Nucleic Acids Research, Vol. 20, No. 11 2827–2834

Zfp-37, a new murine zinc finger encoding gene, is expressed in a developmentally regulated pattern in the male germ line

Patrick S.Burke[1] and Debra J.Wolgemuth[1,2,3]*
[1]Departments of Genetics and Development, [2]Obstetrics and Gynecology and [3]The Center for Reproductive Sciences, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA

Received February 5, 1992; Revised and Accepted April 24, 1992 EMBL accession no. X64413

ABSTRACT

To begin to examine the function in the mouse testis of genes containing the zinc finger motif, we have screened an adult mouse total testis cDNA library with probes to a conserved region of zinc fingers. We have isolated cDNAs for a new murine zinc finger encoding gene that has been designated *Zfp-37*. Northern blot hybridization analysis revealed *Zfp-37* transcripts at high levels in the testis, the only adult tissue in which *Zfp-37* expression was observed. *Zfp-37* was also expressed at lower levels in the mid-gestation embryo and placenta. The major testicular transcripts are 2.3 and 2.6 kb. A 4.0 kb transcript was detected at lower levels in the testis as well as in embryo and placenta. Northern blot and *in situ* hybridization analysis revealed that expression of *Zfp-37* was most abundant in germ cells which have completed meiosis and are undergoing the complex morphogenetic changes of spermiogenesis. The pattern of expression of *Zfp-37* and the presence of the zinc finger domain suggest that *Zfp-37* may have a role in regulating spermiogenesis.

INTRODUCTION

Many of the genes that control development and cellular differentiation encode transcription factors. Transcription factors often have one of a number of common motifs in their protein products, including homeoboxes (1), helix loop helices (2), and zinc fingers (3, 4). These domains bind to specific DNA sequences, while other domains of the protein effect function (5, 6). The conservation of these protein domains and the DNA sequences encoding them has facilitated the isolation of potential developmental control genes by molecular cloning techniques.

The zinc finger is a conserved 28–30 amino acid motif first identified as a repeated unit of the *Xenopus* transcription factor TFIIIA (7, 8). Individual zinc fingers of this class (C_2H_2) are characterized by paired cysteine and histidine residues at conserved positions, which coordinate a Zn^{2+} and cause the intervening amino acids to loop out and form the secondary structure (9, 10). A phenylalanine and a leucine also occupy conserved positions in each zinc finger, but it is three other amino acids within the structure that make specific contacts within the major groove of DNA (10, 11). C_2H_2 zinc fingers are usually present in tandem arrays, up to 37 units (12), and show a high degree of conservation in six of the seven amino acids that typically link the zinc fingers with the consensus peptide sequence TGEKPY/F (13, 14).

Zinc finger coding sequences have been found in a number of the genes that control *Drosophila* development. Mutants for the *Drosophila* gap gene *Krüppel* (15) lack all thoracic and several abdominal segments (16), while another zinc finger encoding *Drosophila* gap gene, *hunchback*, controls development of head structures (17, 18). The *Drosophila* genes *serendipity β* and *serendipity δ*, which result in an embryonic lethal phenotype when mutant, also contain zinc fingers (19), as does *ovo*, a *Drosophila* gene that is required for the development of the female germ line (20).

A role for zinc finger genes in the process of development in the mouse has not been demonstrated *in vivo*, but a number of zinc finger encoding genes display patterns of expression consistent with such a role. Most striking is the pattern of expression of *Krox-20* in the early post-implantation embryo, where its expression exactly corresponds to the boundaries of two neuromeres (21), while in the adult transcripts are detected in the thymus, spleen, and testis (22). Krox-20 protein has been shown to bind to sequences in the 5′ flanking genomic region of the homeobox gene *Hox 1.4 in vitro* and can utilize this site to activate transcription *in vitro* (23). The zinc finger encoding genes *Zfy-1*, *Zfy-2*, *Zfp-29*, *Zfp-35* and *Rex-1* also have sites of expression in the embryo as well as the adult mouse. In the adult mouse, expression of these genes is observed exclusively in the male germ cells, which suggests that these genes can potentially play a role in controlling the complex events of spermatogenesis (24, 25, 26, 27, 28, 29).

* To whom correspondence should be addressed at Department of Genetics and Development, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA

Mouse spermatogenesis provides an interesting model system for studying genes that regulate development and differentiation within a particular cell lineage. The adult testis contains germ cells, undergoing mitosis, meiosis and the complex morphological changes of spermiogenesis as well as highly specialized somatic cells which surround the germ cells. Most of these various testicular cells and stages of their development can be studied using a number of established genetic, molecular, and cell separation techniques.

To begin to identify the role of transcription factors, particularly those of the zinc finger class, in controlling spermatogenesis, an adult mouse testis cDNA library was screened using oligonucleotides designed based on the amino acid sequence of the conserved hexapeptide that links C_2H_2 zinc fingers (13, 14). Reported here is the structure and pattern of expression of a cDNA that encodes a presumptive protein containing 12 zinc fingers. In the adult mouse expression was limited to the testis, where transcripts of 2.3 kb and 2.6 kb were detected. Low levels of a 4.0 kb mRNA were observed in mid-gestation embryo samples. Within the testis, expression was localized to cells that have completed meiosis, suggesting that this gene may be involved in controlling the complex morphogenetic process of spermiogenesis.

MATERIALS AND METHODS

Source of tissues and cells

Swiss Webster mice (Charles River, Wilmington, DE), were used as the source of adult tissues or cell populations for all experiments except those using the mouse mutant strains. Adult tissues were obtained from mice aged 3 to 6 weeks (w). For embryo collection, the day the vaginal plug was observed was day (d) 0.5 post coitum (p.c.) RNAs from testes of d7 post natal (p.n.) and d17 p.n. mice (day of birth=d1) were obtained from C57/6J mice (generously provided by J. Lee, Columbia University). The mouse mutant strains *atrichosis (at)* (ATEB/Le *a/a dat/deb*) and *white spotting (W) (WBB6F1/j W/W^v)* were obtained from The Jackson Laboratory (Bar Harbor, ME). Testicular cells were separated on a 2–4% BSA gradient as described by Wolgemuth *et al.* (30). Pools of pachytene spermatocytes, round spermatids, and residual bodies and cytoplasmic fragments are isolated with typical purities of 73%, 81%, and 76% respectively. The preparations in these experiments contained higher levels of contaminating spermatid syncitia [estimated at 10–20% compared to the usual 5% (30)] in the pachytene spermatocyte fractions. Dissected tissue specimens and cell populations were frozen in liquid nitrogen prior to RNA isolation.

