Antibody genes, Oncogenes and Antisense Genes

Thesis by

Stuart Kilsu Kim

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1985

(Submitted Dec. 1984)

For Jonathan

Acknowledgements

I'd like to acknowledge some of the people who helped do the experiments. Sunny Phil Patten and Eric Sinn, the human ball of energy, worked on somatic hypermutation. Lisa Flitz, Susanna Chan, Paul Mueller and Kathryn Calame assisted later on.

I'd also like to thank some of the characters who have made life at Caltech rewarding. From the early days, Martha Zuniga, Mitch Kronenberg, Jerry Siu and Tim Hunkapiller first come to mind. Jamie Lee was a good friend from our overly serious softball team and gentleman Chuck Reel was a lunchtime buddy. Mark Garfinkel has been helpful with many scientific discussions. Chris Martin and Jan Hoh have been good friends and made 255 S. Hill an excellent place to live. Finally, I'd like to specially thank George Gaines for the bike rides, the Vegas trips, the aerobics showdowns, the Federation, the Hungarian Beauty Pageant, and all the rest.

I'd like to thank Anne Villeneuve for her love.

I've come to know some truly excellent scientists here at Caltech and much of my education has been emulation. In addition to being a fine scientist, Lee Hood is a good leader and a fair judge. He was one of the rare good guys back when antibody genes were in vogue. Norman Davidson has a pure interest in how nature works. I think that if one could capture just some of his curiosity and interest in things biological, one would be assured of a rewarding academic career. Finally, I'd like to specially thank three scientists who have helped the most: Mark Davis, Stewart Scherrer and Barbara Wold. Mark was my mentor for my first two years. He taught me most of what I know about cloning and antibody genes. Stewart is a good friend and a great scientific coach. I cannot thank him enough for the helpful

advise for so many of my experiments. In addition, I'd like to mention his idealism on the reasons for pursuing academic science; I hope I remember this in the future. Most of all, I'd like to thank Barbara. Although difficult and often daring, her experiments are usually insightful and sometimes rewarding.

Abstract

There are a number of mechanisms involved in producing a diversity of antibodies including multiple germline genes, somatic gene rearrangement, somatic hypermutation and combinatorial association. By the process of somatic hypermutation, one immunoglobulin gene in the germline can be mutated to produce many different genes in B cells. In chapter 2, this process is characterized. It was found that phosphorylcholine binding antibodies are encoded by one germline V_H gene segment. In B cells, this V_H gene segment may have extensive point mutations, many of which are silent, indicating the presence of some somatic hypermutational mechanism. Only the V_H gene was found to be mutated indicating that the mutational mechanism was specific for V_H genes.

One way to study somatic immunoglobulin gene rearrangements, presented in chapter 3, might be to characterize rearrangements which are not easily explained. Immunoglobulin gene rearrangement was thought to exclusively involve immunoglobulin genes. However, some immunoglobulin genes can reproducibly rearrange with other DNA sequences. Insight into the basis of these rearrangements was uncovered by identifying the chromosomal origin of the non-immunoglobulin rearranging DNA. This DNA originated on chromosome 15 whereas the immunoglobulin gene originated on chromosome 12. The juxtaposition of these sequences is common in plasmacytomas but rare or absent in normal B cells suggesting that it is involved in tumorigenesis. For example, it may be that aberrant immunoglobulin rearrangements can activate a cellular oncogene resulting in a plasmacytoma. This possiblity was supported by results from other laboratories when it was found that the non-immunoglobulin rearranging DNA contained the cellular homologue of the myc oncogene.

To understand lymphocyte tumorigenesis, it would be useful to understand the function of the c-myc gene product in normal and transformed cells. One way to begin is to determine which types of cells express the c-myc gene. This approach was employed in chapter 5 and it was found that the c-myc gene is expressed in dividing, but not resting, lymphocytes. One possible function for the c-myc gene product is that it functions in cellular proliferation.

Another way to study the function of the c-myc gene product would be to prevent expression of the rearranged c-myc gene in plasmacytomas. For example, it may be possible to inhibit the synthesis of the c-myc gene product by antisense c-myc RNA. If the antisense RNA can hybridize to the c-myc RNA in vivo, synthesis of myc protein may be prevented. A test case, in which antisense TK RNA is used to inhibit TK expression, is presented in chapter 6. In L cells, high levels of antisense TK RNA expression were capable of inhibiting TK activity. The mechanism of inhibition involves RNA:RNA hybridization since double stranded RNA was formed. If this test case can be applied to other instances, it may be possible to use antisense RNA to inhibit the synthesis of a particular gene product and thus study its cellular function.

Table of Contents

Introduction

About ten years ago, it became possible to isolate genes from higher eukaryotes as molecular clones. Of the universe of genes, what are the criteria for choosing which ones to study? What are the technical limitations governing which experimental approaches will yield fruitful information? This introduction outlines some of the ways in which genes from higher eukaryotes, especially mammals, have been studied at the molecular level.

One general approach is to isolate molecular clones of genes which encode well studied proteins. For example, the genes encoding beta-globin, histone and the immunoglobulin lambda chain were among the first molecular clones to be identified from higher eukaryotes (Maniatis et al., 1976; Goldberg and Hogness, pers. comm. 1978; Tonegawa et al., 1977). The fact that these genes are abundantly expressed in certain cells simplified their isolation as cDNA clones. Once isolated, questions could be answered regarding their structure, expression and evolution.

An ever expanding array of genes can now be isolated as molecular clones including genes whose mRNA and gene product are present in scarce quantities. Eventually, it may be possible to isolate molecular clones of genes for the vast majority of proteins with identified functions. For example, in the microchemical facility, it is planned to routinely sequence trace quantities of proteins and, from this data, to synthesize nucleic acid probes to enable the cloning of their genes (Hunkapiller et al., 1984). The bottleneck of scientific advancement may have shifted, from isolation of molecular clones, which is now becoming routine, to the biochemical identification of proteins with a given function, which is becoming more complex.

Another approach is to study the molecular genetics of genes which encode unidentified proteins. Such an approach is complementary to the one discussed above since biochemical analysis remains difficult or impossible in many cases. Of particular interest is the question of how one might study the cellular function of a gene from a higher eukaryote starting only with a molecular clone. In chapter 3, a molecular clone of a rearranged c-myc gene was isolated from a murine plasmacytoma. One pressing question was to determine the role of this oncogene in B lymphocyte development and tumorigenesis. At the time, practically nothing was known about the c-myc protein or its function. Since cases like this are becoming more frequent, it would be instructive to discuss some of the available experimental avenues.

Gene expression which is regulated in a specific way would show when that gene functions and might also suggest the nature of that function. For example, in chapter 5, it was found that c-myc expression correlated with cellular proliferation suggesting that it may be involved in growth control. In other instances, genes specific to neurons have been studied (Sutcliffe et al., 1983; Anderson and Axel, pers. comm.) as well as those which are specific to T lymophocytes (Hedrick et al., 1984a; Yanagai et al., 1984). These genes may be involved in neural or immune function respectively.

In many cases, availability of a molecular clone of a gene can assist biochemical analysis of its gene product. In some cases, one might even begin a biochemical analysis of previously unidentified proteins. For example, using information obtained from molecular clones of the Ubx gene, which is encoded in the bithorax complex in *Drosophila melanogaster*, antisera specific to its protein product were prepared (White and Wilcox, 1985; Beachy et al., 1985). Immunocytochemical studies using these antisera suggested that this protein is localized to the nucleus. The possibility that the Ubx protein functions in the

nucleus enables one to make educated guesses as to its function, including its involvement in regulating gene expression.

It is also possible to employ molecular clones to specifically alter *in vivo* gene expression. Gene expression can be increased using DNA mediated gene transfer; gene activity can be decreased using antibodies or antisense RNA. Molecular approaches such as these may be especially useful in those cases which are not easily amenable to classical genetics.

Using DNA mediated gene transfer, DNA can be introduced into mammalian cells and the newly introduced genes can be expressed (Wigler et al., 1977). Gene transfer has a multitude of uses one of which is to introduce molecular clones into cells and, from the resulting phenotype, to deduce the function of the encoded gene product. However, this approach is limited in that it requires that the gene of interest have a dominant effect. There are many genes for which this might not be the case. Some genes are expressed constitutively so that introduction of extra copies of these genes, by gene transfer, may not cause a change in cellular behavior, even if the newly introduced genes are overexpressed. Also, some genes only function when expressed in concert with other genes. For example, they may participate in a pathway or the functional complex may be composed of several proteins. In these cases, expression of one gene would not be effective in lieu of expression from other genes in the pathway or the complex. Thus, there may be many cases in which the use of gene transfer to study gene function may be fruitless.

A complementary approach is to study the effect of diminished gene activity. One possibility is to employ antibodies which can bind and neutralize a particular gene product *in vivo* (Mabuchi and Okuno, 1977; Yamaizumi et al., 1978; Antman and Livingstone, 1980; Lin and Feramisco, 1981; Burke and Warren, 1984). For example, in early frog embryos, microinjection of mRNA encoding anti-gap

junction antibodies caused certain cellular lineages to fail to develop suggesting that intercellular communication via gap junctions is crucial for early development. In using antibodies in this way, the following points are relevant. Antibodies may be compartmentalized when introduced into cells and therefore may not bind proteins present in other cellular compartments. Also, introduction of antibodies into one or a few cells precludes biochemical analysis of a large population of cells.

Another way to reduce gene activity is to use antisense RNA. If the coding region of a gene is inverted, the opposite strand of DNA is transcribed resulting in complementary or antisense RNA. Potentially, this RNA can hybridize to sense RNA *in vivo* and prevent translation. There are precedents in *E. coli* which suggest that this may be so. Replication of the Col E1 plasmid, which requires an RNA primer, is regulated by antisense RNA (Tomizawa et al., 1981; Tomizawa, 1984). RNA complementary to the primer is formed by transcription of the noncoding DNA strand. Plasmids which produce this antisense RNA are capable of inhibiting replication in *trans* suggesting that hybridization of antisense RNA with primer RNA has interfered with replication. Second, the transposase gene from IS10 can be bidirectionally transcribed forming sense and antisense RNA (Simons and Kleckner, 1983). The presence of antisense RNA inhibits transposase expression. Third, high osmotic pressure induces the expression of OMP C mRNA and antisense OMP F RNA (Mizuno et al., 1984). As a result, OMP F expression is reduced. In these last two examples, it appears that antisense RNA is capable of hybridizing to sense RNA and thus preventing protein synthesis.

In eukaryotes, a small nuclear RNA, U1, has sequences complementary to those found at the 5' ends of introns. It is thought that base pairing between U1 and pre-mRNA is involved in RNA splicing (Lerner et al., 1980; Rogers and Wall,

1980). In this example, double stranded RNA is a requirement for, rather than an inhibitor of, biological function.

These examples suggest that, under certain circumstances, antisense RNA can hybridize to sense RNA *in vivo*. In at least some cases, hybridization may be affected by protein interaction and may not be due solely to binding of complementary nucleic acids (Tomizawa and Som, 1984). Nevertheless, these precedents suggest that antisense RNA could be used to specifically reduce gene activity. This possibility was first tested by synthesizing antisense oligonucleotides and observing their effect on corresponding sense RNAs. Oligonucleotides complementary to the *E. coli* 16S ribosomal RNA, at the point where it binds the Shine-Dalgarno sequence, were observed to inhibit translation *in vitro* and colony formation *in vivo* (Jayaraman et al., 1981; Taniguchi and Weissman, 1978). One interpretation of these results is that oligonucleotide binding prevented the 16S ribosomal RNA from initiating translation. Similarly, oligonucleotides complementary to a short region from Rous sarcoma virus were capable of inhibiting translation *in vitro* and mammalian cell growth and transformation *in vivo* (Stephenson and Zamecnik, 1978; Zamecnik and Stephenson, 1978).

Antisense genes have been constructed which produce complementary RNA capable of specifically reducing gene activity. In *E. coli*, genes which express antisense OMP A, OMP C, LPP and lac Z RNAs have been studied (Coleman et al., 1984, Pestka et al., 1984). All of these genes were expressed from the lac Z promoter and, when induced, were capable of reducing gene activity from the corresponding sense genes as much as twenty fold. Different antisense genes were effective to varying extents. Although in some aspects this variability is not well understood, it seems important for the antisense RNA to cover the AUG start codon. This observation suggests that antisense RNA impedes the initiation of translation. These important experiments suggest that antisense RNA may potentially affect a broad range of bacterial genes.

In frog oocytes, microinjection of antisense beta-globin RNA was capable of reducing the synthesis of beta-globin from microinjected mRNA (Melton, 1985). Since double stranded beta-globin RNA is detected, RNA:RNA hybridization is almost certainly involved. Experimental approaches such as this may be useful to study the function of genes which are expressed in early embryogenesis. However, the presence of antisense RNA, and thus the effect of reduced gene activity, is only transient. Therefore, this approach may not be applicable for those cellular phenotypes which can only be observed over a period of time. Furthermore, since antisense RNA inhibits protein synthesis rather than the gene product itself, the block in translation must last long enough for any preexisting protein to either dilute or degrade. Otherwise, the effect of antisense RNA will disappear before the amount of gene product diminishes significantly. Finally, as with the antibody approach, microinjection of individual cells requires that any effect be apparent in small numbers of cells.

Elegant work by Izant and Weintraub (1984) has shown that antisense RNA can function to reduce gene activity in mouse L cells. Antisense thymidine kinase (TK) RNA, produced by either microinjected or chromosomal antisense genes, can transiently inhibit the expression of a microinjected TK gene. Although preliminary, this experiment suggests that the study of mammalian gene function with antisense genes may be possible. However, to realize this possibility, it is important to show that antisense RNA can permanently reduce gene activity from endogenous as well as microinjected genes. Since TK activity was assayed transiently, it is not known whether the effect of antisense RNA is transient or permanent. Furthermore, whereas the activity of microinjected TK genes is significantly reduced by antisense RNA, the activity of chromosomally integrated TK genes is not (J. Izant and H. Weintraub, pers. comm.). Although the reason for this difference is unknown, it severely limits this type of approach.

In chapter 6, the activity of a TK gene, integrated into the genome of a mouse L cell, is permanently reduced by the expression of antisense TK RNA. In this study, a high level of antisense RNA was necessary to stably reduce TK gene activity. This high level was obtained by expressing antisense TK sequences with sense dihydrofolate reductase (DHFR) sequences as part of a chimeric transcript. Overexpression of DHFR can be selected with the drug methotrexate which, in this instance, also results in overexpression of antisense TK RNA. Thus, at least in tissue culture cells, it may be possible to observe the effect of diminished activity of any gene for which a molecular clone is available. In conjunction with DNA mediated gene transfer, antisense RNA could be used to control the level of activity for a gene of interest. In a sense, such an approach mimics that of classical genetics in which mutations resulting in a gain or a loss of function are analyzed. One advantage of a molecular approach over a genetic approach is that, in the former, gene expression can be altered in a precisely controlled fashion. For example, gene activity might be conditionally reduced by expressing antisense RNA from an inducible promoter. Also, some transcription units produce multiple mRNAs. In these cases, appropriate antisense RNAs could be used to hybridize to a specific mRNA and could thus help discern the function of the particular protein encoded in that mRNA.

References

Antman, K. and Livingstone, D. 1980. Intracellular neutralization of SV40 tumor antigens following microinjection of specific antibody. Cell **19:** 627-635.

Beachy, P., Helfand, S. and Hogness, D. 1985. Segmental distribution of bithorax complex proteins during *Drosophila* development. Nature **313:** 545-551.

Burke, B. and Warren, G. 1984. Microinjection of mRNA coding for an anti-golgi antibody inhibits intracellular transport of a viral membrane protein. Cell **36:** 847-856.

Coleman, J., Green, P. and Inouye, M. 1984. The use of RNAs complemantary to specific mRNAs to regulate the expression of individual bacterial genes. Cell **37:** 429-436.

Hedrick, S., Cohen, D., Nielson, E. and Davis, M. 1984a. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature **208:** 149-153.

Hunkapiller, M., Kent, S., Caruthers, M., Dreyer, W., Firca, J., Griffin, C., Horvath, S., Hunkapiller, T., Tempst, P. and Hood, L. 1984. A microchemical facility for the analysis and synthesis of genes and proteins. Nature **310:** 105.

Izant, J. and Weintraub, H. 1984. Inhibition of thymidine kinase gene expression by antisense RNA: A molecular approach to genetic analysis. Cell **36:** 1007-1015.

Jayaraman, K., McParland, K., Miller, P. and Ts'o, P. 1981. Selective inhibition of *E. coli* protein synthesis and growth by non-ionic oligonucleotide complementary to the 3' end of 16S RNA. Proc. Nat. Acad. U.S.A. **78:** 1537-1541.

Lerner, M., Boyle, R., Mount, S., Wolin, S. and Steitz, J. 1980. Are

snRNPs involved in splicing? Nature **283:** 220-224.

Lin,J. and Feramisco, J. 1981. Disruption of the *in vivo* distribution of the intermediate filaments in living fibroblasts through the microinjection of a specific antibody. Cell **24:** 185-193.

Mabuchi, I. and Okuno, M. 1977. The effects of myosin antibody on the division of starfish blastomeres. J. Cell Biol. **74:** 251-264.

Maniatis, T., Kee, S., Efstradiatis, A. and Kafatos, F. 1976. Amplification and characterization of a beta-globin gene synthesized *in vitro*. Cell **8:** 163-182.

Melton, D. 1985. Injected antisense RNAs specifically block messenger RNA translation *in vivo*. Proc. Nat. Acad. U.S.A. **82:** 144-148.

Mizuno, T., Chou, M., and Inouye, M. 1984. A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). Proc. Nat. Acad. U.S.A. **81:** 1966-1970.

Pestka, S., Daugherty, B., Jung, V., Hotta, K. and Pestka, R. 1985. Anti-mRNA: specific inhibition of translation of single mRNA molecules. Proc. Nat. Acad. Sci. U.S.A. **81:** 7525-7528.

Rogers, J. and Wall, R. 1980. A mechanism for RNA splicing. Proc. Nat. Acad. U.S.A. **77:** 1877-1879.

Sutcliffe, J., Milner, R., Shinnick, T. and Bloom, F. 1983. Identifying the protein products of brain specific genes with antibodies to chemically synthesized peptides. Cell **33:** 671-682.

Simons, R. and Kleckner, N. 1983. Translational control of IS10 transposition. Cell **34:** 683-691.

Stephenson, M. and Zamecnik, P. 1978. Inhibition of Rous sarcoma virus RNA translation by a specific oligodeoxynucleotide. Proc. Nat. Acad. U.S.A. **75:** 285-288.

Taniguchi, T. and Weissman, C. 1978. Inhibition of Q_{beta} RNA 70S ribosome initiation complex formation by an oligodeoxynucleotide complementary to the 3' terminal region of *E. coli* 16S ribosomal RNA. Nature **275:** 770-773.

Tomizawa, T., Itoh, T., Selzer, G, and Som, T. 1981. Inhibition of Col E1 RNA primer formation by a plasmid specified small RNA. Proc. Nat. Acad. U.S.A. **78:** 1421-1425.

Tomizawa, T. 1984. Control of Col E1 plasmid replication: the process of binding of RNA1 to the primer transcript. Cell **38:** 861-870.

Tomizawa, T. and Som, T. 1984. Control of Col E1 plasmid replication: Enhancement of binding of RNA1 to the primer transcript by the Rom protein. Cell **38:** 871-878.

Tonegawa, S., Brack, C., Hozumi, N. and Schuller, R. 1977. Cloning of an immunoglobulin variable region gene from mouse embryo. Proc. Nat. Acad. U.S.A. **74:** 3518.

White, R. and Wilcox, M. 1985. Protein products of the bithorax complex in *Drosophila*. Cell **39:** 163-171.

Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Cheng, Y. and Axel R. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell **11:** 223-232.

Yamaizumi, M., Uchida, T., Okada, Y. and Furosawa, M. 1978. Neutralization of diphtheria toxin in living cells by microinjection of antifragment A contained within resealed erythrocyte ghosts. Cell **13:** 227-232.

Yanagai, Y., Yoshikai, Y., Leggett, K., Clark, S., Aleksander, I. and Mak, T. 1984. A human T cell specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature **308:** 145-149.

Zamecnik, P. and Stephenson, M. 1978. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc. Nat. Acad. U.S.A. **75**: 280-284.

Cell, Vol 27, 573–581, December 1981 (Part 2), Copyright © 1981 by MIT

Antibody Diversity: Somatic Hypermutation of Rearranged V_H Genes

Stuart Kim, Mark Davis,* Eric Sinn, Phil Patten and Leroy Hood
Division of Biology
California Institute of Technology
Pasadena, California 91125

Summary

The immune response to phosphorylcholine in BALB/c mice has been well characterized. Amino acid sequence analyses of heavy-chain variable (V_H) regions from 19 myeloma and hybridoma immunoglobulins binding phosphorylcholine show that 10 are identical (the prototype T15 V_H sequence) and 9 are distinct variants differing by one to eight residues. A T15 V_H DNA probe was used to isolate four closely related members of the T15 V_H gene family, including one encoding the T15 V_H sequence, from a sperm genomic library. A comparison of the protein and germline V_H sequences suggested that most of the immune response to phosphorylcholine is derived from the T15 germline V_H gene segment. The variant heavy chains from the M167 and M603 α immunoglobulins differ in their V_H protein sequences from T15 by eight and three residues, respectively. We analyzed the somatic variability in and around the coding regions of these two variant V_H genes by comparing them with the corresponding regions of the appropriate germline gene segments. The somatic variation has three properties: it is extensive and is found in flanking as well as coding sequences (for example, at least 44 substitutions for the M167 sequence and 10 substitutions for the M603 sequence); in the coding regions, it includes many silent as well as replacement substitutions; and it is focal in nature and centered around the rearranged V_H genes. Although the mutations extend into the neighboring upstream and downstream flanking sequences, sequences approximately 5 kb upstream and downstream from the V_H genes show no substitutions. Moreover, the associated heavy-chain constant genes (C_α) from both variant α genes are unaltered, indicating that a closely linked and coexpressed gene is unmutated. We conclude that this somatic variation is generated by a special hypermutational mechanism highly localized in its site of execution and highly restricted in its time of operation during B-cell development.

Introduction

The somatic variation of antibody genes plays an integral role in permitting a vertebrate such as the mouse to synthesize perhaps 10^8 distinct types of antibody molecules. We analyze the diversity patterns generated by a special somatic mutational mechanism.

The antibody molecule is comprised of light (L) and heavy (H) chains, which fold into variable (V) domains that bind antigen and constant (C) domains that encode effector functions (for a review see Early and Hood, 1981). In turn, the variable regions of the L and the H chains from different immunoglobulins exhibit three areas of extreme variability termed the hypervariable segments (Wu and Kabat, 1970; Capra and Kehoe, 1974), which fold to constitute the walls of the antigen-binding site (Amzel et al., 1974).

Immunoglobulins are encoded by three families of genes—two for L chains, λ and κ, and one for H chains. The V regions are encoded by gene segments separated from one another in the germline—V_L and J_L (joining) as well as V_H, D (diversity) and J_H—which are rearranged and joined during B-cell differentiation to generate V_L and V_H genes, respectively (Brack et al., 1978; Sakano et al., 1979; Seidman et al., 1979; Early et al., 1980; Sakano et al., 1980). The C regions are encoded by separate genes. In the mouse, there are eight C_H genes which encode the various classes and subclasses of immunoglobulins (for example, C_μ–IgM, C_α–IgA, and so forth). Initially, the V_H gene is always associated with the C_μ gene, and IgM molecules are expressed (Cooper et al., 1976; Raff, 1976). Subsequently, antigen may trigger B-cell differentiation and the second type of DNA rearrangement, the class switch, occurs when the C_μ gene is replaced by another C_H gene, and this results in the expression of a different immunoglobulin class (Davis et al., 1980a; Kataoka et al., 1980; Lawton et al., 1980; Maki et al., 1980).

A somatic mutation mechanism has been postulated to explain the diversity scattered throughout certain V_λ (Weigert et al., 1970; Weigert and Riblet, 1976; Bernard et al., 1978; Brack et al., 1978), V_κ (Valbuena et al., 1978; Weigert et al., 1978; Pech et al., 1981), and V_H (Hood et al., 1976; Bothwell et al., 1981; Crews et al., 1981; Gearhart et al., 1981) genes. We have studied the role somatic mutation plays in the immune response to the simple hapten, phosphorylcholine. Protein sequencing studies of V_H regions (Hood et al., 1976; Gearhart et al., 1981) combined with DNA sequencing studies of germline V_H genes (Crews et al., 1981) have demonstrated several features of somatic mutation in the phosphorylcholine system. First, ten of nineteen V_H segments, that portion of the V_H region encoded by the V_H gene segment, have identical protein sequences (this sequence is denoted T15), whereas the nine remaining variants are all unique and differ from the prototype T15 sequence by one to eight amino acid residues. Second, only four germline V_H gene segments closely homologous to a T15 V_H gene probe were found in the BALB/c genome. One germline V_H gene encodes the

* Present address: Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205.

T15 sequence. All nine variant V_H segments appear to arise from the germline T15 V_H gene segment because the variants are related much more closely in sequence to the T15 V_H gene segment than to any one of the other three T15-like V_H gene segments. Thus it appears that the entire immune response to phosphorylcholine is derived from a single germline V_H gene segment and that somatic mutation occurs frequently (9/19 examples). Third, somatic mutation correlates with the class switch because the variant V_H regions only appear associated with immunoglobulins of the IgG or IgA class, but not with immunoglobulins of the initially expressed IgM class.

Since we had analyzed the entire family of germline T15 V_H genes and many protein V_H regions, our next approach was to clone and sequence two variant T15 V_H genes, M167 and M603, whose V_H segments differ at the protein level from T15 by eight and three residues, respectively. We report the cloning and sequence analysis of the V_H gene from the phosphorylcholine-binding IgA myeloma M167. The M167 V_H gene, its flanking regions and the 5′ portion of its C_α gene were compared to the corresponding regions of germline (sperm) DNA and to a second variant phosphorylcholine-binding IgA myeloma M603, which we had previously cloned (Early et al., 1980). We show that the M167 and M603 V_H genes are derived from the germline T15 V_H gene segment. The somatic variation is extensive and is sharply localized in and around the rearranged V_H genes. We conclude that the somatic variation is produced by a special hypermutational mechanism.

Results and Discussion

Isolation of the M167 V_H Gene

Because the four V_H gene segments of the T15 gene family crosshybridize strongly with the T15 V_H probe, a C_α probe was employed to isolate the rearranged M167 V_H gene since the M167 tumor synthesizes an immunoglobulin of the IgA class. Using this probe in a Southern blot analysis of Eco RI–cleaved M167 DNA, two major bands were observed (13 kb, 4.6 kb; Figure 1A). The 13 kb band represents a nonproductively rearranged C_α gene (K. Calame and S. Kim, unpublished data), and the 4.6 kb band was the rearranged and expressed M167 C_α gene (see below).

A library of 2.0×10^6 phage was constructed from M167 myeloma tumor DNA partially digested with Hae III or Alu I and inserted into the vector Charon 4A. The library was screened using a C_α probe, and five clones were obtained. One clone was representative of the 13 kb Eco RI band (data not shown). The remaining four clones contained a 4.6 kb Eco RI band. Southern blot, restriction map and DNA sequence analyses indicated that three of these clones also contained the M167 V_H gene, whereas the fourth did not. Restriction map data derived from one clone, ChM167α10, containing the rearranged M167 V_H gene are presented in Figure 2.

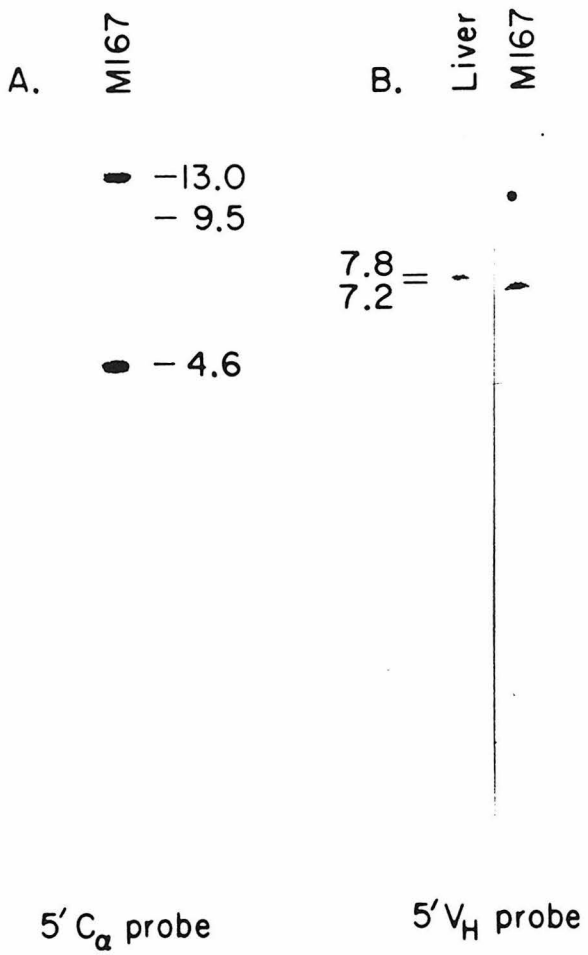


Figure 1. Southern Blot Analyses of M167 and Liver DNAs

Approximately 15 μg of DNA was digested with Eco RI, fractionated on a 0.8% agarose gel, transferred to nitrocellulose paper and hybridized with a 5′ C_α probe constructed by subcloning the 5.1 kb Eco RI fragment from ChM603α6 into pBR322 (A) or with a 5′ V_H probe constructed by subcloning the 1.4 kb Bam HI fragment from ChM167α10 into pBR322 (B). The faint band in A migrating at 9.5 kb comigrates with the germline C_α gene and presumably originates from host-cell DNA contamination. The difference in band size shown in B was corroborated by a Southern blot analysis of Hind III–digested DNAs in which liver DNA showed a single band at 7.0 kb and M167 DNA shows a single band at 4.0 kb using the same 5′ V_H probe (data not shown).

