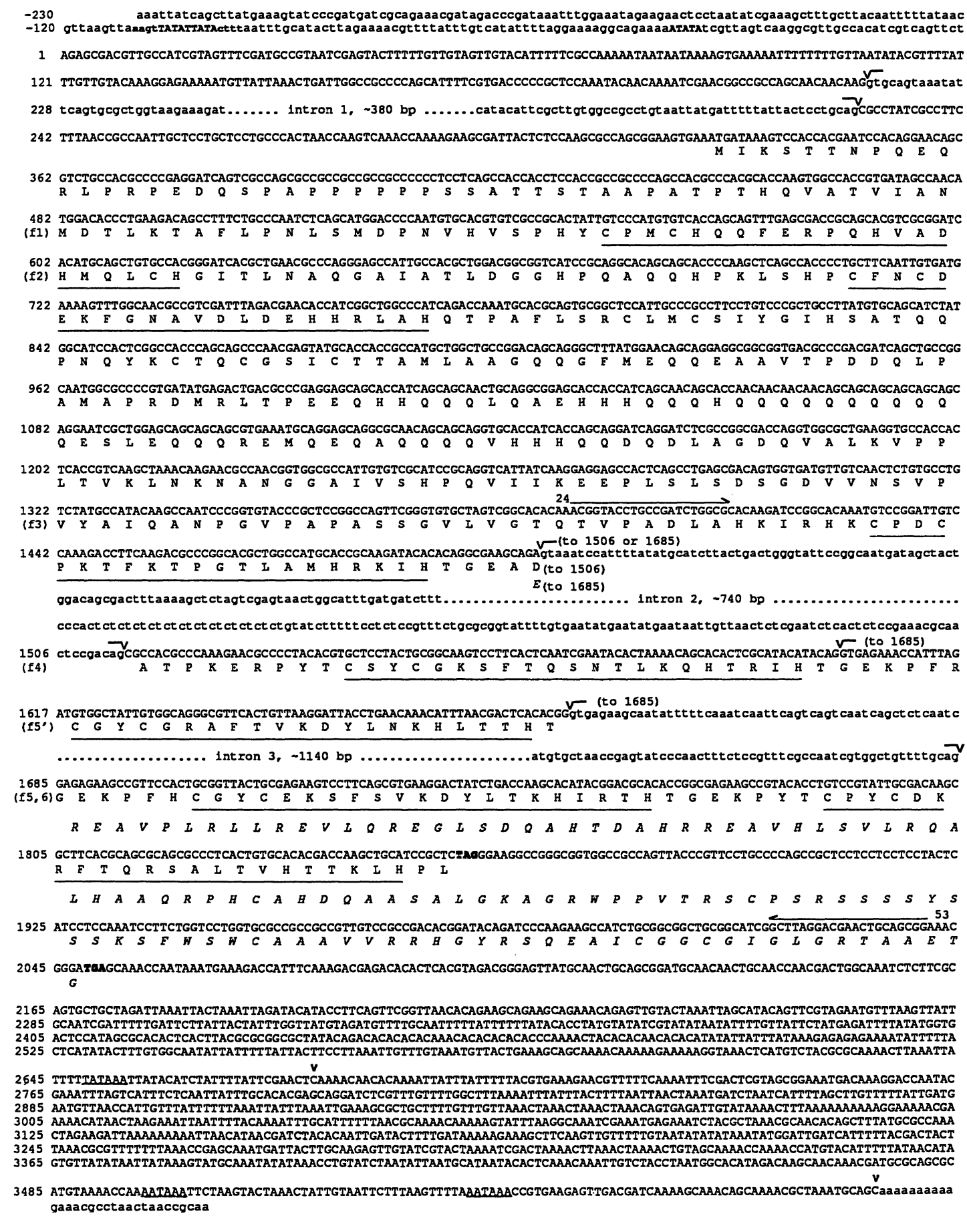
Transcripts encoding the *Drosophila* putative transcription factor CF2 are subject to developmentally regulated alternative splicing, and they encode protein isoforms that differ in the number of zinc fingers. One testis-specific RNA encodes an isoform that includes three zinc fingers and a frame-shifted segment. Two other transcripts encode isoforms with six and seven zinc fingers which bind to distinct promoters and DNA target sequences. Thus, because of alternative splicing, a single gene appears to encode distinct DNA-binding proteins, each capable of regulating different gene sets in different tissues and developmental periods.

Transcription factors can serve as developmental switches, that is, they turn specific sets of genes on or off at appropriate locations and at different times during development (*1*). Transcription factors cannot be regulated only by transcription, because such regulation would require an infinity of factors. Their posttranscriptional modulation occurs by alternative splicing (*2*), intracellular compartmentalization (*3*), interaction with inhibitors (*4*), translational regulation (*5*), or posttranslational modification such as phosphorylation (*6*). We now present evidence that posttranscriptional regulation of a Cys_2-His_2 zinc finger factor can occur by alternative splicing within the DNA-binding domain, thus utilizing the modularity of

T. Hsu, J. A. Gogos, S. A. Kirsh, Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138.
F. C. Kafatos, Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138, and also the Institute of Molecular Biology and Biotechnology, FORTH and Department of Biology, University of Crete, Heraklion 711 10, Crete, Greece.

*Present address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114.
†To whom correspondence should be addressed.