Probes

Three oligonucleotides were designed based on the conserved amino acid sequence of the linker region of C_2H_2 class zinc finger genes (13, 14). These 21-mers were synthesized on an Applied Biosystems 380B DNA Synthesizer (Applied Bisosystems, Foster City, CA): ATA-CAT-ACC-GGC-G-AG-AAG-ACC, ATA-CAC-ACA-GGT-GAA-AAA-ACA, and ATC-CAC-ACT-GGA-GAG-AAA-ACC, were used to screen the initial cDNA library. An EcoRI fragment, free of zinc fingers, of a clone isolated from the initial library screening was subcloned into pBluescript KS(+), named Z91.0, and used for the subsequent library screening and expression studies. Oligonucleotides were end-labeled with ^{32}P[ATP] (DuPont, Wilmington, DE) using T4 polynucleotide kinase (Boehringer Mannheim, Indianapolis, IN). Probes for Northern blot hybridization analyses were labeled with ^{32}P[dCTP] (DuPont) using Amersham's Multiprime DNA labeling systems (Amersham, Arlington Heights, IL) to specific activity greater than 10^9 cpm/mg or antisense RNA was transcribed from linearized plasmid (pZ91.0) in the presence of ^{32}P[UTP] (DuPont), according to protocols suggested by the manufacturer (Promega Biotech, Madison, WI). For *in situ* hybridization analysis sense and antisense RNAs were transcribed from linear plasmids (pZ91.0) according to the manufacturer's instructions in the presence of ^{35}S[UTP] (DuPont) to a specific activity greater than 2×10^9 cpm/pmole. Prior to use for *in situ* hybridization the probes were hydrolyzed to 0.1 kb fragments according to the protocol of Cox *et al.* (31).

Library screening

2.5×10^5 recombinant plaques of an adult mouse total testis cDNA library in λgt10 [provided by Dr. D. Meijer (32)] were screened in duplicate essentially as described in Sambrook *et al.* (33). Plaques were lifted onto nitrocellulose filters, and hybridized with ^{32}P labeled zinc finger 21-mers at 55°. The filters were washed to a final stringency of 2×SSC (1×SSC: 150 mM NaCl/15 mM NaCitrate)/0.25% SDS at room temperature (RT) and exposed to autoradiographic film. Positives were subjected to two rounds of plaque purification. The cDNA inserts from isolated phages were subcloned into pTZ 19R (United States Biochemical, Cleveland, OH) or pBluescript KS(+) (Stratagene, La Jolla, CA). Southern hybridization analysis of the restricted phage λ clones with the zinc finger 21-mers suggested that most of the released fragments contained zinc finger encoding sequences. Oligonucleotide probes specific for the testis-expressed zinc finger encoding genes *Zfy1/Zfy2* (24), *Krox-20* (13), and *mkr5* (14) did not hybridize with any of the isolated clones. One phage λ clone, λZ9221, yielded two restriction fragments upon EcoRI digestion (Fig.1). The fragment which did not hybridize with the zinc finger 21-mers, Z91.0 (Fig. 1), was used to obtain

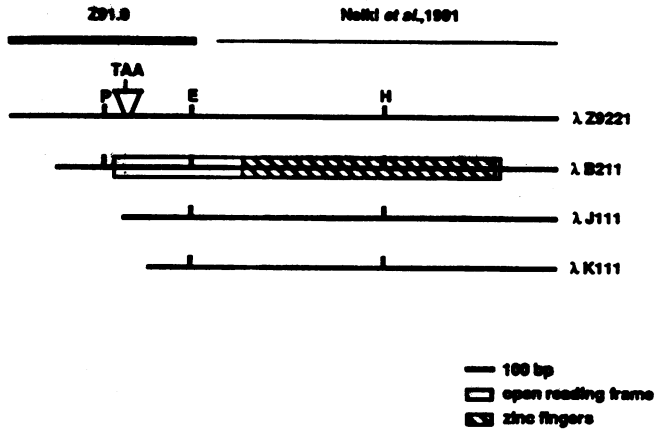

Figure 1. Diagram of cDNA Clones. Schematic representation of cDNA clones of *Zfp-37*. The fragment used as a probe in this study, Z91.0, is indicated above the clone from which it was derived. The cDNA clones are co-linear. The open reading frame is only shown on the longest cDNA. Also indicated is the position of homology between these cDNAs and a partial cDNA sequence which was previously published (Nelki *et al.*, 1991). P=PstI, E=EcoRI, H=HindIII.

three additional independent clones from an adult mouse total testis cDNA library (provided by D.L.Chapman, Columbia University) made in λZAP (Stratagene). 1.0×10^6 recombinant plaques were lifted onto nitrocellulose filters and treated as described above. Hybridization was performed with ^{32}P multiprimed labeled Z91.0 essentially as described above but at 65° for 16−20 hours (h), then washed at a final stringency of 0.1×SSC/0.1% SDS at 65°. Plaque purified clones were *in vivo* excised according to the protocol provided by the manufacturer. The resulting plasmid contained the cDNA of interest inserted between the EcoRI and XhoI site of pBluescript SK(−).

DNA sequencing

Clones were sequenced using the double-stranded sequencing procedure for the Perkin Elmer Cetus AmpliTaq Sequencing Kit (Perkin−Elmer, Norwalk, CT) according to the manufacturer's protocol. DNA sequence analysis was also performed on an Applied Biosystems Model 373A DNA sequencer (Applied Bisosystems, Foster City, CA). Sequences were analyzed using IBI Sequence Analysis Program (34) and the Sequence Analysis for the VAX (35).

RNA isolation and Northern blot hybridization analysis

Total RNA was isolated according to the method of Cathala *et al.* (36) or Auffray and Rougeon (37). Poly(A)$^+$ RNA was selected by oligo dT cellulose chromatography according to Aviv and Leder (38). RNA samples were electrophoresed on denaturing 0.8% agarose/2.2 M formaldehyde gels, blotted onto Nytran (Schleicher and Schuell, Keene, NH) membranes essentially according to Sambrook *et al.* (33). Even loading of the samples was visualized by ethidium bromide staining. Hybridization with multiprime labeled DNA probe was essentially according to Sambrook *et al.* (33). Blots were washed at a final stringency of 0.1×SSC/0.1% SDS at 65°. Hybridization with riboprobe was performed according to Krumlauf *et al.* (39). Filters were washed at a final stringency of 0.2×SSC/1% SDS at 80° for 1−2 h. Filters were exposed to autoradiographic films with intensifying screens at −70° for 1 d to 1 w.