The M167 and M603 V_H Genes Are Derived from the Germline T15 V_H Gene Segment

The amino acid sequences encoded by the four genes of the T15 V_H gene family do not include the M167 or M603 V_H sequences, although the two myeloma sequences are much closer to the T15 sequence than to the three other T15-like sequences. Furthermore, restriction map, heteroduplex and DNA sequence analyses (see below) establish that the 5′ flanking

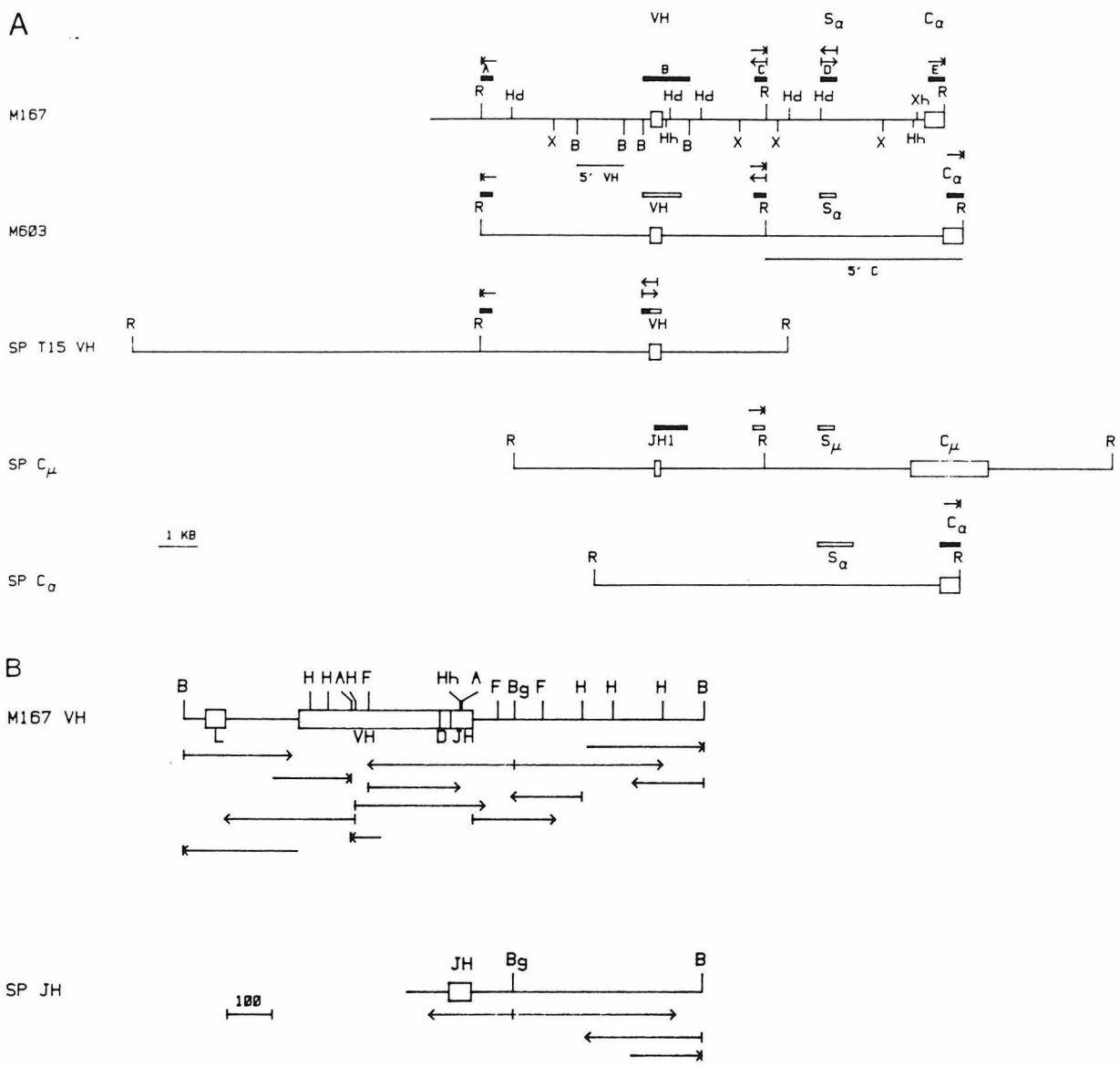


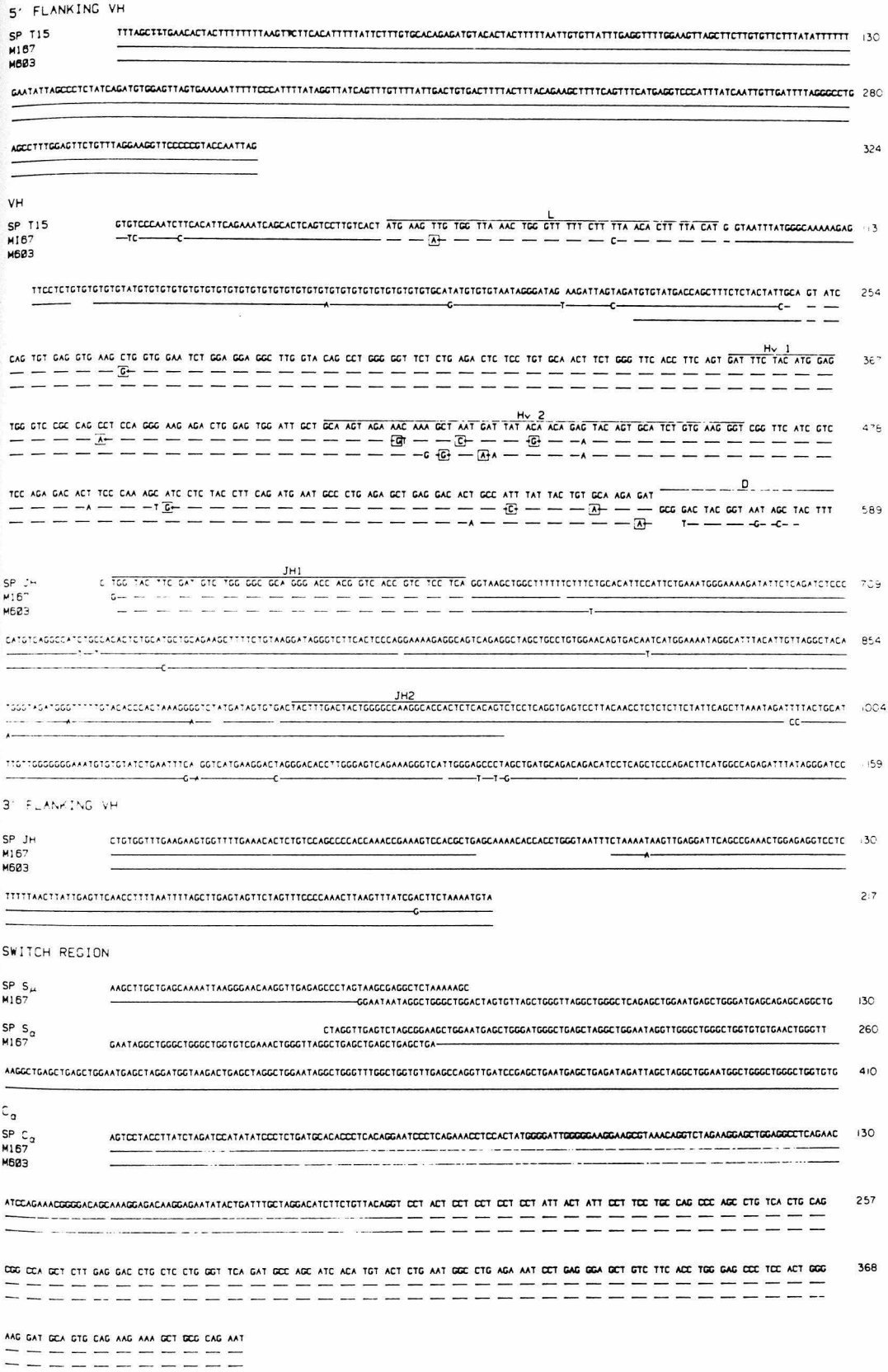
Figure 2. Restriction Maps and Sequence Strategies for the Immunoglobulin Genomic Clones

(A) Restriction maps of various T15-related clones. The M167 clone (ChM167α10), the M603 clone (ChM603α6), the germline T15 V_H clone (λV_1), the germline C_μ gene (ChSpμ27) and the germline C_α gene (ChSpα29) are presented. The λV_1, ChM167α10 and ChM603α6 clones have identical Eco RI, Xba I, Hind III, and Bam HI sites upstream from the V_H gene segment (see Davis et al., 1980a; Crews et al., 1981). The regions of each clone sequenced in this report are indicated by a shaded box above the restriction map. Sequences reported previously are indicated by an open box (Early et al., 1980; Davis et al., 1980b; Crews et al., 1981). The regions used as probes in Figure 1 are indicated by a line below the restriction map. A: Ava II. B: Bam HI. Bg: Bgl II. F: Fnu4HI. H: Hinf I. Hd: Hind III. Hh: Hha I. R: Eco RI. X: Xba I. Xh: Xho I. a: 5′ flanking V_H sequence. b: V_H sequence. c: 3′ flanking V_H sequence. d: switch sequence. e: C_α sequence.

(B) Sequencing strategies. Arrows ending at a restriction site indicate 3′ end-labeling and arrows beginning at a restriction site indicate 5′ end-labeling. The antisense strand of the M167 switch site was sequenced by 5′ end-labeling at the Sau 3A site at position 357 in Figure 3. The antisense strand of the germline T15 V_H gene segment was sequenced by 5′ end-labeling at the Hinf I site at position 276 in Figure 3. Expanded maps of the M167 V_H region and the germline J_H region, as well as their sequencing strategies, are depicted.

regions of both the M167 and the M603 V_H genes are identical, apart from some somatic variation, to the 5′ flanking region of the germline T15 V_H gene segment, but are distinct from the 5′ flanking regions of the other three T15-like V_H gene segments (Figure 3, and data not shown). The above comparisons show that the M167 and M603 V_H genes are related much more closely to the germline T15 V_H gene segment than to the other V_H gene segments in the T15-like V_H gene family.

Southern blot analysis indicated that the 5′ flanking regions of the M167 and M603 V_H genes arose from a germline V_H gene segment originally located on a 7.8 kb Eco RI fragment. When Eco RI–restricted DNA


5' FLANKING VH
SP T15
M167
M603
VH
L
Hv 1
Hv 2
D
JH1
JH2
3' FLANKING VH
SP JH
SWITCH REGION
SP Sμ
SP Sα
Cα
SP Cα

was blotted and probed with a 1.4 kb Bam HI fragment derived from the 5′ flanking side of the M167 V_H gene, liver DNA revealed a single band, 7.8 kb, which comigrated with the Eco RI fragment containing the T15 V_H gene segment from the clone λV_1 (Crews et al., 1981), whereas M167 DNA revealed a single band, 7.2 kb, which comigrated with the Eco RI fragment containing the M167 V_H gene from the clone ChM167α10 (Figure 1B and data not shown). These observations suggest that the 7.8 kb germline T15 V_H Eco RI fragment was rearranged in the tumor line to yield a 7.2 kb M167 V_H gene. Since the 5′ flanking regions of the M603 and M167 V_H genes are essentially identical by all criteria tested, it can be inferred that the M603 V_H gene also arose from a germline 7.8 kb Eco RI fragment.

Three lines of evidence indicate that only one T15-like V_H gene segment on a 7.8 kb Eco RI fragment is contained in the BALB/c genome. First, if there were multiple V_H genes on fragments 7.8 kb in length, we would not expect all of them to be identically rearranged. However, Figure 2 indicates that the 7.8 kb band is missing in M167 DNA and is replaced by a 7.2 kb band. This observation implies that the T15 V_H gene on the other heavy-chain chromosome in the M167 tumor cell line has been deleted. Second, four independent clones containing the 7.8 kb Eco RI fragment have been sequenced and, in each case, the V_H gene segments encoded the prototype T15 sequence rather than the variant M603 or M167 sequences (Early et al., 1980; Crews et al., 1981; S. Crews, unpublished observations). It is therefore statistically improbable that there is more than one V_H gene segment on a 7.8 kb Eco RI fragment. Third, as mentioned above, the rearranged M167 and M603 V_H genes and the germline T15 V_H gene segments have essentially identical 5′ flanking regions by restriction map, heteroduplex and DNA sequence analyses. This flanking sequence identity suggests that the M167 and M603 V_H genes are derived from the germline T15 V_H gene segment.

The M167 and M603 V_H Genes Appear to Have Arisen by Somatic Mutation and Not from Myeloma Artifacts or Genetic Polymorphism

Two trivial explanations for the origin of the variant M167 and M603 sequences are mutational drift in the myeloma cell lines and genetic polymorphisms in the inbred BALB/c strain of mice. Mutational drift in the myeloma cell lines appears improbable for several reasons. First, it is not clear why the variants (for example, M167) would have a selective advantage over the wild-type cells, presumably T15, since the extensive amino acid differences observed (eight for M167) would require multiple rounds of mutation and selection. Second, one particular myeloma cell line, T15, has maintained its V_H amino acid sequence over a span of some ten years, indicating that in this myeloma cell line little or no mutation has occurred (L. Hood, unpublished observation). Third, many hybridoma cell lines that have been passaged for only short periods exhibit variant V_H sequences, indicating that mutations probably arose in functional B cells before cell fusion (Gearhart et al., 1981). Finally, when flanking sequences 5 kb upstream and 5 kb downstream from the M167 and M603 V_H genes are compared with their germline counterparts, no mutations are observed in more than 1000 bp compared (Figure 3). If random mutations were widespread in the myeloma cell lines and not subject to selection, these flanking regions would be expected to exhibit mutations to the same extent as those seen in the coding region. Thus we conclude that the variations seen in the M167 and M603 cell lines are not a consequence of mutational drift in the myeloma cell lines.

It is improbable that genetic polymorphism in differing BALB/c mice can account for the variation seen between the M167, M603 and germline T15 V_H gene segments. First, if this hypothesis were correct, the difference between the putative M167 and T15 V_H alleles (44 substitutions in 1054 bases compared [see below]) would be far greater than the polymorphism seen between the BALB/c and C3H strains in the same area (0/301) (Kurosawa et al., 1981). Second, it is not clear why the polymorphic substitutions would localize themselves in and around the rearranged V_H gene and not extend to its flanking regions (see below).

These data strongly suggest that the M167 and M603 V_H genes are not present in the BALB/c genome and that the origin of the mutation cannot be due either to mutational drift in myeloma cell lines or to genetic polymorphism. Thus the M167 and M603 V_H genes are somatic variants of the germline T15 V_H

Figure 3. A Comparison of Various Homologous Nucleotide Sequences from Somatic, M167 and M603, and Germline DNAs

DNA sequences were determined by the method of Maxam and Gilbert (1980). The germline sequence is always listed on top. Nucleotide identities are indicated by a line, nucleotide substitutions are denoted by the appropriate letter, insertions or deletions are indicated by a space and replacement substitutions are indicated by a box. Exons are indicated by the triplet spacing. The C_μ exon codes for the hinge and C_H2 regions. The 24 nucleotide deletion in the M167 3′ flanking V_H region may be a cloning artifact since small deletions in the same area have appeared in the germline sequence reported by Gough and Bernard (1981) and in a clone derived from the myeloma T15 (S. Kim, unpublished observations). The switch region for M167 is shown. The S_μ site is identical to the M603 S_μ site. The S_α site occurs in a 30 nucleotide repeat often associated with α switching (Davis et al., 1980a). There are 142 bases between the S_μ and S_α sites, which are of unknown origin. The S_α sequence differs from that reported previously (Davis et al., 1980b) at positions 206, 249, 270, 275 and 278 owing to previous errors in data transcription. The M603 V_H, T15 V_H, S_μ, and S_α sequences have been reported previously (Early et al., 1980; Davis et al., 1980b).

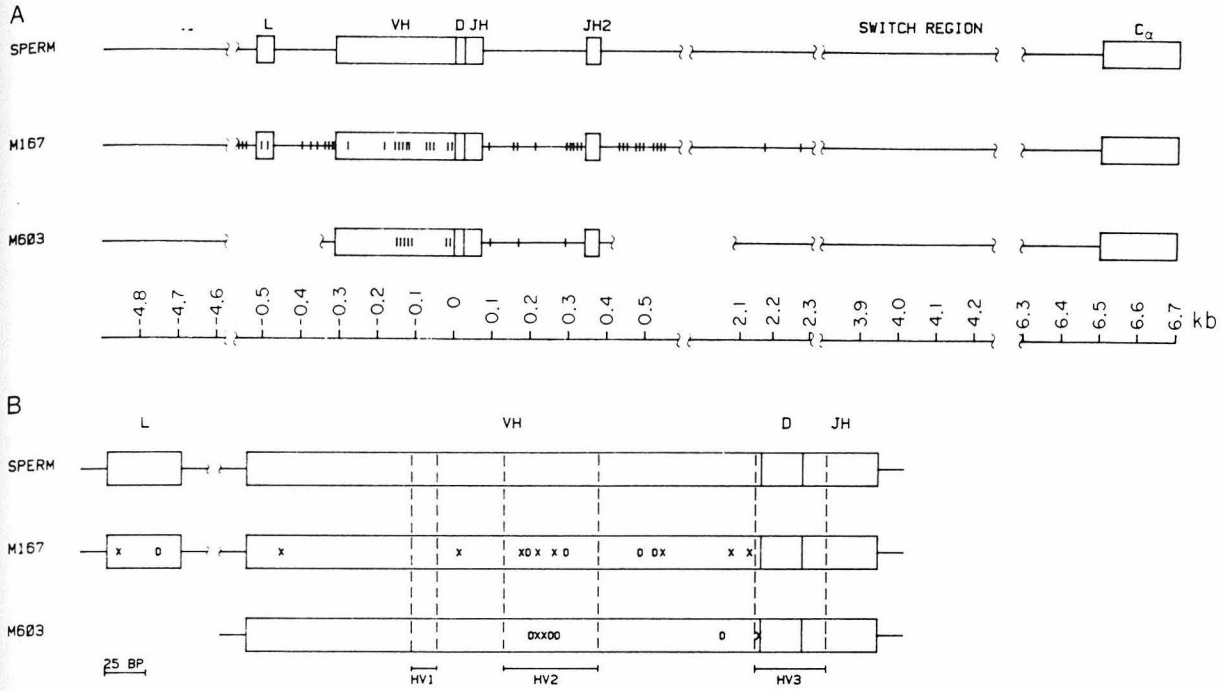


Figure 4. Mutations Occurring in and around Two Variant T15 V_H Genes

(A) The location of the nucleotide substitutions, insertions or deletions in the M167 and M603 clones are indicated by a vertical line. The exact locations of the sequence regions shown are indicated in Figure 2A. When the mutations are adjacent, the lines are moved apart until they are distinct. (B) The point mutations in the coding regions include both silent (0) and amino acid replacement (X) substitutions.

gene segment. Let us now consider the nature of the mutational pattern that has been imposed on these two somatically derived V_H genes.

Somatic Variation Is Extensive and Localized in and around the Rearranged V_H Gene

To determine the extent and distribution of nucleotide substitutions in and around the variant V_H genes, we compared the nucleotide sequences of the M167 and M603 V_H genes and their flanking sequences, regions located 5 kb upstream and 5 kb downstream from the V_H genes, and portions of the 5′ end of the C_α genes to their germline counterparts (Figure 3). Figure 4A diagrammatically depicts the distribution and number of mutations scattered throughout these regions.

The pattern of somatic variation in the variant V_H genes exhibits several interesting characteristics. First, it is extensive. The M167 V_H gene has 44 substitutions in the 1054 nucleotides analyzed (3.8% variation), which includes the coding regions (L, V_H and J_H) as well as the adjacent flanking regions (a portion of the 5′-untranslated region, the L-V_H intron and a portion of the J_H-C_α intron). We count sequence deletions or insertions as well as nucleotide changes as substitutions. Substitutions probably occur in the D-gene segment, although this cannot be verified until the germline M167 D-gene segment is isolated. The M603 V_H gene shows fewer mutations with ten substitutions in 714 bp analyzed (1.4% variation). Second, the variation is sharply localized in and around the variant V_H genes. A 200 bp sequence 2 kb to the 3′ side of the M167 V_H gene has only two substitutions, whereas a 200 bp sequence from the same region in the M603 V_H gene has no substitutions. Moreover, 300–500 bp sequences 5 kb upstream and 5 kb downstream from both the M167 and M603 V_H genes have no substitutions at all. The rate of mutation 2.3 kb downstream from the rearranged M167 and M603 V_H genes is therefore significantly less than in and around the V_H gene itself, and there appears to be no mutation at all 5 kb to either side of the V_H genes. Finally, the 5′ portion of the M167 and M603 C_α genes show no substitutions, indicating that a closely linked and coexpressed gene is not subjected to somatic variation. Thus the somatic variability is localized in and around the rearranged V_H genes. Third, many silent as well as replacement substitutions occur in the coding regions of the variant V_H genes. (A silent substitution is a base change that does not change the amino acid residue encoded by a particular codon.) The M167 V_H coding region differs from the germline T15 V_H coding region by 14 nucleotides, and five of these are silent substitutions. The M603 V_H coding region differs by seven nucleotides, four of which are silent. Thus, in the coding regions, silent mutations are almost as common as replacement substitutions. Finally, a detailed analysis of the types of base substitutions occurring in the variant M167 and

M603 V_H genes does not shed any light on the mechanism of somatic variation (Table 1). For example, no particular base has a greater tendency to mutate than do the others. Transversions as well as transitions occur. Base changes occur about equally at all of the codon positions. However, two interesting points emerge from this analysis. First, the mutational mechanism is capable of generating small deletions and insertions as well as base substitutions (for example, eight deletions and insertions in M167). This observation is surprising and somewhat disturbing because single base insertions or deletions in the coding region would alter the downstream reading frame of the V_H gene. Second, about 50% of the mutations in the coding regions occur in hypervariable region 2 (10/21), whereas this region accounts for only 18% of the V_H gene segment. This concentration of variability probably reflects the result of selection for sequence variations in these regions.

The Mechanism of Somatic Mutation

Several patterns have emerged from this and other studies on the phosphorylcholine immune response that have interesting implications for the mechanism of somatic variation (Crews et al., 1981; Gearhart et al., 1981).

First, a special mutational mechanism must be capable of generating extensive diversity—a 3.8% nucleotide substitution rate in the case of the M167 V_H gene and its adjacent flanking sequences. It appears improbable that mutational changes of this extent can arise solely by intense selection for variant sequences appearing at an ordinary mutational rate because of the number of mutations, their focal nature and the fact that 35 of 44 mutations are silent in the M167 V_H gene and surrounding regions and, accordingly, cannot be fixed by selection. Thus we conclude that some form of special hypermutational mechanism generates multiple substitutions, both silent and replacement, and the constellation of these substitutions is then selected presumably by virtue of increased antigen-binding affinity, release from idiotype suppression or some other unidentified selective force (see Crews et al., 1981).

Second, somatic mutation occurs in the vicinity of the rearranged V_H gene and clearly diminishes as one moves away from the gene until the mutation is infrequent or nonexistent 4–5 kb away. Thus the hypermutational mechanism is localized in and round the V_H gene. The focal nature of the mutational mechanism suggests that it recognizes the rearranged V_H gene. For example, rearranged V or J gene segments may be recognized by an enzyme which then cleaves a single DNA strand in or near the rearranged gene segment. Subsequently, repair of this strand may occur with errors. The error frequency presumably declines as the distance from the cleavage site increases. Indeed, since the C_α genes associated with M167 and M603 V_H genes are not mutated, the mutational mechanism appears specific for the rearranged V_H genes. A general mechanism of this sort was previously proposed by Brenner and Milstein (1967).

Third, in the phosphorylcholine system there is a striking correlation between somatic mutation and the class switch (Gearhart et al., 1981). The V_H regions associated with IgM (C_μ) molecules always have the germline T15 sequence, whereas the variant V_H regions always are associated with IgG (C_γ) or IgA (C_α) molecules that have undergone a class switch. Likewise, the V_L regions from IgG or IgA molecules generally show variation, whereas the V_L regions from IgM molecules do not. The somatic mutation mechanism is not obligatorily coupled to the class-switching mechanism because IgA molecules can have germline V_H and V_L regions (Gearhart et al., 1981).

Table 1. Numbers of Transitions, Transversions, Insertions and Deletions in the M167 and M603 V_H Genes as Compared with the Germline T15 V_H Gene Segment

V_H Gene	Myeloma Mutational Type								Totals
	Transitions								
	A → G	G → A	C → T	T → C					
M167	4	4	5	9					
M603	1	3	1	1					28
	Transversions								
	A → C	A → T	G → C	G → T	C → A	C → G	T → A	T → G	
M167	1	1	0	3	1	3	4	0	
M603	0	0	0	0	0	1	3	0	17
	Deletions								
	A → -	G → -	C → -	T → -	GTGT → -		TA → -		
M167		2	0	3	1		1		
M603	0	0	0	0					7
	0								
	Insertions								
	- → A	- → G	- → C	- → T					
M167	1	0	0	1					2

There is one important caveat to our suggestion of a hypermutational mechanism. Perhaps the M167 V_H gene has undergone extensive mutation in an atypical manner. We believe this is improbable for several reasons. First, if atypical mutations have occurred, why are they so sharply localized around the M167 V_H gene? Second, the M603 V_H gene also shows substitutions in its 3′-flanking sequence, although there are far fewer substitutions in this case. Clearly, an analysis of other variant V_H genes differing by six or more residues from their germline counterparts should be carried out.

Antibody diversity appears to arise in two stages of B-cell development. Germline gene segments are joined in a combinatorial fashion to generate a large baseline repertoire created independent of external antigenic stimulation. In addition, junctional somatic variation is generated at this early stage of B-cell differentiation by the flexible joining of the various gene segments (Max et al., 1979; Sakano et al., 1980; Weigert et al., 1980; Hood et al., 1981) that constitute the V gene. Subsequently, at a late stage in B-cell development, antigen may trigger B cells to undergo differentiation which leads to class switching and perhaps to the activation of a special somatic hypermutational mechanism. Thus antigen may trigger the generation of variants by somatic mutation in the V_L and V_H genes. It is attractive to postulate that antigen then selects those variant lymphocytes that have higher affinity antibody receptors for clonal expansion, thus leading to the well known affinity maturation of the immune response. However, our analysis of the phosphorylcholine system demonstrates that variant immunoglobulins do not necessarily have higher affinities for phosphorylcholine than do the germline antibodies (Gearhart et al., 1981). This raises the possibility that other forms of selection may operate in the vertebrate immune response (see Crews et al., 1981, for discussion). B cells at the early and late developmental stages probably undergo their differentiation events in the central and peripheral lymphoid organs, respectively (Baltimore, 1981).

Experimental Procedures

Construction and Screening of the M167 Library

High molecular weight DNA from the M167 myeloma cell line was prepared as described (Blin and Stafford, 1976). The DNA was digested partially with the restriction enzymes Hae III or Alu I, size-fractionated on a 10% to 40% sucrose gradient, ligated to synthetic Eco RI linkers and inserted into the vector Charon 4A (Blattner et al., 1977; Maniatis et al., 1978). The library was packaged in vitro by the procedure of Hohn and Murray (1977) using the strains of Sternberg (1977). The library was screened, before amplification, by the procedure of Benton and Davis (1977) using a ^{32}P-labeled C_α probe constructed by subcloning the 5.1 kb Eco RI fragment from ChSpα6 into the Eco RI site of pBR322 (see Figure 2).

Genomic Blot Hybridizations

Myeloma DNAs were prepared by the method of Blin and Stafford (1976). Sperm DNA was a gift from T. Hunkapiller, and liver DNA was a gift from M. Steinmetz. Approximately 15 μg of DNA was digested to completion with the appropriate restriction enzyme, size-fractionated on horizontal 0.8% agarose gels, and transferred to nitrocellulose paper by the method of Southern (1975). Hybridizations were carried out in 1.0 M NaCl, 0.2 M Tris–HCl (pH, 7.8), 10× Denhardt's (Denhardt, 1966), 0.1% SDS, 0.1% sodium pyrophosphate, 50 μg/ml poly(rA), 20 μg/ml denatured salmon sperm, and labeled probe at a concentration of 10^7 cpm of hybridizable sequence per milliliter at 68°C for 12 hr with agitation. The blots were washed twice in 1.0 M NaCl, 5× Denhardt's, 0.5 M Tris–HCl, 0.1% SDS, 0.1% sodium pyrophosphate for 30 min at 68°C, and then twice in 0.15 M NaCl, 0.015 M sodium citrate, 0.1% SDS, 0.1% sodium pyrophosphate for 30 min at 68°C. Autoradiography was accomplished at −70°C with an intensifying screen.

DNA Sequencing

Restriction fragments were 5′ endlabeled with ^{32}P-γ-ATP and polynucleotide kinase after treatment with bacterial alkaline phosphatase. DNA polymerase I (large fragment) and ^{32}P-α-dNTPs allowed 3′ endlabeling. The method of Maxam and Gilbert (1980) as modified by Smith and Calvo (1980) was employed for DNA sequence analysis.

Acknowledgments

We are grateful to Stephen Crews, Michael Steinmetz and Henry Huang for valuable advice and discussions. We would like to acknowledge the excellent technical help of Karyl Minard. We are indebted to Steve Crews, Henry Huang and Johanna Griffin for the gifts of cloned DNAs; Tim Hunkapiller for the generation of the computerized data acquisition center; and Bernita Larsh for preparation of the manuscript. This work was supported by grants from the National Institutes of Health and a National Research Service Award from the National Institute of General Medical Sciences to S. K.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 18, 1981; revised October 13, 1981

References

Amzel, L., Poljak, R., Saul, F., Jarga, J. and Richards, F. (1974). The three dimensional structure of a combining-site ligand complex of immunoglobulin NEW at 3.5 Å resolution. Proc. Nat. Acad. Sci. USA *71*, 1427–1430.

Baltimore, D. (1981). Somatic mutation gains its place among the generators of diversity. Cell *26*, 295–296.

Benton, W. D. and Davis, R. W. (1977). Screening λgt recombinant clones by hybridization to single plaques in situ. Science *196*, 180–182.

Bernard, O., Hozumi, N. and Tonegawa, S. (1978). Sequences of mouse immunoglobulin light chain genes before and after somatic changes. Cell *15*, 1133–1144.

Blattner, F., Williams, B., Blechl, A., Denniston-Thompson, K., Faber, H., Furlong, L., Grunwald, D., Kieber, D., Moore, D., Schumme, J., Sheldon, E. and Smithies, O. (1977). Charon phages: super derivatives of bacteriophage lambda for DNA cloning. Science *196*, 161–169.

Blin, N. and Stafford, D. W. (1976). A general method for isolation of high molecular weight DNA from eukaryotes. Nucl. Acids Res. *3*, 2303–2308.

Bothwell, A. L. M., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. (1981). Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a γ2a variable region. Cell *24*, 625–637.

Brack, C., Hirama, M., Lenhard-Schuller, R. and Tonegawa, S. (1978). A complete immunoglobulin gene is created by somatic recombination. Cell *15*, 1–14.

Brenner, S. and Milstein, C. (1966). Origin of antibody variation. Nature *211*, 242–243.

Capra, J. D. and Kehoe, J. M. (1974). Variable region sequences of five human immunoglobulin heavy chains of the V_HIII subgroup: definitive identification of four heavy chain hypervariable regions. Proc. Nat. Acad. Sci. USA *71*, 845–848.

Cooper, M., Kearney, J., Lydyard, P., Grossi, C. and Lawton, A. (1976). Studies of generation of B-cell diversity in mouse, man, and chicken. Cold Spring Harbor Symp. Quant. Biol. *41*, 139–145.

Crews, S., Griffin, J., Huang, H., Calame, K. and Hood, L. (1981). A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. Cell *25*, 59–66.

Davis, M., Calame, K., Early, P., Livant, D., Joho, R., Weissman, I. and Hood, L. (1980a). An immunoglobulin heavy-chain gene is formed by at least two recombinational events. Nature *283*, 733–739.

Davis, M. M., Kim, S. K. and Hood, L. (1980b). DNA sequences mediating class switching in alpha immunoglobulin genes. Science *209*, 1360–1365.

Denhardt, D. T. (1966). A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. *23*, 641–646.

Early, P. and Hood, L. (1981). Mouse immunoglobulin genes. In Genetic Engineering Principles and Methods, *3*, J. Setlow and A. Hollaender, eds. (New York: Plenum Press), pp. 157–188.

Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980). An immunoglobulin heavy chain variable region is generated from three segments of DNA: V_H, D and J_H. Cell *19*, 981–992.

Gearhart, P. J., Johnson, N. D., Douglas, R. and Hood, L. (1981). IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. Nature *291*, 29–34.

Gough, N. and Bernard, O. (1981). Sequences at the joining region genes for immunoglobulin heavy chains and their role in generation of antibody diversity. Proc. Nat. Acad. Sci. USA *78*, 509–513.

Hohn, B. and Murray, K. (1977). Packaging recombinant DNA molecules into bacteriophage particles in vitro. Proc. Nat. Acad. Sci. USA *74*, 3259–3263.

Hood, L., Loh, E., Hubert, J., Barstad, P., Eaton, B., Early, P., Fuhrman, J., Johnson, N., Kronenberg, M. and Schilling, J. (1976). The structure and genetics of mouse immunoglobulins: an analysis of NZB myeloma proteins and sets of BALB/c myeloma proteins binding particular haptens. Cold Spring Harbor Symp. Quant. Biol. *41*, 817–836.

Hood, L., Davis, M., Early, P., Calame, K., Kim, S., Crews, S. and Huang, H. (1981). Two types of rearrangements in immunoglobulin genes. Cold Spring Harbor Symp. Quant. Biol. *45*, 887–898.

Kataoka, T., Kawakami, T., Takahashi, N. and Honjo, T. (1980). Rearrangement of immunoglobulin γ1-chain gene and a mechanism for heavy chain class switch. Proc. Nat. Acad. Sci. USA *77*, 919–923.

Kurosawa, Y., Von Boehmer, H., Haas, W., Sakano, H., Trauneker, A. and Tonegawa, S. (1981). Identification of D segments of immunoglobulin heavy chain genes and their rearrangements in T lymphocytes. Nature *290*, 565–570.

Lawton, A., Kearney, J. and Cooper, M. (1980). Immunoglobulin diversity. Regulation of expression of immunoglobulin genes during primary development of B cells. In Biological Regulation and Development, *2*, Organization and Cell Function, R. Goldberger, ed. (New York: Plenum Press), pp. 437–466.

Maki, R., Trauneker, A., Sakano, H., Roeder, W. and Tonegawa, S. (1980). Exon shuffling generates an immunoglobulin heavy chain gene. Proc. Nat. Acad. Sci. USA *77*, 2138–2142.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. and Efstratiadis, A. (1978). The isolation of structural genes from libraries of eucaryotic DNA. Cell *15*, 687–701.

Max, E., Seidman, J. and Leder, P. (1979). Sequence of five potential recombination sites encoded close to an immunoglobulin κ constant region gene. Proc. Nat. Acad. Sci. USA *76*, 3450–3454.

Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Pech, M., Hochtl, J., Schnell, H. and Zachau, H. (1981). Differences between germline and rearranged immunoglobulin sequences suggest a localized mutational mechanism. Nature *291*, 668–670.

Raff, M. C. (1976). Development and modulation of B lymphocytes: studies on newly formed B cells and their putative precursors in the hemopoietic tissues of mice. Cold Spring Harbor Symp. Quant. Biol. *41*, 159–162.

Sakano, H., Huppi, K., Heinrich, G. and Tonegawa, S. (1979). Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature *280*, 288–294.

Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. and Tonegawa, S. (1980). Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. Nature *286*, 676–683.

Seidman, J. G., Max, E. E. and Leder, P. (1979). A κ immunoglobulin gene is formed by site-specific recombination without further somatic mutation. Nature *280*, 370–375.

Smith, D. R. and Calvo, J. M. (1980). Nucleotide sequence of the E. coli gene coding for dihydrofolate reductase. Nucl. Acids Res. *8*, 2255–2274.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Sternberg, N., Tiemeier, D. and Enquist, L. (1977). In vitro packaging of a λ Dam-vector containing Eco RI fragments of E. coli and phage P1. Gene *1*, 255–280.

Valbuena, O., Marcu, K., Weigert, M. and Perry, R. (1978). Multiplicity of germline genes specifying a group of related mouse κ chains with implications for the generation of immunoglobulin diversity. Nature *276*, 780–784.

Weigert, M. and Riblet, R. (1976). Genetic control of antibody variable regions. Cold Spring Harbor Symp. Quant. Biol. *41*, 837–846.

Weigert, M., Cesari, H., Yonkovich, S. and Cohn, M. (1970). Variability in the lambda light chain sequences of mouse antibody. Nature *228*, 1045–1047.

Weigert, M., Gatmaitan, L., Loh, E., Schilling, J. and Hood, L. (1978). Rearrangement of genetic information may produce immunoglobulin diversity. Nature *276*, 785–790.

Weigert, M., Perry, R., Kelly, D., Hunkapiller, T., Schilling, J. and Hood, L. (1980). The joining of V and J gene segments creates antibody diversity. Nature *283*, 497–499.

Wu, T. T. and Kabat, E. A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chain and their implications for antibody complementarity. J. Exp. Med. *132*, 211–250.

Chapter 2

Appendix

The function of somatic hypermutation

One of the most interesting problems of immunology has been to understand how antigen recognition occurs. Classic studies by Landsteiner proved that the immune system is capable of distinguishing subtle differences in chemical structure suggesting that the total number of antigens which can be specifically recognized must be astronomical. In the humoral arm of the immune system, antigen recognition is accomplished with immunoglobulins. Thus, the key issue concerns how a specific immunoglobulin can be produced with the ability to bind a single antigen from an enormous array of potential antigens.

One explanation is that a large number of immunoglobulins are made so that at least one of them is capable of binding essentially any chemical molecule. Those B cells which produce appropriate immunoglobulins are induced to expand clonally. Thus the response to antigen is amplified and immunity is conferred. A problematical aspect of this clonal selection model is that the number of potential antigens, and thus the number of required immunoglobulins, are both astronomical. The molecular genetics of antibody genes have been well studied and it is clear that there are mechanisms for generating a large number of different immunoglobulins from a limited genome (for a review, see Tonegawa, 1983). A separate issue is whether the physical size of a mouse is sufficient to house enough B cells to express all of these immunoglobulins. There are less than 10^9 lymphocytes in a mouse; the immunoglobulin repertoire cannot exceed this maximum at any one time.

