Gene Regulation During Spermatogenesis: Transcriptional and Translational Control of Phosphoglycerate Kinase 2 in Transgenic Mice

Thesis by
Murray O. Robinson

In Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1992

(defended June 18, 1991)

ACKNOWLEDGEMENTS

Of the many people I'd like to thank for making this process more worthwhile, more enlightening, more entertaining, and more fun, I will first start in Baltimore, at Johns Hopkins, with George Scangos, who provided the first inspiration in a great lab environment complete with mice. I'm especially grateful to Chuck Bieberich, who served as mentor and friend, and who transmitted the dharma of microinjection with proper superstition and karma. I'd like to thank Mel Simon, my advisor, for delivering me from the accent of Baltimore, and for allowing me the freedom to make so many mistakes, and mice. I'd also like to thank everyone in my lab, especially Andrew Pakula, Tom Wilkie and Joyce Ito, Cindy Chang and Jesus del Mazo, Kelly Thomas and Ryn Miake-Lye, et al. Outside our lab, thanks go to Chris Gomez for his tireless enthusiasm, Paul Garrity for his tutelage and patience, Jeff Miner for his friendship and inspiration, and Faiz and Kevin for years of adventures and ambition. I'd also like to thank my friends outside Caltech for providing relief from Pasadena and for keeping my priorities straight. I'd especially like to thank Marc for his friendship, perspective, and advice, not to mention countless shows from Trancas to Miami. Thanks go to Katie, Riley, Kelly, Scott and Johnna for their support and genes. Lastly, I'd like to thank Beverly for her patience, understanding, and love.

"Modern science classifies the world. . .
not into different groups of objects
but into different groups of connections. . .
The world thus appears to be a complicated tissue
of events in which connections of differents kinds
alternate or overlap or combine
and thereby
determine the structure of
the whole."

Werner Heisenberg

ABSTRACT

The work presented in this thesis is an analysis of the transcriptional and translational control of the testis-specific gene for phosphoglycerate kinase (Pgk-2) in an effort to understand the mechanisms of gene regulation during spermatogenesis.

To address transcriptional control (Chapter two), a genomic fragment containing the human gene for PGK-2 was expressed in transgenic mice, and then a 323 bp region of the promoter was shown to be necessary for tissue-specific, developmentally regulated expression.

In Chapter three, a novel quantitation technique based on the polymerase chain reaction was developed and used to measure the accumulation of transgenic and endogenous Pgk-2 transcripts. The human *PGK-2* transgene, the *PGK-2*/CAT transgene and the endogenous *Pgk-2* gene all had similar levels and patterns of expression. Transcript levels of round spermatid-expressed protamine 2 were tenfold higher and showed delayed accumulation kinetics, implying that the peak expression of the Pgk-2-promoted genes occurred in pachytene spermatocytes.

In Chapter four, the translational regulation of the Pgk-2 message was investigated, and the cis-acting sequences for proper translation of the *PGK-2* message were shown to reside in the 5' untranslated region of the gene. Transgenes containing

the *PGK-2* promoter, the CAT coding sequence, and either the *PGK-2* 3' or SV40 3' sequences were both shown to behave like the translationally regulated endogenous *Pgk-2* locus. Polysomal profiles of the *PGK-2*/CAT/SV40 3' construct demonstrated that the message shifts from a translationally inactive state to a translationally active state between *post natal* day 20 and day 33 . This developmental shift occurs with the same timing as that of the endogenous *Pgk-2* message.

The Appendix investigates the relative levels of a number of transcripts in separated spermatogenic cells. Notably, the levels of *Pgk-2* and the X-linked *Pgk-1* were analyzed during spermatogenesis. *Pgk-1* transcript levels diminish in the spermatogonial stages, whereas the *Pgk-2* transcript first appears in early meiotic cells. This pattern of gene activity is consistent with the hypothesis that *Pgk-2* expression substitutes for the inactivated *Pgk-1* locus during the later stages of spermatogenesis.

TABLE OF CONTENTS

Chapter 1

Introduction:
Gene expression during spermatogenesis

Overview

Spermatogenesis, the developmental maturation of male germ cells, is a fundamental and complex process. Common to males in most higher organisms, this process results in the formation of a highly specialized haploid cell that must package and deliver an individual's genetic information to its future progeny.

In many species, including mammals, spermatogenesis is a continuous developmental process that requires the maintenance of a stem-cell population, commitment of cells in that population to differentiation, and the subsequent developmental maturation of those committed cells. At each step of this process there must occur specific signalling events that relay information to the developing cells, signals that instruct the germ cell to proliferate, to commit, to enter meiosis, and finally to terminally differentiate. Little is known about the nature of these signalling events, and it is these signalling events that capture our interest. To learn more about the controlling events of this developmental process, one must first study the downstream events. This thesis is an attempt to begin to characterize the downstream events of spermatogenesis at the molecular level.

Spermatogenesis in mammals occurs with a readily observable spatial organization within the seminiferous tubules of the testis (Leblond and Clermont 1952; Perey et al. 1961; Roosen-Runge 1952). As cells progress through

spermatogenesis, they migrate from the outer membrane of the tubule towards the central lumen where the mature spermatozoa are released (Oakberg 1956a). Throughout this process, dynamic changes occur in the shape and structure of the developing germ cell, providing a basis for defining stages of the process based on morphological criteria (Leblond and Clermont 1952). Because of its spatially defined development and easily identifiable morphological stages, spermatogenesis has been well characterized at the cellular and the ultrastructural level.

Morphological characterization in the mouse.

Most of the morphological characterization of germ cells has been done in rodents (Clermont 1972). As the mouse is the experimental animal of this work, the process in the mouse will be described. Studies on X-irradiated mice (Oakberg 1956b; Oakberg 1957) and incorporation of tritiated thymidine into germ cells (Clermont and Trott 1969) suggested that the duration of a single cell's development is approximately 35 days. This cycle is commonly divided into three phases: the spermatogonial (or mitotic) phase, meiosis and spermiogenesis.

Spermatogonia. Because the testis is producing spermatogonia throughout the adult life of the organism, a self-renewing stem-cell population forms the progenitor pool. First observed in the seminiferous epithelium in large numbers by about day 5 after birth (Huckins 1963; Nebel et al. 1961), these

cells mature into a morphologically defined type A spermatogonium that is believed to be self-renewing (Monesi 1962; DeRooij 1973; for a review, see Bellve 1979). Following a committing step, spermatogonia progress into a dividing type B spermatogonium, undergo a finite number of divisions, then leave the mitotic cycle as preleptotene spermatocytes. The cell divisions of these mitoses are incomplete and result in cytoplasmic bridges, approximately one micron in diameter, connecting up to an estimated 128 spermatogonia (Dym and Fawcett 1971). These connections remain throughout the development of the germ cell and are thought to allow the passage of macromolecules across cells to help maintain gametic equivalence among meiotic partners (Braun et al. 1989a).

Meiosis. Lasting 13 of the 35 days of spermatogenesis (Oakberg 1956b; Oakberg 1957), cells enter meiosis as the preleptotene spermatocytes begin a long S phase of DNA replication. Concurrently, the cells migrate towards the lumen and cross the tight junctions formed by the supporting sertoli cells, passing across the blood-testis barrier (Dym and Fawcett 1970). This point is noteworthy as the germ cell now becomes isolated from components or signals in the systemic circulation, depending only on the sertoli cell for extracellular signals. The long prophase of the first meiotic division accomplishes condensation of chromatin and pairing of homologous chromosomes into the synaptonemal complex. The pachytene

stage of prophase lasts for approximately one week (Oakberg 1956b; Oakberg 1957), and the pachytene cell's longevity and large size make this period of meiosis particularly amenable to analysis. Following a rapid reduction division, each spermatocyte develops into four round spermatids.

Spermiogenesis. The haploid stages of differentiation, termed spermiogenesis, direct the developing germ cell through a dramatic morphological transformation (for a review see Clermont 1972). The early haploid spermatids (round spermatids) begin to assemble tail structures, initiating at one of the centrioles, and develop into a typical nine + two microtubule axoneme structure that will become the flagella. Concurrently, the golgi of the round spermatid begins to aggregate into a vesicular like acrosome. Next, the nucleus begins to elongate and condense and the cytoplasm begins to separate from the nucleus of the streamlining spermatid. Mitochondria migrate to the tail region and form in a helical array around the middle of the flagellum to comprise the midpiece. The cytoplasm is eventually sloughed off to become the residual body, and the mature spermatozoan is deposited into the lumen of the seminiferous tubule.

Developmental synchrony of spermatogenesis. In the mouse, spermatogenesis begins soon after birth and the first wave occurs synchronously (Bellve et al. 1977a; Nebel et al. 1961). At day 6 large numbers of type A spermatogonia are present in the seminiferous epithelium. By day 8, committed

type B spermatogonia appear. These committed germ cells progress through spermatogenesis at the same time, allowing the correlation of the age of the animal with the presence of a particular developmental stage. At day 10, the first preleptotene spermatocytes appear, and by day 13, the first meiotic cells are observed. Day 20 marks the appearance of the first haploid, round spermatids. This developmental synchrony varies somewhat among strains of mice and with the weight of the mouse. However, it has been used quite extensively as a correlative marker for biochemical analysis of spermatogenesis during development. Another technique in which the cells of the seminiferous tubule are dispersed and separated on the basis of their sedimentation rate at unit gravity has also been used extensively to characterize macromolecules at the different stages of development (Bellve et al. 1977a; Bellve et al. 1977b).

Biochemical activities during spermatogenesis.

Transcription. A number of autoradiographic studies have been performed to assay overall transcription levels throughout spermatogenesis. Pulse-chase experiments with tritiated uridine show the relative rates of transcription during development. (Monesi et al. 1978). Uridine incorporation is highest (on a per cell basis) during middle and late pachytene stages. There is a drastic drop in transcription during the meiotic cell divisions; then the rate increases slightly in round

spermatids. Midway into spermiogenesis, concurrent with elongation of the spermatid, uridine incorporation drops to undetectable levels (Kierszenbaum and Tres 1975). The cessation of transcription midway through spermiogenesis requires regulatory machinery to store and translate messages important for the terminal differentiation of spermatozoa. An important aspect of regulation in spermatogenesis, translational regulation, will be discussed in a later section.

X-chromosome inactivation. Cytological and autoradiographic evidence suggests that the single X chromosome becomes inactivated in meiotic spermatocytes. Cytologically, the X and Y chromosomes become condensed and a structure termed the sex vesicle forms around the X-Y bivalent (Solari 1970). A number of autoradiographic studies have shown that the meiotic X chromosome does not incorporate tritiated uridine (reviewed in Monesi et al. 1978). A study of the X-linked enzyme G6PD in the testis revealed that enzyme activity decreases early in development, consistent with a shutting off of the gene (Erickson 1976). On the basis of spermatogenesis defects in mice containing translocations of the X chromosome, Lifschitz and Lindsley (Lifschytz and Lindsley 1972) have proposed that X-chromosome inactivity is essential for spermatogenic development. The male sterility found in mice (and humans) containing an extra X chromosome (XXY) is one mutation cited to support this view. Despite the general acceptance of this

hypothesis, there is little direct supportive evidence. The analysis of the X-linked *Pgk-1* and *HPRT* mRNA levels during spermatogenesis (see Appendix (Singer-Sam et al. 1990)) was designed to address this question.

Translation. Experiments to assay the protein products during spermatogenesis have been carried out primarily to address the possibility of haploid gene expression, a phenomenon now well established. Studies therefore have concentrated on the differences between protein products expressed in meiotic cells versus haploid spermatids. Studies using antibodies (Millette and Bellve 1977) and using two-dimensional gel electrophoresis (Millette and Moulding 1981) have both shown the appearance of many novel proteins following meiosis. In addition to demonstrating the presence of haploid-expressed proteins, these studies enumerate the complex and distinct expression patterns of each of these stages of development.

***In vitro* cell culture attempts.**

For a number of reasons, spermatogenesis represents an excellent system for studying development and differentiation. The morphological changes in the cell are dramatic, well ordered, and well defined. Germ-cell differentiation is an ongoing, self-renewing process that occurs throughout the adult life of the animal. Development of spermatozoa proceeds through both meiotic and haploid states, and thus provides a

chance to observe and study these unique states. The unique conditions and cellular environment that supports spermatogenesis, however, has made the elucidation of its developmental pathway quite difficult. Extensive efforts (Gerton and Millette 1984; Tamaru et al. 1990; O'Brien personal comm.) at *in vitro* culture of spermatogenic cells has met with little success, so biochemical or genetic manipulations of these cells must be preformed in the animal. Only recently has a rudimentary *in vitro* transcription system derived from testis extracts been described (Bunick et al. 1990), and this system has yet to demonstrate the specificity observed *in vivo*. Naturally occurring or laboratory-induced defects in spermatogenesis have been studied in an attempt to learn more about the genetics of spermatogenesis, but information from this approach has been limited.

Genetics of spermatogenesis

The analysis of mutations is a useful tool to probe complex developmental systems. In mice, many naturally occurring and experimentally induced mutations exist, and these mutations are often indespensable in elaborating the biological process in which they are involved (see Green 1981). For spermatogenesis, however, any interesting phenotype is by definition difficult to breed; therefore, studying mutations is a very arduous task. As a result, no comprehensive search or screen for mutations has been undertaken. Of the many

mutations that affect fertility, there are at least two dozen mutations that appear to affect spermatogenesis directly (other mutations involve sex determination, development of germ-cell primordia, hormonal control of germ cell development) (Handel 1987). These mutations are almost all pleiotropic, the primary phenotype having no obvious relation to spermatogenesis. Additionally, mutations in spermatogenesis often result in underdeveloped testes, with few or no germ cells, making biochemical analysis or cell separation very difficult. The various mutations (reviewed in Handel 1987) that have been analyzed can be grouped into several loose classes: those with generalized reduction or absence of germ cells, those affecting structural development, and those with a block in development. None of these mutations have a candidate gene defect, and few have been characterized past morphological analysis.

Recently, a mutation was described that was obtained as an insertional mutation from a transgenic mouse line (MacGrevor et al. 1990). This mutation, called *symplastic spermatids* (*sys*), causes disorganization of the tubule structure, resulting in syncitial grouping of round spermatids, and the absence of further development. This mutation is promising in that it is likely affecting a single gene, has no other observed phenotype, and the presence of the transgene serves as a marker to clone the candidate gene.

Gene expression during spermatogenesis.

Since the advent and improvement of gene cloning and analysis techniques during the 1980s, the expression patterns of hundreds of genes have been characterized in mammals. An intriguing finding, one that has not been discussed formally, has been the varied and unexpected gene expression found in the testis. Two schools of thought have developed to address this finding. The prevailing view (at least among those studying spermatogenesis) is that the developmental complexity of spermatogenesis must require an equally complex and comprehensive expression of genes to direct and participate in spermatogenesis. Alternatively, some argue that the abundant gene expression in the testis is not a well-regulated process to direct differentiation, but rather a relaxation of transcriptional control, a consequence of the unique and dramatic changes the germ-cell chromatin must undergo during differentiation.

Stage-specific expression of genes has been investigated by a few researchers. One study (Kleene et al. 1983), addressing the possibility of haploid-expressed genes, found a number of genes that were expressed at least tenfold higher in haploid cells than in meiotic cells. One of these cDNAs was later found to encode the mouse protamine 1 gene (*Prm1*), which is first expressed in haploid, round spermatids (Kleene et al. 1984). Another group (Thomas et al. 1989) isolated and characterized the expression patterns of eighteen genes specific

to a testis cDNA library and found that they fell into four general stages: expression throughout spermatogenesis, expression beginning in preleptotene cells (or developmental day 12), expression initiating in pachytene cells (developmental day 16) and expression commencing in haploid cells (day 24). This grouping of expression patterns suggests coordinate regulation of genes during spermatogenesis and implies the presence of a coordinated regulatory hierarchy. Expression of many other genes has been characterized, and the majority fall into one of these groups, although nonuniformity in the methods used for characterization make it difficult to assign characterized genes to any one group.

Enzymes and structural genes.

The first genes whose products were identified to be testis-specific were those encoding assayable isozymes or structural components of spermatogenic cells. A testis-specific isozyme for phosphoglycerate kinase (PGK) was identified in 1973 (VandeBerg et al. 1973) and will be discussed later. A testis-specific activity for lactate dehydrogenase (LDH-C) was also found (Goldberg and Hawtrey 1967), resulting in its cloning and subsequent transcriptional characterization (Thomas et al. 1990).

Testis-specific gene expression has also been found in genes encoding the structural components of the highly specialized structure of the spermatozoan. The protamines and the transition proteins, small, basic proteins that replace the

histones during late spermiogenesis, are all expressed exclusively in the haploid, round spermatids (Kleene et al. 1984; Yelick et al. 1989; Yelick et al. 1987; Kleene and Flynn 1987). All of these genes appear to be subject to translational regulation, as their protein products are not observable until mid to late spermiogenesis (a subject to be covered later). Genes that are thought to provide specific structural requirements for the developing spermatozoan include a specific alpha tubulin, expressed in meiotic and haploid cells (Hecht et al. 1988) and a haploid-expressed gamma actin that appears to be expressed in smooth muscle as well (Kim et al. 1989).

Other testis-specific genes do not have an obvious reason for their localized expression, including the testis-specific cytochrome C gene, expressed in meiotic and haploid cells (Virbasius and Scarpulla 1988). It is important to note that the expression of this gene does not replace, but overlaps the expression of the somatic cytochrome C in testis. Also found is a testis-specific, heat-shock protein gene (HSP70.2) expressed in pachytene spermatocytes and round spermatids (Zakeri et al. 1988b).

An interesting, although puzzling set of genes expressed during spermatogenesis are several genes considered to have neuronal-specific functions. Nerve growth factor mRNA and protein and nerve growth factor receptor mRNA have been found in pachytene cells and round spermatids (Ayer-LeLievre

et al. 1988). Pro-opiomelanocortin (POMC) and proenkephalin, both opioid peptide precursors are expressed in meiotic cells (POMC) and meiotic and haploid cells (proenkephalin) (Kilpatrick et al. 1987; Kilpatrick et al. 1990). Transcripts for the key enzyme in the synthesis of acetylcholine, choline acetyltransferase have been found in meiotic and haploid cells in the rat and in humans, and protein has been found in mature human spermatozoa (Ibanez et al. 1991). The function of any of these gene products in male germ cells is unclear, although it is interesting to note that a number of pleiotropic mouse mutations that affect spermatogenesis (*hydrocephalic-polydactyl*, *purkinje cell degeneration*, *quaking*, and *wobbler*) also exhibit a neuronal phenotype.

Oncogenes, growth factors, and regulatory genes. A number of genes in mammals have been identified by their ability to cause neoplastic transformation. These genes, often mutated forms of the normal cellular counterpart, have been termed oncogenes. Extensive analysis of the protein products of many of these genes suggests that the normal cellular counterparts of oncogenes are important in the regulation of cell growth and differentiation. The complex stem-cell renewal and development of germ cells provide many potential regulation points for the action of proto-oncogenes. A number of studies have noted spermatogenic expression of RNA and/or protein products from many oncogenes, growth factors and transcriptional regulatory genes. These genes are of particular

interest because they may function in the control of germ-cell development.

One class of oncogenes expressed during spermatogenesis includes those that possess protein kinase activity. The *c-raf* gene, a serine/threonine protein kinase, is expressed throughout spermatogenesis, with levels peaking during the pachytene stage (Wolfes et al. 1989). Another serine/threonine protein kinase *c-mos* is expressed as a unique-sized transcript of 1.7 kb as opposed to the somatic 1.4 kb transcript (because of an alternate 5' transcription start site (Propst et al. 1987)) in haploid spermatids (Mutter and Wolgemuth 1987). *Pim-1*, a serine/threonine kinase, is also found in haploid cells as a testis-specific transcript of 2.4 kb, shorter than the somatic size of 2.8 kb (Sorrentino et al. 1988). The expression of one more tyrosine kinase gene, *c-abl* is similar to *pim-1* and *c-mos* with a testis-specific transcript of 4.7 kb appearing in mid spermiogenesis (Ponzetto and Wolgemuth 1985). The functions of these genes during spermatogenesis are unknown; however, both copies of the *pim-1* gene have been inactivated in ES cells (teRiele et al. 1990) allowing the effect of the *pim-1* gene during spermatogenesis to be evaluated in chimeric mice.

A second class of oncogenes expressed during spermatogenesis includes the nuclear proteins, known or suspected to be transcription factors. Messages for *c-fos*, *c-jun* and *c-myc* have all been found to be expressed at high levels in

type B spermatogonia, then decrease rapidly in level as cells enter meiosis (Wolfes et al. 1989). This expression pattern is consistent with the role of these proto-oncogenes in the cell cycle (Ransone and Verma 1990) and in undifferentiated cells (Cole 1986). The 5' half of the *ret* transforming gene is a zinc finger containing gene termed *rfp* (ret finger protein), and unlike *c-myc*, *c-fos* and *c-jun*, *rfp* is expressed only in haploid, round spermatids (Takahashi et al. 1988). Such apparently specific expression makes this gene a candidate for a haploid specific transcription factor. Another zinc finger protein, *Zfp-35*, has been found to be expressed at high levels during the pachytene stage (Cunliffe et al. 1990), and expression of the homeobox-containing gene *Hox 1.4* has been identified in meiotic and haploid spermatogenic cells (Wolgemuth et al. 1987). This gene and another, *int-1* (or *Wnt-1*), can be included in the group of genes described above that are expressed in the testis and neuronal cells. The *int-1* gene, whose expression has been recently implicated in the formation of midbrain and cerebellar structures (McMahon and Bradley 1990), is found only in developing neuronal tissues and midway into spermiogenesis (Shakleford and Varmus 1987).

One last class of proto-oncogenes, the *ras* family, has also been analyzed during spermatogenesis (Wolfes et al. 1989). *H-ras* and *N-ras* are expressed throughout spermatogenesis with levels peaking in pachytene cells and round spermatids,

whereas *K-ras* expression doesn't begin until the preleptotene stage. Analysis of ras proteins in the same study is consistent with the combined RNA expression patterns of the three genes.

Processed pseudogenes. Another interesting group of genes expressed during spermatogenesis are the functional processed pseudogenes. *Zfa* (Zinc finger autosomal) (Ashworth et al. 1990), pyruvate dehydrogenase E1 alpha subunit (PDHA2) (Dahl et al. 1990 and H. Dahl, pers. comm.) and phosphoglycerate kinase 2 (*Pgk2*) (McCarrey and Thomas 1987) are three of the few functional, processed pseudogenes known. All three are derived from intron-containing genes on the X-chromosome and all are expressed in meiotic and haploid cells in the testis. The intron-containing counterparts of these genes are expressed in all somatic cells (Schneider-Gadicke et al. 1989), and therefore the postulated inactivation of the X chromosome may require that the presumably essential functions of these three genes are provided for by the autosomal, processed genes. The nature of the processed gene as a reverse transcribed message makes it difficult to explain how each gene acquired a functional promoter. Perhaps the primordial processed pseudogenes were fortuitously expressed in a number of cell types, and subsequent changes causing spermatogenic inactivation of the X chromosome exerted selection pressure to maintain germ cell-specific expression.

Translational control of gene expression during spermatogenesis.

Midway through spermiogenesis, at the elongating spermatid stage, transcription of the haploid genome ceases (Kierszenbaum and Tres 1975), yet significant morphological changes must still occur within the developing spermatid. These structural changes require the ordered appearance of many proteins to complete spermiogenesis, and therefore it has been proposed that the cell must transcribe, store and translate messages required for the late stages of germ-cell development. Little is known about the mechanism or regulation of this process, although recently, evidence supporting this assertion has accumulated.