In situ hybridization experiments

In situ hybridization using ^{35}S radiolabeled RNA probes on sections of paraffin embedded tissues was done essentially as described by Jaffe *et al.* (40). Following hybridization and

```
   1  GGCACGAGGGAGCTGGTGTTTTAGGAATGGGAGCAGCTGGAACCTGTGCAGAGAGATGTGTACAAGGATACGAAGCTAGAGAACTGCAGC

  91  AATCCAGCCTCCATGGGAAATCAAGATCCCAAACAAGACATAGTCTCCGTTTTGGAAGAAGAAGAGCCATCATCGGGAAAGGGGAAAAAA
                 MetGlyAsnGlnAspProLysGlnAspIleValSerValLeuGluGluGluGluProSerSerGlyLysGlyLysLys

 181  GCCAGCCCAAGTAGTCTGAAAAAAATAGCAAGGCCCAAGACAGCAGGAACAAGTGCAAAACTCCAACAAGATGATGAGCATAGGGAGGAA
      AlaSerProSerSerLeuLysLysIleAlaArgProLysThrAlaGlyThrSerAlaLysLeuGlnGlnAspAspGluHisArgGluGlu

 271  AAGCAGAAGACCCAAAGCAAACTTACTAAGGAAGTGACACTCAGGAAGAAAAGTTCCAACAGCAAGAAAAGCAGTGAGTATGGTTTGTTG
      LysGlnLysThrGlnSerLysLeuThrLysGluValThrLeuArgLysLysSerSerAsnSerLysLysSerSerGluTyrGlyLeuLeu

 361  GAGAACAAAAGTCTCCACTCAAAACACACTCCTTCCGAGAAAAAACTGCTTAAGTCCAGTTCCCGTGGGAAGAACTCGAATCAGAATTCA
      GluAsnLysSerLeuHisSerLysHisThrProSerGluLysLysLeuLeuLysSerSerSerArgGlyLysAsnSerAsnGlnAsnSer

 451  GACTCTCTGAAAAAGAAACCTGACACAGCTAATGAACACAGGAAATCACTCAGCCATTCTGCATCTGATGTGAACAAAGATGAAATTCCA
      AspSerLeuLysLysLysProAspThrAlaAsnGluHisArgLysSerLeuSerHisSerAlaSerAspValAsnLysAspGluIlePro

 541  ACTAGAAAGAAATGCGACAAGTTACCCAACAATAAGTTGTCTGATAAAGGTGACAAAAACCAAACCAGCAAAAAATGTGAGAAAGTATGC
      ThrArgLysLysCysAspLysLeuProAsnAsnLysLeuSerAspLysGlyAspLysAsnGlnThrSerLysLysCysGluLysValCys

 631  CGTCATAGTGCATCCCATACCAAGGAAGACAAAATTCAGACCGGGGAGAAACGGAAATCACACTGCCGTACTCCATCTAAACCTGAAAAA
      ArgHisSerAlaSerHisThrLysGluAspLysIleGlnThrGlyGluLysArgLysSerHisCysArgThrProSerLysProGluLys

 721  GCCCCAGGTTCTGGGAAACCTTATGAATGTAACCACTGTGGGAAGGTCCTCAGCCATAAACAGGGACTCCTTGACCATCAAAGAACTCAC
      AlaProGlySerGlyLysProTyrGluCysAsnHisCysGlyLysValLeuSerHisLysGlnGlyLeuLeuAspHisGlnArgThrHis

 811  ACTGGGGAGAAACCATATGAATGTAATGAATGTGGGATAGCTTTCAGCCAGAAGTCCCACCTTGTTGTACATCAGAGAACTCACACTGGG
      ThrGlyGluLysProTyrGluCysAsnGluCysGlyIleAlaPheSerGlnLysSerHisLeuValValHisGlnArgThrHisThrGly

 901  GAAAAACCATACGAGTGTGAACAGTGTGGCAAAGCACACGGACATAAACATGCCCTCACTGACCATCTAAGAATCCATACTGGAGAAAAG
      GluLysProTyrGluCysGluGlnCysGlyLysAlaHisGlyHisLysHisAlaLeuThrAspHisLeuArgIleHisThrGlyGluLys

 991  CCCTACAAATGTAATGAATGTGGCAAAACGTTTAGACACAGCTCAAACCTTATGCAACACCTAAGATCTCACACGGGTGAGAAGCCGTAT
      ProTyrLysCysAsnGluCysGlyLysThrPheArgHisSerSerAsnLeuMetGlnHisLeuArgSerHisThrGlyGluLysProTyr

1081  GAATGTAAGGAATGTGGCAAATCCTTTAGATATAATTCATCTCTTACTGAACATGTGAGAACACACACAGGTGAAATACCATACGAATGT
      GluCysLysGluCysGlyLysSerPheArgTyrAsnSerSerLeuThrGluHisValArgThrHisThrGlyGluIleProTyrGluCys

1171  AACGAATGTGGCAAAGCTTTCAAGTATGGCTCATCCCTGACTAAACATATGCGGATTCATACAGGGGAGAAACCATTTGAATGTAATGAA
      AsnGluCysGlyLysAlaPheLysTyrGlySerSerLeuThrLysHisMetArgIleHisThrGlyGluLysProPheGluCysAsnGlu

1261  TGTGGGAAAACTTTTAGCAAAAAGTCACACCTAGTTATACATCAAAGAACTCATACAAAGGAGAAACCTTATAAATGTGATGAGTGTGGG
      CysGlyLysThrPheSerLysLysSerHisLeuValIleHisGlnArgThrHisThrLysGluLysProTyrLysCysAspGluCysGly

1351  AAAGCCTTTGGACATAGCTCATCTCTTACCTACCATATGAGAACTCATACAGGTGACTGCCCCTTTGAATGTAATCAATGTGGTAAGGCC
      LysAlaPheGlyHisSerSerSerLeuThrTyrHisMetArgThrHisThrGlyAspCysProPheGluCysAsnGlnCysGlyLysAla

1441  TTTAAACAGATTGAAGGCCTTACCCAACACCAGAGAGTTCACACAGGGGAGAAACCCTATGAGTGTGTTGAATGTGGGAAAGCCTTTAGT
      PheLysGlnIleGluGlyLeuThrGlnHisGlnArgValHisThrGlyGluLysProTyrGluCysValGluCysGlyLysAlaPheSer

1531  CAGAAGTCACACCTCATCGTACACCAGAGACCTCATACAGGGGAGAACCCCTTTGAATGTTATGAGTGTGGAAAAGCCTTCAATGCAAAA
      GlnLysSerHisLeuIleValHisGlnArgProHisThrGlyGluAsnProPheGluCysTyrGluCysGlyLysAlaPheAsnAlaLys

1621  TCACAACTTGTTATTCATCAGAGATCCCACACTGGAGAGAAACCCTATGAATGTATTGAATGTGGTAAAGCCTTCAAGCAAAATGCCTCT
      SerGlnLeuValIleHisGlnArgSerHisThrGlyGluLysProTyrGluCysIleGluCysGlyLysAlaPheLysGlnAsnAlaSer

1711  CTTACCAAACATATGAAAATTCACTCAGAAGAACAATCTGAGGAAGAAGATTAATGTAGGAAACCTGACGAACTGACTGGTTTGGTATTA
      LeuThrLysHisMetLysIleHisSerGluGluGlnSerGluGluGluAspEnd

1801  TTTAACCTTAAAGATGTTCTCAATTTGATGATGTTAGAATATCTTTTTTTAGGAAATCATTCCTGGTGATACATGAGAGAATTTGAATAT

1891  GGATCTTTACAATAAGATGGTAATAAAATTTAACCTTGATCCCAAAAAAAAAAAAAAAAA  1958
```

