Genetic Recombination in B Cell Ontogeny: A Mosaic of Complementary Systems

by

William Don Harriman

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Immunology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

This work is dedicated to my sons, their mother, and their grandparents, for keeping my energy high, and to my mentor, Professor Matthias Wabl, for focusing this energy toward a productive end.

Preface

This thesis work has been done under the supervision and guidance of Professor Matthias Wabl in the Department of Microbiology and Immunology at the University of California, San Francisco.

Thesis committee members:

Matthias Wabl, Chair
Matija Peterlin
Mark Feinberg

Qualifying exam committee members:

Joel Goodman, Chair
Geoff Davis
Bill Seaman
Tris Parslow
Phil Coffino

I would like to thank current and former members of the Wabl Lab, including Hans-Martin Jäck, Jürgen Bachl, Marilia Cascalho, and Heiner Völk, for many insightful discussions, and for providing a friendly working environment.

Also, a special thanks to my father, Robert D. Harriman, for technical assistance on the production of this document and others.

Abstract

The body of this thesis work entails the experimental demonstration that terminal deoxynucleotidyl transferase, TdT, can add nucleotides at many sites in the genome, potentially at any DNA double strand break. Through gene transfer, fibroblasts were made to express the normally lymphocyte-specific polymerase. The expression was controlled through an inducible promoter, thus allowing reasonable interpretation of subtle differences, since the factor of clonal variation is minimized. It is shown that in the presence of the chain terminator ddA, TdT can block the short-term repair of radiation induced double strand breaks. This implies that TdT can interact with repair related factors and provides an additional link between the overlapping systems of lymphocyte-specific recombination and general DNA repair.

Table of Contents

Figures

Introduction

The role of a B cell in the mammalian immune response is mediated through its antigen-specific molecule: the antibody, or immunoglobulin. Antibodies can neutralize antigens directly through binding and immobilizing, or by marking antigen-associated cells or particles for destruction by other components of the immune system. When displayed as a membrane protein at the surface of a B cell, the antibody molecule serves as the antigen receptor, through which functional signaling can occur. Antigen-bound antibody can also be internalized, processed, and presented in the context of MHC to T cells, which can in turn provide signals to augment the B cell response. Antibodies, along with MHC molecules and T cell antigen receptors, comprise the three classes of molecules used by the immune system to specifically recognize antigens. Of these three, antibodies are distinguished by the widest range of antigenic structures they can recognize, by the greatest ability to distinguish between antigens, and by the greatest strength of binding to antigen.

All immunoglobulins have a common core structure of two identical light chains and two identical heavy chains. One light chain is attached to each heavy chain, and the two heavy chains are attached to each other. They can be secreted as monomers, dimers, trimers, or pentamers, depending on the class, or isotype. The major isotypes are IgM, IgD, IgG, IgA and IgE, which are encoded by the μ, δ, γ, α and ε heavy chain genes, respectively. Light chains have two isotypes, κ and λ. The portion of the immunoglobulin molecule which defines its isotype is designated the constant region, due to its limited diversity. A single B cell produces antibodies of only one class at a time (with the exception of IgM+IgD), but can change the isotype of immunoglobulin it

produces, a process known as immunoglobulin class switching. Both heavy and light chains also contain highly variable regions which associate to generate the antigenic specificity of the immunoglobulin molecule.

In the germline of vertebrates there are no functional genes encoding the polypeptide chains of immunoglobulins, only gene segments. During B cell differentiation, selected segments are joined to yield the functional heavy and light chain genes. Two types of DNA rearrangements, each of which may result in either deletions or inversions, join together gene segments that may be far away from each other in the germline (1, 2). The first type, which is utilized in the generation of functional variable region genes, is site specific and is mediated by a recombinase that recognizes known signal sequences flanking the exon sequences to be joined (3). The second type is rather imprecise; the recombination break points vary widely within an intron. Because the intron sequences will be deleted by RNA splicing, this imprecision does not impair the final outcome of the process. Immunoglobulin heavy chain class switching results from this second type of rearrangement. The gene encoding the immunoglobulin heavy chain is reconstructed so that the gene segment encoding a given constant region is replaced by another constant region gene segment located further downstream in the locus. Therefore, a B cell clone can alter the effector function of the immunoglobulin produced (dictated by its isotype) without affecting its specificity (dictated by its variable region) (4).

Much is known about the nature of the DNA sequences which can undergo either VDJ- or switch- rearrangement, and of the products which are generated through such recombination. But surprisingly, very little is known about the molecular components, referred to as the VDJ- or switch-

recombinase, that catalyze these reactions. Both of these recombinases must perform several of the same functions, including binding to DNA, cutting it, and ligating it. The two enzymes obviously differ in their DNA recognition sequences. For the VDJ-recombinase, these are the heptamer-nonamer recombination signal sequences, which flank the gene segments (3). For the switch recombinase the recognition sequences are not known, but the simplest assumption is that they are part of the switch region. Since the cutting and ligating activities are common to both enzymes, they may even be executed by the same proteins, but the binding component must be different.

One of the enzymatic properties of both recombinases, DNA ligation, is shared with a more general, non-lymphocyte-specific, activity: repair of DNA double strand breaks. This common feature seems to go beyond analogous activities alone since there is evidence that some component(s) may actually be utilized by both systems. In fibroblasts isolated from scid mice, which are defective in the ligation step of VDJ recombination, there is also a defect DNA double strand break repair, as evidenced by increased sensitivity to ionizing radiation (5, 6). Additionally, there are other mutants (xrs and XR-1), belonging to separate complementation groups from each other and from scid, which show both radiosensitivity and defects in VDJ recombination (7-10). Other radiosensitive cell lines that can rejoin double strand breaks (such as ataxia-telangiectasia cells) have normal abilities to carry out VDJ recombination (11). Thus, a recombination defect is not characteristic of all radiosensitive mutants but is associated specifically with double strand break rejoining. These results suggest a mechanistic overlap between these two processes and indicate that the site-specific double strand breaks introduced during VDJ recombination may be processed, at least initially, by a mechanism similar to that operating on

double strand breaks produced by radiation. They also give rise to a slightly more complicated picture of the recombinase, where some components may be sequence and tissue specific while others are shared between systems.

One of the lymphocyte specific factors which is known to act at the site of VDJ recombination is the template independent DNA polymerase terminal deoxynucleotidyl transferase (TdT) (12-15). It is not required at any step of the VDJ recombination process, but rather is an accessory protein which, when present, can add nucleotides at the recombination breakpoint and generate additional diversity for the newly rearranged VDJ gene segment. These added nucleotides are referred to as "N regions". TdT is expressed only very early in B and T cell ontogeny, and is in fact downregulated by the time pre-B cells rearrange their light chain genes, resulting in the lack of N regions at these loci (16). Accordingly, mouse pre-B cells transfected with a TdT gene have been reported to contain N region insertions in a substrate containing gene segments of an immunoglobulin light chain (17). Also, fibroblasts transfected with genes expressing RAG-1, RAG-2, and TdT were shown to recombine a transfected substrate and in the process to form N regions (18, 18a). If RAG-1 and RAG-2 were transfected without TdT , there were no N regions. Finally, TdT deficient mice have recently been produced, and they display no N regions in their rearranged immunoglobulin heavy chain genes (19).

In vitro, TdT can add nucleotides to any 3'OH group on single or double strand DNA, but the chemical nature of its in vivo substrate remains a mystery. Does TdT seek the appropriate termini independently of the recombinase, or is it actively recruited as part of a larger enzyme complex (i.e. the recombinase)? Does it require assistance from another protein in order to generate a suitable

substrate? Some researchers report an endonucleolytic nicking activity associated with TdT, in which case it would be able to generate its own 3'OH substrate (20). Others have even suggested that TdT possesses complete recombinase activity (21). It is unlikely that severed DNA ends simply sit free for extended periods of time without being bound by a protein or covalently modified in some way (22). One model of VDJ recombination involves the formation of a hairpin structure at the coding ends prior to endonucleolytic cleavage and ligation (23). It has further been suggested that failure to open and resolve such a hairpin structure is the cause of the scid defect, thereby implying that such a structure may be relevant in radiation induced double strand breaks (24). Broken DNA ends, covalently modified or otherwise, may also be bound by an end-binding protein such as Ku80 (25). Ku80 is the product of the XRCC5 gene, which is defective in the radiation sensitive xrs mutants (26).

The question to be addressed in this study is whether TdT activity can be directed to radiation induced double strand breaks in vivo. If so, it would suggest that the radiation induced breaks are not immediately modified in a way which precludes nucleotide addition by TdT, or alternatively, that TdT associates with a protein or complex which can transiently generate a free 3'OH at the DNA terminus, thereby creating a substrate for polymerization prior to ligation. In the former scenario, under conditions of extreme stress (i.e. irradiation) the cellular repair machinery may become exhausted and be unable to immediately "cap" all termini, which would provide a window for TdT mediated nucleotide addition. However, this would not account for the for the fact that TdT can add nucleotides during normal VDJ recombination where the repair machinery presumably would not be exhausted. So in either case, TdT

would have to penetrate a blocked terminus in order to find a suitable substrate. It would be unlikely that TdT would have activity at two separate steps of VDJ recombination, therefore if TdT can add nucleotides at double strand breaks in vivo, it suggests that during VDJ recombination TdT is active at the step which is shared with DNA repair: ligation. Assuming that the site specific components of the VDJ recombination machinery function at the early steps of recombination, binding and cutting, and that the final step of ligation is left to the general repair machinery, there would be the interesting situation of a lymphocyte specific polymerase, TdT, necessarily associating with general repair machinery in order to execute its known physiological function.

During the course of this work it became necessary to modify some standard laboratory techniques in order to facilitate the research described herein. While technical modifications are accepted as matter of course in many scientific endeavors, they can sometimes represent such significant improvements over what was previously accepted that they warrant communication to the scientific community. Hopefully, the two technical papers included in this dissertation are examples of such advances. The first paper describes an assay for terminal deoxynucleotidyl transferase activity which combines steps for specific purification and subsequent enzymatic quantification. It should prove useful to investigators who study the function of TdT, especially to those who perform gene transfer experiments. The second technical paper describes a technique for the quantification of damaged cellular DNA which has been resolved through CHEF gel electrophoresis, a technique that has emerged as an important means for directly assessing DNA double strand breaks and their repair. Previously described quantification methods are laborious and/or require expensive equipment. This new

technique is a vast improvement in both respects, and will ostensibly be a boon to expanding research in the exciting field of DNA repair. Both techniques were employed in the paper which is the culmination of my thesis work, showing that TdT is enzymatically active upon DNA double strand breaks in vivo.

Also included in this dissertation is a scientific review of current knowledge in the area of immunoglobulin class switch recombination, which has obvious relevance as background to my current work. Finally, I have included as an Appendix an article, "Can we outlive Methuselah?", which was written with a popular audience in mind. It is not scientifically rigorous, but it nicely outlays a theory of aging based upon the accumulation of mutation. In fact, at the outset, my thesis work was built around the prospect of TdT potentially increasing the cellular mutation rate, thereby accelerating the organismal rate of aging. A mutator phenotype can be achieved in various ways, from expression of an error-prone polymerase to aberrant expression of any of a possibly great number of repair-related genes. The latter case seems to applicable to colon cancer, where defective mismatch repair genes have been implicated in the elevation of mutation rates (27, 28). Nonetheless, to date, TdT is the only known mammalian protein whose presumed function is to cause, not prevent, mutations.

References

1. Shimizu, A., Honjo, T. 1984. Immunoglobulin class switching. Cell 36: 801-03

2. Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature 302: 575-81

3. Schatz, D. G., Oettinger, M. A., Schlissel, M. S. 1992. V(D)J recombination: molecular biology and regulation. Annu. Rev. Immunol. 10: 359-83

4. Harriman, W., Volk, H., Defranoux, N., Wabl, M. 1993. Immunoglobulin class switch recombination. Annu. Rev. Immunol. 11: 361-84

5. Fulop, G. M., Phillips, R. A. 1990. The scid mutation in mice causes a general defect in DNA repair. Nature 347: 479-82

6. Hendrickson, E. A., Qin, X. Q., Bump, E. A., Schatz, D. G., Oettinger, M., Weaver, D. T. 1991. A link between double-strand break-related repair and V(D)J recombination: the scid mutation. Proc. Natl. Acad. Sci. U.S.A. 88: 4061-5

7. Jeggo, P. A., Tesmer, J., Chen, D. J. 1991. Genetic analysis of ionising radiation sensitive mutants of cultured mammalian cell lines. Mutat. Res. 254: 125-33

8. Jeggo, P. A., Hafezparast, M., Thompson, A. F., Broughton, B. C., Kaur, G. P., Zdzienicka, M. Z., Athwal, R. S. 1992. Localization of a DNA repair gene (XRCC5) involved in double-strand-break rejoining to human chromosome 2. Proc. Natl. Acad. Sci. U.S.A. 89: 6423-7

9. Whitmore, G. F., Varghese, A. J., Gulyas, S. 1989. Cell cycle responses of two X-ray sensitive mutants defective in DNA repair. Int. J. Radiat. Biol. 56: 657-65

10. Pergola, F., Zdzienicka, M. Z., Lieber, M. R. 1993. V(D)J recombination in mammalian cell mutants defective in DNA double-strand break repair. Mol. Cell. Biol. 13: 3464-71

11. Hsieh, C. L., Arlett, C. F., Lieber, M. R. 1993. V(D)J recombination in ataxia telangiectasia, Bloom's syndrome, and a DNA ligase I-associated immunodeficiency disorder. J. Biol. Chem. 268: 20105-9

12. Chang, L. M., Bollum, F. J. 1971. Deoxynucleotide-polymerizing enzymes of calf thymus gland. V. Homogeneous terminal deoxynucleotidyl transferase. J. Biol. Chem. 246: 909-16

13. Chang, L. M., Brown, M., Bollum, F. J. 1973. Induction of DNA polymerase in mouse L cells. J. Mol. Biol. 74: 1-8

14. Bollum, F. J., Brown, M. 1979. A high molecular weight form of terminal deoxynucleotidyl transferase. Nature 278: 191-2

15. Nakamura, H., Tanabe, K., Yoshida, S., Morita, T. 1981. Terminal deoxynucleotidyltransferase of 60,000 daltons from mouse, rat, and calf thymus. Purification by immunoadsorbent chromatography and comparison of peptide structures. J. Biol. Chem. 256: 8745-51

16. Sakano, H., Huppi, K., Heinrich, G., Tonegawa, S. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 280: 288-94

17. Landau, N. R., Schatz, D. G., Rosa, M., Baltimore, D. 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. 7: 3237-43

18. Kallenbach, S., Doyen, N., Fanton, d. M., Rougeon, F. 1992. Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. Proc. Natl. Acad. Sci. U.S.A. 89: 2799-803

18a. Komori, T., Okada, A., Stewart, V., Alt, F. W. 1993. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes [published erratum appears in Science 1993 Dec 24;262(5142):1957]. Science 261: 1171-5

19. Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., Mathis, D. 1993. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science 261: 1175-8

20. Spigelman, Z., Duff, R., Beardsley, G. P., Broder, S., Cooney, D., Landau, N. R., Mitsuya, H., Ullman, B., McCaffrey, R. 1988. 2',3'-Dideoxyadenosine is selectively toxic for TdT-positive cells. Blood 71: 1601-8

21. Kunkel, T. A., Gopinathan, K. P., Dube, D. K., Snow, E. T., Loeb, L. A. 1986. Rearrangements of DNA mediated by terminal transferase. Proc. Natl. Acad. Sci. U.S.A. 83: 1867-71

22. Lieber, M. R. 1992. The mechanism of V(D)J recombination: a balance of diversity, specificity, and stability. Cell 70: 873-6

23. Roth, D. B., Menetski, J. P., Nakajima, P. B., Bosma, M. J., Gellert, M. 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell 70: 983-91

24. Harrington, J., Hsieh, C. L., Gerton, J., Bosma, G., Lieber, M. R. 1992. Analysis of the defect in DNA end joining in the murine scid mutation. Moll. Cell. Biol. 12: 4758-68

25. Zhang, W. W., Yaneva, M. 1992. On the mechanisms of Ku protein binding to DNA. Biochem. Biophys. Res. Commun. 186: 574-9

26. Taccioli, G. E., Gottlieb, T. M., Blunt, T., Priestley, A., Demengeot, J., Mizuta, R., Lehmann, A. R., Alt, F. W., Jackson, S. P., Jeggo, P. A. 1994. Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. Science 265: 1442-5

27. Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D., Perucho, M. 1993. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 363: 558-61

28. Bhattacharyya, N. P., Skandalis, A., Ganesh, A., Groden, J., Meuth, M. 1994. Mutator phenotypes in human colorectal carcinoma cell lines. Proc. Natl. Acad. Sci. U.S.A. 91: 6319-23

CHAPTER 1

IMMUNOGLOBULIN CLASS SWITCH RECOMBINATION

as appearing in

ANNUAL REVIEW OF IMMUNOLOGY
1993

Abstract

A B lymphocyte that produces the immunoglobulin heavy (H) chain m may switch to the production of another H chain class: γ, ε, or α. Since the new heavy chain retains the original variable (V) region, antigenic specificity is maintained. The switch is accompanied by a large deletion of DNA at the heavy chain locus. In order to explain how this deletion is generated, three models have been proposed: recombination between homologs, unequal sister chromatid exchange, and looping out and deletion. While none of the predicted other recombination products of the first two models have been found, both byproducts of looping out—inversions and circular DNA—have been isolated. Thus looping out and deletion appears to be the appropriate model to explain the genetic events leading to the immunoglobulin heavy chain class switch. It is thought that one requirement for switching may be transcription of the constant (C) region to which the cell switches. The switch rearrangement is catalyzed by a switch recombinase, and the isolation of the components of this putative enzme system is in progress. Although the switch deletion is an accepted fact, the discussion is enlivened by scenarios for switching without DNA rearrangement; such suggestions include processing at the RNA level and trans-splicing.

Introduction

According to Karl Popper (1), a scientific theory cannot be verified; it can only be falsified. This view devalues "positive" evidence for a model and stresses the virtues of rejecting a model. But science is a social enterprise (2), and rejecting a model does not win one any friends unless the model is one's own, while positive evidence need not make one any enemies. Furthermore, it is more gratifying to legitimize one model than to banish others. In this review we will explore both routes for selecting a model to explain the events in the genome of a B lymphocyte that result in a switch of the immunoglobulin heavy chain class produced by that cell.

In the mouse, the early immune response is dominated by the expression of IgM, which contains the heavy (H) chain μ; later there appear IgG_3, IgG_1, IgG_{2b}, IgG_{2a}, IgA, and (rarely) IgE, which contain the H chains γ_3, γ_1, γ_{2b}, γ_{2a}, α, and ε, respectively. The organization of the C gene segments at the H-chain locus is shown in Fig. 1 (3). Except for C_δ, each is preceded by a switch (S) region of several kilobases within which the breakpoints for the switch rearrangement are usually found.

Generally, the switch takes place after a committed B lymphocyte has been stimulated by antigen or mitogen (4-8). But *a priori*, cells may switch from μ to other isotypes, as soon as μ chain is synthesized, i.e., as early as the pre-B cell stage, before light chain is produced. Indeed, Abelson virus-transformed pre-B cells undergo class switching (9); in these cells at least, the switch process can even take place before VDJ rearrangement is complete (Wabl, unpublished).

The switch rearrangement process does not differentiate active and silent alleles. This is not surprising, since without reading the sequence, a non-

productive allele can often not be distinguished from a productive one. For instance, a silent allele may be non-productive because VDJ rearrangement has resulted in a frameshift, and that will make itself apparent only upon translation in the cytoplasm. In the nucleus, such a non-productive allele will be transcribed as well as a productive one.

Switching via DNA Recombination

In the germ line of vertebrates there are no functional genes encoding the polypeptide chains of immunoglobulins, only gene segments. During B-cell differention, selected segments are joined to yield the functional H and light chain genes. Two types of DNA rearrangements, each of which may result in either deletions or inversions, join together gene segments that may be far away from each other in the germ line (reviewed in 10, 11). The first type of rearrangement is rather precise; it is site specific and is mediated by a recombinase that recognizes known signal sequences flanking the exon segments to be joined. Exons encoding the variable (V) regions of the immunoglobulin chains and thus the antibody specificity are generated by such a process (12). The second type is rather imprecise; the recombination break points vary widely within an intron. Because the intron sequences will be deleted by RNA splicing, this imprecision does not impair the final outcome of the process. The immunoglobulin H chain class switch results from this second type of arrangement. The gene encoding the immunglobulin H chain is reconstructed so that the gene segment encoding a given C region is replaced by another C region gene segment. located further downstream in the locus. Therefore, a B cell clone can alter the effector function of the immunoglobulin produced (dictated by its CH isotype) without affecting its specificity (dictated by its V region). After it was found that the immunoglobulin class switch results

from the deletion of DNA segments (13), three different mechanisms were proposed to account for this deletion: (i) unequal recombination between homologs; (ii) unequal recombination between sister chromatids; and (iii) looping out and deletion. These are the simplest mechanisms that can account for any deletion of DNA in diploid cells. All three mechanisms may operate in B lymphocytes to achieve the switch deletion, and immunologists, as in their wont, love to point out the charms of their favorite. In the following we discuss both the positive and negative evidence for these three models.

Recombination Between Homologs

To produce a deletion, the exchange between homologous chromosomes must be unequal. If recombination occurs during or just prior to mitosis, there are four strands that can cross over. Random segregation of the four strands would result in four different types of daughter cells (Fig. 2). In mice, the order at the H locus is centromere, C, V (14). Allelic exclusion dictates that only one of the two homologs encodes a functional H chain; thus, let V_A be the functionally rearranged variable region allele and V_B the silent, non-functional homolog. Two types of daughter cell should have two identical V segments. Cell 1 should produce both a γ and a μ chain from two active alleles and should lose one C_μ segment. The other two types of daughter cell remain heterozygous with respect to V. The γ-producing cell 3 has gained on its silent homolog the C_μ that was lost by the active homolog. The μ-producing cell 4 is the only one that is inconspicuous with respect to both genetic make up and H chain production. From the above, it follows that half of the γ-producing cells should also produce μ (cell 1); the other half should contain two C_μ segments on the silent homolog (cell 3). Furthermore, among the descendants of the cell in which the switch took place, there should be half as many non-productive

cells containing three C_μ segments (cell 2) as there are γ-producing cells. *In vivo*, non-producing cells should be selected against, but *in vitro* they might even have a growth advantage. In any event, none of these conspicuous cells have been found. They were absent in a μ-producing pre-B-cell line (15), the cells of which continually switch to γ_{2b} (9, 16). In this line there is also direct evidence that the switch rearrangement occurs in *cis* (17). If anywhere near half of the cells that produce γ, ε, or α in long-term cultures contain two active V alleles (and as a consequence also produce μ chain), and if as many cells contain three C_μ gene segments, they ought to have been noticed during the analysis of hundreds or thousands of hybridomas and myelomas that were generated in the labs around the world, by ELISA, Southern blotting, or cloning and sequencing.

Serology provides another strong argument against recombination between homologs being the major mechanism of the switch deletion. In rabbits (18-21) and in mice (22), V and C region allotypes, *i.e.*, serologically defined allelic products, are generally both encoded by the same homolog. But according to one report (20), as many as 8% of the molecules may have been of a recombinant type. Even if one took this figure at face value, the contribution of recombination between homologs to the switch deletion would be minor. Moreover, homologous (*i.e.*, equal) mitotic recombination between the V and C loci would not contribute to the switch rearrangement process.

Although this mechanism of recombination is not the major source of switch deletions, it must occur at some low frequency. Indeed, illegitimate recombination between the switch regions and regions on other, non-homologous, chromosomes is known to occur sometimes, *e.g.*, in the translocations of the oncogene *myc* to the immunoglobulin locus (reviewed in 23). However, class switching of a randomly integrated μ transgene has been

reported to occur quite often (24). This might be due to trans-splicing of RNA (which is discussed below), but it might also be due to interchromosomal recombination (24a); it was this latter interpretation that rekindled an interest in the mechanism generating the switch deletion. This observation is interesting in its own right, although a μ transgene, being deprived of all *cis* sequences to which it would otherwise switch, hardly reflects the physiological situation[1] Furthermore, the high frequency of switched cells might also be due to selection by antigen rather than to any intrinsically high rate of the process. In B cells from these mice, translocations to the H chain locus on chromosome 12 ought to be cytologically observable.

Recombination Between Sister Chromatids

Although sister chromatid exchange (SCE) is frequent in somatic cells, there is no known physiological consequence. Since it would be gratifying to assign to it a role in generating antibody diversity, the proposal that unequal SCE might be responsible for producing the switch deletion was immediately welcomed. This mechanism was proposed to explain the fact that the major intron in a γ_1-producing myeloma containig a piece of the S_α region (which is the most 3' switch region) between S_μ and $S_{\gamma 1}$ (26). It would take at least two unequal exchanges to account for the observation, and one can imagine the participation of a number of other mechanisms, including translocation and episomal reintegration. It was also proposed that unequal SCE may have created γ_{2a} variants from a γ_{2b}-producing myeloma (27).

It is inherent in this model that one cannot easily disprove it. Nor is it easy to obtain direct evidence supporting it except to catch it in *flagrante delicto.* In an attempt to do this cytologically, we measured SCE frequencies in lipopolysaccharide-stimulated, mouse B lymphocytes (Wabl, unpublished).

