MOLECULAR AND CELLULAR BIOLOGY, Feb. 1990, p. 681–688
0270-7306/90/020681-08$02.00/0

Vol. 10, No. 2

# Mouse Zfx Protein Is Similar to Zfy-2: Each Contains an Acidic Activating Domain and 13 Zinc Fingers

GRAEME MARDON,[1,2] SHIUH-WEN LUOH,[1,2] ELIZABETH M. SIMPSON,[1,2] GRACE GILL,[3]†
LAURA G. BROWN,[1,2] AND DAVID C. PAGE[1,2]*

*Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142[1]; Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139[2]; and Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138[3]*

Received 29 September 1989/Accepted 13 November 1989

**The *Zfy* gene is located on the Y chromosome of placental mammals and encodes a zinc finger protein which may serve as the primary sex-determining signal. A related gene, *Zfx*, is similarly conserved on the X chromosome. Unlike that in most mammals, the mouse genome contains four homologous zinc finger loci: *Zfy-1*, *Zfy-2*, *Zfx*, and *Zfa* (on an autosome). We report that, in contrast to the mouse *Zfy* genes, *Zfx* is widely transcribed in embryos, newborns, and adults, both male and female. Moreover, *Zfx* transcripts contain long 3′ untranslated sequences which are phylogenetically conserved. *Zfa* is a processed gene derived from *Zfx*. An analysis of cDNA clones demonstrated that *Zfx* encodes a 799-amino-acid protein that is 70% identical to the mouse Zfy-1 and Zfy-2 proteins. Zfx, Zfy-1, and Zfy-2 contain highly acidic amino-terminal domains and carboxy-terminal regions containing 13 zinc fingers. When fused to the DNA-binding domain of GAL4, the acidic domains of Zfx and Zfy-2 activated transcription in yeast cells.**

In humans and mice, the sex of a developing embryo is determined by one or more genes on the Y chromosome (8, 16, 40). The *ZFY* gene, located in the sex-determining region of the human Y chromosome, may serve as the primary sex-determining signal (33). *ZFY* encodes a protein with 13 Cys-Cys–His-His zinc fingers, a nucleic acid binding motif first described in *Xenopus* transcription factor IIIA (4, 29). Homologs of *ZFY* are found on the Y chromosomes of all placental mammals examined. A related gene, *ZFX*, is also conserved among mammals and has been mapped to the X chromosome in both humans (33) and mice (30, 31; D. C. Page, C. M. Disteche, E. M. Simpson, A. de la Chapelle, M. Anderson, T. Alitalo, L. G. Brown, and P. Green, Genomics, in press). The zinc finger domains of human ZFY and ZFX are 97% identical, suggesting that the two proteins may recognize the same DNA or RNA sequence (39). *ZFY* and *ZFX* may both function in sex determination (33).

Two zinc finger genes, *Zfy-1* and *Zfy-2*, are located in the sex-determining region of the mouse Y chromosome (33). This is the result of an intrachromosomal duplication that occurred during mouse evolution (24). The two mouse *Zfy* genes may encode functionally redundant products; both genes are transcribed in the adult testis, and the predicted proteins are 95% identical. Both Zfy proteins are composed of two domains: an amino-terminal region that is highly acidic and a carboxy-terminal domain that contains 13 zinc fingers (1, 25). The combination of an acidic activating region and a DNA-binding motif is characteristic of several eucaryotic transcription factors, e.g., yeast GAL4 and GCN4 and the mammalian glucocorticoid receptor (11–13, 23). Thus, the putative sex-determining genes *Zfy-1* and *Zfy-2* appear to encode sequence-specific activators of transcription.

In this report, an analysis of cDNA clones demonstrates that *Zfx* encodes a protein that is 70% identical to mouse Zfy-1 and Zfy-2. Like the Zfy proteins, Zfx is predicted to have a two-domain structure, with an acidic amino-terminal region and 13 zinc fingers in the carboxy-terminal half. Zfx and Zfy proteins may both function in the mouse as sequence-specific transcriptional activators, as substantiated here by the ability of their acidic domains to activate transcription in yeast cells.

## MATERIALS AND METHODS

**Northern (RNA) and Southern blotting.** Poly(A)$^+$ RNAs were prepared from FVB/N mouse tissues or whole embryos as previously described (25). DNA inserts of plasmids were purified, radiolabeled by random-primer synthesis, and hybridized to RNA or DNA blots for 16 h at 37°C (low stringency), 42°C (medium stringency), or 47°C (high stringency) as previously described (25, 33). The nylon membranes were washed three times for 30 min each at 50°C (low stringency), 60°C (medium stringency), or 65°C (high stringency) in 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate and exposed at −80°C for 1 to 7 days to X-ray film backed by an intensifying screen.

**Cloning and nucleotide sequence analysis.** cDNA libraries were prepared from FVB/N newborn liver poly(A)$^+$ RNA as previously described (25). Following ligation to *Eco*RI adaptors, the double-stranded cDNA was size fractionated by agarose gel electrophoresis. After electroelution, the 5- to 10-kilobase (kb) fraction was ligated to λZAP (Stratagene). An unamplified library of $5 \times 10^5$ recombinants was screened by using the human genomic insert of plasmid pDP1007 as a probe. Six cDNA clones were identified, each containing an insert of 5.7 to 6.8 kb. Comparative restriction mapping suggested that all six clones derive from the same locus. The cDNA inserts of two clones were transferred into the *Eco*RI site of Bluescript (Stratagene), generating plasmids pDP1115 and pDP1119.

Plasmids pDP1193 and pDP1194 contain 1.9- and 2.1-kb *Eco*RI fragments of *Zfx* and *Zfa*, respectively (see Fig. 2B, probe 1), cloned into the *Eco*RI site of Bluescript. Both inserts derive from FVB/N genomic mouse DNA.

* Corresponding author.
† Present address: Department of Biochemistry, University of California, Berkeley, CA 94720.

Nucleotide sequencing was carried out by dideoxy-chain termination (37), using synthetic oligonucleotide primers.