A number of genes that undergo translational regulation during spermatogenesis have been identified. The protamines and the transition proteins, all transcribed in haploid cells, are subject to translational regulation as assayed by compartmentalization of messages in polysome gradients (Kleene et al. 1984; Heideran and Kistler 1987; Kleene 1989; Yelick et al. 1989). The message for PGK-2, which first appears in meiotic spermatocytes is also translationally regulated (Gold et al. 1983), as are the messages for the testis-specific heat-shock protein HSP70.2 and the 4.7kb testis-specific *c-abl* transcript (Zakeri et al. 1988a). In the case of the protamines, some progress has been made toward the identification of a mechanism for the observed translational repression. In

transgenic mice, the 3' untranslated region of *Prm1* controls the regulated translation of a protamine/growth hormone fusion gene in spermatids (Braun et al. 1989b). Another study has identified a sequence in the 3' end of *Prm2* that is conserved among the protamines and transition proteins and has protein-binding activity (Kwon and Hecht 1991). This sequence is postulated to bind a "masking protein" that serves to inhibit the translation of *Prm2* in round spermatids. A candidate "masking protein" has also been found during spermatogenesis. A cDNA with homology to the RNA helicase, eIF 4A, is expressed only in meiotic and haploid cells, and is postulated to have some role in the regulation of translation during spermatogenesis (Leroy et al. 1989). It will be interesting to see if this or similar proteins have sequence-specific RNA binding activity, and whether the binding activity can be modulated during development.

Phosphoglycerate kinase.

Phosphoglycerate kinase (PGK) catalyzes the formation of 3-phosphoglycerate from 1,3-bisphosphoglycerate in the glycolytic pathway. Located on the X chromosome (Kozak et al. 1974), *Pgk-1* is expressed in all somatic cells and provides the essential enzymatic activity for glycolysis. First identified in 1973 as a testis-specific isozyme in the mouse (VandeBerg et al. 1973), PGK-2 protein is present in haploid, round and elongating spermatids (VandeBerg et al. 1976; Kramer 1981).

The human gene for PGK-2 was cloned in 1985 (Tani et al. 1985; McCarrey and Thomas 1987), and the mouse gene in 1987 (Boer et al. 1987). Expression analysis determined that *Pgk-2* message is first found in early meiotic spermatocytes and increases to peak expression during the late pachytene stage (Robinson and Simon 1991; Singer-Sam et al. 1990). Expression continues in round spermatids, and although *in situ* data suggest that expression ceases in late round spermatids (Goto et al. 1990 and unpublished observations), northern analysis detects the presence of *Pgk-2* message in residual bodies (Thomas et al. 1989). *In vitro* translation and PGK-2 immunoprecipitation of polysome fractions suggested that the message was subject to translational regulation (Gold et al. 1983).

Because of its tissue-specific, developmentally regulated expression, and its additional regulation at the translational level, *Pgk-2* is an excellent gene for molecular analysis. The work presented here initiates the determination of cis-acting elements that control the transcriptional and translational control of *Pgk-2*.

REFERENCES

Ashworth, A., B. Skene, S. Swift and R. Lovell-Badge. 1990. Zfa is an expressed retroposon derived from an alternative transcript of the Zfx gene. *EMBO* **9**: 1529-1534.

Ayer-LeLievre, C., L. Olson, T. Ebendal, F. Hallbook and H. Persson. 1988. Nerve growth factor mRNA and protein in the testis and epididymis of mouse and rat. *Proc. Natl. Acad. Sci. USA* **85**: 2628-2632.

Bellve, A. R. 1979. The molecular biology of mammalian spermatogenesis. *In* Oxford reviews of reproductive biology. C. A. Finn. Oxford University Press, London. pp. 159-261.

Bellve, A. R., J. C. Cavicchia, C. F. Millette, D. A. O'Brien, Y. M. Bhatnagar and M. Dym. 1977a. Spermatogenic cells of the prepuberal mouse. *J. Cell Biol.* **74**: 68-85.

Bellve, A. R., C. F. Millette, Y. M. Bhatnagar and D. A. O'Brien. 1977b. Dissociation of the mouse testis and characterization of isolated spermatogenic cells. *J. Hist. and Cyt.* **25**: 480-494.

Boer, P. H., C. N. Adra, Y. Lau and M. W. McBurney. 1987. The testis-specific phosphoglycerate kinase gene pgk-2 is a recruited retroposon. *Mol. Cell. Biol.* **7**: 3107-3112.

Braun, R. E., R. R. Behringer, J. J. Peschon, R. L. Brinster and R. D. Palmiter. 1989a. Genetically haploid spermatids are phenotypically diploid. *Nature* **337**: 373-376.

Braun, R. E., J. J. Peschon, R. R. Behringer, R. L. Brinster and R. D. Palmiter. 1989b. Protamine 3'-untranslated sequences regulate temporal control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes Dev.* **3**: 793-802.

Bunick, D., P. A. Johnson, T. R. Johnson and N. B. Hecht. 1990. Transcription of the testis-specific mouse protamine 2 gene in a homologous *in vitro* transcription system. *Proc. Natl. Acad. Sci. USA* **87**: 891-895.

Clermont, Y. 1972. Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. *Physiol. Rev.* **52**: 198-236.

Clermont, Y. and M. Trott. 1969. Duration of the cycle of the seminiferous epithelium in the mouse and hamster determined by means of tritiated thymidine and radioautoradiography. *Fert. Steril.* **20**: 805-817.

Cole, M. D. 1986. The myc oncogene: its role in transformation and differentiation. *Ann. Rev. Genet.* **20**: 361-384.

Cunliffe, V., P. Koopman, A. McLaren and J. Trowsdale. 1990. A mouse zinc finger gene which is transiently expressed during spermatogenesis. *EMBO* **9**: 197-205.

Dahl, H. H. M., R. M. Brown, W. M. Hutchison, C. Maragos and G. K. Brown. 1990. A testis-specific form of the human pyruvate dehydrogenase E1 alpha subunit is coded for by an intronless gene on chromosome 4. *Genomics* **8**: 225-232.

DeRooij, D. G. 1973. Spermatogonia stem cell renewal in the mouse. *Cell Tissue Kinet.* **6**: 281-287.

Dym, M. and D. W. Fawcett. 1970. The blood-testis barrier in the rat and the physiological compartmentation of the seminferous epithelium. *Biol. Reprod.* **3**: 308-326.

Dym, M. and D. W. Fawcett. 1971. Further observations on the numbers of spermatogonia, spermatocytes and spermatids connected by intracellular bridges in the mammalian testis. *Biol. Reprod.* **4**: 195-215.

Erickson, R. P. 1976. Glucose-6-phosphate dehydrogenase activity changes during spermatogenesis: possible relevance to X-chromosome inactivation. *Dev. Biol.* **53**: 134-137.

Gerton, G. L. and C. F. Millette. 1984. Generation of flagella by cultured mouse spermatids. *J. Cell Biol.* **98**: 619-628.

Gold, B., H. Fujimoto, J. M. Kramer, R. P. Erickson and N. B. Hecht. 1983. Haploid accumulation and translational control of phosphoglycerate kinase-2 messenger RNA during mouse spermatogenesis. *Dev. Biol.* **98**: 392-399.

Goldberg, E. and C. Hawtrey. 1967. The ontogeny of sperm-specific lactate dehydrogenase in mice. *J. Exp. Zool.* **164**: 411-418.

Goto, M., T. Koji, K. Mizuno, M. Tamaru, S. Koikeda, P. K. Nakane, N. Mori, Y. Masamune and Y. Nakanishi. 1990. Transcription switch of two phosphoglycerate kinase genes during spermatogenesis as determined with mouse testis sections *in situ*. *Exp. Cell Res.* **186**: 273-278.

Green, M. C. 1981. Genetic variants and strains of the laboratory mouse. Gustav Fischer Verlag, New York.

Handel, M. A. 1987. Genetic control of spermatogenesis in mice. *In* Spermatogenesis: Genetic Aspects. W. Hennig. Springer-Verlag, Berlin. pp. 1-62.

Hecht, N. B., R. J. Distel, P. C. Yelick, S. M. Tanhauser, C. E. Driscoll, E. Goldberg and K. S. K. Tung. 1988. Localization of a highly divergent mammalian testicular alpha tubulin that is not detectable in brain. *Mol. Cell. Biol.* **8**: 996-1000.

Heideran, M. A. and W. S. Kistler. 1987. Transcriptional and translational control of the message for transition protein 1, a major chromosomal protein of mammalian spermatids. *J. Biol. Chem.* **262**: 13309-13315.

Huckins, C. 1963. Changes in gonocytes at the time of initiation of spermatogenesis in the rat. *Anat. Rec.* **145**: 243.

Ibanez, C. F., M. Pelto-Huikko, O. Soder, E. M. Ritzen, L. B. Hersh, T. Hokfelt and H. Persson. 1991. Expression of choline acetyltransferase mRNA in spermatogenic cells results in an accumulation of the enzyme in the postacrosomal region of mature spermatozoa. *Proc. Natl. Acad. Sci. USA* **88**: 3676-3680.

Kierszenbaum, A. L. and L. L. Tres. 1975. Structural and transcriptional features of the mouse spermatid genome. *J. Cell Biol.* **65**: 258-270.

Kilpatrick, D. L., K. Borland and D. F. Jin. 1987. Differential expression of opioid peptide genes by testicular germ cells and somatic cells. *Proc. Natl. Acad. Sci. USA* **84**: 5695-5699.

Kilpatrick, D. L., S. A. Zinn, M. Fitzgerald, H. Higuchi, S. L. Sabol and J. Meyerhardt. 1990. Transcription of the rat and mouse

proenkephalin genes is initiated at distinct sites in spermatogenic and somatic cells. *Mol. Cell. Biol* **10**: 3717-3726.

Kim, E., S. H. Waters, L. E. Hake and N. B. Hecht. 1989. Identification and developmental expression of a smooth muscle gamma actin in postmeiotic male germ cells of mice. *Mol. Cell. Biol.* **9**: 1875-1881.

Kleene, K. C. 1989. Poly(A) shortening accompanies the activation of translation of five mRNAs during spermatogenesis in the mouse. *Development* **106**: 367-373.

Kleene, K. C., R. J. Distel and N. B. Hecht. 1983. cDNA clones encoding cytoplasmic poly A+ RNAs which first appear at detectable levels in haploid phases of spermatogenesis in the mouse. *Dev. Biol.* **98**: 455-464.

Kleene, K. C., R. J. Distel and N. B. Hecht. 1984. Translational regulation and deadenylation of a protamine mRNA during spermatogenesis in the mouse. *Dev. Biol.* **105**: 71-79.

Kleene, K. C. and J. F. Flynn. 1987. Characterization of a cDNA clone encoding a basic protein, TP2, involved in chromatin condensation in the mouse. *J. Biol. Chem.* **262**: 17272-17277.

Kozak, L. P., G. K. McLean and E. M. Eicher. 1974. X-linkage of phosphoglycerate kinase in the mouse. *Biochem. Genet.* **11**: 41-47.

Kramer, J. M. 1981. Immunofluorescent localization of PGK-1 and PGK-2 isozymes within specific cells of the mouse testis. *Dev. Biol.* **87**: 30-36.

Kwon, Y. K. and N. B. Hecht. 1991. Cytoplasmic protein binding to highly conserved sequences in the 3' untranslated region of mouse protamine 2 mRNA, a translationally regulated transcript of male germ cells. *Proc. Natl. Acad. Sci. USA* **88**: 3584-3588.

Leblond, C. P. and Y. Clermont. 1952. Definition of the stages of the cycle of the seminiferous epithelium in the rat. *Am. J. Anat. Sci.* **55**: 548-573.

Leroy, P., P. Alzari, D. Sassoon, D. Wolgemuth and M. Fellous. 1989. The protein encoded by a murine male germ cell-specific transcript is a putative ATP-depedent RNA helicase. *Cell* **57**: 549-559.

Lifschytz, E. and D. F. Lindsley. 1972. The role of X-chromosome inactivation during spermatogenesis. *Proc. Natl. Acad. Sci. USA* **69**: 182-186.

MacGrevor, G. R., L. D. Russell, M. E. VanBeek, G. R. Hanten, M. J. Kovak, C. A. Kozak, M. L. Meistrich and P. A. Overbeek. 1990. Symplastic spermatids (sys): a recessive insertional mutation in mice causing a defect in spermatogenesis. *Proc. Natl. Acad. Sci. USA* **87**: 5016-5020.

McCarrey, J. R. and K. Thomas. 1987. Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. *Nature* **326**: 501-505.

McMahon, A. P. and A. Bradley. 1990. The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**: 1073-1085.

Millette, C. F. and A. R. Bellve. 1977. Temporal expression of membrane antigens during mouse spermatogenesis. *J. Cell Biol.* **74**: 86-97.

Millette, C. F. and C. T. Moulding. 1981. Cell surface marker proteins during mouse spermatogenesis: two dimensional electrophoretic analysis. *J. Cell Sci.* **48**: 367-382.

Monesi, V. 1962. Autoradiographic study of DNA synthesis and the cell cycle in spermatogonia and spermatocytes of mouse testis using tritiated thymidine. *J. Cell Biol.* **14**: 1-18.

Monesi, V., R. Geremia, A. D'Agostina and C. Boitani. 1978. Biochemistry of male germ cell differentiation in mammals: RNA synthesis in meiotic and postmeiotic cells. *Curr. Top. Dev. Biol.* **12**: 11-36.

Mutter, G. L. and D. J. Wolgemuth. 1987. Distinct developmental patterns of c-mos proto-oncogene expression in female and male mouse germ cells. *Proc. Natl. Acad. Sci. USA* **84**: 5301-5305.

Nebel, B. P., A. P. Amarose and E. M. Hackett. 1961. Calendar of gametogenic development in the prepuberal male mouse. *Science* **134**: 832-833.

Oakberg, E. F. 1956a. A description of spermatogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. *Am. J. Anat.* **99**: 391-413.

Oakberg, E. F. 1956b. Duration of spermatogenesis in the mouse and timing of stages of the cycle of the seminiferous epithelium. *Am J. Anat.* **99**: 507-516.

Oakberg, E. F. 1957. Duration of spermatogenesis in the mouse. *Nature* **180**: 1137-1138.

Perey, B., Y. Clermont and C. P. Leblond. 1961. The wave of the seminiferous epithelium in the rat. *Am J. Anat.* **108**: 47-77.

Ponzetto, C. and D. J. Wolgemuth. 1985. Haploid expression of a unique c-abl transcript in the mouse male germ line. *Mol. Cell. Biol.* **5**: 1791-1794.

Propst, F., M. P. Rosenberg, A. Iyer, K. Kaul and G. F. VandeWoude. 1987. c-mos proto-oncogene RNA transcripts in mouse tissues: structural features, developmental regulation, and localization in specific cell types. *Mol. Cell. Biol.* **7**: 1629-1637.

Ransone, L. J. and I. M. Verma. 1990. Nuclear proto-oncogenes fos and jun. *Ann. Rev. Cell Biol.* **6**: 128-146.

Robinson, M. O. and M. I. Simon. 1991. Determining transcript number using the polymerase chain reaction: Pgk-2, mP2, and PGK-2 transgene mRNA levels during spermatogenesis. *Nucl. Acids Res.* **19**: 1557-1562.

Roosen-Runge, E. C. 1952. Kinetics of spermatogenesis in mammals. *Ann. N. Y. Acad. Sci.* **55**: 574-584.

Schneider-Gadicke, A., P. Beer-Romero, L. G. Brown, G. Mardon, S. Luoh and D. C. Page. 1989. Putative transcription activator

with alternative isoforms encoded by human ZFX gene. *Nature* **342**: 708-711.

Shakleford, G. M. and H. E. Varmus. 1987. Expression of the proto-oncogene int-1 is restricted to postmeiotic male germ cells and the neural tube of mid-gestational embryos. *Cell* **50**: 89-95.

Singer-Sam, J., M. O. Robinson, A. R. Bellve, M. I. Simon and A. D. Riggs. 1990. Measurement by quantitative PCR of changes in HPRT, PGK-1, PGK-2, APRT, MTase and Zfy gene transcripts during mouse spermatogenesis. *Nucl. Acids Res.* **18**: 1255-1259.

Solari, A. J. 1970. The spatial relationship of the X and Y chromosomes during meiotic prophase in mouse spermatocytes. *Chromosoma* **29**: 217-236.

Sorrentino, V., M. D. McKinney, M. Giorgi and R. Geremia. 1988. Expression of cellular proto-oncogenes in the mouse male germ line: a distinctive 2.4 kb pim-1 transcript is expressed in haploid postmeiotic cells. *Proc. Natl. Acad. Sci. USA* **85**: 2191-2195.

Takahashi, M., Y. Inaguma, H. Hiai and F. Hirose. 1988. Developmentally regulated expression of a human finger

containing gene encoded by the 5' half of the ret transforming gene. *Mol. Cell. Biol.* **8**: 1853-1856.

Tamaru, M., Y. Nagao, M. Taira, M. Tatibana, Y. Masamune and Y. Nakanishi. 1990. Selective activation of testis-specific genes in cultured rat spermatogenic cells. *Biochem. Biophys. Acta* **1949**: 331-338.

Tani, K., J. Singer-Sam, M. Munns and A. Yoshida. 1985. Molecular cloning and structure of an autosomal processed gene for human phosphoglycerate kinase. *Gene* **35**: 11-18.

teRiele, H., E. R. Maandag, A. Clarke, M. Hooper and A. Berns. 1990. Consecutive inactivation of both alleles of the pim-1 proto-oncogene by homologous recombination in embryonic stem cells. *Nature* **348**: 649-651.

Thomas, K., J. DelMazo, P. Eversole, A. Bellve, Y. Hiraoka, S. S. L. Li and M. Simon. 1990. Developmental regulation of expression of the LDH multigene family during mouse spermatogenesis. *Development* **109**: 483-493.

Thomas, K. H., T. W. Wilkie, P. Tomashefsky, A. R. Bellve and M. I. Simon. 1989. Differential gene expression during mouse spermatogenesis. *Biol. Reprod.* **41**: 729-739.

VandeBerg, J. L., D. W. Cooper and P. J. Close. 1973. Mammalian testis phosphoglycerate kinase. *Nature New Biol.* **243**: 48-50.

VandeBerg, J. L., D. W. Cooper and P. J. Close. 1976. Testis specific phosphoglycerate kinase B in mouse. *J. Exp. Zool.* **198**: 231-240.

Virbasius, J. V. and R. C. Scarpulla. 1988. Structure and expression of rodent genes encoding the testis-specific cytochrome c. *J. Biol. Chem.* **263**: 6791-6796.

Wolfes, H., K. Kogawa, C. F. Millette and G. M. Cooper. 1989. Specific expression of nuclear proto-oncogenes before entry into meiotic prophase of spermatogenesis. *Science* **245**: 740-743.

Wolgemuth, D. J., C. M. Viviano, E. Gizang-Ginsberg, M. A. Frohman, A. L. Joyner and G. R. Martin. 1987. Differential expression of the mouse homeobox-containing gene Hox 1.4 during male germ cell differentiation and embryonic development. *Proc. Natl. Acad. Sci. USA* **84**: 5813-5817.

Yelick, P. C., R. Balhorn, P. A. Johnson, M. Corzett, J. A. Mazrimas, K. C. Kleene and N. B. Hecht. 1987. Mouse protamine 2 is synthesized as a precursor, whereas mouse protamine 1 is not. *Mol. Cell. Biol.* **7**: 2173-2179.

Yelick, P. C., Y. K. Kwon, J. F. Flynn, A. Borzorgzadeh, K. C. Kleene and N. B. Hecht. 1989. Mouse transition protein 1 is translationally regulated during the postmeiotic stages of spermatogenesis. *Mol. Reprod. Develop.* **1**: 193-200.

Zakeri, Z. F., C. Ponzetto and D. J. Wolgemuth. 1988a. Translational regulation of the novel haploid-specific transcripts for the c-abl proto-oncogene and a member of the 70 kDa heat-shock protein gene family in the male germ line. *Dev. Biol.* **125**: 417-422.

Zakeri, Z. F., D. J. Wolgemuth and C. R. Hunt. 1988b. Identification and sequence analysis of a new member of the mouse HSP70 gene family and characterization of its unique cellular and developmental pattern of expression in the male germ line. *Mol. Cell. Biol.* **8**: 2925-2932.

Chapter 2

Transcriptional regulatory regions of testis-specific *PGK-2* defined in transgenic mice

(This chapter was published in *Proceedings of the National Academy of Sciences, USA*, **86**: 8437-8441, 1989)

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8437–8441, November 1989
Developmental Biology

Transcriptional regulatory regions of testis-specific *PGK2* defined in transgenic mice

(spermatogenesis/translational regulation/luciferase)

MURRAY O. ROBINSON*, JOHN R. MCCARREY†, AND MELVIN I. SIMON*

*Division of Biology, California Institute of Technology, Pasadena, CA 91125; and †Division of Reproductive Biology, The Johns Hopkins University School of Public Health, Baltimore, MD 21205

Contributed by Melvin I. Simon, August 7, 1989

ABSTRACT **The gene encoding testis-specific phosphoglycerate kinase 2 (PGK; ATP:3-phospho-D-glycerate 1-phosphotransferase, EC 2.7.2.3) is expressed only in meiotic and haploid male germ cells. Transgenic mice containing an 8-kilobase human genomic *PGK2* gene express the human gene in a tissue-specific and developmentally regulated manner. To determine the nature and location of sequences controlling this expression, transgenic mice with various lengths of the human *PGK2* 5′ region fused to the chloramphenicol acetyltransferase (CAT) gene were analyzed for expression. A 323-base-pair region 5′ to the coding region was found to contain information essential for both tissue-specific and developmentally regulated expression of the CAT reporter gene. Transgenic mice containing a *PGK2*/luciferase-coding construct were compared with mice containing an equivalent CAT construct. Luciferase gene expression was also testis-specific and was more sensitive than CAT gene expression, but otherwise regulation of the two reporter genes was similar in the germ cells of transgenic mice. Translation of both *PGK2*/CAT and *PGK2*/luciferase fusion genes was seen concurrently with the first detectable transcripts.**

Spermatogenesis involves a complex developmental program including the progression of cells through mitotic, meiotic, and haploid stages (for reviews, see refs. 1 and 2). Mitotically dividing spermatogonia enter meiosis as spermatocytes and then go through two cell divisions to become round spermatids. These haploid cells undergo elongation, nuclear condensation, and other morphological changes, resulting in mature spermatozoa.

In the mouse, spermatogenesis is an ongoing process that begins soon after birth. Initial appearance of spermatogenic cell types occurs synchronously, and the sequence of events can be followed as a function of increasing age in a variety of strains of inbred mice (3); At day-9 or -10 *post partum*, spermatogonia first enter meiosis, by day-14 or -15 the first pachytene spermatocytes are seen, and by day-22, haploid round spermatids appear (3). The precise regulation of germ-cell differentiation requires a controlled program of stage-specific gene expression. To examine this process at the molecular level, the mechanisms involved in the transcriptional control of germ-cell-specific genes must be understood. A well-characterized gene product expressed during spermatogenesis and, therefore, a good candidate for molecular analysis is the testis-specific isozyme for the glycolytic enzyme phosphoglycerate kinase (PGK) (ATP:3-phospho-D-glycerate 1-phosphotransferase, EC 2.7.2.3) (4–8).

In mammals PGK is encoded by two genes, *Pgk-1* and *Pgk-2* (*PGK1* and *PGK2* in human). *Pgk-1* is X-chromosome-linked and expressed in somatic cells (6, 9, 10), whereas the autosomal *Pgk-2* gene is expressed only in meiotic and haploid male germ cells (11, 12). *Pgk-2* transcription is first seen in pachytene spermatocytes, with message levels increasing during later stages of spermatogenesis (12–14). PGK-2 protein is not detected until 7 days later in haploid round spermatids (6–8) and is coincident with the appearance of *Pgk-2* message on polysomes (12), suggesting additional regulation at the translational level. Accordingly, analysis of cis-acting regulatory elements of *Pgk-2* may provide a basis for understanding the control of gene expression during spermatogenesis.

Human (13, 15, 16) and mouse (17) *Pgk-2* genomic sequences have been cloned, and both genes lack introns and share extensive homology. The sequences 5′ to the coding region are 66% homologous (17), and both genes exhibit similar promoter motifs, consisting of the consensus sequences for the SP1 transcription factor (GC box) and for nuclear factor 1 (CAAT box) (18, 19) (Fig. 1), rather than the TATA consensus sequence found adjacent to the start of the coding region in many genes (21). Given the reasonable extent of homology, the ability to compare mouse and human sequences for potential control elements, and the ease of detecting the human gene in a mouse background, we chose to define the extent of sequences required for appropriate expression of the human *PGK2* gene. We used transgenic mice to assay for cis-acting elements that control stage and cell type-specific gene expression.