The exchanges can be directly visualized by a cytogenetic procedure. Figure 3 shows a photomicrograph of a metaphase, in this case of hamster chromosomes, with differentially stained sister chromatids. The chromatid exchanges lead to a harlequin-type pattern of staining (left), which becomes more fidgety when (some) mutagens are added (right). In LPS-stimulated mouse spleen cell cultures of high cell density ($2x10^6$ cells/ml, or more), only IgM (together with IgD) is produced, while in cultures of lower cell density ($2.5x10^5$ cells/ml), the other six classes and subclasses are also produced. The SCE frequencies in both types of cultures were determined by differentially staining the sister chromatids in an attempt to correlate it with the immunoglobulin class switch. In C57BL/6 mice the number of SCE in cultures of high cell density was 0.227 ± 0.016 per chromosome, as compared to 0.378 ± 0.124 at low cell density. Similar values were obtained for the Rb(4;12) mouse strain. Although the ability of lymphocytes to switch to isotypes other than μ can be correlated with increased SCE frequency, there is no dramatic increase of SCE on chromosome 12 (where the mouse heavy chain linkage group is located) over the other chromosomes: In low cell density cultures the number for chromosome 12 was 0.460 (0.409–0.525 at 95% confidence limits), as compared to 0.374, determined for over 5000 chromosomes other than chromosome 12. Chromosome 12 has medium size when compared to the other 19 chromosomes of the mouse; the slight difference in SCE frequency, as compared to the average chromosome, may be due to a slight deviation from the average size. This nominal level of SCE on chromosome 12 may be sufficient to generate a sigificant number of switch recombinants, yet it is clear that chromosome 12 does not stick out as being a special target for SCE. While this result by no means disproves the unequal sister chromatid exchange model for the class switch, it also does not increase enthusiasm for it.

The switch from μ to another H-chain isotype by sister chromatid exchange would have to occur at mitosis, when sister chromatids exist, and it would have to occur on the active homolog (Fig. 4). The daughter cell containing the switched γ gene (cell 2) bears in its genome no mark to betray how it was generated. However, there is such a mark in its sister (cell 1), since the active homolog of cell 1 should contain two C_μ segments—what one cell loses, the other one must gain—in addition to the C_μ on the silent homolog; *i.e.*, there are three C_μ segments in the genome. Unequal exchange on the silent homolog would also leave half of the μ^+ cells with three C_μ segments. Despite intensive searching, this configuration has not been found in the pre-B cell line that switches from μ to γ_{2b} (15), nor in any other instance.

Looping-out and Deletion

Looping-out is the third way to produce a deletion (Fig. 5) and, as Sherlock Holmes pointed out, if all other speculations have been excluded, the remaining hypothesis must be the correct one (28). Looping out and deletion are also involved in the mechanism used to rearrange the V segments at the T-cell receptor and immunoglobulin loci (29-34). After looping out and cutting, the four free DNA ends created in this way can be religated in three different ways to produce (i) the original configuration (no switch), (ii) an inversion of the looped out sequences (loss of H-chain expression), or (iii) deletion of the looped sequences from the chromosme (switch). In the pre-B cell line we found many cells with an inversion between two switch regions (Jäck et. al., 1988). This was the first "positive" evidence that the switch rearrangement involves looping out. If these sequences were instead deleted, the ends of the excised 65 to 200 kb of DNA could theoretically be ligated to form a circle; this switch circle would contain C_μ, including the 3' part of S_μ, and all of the C

regions between C_μ and the C_H to which the cell had switched. The breakpoint would be located somewhere between the 3' part of $S\mu$ and the 5' part of S_H; these two switch regions, characteristic for the switch circle, would be joined in an order that is reciprocal to the configuration remaining on the chromosome.

Recently, such switch circles were isolated (Fig. 6), cloned, and partially sequenced from lipopolysaccharide-stimulated mouse spleen cells (35, 36). The alkaline lysis method, originally developed for the isolation of large Ti plasmids in bacteria and subsequently modified by Griffin et al. (37) and Carroll et al. (38) for Epstein-Barr virus genomes (180 kb) and large episomes in mammalian cells, was efficient enough to allow isolation of rare circles (35). No contamination by chromosomal DNA was detected by Southern blot analyses or cloning into λ phages. Furthermore, from the intensities of the Southern blot bands, it was estimated that a very good yield is obtained with this simple and fast method (0.25 to 0.5 circle per switching event, if one assumes that the circles do not replicate). These results were quickly reproduced by Sakano et al. (39).

As none of the genomic configurations predicted by recombination between homologs or sister chromatids could be demonstrated, while both products of looping-out—inversions and circular DNA—could be isolated, it would seem that the chapter on the question of which mechanism produces the switch deletion can be closed.

Do the switch circles replicate?

The origins of replication on animal chromosomes are estimated to be 50 to 330 kb apart (40), so it would not be surprising if there were an origin somewhere within the approximately 200 kb of DNA containing the segments encoding the various the C locus of the immunoglobulin H chain C regions. In

a pre-B cell line in which cells had switched to γ_{2b} on both alleles and, thereby, had deleted $C\mu$ from the chromosome, no $C\mu$ sequences were found by polymerase chain reaction (U. von Schwedler and M. Wabl, unpublished observation); thus, the circles excised during a switch to γ_{2b} and containing $C\mu$-$C\delta$-$C\gamma_3$-$C\gamma_1$ do not seem to replicate, at least not at the same rate as the chromosomes. Switch circles containing the $C\gamma_{2b}$, $C\gamma_{2a}$, and/or $C\varepsilon$ exons would be much larger, and might well contain an origin of replication that also functions extrachromosomally. In non-B cells, the immunoglobulin H locus functions as a single replicon with its origin downstream of $C\alpha$ (41); however, additional origins of replication may be activated in B cells.

Although the intervening sequences between the C_H exons have been sequenced, one does not know whether any additional genes are located in this region. For example, an open reading frame has been found between the $C\mu$ and $C\delta$ (42). If the maintenance of expression of these putative genes were to be crucial for the switched cell, there would be a strong selection the continuous presence of some switch circles.

What decides between deletions and inversions?

As explained above, deletions and inversions are two different ways to resolve the looping out of DNA for the recombination process. In some hybridomas the active allele has a switch deletion, and the silent allele has a switch inversion (43). In the Abelson virus-transformed pre-B cell line, where this phenomenon was studied and where there is no selection, there are almost as many inversions as there are deletions (17). The reason for this is unknown, and it is only paraphrasing the experimental observation to say that the switch recombinase is a sloppy enzyme.

Homologous vs. site specific recombination

It has been argued that the repetitive sequences of the S regions facilitate homologous recombination (44), or focus the putative switch recombinase (45). Thus, the switch regions might help to juxtapose DNA by virtue of their homologous sequences, or they might be recognized by and bind to a component of the switch recombinase. The fact that there were found switch inversions makes non-homologous, site-specific recombination more likely, because homology is probably not maintained in the other reading direction. But one does not know whether or not the switch regions contain inversions in their genomic configuration. And, of course, it is very difficult to exclude homologous recombination, if one sets the degree of homology low. General eukaryotic recombinases can join DNA sequences with as little as 13 base pairs of continuous identical sequences, or sets of three smaller patches separated by deletions and/or looping outs (46, 47). The two participating switch regions do not pair according to maximum homology prior to recombination (48). But how many base pairs of homology would be needed for switching? At minimum one, in which case the controversy resolves into almost a semantic one.

The sequences of the mouse S regions consist of repetitive elements which vary in length from 10 to 80 base pairs. S regions span anywhere from 1 kb ($S\varepsilon$) to 10 kb ($S\gamma_{2b}$). Each switch region has its own charcteristic short and long repeats, yet all have enough homology to cross-hybridize to each other to some extent. As compared to $S\mu$, the degree of homology is highest with $S\varepsilon$, and then decreasing in the following order: $S\varepsilon > S\alpha > S\gamma_3 > S\gamma_1 > S\gamma_{2b} > S\gamma_{2a}$ (49, 50). It is found that recombination with the switch region most homologous to $S\mu$ is the least frequent: $S\varepsilon$. This would speak against homologous recombination; but $S\varepsilon$ is also the shortest, and cellular selection might bias the frequencies. One

can also argue that first the locus has to be "open", and only then would homology become important.

Transcription as a requirement

In both man and mouse switching is preceded by transcription of the unrearranged C gene segment to which the cell will be switching (51-65). Since they appear not to be translated, these germ line transcripts are also called "sterile transcripts", however, an open reading frame has been found in the germline μ (66), γ_{2b} (54), and α (67) transcript. The transcripts have multiple initiation sites, about 2 kb 5' to the S region, and proceed through the S region and C segment (54, 56-58, 60, 61, 68, 69); thus they lack the VDJ segment. The question is, are these transcripts a prerequesite for the switch recombination or are they fellow travelers of it? An argument for their being required is the strong correlation between their induction by a particular interleukin and the specific class switch directed by that interleukin; e.g. IL-4, which directs spleen cells to switch predominantly to IgG_1 or IgE, also induces germline γ_1 and ε transcripts (55-57, 64, 65, 68-72).

In yeast it has been shown that transcription by DNA polymerase I or II can stimulate recombination more than 15 fold (73-75). Recently, Leung and Maizels (76) have directly tested the effect of transcription on class switching in mouse cells. The presence of an upstream transcriptional control region increases the recombination frequency of an extrachromosomal substrate at least 10 fold.

Switch Recombinase

Now that the over-all mechanism of switching has been solved, an important — if not *the* important — immunological question remains: how is

the class switch directed? Because the difference between switching to IgG and switching to IgE is the difference between immunity and allergy, this question also has great clinical relevance. An authoritative review on the role of interleukins, *i.e.*, on the role of extracellular factors, was provided by Finkelman et al., (77). A detailed answer to the question will also require the identification of the intracellular molecular components responsible for the immunoglobulin H chain class switch. We assume that the switch rearrangement process, like other biochemical reactions within the cell, is facilitated by an enzyme(s). This putative enzyme has been named *switch recombinase.* Nothing is known about its components. The switch recombinase must perform several of the same functions as the VDJ recombinase, which generates the VDJ exon from three different germline gene segments. These functions include binding to DNA, cutting it, and ligating it. The two enzymes obviously differ in their DNA recognition sequences. For the VDJ-recombinase, these are the heptamer-nonamer recombination signal sequences, which flank the gene segments (reviewed in 12). For the switch recombinase the recognition sequences are not known, but the simplest assumption is that they are part of the switch region. Since the cutting and ligating activities are common to both enzymes, they might even be executed by the same proteins, but the binding component must be different. Although no breakthrough similar to the one for the VDJ-recombinase — the characterization of the *scid* mutation and the cloning of RAG-1 and RAG-2 — has been made, the regions where most of the recombination break points are located (the switch regions) have been identified long time ago (10). Some sequence requirements for the recombination process have been established (76, 78), and some binding proteins thought to be involved in the switch rearrangement process have been described.

Binding proteins

A popular hypothesis on directed class switching proposes that specificity is provided by the so-called "sterile transcripts" originating 5' to the repetitive elements of the S regions (see above). This hypothesis is based on the observation that the VDJ recombinase is the same for B and T cells, yet only the genes for the appropriate antigen receptors are, in general, rearranged. It was argued that either the immunoglobulin locus or the T cell antigen receptor locus is "accessible" in a given cell, but not both. Logically, there is no difference between the assumption of a core recombinase plus several different factors guiding the recombinase to a given locus and the proposal for different recombinases working at the various loci. For the accessibility hypothesis to be more than a tautology, it must specify what makes the loci accessible. If specific transcription provides the specificity of rearrangement, then the transcription factors may be considered part of the putative switch recombinase complex. The same may be said about the factors of a putative (de)methylation system that would be responsible for the accessibility of recombination sites (79-81). Although accessibility of this region may make it a (specific) target for recombination, there must be, in addition, specific factors that direct the recombination to two different switch regions, but not to other—also transcribed—sequences that are very similar to each other. These factors need not be specific for a particular switch region, but they ought to be specific for a generic switch region.

It is unclear whether a given recombinase, or its binding component, recognizes all switch regions or only a subset thereof. It is also not clear whether or not the essential elements of the switch regions reside in the portion with the repetitive sequences; but the most straight-forward assumption is that the switch region itself is the binding site. It was reported

that the γ_1 switch region contains repetitive octamer-like sequences (82) and that the sequences flanking those repeats are important for protein binding. Several other laboratories have described similar factors (83-86). Some of them are LPS-inducible and bind outside the (repetitive) switch regions.

Such binding factors were commonly identified with the gel mobility shift technique. In conjunction with protection and methylation interference assays this technique was often used to identify DNA sequences that are essential for a specific protein-DNA interaction. However, the gel shift assay is not sensitive enough to exclude with certainty small but functional levels of a protein. Therefore, the moment of truth, in regard to tissue specificity, for binding factors identified by gel shift assays comes when they are cloned and when different tissues are assayed for mRNA encoding the factors. Only then does it become clear whether or not their expression is restricted to the appropriate B cell differentiation stage(s). At the time this review was written, no such cloning has been reported.

A more tedious task is it to demonstrate that such factors directly contribute to the switch rearrangement process, *i.e.*, they are part of the switch recombinase complex. But once a binding protein is cloned there are obvious ways to assess whether it is at all a necessary component for class switching. For example, knocking out the gene encoding the binding protein in the cell line that continually undergoes class switching would provide an answer. Or the cDNA can be cloned into a mammalian expression vector, and then stably or transiently transfected into a cell line that does not switch its immunoglobulin isotype, *e.g.*, the 3T3 fibroblast line, a plasmacytoma, a hybridoma, and a B lymphoma. It then can be tested whether or not the cells switch, or whether a cell line switching to one isotype now (exclusively) switches (also) to the isotype targeted by the binding protein.

Cloning by function

The stunning (and, so far, only) success in cloning (parts of) the VDJ recombinase came from cloning by function (87, 88). Although not many researchers would have bet on this particular approach, it was facilitated by the extensive work that had been done on recombination substrates with their defined heptamer–nonamer signal sequences. Difficulties with designing an appropriate switch substrate stem from the fact that there is no agreement on the signal sequences for the switch recombination. Therefore, large fragments may have to be used. Unfortunately, such substrates may be difficult to generate due to the highly repetitive nature of the switch regions; classic cloning of the entire region often results in deletion of part of the sequences, and cloning of smaller segments is problematic since there is a lack of suitable restriction sites. However, a retroviral vector containing S_{μ} amd $S_{\gamma 2b}$ was used to monitor switching in the Abelson virus-transformed pre-B cell line 18-81 (78). The switch frequencies reported with this substrate were moderate, but in another vector with switch regions lifted off from the germ line there were no problems with "spontaneous" deletions, and switch frequencies were high (76). The experimental approaches described above focus on different aspects and functions of the putative switch recombinase. The isolation and characterization of a specific switch region binding protein must be complemented by a functional assay that demonstrates that the protein is part of the recombinase. And the isolation of factors that promote recombination in a switch substrate must be followed by an analysis that differentiates between transcription factors and enzymes. Success with a set of experiments designed to isolate the one factor may make inroads into the hunt for the other component. For example, the first specific binding protein can be used to pull out further proteins, which are bound to the first one and which are also components of the recombinase.

Switching without DNA recombination

Alternative splicing of long nuclear RNA

As soon as introns were discovered, it was suggested that immunoglobulin class switching is due to alternative splicing of long nuclear RNA (89); and for hybridomas it was almost immediately ruled out that RNA splicing is the mechanism of producing mRNA encoding an isotype other than μ (89a). However, alternative splicing is thought to be the mechanism by which small resting B lymphocytes produce both IgM and IgD (90, 91), although the predicted 28 kb RNA has yet to be isolated.

Alternative splicing is also an attractive idea to explain the existence of the lymphocytes producing two isotypes that are generated after stimulation with bacterial lipopolysaccharide *in vitro*. After six hours of stimulation of mouse spleen cell cultures at cell concentrations lower than 10^6 per ml (6), many B lymphoblasts begin to produce IgM as well as immunoglobulin of another isotype. But are the mRNA's encoding each isotype still being produced in these cells, or is the μ chain translated from still undegraded mRNA that was transcribed *before* an isotype switch? Attempts in several laboratories to rescue the double-producing lymphoblasts by fusion to cells of a similar differentiation stage, *e.g.*, to a hamster B cell lymphoma synthesizing IgG, have failed; each hybrid cell produced either IgM or another isotype, but not both. However, one can argue that the fusion event forces the previously double-producing cell to produce only one isotype. In any event, proof for the alternative splicing of a long transcript (approximately 200 kb for mouse α chain) would clearly have to include proof for the existence of such a long transcript. Although it was demonstrated that sorted spleen cells producing both IgM and IgG_1 contain RNA segments transcribed from intervening sequences between the segments encoding the various H chain C regions (92),

there seem to be no other papers following up on this line of experimentation. We note that switching in LPS-stimulated spleen cells is accompanied by the classical switch deletion of DNA (93).

Unless continuous production of two isotypes (other than IgM and IgD) is documented (*e.g.*, by sorting, stripping, and allowing resynthesis), the absence of DNA rearrangement is only a weak argument for alternative splicing of a long transcript. For instance, it has been shown that most, perhaps all, resting B cells that display both IgM and IgE on the surface do so because IgE is absorbed onto the membrane (94). And a population of sorted cells will contain many cells with unique histories of recombination, which will preclude the detection of a specific restriction fragment length polymorphism indicative of switch recombination. Thus the argument must rely on not seeing any less than exactly two unswitched alleles. In cell lines that produce an isotype other than μ, the absence of DNA rearrangement would seem to be a better argument for the existence of alternative splicing. However, the literature is replete with examples of failure to find DNA rearrangement because of experimental errors.

Since there exist genes that span well over 100 kb and that are, presumably, transcribed as a single continuous RNA molecule, there is no reason *a priori* to dismiss the notion of a transcript long enough to allow alternative splicing as a mechanism for the immunoglobulin class switch. However, there still seems to be no convincing demonstration that such a mechanism can account for the immunoglobulin class switch *in vitro* or *in vivo*, even in special cases.

Trans-splicing

Trans-splicing as the mechanism for some switch events is a novel and exciting suggestion. A variant of the BCL1 cell line, which produces both

IgM and IgD, has switched to IgG_1, while maintaining IgM production (95). A detailed Southern blot analysis of the entire C region locus failed to find a switch rearrangement (95). However, the analysis is complicated by the fact that BCL1 has lost the silent homolog of chromosome 12 and duplicated the active homolog; *i.e.*, there are two alleles containing identical, functionally rearranged VDJ exon linked to the segments encoding the H chain C regions. While there is no rearrangement on the one allele, which then presumably encodes the μ chain, there is a large deletion on the duplicated allele just 3' of the VDJ exon. Since no $C\gamma_1$ segment could be found to be linked to this VDJ exon, it was concluded that this allele does not encode the γ_1 chain (Tucker). It would be interesting to switch the BCL1 cell line repeatedly in independent switching events. If the deletion on the one—"inactive" allele is strictly correlated with the switching, this would indicate the possibility that it constitutes indeed a switch deletion. giving rise to γ_1 expression. Conversely, if μ and γ_1 are encoded on one, unrearranged allele, they should segregate together when chromosomes are lost.

In a follow up study, Nolan-Willard et al. (96) reported that the IgG_1-producing BCL1 subclone synthesizes pre-mRNA that contains both μ and γ_1 sequences. They thus propose that the simultaneous μ and γ_1 RNA synthesis is accomplished by discontinuous transcription followed by either trans-splicing or ligation of μ pre-mRNA to sterile γ_1 transcripts. While this report shows unassailable sandwich hybridization blots, proof for the presence of the putative pre-mRNA would require a cDNA sequence.

In mice with a human μ transgene lacking other C gene segments, stimulated spleen cells were reported to synthesize both human μ and mouse γ_1 chains (97). A problem of allelic inclusion was ruled out by the authors. Since the transgene was not integrated on chromosome 12, the authors also

excluded an intrachromosomal deletion as an explanation. Rearrangement of neither the transgene nor the endogenous H chain locus was found in the sorted μ plus γ producing blast cells. Of 12 lymphomas screened, one did not produce μ, but apparently contained an mRNA encoding the human transgenic VDJ linked to mouse $C_{\gamma 2a}$ segment. This led the authors to conclude that trans-splicing is responsible for the class-switch in this particular case. However, the trans-spliced mRNA was detected by PCR, which is prone to many artifacts, including template switching. As stated above, isolation of cDNA encoding a trans-spliced mRNA would make one feel much more comfortable with the unorthodox conclusion. There should be no difficulty in doing that in a lymphoma line.

We have to await confirmation of these intriguing findings before such a mechanism is awarded a suitable place in the panopticum of events that *can* happen. However, in its simplest form, such a mechanism cannot account for the majority of switch events, at least in rabbits and in mice, where there are convenient serological markers for both V region and C region allotypes (see above). In heterozygotes, over 90% of serum IgG does not show scrambled markers; *i.e.*, when the V region of an immunoglobulin H chain is of haplotype *a*, so is the C region of that molecule, and when the V region is of haplotype *b*, so is the C region. Were trans-splicing responsible for the class switch, we would expect a V region of haplotype *a* to be associated with a C region of haplotype *b* in about fifty percent of the H chain molecules.

Concluding remarks

The ideas that are stronger than experiments are called paradigms (Kuhn, 1962). From this definition it follows that one cannot disprove a paradigm with experiments, one can only try to replace it with another

paradigm. It is interesting to note that paradigm articulation seems to proceed differently in the similar problems of VDJ rearrangement and switch rearrangement.

Although unequal sister chromatid exchange had been suggested for as the mechanism for rearrangement of V_κ to J_κ, this mechanism was dropped in favor of the looping out and deletion mechanism when the circular excision products stemming from the T cell antigen receptor V segment rearrangements were isolated. For the generation of the chicken V exons, the demonstration that signal joints could be found in Bursa cells by PCR was accepted as a proof for the looping out and deletion mechanism, even though the very same joints are predicted by the other two competing models.

However, for the heavy chain class switch the models seem to coexist happily together. To the best of our knowledge, no one is making a V_H transgenic mouse to see whether the exogenous V_H can recombine with an endogenous D and thus collect evidence for recombination between homologs or trans-splicing. Moreover, a protein that binds to a switch region but that is found in all cells, and even in *Drosophila*, would probably not create the same excitement as the ubiquitous heptamer binding protein implicated in the VDJ recombination process.

The discovery of the switch deletion, reported fifteen years ago by Honjo in a landmark paper (13), was easily accepted because it was preceded by analogous findings for the generation of the V exons. Indeed, except for some variations on the theme, it can be said that until now research on switch rearrangement has followed in the wake of research on VDJ rearrangement. We predict that there will be a role reversal in their application to clinical immunology. For example, when the putative intracellular factors guiding the switch recombinase specifically to $S\varepsilon$ are cloned, they can be used to screen for

small molecules that bind to them and thus may be pharmacologically active in preventing the switch to IgE. Such small molecules would act as universal anti-allergens.

References

1. Popper, K. R. S. 1968. Conjectures and refutations: the growth of scientific knowledge. Harper torchbooks TB 1376., New York: Harper & Row

2. Fleck, L. 1935. Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv. Basel: Schwabe & Co.

3. Shimizu, A., Takahashi, N., Yaoita, Y., Honjo, T. 1982. Organization of the constant-region gene family at the mouse immunoglobulin heavy chain. Cell 28: 499-506

4. Nossal, G. J. V., Szenberg, A., Ada, G. L., Austin, C. M. 1964. Single cell studies on 19S antibody production. J. Exp. Med. 119: 485-502

5. Pernis, B., Forni, L., Luzatti, A. L. 1976. Synthesis of multiple immunoglobulin classes by single lymphocytes. Cold Spring Harbor Symp. Quant. Biol. 41: 175-83

6. Kearney, J. F., Lawton, A. R. 1975. B lymphocyte differentiation induced by lipopolysaccharide. I. Generation of cells synthesizing four major immunoglobulin classes. J. Immunol. 115: 671-77

7. Wabl, M. R., Forni, L., Loor, F. 1978. Switch in immunoglobulin class production observed in single clones of committed lymphocytes. Science 199: 1078-80

8. Coutinho, A., Benner, R., Björklund, M., Forni, L., Holmberg, D., Ivars, F., Martinez, A. C., Petterson, S. 1982. A "trans" perspective on the control of immunoglobulin C gene expression. Immunol. Rev. 67: 87-114

9. Burrows, P. D., Beck, G. B., Wabl, M. R. 1981. Expression of μ and γ immunoglobulin heavy chains in different cells of a cloned mouse lymphoid line. Proc. Natl. Acad. Sci. USA 78: 564-68

10. Shimizu, A., Honjo, T. 1984. Immunoglobulin class switching. Cell 36: 801-03

11. Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature 302: 575-81

12. Schatz, D. G., Oettinger, M. A., Schlissel, M. S. 1992. V(D)J recombination: molecular biology and regulation. Annu. Rev. Immunol. 10: 359-83

13. Honjo, T., Kataoka, T. 1978. Organization of immunoglobulin heavy chain genes and allelic deletion model. Proc. Natl. Acad. Sci. USA 75: 2140-44

14. Seldin, M. F., Howard, T. A., D'Eustachio, P. 1989. Comparison of linkage maps of mouse chromosome 12 derived from laboratory strain intraspecific and Mus spretus interspecific backcrosses. Genomics 5: 24-28

15. Wabl, M., Meyer, J., Beck-Engeser, G., Tenkhoff, M., Burrows, P. D. 1985. Critical test of a sister chromatid exchange model for the immunoglobulin heavy-chain class switch. Nature 313: 687-89

16. Burrows, P. D., Beck-Engeser, G. B., Wabl, M. R. 1983. Immunoglobulin heavy-chain class switching in a pre-B cell line is accompanied by DNA rearrangement. Nature 306: 243-46

17. Jäck, H. M., McDowell, M., Steinberg, C. M., Wabl, M. 1988. Looping out and deletion mechanism for the immunoglobulin heavy-chain class switch. Proc. Natl. Acad. Sci. USA 85: 1581-85

18. Kindt, T. J., Mandy, W. J., Todd, C. W. 1970. Association of allotypic specificities of group a with allotypic specificities A11 and A12 in rabbit immunoglobulin. Biochemistry 9: 2028-32

19. Landucci-Tosi, S., Mage, R. G., Dubiski, S. 1970. Distribution of allotypic specificities A1, A2, A14, and A15 among immunoglobulin G molecules. J. Immunol. 104: 641-47

20. Knight, K. L., Hanly, W. C. 1975. Genetic control of a chain of rabbit IgA: allotypic specificities on the variable and the constant regions. Contemp. Top. Mol. Immunol. 4: 55-58

21. Pernis, B., Forni, L., Dubiski, S., Kelus, A. S., Mandy, W. J., Todd, C. W. 1973. Heavy chain variable and constant region allotypes in single rabbit plasma cells. Immunochemistry 10: 281-85

22. Gearhart, P. J., Hurwitz, J. L., Cebra, J. J. 1980. Successive switching of antibody isotypes expressed within the lines of a B-cell clone. Proc. Natl. Acad. Sci. USA 77: 5424-28

23. Croce, C. M. 1987. Role of chromosomal translocations in human neoplasia. Cell 49: 155-56

24. Durdik, J., Gerstein, R. M., Rath, S., Robbins, P. F., Nisonoff, A., Selsing, E. 1989. Isotype switching by a microinjected mu immunoglobulin heavy chain gene in transgenic mice [published erratum appears in Proc Natl Acad Sci USA 1989 Aug;86(16):6312]. Proc. Natl. Acad. Sci. USA 86: 2346-50

24a. Gerstein, R. M., Frankel, W. N., Hsieh, C. L., Durdik, J. M., Rath, S., Coffin, J. M., Nisonoff, A., Selsing, E. 1990. Isotype switching of an immunoglobulin heavy chain transgene occurs by DNA recombination between different chromosomes. Cell 63: 537-48

25. Liesegang, B., Radbruch, A., Rajewsky, K. 1978. Isolation of myeloma variants with predefined variant surface immunoglobulin by cell sorting. Proc. Natl. Acad. Sci. USA 75: 3901-05

26. Obata, M., Kataoka, T., Nakai, S., Yamagishi, H., Takahasi, N., Yamawaki-Kataoka, Y., Nikaido, T., Shimizu, A., Honjo, T. 1981. Structure of a rearranged g1 chain gene and its implication to immunoglobulin class-switch mechanism. Proc. Natl. Acad. Sci. USA 78: 2437-41

27. Tilley, S. A., Birshtein, B. K. 1985. Unequal sister chromatid exchange. A mechanism affecting Ig gene arrangement and expression. J. Exp. Med. 162: 675-94

28. Doyle, A. C. S. 1927. The Complete Sherlock Holmes; with a preface by Christopher Morley. Deluxe ed. New York: Garden City Publishing Co.