**Construction of GAL4 fusion genes.** The acidic amino-terminal domains of Zfx and Zfy-2 were fused, in frame, to the DNA-binding domain of GAL4 (amino acids 1 to 148). First, sequences between the *Hin*dIII and *Bam*HI sites of Bluescript KS (−) were replaced by an adaptor containing an *Nco*I site:

5′-AGCTTCGCCATGGAG
AGCGGTACCTCCTAG-5′

The *Nco*I-*Pst*I (with *Xba*I linker) fragment of *Zfy-2* cDNA pDP1122 (25) or *Nco*I-*Pvu*II (with *Xba*I linker) fragment of *Zfx* cDNA pDP1115 was cloned into the *Nco*I and *Xba*I sites of the modified Bluescript vector. *Cla*I-*Xba*I fragments from the resulting clones were inserted into the *Cla*I and *Xba*I sites of pGG25ΔX, a single-copy yeast plasmid which expresses GAL4(1-148) from an ADH1 promoter. The resulting GAL4-Zfx and GAL4–Zfy-2 fusion plasmids are, respectively, pDP1199 and pDP1200. The net charge on these portions of Zfx (residues 1 to 330) and Zfy-2 (residues 1 to 314) was calculated as the sum of acidic residues (glutamic and aspartic acid) minus the sum of basic residues (arginine and lysine).

**Yeast transformation and assay of β-galactosidase activity.** GGY1:SV15 is a ΔGAL4 yeast strain bearing an integrated GAL1:*lacZ* fusion with a single near-consensus GAL4-binding site upstream (9). pDP1199 and pDP1200, as well as plasmids expressing intact GAL4 (pGG22) or the GAL4 DNA-binding domain alone (pGG25ΔX), were introduced into GGY1:SV15 after treatment with lithium acetate (15). Yeast cells were grown in selective media lacking histidine but containing 2% galactose, 3% glycerol, and 2% lactic acid, pH 6. β-Galactosidase assays (9) were performed in triplicate; the standard error was less than 15%. Immunoprecipitation of the various GAL4 derivatives with an anti-GAL4(1-147) serum revealed no significant differences in the amounts of the proteins (data not shown).

## RESULTS

**Human *ZFY* detects transcripts in both female and male mice.** The genomic insert of plasmid pDP1007 encodes the zinc finger domain of human ZFY (33). When hybridized to mouse genomic DNAs at low stringency, pDP1007 detects related sequences on the Y chromosome (*Zfy-1* and *Zfy-2*), the X chromosome (*Zfx*), and autosome 10 (*Zfa*) (30–33; Page et al., in press). At moderate stringency, pDP1007 hybridizes only to *Zfx* and *Zfa*. When hybridized to Northern blots at moderate stringency, pDP1007 detects transcripts of 6 and 7.5 kb in mouse embryo RNAs (Fig. 1). These transcripts, which are present in both males and females, must derive from either *Zfx* or *Zfa*. Similar results were obtained with newborn male liver RNA (results not shown).

In order to further characterize these transcripts, a cDNA library prepared from newborn male mouse liver RNA was screened with human *ZFY* probe pDP1007 under the same conditions used for Northern analysis. One of the cDNA clones isolated, pDP1115, contains an insert of 6.8 kb (Fig. 2A).

**A cDNA clone derived from mouse *Zfx*.** Using Southern analysis, we determined that cDNA clone pDP1115 is derived from mouse *Zfx*. The 5′-most 2.7 kb of the cDNA (Fig. 2A, probe 1) detects 1.9- and 2.1-kb *Eco*RI fragments in genomic mouse DNA (Fig. 2B). The 1.9-kb *Eco*RI fragment hybridizes with twice the intensity in females as in males,

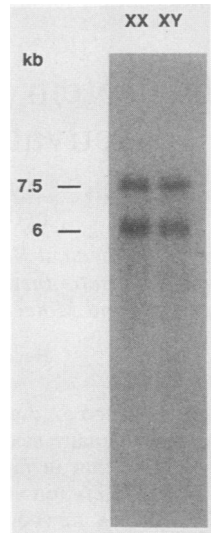


FIG. 1. Human *ZFY* detects transcripts in male and female mouse embryos. Human *ZFY* probe pDP1007 was hybridized at moderate stringency to poly(A)$^+$ RNAs prepared from female (XX) and male (XY) whole embryos at 14 days postcoitum. The sizes of transcripts detected are indicated.

indicating an X chromosomal location, while the 2.1-kb *Eco*RI fragment hybridizes with equal intensity in both sexes, suggesting an autosomal origin. The same 1.9- and 2.1-kb *Eco*RI fragments are detected with human *ZFY* probe pDP1007 (results not shown). The 1.9-kb *Eco*RI fragment is derived from mouse *Zfx*, while the 2.1-kb fragment is from mouse *Zfa* (30; Page et al., in press) (Fig. 2B). This result suggests that cDNA clone pDP1115 is derived from either *Zfx* or *Zfa*. The additional sequences detected by probe 1 probably correspond to other exons represented in the mouse cDNA clone but absent in human *ZFY* genomic clone pDP1007. A portion of the 3′ untranslated (UTR) sequence of the cDNA (Fig. 2A, probe 2) also detects two *Eco*RI fragments in mouse genomic DNA: a 5.3-kb fragment from the X chromosome and an 8.2-kb fragment that is autosomal. The 3′-most portion of the cDNA (Fig. 2A, probe 3) detects only the 5.3-kb *Eco*RI fragment, demonstrating that cDNA clone pDP1115 is derived from *Zfx* (Fig. 2B).

Using Northern analysis, we determined that *Zfx* cDNA clone pDP1115 corresponds to the 7.5-kb transcript detected in whole embryo RNA (Fig. 1). Probes 1 and 2 detect the same 6- and 7.5-kb transcripts in mouse poly(A)$^+$ RNA as those seen with human *ZFY* probe pDP1007 (Fig. 1 and 2C). Probe 3, however, detects only the 7.5-kb RNA (Fig. 2C), demonstrating that cDNA clone pDP1115 corresponds to that particular transcript. As probe 3 is an X chromosome-specific probe (Fig. 2B), we conclude that the 7.5-kb transcript derives from mouse *Zfx*. Southern and Northern hybridizations with probes prepared from other portions of *Zfx* cDNA clone pDP1115 suggest that the 6-kb transcript is also derived from mouse *Zfx* (results not shown).

**Mouse *Zfx* is widely transcribed.** Transcripts of 6 and 7.5 kb are detected by *Zfx* cDNA probe 1 in all mouse tissues examined, including the following, both male and female: whole embryos at 12, 14, 16, and 18 days postcoitum; and newborn and adult brain, gonad, heart, kidney, liver, lung, and spleen (Fig. 3). In general, the two transcripts are more abundant in whole embryos and adult tissues than in newborn tissues. However, relatively high levels of *Zfx* transcripts are present in newborn testis.