Here we show that a human *PGK2* gene fragment is appropriately expressed in the mouse and that all or part of a 323-base-pair (bp) region 5′ to the coding region of the human *PGK2* gene is necessary to direct testis-specific expression of reporter genes.

MATERIALS AND METHODS

Gene Constructs. An 8-kilobase (kb) human genomic *PGK2* *Eco*RI fragment (Fig. 2*A*). that contains the *PGK2* coding sequence flanked by 2.5 kb 5′ and 4.25 kb 3′ to the gene was isolated for microinjection from a genomic clone generated by Szabo *et al.* (15). The *PGK2* 5′ region was isolated by subcloning a 2.2-kb *Sst* I fragment containing 1.4 kb of the 5′ region and the first 0.8 kb of the coding region into mp18, converting the sequences across the initiation codon into a *Hin*dIII site by site-directed mutagenesis, and isolating the resulting 1.4-kb *Hin*dIII fragment. The *PGK2*/chloramphenicol acetyltransferase (CAT) fusion gene was constructed by inserting the 1.4-kb fragment into the unique *Hin*dIII site of the plasmid pSVOCAT (22). This *PGK2*/CAT plasmid was digested with either *Xba* I, *Pst* I, or *Ava* II at the 5′ end and *Bam*HI at the 3′ end, and the resulting 3.1-kb, 2.2-kb, and 1.9-kb *PGK2*/CAT fragments (Fig. 2*B*) were each isolated for microinjection. The *PGK2*/luciferase fusion gene (Fig. 2*C*) was constructed by inserting the 1.4-kb *Hin*dIII 5′ region

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*"

Abbreviations: PGK-2, phosphoglycerate kinase 2; CAT, chloram-

```
     PstI▼
-520 CTGCAGAGGT TTTTACATAT CAAAATGGTT AAGATTTGAC ATGAATGAGG
-470 TGTATGTAGG TTTGCGGGGG TGGGGGTGGA GTTCTTTTTG TTTTGTTTTT
-420 AATAATAAAA GCAACTGTTA ACCGAGCTGT GGGGTGGGGG CAAAAAGAGC
-370 CAGAAGCGGC GCACACCTCA GGACTATTCT TGTTTTTTTA GAACCATTCT
-320 TATTCTGGGG TTTCTTCACC TACCCAAGTC TCGCCTGAAG CCAGGTACAG
-270 CTCTATTCCA CTACATGACC CTCTGCCCAG GAAGTTGGAA TCTTCACCTA
                          AvaII▼
-220 GCAACACAGT TCAGATCGAG ATTGACAGGA CCATGAGCCA ATCACAAAGC
-170 TAGATTTGCT TTCAGTCTAA CAGTGGCCGT TGTGCTGGAG ACAGTAGGAG
-120 AAGAAAGGGG CGGGACAAGG GCAAAGGCGT TAGAAGTCAC CACCGACCCA
 -70 GCCCCTCAAC AGCAAGTTGG TTCTTCAGCA TTAAGATCCA GGTGTCAGCC
                      +1
 -20 TATGTCTTTA TATTGTCAAG ATG
```

FIG. 1. The nucleotide sequence of the 5′ region of human *PGK2*. Nucleotides are numbered relative to the first nucleotide of the translation start codon (+1) and extend from the *Pst* I site at −515 bp relative to the initiation codon. The CAAT consensus sequence (position −184 bp) and the GC consensus sequence (position −114 bp) are boxed. Restriction sites used for transgenic constructs are indicated. The sequence was taken from McCarrey (20).

fragment into the unique *Hin*dIII site of the plasmid pSVOAL (23) and isolating the 3.5-kb *Xba* I–*Bam*HI fragment for microinjection.

Transgenic Mice. Each of the five constructs was microinjected into fertilized eggs from superovulated (C57BL6J/DBA2J) F_1 females mated to males of the same cross. Surviving eggs were then reimplanted into pseudopregnant (C57BL6J/DBA2J) F_1 females. Microinjection was performed essentially as described (24). Founder transgenic mice were identified by Southern blot hybridization using a 500-bp human *PGK2* ^{32}P-labeled probe that spans the initiation codon. Subsequent generations were identified by either dot-blot hybridization or amplification of a 780-bp *PGK2*/CAT fragment from 0.5 μg of genomic DNA using the polymerase chain reaction (25).

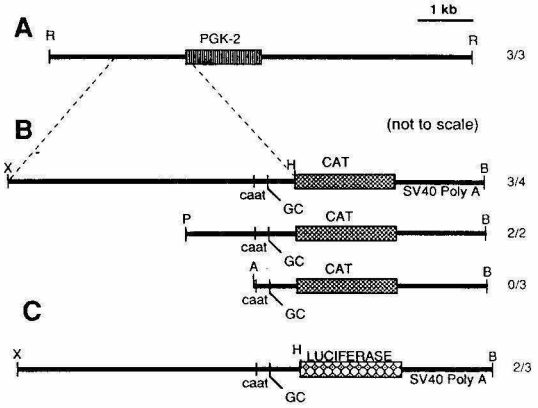


FIG. 2. Constructs injected into fertilized mouse embryos. (*A*) Structure of the genomic human *PGK2* gene *Eco*RI fragment. This 8-kb fragment contains 2.5 kb 5′ to the ATG initiation codon, a coding region of 1.25 kb contained within a single exon, and 4.25 kb 3′ to the gene. (*B*) *PGK2*/CAT fusion constructs. 1.4 kb of the human *PGK2* upstream region was placed into the *Hin*dIII site of the plasmid pSVOCAT, which contains the gene for CAT and simian virus 40 (SV40) poly(A) addition sequences. Restriction enzymes used to isolate each fragment are shown, as are positions of the CAAT consensus sequence (caat) and the SP1 binding-site consensus sequence (GC). (*C*) Structure of the *PGK2*/luciferase gene. The 1.4-kb promoter region was inserted into the *Hin*dIII site of pSVOAL that contains the luciferase-coding region and the simian virus 40 poly(A) addition sequences. The number of lines expressing the transgene versus the number of lines tested is shown at right. A, *Ava* II; B,

RNA Analysis. Total RNA was extracted by tissue homogenization with 6 M urea/3 M LiCl, as described (26). Northern (RNA) analysis was done as described (27). The human message was analyzed by using the same probe as that used for Southern analysis. The endogenous mouse message was assayed using a ^{32}P-labeled probe from the 3′ untranslated region of the mouse *Pgk-2* gene.

CAT Enzyme Assays. CAT assays were performed as described (28). Protein extracts were prepared from tissues of hemizygous male mice, and 30 μg of total protein was incubated for 15 min to quantitate activity. The limit of detection of CAT activity was determined by assaying serial dilutions of commercially obtained CAT enzyme (Boehringer Mannheim) of known activity. Calculations of specific activity and conversion to molecules of enzyme were based on a native M_r of 80,000 (29) and a specific activity of 100,000 units/mg (Boehringer Mannheim), such that 1 unit of activity corresponded to 6.2×10^{10} molecules of active tetramer. One unit was defined as the conversion of 1 nmol of [^{14}C]chloramphenicol to the 3′-acetyl form in 1 min at 25°C.

Luciferase Enzyme Assays. Tissues were homogenized in a Dounce homogenizer in phosphate-buffered saline, pH 8.0, cellular debris was pelleted for 5 min at 12,000 × *g*, and supernatants were assayed immediately. Extracts (10 μg of total protein) were assayed in 100 mM Tris·HCl, pH 7.5/15 mM $MgSO_4$, 5 mM ATP; 1 mM D-luciferin was injected immediately before recording with a Monolight 2001 luminometer (Analytical Luminescence Laboratory, San Diego, CA). Peak light emission was measured, as it is both sensitive and linear over several orders of magnitude. Calculations of enzyme activity and number of luciferase molecules were determined by using an M_r of 60,000 (23). Specific activity and limit of detection measurements were determined based upon serial dilutions of a solution of crystallized firefly luciferase (Boehringer Mannheim) into testis extracts of C57BL6J/DBA2J nontransgenic mice. It was determined that 10^9 luminometer units (lu) were equivalent to 1 mg of active luciferase protein, and it was then calculated that one lu corresponds to 5×10^5 luciferase molecules.

RESULTS

Expression of Endogenous and Human *PGK2* in Transgenic Mice. To determine the pattern of expression of the human *PGK2* gene in transgenic mice, an 8-kb *Eco*RI genomic fragment was purified and injected into single-cell fertilized mouse embryos. This genomic fragment includes the entire *PGK2* coding sequence flanked by 2.5 kb upstream and 4.25 kb 3′ to the gene (Fig. 2*A*). Six founder mice were identified by hybridization to a human *PGK2* probe. Lines were established from three of the six founder mice, and each was subsequently analyzed.

Murine *Pgk-2* expression has been shown to be testis-specific (4, 6) and developmentally regulated (7, 11, 14). To determine the tissue specificity of the human *PGK2* transgene, total RNA was prepared from various tissues of adult males from each of the three transgenic lines (designated P4, P14, and P38), and expression of the transgene was analyzed using Northern (RNA) blots. Expression of the human *PGK2* RNA was testis-specific in all three lines tested. Line P14 was bred to obtain animals that were homozygous for human *PGK2*, and expression of the transgene in this line was followed (Fig. 3*A*). Testes from homozygous mice contained approximately twice the amount of transgene RNA per μg of total RNA as the hemizygous testis, consistent with regula-

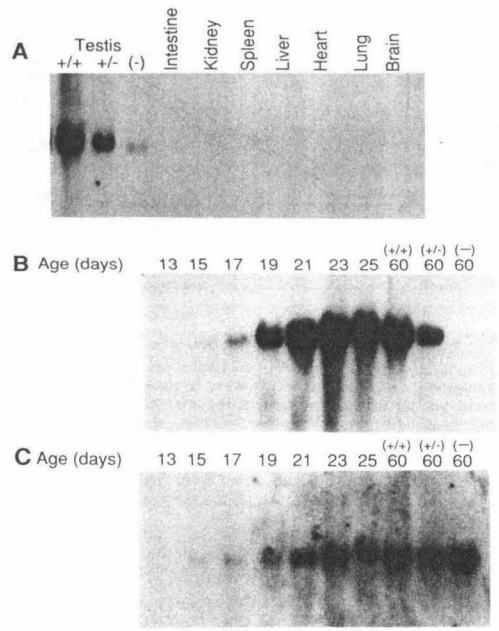


FIG. 3. Expression of *PGK2* in transgenic mice. (*A*) Northern analysis of various tissues from transgenic line P14. Total RNA (20 μg) from each of the tissues indicated was loaded per lane and probed with human *PGK2* sequences that share 70% homology to the mouse *Pgk-2* message. The faint band in the nontransgenic (−) testis lane is presumed to be crosshybridization to the endogenous mouse *Pgk-2* message. (*B*) Developmental expression of human *PGK2* in homozygous transgenic line P14. Total RNA was prepared from testes of line P14 mice sacrificed between day-11 and day-60 (date of birth = day-0). Twenty micrograms of total RNA per lane from the days indicated was loaded and hybridized to the same probe as in *A*. (*C*) Developmental expression of mouse *Pgk-2*. The filter in *B* was stripped and reprobed with mouse *Pgk-2* 3′ untranslated sequences. This probe is 50% homologous to the human *PGK2* sequence.

lines tested, however, there was no observed correlation between the number of integrated copies of the human gene and the level of human *PGK2* message (unpublished results, M.O.R.).

To determine the time during onset of spermatogenesis when the human *PGK2* transcript first appeared, RNA was prepared from testes of line P14 homozygous mice between day-11 and day-60. Transgene RNA was first detected at day-15 and increased to a high level by day-60 (Fig. 3*B*). Day-15 typically correlates with the appearance of pachytene spermatocytes (3). The presence of cytologically recognizable pachytene cells in this line at day-15 was confirmed by direct microscopic examination of dissociated testis cells. To determine whether the human transgene showed the same pattern of developmental expression as the endogenous mouse *Pgk-2* gene, the Northern blot was hybridized to a mouse 3′ specific probe at high stringency such that no crosshybridization to the human transgene was seen. The pattern of expression of mouse *Pgk-2* was similar to the pattern seen with the human probe (Fig. 3*C*). Therefore, it was concluded that expression of the human *PGK2* transgene is regulated in a manner similar to that of the endogenous *Pgk-2* gene.

Promoter Analysis of Human *PGK2*. Having established that the transgene functions in a tissue-specific and developmentally regulated manner, various deletions of the 5′ noncoding region were tested to map the cis-acting elements involved in the transcriptional regulation of *PGK2*. A 1.4-kb *Hin*dIII fragment containing the promoter region of *PGK2* that extends to, but does not include, the translation initiation codon, was inserted into a unique *Hin*dIII site in pSVOCAT (22). The resulting *PGK2*/CAT fusion gene places the CAT gene under the control of the sequences 5′ to the human *PGK2* coding region and uses the 3′ untranslated region of the simian virus 40 large tumor antigen gene to provide poly(A) addition sequences. This construct was then used to generate three different injection fragments, each containing the CAT gene with different lengths of the *PGK2* 5′ region. Fragment lengths of 1.4 kb, 515 bp, and 192 bp were generated by digesting the 5′ end of the fusion gene with *Xba* I, *Pst* I, or *Ava* II, respectively (Fig. 2*B*). The 192-bp site was chosen because it marked the 5′ end of the CAAT consensus sequence, which functions as a transcriptional activator in many cell types (18) and was, therefore, considered to be the boundary of the minimal promoter.

Expression of *PGK2*/CAT Transgenes. Thirty mice born from the 1.4-kb *PGK2*/CAT injections were analyzed using Southern blot; six founder mice were identified, and lines were derived from four of the founders (designated B4, B17, B22, and B29). Two lines of mice (PS13 and PS28) were established from the 515-bp *PGK2*/CAT construct, and three lines (AV25, AV36, and AV41) were made from the 192-bp *PGK2*/CAT fusion gene. Tissues of adult males from each of the nine lines were tested for CAT activity by thin-layer chromatography, and activity was measured using a scintillation counter (Table 1). Tissues assayed were brain, heart, lung, liver, kidney, intestine, muscle, and testis. Testis-specific expression of the transgene was seen in three lines of the 1.4-kb fragment construct (B17, B22, and B29), whereas line B4 did not express the transgene. Both transgenic lines containing the 515-bp *PGK2* fragment also had CAT activity restricted to the testes. By contrast, none of the three lines containing the 192-bp fragment constructs exhibited CAT activity in any tissue assayed. These results show that, although efficient tissue specific expression could be directed with 515 bp of upstream *PGK2* sequences, expression was eliminated by removing 323 bp between position −515 relative to the initiation codon and the minimal promoter elements that begin at −192 bp.

Table 1. CAT activity in testes of transgenic PGK-2/CAT lines

	1.4-kb promoter				515-bp promoter		192-bp promoter		
	B4	B17	B22	B29*	PS13	PS28	AV25	AV36	AV41
CAT activity†	<0.03	1.04	0.59	1.37	0.44	0.78	<0.03	<0.03	<0.03
Tissue expressing enzyme	None	T	T	T	T	T	None	None	None

T, testes. B4–AV41 are transgenic lines.

*Line B29 expressed the following CAT activity, in milliunits (mU) per μg, in testis: day-11, <0.03 mU; day-13, 0.05 mU; day-14, 0.08 mU; day-15, 0.15 mU; and day-16, 0.14 mU.

†Expressed as milliunits per μg.

To assay for the presence of regulatory elements between 1.4 kb and 515 bp that might affect the level of expression, CAT activity was quantitated for all lines of mice carrying these fragments. Constructs containing the 1.4-kb fragment had ≈2-fold greater mean activity than those with the 515-bp fragment (Table 1), a difference in level of activity that has been seen between lines of a single construct. Therefore, no cis-acting elements appear to exert any additional effect in the region between the 1.4-kb fragment and the 515-bp fragment.

The developmental regulation of the *PGK2*/CAT fusion genes was assayed by monitoring the appearance of CAT enzyme activity in extracts prepared from hemizygous testes of line B29 (1.4-kb fragment) at days 11, 13, 14, 15, and 16. CAT activity was first detected at a low level in extracts from day-13 testes (footnote to Table 1), indicating that *PGK2*/CAT transcript is present at day-13 and is likely to coincide with the first appearance of pachytene spermatocytes. Line PS28 (515-bp fragment) was also analyzed with respect to developmental appearance of CAT activity, and a similar pattern of expression was seen (data not shown). Previous experiments have shown that mouse PGK-2 protein first appears in spermatids (7, 12). However, CAT protein is present before the appearance of spermatids, and developmental Northern blot analysis of line B29 first detects *PGK2*/CAT message at day-15 (unpublished work, M.O.R.). This result shows that the transgene is not only transcriptionally active, but also competent for translation of the CAT reporter gene.

Luciferase Expression in the Testis. To increase sensitivity and to determine whether the promoter could function independently of the reporter gene, a parallel study was conducted with transgenic mice in which the 1.4-kb upstream region of *PGK2* was fused to the firefly luciferase gene. The 1.4-kb *PGK2* *Hin*dIII fragment was inserted into the unique *Hin*dIII site of pSVOAL (23), which contains the luciferase-coding region and simian virus 40 poly(A) addition sequences, and the resultant fusion gene (Fig. 2*C*) was isolated and injected into mouse embryos. Three founder mice were identified, and breeding lines (PL4, PL5, and PL12) were established for each.

Expression of the *PGK2*/luciferase fusion gene was examined by making protein extracts from various tissues of hemizygous adult males of each line and assaying each sample (10 μg of total protein) in a luminometer for luciferase activity (Table 2). In the presence of the substrate D-luciferin and ATP, luciferase emits a transient peak of yellow light with a decaying bioluminescence lasting several minutes (30). *PGK2*/luciferase expression, measured by luciferase-catalyzed peak photon production, was found to be testis-specific. Testis extracts prepared from lines PL4 and PL12 gave signals of 1400 and 1600 lu per μg, respectively, whereas all other tissues assayed (brain, heart, lung, liver, kidney, intestine, muscle, and ovary) gave background readings (<10 lu). Line PL5 did not express the transgene in any tissue assayed. To standardize the luciferase activity assay, aliquots of commercially prepared pure firefly luciferase were tested in the same luminometer. Peak light emissions in the linear range were recorded from 500 pg–50 fg; 50 fg corresponded to 10 lu, the limit of reliable detection in the luminometer. Adding testis extracts from wild-type mice to pure luciferase did not affect the activity, suggesting that no inhibitory effect is present in the testis extract.

Table 2. Luciferase activity in transgenic *PGK2*/luciferase lines

Transgenic line	PL4	PL5	PL12*
Luciferase activity†	1400	<10	1600
Tissue expressing enzyme	T	None	T

T, testes.

*Line PL12 expressed the following luciferase activity, in lu per μg of protein, in testis: day-9, <10 lu; day-12, <10 lu; day-13, 49 lu; and day-15, 149 lu.

†Expressed as lu per μg.

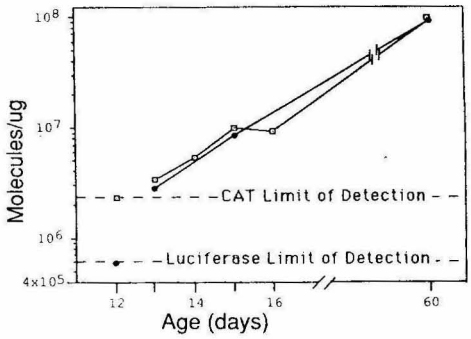


FIG. 4. Comparison of CAT and luciferase protein levels during development. Enzyme activities were converted to molecules of protein per μg of total protein in the extract to allow for direct comparison. Molecules of reporter protein are plotted as a function of the age of the mouse from which the protein extract was prepared. Limit of detection for each respective reporter protein is indicated by a dotted line. □, CAT protein level from line B29; ●, luciferase protein level from line PL12. All points represent single determinations.

Developmental appearance of luciferase protein was examined for transgenic line PL12. Analysis of extracts of prepuberal testes from males of this line on days 9, 12, 13, and 15 revealed luciferase activity first appearing at day-13 (footnote to Table 2), consistent with the appearance of CAT activity in lines B29 and PS28. Apparently *PGK2*/luciferase message is also competent for translation in pachytene cells.

To directly compare the sensitivity and accumulation of each reporter enzyme, CAT and luciferase activities were converted to molecules of enzyme by using published and determined specific activities. Serial dilutions of commercially obtained CAT enzyme into nontransgenic testis extracts established the limit of detection at 0.03 milliunits, or 2×10^6 molecules of CAT tetramer. The limit of detection of luciferase enzyme was determined to be 10 lu or 5×10^5 molecules of luciferase protein, a 4-fold increase in sensitivity over CAT detection levels. Plotting the activities of reporter molecules in the transgenic lines B29 (CAT) and PL12 (luciferase) as a function of development (Fig. 4) demonstrates that the reporter genes function in a similar fashion with respect to accumulation of protein. Therefore, based on identical tissue specificity, developmental appearance, and accumulation kinetics, the control elements present in these constructs appear not to be affected by the sequences present in the reporter genes used.

DISCUSSION

Deletion analysis of the human *PGK2* gene in transgenic mice serves to define the extent of the sequences that encode transcriptional control elements. A 323-bp region between the functional 515-bp upstream region and the nonfunctional 192-bp sequence contains information essential for directing testis-specific and developmentally regulated expression. The 192-bp fragment contains sequence elements that are probably essential for promoter function—e.g., the CAAT box and GC box sequences (Fig. 1); however, these elements clearly require the upstream sequences to function. This fact is based on the observation that although none of the three transgenic lines containing the 192-bp *PGK2* fragment

showed any expression, 10 of 12 transgenic lines containing larger upstream regions exhibited testis-specific expression of the *PGK2* transgenes. Possibly elements located upstream of the 1.4-kb location normally exert effects on the levels of transcription of human *PGK2*; however, we have shown here that such elements are not essential for proper expression.

Another germ-cell-specific gene the control regions of which have been mapped is the mouse protamine 1-encoding gene (*Prm-1*) (31). Although *Prm-1* transcription initiates in round spermatids and not in pachytene cells, comparison of the regulatory regions of *Pgk-2* and *Prm-1* provides candidate sequence elements that may be involved in the control of tissue-specific expression. Both *Prm-1* and human *PGK2* 5′ regions contain two copies each of the sequence GGGTGG-GG, present in one orientation in *PGK2* and in the opposite orientation in the *Prm-1* gene (20). It is also found in the upstream region of the *Prm-2* gene in the same orientation as *PGK2* (20). This sequence could have a role in tissue-specific gene expression; however, it should be noted that this sequence has not been identified in the available upstream sequence in the mouse *Pgk-2* gene.

A study of genes expressed at different times during spermatogenesis has resulted in the identification and cloning of several cDNAs corresponding to transcripts that display a pattern of expression nearly identical to that of *Pgk-2* (14). If *Pgk-2* is one of a number of coordinately expressed genes, then cis-acting elements responsible for the regulation of *Pgk-2* expression may be reiterated in the control regions of these other genes. Elucidation of common cis-acting regulatory elements would establish a mechanism for coordinate gene regulation at this time during spermatogenesis.

We have shown that the *PGK2*/CAT and the *PGK2*/luciferase fusion genes are translated concurrently with the first observable *PGK2* and *PGK2* fusion gene transcripts. Previous reports, however, suggest that translation of the endogenous message is delayed (7, 12). This difference can be the result of a number of effects. Possibly the *PGK2*/CAT fusion genes are missing regulatory elements required for translational control of the message, either because of species differences between human and mouse *Pgk-2* or because the fusion gene is lacking the coding region and the 3′ untranslated sequences of the *PGK2* message. Alternatively, it is possible that although *PGK2* and *PGK2* fusion gene messages are translationally controlled, a low rate of translation exists in pachytene cells that is only detectable by using CAT and luciferase reporter genes. PGK-2 protein, however, would not be observed until the increased translational efficiency in later cell types permitted detection.

We thank Nathan Ellis for the gift of the mouse *Pgk-2* 3′ clone and Jessie Dausman for assistance with the transgenic mice. This work was supported by a grant from the American Cancer Society to M.I.S. and from the Markey Foundation Grant for Developmental Biology to the California Institute of Technology as well as a Program Project Grant AG 07687 to M.I.S. Support was also provided by a National Institutes of Health Grant HD 23126 to J.R.M., who is a Career Development Fellow HD 00829.