29. Fujimoto, S., Yamagishi, H. 1987. Isolation of an excision product of T-cell receptor alpha-chain gene rearrangments. Nature 327: 242-43

30. Okazaki, K., Sakano, H. 1988. Thymocyte circular DNA excised from T cell receptor alpha-delta gene complex. EMBO J. 7: 1669-74

31. Toda, M., Fujimoto, S., Iwasato, T., Takeshita, S., Tezuka, K., Ohbayashi, T., Yamagishi, H. 1988. Structure of extrachromosomal circular DNAs excised from T-cell antigen receptor alpha and delta-chain loci. J. Mol. Biol. 202: 219-31

32. Toda, M., Hirama, T., Takeshita, S., Yamagishi, H. 1989. Excision products of immunoglobulin gene rearrangements. Immunol. Lett. 21: 311-16

33. Abe, M., Shiku, H. 1989. Isolation of an IgH gene circular DNA clone from human bone marrow. Nucl. Acid. Res. 17: 163-70

34. McCormack, W. T., Tjoelker, L. W., Carlson, L. M., Petryniak, B., Barth, C. F., Humphries, E. H., Thompson, C. B. 1989. Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. Cell 56: 785-91

35. von Schwedler, U., Jäck, H. M., Wabl, M. 1990. Circular DNA is a product of the immunoglobulin class switch rearrangement. Nature 345: 452-56

36. Iwasato, T., Shimizu, A., Honjo, T., Yamagishi, H. 1990. Circular DNA is excised by immunoglobulin class switch recombination. Cell 62: 143-49

37. Griffin, B. E., Bjork, E., Bjursell, G., Lindahl, T. 1981. Sequence complexity of circular Epstein-Barr virus DNA in transformed cells. J. Virol. 40: 11-19

38. Carroll, S. M., Gaudray, P., de Rose, M. L., Emery, J. F., Meinkoth, J. L., Nakkim, E., Subler, M., von Hoff, D. D., Wahl, G. M. 1987. Characterization of an episome produced in hamster cells that amplify a transfected CAD gene at high frequency: Functional evidence for a mammalian replication origin. Mol. Cell. Biol. 7: 1740-50

39. Matsuoka, M., Yoshida, K., Maeda, T., Usuda, S., Sakano, H. 1990. Switch circular DNA formed in cytokine-treated mouse splenocytes: evidence for intramolecular DNA deletion in immunoglobulin class switching. Cell 62: 135-42

40. Edenberg, H. J., Huberman, J. A. 1975. Eucaryotic chromosome replication. Ann. Rev. Genet. 9: 245-84

41. Brown, E. H., Iqbal, M. A., Stuart, S., Hatton, K. S., Valinsky, J., Schildkraut, C. L. 1987. Rate of replication of the murine immunoglobulin heavy-chain locus: evidence that the region is part of a single replicon. Mol. Cell. Biol. 7: 450-57

42. Richards, J. E., Gilliam, A. C., Shen, A., Tucker, P. W., Blattner, F. R. 1983. Unusual sequences in the murine immunglobulin m-d heavy-chain region. Nature 306: 483-87

43. Siebenkotten, G., Esser, C., Wabl, M., Radbruch, A. 1992. The murine IgG1/IgE class switch program. Eur. J. Immunol. 22: 1827-34

44. Takahashi, N., Kataoka, T., Honjo, T. 1980. Gene 11: 117-27

45. Davis, M. M., Kim, S. K., Hood, L. 1980. Immunoglobulin class switching: developmentally regulated DNA rearrangement during differentiation. Cell 22: 1-2

46. Hsieh, P., Camerini, O. C., Camerini, O. R. 1990. Pairing of homologous DNA sequences by proteins: evidence for three-stranded DNA. Genes Dev. 4: 1951-63

47. Mills, F. C., Brooker, J. S., Camerini, O. R. 1990. Sequences of human immunoglobulin switch regions: implications for recombination and transcription. Nucl. Acid. Res. 18: 7305-16

48. Petrini, J., Dunnick, W. A. 1989. Products and implied mechanism of H chain switch recombination. J. Immunol. 142: 2932-35

49. Stanton, L. W., Marcu, K. B. 1982. Nucleotide sequence and properties of the murine γ3 immunoglobulin heavy chain switch region: implications for successive Cg gene switching. Nucl. Acid. Res. 10: 5993-97

50. Nikaido, T., Yamawaki-Kataoka, Y., Honjo, T. 1982. Nucleotide sequences of switch regions of immunoglobulin Ce and Cg genes and their comparison. J. Biol. Chem. 257: 7322-29

51. Stavnezer-Nordgren, J., Sirlin, S. 1986. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. EMBO J. 5: 95-102

52. Yancopoulos, G., DePinho, R., Zimmerman, K., Lutzker, S., Rosenberg, N., Alt, F. 1986. Secondary rearrangement event in pre-B cells: VHDJH replacement by a LINE-1 sequence and directed class-switching. EMBO J. 5: 3259-66

53. Lutzker, S., Rothman, P., Pollock, R., Coffman, R., Alt, F. W. 1988. Mitogen- and IL-4-regulated expression of germ-line Ig gamma 2b transcripts: evidence for directed heavy chain class switching. Cell 53: 177-84

54. Lutzker, S., Alt, F. W. 1988. Structure and expression of germ line immunoglobulin gamma 2b transcripts [published erratum appears in Mol Cell Biol 1988 Oct;8(10):4585]. Mol. Cell. Biol. 8: 1849-52

55. Rothman, P., Lutzker, S., Cook, W., Coffman, R., Alt, F. W. 1988. Mitogen plus interleukin 4 induction of C epsilon transcripts in B lymphoid cells. J. Exp. Med. 168: 2385-89

56. Stavnezer, J., Radcliffe, G., Lin, Y. C., Nietupski, J., Berggren, L., Sitia, R., Severinson, E. 1988. Immunoglobulin heavy-chain switching may be directed by prior induction of transcripts from constant-region genes. Proc. Natl. Acad. Sci. USA 85: 7704-08

57. Esser, C., Radbruch, A. 1989. Rapid induction of transcription of unrearranged S gamma 1 switch regions in activated murine B cells by interleukin 4. EMBO J. 8: 483-88

58. Sideras, P., Mizuta, T. R., Kanamori, H., Suzuki, N., Okamoto, M., Kuze, K., Ohno, H., Doi, S., Fukuhara, S., Hassan, M. S. 1989. Production of sterile transcripts of C gamma genes in an IgM-producing human neoplastic B cell line that switches to IgG-producing cells. Int. Immunol. 1: 631-42

59. Gerondakis, S. 1990. Structure and expression of murine germ-line immunoglobulin epsilon heavy chain transcripts induced by interleukin 4. Proc. Natl. Acad. Sci. USA 87: 1581-85

60. Jabara, H. H., Schneider, L. C., Shapira, S. K., Alfieri, C., Moody, C. T., Kieff, E., Geha, R. S., Vercelli, D. 1990. Induction of germ-line and mature C epsilon transcripts in human B cells stimulated with rIL-4 and EBV. J. Immunol. 145: 3468-73

61. Lehman, D. A., Nomura, D. Y., Coffman, R. L., Lee, F. D. 1990. Molecular characterization of germ-line immunoglobulin A transcripts produced during transforming growth factor type beta-induced isotype switching. Proc. Natl. Acad. Sci. USA 87: 3962-66

62. Qiu, G., Gauchat, J. F., Vogel, M., Mandallaz, M., A., d., Stadler, B. M. 1990. Human IgE mRNA expression by peripheral blood lymphocytes stimulated with interleukin 4 and pokeweed mitogen. Eur. J. Immunol. 20: 2191-99

63. Rothman, P., Lutzker, S., Gorham, B., Stewart, V., Coffman, R., Alt, F. W. 1990. Structure and expression of germline immunoglobulin gamma 3 heavy chain gene transcripts: implications for mitogen and lymphokine directed class-switching. Int. Immunol. 2: 621-27

64. Rothman, P., Chen, Y. Y., Lutzker, S., Li, S. C., Stewart, V., Coffman, R., Alt, F. W. 1990. Structure and expression of germ line immunoglobulin heavy-chain epsilon transcripts: interleukin-4 plus lipopolysaccharide-directed switching to C epsilon. Mol. Cell. Biol. 10: 1672-79

65. Severinson, E., Fernandez, C., Stavnezer, J. 1990. Induction of germ-line immunoglobulin heavy chain transcripts by mitogens and interleukins prior to switch recombination. Eur. J. Immunol. 20: 1079-84

66. Lennon, G. G., Perry, R. P. 1985. Cm-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5' non-translatable exon. Nature 318: 475-78

67. Radcliffe, G., Lin, Y. C., Julius, M., Marcu, K. B., Stavnezer, J. 1990. Structure of germ line immunoglobulin alpha heavy-chain RNA and its location on polysomes. Mol. Cell. Biol. 10: 382-86

68. Berton, M. T., Uhr, J. W., Vitetta, E. S. 1989. Synthesis of germ-line gamma 1 immunoglobulin heavy-chain transcripts in resting B cells: induction by interleukin 4 and inhibition by interferon gamma. Proc. Natl. Acad. Sci. USA 86: 2829-33

69. Gauchat, J. F., Lebman, D. A., Coffman, R. L., Gascan, H., deVries, J. 1990. Structure and expression of germline epsilon transcripts in human B cells induced by interleukin 4 to switch to IgE production. J. Exp. Med. 172: 463-73

70. Snapper, C. M., Finkelman, F. D., Stefany, D., Conrad, D. H., Paul, W. E. 1988. IL-4 induces co-expression of intrinsic membrane IgG1 and IgE by murine B cells stimulated with lipopolysaccharide. J. Immunol. 141: 489-98

71. Gauchat, J. F., Aversa, G., Gascan, H., deVries, J. 1992. Modulation of IL-4 induced germline epsilon RNA synthesis in human B cells by tumor necrosis factor-alpha, anti-CD40 monoclonal antibodies or transforming growth factor-beta correlates with levels of IgE production. Int. Immunol. 4: 397-406

72. Xu, M. Z., Stavnezer, J. 1992. Regulation of transcription of immunoglobulin germ-line gamma 1 RNA: analysis of the promoter/enhancer. EMBO J. 11: 145-55

73. Keil, R. L., Roeder, G. S. 1984. Cis-acting, recombination stimulating activity in a fragment of ribosomal DNA of S. cerevisiae. Cell 39: 377-86

74. Voelkel-Meiman, K., Keil, R. L., S., R. G. 1987. Recombination-stimulating sequences in

yeast ribosomal DNA correspond to sequences regulating transcription by RNA polymerase I. Cell 48: 1071-79

75. Thomas, B. J., Rothstein, R. 1989. Elevated recombination rates in transcriptionally active DNA. Cell 56: 619-30

76. Leung, H., Maizels, N. 1992. Transcriptional regulatory elements stimulate recombination in extrachromosomal substrates carrying immunoglobulin switch-region sequences. Proc. Natl. Acad. Sci. USA 89: 4154-58

77. Finkelman, F. D., Holmes, J., Katona, I. M., Urban, J. J., Beckmann, M. P., Park, L. S., Schooley, K. A., Coffman, R. L., Mosmann, T. R., Paul, W. E. 1990. Lymphokine control of in vivo immunoglobulin isotype selection. Annu Rev Immunol 8: 303-33

78. Ott, D. E., Alt, F. W., Marcu, K. B. 1987. Immunoglobulin heavy chain switch region recombination within a retroviral vector in murine pre-B cells. EMBO J. 6: 577-84

79. Kochanek, S., Radbruch, A., Tesch, H., Renz, D., Doerfler, W. 1991. DNA methylation profiles in the human genes for tumor necrosis factors a and b in subpopulations of leukocytes and in leukemias. Proc. Natl. Acad. Sci. USA 88: 5759-63

80. Burger, C., Radbruch, A. 1990. Protective methylation of immunoglobulin and T cell receptor (TcR) gene loci prior to induction of class switch and TcR recombination. Eur. J. Immunol. 20: 2285-91

81. Esser, C., Radbruch, A. 1990. Immunglobulin class switching: molecular and cellular analysis. Ann. Rev. Immunol. 8: 717-35

82. Schultz, C. L., Elenich, L. A., Dunnick, W. A. 1991. Nuclear protein binding to octamer motifs in the immunoglobulin g1 switch region. Int. Immunol. 3: 109-16

83. Waters, S. H., Saikh, K. U., Stavnezer, J. 1989. A B-cell-specific nuclear protein that binds to DNA sites 5' to immunoglobulin Sa tandem repeats is regulated during differentiation. Mol. Cell. Biol. 9: 5594-601

84. Wuerffel, R. A., Asher, T. N., Kenter, A. L. 1990. Detection of an immunoglobulin switch region-specific DNA-binding protein in mitogen-stimulated mouse splenic B cells. Mol. Cell. Biol. 10: 1714-18

85. Williams, M., Maizels, N. 1991. LR1, a lipopolysaccharide-responsive factor with binding sites in the immunoglobulin switch regions and heavy-chain enhancer. Genes Dev. 5: 2353-61

86. Liao, F., Giannini, S. L., Birshtein, B. K. 1992. A nuclear DNA-binding protein expressed during early stages of B cell differentiation interacts with diverse segments within and 3' of the Ig H chain gene cluster. J. Immunol. 148: 2909-17

87. Schatz, D. G., Oettinger, M. A., Baltimore, D. 1989. The V(D)J recombination activating gene, RAG-1. Cell 59: 1035-48

88. Oettinger, M. A., Schatz, D. G., Gorka, C., Baltimore, D. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248: 1517-23

89. Gilbert, W. 1978. Why genes in pieces? Nature 271: 501

89a. Shulman, M. J., Köhler, G. 1978. Immunoglobulin mu and gamma heavy chain classes are not determined by class-specific RNA-splicing enzymes. Nature 274: 917-9

90. Knapp, M. R., Liu, C.-P., Newell, N., Ward, R. B., Tucker, P. W., Strober, S., Blattner, F. 1982. Simultaneous expression of immunoglobulin m and d heavy chains by a cloned B-cell lymphoma: a single copy of the VH gene is shared by two adjacent CH genes. Proc. Natl. Acad. Sci. USA 79: 2996-3002

91. Maki, R., Roeder, W., Traunecker, C., Sidman, C., Wabl, M., Raschke, W., Tonegawa, S. 1981. The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. Cell 24: 353-65

92. Perlmutter, A. P., Gilbert, W. 1984. Antibodies of the secondary response can be expressed without switch recombination in normal mouse B cells. Proc. Natl. Acad. Sci. USA 81: 7189-93

93. Radbruch, A., Sablitzky, F. 1983. Deletion of Cμ genes in mouse B lymphocytes upon stimulation with LPS. EMBO J. 2: 1929-35

94. Katona, I. M., Urban, J. F., Finkelman, F. D. 1985. B cells that simultaneously express surface IgM and IgE in Nippostrongylus brasilensis - infected SJA/9 mice do not provide evidence for isotype switching without gene deletion. Proc. Natl. Acad. Sci. USA 82: 511-15

95. Chen, Y.-W., Word, C., Dev, V., Uhr, J., Vitetta, E., Tucker, P. 1986. Double isotype production by a neoplastic B cell line. II. Allelically excluded production of μ and γ1 heavy chains without CH gene rearrangement. J. Exp. Med. 164: 562-79

96. Nolan-Willard, M., Berton, M. T., Tucker, P. 1992. Coexpression of mu and gamma 1 heavy chains can occur by a discontinuous transcription mechanism from the same unrearranged chromosome. Proc. Natl. Acad. Sci. USA 89: 1234-38

97. Shimizu, A., Nussenzweig, M. C., Han, H., Sanchez, M., Honjo, T. 1991. Trans-splicing as a possible molecular mechanism for the multiple isotype expression of the immunoglobulin gene. J. Exp. Med. 173: 1385-93

98. Kuhn, T. S. 1962. The structure of scientific revolutions. Chicago: U. Chicago Press

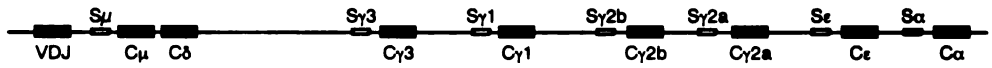


Figure 1 Organization of the immunoglobulin heavy-chain locus in a B lymphocyte.VDJ, gene segment encoding the variable region; Cμ, Cδ, $C\gamma_3$, $C\gamma_1$, $C\gamma_{2b}$, $C\gamma_{2a}$, Cε, Cα, gene segments encoding the constant region of the respective immunoglobulin heavy chains; S, switch regions located 5' to the C gene segments.

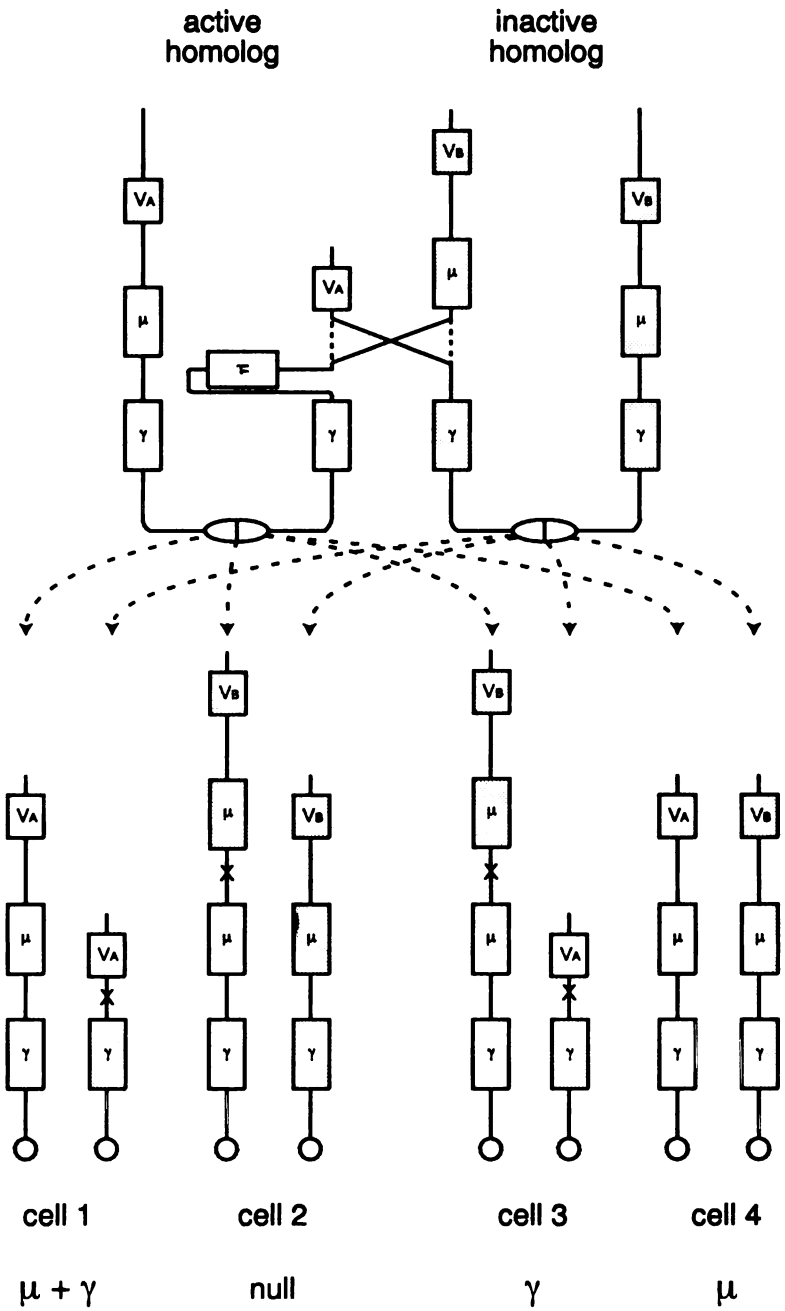


Figure 2 Unequal exchange between homologous chromosomes leading to class switch.VA, variable region of the active homolog; VB, variable region of the inactive homolog; m, m constant gene segment; γ, γ constant gene segment representing all other constant region segments (except δ); X, recombination point; cell 1 to 4, types of daughter cells of the switching cell; the type of heavy chain a cell expresses is given underneath. The split ovals represent the centromeres, holding the two chromatids of each homolog together. In the mitosis following the switch the four chromatids segregate randomly, creating four types if daughter cells with different genomic configurations. Each daughter cell has two homologs, with one chromatid each. Shaded boxes are the gene segments originally present on the inactive homolog.

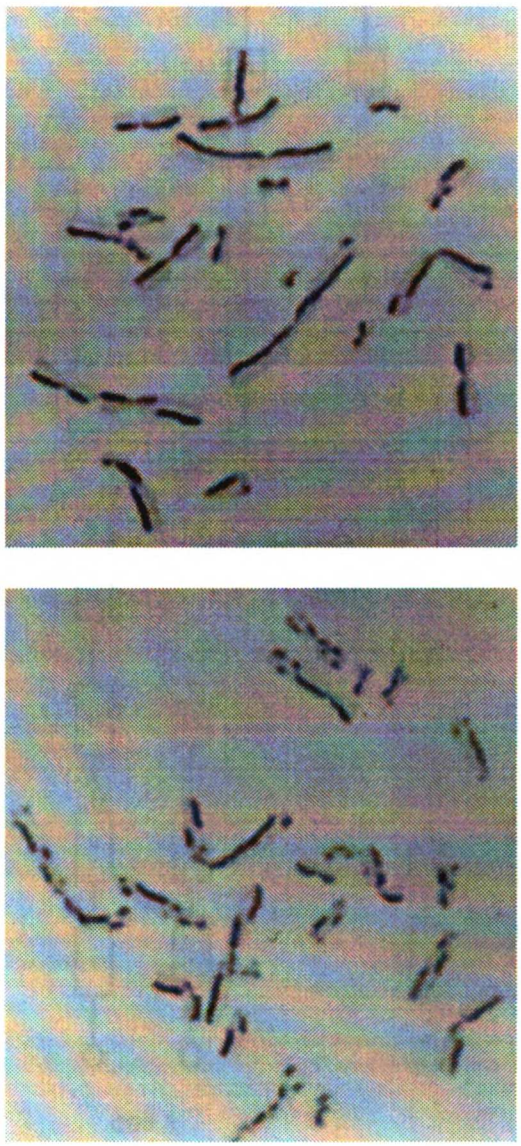


Figure 3 Photomicrograph of metaphase chromosomes with their sister chromatids stained differentially. Top: regular hamster chromosomes. Bottom: after exposure of the cell to a mutagen. Courtesy of Dr. S. Wolff, San Francisco.