Figure 2. Sequence of *Zfp-37* cDNA. Complete sequence of λB211. The poly(A)$^+$ signal and putative translation initiation sequence are underlined. Zinc finger encoding sequences are indicated within the box.

washing, the slides were counter stained with hematoxylin and eosin, and coverslipped using Pro-texx mounting media (American Scientific Products, McGraw Park, IL). Photomicrographs were taken using Kodak Technical Pan film on a Leitz photomicroscope using bright field and epiluminesence optics.

RESULTS

A zinc finger gene expressed in the testis can encode a protein with 12 zinc fingers

The cDNAs diagrammed in Figure 1 were isolated from adult mouse total testis cDNA libraries using oligonucleotides designed against the sequence of the linker region of C_2H_2 type zinc fingers (13, 14) as described in Materials and Methods. The sequence for the longest cDNA (λB211, Fig. 2) encoding the putative open reading frame is 1950 bp and includes a polyadenylation signal and poly(A) tail. Also present is an ATG, which is within a sequence that is in good agreement with the consensus translation initiation sequence for vertebrates (GCCRCCATGG) (41), and an open reading frame of 1661 bp. Conceptual translation of this open reading frame yielded a protein of 553 amino acids with 12 zinc fingers of the C_2H_2 type. The zinc finger sequences, indicated by a box in Figure 2 and aligned in Figure 3, are located in a tandem array in the 3′ portion of the clone from base pair 749 to 1736.

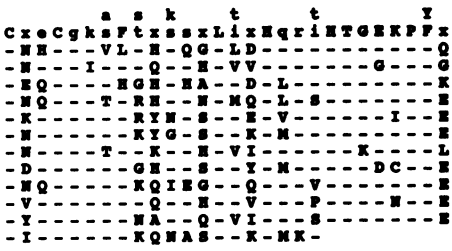

Figure 3. Zinc Fingers of *Zfp-37*. The 12 zinc fingers of the conceptual translation of λB211 are shown aligned beneath the consensus sequence of mouse C_2H_2 zinc fingers. This consensus was obtained by comparison of approximately 120 individual zinc finger peptide sequences (Chavrier *et al.*, 1988; Chowdhury *et al.*, 1988; Cunliffe *et al.*, 1990; Denny and Ashworth, 1991). Upper case letters indicate the most highly conserved residues, and the lower case letters indicate more variable positions. In positions where either of two amino acids are likely to be present, both are indicated. Residues identical to those of the consensus are indicated with a dash.

The original isolate (λZ9221) contains a 99 bp insertion relative to λB211 and λJ111 (Fig. 1). This artifact contained a stop codon in frame between the ATG and the zinc finger encoding sequences. The insertion, which was present in the probe used in the present study, was not responsible for the hybridization to any of the transcripts detected by Northern analysis. Probes from regions of λZ9221 that did not include the 99 bp, as well as the first 444 bp of 5′ sequence of λB211, detected the same transcripts as Z91.0 (data not shown).

A partial cDNA sequence of 1434 bp (42) had been identified which showed 99.6% homology to the 3′ region of λB211, including the zinc finger encoding sequences (Fig. 1). This cDNA was smaller than any cDNAs isolated and any transcripts detected in the present study. The cDNA of Nelki *et al.* (42) included an extra base pair (position 121), relative to the clones in this study (position 630 of λB211). Without this extra base, the open reading frame extended beyond the 5′ end of the previously reported cDNA sequence (42); the previously postulated initiator codon as well as the stop codons upstream of the postulated ATG were eliminated (42). There were two other differences between our sequence data and that reported by Nelki *et al.* (42). At base 877 of λB211, a G in place of a C changes a leucine to a valine within one zinc finger of the presumptive protein, and at position 1903 of λB211, there was an extra A in the 3′ untranslated region. The presence of an initiator methionine within a sequence favorable to translation initiation (41) in B211 and the observation that the smallest