1. Handel, M. A. (1987) in *Spermatogenesis Genetic Aspects*, ed. Hennig, W. (Springer, Berlin), pp. 1–62.
2. Hecht, N. B. (1986) in *Experimental Approaches to Mammalian Embryonic Development*, eds. Rossant, J. & Pederson, R. (Cambridge Univ. Press, Cambridge, England), pp. 151–193.
3. Bellve, A. R., Cavicchia, J. C., Millette, C. F., O'Brien, D. A., Bhatnagar, Y. M. & Dym, M. (1977) *J. Cell Biol.* **74**, 68–85.
4. Vandeberg, J. L., Cooper, D. W. & Close, P. J. (1973) *Nature (London) New Biol.* **243**, 48–50.
5. Chen, S., Donahue, R. P. & Scott, R. (1976) *Fertil. Steril.* **27**, 699–701.
6. Vandeberg, J. L., Cooper, D. W. & Close, P. J. (1976) *J. Exp. Zool.* **198**, 231–240.
7. Kramer, J. M. & Erickson, R. P. (1981) *Dev. Biol.* **87**, 37–45.
8. Kramer, J. M. (1981) *Dev. Biol.* **87**, 30–36.
9. Kozak, L. P., McLean, G. K. & Eicher, E. M. (1974) *Biochem. Genet.* **11**, 41–47.
10. Beutler, E. (1969) *Biochem. Genet.* **3**, 189–195.
11. Erickson, R. P., Kramer, J. M., Rittenhouse, J. & Salkeld, A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6086–6090.
12. Gold, B., Fujimoto, H., Kramer, J. M., Erickson, R. P. & Hecht, N. B. (1983) *Dev. Biol.* **98**, 392–399.
13. McCarrey, J. R. & Thomas, K. (1987) *Nature (London)* **326**, 501–505.
14. Thomas, K. T., Wilkie, T. W., Tomashefsky, P., Bellve, A. R. & Simon, M. I. (1989) *Biol. Reprod.*, in press.
15. Szabo, P., Grzeschik, K. & Siniscalco, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3167–3169.
16. Tani, K., Singer-Sam, J., Munns, M. & Yoshida, A. (1985) *Gene* **35**, 11–18.
17. Boer, P. H., Adra, C. N., Lau, Y. & McBurney, M. W. (1987) *Mol. Cell. Biol.* **7**, 3107–3112.
18. Kadonaga, J. T., Jones, K. A. & Tjian, R. (1986) *Trends Biochem. Sci.* **11**, 20–23.
19. McKnight, S. & Tjian, R. (1986) *Cell* **46**, 795–805.
20. McCarrey, J. R. (1987) *Gene* **61**, 291–298.
21. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
22. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
23. de Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, S. (1987) *Mol. Cell. Biol.* **7**, 725–737.
24. Hogan, B., Costantini, F. & Lacy, E. (1986) *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 152–182.
25. Mullis, K. B. & Faloona, F. A. (1987) *Methods Enzymol.* **155**, 335–350.
26. Auffray, C. & Rougan, F. (1980) *Eur. J. Biochem.* **107**, 303–324.
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
28. Overbeek, P. A., Chepelinksy, A. B., Khillan, J. S., Piatogorsky, J. & Westfall, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7815–7819.
29. Shaw, W. V. (1975) *Methods Enzymol.* **43**, 737–755.
30. Deluca, M. & McElroy, W. D. (1978) *Methods Enzymol.* **57**, 3–15.
31. Peschon, J. J., Behringer, R. R., Brinster, R. L. & Palmiter, R. D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5316–5319.

Chapter 3

Determining transcript number using the polymerase chain reaction: *Pgk-2*, *mP2*, and *PGK-2* transgene mRNA levels during spermatogenesis

(This chapter was published in *Nucleic Acids Research* **19**: 1557-1562, 1991)

Nucleic Acids Research, Vol. 19, No. 7 1557

Determining transcript number using the polymerase chain reaction: *Pgk-2*, *mP2*, and *PGK-2* transgene mRNA levels during spermatogenesis

Murray O.Robinson* and Melvin I.Simon
Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

Received December 11, 1990; Revised and Accepted February 27, 1991

ABSTRACT

We describe a technique that uses reverse transcription and the polymerase chain reaction (pcr) to rapidly quantitate numbers of specific mRNA transcripts from nanogram quantities of total cellular RNA. Linearity of input molecules to output signal was maintained by limiting the cycle number and the amount of input RNA and by minimizing the number of manipulations. Absolute levels of specific transcripts were determined by the inclusion of a separate standard curve composed of serially diluted *in vitro* transcribed RNA run alongside the experimental samples. This allowed rapid quantitation of many samples simultaneously. We applied this technique to measuring the expression of phosphoglycerate kinase 2 (Pgk-2) transgenes in the mouse testis during development. A human *PGK-2* transgene, a *PGK-2*/CAT transgene, and the endogenous *mPgk-2* gene all displayed similar patterns and levels of expression, consistent with the conclusion that peak RNA accumulation occurs in pachytene spermatocytes. Mouse protamine 2 (*mP2*) is expressed at a level approximately tenfold higher than *Pgk-2* and displays a different pattern of expression consistent with initiation of transcription occuring in haploid round spermatids.

INTRODUCTION

Quantitative analysis of RNA is central to the understanding of the mechanisms that regulate gene activity. Northern analysis is currently the most widely used method for analyzing RNA levels, however, northern analysis requires microgram amounts of RNA, is time consuming and is of limited quantitative use. Techniques based on solution hybridization of RNA are more sensitive and quantitative, however, the increased time and effort involved limit their widespread use. The advent of the polymerase chain reaction (pcr) (1) has resulted in the development of a number of quantitative (2–5) and semi-quantitative (6–8) procedures for the analysis of steady-state RNA levels. These protocols all include the addition of reverse transcriptase to the samples to convert the RNA into a DNA copy providing a suitable substrate for taq polymerase. The pcr approach allows increased sensitivity, ease of use, and requires only small amounts of RNA.

Protocols designed to quantitate the products of the pcr must account for the exponential nature of the cycling reaction. The product (N) of pcr amplification can be defined by the function $N=N_o(1+\text{eff.})^n$, and is initially linear with respect to input molecules (N_o). However, as cycle number (n) and total product (N) increase, the efficiency (eff.) of the pcr is reduced as the concentration of one of the substrates eventually becomes limiting, resulting in loss of linearity. Furthermore, small changes in the efficiency of the reaction result in large differences in the amount of total product. This leads to non-uniformity in the level of pcr product from tube to tube or experiment to experiment. Accordingly, efforts to use the pcr to measure RNA levels have included internal competitive or noncompetitive templates to control for such differences in amplification. Although useful for standardization, each of these templates requires additional cloning steps, and the use of an internal standard requires several reactions to measure each sample.

Another approach (9) takes advantage of the theoretical linearity of input molecules versus pcr product. Using this approach, the relative levels of testis-specific messages were measured by limiting both the cycle number and the amount of starting material such that the efficiency of the reaction remained constant. Although this approach is useful for determining the relative levels of a specific mRNA in different samples, it is not useful for comparing the level of different messages, and it is difficult to determine whether the message levels detected are of physiological importance (5,10). We have extended the utility of this approach through the use of an externally generated standard curve. The standard curve is used both to determine the range over which the reaction is linear and to determine the level of signal that corresponds to a specific number of RNA molecules. Here we present a novel quantitation protocol (modified from reference 9) that provides rapid and accurate analysis of absolute RNA levels from many samples and allows quantitative comparison of different messages.

We are interested in gene expression during spermatogenesis, and have been studying the transcriptional control of the gene for Phosphoglycerate Kinase (*Pgk-2*), expressed only in meiotic

* To whom correspondence should be addressed

1558 Nucleic Acids Research, Vol. 19, No. 7

and post meiotic male germ-cells. In the mouse, the initial stages of spermatogenesis occur synchronously and begin shortly after birth. Germ-cells first enter meiosis around day 12, and divide to form haploid round spermatids by about day 21. Thus, the age of the prepuberal mouse can be correlated with the developmental progression of germ-cells in the testis. Previously, we showed that a human *PGK-2* transgene is expressed in a tissue-specific and developmentally controlled manner in mice, and that a *PGK-2*/CAT fusion gene demonstrates tissue-specific CAT activity that is first detected at *post natal* day 13 (11). Due to the limits of quantitation and low yields of RNA from immature testes we were unable to accurately measure and compare RNA levels from these transgenes. Here, we have used this technique to show that the human *PGK-2* transgene, the *PGK-2*/CAT fusion gene and the endogenous mouse *Pgk-2* gene all share the same pattern of expression, and the mRNA levels vary less than fourfold among the three genes. Furthermore, the developmental expression profiles of the *PGK-2* genes are consistent with high levels of expression in pachytene cells, a cell-type that was previously thought to contain only low levels of *Pgk-2* message (12).

MATERIALS AND METHODS

RNA purification

Tissue samples were obtained from transgenic mice homozygous for either the *PGK-2*/CAT transgene or the human genomic *PGK-2* transgene. Animals were sacrificed on the days indicated (day 0 = date of birth), and the tissue was frozen at −70°C prior to RNA extraction. RNA was prepared by homogenization of the tissues using a polytron, and then precipitation in 3 M LiCl, 6 M Urea as described (13).

RNA transcriptions

In vitro transcribed RNAs for standardization of the pcr were prepared with low specific activity according to the procedure for large scale transcriptions provided by the supplier (Bethesda Research Laboratories, Gaithersburg, MD). The concentration of the RNAs were determined by both TCA precipitable radioactivity and absorption at 260 nm. The concentration and integrity of the RNAs were confirmed by agarose gel electrophoresis in the presence of formaldehyde. All RNA transcriptions were performed using fragments of the genes of interest cloned into Bluescript pBPIIKS vectors (Stratagene, San Diego, CA) containing T3 and T7 promoters (Fig. 1). *Pgk-2*/CAT RNA was transcribed with T3 polymerase from a linearized vector containing the 780 bp CAT coding region cloned into the *Hind* III and *Bam* HI sites of the pBPIIKS polylinker. *PGK-2* RNA was transcribed with T7 polymerase from a linearized vector containing a 1.6 kb *Hinc* II fragment from the *PGK-2* coding region cloned into the *Hinc* II site. *Pgk-2* RNA was transcribed with T7 polymerase from a linearized template containing a 481 bp *Hinc* II fragment from the 3′ end of the *Pgk-2* coding sequence. Mouse Protamine 2 (*mP2*) RNA was transcribed with T3 polymerase from a linearized vector containing a 269 bp *Pst* I/*Eco* RI fragment from the 3′ untranslated region of *mP2* cloned into the same sites of pBPIIKS. All *in vitro* transcribed RNAs were diluted to 1×10^9 molecules per μl and stored as ethanol precipitates at −20°C.

Pcr primers

All primers used for pcr amplification contained between 21 and 25 nucleotides and had calculated melting temperatures (Tm) of 64° C or above, as determined by Tm = (4× each cytosine or guanine + 2× each adenine or thymidine)°C. The *PGK-2*/CAT, human *PGK-2*, mouse *Pgk-2*, and *mP2* primers amplified regions of 281 bp, 358 bp, 356 bp, and 208 bp respectively. The sequences of each pair of primers are shown in Figure 1.

Reverse transcription/pcr

Both reverse transcription and pcr were performed in the same tube in a single buffer. Quantitation of experimental RNA samples was as follows: a suspension of 100 ng ethanol precipitated RNA was aliquoted to a 0.67 ml centrifuge tube, glycogen (20 μg, Boehringer Mannheim, West Germany) was added and the sample was centrifuged in a microfuge at 12,000 rpm for 10 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 1× pcr buffer (50 mM KCl, 10 mM Tris pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin; Cetus, Emeryville, CA) containing dNTPs at 200 μmole each, 2 μCi alpha ^{32}P labeled dATP (3000 Ci/mmole, Amersham, Arlington, IL), and 100 pmole each primer. Each sample was then incubated at 65°C for 3 minutes and then placed on ice for 2 minutes. Moloney Murine Leukemia Virus Reverse Transcriptase (20 units, Pharmacia, Sweden), Amplitaq Polymerase (5 units, Cetus), and RNasin (35 units, Promega, Madison, WI), were mixed and then added to the sample, and 50 μl mineral oil was layered on top. Each set of tubes was immediately transferred to an Ericomp Thermocycler (Ericomp, San Diego, CA) and incubated for 10 minutes at 37°C, then 19−25 cycles of 92°C for 20 seconds, annealing temperature for 10 seconds, and 72°C for 30 seconds. Annealing temperatures were set at the melting temperature of the lowest Tm oligonucleotide.

In vitro transcribed RNA standards were treated similarly, except that the ethanol precipitated T3 or T7 RNA transcripts

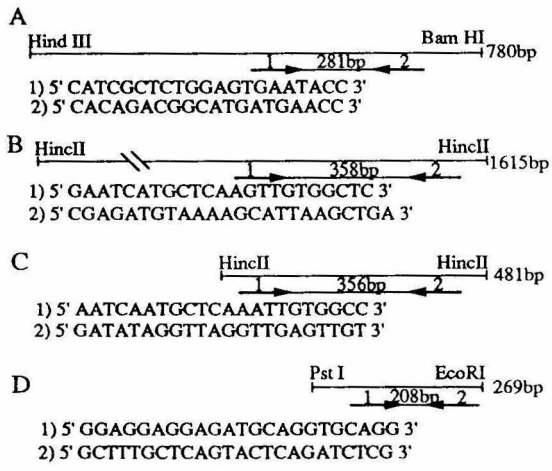


Figure 1. Schematic representation of subclones and sequence of primers used to quantitate RTpcr. All subclones are flanked by restriction sites used to clone into pBluescript BPIIKS vectors. The arrows denote the approximate positions of the 5′ and 3′ pcr primers. The sizes of the subclones and the pcr products are indicated, and the sequences of the 5′ and 3′ oligonucleotide primers (1 and 2 respectively) are shown below each subclone. Subclones are oriented with the start of transcription to the left. A, *PGK-2*/CAT; B, human *PGK-2*; C, mouse *Pgk-2*; D, *mP2*.

were first resuspended in diethyl pyrocarbonate treated H_20, serially diluted as required, then reprecipitated in the presence of 20 μg glycogen and 100 ng liver RNA. After the pcr, formamide loading buffer (0.1% w/v xylene cyanol, 0,1%w.v bromophenol blue, 10 mM Na_2EDTA, 95% v/v formamide) was added, the samples were heated to 65°C for 3 minutes, and then electrophoresed in a 6% acrylamide/6 M urea gel at 200 volts for 1.5 hours. The gel was then dried down and exposed to a Molecular Dynamics Phosphorimager plate (Molecular Dynamics, Sunnyvale, CA) for one to ten hours and quantitated using the phosphorimager. Similar results were obtained by excising the bands from the dried gel and quantitating by liquid scintillation counting.

RESULTS

The protocol for the combined reverse transcription and polymerase chain reaction (RTpcr) is predicated on the theoretical linear relationship of input molecules to total product. By limiting the amount of starting material and the number of cycles of pcr, we could keep all substrates in excess, insuring uniform amplification. Because the final yield of the pcr is sensitive to small differences in efficiency, another consideration was to minimize the number of manipulations involved. By combining all components of both the reverse transcriptase and the pcr in a single reaction mixture (pcr buffer, dNTPs, primers, reverse transcriptase and taq polymerase) we could achieve amplification of the RNA. It was then essential to demonstrate that the reaction was linear, sensitive and reproducible.

Linearity of RTpcr

To examine the linearity of the RTpcr, eleven samples of twofold dilutions of T3 transcribed CAT RNA were mixed with 100 ng C57B6/DBA2 mouse liver RNA. The range of the experiment was from 4,900 molecules to 5,000,000 molecules of CAT RNA. Each sample was prepared as described in Materials and Methods and both the pcr mix and the enzyme mix were aliquoted from a common stock. Samples with fewer input molecules required more cycles to obtain a detectable signal. Accordingly, samples containing from 4,900 to 78,000 CAT RNA molecules were run for 25 cycles, and samples containing 78,000 to 5,000,000 molecules were run for 21 cycles. The signals from the 25 cycle reactions were normalized to the 21 cycle values using the overlapping 78,000 molecule samples. Approximately half of the resulting products were run on a 6% polyacrylamide gel, dried, and quantitated using the Molecular Dynamics phosphorimager system. As shown in figure 2A, there is a linear relationship between input RNA molecules and signal when the number input RNA molecules is between 19,000 and 5,000,000. The signal begins to level off as the concentration of CAT RNA molecules is decreased below this range. A control reaction containing no added CAT RNA molecules also gives a faint signal (data not shown), suggesting that this constant signal at very low RNA concentrations is due to contamination of the pcr solutions with a total of approximately 10,000 CAT RNA or DNA sequences per reaction.

Variability between experiments

To determine the variability of the RTpcr, CAT RNA dilution curves were repeated over a period of one month, using 21 cycle reactions and identical thermocycler times and temperatures, but using different prepared solutions, different tubes of enzyme stock, and different lots of ^{32}P labeled dATP. Figure 2B shows three such CAT dilution series, each normalized at the 78,000 molecule point to account for differences in exposure time. Although each experiment shows a linear relationship between input molecules and signal, the slopes of the three lines vary as much as twofold; the slope of each line representing the relative efficiency of amplification. It is important to note that although the efficiency of the reaction may vary from day to day (presumably due to differences in the concentrations or specific activities of the components), in any given experiment, linearity of input molecules to signal is observed. Because a high degree

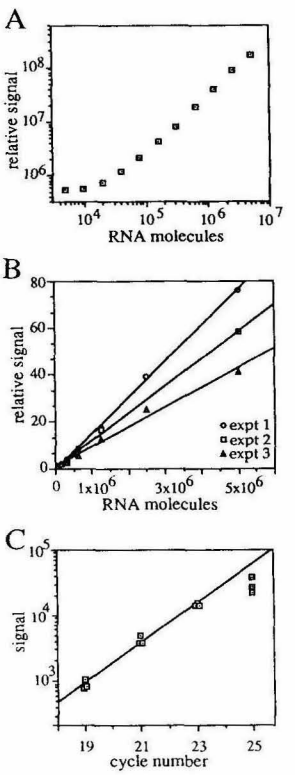


Figure 2. Quantitation of Reverse Transcription/pcr. A) Linearity of RTpcr. Eleven samples of twofold serial dilutions of T3 transcribed CAT RNA were mixed with 100 μg liver RNA and subjected to the RTpcr protocol. Half of the reaction mixture was run on a 6% polyacrylamide gel, dried, quantitated using a Molecular Dynamics phosphorimager and plotted as signal versus RNA molecules. Samples containing 4,900 to 78,000 molecules were run for 25 cycles; samples containing 78,000 to 5,000,000 molecules were run for 21 cycles. The two sets of points were normalized to the 21 cycle curve by the overlapping 78,000 molecule values. B) Experiment variability. Over a period of one month, three sets of serial twofold or fourfold serial dilutions of T3 transcribed CAT RNA molecules were mixed with 100 ng liver RNA, subjected to RTpcr for 21 cycles and quantitated as in A. A line drawn through each set of points describes the linear fit of the data by the least squares technique. C) Tube to tube uniformity. Triplicate samples of 5 million molecules of CAT T3 RNA mixed with 100 ng liver RNA were subjected to RTpcr for the number of cycles indicated and quantitated as in A. A line drawn through the 19, 21, and 23 cycle points describes the exponential increase in product with increasing cycle number.

1560 Nucleic Acids Research, Vol. 19, No. 7

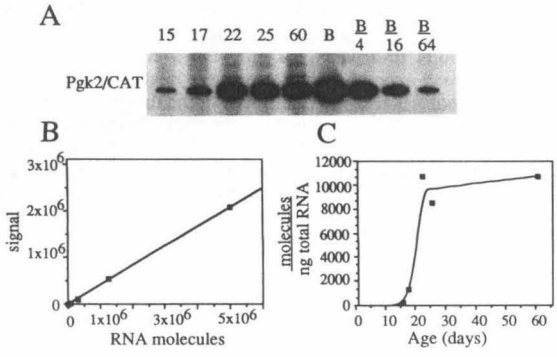


Figure 3. Quantitation of *Pgk-2*/CAT transgene RNAs in the testis. A) Autoradiograph of a dried 6% polyacrylamide gel showing the products of RT pcr from a developmental timecourse. Total RNA samples (100 ng) from the testes of *Pgk-2*/CAT transgenic mice of the ages indicated were amplified for 21 cycles using CAT primers. On the right is a standard curve of fourfold serial dilutions of T3 transcribed CAT RNA run concurrently to allow absolute quantitation of the *Pgk-2*/CAT transgene RNAs. The letter B = 5×10^6 molecules. B) Plot of signal versus CAT RNA molecules of the standard curve in A. The equation of the line drawn through the points allows the conversion of the developmental timecourse signal into numbers of RNA molecules. C) Appearance of *Pgk-2*/CAT RNA during development. The signal in A was converted to numbers of CAT RNA molecules using the standard curve and is presented as molecules CAT RNA per ng total RNA as a function of development.

of variability was seen, we were unable to use a single standard curve for multiple experiments, and were required to include a new standard curve to run alongside each set of experiments.

Uniformity among reaction vessels

The efficiency of the pcr is affected by variables such as incubation time, reaction temperature, and the concentrations of the substrates. Variation in the efficiency of pcr among individual tubes in an experiment would lead to unequal amplification from tube to tube, making quantitation impossible. The linearity of the standard curves presented above (each data point representing a separate tube) suggests that each sample is amplifying with the same efficiency. To test this directly, triplicate samples of 5,000,000 molecules of CAT RNA were each amplified for 19, 21, 23 or 25 cycles and processed as described above. Figure 2C shows a semilog plot of signals corresponding to the 12 samples. The variation in signal among triplicate samples is 20% or less. The efficiency of amplification (calculated by $N=N_o(1+\text{eff.})^n$) at 19, 21 and 23 cycles is nearly 100%, however, by 25 cycles the efficiency of amplification is clearly decreasing. Thus, we can conclude that while in the exponential range of the reaction, the efficiency of amplification is uniform, both for adjacent tubes at a particular cycle number and for reactions amplified for sequential numbers of cycles.

PGK-2/CAT RNA developmental timecourse

Having demonstrated that we can select conditions under which the reaction is linear with *in vitro* transcribed standards, we next applied the technique to experimental samples. We have shown previously that CAT activity expressed from a *PGK*-2/CAT transgene is found only in the mouse testis and is first detected at *post natal* day 13 (11). To analyze the developmental pattern

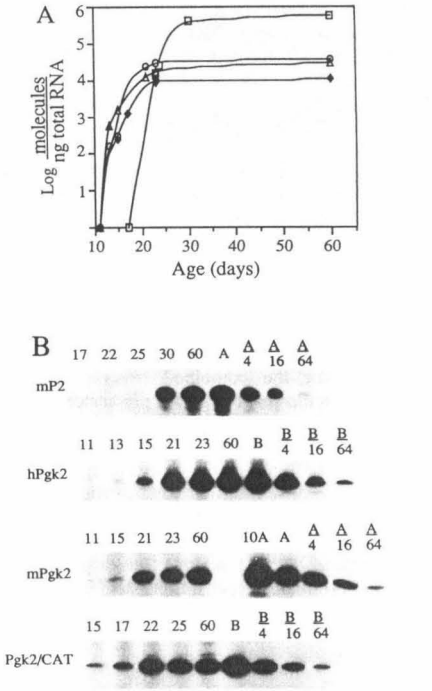


Figure 4. Developmental accumulation and quantitation of *Pgk-2*, *mP2*, *PGK-2* transgene and *PGK-2*/CAT transgene RNAs. One hundred ng total testis RNA isolated from mice at various days was subjected to the RTpcr protocol for various cycles. A) Semilog plot of the message levels of the four genes analyzed. *mP2*, open squares; *PGK-2*, open circles; *mPgk-2*, open triangles; *PGK-2*/CAT, closed diamonds. B) Autoradiographs of pcr products used to derive points in A. The target gene amplified in each experiment is indicated. Numbers refer to the developmental age in days of each RNA sample. The letters refer to the number of *in vitro* transcribed RNA molecules used in each standard curve. A, 1×10^7 in the *mPgk-2* experiment, 1×10^8 in the *mP2* experiment; B, 5×10^6. Note that in the *mPgk-2* experiment, the 10A sample (1×10^8 molecules) is out of the linear range and was not used in the standard curve calculation.

of expression, CAT RNA levels were measured in samples of testis RNA isolated from transgenic *PGK-2*/CAT mice of various ages. *Post natal* days 15 to day 25 were monitored closely as this represents the period of most dynamic change in the expression of the *PGK-2* promoter.