Fig. 1. Nucleotide sequence of the *CF2* gene coding strand and derived amino acid sequences. Exon sequences are numbered consecutively from the first of three mRNA 5′ ends (·) detected by primer extension (*8*). In the 5′ untranscribed region, the TATA box (ATATA) and a partial match to the CF2-I protein binding consensus (TATATTATA) are shown in upper case; bold lower case indicates nucleotides flanking the consensus, which are encompassed in CF2-I protein footprints (Fig. 3B). Portions of the intron sequences near the splice junctions are shown in lower case but not numbered. (↘) indicates splice donor and (→↘) splice acceptor sites. Possible zinc finger motifs in the peptide sequences are underlined, and designations (f1, f2) on the left mark their beginnings. The additional line of peptide sequence (italicized) from position 1505 on is the frame-shifted CF2-III peptide. Arrows above the nucleotide sequence indicate the oligonucleotide primers used in PCR reactions (Fig. 2, C and D). Underlined residues in the 3′ untranslated region indicate possible polyadenylation signals, which are consistent with the 3′ ends of cDNA clones (vertical arrow heads) and the results of S1 protection assays (*8*). The designations for amino acid residues are in the standard single-letter code: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.


-230 aaattatcagcttatgaaagtatcccgatgatcgcagaaacgatagacccgataaatttggaaatagaagaactcctaatatcgaaagctttgcttacaatttttataac
-120 gttaagttaaagtTATATTATActttaatttgcatacttagaaaacgttttatttgtcatattttaggaaaaggcagaaaaATATAtcgttagtcaaggcgttgccacatcgtcagttct
1 AGAGCGACGTTGCCATCGTAGTTTCGATGCCGTAATCGAGTACTTTTTGTTGTAGTTGTACATTTTTCGCCAAAAATAATAATAAAAGTGAAAAATTTTTTTTGTTAATATACGTTTTAT
121 TTGTTGTACAAAGGAGAAAAATGTTATTAAACTGATTGGCCGCCCCAGCATTTTCGTGACCCCCGCTCCAAATACAACAAAATCGAACGGCCGCCAGCAACAACAAGGtgcagtaaatat
228 tcagtgcgctggtaagaaagat..... intron 1, ~380 bp .....catacattcgcttgtggccgcctgtaattatgatttttattactcctgcagCGCCTATCGCCTTC
242 TTTAACCGCCAATTGCTCCTGCTCCTGCCCACTAACCAAGTCAAACCAAAAGAAGCGATTACTCTCCAAGCGCCAGCGGAAGTGAAATGATAAAGTCCACCACGAATCCACAGGAACAGC
M I K S T T N P Q E Q
362 GTCTGCCACGCCCCGAGGATCAGTCGCCAGCGCCGCCGCCGCCCCCCTCCTCAGCCACCACCTCCACCGCCGCCCCAGCCACGCCCACGCACCAAGTGGCCACCGTGATAGCCAACA
R L P R P E D Q S P A P P P P P P S S A T T S T A A P A T P T H Q V A T V I A N
482 TGGACACCCTGAAGACAGCCTTTCTGCCCAATCTCAGCATGGACCCCAATGTGCACGTGTCGCCGCACTATTGTCCCATGTGTCACCAGCAGTTTGAGCGACCGCAGCACGTCGCGGATC
(f1) M D T L K T A F L P N L S M D P N V H V S P H Y C P M C H Q Q F E R P Q H V A D
602 ACATGCAGCTGTGCCACGGGATCACGCTGAACGCCCAGGGAGCCATTGCCACGCTGGACGGCGGTCATCCGCAGGCACAGCAGCACCCCAAGCTCAGCCACCCCTGCTTCAATTGTGATG
(f2) H M Q L C H G I T L N A Q G A I A T L D G G H P Q A Q Q H P K L S H P C F N C D
722 AAAAGTTTGGCAACGCCGTCGATTTAGACGAACACCATCGGCTGGCCCATCAGACCAAATGCACGCAGTGCGGCTCCATTGCCCGCCTTCCTGTCCCGCTGCCTTATGTGCAGCATCTAT
E K F G N A V D L D E H H R L A H Q T P A F L S R C L M C S I Y G I H S A T Q Q
842 GGCATCCACTCGGCCACCCAGCAGCCCAACGAGTATGCACCACCGCCATGCTGGCTGCCGGACAGCAGGGCTTTATGGAACAGCAGGAGGCGGCGGTGACGCCCGACGATCAGCTGCCGG
P N Q Y K C T Q C G S I C T T A M L A A G Q Q G F M E Q Q E A A V T P D D Q L P
962 CAATGGCGCCCCGTGATATGAGACTGACGCCCGAGGAGCAGCACCATCAGCAGCAACTGCAGGCGGAGCACCACCATCAGCAACAGCACCAACAACAACAACAGCAGCAGCAGCAGCAGC
A M A P R D M R L T P E E Q H H Q Q Q L Q A E H H H Q Q Q H Q Q Q Q Q Q Q Q Q Q
1082 AGGAATCGCTGGAGCAGCAGCAGCGTGAAATGCAGGAGCAGGCGCAACAGCAGCAGGTGCACCATCACCAGCAGGATCAGGATCTCGCCGGCGACCAGGTGGCGCTGAAGGTGCCACCAC
Q E S L E Q Q Q R E M Q E Q A Q Q Q Q V H H H Q Q D Q D L A G D Q V A L K V P P
1202 TCACCGTCAAGCTAAACAAGAACGCCAACGGTGGCGCCATTGTGTCGCATCCGCAGGTCATTATCAAGGAGGAGCCACTCAGCCTGAGCGACAGTGGTGATGTTGTCAACTCTGTGCCTG