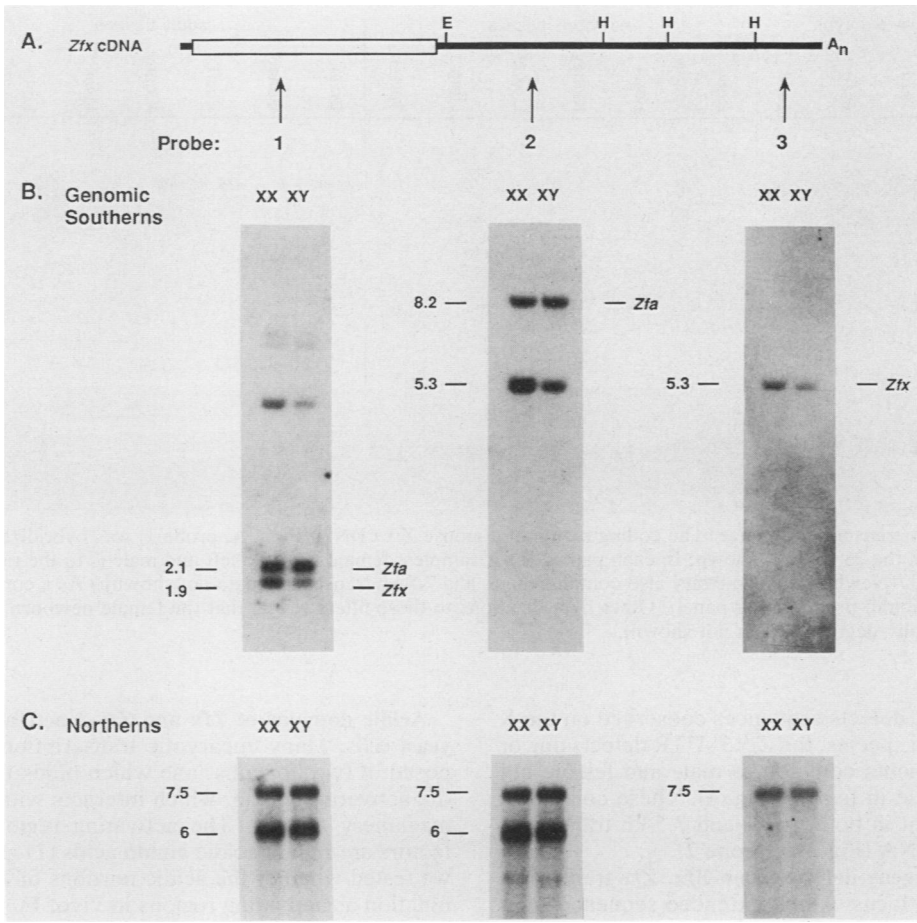


FIG. 2. A cDNA clone derived from mouse *Zfx*. (A) *Zfx* cDNA clone pDP1115, 6.8 kb in length, is shown schematically. The single long open reading frame is depicted as an open box. A polyadenosine run ($A_n$) is found at the 3′ end of the cDNA. *Eco*RI (E) and *Hin*dIII (H) restriction sites are indicated. (B) Restriction fragments 1 (2.7 kb), 2 (1.8 kb), and 3 (0.6 kb) were gel purified from an *Eco*RI-*Hin*dIII digest of pDP1115 and hybridized at high stringency to *Eco*RI-digested mouse genomic DNAs (BALB/c). Exposures of 2 to 5 days are shown. The sizes of fragments detected are indicated in kilobases. *Zfa* and *Zfx* refer to the zinc finger homologs on mouse chromosomes 10 and X, respectively. (C) Probes 1, 2, and 3 were hybridized at high stringency to poly(A)$^+$ RNAs prepared from whole embryos at 14 days postcoitum. Filters were exposed for 1 to 4 days. Transcripts of 6 and 7.5 kb are indicated.

**Mouse *Zfx* and *Zfy* genes encode similar proteins.** Nucleotide sequence analysis of mouse *Zfx* cDNA clone pDP1115, 6.8 kb in length, revealed a single, long open reading frame (Fig. 4). The first AUG in this frame (Fig. 4, position 1) occurs in a sequence context that is favorable for initiation of translation (17). Beginning at this putative initiation codon, the open reading frame encodes a protein 799 amino acids in length, with a predicted molecular weight of 90,000. The AUG codon is preceded by a 5′ leader of 145 nucleotides containing stop codons in all three reading frames. A putative 3′ untranslated sequence of 4,163 bases is followed by 65 adenosines. A canonical AATAAA polyadenylation signal (7) occurs 26 nucleotides 5′ of the poly(A) track. Several other potential polyadenylation signals are also present within the 3′ UTR.

Comparative analysis of a second cDNA clone suggests that *Zfx* transcripts undergo alternate splicing and polyadenylation, which may account for the difference in lengths of the 6- and 7.5-kb *Zfx* transcripts (Fig. 2C). The nucleotide sequence of cDNA clone pDP1119, 5.7 kb in length, is identical to that of *Zfx* cDNA pDP1115 from nucleotides −31 to +5458 (Fig. 4). However, sequences 5′ of position −31 in the two clones appear unrelated, perhaps the result of alternate splicing. As a result of alternate polyadenylation, the 3′ UTRs of these two cDNAs differ in length by 1.1 kb. The significance of alternate processing of *Zfx* transcripts remains to be determined.

The predicted Zfx amino acid sequence is very similar to both mouse Zfy-1 and Zfy-2 (1, 25) over the entire length of the protein (Fig. 5). The Zfx and Zfy proteins have nearly identical amino and carboxy termini. Like the Zfy proteins, Zfx has two large domains: a highly acidic amino-terminal portion, with nearly 25% aspartic or glutamic acid residues, and a carboxy-terminal region encompassing 13 zinc fingers. The acidic domains of Zfx and Zfy are 60% identical, with seven insertions or deletions of 1 to 11 amino acids. In contrast, the zinc finger domains are 80% identical, with no insertions or deletions. In Zfx, as in the Zfy proteins, a short basic sequence is located between the acidic and zinc finger domains (Fig. 5). Similar basic regions serve as nuclear localization signals in several proteins (5, 6, 19, 22).