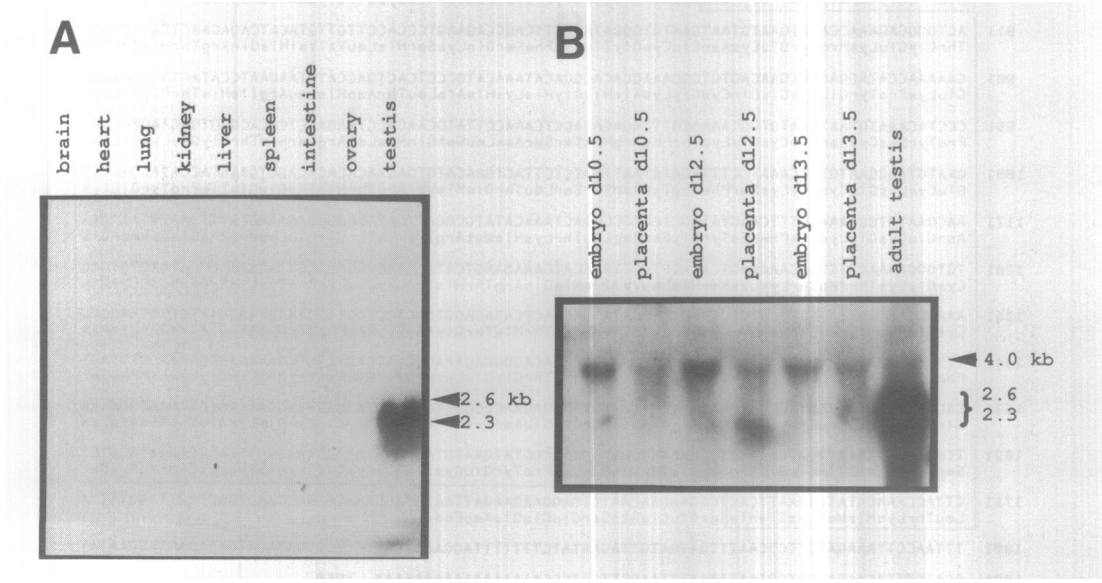

Figure 4. Northern Blot Hybridization Analysis of Adult Tissues, Embryo and Placenta. a) Northern blot of total RNA from adult tissues using random prime labeled Z91.0. 20 μg total RNA was loaded in each lane. Exposure: 1 week. b) Northern blot of total RNA from embryo, placenta, and testis using random prime labeled Z91.0. Each lane contained 20 μg total RNA. Equal loading was judged by visualizing the ethidium bromide stained gel (data not shown). Exposure: 1 week.

transcript detected with probes from λB211 is 2.3 kb (see below) indicates that λB211, not the previously reported cDNA clone (42), is the more complete cDNA. The zinc finger encoding gene represented by this cDNA (λB211) and that of Nelki *et al.* (42) has been assigned the name *Zfp-37* (Dr. D. Doolittle, The Jackson Laboratory, personal communication).

Distribution of *Zfp-37* expression

To determine the pattern of expression of *Zfp-37*, total RNA from a panel of adult tissues was analyzed by Northern blot hybridization using Z91.0 as probe. *Zfp-37* transcripts were detected only in the testis (Fig. 4a). This signal could be further resolved as transcripts of 2.3 kb and 2.6 kb (see Fig. 5b and 5c). No message was detected in brain, heart, lung, kidney, liver, spleen, small intestine, or ovary (Fig. 4a) nor in uterus or stomach (data not shown). *Zfp-37* expression was also examined by Northern blot hybridization analysis of RNA from mid-gestation embryos and placentae. Low levels of hybridization to a 4.0 kb mRNA were observed in all samples tested (Fig. 4b), being more abundant in embryo than placenta RNA. This 4.0 kb band was occasionally observed in testis RNA as well (Fig. 4b), but was never observed in any other adult tissue.

Localization of the cellular site of expression of *Zfp-37* in the testis

The adult testis contains both somatic and germinal cells. The somatic cells, which include Sertoli, Leydig, and peritubular myoid cells, produce hormones and provide support for the developing germ cells. Germ cells in various stages of spermatogenesis, from spermatogonia to spermatozoa, constitute the majority of cells (>90%) in the adult testis. To determine whether *Zfp-37* was expressed in the somatic, germinal, or in both cell types, RNA was isolated from testes of germ cell deficient mutant mouse strains. The testes of adult mice homozygous for mutations at the *W* (*white spotting*) or *at* (*atrichosis*) locus are essentially devoid of germ cells yet contain the full array of somatic cell types. Heterozygous littermates contain the full complement of germ and somatic cells and are fertile (43, 44). Northern blot hybridization analysis using Z91.0 as a probe detected expression only in RNA isolated from the testes of heterozygous (*W*/+) or (*at*/+) mice (Fig. 5a), indicating that the 2.3 kb and 2.6 kb transcripts were limited to testes containing germ cells. *Zfp-37* is therefore likely expressed only in germ cells.

Testes that are enriched for pre-meiotic, meiotic, and post-meiotic germ cells can be obtained from mice at specific stages of post-natal development (45). From birth until d7 p.n. development, the testes of mice contain somatic cells and mitotically dividing spermatogonia. Testes of mice at d17 p.n. contain all the somatic cell types, spermatogonia, and meiotic spermatocytes. The testes of adult mice contain post-meiotic spermatids and spermatozoa as well as spermatocytes, spermatogonia, and somatic cells. RNA was prepared from testes obtained from d7 p.n., d17 p.n., and adult mice and analyzed by Northern blot hybridization with antisense riboprobe transcribed from pZ91.0 (Fig. 5b). Transcripts of 2.3 kb and 2.6 kb were observed only in the sample from adult testis, indicating that these mRNAs were expressed post-meiotically, and not in spermatogonia or spermatocytes.

To determine in which post-meiotic germ cell types *Zfp-37* is expressed and whether the 2.3 kb and 2.6 kb transcripts exhibit a differential distribution in spermiogenic cells, RNA was prepared from pooled fractions of enriched populations of meiotic prophase spermatocytes (predominantly of the pachytene stage), early (round) spermatids, and cytoplasmic fragments and residual bodies. Cytoplasmic fragments arise as mechanically sheared