A second explanation is that antigen causes preimmune antibodies to change their binding specificities resulting in tighter binding. One advantage of this instructionist model is that the postulated size of the immunoglobulin repertoire is much smaller. However, a disadvantage is that, *a priori*, it is unclear how antigen can induce changes in binding specificity.

How does somatic hypermutation fit into the framework of antigen recognition described above? Somatic hypermutation is thought to occur as follows

(Gearhart et al., 1979; Crews et al., 1980; Kim et al., 1980). After the formation of a baseline repertoire, introduction of antigen causes appropriate B cells to selectively proliferate. As they do so, antibody genes undergo hypermutation and express receptors with either improved or lessened binding affinities. A second round of immunological selection can then select for B cells which bind antigen more tightly. Thus, hypermutation usually correlates with antibody class; IgG, but not IgM, antibodies are mutated (Gearhart et al., 1979; Bothwell et al., 1981). However, there are some exceptions, exemplified by antibodies which bind dextran (Schilling et al., 1980), in which IgM antibodies have mutations. In these cases, it is possible that antigenic encounter may induce hypermutation but not class switching. Alternatively, preimmune lymphocytes may undergo hypermutation.

Given that the above interpretation of somatic hypermutation is correct, the purpose of this appendix is to discuss its function. One view is that somatic hypermutation functions to generate antibody diversity. It is clear that many different types of antibodies are formed from germline sequences by this process. However, there are some observations which suggest that somatic mutation may do more than simply generate antibody diversity. First, other mechanisms for generating antibody diversity - i.e., multiple genes, combinatorial joining, junctional diversity, and combinatorial association- can potentially generate up to 10^{10} to 10^{13} different antibodies. At any one time, there are only 10^{6} to 10^{8} different clones of B cells in a mouse (Sigal and Klinman, 1979; Eichman, 1979). Therefore, a mouse can only express a subset of its potential antibodies. Even over its entire lifetime, the entire potential repertoire of immunoglobulins may not be expressed. In other words, additional mechanisms to generate antibody diversity may not be efficacious since the increased potential for diversity may not be expressed by an individual.

Second, it appears that somatic hypermutation primarily functions late in B cell ontogeny. Mechanisms which generate antibody diversity usually operate early in B cell ontogeny. This discrepancy suggests that somatic hypermutation may have additional functions which cause it to be developmentally regulated in this different fashion.

Third, the number of different antibodies due to hypermutation is inconsequential in comparison to those due to other means. For example, a mouse might have on the order of 10^8 different antibodies. Following antigenic challenge, on the order of dozens or hundreds of different antibody sequences may appear. When compared to the vast number already present, this number is insignificant. This conclusion is substantiated even when all stimulated B lymphocytes, responding to different antigens, are considered. Although there are some exceptions, mutation usually occurs in genes that have undergone class switch rearrangement. Since most B lymphocytes are of the IgM class and have not undergone class switching, most immunoglobulin genes would be expected to be unmutated.

Another view is that, through hypermutation, antigen contact induces the cell to alter the binding specificity of its immunoglobulin. Hypermutation most likely operates randomly so that immunoglobulins with decreased as well as increased affinities are produced. Those immunoglobulins with increased antigenic affinity can be selected immunologically.

As discussed below, this increase in affinity is likely to be immunologically significant. Hypermutation appears to act only on immunoglobulin V genes and only in mature B lymphocytes. Such a high level of regulation suggests that it must have an important function. Also, hypermutation is involved in many immune responses; variable region genes which undergo hypermutation include the lambda V region, the phosphorylcholine binding kappa V region, and the dextran, NP^b, arsonate, and phosphorylcholine binding heavy V regions (Brack et al., 1978; Bernard et al., 1978;

Selsing and Storb, 1981; Gershenfeld et al., 1982; Bothwell et al., 1981; Crews et al., 1980). The frequent involvement of hypermutation in the immune response indicates that it may have a fundamental role in producing immunoglobulins with antigenic specificity. Finally, one effect of hypermutation is probably detrimental, since random changes in immunoglobulin sequence may inactivate some B lymphocytes responding to antigen. It seems likely that the beneficial effect of hypermutation, to produce immunoglobulins with increased binding affinity, outweighs this detrimental effect. Otherwise, hypermutation would be selectively lost in phylogeny.

A corrollary to the conclusion that hypermutation plays an important role in the immune response is that without it, the immune repertoire may be too small. In other words, clonal expansion of exact replicas of preexisting B cells may not result in an effective immune response. The instructionist theories may have been correct to the extent that they postulated that new immunoglobulins must be made in response to antigen. These theories were probably incorrect in postulating how this might occur, namely with conformable receptors. However, in my opinion, this is inconsequential compared to the major issue which is whether enough different receptors can be expressed so that all potential antigens can be bound with sufficient affinity to elicit an immune response. Regarding this point, the instructionist theories may have been insightful.

The hypothesis that hypermutation functions partly to alter binding specificity in response to antigen makes several predictions. First, the rate of mutation should increase in those B cells which have been stimulated by antigen. Even though the frequency of mutation is much higher for B cells late, rather than early, in ontogeny, the rate of mutation is not known. One possibility is that the mutation rate is constant but that mutations accumulate over time. For phosphoryl-choline binding immunoglobulins, one out of 40 antibodies of the IgM

class and 15 out of 20 antibodies of the IgG class have mutated sequences (Perlmutter et al., 1984). In the absence of immunological selection, IgG^+ cells would have to be at least 100 times older than IgM^+ cells in order to have accumulated so many more mutations. Assuming that IgM^+ cells average 1 week in age (a minimal estimate), it would take longer than the lifetime of the mouse for IgG^+ cells to accumulate mutations so frequently. Thus, the notion that the increase in mutational frequency simply reflects differences in cellular age is probably untenable.

Rather than increased mutational rate, it is possible that immunologic selection is the cause of increased mutational frequency. Immunologic selection could cause those B cells with appropriately mutated immunoglobulin genes to proliferate. In this case, although the mutation rate of immunoglobulin genes would still be abnormally high, it would be constant throughout the immune response.

A second prediction is that the baseline repertoire is insufficient to reliably mount an effective immune response. In mice, it appears that the size of the baseline repertoire is limited by the number of B lymphocytes. Perhaps in a larger animal, the increased number of B cells would be reflected in an increased repertoire size. If the repertoire were sufficiently large, those immunoglobulins that are produced as a result of somatic hypermutation in mice might be already present in the preimmune animal. Thus, in a larger animal, somatic hypermutation may play a smaller role.

It is interesting to compare Darwin's theory of evolution, "survival of the fittest", with B cell development, "survival of the fittingest". Evolution results in the appearance of many more organisms which can overcome environmental selective pressures. An immune response results in the appearance of many more B cells which can respond to antigenic challenge. Some of these antigen-reactive B cells may be the result of hypermutation, i.e., targeted mutagenesis in response to

antigenic challenge. It is interesting to raise the question as to whether evolution might not also use a similar mechanism. There may be genes which are of central phylogenetic importance in the sense that they are frequently selected for speciation. It is possible that these genes, like immunoglobulin genes, are targeted for mutagenesis.

References

Bernard, O., Hozumi, N., and Tonegawa, S. 1978. Sequence of mouse immunoglobulin light chain genes before and after somatic changes. Cell **15:** 1133-1144.

Bothwell, A., Paskind, M., Reth, M., Imanishi-Kari, T., Rajewsky, K. and Baltimore, D. 1981. Heavy chain variable region contribution to the NP^b family of antibodies: somatic mutation evident in a gamma2a variable region. Cell **24:** 625-637.

Brack, C., Hirama, M., Lenhard-Schuller, R. and Tonegawa, S. 1978. A complete immunoglobulin gene is created by somatic recombination. Cell **15:** 1-14.

Crews, S., Griffin, J., Huang, H., Calame, K. and Hood, L. 1981. A single V_H gene segment encodes the immune response to phosphorylcholine: somatic mutation correlates with the class of the antibody. Cell **25:** 59-66.

Eichmann, K. 1978. Expression and function of idiotypes on lymphocytes. Adv. Immunol. **26:** 195-254.

Gearhart, P., Johnson, N., Douglas, R. and Hood, L. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. Nature **291:** 29-34.

Gershenfeld, H., Tsukamoto, A., Joho, R. and Weissman, I. 1982. Somatic diversification is required to generate the V_{kappa} genes of MOPC 511 and MOPC 167 myeloma proteins. Proc. Nat. Acad. Sci. U.S.A. **78:** 7674-7677.

Kim, S., Davis, M., Sinn, E., Patten, P. and Hood, L. 1981. Antibody diversity: somatic hypermutation of rearranged V_H genes. Cell **27:** 573-581.

Perlmutter, R., Crews, S., Douglas, R., Sorenson, G., Johnson, N., Nivera, N., Gearhart, P. and Hood, L. 1984. Adv. Immunol. **35:** 1-35.

Shilling, J., Clevinger, B., Davie, J. and Hood, L. 1980. Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangements in heavy chain V region gene segments. Nature **282:** 35-40.

Sigal, N. and Klinman, N. 1978. The B cell clonotype repertoire. Adv. Immunol. **26:** 255-341.

Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature **302:** 578-581.

Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6994–6998, November 1982
Immunology

Molecular cloning of translocations involving chromosome 15 and the immunoglobulin C_α gene from chromosome 12 in two murine plasmacytomas

(class switching/oncogenesis)

KATHRYN CALAME*, STUART KIM†, PETER LALLEY‡, ROBERT HILL§, MARK DAVIS†¶, AND LEROY HOOD†

*Department of Biological Chemistry, University of California School of Medicine, Los Angeles, California 90024; †Division of Biology, California Institute of Technology, Pasadena, California 91125; ‡Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830; and §Department of Cell and Tumor Biology, Roswell Park Memorial Institute, Buffalo, New York 14263

Contributed by Leroy Hood, August 17, 1982

ABSTRACT **Expression of IgA by plasmacytomas occurs as a result of a DNA rearrangement that brings the variable region gene, V_H, a few kilobases 5′ to the constant region gene, C_α. In this study, we show that the allelic nonexpressed C_α gene also is rearranged in most plasmacytomas. Cloning, restriction mapping, heteroduplex analyses, and sequence analyses of the nonproductively rearranged C_α genes from two plasmacytomas, M603 and M167, have demonstrated that the nonproductive rearrangement occurs within the α switching region, S_α. In each case, the same DNA sequence has been joined to the 5′ side of C_α and we have termed this DNA "NIRD" (for nonimmunoglobulin rearranged DNA). Southern blotting analyses of genomic DNAs from various IgG-, IgM-, or IgA-producing plasmacytomas suggest that NIRD is rearranged in almost all plasmacytomas. However, NIRD rearranges to the S_α region only in IgA-producing cells, not in IgM or IgG producers. Cytogenetic evidence has shown that T(12;15) translocations are common in murine plasmacytomas. Immunoglobulin heavy chain genes are located on chromosome 12, and the translocation breakpoint in plasmacytomas occurs near the immunoglobulin genes. NIRD has been mapped to chromosome 15 by Southern blotting analysis of mouse–hamster cell lines, suggesting that the nonproductively rearranged C_α clones represent the T(12;15) translocations identified cytogenetically. Therefore, we have identified a region of DNA on chromosome 15 that is commonly rearranged in transformed mouse lymphocytes. We speculate on the significance of NIRD in neoplastic transformation of mouse lymphocytes.**

Murine plasmacytomas (also referred to as myelomas) have been observed to have specific chromosomal translocations in which the distal portion of chromosome 15 is translocated to either chromosome 12 or chromosome 6 (1). The immunoglobulin heavy chain and κ light chain gene families are located on chromosomes 12 and 6, respectively, and the translocational breakpoint on chromosome 12 occurs at the same region of the chromosome as do the immunoglobulin genes. Trisomy of chromosome 15 occurs in most murine T-cell and some B-cell leukemias. Based on these observations, it has been suggested that a cellular oncogene located on the distal portion of chromosome 15 may become abnormally activated, through translocation or triplication, and result in the neoplastic transformation of murine lymphoid cells (1).

One explanation for the involvement of the immunoglobulin loci in the chromosomal translocations is that immunoglobulin genes undergo DNA rearrangements in B cells. These rearrangements may predispose the immunoglobulin genes to undergo chromosomal translocation. Conceivably, translocation of immunoglobulin genes, which are actively transcribed in B cells, could increase the rate of transcription of genes residing near the recombinational breakpoint. There are two types of immunoglobulin gene rearrangements (reviewed in ref. 2). First, separate variable (V), diversity (D), and joining (J) gene segments are rearranged to create a functional V gene. These V-D-J or V-J joining events also are denoted V gene formations for the heavy and light chains, respectively. The second type of rearrangement, heavy chain or C_H switching, occurs among the closely linked C_H genes of which there are eight in the mouse ($5'C_\mu$-C_δ-$C_{\gamma 3}$-$C_{\gamma 1}$-$C_{\gamma 2b}$-$C_{\gamma 2a}$-C_ε-$C_\alpha 3'$). The C_H genes determine the class or isotype of the antibody—e.g., C_μ makes IgM, C_α makes IgA, etc. All C_H genes except for C_δ have repetitive elements to their 5′ sides that are denoted switch (S) regions. For example, if the B cell is to switch from its initial expression of IgM to the synthesis of IgA, then the S_μ and S_α regions are joined, deleting the intervening DNA and placing V_H near C_α.

Immunoglobulin gene expression in B cells exhibits allelic exclusion; that is, in an individual B cell, only one heavy chain allele and one light chain allele are expressed. The chromosome containing the expressed immunoglobulin gene undergoes a productive DNA rearrangement. Often, the chromosome carrying the unexpressed gene also is rearranged and these are denoted nonproductive rearrangements.

In the study reported here, we analyzed the nature of nonproductive C_α rearrangements in murine plasmacytomas. Our results show that most IgA-producing plasmacytomas have a nonproductively rearranged C_α on a 13-kilobase (kb) *Eco*RI fragment. Analysis of nonproductively rearranged C_αs from two plasmacytomas, M603 and M167, demonstrated that the nonproductive rearrangements occurred within the S_α region in both cases and that the same 5′ sequence had been joined to C_α. We have denoted this sequence "nonimmunoglobulin rearranging DNA" or NIRD. Southern blotting analyses indicated that, whereas most IgA-expressing tumors have a NIRD-S_α rearrangement, IgG- or IgM-expressing tumors have different NIRD rearrangements. Thus, in IgA-producing plasmacytomas, both S_α regions are rearranged whereas in IgG-producing plasmacytomas, neither S_α region is rearranged.

Most significantly, we also have demonstrated that the NIRD sequence is located on mouse chromosome 15. Because the C_α gene originates from chromosome 12, these results demonstrate that the nonproductively rearranged C_α clones from M603 and M167 represent the junction of a translocation involving chro-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: V, variable; D, diversity; J, joining; S, switch; kb, kilobase(s); NIRD, nonimmunoglobulin rearranging DNA; bp, base pair(s).

¶ Present address: Laboratory of Immunology, National Institutes of Health, Bethesda, MD 20205.

mosome 15 and chromosome 12. Furthermore, the results of the Southern blotting analyses mentioned above define a small region on chromosome 15 that is almost always rearranged in murine plasmacytomas.

METHODS

Construction and Isolation of the CHM603α30 and CHM167α7 Clones. CHM603α30 was isolated from an *Eco*RI partial library of plasmacytoma M603 DNA in Charon 4A by screening with a C_α cDNA probe (3). CHM167α7 was isolated using a 5′ C_α genomic clone probe to screen a *Hae*III/*Alu* I-digested *Eco*RI linker library of plasmacytoma M167 DNA in Charon 4A (4).

DNA Blots. Ten micrograms of total genomic DNA was digested to completion with restriction endonuclease, size-separated on a 1% (wt/vol) agarose gel, and transferred to nitrocellulose filters (5). Filters were prehybridized, hybridized, and washed by using the conditions of Wahl *et al.* (6). Washed filters were exposed to preflashed Kodak XAR-5 film at −70°C with a Dupont Cronex Lightning Plus intensifying screen for 1–3 days.

Somatic Cell Hybrids. Somatic cell hybrids between Chinese hamster and mouse spleen cells (BALB/c) were generated and maintained as described (7). The mouse chromosome content of the hybrid clones was determined enzymatically and cytogenetically. Each hybrid clone was analyzed for the expression of 25 enzymes representing linkage group assigned to 16 of the 19 autosomes and the X chromosome as described (8, 9). Enzymatic, cytogenetic, and Southern blotting analyses were carried out on parallel cultures of each hybrid clone.

Preparation of NIRD Probe. The 440-base-pair (bp) *Msp* I fragment from α30 (see Fig. 2) was gel purified and subcloned into the *Cla* I site of pBR322.

Other Analyses. Heteroduplex analysis was carried out as described by Davis *et al.* (10). DNA sequence analysis was performed by the partial cleavage method of Maxam and Gilbert (11).

RESULTS

Most IgA-Producing Plasmacytomas Have Similar Nonproductive C_α Gene Rearrangements. In preliminary experiments, we analyzed genomic Southern blots of DNA from eight independently arising IgA-producing plasmacytomas with a 5′ C_α probe (unpublished data). As expected, sperm DNA digested with *Eco*RI revealed a germ-line 5′ C_α band at 9.5 kilobases (kb) whereas the different plasmacytomas gave variable-size fragments which corresponded to the productively rearranged C_α gene. Unexpectedly, however, a 13-kb *Eco*RI fragment was present in seven plasmacytomas (H8, M511, M167, M603, S63, T15, and Y5236) but was absent from one (W3207). This finding suggested that similar or possibly identical nonproductive C_α gene rearrangements occurred repeatedly and prompted us to clone and analyze examples of such nonproductively rearranged C_α genes.

The construction and screening of genomic libraries from the DNA of the plasmacytomas M603 and M167 have been described elsewhere (3, 4). Clones representative of the 13-kb *Eco*RI fragment from M603 (CHM603α30) and M167 (CHM167α7) are shown in Fig. 1.

Nonproductive C_α Clones α7 and α30 Are Identical by Heteroduplex Analysis. Heteroduplex analyses of the M603 nonproductively rearranged C_α clone α30 with the germ-line C_α clone α29 indicated that the two clones are homologous throughout the C_α gene region and for about 2.6 kb 5′ to the C_α gene (Fig. 1). Thus, there has been a rearrangement 2.6 kb

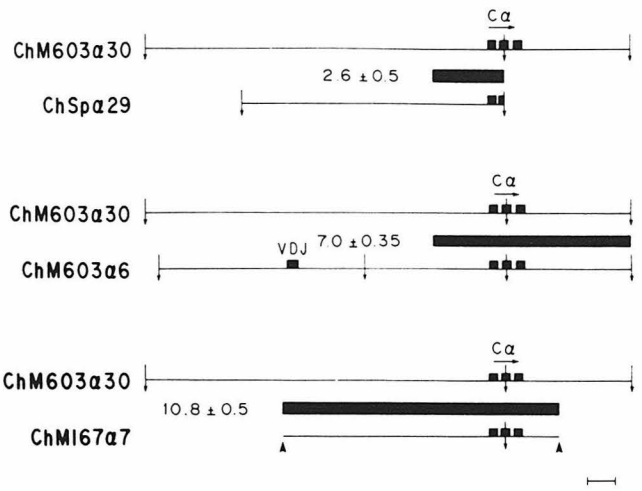


FIG. 1. Comparison of CHM603α30 to other C_α clones by heteroduplex analysis. Homology between α30 and CHSpα29 (a germ-line C_α clone), CHM603α6 (containing the expressed C_α gene from plasmacytoma M603), and CHM167α7 is shown by solid bars (sizes shown in kb). Natural *Eco*RI sites are indicated by arrows; linker *Eco*RI sites are indicated by triangles. CHM167α7 contains 3.6 kb of additional DNA 3′ to the C_α gene which is not shown in the figure because our blotting studies indicate that it is a cloning artifact and it does not occur 3′ to the C_α gene in the M167 genome (data not shown). Scale marker = 1 kb.

5′ to the C_α gene in α30 and the rearrangement occurs within the S_α region. Heteroduplex analyses comparing the α30 clone to the productively rearranged M603 C_α clone α6 (3) revealed that the 5′ half of α30 is not homologous to the 5′ portion of α6 and thus demonstrate that in α30 the sequences upstream from the S_α sequences are neither J_H sequences nor M603 V_H sequences. However, when the two nonproductively rearranged C_α clones, α30 and α7, were compared to one another by heteroduplex analysis, homology was observed not only throughout the C_α genes but also in their 5′ flanking sequences for at least 7.5 kb. These results therefore demonstrate that the nonproductively rearranged C_α genes in the M603 and M167 plasmacytomas have identical or similar DNA sequences joined to the C_α gene. This sequence, of unknown origin, is denoted NIRD.

NIRD Sequences Are Joined to the S_α Region. In order to characterize the sites at which NIRD sequences and the germline C_α flanking sequences were joined, we determined the DNA sequences at the boundaries between these two DNAs. Fine structure restriction mapping of the two clones in a 1.5-kb region surrounding the S sites (Fig. 2) confirmed the homology observed by heteroduplexing and led to the sequence-determination strategies shown. The sequences illustrated in Fig. 3 demonstrate that the S_α recombination point in the two clones differs by 94 bp. In addition, clone α7 contains approximately 70 bp more NIRD sequences than does clone α30. Thus, although clones α7 and α30 are homologous by heteroduplex analyses, they join NIRD and germ-line C_α flanking sequences at slightly different points.

Because clones α7 and α30 both have recombinational breakpoints within the S_α region, it is likely that the rearrangement was due to a nonproductive C_H switching event. Indeed, the S_α recombination breakpoint from α30 lies only 2 bp away from the S_α site used in generating the expressed α gene in the plasmacytoma T15 (12).

Most IgA-Producing Plasmacytomas Have Nonproductive α Genes That Contain NIRD. The identity of the NIRD sequence in the M167 and M603 nonproductively rearranged C_α

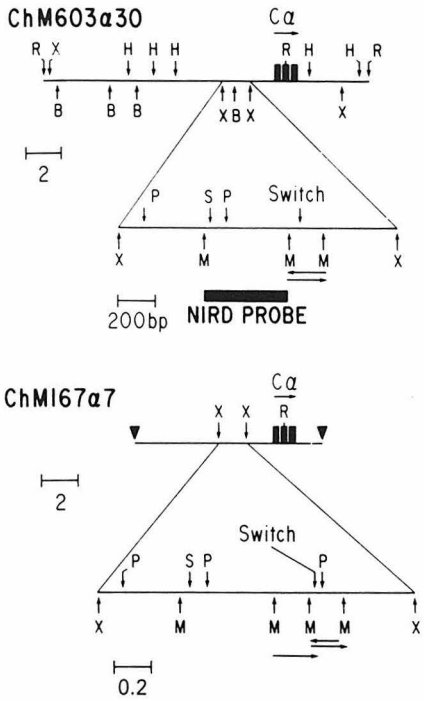


FIG. 2. Restriction map of α30 and α7 in the vicinity of the S region. The regions subjected to sequence analyses are indicated by arrows. Both clones contain one additional *Msp* I site that is not mapped. The solid bar indicates the 440-bp *Msp* I fragment used as a NIRD probe. The portion outside of the 1.7-kb *Xba* fragment in α7 was not mapped. R, *Eco*RI; X, *Xba* I; P, *Pst* I; M, *Msp* I; S, *Sac* I; B, *Bam*HI; H, *Hin*dIII. Scale markers are in kb.

genes led us to subclone a DNA fragment from this region to use as a probe in the analysis of other IgA-producing plasmacytomas to determine whether their nonproductive C_α rearrangements use NIRD sequences. A 440-bp *Msp* I restriction fragment was subcloned to use as a hybridization probe (Fig. 2). This subclone is referred to as the NIRD probe.

Fig. 4 shows the results of Southern blot analyses of genomic DNAs from five IgA-producing plasmacytomas. Embryo and liver DNAs contained NIRD on a 16-kb *Eco*RI fragment whereas the five IgA-producing plasmacytomas (M167, H8, T15, W3082, and M511) contained NIRD on the 16-kb fragment and on an additional 13-kb *Eco*RI fragment. Because the DNA from the H8, T15, and M511 tumors revealed a 13-kb *Eco*RI band when hybridized with either a NIRD or a C_α probe, we

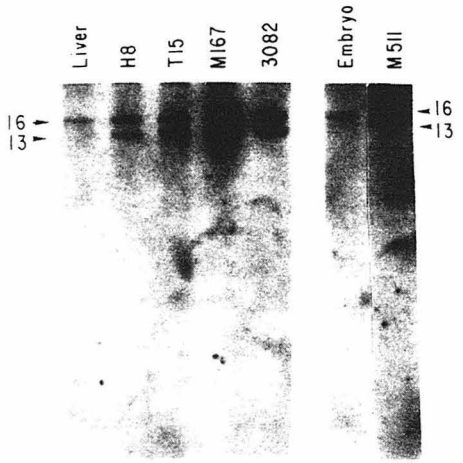


FIG. 4. Genomic blots of germ-line and plasmacytoma DNAs with the NIRD probe. DNAs were digested with *Eco*RI.

conclude that, like M603 and M167, they have NIRD sequences 5′ to the nonproductively rearranged C_α gene. In the M511 case, the juxtaposition of NIRD with C_α was further shown by using Southern blotting analysis of *Hin*dIII- and *Xba* I-digested DNAs (data not shown). By analogy, the above blotting results suggest that S63 and Y5236, which have a rearranged 13-kb *Eco*RI C_α band, and W3082, which has a rearranged 13-kb *Eco*RI NIRD band, have the NIRD/C_α rearrangement. Finally, comparison of the NIRD restriction sites from α7 or α30 to the restriction sites from nonproductive C_α gene clones from the IgA-producing plasmacytomas S107 (13) or J558 (14) showed that these clones also have a NIRD/S_α rearrangement. Thus, it appears that as many as 10 of 11 IgA-producing plasmacytomas may have nonproductively rearranged S_α regions joined to NIRD. In other words, rearrangement of NIRD to the S_α region is a frequent event in nonproductive C_H switching in IgA-producing tumors. Moreover, the rearrangements in each IgA-producing plasmacytoma appear to have occurred to nearly the same area within S_α and NIRD because the size of the restricted site fragments spanning the recombination breakpoint varied only slightly.

NIRD Rearrangements Occur in Non-IgA-Producing Cell Lines but Are Not Closely Associated with the Allelic Counterpart of the Expressed C_H Gene. The status of NIRD sequences in cells expressing isotypes other than IgA was deter-

```
                                                                    Sα
                                                                    |
M603      GCGGGATTGGCGGGCGGCGACCTCCGTTGTCGGTCCCAGGCCTCCAGAAATGTACCAAGCTAAATTTAAATGCCTCTCAGAGACTGGTTGAGCTGAACTA      100
SP Sα     CTGGGATGAGACAGGCTGACTGCAGGAGGAAGACTGGAAGGGCTGGCTGAGCTAGACTAGGCTGGGCTGAGCTGGAATGAGCTG ——————————————      1213

          GTATAAACTTGGCTAGGCTACAATGGATTGAGCTGAGCTAGACTTAGGGTGGAATGGGCTGAACAAGGCTGAGCTTACCTAGACGC                  186
          ————————————————————————————————————————————————————————————————————————————————————                  1299

                                                Sα
                                                |
M167      TTTCTTTGGCCACCACCGTAAGCGACCTCCCGGTTTGACCCAAACTAAGCTGGGATGAGACAAGCTGGACTGCAGGAGGAAGACTGGAAGGGCGTGACTG      100
SP Sα     TGAGCTGAGCTGAGCTGGGCTAAGCTGGGATGGACTAGGAT ——————————————————————————————————————————————————————      1164

          AGCTAGACTAGGCTGGGCTGAGCTGGAATGAGCTGGGTTGAGCTGAACTAGTATAAACTTGGCTAGGCTACAATGGATTGAGCTGAGCTAGACTTAGGGT      200
          ————————————————————————————————————————————————————————————————————————————————————————————      1264

          GGAATGGGCTGAACAAGGCTGAGCTTACCTAGACC                                                                    235
          ——————————————————————————————————                                                                     1299
```

FIG. 3. DNA sequence of germ-line S_α (α29), M603 rearranged S_α (α30), and M167 rearranged S_α (α7) DNAs at the S sites. Nucleotide identities are indicated by a line. Numbering of germ-line S_α sequences is from ref. 12.

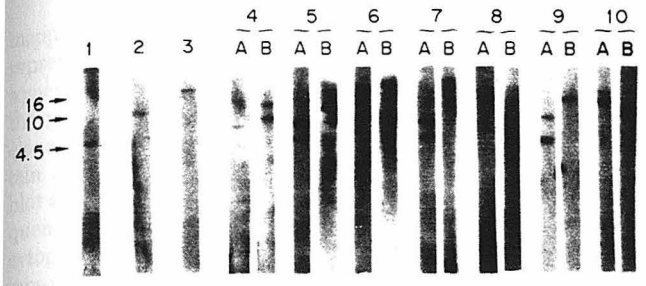


FIG. 5. Genomic blots of non-IgA-producing plasmacytomas. Lanes: 1–3, liver DNA; 4–10, plasmacytoma DNAs; A, probed with the NIRD probe; B, same filter probed with C_H cDNA probes after washing for 3 hr in 5 mM Tris, pH 8.0/0.2 mM EDTA/0.05% sodium pyrophosphate at 68°C. Details: lanes 1–3, liver DNA digested with *Hin*dIII, *Kpn* I, and *Eco*RI, respectively, and probed with the NIRD probe; 4B, E1 (IgG2b) DNA digested with *Hin*dIII, probed with a $C_{\gamma 2b}$ probe; 5B and 6B, IB7 (IgG3) DNA digested with *Hin*dIII and *Eco*RI, respectively, probed with a $C_{\gamma 3}$ probe; 7B and 8B, P3K (IgG1) DNA digested with *Hin*dIII and *Eco*RI, respectively, probed with a $C_{\gamma 1}$ probe; 9B and 10B, HPCM2 (IgM) DNA digested with *Eco*RI and *Kpn* I, respectively, and probed with a C_μ probe.

mined. Genomic DNA from an IgM-producing hybridoma (M2), an IgG3-producing hybridoma (IB7), an IgG1-producing plasmacytoma (P3K), and an IgG2b-producing plasmacytoma (E1) was analyzed by Southern blotting. Fig. 5 shows the results of several enzyme digestions in which the same filter was hybridized first with the NIRD probe, washed, and then hybridized with a cDNA probe corresponding to the isotype expressed by the cells from which the DNA was prepared. In all cases, a rearranged NIRD band was observed in addition to the germline NIRD band seen in liver DNA.

The parental cell line used in hybridoma fusions, SP2, exhibits a 13.0-kb *Eco*RI band and a 6.6-kb *Hin*dIII band when hybridized to a NIRD probe (14). Both M2 and IB7 exhibited the same size *Eco*RI and *Hin*dIII bands as SP2 when hybridized to a NIRD probe (Fig. 5; unpublished data). This similarity suggests that the rearranged NIRD bands in hybridomas M2 and IB7 were donated by the parental cell line, SP2.

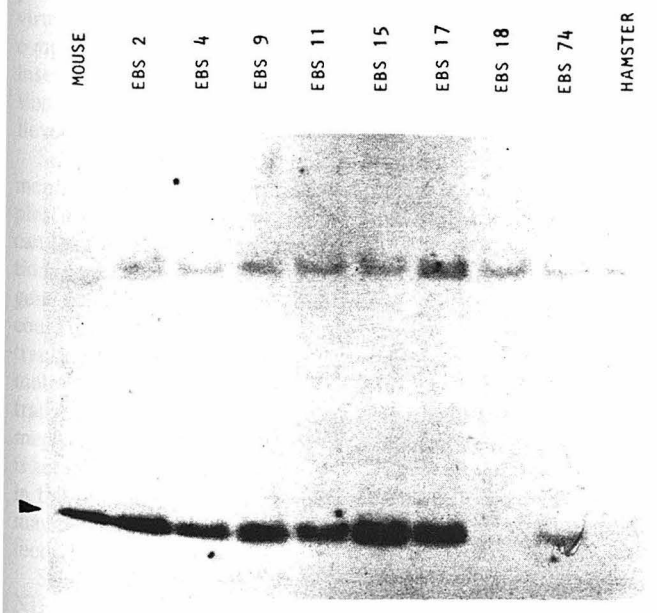


FIG. 6. Genomic blot of somatic hamster–mouse cell hybrid DNAs. Genomic DNA from mouse liver, hamster, and hamster–mouse somatic cell hybrids was digested with *Hin*dIII and hybridized to a 1.2-kb *Hin*dIII NIRD probe subcloned into pBR322 from α30. Arrowhead, 1.2-kb *Hin*dIII band in mouse liver DNA.

Table 1. Segregation of NIRD in mouse–hamster hybrids

Chromosome	Marker enzyme	Concordant	Discordant
1	PEP-3	3	5
2	AK-1/SODH/ACP-2	5	3
3	*	5	3
4	PGD/PGM-2	3	5
5	PEP-7/PGM-1	2	6
6	TPI	3	5
7	LDH-1/GPI/PEP-4	6	2
8	GR/APRT	4	4
9	ME/MPI	4	4
10	PEP-2/HK-1	4	4
11	GLK	1	7
12	ACP-1	6	2
13	*	6	2
14	ES-10	1	7
15	*	8	0
16	SOD-1	4	4
17	GLO	6	2
18	PEP-1	2	6
19	GOT	5	3
X	HPRT	7	1

The marker enzymes, their chromosomal assignments, and the procedures used to separate the Chinese hamster and mouse enzymes have been described (8, 9). Marker enzyme analysis and karyotype analysis agreed for each chromosome. Enzyme, karyotype, and Southern blotting analyses were performed on the same passage for each hybrid cell line.

* Chromosomes 3, 13, and 15 have no enzyme markers; their concordance/discordance was based on karyotype analysis alone.

In no case, however, was the rearranged NIRD band found on the same size restriction fragment as the relevant C_H gene. We also have found that NIRD is not rearranged to the C_α gene in these cells because the C_α gene and S_α region remain in germline configuration (data not shown). These results show that NIRD is rearranged in almost all plasmacytomas and hybridomas analyzed in our study regardless of which isotype is expressed (IgM, IgG, or IgA). Although in IgA-producing cells NIRD rearranges to the S_α region, in IgG- or IgM-producing cells NIRD rearranges neither to S_α nor to the S region used for rearrangement on the expressing chromosome.