**The 3′ UTR of *Zfx* is conserved in mammals.** Nucleotide sequence analysis of *Zfx* cDNA clone pDP1115 revealed no open reading frame greater than 125 codons in the more than 4 kb following the zinc finger coding region. Nonetheless, when hybridized to genomic DNAs from a variety of mam-

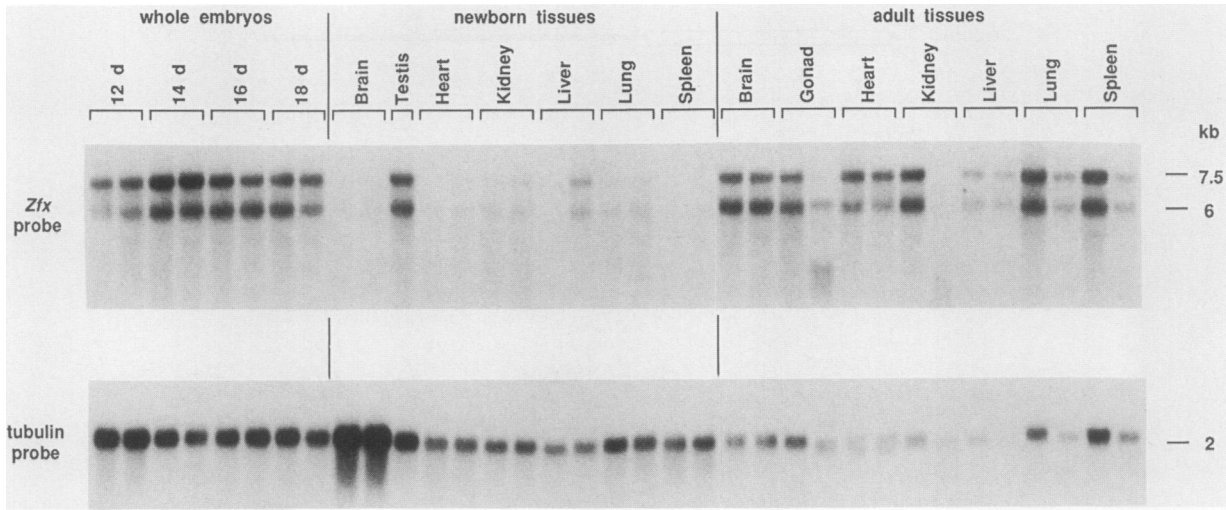


FIG. 3. *Zfx* is widely transcribed in mice. The coding region of a mouse *Zfx* cDNA (Fig. 2A, probe 1) was hybridized at high stringency to poly(A)$^+$ RNAs from the 35 sources shown. In each pair of RNA samples, female is to the left and male is to the right. Transcripts of 6 and 7.5 kb are indicated. (Newborn mouse ovary also contains the 6- and 7.5-kb transcripts [data not shown].) As a control, the filters were hybridized with an α-tubulin probe (lower panel). Other hybridizations to these filters reveal that the female newborn liver and male adult kidney RNAs are partially degraded (data not shown).

mals, the *Zfx* 3′ UTR detects sequences conserved on the X chromosome. In each species, the *Zfx* 3′ UTR detects one or two restriction fragments common to male and female but about twice as intense in female (Fig. 6). These conserved sequences are present in both the 6- and 7.5-kb transcripts detected in mouse RNA (Fig. 2C, probe 2).

***Zfa* is a processed gene derived from *Zfx*.** *Zfx* transcripts and the *Zfa* genomic locus exhibit extended sequence similarity, with sequences derived from the *Zfx* 5′ UTR, coding region, and 3′ UTR cross-hybridizing to *Zfa* at high stringency (Fig. 2B) (unpublished data). Nonetheless, in comparison with its homologs in mice and humans, the *Zfa* locus is condensed; *Zfx* cDNAs hybridize to sequences spanning at least 4.0 kb but no more than 11.5 kb of *Zfa* genomic DNA. In contrast, the human *ZFY* and *ZFX* genes span 50 and 70 kb, respectively, and the mouse *Zfy-1*, *Zfy-2*, and *Zfx* loci are of comparable size (33, 39; E. M. Simpson and S.-W. Luoh, unpublished data; A. Schneider-Gädicke, P. Beer-Romero, L. G. Brown, G. Mardon, S.-W. Luoh, and D. C. Page, Nature [London], in press).

Given the condensed size of the *Zfa* locus—and the fact that an autosomal homolog of *Zfx* is not found in most placental mammals—we hypothesized that *Zfa* might be a processed gene derived from a *Zfx* transcript. Indeed, by nucleotide sequence comparison, we found that the *Zfa* genomic locus lacks at least one intron present in *Zfx* (Fig. 7). In *Zfx*, the alignment of genomic and cDNA sequences reveals an intron immediately 5′ to the zinc finger exon. (Human *ZFX* and *ZFY* transcripts are spliced at precisely the same site [39].) On both sides of this splice, the *Zfa* genomic sequence is nearly identical to that of the *Zfx* cDNA; the intron has been precisely excised in *Zfa*. On the basis of the small size of *Zfa*, the lack of at least one intron, and sequence similarity to *Zfx* cDNAs, we conclude that the *Zfa* locus derives from a processed *Zfx* transcript. The absence of such an autosomal locus in most placental mammals suggests that this retroposition occurred during rodent evolution. Since some processed genes encode functional proteins (e.g., human PGK-2 [26]), we cannot rule out a functional role for *Zfa*.

**Acidic domains of Zfx and Zfy-2 activate transcription in yeast cells.** Many eucaryotic transcription factors are composed of two domains, one which binds DNA and another, the activating region, which interacts with the transcription machinery (3, 35). The activating regions of many such factors are rich in acidic amino acids (11–13, 23). Therefore, we tested whether the acidic portions of Zfx or Zfy-2 could function as activating regions in vivo. Fusion proteins composed of the acidic region of either Zfx (residues 1 to 330; net charge, −77) or Zfy-2 (residues 1 to 314; net charge, −68) attached to the DNA-binding region of yeast GAL4 (residues 1 to 148) were assayed for their ability to stimulate transcription in yeast cells. The β-galactosidase activity produced from a GAL1:lacZ fusion gene bearing a single GAL4-binding site upstream is a measure of transcriptional activation by a given GAL4 derivative. Both the GAL4-Zfx and GAL4–Zfy-2 fusion proteins activate transcription, whereas the DNA-binding region of GAL4 alone does not (23) (Fig. 8). The GAL4–Zfy-2 fusion activates transcription nearly as well as wild-type GAL4, while the GAL4-Zfx fusion is about 20-fold less effective.

## DISCUSSION

**Mouse *Zfx* and *Zfy* genes encode similar proteins.** The existence of two *Zfy* genes on the mouse Y chromosome is the result of an intrachromosomal duplication during rodent evolution (24). The Zfy-1 and Zfy-2 proteins (1, 25) are 95% identical (Fig. 5) and may be functionally redundant. The mouse Zfx protein is 70% identical to the mouse Zfy proteins and exhibits the same two-domain structure: the amino-terminal half is acidic, while the carboxy-terminal half contains 13 Cys-Cys–His-His zinc fingers (Fig. 5).