L T V K L N K N A N G G A I V S H P Q V I I K E E P L S L S D S G D V V N S V P
1322 TCTATGCCATACAAGCCAATCCCGGTGTACCCGCTCCGGCCAGTTCGGGTGTGCTAGTCGGCACACAAACGGTACCTGCCGATCTGGCGCACAAGATCCGGCACAAATGTCCGGATTGTC
(f3) V Y A I Q A N P G V P A P A S S G V L V G T Q T V P A D L A H K I R H K C P D C
1442 CAAAGACCTTCAAGACGCCCGGCACGCTGGCCATGCACCGCAAGATACACACAGGCGAAGCAGAgtaaatccattttatatgcatcttactgactgggtattccggcaatgatagctact
P K T F K T P G T L A M H R K I H T G E A D
ggacagcgactttaaaagctctagtcgagtaactggcatttgatgatctttt..... intron 2, ~740 bp .....
cccactctctctctctctctctctctctctctgtatctttttcctctccgtttctgcgcggtattttgtgaatatgaatatgaataattgttaactctcgaatctcactctccgaaacgcaa
1506 ctccgacagCGCCACGCCCAAAGAACGCCCCTACACGTGCTCCTACTGCGGCAAGTCCTTCACTCAATCGAATACACTAAAACAGCACACTCGCATACATACAGGTGAGAAACCATTTAG
(f4) A T P K E R P Y T C S Y C G K S F T Q S N T L K Q H T R I H T G E K P F R
1617 ATGTGGCTATTGTGGCAGGGCGTTCACTGTTAAGGATTACCTGAACAAACATTTAACGACTCACACGGgtgagaagcaatatttttcaaatcaattcagtcagtcaatcagctctcaatc
(f5') C G Y C G R A F T V K D Y L N K H L T T H T
..... intron 3, ~1140 bp .....atgtgctaaccgagtatcccaactttctccgtttcgccaatcgtggctgttttgcag
1685 GAGAGAAGCCGTTCCACTGCGGTTACTGCGAGAAGTCCTTCAGCGTGAAGGACTATCTGACCAAGCACATACGGACGCACACCGGCGAGAAGCCGTACACCTGTCCGTATTGCGACAAGC
(f5,6) G E K P F H C G Y C E K S F S V K D Y L T K H I R T H T G E K P Y T C P Y C D K
R E A V P L R L L R E V L Q R E G L S D Q A H T D A H R R E A V H L S V L R Q A
1805 GCTTCACGCAGCGCAGCGCCCTCACTGTGCACACGACCAAGCTGCATCCGCTCTAGGGAAGGCCGGGCGGTGGCCGCCAGTTACCCGTTCCTGCCCCAGCCGCTCCTCCTCCTCCTACTC
R F T Q R S A L T V H T T K L H P L
L H A A Q R P H C A H D Q A A S A L G K A G R W P P V T R S C P S R S S S S Y S
1925 ATCCTCCAAATCCTTCTGGTCCTGGTGCGCCGCCGCCGTTGTCCGCCGACACGGATACAGATCCCAAGAAGCCATCTGCGGCGGCTGCGGCATCGGCTTAGGACGAACTGCAGCGGAAAC
S S K S F W S W C A A A V V R R H G Y R S Q E A I C G G C G I G L G R T A A E T
2045 GGGATGAGCAAACCAATAAATGAAAGACCATTTCAAAGACGAGACACACTCACGTAGACGGGAGTTATGCAACTGCAGCGGATGCAACAACTGCAACCAACGACTGGCAAATCTCTTCGC
G
2165 AGTGCTGCTAGATTAAATTACTAAATTAGATACATACCTTCAGTTCGGTTAACACAGAAGCAGAAGCAGAAACAGAGTTGTACTAAATTAGCATACAGTTCGTAGAATGTTTAAGTTATT
2285 GCAATCGATTTTTGATTCTTATTACTATTTGGTTATGTAGATGTTTTGCAATTTTTATTTTTTATACACCTATGTATATCGTATATAATATTTTGTTATTCTATGAGATTTTATATGGTG
2405 ACTCCATAGCGCACACTCACTTACGCGCGGCGCTATACAGACACACACACAAACACACACACCCAAAACTACACACAACACACATATATTATTTATAAAGAGAGAGAAAATATTTTTA
2525 CTCATATACTTTGTGGCAATATTATTTTTATTACTTCCTTAAATTGTTTGTAAATGTTACTGAAAGCAGCAAAACAAAAAGAAAAAGGTAAACTCATGTCTACGCGCAAAACTTAAATTA
2645 TTTTTATAAATTATACATCTATTTTATTCGAACTCAAAACAACACAAAATTATTTATTTTACGTGAAAGAACGTTTTTCAAAATTTCGACTCGTAGCGGAAATGACAAAGGACCAATAC
2765 GAAATTTAGTCATTTCTCAATTATTTGCACACGAGCAGGATCTCGTTTGTTTTGGCTTTAAAATTTATTTACTTTTAATTAACTAAATGATCTAATCATTTTAGCTTGTTTTTATTGATG
2885 AATGTTAACCATTGTTTATTTTTTAAATTATTTAAATTGAAAGCGCTGCTTTTGTTTGTTAAACTAAACTAAACTAAACAGTGAGATTGTATAAAACTTTAAAAAAAAAAGGAAAAACGA
3005 AAAACATAACTAAGAAATTAATTTTACAAAATTTGCATTTTTTAACGCAAAACAAAAAGTATTTAAGGCAAATCGAAATGAGAAATCTACGCTAAACGCAACACAGCTTTATGCGCCAAA
3125 CTAGAAGATTAAAAAAAAATTAACATAACGATCTACACAATTGATACTTTTGATAAAAAGAAAGCTTCAAGTTGTTTTTGTAATATATATAAATATGGATTGATCATTTTTACGACTACT
3245 TAAACGCGTTTTTTTAAACCGAGCAAATGATTACTTGCAAGAGTTGTATCGTACTAAAATCGACTAAAACTTAAACTAAAACTGTAGCAAAACCAAAACCATGTACATTTTTATAACATA
3365 GTGTTATATAATTATAAAGTATGCAAATATATAAACCTGTATCTAATATTAATGCATAATACACTCAAACAAATTGTCTACCTAATGGCACATAGACAAGCAACAAACGATGCGCAGCGC
3485 ATGTAAAACCAAAATAAATTCTAAGTACTAAACTATTGTAATTCTTTAAGTTTTAAATAAACCGTGAAGAGTTGACGATCAAAAGCAAACAGCAAAACGCTAAATGCAGCaaaaaaaaaa
gaaacgcctaactaaccgcaa

DNA recognition by each zinc finger.