NIRD Originates from Chromosome 15 as Shown by Somatic Cell Hybrid Mapping. Chinese hamster–mouse somatic cell hybrids that selectively lose mouse chromosomes can be used to map mouse sequences to individual chromosomes. Different hamster–mouse hybrid cell lines contain different arrays of mouse chromosomes as determined by enzyme and karyotype analysis. DNA was prepared from these hybrid cells and analyzed by Southern blots for the presence of NIRD sequences. NIRD sequences mapped concordantly with mouse chromosome 15 but discordantly with all other mouse chromosomes (Fig. 6; Table 1). Since the C_α genes originate from chromosome 12 and the NIRD sequences originate from chromosome 15, the α30 and α7 clones, from the plasmacytomas M603 and M167, span a translocation involving chromosome 15 and chromosome 12.

DISCUSSION

This paper describes clones isolated from two IgA-producing murine plasmacytomas in which the same DNA sequence is joined to the 5′ side of the C_α gene by a nonproductive C_α switch rearrangement. This sequence, termed NIRD, has been

mapped to chromosome 15, thus showing that these two clones represent the junction of a translocation involving chromosomes 12 and 15. Nonproductively rearranged C_α genes from the plasmacytoma tumors S107 (13) and J558 (14) have been cloned and, by restriction site similarity, these clones also contain a NIRD/S_α rearrangement. Furthermore, by Southern blot analysis presented here and elsewhere (13, 14), NIRD sequences are rearranged in 40 of 48 plasmacytomas. Previous cytogenetic evidence has shown that the distal portion of chromosome 15 is frequently translocated to either chromosome 12 or chromosome 6 in murine plasmacytomas (15, 16). The simplest explanation of these observations is that the α30 and α7 clones represent the T(12;15) translocations and that, in murine plasmacytomas, the NIRD sequence on chromosome 15 is frequently translocated, along with the distal portion of the chromosome, to the heavy chain immunoglobulin locus on chromosome 12 or to the κ chain immunoglobulin locus on chromosome 6.

Several lines of evidence suggest that NIRD rearrangement is limited to plasmacytomas. First, thymus, liver, and embryo DNAs show only the germ-line NIRD band (Fig. 4; unpublished data). Second, NIRD rearrangement is not observed in *Eco*RI-digested DNAs from a helper T-cell line (HT-1), a cytotoxic T-cell line (CTLL16), or two suppressor T-cell hybridomas (B3B5 and C4#4) (unpublished data). Third, except for the NIRD rearrangement donated by the parental cell line, SP2, NIRD rearrangement was not observed in 10 hybridomas, suggesting that NIRD is not rearranged in normal B lymphocytes (Fig. 5; ref. 14). Thus, these results suggest that NIRD rearrangement is specific for plasmacytomas.

One plausible explanation for NIRD rearrangement in plasmacytomas is that NIRD may contain an oncogene residing on chromosome 15 which can be activated by translocation to chromosome 12. For example, when B cells undergo C_H switching, DNA rearrangement of S regions may predispose the immunoglobulin heavy chain genes to translocation. In those cases in which an S region joins to NIRD, the NIRD oncogene may become activated, contributing to neoplastic transformation. A precedent for this hypothesis can be found in avian leukosis virus-transformed cells (17). In this case, a cellular oncogene, c-*myc*, is expressed at a higher level when it is activated by the insertion of a retrovirus, avian leukosis virus. Increased expression of c-*myc* by retrovirus-mediated promotor insertion is believed to cause neoplastic transformation in these cells.

Several points are worth noting regarding NIRD rearrangements. First, although promotor insertion seems to be the simplest explanation, it is not known how T(12;15) translocations can lead to neoplastic transformation. For example, translocations could derepress the NIRD gene. Alternatively, the NIRD gene could be a regulatory gene whose activation or repression could control oncogenes residing elsewhere. Furthermore, the translocated immunoglobulin gene is missing its normal promotor that resides with the V_H gene. Thus, even if T(12;15) translocations activated an oncogene residing at NIRD, the mechanism of activation is not clear. Perhaps the translocation is actually reciprocal so that a 5′ V_H-NIRD 3′ sequence resides in the plasmacytomas. Alternatively, the translocated C_α gene may activate nearby sequences via enhancer or chromatin effects (18).

Second, because some plasmacytomas show an unrearranged NIRD band on Southern blots, there may be a different mechanism for transformation in these cells. Alternatively, rearrangements may occur on chromosome 15 that are undetected by the NIRD probe. For example, by Southern blotting analysis, MOPC315 is reported to have no NIRD rearrangement (14) whereas a translocation involving chromosome 15 (D3/E) is detected cytogenetically (15).

Third, because chromosome 15 is observed to translocate to chromosome 6 (16) (which contains the immunoglobulin κ chains) and because light chains do not undergo C_H switching, it is possible that translocations may be induced by *V-J* or *V-D-J* joining as well as by C_H switching.

Fourth, the type of NIRD rearrangement found in plasmacytomas seems to correlate with the expressed isotype. For example, cloning and Southern blotting analysis suggest that NIRD is joined to S_α in 10 of 11 IgA-producing plasmacytomas. However, NIRD is not joined to S_α in four IgM- and IgG-producing lines.

Fifth, because trisomy of chromosome 15 is observed in murine T- and B-cell leukemias (3), it is possible that NIRD is involved in leukemogenesis. In addition, in human Burkitt lymphoma, translocations involving chromosome 8 and chromosome 2, 14, or 22 are known to occur (3). Because chromosome 2 contains the κ immunoglobulin genes, chromosome 14 contains the heavy chain genes, and chromosome 22 contains the λ light chain genes, it is possible that an analogous situation exists in these tumors in which a sequence similar to NIRD resides on human chromosome 8.

Finally, NIRD rearrangement in murine plasmacytomas may not be involved in oncogenesis at all. These rearrangements may have significance in a way not yet identified.

While this manuscript was in preparation, Marcu and coworkers (19) reported that a similar T(12;15) chromosomal translocation event occurs in the murine plasmacytoma J558.

The authors wish to thank Roy Riblet for the gift of NZB plasmacytoma DNAs and Mitch Kronenberg for the gift of T-cell hybridoma DNAs. We thank R. Wall for supplying the P3K and E1 cell lines. We thank Steve Crews, Rick Barth, Roger Perlmutter, and Barbara Wold for helpful advice and discussions. The assistance of Connie Katz, Bernita Larsh, and Kathy Kanamori in preparing the manuscript is gratefully acknowledged. This work was supported by National Institutes of Health grants to K.C. and to L.H.; S.K. and M.D. were supported by a National Institutes of Health training grant.

1. Klein, G. (1981) *Nature (London)* **294,** 313–318.
2. Early, P. & Hood, L. (1981) in *Genetic Engineering, Principles and Methods*, eds. Setlow, J. & Hollander, A. (Academic, New York), Vol. 3, pp. 157–188.
3. Davis, M., Early, P., Calame, K., Livant, D. & Hood, L. (1979) in *Eukaryotic Gene Regulation*, ICN-UCLA Symposium on Molecular and Cellular Biology, eds. Axel, R., Maniatis, T. & Fox, C. F. (Academic, New York), Vol. 14, pp. 393–406.
4. Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. (1981) *Cell* **27,** 573–581.
5. Southern, E. (1975) *J. Mol. Biol.* **98,** 503–527.
6. Wahl, G. M., Stern, M. & Stark, G. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3683–3687.
7. Minna, J. D., Marshall, T. H. & Shaffer-Berman, P. U. (1975) *Somatic Cell Genet.* **1,** 355–369.
8. Lalley, P. A., Franke, U. & Minna, J. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2382–2386.
9. Lalley, P. A., Minna, J. & Franke, U. (1978) *Nature (London)* **274,** 160–162.
10. Davis, R., Simon, M. & Davidson, L. (1971) *Methods Enzymol.* **12,** 413–428.
11. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–559.
12. Davis, M., Kim, S. & Hood, L. (1980) *Science* **209,** 1360–1365.
13. Kirsch, I. R., Ravetch, J. V., Kwan, S. P., Max, E. E., Ney, R. L. & Leder, P. (1981) *Nature (London)* **293,** 585–587.
14. Harris, L., Lang, R. & Marcu, K. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4175–4179.
15. Shepard, J. S., Pettengill, O. S., Wurster Hill, D. H. & Sorenson, G. D. (1978) *J. Natl. Cancer Inst.* **61,** 255–258.
16. Ohno, S., Babonits, M., Wiener, F., Spira, J., Klein, G. & Potter, M. (1979) *Cell* **18,** 1001–1007.
17. Hayward, W. S., Neel, B. G. & Astrin, S. M. (1981) *Nature (London)* **290,** 475–480.
18. Weisbrod, S. (1982) *Nature (London)* **297,** 289–295.
19. Harris, L. J., D'Eustachio, P., Ruddle, F. H. & Marcu, K. B. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6622–6626.

Chapter 4

Rearrangment of switch regions in a Pre-B cell line

Introduction

An immunoglobulin heavy chain is composed of a variable region, V_H, which helps bind antigen, and a constant region, C_H, which carries out effector functions. Each C_H region may have a different effector function. For example, IgM appears as a pentamer and can react to low concentrations of antigen. IgG is effective at activating complement. IgE can stimulate mast cells and IgA can cross the intestinal epithelium to function within the gut.

Antigenic challenge stimulates appropriate B cells, which originally express IgM, to respond in two ways. First, they proliferate and secrete IgM. This response is effective in clearing antigen. Second, they undergo class switching by producing immunoglobulins with the same V_H region but with a different C_H region (Cooper et al., 1977; Pernis, 1977). Thus, the capability to bind a particular antigen is retained while the effector response is modified. The type, amount and location of antigen may affect class switching. For example, IgA expression is observed from B cells in the Peyer's Patch whereas B cells expressing IgG are observed in peripheral lymph nodes.

Immunoglobulin class switching involves a DNA rearrangement (fig. 1A, Davis et al., 1979; Kataoka et al., 1980; Sakano et al., 1980). Initially, a V_H gene segment is linked to a C_u gene segment. Class switching causes a deletion beginning in the S_u region and ending in another S region. As a result, the C_u gene segment is deleted and replaced by another C_H gene segment.

In order to characterize the mechanism of class switching, the sequence of the switch regions and the endpoints of the rearrangements have been studied. Switch regions consist of many small repeats (Dunnick et al., 1980; Takahashi et al., 1980; Kataoka et al., 1981; Davis et al., 1980; Sakano et al., 1980; Obata et al., 1981). In the S_u and S_{alpha} regions, which have been the most extensively studied, some independently derived rearrangement breakpoints occur within a few

nucleotides of each other. Since switch regions extend for greater than one thousand base pairs, the close proximity of these breakpoints may be significant, suggesting that there may be hotspots for rearrangement within the switch regions.

Interpretation of these results has been difficult. The presence of repeats in switch regions has been interpreted to signify that class switching occurs by homologous recombination (Honjo et al., 1981). Alternatively, these repeats may mean that sequences which are recognized by specific class switching enzymes are reiterated to improve the efficiency of the rearrangement (Davis et al., 1980; Sakano et al,. 1980; Piccoli et al., 1984; Marcu et al., 1982). Various sequences have been found at or near rearrangement breakpoints and have been proposed to mediate the rearrangements (Davis et al., 1980; Sakano et al., 1980; Honjo et al., 1981; Marcu et al., 1982).

The short recognition sequences proposed by Marcu et al. (1982) and Honjo et al. (1980) are both heavily repeated within S regions. A section of all S regions consists of long stretches in which these simple sequences are tandemly repeated whereas other sections display an interspersed pattern of repeats. These repeats are reiterated to such an extent that it is not surprising that rearrangement breakpoints should, by chance, fall near one of them suggesting that their ubiquitous proximity to switch sites is not significant. Furthermore, the following two observations suggest that these sequences may not mediate rearrangement. First, the rearrangement breakpoints in the S_u, $S_{gamma2b}$ and S_{alpha} regions do not occur in those sections which contain the highest concentrations of the putative recognition sequences (Davis et al., 1980; Sakano et al., 1980; Kataoka et al., 1980). Second, in the S_u and S_{alpha} regions, the rearrangement breakpoints tend to cluster whereas the putative recognition sequences are dispersed. This observation is not consistent with the idea that all of the repeats are functionally capable of mediating class switching.

Class switching has been studied in the preB cell line 1881 (Siden et al., 1979). This cell line was established by transforming murine bone marrow cells with Abelson murine leukemia virus and was found to have the following properties. Immunoglobulin expression can be stimulated by a mitogen, LPS (Rice and Baltimore, 1983). Rearrangement of kappa light chain genes occurs spontaneously *in vitro* (Riley et al., 1981). Point mutations occur at a much higher frequency in the V_H gene segment than in non-immunoglobulin genes suggesting that the V_H genes are being modified by somatic hypermutation (Wabl et al., 1984; M. Wabl, pers. comm.). Finally, 1881 can switch expression from C_u to $C_{gamma2b}$. Up to 30% of a population can express IgG_{2b} and this heterogeneity can be maintained upon cloning. Southern blot analysis of cells expressing $C_{gamma2b}$ has revealed that a class switch rearrangement has occurred (Burrows et al., 1983).

We have attempted to help clarify the mechanism of class switching by taking advantage of the switching activity in 1881. In this report, S_u and $S_{gamma2b}$ regions were introduced into the 1881 cell line by DNA mediated gene transfer using a selectable marker, SV2GPT (Mulligan and Berg, 1981). The switch regions were placed on either side of the EcoGPT gene so that class switch rearrangement would delete the EcoGPT gene. When GPT^- revertants from GPT^+ transformants were selected, it was found, in four out of four cases, that the EcoGPT gene had been deleted but that the deletion endpoints were not in switch region DNA. Since class swtich rearrangements would be expected to result in deletion endpoints in switch regions, it appears that the EcoGPT gene was deleted randomly.

Results

Experimental protocol

In an attempt to further analyze the mechanism of immunoglobulin class switching, S_u and $S_{gamma2b}$ region DNAs were introduced into the preB cell line 1881, which continually switches from IgM to IgG_{2b} expression *in vitro*. If the endogenous switching activity in 1881 can recognize the newly introduced switch region sequences, then these sequences should recombine and thereby delete the region between them. As a result, the selectable marker EcoGPT, which was inserted between the switch regions, would be lost. Thus, loss of EcoGPT expression might indicate that the newly introduced switch regions had undergone class switch rearrangement.

The EcoGPT gene encodes xanthine guanine phospho-ribosyl transferase (XGPRT) and confers a GPT^+ phenotype (Mulligan and Berg, 1981). It can be used as a selectable marker in mammalian cells and it is possible to select cells which either possess or lack EcoGPT gene activity. This bacterial enzyme is used in the purine salvage pathway. In the absence of *de novo* purine biosynthesis, cellular survival depends on the salvage of hypoxanthine, xanthine or guanine from the medium. The bacterial enzyme XGPRT can utilize xanthine whereas the endogenous mammalian enzyme HGPRT cannot. Therefore, if xanthine is the only purine present in the medium, mammalian cells with EcoGPT activity can be selected by blocking *de novo* purine biosynthesis with mycophenolic acid (MPA). Conversely, since GPT^+ cells are sensitive to 6-thioxanthine or 8-azaxanthine, the lack of EcoGPT gene activity can be selected using these xanthine analogs (J. F. Nicholas, pers. comm.). Before class switching, transformed 1881 cells should be GPT^+ and can be selected in medium containing MPA and xanthine; after class switching, the cells should be GPT^- and can be selected in medium containing 6-thioxanthine (fig. 1A).

The immunoglobulin class switching vector, pugamma9, is shown in figure 1B. Some of its components are described below. First, the vector contains the S_u region on a 3.8 kb Xba I fragment. This fragment contains all known S_u sites. Second, an 0.8 kb Xba I to Eco RI fragment that contains most of the $S_{gamma2b}$ region was included. Only the switch site from the plasmacytoma MOPC41, which occurs outside of this fragment toward the 3' side, is missing (Sakano et al., 1980; Takahashi et al., 1980).

Third, the immunoglobulin enhancer sequence was inserted downstream of the EcoGPT gene since it is required for EcoGPT expression in B cells (Gillies et al., 1983; Queen and Baltimore, 1983; Banerji et al., 1980). Omission of the immunoglobulin enhancer from this vector reduced the frequency of GPT^+ transformants one hundred to one thousand fold (see below).

Fourth, the gamma2b repeat sequence was used as a control for homologous recombination. The gamma2b repeat region has many repeats which are also present, although to a lesser extent, in the $S_{gamma2b}$ region. Deletion of the EcoGPT gene via homologous recombination should have an upstream breakpoint in the gamma2b repeat, which is homologous to the $S_{gamma2b}$ region, rather than in the S_u region, which is less homologous. Therefore, switch rearrangements may be distinguished from homologous recombination since the former should have breakpoints in the S_u region whereas the latter should have endpoints in the gamma2b repeat region.

Fifth, the shuttle vector, pSVOD, allows the switch vector, integrated in the B cell genome, to be rapidly reisolated in *E. coli*. The vector includes pBR322-derived sequences as well as the SV40 origin of replication. The SV40 large T antigen can cause overreplication of SV40 sequences which are integrated in the chromosome. As a result, the SV40 sequences appear as extrachromosomal DNA (Botchan et al., 1979). Cos cells constitutively express SV40 large T antigen

(Gluzman, 1981). When these cells are fused to cells carrying an integrated copy of SV40, the integrated copy overreplicates and appears extrachromosomally (Botchan et al., 1980). Similarly, fusion of Cos cells with 1881 cells which have been transformed with pugamma9 should cause the switch vector to appear extrachromosomally. Extrachromosomal DNA can be easily separated from genomic DNA by a Hirt extraction (Hirt, 1967) and the purified DNA can be used to transform *E. coli*.

Transformation of 1881.1 with pugamma9

An HGPRT$^-$ mutant of 1881, 1881.1, was selected using 6-thioguanine. This cell line was transfected with 5 μg of pugamma9, 5 μg of B6beta2u (see below) and 90 μg of genomic 1881.1 carrier DNA as a calcium phosphate precipitate (Oi et al., 1983; Rice and Baltimore, 1983). Cells were plated in microtiter wells and GPT$^+$ cells were selected. Transformants appeared after two to three weeks at a frequency of 5×10^{-6} clones/μg vector/cell. Four clones were arbitrarily chosen for further analysis: F1, J1, I1 and I2.

The B6beta2u DNA contains the genomic beta2 microglobulin gene from the C57BL/6 mouse strain (Parnes and Seidman, 1982) and was included in the transformation as a second marker. Beta2 microglobulin produced by C57BL/6 and Balb/c mice can be distinguished by monoclonal antibodies. Since 1881.1 is of Balb/c origin, expression of the C57BL/6 beta2 microglobulin gene could be observed in the transformants. Rearrangements induced by class switching should not alter expression of this second marker. However, random deletion of EcoGPT may include the second marker as well. Thus, the C57BL/6 beta2 microglobulin gene was employed to distinguish class switch events from random deletions. Unfortuneately, none of the transformants expressed the C57BL/6 beta2 microglobulin gene as evidenced using monoclonal antibodies in a radioimmune

assay (data not shown). As discussed below, it is likely that co-transformation of 1881 cells is inefficient and accordingly, cells selected for GPT expression rarely have B6beta2u expression.

Southern blotting experiments were performed to prove that the GPT^+ transformants contained copies of the pugamma9 switch vector. Genomic DNA from 1881.1 and the four transformants was cut with Eco RI and analyzed using pSV2GPT DNA as a probe (fig. 2A). No pSV2GPT sequences were present in 1881.1 DNA whereas F1, J1, I1 and I2 DNAs all exhibited one or more bands. Thus, these cell lines contain copies of the switch vector.

Next, $pS_{gamma2b}$ was used as a probe in an experiment similar to the one described above in order to determine the approximate copy number of the vector in the transformants (fig. 2B). The lane with 1881.1 DNA shows bands at 14, 12 and 8.5 kb, which are probably due to rearrangement of the endogenous $S_{gamma2b}$ region, as well as the germline $S_{gamma2b}$ and $S_{gamma2a}$ bands at 6.7 kb. Lanes with DNA from the four transformants all show one or more bands in addition to the bands seen in 1881.1 DNA. The extra bands at 3.5 kb and at 0.8 kb are probably due to the presence of the switch vector since pugamma9 has plasmid and $S_{gamma2b}$ sequences on these sized fragments. These bands are similar in intensity to the bands at 6.7 kb, due to the endogenous $S_{gamma2b}$ and $S_{gamma2a}$ regions. Therefore, only one or a few copies of the switch vector are likely to have integrated in the transformants.

Analysis of the transformants

Different cell clones may have different class switching activities (Burrows et al., 1982) and the particular conformation of pugamma9 in each transformant may affect its ability to rearrange. Accordingly, the transformants may rearrange the switch vector at different rates. First, the number of cells positive for IgG was

measured in all four transformants and revealed that all had between 20% - 40% IgG^+ cells. This observation suggests that all four transformants retained approximately equal levels of class switching activity. Next, the frequency of reversion from GPT^+ to GPT^- was measured to see which transformants might rearrange pugamma9 frequently. Since these cells are $HGPRT^-$, cells which are also $XGPRT^-$ should be resistant to 6-thioguanine. The transformants were grown in non-selective medium for seven days and then plated at various dilutions into microtiter wells in medium containing 6-thioguanine. Colonies appeared after seven days and were grown as revertant clones. The reversion frequencies varied from approximately 10^{-3} to 10^{-5} (table 1). The cell line J1 had a high reversion frequency, 1.1×10^{-3}, and was therefore analyzed further.

The genomic arrangement of the switch vector in J1 was studied by reisolating integrated copies as molecular clones. A library was constructed in the lambda insertion vector embl 3 with DNA partially digested with the restriction enzyme Sau 3A. If the insert in the lambda vector contains the pBR322 origin of replication and beta-lactamase gene, then it can replicate in *E. coli* lysogenic for lambda and confer ampicillin resistance. Accordingly, clones containing functional pSVOD sequences were isolated as phasmids by infecting a lambda lysogen, SS13, and selecting with ampicillin. Those clones that also contained SV2GPT sequences were identified by hybridization to an SV2GPT probe. Two copies of SV2GPT were found integrated in the genome of J1. Many clones of each copy of SV2GPT were isolated; lambdaJ1-4 and lambdaJ1-2 are representatives of each of the copies (fig. 3).

The copy of SV2GPT contained in lambdaJ1-2 is probably not expressed. Restriction map analysis of lambda J1-2 indicated that flanking DNA derived from the host cell's genome joined the switch vector DNA within a 3.9 kb Hind III fragment which contains the EcoGPT gene. It is possible that the EcoGPT

transcription unit has been truncated. Even if this is not so, the breakpoint must certainly delete the immunoglobulin enhancer which was located downstream of the 3.9 kb Hind III fragment and is completely missing. Since the immunoglobulin enhancer is required for expression, this gene is probably inactive.

The copy of pugamma9 contained in the clone lambdaJ1-4 is probably expressed since both the EcoGPT gene and the immunoglobulin enhancer are intact. In addition, the upstream S_u and downstream $S_{gamma2b}$ regions are present. Thus, this cell line is a reasonable candidate to determine whether class switching can rearrange newly introduced switch regions and thereby delete the EcoGPT gene.

Analysis of the revertants

Five GPT$^-$ revertant clones of J1 were isolated using 6-thioguanine: J1.2, J1.5, J1.6, J1.7 and J1.8. The context of the pugamma9 vector in these clones was analyzed by genomic blotting and molecular cloning to see if a class switch rearrangement had occurred. Genomic DNAs were isolated, cut with Hind III, and analyzed by Southern blotting using SV2GPT as a probe. As shown in figure 4A, the parental J1 cell line contains a 3.7 kb band, containing the non-expressed EcoGPT gene, and a 3.9 kb band, containing the expressed gene. All five clones have bands at 3.7 kb indicating that the non-expressed copy of EcoGPT has been retained. Three clones, J1.5, J1.6 and J1.7, have deleted the band at 3.9 kb indicating that the expressed EcoGPT gene has been deleted. Two clones, J1.2 and J1.8, are missing the 3.9 kb band but contain an extra band at 9.5 and 11 kb, repectively. In these cases, it appears that a rearrangement has occurred within the 3.9 kb Hind III fragment which contains the expressed EcoGPT gene.

These revertants were further analyzed using genomic DNA that was cut with Eco R1, blotted and hybridized to a pSVOD probe. The parental J1 DNA has a

band at 3.5 kb, which is closely linked to the expressed EcoGPT gene, and a band at 3.3 kb, which is near the non-expressed EcoGPT gene. In all clones analyzed (J1.2, J1.5, J1.6 and J1.7), the 3.5 kb band is deleted (fig. 4B). In clones J1.5, J1.6 and J1.7, class switch rearrangement would be expected to have a breakpoint in the $S_{gamma2b}$ region resulting in a deleted EcoGPT gene and an unrearranged pSVOD fragment. Since both sequences are deleted, the deletion probably does not have a breakpoint in the $S_{gamma2b}$ region and therefore, is not due to class switching. In clones J1.2 and J1.8, a deletion downstream of the EcoGPT gene has removed the pSVOD sequences and one of the deletion endpoints occurs within or near the gene.

Next, EcoGPT expression in the revertant clones was analyzed by Northern blot analysis. Poly A^+ RNA was isolated, blotted and hybridized to an SV2GPT probe. As expected, in lanes containing J1.5, J1.6 and J1.7 RNA, which have deleted EcoGPT genes, no hybridization was observed (fig. 5). However, in lanes containing J1.2 and J1.8 RNAs, hybridization was observed. In the lane containing RNA from J1.2, the band was much less in intensity than in the lane containing J1 RNA but was of the same size. In J1.8, the intensity was similar but the RNA was different in size.

The ability to restore EcoGPT activity in the revertants was tested. After seven days in nonselective media, $3x10^6$ cells from each clone were placed in microtiter wells in medium containing MPA and xanthine. No GPT^+ cells were obtained from the J1.5, J1.6 and J1.7 clones consistent with the evidence suggesting the the EcoGPT gene had been deleted. The J1.2 and J1.8 cells produced GPT^+ clones at an approximate frequency of 10^{-3} and 10^{-4} respectively.

In summary, genomic Southern blotting, Northern blotting and revertant analyses suggest that the J1.5, J1.6 and J1.7 revertants have a large deletion including the EcoGPT gene and pSVOD sequences. In these cells, EcoGPT RNA is not expressed and GPT^+ cells are never observed. Since the deletion probably does

not have a breakpoint in the $S_{gamma2b}$ region, it is likely that the rearrangement is not due to class switching. The J1.2 and J1.8 clones have a deletion including the pSVOD sequences and ending near or within the EcoGPT gene. In J1.2, much less EcoGPT RNA is expressed whereas in J1.8, the RNA is different in size. In both these cell clones, it is possible to select cells in which the EcoGPT gene has been reactivated. It is possible that a class switch rearrangement, with one endpoint in the $S_{gamma2b}$ region and the other endpoint in downstream sequences, could have caused these deletions.

In order to determine whether the deletion in J1.8 had a breakpoint in the $S_{gamma2b}$ region, molecular clones of the rearranged EcoGPT gene were isolated. Genomic DNA from J1.8 was cut with Hind III and fragments 11 kb long were cloned in the lambda vector lambdaNM762. A map of the rearranged fragment is shown in figure 6. When the rearranged EcoGPT gene from J1.8 is compared to the unrearranged gene from J1, it can be seen that the rearrangement has most likely occurred at the 3' end of the gene between the Bam HI and Eco R1 sites. The Bam HI site marks the end of the EcoGPT transcription unit and the Eco R1 site marks the beginning of the immunoglobulin enhancer. Since the rearrangement breakpoint is probably downstream of the EcoGPT transcription unit, it is likely that loss of gene activity was caused by deleting the immunoglobulin enhancer. Interestingly, this mutation alters the size of the EcoGPT RNA and not the amount (fig. 5). Since the breakpoint must be in SV40 sequences derived from SV2GPT, it appears unlikely that this rearrangement is due to class switching. Rather, it is more likely that random deletions, as in J1.5, J1.6 and J1.7, have occurred.

Discussion

Rearrangement of pugamma9 sequences in 1881.1 cells

In this report, a cell line with endogenous class switching activity, 1881.1, was transformed with the immunoglobulin class switching vector pugamma9. Molecular cloning and genomic Southern blotting analyses of one transformant, J1, indicated that one expressed copy of EcoGPT had been integrated and that S_u and $S_{gamma2b}$ regions were located immediately upstream and downstream of this gene respectively. Transformants which had reverted to GPT^- were selected and two types of mutation were observed. In the first type, as exemplified by the J1.5, J1.6 and J1.7 revertant cell lines, a large deletion removed the EcoGPT gene and adjoining pBR322 sequences. This observation indicates that the endpoints of this deletion cannot be in the $S_{gamma2b}$ region and therefore, this deletion is probably not caused by a class switch rearrangement. In the second type, seen in the J1.2 and J1.8 cell lines, a deletion including the pSVOD sequences but ending within or near the EcoGPT gene was observed. The deletion endpoint near the EcoGPT gene in J1.8 was analyzed further. Restriction map analysis of a molecular clone containing this deletion endpoint revealed that the breakpoint occurred in SV40 sequences derived from the 3' end of the EcoGPT gene and not in immunoglobulin enhancer or switch region sequences. Again, this result implies that the deletion was not due to immunoglobulin class switching.

An underlying assumption in this report is that switch regions, which contain the rearrangement breakpoints, also contain the sequences which mediate the rearrangements. It is possible that the rearrangement breakpoint may lie at a distance from the sequences that promote rearrangement. Precedent for such a recognition mechanism include endonucleolytic cleavage by type I restriction endonucleases and recombination mediated by chi sequences in lambda phage (Stahl,

1979). In both cases, recognition of a specific sequence results in a rearrangement oftentimes many kilobases away. If this mechanism were true for class switching, then the recognition sequences could be separate from the rearrangement breakpoints. The switch regions in pugamma9 might not contain the sequences necessary to promote class switching. Even if it did contain these sequences, class switching could not be distinguished from random deletion since both rearrangements could have breakpoints in flanking DNA.

In this experiment, a second marker, the C57Bl/6 beta2 microglobulin gene, was cotransfected with pugamma9 in an effort to distinguish random deletions from specific class switch rearrangements. If the beta2 microglobulin gene integrated near pugamma9, then it should be deleted by large deletions but not by class switch rearrangements. This approach failed for two reasons. First, none of the four transformants expressed the beta2 microglobulin gene. In a similar experiment, the TK gene was cotransfected with pugamma9 and, as before, GPT^+ cells did not express the second marker. Specifically, a TK^- derivative of 1881.1, A1 (Burrows et al., 1983), was transformed with 5 μg pugamma9, 5 μg pTK and 90 μg carrier DNA. Out of fifty GPT^+ clones tested, none were TK^+. The lack of TK^+ transformants was not due to a failure of the TK promoter to be expressed in B cells. Essentially the same results were obtained when a recombinant gene was used in which TK coding sequences were expressed from a metallothionein promoter. It is unknown why the beta2 microglobulin, the pTK and the pmetTK genes are not expressed in the GPT^+ transformants. Perhaps cotransfer of unlinked genes in B cell transformation is inefficient or only a small fraction of integrated genes are expressed.

Second, even if a second marker were to be cotransfected with pugamma9, many random deletions could not be distinguished from class switch rearrangement. In order to distinguish these two types of rearrangement, most

random deletions would have to be large so that the second marker would usually be lost. There are many cases in this experiment which show that this is not so. For example, two out of five random deletions had breakpoints near or within the EcoGPT gene. Most likely, these deletions are small rather than large since deletion endpoints are easily found. Also, in these two cases, if the second marker were to be located upstream of the EcoGPT gene, it would not have been affected by the rearrangement. In the transformant J1, a non-functional copy of pugamma9 was integrated and this copy was absent in only one of the five revertant clones of J1. This result shows that other sequences which are integrated along with the expressed copy of pugamma9 are not usually absent when pugamma9 is removed by deletion. In summary, these results suggest that random deletions which remove the EcoGPT gene may be small and will not dependably remove a second, linked marker.

Future studies based on the preliminary work presented here may benefit from analyzing other transformants or by isolating cells which have recently undergone class switching. Only one transformant, J1, was studied in detail; other transformed clones may rearrange the pugamma9 vector by class switching. The EcoGPT reversion frequency for clone I2 is high ($1x10^{-3}$) and would thus be a good candidate for further analysis.

The high frequency of IgG^+ cells in a population of 1881 cells may possibly be the result of a high rate of class switching. Alternatively, this frequency may be due to selective proliferation of IgG^+ cells or to their gradual accumulation in the population over time. Thus, it may not be valid to assume that the 1881 cell line has a high rate of IgM to IgG_{2b} class switching only because many of its cells express IgG_{2b}. Even if the rate of class switching in 1881 cells is low, it should be possible to select cells which have recently switched from IgM to IgG expression. From a population of cells which initially expressed IgM exclusively, IgG^+ cells

could be selected. In these cells, it is possible that the pugamma9 vector might rearrange more readily.

Experimental techniques

Some of the techniques employed in this report are novel and merit discussion. The immunoglobulin enhancer region was needed for expression of the EcoGPT gene in the preB cell 1881. The vector pugamma8 differs from pugamma9 in that it lacks the immunoglobulin enhancer. Using 5 μg pugamma9 and 95 ug carrier DNA, a transformation efficiency of $5x10^{-6}$ clones/μg pugamma9/cell was routinely observed. Under analagous conditions, pugamma8 transformed 1881.1 at a frequency of less than $2x10^{-10}$ clones/ug pugamma8/cell. This difference in transformation frequency is probably due to the inability of EcoGPT to be expressed without an enhancer in B cells. If pugamma8 is not diluted with carrier DNA during transformation, a transformation frequency of approximately $1x10^{-9}$ clones/ug pugamma8/cell is observed. Southern blot analysis of these transformants revealed that up to fifty copies of pugamma8 had been integrated. It appears that fifty SV40 promoters lacking immunoglobulin enhancers, each inefficiently expressed, could collectively express enough EcoGPT to pass selection. The reduction in transformation frequency when carrier DNA is used may reflect the inability of pugamma8 to integrate in high copy numbers when heavily diluted with carrier DNA. These results are consistent with other studies which suggest that an immunoglobulin enhancer is required for some promoters to function in B cells (Banerji et al., 1983; Queen and Baltimore, 1983; Gillies et al., 1983).