Acidic and DNA-binding domains are found in combination in many eucaryotic transcription activators (11–13, 23). By analogy, both Zfx and Zfy will probably function as sequence-specific activators of transcription. When fused to the GAL4 DNA-binding domain, the acidic regions of both Zfx and Zfy-2 indeed activate transcription in yeast cells

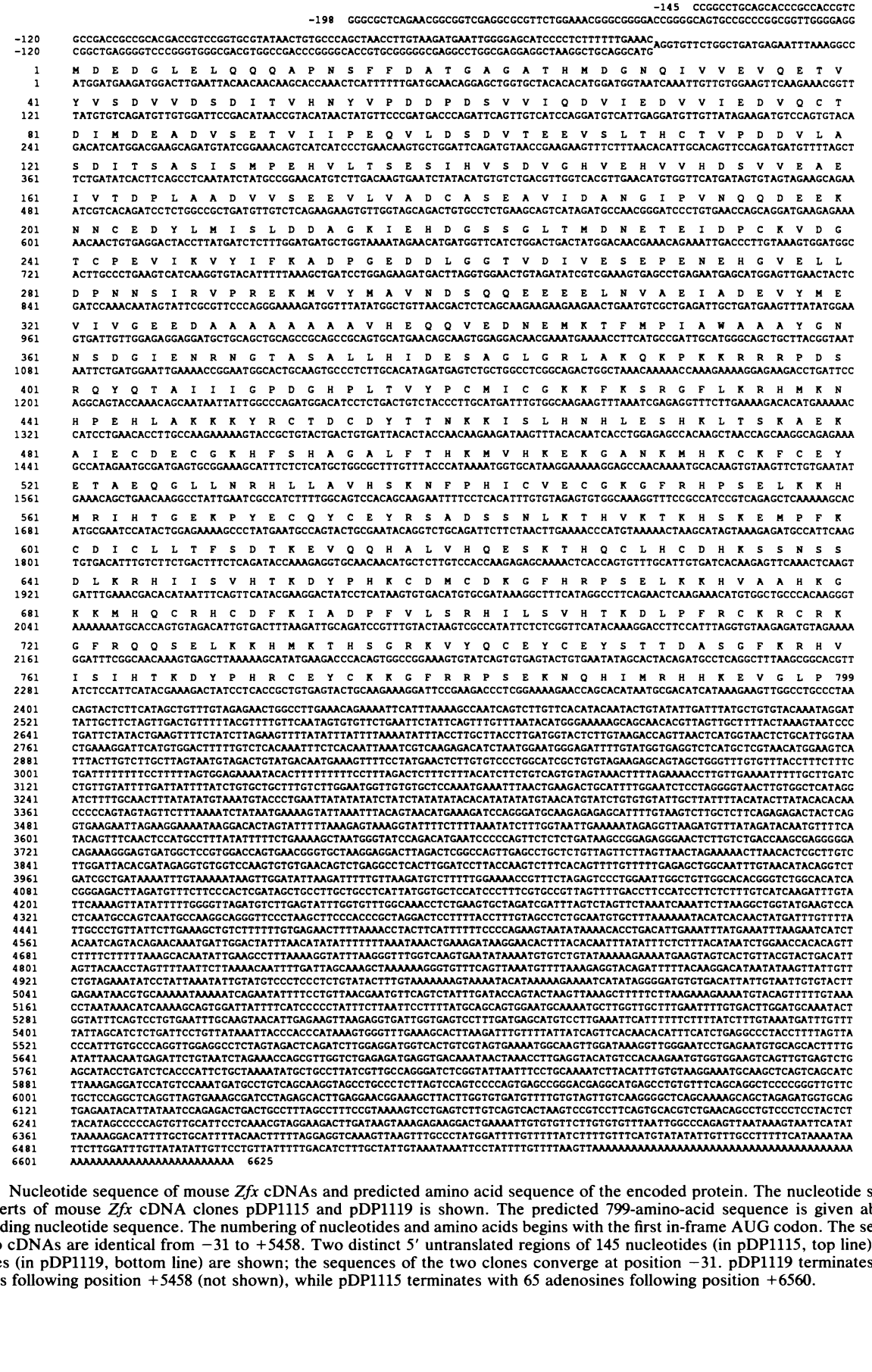

FIG. 4. Nucleotide sequence of mouse *Zfx* cDNAs and predicted amino acid sequence of the encoded protein. The nucleotide sequence of the inserts of mouse *Zfx* cDNA clones pDP1115 and pDP1119 is shown. The predicted 799-amino-acid sequence is given above the corresponding nucleotide sequence. The numbering of nucleotides and amino acids begins with the first in-frame AUG codon. The sequences of the two cDNAs are identical from −31 to +5458. Two distinct 5′ untranslated regions of 145 nucleotides (in pDP1115, top line) and 198 nucleotides (in pDP1119, bottom line) are shown; the sequences of the two clones converge at position −31. pDP1119 terminates with 40 adenosines following position +5458 (not shown), while pDP1115 terminates with 65 adenosines following position +6560.