The chorion transcription factor CF2 was isolated as a potential transcriptional regulator of a chorion (eggshell) gene of *Drosophila melanogaster* and is a zinc finger protein (*7*). As would be expected, immunocytochemistry with a CF2 antibody reveals that, consistent with its proposed function, CF2 is present in the nuclei of follicle cells, which surround the oocyte and produce the chorion (*8*). Localized CF2 is not detected in the early embryo but appears during later embryonic stages in epidermal nuclei of all thoracic and abdominal segments (*8*). It is unknown if any of four embryonic lethal complementation groups mapping at 25A5–8 correspond to the *CF2* locus (*9*).

The original CF2 clone isolated from a cDNA library derived from embryos is incomplete but includes a COOH-terminal, putative DNA binding domain of four Cys_2-His_2 zinc finger motifs (*7*). This incomplete CF2 protein can bind specifically to several different chorion gene promoters from both *Drosophila* and silkmoths (*7, 10*). To further characterize CF2 structure and function, we isolated additional cDNA clones from cDNA libraries derived from embryonic and ovarian tissues. In addition, an approximately 16-kb genomic fragment, encompassing the cDNA sequences and 10 kb of 5′ flanking DNA, was isolated and characterized from a *Drosophila* wild-type (Oregon-R) library isogenic for the *dp cn bw* second chromosomes (Fig. 1).

Two polyadenylated CF2 RNA species, 2.65 and 3.6 kb, are detected throughout development, but most abundantly in egg chambers (*7*). Northern (RNA) hybridizations indicate that both RNA species share sequences at their 5′ ends but differ at the 3′ ends (*8*). Comparative polymerase chain reaction (PCR) mapping of cDNA and genomic DNA (*8*), combined with direct sequencing of genomic clones, shows that there are no introns within the 3′ untranslated region. On the basis of cDNA sequence information, Northern blots, and S1 protection assays (*8*), it appears that the 3.6-kb transcript ends after two AATAAA motifs, approximately 1.73 kb downstream of the protein termination codon; and the 2.65-kb transcript is truncated at the 3′ end, after a modified TATAAA motif approximately 0.82 kb downstream of the termination codon (Fig. 1). Three closely spaced RNA 5′ ends were mapped by primer extension to lie approximately 70 nucleotides upstream of the longest cDNA clone (Fig. 1). A single intron, intron 1, interrupts the 5′ untranslated region (Figs. 1 and 2A), and two other introns are found in the protein-encoding region (below).

Near the NH_2-terminus and in the middle of the protein, the CF2 coding region contains proline-rich and glutamine-rich segments (Fig. 1). These features are encountered frequently in transcription factors and may correspond to activation domains (*11*). Two zinc fingers are located near the NH_2-terminus (fingers 1 and 2, Fig. 2, A and B), in addition to the four zinc finger motifs (fingers 3, 4, 5, and 6) previously identified near the COOH-terminus (*7*). The significance of fingers 1 and 2 is unclear, because the DNA recognition properties of CF2 are unaltered by their presence or absence (*12*). On the other hand, binding-site selection experiments (*12*) have failed to identify a recognized sequence for figure 3. Thus, the first three zinc fingers do not appear to contribute to specific DNA sequence recognition, perhaps because the long and presumably flexible linker sequences separating them (29, 219, and 14 amino acids between fingers 1 and 2, 2 and 3, and 3 and 4, respectively; Fig. 2B) may prevent formation of a continuous DNA recognition domain. Because individual fingers are not expected to have a strong enough interaction with DNA to impart specificity (*12, 13*), fingers 1, 2, and 3 might serve as "DNA-holding" structures, stabilizing the specific protein-DNA interactions that are dictated by the COOH-terminal finger domain.

The COOH-terminal zinc finger region is interrupted by two moderately sized introns (about 0.74 and 1.2 kb). Intron 2 lies between 3 and 4, whereas intron 3 is between fingers 4 and 5 (Figs. 1 and 2A). This creates the potential for alternative splice acceptor selection that would produce two CF2 RNA isoforms, each encoding different numbers of zinc fingers. In addition, the sequence of the proximal end of intron 3 suggests that it may encode an additional zinc finger downstream of finger 4, followed by an alternative splice donor consensus sequence (Figs. 1 and 2A). This would create an additional protein isoform, with an extra zinc finger inserted between fingers 4 and 5. Indeed, RNA splice variants consistent with the existence of all three of these isoforms were detected by PCR analysis and by cDNA cloning. These