In a single experiment, 100 nanogram aliquots of RNA from mouse testes at *post natal* days 15, 17, 22, 25, and 60 were amplified at 21 cycles using primers corresponding to sequences in the CAT coding region. Included alongside the experimental points in separate tubes was a standard curve consisting of fourfold serial dilutions of T3 transcribed CAT RNA starting at 5,000,000 molecules. The standard curve was amplified simultaneously using the same primers under identical conditions. Figure 3A shows an autoradiogram of the resulting pcr products. The points of the standard curve, designed to encompass the range of the experimental points, were quantitated, and demonstrated that the reaction was in the linear range (Figure 3B). The standard curve was used to convert the signal from the experimental points (d15-d60) to molecules of CAT RNA per nanogram total RNA.

The resulting plot of the developmental accumulation of *PGK-2*/CAT RNA (Figure 3C) shows the level of message increasing from 250 molecules/ng total RNA at day 15 to 10,000 molecules/ng at day 60. The majority of the RNA accumulates between days 17 and 25 and correlates with the appearance and proliferation of pachytene spermatocytes in the developing testis (14). The observation that the level of specific message per nanogram total RNA remains at a high level from day 25 to day 60, a period during which the proportion of pachytene cells in the testis is decreasing, suggests that *PGK-2*/CAT RNA is also being expressed at significant levels in spermatids, the predominant cell-type in the adult mouse testis (14).

Mouse *Pgk-2*, Protamine 2 (*mP2*), and human *PGK-2* developmental accumulation

To determine if the *Pgk-2*/CAT transgene expression reflects that of the human transgene and endogenous *Pgk-2* genes, we analyzed the expression patterns of both the human genomic *PGK-2* transgene and the endogenous *Pgk-2* locus and compared the expression levels and patterns to that of *PGK-2*/CAT. A measure of developmental accumulation of these genes would determine more precisely the regulation of the *Pgk-2* genes in the testis and would uncover any differences in the regulation of the three genes. *mP2* was also analyzed, as its expression pattern begins in the haploid round spermatid stage of spermatogenesis (15). The expression pattern of *mP2* is important as it will delineate the appearance and proliferation of round spermatids, and will thus serve as a reference point for the *Pgk-2* expression patterns.

Each of the three genes, mouse *Pgk-2*, *mP2*, and the human *PGK-2* transgene were analyzed by RTpcr using a standard curve for each obtained using known amounts of T3 or T7 transcribed RNAs. The standard curves demonstrated linearity of amplification and allowed absolute quantitation of the expression levels as described above. RNA isolated from mice containing the human *PGK-2* transgene was used for both the endogenous *mPgk-2* gene and the *PGK-2* transgene. RNA isolated from mice containing the *PGK-2*/CAT transgene was used for the *mP2* experiment. Figure 4 shows developmental RNA accumulation curves for all four genes analyzed. Among the Pgk-2 genes, the highest level of expression is found in the human *PGK-2* transgene, and the lowest in the *PGK-2*/CAT transgene. Levels of expression of transgenes are in part a function of their site of chromosomal integration, and therefore, correlating the levels of expression with the presence or absence of transcriptional signals is difficult. The transgenic lines chosen for this analysis were the highest expressing lines tested (11), and despite concerns over position of integration effects, the three Pgk-2 promoted RNA levels varied less than fourfold. More importantly, examination of the RNA accumulation of the three genes demonstrates that all three genes are expressed with very similar developmental timecourses. In contrast, the *mP2* developmental analysis shows a very different pattern of RNA accumulation. Expression of *mP2* is first detected at day 22 and accumulation reaches 65% of maximum levels by day 30; this is consistent with *mP2* gene expression inititating in round spermatids.

DISCUSSION

Here we describe a procedure for the quantitative analysis of RNA message levels using reverse transcription and the polymerase chain reaction. We have shown that by careful selection of cycle number and quantity of input RNA we can readily quantitate RNA levels. Because we have demonstrated uniform amplification from tube to tube, we are able to include a standard curve alongside the experimental points to determine absolute levels of mRNA. This protocol is rapid and simple: up to twelve samples of RNA have been quantitated in less than seven hours with sufficient resolution to detect at least twofold differences in the amount of RNA. Incorporation of ^{32}P labeled dATP into the pcr product increases sensitivity, allows simple quantitation, and eliminates the need for Southern transfer. The linear range of the assay is demonstrated here to be greater than two orders of magnitude. By using less input RNA or fewer cycles, this linear range can be extended greatly. At very low levels of input RNA the signal begins to level off as the number input molecules continues to decrease, and we believe this is due to contaminating RNA or DNA. We have not attempted to stringently eliminate contamination in our system, although, by reducing the level of contamination we would expect to increase the sensitivity of the assay.

The linearity achieved with this method may be attributed to several factors. First, limiting cycle number and input RNA molecules keeps all other substrates in excess. Designing the protocol to use one tube, one buffer for all enzymatic reactions also minimizes variability. The use of the 3′ pcr primer (as opposed to a poly dT oligomer) to prime the reverse transcriptase reaction eliminates the potential for differential priming efficiencies of certain messages and insures uniform cDNA synthesis. Additionally, maintaining the annealing temperature at the Tm of the primers reduces the level of background bands, and may therefore increase linearity by not depleting the substrates.

Although this protocol uses unaltered *in vitro* transcribed RNA molecules as the template, these molecules differ from the target messages not only in length but in the absence of modifications that natural messages contain (5′ capping and 3′ poly A tail). These differences may affect the efficiency of amplification in early steps of the pcr and could affect the accuracy of the quantitation. Additionally, care must be taken to assure that all reactions amplify with equal efficiency. Therefore, a standard curve in the linear range that brackets the experimental points is essential for accurate quantitation. Another variable that could affect the efficiency of the reaction is the source and preparation of the RNA sample. We have amplified standard amount of CAT RNA in samples of total RNA from various mouse tissues and shown that all samples amplify with similar efficiencies (data not shown). However, RNA preparations of different purity or residual contaminants might be expected to have different amplification efficiencies.

Several other pcr protocols have been published for quantitating RNA levels. These protocols rely on internal standards that must be modified to produce an amplified product that may be distinguished from the target pcr product. Such modifications include addition of a restriction site (2,4), addition of an intron (4) and the construction of an artificial amplification template (3). The use of such internal standards requires not only the additional steps to prepare these altered template, but also necessitates performing several pcr reactions to quantitate each sample of RNA.

In our analysis of the transcriptional control of the *Pgk-2* gene during spermatogenesis, RTpcr has enabled us to readily measure the steady-state levels of *Pgk-2* RNA throughout the development of the germ-cell. By determining the initiation of transcription

and quantitating the accumulation of message of the three genes, we show here that the *PGK-2*/CAT transgene, the *PGK-2* transgene, and the mouse *Pgk-2* gene all function in an identical manner. We assume therefore that the transcriptional machinery in the developing germ cell recognizes the same cis-acting sequences in the 1.4 kb human 5′ sequence as in the endogenous locus.

The developmental curves presented here also serve as markers for the appearance and proliferation of specific cell-types in the developing testis. As the appearance and accumulation of *mP2* RNA is a marker for the appearance and proliferation of round spermatids, we suggest that the peak of *Pgk-2* expression occurs earlier, during the pachytene stage of germ cell development. Noting that the proportion of pachytene cells peaks at approximately day 21, then decreases in the adult (14), the continued high levels of expression of *Pgk-2* RNA during development suggest that continued expression occurs in haploid round spermatids. This conclusion is supported by *in situ* data (16 and our unpublished observations).

Singer-Sam et. al. (9), using a related technique with separated spermatogenic cell populations, showed that round spermatids had the highest level of *Pgk-2* RNA. This apparent discrepancy may be reconciled by the observation that round spermatids contain roughly one fourth as much RNA as pachytene spermatocytes (17), causing the *Pgk-2* specific RNA to appear higher in round spermatids when calculated as a function of total RNA. Analysis on a per cell basis (as done in *in situ* experiments), or as a function of total testis RNA, however, shows the highest levels of *Pgk-2* RNA to be in pachytene cells.

This type of analysis is very useful for determining the accumulation of specific messages during development and correlating those messages with the appearance and proliferation of specific cell-types. These expression curves allow us to plot coordinate regulation of both messages and protein products as a function of development and therefore to learn more about the control of development in the testis.

ACKNOWLEDGEMENTS

We would like to thank Andrew Pakula and Tom Wilkie for critical review of the manuscript. This work was supported by N.I.H. grant AG 07687.

REFERENCES

1. Mullis, K. B. and Faloona, F. A. (1987) Methods Enzymol. **155**, 335–350.
2. Becker-Andre, M. and Hahlbrock, K. (1989) Nucleic Acids Res. **17**, 9437–9446.
3. Wang, A. M., Doyle, M. V., and Mark, D. F. (1989) Proc. Natl. Acad. Sci. **86** 9717–9721.
4. Gilliland, G., Perrin, S., Blanchard, K. and Bunn, H. F. (1990) Proc. Natl. Acad. Sci. **87**, 2725–2729.
5. Chelly, J., Kaplan, J., Maire, P., Gantron, S. and Kahn, A. (1988) Nature **333**, 858–860.
6. Rappolee, D. A., Mark, D., Banda, M. J. and Werb, Z. (1988) Science **241**, 708–712.
7. Frye, R. A., Benz, C. C. and Liu, E. (1989) Oncogene **4**, 1153–1157.
8. Brenner, C. A., Tam, A. W., Nelson, P. A., Engleman, E. G., Suzuki, N., Fry, K. E. and Larrick, J. W. (1989) Biotechniques **7**, 1096–1103.
9. Singer-Sam, J., Robinson, M. O., Bellve, A. R., Simon, M. I. and Riggs, A. D. (1990) Nucleic Acids Res. **18**, 1255–1259.
10. Chelly, J., Concordet, J., Kaplan, J. and Kahn, A. (1989) Proc. Natl. Acad. Sci. **86**, 2617–2621.
11. Robinson, M. O., McCarrey, J. R. and Simon, M. I. (1989) Proc. Natl. Acad. Sci. **86**, 8437–8441.
12. Gold, B., Fujimoto, H., Kramer, J. M., Erickson, R. P. and Hecht, N. B. (1983) Dev. Biol. **98**, 392–399.
13. Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. **107**, 303–324.
14. Bellve, A. R., Cavicchia, J. C., Millete, C. F., O'Brien, D. A., Bhatnagar, Y. M. and Dym, M. (1977) J. Cell Biol. **74**, 68–85.
15. Kleene, K. C., Distel, R. J. and Hecht, N. B. (1984) Dev. Biol. **105**, 71–79.
16. Goto, M., Koji, T., Kiyonobu, M., Tamaru, M., Koikeda, S., Nakane, P. K., Mori, N., Masamune, Y. and Nakanishi, Y. (1990) Exp. Cell Res. **186**, 273–278.
17. Kleene, K. C., Distel, R. J. and Hecht, N. B. (1983) Dev. Biol. **98**, 455–464.

Chapter 4

The 5' untranslated region of *PGK-2* is sufficient to confer translational regulation onto a heterologous mRNA in transgenic mice

The 5' untranslated region of *PGK-2* is sufficient to confer translational regulation onto a heterologous mRNA in transgenic mice

Murray O. Robinson and Melvin I. Simon

Divsion of Biology
California Institute of Technology
Pasadena, CA 91125

Keywords: CAT, Spermatogenesis, germ cells

ABSTRACT

During spermatogenesis, a number of mRNAs, including the message for phosphoglycerate kinase 2 (Pgk-2) are subject to developmentally controlled translational repression. Here we show that all of the elements required for translational control of the *PGK-2* gene in transgenic mice are present in a fusion mRNA containing only the 5' untranslated region of the *PGK-2* message. CAT fusion genes containing the *PGK-2* 5' region and either the *PGK-2* 3' region or the SV40 3' region exhibited a similar lag between RNA and protein accumulation during testis development, a lag that mimics the observed accumulation of RNA and protein in the endogenous *Pgk-2* gene. Furthermore, transgene messages containing the *PGK-2* 5' untranslated region, the CAT coding sequence and the SV40 small t antigen 3' untranslated region exhibited a shift from the nonpolysomal fraction to the polysomal fraction of testis cytoplasmic extracts during the developmental maturation of the testis. This developmentally regulated recruitment onto polysomes occurs with the same timing as the polysomal shift observed with the endogenous *Pgk-2* message.

INTRODUCTION

The regulated translation of an mRNA has recently emerged as an important aspect of the control of gene expression in eukaryotes. A growing number of examples of regulation at the translational level have been uncovered over the past few years. Examples include the iron-dependent regulation of ferritin translation (1); the activation of maternal RNAs in many organisms including, xenopus (2), and mouse (3) oocytes; cell-cycle control of ribosomal protein translation (4 and references therein), steroid regulation of translation in the myelin sheath (5), and a number of examples during spermatogenesis in trout (6) drosophila (7) and mouse (8, 9, 10, 11, 12, 13).

Spermatogenesis is a highly regulated process consisting of mitotic, meiotic and haploid stages of differentiation (for reviews see 14, 15). During the later stages of this process, the nucleus elongates and condenses, and the DNA-binding histones are replaced by protamines, effectively eliminating transcription (16). The terminal differentiation steps of spermatogenesis are thought to be facilitated by proteins whose messages were transcribed previously and held in an untranslatable state. This requirement for activation of translationally repressed messages has led to the identification of a number of messages subject to translational regulation. These genes include the transition proteins (10, 12, 13), the protamines (9, 13), a testis-specific, heat shock protein and a testis-specific transcript of c-Abl (11) and the testis-specific isozyme PGK-2 (8).

The gene for the testis-specific isozyme, *Pgk-2* is transcribed exclusively in meiotic and haploid male germ cells (17, 18), and is subject to translational regulation. Pgk-2 RNA is first observed in the leptotene/zygotene stages of meiosis (18, 19), whereas Pgk-2 protein is not observed until the round spermatid stage, some seven days later (20, 21). Gold et al. (8), using *in vitro* translation of fractionated cytoplasmic RNAs showed that Pgk-2 RNA expressed in pachytene cells was found in the nonpolysomal fraction where it is presumably not translated. Subsequently, the mouse (22) and human (23) genes for Pgk-2 (*Pgk-2* for the mouse and *PGK-2* the human) were cloned, allowing a molecular analysis of the expression and regulation of the Pgk-2 gene.

Using the Human *PGK-2* gene, we have previously demonstrated that a 515 bp fragment of the 5' region is sufficient to direct testis-specific expression of a reporter gene (24), and that the developmental accumulation of the endogenous and the transgene messages are similar (25). Here, we investigated the translational regulation of the *PGK-2* gene to localize cis-acting regions of the gene responsible for the observed translational lag of the endogenous gene. We show that a transgene containing only the 5' untranslated region (UTR) of the human PGK-2 gene is sufficient to confer correct developmentally controlled translational regulation onto a CAT reporter gene.

MATERIALS AND METHODS

Gene Constructs and Transgenic Mice. The construction of the *PGK-2*/CAT/SV40 plasmid has been described previously (24),

and was isolated from vector sequences for microinjection by digestion with *Xba* I and *Bam* HI. The *PGK-2*/CAT/*PGK-2* 3' plasmid was constructed by first cloning the *Hind* III to *Mbo* I fragment (isolated from the first vector) containing the CAT coding region into the *Hind* III and *Bam* HI sites of pBPIIKS+ (Stratagene, San Diego, CA). Next, a 1.4 kb *Hind* III fragment containing the entire 5' UTR of the *PGK-2* gene was inserted into the unique *Hind* III site 5' to the CAT coding region. Next, the *PGK-2* 3' end was isolated as a 1.3 kb *Xba* I to *Bam* HI restriction fragment isolated from an 8 kilobase (kb) human genomic *PGK-2* fragment This fragment, containing 53 nucleotides 5' to the termination codon and the entire 3' untranslated region, was cloned into the *Xba* I and *Bam* HI sites of pSP72 (Promega, Madison, WI) vector. This fragment was reisolated as an *Spe* I to *Sac* I fragment and cloned into the same sites 3' to the CAT coding sequence in the *PGK-2*/CAT vector, providing the 3' UTR and poly A addition sequences for the CAT coding sequence. This fusion gene was isolated from vector sequences by digestion with *Xho* I and *Sac* I. Production and screening of transgenic mice was as described previously (24).

RNA and RTpcr. Total RNA was extracted after tissue homogenization in 6 M urea/3M LiCl as described (26). RTpcr was performed as described (25) using 100 ng total RNA per assay. Amplification in the exponential range was monitored by the inclusion of an external standard curve using *in vitro* transcribed CAT RNA.

CAT assays. CAT assays were performed as described (27). Protein extracts were prepared by homogenization of tissues from homozygous male mice. Linearity of CAT activity was insured by monitoring the reactions over a thirty-minute period. CAT activity was calculated from the rate of each linear reaction.

PGK isozyme electrophoresis. Protein extracts were prepared by homogenization of tissues in a dounce homogenizer at 0 ^{o}C. The mouse sperm sample was prepared by homogenization of the epididymus from an adult C57BL6/DBA2 mouse. Electrophoresis and staining were performed as described (28) using a Gelman Semi-Micro II electrophoresis chamber and Sepraphore III cellulose acetate strips (Gelman, Ann Arbor, MI)

Polysome Gradients. Polysome gradients were performed essentially as described (9). Four mice were used for the day 20 sample, two for day 33 and one for the adult sample. Briefly, cytoplasmic extracts were prepared by homogenization of decapsulated testes in 1ml 100mM KCl, 20 mM Hepes, pH 7.6, 1.5mM MgCl2 (HKM buffer) with 0.1% diethyl pyrocarbonate, 0.5% Triton N-101. Nuclei and cell debris were pelleted by centrifugation at 12,000g for 5 minutes. The resulting cytoplasmic extract (0.4 ml) was layered onto a 10-40 % linear sucrose gradient containing HKM buffer with a 0.5 ml 60% sucrose cushion. Identical gradients were prepared in which 10 mM EDTA was substituted for the MgCl2. Samples were centrifuged at 288,000g for 100 minutes. Fractions were collected from the bottom of the tube using a peristaltic pump, and twenty fractions of 550 µl each were collected, the A260 was

determined in a spectrophotometer, and the fractions were ethanol precipitated. For confirmation of the A260 profile an aliquot of each fraction was electrophoresed in the presence of formaldehyde and stained with ethidium bromide to visualize the ribosomal RNA bands. The position of the monosomes within each gradient was determined by the appearance of strong ribosomal RNA bands at fraction 13 in each gradient. For the RNA quantitation, 20 µl precipitate from each fraction was centrifuged (12,000g for 10minutes) in the presence of 10 µg glycogen, then resuspended and processed for RTpcr. Samples were amplified using primers and cycle numbers described previously (25).

RESULTS

A number of translationally regulated genes have been analyzed in molecular detail (1, 4, 29, 30, 31) In these genes, the sequences controlling the translational regulation have been localized to the 5' or 3' UTRs. During spermatogenesis, translational regulation of the protamine 1 (*Prm1*) gene appears to require the 3' untranslated sequences (32). Additionally, a protein/RNA interaction has been identified in the 3' UTR of protamine 2 (*Prm2*) that includes a sequence that is well conserved among several haploid-expressed translationally regulated genes (33). Given this information we constructed two *PGK-2*/CAT fusion genes to determine whether the noncoding sequences of the *PGK-2* gene had an effect on the translation of the message. The first construct contained both the 5' and the 3' *PGK-2* untranslated sequences flanking the CAT gene; the

second construct, used previously (24), contained the SV40 small T antigen 3' UTR in place of the *PGK-2* 3' sequences (Figure 1). The 5' sequences of the fusion genes contain 1.4 kb of upstream sequence and included the entire 5' untranslated sequence of the *PGK-2* gene. This was accomplished by converting ATG start codon of the PGK-2 sequence to a *Hind* III site and cloning the sequence into the unique *Hind* III site of pSVOCAT, 29 bp upstream of the CAT initiation codon.

Transgenic mice provide a useful system to investigate the role of cis-acting elements in the control of gene regulation, and for spermatogenesis, the lack of a culture system and the requirement for monitoring a developmental process dictated that we assay the constructs *in vivo*. Both of these constructs were introduced into transgenic mice, breeding lines for each construct were established, and one line from each construct was selected for further analysis. One transgenic line, B17, containing the promoter/CAT fusion with the SV40 3' UTR has been characterized previously, and its expression has been shown to be testis-specific (24). The transgenic line TR6, containing the promoter/CAT fusion with the *PGK-2* 3' end was first analyzed by CAT assays of various tissues to establish that the expression of the transgene was tissue-specific (data not shown). These two transgenic lines were next analyzed for the expression of the transgene RNAs and protein products during development.

The first wave of spermatogenesis in the mouse is synchronous and begins soon after birth. This regulated development can be used to correlate the age of the mouse with the appearance and

proliferation of specific spermatogenic cell-types (34). To assay the accumulation of transgenic RNA during spermatogenesis, we extracted RNA from the testes of mice from the TR6 and the B17 lines at various ages after birth. One hundred ng total RNA was then processed using quantitative RTpcr (25) and the signal from each sample was plotted as percent of adult RNA expression. As Figure 2 shows, the RNA accumulation of both the TR6 construct and the B17 construct throughout testis development are similar, suggesting that the 5' promoter region alone is sufficient to confer the transcriptional control of the message. RTpcr analysis of the endogenous *Pgk-2* message also shows a similar pattern of RNA accumulation (25), a pattern consistent with high-level expression in pachytene spermatocytes. Next, the developmental accumulation of the CAT protein was assayed. As in the RNA analysis, extracts of testes from both the TR6 and the B17 lines were prepared at various days after birth. CAT activity from these extracts were assayed by the ^{3}H Acetyl CoA method (27), a method that is easily quantifiable. Plotted as percent adult activity, the accumulation of CAT activity during development is also shown in Figure 2. The accumulation of CAT activity for both the TR6 and the B17 lines is similar, but both show a different pattern than that of the message. The accumulation of CAT activity appears to lag behind the RNA accumulation curve by several days. The accumulation of CAT activity in both of these lines is similar to the published appearance of endogenous PGK-2 activity during development of CD-1 mice (20). However, unlike several reports that first report PGK-2 activity at about 22 days (consistent

with the appearance of haploid, round spermatids) (21), we detected CAT activity as early as day 13, a developmental time consistent with expression in meiotic spermatocytes. We suspected that this difference was not due to differences in the regulation between the endogenous locus and the transgenes, but rather was due to an increased sensitivity of detection of CAT activity versus PGK-2 protein. To address this discrepancy, we used a recently developed isozyme activity assay for PGK-2 enzyme (28), and used extracts from B6/DBA2 strain mice (the same background as the transgenic strains). Extracts from the testes of mice at ages 14, 15, 19, 22, 23, and 25 days were electrophoresed on a cellulose acetate membrane and stained for PGK enzyme activity. A faint band of PGK-2 activity was observed by day 14 or 15 (Figure 3). Note that the more intensely staining band co-migrates with PGK-1 activity. PGK-1 activity has been found previously in testes extracts, albeit at nearly undetectable levels (20). Because the PGK-2 activity detected here is much less than the observed PGK-1 activity, we believe that this staining procedure is more sensitive than previous methods that failed to detect PGK-2 activity this early in development. The identification of PGK-2 enzyme activity at day 14, and consistent with the expression of CAT activity, suggests that at least some translation of both the endogenous PGK-2 message and both transgenic CAT messages is occurring in pachytene spermatocytes. Western blot analysis using the same extracts and an antibody against PGK-2 gave a similar result (data not shown). This result suggests that the endogenous gene and both transgenes are all

regulated in a similar manner, both in terms of transcription and translation. This similarity suggests that the *PGK-2*/CAT transgenes are also translationally regulated, and furthermore, that the 3' UTR is dispensable for controlling the observed lag between message and protein accumulation.