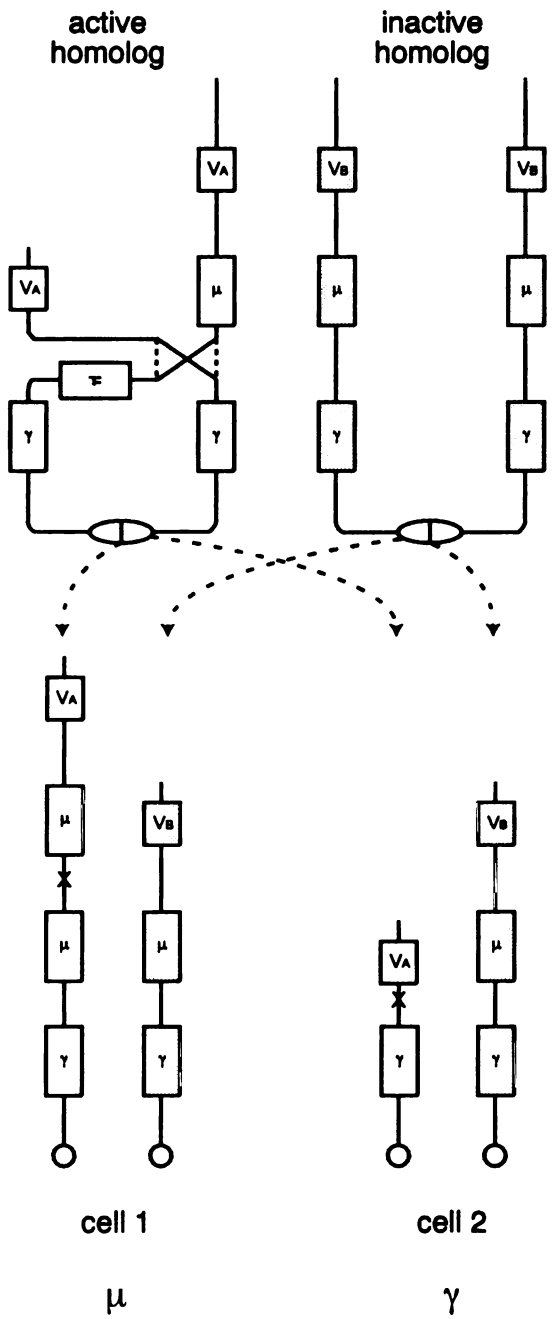


Figure 4 Unequal exchange between sister chromatids leading to class switch. For explanation of symbols, see Fig. 2. Segregation during mitosis creates two types of daughter cells: cell 2 has switched to γ chain production, cell 1 still produces μ chain but contains three Cμ gene segments.

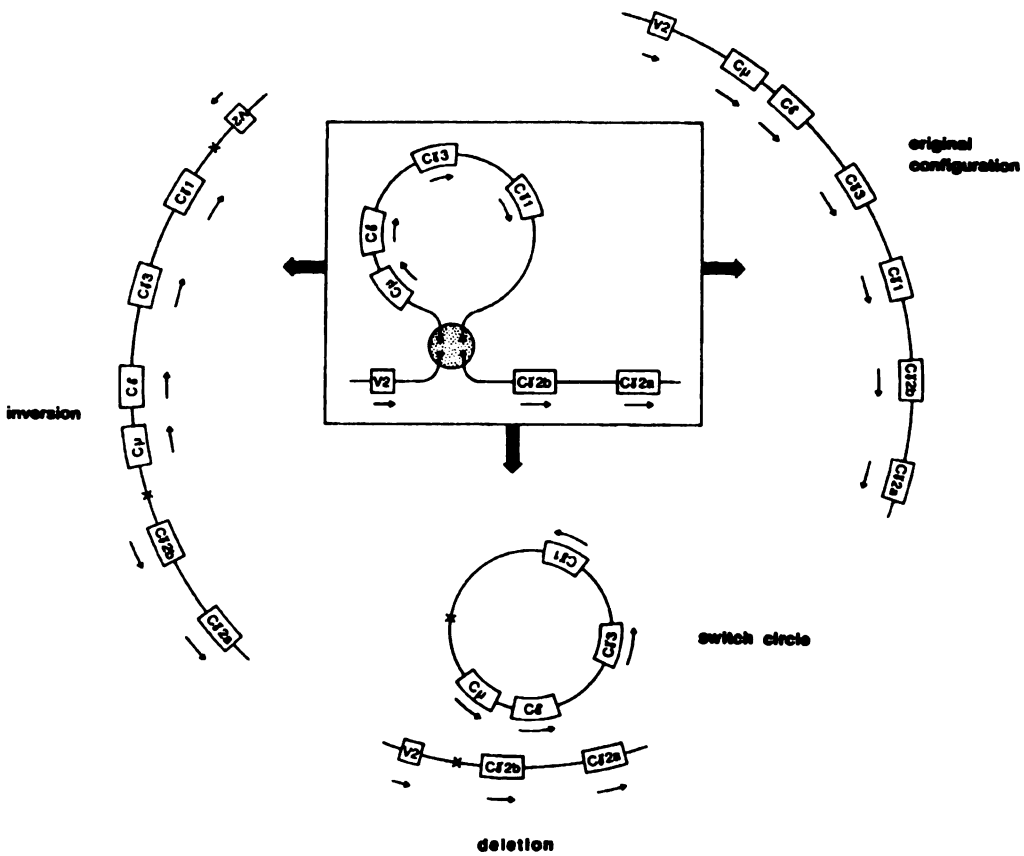


Figure 5 Looping out at the heavy-chain locus leading either to an inversion or to the original configuration or to deletion of a switch circle. X, recombination (break) point. Arrows under the gene segment define the direction of transcription. The stippled circle in the middle represents the putative enzyme complex holding the four free DNA ends together, which can be ligated in three different ways to create the chromosomal configurations shown on the sides and below. The inversion aborts any heavy chain production owing to the wrong transcriptional orientation, while the deletion of the switch circle allows expression of γ_{2b} chain.

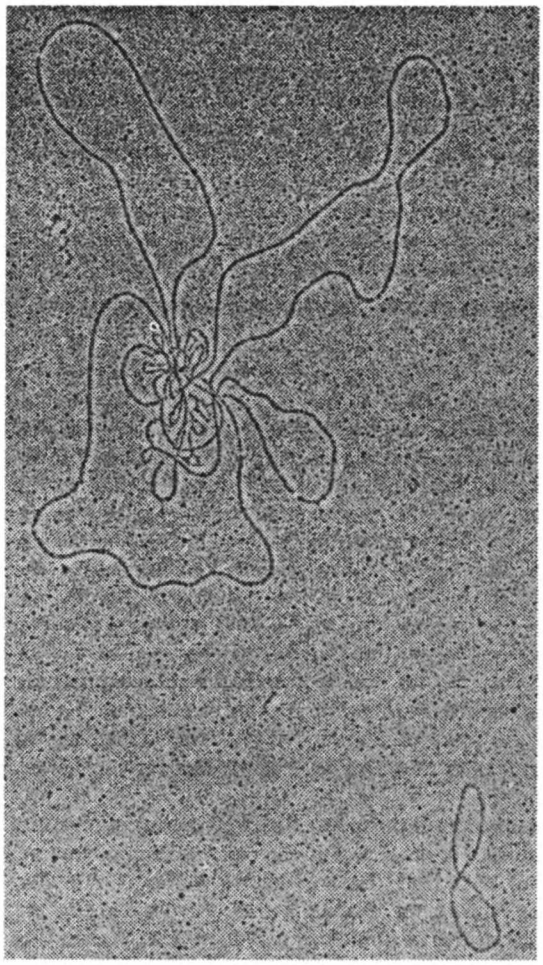


Figure 6 Electron micrograph of an isolated circular DNA molecules. The DNA molecule in the center is larger than 100 kb. The figure-eight shaped small molecule at the lower right represents a pBR 322 molecule and serves as a size marker.

CHAPTER 2

A RAPID ASSAY FOR DETECTING CELLULAR TdT ENZYMATIC ACTIVITY

as appearing in

JOURNAL OF IMMUNOLOGICAL METHODS
1995

Abstract

We have developed a solid-phase assay for the quantification of terminal deoxynucleotidyl transferase (TdT) enzymatic activity in crude cellular extracts. Affinity-purified, polyclonal anti-TdT antibodies are bound to the wells of a microtiter plate, and TdT in extracts is then bound to the immobilized antibodies. The enzymatic activity of the antibody-bound TdT is measured directly in the wells of the microtiter plate. This method yields highly reproducible values, even with samples of low activity. Because it is also technically very simple, it is ideal for determining enzymatic activity for large numbers of clones with limited numbers of cells.

Introduction

Because it is expressed only in T and B lymphocytes, terminal deoxynucleotidyl transferase (TdT) is of special interest for immunology. TdT is expressed only for a brief period of time in the normal course of development of lymphoid cells. It is responsible for the insertion of N regions in immunoglobulin heavy chains (Kallenbach et al., 1991; Gilfillan et al., 1993; Komori et al., 1993) and in T cell receptor chains (Gilfillan et al., 1993; Komori et al., 1993).

The purification of TdT from calf thymus was described over two decades ago (Chang and Bollum, 1970; Chang and Bollum, 1971). Subsequently researchers worked out scaled-down versions of this purification protocol in order to assess enzymatic activity for fewer cells (Coleman, 1977; Beutler and Kuhl, 1978; Mertelsmann et al., 1978). Scaling down was important if TdT activity were to be used routinely as a diagnostic tool in the identification of leukemia patients with acute lymphoblastic leukemia (ALL) (McCaffrey et al., 1975). Although the scaled-down procedures were theoretically feasible, none became universally accepted due to numerous technical difficulties and to the lack of reproducibility that was encountered in practice (Fleisher et al., 1987). In addition, relatively large numbers of cells ($>1x10^7$) were still required. The diagnosis of ALL became greatly facilitated with the availability of immunological reagents that reduced the TdT assay to simple immunofluorescence. However, immunofluorescence techniques are not sufficient for studying the *function* of TdT as opposed to the *presence* of a, possibly mutated, TdT protein chain. During the course of a program of gene transfer experiments, we needed to screen for and verify enzymatic activity in a large number of individual clones transfected with a TdT construct driven by an inducible expression system. We therefore developed a purification and

assay protocol that combines the specificity of immunological purification with a standard enzyme assay. This procedure is shown schematically in Fig. 1.

Materials and Methods

TdT assay. Immunlon-1 microtiter plates were incubated with 100 µl of a 1 mg/ml solution of purified protein A (Pharmacia) for 4 hr at room temperature, washed with PBSF (PBS + 1 mg/ml BSA fraction V from Sigma + 1% NaN_3), and incubated with PBSF for 2 hr at room temperature. The plates were again washed with PBSF and incubated overnight at room temperature with 100 µl of a 1:10 dilution of affinity-purified anti-TdT polyclonal antibody (Supertechs). The diluted antibody was then recovered for reuse, the plates were washed with PBSF, 100 µl of crude cellular extract was added, and the plates were incubated 4 hr at 4 °C. The extracts were then discarded, the plates were washed with TdT assay buffer (50 mM Tris pH 8.0, 0.8 mM $MnCl_2$, 0.1 mM DTT), 0.2 µg poly $dA_{40\text{-}60}$ (Pharmacia) and 1 µCi ^{33}P-dATP (Amersham) in 100 µl assay buffer were added, and the plates were incubated 3 hr at 37 °C. Triplicate samples of 10 µl were removed from each well and spotted onto Whatman GFC filters. After air drying, the filters were washed 4 times in ice-cold TCA solution (5% trichloroacetic acid, 20mM sodium pyrophoshate) followed by a single 70% ethanol wash. Incorporated label was measured by scintillation counting. Affinity-purified bovine TdT (Boehringer Mannheim) was used as a standard for TdT activity.

Cells and cell extracts. Cell lines used in these experiments were MOLT-3 and RPMI 8402 (both human T-ALL cell lines known to be TdT positive), 3T3 mouse fibroblasts, and D7.17, a 3T3 transfectant with inducible TdT expression (Harriman and Wabl, unpublished). Lysates were made by washing 1 to 10 million cells in PBS, resuspending in 100 µl PBSF + 0.5% Triton X-100, vortexing briefly, and incubating on ice 20 min, followed by another brief vortexing. Lysates were then spun 10 min at 13,000 rpm in a microfuge at 4 °C and supernatants were transferred directly to microtiter plates prepared as described above.

Results and Discussion

Initially the solid-phase TdT assay was tested with pre-purified TdT at various dilutions in the absence of cellular extract in order to obtain a standard curve. This standard curve was then compared to standard curves obtained for a typical aqueous phase TdT assay. As can be seen in Fig. 2, when the reduced intermolecular interactions of the solid-phase reaction are compensated for by

increasing the substrate oligo 10-fold and prolonging the reaction time, a quite satisfactory standard curve is obtained. In fact, the solid-phase assay turned out to be just as sensitive as its aqueous-phase counterpart. Furthermore, as can be seen in Fig. 3, the presence of crude cellular extract (of $1x10^7$ TdT-negative cells in PBSF + 0.5% Triton) during the binding step did not affect the efficiency of the outcome at any point. This is a significant advantage of the solid phase assay, because small amounts of crude extract or even contaminants that co-elute with TdT purified from an ion-exchange column can severely inhibit a TdT enzymatic reaction or cause background problems (Beutler and Kuhl, 1978). Moreover, as detergent did not interfere with the binding, it was possible to lyse the cells as described, without sonication or ultracentrifugation.

In order to determine levels of TdT activity in the cell lines, the assay was run as described above, and the solid-phase TdT standard curve was interpolated to obtain the number of units of TdT from the raw incorporation data. The results are shown in Fig. 4. The amount of TdT activity found, ca. 3 units per $1x10^8$ cells, in the RPMI 8402 cells is in good agreement with previously published results (Spigelmann et al., 1988; McCaffrey et al., 1975). A similarly good agreement was found with MOLT-3 cells . The untransfected 3T3 cells showed incorporation levels at or below background, while the transfected cell line, D7.17, had detectable levels of TdT in both the induced and uninduced states. This result is in agreement with RNA expression analysis and SDS gels, both of which show some leakiness in the system and a 5 to 10-fold increase after induction (Harriman and Wabl, unpublished).

TdT activity in RPMI 8402 cells was about five times higher than in the induced D7.17 cells. This result illustrates the importance of using actvity as well as serological assays, for immunoprecipitations performed with the

same antibody preparation used in this study showed roughly equal amounts of protein expression in RPMI 8402 and induced D7.17 cell lines (Harriman and Wabl, unpublished). We have not yet tracked down the reason for the discrepancy. Perhaps the human TdT of RPMI 8402 cells is bound with better kinetics under these conditions than the mouse TdT of the D7.17 cells. Alternatively, and more likely, human TdT might simply be more processive than mouse TdT under these assay conditions. It would be ideal to have a recombinant human or mouse TdT as a standard to eliminate any possible bias in the antibody preparation that could skew the standard curve, which was generated with purified bovine TdT in this study. Nonetheless, it is clear that TdT activity of at least three species can be assayed with this protocol.

References

Beutler, E. and Kuhl, W. (1978) An assay for terminal deoxynucleotidyl transferase in leukocytes and bone marrow. Am. J. Clin. Pathol. 70, 733-737.

Chang, L. M. and Bollum, F.J. (1970) Doxynucleotide-polymerizing enzymes of calf thymus gland. IV. Inhibition of terminal deoxynucleotidyl transferase by metal ligands. Proc. Natl. Acad. Sci. U.S.A. 65, 1041-1048

Chang, L. M. and Bollum, F.J. (1971) Deoxynucleotide-polymerizing enzymes of calf thymus gland. V. Homogeneous terminal deoxynucleotidyl transferase. J. Biol. Chem. 246, 909-916.

Coleman, M. S. (1977) A critical comparison of commonly used procedures for the assay of terminal deoxynucleotidyl transferase in crude tissue extracts. Nucleic Acids Res. 4, 4305-4312.

Gilfillan, S., Dierich, A., Lemeur,M., Benoist, C. and Mathis D. (1993) Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science, 261, 1175-1178.

Fleisher M., Stankievic, R., Schwartz, D. and Schwartz, M.K. (1987) Solid-phase enzyme immunoassay of terminal deoxynucleotidyl transferase evaluated. Clinical Chem. 33, 293-296.

Kallenbach, S., Doyen, N., Fanton d'Andon, M. and Rougeon,. F. (1992) Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. Proc. Natl. Acad. Sci. U.S.A. 89, 2799-2803.

Komori, T., Okada, A., Stewart, V., Alt, F.W. (1993) Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. Science 261, 1171-1175.

McCaffrey, R., Harrison T.A., Parkman R. and Baltimore, D. (1975) Terminal deoxynucleotidyl transferase activity in human leukemic cells and in normal human thymocytes. New England J. Med. 292, 775-780.

Mertelsmann, R., Mertelsmann, I., Koziner, B., Moore, M., Carkson, B. (1978) Improved biochemical assay for terminal deoxynucleotidyl transferase in human blood cells. Leukemia Res. 2, 57-69.

Spigelman, Z., Duff, R., Beardsley, G.P., Broder, S., Cooney, D., Landau, N.R., Mitsuya, H., Ullman, B., McCaffrey, R. (1988) 2',3'-Dideoxyadenosine is selectively toxic for TdT-positive cells. Blood 71, 1601-1608.

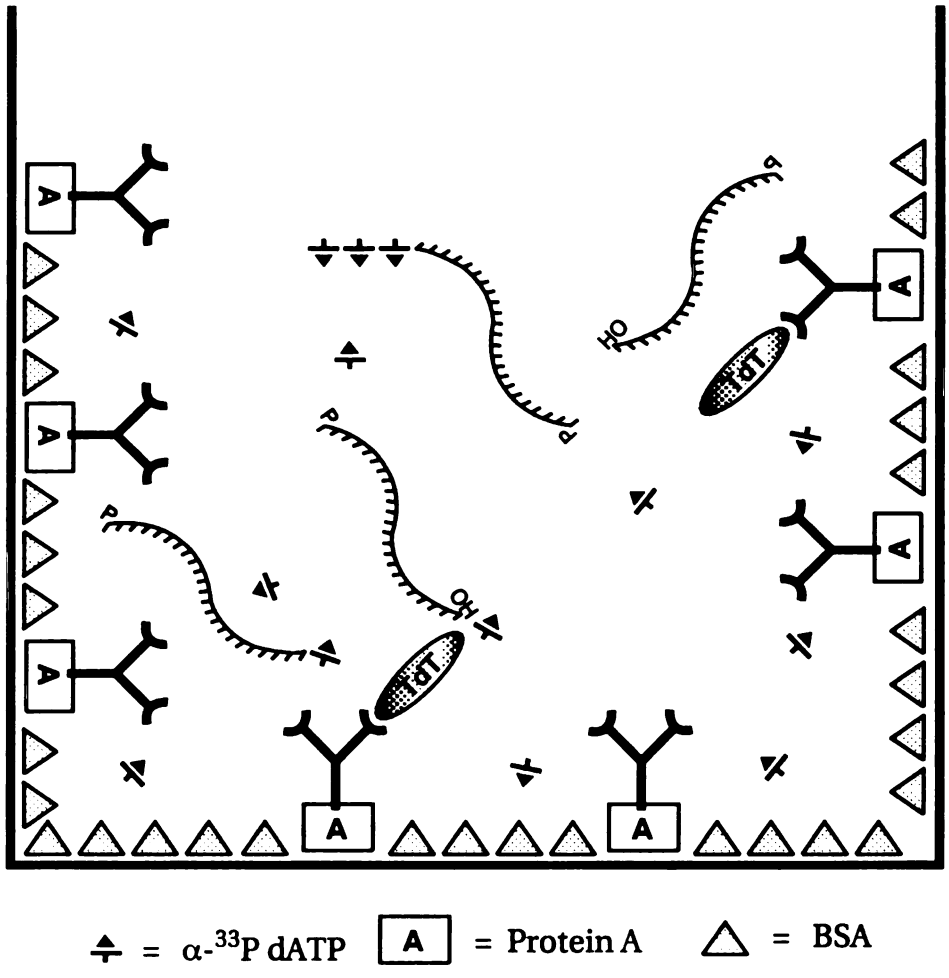


Figure 1 Schematic representation of solid-phase enzymatic assay

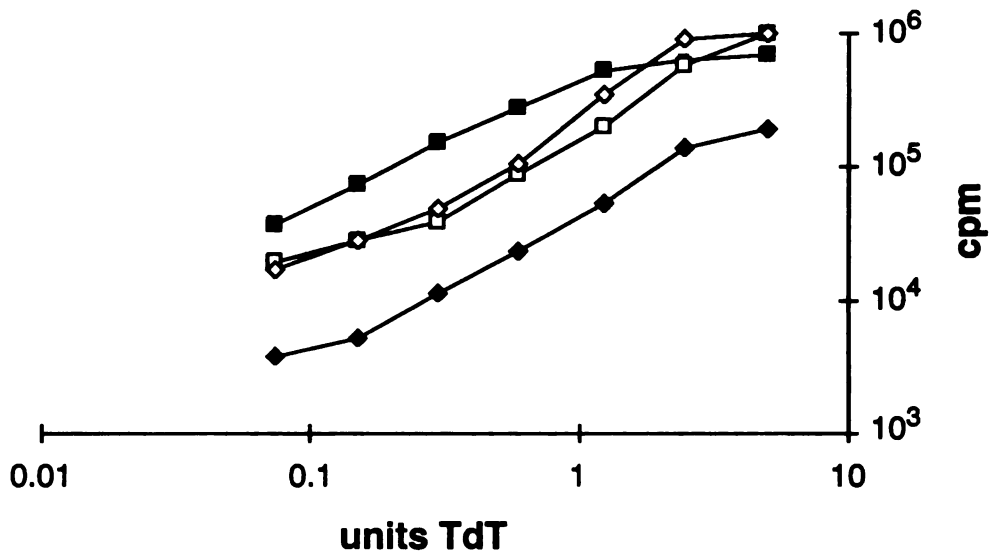


Figure 2 A comparison of solid-phase and aqueous-phase TdT enzymatic reactions. Aqueous-phase assays are shown with 0.02 μg poly-dA_{40-60} at 1 hr (open squares) and 3 hr (open diamonds). Solid-phase assays are shown at 3 hr with 0.02 μg (closed diamonds) and 0.2 μg (closed squares) poly-dA_{40-60}. Values indicated are averages of triplicate samples, minus background. Absolute background levels were 9000 ± 500 cpm and 2400 ± 300 for aqueous and solid-phase assays, respectively. The standard error was less than 5% for all points above 1 unit TdT, and less than 10% for all points below 1 unit. Abscissae: units of pre-purified TdT used in the assay. Ordinates: Counts per minute of isotope incorporated in the assay.

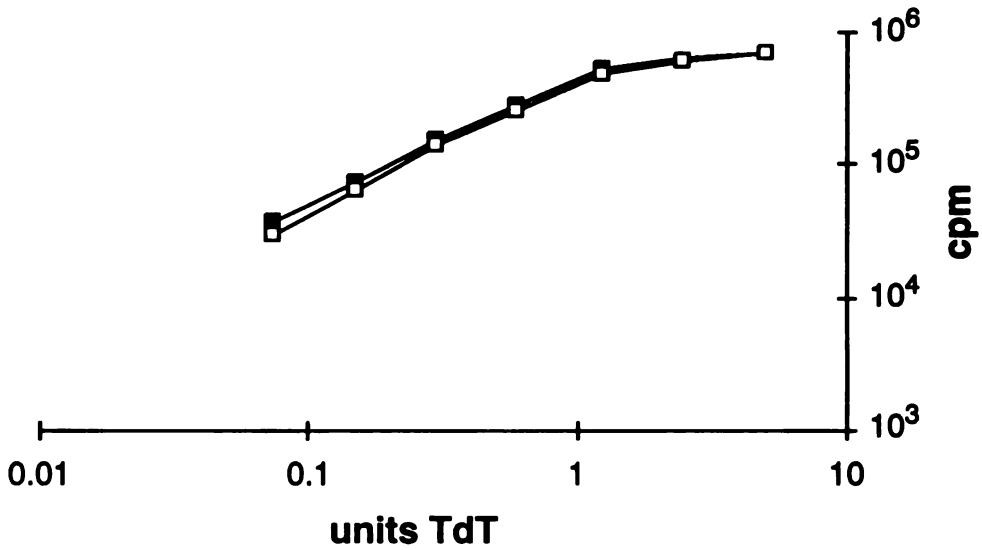


Figure 3 Lack of inhibition of the solid-phase assay by cellular extract. Open squares represent a solid-phase TdT assay carried out in the presence of an extract of 1 x 10^7 TdT-negative 3T3 cells. Solid squares represent an otherwise identical assay carried out in the absence of the 3T3 extract. Values indicated are averages of triplicate samples, minus background. Absolute background was measured at 2400 ± 300. The standard error was less than 5% for all points above 1 unit TdT, and less than 10% for all points below 1 unit. Abscissae: units of pre-purified TdT used in the assay. Ordinates: Counts per minute of isotope incorporated in the assay.