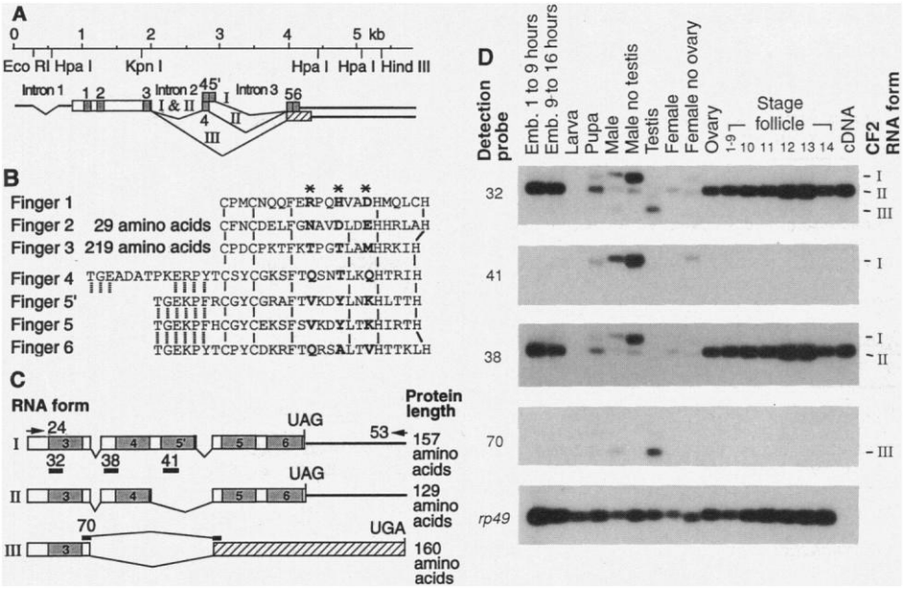


Fig. 2. (**A**) *CF2* genomic structure and alternative splicing events. The scale shows distances from the transcription start site (Fig. 1). Rectangles, protein-encoding regions; stippled numbered boxes, zinc finger motifs; hatched boxes, frame-shifted peptide of CF2-III. Roman numerals indicate the isoforms that are generated by each alternative splicing event. (**B**) Amino acid sequence comparison of the CF2 zinc fingers and preceding linker sequences. The distances between fingers 1 and 2, and fingers 2 and 3, are indicated. Conserved residues found in most zinc finger proteins are aligned by vertical lines (solid in the finger region, dashed in the linker). Amino acid residues implicated in DNA recognition are indicated with asterisks and boldface (*12, 13*). (**C**) Regions amplified by PCR from the three CF2 RNA forms. The introns are not shown in scale. Two short opposing arrows (No. 24 and No. 53) mark the two oligonucleotide primers that were used in the PCR reactions (D). Four short solid lines (Nos. 32, 38, 41, and 70) indicate the four oligonucleotide probes used for identifying CF2 RNA forms (*16*) (Fig. 2D). Primer No. 70 encompasses the splice junction that generates CF2-III RNA; therefore, forms I and II could not be detected by this probe under our hybridization conditions. (**D**) Detection of CF2 mRNA forms by PCR (*16*). Total RNA was prepared as previously described (*17*). PCR primers flanking the CF2 COOH-terminal zinc finger domain are indicated in Figs. 1 and (C), and primers specific for the *rp49* gene were included to provide an internal control (*16*). The PCR products were fractionated by gel electrophoresis and were Southern (DNA) blotted with RNA form-specific oligonucleotide detection probes (*16*) (C). The RNA samples used are indicated on the top of each lane, and the detection probes are shown on the left. Interpretation of the CF2 RNAs detected is indicated on the right.

mRNAs proved to be developmentally regulated (Fig. 2D).

A pair of oligonucleotide primers flanking the COOH-terminal finger domain (oligonucleotides 24 and 53; Figs. 1 and 2C) were used to amplify (by PCR) CF2 RNA sequences from various developmental stages and tissues. The PCR products were size fractionated by gel electrophoresis and Southern (DNA) blotted with specific oligonucleotide probes, revealing three distinct RNA spliced forms that were named CF2-I, CF2-II, and CF2-III in order of decreasing sizes (Fig. 2D). Fragments amplified by PCR representing all three forms were cloned and at least two isolates for each form were sequenced. Their primary structures confirmed the splice junctions predicted by genomic sequencing, establishing that these PCR products were not artefactual. Both form I and II RNAs splice out the second intron but splice the third intron from two alternative splice donors (Fig. 2, A and C). The CF2-II mRNA uses the upstream splice donor and consequently encodes a DNA binding domain consisting of fingers 4, 5, and 6. The CF2-I mRNA uses the downstream splice donor, thereby adding another zinc finger motif between fingers 4 and 5. We have named this additional finger 5′, because it resembles finger 5 in residues that are thought to be critical for DNA recognition (*13*) (Fig. 2B). Developmental analysis of the PCR products (Fig. 2D) led to the following observations. (i) Embryos contain almost exclusively CF2-II RNA. This RNA is present before hour 9 of embroynic development, although the protein is not detectable until hour 13 (stage 15). These data suggest that translational control (8) may exist. It is not known whether the early RNA is maternal or zygotic in origin. (ii) No CF2 RNA is detectable in larvae. (iii) In pupae, CF2 RNA reappears; both form I and form II RNAs are present; form III RNA is barely detectable. (iv) Although the PCR technique is not necessarily quantitative, it appears that in the adult somatic tissue (after removal of the gonads), and especially in the male, CF2-I is enhanced relative to CF2-II. No form III RNA is detectable in the adult somatic tissues. (v) The gonads show distinctive CF2 mRNA composition. The ovary and staged egg chambers contain form II RNA. In contrast, the testis contains form III RNA.

The splicing event that generates CF2-III RNA connects the finger 3 coding region to the region that encodes fingers 5 and 6 (Fig. 2, A and C), thus bypassing the exon that encodes fingers 4 and 5′. This splicing event puts the acceptor sequence out of frame and consequently, the CF2-III RNA encodes a polypeptide that also lacks fingers 5 and 6. Instead, it contains a novel 122-residue peptide encoded by mRNA sequences downstream of the splice junction; the last 64 of these amino acid residues are encoded by nucleotides beyond the termination codon of forms I and II (Figs. 1 and 2C). The predicted sequence of this frame-shifted peptide does not resemble any protein in the databases and does not contain recognizable DNA-binding motifs. In any case, its prevalence and specificity in the testis suggests that this RNA form is not simply a by-product of unusual splicing.

Prior to this study, all of our CF2 cDNA clones had been obtained from cDNA libraries derived from embryonic and ovarian tissues, and they correspond to form II RNA (*7*, *14*). Subsequently, libraries derived from adult males and testis were used to isolate form I and III cDNA clones. These clones confirmed the splicing pattern inferred from the PCR products, including the frameshift in form III (*8*).

The differences in the COOH-terminal half of the protein isoforms predicted from these cDNAs have important functional implications. It is not yet known if CF2-III has sequence-specific DNA-binding activity. In vitro binding experiments indicate that protein isoforms I and II recognize different DNA sequences (Fig. 3A), suggesting that in vivo these isoforms regulate distinct genes.