Having observed a lag between the RNA and CAT activity for the transgene, and having demonstrated that this lag is similar to that observed for the endogenous *Pgk-2* locus, we next addressed the nature of the lag. It is possible that inefficient translation of the RNA coupled with a stable protein could account for the difference in accumulation rates of the RNA and the protein. Alternatively, in pachytene cells, the RNAs may be found in the non polysomal fraction and then shift into the polysomal fraction accompanying their active translation later in development.

To address the possibility of inefficient translation, we considered the sequences around the initiation codon. Kozak (35) has determined a consensus sequence that plays an important role in the efficiency of translation of messages. Comparing the sequences of both the *PGK-2*/CAT transgene and the endogenous *Pgk-2* gene to the Kozak consensus sequence (Figure 1B) revealed that both genes were in agreement with the important base pairs required for efficient translation. Kozak's mutational analysis shows that an A in position -3 and a G in position +1 are most important for efficient translation, and these positions are present in both genes examined. However, there are other factors involved in efficient translation of messages, including 5' and 3' UTR sequences (36, 37), and it is also

possible that male germ-cell requirements for translation do not conform to the Kozak rules.

To address the translational efficiency of the transgene message (containing only the 5' UTR of the human *PGK-2* gene) and to confirm that of the endogenous *Pgk-2* message, we next assayed each message's association with the polysomal fraction of cytoplasmic testis extracts. Monitoring RNA compartmentalization as a function of development would indicate whether a change in the efficiency of translation had taken place. A shift from nonpolysomal to polysomal compartments during the development of the testis would signal an increase in the translational efficiency, whereas no change in the polysomal profile would suggest constant, inefficient translation of the message.

We chose to fractionate cytoplasmic extracts from developmental days 20, 33 and adult. At day 20, the RNA accumulation is near adult levels, whereas the protein accumulation is roughly 5% of adult levels (see Figure 2); day 33 is soon after the CAT protein levels begin to rise rapidly, and the adult sample shows the steady-state levels of both RNA and protein. Cytoplasmic extracts were made from the testes of mice at each of the days examined, layered onto 10 to 40% sucrose gradients containing a 60% sucrose cushion, then centrifuged at 288,000xg for 100 minutes. Fractions were then collected from the bottom of each gradient and the A260 of each fraction was measured. A260 profiles of the three gradients (Figure 4A) show the relative distribution of the polysomes and monosomes. Profiles of the day 20 and day 33 gradients are

very similar, whereas the adult gradient, although essentially similar, exhibits an increase in the number of smaller polysomes. An aliquot of each fraction was also electrophoresed in a denaturing agarose gel to confirm the position of the monosome peak. Identical gradients were run with EDTA substituted for Magnesium to dissociate the polysomes, and one such gradient is shown. To further characterize the polysome gradients, quantitative RTpcr was performed on fractions from the gradients using B Actin primers to demonstrate the polysomal profile of a message that is not subject to translational regulation (10) (Figure 4B). To demonstrate quantitative signal from the RTpcr, all experiments were either performed at two different cycle numbers and then the relative signals for each experiment were shown to be the same, or the highest expressing fraction was serially diluted to demonstrate that the RTpcr was in the linear range. The polysomal profiles of the actin gradients show a relatively consistent distribution in both polysomal and nonpolysomal fractions at all three developmental timepoints. We next used the same fractions to compare RNA distribution of the endogenous *Pgk-2* locus. Figure 5A shows the relative distribution of the *Pgk-2* message in polysomal and nonpolysomal compartments. At day 20, nearly all of the message present is found in the nonpolysomal fractions (fractions 16-18). A small amount of message is seen in the polysomal fractions as would be expected because of the presence of the PGK-2 isozyme at day 20 (Figure 3). By day 33, however, a clear shift is observed as a substantial fraction of the message is observed in the polysomal fraction. By day 33,

large numbers of haploid round spermatids are present as are a number of elongating and condensing spermatids. This observation fits well with the increase in PGK-2 protein accumulation reported after day 30 in CD-1 mice (20). The adult polysome profile shows even more RNA in the polysomal fraction and is consistent with a large number of elongating and condensing spermatids present in the adult testis. To examine the polysome profile of the B17 transgene message, the same polysome gradients were analyzed by RTpcr using CAT primers (Figure 5B). At day 20, the CAT message is found primarily in the nonpolysomal fractions, and again by day 33, a shift was observed from the nonpolysomal to the polysomal fractions. The polysomal distribution of CAT message at day 60 is quite similar to the day 33 distribution, but perhaps even more signal is present in the polysomal fraction. Note that the majority of the CAT RNA in the day 33 and the adult gradients is associated with smaller polysomes than the *Pgk-2* message. The expected size of each message, (Pgk-2 is 1.6 kb and PGK-2/CAT is 1.4 kb) is quite similar, so we believe that other factors such as intrinsic translation efficiencies (those specified by Kozak's rules), or other variables affecting the efficiency of translation (36, 37) may be affecting the number of ribosomes associated with each message.

It is also not clear, however, why when compared to *Pgk-2* there is so little CAT message present in the nonpolysomal fraction at day 60. It is possible that factors such as differential stability of the CAT versus the *Pgk-2* mRNA may affect the percentage of each message in the polysomal fractions. It is evident, however, that CAT

transcripts from the B17 line (*PGK-2*/CAT/SV40) exhibit a shift from the nonpolysomal to the polysomal fraction in a developmental time course similar to that of the endogenous *Pgk-2* message.

DISCUSSION

We have shown here that an mRNA containing the 5' UTR of the human *PGK-2* gene is subject to the same translational regulation as the endogenous mouse *Pgk-2* message. We first demonstrated that RNA and protein accumulation in both the *PGK-2*/CAT transgene and the *Pgk-2* gene followed a similar lag; then we correlated this lag with the developmentally timed recruitment of both the transgene and the endogenous message onto polysomes. We suggest from these results that the 5' UTR of the *Pgk-2* message is sufficient for controlling its translational regulation. It is possible that sequences in the CAT gene or in the SV40 small t antigen 3' UTR (and *PGK-2* 3' UTR) have a role in the observed translational control, but because of the similar regulation of both genes, we feel that this is unlikely.

Measuring RNA and protein levels throughout the development of the testis enables us to compare easily the relative amounts of the various molecules and to correlate them with the appearance and proliferation of specific cell types. Using this approach, we first demonstrated that the transgene protein accumulation lags behind that of its message by a developmental time of over one week. Examination of the timing of RNA accumulation during development suggests that the transgene message and the endogenous message accumulate to high levels in pachytene cells. This is based on the

predicted expansion of the pachytene cell population from day 15 to day 22 (34) the period of rapid accumulation of *Pgk-2* message. Additionally, *Prm2* mRNA, a marker for presence of round spermatids, does not accumulate rapidly until after day 25 (25). The transgene protein levels, as measured by CAT activity, don't begin to rise rapidly until the expansion of round spermatids (34,38) and continue to accumulate throughout the development of the testis. This accumulation curve is similar to that observed for the endogenous PGK-2 protein (20). Unlike previous reports of expression of the endogenous protein (21), we observed both CAT activity and PGK-2 protein activity in testis extracts likely to be containing only meiotic cells. This demonstrates that some translation of both the transgene and the endogenous message is occurring in pachytene cells.

It is interesting that the the cis-acting sequences for translational control map to the 5' UTR of *Pgk-2*. The cis-acting elements for *Prm-1* translational control mapped to its 3'UTR (32) and in *Prm-2* RNA binding activity was mapped to a sequence element in the 3'UTR, a sequence conserved among *Prm-1*, *Prm-2*, and *TP-1* (33) We did not find this sequence in the *Pgk-2* message, and for translational regulation of *Pgk-2*, the 3' UTR proved to be dispensable. Additionally, *Prm-1*, *Prm-2*, *TP-1*, and an unidentified mRNA (13) undergo deadenylation concurrent with translation, a phenomenon not observed in the *Pgk-2* message (24). It is also possible that *Pgk-2* translational regulation is temporally different from that of the haploid-expressed group. A polysome gradient from

26 day-old mice shows *PGK-2*/CAT RNA associated with polysomes, but shows *Prm-2* message present only in the nonpolysomal fraction (M.O.R., unpublished), suggesting that the activation of PGK-2 translation is earlier than that of *Prm-2*. These differences between *Pgk-2* and the other translationally regulated messages in testis suggest that the mechanism of regulation of *Pgk-2* might be different from that of the haploid-expressed genes.

Previous work establishing *Pgk-2* as a translationally regulated message relied on *in vitro* translation to monitor RNA levels in the spermatogenic cells (8). In light of recent studies using the cloned *Pgk-2* gene (18, 25), the levels of *Pgk-2* message in pachytene cells as assayed by *in vitro* translation may have been underestimated. There is evidence that the translational efficiency of some messages is regulated by covalent modification of the mRNA (39). If this were the case for *Pgk-2*, the observed low levels of message in pachytene cells might be due to inefficient *in vitro* translation of pachytene cell mRNA. Thus, the relative translational efficiency of deproteinated *Pgk-2* RNA in pachytene and haploid cells might yield information as to the mechanism of its translational regulation.

Examples of steroid regulation of translation (5) present an intriquing possibility for genes expressed during spermatogenesis. It is known that spermatogenesis is dependent on steroid hormones. Among many potential effects of hormone action, one effect could possibly be activation or repression of translation. Verdi and Campagnoni (5) have identified a sequence in the 5' UTR of the Myelin Basic Protein mRNA that promotes hormone-dependent

translation of this message. Sequence comparison in the 5'UTR of *Pgk-2* shows no overt conservation to this sequence. Nonetheless, the possibility of hormone stimulation or repression of translation in spermatogenesis might be further investigated.

Efforts in determining the mechanism of translational regulation of the *Pgk-2* message might now focus on the region of the 5' UTR. Protein/RNA binding assays using the 5' UTR sequences might lead to the identification of a binding activity. It is interesting to note that the presence of extra copies of *PGK-2* message in these experiments did not appear to affect the translational control of the endogenous message. Experiments intended to titrate out binding activity by overexpression of the 5' UTR of *Pgk-2* in transgenic mice might yield information as to the mechanism and function of translational control of the *Pgk-2* gene.

REFERENCES

1. Klausner, R. D. & Harford, J. B. (1989) *Science*. **246,** 870-872.

2. Richter, J. D. (1987) in *Translational Regulation of Gene Expression*, ed. Ilan. (Plenum, New York), pp. 111-139.

3. Davidson, E. H. (1986) *Gene Activity in early development* (Academic Press, Orlando).

4. Levy, S., Avni, D., Hariharan, N., Perry, R. P. & Meyuhas, O. (1991) *Proc. Natl. Acad. Sci. USA*. **88,** 3319-3323.

5. Verdi, J. M. & Campagnoni, A. T. (1990) *J. Biol. Chem*. **265,** 20314-20320.

6. Sinclair, G. D. & Dixon, G. H. (1982) *Biochemistry*. **21,** 1869-1877.

7. Schafer, M., Kuhn, R., Bosse, F. & Shafer, U. (1990) *EMBO*. **9,** 4519-4525.

8. Gold, B., Fujimoto, H., Kramer, J. M., Erickson, R. P. & Hecht, N. B. (1983) *Dev. Biol*. **98,** 392-399.

9. Kleene, K. C., Distel, R. J. & Hecht, N. B. (1984) *Dev. Biol*. **105,** 71-79.

10. Heideran, M. A. & Kistler, W. S. (1987) *J. Biol. Chem*. **262,** 13309-13315.

11. Zakeri, Z. F., Ponzetto, C. & Wolgemuth, D. J. (1988) *Dev. Biol*. **125,** 417-422.

12. Yellick, P. C., Kwon, Y. K., Flynn, J. F., Borzorgzadeh, A., Kleene, K. C. & Hecht, N. B. (1989) *Molecular Reproduction and Development*. **1,** 193-200.

13. Kleene, K. C. (1989) *Development*. **106,** 367-373.

14. Handel, M. A. (1987) in *Spermatogenesis: Genetic Aspects*, ed. Hennig. (Springer-Verlag, Berlin), pp. 1-62.

15. Hecht, N. B. (1986) in *Experimental Approaches to Mammalian Embryonic Development*, ed. Rossant and Pederson. (Cambridge University Press, Cambridge, England), pp. 151-193.

16. Kierszenbaum, A. L. & Tres, L. L. (1975) *J. Cell Biol.* **65,** 258-270.

17. Erickson, R. P. & Kramer, J. M. (1980) *Proc. Natl. Acad. Sci. USA.* **77,** 6086-6090.

18. Goto, M., Koji, T., Mizuno, K., Tamaru, M., Koikeda, S., Nakane, P. K., Mori, N., Masamune, Y. & Nakanishi, Y. (1990) *Exp. Cell Res.* **186,** 273-278.

19. Singer-Sam, J., Robinson, M. O., Bellve, A. R., Simon, M. I. & Riggs, A. D. (1990) *Nucl. Acids Res.* **18,** 1255-1259.

20. Kramer, J. M. & Erickson, R. P. (1981) *Dev. Biol.* **87,** 37-45.

21. VandeBerg, J. L., Cooper, D. W. & Close, P. J. (1976) *J. Exp. Zool.* **198,** 231-240.

22. Boer, P. H., Adra, H. N., Lau, Y. & McBurney, M. W. (1987) *Mol. Cell Biol.* **7,** 3107-3112.

23. McCarrey, J. R. & Thomas, K. (1987) *Nature.* **326,** 501-505.

24. Robinson, M. O., McCarrey, J. R. & Simon, M. I. (1989) *Proc. Natl. Acad. Sci. USA.* **86,** 8437-8441.

25. Robinson, M. O. & Simon, M. I. (1991) *Nucl. Acids Res.*

26. Auffray, C. & Rougeon, R. (1980) *Eur. J. Biochem.* **107,** 303-324.

27. Neumann, J. R., Morency, C. A. & Russian, K. O. (1987) *Biotechniques.* **5,** 444-447.

28. Monk, M. (1987) in *Mammalian Development: a Practical Approach*, ed. Monk. (IRL Press, Washington, D.C.), pp. 139-161.

29. Vassalli, J., Huarte, J., Belin, D., Gubler, P., Vassalli, A., O'Connell, M. L., Parton, L. A., Rickles, R. J. & Strickland, S. (1989) *Genes Dev.* **3,** 2163-2171.

30. Standart, N., Dale, M., Stewart, E. & Hunt, T. (1990) *Genes Dev.* **4,** 2157-2168.

31. McGrew, L. L. & Richter, J. D. (1990) *EMBO.* **9,** 3743-3751.

32. Braun, R. E., Peschon, J. J., Behringer, R. R., Brinster, R. L. & Palmiter, R. D. (1989) *Genes Dev.* **3,** 793-802.

33. Kwon, Y. K. & Hecht, N. B. (1991) *Proc. Natl. Aca. Sci. USA.* **88,** 3584-3588.

34. Bellve, A. R., Cavicchia, J. C., Millette, C. F., O'Brien, D. A., Bhatnagar, Y. M. & Dym, M. (1977) *J. Cell Biol.* **74,** 68-85.

35. Kozak, M. (1986) *Cell.* **44,** 283-292.

36. Jackson, R. J. & Standart, N. (1990) *Cell.* **62,** 15-24.

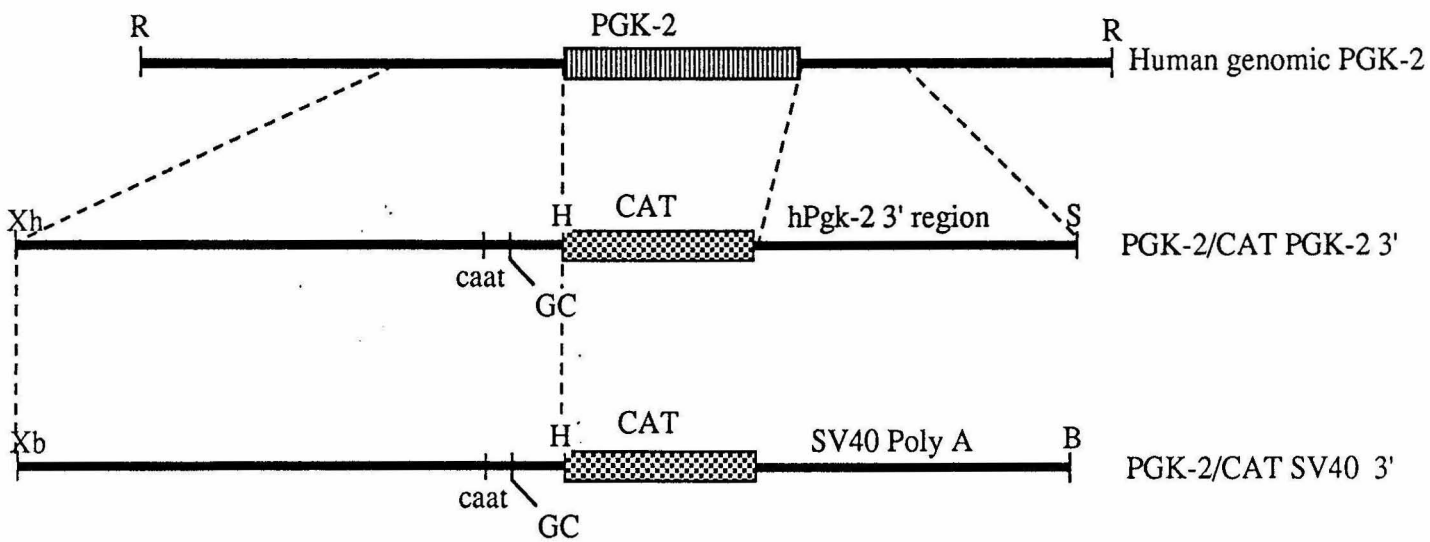
37. Falcone, D. & Andrews, D. W. (1991) *Mol. Cell Biol.* **11,** 2656-2664.

38. Nebel, B. P., Amarose, A. P. & Hackett, E. M. (1961) *Science.* **134,** 832-833.

39. Caldwell, D. C. & C. P. Emerson, J. (1985) *Cell.* **42,** 691-700.

Figure 1. Gene constructs. A) Schematic drawing of human *PGK-2* genes injected into mice. The human genomic *PGK-2* is shown to denote the origin of the *PGK-2*/CAT sequences used in the transgenes. caat, Cat transcription factor consensus sequence; GC, Sp1 transcription factor consensus sequence; B, *Bam* HI; H, *Hind* III; R, *Eco* RI; S, *Sac* I; Xb, *Xba* I, Xh, *Xho* I. B) Sequence around the initiation codon. The endogenous *Pgk-2* sequence and the *PGK-2*/CAT sequence are shown next to the Kozak consensus sequence for comparison. Initiation codons are underlined.

A



B

	+1
Kozak Consensus	C C A C C A T G G
mPgk-2	T C A A G A T G G
PGK-2/CAT	C T A A A A T G G

Figure 2. Developmental accumulation of transgene RNA and protein. RTpcr was performed on RNA samples from testes of mice removed on various days up to postnatal day 60. The resulting signal was quantitated and is plotted as percent adult level. TR6 RNA (open circles); B17 RNA (filled circles). CAT assays were performed on protein extracts made from testes of mice of the indicated ages. The activity is plotted as percent adult level. TR6 CAT activity (open squares); B17 CAT activity (filled squares). The adult level for both the RNA and CAT activity was determined by the average of two or more individual samples..

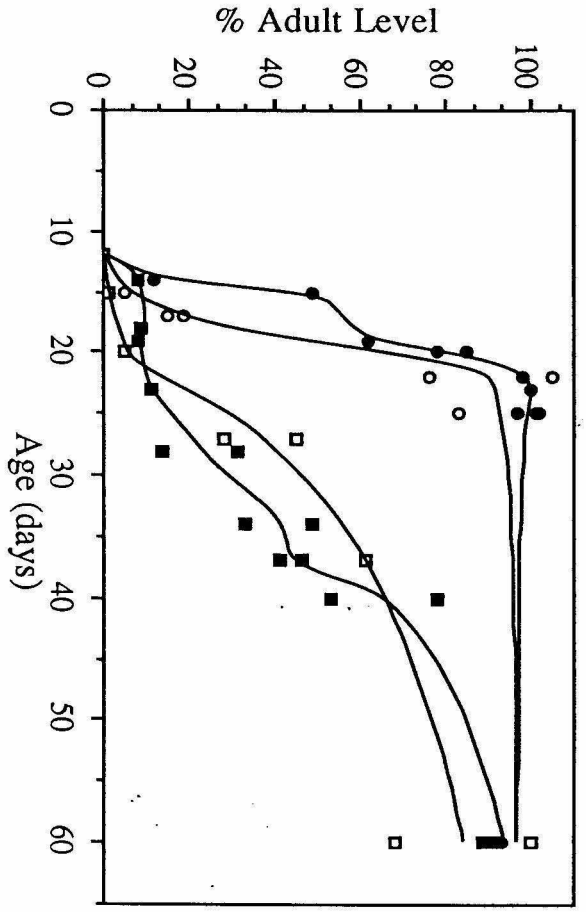
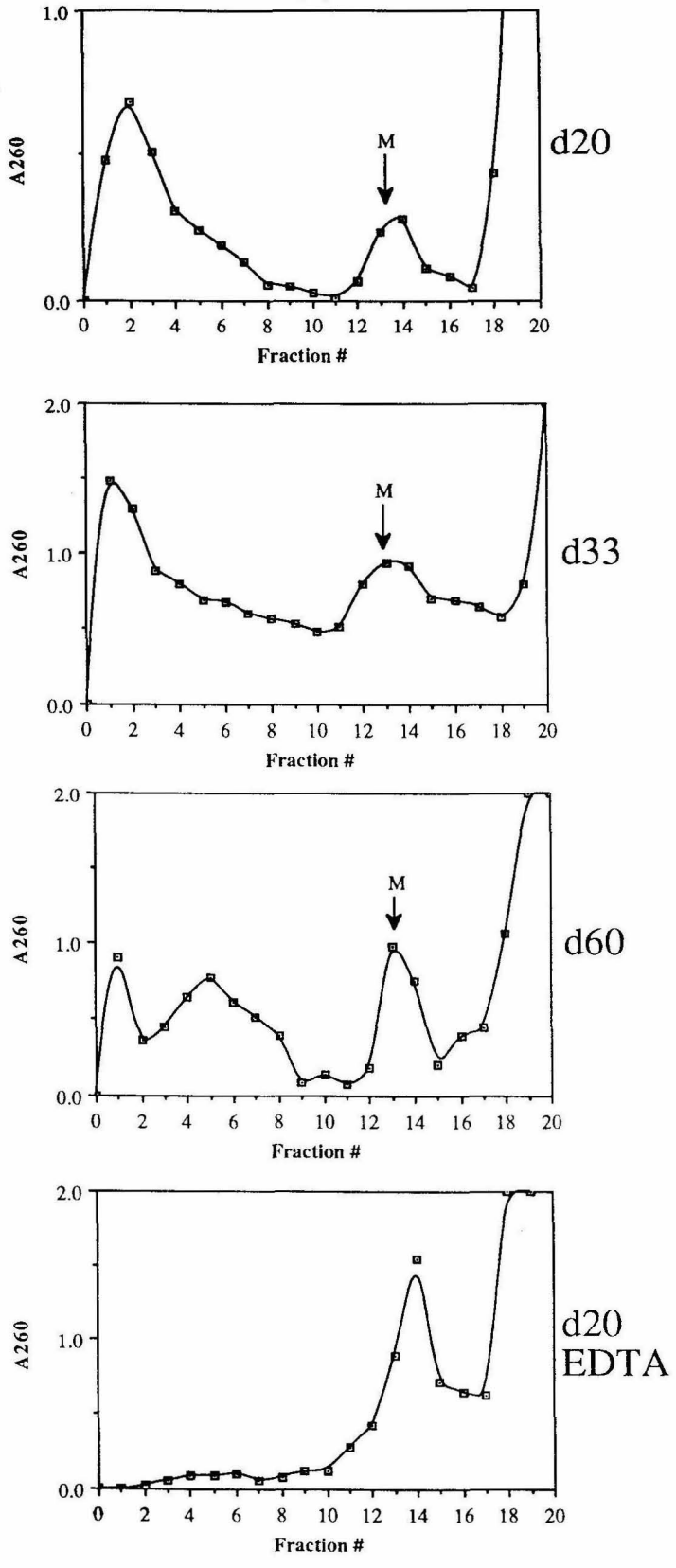

% Adult Level
0
20
40
60
80
100
Age (days)
0
10
20
30
40
50
60

Figure 3. PGK isozyme electrophoresis. Protein extracts from the testes of mice at the ages (in days) indicated, and from a sample enriched for mouse spermatozoa (M), were electrophoresed on cellulose acetate followed by staining for PGK activity. The positions of the PGK isozymes are indicated. The origin is towards the top of the figure.