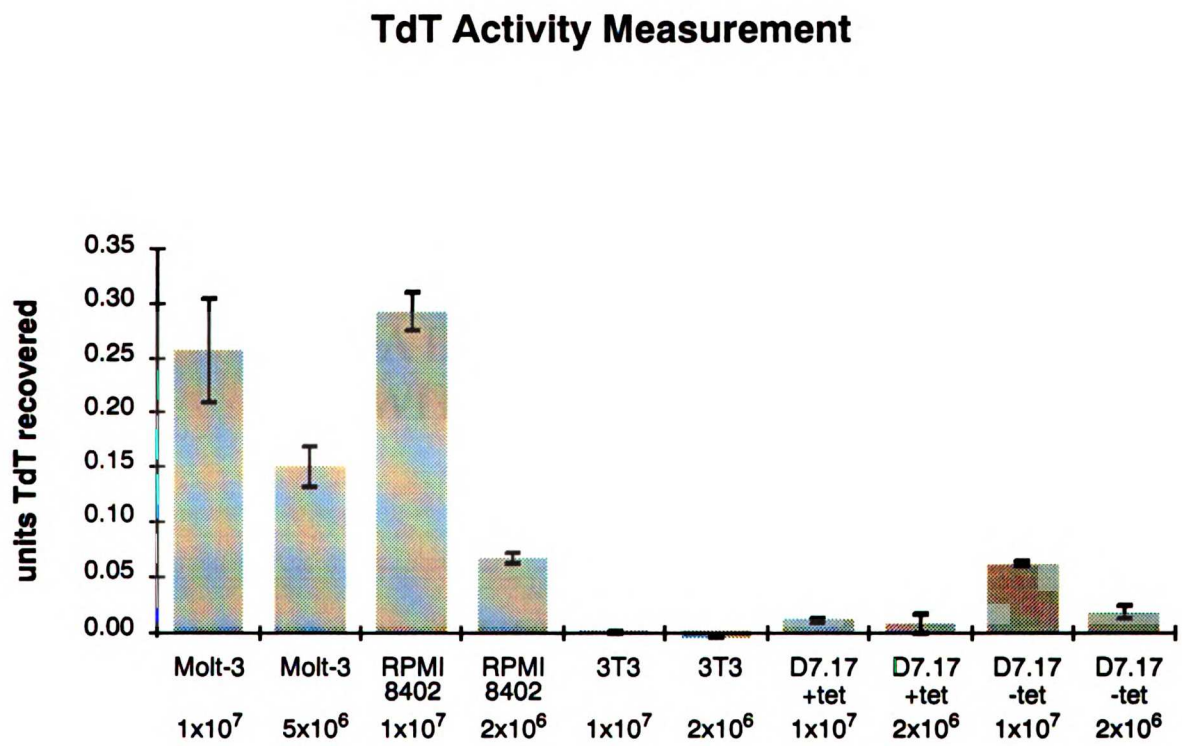


Figure 4 TdT activity in cell lines. TdT was assayed in extracts of various cell lines at the specified cell numbers. The values are units of TdT calculated from a solid-phase standard curve like that shown in Fig 3. MOLT-3 and RPMI 8402 are TdT-positive, human T-ALL lines. The TdT-negative, mouse fibroblast line 3T3 was stably transfected with a tetracycline-inducible TdT construct to obtain D7.17.

CHAPTER 3

A VIDEO TECHNIQUE FOR THE QUANTIFICATION OF DNA IN GELS STAINED WITH ETHIDIUM BROMIDE

as submitted to

ANALYTICAL BIOCHEMISTRY

1995

Abstract

We have developed a method for the comparison of quantitative differences between nucleic acids samples run on agarose gels. It is useful for the analysis of digested genomic DNA, RNA preparations, and CHEF gel analysis of damaged chromosomal DNA. The technique utilizes a standard color (RGB) video camera, a microcomputer, and commercially available software. While not as elaborate as other image analysis systems, our method is suitable for many purposes and can be set up for a relatively low cost.

Introduction

Various pulsed-field electrophoretic techniques have been used to separate high molecular weight DNA (200 kb to 10 Mb) which have previously been difficult to visualize [1] [2]. Since agents which induce DNA double-strand breaks in eukaryotic cells often produce fragments in this range, pulsed-field gel electrophoresis (PFGE) has been employed extensively to quantify DNA damage and repair. Generally, the extent of DNA damage is measured by a comparison of the amount of DNA which has migrated into the lane to the amount of DNA which is retained in the well (plug) [3]. Cellular DNA can be labeled radioactively, and after electrophoresis portions of the agarose gel can be excised and the embedded DNA quantitated with a scintillation counter. Measurements can be expressed in relative terms, such as percent released from plug [4][5], and migration distance [2][6][7]. This method has its drawbacks in the sense that cells need to be radiolabeled (which is difficult in some cases), and that the gel is necessarily destroyed, preventing further processing such as southern blot analysis.

Consequently, investigators are increasingly turning to fluorescence measurements as a means for DNA quantitation in PFGE. It has been shown that ethidium bromide stained DNA illuminated with UV light fluoresces proportionally to DNA concentration, and the fluorescence can be measured either by film or camera. Film has the disadvantage that the extent of exposure is not linearly proportional to fluorescence intensity [8][9]. On the other hand, a variety of digital cameras have been shown to give a good linear response [10][11][12]. Results from a fluorescence/video-based PFGE analysis can correlate well with those obtained through radioactivity analysis, provided that certain correction factors are used [11]. Correction factors are necessary to address two inherent difficulties in video image analysis of EtBr stained gels:

fluorescent light scattering and saturation. Other problems, such as non-uniform staining, high background, and non-uniform UV illumination, are equipment and procedure related. These problems need to be addressed on a case-by-case basis, and have been discussed in depth elsewhere [9] [11]. Regarding scattering, Dewey et al. have empirically derived a formula which corrects for plug fluorescence "spilling over" into the lane, and they generate a good agreement between their video quantitation and radioactivity quantitation results. Saturation problems were addressed by the utilization of neutral density filters to reduce the overall transmittance of light.

Our approach is from a considerably different angle. Instead of trying to simulate the situation with radioactivity by summing up all of the fluorescence within a given plug or lane, we simply measure the intensity of the median-value pixel of the plug or lane in question. This value is not significantly altered by variations in the fluorescence of adjacent lanes and can provide a very wide range of measurable DNA concentrations from a single scan. With two scans, virtually the entire visible concentration range of ethidium bromide stained DNA can be accurately measured.

Materials and Methods

Agarose Gels Running Conditions, Ethidium Bromide Staining. For digested genomic DNA analysis, 100µg mouse tail DNA was digested overnight with EcoRI, extracted twice with phenol and once with 1:1 phenol/chloroform, and precipitated with ethanol. The pellet was resuspended in 100 µl TE buffer and an aliquot was taken for A280 measurement and subsequent quantitation. Various dilutions were made from the digest and mixed with DNA loading buffer (final concentration 2.5% Ficoll, .025% bromphenol blue). Samples were loaded in triplicate onto 15x15 cm 0.8% agarose (Boehringer Mannheim) gels approximately 1 cm thick, the wells having been formed with a 15-well comb. The gels were run in 1x TBE buffer in a standard electrophoresis chamber at 60 V for 4 hr. Gels were stained with 3µg/ml EtBr in TBE for 2 hours, destained in TBE 4 hr, with the buffer being changed twice.

Cultured NIH 3T3 cells were used for CHEF gel analysis. Exponentially growing cells

were harvested, washed in PBS, resuspended at various concentrations in PBS at 22°C, and mixed 1:1 with 1.5% low-gelling temperature agarose (FMC SeaPlaque), and injected into 3mm diameter plastic tubing, and placed on ice. After solidifying, the embedded cells were blown out from the tubing and cut into 5 mm cylindrical plugs and irradiated on ice with a Cs-137 source at a dose rate of 2 Gy/min. Plugs were then either immediately immersed in 0.5 ml of cell lysis buffer (0.5M EDTA, 0.01M Tris pH 8.0, 2% NLS, 0.1 mg/ml Proteinase K) and incubated O/N at 42°C, or placed in media at 37°C for 2 hr prior to lysis. The next day RNase A was added to a concentration of 0.1 mg/ml and the samples were incubated 1 hr more at 22°C. Plugs were then placed into wells of a 0.8% agarose gel, which was run in a CHEF apparatus (BioRad CHEF-DR II) with 0.5x TBE buffer at 4°C. Generally, gels were run for 50 hr at 60V with an initial and final pulse time of 75 min. After running, gels were stained with 3µg/ml EtBr in 0.5x TBE for 4 hours, destained in 0.5x TBE O/N, with two buffer changes.

UV Illumination, Camera, Computer Hardware and Software. Gels were placed onto a UV transilluminator (Fotodyne Foto/Prep I, model 3-3500) situated approximately 10 inches below the lens of a vertically adjustable mounted video recorder (Sony CCD V101). An UV absorbing orange filter was placed in front of the lens. The 72 dpi S-Video output of the camera was fed into a RasterOps 364 Video Card running inside a Apple Macintosh IIci with 8 MB RAM. The gel could be visualized on the computer monitor, and after adjusting size, focus and brightness (by manually adjusting the iris on the camera), an image was "grabbed" with FrameGrabber 3.0, software which was shipped with the video card. Two images, one bright and one dark, were scanned for each gel. The images were saved as 24-bit color TIFF's. They were then opened in Adobe Photoshop 2.51 (in RGB mode) and cropped to the borders of the gel, yielding dimensions of approximately 600x300 pixels for a 14x7 cm gel area. The dark scan was used for plug quantitation and the bright scan for lane quantitation. A 5x15 pixel box was saved as a selection path, this path was centered upon the area corresponding to each plug (fig 1). Each path then was individually selected and the "Histogram" command was executed. The histogram window displays a chart showing the distribution of selected pixels with respect to their color intensities (on a scale of 0 to 255). This window also displays average pixel value and median pixel value for each color. The median pixel value, or MPV, of the Red (R) channel was then recorded for each selection. For the lanes of the CHEF gels, and for the digested DNA gels, larger selection boxes were taken- generally starting from 3-5 pixels out from the lane and extending through the area of the readily visible stained DNA (fig 1). The boxes were 10 to 15 pixels wide and from 40 to 80 pixels long.

Results

System Calibration

An EtBr stained gel with no DNA loaded was used to determine whether an uneven background would be a problem when collecting R-MPV. Plug-box sized selections and lane-box sized selections were made at various areas of the gel and R-MPV was recorded. Consistent measurements (±4 fluorescence units) were made at all iris settings and in all locations, provided the selection was 1 cm (20 pixels) or more from any gel edge. Next the potential problems of plug-to-lane scattering and adjacent lane scattering were addressed. Triplicate samples of $6x10^6$ cells per/ml were irradiated with 50 Gy and run under the CHEF conditions described above. After staining, lane R-MPV was recorded at 87 ± 6, then the plugs were removed and R-MPV was recorded at 89 ± 5. Thus plug-to-lane scatter did not significantly affect the R-MPV readings. Additionally, adjacent-lane scatter was also determined not to pose a serious problem, as background measured in lane and plug areas of an adjacent empty lane was not significantly increased in comparison to an empty lane four lanes away from any sample.

When first trying to determine whether the use of our equipment would be feasible for estimating DNA concentration in general, we calibrated the system by running purified digested genomic DNA by standard electrophoresis. The logic was that purified DNA is easily quantitated by A_{280} measurement and that gel loading would be more consistent than could be achieved with plug-embedded DNA. We allowed the DNA to migrate only a short distance so that it would best mimic the appearance of irradiated cell DNA after a 50 hr CHEF run. Fig 2 shows the plot of MPV of each R, G and B channels for lanes containing digested genomic DNA ranging from 0.1 to 10 µg. Also shown is the Gray channel, which is the summation of the RGB channels,

and is what would be measured with a black and white camera. It is clear that the red channel delivers the best range and sensitivity (which is logical considering the color of EtBr-stained DNA), and was therefore the channel used for quantitation in our study.

The gel described above was scanned a total of six times, each with a slightly different iris adjustment which ranged from scan 1 (severely underexposed) to scan 6 (severely overexposed). The R-MPV is plotted versus amount of DNA loaded in Fig 3. Scans 1 and 2 show markedly reduced sensitivity while scan 6 shows saturation at the highest DNA concentration. Linearity was noted among all the scans and the range of the middle scans (4&5) was remarkable. It was also noted that at the highest concentration point in scan 5, a significant number (ca 25%) of pixels were saturated, but the median pixel was not, and linearity was maintained.

CHEF Gel Analysis

We next set out to determine whether the same proportionality which existed between the R-MPV and digested genomic DNA also held for damaged DNA extracted from lysed cells through CHEF electrophoresis. It was also important to determine whether plug-contained DNA could be quantitated by the R-MPV technique. Fig. 4a shows R-MPV for serial dilutions of purified genomic DNA embedded in agarose plugs and placed into the slots of a gel and EtBr stained without electrophoresis. Fig 4b shows a similar plot but with dilutions of lysed cells rather than purified DNA. A plot representing this same gel after CHEF electrophoresis is shown in fig 4c, where both plug and lane values are seen.

Interpretation of R-MPV Data

It became clear that a single scan would not have enough range to be accurate for plug quantitation and yet sensitive enough to measure low yield lanes. Therefore, for each CHEF gel two scans were taken: one for plugs and one for lanes. The next step was to link the two scans so that we could ultimately obtain in the common unit of Percent Released (%Rel). We envisioned a formula such as:

$$\%Rel = \text{Lane R-MPV}/(kc)\text{Plug R-MPV} + \text{Lane R-MPV}$$

where k represents a constant linking the two scans and c a constant reflecting differences in fluorescence dynamics between the plug and lane. To determine whether a k value could be assigned reliably, we compared individually each of the six scans represented in fig 1 to each of the others in the group and calculated the k values based upon each concentration data point (except those with R-MPV's of 0 or 255). Average k values ± sd are shown for each scan comparison in fig 5. Clearly, reasonable estimates of k values can be obtained regardless of which point is used for the calculation, although points which are not immediately proximal to the detection limits of either scan would be preferred.

The c value calculation, however, proved to be more problematic. It varied with respect to number of pixels selected for histogram analysis, amount of DNA loaded, and DNA migration distance, among other factors. Apparently the only way to produce a %Rel value with our system would be to include known concentration standards in each gel. This would be technically feasible, but would diminish the simplicity we desired from our system.

We thus attempted to create an index in which the k and c values

would not need to be calculated yet could still provide some relevant information on DNA damage and repair. Knowing that R-MPV is proportional to DNA concentration in both plugs and lanes, we can compare the plug R-MPV of one sample to the plug R-MPV of another. Likewise, R-MPV from the lanes of different samples are comparable. In addition, the ratio of lane R-MPV: plug R-MPV is comparable between samples. A "Damage Index" could be set up as the following:

$$DI_x = \text{Lane R-MPVx}/\text{Plug R-MPVx}$$

where DIx represents the Damage Index at certain irradiation point x (fig 6). For comparisons between samples on a different gels scanned at different times, calculation of k (but not c) would be required. A plot of the DI versus radiation dose for NIH 3T3 cells, before and after repair, is shown in fig 7. The slopes of these curves are a reflection of double strand break induction and repair, respectively. Such analysis should prove useful for comparisons between different cell samples.

Discussion

We demonstrate in this study that a parameter of a video image, R-MPV, can be used to show quantitative relationships between different DNA samples which have been electrophoresed on an agarose gel and stained with ethidium bromide. With the inclusion of appropriate concentration standards, DNA or RNA samples of unknown concentration can be quantitated easily before southern or northern hybridization. Such information would be useful in the determination of gene copy number and gene expression levels. The most significant application for the system described here would however be in

the area of the estimation of DNA double-strand breaks and their repair, particularly because of the potential difficulties in quantitating such DNA.

The system described here has its limitations, one of which is that molecular weight data is not produced. While such data would be theoretically obtainable by dividing our "lane box" into numerous horizontal boxes, we undoubtedly would run into scattering problems for high molecular weight DNA migrating just outside the plug, and would have to employ correction factors as described (11). Scattering is not a problem in our system, presumably because the median pixel lies outside of the area affected by plug-to-lane scattering. DNA is simply assigned to either plug or lane with no further delineation. Additionally, the system is based on relative comparisons of like areas between samples, not on the summation of all fluorescence emanating from a given sample. Therefore our results are not presented in terms of the standard "Percent Released", but rather in the relative terms of a Damage Index, DI. Percent released is a convenient value when calculated by measuring radiolabeled DNA, but less so for video-based analysis, where complex corrections need to be employed. In many instances the Percent Released provides no more useful information than the Damage Index described here, which is better suited for video analysis.

Despite its limitations, this system has proven itself useful for DNA double strand break analysis in our laboratory, and is reliable, simple, and inexpensive.

References

1. Ahn, S.Y., B. Nevaldine, and P.J. Hahn, Direct measurement by pulsed-field gel electrophoresis of induction and rejoining of X-ray-induced double-strand breaks in cultured mouse cells. Int J Radiat Biol, 1991. 59(3): p. 661-75.

2. Lawrence, T.S., et al., The use of biphasic linear ramped pulsed field gel electrophoresis to quantify DNA damage based on fragment size distribution. Int J Radiat Oncol Biol Phys, 1993. 27(3): p. 659-63.

3. Stamato, T.D. and N. Denko, Asymmetric field inversion gel electrophoresis: a new method for detecting DNA double-strand breaks in mammalian cells. Radiat Res, 1990. 121(2): p. 196-205.

4. Kysela, B.P., B.D. Michael, and J.E. Arrand, Relative contributions of levels of initial DNA damage and repair of double strand breaks to the ionizing radiation-sensitive phenotype of the Chinese hamster cell mutant, XR-V15B. Part I. X-rays. Int J Radiat Biol, 1993. 63(5): p. 609-16.

5. Iliakis, G.E., et al., Detection of DNA double-strand breaks in synchronous cultures of CHO cells by means of asymmetric field inversion gel electrophoresis. Int J Radiat Biol, 1991. 59(2): p. 321-41.

6. Ager, D.D., et al., Measurement of radiation-induced DNA double-strand breaks by pulsed-field gel electrophoresis. Radiat Res, 1990. 122(2): p. 181-7.

7. Garwood, D.P., L.L. Thompson, and W.C. Dewey, Use of pulsed-field gel electrophoresis to measure X-ray-induced double-strand breaks in DNA substituted with BrdU. Radiat Res, 1991. 128(2): p. 210-5.

8. Freeman, S.E., L.L. Larcom, and B.D. Thompson, Electrophoretic separation of nucleic acids: evaluation by video and photographic densitometry. Electrophoresis, 1990. 11(5): p. 425-31.

9. Ribeiro, E.A., L.L. Larcom, and D.P. Miller, Quantitative fluorescence of DNA-intercalated ethidium bromide on agarose gels. Anal Biochem, 1989. 181(2): p. 197-208.

10. Sutherland, J.C., et al., Electronic imaging system for direct and rapid quantitation of fluorescence from electrophoretic gels: application to ethidium bromide-stained DNA. Anal Biochem, 1987. 163(2): p. 446-57.

11. Dewey, W.C., et al., A charge-coupled-device camera image analysis system for quantifying DNA distributions in agarose gels after pulsed-field gel electrophoresis. Radiat Res, 1994. 140(1): p. 37-47.

12. Rosemann, M., et al., An image analysis technique for detection of radiation-induced DNA fragmentation after CHEF electrophoresis. Int J Radiat Biol, 1993. 64(2): p. 245-9.

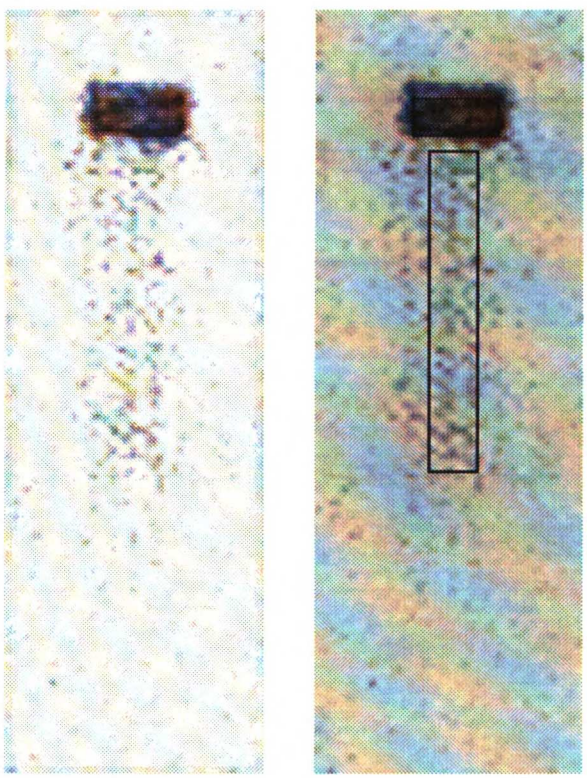
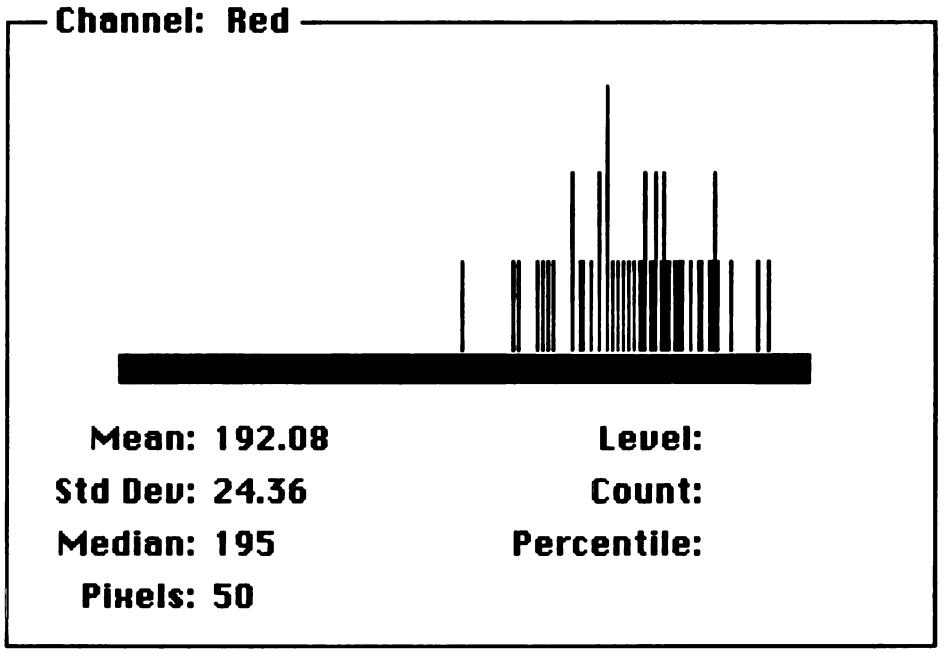
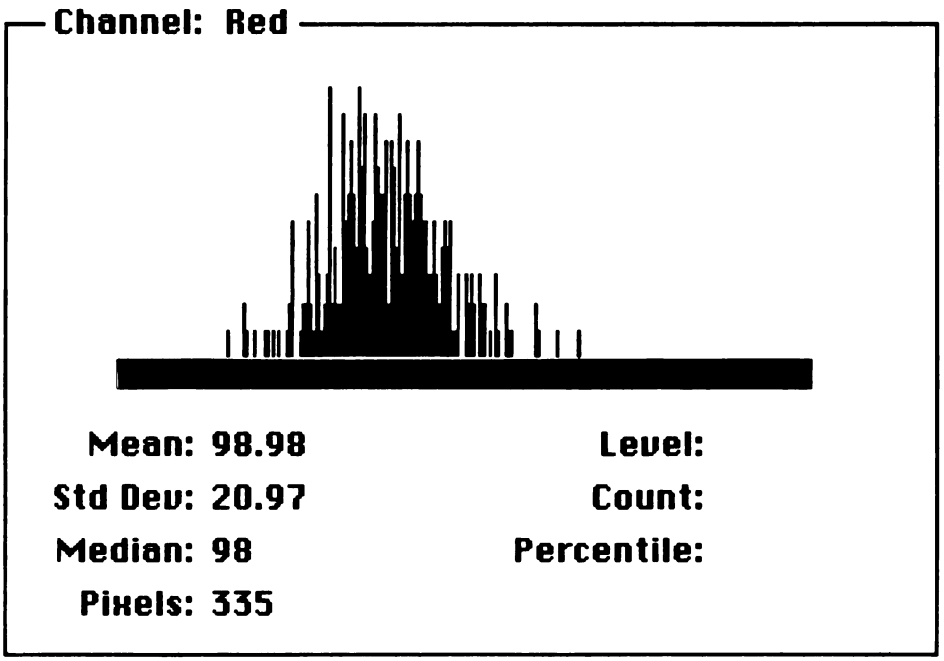


Figure 1 A sample section of the digitized image captured as described in methods. For display purposes the red channel was separated and then inverted so that red shows here as black. (a) CHEF gel sample of $3x10^6$/ml 3T3 cells after 50 Gy irradiation. (b) The same sample with selection paths marked. Also shown are the histograms generated from the plug selection (c) and lane selection (d) of the displayed sample.