We present evidence (*12*) that, as in zinc finger proteins Zif268 and Krox-20 (*13*), each finger of the COOH-terminal domain of isoform CF-II recognizes a contiguous nucleotide triplet in an antiparallel manner. The optimal trinucleotide targets are GTA for finger 6, TAT for finger 5, and ATA for finger 4, and the consensus binding site for isoform II is GTA.TAT.ATA (dots delimit each nucleotide triplet recognized by one zinc finger). Indeed, CF2-II protein shows a strong affinity for this predicted site in vitro (Fig. 3A). Accordingly, CF2-II protein binds tightly to a nearly perfect consensus site (GTg.TAT.ATA) in the *s15* chorion promoter, whereas CF2-I binds hardly at all (Fig. 3B). Note that in choriogenic follicles, CF2-II RNA is the abundant form (Fig. 2D).

In contrast, the CF2-I protein isoform binds to a distinct consensus sequence, GTA.TAT.TAT.ATA (Fig. 3A). The internal duplication of TAT in this sequence was predictable because the extra finger 5′ has the same critical residues implicated in DNA recognition as finger 5 (Fig. 2B). A candidate target for the CF2-I protein is the CF2 promoter itself. A possible recognition sequence tTA.TAT.TAT.Act, similar to the form I consensus, GTA.TAT.TAT.-ATA, lies approximately −100 relative to the *CF2* transcription initiation site (Fig. 1). DNase I (deoxyribonuclease) protection assays confirmed that protein isoform I binds

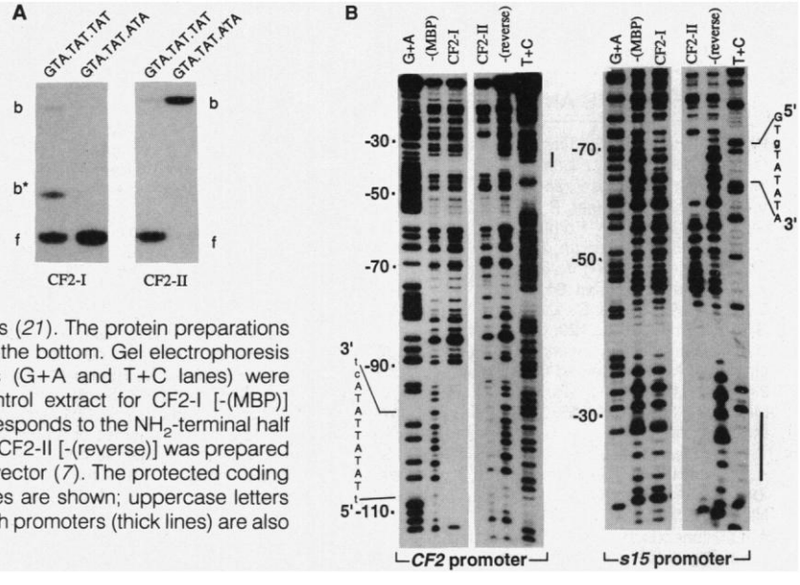


Fig. 3. (**A**) Differential binding specificities of two CF2 protein isoforms (*19*). Partially purified bacterial extracts containing either CF2-I or CF2-II were used to bind synthetic double-stranded oligonucleotides containing the target sequences indicated at the top of each lane (*20*). The free (f) and bound (b) oligonucleotides are indicated. The additional bound species (b*) found in the CF2-I binding reaction, which shows properties similar to b, may correspond to a breakdown product of the CF2-I fusion protein. (**B**) Different promoter specificities of the CF2 isoforms. The 5′ end-labeled DNA fragments of either the CF2 promoter (^{32}P-labeled at −145 of the coding strand) or the s15 promoter (^{32}P-labeled at −13 of the noncoding strand) were used in DNase I protection assays (*21*). The protein preparations used are indicated at the top, and the promoter fragments at the bottom. Gel electrophoresis and the preparation of Maxam-Gilbert sequencing ladders (G+A and T+C lanes) were performed by standard procedures (*18*). The negative control extract for CF2-I [-(MBP)] included partially purified maltose-binding protein, which corresponds to the NH_2-terminal half of the CF2-I fusion protein (*19*); the negative control extract for CF2-II [-(reverse)] was prepared with the use of an inverted CF2-II fragment in the expression vector (*7*). The protected coding strand sequences that match the consensus CF2 binding sites are shown; uppercase letters are those identical to the consensus. The TATA elements of both promoters (thick lines) are also protected by CF2-II.

strongly to this site, but that isoform II binds only poorly (Fig. 3B). The appearance of CF2-I RNA after the larval period coincides with the renewal of CF2 transcription (Fig. 2D). It may be that one function of isoform I is to autoregulate CF2 expression during the pupal and adult stages. In addition, the CF2-II isoform also binds to the TATA box region of both the s15 and the CF2 promoters (Fig. 3B). The significance of this TATA binding is not known.

Recently, a few examples of alternative splicing resulting in altered binding specificities of zinc finger proteins have been documented, notably for the *Drosophila Broad-complex* and *tramtrack* genes, and the human Wilms tumor gene, wtl. However, in these cases, the alternative splicing events result either in the replacement of entire zinc finger domains consisting of two or more fingers (*22*) or in the creation of different lengths of linker sequence between fingers (*23*). Therefore they are not analogous to the type of alternative splicing reported here. The only known case analagous to CF2 splicing may be the splicing of murine proto-oncogene Evi-1, for which the impact of alternative splicing on DNA binding is unknown (*24*).

Our data and that of Gogos *et al.* (*12*) support the concept of the modularity of zinc finger motifs in the formation of DNA binding domains. In the regulation of CF2, this modularity is exploited in vivo during development to generate alternatively spliced RNAs that encode protein isoforms differing in the number of zinc fingers and DNA binding specificities. One of the alternative splicing events is accompanied by a frameshift, which radically decreases the number of zinc finger motifs in the protein. Such in vivo modulations of zinc finger motifs may be a general and efficient mechanism for controlling posttranscriptionally the spectrum of regulatory processes that are served by a single transcriptional regulatory gene.