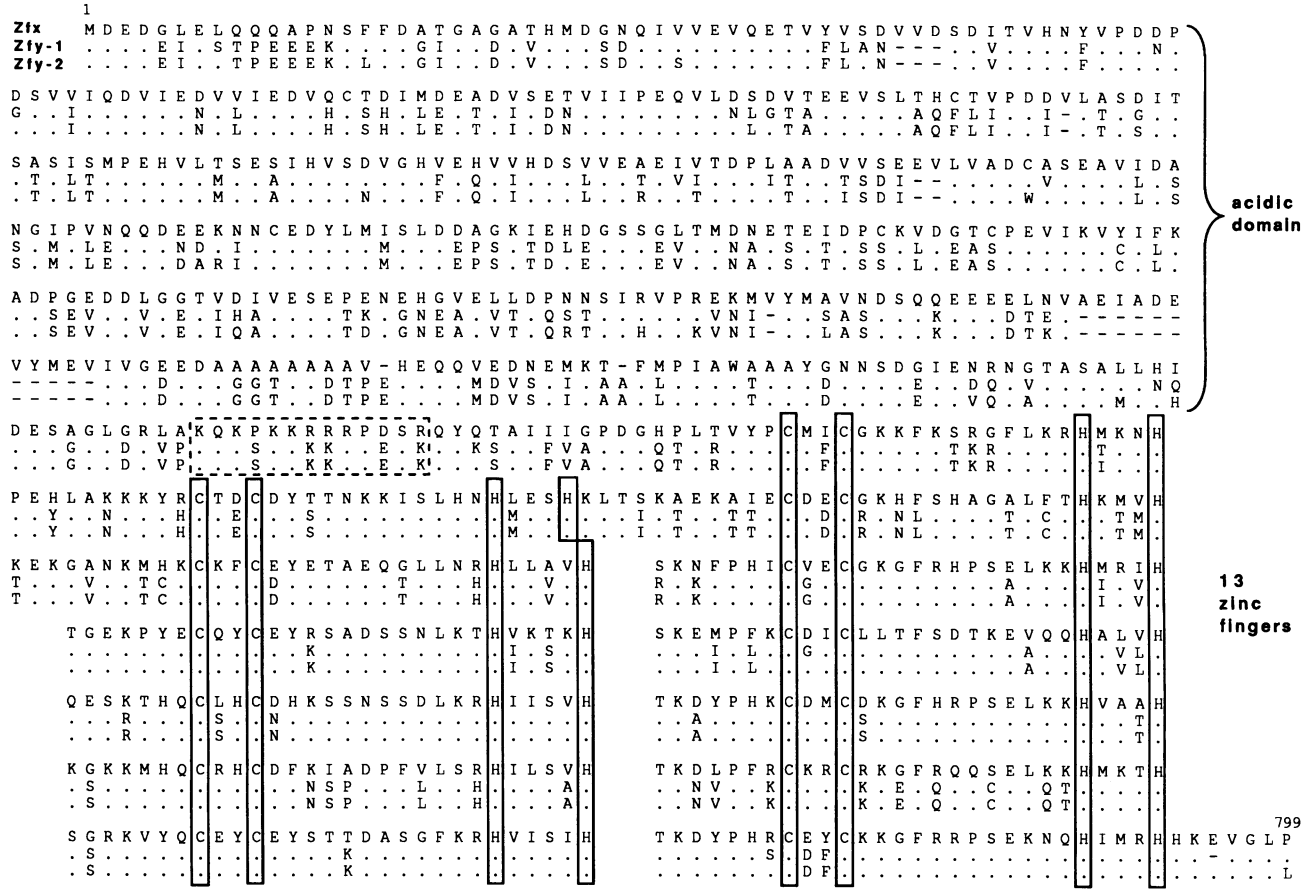


FIG. 5. Predicted amino acid sequence of mouse Zfx is 70% identical to mouse Zfy-1 and Zfy-2. The predicted amino acid sequences of mouse Zfx (799 residues), mouse Zfy-1 (782 residues [1]), and mouse Zfy-2 (783 residues [25]) are compared. Dots represent identity to Zfx, while dashes indicate gaps in one sequence compared with another. A short basic sequence (dashed box) is located between large acidic and zinc finger domains. In the zinc finger region, the sequences are aligned as six and one-half repeats of a 57-amino-acid unit (each composed of two fingers and two linkers), as previously reported for human ZFY and ZFX and mouse Zfy-2 (25, 33, 39) and which we note is also present in mouse Zfy-1 (1). The invariant cysteines (C) and histidines (H), characteristic of zinc finger domains, are boxed.

(Fig. 8). The machinery with which transcription factors interact appears to be conserved between yeasts and mammals; the intact mammalian estrogen and glucocorticoid receptors will, for example, activate transcription in yeast cells, given appropriate binding sites (28, 38). Therefore, we believe that our yeast experiments shed light on the function of the Zfx and Zfy proteins in mice.

The sequence similarity between the products of *Zfx* and *Zfy* reinforces previous evidence (33, 39) that these X- and Y-chromosomal genes evolved from a single, common ancestral gene prior to the radiation of placental mammals. A comparison of human and mouse zinc finger sequences suggests that Zfx is more conserved than is Zfy. In this domain, the ZFX and Zfx proteins are 99.5% identical, differing by just two amino acid substitutions, while the ZFY and Zfy proteins are only 80% identical. However, human ZFY is 97% identical to ZFX and Zfx. Thus, the mouse Zfy proteins have diverged from the other X- and Y-linked human and mouse zinc finger proteins.

A comparison of the amino-terminal portions of the human ZFX and mouse Zfx proteins reveals 92% amino acid identity, with just two gaps in the aligned sequences (Schneider-Gädicke et al., in press). As a result of alternative splicing, the human *ZFX* gene also encodes an isoform with a truncated acidic domain (Schneider-Gädicke et al., in press). Although we have no direct evidence of alternative splicing within the coding region of mouse *Zfx*, we note that mouse *Zfx* contains an internal ATG (Fig. 4, codon 227) that could allow production of a similar truncated isoform. (A methionine codon is also present in homologous positions in Zfy-1 and Zfy-2 [Fig. 5].)

**Are the *Zfx* and *Zfy* genes functionally distinct?** Although *Zfx* and *Zfy* encode closely related proteins, the functions of the genes are not necessarily identical. First, their transcription patterns are quite distinct. *Zfy* transcripts have been detected only in the adult testis (25, 31). In contrast, *Zfx* appears to be widely transcribed in embryos, newborns, and adults, both female and male (Fig. 3).

Second, *Zfx* and *Zfy* differ dramatically in their 3′ UTRs. While the 3′ UTRs of *Zfy* transcripts are less than 200 bases long, *Zfx* transcripts contain 3′ UTRs of 3 to 4 kb (Fig. 4). The mouse *Zfx* 3′ UTR cross-hybridized to X-linked, presumably *Zfx*-associated sequences in all placental mammals tested (Fig. 6). In fact, nucleotide sequence analysis of a human *ZFX* cDNA reveals that its 3′ UTR is 83% identical to 2.7 kb of the mouse *Zfx* 3′ UTR (Schneider-Gädicke et al., in press). Conservation of long segments of 3′ UTRs has been observed in several other genes (14, 18, 20, 21, 34). Though