Further support for the need of an immunoglobulin enhancer comes from analysis of the J1.8 revertant. In J1.8, a rearrangement has probably occurred downstream of the EcoGPT gene thereby deleting the immunoglobulin enhancer.

Thus, it appears that an expressed EcoGPT gene (i.e., in J1) became dormant when the immunoglobulin enhancer was deleted (i.e., in J1.8). An important caveat to this conclusion is that analysis was performed only by restriction mapping. The integrity of the EcoGPT gene was not proven by sequence data.

The published selection medium was modified (Mulligan and Berg, 1981). The concentration of MPA was decreased from 25 μg/ml to 1 μg/ml. Even at this lower concentration, most non-transformed B cells died within 48 hr after the addition of MPA. Also, hypoxanthine or adenine were added to the selection medium since it was found that GPT^+ transformants required either of these purines for growth. Presumably, MPA inhibits some step in *de novo* purine synthesis. Typically, adenine rather than hypoxanthine was added because hypoxanthine can competetively inhibit the utilisation of xanthine by XGPRT.

A method was developed for selecting cells which do not express the bacterial enzyme XGPRT but may express the mammalian enzyme HGPRT. The bacterial but not the mammalian enzyme is capable of utilizing xanthine and thus the xanthine analogs 6-thioxanthine and 8-azaxanthine. These analogs become toxic when XGPRT, but not HGPRT, is present and allow one to select mammalian cells which have lost EcoGPT expression. The ability to select for mammalian cells which either have or lack a marker can be useful in many cases, including the one presented here.

However, caveats to the use of the selection scheme discussed above, shown in the evidence presented below, are that not all cells may be sensitive to xanthine analogs and relatively high levels of XGPRT expression may be required for sensitivity. In 1881.1, GPT^+ cells are sensitive to 6-thioxanthine, 8-azaxanthine or 6-thioguanine. However, twenty times more 6-thioxanthine or 8-azaxanthine (100 μg/ml) than 6-thioguanine (5 μg/ml) are required for cell death suggesting that the xanthine analogs are less toxic than the guanine analog.

Cells which express low amounts of XGPRT can grow in the presence of MPA as well as 6-thioxanthine or 8-azaxanthine. A gene consisting of the metallothionein promoter, intron and polyadenylation site as well as the XGPRT coding region was constructed. This gene, metGPT, can transform 1881.1 cells at a frequency ($2x10^{-6}$ clones/μg pmetGPT/cell) which is only slightly lower than that of pugamma9 showing that the metallothionein promoter does not require the immunoglobulin enhancer for expression in the preB cell 1881. The pmetGPT transformants appear after three weeks and subsequently grow more slowly than pugamma9 transformants suggesting that the pmetGPT gene is expressed less efficiently than the EcoGPT-immunoglobulin enhancer gene. Cells that are transformed with pmetGPT are not sensitive to either 6-thioxanthine or 8-azaxanthine (data not shown). It appears that a low level of expression of XGPRT, such as that produced by pmetGPT, allows resistance to both MPA, thus giving a GPT^+ phenotype, as well as xanthine analogs, thus giving a "GPT^-" phenotype.

Some cells are impervious to 6-thioxanthine or 8-azaxanthine regardless of whether they express XGPRT. L cells were transformed with pugamma8. The SV40 promoter should transcribe efficiently in a mouse fibroblast and yet, these cells were resistant to 6-thioxanthine or 8-azaxanthine.

Materials and Methods

Cell culture

The 1881 and A1 cell lines were a gift from Dr. D. Baltimore and Dr. M. Wabl respectively. The 1881 cell line was carried in RPMI 1640, 10% fetal calf serum and 50 μM betamercaptoethanol (RBF_{10} medium). For freezing, cells were pelleted and resuspended in freezing medium (Dulbecco's modified Eagle's meduim, 10% calf serum, 10% dimethylsulfoxide). The cells were then sequentially placed

for 5 minutes at 4°C, .5 to 2 hrs at -20°C, overnight at -70°C and stored permanently in liquid nitrogen.

Selection of mutants

Cells lacking HGPRT were selected using 6-thioguanine. Cells in log phase growth (1-2x10^6 cells/ml) were pelleted and resuspended at 10^6 cells/well in 24 well microtiter dishes in medium containing 5μg/ml 6-thioguanine. Half of the medium was replaced every third day. After two weeks, 13 clones appeared from 5.5x10^7 cells (2.3x10^{-7}) and were designated 1881.1 to 1881.13. The frequency of reversion to $HGPRT^+$ was tested by selecting in HAT media (15 μg/ml hypoxanthine, 1 μg/ml aminopterin, 5 μg/ml thymidine). Cells were grown in neutral media for seven days and then plated in 24 well dishes. Clones appeared after two weeks: 1881.1 produced 11 colonies from 10^8 cells (1.1x10^{-7}); 1881.3 produced 7 colonies from 5x10^7 cells (1.4x10^{-7}); 1881.4 produced 2 colonies from 2x10^8 cells (1x10^{-8}). Other $HGPRT^-$ cell lines were not tested. Finally, GPT^- revertants can be selected from GPT^+ transformants by resistance to 6-thioxanthine (100 μg/ml) or 8-azaxanthine (100 μg/ml). If the cell is also $HGPRT^-$, 6-thioguanine (5 μg/ml) can also be employed.

Transformation

Cells were transformed by the procedure of Graham and van der Eb (1973) as modified by Oi et al. (1983). Vigorous cell growth was important for high transformation efficiency. Cells in early log phase growth (1-2x10^6 cells/ml, 4x10^8 cells total) were placed in centrifuge tubes with a 1 ml underlay of calf serum. The cells were pelleted (100 rpm x 10 minutes), resuspended in 50 ml of warm Dulbecco's modified Eagle's medium, and pelleted onto a 1 ml cushion of calf serum as before. The washed cells were resuspended in a DNA calcium phosphate mixture

that was prepared as follows. To 2.5 ml of a hepes buffered solution (50 mM hepes, 280 mM sodium chloride, pH 7.10) was added 50 µl of a 70 mM solution of sodium phosphate pH 6.8. In a separate tube, 300 µl of sterile 2M calcium chloride was added to 100 µg of DNA (e.g. 5 µg pugamma9, 5 µg B6beta2u, and 90 µg genomic 1881.1 carrier DNA) dissolved in 2.2 ml sterile water. The first solution was added dropwise to the second solution while bubbling. After thirty minutes at room temperature, the DNA had formed a fine white precipitate and was added to the cells directly. The resuspended cells were incubated at 37°C for 30 minutes, diluted with 100 ml of 10% fetal calf serum in Dulbecco's modified Eagle's medium and incubated at 37°C for 24 hours. Next, the cells were pelleted through a 1 ml cushion of calf serum and plated in RBF_{10} at $5x10^5$ cells/well in 24 well dishes. After one day, the medium was replaced with media containing 1 µg/ml MPA, 250 µg/ml xanthine and 15 µg/ml adenine. Half of the medium was replaced every 3 days. Colonies appeared after two to three weeks.

Radioimmune assay

The procedure of Pierres et al. (1981) was followed. Cells were pelleted and resuspended at 10^7 cells/ml in dilution buffer (0.2% fetal calf serum, .02% sodium azide in PBS). Monoclonal anti-C57Bl/6 beta2 microglobulin (NEN) was diluted 1:100 into dilution buffer and spun in an eppendorf centrifuge for 15 minutes to pellet insoluble matter. In a 96 well microtiter dish were mixed 20 µ of cells ($2x10^5$ cells) and 50 µl of diluted antibody. The cells were incubated for 1 hr at 4°C and pelleted (2400 rpm x 3 minutes, Beckman J6 centrifuge). The pellet was washed three times in dilution buffer by resuspending the pellet in 100 ul of dilution buffer and pelleting as before. The washed pellet was resuspended in 50 ul of dilution buffer containing 10^5 cpm of ^{125}I-protein A (NEN). The cells were washed three times in dilution buffer as before and then counted. Spleen cells from C57Bl/6 mice were used as positive controls; 1881 cells were used as negative controls.

Molecular cloning

Normal procedures were done in normal ways. Genomic DNA from J1 was partially digested with Sau3A and cloned into Eco R1 and Bam HI cut Embl3 DNA. After packaging, $3x10^6$ independent recombinant phage were obtained. Clones containing a pBR322 origin of replication and beta-lactamase gene were isolated in one step as phasmids. Approximately 10^7 pfu from separate aliquots of the amplified library was used to infect *E. coli* strain SS13, which is lysogenic for phage lambda. Colonies from ampicillin plates were picked, streaked and individual colonies were grown to stationary phase in L broth supplemented with 50 ug/ml ampicillin. A portion of the culture was lysed by the addition of a small amount of chloroform. The phage titer of the supernatant was typically 10^8 pfu/ml. Subsequent analysis of the phage indicated that all contained switch vector sequences. Clones lacking functional pBR322 sequences were obtained by screening nitrocellulose filters with SV2GPT as a probe. Genomic DNA from the revertant J1.8 was completely cut with Hind III and 11 kb fragments were isolated from a low melting temperature agarose gel. The fragments were cloned in lambdaNM762 and screened using SV2GPT as a probe.

References

Banerji, J., Olson, L. and Schaffner, W. 1983. A lymphocyte-specific enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33:** 729-740.

Botchan, M., Topp, W. and Sambrook, J. 1979. Studies on SV40 excision from cellular chromosomes. Cold Spring Harbor Symp. Quant. Biol. **43:** 709-719.

Botchan, M., Stringer, J., Mitchison, T. and Sambrook, J. 1980. Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell **20:** 143-152.

Burrows, P., Beck-Engeser, G. and Wabl, M. 1983. Immunoglobulin heavy-chain class switching in a pre-B cell line is accompanied by a DNA rearrangement. Nature **306:** 243-246.

Cooper, M., Kearney, J., Lydyard, P., Grossi, C. and Lawton, A. 1977. Studies of generation of B-cell diversity in mouse, man and chicken. Cold Spring Harbor Symp. Quant. Biol. **41:** 139-145.

Davis, M., Kim, S. and Hood, L., 1980. DNA sequences mediating class switching in alpha immunoglobulins. Science **209:** 1360-1365.

Davis, M., Calame, K., Early, P., Livant, D., Joho, R., Weissman, I. and Hood, L. 1979. Immunoglobulin heavy chain gene is formed by two recombinational events. Nature **283:** 733-739.

Dunnick, W., Rabbitts, T. and Milstein, C. 1980. An immunoglobulin deletion mutant with implications for the heavy chain switch and RNA splicing. Nature **286:** 669-675.

Gillies, S., Morrison, S., Oi, V. and Tonegawa, S. 1983. A tissue specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33:** 717-728.

Gluzman, Y. 1981. SV40 transformed simian cells support the replication of early SV40 mutants. Cell **23:** 175-182.

Graham, F. and Van der Eb, A. 1973. J. Virol. **52:** 455-456.

Hirt, D. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:** 365-369.

Honjo. T., Nakai, S., Nishida, Y., Kataoka, T., Yamawaki-Kataoka, Y., Takahashi, N., Obata, M., Shimuzu, A., Yaoita, Y., Nikaido, T. and Ishida, N. 1981. Rearrangement of immunoglobulin genes during differentiation and evolution. Immunol. Rev. **59:** 33-67.

Kataoka, T., Kawakami, N., Takahashi, N. and Honjo, T. 1980. Rearrangement of immunoglobulin gamma-1 chain gene and mechanism for heavy chain class switch. Proc. Nat. Acad. Sci. U.S.A. **77:** 919-923.

Kataoka, T., Miyata, T. and Honjo, T. 1981. Repetitive sequences in class switch recombination regions of immunoglobulin heavy chain gene. Cell **23:** 357-368.

Marcu, K., Lang, R., Stanton, L, and Harris, L. 1982. A model for the molecular requirements of immunoglobulin heavy chain class switching. Nature **298:** 87-89.

Mulligan, R. and Berg. P. 1981. Selection for animal cells that express the *E. coli* gene coding for xanthine guanine phosphoribosyl transferase. Proc. Nat. Acad. Sci. U.S.A. **78:** 2072-2076.

Murray, N., Brammar, W. and Murray, K. 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **150:** 53.

Nikaido, T., Nakai, S. and Honjo, T. 1981. Switch region of immunoglobulin C_u gene is composed of simple tandem repetitive sequences. Nature **292:** 845-848.

Obata, M., Kataoka, T., Nakai, S., Yamagishi, H., Takahashi, N., Yamawaki-Kataoka, Y., Nikaido, T., Shimuzu, A. and Honjo, T. 1981. Structure of a rearranged gamma-1 gene and its implication to immunoglobulin class switch mechanism. Proc. Nat. Acad. Sci. U.S.A. **78:** 2437-2441.

Oi, V., Morrison, S., Herzenberg, L. and Berg, P. 1983. Immunoglobulin gene expression in transformed lymphoid cells. Proc. Nat. Acad. Sci. U.S.A. **80:** 825-829.

Parnes, J. and Seidman, J. 1982. Structure of wild type and mutant mouse beta 2-microglobulin genes. Cell **29:** 661-669.

Pernis, B., Forni, L. and Luzzati, K. 1977. Synthesis of multiple immunoglobulin classes by single lymphocytes. Cold Spring Harbor Symp. Quant. Biol. **41:** 75-83.

Piccoli, S., Caimi, P. and Cole, M. 1984. A conserved sequence at c-myc oncogene chromosomal translocation breakpoints in plasmacytomas. Nature **310:** 327-330.

Pierres, M., Devaux, C., Dosseto, M. and Marchetto, S. 1981. Clonal analysis of B and T cell reponses to I_a antigens. Immunogenetics. **14:** 481-485.

Queen, C. and Baltimore, D. 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell. **33:** 741-748.

Rice, D. and Baltimore, D. 1983. Regulated expression of an immunoglobulin kappa gene introduced into a mouse lymphoid cell line. Proc. Nat. Acad. Sci. U.S.A. **79:** 7862-7865.

Riley, S., Brock, E. and Kuehl, W. 1981. Induction of light chain expression in a pre-B cell line by fusion to myeloma cells. Nature **289:** 804-806.

Sakano, H., Maki, A., Kurosawa, Y., Roeder, W. and Tonegawa, S. 1980. The two types of somatic recombination necessary for generation of complete immunoglobulin heavy chain genes. Nature **286:** 676-683.

Siden, E., Baltimore, D., Clark, D. and Rosenberg, N. 1979. Immunoglobulin synthesis by lymphoid cells transformed *in vitro* by Abelson murine leukemia virus. Cell. **16:** 389-396.

Stahl, F. 1979. Special sites in generalized recombination. Ann. Rev. Genet. **13:** 7.

Takahashi, N., Kataoka, T. and Honjo, T. 1980. Nucleotide sequence of class-switch recombination region of the mouse immunoglobulin gamma2b chain gene. Gene **11:** 117-127.

Wabl, M., Beck-Engeser, G. and Burrows, P. 1984. Allelic inclusion in the preB cell line 18-81. Proc. Nat. Acad. Sci. U.S.A. **81:** 867-870.

Table 1

Frequency of reversion of GPT^+ transformants

Cell line	clones/cells	frequency
F1	$13/10^6$	1.5×10^{-5}
J1	$11/10^4$	1.1×10^{-3}
I1	$31/10^5$	3.1×10^{-4}
I2	$19/10^4$	1.9×10^{-3}

Figure Legends

1A. Experimental protocol. left: Immunoglobulin class switching normally rearranges S_u and $S_{gamma2b}$ regions thereby deleting the C_u gene segment and switching from IgM to IgG_{2b} expression. right: Rearrangement of the switching vector will delete the SV2GPT gene and switch from GPT^+ to GPT^-.

1B. Restriction map of pugamma9. The germline C_u gene segment, the germline $C_{gamma2b}$ gene segment and the pSVOD vector, are shown with the immunoglobulin class switching vector pugamma9. R = Eco R1, H = Hind III, B = Bam H1, Sl = Sal I, S = Sac I, O = Bgl II, X = Xba I.

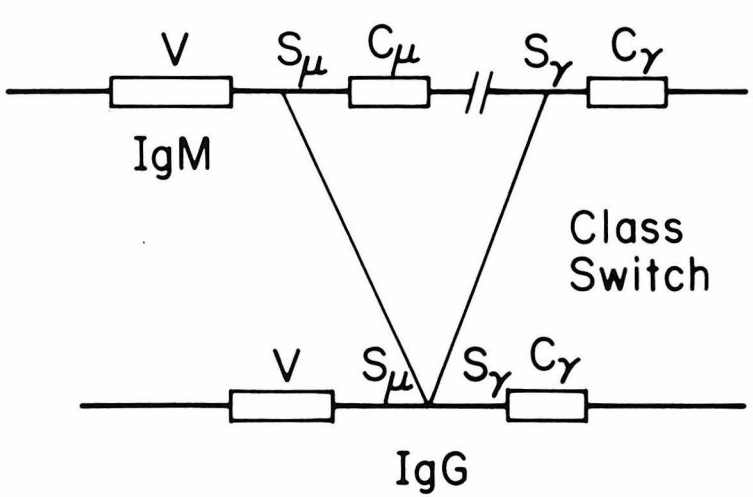
2. Southern blotting analysis of GPT^+ transformants. 20 μg of genomic DNA was cut with Eco R1, electrophoresed on 0.8 agarose gels and blotted onto nitrocellulose. The filters were hybridized to probe (10 ng/ml) in formamide hybridization buffer. Location of bands (in kb) of size markers appear on the left. A. The pSV2GPT plasmid was used as a probe. B. A 3.8 kb Eco R1 fragment, containing the gamma2b repeat and the $S_{gamma2b}$ region was used as probe. The bands due to the endogenous $S_{gamma2b}$ region, the vector-derived $S_{gamma2b}$ region and the pSVOD vector are indicated on the right. Although not seen here, in other experiments the F1 and J1 cell lines showed an Eco R1 band at 3.5 kb when hybridized to pSVOD.

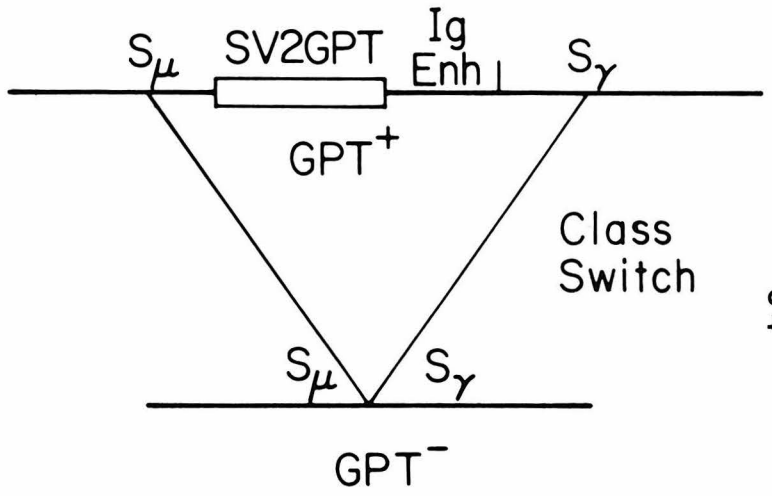
3. Restriction map of lambdaJ1-4 and lambdaJ1-2. Only the inserts are shown. See figure 1B for restriction site symbols.

4. Southern blot analysis of the revertants. Southern blots were performed as described in figure 2. Source of DNA appears at the top. Restriction enzyme and probe appear at the bottom.

5. Northern blot analysis of the revertants. 5 μg/lane of poly A^+ RNA was electrophoreses on 1% agarose, 5.5% formaldehyde gels, blotted onto nitrocellulose and hybridized to SV2GPT probe. Source of RNA appears at the top.

6. Restriction map of lambdaJ1.8-2. Genomic DNA from J1.8 was cut with Hind III and 11 kb fragments were purified from a low melting temperature agarose gel. These fragments were cloned in lambdaNM762 and screened using SV2GPT as probe. The partial restriction map of the insert in lambdaJ1.8-2 is shown.


V
S_μ
C_μ
S_γ
C_γ
IgM
Class
Switch
V
S_μ
S_γ
C_γ
IgG


S_μ
SV2GPT
Ig
Enh
S_γ
GPT^+
Class
Switch
S_μ
S_γ
GPT^-

64

fig. 1A

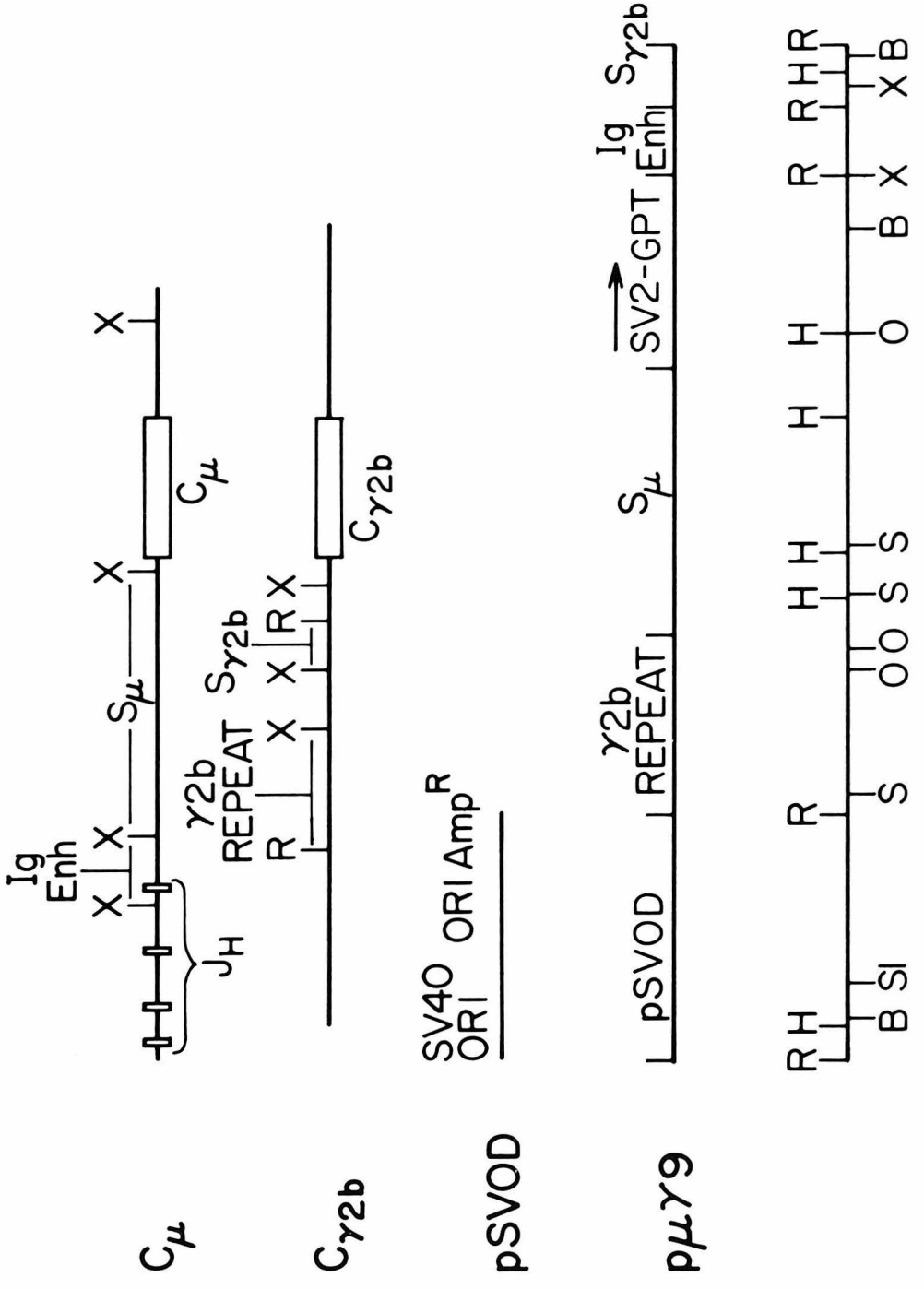

Cμ
Ig Enh
X
X
JH
Sμ
X
Cμ
γ2b REPEAT
R
X
X
Sγ2b
R X
Cγ2b
Cγ2b
SV40 ORI
ORI
Amp R
pSVOD
pSVOD
γ2b REPEAT
Sμ
SV2-GPT
Ig Enh
Sγ2b
pμγ9
R H
B Sl
R
S
H H
O O S S
H H
O
R R H R
B X
R X
X B
1kb

fig. 1B

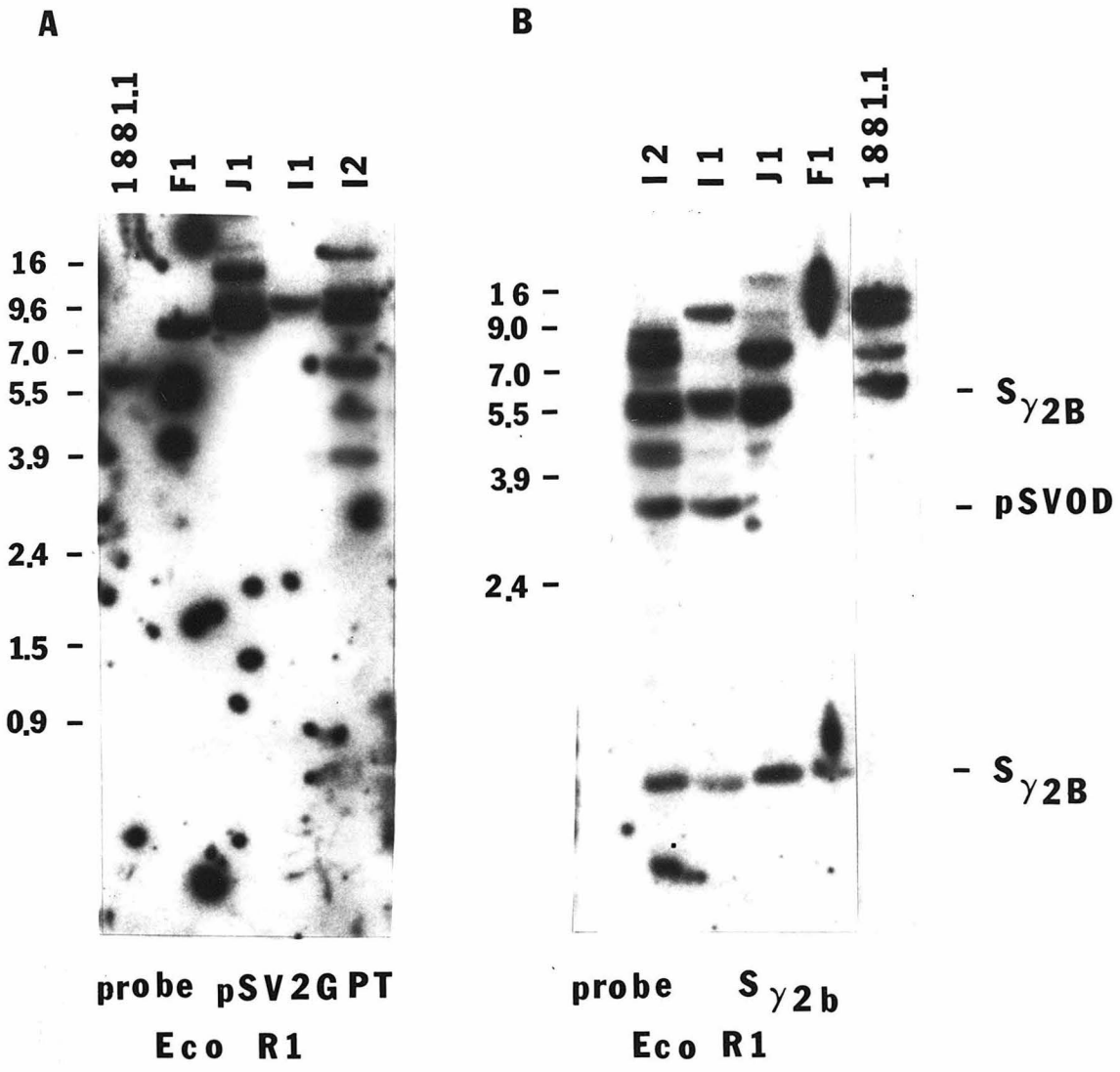

A
B
1881.1
F1
J1
I1
I2
16
9.6
7.0
5.5
3.9
2.4
1.5
0.9
9.0
- Sγ2B
- pSVOD
- Sγ2B
probe pSV2GPT
Eco R1
probe Sγ2b
Eco R1

fig. 2

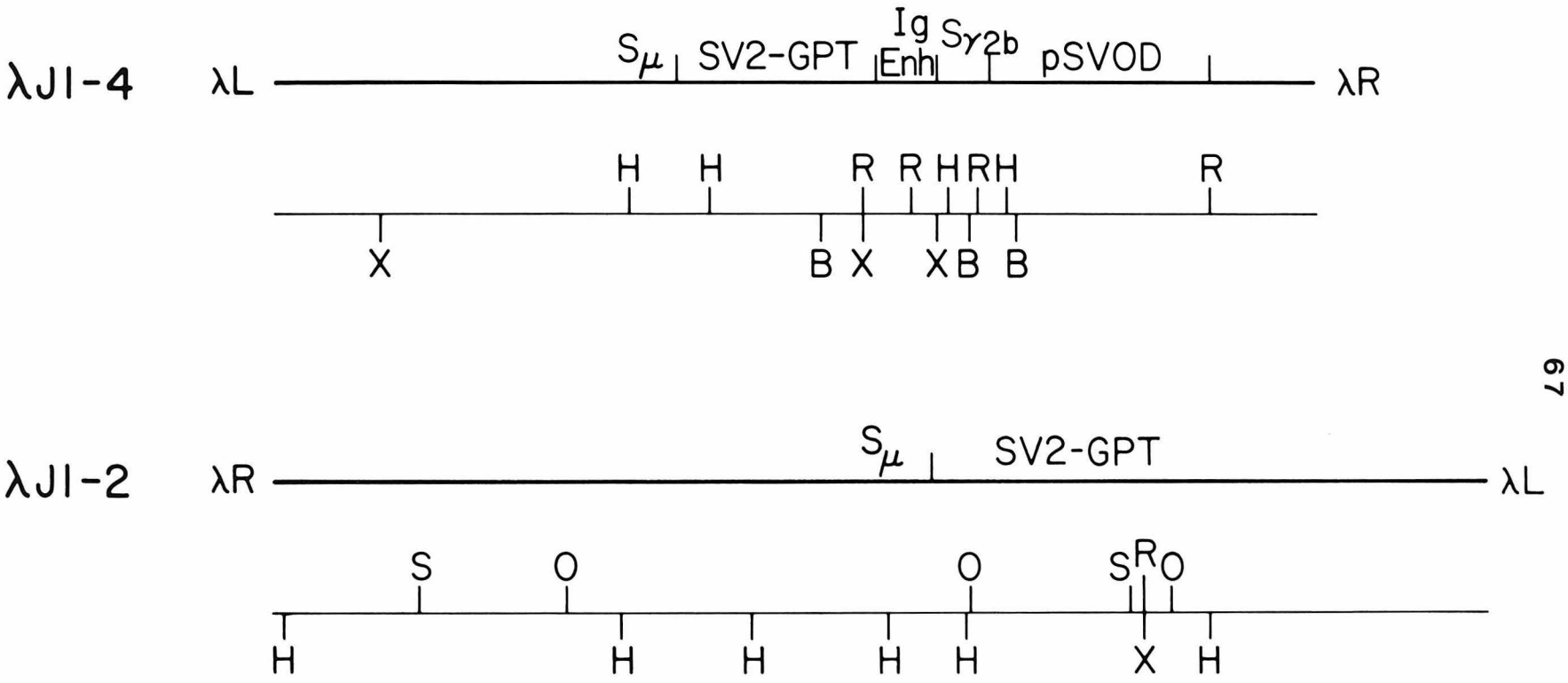
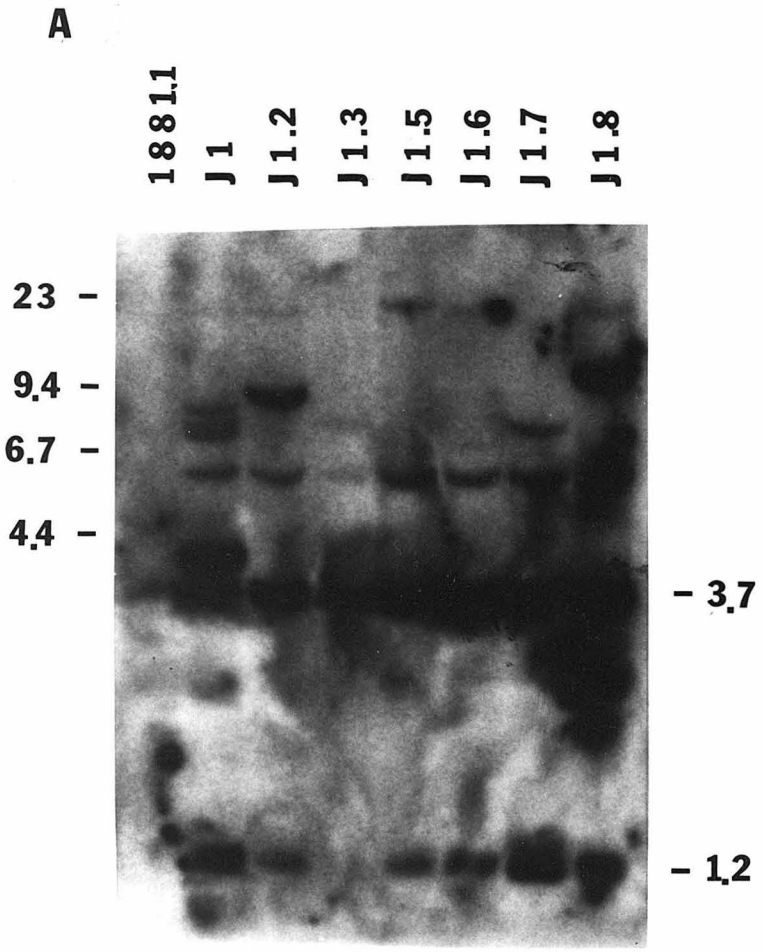