REFERENCES AND NOTES

1. P. J. Mitchell and R. Tjian, *Science* **245**, 371 (1989); I. Herskowitz, *Nature* **342**, 749 (1989); K. Moses, *Trends Genet.* **7**, 250 (1991).
2. V. L. Stroeher, C. Gaiser, R. L. Garber, *Mol. Cell. Biol.* **8**, 4143 (1988); K. Kornfeld *et al.*, *Genes Dev.* **3**, 243 (1989); C. Roman, L. Cohn, K. Calame, *Science* **254**, 94 (1991); N. S. Foulkes, E. Borrelli, P. Sassone-Corsi, *Cell* **64**, 739 (1991); P. R. DiBello, D. A. Withers, C. A. Bayer, J. W. Fristrom, G. M. Guild, *Genetics* **129**, 385 (1991).
3. D. Picard and K. R. Yamamoto, *EMBO J.* **6**, 3333 (1987); P. A. Baeuerle and D. Baltimore, *Science* **242**, 540 (1988); C. A. Rushlow, K. Han, J. L. Manley, M. Levine, *Cell* **59**, 1165 (1989); R. Steward, *ibid.*, p. 1179.
4. V. R. Baichwal and R. Tjian, *Cell* **63**, 815 (1990); S. A. Johnston, J. M. Salmeron, Jr., S. S. Dincher, *ibid.* **53**, 143 (1987); Y. Nakabeppu and D. Nathans, *ibid.* **64**, 751 (1991).
5. A. G. Hinnebusch and P. P. Mueller, in *Translational Regulation in Gene Expression*, J. Ilan, Ed. (Plenum, New York, 1987), pp. 397–412; W. Driever and C. Nüsslein-Volhard, *Cell* **54**, 83 (1988); B. Wightman, T. R. Bürglin, J. Gatto, P. Arasu, G. Ruvkun, *Genes Dev.* **5**, 1813 (1991).
6. P. Raychandhuri, S. Bagchi, J. R. Nevins, *Genes Dev.* **2**, 1686 (1989); J. R. Cherry, T. R. Johnson, C. Dollard, J. R. Shuster, C. L. Denis, *Cell* **56**, 409 (1989); L. M. Mylin, M. Johnson, J. E. Hopper, *Mol. Cell. Biol.* **10**, 4623 (1990); S. P. Jackson, J. J. MacDonald, S. Lees-Miller, R. Tjian, *Cell* **63**, 155 (1990); W. J. Boyle *et al.*, *ibid.* **64**, 573 (1991).
7. M. J. Shea, D. L. King, M. J. Conboy, B. D. Mariani, F. C. Kafatos, *Genes Dev.* **4**, 1128 (1990).
8. T. Hsu, J. A. Gogos, D. L. King, F. C. Kafatos, unpublished results.
9. J. Szidonya and G. Reuter, *Genet. Res. Cambridge* **51**, 197 (1988).
10. P. Tolias, S. A. Mitsialis, J. A. Gogos, F. C. Kafatos, unpublished results.
11. N. Tanese, B. F. Pugh, R. Tjian, *Genes Dev.* **5**, 2212 (1991); P. J. Mitchell and R. Tjian in (*1*).
12. J. A. Gogos, T. Hsu, J. Bolton, F. C. Kafatos, *Science* **257**, 1951 (1992).
13. N. P. Pavlevitch and C. O. Pabo, *ibid.* **252**, 809 (1991); J. Nardelli, T. J. Gibson, C. Vesque, P. Charnay, *Nature* **349**, 175 (1991).
14. N. H. Brown and F. C. Kafatos, *J. Mol. Biol.* **203**, 425 (1988); K. Zinn, L. McAllister, C. Goodman, *Cell* **53**, 577 (1988); W. Steinhauer, R. Walsh, L. Kalfayan, *Mol. Cell. Biol.* **12**, 5726 (1989).
15. J. C. Martínez-Cruzado, C. Swimmer, M. G. Fenerjian, F. C. Kafatos, *Genetics* **119**, 663 (1988).
16. In addition to the CF2 sequence indicated in Fig. 1, PCR primer No. 24 carries 5′-CCGGATCC-3′ (Bam HI site) at the 5′ side, and primer No. 53 contains 5′-CGCTCGAG-3′ (Xho I site) at the 5′ side. Another pair of primers, (5′-GTATCGACAACAGAGTCGGTCGC-3′ and 5′-TTGGTGAGCGGACCGACAGCTGC-3′) specific for the ribosomal protein 49 gene [*rp49*; P. O'Connell and M. Rosbash, *Nucleic Acids Res.* **12**, 5495 (1984)] were included in the PCR reaction as an internal control; a mixture of these primers was also used for *rp49* detection. Probes for specific CF2 isoform detection were: (numbers in parentheses indicate the positions of the 5′ and 3′ ends as in Fig. 1): No. 32 (noncoding strand), 5′-(1449)GGTC-TTTGGACAATCCGG(1432)-3′; No. 38 (coding strand), 5′-(1519)GAACGCCCCTACACGTGC-(1536)-3′; No. 41 (coding strand), 5′-(1623)CTATT-GTGGCAGGGCGTTCAC(1643)-3′; No. 70 (coding strand), 5′-(1496)GCGAAGCAGAGAGAGA-AGCC(1694)-3′. For determination of isoform developmental specificity by PCR, total RNA samples (1 μg) from various developmental stages and tissues and 50 ng of a CF2 cDNA clone in Bluescript plasmid (from Stratagene) were used as a positive control. Each RNA sample was heat-denatured at 90°C and reverse transcribed in a solution (10 μl) containing 50 mM tris-HCl (pH 8.3), 40 mM KCl, 6 mM $MgCl_2$, 1 mM DTT (dithiothreitol), BSA (bovine serum albumin) (0.1 mg/ml), random primers (10 μM; Boehringer-Mannheim), the four dNTPs at 1 mM each, and 100 units of MMLV (Moloney murine leukemia virus) reverse transcriptase (Gibco/BRL). After incubation at room temperature for 5 min and then at 37°C for 1 hour, the reaction was terminated by heating at 90°C for 5 min. The following were added to each sample: 40 μl of 10 mM tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin (w/v), each of the CF2 primers at 0.5 μM, each of the *rp49* primers at 0.025 μM, and 2.5 units of Taq DNA polymerase (Perkin-Elmer Cetus). The conditions for the PCR reactions were 30 cycles at 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min plus 2 s for each successive cycle. The final PCR mixtures (5 μl) were fractionated through a 3% agarose gel (3:1 mixture of NuSieve GTG low melting agarose of American Biochemicals and regular agarose from Gibco/BRL) in 40 mM tris-acetate (pH 8), 1 mM EDTA. The gel was rinsed in 0.4 M NaOH, 1.5 M NaCl and nucleic acids were transferred to nylon membrane (Bio-Rad) by diffusion in 0.4 M NaOH. Hybridization with 5′ end-labeled detection probes was performed in 2X SSC [(0.3 M NaCl, 0.03 M sodium citrate (pH 7)] containing 1% BSA, 1% SDS, and ^{32}P-labeled probes at T_m − 10°C (T_m = melting temperature) at 2 ng/ml. The filter was washed in 0.5× SSC containing 0.1% SDS twice for 10 min each at room temperature and twice for 10 min each wash at the hybridization temperature. Autoradiography was performed without an intensifying screen. The sizes of the observed fragments were consistent with expectations: CF2-I, 668 bp; CF2-II, 584 bp; CF2-III, 489 bp; and *rp49*, 259 bp.
17. S. A. Mitsialis and F. C. Kafatos, *Nature* **317**, 453 (1985).
18. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
19. The CF2-I DNA fragment generated from a PCR reaction was cloned as a BAM HI–Pst I fragment into the expression plasmid vector pMal-cRI (New England Biolabs), producing an in-frame fusion with the maltose-binding protein (MBP) product of the upstream *malE* gene. This fusion protein was isolated and column-purified according to New England Biolabs and was dialyzed against 25 mM Hepes (pH 7.7), 0.1 M KCl, 1 mM DTT, 25 μM $ZnSo_4$, 1 mM EDTA, 20% glycerol, and 0.1% Triton X-100. MBP to be used as negative control for CF2-I in DNase I protection assays (Fig. 3B) was prepared similarly from TB1 cells (New England Biolabs) carrying pMal-cRI vector. Preparation of CF2-II extract and its negative counterpart have been described (*7*).
20. To prepare double-stranded binding targets, a 5′-oligonucleotide was end-labeled with ^{32}P 5′-CTTCCCGGGAATTC-3′, was annealed to either 5′-CCGCTCGAGGATCCGTATATTATGAATTCCCGGGAAG-3′ or 5′-CCGCTCGAGGATCCGTATATATAGAATTCCCGGGAAG-3′, the central portions of which (underlined) correspond to the predicted binding sites for CF2-I and CF2-II, respectively; these proteins do not bind the flanking sequences (*12*). The annealed short oligonucleotide was extended with the use of Klenow enzyme (New England Biolabs), and the full length double-stranded products were purified by gel electrophoresis. Binding was conducted with 2.5 μl of protein extract, 20,000 cpm of labeled targets, poly(dI-dC)·poly(dI-dC) (0.2 mg/ml) in 5-μl reactions containing 15 mM Hepes (pH 7.7), 75 mM KCl, 0.5 mM DTT, 0.5 mM EDTA, 10% glycerol, 0.05% Triton X-100, and 0.5 mM $ZnSO_4$. The mixture was incubated at room temperature for 40 min and fractionated through a 6% polyacrylamide gel containing 5% glycerol in 0.25× TBE buffer [0.0225 M tris-borate (pH 8.3), 0.5 mM EDTA] at 4°C. The gel was dried and autoradiography was performed.
21. The DNase I protection assays were carried out essentially as described (*7*) but in 15 μl of binding mixture containing 10,000 cpm of ^{32}P 5′ end-labeled DNA fragment, 7.5 μl of protein extract, 15 mM Hepes (pH 7.7), 75 mM KCl, 0.5 mM DTT, 0.5 mM EDTA, 10% glycerol, poly(dI-dC)·poly(dI-dC) (0.2 mg/ml), 0.05% Triton X-100, 1.8% polyvinyl alcohol, and 0.5 mM $ZnSO_4$.
22. P. R. DiBello, D. A. Withers, C. A. Bayer, J. W. Fristoom, G. M. Guild, *Genetics* **129**, 385 (1991); D. Read and J. L. Manley, *EMBO J.* **11**, 1035 (1992).
23. W. A. Bickmore *et al.*, *Science* **257**, 235 (1992).
24. D. Bordereaux, S. Fichelson, P. Tambourin, S. Gisselbrecht, *Oncogene* **5**, 925 (1990).

25. Supported by an NSF grant (F.C.K.), an NIH postdoctoral fellowship (T.H.), and a Gerondelis Foundation predoctoral fellowship (J.A.G.). We thank M. J. Shea for advice, D. L. King for unpublished immunostaining information, A. Khoury for oligonucleotide synthesis, T. Maniatis for critical reading of the manuscript, W. Gelbart for a gift of genomic library, T. Carr for an adult male cDNA library, T. Hazelrigg for a testis cDNA library, S. Pedreira for secretarial assistance, and B. Klumpar for photography. The CF2 sequence has been submitted to Genbank under accession number M97196.

26 March 1992; accepted 28 July 1992