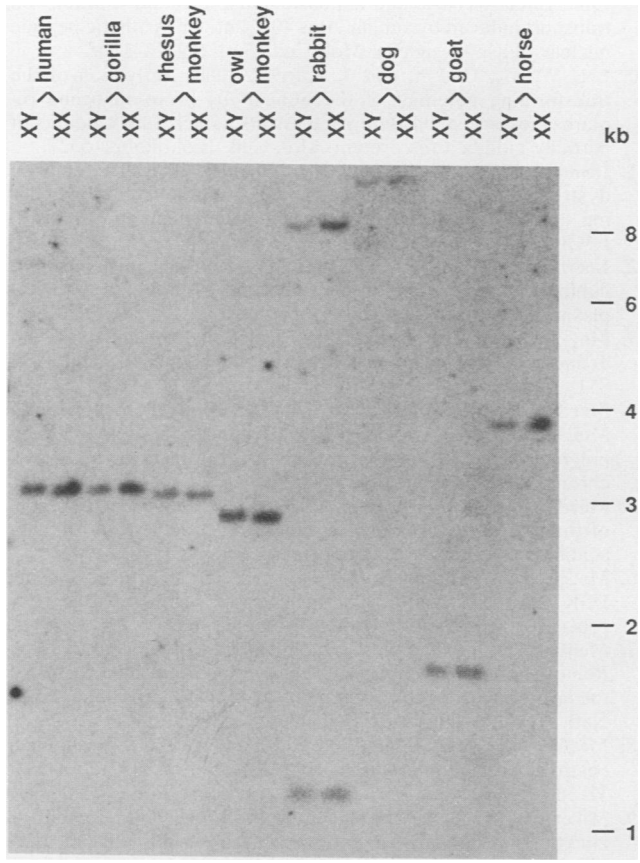


FIG. 6. The 3′ UTR of *Zfx* is conserved in mammals. A portion of the 3′ UTR of a mouse *Zfx* cDNA (Fig. 2A, probe 2) was hybridized at low stringency to *Eco*RI-digested male (XY) and female (XX) genomic DNAs from a variety of mammals. A scale in kilobases is shown at the right.

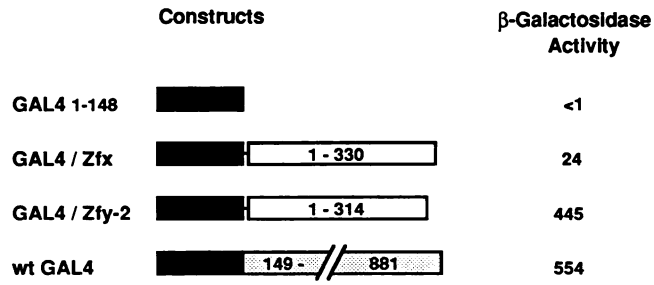


FIG. 8. The acidic domains of Zfx and Zfy-2 activate transcription in yeast cells. Fusion proteins composing the DNA-binding region of GAL4 (residues 1 to 148) (black boxes) and the acidic, amino-terminal domains of either mouse Zfx (residues 1 to 330) or Zfy-2 (residues 1 to 314) were assayed for their ability to activate transcription in yeast cells. The activating domain (residues 149 to 881) of wild-type (wt) GAL4 is shown as a stippled box. Single-copy plasmids encoding each of the indicated GAL4 derivatives were introduced into yeast cells deleted for GAL4 and bearing an integrated GAL1:*lacZ* fusion gene with a single GAL4-binding site upstream. The relative β-galactosidase activity stimulated by each GAL4 derivative is shown. Immunoprecipitation of the various GAL4 derivatives with an anti-GAL4(1-147) serum revealed no significant differences in the amounts of the proteins (results not shown).

the c-*fos* 3′ UTR has been implicated in transforming potential and message stability (20, 27, 36) and the protamine 3′ UTR regulates translation (2), the function of conserved 3′ UTR domains is generally not well understood.

Third, as assayed in yeast cells, the acidic domains of the Zfx and Zfy proteins differ in the degree to which they activate transcription. Mutations in an activating region of GAL4 have similar effects on activation in mammalian and yeast cells, suggesting that determinants of activation potency are conserved in yeasts and mammals (G. Gill, unpublished data). We have shown, in yeast cells, that a GAL4–Zfy-2 fusion protein activates transcription 18 times as effectively as a similar GAL4-Zfx fusion (Fig. 8). Mutational analysis of a GAL4-activating region demonstrated a strong correlation between the acidity of the region and potency of activation (9). However, the amino-terminal portion of Zfy-2 is a more potent activator despite being less acidic than the corresponding portion of Zfx (net charge, −68 versus −77, respectively). Thus, other structural features must affect the efficiency of activation (10). Even if Zfx and Zfy bind to the same DNA sequences and regulate the same genes, as suggested for the human homologs (39), their regulatory effects may be strikingly different. Such issues will be more directly accessible once a binding site has been identified.

**Sex-determining function for *Zfx* and *Zfy*?** If *Zfx* and *Zfy* are involved in gonadal sex determination, then it might be expected that these genes are expressed in mid-gestation embryos, when testis differentiation is first detectable histologically. Such transcription is observed for *Zfx* but has not been reported for *Zfy*. Of course, these observations do not prove or disprove a role for either gene in primary sex determination. *Zfx* is also transcribed in many adult tissues, implying that its functions are not limited to embryos. Similarly, the mouse *Zfy* genes are transcribed in the adult (but not newborn) testis, suggesting a role in male reproduction (25).

If the *Zfy* genes function in primary sex determination, then they must be functionally distinct from the *Zfx* gene. Although the mouse *Zfy* and *Zfx* genes appear to encode similar proteins, three findings suggest that the genes are not functionally interchangeable. First, the transcription patterns of *Zfy* and *Zfx* differ dramatically. Second, mouse *Zfx* transcripts contain long, conserved 3′ untranslated sequences that may have functional consequences. Third, Zfy may be far more potent than Zfx as an activator of transcrip-

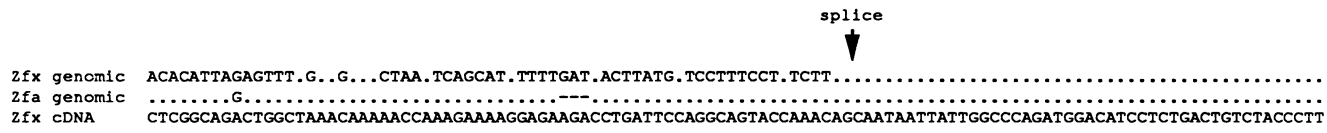


FIG. 7. *Zfa* lacks an intron present in *Zfx*. Partial nucleotide sequences of genomic clones from *Zfx* (pDP1193) and *Zfa* (pDP1194) are compared with the corresponding *Zfx* cDNA sequence. Dots represent identity to the *Zfx* cDNA, while dashes indicate gaps in one sequence compared with another. The arrow indicates the 3′ end of an intron present in the *Zfx* genomic clone but absent from both the *Zfx* cDNAs and the *Zfa* genomic clone. This splice occurs after nucleotide +1216 in the *Zfx* cDNAs.

tion. The functional relationship of these genes and their putative role in sex determination are being tested in transgenic mice.

## ACKNOWLEDGMENTS

We thank our colleagues, especially Douglas Vollrath, for critical review of the manuscript and Rebecca Mosher for preparation of mammalian genomic DNAs.

This study was supported by grants from the National Institutes of Health, the Whitaker Health Sciences Fund, the Lucille P. Markey Charitable Trust, and the Searle Scholars Program of The Chicago Community Trust. E.M.S. was supported by a fellowship from the Medical Research Council of Canada. G.G. was supported by Public Health Service grant GM32308 to Mark Ptashne from the National Institutes of Health.