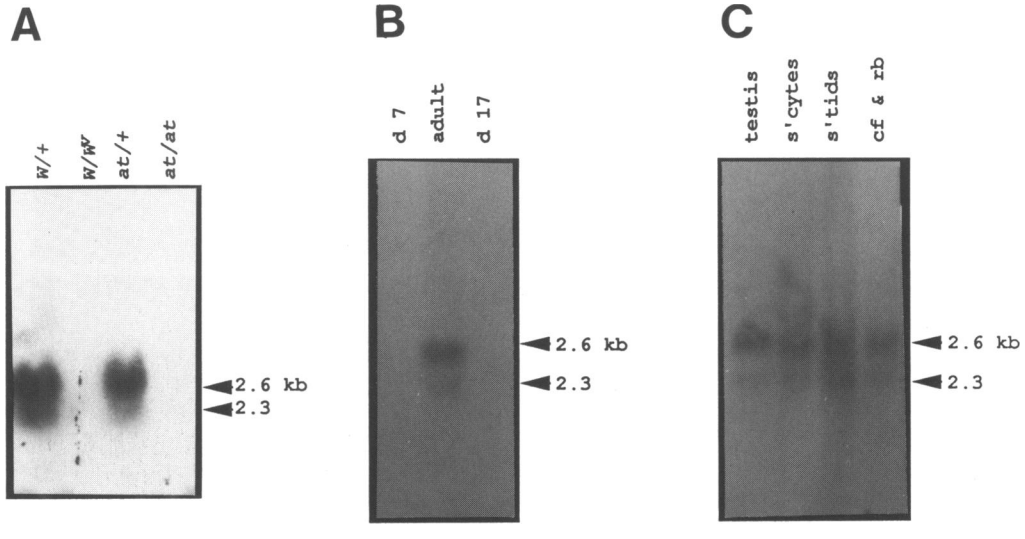

Figure 5. Northern Blot Hybridization Analysis of Mutant Testes, Pre-pubertal Testes and Pools of Separated Germ Cells. a) Northern blot of total RNA from testes of *W*/+ or *at*/+ (wild type. Lanes 1 and 3), or *W*/*W*^{v} or *at*/*at* (germ cell deficient. Lanes 2 and 4) mice using random primed labeled Z91.0. Each lane contained 20 μg total RNA. Exposure: 1 week. b) Northern blot of poly $(A)^+$ RNA from testes at discrete developmental stages using labeled antisense riboprobe from pZ91.0. 5 μg poly$(A)^+$ RNA in each lane. Exposure: 1 week. c) Northern blot of RNA from total testis and pools of purified germ cells using labeled antisense riboprobe from pZ91.0. The lanes contained 20 μg total testis RNA, or 15 μg total RNA from each indicated pool of cells. s'cytes=pachytene spermatocytes, s'tids=round spermatids, cf & rb=cytoplasmic fragments and residual bodies. Equal loading was judged by visualizing the ethidium bromide stained gel (data not shown). Exposure: 1 week.

cytoplasm from the late (elongating) spermatids. Together with the residual bodies, cytoplasmic fragments represent the cytoplasm from late spermatids (30). Northern blot hybridization analysis using antisense riboprobe transcribed from pZ91.0 (Fig. 5c) revealed that both the 2.3 kb and 2.6 kb transcripts were present in all samples, but most abundantly in the round spermatid fraction. Furthermore, each transcript appeared to be of the same relative abundance from sample to sample. In these experiments, the fractions pooled as pachytene spermatocytes contained up to 20% contaminating spermatid syncitia; more typically, this procedure yields fractions with approximately 5% contaminating cells (30). This contamination likely accounted for the presence of *Zfp-37* transcripts in the pachytene spermatocyte fraction.

In situ hybridization analysis confirmed the localization of *Zfp-37* expression to spermatids. Hybridization signal was most intense over the post-meiotic round spermatids, and elongating spermatids (Fig. 6a, and b). The abundant expression of *Zfp-37* in the round spermatids and elongating spermatids is visualized in the higher magnification photomicrographs (Fig. 6c and 6d). The larger cells that occupy the majority of the tubule in Figure 6d are pachytene spermatocytes. These cells and the spermatogonia, located in the basal region of the seminiferous epithelium (Fig. 6a-d), did not exhibit signal above background. Hybridization with sense riboprobe transcribed from linearized pZ91.0 (Fig. 6f) displayed no specific labeling. These results confirm the Northern blot hybridization analysis of RNA isolated from testes from d7 p.n., d17 p.n., and adult mice, and suggested that the hybridization signal detected by Northern analysis in the pachytene spermatocyte fraction (Fig. 5c) was indeed due to the contaminating early spermatid syncitia. No signal above background was observed over the somatic cells or within the germ cell deficient tubules in sections of *at/at* testis (Fig 6e).

DISCUSSION

We have isolated four clones from adult mouse testis cDNA libraries that represent the murine zinc finger encoding gene designated *Zfp-37*. Clone λB211 contains an ATG within a sequence that is favorable to translation initiation in vertebrates