1 c


Channel: Red
Mean: 192.08
Std Dev: 24.36
Median: 195
Pixels: 50
Level:
Count:
Percentile:

1 d


Channel: Red
Mean: 98.98
Std Dev: 20.97
Median: 98
Pixels: 335
Level:
Count:
Percentile:

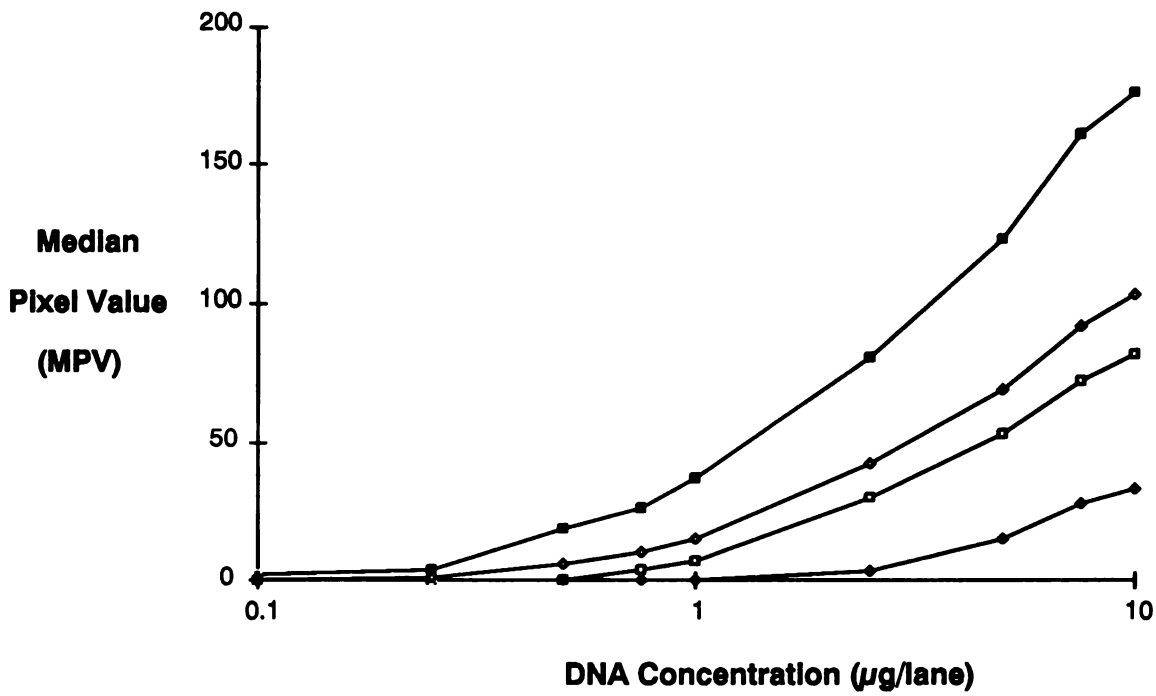


Figure 2 The Red channel provides superior dynamic range and sensitivity for the quantitation of ethidium bromide stained DNA. Video image analysis of serial dilutions of restriction enzyme digested genomic DNA electrophoresed, stained and scanned as described. DNA concentration is shown to be proportional the median pixel value of each channel. The Red channel is indicated by closed squares; Green by open squares; Blue by closed diamonds; composite grayscale channel by open diamonds.

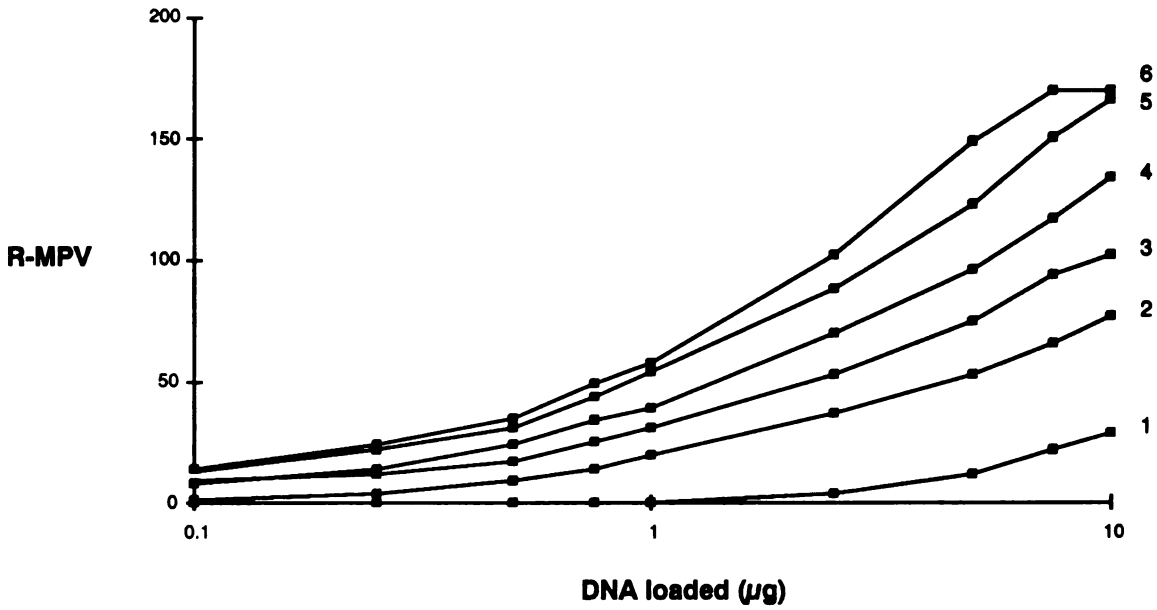


Figure 3 Proportionality is maintained over a wide range of scanning intensities. Video image analysis of serial dilutions of restriction enzyme digested genomic DNA electrophoresed, stained and scanned as described. A single gel was scanned at six different iris settings, resulting in scans with incremental levels of brightness, ranging from underexposed (scan 1) to overexposed (scan 6).

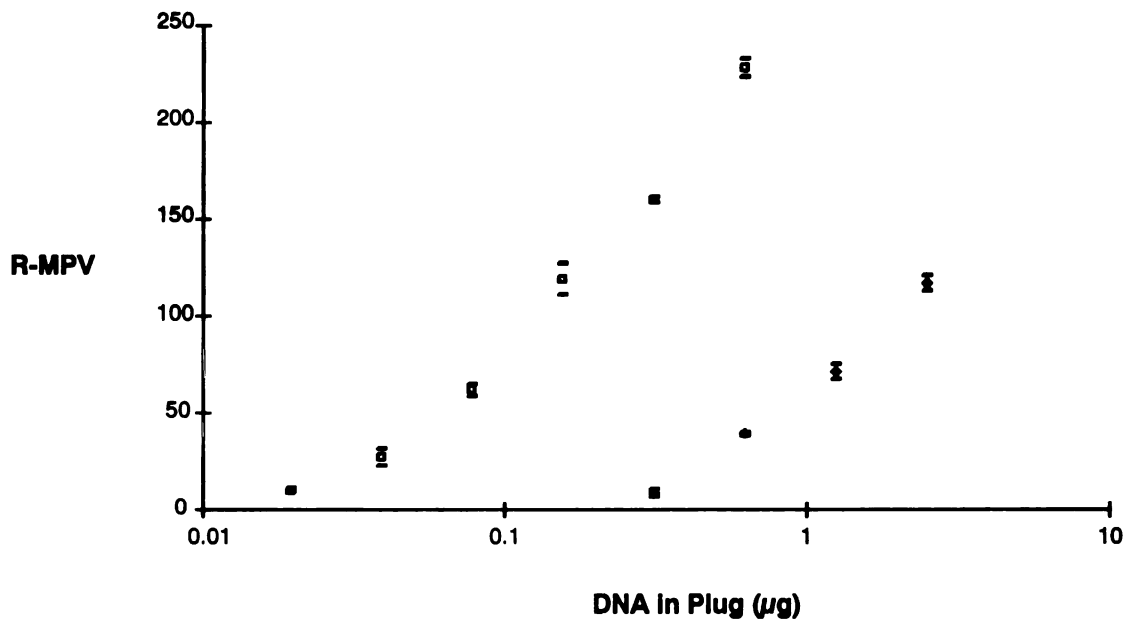


Figure 4a Red channel median pixel value, R-MPV, is proportional to agarose-embedded DNA concentration. Two scans were taken: one high-sensitivity scan (squares) and one low-sensitivity scan (diamonds). Values shown are with background subtracted and error bars.

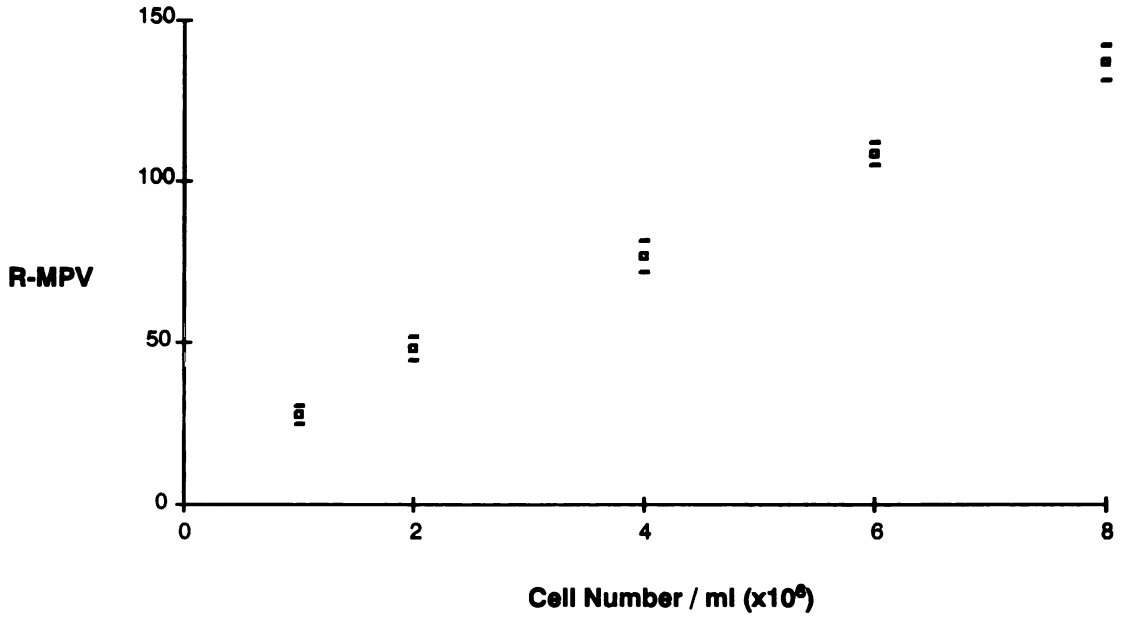


Figure 4b Red channel median pixel value is proportional to agarose-embedded NIH 3T3 cell number. Values shown are with background subtracted and error bars.

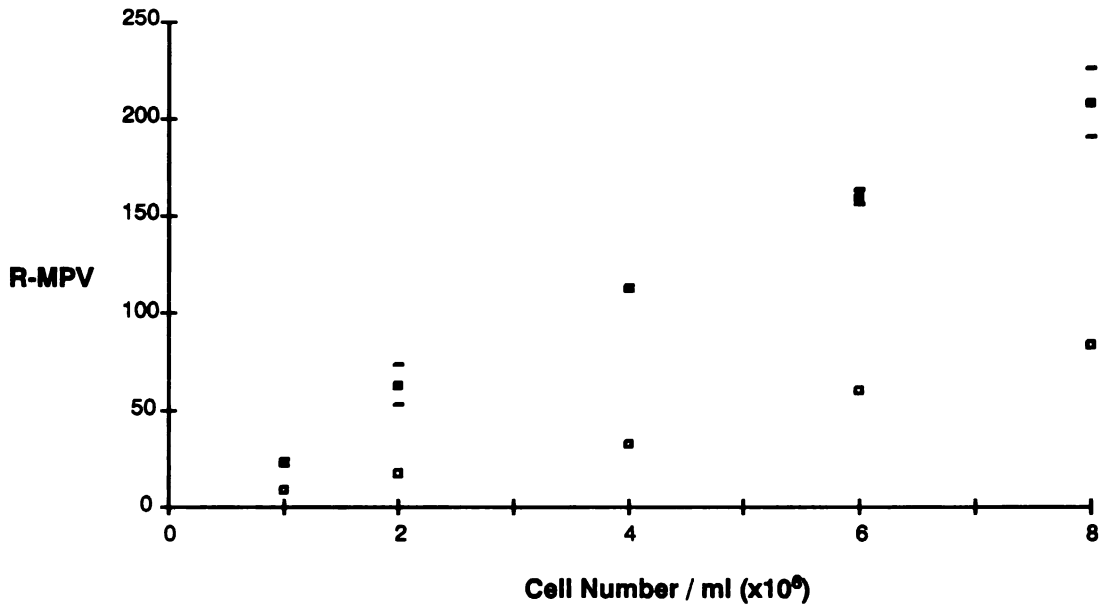


Figure 4c R-MPV analysis can be used to assess damaged DNA fractionated by CHEF electrophoresis. Red channel median pixel value of plugs (closed squares) and lanes (open squares) versus agarose-embedded NIH 3T3 cell number. Cells were irradiated with 50 Gy prior to CHEF electrophoresis. The standard errors of the lane values were never larger than the range depicted by each icon.

3	4	5	6	
0.66± .07	0.51 ± .05	0.39 ± .05	0.32 ± .04	2
	0.77 ± .03	0.59 ± .01	0.51 ± .01	3
		0.74 ± .03	0.65 ± .02	4
			0.89 ± .02	5

Figure 5 k value determination. The k value refers to a constant which allows quantitative information to be compared between different scans. In this case the six scans described in figure 3 were compared to each other over every measurable data point. For example, the k value of "6:5" was obtained by dividing the R-MPV of each DNA concentration point of scan 5 by its corresponding R-MPV in scan 6. Shown are average k values.

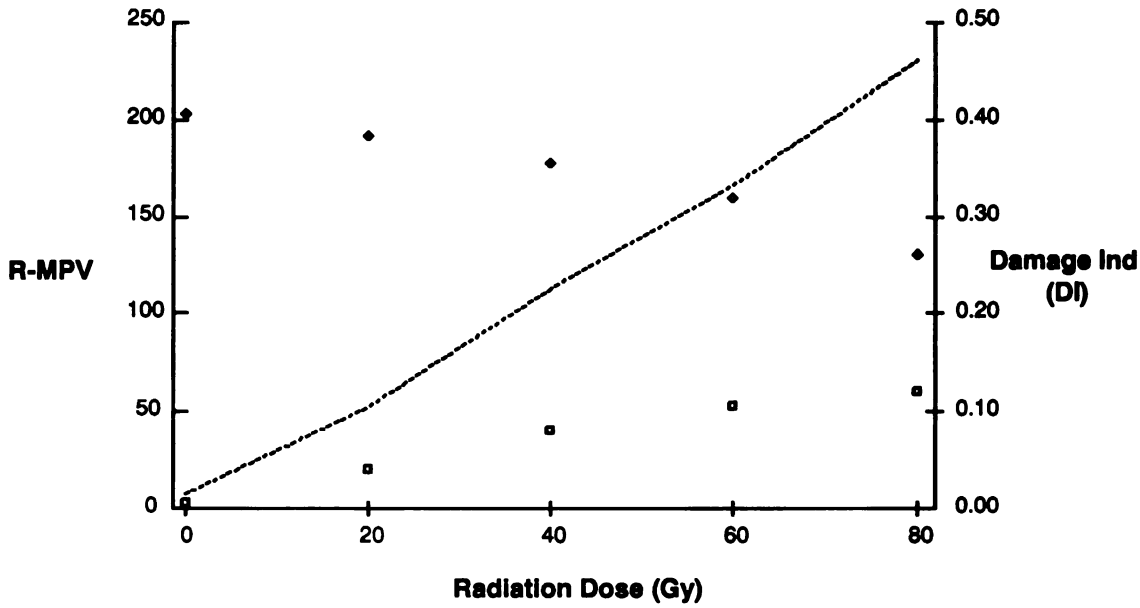


Figure 6 DNA distribution in plugs and lanes as a function of radiation dose. NIH 3T3 cells were embedded in agarose at a concentration of $3x10^6$ per ml and irradiated on ice. Cells were then lysed and the DNA was fractionated on a CHEF gel. R-MPV's were obtained for plugs (diamonds) and lanes (squares) over a range of radiation doses. The dashed line represents the Damage Index, which is a ratio of lane R-MPV: plug R-MPV.

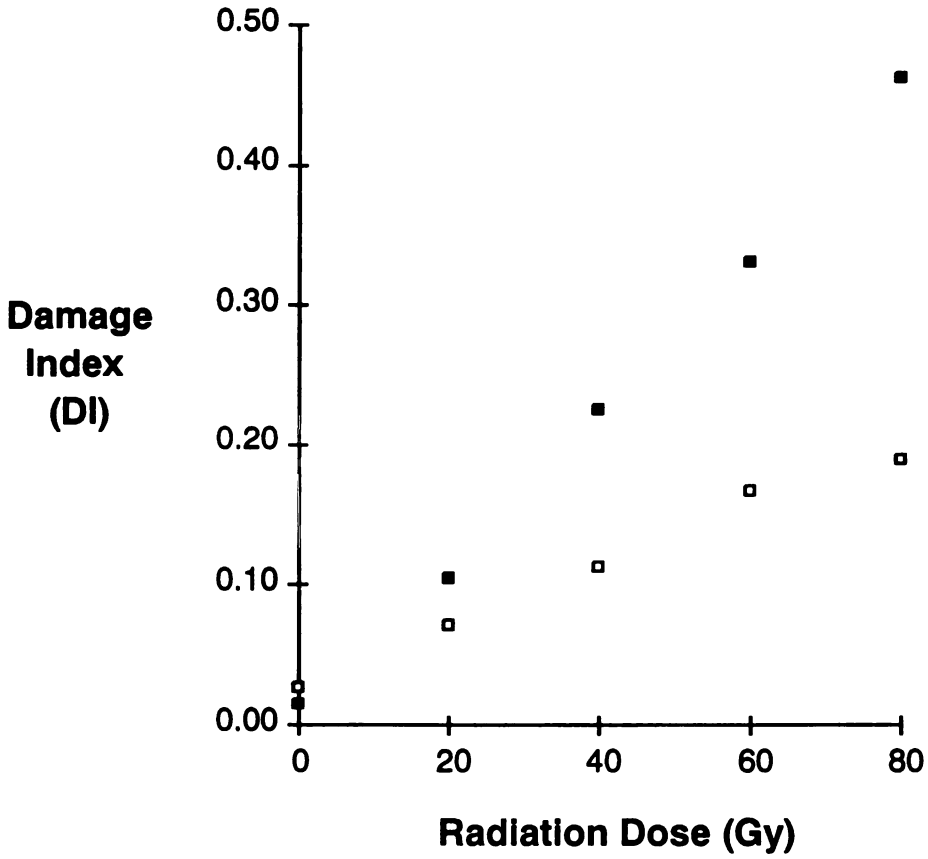


Figure 7 Damage Index comparison can be used to visualize DSB repair. NIH 3T3 cells were embedded in agarose at a concentration of $3x10^6$ per ml and irradiated on ice. Cells were then either lysed immediately (closed squares), or allowed to repair in media for one hour prior to lysis (open squares).

CHAPTER 4

NUCLEOTIDE ADDITION TO DOUBLE STRAND DNA BREAKS IN VIVO BY TERMINAL TRANSFERASE

as submitted to

EXPERIENTIA

Abstract

We have established fibroblast lines with inducible expression of the enzyme terminal deoxynucleotidyl transferase (TdT). TdT is a DNA polymerase that is unique in that it catalyzes without template direction the addition of deoxyribonucleotides at the 3'-hydroxyl ends of DNA primers (single strand or double strand) in vitro. *This activity does not affect growth rate or viability of the TdT induced cells, but can affect DNA repair. In experiments in which we irradiated TdT expressing and non-expressing fibroblasts and cultured them with chain terminators we determined that double strand DNA breaks can serve as a substrate for TdT driven polymerization* in vivo.

Introduction

Terminal deoxynucleotidyl transferase (TdT) consists of a single polypeptide chain of molecular weight 58,000 (refs. 1-4); it is found in pre-B and pre-T cells and their transformed counterparts[5-9]. It is a DNA polymerase that is unique in that it catalyzes without template direction the addition of deoxyribonucleotides at the 3'-hydroxyl ends of DNA primers[2,10]. Deoxyguanosine residues are predominant in the added nucleotide segment[11]. A genomic clone encoding human[12] TdT and cDNA clones encoding mouse[13,14], human[15], and bovine[16] TdT have been isolated.

In vivo TdT adds random nucleotides to the free 3' ends of DNA that are generated during immunoglobulin gene rearrangement, thereby creating the so-called N regions[17,18]. Mouse pre-B cells transfected with a TdT gene have been reported to contain N region insertions in a substrate containing gene segments of an immunoglobulin light chain[19]; unlike the immunoglobulin heavy chain, the immunoglobulin light chains in normal lymphocytes do not usually contain N regions[20]. The paucity of N regions in immunoglobulin light chains is presumably due to the short time period in which TdT is expressed; light chains are almost always rearranged later than heavy chains. When transfected into fibroblasts, genes expressing RAG-1, RAG-2, and TdT were shown to recombine a transfected substrate and in the process to form N regions. If *RAG-1* and *RAG-2* were transfected without *TdT*, there were no N regions[14], which proves that TdT is responsible for the insertion of N regions. It has also been shown that few N nucleotides were added to rearranged immunoglobulin and T cell receptor genes in adult mice whose TdT gene had been rendered nonfunctional; moreover, there is much more homology-directed recombination in such mice[21,22].

It is not completely clear whether or not TdT also acts in cells other than lymphocytes. It is thought that the enzyme is not expressed in cells other

than lymphocytes; although low activities cannot be excluded, no mRNA could be detected by conventional techniques. If low levels of TdT are present in other cell types, its physiological function may differ from that in lymphocytes.

In vitro, a gapped molecule is a good substrate for TdT[23]. Such gaps occur during excision repair, *e.g.*, after UV irradiation. It is not clear what the substrate for TdT is *in vivo*. Double-strand breaks, single-strand breaks resulting in gaps, or simply nicks might be substrates. Because in lymphocytes TdT is responsible for N-region addition during the rearrangement of immunoglobulin and T cell receptor genes, and because rearrangement involves double strand breaks, a simple assumption would be that double strand breaks serve as a substrate.

Our aim was to investigate whether TdT can be enzymatically active at sites other than those generated through the lymphocyte specific recombinational machinery and whether double strand breaks can be a substrate. Because of the difficulty experienced by our group and others in obtaining a stably transfected cell line constitutively expressing TdT, we utilized a tetracycline-dependent inducible expression system[24] to express TdT in mouse fibroblasts.

This control system consists of two elements. One element, a fusion protein consisting of the *tet* repressor of *E. coli* and the activating domain of virion protein 16 of herpes simplex virus, is constitutively expressed; it constitutes a tetracycline-controlled transactivator (tTA). The tTA stimulates transcription of the other element, a minimal promoter sequence derived from the human cytomegalovirus promoter IE combined with *tet* operator sequences. These promoters are silent in the presence of the tTA plus tetracycline, which prevents the tetracycline-controlled tTA from binding to *tet* operator sequences.

Materials and Methods

Plasmid constructions. The murine TdT cDNA was excised as a 1.7 kb Eco RI fragment of pTdT14 and cloned into the Eco RI of the polylinker of pUHD10-3 (ref. 24) to generate ptet-TdT. This fragment contains coding sequence only of the TdT gene and none of the CMV promoter sequences from the original pTdT plasmid. The plasmid containing the tTA gene and a neomycin resistance gene, pUHD15-1 neo, was used with no modifications. For induction of TdT expression G418 resistant clones were grown in 12 well dishes until they reached confluence (ca. $1x10^6$ cells) and then were trypsinized and plated into duplicate plates with or without tetracycline. After 16-18 hr the cells were harvested and tested for TdT expression by immunofluorescence.

RNA dot blots. RNA was prepared from $1x10^6$ cells by the RNAzol B method (TelTest, Inc.) Cells were lysed with the RNAzol B reagent, mixed with 1/10 volume chloroform, microfuged 15,000 rpm 15 min 4°C, the aqueous phase was retrieved and the RNA was precipitated with isopropanol. The RNA samples were quantitated by their A_{280} measurements. For the dot blots, the RNA's (10ml) were mixed with formamide (25ml) and formaldehyde (10ml) and 10x SSC (10ml) and incubated 10 min at 55°C. Two-fold serial dilutions of the samples were made in 10x SSC, and the RNA's were bound to a nitrocellulose membrane by aspirating the samples through a manifold apparatus. Filters were rehydrated in 2x SSC and prehybridized in 50% formamide, 5x SSC, 0.02M Tris (pH 7.6), 1x Denhardt's solution, 10% dextran sulfate, and 0.02 mg/ml sonicated salmon sperm DNA, for 4 hours at 42°C in a hybridization oven, followed by an overnight hybridization with a 32P -labeled DNA probe. The TdT probe used was the 1.7 kb Eco RI fragment of pTdT, and was labeled by the Random Primed DNA Labeling Kit of Boehringer Mannheim.

SDS gels. $2x10^6$ cells were methionine-starved for 1 hr in met-depleted RPMI medium and then incubated with 50 mCi 35S-labeled methionine for 2 additional hours. The cells were lysed in 1x NET lysis buffer containing 0.5% Triton X-100, and the lysates were pre-cleared with protein A/sepharose beads and then incubated with affinity purified polyclonal rabbit a-bovine TdT (Supertechs) at 4°C for 1 hr., followed by a 30 min. incubation with protein A/sepharose beads (Pharmacia) at 4°C on a rocking platform. The beads were washed and then boiled in SDS and β-ME and briefly spun. The supernatants were applied to a 10% polyacrylamide gel (4% stacking gel) and electrophoresed for 2 hr at 100V and then 3 hr at 200V. Gels were fixed in 10% methanol + 10% acetic acid, soaked in Amplify (Amersham), dried in a gel dryer and exposed to film.

CHEF gel running conditions and analysis. All cultures were maintained at like densities and under identical conditions with the exception of the induced culture, where the tetracycline (0.2 mg/ml) was removed 16-18 hr prior to harvesting. Also added at this time was dideoxyadenosine (ddA, Sigma), in the experiments where it was used, to a final concentration of 250 μM, along with coformycin (Calbiochem), to a final concentration of 30 μM28. At harvest, cells were trypsinized, washed in PBS, resuspended at 5 x 10^7 cells/ml in PBS at 22°C, mixed 1:1 with molten 1.5% low-gelling temperature agarose (FMC SeaPlaque), injected into 3 mm plastic tubing, and placed on ice. After solidifying, the embedded cells were blown out from the tubing, cut into 5 mm cylindrical plugs, and irradiated on ice with a Cs-137 source at a dose rate of 2 Gy/min. Plugs were then either immediately immersed in 0.5 ml lysis buffer (0.5M EDTA, 0.01M Tris pH 8.0, 2% N-Lauroylsarcosine, 0.1 mg/ml Proteinase K) and incubated O/N at 42°C, or placed in media at 37°C for 2 hr prior to lysis. The next day RNase A was added to a concentration of 0.1 mg/ml, and the samples were incubated 1 hr at 22oC. Plugs were then placed into wells of an 0.8% agarose gel, which was run in a CHEF apparatus (BioRad CHEF-DR II) with 0.5 x TBE buffer at 4°C. Gels were run for 50 hr at 60V with an initial and final pulse time of 75 min. After running, gels were stained with 3 μg/ml EtBr in 0.5x TBE for 4 hr, destained in 0.5 x TBE O/N, with two buffer changes.