λJ1-4
λL
Sμ
SV2-GPT
Ig Enh
Sγ2b
pSVOD
λR
H H R R H R H R
X B X X B B
λJ1-2
λR
Sμ
SV2-GPT
λL
S O O S R O
H H H H H X H

fig. 3


A
1881.1
J1
J1.2
J1.3
J1.5
J1.6
J1.7
J1.8
23
9.4
6.7
4.4
3.7
1.2
probe SV2GPT
HIND III

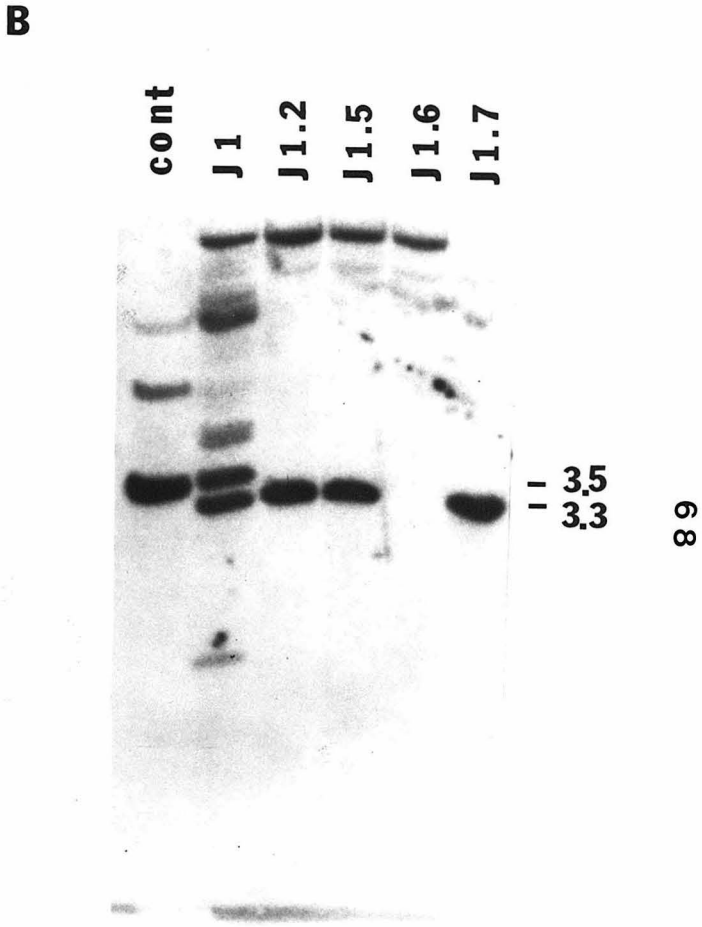
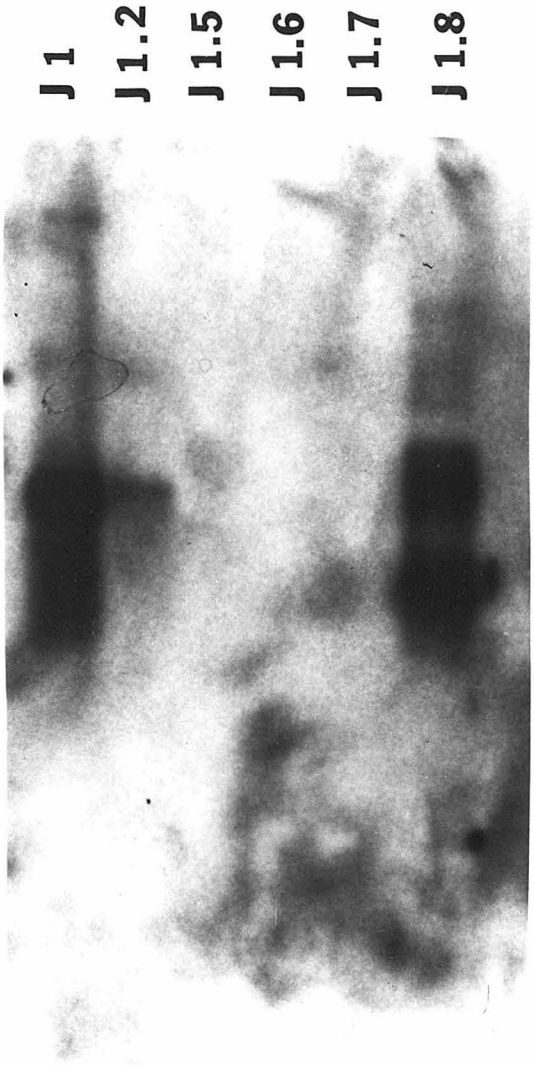
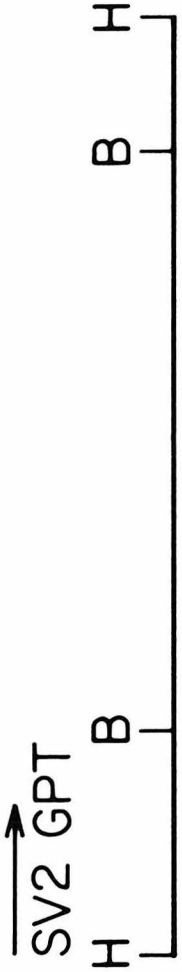

B
cont
J1
J1.2
J1.5
J1.6
J1.7
3.5
3.3
probe pSVOD
Eco RI

fig. 4



probe SV2G PT



λJ1.8-2

fig. 6

Chapter 5

Growth regulation of c-myc gene expression in normal lymphocytes

Introduction

Most murine plasmacytomas contain translocations between chromosomes 12 and 15 (Harris et al., 1982; Calame et al., 1982; Klein, 1981). The breakpoint on chromosome 12 is in one of the switch regions of the immunoglobulin heavy chain genes and the breakpoint on chromosome 15 is near or within the c-myc gene (Marcu et al., 1983; Crews et al., 1982; Adams et al., 1983; Shen-Ong et al., 1982; Taub et al., 1982). The c-myc gene is the cellular homologue of the oncogene found in the genome of the myelocytomatosis 29 retrovirus. Rearrangements of this gene are not observed in normal B lymphocytes. Evidently immunoglobulin class switching, which produces breakpoints in the switch regions, can uncover the oncogenic potential of c-myc by an aberrant rearrangement which juxtaposes the immunoglobulin switch region with the c-myc gene.

Structural studies of the normal and oncogenic c-myc genes have determined the following. First, the normal c-myc gene has three exons. Although the first exon is untranslated, it may regulate gene expression at the translational level since it contains sequences complementary to sequences near the AUG start codon in the second exon (Saito et al., 1983). As a result of base pairing between these two RNA sequences, a stem loop structure may form under certain circumstances which may regulate translation. Second, the c-myc gene is usually rearranged in murine plasmacytomas and the rearrangement breakpoint usually removes the first exon (Stanton et al., 1983; Crews et al., 1982; Adams et al., 1983; Saito et al., 1983; Taub et al., 1982). The remaining exons are expressed by cryptic promoters present in the first intron. In some cases, these promoters express high levels of c-myc suggesting that an increased quantity of gene product is crucial in tumorigenesis (Shen-Ong et al., 1982; Adams et al., 1983; Marcu et al., 1983; Mushinski et al., 1983). Finally, the amino acid sequence of the c-myc gene product is occasionally

altered by somatic point mutations in the gene (Rabbitts et al., 1984; Saito et al., 1983; Taub et al., 1983). These mutations may possibly be due to the same mechanism which mutates antibody genes. This result suggests that altered type of gene product may be important in tumorigenesis.

In addition to these structural studies, questions can be asked about the function of c-myc in both normal development and tumorigenesis. The transformed phenotype may not be due solely to the oncogenic activation of c-myc since murine plasmacytomas may also contain mutated c-mos, c-ras, and B-lym genes (for a review, see Varmus, 1984). One possibility is that the transformed phenotype is due to the concurrent action of several oncogenes and if so, the component of the phenotype due to c-myc is unknown. Another possibility is that, through a "hit and run" mechanism, some oncogenes were used early in oncogenesis but are no longer functional in the final tumor. If used early, the c-myc gene may not have any function in the final tumor; alternatively, c-myc may be the only oncogene left functioning. It would be extremely useful to observe either a normal cell with the activated c-myc gene alone (e.g. Land et al., 1983) or a plasmacytoma deprived only of the activated c-myc gene (e.g. using antisense RNA as in chapter 6).

In the work presented here, we have studied c-myc gene expression in normal B lymphocytes in order to possibly gain insight into its cellular function. First, the pattern of c-myc expression in proliferating and non-proliferating lymphocytes was analyzed. It was shown that c-myc expression is increased when normal B and T lymphocytes are stimulated to proliferate. Since the expression of this oncogene correlates with proliferation, it may possibly function in regulating normal cellular proliferation. When expressed from novel promoters in plasmacytomas, uncontrolled proliferation may result.

One way in which c-myc could regulate cellular proliferation is that it could signal for continued passage through the cell cycle. For example, the c-myc gene product may be required at a particular point in the cell cycle in order for the cell to enter the next phase. If the gene product acts as a "gate", thereby deciding whether or not the cell can continue through the cell cycle, its expression may be limited to the particular phase of the cycle in which it functions. Accordingly, expression of c-myc in cells enriched for various phases of the cell cycle was analyzed. In normal, proliferating B cells, expression of the c-myc, immunoglobulin C_u and actin genes were all similar in different phases of the cell cycle. Thus, it appears that c-myc gene expression is not controlled within the cell cycle.

Results

The bacterial mitogen, lipopolysaccharide (LPS), and the plant lectin, concanavalin A (Con A), cause polyclonal proliferation of B and T lymphocytes respectively. As shown in figure 1, c-myc expression increases when lymphocytes are stimulated with either LPS or Con A. Spleen cells were stimulated for 0, 1, 2, 3 or 4 days in 50 ug/ml LPS or 10 ug/ml Con A. Poly A^+ RNA was isolated from these cells, electrophoresed on a 1% agarose, 5.5% formaldehyde gel, blotted onto nitrocellulose and hybridized to a c-myc cDNA probe. In either LPS-stimulated B cells or Con A-stimulated T cells, expression of c-myc was at least 20 fold higher than in non-stimulated spleen cells. This result indicates that c-myc expression is growth regulated.

Sometimes, genes that are expressed in growing but not resting cells are also preferentially expressed at a specific point within the cell cycle of proliferating cells. Examples include the human chromosomal thymidine kinase gene and the murine dihydrofolate reductase gene (Bradshaw, 1983; Leys and Kellems, 1981). Furthermore, cell cycle control of normal c-myc expression might explain why the rearranged c-myc gene may become oncogenic even when expressed at low levels. In many plasmacytomas, the rearranged c-myc gene is poorly expressed raising questions regarding the cause of activation of this oncogene. One possibility is that the normal promoter is under cell cycle control whereas the cryptic promoter is not. In this model, any expression of c-myc, at the wrong time in the cell cycle, can lead to continued passage through the cell cycle and thus uncontrolled cellular growth.

In order to test this model, lymphocytes enriched for various phases of the cell cycle were prepared and c-myc RNA expression was analyzed. Serum starvation, metabolic inhibitors, and elutriation are techniques capable of enriching for cells in a given part of the cell cycle. We chose elutriation because only it leaves cell growth undisturbed. Elutriation separates cells primarily by size. Non-stimulated lymphocytes, stimulated lymphocytes and macrophages are each of a distinct size. Since cell size increases constantly through the cell cycle, the subpopulation of stimulated lymphocytes can be further divided so that each each fraction contains cells highly enriched for a particular phase of the cell cycle. For each fraction, the phase in the cell cycle of individual cells was determined by staining with the dye mithramycin, which stains DNA quantitatively. Subsequent analysis by flow cytometry can accurately determine the DNA content of single cells in each fraction.

As seen in figure 2, cells present in earlier fractions of elutriated cells contain half as much DNA as cells present in later fractions. This observation indicates that the earlier fractions contain cells in the G1 phase and the later fractions contain cells in the G2 and M phases. Cells from intermediate fractions, with a DNA content between 1n and 2n, are in the S phase. RNA was prepared from the fractions predominantly containing cells in the G1, S and G2 + M phases. The presence of c-myc RNA was assayed by Northern blotting. Total RNA was fractionated on a 1% agarose, 5.5% formaldehede gel, blotted onto nitrocellulose and hybridized to a c-myc cDNA probe. As internal controls, the immunoglobin C_u and the cytoplasmic actin RNA levels were measured. As can be seen in figure 3, these genes all show a comparable pattern of expression in which the band in the lane containing G2 + M phase RNA is strong and the band in the lane containing G1 phase RNA is weak. Either these genes are expressed predominantly in the G2 + M

phases or unequal amounts of RNA were inadvertantly added to the three lanes. In either case, the pattern of c-myc expression does not differ from either C_u or actin expression throughout the cell cycle.

Discussion

One of the results presented in this report shows that c-myc gene expression is increased when splenocytes are stimulated to proliferate with either LPS or Con A. It appears that c-myc preferentially functions in proliferating rather than in non-proliferating lymphocytes. These results were subsequently confirmed elsewhere (Kelly et al., 1983; Campisi et al., 1984). Explicitly, from these experiments it has not been shown that c-myc gene expression causes proliferation.

However, the causative role of c-myc in proliferation has been strongly suggested by the following experiment. If normal fibroblasts are transformed with the rearranged c-myc gene, they become immortalized as they are capable of indefinite proliferation (Land et al., 1983). This result implies that c-myc can cause cellular proliferation and thus acts at the "establishment" step in tumorigenisis.

The c-myc gene product might function by controlling the cell cycle; cells in the G0 phase are induced to enter G1 and cells in the G1 phase are allowed to continue to the S phase. If this model is correct, then c-myc might function at a specific phase of the cell cycle and its expression might be limited to that

particular point. One advantage of this model is that it could explain the following observation. Some plasmacytomas have rearranged c-myc genes which are expressed at low levels (e.g. 1 RNA/cell for 38C13 or M167, unpub. obs.) suggesting that some c-myc genes can be oncogenic even if expressed at low levels. Although there are alternatives, an attractive possibility is that c-myc is normally expressed at only a certain time in the cell cycle and that any amount of aberrant expression at other times allows the cell to proliferate in an uncontrolled manner. In this case, rather than overexpression, the cause of oncogenic activation would be expression at the wrong time in the cell cycle.

In order to test this hypothesis, the pattern of normal c-myc gene expression was analyzed at different points of the cell cycle in proliferating lymphocytes. When RNA from fractions enriched for certain phases of the cell cycle was analyzed by Northern blot analysis, it was observed that c-myc expression did not differ from either C_u or actin expression. All three sequences hybridized more strongly to RNA from cells enriched in the G2 + M phase than to RNA from cells enriched in the G1 phase. Either all three genes are expressed predominantly in the G2 or M phases or more RNA was inadvertantly loaded onto this lane. In either case, it does not appear that the level of c-myc gene expression within the cell cycle differs from the level of expression of these other genes.

One possible caveat to these results is that these experiments were performed on populations of spleen cells. It is possible that the difference in c-myc expression is not due to the difference between growing versus non-growing cells but rather due to the selective outgrowth of cells which always express c-myc. This possibility is unlikely since cells which are stimulated by LPS or Con A already comprise a significant fraction (40%) of splenocytes so that their selective outgrowth cannot cause the twenty fold difference observed for c-myc expression.

Similarly, it can be argued that the cell cycle experiments utilized a population of proliferating B cells of an unknown number of types. Mithramycin staining indicated that these fractions contained cells most of which were in the same phase of the cell cycle. Thus, c-myc cannot be under cell cycle control in all cells. However, it is possible that cell cycle regulation of c-myc expression may occur in a fraction of these cell types.

As can be seen in figure 2, although the elutriated fractions were enriched for certain phases of the cell cycle, they were not pure. Analysis of the pooled fractions showed that the "G1 phase" population contained 90% of its cells in the G1 phase, the "S phase" population contained 65% of its cells in the S phase ($\tilde{6}$ fold enrichment), and the "G2 + M phase" population contained 40% of its cells in the G2 or M phases ($\tilde{8}$ fold enrichment). A large variation in the quantity of c-myc mRNA could be easily observed in these enriched fractions. However, this experiment was not sufficiently sensitive to detect small variations in expression.

Materials and Methods

Isolation of a c-myc cDNA clone

A c-myc cDNA clone was isolated in the lambda insertion vector lambdagt10. PolyA RNA from the plasmacytoma M167 was primed with oligo dT and converted

into double stranded cDNA by conventional techniques. Eco R1 linkers were added and the cDNA was cloned into Eco R1 digested lambdagt10 DNA. Out of 5 x 10^6 independent recombinants, 12 c-myc cDNA clones were isolated using a genomic c-myc gene probe. Thus, the rearranged and normal c-myc alleles in the M167 plasmacytoma are expressed at a low level. The recombinant lambdamyc29 had the largest c-myc insert and this insert was subcloned into pBR325 at the Eco R1 site. The insert is 1.55 kb long and contains exons 2 and 3 of the c-myc gene. This clone was used as a probe in the following experiments.

Mitogenic stimulation of lymphocytes

Mitogenic stimulation of splenocytes was accomplished by a modification of the procedure of Kearney and Lawton (1975). Balb/c mice were sacrificed by cervical dislocation. The mice were rinsed in ethanol and their spleens were removed using scissors and forceps previously dipped in ethanol. The spleens were immediately placed in 10 ml of Dulbecco's phosphate buffered saline (PBSA) in a 50 ml conical centrifuge tube on ice. One spleen contains approximately 10^8 B cells and 10^8 T cells. A single cell suspension was made as follow. A 200 mesh stainless steel screen was sterilized and placed over a 60 mm tissue culture dish containing 5 ml of PBSA, 1 mM EDTA (EDTA was added to prevent coaggulation of the cell suspension). The spleen was placed on the wire mesh and cut into 1 cm cubes with sterile scissors and forceps. The blunt end of a sterile 5 ml syringe was used to pass single splenocytes through the mesh. The PBSA was placed in a sterile tube and the cells were collected by centrifugation. The cells were washed once more in PBSA, 1 mM EDTA. The washed cells were placed in RPMI 1640, 10% fetal calf serum, 50

uM betamercaptoethanol (RBF_{10} medium). Viable cells were counted in 0.2% trypan blue. The cells were cultured at a density of 10^6 splenocytes/ml. LPS (Sigma) or Con A (Sigma) was added at 50 ug/ml or 10 ug/ml respectively.

Cell cycle synchonization by elutriation

Elutriation of lymphocytes was accomplished by a modified procedure of Rothenberg (1982). The splenocytes from 20 spleens were stimulated with 50 ug/ml LPS in vitro for three days yeilding approximately 4×10^9 B lymphocyte blasts. These cells were pelleted, washed in PBSA, pelleted and suspended at 4×10^8 cells/ml in Ca^{2+}- and Mg^{2+}- free Hank's balanced salts solution supplemented with 5 mM sodium azide and 2.5 mg/ml bovine serum albumen (BSS/BSA). The cells were fractionated on a JE-6 elutriator rotor (Beckman) spinning at 2850 RPM. The variable flow rate was controlled by a Cole-Parmer Masterflex pump (Beckman). Red blood cells and unstimulated lymphocytes were elutriated with a flow rate between 11.8 ml/min. and 19.7 ml/min. Stimulated B lymphocytes appeared when the flow rate was between 23.8 ml/min. and 42.2 ml/min. These cells were further fractionated into G1, S and G2 + M phases which elutriated at 23.8 ml/min. - 26.5 ml/min., 28.8 ml/min. - 34.1 ml/min and 36.1 ml/min. - 37.5 ml/min respectively. Macrophages and other large cells elutriated with flow rates of greater than 42.8 ml/min. The cells were collected in sterile tubes on ice. The DNA content of individual cells in each fraction was measured by staining the cells with mithramycin and analysis by flow cytometry (Crissman and Tobey, 1974). A small portion of each fraction was pelleted and resuspended in 25% aqueous ethanol containing 11 ug/ml mithramycin and 15 mM $MgCl_2$. After 15 minutes, the cells were analyzed by flow cytometry on the fluorescence activated cell sorter.

Northern blotting

Total RNA was prepared by the method of Chirgwin (1979) except that the initial homogenates were sonicated to shear chromosomal DNA which improved RNA recovery and purity. Total RNA (20 ug) was fractionated on 1% agarose, 5.5% formaldehyde gels . The RNA was blotted onto nitrocellulose, hybridized to radiolabelled pmyc29 probe, washed and autoradiographed.

References

Adams, J. M., Gerondakis, S., Webb E., Corcoran, L., and Cory, S. 1983. Cellular c-myc oncogene is altered by chromosome translocation to an immunglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. Proc. Nat. Acad. Sci. U.S.A. 80: 1982-1986.

Bradshaw, H. 1983. Proc. Nat. Acad. Sci. U.S.A. **80:** 5588.

Calame, K., Kim, S., Lalley, P., Hill, R., Davis, M., and Hood, L. 1982. Molecular cloning of translocations involving chromosome 15 and the immunoglobulin C alpha gene from two murine plasmacytomas. Proc. Nat. Acad. Sci. U.S.A. 79: 6994-6998.

Campisisi, J., Gray, H. E., Pardee, A. B., Dean, M., and Sonnenshein, G. E. 1984. Cell cycle control of c-myc but not c-ras expression is lost following chemical transformation. Cell 36: 241-247.

Chirgwin, J., Przybala, A., MacDonald, R. and Rutter, W. 1979. Isolation of biologically active RNA from sources enriched in ribonucleases. Biochem. 18: 5295-5299.

Crews, S., Barth, R., Hood, L., Prehn, J. and Calame, K. 1982. Mouse c-myc oncogene is located on chromosome 15 and translocated to chromosome 12 in plasmacytomas. Science 218: 1319-1321.

Crissman, H. A., and Tobey, R. A. 1974 Cell cycle analysis in 20 minutes. Science 184: 1297-1298.

Harris, L. J., D'Eustachio, P., Ruddle, F. H., and Marcu, K.B. 1982. DNA sequence associated with chromosometranslocations in mouse plasmacytomas. Proc. Nat. Acad. Sci. U.S.A. 79: 6622-6626.

Kearney, J. F., and Lawton, A. R. 1975. B lymphocyte differentiation induced by lipopolysaccharide. J Imm. 115: 671-676.

Kelly, K., Cochran, B. H., Stiles, C. D., and Leder, P.1983. Cell specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor. Cell 35: 603-610.

Klein, G. 1981. The role of gene dosage and genetic transposition in carcinogenesis. Nature 294: 313-318.

Land, H., Parada, L., and Weinberg, R. A. 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 304: 596-602.

Leys, J. and Kellems, R. 1984. Mol. Cel. Biol. 1: 961.

Marcu, K. B., Harris, L. J., Stanton, L. W., Erikson, J., Watt, R. and Croce, C. M. 1983. Transcriptionally active c-myc oncogene is contained within NIARD, a DNA sequence associated with chromosome translocations in B cell neoplasia. Proc. Nat. Acad. Sci. U.S.A. 80: 519-523.

Mushinski, J. F., Baker, S. R., Potter, M., Reddy, E. P. 1983. Increased expression of myc-related oncogene mRNA characterizes most Balb/c plasmacytomas induced by Pristane or Abelson murine leukemia virus. Proc. Nat. Acad. Sci. U.S.A. 80: 1073-1077.

Rabbitts, T. T., Forster, A., Hamlyn, P., and Baer, R. 1984. Effect of somatic mutation within translocated c-myc genes in Burkitt's lymphoma. Nature 309: 592-597.

Rothenberg E. 1982. A specific biosynthetic marker for immature thymic lymphoblasts. Active synthesis of thymus leukemia antigen restricted to proliferating cells. J. Exp. Med. 155: 140.

Saito, H., Hayday, A., Wiman, K., Howard, W. and Tonegawa, S. 1983. Activation of the c-myc gene by translocation: A model for translational control. Proc. Nat. Acad. Sci. U.S.A. 80: 7476-7480.

Shen-Ong, G. L. C., Keath, E. J., Piccoli, S. P., and Cole, M. D. 1982. Novel myc oncogene RNA from abortive immunoglobulin gene recombination in mouse plasmacytomas. Cell 31: 443-452.

Stanton, L. W., Watt, R., and Marcu, K. B. 1983. Translocation breakage and truncated transcripts of c- myc oncogene in murine plasmacytomas. Nature. 303: 401-406.

Taub, R. Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S. and Leder, P. 1982. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and mouse plasmacytoma cells. Proc. Nat. Acad. Sci. U.S.A. 78: 7837-7841.

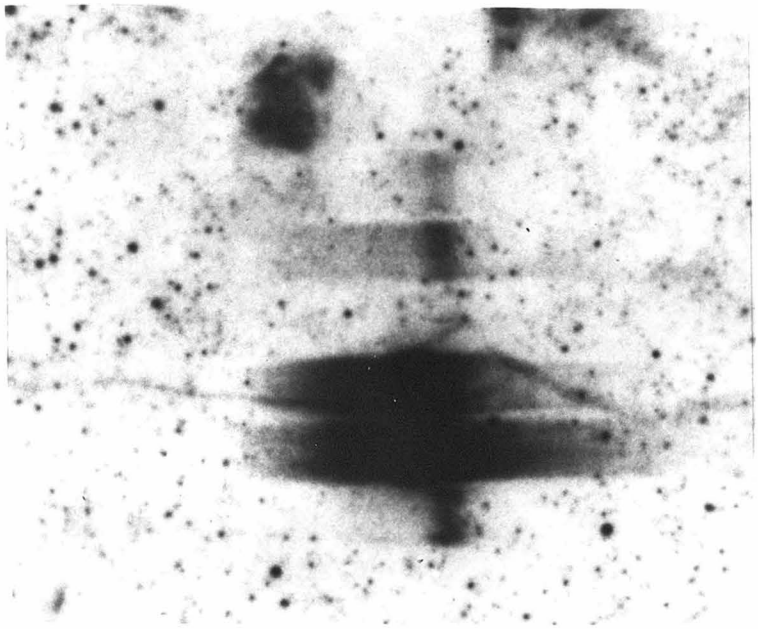
Varmus, H. 1984. The molecular genetics of cellular oncogenes. Ann. Rev. Genet. 18: 553-612.

Figure legends

fig. 1. Expression of c-myc in LPS and Con A stimulated lymphocytes. Lymphocytes were stimulated <u>in vitro</u> for the time shown with either 50 ug/ml LPS or 10 ug/ml Con A. RNA was prepared and 5 ug Poly A^+ RNA was analyzed by Northern blotting. Also included are RNAs from a pre B cell (1881.1), a plasmacytoma (SP2) and a lymphoma (38C13). The blots were hybridized to a c-myc cDNA clone (pmyc29).

fig. 2. Flow cytometry analysis of elutriated cells. Elutriation was used to enrich LPS stimulated B cells in different phases of the cell cycle. The DNA from the enriched fractions was stained with mithramycin and analyzed by flow cytometry. A. fraction enriched for G1 cells. B. fraction enriched for S cells. C. fraction enriched for G2 and M phase cells.

fig. 3. Northern blot analysis of c-myc, C_u, and actin gene expression in different phases of the cell cycle. RNAs from the fractions shown in fig. 2 were prepared. 20 ug of total RNA was electrophoresed on a formaldehyde gel and blotted. The origin of the RNA is shown at the top; tot refers to the total population of cells before elutriation. The filters were sequentially hybridized with c-myc, C_u and actin as shown at the bottom.



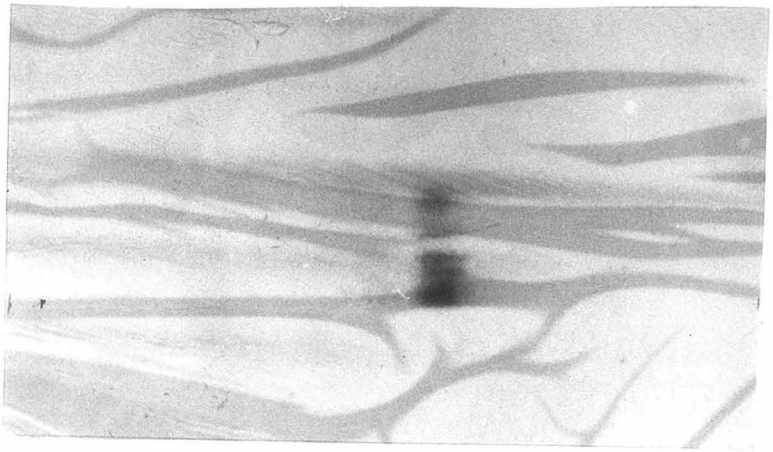


fig 1

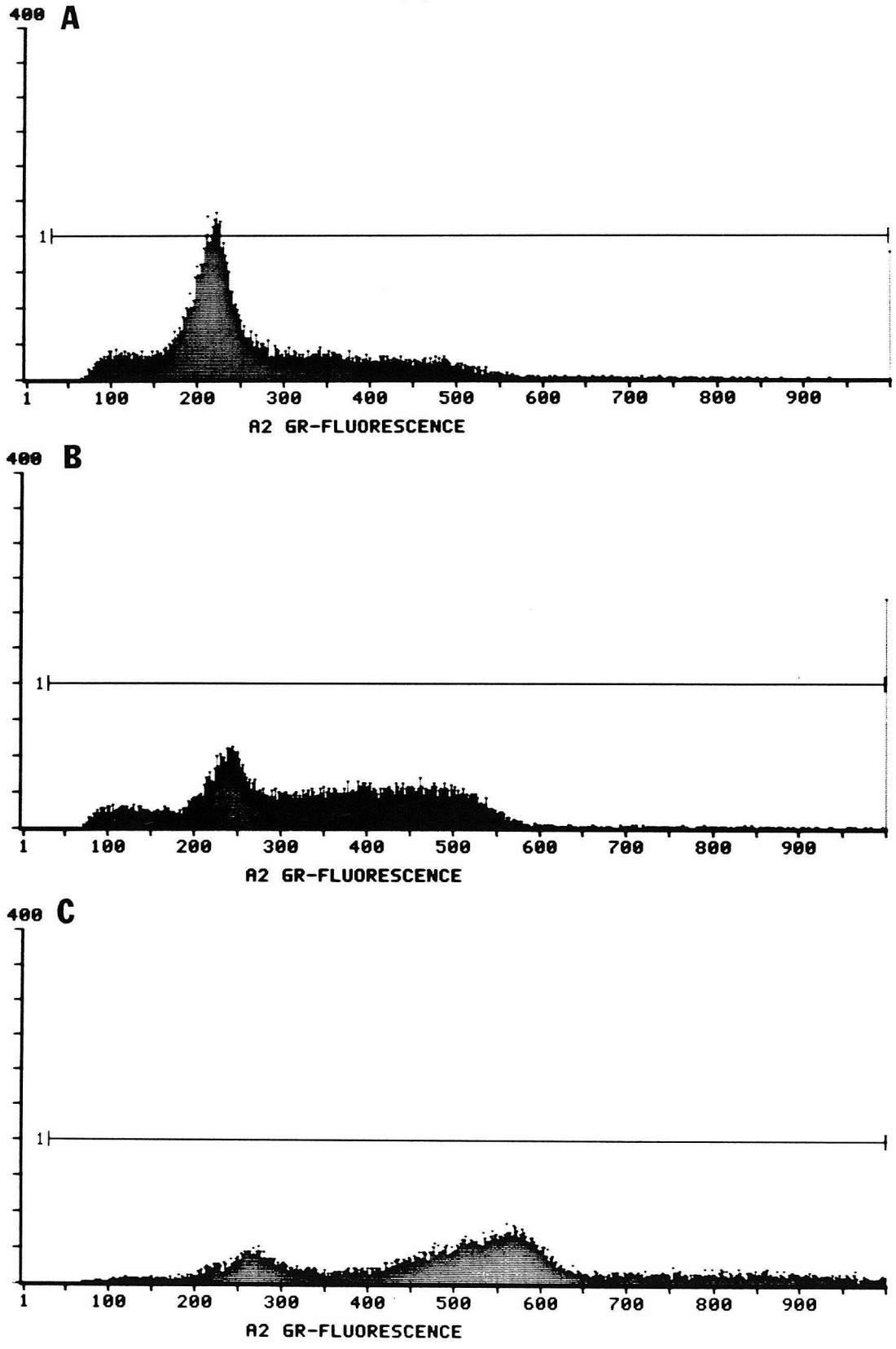

400
A
1
1
100
200
300
400
500
600
700
800
900
A2 GR-FLUORESCENCE
400
B
1
1
100
200
300
400
500
600
700
800
900
A2 GR-FLUORESCENCE
400
C
1
1
100
200
300
400
500
600
700
800
900
A2 GR-FLUORESCENCE

fig. 2

fig 3

Chapter 6

Stable Reduction of Thymidine Kinase Activity in Cells Expressing High Levels of Antisense RNA

Summary

We have investigated the feasibility of using antisense RNA to diminish the level of a given gene product in a specific and stable fashion. The experiments are based on the idea that antisense RNA might be capable of hybridizing to mRNA (or its precursor) in vivo, thereby rendering it biologically inactive. As a test case, we have attempted to reduce thymidine kinase (TK) activity in cultured mammalian cells. Antisense TK RNA was expressed as part of a chimeric dihydrofolate reductase-antisense TK transcript. These antisense genes were introduced by DNA cotransformation into a mouse L cell line which expresses thymidine kinase constitutively. These transformants were then selected for progressively higher resistance to methotrexate, provided by overproduction of dihydrofolate reductase (DHFR). The result of this selection procedure was a concomitant increase in antisense TK RNA concentration, owing to its presence on the same transcription unit as DHFR. In this manner, cells expressing 5,000 to 10,000 antisense molecules per cell were obtained.

In several cell lines expressing high levels of antisense TK RNA, thymidine kinase enzyme activity was reduced by eighty to ninety percent. One such line was characterized at the molecular level, and it was found that RNA:RNA duplexes could be detected in the nuclear RNA fraction. In addition, sense TK RNA is present in these cells at a level that is equal to or greater than the level in parental

TK^+ cells, but it is predominantly localized to the nucleus. These results suggest that antisense RNA hybridizes with sense TK RNA in the nucleus. The duplexes fail to enter the cytoplasm with normal efficiency, and as a consequence, the amount of TK message available for translation is reduced.

Introduction

Molecular cloning and DNA sequencing procedures have permitted the isolation and characterization of a formidable array of genes from higher eukaryotes. Despite the detailed structural knowledge provided by sequence analysis of such genes, the functions of their protein products remain, in many cases, incompletely understood and in some instances entirely unknown. To analyze further the cellular function of a cloned gene, several approaches may be taken. These include altering either the type or the amount of gene product expressed in vivo.