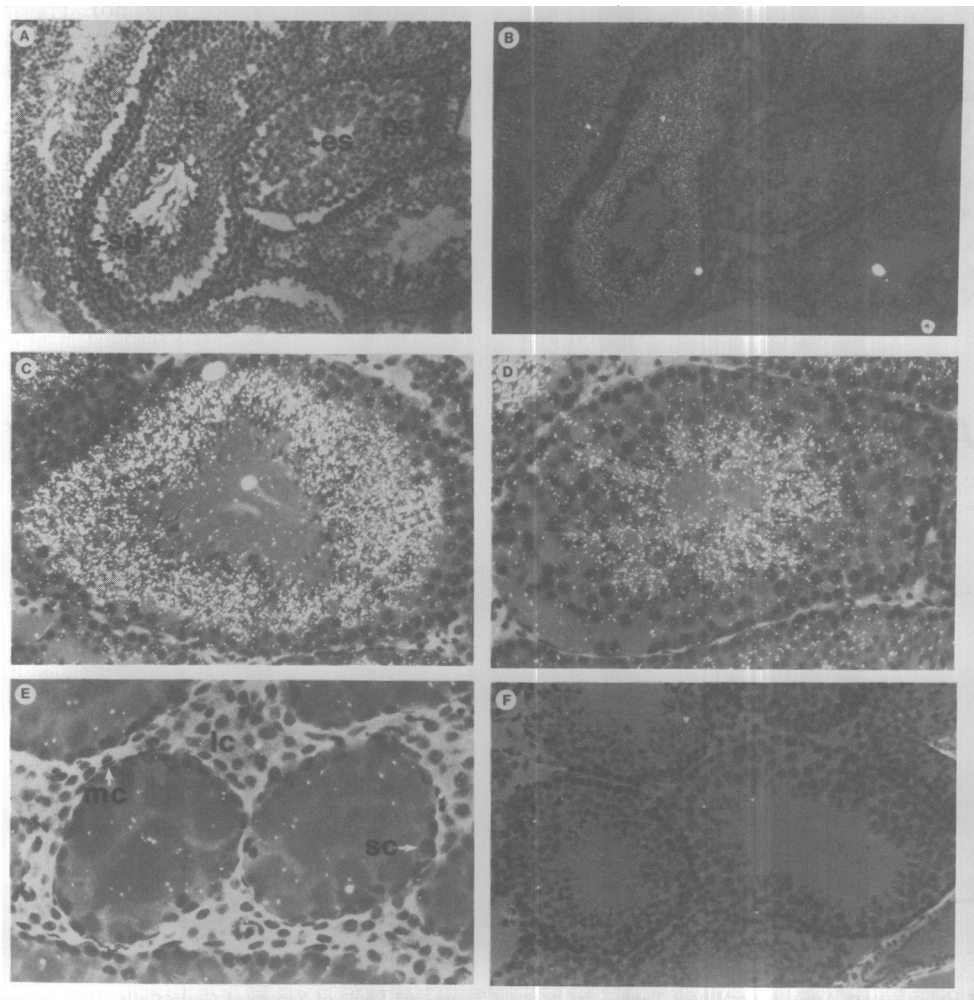

Figure 6. *In Situ* Hybridization Analysis of Testicular Sections. a) bright field and b) epiluminesence optics photomicrographs of histological sections of *W/+* testis using ^{35}S labeled antisense riboprobe from pZ91.0. 65×magnification. c) and d) epiluminesence optics photomicrographs of *W/+* testis using labeled antisense riboprobe from pZ91.0. 160×magnification. e) epiluminesence optics photomicrograph of *at/at* testis using labeled antisense riboprobe from pZ91.0. 200×magnification. f) epiluminesence optics photomicrograph of *at/+* testis using labeled sense riboprobe from pZ91.0 (control). 125×magnification. sg=spermatogonia, ps=pachytene spermatocytes, rs=round spermatids, es=elongating spermatids, mc=peritubular myoid cell, lc=Leydig cell, sc=Sertoli cell.

(41) and in the same reading frame with sequence encoding 12 zinc fingers. These observations strongly support the possibility that the clone contained the entire open reading frame. Initial database searches with sequences upstream of the zinc fingers of λB211 did not reveal identity with known sequences, while the zinc finger encoding region showed as much as 65% homology with other zinc finger encoding genes.

The most abundant transcript detected by Northern analysis was the 2.6 kb species that was limited to the cells undergoing the complex morphogenetic changes of spermiogenesis. A second transcript of 2.3 kb was also observed in spermatids, at decreased but constant level relative to the 2.6 kb mRNA. Although it is possible that the size difference was due to changes in poly(A) tail length that has been associated with translation initiation and observed for other genes expressed in the germ cells (46; C.M.Viviano and D.J.Wolgemuth, in preparation) no change in relative abundance was observed between the 2.6 kb and 2.3 kb transcripts between early and late spermatids. It is also possible that the 2.3 kb, as well as the faint 4.0 kb, transcripts could arise as products of alternative processing, from alternative promoters of the same gene, or as the products of a gene or genes related to *Zfp-37*. We are in the process of isolating and characterizing genomic clones to make this determination.

The first 216 amino acids of the presumptive protein are 24% lysine and arginine residues with a predicted charge of (+22) (34). The amino terminus is also serine/threonine rich (21%) and may be phosphorylated, since many of these residues are present within consensus peptide sequences for several kinases (47), including a cyclic AMP-dependent protein kinase which has been shown to be active in germ cells (48). A threonine residue at amino acid 200, immediately preceding the zinc fingers, is within a consensus sequence for phosphorylation by *cdc2* (S/T-X-P-R/K) (49, 50).

The remainder of the protein consists of 12 zinc fingers of the C_2H_2 type in a tandem array. A consensus peptide sequence for mouse C_2H_2 zinc fingers was dervived by comparing peptide sequences from individual and consensus zinc fingers presented for a number of these genes (13, 14, 27, 28, 51) (Fig. 3). The zinc fingers of ZFP-37 showed good agreement with this consensus sequence (Fig. 3). It is interesting to note that the ZFP-37 putative zinc fingers displayed the most variability at positions which make specific base contacts in the major groove of the target DNA (10, 11). After this domain is a very short carboxy terminus in which 6 of the 9 amino acids are acidic residues.

The presumptive protein product of *Zfp-37* contains zinc fingers and is thus a member of a class of proteins capable of binding specific DNA sequences and regulating gene expression. It will be interesting to determine and characterize the genomic target, or targets, of ZFP-37 to elucidate the function of this particular zinc finger gene. Identification of target genes will also permit us to test an hypothesis that has been proposed for predicting zinc finger DNA binding sites (52). A predicted DNA binding site was formulated according to the rules suggested by Klevit (52), and the sequence (GNNAGNGNNGAANNNNNGGGN-GNNANAAGNNANANG) was used to search the GenBank and EMBL databases. No homology was detected, but relatively few of the sequences in the databases represent genomic as opposed to cDNA sequences. Also, the DNA sequence of the native binding site, even if it agrees closely with the predicted site, may be discontinuous, as appears to be the case for TFIIIA (52, 53). This could render a simple homology search inconclusive.

The distribution of the transcripts of *Zfp-37* is consistent with its product being involved in the events of spermiogenesis. During spermiogenesis the cell undergoes a number of complex structural changes. The morphology of the round spermatid changes dramatically: head and flagellar structures are formed, the histones are replaced by protamines, the nucleus condenses, the majority of the cytoplasm is sloughed off and the cell becomes transcriptionally and translationally inactive. The identification of *Zfp-37* provides an additional candidate in the genetic hierarchy governing this complex developmental pathway, and gives us a tool to resolve other steps in the process.

ACKNOWLEDGMENTS

We would like to thank: Dr. Dies Meijer for providing one of the testis cDNA libraries used in this study, Deborah Chapman for the second testis cDNA library, critical reading of the manuscript and many stimulating discussions, Leah Jaffe and Dr. Jamie Lee for help with the *in situ* hybridization procedure, Carol Gruppi for providing the Northern blot of embryo and placenta RNAs, Dr. Don Doolittle for assistance with the nomenclature of *Zfp-37*, and Dr. Martin Winer for his comments on the manuscript. This work was supported in part by NIH grant R01 HD18122 and T32 GM 07088.