UV illumination, camera, computer hardware and software. Gels were placed onto a UV transilluminator (Fotodyne Foto/Prep I, model 3-3500) situated ca. 25 cm below the lens of a vertically adjustable mounted video recorder (Sony CCD V101). A UV absorbing orange filter was placed in front of the lens. The 72 dpi S-Video output of the camera was fed into a RasterOps 364 Video Card running inside a Macintosh IIci with 8 MB RAM. The gel could be visualized on the monitor, and after adjusting size, focus, and brightness (by manually adjusting the iris on the camera), an image (TIFF) was captured with FrameGrabber 3.0 software, which accompanied the video card. Specific areas of the image, corresponding to lanes or plugs of the various samples, were assessed for their DNA content by the measurement of the median pixel value of the red channel (R-MPV). This technique is described in greater detail elsewhere.25

Results and Discussion

We cloned a murine TdT cDNA into the multi-cloning site of the expression vector which contains the core promoter linked to a repetitive

bacterial tetracycline operon element. When this construct was co-transfected with a plasmid containing a gene for a constitutively expressing tTA, some 3T3 clones were obtained, which showed tetracycline dependent inducible expression of TdT. One clone, tet-TdT, contained upon induction >80% TdT positive cells, as judged by immunofluorescence. The induction factor in TdT cells was approximately 8-10 fold, and was maintained over long periods of cell culture. The kinetics of induction/de-induction is shown in Fig. 1. It took 16-24 hr until full level mRNA expression levels were reached, and after adding back tetracycline (i.e., repressing expression) the induced mRNA was gone after 4 hr, when background levels were reached. In our clones, there was some leakiness observed in the presence of tetracycline.

By SDS PAGE we showed that the induced TdT mRNA was properly translated. A 60 kD band (Fig. 2, lane 6) is present when biosynthetically labeled material was precipitated with an anti-TdT antiserum, but not when non-immune serum was used (lane 5). The amount was similar to the one of the TdT constitutively expressed in (human) RPMI 8402 cells (lane 8). A faint 60 kD band is also seen in lane 4, where immunoprecipitated material from uninduced cells was run. We also confirmed that the TdT was active by using a solid-phase assay we have developed for the quantification of TdT enzymatic activity in crude cellular extracts[26].

We then compared the induced to the non-induced cells in regard to differences in growth rate and ability of repairing radiation-induced DNA strand breaks. No significant alteration of growth rate was observed upon induction of TdT expression. The calculated cell doubling times, based on three experiments spanning one week each, were 22.5±0.4 hr and 22.7±0.3 hr for induced and uninduced cultures, respectively. The parental NIH 3T3 had a

calculated doubling time of 18.5±0.5 hr. There was concern that perhaps in the induced culture the higher expressing cells were being overgrown by the lower expressing cells, and thus masking a difference that would not otherwise be noticed. However, this did not seem to be the case, as a long term (3 week) induced culture was compared to a 16 hr induced culture and their TdT expression was judged to be identical by both immunofluorescence and RNA expression analysis. It is possible that the base-line expression of TdT in the uninduced line is responsible for the slower growth of the transfectant, as compared to the parent line, and that higher expression of the TdT does not add to the effect; or that slower growth is due to the expression of neomycin phosphotransferase.

It is difficult to predict what kind of effect, if any, TdT would have on repair of radiation-induced DNA breaks. TdT expression might hinder DNA repair soon after the damage was incurred, but eventually lesions might be repaired, resulting in no observable effect on cell survival. When we measured clonogenic survival after exposure to low to moderate doses of radiation (0-10 Gy), and also short-term survival (3 days) after exposure to higher doses (10-100 Gy), we saw no significant difference, in either experiment, between induced and uninduced cells. Thus we quantified damaged DNA, either with or without repair, by CHEF gel analysis. Under the conditions used DNA fragments of less than 10 Mb migrate from the well and into the gel. Larger fragments are retained in the well. By comparing the amount of DNA in the lane to that in the well, it is possible to quantify the extent of DNA damage. Cells were or were not pre-loaded with ddA for 16 hr (during tetracycline induction), irradiated, and prepared for CHEF analysis, or allowed to recover for 2 hr and then prepared for analysis. DNA damage was measured over a range of radiation doses (0-80 Gy), with no (Fig. 3A) or with a 2 hr recovery time

(repair) (Fig. 3B). With or without repair, and in the absence of the chain terminator ddA, there was no significant difference between induced (filled triangles) and uninduced (filled diamonds) cells and the untransfected control (filled squares).

The substance ddA is a chain terminating nucleotide analog which can be utilized by TdT and has been shown to be selectively toxic to TdT expressing cells[27]. In the presence of ddA, the transfected cells suffered more DNA damage than the parental 3T3 cells, but did not show additional damage upon induction. We do not know the reason for the discrepancy of the findings that ddA is toxic to TdT positive lymphoid cells but not to our transfected fibroblasts. The authors of the original work suggest that TdT has an endonuclease activity, demonstrated by its ability to relax supercoiled DNA in vitro, and therefore can "create" its own substrate. If this were the case our TdT cells should display ddA toxicity, which they do not. A simple explanation would be that some component of the recombinase that is required for this TdT activity is lacking in our cells. It has been suggested that the recombinase may cut at many more sites than just the antigen receptor loci[28]. On this hypothesis the recombinase would generate the breaks by which TdT would mediate ddA toxicity. We disfavor this hypothesis, because a number of TdT positive cells lines that show ddA toxicity are not recombinationally active[29].

Whatever the reason for the differences discussed above, we did observe a difference in 2 hr repair in the presence of ddA, between the induced and uninduced cells. The induced cells (Fig. 3B, open triangles) retained a large proportion of double strands unrepaired as compared to uninduced (open diamonds) or NIH 3T3 cells (open squares). We conclude from this observation that TdT can add nucleotides to radiation induced double strand breaks. This finding has implications for the nature of double strand breaks as well as for the

chronology of the repair process. It indicates that not all broken DNA molecules with free 3' OH groups are immediately blocked by an end-binding protein (such as Ku 80, ref. 30), or immediately covalently modified into a "hairpin"[28] or other structure. These events may well occur, but under conditions of massive stress (i.e. irradiation) large numbers of "free" ends are available as substrate to TdT. It should be noted that purified TdT can utilize as substrate broken DNA ends of apoptotic cells which have been fixed[31], also suggesting that the ends are not covalently modified.

Acknowledgment. This work was supported by NIH grant R01 GM37599

References

1. Bollum, F.J., in: P.D. Boyer (ed.), The Enzymes, Vol. 10., p. 145, Academic Press, Inc., New York (1974).

2. Bollum, F.J. and Brown, M., Nature 278 (1979) 191.

3. Nakamura, H., Tanabe, K., Yoshida, S., and Morita, T., J. Biol. Chem. 256 (1981) 8745.

4. Silverstone, A., Sun, L., Witte, O.N., and Baltimore, D., J. Biol. Chem. 255 (1980) 791.

5. Chang, L.M.S., Biochem. Biophys. Res. Commun. 44 (1971) 124.

6. Coleman, M.S., Hutton, J.J., de Simone, P., and Bollum, F.J., Proc. Natl. Acad. Sci. U.S.A. 71 (1974) 4404.

7. Kung, P.C., Silverstone, A.E., McCaffrey, R.P., and Baltimore, D., J. Exp. Med. 141 (1975) 855.

8. McCaffrey, R., Smoler, D.F., and Baltimore, D., Proc. Natl. Acad. Sci. U.S.A. 70 (1970) 521.

9. Silverstone, A., Cantor, H., Goldstein, G., and Baltimore, D., J. Exp. Med. 144 (1976) 543.

10. Kato, K.I., Goncalves, J.M., Houts, G.E., and Bollum, F.J., J. Biol. Chem. 242 (1967) 2780.

11. Basu, M., Hegde, M.V., and Modak, M.J., Biochem. Biophys. Res. Comm. 111 (1983) 1105.

12. Riley, L.K., Morrow, J.K., Danton, M.J., Coleman, M.S., Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 2489.

13. Landau, N.R., St. John, T.P., Weissman, I.L., Wolf, S.C., Silverstone, A.E., and Baltimore, D., Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 5836.

14. Kallenbach, S., Doyen, N., Fanton, d.M., and Rougeon, F., Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 2799.

15. Peterson, R.C., Cheung, L.C., Mattaliano, RJ., Chang, L. M.S., and Bollum, F.J.. Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 4363.

16. Koiwai, O., Yokota, T., Kageyama, T., Hirose, T., Yoshida, S., and Arai, K., Nucleic Acid Res. 14 (1986) 5777.

17. Kurosawa, Y., and Tonegawa, S., J. Exp. Med. 155 (1982) 201.

18. Barth, R.K., Kim, B.S., Lan, N.C., Hunkapiller, T., Sobiek, N., Winoto, A., Gershenfeld, H., Okada, C., Hansburg, D., Weisman, I. L., and Hood, L., Nature 316 (1985) 517.

19. Landau, N. R., Schatz, D. G., Rosa, M., and Baltimore, D., Mol. Cell. Biol. 7 (1987) 3237.

20. Sakano, H., Huppi, K., Heinrich, G., and Tonegawa, S., Nature 280 (1979) 288.

21. Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., and Mathis, D., Science 261 (1993) 1175.

22. Komori, T., Okada, A., Stewart, V., and Alt, F.W., Science 261 (1993) 1171.

23. Kunkel, T. A., Gopinathan, K.P., Dube, D.K., Snow, E.T., and Loeb, L.A., Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 1867.

24. Gossen, M., and Bujard, H., Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 5547.

25. Harriman and Wabl, submitted.

26. Harriman and Wabl, J. Immunol. Methods, in press.

27. Spigelman, Z., Duff, R., Beardsley, G.P., Broder, S., Cooney, D., Landau, N.R., Mitsuya, H., Ullman, B., and McCaffrey, R., Blood 71 (1988) 1601.

28. Lieber, M.R., Cell 70 (1992) 873.

29. Rathbun, G., Oltz, E.M., and Alt, F.W., Int. Immunol. 5 (1993) 997.

30. Taccioli, G.E., Gottlieb, T.M., Blunt, T., Priestley, A., Demengeot, J., Mizuta, R., Lehmann, A.R., Alt, F.W., Jackson, S.P., and Jeggo, P.A., Science 265 (1994) 1442.

31. Gorczyca, W., Gong, J., and Darzynkiewicz, Z., Cancer Res. 53 (1993) 1945.

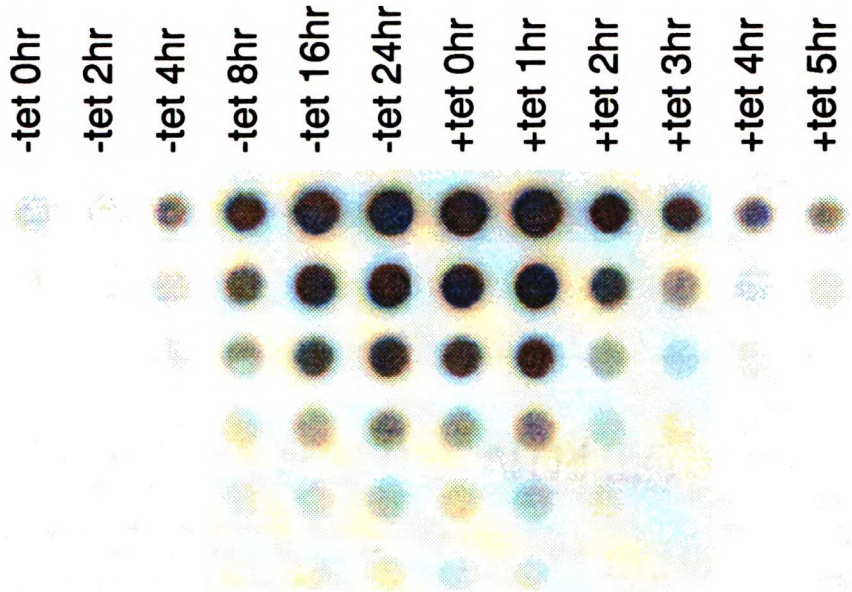


Figure 1 Kinetics of induction and de-induction of TdT expression in tet-TdT cells shown by RNA dot blot analysis. RNA was isolated from 10^6 cells at the indicated times following the removal of tetracycline from the medium. After 24 hr tetracycline was added back into the cultures and RNA was prepared.

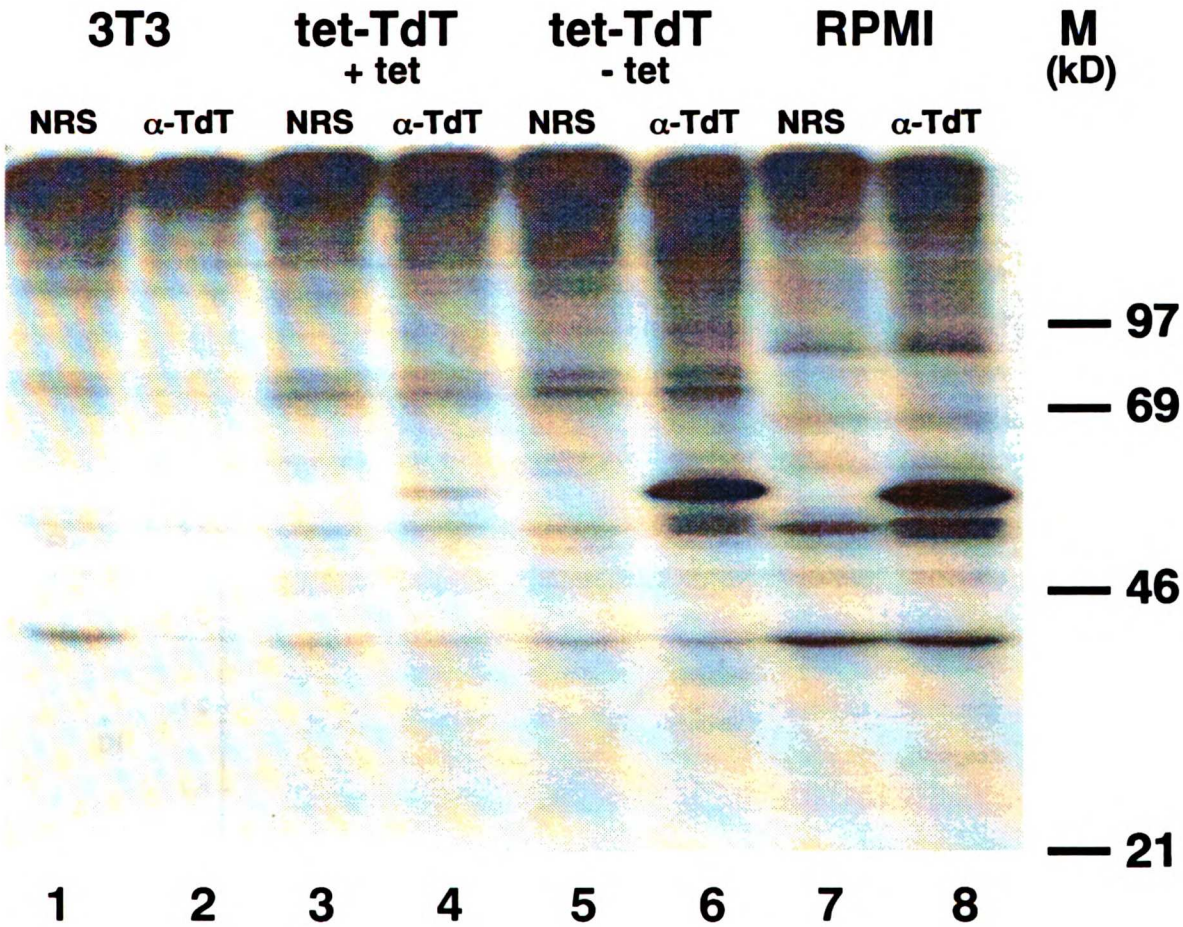


Figure 2 SDS PAGE of ^{35}S methionine labeled immunoprecipitates of induced (- tet) and uninduced (+ tet) cells. 3T3, NIH 3T3 cells; tet-TdT, NIH 3T3 cells transfected with inducible TdT and transactivator gene; RPMI, RPMI 8402 cells; NRS, non-immune rabbit serum; α-TdT, rabbit anti-TdT antiserum. The murine TdT expressed in the transfectant, at 60 kD, is slightly larger than the 59 kD human TdT of the RPMI 8402 cells.

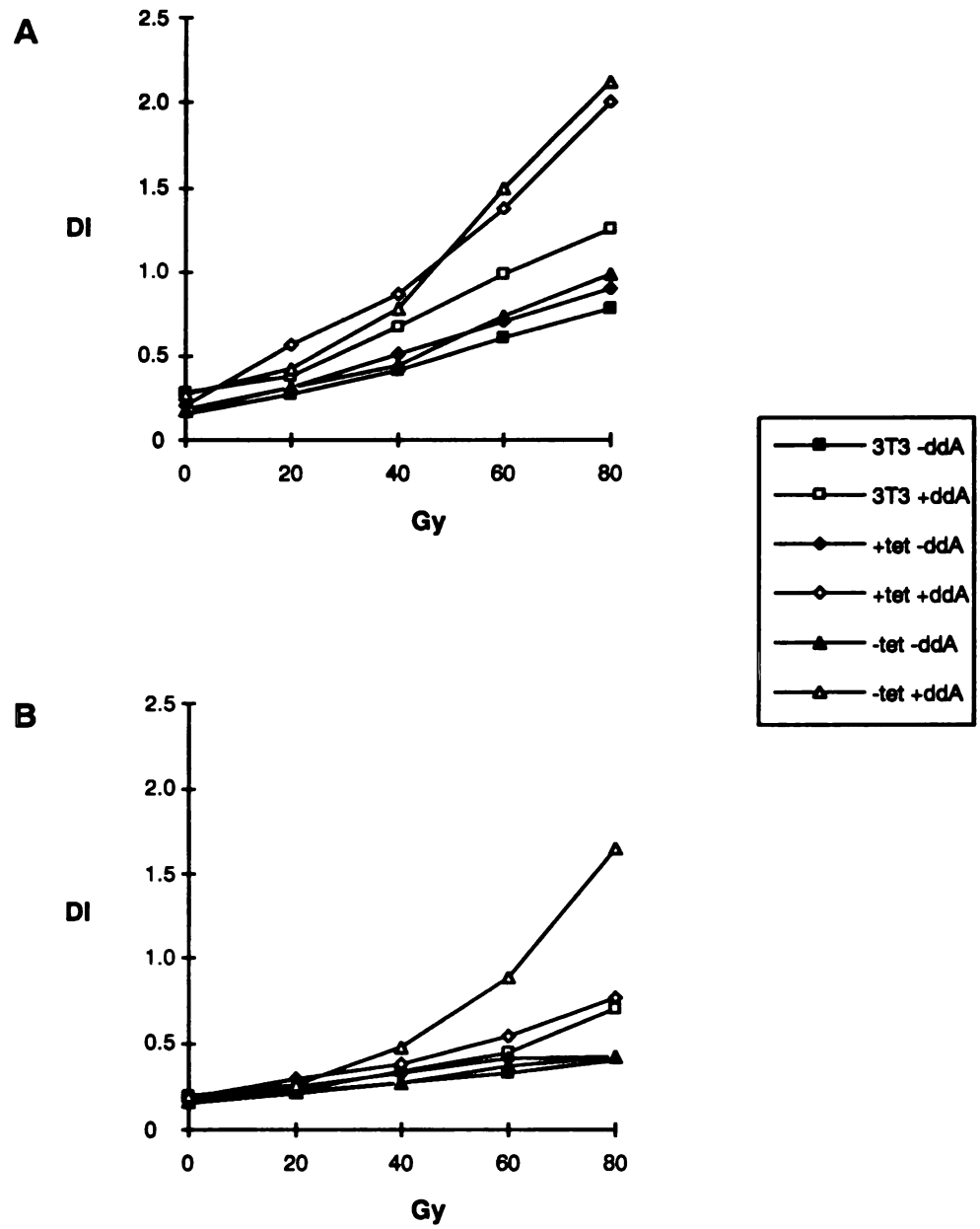


Figure 3 Extent of DNA damage, indicated by the Damge Index (DI), for γ-irradiated cells, in the presence (open symbols) or absence (filled symbols) of ddA. + tet, uninduced cells; - tet, induced cells. (A) DNA damage was assessed immediately after irradiation; (B) after a 2 hr recovery period.

APPENDIX

CAN WE OUTLIVE METHUSELAH?

as appearing in

CYBERNETICS AND SYSTEMS
1994

Introduction

Although almost no one wants to be old, most of us certainly want to get old. These desires are not as irrational or self-contradictory as might appear at first glance. For there are, in fact, two quite separate issues involved here. There is the normal process of slow but relentless corporal deterioration that we call aging. And then there is the process we call death, ceasing to exist, at least corporally. To want to slow down the process of aging (not "to be old") and to want to stave off death as long as possible ("to get old") are not in fact contradictory. Although these two issues are probably linked to each other at some deep level, it is useful to muse about them separately. What causes aging, and can we stop the process or at least slow it down? What determines our lifespan, and can we extend it for much longer than a hundred years? For the purposes of this essay, we will deal with human beings strictly from a biological viewpoint, *i.e.*, as a particular mammalian species. There are indeed qualities that set human beings apart from all other living creatures, but we do not feel these are relevant to the present discussion.

To most of us, it seems natural that, barring accidents or disease, we die of "old age". We view death as the natural end of the aging process, and thus if we could explain aging, it seems we would not have to explain death. For many animals, including human beings, this view would seem to have considerable merit, at least superficially, for it is based on empirical observation. But there are many other organisms that die after a certain time without the progressive deterioration that characterizes aging. Indeed, in the language of every day life, we talk about the "clock" of life running down. Some animals breed only once and then die. For instance, the male Pacific salmon senesces catastrophically after his single attempt at breeding [1-3]. Or after having spent a year in the water as a larva, an adult mayfly lives less than a week [1-3]. Thus,

for these animals, the clock metaphor seems more appropriate than it does for human beings.

In Figure 1A, there is a clock, a timer that we use in the laboratory. In Figure 1B, the clock has been wound up, set to ring in three minutes or so, and is running down. In Figure 1C, the time is up, and the clock is ringing. In fact, the clock looks the same in all three panels. In other words, the clock did not age while it was running down. All the parts of the run-down clock in Figure 1C are still in mint condition, and the clock may be wound up again and function just as well as it did the first time. In contrast to this, when a person dies from a heart failure, for instance, not only is that organ defective, but many other parts of his or her body are in bad shape. Thus, the clock metaphor for lifespan does not include aging.

But there are regularities in the process of aging. For example, with increasing age the hair gets gray and starts thinning. And in men the thinning is obviously not random. Male pattern baldness may start with a receding hairline, or it may start as a bald patch on the crown of the head like a monk's tonsure — the pattern is inherited, encoded in the genes. And the gray hairs tend to show up first at the temples. Although the rate of graying varies somewhat, it eventually happens to almost every man. Except for Ronald Reagan, all men have gray or white hair by the age of seventy, if they have any hair at all.

The Age of the Computer has replaced the Age of the Machine, and Informatics, not Thermodynamics, is the Science of our time. To many people, the regularity and steadiness with which obvious signs of aging occur seems to be proof that aging is a natural extension of a "program" that a fertilized egg must execute to become a fetus, then a baby, a child, a teenager, and finally an adult. There is little doubt about the utility of the program metaphor for the development of an adult from an egg. The source code for the program is in the

DNA, and we even know how to read it. As the compiler and link editor are at least partly in the cytoplasm of the egg, we know somewhat less about them and, as a consequence, about the object code. But although aging temporally follows the clearly programmed events leading up to the adult, there is no logical justification for assuming that aging must have the same underlying cause. Yes, there is surely no better time to age than adulthood! But it must be kept in mind that aging might begin as soon as the fertilized egg divides, but that the aging-induced changes might be too incremental to be noticed during the swift, dramatic changes that take place during early development.

There are other metaphors for aging besides clocks and computer programs. Rather than describing our limited life span by the metaphor of the "clock that runs down" the "old car that quits on the highway" may describe the situation much better. We buy a new car and expect it to last for five or ten years, depending on the make and amount of use. If we take good care of it, replace broken parts, and service it at regular intervals, the car may well last substantially longer. But no one expects a car to last for, say, fifty years, assuming that it is used regularly. Like the parts of our body, some parts of the cars are likely to "age" faster others. In the old Volkswagen beetle, the clutch was the first subsystem to go. And, while the body might last for more than 200 thousand kilometers, the engine would usually quit around 100 thousand kilometers. Surely, the VW engineers did not program the parts to "age" in that predictable progression. In fact, the concept of programmed differential aging makes little sense, for automobiles as well as for people. Oliver Wendell Holmes wrote about a "wonderful one-hoss shay" [4], the parts of which were so designed that they all wore out at exactly the same time; the carriage ran perfectly for a long time before collapsing one day into a heap of unusable parts. Now that is a sensible aging "program"! If we were so designed, hospitals

would be reduced to places for treating the victims of accidents or infectious disease, and there would be no explosion in health-care costs; moreover, pension plans and old age asylums would be superfluous.