In DNA mediated gene transfer experiments, the expression of a gene can be increased in cells which do not otherwise express it or express it at a low level. These experiments can be instructive when the addition of gene function gives rise to a detectable phenotypic change in the recipient cell or organism.

An important complementary approach is to specifically eliminate the gene product of interest or block its function. In addition to classical genetic methods, the task of establishing the null or hypomorphic phenotype has been undertaken using two molecular techniques. One elegant molecular approach employs antibodies which, when introduced into cells, appear to specifically inhibit the function of the protein(s) against which they are directed (Mabuchi and Okuno,

1977; Yamaizumi et al., 1978; Antman and Livingstone, 1980; Lin and Feramisco, 1981; Burke et al. 1984; Warner et al., 1984). However, this technique may be limited in its generality. In some cases, the immune reagent may be ineffective due to differences in subcellular compartmentalization of antigen and antibody. Moreover, when antibody binding does occur, it may fail to neutralize protein function.

A second molecular approach involves antisense nucleic acids. In this case, the hypothesis is that sufficiently high intracellular concentrations of nucleic acids complementary to a given messenger RNA species will hybridize with that mRNA (or its precursor) in vivo, and that such hybridization will ultimately inhibit synthesis of the protein product.

There are a number of observations that suggest that this hypothesis may be correct. For example, in E. coli, expression of the IS10 transposase gene (Simons and Kleckner, 1983) and the Omp F gene (Mizuno et al., 1984) are apparently regulated at the translational level by antisense RNAs. Also, the activity of an RNA which serves to prime replication for the Col E1 plasmid appears to be reduced by annealing with a complementary RNA (Tomizawa et al., 1981). These results indicate that antisense RNAs may exert regulatory effects by hybridizing with their complements in vivo. Nevertheless, it is important to recognize that proteins may play an important part in mediating these hybridization events (e.g. see Tomizawa 1984).

It is then of interest to test whether antisense nucleic acids can generally be employed to reduce the activity of a specific gene. In E. coli, expression of antisense IPP, Omp A, Omp C and lac Z RNAs reduced the cellular concentration of

the corresponding proteins (Coleman et.al. 1984; Pestka et al., 1984). In frog oocytes, injection of antisense globin RNA into the cytoplasm inhibited synthesis of globin from microinjected beta globin messenger RNA (Melton, 1985). It was shown that RNA:RNA duplexes formed in the oocyte after microinjection, suggesting that RNA hybridization was instrumental in establishing the observed block in globin protein synthesis. One important implication of these results is that the introduction of a pulse of antisense RNA may be useful in preventing the onset of expression from a particular gene in early embryogenesis. For example, microinjection of antisense kruppel RNA appears to cause a defect in Drosophila embryo segmentation which is analogous to defects caused by mutations in the kruppel gene (Rosenberg et al., 1985).

In mammalian cells, an important initial observation was reported by Izant and Weintraub (1984). Microinjection of TK genes into mammalian cells typically results in a burst of TK synthesis which can be monitored by a single cell autoradiographic assay for incorporation of ^{3}H-thymidine. If the recipient cells also contained an excess of genes coding for antisense TK RNA, a large diminution in TK expression from the injected thymidine kinase genes was observed. The observed reduction in transient thymidine kinase expression was shown to be dependent on sequence complementarity between sense and antisense RNA genes, as would be predicted if the reduction in enzyme activity is mediated by antisense RNA hybridization. Since this experimental approach is confined to small cell numbers, direct investigation of the mechanism of enzyme depression is precluded. Furthermore, if this approach is to be generally applicable, the effect of antisense RNA should not be limited to a transient reduction of the activity of microinjected genes.

Our objectives are to develop a general protocol for the stable reduction of a given gene product in mammalian cells. Toward these objectives, we believe that the following points are relevant. First, an approach based on production of antisense RNAs from genes stably located in each cell's nucleus is desirable. An advantage of endogenous antisense production is that any effect that it may generate will be evident in a large population of cells, permitting examination at the biochemical level. Second, it will be important to be able to reduce the activity of resident chromosomal genes because these will most often be the targets of interest for studies utilizing antisense RNA. Third, it is expected that antisense RNA should inhibit the synthesis of specific proteins. Therefore, inhibition must be maintained long enough to allow for the disappearance of any pre-existing protein. This means that a pulse of antisense RNA will be ineffective in any case in which the antisense RNA degrades more rapidly than the protein product. Finally, it would be advantageous to express the antisense RNA from an inducible promoter so that the reduction in protein level could be made conditional.

In the work presented here, sense and antisense Herpes virus TK genes were stably introduced into mouse L cells where they expressed TK mRNA and antisense TK RNA respectively. At high concentrations of antisense TK RNA, thymidine kinase activity was stably diminished in several cell lines. Analysis of RNA extracted from one such cell line revealed that double stranded TK RNA is present in the nuclear compartment. Furthermore, the level of sense TK RNA in the cytoplasm is much lower for these cells than in parental TK^+ cells. Based on these observations, a model for the mechanism of TK diminution by antisense RNA is discussed.

Results

Initial antisense genes are not expressed when transfected into cells

One straightforward way to synthesize antisense RNA is to invert a coding region sequence relative to its promotor. In preliminary experiments, antisense dihydrofolate reductase or thymidine kinase sequences were inserted between the metallothionein promoter and a metallothionein intron and polyadenylation addition site. The metallothionein gene was chosen because transcription can be induced by heavy metals, so that the expression of antisense transcripts could be regulated. Each of these recombinant genes was introduced into TK^+ L cells by DNA mediated gene transfer. Analysis of the L cell transformants containing either of these antisense genes suggested that neither TK nor DHFR activity was significantly reduced upon induction with heavy metals. Surprisingly, little or no antisense DHFR or TK RNA was detected in RNA blotting experiments (S. Kim, unpub. obs.).

In these initial experiments, a major problem was a failure to accumulate significant quantities of antisense RNA. It is known that numerous chimeric genes containing the metallothionein promotor direct the synthesis of high levels of messenger RNA under induction conditions similar to those tested here (Mayo et al., 1982). A particularly relevant example is vector pMD (figure 1), which contains the complete DHFR protein coding sequence in sense configuration flanked by promotor and 3' sequences from the metallothionein gene. Unlike the analogous gene containing DHFR sequences in the antisense orientation, substantial levels of mRNA are produced upon heavy metal induction of cells containing the pMD plasmid (P. Mueller and B. Wold, to be presented elsewhere). Taken together, these results suggest that the failure to accumulate antisense transcripts may be

due to post-transcriptional events such as aberrant processing, transport or turnover.

Design of Amplifiable Antisense RNA Genes

As an alternative to the simple antisense genes described above, a second generation of plasmids were designed which we expected would ultimately produce higher steady state levels of antisense RNA. These genes contain the dihydrofolate reductase protein coding region (in the sense orientation) followed by antisense TK sequences (figure 1). RNA transcribed from these chimeric genes is expected to be bifunctional, directing the synthesis of an enzyme, DHFR, and providing antisense TK RNA. The level of DHFR can be significantly increased, usually by means of gene amplification, by selecting cells resistant to the drug methotrexate (Mtx; for a review, see Stark and Wahl, 1984). Since antisense TK sequences appear on the same RNA as DHFR sequences, selection for cells resistant to high levels of Mtx should result in the production of large amounts of antisense TK RNA. If antisense TK RNAs are capable of hybridizing with sense TK RNAs in vivo, then it would be expected that the level of TK activity will decrease as the concentration of antisense TK RNA increases.

In these experiments, three antisense genes were employed (figure 1). The parental plasmid, pMD, contains a gene consisting of a metallothionein promoter, DHFR coding region, metallothionein intron and metallothionein poly A addition site. DNA fragments containing thymidine kinase sequences were inserted into the 3' untranslated region in inverted orientation relative to the DHFR sequences. Thus, transcription originating from the metallothionein promoter would be expected to produce a messenger RNA coding for reductase and containing

antisense TK sequences. The first plasmid, pMDKT1, contained a 1.1 kb fragment derived from the TK gene. This fragment, also used by Izant and Weintraub (1984), includes 53 base pairs encoding the 5' untranslated region of the TK mRNA and all of the protein coding sequence except the carboxy terminal 23 bases. The second vector, pMDKT2, was constructed by deleting a central fragment (denoted segment B) from the TK sequences in pMDKT1 leaving behind a 301 bp fragment from the 5' portion (segment A) and a 225 bp sequence from the 3' portion of the TK gene (segment C). The third vector, pMDKT3, has been further deleted. It retains only the terminal 225 base pair fragment (segment C) corresponding to the carboxy terminus of thymidine kinase, and possesses a reduced 3' untranslated region from the metallothionein I gene. The transcripts expected to be produced from each of these genes are also indicated in figure 1.

Introduction of antisense vectors into TK^+ L cells

In order to construct a parental thymidine kinase positive cell line in which antisense TK genes would subsequently be tested, DNAs containing TK and adenine phosphoribosyl transferase (APRT) genes were introduced into L $APRT^-$ TK^- cells by DNA mediated gene transfer (Wigler et al. 1977). A mixture of 1 ug pTK DNA, 0.1 ug pAPRT DNA, and 20 ug carrier DNA was added as a calcium phosphate precipitate per 100 mm plate. Genes that are transfected into cells in culture are not as stable as endogenous genes; the rate of phenotypic reversion for cells containing a gene introduced by DNA transfer can be two to five orders of magnitude higher than the mutation rate for endogenous genes (Pellicer et al., 1980). For example, transfected TK genes may revert at a frequency of 10^{-1} to 10

$^{-4}$ revertants/cell. A high frequency of spontaneous reversion of thymidine kinase positive cells to the TK negative phenotype would obviously complicate interpretation of any reduction in TK activity in the presence of antisense TK RNA. To ameliorate this problem, the APRT gene was included. If a second gene is cotransfected with TK, i.e. APRT, both genes appear to integrate into the same chromosomal location and selection for the second gene often reduces the frequency of spontaneous reversion of the first gene (Roberts and Axel, 1982). Accordingly, cells expressing APRT were selected and ten clones were isolated. Seven of the APRT$^+$ clones were also TK$^+$. One of the seven clones reverted from APRT$^+$, TK$^+$ to Aprt$^+$, TK$^-$ at a frequency less than 10^{-5} revertants/cell in three independent experiments. This frequency of reversion is sufficiently low that thymidine kinase activity is expected to be unaffected by the accumulation of spontaneous TK revertants. Therefore, this clone, designated P, was used as the parental cell line into which antisense TK vectors were introduced. In all experiments that follow, P cells were grown in media which imposes selection for APRT.

Next, the pMD control vector and the antisense plasmids were each introduced into the parental cell line by DNA mediated gene transfer. To each 100 mm plate of parental cells was added 1 ug pNEO3 DNA, 10 ug carrier DNA and 10 ug of pMD, pMDKT1, pMDKT2, or pMDKT3 DNA as a calcium phosphate precipitate. The pNEO3 vector (B. Wold and R. Axel, unpub. results) expresses a bacterial enzyme which allows mammalian cells to grow in the presence of the drug G418 (Southern and Berg, 1982). Transformants were selected by growth in G418 yielding several hundred colonies per plate. Transformants from each plate were subsequently pooled. Five plates of parental cells were transfected with pMDKT1 resulting in five pools of transformants designated KT1-A to KT1-E. Similarly,

cells transfected with pMDKT2 yielded KT2-A to KT2-E, cells transfected with pMDKT3 gave KT3-A to KT3-E and cells transfected with pMD were designated KT0-A to KT0-C.

Low levels of antisense TK RNA do not affect thymidine kinase activity

As expected, all of the transformants exhibited increased levels of resistance to methotrexate relative to parental cells indicating that the newly transfected genes were capable of expressing dihydrofolate reductase, as expected. Prior to transfection, the parental cell line was resistant to 10 nM Mtx whereas cells containing the DHFR chimeric genes were immediately resistant to elevated levels of methotrexate: KT1 cells were resistant to 50 nM MTx, KT2 cells were resistant to 150 nM MTx and KT3 and KT0 cells were resistant to 300 nM Mtx. The varying levels of resistance to methotrexate indicate that differing amounts of DHFR are expressed in cells transfected with each vector. KT1 cells were resistant to the lowest levels of Mtx suggesting that the inclusion of the full 1.1 kb antisense TK sequence in pMDKT1 interfered with DHFR expression. Conversely, KT3 cells were resistant to as much Mtx as KT0 cells indicating that the inclusion of the 225 bp fragment in pMDKT3 did not significantly affect DHFR expression.

Does expression of antisense TK RNA in these primary transformant populations have any effect on the thymidine kinase phenotype? To answer this question, TK enzyme activity was quantitated in KTO, KT1, KT2 and KT3 cells. The data presented in table 1, accurate to within 10%, show that the level of TK activity in these cells was unchanged relative to the parental thymidine kinase positive cell line (see Materials and Methods).

Selection for methotrexate resistance results in increased levels of antisense TK RNA

It was expected that selection for Mtx resistance would often yield cells that express elevated levels of dihydrofolate reductase, provided by high levels of DHFR-antisense-TK RNA. Resistance to Mtx was gradually increased by successive rounds of selection in twofold greater drug concentrations (Alt et.al. 1978). To establish that selection for methotrexate resistance resulted in elevated levels of antisense containing transcripts, RNA was extracted from KT3-B-.3, KT3-B-2 and KT3-B-64 cells, which are resistant to 0.3 uM Mtx, 2 uM Mtx and 64 uM Mtx respectively (the level of resistance to Mtx in uM units is indicated following the dash in each cell name). These RNAs were analyzed for the presence of antisense TK RNA using an RNAse protection assay. Radioactively labeled TK RNA was synthesized in vitro using an SP6 transcription system (Melton et al., 1984). This probe was hybridized to cellular RNA and unhybridized probe was digested with RNAses A and T1, which are specific for single stranded RNA. The protected probe was melted and analyzed by fractionation on a denaturing polyacrylamide gel. As seen in figure 3, approximately fifty to one hundred times more antisense TK RNA is made in cells resistant to 64 uM Mtx than in cells prior to selection for Mtx resistance. Similarly, in KT2-E cells, it was found that selection for Mtx resistance resulted in a large increase in the expression of antisense TK RNA (data not shown). Therefore, as expected, selection for resistance to methotrexate resulted in increased concentrations of DHFR-antisense-TK RNA in cells transfected with either pMDKT2 or pMDKT3. Examination of genomic DNA isolated from parental cells, initial transformants and

cells resistant to 64uM Mtx showed that the predominent means of generating higher levels of DHFR-antisense-TK RNA was by amplification of gene number (data to be presented elsewhere). However, in two independent cases, cells transfected with pMDKT1 did not produce more antisense TK RNA in response to Mtx selection (data not shown). Since these cells did not show a reduction in TK activity and since the concentration of antisense TK RNA was not increased by methotrexate selection, they were not studied further.

KT2 and KT3 cells which are resistant to high concentrations of Mtx have lower TK activity

Since stepwise selection for methotrexate resistance resulted in a gradual increase in the concentration of antisense TK RNA, it was anticipated that at some point the concentration of antisense RNA may become great enough to inhibit thymidine kinase expression. To test this possibility, the level of TK enzyme activity in KT3 cells was determined at each step of increased Mtx resistance. As can be seen in figure 2, there is much less TK activity in KT3-B-64 or KT3-D-64 cells than in parental cells. Beginning with cells resistant to 4 uM Mtx, the level of TK activity continually decreased until only 15% or 18% of the enzyme activity present in parental cells remained in KT3-B-64 and KT3-D-64 cells respectively. The level of TK activity for other transformants before and after Mtx selection is shown in table 1. Two out of four independent pools of KT3 cells showed a significant decrease in TK activity in cells resistant to 64 uM Mtx. Likewise, two out of three pools of KT2 cells exhibited reduced TK activity in cells resistant to 100 uM Mtx.

Specificity of antisense RNA effects

Control experiments showed that the reduction in TK activity in KT2 and KT3 cells apparently depends on the presence of antisense TK sequences. The plasmid pMD (figure 1) is identical to the antisense genes in KT2 and KT3 cells except that it contains no antisense thymidine kinase sequences. This plasmid was introduced into parental TK^+ cells and transformants resistant to methotrexate were selected as before. In cells resistant to 64 uM methotrexate, there was no detectable reduction in enzyme activity (table 1 and fig. 2). This result suggests that the reduction in TK activity observed in several KT2 and KT3 populations requires antisense TK sequence and is not an artifact of methotrexate selection.

A second set of control experiments indicated that the diminution in thymidine kinase activity in KT2 and KT3 cells is specific. The level of hexokinase activity was determined. Hexokinase catalyzes the first step in glycolysis and was chosen because it is easily assayed. As can be seen in table 2, its activity is essentially unchanged in cell clones which have as much as a nine fold reduction in their level of TK activity.

Double stranded RNA is formed in KT3-B-64 cells

The results presented so far suggest that TK activity can be specifically reduced in cells which express high concentrations of antisense TK RNA. We have examined one cell population, KT3-B-64, in greater detail. If the reduction in the level of thymidine kinase activity is caused by hybridization of sense TK RNA to antisense TK RNA, then double stranded RNA might be present at detectable levels in these cells. Accordingly, an experiment designed to assay for the presence of double stranded TK RNA was performed. Double stranded RNA was prepared by treating RNA samples with RNAses A and T1. The ribonucleases were removed by digestion with proteinase K and the RNAse resistant fraction was denatured and hybridized to an RNA probe for sense TK RNA. After hybridization, unreacted probe was digested by RNAse A and T1 followed by fractionation on a denaturing polyacrylamide gel. As expected, the TK^+ parental cell line and TK^+ cells expressing low levels of antisense RNA contain no detectable double stranded TK RNA (figure 4). However, a band was observed when RNA samples from the TK diminished population were analyzed.

Several control experiments show that the protected band in figure 4, lane 3, is due to the presence of RNA:RNA hybrids rather than contaminating genomic DNA. It is expected that RNA:RNA duplexes will be labile to ribonuclease digestion if they are first heat denatured. They also are expected to be labile to alkaline hydrolysis. As expected, the 225 nucleotide band observed in KT3-B-64 RNA is sensitive to both of these pretreatments (data not shown). Furthermore, RNA hybrids should be resistant to exhaustive DNAase digestion. The band is unaffected by prior treatment of a KT3-B-64 RNA sample with high levels of DNAse I (data not shown). We therefore conclude that the observed band is due to RNA duplexes in KT3-B-64 RNA.

Results of an important control experiment presented in figure 4, lane 1, argue that the RNA duplexes detected in KT3-B-64 RNA are formed in vivo and are not a product of reassociation during RNA extraction and handling. No double stranded RNA was observed when antisense TK RNA was synthesized in vitro and added to parental TK^+ cells at the start of RNA preparation. The amount of antisense TK RNA that was added to P cells was equivalent to the amount of antisense TK RNA in KT3-B-64 cells (50 ng antisense TK per 100 ug of P cell total RNA). This experiment suggests that the double stranded TK RNA in KT3-B-64 cells did not form during RNA preparation and is probably present in vivo. Therefore, the mechanism of inhibition of TK activity may involve the hybridization of antisense TK RNA to sense TK RNA. The size of the double stranded RNA, 225 base pairs, is the same as the size of antisense TK sequences contained in the DHFR-antisense-TK gene of KT3 cells (pMDKT3), suggesting that the entire antisense TK RNA sequence is present in the double stranded RNA.

Double stranded RNA is found in the nucleus but not in the cytoplasm

One possibility is that hybridization of antisense TK RNA with sense TK RNA occurs in the cytoplasm, and reduces thymidine kinase levels by inhibiting translation of TK mRNA. Another possibility is that hybridization occurs in the nucleus, resulting in inhibition of RNA processing or transport. In order to distinguish between these two possibilities, KT3-B-64 cells were fractionated and, from an equivalent number of cells, nuclear and cytoplasmic RNAs were prepared. The amount of double stranded TK RNA in each sample of total RNA was measured using an RNAse protection assay as outlined above. As shown in figure 4, lanes 4 and 5, at least 95% of the double stranded TK RNA is confined to the nuclear

fraction. The amount of hybrid detected in nuclear RNA samples can account for virtually all of the double stranded RNA found in RNA prepared from the entire cell. These observations suggest that RNA hybridization may occur predominantly in the nucleus.

In KT3-B-64 cells, sense TK RNA is present in the nucleus but not in the cytoplasm

In order to further investigate the cause for diminished TK enzyme activity, sense TK RNA in parental cells and KT3-B-64 cells was analyzed. RNA samples were melted and hybridized to an RNA probe for sense TK RNA. Unhybridized probe was digested with RNAse A and T1 and the protected probe was analyzed by fractionation on a nondenaturing polyacrylamide gel. Since the RNA sample was melted before hybridization, this procedure should detect sense TK RNA whether it is single or double stranded. As can be seen in figure 5, KT3-B-64 cells show similar, if not increased, levels of sense TK RNA when compared to parental cells. However, KT3-B-64 cells have eight times less TK activity suggesting that the sense TK RNA is inefficiently utilized. In order to explore the basis of this poor expression, the abundance of the sense TK RNA in the nuclear and cytoplasmic fractions, from equivalent numbers of cells, was analyzed. It was found that in KT3-B-64 cells, greater than 95% of the sense TK RNA is localized to the nucleus, whereas in parental cells, sense TK RNA is primarily cytoplasmic (fig. 5 and data not shown). This result suggests that the decrease in thymidine kinase activity is caused by a lack of sense TK messenger RNA in the cytoplasm where it can be translated.

Only 225 base pairs of TK sequence are in hybrid form in the nuclear RNA of KT3-B-64 cells, because there are only 225 bases of antisense sequence in this particular antisense gene. However, when the nuclear sense TK RNA is examined, the majority is at least 1.1 kilobases in length (this corresponds to the full TK sequence assayed by the RNA probe, and constitutes ninety percent of a full-length TK transcript). Therefore, once hybridized, the free ends of the single stranded sense TK RNA appear not to be degraded much more rapidly than the double stranded "core". Quantitation of sense TK and duplex TK RNA was performed by excising the relevant bands from gels like those shown in figures 4 and 5, and measuring the amount of labelled probe. Given the specific activity of the RNA probe, it was calculated that there are 50 to 100 molecules of duplex thymidine kinase RNA per cell. Within the limits of the measurement, this accounts for all of the sense TK transcript in KT3-B-64 nuclei, and implies that the majority of these RNAs are in hybrid form.

Discussion

The results presented here suggest that high levels of cellular antisense TK RNA can stably reduce TK activity. Antisense TK sequences were inserted into the 3' untranslated region of a DHFR gene. One advantage of this approach is that the level of antisense TK RNA could be increased by selecting cells which are resistant to Mtx. Resistance to high concentrations of Mtx usually involves overexpression of DHFR RNA which, in this experiment, also results in an increase in the amount of linked antisense TK RNA. Accordingly, vectors capable of producing DHFR-antisense-TK RNA were introduced by gene transfer into mouse L cells and Mtx resistant cells were selected. It was found that KT3-B-64 cells, which are Mtx resistant, produced approximately one hundred times more antisense TK RNA than KT3-B-0.3 cells, which are Mtx sensitive. Therefore, selection for Mtx resistance greatly increased the concentration of DHFR antisense TK RNA. As expected, the predominant mechanism of increased expression was amplification of the copy number of the DHFR-antisense-TK genes (data not shown).

Thymidine kinase activity was measured in cells which had differing concentrations of antisense TK RNA. In several independant cell populations, it was found that the level of TK activity diminished as the concentration of antisense TK RNA increased. This result is consistent with the view that high levels of antisense TK RNA may be responsible for specifically inhibiting TK synthesis.

The mechanism by which TK activity is reduced in the presence of antisense RNA was studied by analyzing the types and amounts of RNA present. First, double stranded TK RNA was observed in KT3-B-64 cells, which have decreased TK activity, but was not found in KT3-B-0.3 cells or parental cells, which have full TK

activity (fig 4 and data not shown). The double stranded RNA was predominantly nuclear rather than cytoplasmic. Second, in total RNA, it was observed that KT3-B-64 cells had as much or slightly more sense TK RNA than do P cells. Since KT3-B-64 cells have eight times less TK activity than P cells, it is strongly suggested that the sense TK RNA in KT3-B-64 cells is poorly utilized. In order to investigate this paradox, the abundance of the sense TK RNA was examined in the nuclear and cytoplasmic compartments of P and KT3-B-64 cells. It was found that the sense TK RNA was at least 95% nuclear in KT3-B-64 cells whereas it was at least 75% cytoplasmic in parental cells.

A model for reduction of TK gene activity by antisense RNA

Based on these observations, we prefer the following model for the diminution of TK activity in KT3-B-64 cells. When the concentration of antisense TK RNA becomes sufficiently high, it hybridizes to sense TK RNA. Hybridization occurs primarily in the nucleus and the double stranded RNA is poorly transported to the cytoplasm. As a result, the level of TK enzymatic activity diminishes because relatively less sense TK RNA appears in the cytoplasm where it can be translated.

However, since it is not possible to draw conclusions regarding kinetic processes based soley on steady-state measurements such as those presented here, there are other models which are consistent with the data. For example, the failure to detect cytoplasmic double stranded RNA could be explained by cytoplasmic hybridization followed by either transport of the double stranded RNA

to the nucleus or its rapid degradation. In addition, it is important to recognize that while our results show that the presence of double stranded TK RNA correlates with a decrease in TK activity, they do not show that the double stranded RNA is the cause of the reduction. Another model which formally explains the results is that, for some unknown reason independent of antisense TK RNA, expression from the sense TK gene is reduced and double stranded RNA is formed as a secondary event. We feel that this last model is improbable for the following reasons. The parental cell clone was chosen because of its ability to stably express TK and therefore, spontaneous changes in TK activity in five out of eight independently derived cell populations from that clone are unlikely. Furthermore, random changes in TK activity would not be expected to correlate in independent instances with high levels of antisense TK RNA.

Antisense RNA physiology

The level of accumulation of antisense TK RNA differed depending on the type and amount of antisense TK sequence present in the chimeric gene. Several observations made in the course of this study suggest that, in the pMDKT1 gene, inclusion of a 1.1 kb fragment of antisense TK sequence interfered with the accumulation of its transcripts. First, the initial resistance to methotrexate following DNA mediated gene transfer with pMDKT1 is six times lower than the resistance obtained with pMD (an identical gene in all respects except that it contains no TK sequence). Second, selection for Mtx resistance was much more difficult in cells containing this gene than in cells containing DHFR genes with smaller segments of antisense TK sequence (pMDKT2 and pMDKT3) or no TK

sequence at all (pMD). Finally, when Mtx resistant KT1 cells were obtained, the antisense TK sequences were not expressed at a higher level but appeared as a smear on a Northern blot (data not shown). In contrast, pMDKT2 and pMDKT3 do not show this severe interference with DHFR expression. These genes differ from PMDKT1 in that they lack a central Ava I fragment of the antisense TK gene and produce an RNA which is smaller. The reasons for the difference in DHFR expression between these vectors are incompletely understood. A practical implication is that it may be necessary to identify and remove antisense sequences which interfere with the accumulation of antisense RNA on a case by case basis.

The level of thymidine kinase activity was significantly diminished only when a high level of antisense TK RNA was expressed. The approach taken here was to express antisense TK sequences as part of a DHFR messenger RNA. This allowed us to increase the concentration of antisense TK RNA by selecting for resistance to increasing levels of methotrexate. Although this selection scheme resulted in increased levels of DHFR-antisense-TK RNA in the cytoplasm, a concomitant effect was its increase in the nucleus. The increase in nuclear antisense RNA concentration may have been the cause of the reduction in TK activity since, in this case, it appears that RNA:RNA hybridization occurred in the nucleus. In KT3-B-64 cells, there is a steady state level of at least 5,000 antisense TK RNA molecules per cell, with about 40% localized in the nuclear compartment and 60% in the cytoplasm (figure 3). Thus, in these cells, there is at least 500 fold more antisense TK RNA than sense TK RNA.

The levels of antisense RNA reported here are likely to be significantly higher than the amounts reported in the experiments of Izant and Weintraub (1984). In that study, the cytoplasmic level of antisense TK RNA was equivalent to the

amount of sense TK mRNA. Although the level of nuclear antisense RNA was not determined, if it were high, it could have contributed significantly to the observed effect on TK activity. Nevertheless, the 500 fold difference in cytoplasmic RNAs argues persuasively that there is also a difference in the amount of nuclear antisense RNAs. This difference suggests that, to reduce TK activity, more antisense RNA is required in this report than in the previous one. The origins of this apparent paradox may reside in major differences in experimental approach including: translated versus untranslated antisense RNAs, stable versus transient inhibition and newly microinjected versus stably integrated TK genes.

As shown in table 2, the level of TK activity varied in different cell clones derived from the KT3-B-64 cell population, ranging from 80% in clone KT3-B.1-64 to 11% in clone KT3-B.5-64. One possible explanation for the difference in TK activity is that the levels of nuclear antisense TK RNA may fluctuate between these clones. Another possibility is that, if the position of chromosomes within the nucleus is fixed, RNA hybridization may be more efficient when the antisense TK gene is physically close to the sense TK gene because the local concentration of antisense TK RNA may be greater. Cells resistant to high levels of methotrexate, but with different levels of TK activity, are being studied at the RNA and DNA levels to determine the molecular basis for this difference.

At present, we have only limited data on effects generated by different regions of antisense TK sequence. In two previous studies, one in E. coli and the other in Xenopus oocytes, it was found that antisense RNA was effective only if sequences complementary to the AUG start codon were included (Coleman et al., 1984; Melton 1985). This supports the idea that the principal effect of antisense RNA is exerted at the translational level. Perhaps this is not surprising since, in the Xenopus experiments, both sense and antisense RNAs were injected directly

into the cytoplasm and in E. coli, transcription and translation are tightly coupled. In the animal cell experiments reported here, we have found that antisense sequences derived exclusively from the 3' region of TK protein coding sequence are sufficient to reduce gene activity. Furthermore, a second antisense gene (pMDKT2), which contains sequences spanning the AUG initiation codon in addition to the carboxy terminal sequence of pMDKT3, was neither more nor less effective in diminishing TK activity than was pMDKT3 alone. It remains to be seen what antisense sequences will be most effective in reducing gene activity in this experimental system.

The potential for employing antisense RNA to diminish gene activity

In particular, this report has suggested that antisense TK RNA can significantly diminish thymidine kinase activity; in general, will techniques involving antisense RNA be useful to study gene function? It is possible that the TK gene is a special case and, in general, antisense RNA will be ineffective in inhibiting the expression of other genes. Salient differences between the Herpes virus thymidine kinase gene and other genes include the fact that the TK gene was introduced into the genome by DNA mediated gene transfer, the gene is relatively small (1.3 kb), and it contains no intervening sequences. However, it is interesting to note that, rather than abnormally promote the ease with which antisense RNA might inhibit protein production, these characteristics may actually detract from its effectiveness. For example, the 3' end of TK mRNA, which was hybridized with antisense RNA in these experiments, is the last sequence to be transcribed and must therefore have a somewhat shorter nuclear dwell time than upstream sequences. In addition, the lack of an RNA splicing step may permit more rapid

nuclear transport for TK RNA than for typical spliced RNAs. This line of reasoning suggests that RNA transcribed from 5' sequences of large genes containing many intervening sequences will likely take longer to be exported from the nucleus. Therefore, antisense RNA may be able to hybridize to these sequences more efficiently than to 3' sequences from small genes such as TK.

Another issue regarding the utility of antisense RNA is whether a reduction of tenfold in gene activity is sufficient to cause a hypomorphic phenotype. A ten fold reduction of TK activity did not render these cells resistant to bromodeoxyuridine and therefore TK^- as conventionally defined. Thus, this technique is not yet useful for establishing the complete loss of function of a gene. However, there are some genes which show a phenotypic effect with only a partial loss of function. In genes which exhibit haplo-insufficiency, such as the Notch locus in Drosophila, a twofold difference in gene activity yields an altered phenotype (Wright, 1970). In some cases, a hypomorphic phenotype may reveal subtleties about the function of a gene product which are not always obvious from the null phenotype. For example, a complete lack of the low density lipoprotein receptor grossly affects cholesterol metabolism and is usually fatal in childhood (Goldstein and Brown, 1983). However, a two fold reduction causes one to become prone to premature atherosclerosis and cardiovascular disease. These examples suggests that in mammalian cells, there are likely to be many genes which cause a change in cellular behavior when their activity is reduced ten fold. Future improvements may increase the effectiveness of antisense RNA and may thus increase the range of genes for which an analysis using antisense RNA would be fruitful.

Materials and Methods

Plasmid construction

pMD, a gift from P. Mueller, was constructed from the metallothionein I gene and the DHFR cDNA (Drunam et al., 1980; Nunberg et al., 1980). A Bam HI to Bgl II fragment containing the DHFR cDNA was inserted into the metallothionein I gene which had been previously cleaved by Bgl II, located 9 bp upstream of the ATG start codon, and Alu I, located in the second exon. The Alu I site in the metallothionein gene was converted into a Bgl II site by the addition of Bgl II linkers. pMDKT1 was made by inserting a 1.1 kb Bgl II to Sma I fragment derived from the HSV TK gene into the Bgl II site in pMD after the attachment of a Bam H1 linker to the Sma 1 site. pMDKT2 was made by partially digesting pMDKT1 with Ava I and isolating a plasmid that had the 600 bp Ava I fragment from the middle of the TK sequences deleted. pMDKT3 was made by deleting a 1050bp Ava I fragment from pMDKT1 and recircularizing after treatment with DNA polymerase I Klenow fragment.