PGK-2
14 15 19 22 23 25 M
PGK-1

Figure 4. Characterization of polysome gradients. A) A260 profiles of polysome gradients. Cytoplasmic extracts from the testes of mice were layered onto 10-40% sucrose gradients and centrifuged at 288,000g for 100 minutes. Fractions were collected from the bottom of the gradient (fraction 1 = bottom, fraction 20 = top) and an aliquot was removed for O.D. determination. The ages of the mice (in days) from which the extracts were produced is indicated. Aliquots from each fraction were electrophoresed on denaturing agarose gels to confirm the position of the monosomal fraction. The first fraction to contain significant numbers of monosomes (as determined by the 28S and 18S ribosomal RNA bands) is indicated (M). A gradient is also shown in which the day 20 sample was run in a gradient containing EDTA in place of Magnesium (d20 EDTA). B) B Actin RNA levels in polysome gradients. An aliquot from every other fraction was subjected to RTpcr using B Actin oligonucleotide primers. The resulting signal is plotted as the percent of the total B Actin signal assayed in each gradient. An autoradiograph of each radiolabelled pcr product is shown above each assayed fraction.

A



B

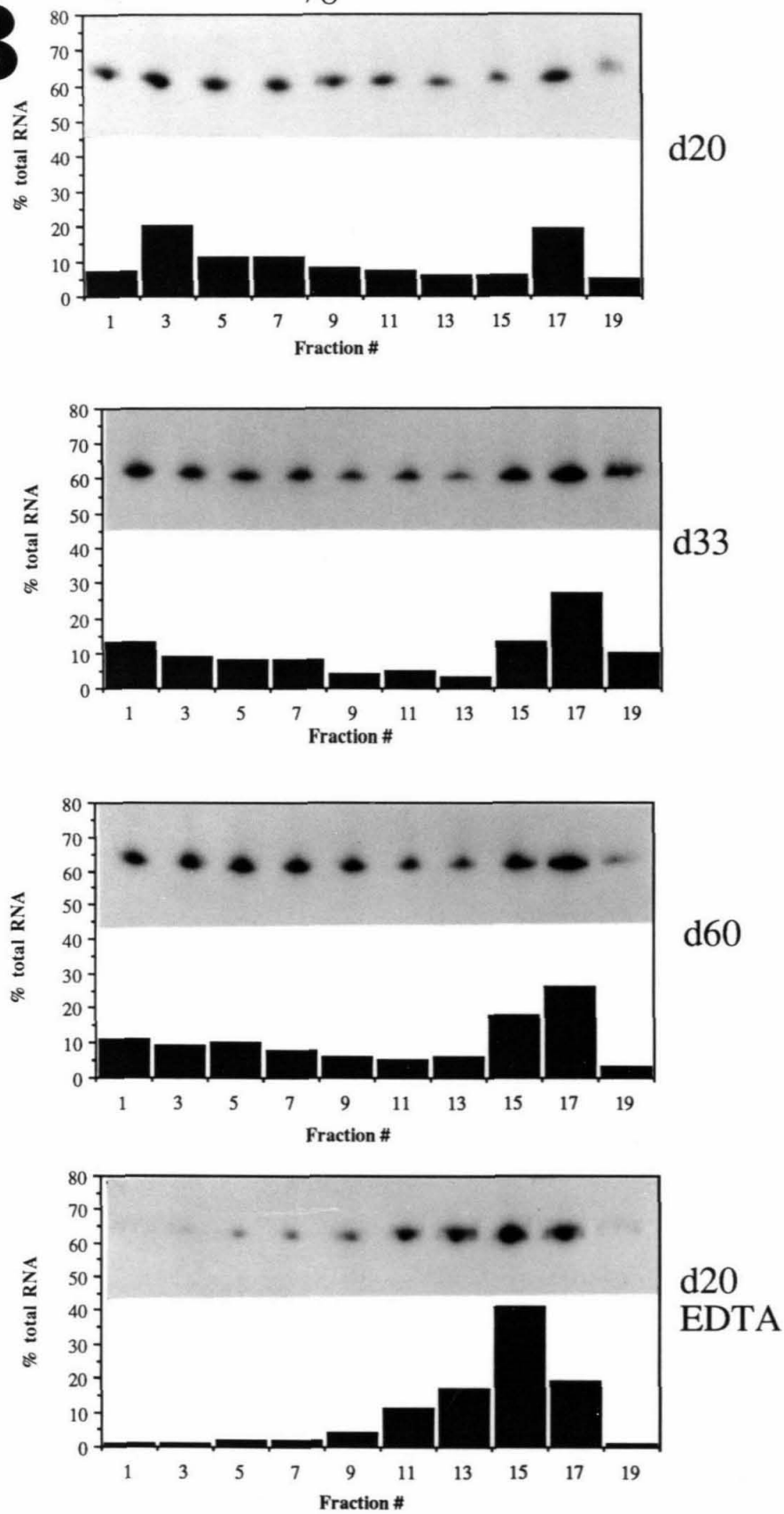
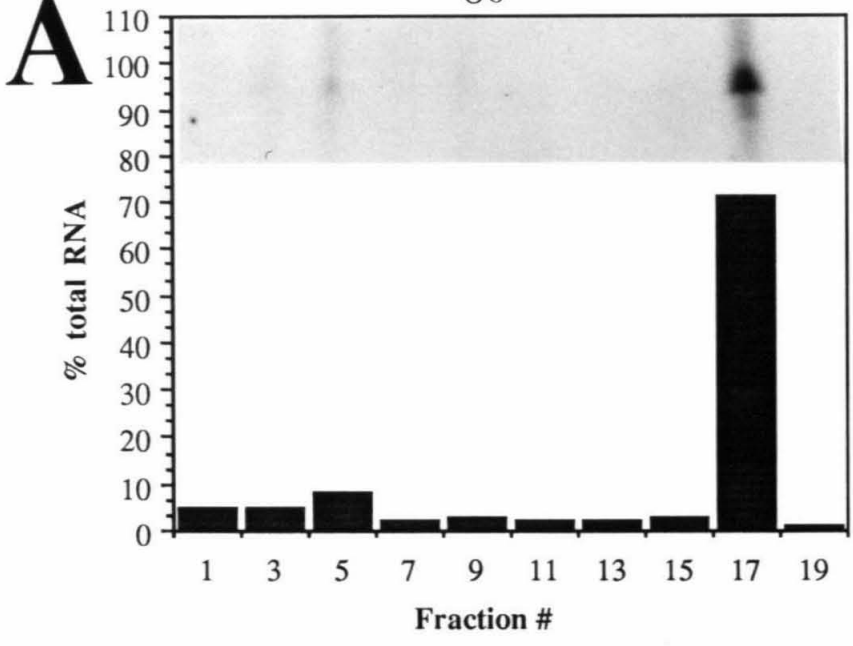
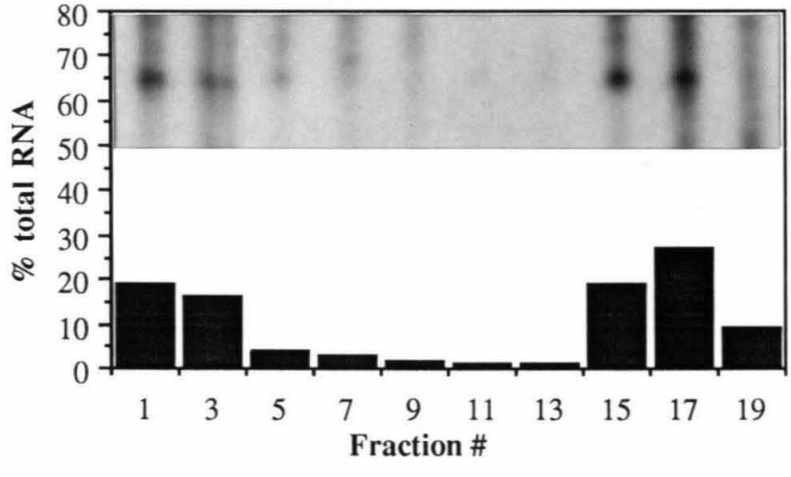


Figure 5. Polysomal profiles of *PGK-2*/CAT and endogenous *Pgk-2*. The same fractions of the gradients in Figure 4 were analyzed by RTpcr for RNA levels. An aliquot of each fraction was amplified using either CAT or *Pgk-2* oligonucleotide primers and the respective signal is plotted as the percent of the total CAT or *Pgk-2* signal from each gradient. The radiolabelled pcr product corresponding to each fraction is shown. A) Mouse *Pgk-2* polysome profiles, B) *PGK-2*/CAT transgene polysome profiles.

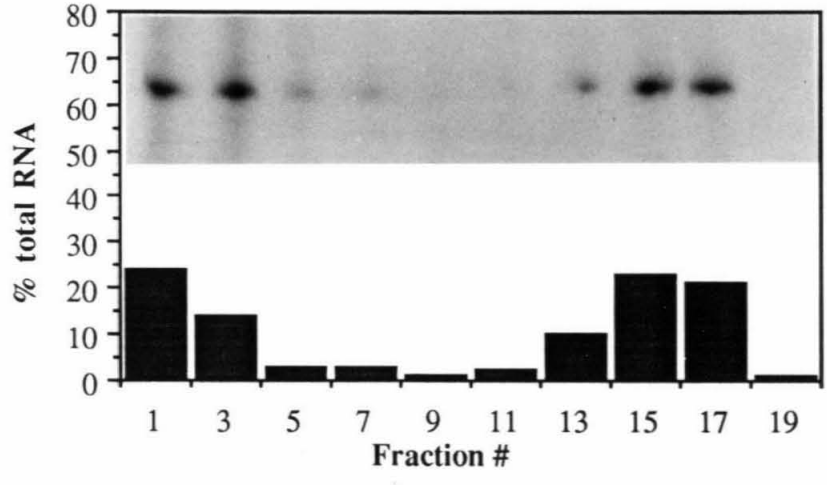
A


110
100
90
80
70
60
50
40
30
20
10
0
% total RNA
1 3 5 7 9 11 13 15 17 19
Fraction #

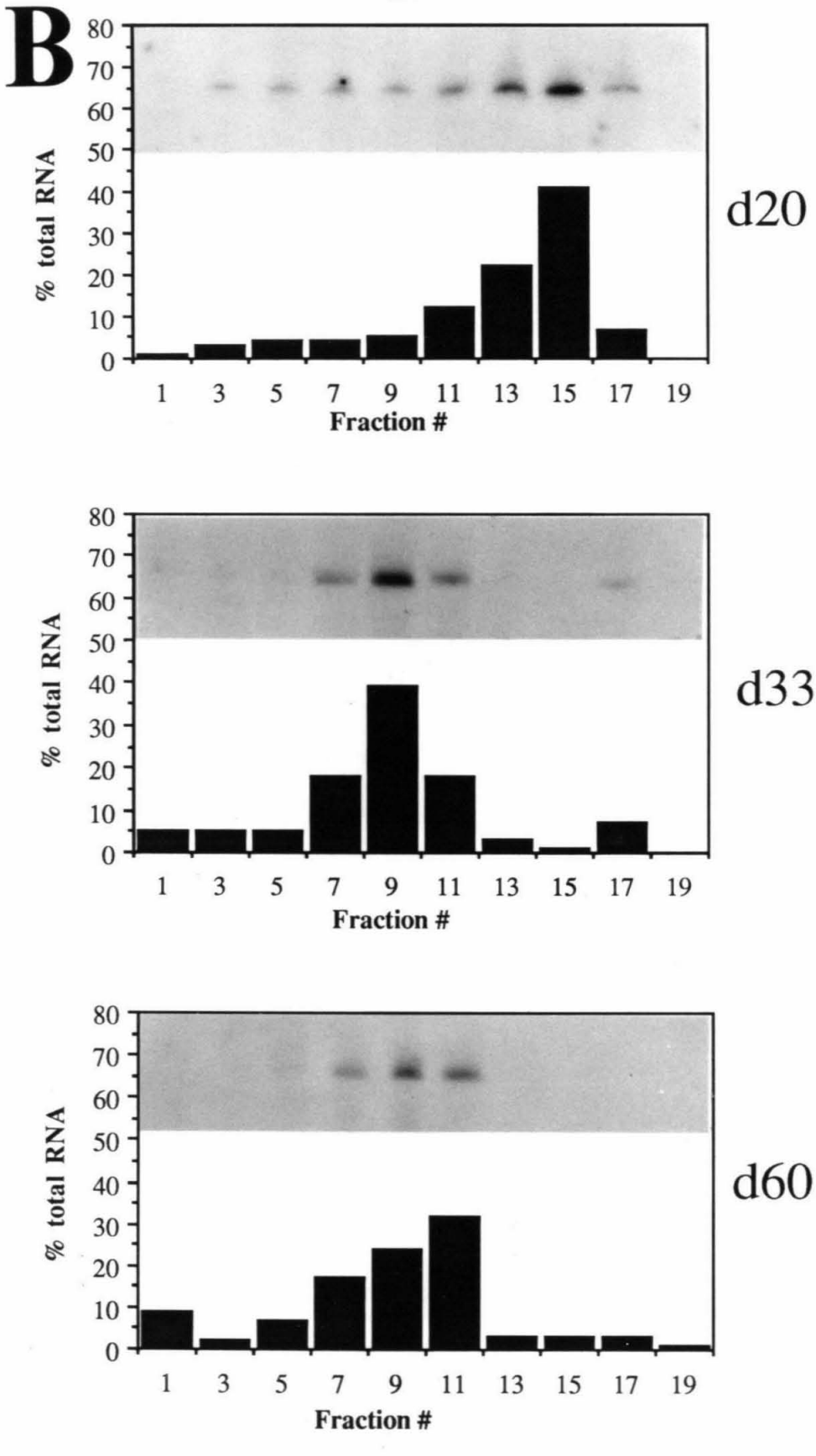
d20


80
70
60
50
40
30
20
10
0
% total RNA
1 3 5 7 9 11 13 15 17 19
Fraction #

d33


80
70
60
50
40
30
20
10
0
% total RNA
1 3 5 7 9 11 13 15 17 19
Fraction #

d60


B
% total RNA
Fraction #
d20
d33
d60

CONCLUSIONS

Transgenic studies.

Introducing foreign genes into the germline of mice is a powerful technique for analyzing the function and expression of genes. This approach to studying gene expression provides an environment very similar to that of the endogenous gene (including *in vivo* developmental signals and integration into the host chromatin), yet provides the ability to manipulate the transgene to assay for functional sequences. Transgenes, however, are unlike the endogenous locus in that they are only very rarely integrated in the correct genomic location (Brinster et al. 1989), and the copy number of the inserted gene is difficult to regulate (unpublished observations). The newly developed techniques for inserting modified genes into the endogenous chromosomal locus will provide a more natural setting for the analysis of gene function; however, currently these techniques are sufficiently labor-intensive as to preclude analyses as extensive as the one presented here.

Use of transgenes to study *PGK-2* gene expression.

Considerations for the analysis of genes in transgenic mice include designing the transgene so that it may be assayed in a simple and straightforward manner. To examine the expression of genomic fragments of any gene, it is essential to be able to distinguish between the transgene and the endogenous locus. At the time we began our analysis, only the human *PGK-2* locus was available. The

decision to use the human gene was made for two reasons. Presumably, the human gene would be different enough from the mouse gene that it could be identified in the mouse. This proved to be true and the expression of the human gene in the mouse was easily assayed (Chapter 2). Additionally, if the human gene was regulated properly, the sequences of both genes could be compared, yielding more information about functional elements. The conservation of regulated expression has led to the identification of candidate control sequences that may now be further scrutinized.

To further address the control of gene expression in the transgenic system, it was essential to isolate the controlling regions of the gene. This is generally accomplished by the use of reporter genes that code for easily assayable proteins. We chose to use the gene encoding chloramphenicol acetyltransferase (CAT) (and additionally a new reporter gene, firefly luciferase), and the use of these genes allowed the isolation of the sequences responsible for the control of expression. However, the use of a protein product to assay for transcriptional control is an indirect method of analysis, as the protein assays measure transcription and translation. This indirect method of expression analysis led us fortuitouslyto the identification of *PGK-2* translational regulation, as the protein accumulation rates were different from the transcript accumulation (Chapters 2 and 3).

These investigations serve to point out important considerations for analysis of gene expression in transgenic mice: first that reporter proteins may be misleading as an indicator of transcriptional activity, and second, transgenes serve as a powerful

tool to investigate cis-acting regions involved in both transcriptional and translational control of gene expression.

PGK-2 expression in the testis.

The results of this work present an interesting picture of the regulation of *PGK-2*. These studies suggest that a relatively small region of the 5' region of the gene is responsible for both its transcriptional and translational regulation. Transcription (implied by analysis of RNA levels) requires all or part of the 515 bp of the promoter of the gene. A subsequent study of *Zfp-35* expression in the testis refers to the study in Chapter 2 and notes a conserved sequence near the 5' end of the *PGK-2* promoter (Cunliffe et al. 1990). A study of the meiotic-specific expression of the proenkephalin gene notes conservation to another sequence in the functionally defined *PGK-2* promoter (Kilpatrick et al. 1990). Sequence comparison is a first step to identify possible control elements in the common expression of genes, although further studies must determine their role in control of gene expression. For *PGK-2*, studies have been initiated to search for the presence of DNA/protein interactions in the regions defined in this thesis.

The documented use of translational regulation in spermatogenesis, and earlier studies on translational regulation of *PGK-2* (Gold et al. 1983) led to our investigation of cis-acting regions involved in its translational control. The identification of the 5' untranslated sequences as the common element in transgenic messages that were subject to translational regulation implicates this

region in the control of translation. However, it has yet to be proven that this region is necessary for the observed effect. Further studies on this region will elucidate the role of the 5' untranslated sequences in the control of gene expression of *PGK-2*.

REFERENCES

Brinster, R. L., R. E. Braun, D. Lo, M. R. Avarbock, F. Oram and R. D. Palmiter. 1989. Targeted correction of a major histocompatibility class II E alpha gene by DNA microinjected into mouse eggs. *Proc. Natl. Acad. Sci. USA* **86**: 7087-7091.

Cunliffe, V., S. Williams and J. Trowsdale. 1990. Genomic analysis of a mouse zinc finger gene, zfp-35, that is up-regulated during spermatogenesis. *Genomics* **8**: 331-339.

Gold, B., H. Fujimoto, J. M. Kramer, R. P. Erickson and N. B. Hecht. 1983. Haploid accumulation and translational control of phosphoglycerate kinase-2 messenger RNA during mouse spermatogenesis. *Dev. Biol.* **98**: 392-399.

Kilpatrick, D. L., S. A. Zinn, M. Fitzgerald, H. Higuchi, S. L. Sabol and J. Meyerhardt. 1990. Transcription of the rat and mouse proenkephalin genes is initiated at distinct sites in spermatogenic and somatic cells. *Mol. Cell. Biol* **10**: 3717-3726.

Appendix

Measurement by quantitative PCR of changes in HPRT, PGK-1, PGK-2, APRT, MTase, and Zfy gene transcripts during mouse spermatogenesis

(This was published in *Nucleic Acids Research,* **18**: 1255-1259, 1989)

This paper was included to present the data on *Pgk-1*, and *Pgk-2* transcript levels in separated germ cells.

Nucleic Acids Research, Vol. 18, No. 5

Measurement by quantitative PCR of changes in HPRT, PGK-1, PGK-2, APRT, MTase, and Zfy gene transcripts during mouse spermatogenesis

Judith Singer-Sam, Murray O.Robinson[1], Anthony R.Bellvé[2], Melvin I.Simon[1] and Arthur D.Riggs
Division of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, [1]Caltech, Division of Biology, Pasadena, CA 91125 and [2]Departments of Anatomy and Cell Biology, Urology and the Center for Reproductive Sciences, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

Received October 12, 1989; Revised and Accepted February 6, 1990

ABSTRACT

A reverse transcriptase-polymerase chain reaction assay (RT-PCR) was used quantitatively to measure accumulated levels of RNA transcripts in total mouse RNAs derived from male germ cells at various spermatogenic stages. RNA levels for two X-linked enzymes, phosphoglycerate kinase (PGK-1) and hypoxanthine phosphoribosyl transferase (HPRT), both decrease during spermatogenesis, although the transcript levels decrease much more rapidly for PGK-1. RNA for the Y-linked ZFY (zinc finger protein) is elevated in all spermatogenic cell fractions tested, being particularly high in leptotene/zygotene spermatocytes and round spermatids. RNA for adenine phosphoribosyltransferase (APRT) increases 5-fold to a peak during late pachynema. RNA for PGK-2, undetectable in spermatogonial cells, increases at least 50-fold by the round spermatid stage. DNA (cytosine-5-)-methyltransferase (MTase) transcript levels are over an order of magnitude higher throughout spermatogenesis than in non-dividing liver cells.

INTRODUCTION

Many studies of gene expression are limited by the sensitivity of standard hybridization techniques, and, for this reason, several groups have recently made use of the polymerase chain reaction for the detection of specific mRNAs after reverse transcription to DNA (RT-PCR) (1, 2, 3, 4, 5, 6). Because of the increasing use of RT-PCR, we investigated the quantitative aspects of the assay. By limiting the amount of input template and the number of cycles of PCR, we obtained a signal which was linearly related to specific RNA levels over at least a thousand-fold range, with reproducibility adequate for any study where changes in RNA levels are greater than two-fold.

We have used the assay to investigate changes in specific RNA transcripts of X-linked and autosomal genes during mouse spermatogenesis. Our work was initiated as a study of the X chromosome inactivation which occurs during mouse spermatogenesis (7, 8, 9, 10, 11). In addition to the two X-linked genes, PGK-1 and HPRT, the Y-linked Zfy genes were also included in the study because they may be involved in spermatogenesis (12).

Three other genes were investigated. The autosomal gene for adenine phosphoribosyltransferase (APRT), was included in the study because, like HPRT, APRT is part of the purine salvage pathway and is expressed ubiquitiously at low levels. Phosphoglycerate kinase-2 (PGK-2), an autosomal, testis-specific isozyme of PGK-1, was included as a control because it is known to be expressed only in meiotic and post-meiotic male germ cells (13). DNA (cytosine-5-)-methyltransferase (MTase) is part of a gene silencing system probably involved in both X chromosome inactivation and genomic imprinting (14, 15, 16). We find that each of the genes studied shows a unique pattern of transcript level changes during spermatogenesis.

MATERIALS AND METHODS

Isolation of spermatogenic cells and purification of RNA

Fractionated germ cells at various stages of development were prepared as described (7, 8) from Swiss Webster mice (Charles River). In brief, seminiferous cords were isolated from testes of prepuberal mice, the germ cells were dissociated by collagenase and trypsin treatment, then seminiferous cells were separated by sedimentation velocity at unit gravity. Types A and B spermatogonia were isolated from 8-day-old mice, leptotene/zygotene and 'p17' pachytene spermatocytes from 17-day-old mice, and 'p60' pachytene spermatocytes, round spermatids and liver cells from mice 60 or more days old. Sertoli cells were isolated from 6-day-old mice.

Examination by light microscopy showed the fractions to be 80% or more of the designated cell type, with the possible exception of the leptotene/zygotene fraction. For each fraction contamination appeared to be with the developmentally adjacent cell types and Sertoli cells. See RESULTS for additional evidence regarding purity of the fractions.