Aging, or more precisely, the rate of aging, might be determined solely by the level of exposure to damaging agents. Cars in Southern California, where there is little rain and also little need to spread salt on the streets to prevent icing, last longer than do cars in New York City. We are all exposed to many substances that damage DNA, which is the source code for building our body, as well as to substances that damage particular tissues and organs. While there is definitive proof that certain substances such as tobacco smoke, asbestos, and nitrosamines can severely reduce lifespan, the arguments that exposure to damaging external agents is generally responsible for the rate of aging are rather banal. For instance, most birds considerably outlive mammals of comparable size, and this has been attributed to birds being exposed to fewer potential hazards because they spend a lot of time in the air [3]. And, indeed, the flightless emu and the ostrich are notably short-lived birds [1, 2]. Moreover, bats have much longer maximum lifespans than do similar-sized rodents [1, 2]. On the ground, turtles with their thick protective shells outlive other reptiles [2].

If we accept that aging is due to damage, why not just replace the damaged parts and remain young and go on forever? Indeed, although it might not be economically feasible, there is no reason why we could not replace each part of a car as it became non-functional, so that the car would run forever. But there are some rather basic differences between a car and a human being, and we can no longer avoid discussing them. We all consist of millions of millions of cells. If you take a piece of skin and place it under the microscope, you, like Antonie van Leeuwenhoek, can see the cells, which are more or less self-

contained units of life. That is, if we disrupt the skin into the cells of which it is made, these skin cells will continue to live if they are fed with the proper nutrients. A cell in culture, as in the body, converts food into energy and into more of the components of which it is made; it may get larger and eventually divide into two cells. A car, of course, has no such unit. It just has a number of parts, some of which are unique, and others which are present in several copies.

Although many of our cells die every day, the death of these cells does not imply death of the individual. On the contrary, cell death is even required for the continuation of life of the individual. One of the functions of the immune system is to kill cells that are infected by viruses or bacteria. And many cells, *e.g.*, the cells in the outer layer of the skin and in the inner layer of the gut, are constantly dying off at a rapid rate, but are being replaced at an equally rapid rate by new ones. Dandruff, which plagues many of us, consists of the remnants of dead scalp cells. So, although we have argued against a program for aging and death of the organism — of you and me — there clearly exists a program for the aging and death of some cells [5]. When this normal, programmed cell death is not sustained, the result may well be a tumor that eventually kills the organism. While the organism dies, the tumor cells are often effectively immortal. Every biomedical scientist has heard of Helen Lane (or at least the acronym HeLa). HeLa cells, which are derived from the cancer that killed Helen Lane more than forty years ago, are still cultured today in laboratories all over the world.

While a surgeon may replace a whole organ, just as an auto mechanic may replace a subsystem of our car, the surgeon's intervention, unlike that of the mechanic, is fraught with dangers and complications. Some animals, like salamanders, can regenerate lost limbs, but our bodies can replace only cells.

Indeed, barring accidents like tumor formation, our cells are replaced systematically as needed, and in that respect we are are more like gothic cathedrals than automobiles. Gothic cathedrals are built of sandstone blocks. These stones must be continually replaced over the centuries, with the result that almost no stone of a catheral today was part of the original structure. Yet the cathedral continues to look essentially the same. Why, then, don't we continue to look the same, even though, with the exception of the nervous system, few of our cells existed when we turned twenty-one? The answer to this question, unlike most of the questions we are discussing here, is quite simple. It is obvious to the stonemason how to replace a stone in a cathedral; and if for some reason, like a major catastrophe, it is not so obvious, he can always consult the blueprint for the cathedral in the parsonage. But it is never obvious how to replace a cell, and the blueprint for it is kept inside the cells themselves. New cells are made by cell division. At the time of cell division, its blueprint must be copied, and the copying process is by no means error-free — many of the errors are innocuous, but occasionally they are very serious. When errors — we biologists call them mutations — are introduced, the new cell might even be "better", in some sense, than the original; for example, it might grow faster. But what may seem "better" for the individual cell, will almost certainly be detrimental to the organism as a whole. A cell that grows faster and fails to die when it should is *ipso facto* a tumor cell.

But tumors are only a part of the problem. The parts of a multicellular organism, just as the members of large organizations, like states or corporations, all have to do their specialized duties or the organism as a whole will cease to function well and eventually cease to function at all. *Nota bene*, organizations age too. During the second World War, American physicists closed ranks to develop the atomic bomb. Their fear of the Nazis overcame

their natural dislike of things military, and they submerged their individuality to the common cause. No one who has had any contact with that generation of physicists fails to be impressed with the speed and efficiency of the Manhattan Project. But as personnel changed after the war ("mutation" is Latin for "change") the aging successor organizations, the U.S. Atomic Energy Commission and the U.S. Department of Energy became less impressive. And the necessity for cell replacement is also only a part of the problem. Mutations can still creep into the DNA of non-dividing cells. It is as if some high school students, as a joke, broke into the parsonage and altered the blueprints there. Our nerve cells, for instance, never divide after we are one year old. They sit there like civil servants who draw their salaries whether they work or not. So, if a mutation renders a nerve cell non-functional, it can never be replaced. Small wonder that senility is one of the manifestations of aging. Higher organisms have developed a large armamentarium of proofreading and error-correction mechanisms to reduce the rate of mutation, *i.e.*, the rate of errors in the the copies of the DNA blueprint. As Manfred Eigen and his colleagues [6] have shown, the maximum size of the blueprint is limited by the rate of errors in copying it, at least for simple creatures like viruses. But reducing the rate of error to zero would mean that entropy would not increase, and even biological evolution has not been clever enough to repeal the Second Law of Thermodynamics. The accumulation of errors over time brings incremental changes in the architecture of an organism, and, as explained above, these changes are unlikely to be to its advantage; the result is aging. We have made several references to DNA, sometimes using the metaphor of DNA as the source code for the program of life, sometimes calling it the blueprint for a cell. DNA is found in every cell of an individual. Thus, there is not a single blueprint stored away somewhere in a particular organ, but each

cell has its own blue-print that is almost identical to that of the other cells of the same individual. DNA as a substance was discovered more than a century ago in a laboratory in the idyllic castle of Tübingen in southern Germany. That DNA is the substance of which genes are made was discovered more than half a century ago in the rather less idyllic venue of the Rockefeller Institute in New York City. Since DNA is a chemical, one can isolate it and obtain it in pure form. Figure 2 is an attempt to take away from you some of the mystery that surrounds DNA. Figure 2A shows a solution of DNA in a dish. While we see the puddle of liquid, we cannot see the DNA any more than we can see the sugar in syrup. But when alcohol is added to the puddle the DNA will separate from the water and can be seen as white precipitate resembling snow flakes in Figure 2B. The fibrous nature of DNA is evident in Figure 2C where it is held between the loops of two bacteriological loops.

With special methods, one can read the source code. The DNA of a single cell consists of two linear sequences — one inherited from the mother and one inherited from the father of the organism — which together have 6 thousand million characters, each of which must be copied before a cell divides. On the average, one character will be copied falsely during each cell division. This is pretty good, but not good enough to keep us looking the same for centuries. While the microchip memory makers have caused us to become blasé about how much information can be stored in a small space, it is perhaps interesting to consider just how densely information is stored in a cell. If we could straighten out the DNA filament of a single cell, it would be one meter long, but only 20 Å in diameter, for a total volume of about 3 cubic microns. Since four letters are used in the DNA code, each character represents 2 bits, and the cell crams 12 Gigabits into 3 cubic microns. Those guys in the silicon foundries still have a long way to go!

We, as the species *Homo sapiens*, have been around for more than a hundred-thousand years. During this time, many errors in the blue prints must have occurred. How then, can we ever be young? Why is a baby not just as old as its parents? The answer to these questions involves (i) the singleness of the fertilized egg, (ii) separation of soma and germline, and (iii) embryogenesis and selection. As individuals, we all start from a fertilized egg, which is a single cell that resulted from a fusion of the mother's egg and the father's sperm. The fertilized egg represents time zero of a person's life, but its outside appearance is neither young nor old in the sense that an adult, multicellular organism is. The singleness is important. Here the metaphor of the multicellular organism as large organization will help us again. An engineer quits his job at a large corporation and starts his own company. In a relatively short time, the new company either goes belly-up or is successful; in the latter case it can then, and only then, afford the luxury of aging. If the the blueprint in the egg contain serious errors, it will not succeed in directing the development of an embryo to an adult; if it is relatively error-free, it will do so, and that adult will begin to age. Early in embryonic development, the cells that will form eggs or sperm (germ cells) are segregated from the cells that will form the rest of the body (soma). Although the strategies used by eggs and sperm are very different, both of these germ cells contain fewer mutations , on the average, than do most other sorts of cell. The progenitors of eggs, undergo a strictly limited number of cell divisions, which limits the number of times the blueprint must be copied, which in turn limits the number of errors introduced. A new-born baby girl has all the oocytes, the direct progenitors of eggs, she will ever have [7]. The progenitors of sperm undergo a large number of divisions, but the rate of error formation is thought to be less than it is in other cells. Because the testicles are in a sack outside of the body, their temperature is lower, and mutation rates are

highly dependent on temperature.

The correctness of the program is tested over and over again during the development of the embryo. Many, perhaps most, embryos do not make it and abort at an early stage, so early that it most often goes unnoticed. That is, there is a selection against deleterious mutations. And if serious errors become apparent during childhood, the afflicted person may never reach the sexual maturity and social acceptance needed to pass his blueprint onto the next generation of children. But could a baby's skin look like that of an an eighty-year old? Yes, in rare cases it does, but such a child dies at a tender age.
But selection is a far more subtle and important force than the above remarks would suggest. Many, perhaps most, deleterious mutations that arise in the cells of an adult are in fact also eliminated as a result of selection.
Unfortunately, we have no equivalent in our bodies to the stonemason, who takes care of the gothic cathedral, no monitor of cellular health that can reliably identify all ill-performing cells. Thus, as adults we cannot put all of our cells through a selection process to allow us to keep only the dependable ones.
Those cells that are responsible for aging are the ones that cannot be counter-selected, a class that includes the non-dividing cells. The cessation of growth in adult human beings aggrevates the problem, for then there are fewer cells that divide, hence fewer that can be counterselected when they are defective. It is probably not a coincidence that the animals that live the longest are the ones that continue to grow as adults; the giant turtles of the Galapagos Islands continue to grow and get extremely large in their old age. It is indeed difficult for nature to design a selection system for a multicellular organism. This is also the problem of large organizations — how can those who remain dedicated to the purposes of the organization be selected rather than those who are dedicated to advancing their own careers? Moreover, the natural selection of

organisms — not cells — is the driving force of organic evolution, but a discussion of that phenomenon would lead us to stray too far from our main theme.

While it is impossible to copy the blueprints of all our many cells without making errors, and while it is thus impossible for anyone to stay in this world forever, just how long could we live? A thousand years instead of a hundred years or less? As we have already mentioned, the rate of aging is highly variable among animal species. For example, a mouse lives about two years, a dog twenty years, a human being three score and ten years; giant turtles purportedly live many hundreds of years. Thus, our question becomes: what determines the average lifespan of individuals of a given species? Errors in copying DNA in cell replacement is probably not the only proximate cause for aging. Various kinds of damage and the extent to which different species are successful in avoiding or counteracting them will determine their lifespan. We have become quite smart in avoiding and fighting environmental damage over the last centuries, and thus human life-expectancy has gone up considerably during this time. Increasing sanitation in cities and hospitals was probably the key element in the century or so before this one. Good eating habits, immunization, avoidance of environmental insults such as excessive sunlight and asbestos, as well as dramatic improvements in medicine and surgery are among the factors that allow us to live even longer.

But could we avoid damage more and more and thus get older and older? At the moment it seems that no matter how well we do, we are not going to get much older than maybe 130 years old, which is rare even in the areas of the Caucasus Mountains where longevity seems to be the greatest. Thus, the maximum life span seems to be determined by an intrinsic characteristic of our DNA. From what we have said before about cell replacement and copying

errors, the fidelity with which the blue-prints of the cells are copied must be a factor, perhaps even the most important one. This fidelity is achieved by a number of proofreading systems inside the cells. The proofreading systems are inherited by the very same blueprint that is present in all cells; they consist of several biocatalysts, called repair enzymes. Some other species have proof-reading systems superior to our own. For instance, a bacterium called *Deinococcus radiodurans* is able to survive the environmental insult of one million rads of gamma radiation, which allows it to live in the cooling system of a nuclear reactor. Maybe Galapagos turtles also have better repair enzymes; to the best of our knowlege they have never been studied. If we could transplant a cocktail of better proofreading enzymes from other species, we might be able to live much longer.

Is it really conceivable that we could transplant biocatalysts in order to slow down aging? In principle, the procedures to introduce foreign blueprints for any biocatalyst into our organism are already in place. These so-called recombinant DNA techniques are astonishingly simple. As you saw in Figure 2, it is easy to purify DNA. Special biocatalysts, called restriction endo-nucleases, are able to cut DNA at specific sites, and indeed, do so every day in our laboratory and in countless other laboratories around the world. The cut and otherwise modified DNA can be reintroduced into cells. Figure 3A shows part of a tape that represents the DNA isolated in Figure 2. This tape contains a nearly endless message, parts of which we want to replace. The part to be replaced is defined by the sequence of its letters, here set apart from the rest by small gaps. In Figure 3B, a piece similar to the one that has been cut out, is ready to be spliced in place of the original one. If it were at the cutting table of a film studio, the new piece would be fixed to the two ends, where the original piece was cut out, with Scotch tape. In the lab we use enzymes instead of

Scotch tape, but the recombined piece of tape in Figure 3 is analogous to recombinant DNA. The final application of recombinant DNA techniques to a person might turn out to be as simple as vaccination. That is not to say that we could substantially prolong life today or tomorrow. But there is no *a priori* reason why this could not be done at some time in the not so distant future. Before concluding this essay, we should warn you that biologists are a contentious lot. While we hope that our colleagues would not find any errors of fact here, not all of them would agree with our emphasis on the role of mutation in aging and dying. Those who have other ideas would certainly point out to you that all of the authors of this essay work on, and hence have a vested interest in, mutation. Nor do only biologists consider themselves experts in this field. George Bernard Shaw was convinced that human lifespan could only be increased by acts of human will. In the introduction to his play "Back to Methuselah", Shaw suggested that 300 years was the goal for which we should strive [8]. Well, Methuselah made it to 969 years. If our molecular biological speculations are correct, we may well be able to outlive him. Let us discuss this issue again in a thousand years or so!

Acknowledgment. Work on hypermutation is supported by N.I.H. grant R01 GM37599 to M.W. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche, Basel, Switzerland.

References

1. Comfort, A. The Biology of Senescence 3rd edn (Churchill Livingstone, Edinburgh and London, 1979).

2. Williams, G.C. Evolution 11, 398-411 (1957).

3. Partridge, L., and Barton, N.H. Nature 362, 305-311 (1993).

4. Holmes, O.W. The One-Hoss Shay, The Chambered Nautilus, and other poems, gay and grave (Houghton Miflin, Boston and New York, 1900).

5. Krammer, P.H., and Debatin, K.-M. Current Biol. 2, 383-385 (1992).

6. Eigen, M. Scientific American 269, 42-49 (1993).

7. Perotta, C.A. Am. J. Anatomy 111, 195 (1962).

8. Shaw, G.B. Back to Methuselah (Constable, London, 1921).

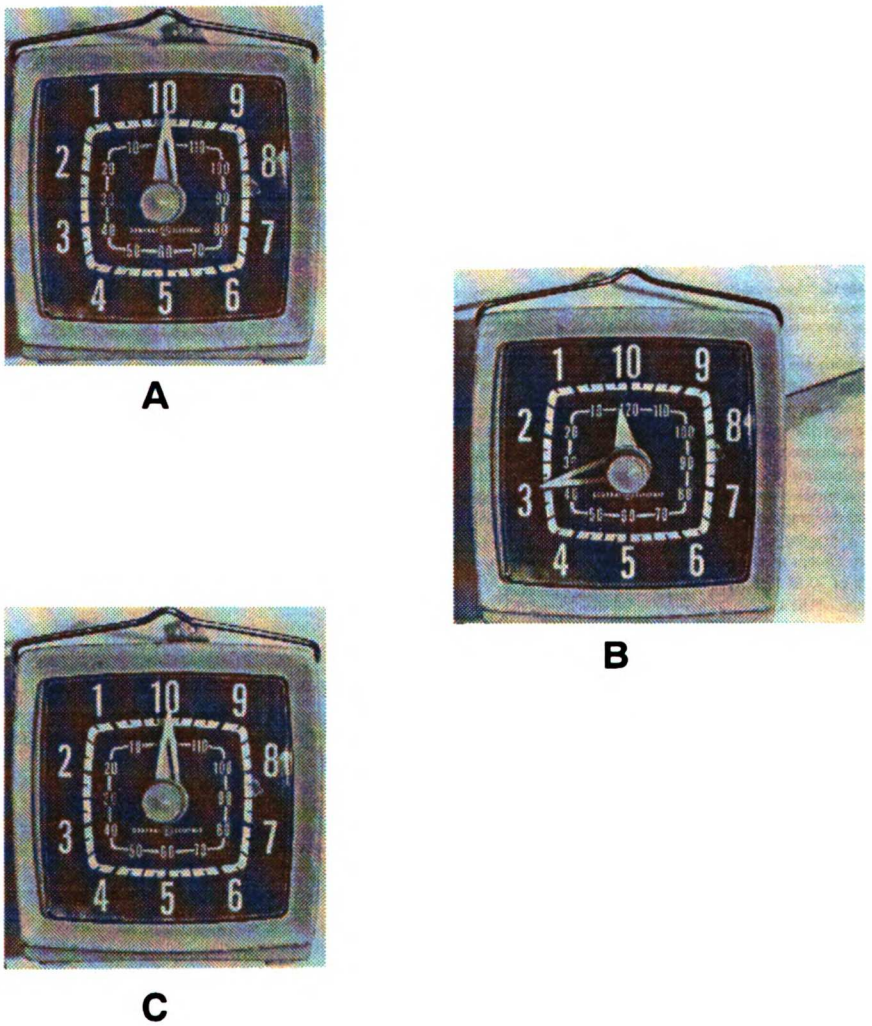


Figure 1 Why the clock metaphor is inappropriate for the aging of clocks. A laboratory timer is set and started (A), runs its time course (B), then rings and stands still (C). Note the close resemblance of (A) and (C).

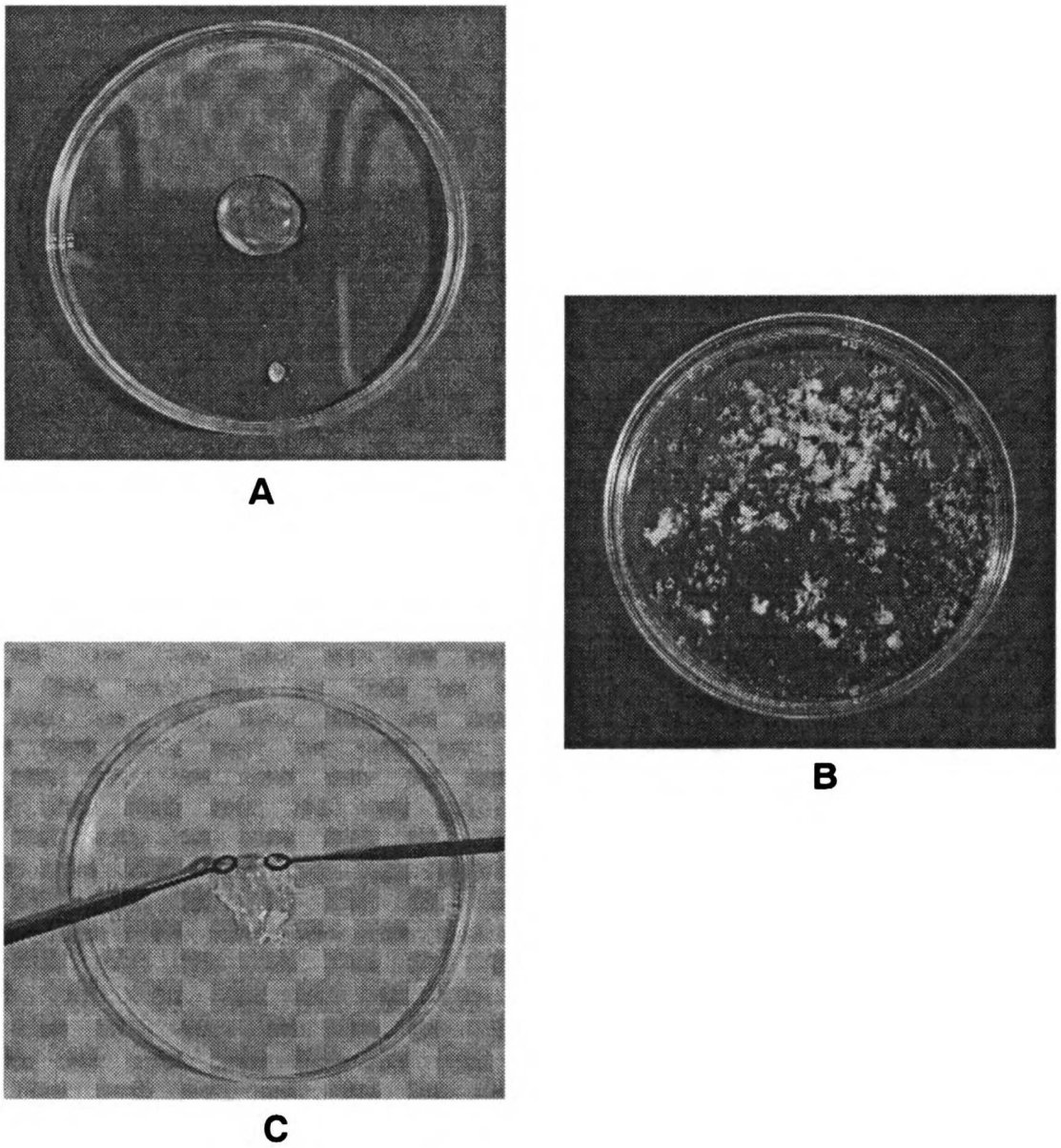


Figure 2 Real, existing, unmetaphorical DNA. To an aqueous solution of DNA (A), alcohol is added, so that the DNA is precipitated (B) and can be pulled into fibers (C).

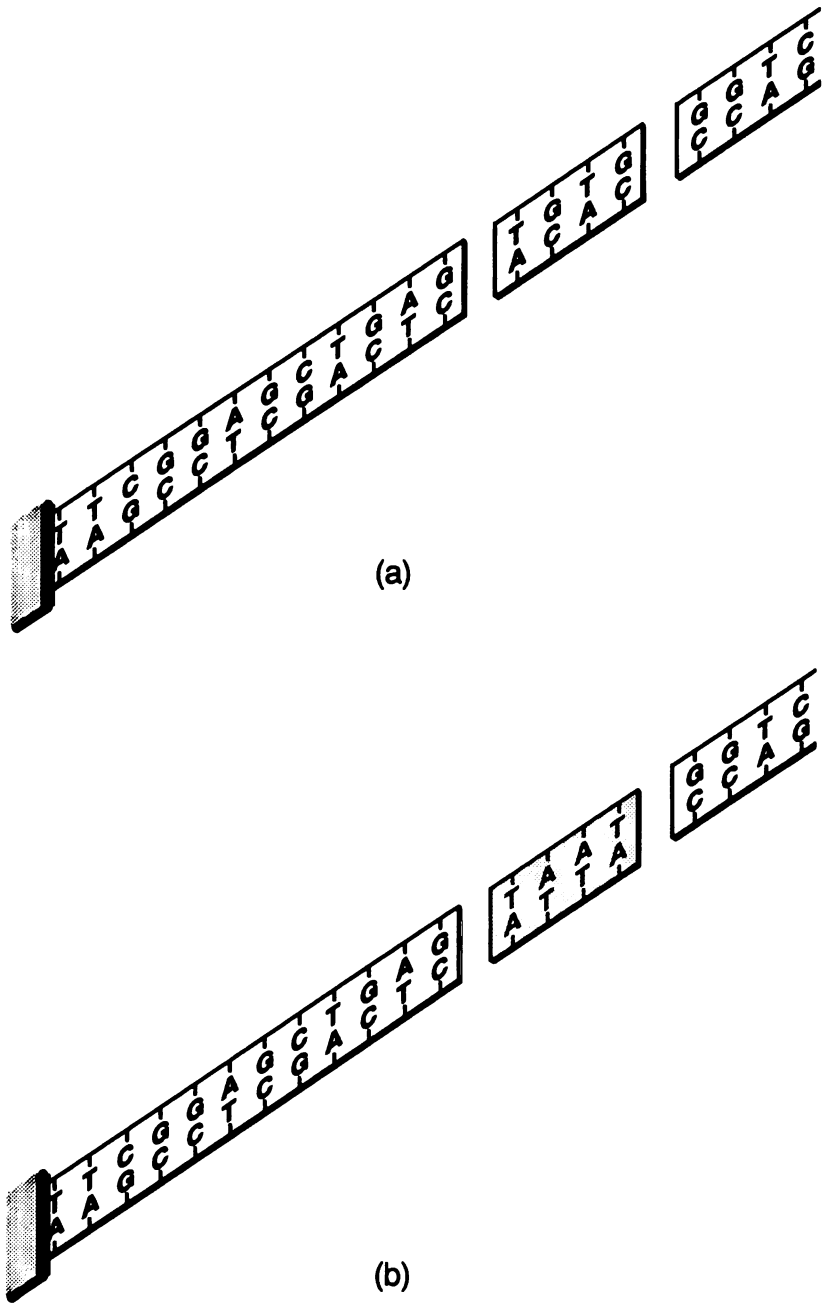


Figure 3 DNA as a tape. A short piece of DNA can be cut out (a) and a new piece spliced in (b).