Cell culture

L cells were maintained in Dulbecco's modification of Eagles Medium supplemented with 10% calf serum, 10 ug/ml streptomycin sulfate and 60 ug/ml penicillin-K (EC_{10} media). APRT selection used EC_{10} media supplemented with 4 ug/ml azaserine and 15 ug/ml adenine. P cells and all cells derived from it were

grown in EC_{10} supplemented with azaserine, adenine and 1x non-essential amino acids (EC_{10}AAN media). Selection for cells resistant to Mtx was accomplished by the procedure of Alt et al., (1976). Briefly, approximately 2 to 5 x 10^5 cells in a 100 mm plate were placed in EC_{10}AAN medium which contained a two fold higher concentration of Mtx. After about two weeks, the plates contained approximately 10 x 10^5 cells and the process was repeated after splitting the cells appropriately. Since thymidine, the substrate used by TK, is not added to EC_{10}AAN media, the absence or presence of TK activity should not affect the growth of cells under Mtx selective pressure. Cells were transformed by the procedure of Wigler et al. (1977)

Enzyme assays

TK activity was assayed by a modification of the procedure of Wigler (1977). The cells from one 100 mm plate near confluence were washed with cold PBS, scraped in 5 ml of cold PBS and pelleted. The pellet was resuspended in 150 ul of cold TK lysis buffer (10 mM Tris HCl pH 7.8, 10 mM KCl, 1 mM $MgCl_2$, 1mM betamercaptoethanol and 20 uM thymidine). The cell suspension was frozen and rethawed four times in dry ice. To the solution was added 25 ul of 1.0 M KCl and the cells were spun for 10 min. in an eppendorf microfuge at 4^o. On ice, 25 ul of reaction buffer (150 mM Tris HCl pH 7.8, 20 mM $MgCl_2$, 25 mM KCl, 1.5 mM betamercaptoethanol, 60 mM ATP and 75 uCi/ml ^{3}H-thymidine 50 Ci/mmol) was added to 50 ul of cellular supernatant and mixed. Each sample was done in duplicate at several different lysate concentrations. As a negative control, lysis buffer plus KCl was used; neither BSA nor an extract made from L $APRT^-$ TK^- cells gave a signal above the background value using lysis buffer alone. The rack containing the samples was placed in a 37^o water bath for 30 minutes. The rack

was placed on ice and the reactions were terminated by spotting 50 ul on DE81 disks. The disks were washed twice in 1 mM ammonium formate, once in water and once in ethanol for five minutes each with shaking at room temperature. The filters were dried, added to 10 ml hydrofluor and counted in a scintillation counter. Under these conditions, the reactions are linear for both amount of cell extract, from 5 ug to 150 ug total cell protein, and time, up to 90 minutes. The four samples usually gave values to within 10% of each other. The ATP concentration in the reaction mix was increased from 15 mM to 60 mM in order to keep the reaction linear for 90 minutes. Protein content was measured by the method of Bradford (1976). One unit is defined as that amount of enzyme required to convert 1 nmole thymidine into TMP in one minute.

Hexokinase activity was measured as previously described (Walker and Parry, 1966).

Synthesis of RNA probes with SP6 polymerase

A 1.1 kb Bgl 2 to Sma 1 fragment from the HSV TK gene was inserted into the Bgl 2 site of pSP6-4T such that sense TK sequences were transcribed and was therefore used as a probe for antisense TK RNA. A 1.3 kb Eco R1 to Sma 1 fragment from the HSV TK gene was inserted between the Eco R1 and the Sma 1 sites of pSP6-4T and was used as a probe for antisense RNA. Both vectors were linearized at their unique Eco R1 site prior to transcription. Synthesis of RNA probes by SP6 polymerase was done essentially as described (Melton et al. 1984). The transcription reaction did not contain bovine serum albumen. The RNA was labeled with 100 uCi of ^{32}P-GTP diluted with unlabeled GTP at a final concentration of 38 uM. Other nucleotides were present at 500 uM. RNA was

separated from unincorporated nucleotides by sepharose G50 chromatography and DNA was removed by selecting for polyA$^+$ RNA on an oligo-dT column (pSP6-4T contains poly A sequences).

RNAse protection

Total cellular RNA was prepared by the method of Chirgwin et al. (1979) except that the DNA in the initial lysate was sheared by sonication. Nuclear and cytoplasmic RNA was prepared by the method of Favaloro et al. (1980). Typically, 3 x 10^8 cells were lysed in an initial volume of 10 ml and yielded 1 mg of total RNA. RNA was prepared quickly and was never left unfrozen longer than necessary.

RNAse protection experiments were performed as described (Zinn et al., 1984). To a sample of RNA to be tested (1 ug to 50 ug) was added 2 x 10^5 cpm of labeled probe (1 x 10^8 cpm/ug) and precipitated in 2 M ammonium acetate and 1 volume isopropyl alcohol. In the experiments described in figure 5, in which sense TK RNA was assayed in the presence of large amounts of antisense TK RNA, 5 x 10^6 cpm of probe was used in order to insure a mass excess of probe. The dried RNA was dissolved in 30 ul of RNA hybridization buffer (80% formamide, 40 mM PIPES pH 6.4, 400 mM NaCl, 1 mM EDTA), melted at 85^o for 10 minutes and hybridized at 45^o overnight. A 300 ul addition of RNAse buffer (10 mM Tris Cl pH 7.8,5 mM EDTA, 300 mM NaCl) containing 40 ug/ml RNAse A and 2 ug/ml RNAse T1 degraded unhybridized probe after incubation at 30^o for 30 minutes. Control experiments showed that essentially all single stranded probe was digested using one tenth of the above RNAse concentration whereas the double stranded RNA was

resistant to thirty times the above RNAse concentration. The RNAses were inactivated by adding 10 ul of 20% SDS and 50 ug of proteinase K incubating at 37^{o} for 15 minutes. The RNA was extracted with phenol and then precipitated three times from ammonium acetate as before after the addition of 25 ug of carrier tRNA. When the protected probe size was expected to be under 1 kb in length, the samples were dissolved in loading buffer (80% formamide, 40 mM Tris borate pH 7.5), melted at 85^{o} for 10 minutes and loaded onto 4% denaturing polyacrylamide gels (Maxam and Gilbert, 1980). Single stranded fragments were more susceptible to degradation by radiolysis than double stranded fragments. Therefore, for fragments greater than 1 kb in length, neutral polyacrylamide gels, using TAE buffer (50 mM Tris acetate, pH 7.5, 1 mM EDTA), were used to analyze the undenatured RNA:RNA duplex.

References

Alt, F. W., Kellems, R. E., Bertino, J. R. and Schimke, R. T. 1976. Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. J. Biol. Chem. **253:** 1357-1370.

Antman, K. and Livingstone, D. 1980. Intracellular neutralization of SV40 tumor antigens following microinjection of specific antibody. Cell **19:** 627-635.

Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Anal. Biochem. **72:** 248-254.

Burke, B. and Warren, G. 1984. Microinjection of mRNA coding for an anti-golgi antibody inhibits intracellular transport of a viral membrane protein. Cell **36:** 847-856.

Chirgwin, J., Przybyla, A., MacDonald, R. and Rutter, W. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonucleases. Biochem. **18:**5295-5299.

Coleman, J., Green, P. and Inouye, M. 1984. The use of RNAs complemantary to specific mRNAs to regulate the expression of individual bacterial genes. Cell **37:** 429-436.

Drunam, D. M., Perrin, F., F. Gannon, and Palmiter, R. 1980. Isolation and characterization of the mouse metallothionein-1 gene. Proc. Nat. Acad. Sci. U.S.A. **77:** 6511-6515.

Favaloro, J., Treisman, R. and Kamen, R. 1980. Transcription maps of polyoma virus specific RNA: analysis by two dimensional S1 gel mapping. Methods Enzymol. **65:** 718.

Goldstein, J. L. and Brown, M. S. 1983. Familial hypercholesterolemia. in The Metabolic Basis of Inherited Disease. Chapter 33. J. B. Stanbury, J. B. Wyngaarden, D. S. Fredrickson, J. L. Goldstein and M. S. Brown, eds. (New York: McGraw-Hill). pp672-712.

Izant, J. G. and Weintraub, H. 1984. Inhibition of thymidine kinase gene expression by antisense RNA: A molecular approach to genetic analysis. Cell **36:** 1007-1015.

Lin,J. and Feramisco, J. 1981. Disruption of the in vivo distribution of the intermediate filaments in living fibroblasts through the microinjection of a specific antibody. Cell **24:** 185-193.

Mabuchi, I. and Okuno, M. 1977. The effects of myosin antibody on the division of starfish blastomeres. J. Cell Biol. **74:** 251-264.

Maxam, A. M. and Gilbert, W. 1980. Sequencing end labeled DNA with base specific chemical cleavages. Methods Enzymol. **65:** 499-560.

Mayo, K., Warren, R. and Palmiter, R. 1982. The mouse metallothienein I gene is transcriptionally regulated by cadmium following transfection into human or mouse cells. Cell **29:** 99-108.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nuc. Acids Res. **12:** 7035- 7056.

Melton, D. A. 1985. Injected antisense RNAs specifically block messenger RNA translation in vivo. Proc. Nat. Acad. Sci. U.S.A. **82:** 144-148.

Mizuno, T., Chou, M., and Inouye, M. 1984. A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). Proc. Nat. Acad. Sci. U.S.A. **81:** 1966-1970.

Nunberg, J. H., Kaufman, R. J., Chang, A. C. Y., Cohen, S. N. and Schimke, R. T. 1980. Structure and genomic organization of the Mouse dihydrofolate reductase gene. Cell **19:** 355-364.

Pellicer, A., Robins, D., Wold, B., Sweet, R., Jackson, J., Lowy, I., Roberts, J., Sim, G., Silverstein, S. and Axel, R. 1980. Altering genotype and phenotype by DNA mediated gene transfer. Science **209:** 1414-1422.

Pestka, S., Daugherty, B., Jung, V., Hotta, K. and Pestka, R. 1985. Anti-mRNA: specific inhibition of translation of single mRNA molecules. Proc. Nat. Acad. Sci. U.S.A. **81:** 7325-7328.

Roberts, J. and Axel, R. 1982. Gene amplification and genetic correction in somatic cells. Cell **29:** 109- 119.

Rosenberg, U. B., Preiss, A., Seifert, E., Hackle, H., and Knipple, D. C. 1985. Production of phenocopies by kruppel antisense RNA injection into Drosophila embryos. Nature **313:** 703-706.

Simons, R. W. and Kleckner, N. 1983. Translational control of IS10 transposition. Cell **34:** 683-691.

Southern, P. and Berg, P. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Molec. and Appl. Genet. **1:** 327-341.

Stark, G. R. and Wahl, G. M. 1984. Gene amplification. Ann. Rev. Biol. **53:** 447-492.

Tomizawa, J., Itoh, T., Selzer, G. and Som, T. 1981. Inhibition of ColE1 RNA primer formation by a plasmid- specified small RNA. Proc. Nat. Acad. Sci. U.S.A. **78:** 1421-1425.

Tomizawa, J. 1984. Control of ColE1 plasmid replication: the processes of binding of RNA I to the primer transcript.Cell **38:** 861-870.

Walker, D. and Parry, M. 1966. Glucokinase. Meth. Enzymol. **9:** 381-388.

Warner, A. E., Guthrie, S. C. and Gilula, N. B. 1984.Antibodies to gap-junctional protein selectively disrupt junctional communication in the early amplhibian embryo. Nature **311:** 127-132.

Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Cheng, Y. and Axel R. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell **11:** 223-232.

Wright, T. R. F. 1970. The genetics of embryogenesis in Drosophila. Adv. Genet. **15:** 261-395.

Yamaizumi, M., Uchida, T., Okada, Y. and Furosawa, M. 1978. Neutralization of diphtheria toxin in living cells by microinjection of antifragment A contained within resealed erythrocyte ghosts. Cell **13:** 227-232.

Zinn, K., Dimaio, D. and Maniatis, T. 1983. Identification of two distinct regulatory regions adjacent to the Human beta-interferon region. Cell **34:** 865-879.

Table 1. TK activity is diminished in cell lines that are resistant to Mtx

cell line	vector	Mtx [a]	TK [b]	n [c]
P	NA	.01	1.0	(>4)
KT0-A	pMD	.30	1.07	(2)
KT0-A-64	pMD	64	1.22	(2)
KT0-B	pMD	.30	1.13	(2)
KT0-B-64	pMD	64	.95	(2)
KT0-C	pMD	.30	1.05	(2)
KT0-C-64	pMD	64	.90	(1)
KT1-B-.05	pMDKT1	.05	1.0	(2)
KT1-B-2	pMDKT1	2	1.0	(4)
KT2-B-12	pMDKT2	12	.37	(2)
KT2-B-100	pMDKT2	100	.19	(>4)
KT2-D-9	pMDKT2	9	1.06	(>4)
KT2-D-100	pMDKT2	100	.17	(>4)
KT2-E-4	pMDKT2	4	.57	(>4)
KT2-E-24	pMDKT2	24	.52	(>4)
KT3-B-.15	pMDKT3	.15	1.03	(2)
KT3-B-64	pMDKT3	64	.15	(>4)
KT3-C-.15	pMDKT3	.15	.99	(2)
KT3-C-64	pMDKT3	64	.56	(3)
KT3-D-.15	pMDKT3	.15	1.04	(2)
KT3-D-64	pMDKT3	64	.18	(>4)
KT3-E-.15	pMDKT3	.15	1.07	(2)
KT3-E-64	pMDKT3	64	.80	(3)

a. Concentration of Mtx (in uM) to which cells are resistant.

b. Relative amount of TK activity normalized to P cells. The TK activity in P cells was 1.58×10^{-2} U/mg protein.

c. number of experiments.

Table 2. TK activity in cell clones with amplified antisense genes

cell line	vector	type[a]	Mtx[b]	TK[c]	hexokinase[d]
P	NA	C	.01	1.0	
KT3-B-64	pMDKT3	P	64	.15	
KT3-B.1-64	pMDKT3	C	64	.78	
KT3-B.2-64	pMDKT3	C	64	.44	
KT3-B.3-64	pMDKT3	C	64	.47	
KT3-B.5-64	pMDKT3	C	64	.11	1.06
KT3-B.6-64	pMDKT3	C	64	.16	1.03
KT3-B.7-64	pMDKT3	C	64	.18	.94
KT3-B.8-64	pMDKT3	C	64	.18	.84
KT3-B.9-64	pMDKT3	C	64	.35	

a. C = clone. P = population

b. Concentration of resistance to Mtx in uM.

c. Relative amount of TK activity normalized to P cells.

d. Relative amount of hexokinase activity normalized to P cell.

Figure Legends

fig. 1. Restriction map of antisense vectors. The maps of pMD, pMDKT1, pMDKT2 and pMDKT3 are shown. DHFR sequences are denoted by a white box. Antisense TK sequences are represented by a black box. The matallothionein promoter is indicated by an arrow. R = Eco R1, S = Sac 1, A = Ava 1, H = Hind III, B = Bam H1, P = Pst 1.

fig. 2. TK activity in KT3 cells resistant to various levels of methotrexate. Resistance to given concentrations of methotrexate, in uM units, is plotted logarithmically along the x axis. Cell lines resistant to various levels of methotrexate were assayed for TK activity. The relative levels of TK activity are shown along the y axis with parental TK^+ cells (1.58×10^{-2} U/mg protein) defined as having 100% activity. Assays were performed in triplicate and are accurate to within 10%. KT3-B cells are represented by dark circles and KT3-D cells are represented by open circles.

fig. 3. Antisense TK RNA analyzed by RNAse protection. Radiolabelled sense TK RNA was synthesized in vitro and hybridized to RNA samples. Unhybridized probe was digested with RNAse A and T1. The protected probe, due to the presence of cellular antisense TK, was melted and fractionated on a 4% denaturing polyacrylamide gel. The cellular origin of the RNAs appear at the top. A. Total RNA (2.5 ug) from KT3-B-0.3, KT3-B-2 and KT3-B-64, representing cells before,

during and after methotrexate selection, was analyzed. A 50 fold exposure, to show the antisense RNA before amplification, is shown at the right. B. Starting from an equivalent number of KT3-B-64 cells, RNA was prepared from nuclear, cytoplasmic and total cellular fractions. Equal portions of each RNA fraction, corresponding to 2.5 ug of total cellular RNA, was analyzed for the presence of antisense RNA.

fig. 4. Detection of double stranded TK RNA. RNA was digested with RNAse A and T1 leaving only double stranded RNA. The RNases were removed by digestion with proteinase K and phenol extraction. The protected RNA was hybridized to radiolabelled antisense TK RNA synthesized in vitro. Unhybridized probe was digested with RNAse A and T1. Protected probe, due to hybridization with the sense TK strand of the protected double stranded RNA, was melted and loaded on a 4% sequencing gel. The origin of the RNAs appears at the top. Con. refers to a control experiment in which 50 ng of antisense TK RNA, synthesized in vitro, was added to P cells containing 100 ug total RNA immediately before RNA isolation. P refers to parental TK^+ cells. The Con., P and KT3-64 lanes used 25 ug of RNA. From an equivalent number of cells, cytoplasmic and nuclear RNA from KT3-B-64 cells was prepared. Equal fractions of each RNA preparation, corresponding to 25 ug of total cellular RNA, was analyzed for the presence of double stranded RNA.

fig. 5. Sense TK RNA analyzed by RNAse protection. RNA was hybridized to radiolabelled antisense TK RNA synthesized in vitro. RNAse protection was performed as described in fig. 2. Origin of the RNA appears at the top. left: 12 ug of RNA was used. right: From an equivalent number of KT3-B-64 cells, RNA from nuclear, cytoplasmic and total cellular fractions was prepared. An equal portion of each RNA sample, corresponding to 12 ug of total cellular RNA, was analyzed for sense TK RNA.

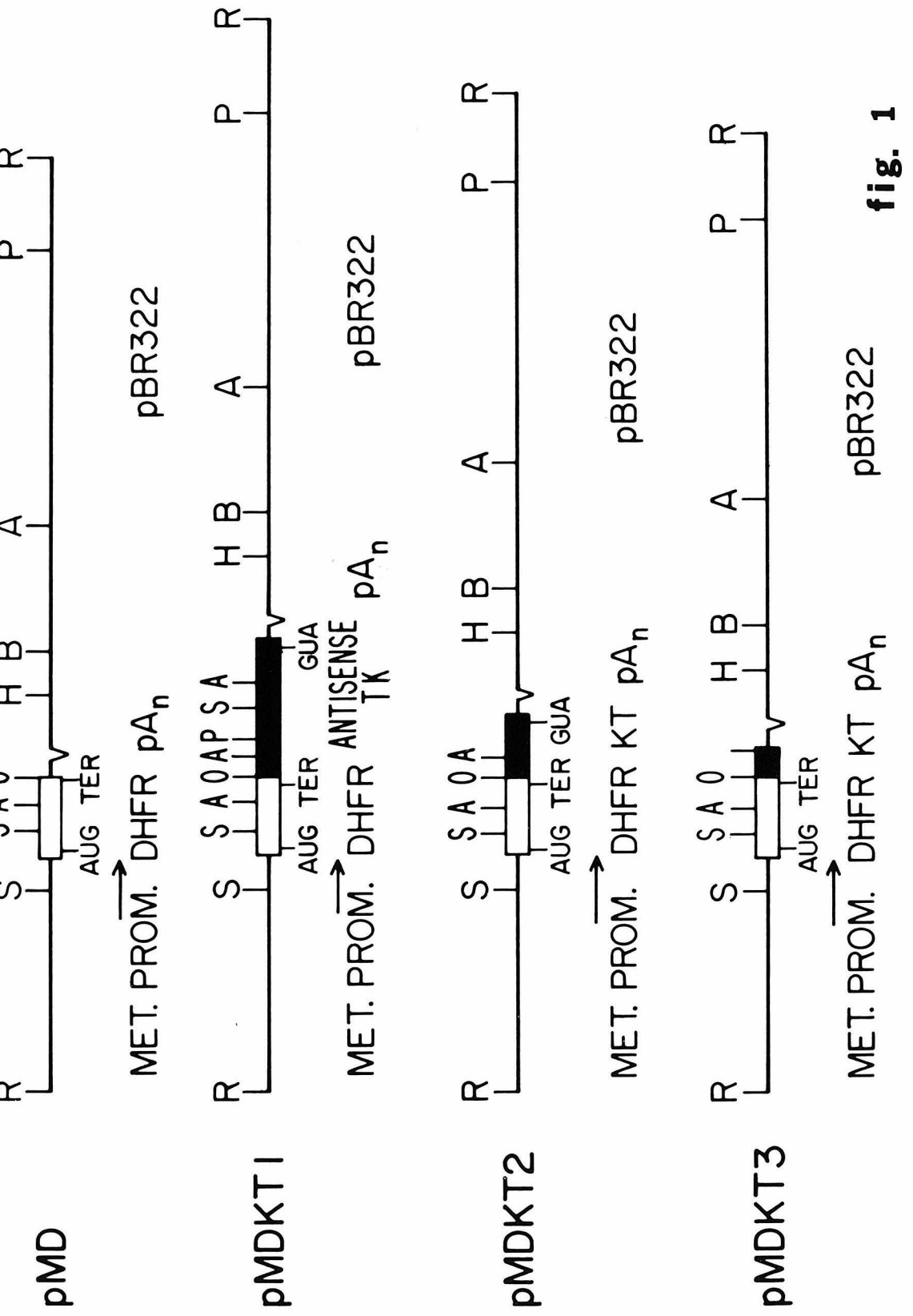
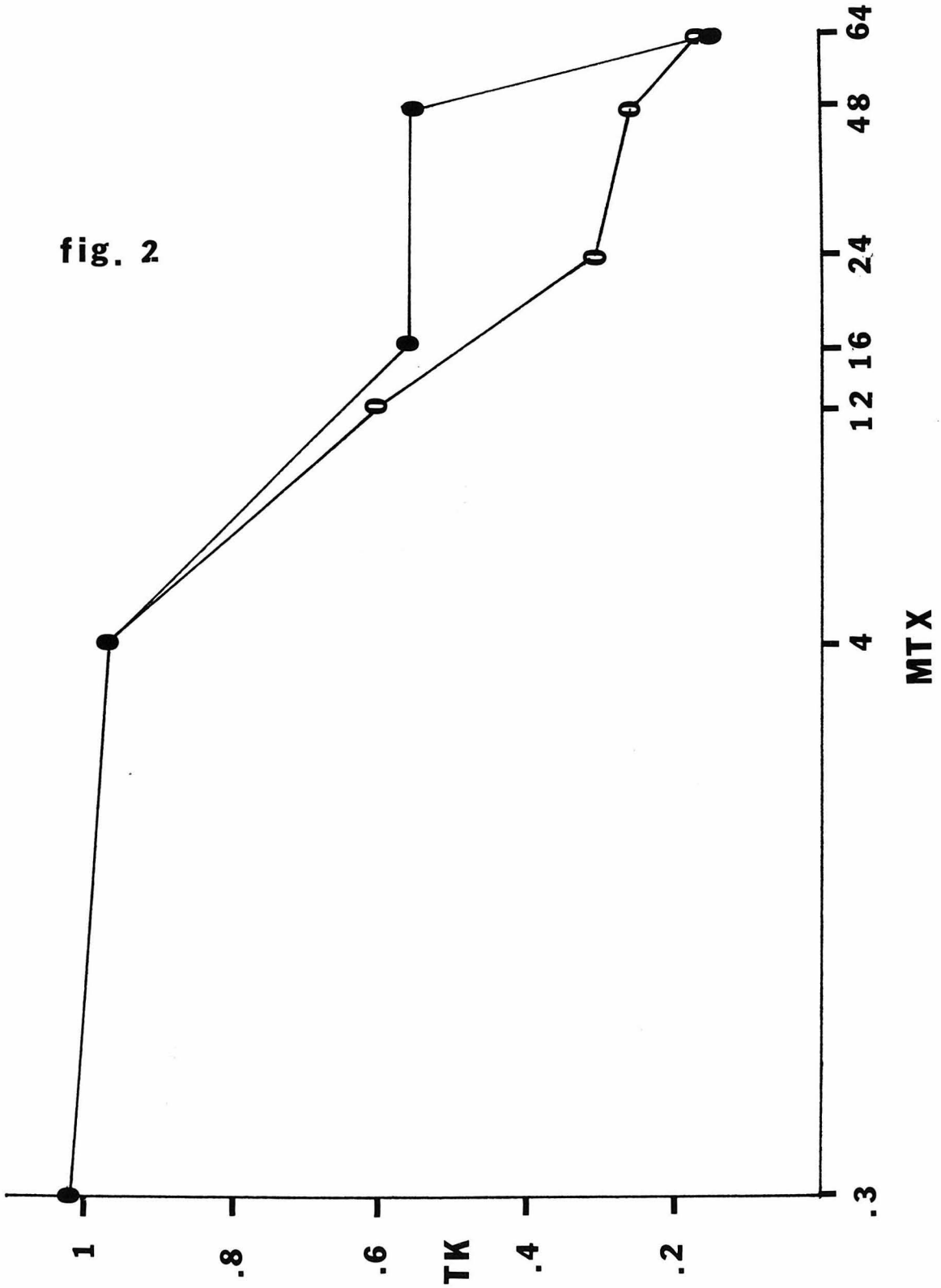

pMD
R
S
S A 0
AUG TER
H B
A
P
R
MET. PROM. DHFR pA$_n$
pBR322
pMDKT1
R
S
S A 0 A P S A
AUG TER
GUA
H B
A
P
R
MET. PROM. DHFR ANTISENSE TK pA$_n$
pBR322
pMDKT2
R
S
S A 0 A
AUG TER GUA
H B
A
P
R
MET. PROM. DHFR KT pA$_n$
pBR322
pMDKT3
R
S
S A 0
AUG TER
H B
A
P
R
MET. PROM. DHFR KT pA$_n$
pBR322

fig. 1

fig. 2



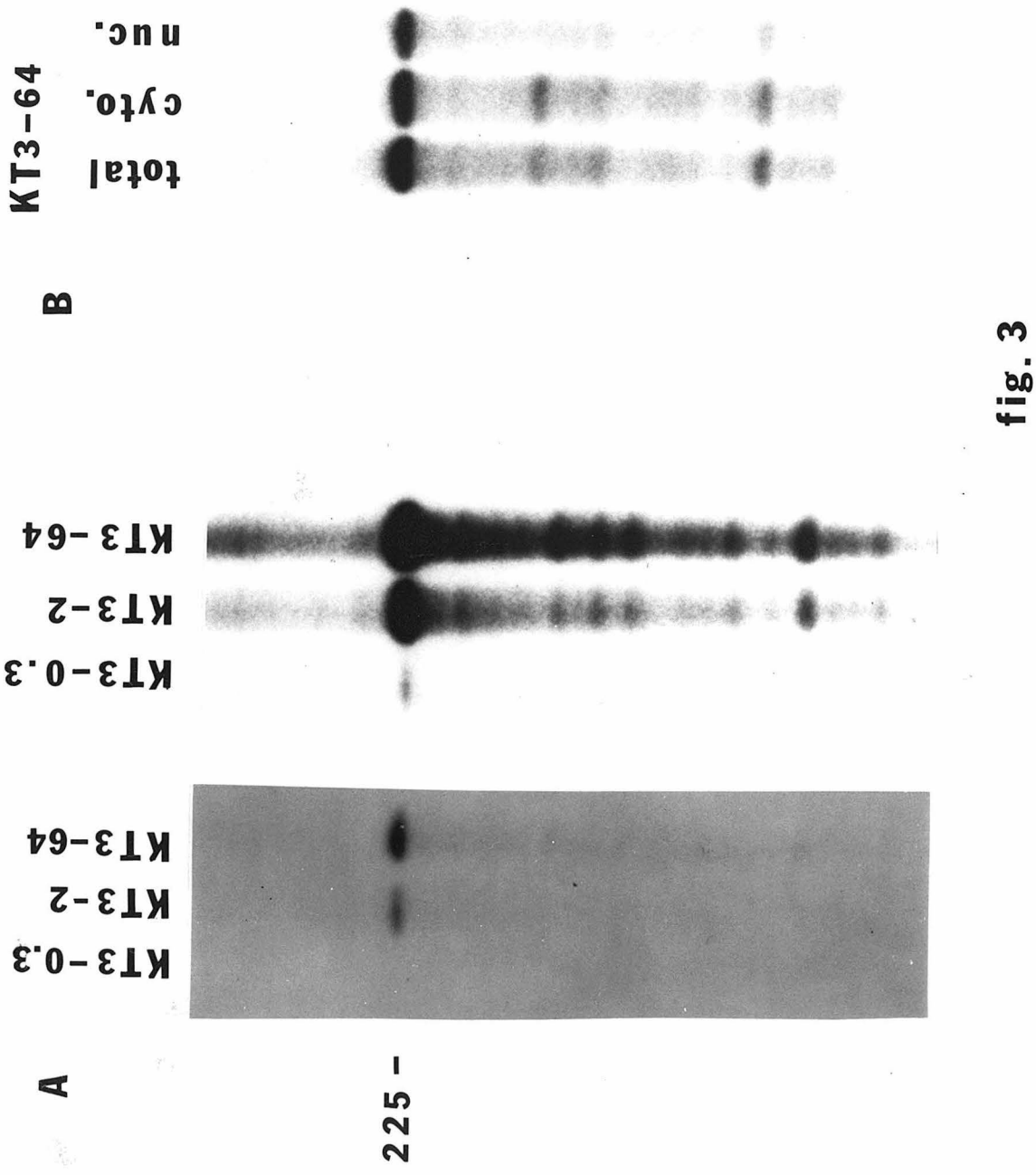
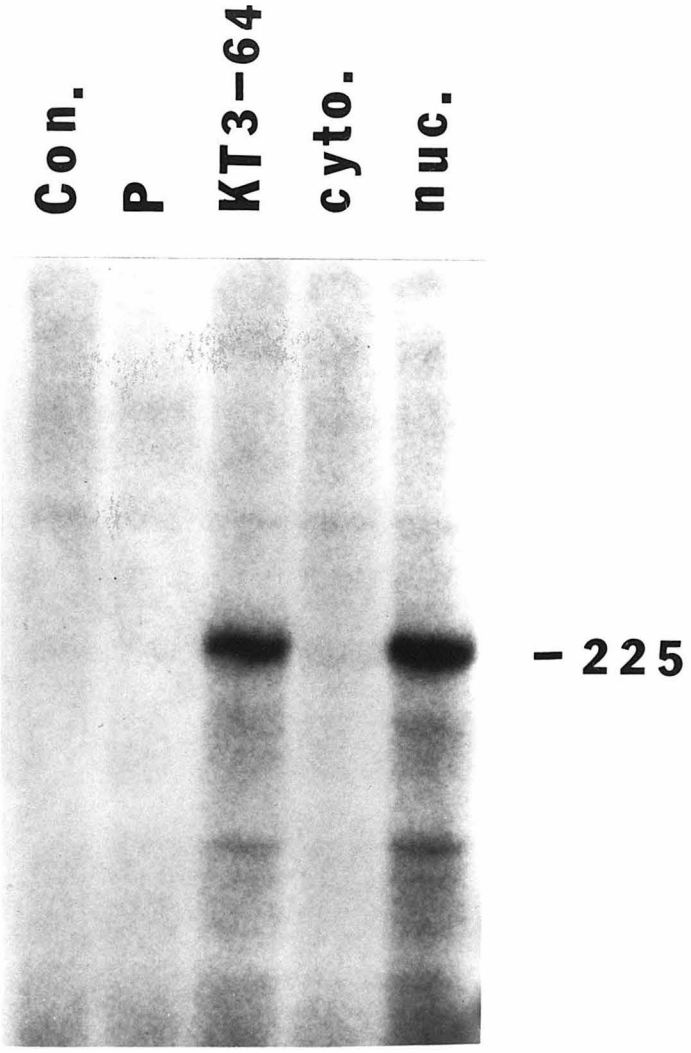

A
225 -
KT3-0.3
KT3-2
KT3-64
B
KT3-0.3
KT3-2
KT3-64
KT3-64
total
cyto.
nuc.

fig. 3


Con.
P
KT3-64
cyto.
nuc.
-225

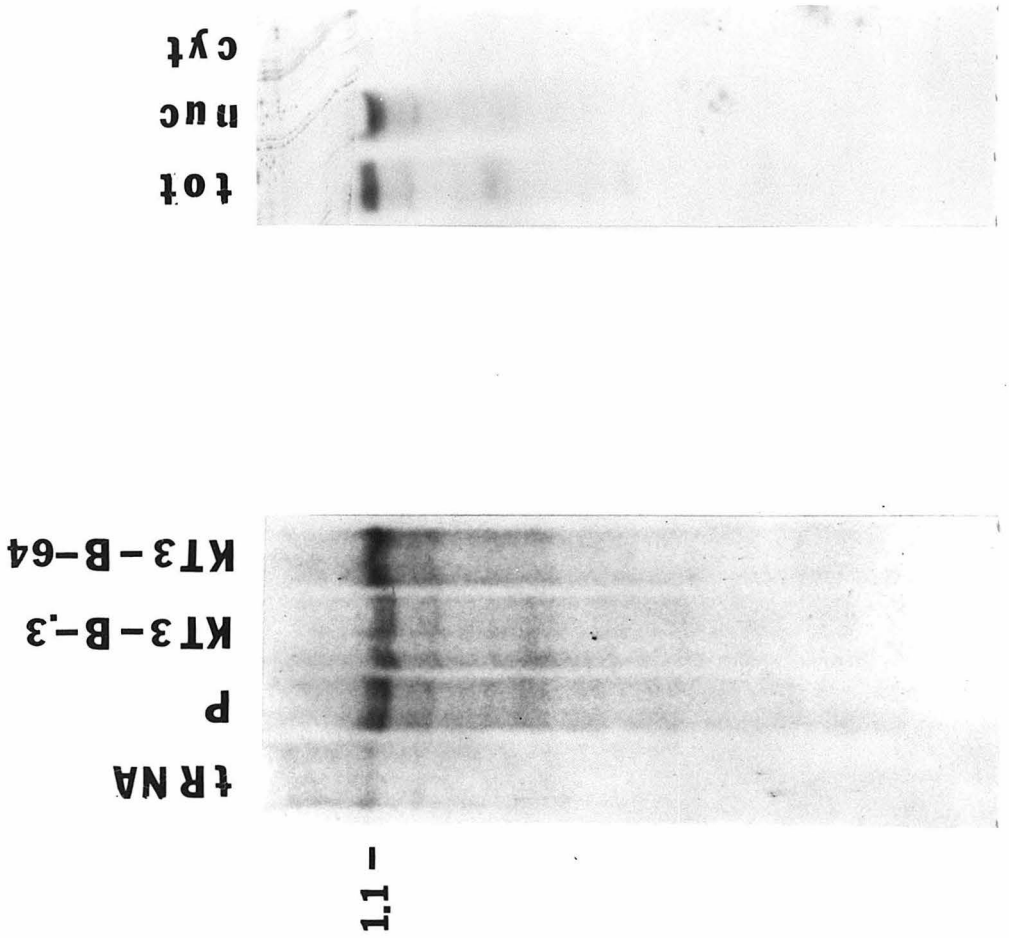


fig. 5