## LITERATURE CITED

1. **Ashworth, A., S. Swift, and N. Affara.** 1989. Sequence of cDNA for murine *Zfy-1*, a candidate for *Tdy*. Nucleic Acids Res. **17**:2864.
2. **Braun, R. E., J. J. Peschon, R. R. Behringer, R. L. Brinster, and R. D. Palmiter.** 1989. Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. Genes Dev. **3**:793–802.
3. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell **43**:729–736.
4. **Brown, R. S., C. Sander, and P. Argos.** 1985. The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett. **186**:271–274.
5. **Dang, C. V., and W. M. F. Lee.** 1988. Identification of the human c-*myc* protein nuclear translocation signal. Mol. Cell. Biol. **8**:4048–4054.
6. **Dingwall, C., and R. A. Laskey.** 1986. Protein import into the cell nucleus. Annu. Rev. Cell Biol. **2**:367–390.
7. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
8. **Ford, C. E., O. J. Miller, P. E. Polani, J. C. de Alameida, and J. H. Briggs.** 1959. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). Lancet **i**:711.
9. **Gill, G., and M. Ptashne.** 1987. Mutants of GAL4 protein altered in an activation function. Cell **51**:121–126.
10. **Giniger, E., and M. Ptashne.** 1987. Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. Nature (London) **330**:670–672.
11. **Godowski, P. J., D. Picard, and K. R. Yamamoto.** 1988. Signal transduction and transcriptional regulation by glucocorticoid receptor-LexA fusion proteins. Science **241**:812–816.
12. **Hollenberg, S. M., and R. M. Evans.** 1988. Multiple and cooperative *trans*-activation domains of the human glucocorticoid receptor. Cell **55**:899–906.
13. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4, of yeast. Cell **46**:885–894.
14. **Hsu, C.-Y., and F. R. Frankel.** 1988. Conserved and unique sequences in the 3'-untranslated region of rat smooth-muscle alpha-actin mRNA. Gene **69**:345–348.
15. **Ito, H., Y. Fukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **53**:163–168.
16. **Jacobs, P. A., and J. A. Strong.** 1959. A case of human intersexuality having a possible *XXY* sex-determining mechanism. Nature (London) **183**:302–303.
17. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
18. **Krawetz, S. A., W. Connor, and G. H. Dixon.** 1987. Cloning of bovine P1 protamine cDNA and the evolution of vertebrate P1 protamines. DNA **6**:47–57.
19. **Lanford, R. E., R. G. White, R. G. Dunham, and P. Kanda.** 1988. Effect of basic and nonbasic amino acid substitutions on transport induced by simian virus 40 T-antigen synthetic peptide nuclear transport signals. Mol. Cell. Biol. **8**:2722–2729.
20. **Lee, W. M., C. Lin, and T. Curran.** 1988. Activation of the transforming potential of the human *fos* proto-oncogene requires message stabilization and results in increased amounts of partially modified *fos* protein. Mol. Cell. Biol. **8**:5521–5527.
21. **Lemaire, C., R. Heilig, and J. L. Mandel.** 1988. The chicken dystrophin cDNA: striking conservation of the C-terminal coding and 3' untranslated regions between man and chicken. EMBO J. **7**:4157–4162.
22. **Loewinger, L., and F. McKeon.** 1988. Mutations in the nuclear lamin proteins resulting in their aberrant assembly in the cytoplasm. EMBO J. **7**:2301–2309.
23. **Ma, J., and M. Ptashne.** 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48**:847–853.
24. **Mardon, G., R. Mosher, C. M. Disteche, Y. Nishioka, A. McLaren, and D. C. Page.** 1989. Duplication, deletion, and polymorphism in the sex-determining region of the mouse Y chromosome. Science **243**:78–80.
25. **Mardon, G., and D. C. Page.** 1989. The sex-determining region of the mouse Y chromosome encodes a protein with a highly acidic domain and 13 zinc fingers. Cell **56**:765–770.
26. **McCarrey, J. R., and K. Thomas.** 1987. Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. Nature (London) **326**:501–505.
27. **Meijlink, F., T. Curran, A. D. Miller, and I. M. Verma.** 1985. Removal of a 67 base pair sequence in the noncoding region of the proto-oncogene *fos* converts it to a transforming gene. Proc. Natl. Acad. Sci. USA **82**:4987–4991.
28. **Metzger, D., J. H. White, and P. Chambon.** 1988. The human oestrogen receptor functions in yeast. Nature (London) **334**: 31–36.
29. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. **4**:1609–1614.
30. **Mitchell, M., D. Simon, N. Affara, M. Ferguson-Smith, P. Avner, and C. Bishop.** 1989. Localization of murine X and autosomal sequences homologous to the human Y located testis-determining region. Genetics **121**:803–809.
31. **Nagamine, C. M., K. Chan, C. A. Kozak, and Y.-F. Lau.** 1989. Chromosome mapping and expression of a putative testis-determining gene in mouse. Science **243**:80–83.
32. **Page, D. C.** 1988. Is *ZFY* the sex-determining gene on the human Y chromosome? Philos. Trans. R. Soc. Lond. B Biol. Sci. **322**:155–157.
33. **Page, D. C., R. Mosher, E. M. Simpson, E. M. C. Fisher, G. Mardon, J. Pollack, B. McGillivray, A. de la Chapelle, and L. G. Brown.** 1987. The sex-determining region of the human Y chromosome encodes a finger protein. Cell **51**:1091–1104.
34. **Ponte, P., S.-Y. Ng, J. Engel, P. Gunning, and L. Kedes.** 1984. Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucleic Acids Res. **12**:1687–1696.
35. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335**:683–689.
36. **Rahmsdorf, H. J., A. Schonthal, P. Angel, M. Litfin, U. Ruther, and P. Herrlich.** 1987. Posttranscriptional regulation of c-*fos* mRNA expression. Nucleic Acids Res. **15**:1643–1659.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Schena, M., and K. R. Yamamoto.** 1988. Mammalian glucocorticoid receptor derivatives enhance transcription in yeast. Science **241**:965–967.
39. **Schneider-Gädicke, A., P. Beer-Romero, L. G. Brown, R. Nussbaum, and D. C. Page.** 1989. *ZFX* has a gene structure similar to *ZFY*, the putative human sex determinant, and escapes X inactivation. Cell **57**:1247–1258.
40. **Welshons, W. J., and L. B. Russell.** 1959. The Y chromosome as the bearer of male determining factors in the mouse. Proc. Natl. Acad. Sci. USA **45**:560–566.