Total RNA from each cell fraction, prepared by a method that removes DNA by differential precipitation (17), was kindly supplied by K. Thomas and P. Tomashefsky (9). The RNA was stored as an ethanol precipitate at −70°C. Prior to each

Table 1. Primers Used in the RT-PCR Assay

Gene	Primers Used	Product Size	Bases Spanned	Ref.
PGK-1	TAGTGGCTGAGATGTGGCACAG GCTCACTTCCTTTCTCAGGCAG	166 bp	1364 – 1529	(40)
PGK-2	ATTGTCATAATGTACTTGCTTCC GTTTCCTCCCTAGGAAGGACTGG	225 bp	1254 – 1478	(41)
HPRT	CGAGGAGTCCTGTTGATGTTGC CTGGCCTATAGGCTCATAGTGC	172 bp	685 – 856	(21)
APRT	CCAGCAGCACTAGGAACTGCTT AGGGTGTGTGGGACGTGTACAA	167 bp	2848 – 3014	(42)
MTase*	AGCCAGTTGTGTGACTTGGAAACCA ACCGTTGGCTTTTTGAGTGAGAGTG	141 bp	17 – 157	(43)
ZFY	AAGATAAGCTTACATAATCACATGGA CCTATGAAATCCTTTGCTGCACATGT	618 bp	1336 – 1953 (12)	(24)

*MTase: DNA (cytosine-5-)-methyltransferase

experiment aliquots of RNA were centrifuged with glycogen (40 μg) added as a carrier, rinsed in 70% ethanol, then resuspended either in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 10 mM NaCl, or directly in the PCR reaction mix.

The RT-PCR assay

Information about the oligonucleotides used as primers for the PCR reaction is given in Table I. Initially, assays were done as previously described (1). In later experiments, RNA was mixed with the PCR reaction mix (10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% (w/v) gelatin, 200 μM dNTPs, and primers at 1 μM each), and RNasin (Promega, 40 U), in a final volume of 50 μl. After the addition of AMV reverse transcriptase (2 U, Life Sciences), and Taq polymerase (2 U, Cetus Perkin-Elmer), the samples were placed in a thermal cycler at 50°C for 8 min, followed by PCR at 95°C for 1 min, 60°C for 2 min, and 73°C for 2.5 min, except in the case of PGK-2, where the renaturation step was done at 47°C. After PCR, a 20 μl aliquot was subjected to electrophoresis in a 2% agarose gel containing 0.4 μg/ml ethidium bromide in 1× TBE buffer (18).

Quantitation

After gel electrophoresis, the amount of PCR product in each lane was determined in either of two ways: 1) The gel was photographed under UV light onto technical film (Kodak, 4415), with several exposures taken to ensure that the signal from each band was in the linear range of film response. Densitometry was done by use of a video densitometer (Biorad, Model 620), and peak areas were analyzed by means of the 1-D Analyst programs (Biorad). 2) The DNA was transferred onto a nylon membrane (Genetrans 45, Plasco, Inc.) by vacuum blotting (1). Hybridization was done as described (1), at 60°C. Each autoradiogram was analyzed by densitometry as described above; alternatively, the radioactivity in each band was determined directly by means of the AMBIS Radioisotope Scanning System II (Automated Microbiology System, Inc.).

RESULTS

Quantitation by the RT-PCR assay

Figure 1 shows a dilution series of total RNA from liver or Sertoli cells, assayed for PGK-1 transcripts by RT-PCR. The signal seen for PGK-1 clearly reflects the amount of RNA loaded. To further

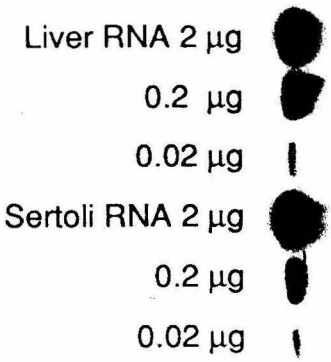


Figure 1. Assay by RT-PCR of PGK-1 transcript levels in mouse liver and Sertoli cells. Samples of total RNA (0.02, 0.2 or 2 μg) were reverse transcribed, then PCR was performed for 20 cycles with the primers for PGK-1 shown in Table 1. The products from the PCR reaction were separated by gel electrophoresis, transferred to a nylon membrane, and hybridized with a ^{32}P-oligonucleotide probe specific for PGK-1 . The probe used was 5' CTCAGCTCGTCTTAC-TGCATCA, beginning at base position 1405 of mouse PGK cDNA (40). The bands seen correspond to the expected size of 166 bp.

investigate quantitation by RT-PCR, filter bound radioactivity was determined; figure 2 shows data accumulated from several experiments such as the one shown in figure 1. A linear correlation between the amount of template RNA and the PCR signal is seen over a thousand-fold range, with an average deviation from the best fit line of 40%. To obtain a linear response to template concentration, the primers, nucleotides, and polymerase must remain in large excess. Therefore, the appropriate number of PCR cycles must be empirically determined for each primer set, since the number of cycles taken to reach saturation depends on primer efficiency and the abundance of the target transcript. For the experiments reported here, total RNA used varied from 0.25–2 μg, and the cycle number varied from 20–26 cycles. For each set of primers, a dilution series was used to determine that the results obtained were in the linear range of the assay. Since the exposure time for the autoradiogram shown in Figure 1 was only 40 min with no enhancing screens, it is clear that the ultimate sensitivity, even

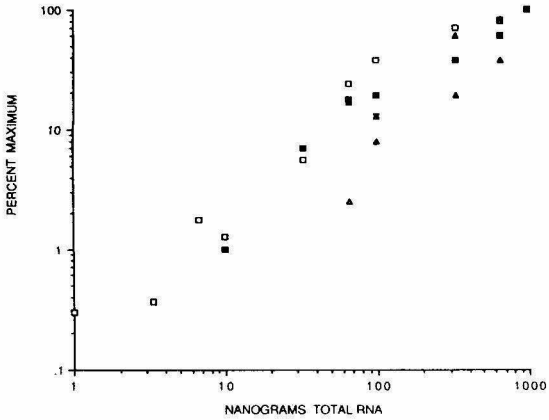


Figure 2. PCR signal as a function of the amount of template RNA. Total mouse liver RNA was diluted in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 10 mM NaCl, and added to the PCR reaction mix in the amounts indicated. RT-PCR, gel electrophoresis, hybridization and autoradiography were performed as indicated in Materials and Methods. The cpm in each band were determined by means of a radioisotope scanning system, and normalized relative to the value obtained for 1 μg of RNA. Each symbol represents an independent experiment. Taq polymerase was added either with reverse transcriptase (open sqares and triangles) or after the reverse transcription step (closed squares and triangles).

without increasing cycle numbers, is much greater than needed for the work reported here.

Measurement of specific RNA transcripts during spermatogenesis

The RT-PCR method was used to assay for specific transcripts in total RNA extracted from cells at various stages of spermatogenesis. Figure 3 shows results of a typical experiment, and Figure 4 summarizes the results obtained for all transcripts assayed. In each case the data were normalized to the stage of maximum expression. For all transcripts studied, the major PCR products were the size expected, and the identity of the product was confirmed by the presence and proper location of restriction enzyme sites and/or hybridization to an internal oligonucleotide probe. As shown in Figs. 1, 3 and 4A, transcripts from the PGK-1 gene decrease to about 5% of somatic cell levels (liver or Sertoli cells) in leptotene/zygotene spermatocytes, and decrease to even lower levels at later stages. In the mixture of type A and B spermatogonia, the level of PGK-1 RNA is already reduced relative to liver. As a control, we measured PGK-2 RNA (Fig 4A). PGK-2 transcripts are first detectable in leptotene/zygotene spermatocytes, earlier than had been previously reported (9), and continued to increase, as expected (13), at least until the round spermatid stage.

Purity of RNA samples

DNA contamination can be ruled out for several reasons. First, the PGK-1 signal decreases to less than 5% of somatic levels after the leptotene/zygotene stage (Fig. 3). Since the PGK-1 primers amplify genomic DNA at least as efficiently as RNA (our unpublished data), it follows that the post-leptotene/zygotene spermatocyte RNA samples are free of detectable DNA. Secondly, a PGK-2 signal is not seen in somatic cells or type A/B spermatogonial preparations; therefore these samples are free of detectable DNA. Since the sensitivity of the assay allows the same samples to be used to measure the other gene targets, it follows that DNA contamination is insignificant. Thirdly, the addition of up to 10% contaminating DNA was found not to affect the results for HPRT.

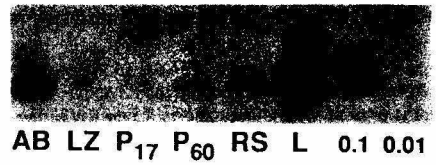


Figure 3. Changes in PGK-1 transcript levels during spermatogenesis. The experimental conditions are the same as those shown in Fig. 1. Liver RNA (L), diluted prior to RT-PCR, is shown in the three right lanes. AB, mixed type A and B spermatogonia; LZ, mixed leptotene and zygotene spermatocytes; P17 and P60, pachytene spermatocytes; R, round spermatids.

Even though the cell populations we isolated are known in general to have varying degrees of contamination with other cell types (7, 8), interference by other RNAs can be ruled out, in large part, by the nature of the results we obtained. PGK-1 is a key enzyme of the glycolytic pathway and is ubiquitously expressed in all somatic cell types (19). Since the leptotene/zygotene samples, which might have been the most contaminated with Sertoli cells, have only about 5% of the PGK-1 content of liver and other somatic cells, it follows that contamination with somatic cells expressing PGK-1 is very low. Sertoli cells have about the same level of PGK-1 RNA as liver cells (see Fig. 1), so contamination with Sertoli cells must be 5% or less. Similar reasoning establishes that leptotene/zygotene samples are also not significantly contaminated with stages producing high levels of PGK-2, i.e., pachytene and round spermatids. In addition, the same five samples we used have also been studied by Thomas et al. (9); they assayed by Northern blotting for a Sertoli cell-specific sulfated glycoprotein transcript and concluded that Sertoli cell contamination was not significant.

Transcript levels for the PGK-1, HPRT and APRT genes

The results shown for PGK-1 in Figures 3 and 4A are consistent with cytological data indicating that the single X chromosome in male spermatogenic cells is inactivated early in meiotic prophase. Transcript levels are reduced at least 50-fold by pachynema. The PGK-1 gene probably begins to be shut off during the spermatogonial stage, since our studies (not shown) indicate that primitive A spermatogonia still contain PGK-1 transcripts in amounts comparable to liver or Sertoli cells. Unexpectedly, HPRT decreases during spermatogenesis much more slowly than does PGK-1. As shown in Figure 4B, HPRT RNA in type A/B spermatogonia may be somewhat reduced relative to liver or Sertoli cells, but significant reduction does not occur until later stages and then the level only falls 5 to 10-fold.

APRT and HPRT both function in the purine salvage pathway, are ubiquitously expressed at low levels, and have similar G+C-rich promoters (20, 21). However, as seen in Fig 4B, the expression patterns are quite different between the two genes, with APRT increasing 5-fold to a peak during pachynema as HPRT is decreasing. It is clear that control of transcript levels, and thus the control of the ratio of transcription to degradation, is different between HPRT and APRT.

Transcript levels for the DNA methyltransferase and Zfy genes

The highest MTase RNA levels are seen in leptotene/zygotene and pachytene spermatocytes, but the main conclusion is that the transcript is present throughout meiosis. Liver RNA, the somatic cell control shown in Fig. 4C, shows no detectable MTase RNA under conditions of the assay. It is known that normal liver has very low levels of MTase; lymphocytes and other mitotically active cells express at least 10-fold higher levels of MTase (22). The levels we find in spermatogenic cells are comparable to those in Sertoli cells and hybridoma cells in culture (not shown).

Fig.4C also shows the results obtained for Zfy. All five fractions tested have increased RNA levels over those found in liver cells, with leptotene/zygotene spermatocytes and round spermatids showing maximum levels. In all cases only one PCR product was found, of the size expected from mRNA of both the *zfy-1* and *zfy-2* genes in *Mus musculus* domesticus strains (12, 23, 24); thus, we may be measuring the sum of the two transcripts.

DISCUSSION

We have described here the quantitative use of RT-PCR to measure changes in specific RNA levels. When applied to DNA, the PCR reaction can be used quantitatively (as little as 4% standard deviation (19)), even without especially prepared internal standards. When applied to RNA, a reverse transcription step is necessary, and this probably introduces some additional variation. However, our studies confirm that the PCR product is linearly related to the template RNA concentration, and the reaction is sufficiently reproducible for the reliable detection of greater than two-fold differences (see Figures 2 and 4). The assay is very sensitive, requiring as little as 1/1000th the amount of total RNA used for Northern blots (Fig. 1; ref.(9)).

RT-PCR was used to measure relative levels of accumulation of several specific transcripts at different stages in spermatogenesis. One advantage of the assay is that it requires so little material that we have been able to use the same samples to ascertain RNA purity and to measure RNA transcript levels for all of the genes studied. We have not measured absolute transcript levels or rates of transcription, although, in the case of the PGK-2 gene, we have been able to observe the *initiation* of transcription.

Two of the RNAs we assayed, PGK-1 and HPRT, are X-linked. In 1965 Monesi (25) showed that the X chromosome was present as a heteropycnotic body by lepotonema of male meiotic prophase. He also showed by autoradiography after ^{3}H-uridine labeling, that both the X and Y chromosomes of mouse spermatocytes appear transcriptionally inactive at leptonema, and remain so throughout prophase. However, the direct demonstration of inactivation of specific X-linked genes expressed at low levels was not feasible. Therefore, the exact timing of X chromosome inactivation during spermatogenesis, and the relationship of X-inactivation to normal sperm development were unknown (26, 27).

In this study we have assayed for transcript levels of three genes on the sex chromosomes. We find that PGK-1 transcript levels decrease rapidly early in meiotic prophase, starting with type A/B spermatogonia, and drop rapidly to very low levels, consistent with previous cytological data. HPRT transcripts also decrease, but at a much slower rate. If transcription ceases, the message for HPRT is very stable, since transcript levels decrease only 5 to 10-fold over a period of 13 days. There are many long-lived mRNAs in meiotic cells (28, 29), and HPRT mRNA could be among them.

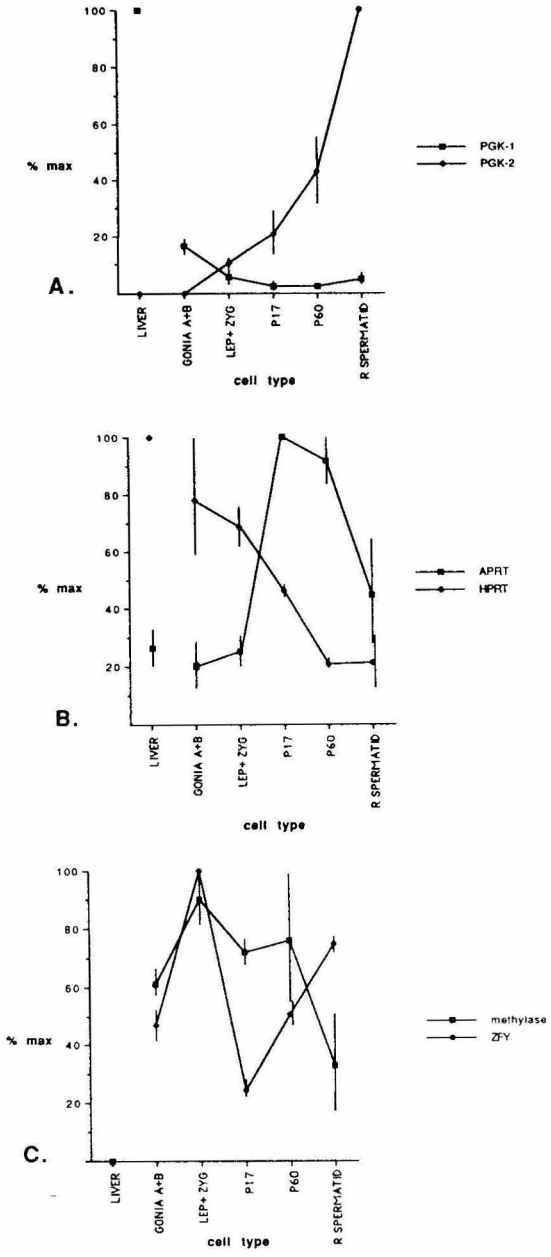


Figure 4. RNA levels during spermatogenesis. Primers and conditions used for the RT-PCR assay are given in Table I and Materials and Methods. Each experiment was done at least in duplicate and the range of values is shown by the error bars. 4A) PGK-1 and PGK-2. 4B) APRT and HPRT. 4C) DNA methylase and Zfy.

We cannot rule out some continued transcription of the HPRT gene. Allsop, et al.(30) showed HPRT protein levels to increase during rat spermatogenesis, and suggested that HPRT may be necessary for spermatogenesis to occur. HPRT$^-$ mice however, undergo normal spermatogenesis (31, 32), so there is no obvious need for stabilization and preservation of HPRT message.

We have found Zfy transcripts to be present in all spermatogenic cells, and particularly high in leptotene/zygotene spermatocytes and round spermatids (Fig. 4C). Two Y-linked Zfy transcripts are known to be expressed in adult mouse testis, *Zfy-1* and *Zfy-2*, the former giving a relatively faint signal (12, 24). With *Mus musculus* musculus strains, the two transcripts are distinguishable by RT-PCR, giving different sized products with the Zfy primers shown in Table I (12, 23). In contrast, with the *Mus musculus* domesticus strains we used for this study, only one PCR product was observed. Two peaks of Zfy expression are seen, and it is tempting to speculate that *Zfy-1* and *Zfy-2* may be transcribed at different times during spermatogenesis. In any event, our finding that Zfy transcripts are present in relative abundance in meiotic spermatocytes and round spermatids is consistent with a role for one or both Zfy genes during sperm development (12, 24, 33, 34).

Recent studies have shown that many genes, perhaps representing 10% of the genome, function differently depending on parental origin (16, 35). This phenomenon of genomic imprinting probably involves DNA methylation. Differences in methylation between male and female gametes have been detected for repetitive DNA sequences (36, 37, 38), and oocyte DNA is globally less methylated than is sperm DNA (36). Therefore, changes in DNA methylase activity during gametogenesis might be important, but only one relevant study has been reported. Reddy and Reddy (39) measured MTase specific activity during maturation of the rat testis, finding activity for the whole testis to be highest at 20–30 days, coinciding with the appearance of spermatids. We find that MTase RNA levels are high in meiotic germ cells, but decreases postmeiotically, so the potential is present for methylation changes relevant to genomic imprinting.

ACKNOWLEDGEMENTS

We thank Robert L. Tanguay for excellent technical assistance, and John Rossi, Paul Salvaterra, Anna Wu, and Philip Tomashefsky for their kind help and suggestions. This work was supported by NIH grant 1 R01 AG08196-01 to ADR, an American Cancer Society grant to MIS, and by NICHHD grant 05077 to ARB.

REFERENCES

1. Murakawa, G.J., Zaia, J.A., Spallone, P.A., Stephens, D.A., Kaplan, B.E., Wallace, R.B. and Rossi, J.J. (1988) DNA, **7**, 287–295.
2. Rappolee, D.A., Wang, A., Mark, D. and Werb, Z. (1989) J. Cell. Biochem., **39**, 1–11.
3. Harbarth, P. and Vosberg, H.-P. (1988) DNA, **7**, 297–306.
4. Doherty, P.J., Huesca-Contreras, M., Dosch, H.M. and Pan, S. (1989) Analyt. Biochem., **177**, 7–10.
5. Becker-André, M. and Hahlbrock, K. (1989) Nucleic Acids Res., **17**, 9437–9446.
6. Gilliland, G., Perrin, S., Blanchard, K. and Bunn, H.F. (1990) Proc. Natl. Acad. Sci. USA, **in press**.
7. 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.
8. Bellvé, A.R., Millette, C.F., Bhatnagar, Y.M. and O'Brien, D.A. (1977) J. Histochem. Cytochem., **25**, 480–494.
9. Thomas, K.H., Wilkie, T.M., Tomashefsky, P., Bellvé, A.R. and Simon, M.I. (1989) Biology of Reproduction, **41**, 729–739.
10. Gartler, S.M. and Riggs, A.D. (1983) Ann. Rev. Genet., **17**, 155–190.
11. Monesi, V., Geremia, R., D'Agostino, A. and Boitani, C. (1978) In Current Topics in Developmental Biology. pp. 11–36.
12. Nagamine, C.M., Chan, K., Kozak, C.A. and Lau, Y.-F. (1989) Science, **243**, 80–83.
13. Gold, B., Fujimoto, H., Kramer, J.M., Erickson, R.P. and Hecht, N.B. (1983) Dev. Biol., **98**, 392–399.
14. Riggs, A.D. (1975) Cytogenet. Cell Genet., **14**, 9–25.
15. Monk, M. (1986) BioEssays, **4**, 204–208.
16. Monk, M. (1988) Genes Dev., **2**, 921–925.
17. Cathala, G., Savouret, J., Mendez, B., West, B.L., Karin, M., Martial, J.A. and Baxter, J.D. (1983) DNA, **2**, 329–35.
18. Maniatis, T., Fritsch, F. and Sambrook, J. (1982) Molecular Cloning. Cold Spring Harbor Laboratory, New York.
19. Singer-Sam, J., Yang, T.P., Mori, N., Tanguay, R.L., LeBon, J.M., Flores, J.C. and Riggs, A.D. (1989) In Clawson, G., Willis, D., Weissbach, A. and Jones, P. (eds.), Nucleic Acid Methylation. 128. Alan R. Liss, New York, pp. 285–298.
20. Dush, M.K., Briggs, M.R., Royce, M.E., Schaff, D.A., Khan, S.A., Tischfield, J.A. and Stambrook, P.J. (1988) Nucl. Acids Res, **16**, 8509–8524.
21. Melton, D.W. (1987) In Maclean, N. (ed.), Oxford Surveys on Eukaryotic Genes. 4. Oxford, pp. 34–76.
22. Zucker, K.E., Riggs, A.D. and Smith, S.S. (1985) J. Cell. Biochem., **29**, 337–349.
23. Nagamine, C.M., Chan, K., Hake, L.E. and Lau, Y.-F.C. (1990) Genes Dev., **in press**.
24. Mardon, G. and Page, D.C. (1989) Cell, **56**, 765–770.
25. Monesi, V. (1965) Exp. Cell Res., **39**, 197–224.
26. Lifschytz, E. and Lindsley, D.L. (1972) Proc. Natl. Acad. Sci. USA, **69**, 182–186.
27. VandeBerg, J.L. (1983) J. Exp. Zool., **228**, 271–286.
28. Hecht, N.B. (1986) In Rossant, J. and Pedersen, R. (eds.), Experimental Approaches to Mammalian Embryonic Development. Cambridge University Press, Cambridge, pp. 151–193.
29. Willison, K. and Ashworth, A. (1987) Trends in Genetics, **3**, 351–355.
30. Allsop, J. and Watts, R.W.E. (1986) Differentiation, **32**, 144–147.
31. Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M. (1987) Nature, **326**, 292–295.
32. Kuehn, M.R., Bradley, A., Robertson, E.J. and Evans, M.J. (1987) Nature, **326**, 295–298.
33. McLaren, A. (1988) Trends Genet., **4**, 153–157.
34. Koopman, P., Gubbay, J., Collignon, J. and Lovell-Badge, R. (1989) Nature, **342**, 940–942.
35. Solter, D. (1988) Annu. Rev. Genet., **22**, 127–146.
36. Monk, M., Boubelik, M. and Lehnert, S. (1987) Development, **99**, 371–382.
37. Monk, M. (1987) Nature, **328**, 203–204.
38. Sanford, J.P., Clark, H.J., Chapman, V.M. and Rossant, J. (1987) Genes Devel., **1**, 1039–1046.
39. Reddy, P.M. and Reddy, P.R. (1988) Biochem Int, **16**, 543–547.
40. Mori, N., Singer-Sam, J., Lee, C.-Y. and Riggs, A.D. (1986) Gene, **45**, 275–280.
41. Boer, P.H., Adra, C.N., Lau, Y.F. and McBurney, M.W. (1987) Mol. Cell. Biol., **7**, 3107–3112.
42. Dush, M.K., Sikela, J.M., Khan, S.A. and Tischfield, J.A. (1985) Proc. Natl. Acad. Sci. USA, **82**, 2731–2735.
43. Bestor, T.J., Laudano, A., Mattaliano, R. and Ingram, V. (1988) J. Mol. Biol., **203**, 971–983.