REFERENCES

1. Gehring, W.J. (1987) *Science* **236**, 1245–1252.
2. Pabo, C.O., and Sauer, R.T. (1984) *Ann. Rev. Biochem.* **53**, 293–321.
3. Klug, A., and Rhodes, D. (1987) *TIBS* **12**, 464–469.
4. Evans, R.M., and Hollenberg, S.M. (1988) *Cell* **52**, 1–3.
5. Smith, D.R., Jackson, I.J, and Brown, D.D. (1984) *Cell* **37**, 645–652.
6. Sigler, P.B. (1988) *Nature* **333**, 210–212.
7. Brown, R.S., Sander, C., and Argos, P. (1985) *FEBS Letters* **186**, 271–274.
8. Miller, J., McLachlan, A.D., and Klug, A. (1985) *EMBO J.* **4**, 1609–1614.
9. Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. (1989) *Science* **245**, 635–637.
10. Pavletich, N.P., and Pabo, C.O. (1991) *Science* **252**, 809–817.
11. Nardelli, J., Gibson, T.J., Vesque, C., and Charnay, P. (1991) *Nature* **349**, 175–178.
12. Ruiz i Altaba, A., Perry-O'Keefe, H., and Melton, D.A. (1987) *EMBO J.* **6**, 3065–3070.
13. Chavrier, P., Lemaire, P.,Revelant, O., Bravo, R., and Charney, P. (1988) *Mol. Cell. Biol.* **8**, 1319–1326.
14. Chowdhury, K., Rohdewohld, H., and Gruss, P. (1988) *Nucleic Acids Res.* **16**, 9995–10011
15. Rosenberg, U.B., Schröder, C., Preiss, A., Kienlin, A., Côtè, S., Riede, I., and Jäckle, H. (1986) *Nature* **319**, 336–339.
16. Wieschaus, E., Nusslein-Volhard, C., and Kluding, H. (1984) *Devel. Biol.* **104**, 172–186.
17. Lehmann, R., and Nüsslein-Volhard, C. (1987) *Devel. Biol.* **119**, 402–417.
18. Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kielin, A., Jones, K., and Jäckle, H. (1987) *Nature* **327**, 383–389.
19. Vincent, A., Colot, H.V., and Rosbash, M. (1985) *J. Mol. Biol.* **186**, 149–166.
20. Mèvel-Ninio, M., Terracol, R., and Kafatos, F.C. (1991) *EMBO J.* **10**, 2259–2266.
21. Wilkinson, D.G., Bhatt, S., Chavrier, P., Bravo, R., and Charnay, P. (1989) *Nature* **337**, 461–464.
22. Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R., and Charnay, P. (1988) *EMBO J.* **7**, 29–35.
23. Chavrier, P., Vesque, C., Galliot, B., Vigneron, M., Dollè, P., Duboule, D., and Charnay, P. (1990) *EMBO J.* **9**, 1209–1218.
24. Mardon, G., and Page, D.C. (1989) *Cell* **56**, 765–770.
25. Koopman, P., Gubbay, J., Collignon, J., and Lovell-Badge, R. (1989) *Nature* **342**, 940–942.
26. Nagamine, C.M., Chan, K., Hake, L.E., and Lau, Y.F.C. (1990) *Genes & Devel.* **4**, 63–74.

27. Cunliffe, V., Koopman, P., McLaren, A., and Trowsdale, A. (1990) *EMBO J.* **9**, 197–205.
28. Denny, P., and Ashworth, A., (1991) *Gene* **106**, 221–227.
29. Rogers, M. B., Hosler, B. A., and Gudas, L. J. (1991) *Development.* **113**, 815–824.
30. Wolgemuth, D.J., Gizang-Ginsberg, E., Engelmyer, E., Gavin, B.J., and Ponzetto, C. (1985) *Gamete Res.* **12**, 1–10.
31. Cox, K.H. DeLeon, D.V., Angerer, L.M., and Angerer, R.C. (1984) *Devel. Biol.* **101**, 485–502.
32. Meijer, P., Hermans, A., von Lindern, M., van Agthoven, T., de Klein, A., Mackenbach, P., Grootegoed, A., Talarico, D., Valle, G. D., and Grosveld, G. (1987) *EMBO J.*. **6**, 4041–4047.
33. Sambrook, J., Fritsch, E.J., and Maniatis, T. (1989) Molecular cloning—a laboratory manual. 2nd Edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.
34. Pustell, J.M. (1988) *Nucleic Acids Res.* **16**, 1813–1820.
35. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
36. Cathala, G., Savouret, J.F., Mendez, B., West, B.L., Karin, M., Martial, J.A., and Baxter, J.D. (1983) *DNA* **2**, 329–335.
37. Auffray, C., and Rougeon, F. (1980) *Biochem* **107**, 303–314.
38. Aviv, H., and Leder, P. (1972) *Proc. Nat. Acad. Sci. USA* **6**, 1408–1412.
39. Krumlauf, R., Holland, P.W.H., McVey, J.H., and Hogan, B.L.M. (1987) *Development* **99**, 603–617.
40. Jaffe, L., Jeannotte, L., Bikoff, E.K., and Robertson, E.J. (1990) *J. Immunol.* **145**, 3474–3482.
41. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8133.
42. Nelki, D., Dudley, K., Cunningham, P., and Akhavan, M. (1990) *Nucleic Acids Res.* **18**, 3655.
43. Mintz, E., and Russell, E.S. (1957) *J. Exper. Zool.* **134**, 207–229.
44. Hummel, K.P. (1964) *Mouse Newsletter* **34**, 31–32.
45. Bellvé, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.M., and Dym, M. (1977) *J. Cell Biol.* **74**, 68–85.
46. Yelick, P.C., Kwon, Y.K., Flynn, J.F., Borzorgzadeh, A., Kleene, K.C., and Hecht, N.B. (1989) *Mol. Reprod. and Devel.* **1**, 193–200.
47. Kennely, P.J. and Krebs, E.B. (1991) *J. Biol. Chem.* **24**, 15555–15558.
48. Conti, M., Geremia, R., and Monesi, V. (1979) *Mol. Cell. Endocrin.* **13**, 137–148.
49. Shenoy, S., Choi, J.K., Bagrodia, S., Copeland, T.D., Maller, J.L., and Shalloway, D. (1989) *Cell* **57**, 763–774.
50. Langan, T.A., Gautier, J., Lohka, M., Hollingsworth, R., Moreno, S. Nurse, P., Maller, J., and Sclafani, R.A. (1989) *Mol. Cell. Biol.* **9**, 3860–3868.
51. Chowdhury, K., Dietrich, S., Balling, R., Guenet, J-L., and Gruss, P. (1989) *Nucleic Acids Res.* **17**, 10427–10438.
52. Klevit, R.E. (1991) *Science* **253**, 1367–1393.
53. Fairall, L., Rhodes, D., and Klug, A. (1986) *J. Mol. Biol.* **192**, 577–591.