DOE/ER/60521--2

DE89 010018

BIOLOGICAL EFFECTS OF DNA REPAIR, INCLUDING MUTAGENESIS

Progress Report for the Period

04/01/88 to 03/31/89

Franklin Hutchinson

Department of Molecular Biophysics and Biochemistry

Yale University

New Haven, Connecticut 06511

April 1989

Prepared for

OFFICE OF HEALTH AND ENVIRONMENTAL RESEARCH

THE U.S. DEPARTMENT OF ENERGY

GRANT DE-FG02-87ER60521

REPORT NO. DOE/ER/60521-2

This report covers the period from April 1, 1988, to March 31, 1989.

The major purpose of the research on this project is to study at the molecular level the gross rearrangements that occur in the DNA of mammalian cells, particularly those rearrangements induced by the action of genotoxic agents such as ionizing radiation. The cells being studied are Chinese Hamster Ovary (CHO)-K1 cells in which a piece of plasmid DNA has been stably incorporated into a chromosome. The piece of plasmid DNA contains both the *Escherichia coli gpt* gene and a *neo* gene, separated by about two kilobases. Mutations involving the *gpt* gene are identified by resistance of the cells to 6-thioguanine. Deletions can be detected quickly by loss of the *neo* gene (loss of resistance to G418). More information about the deletions, and information on other types of rearrangements, can be otained by Southern blots, using plasmid DNA as a hybridization probe.

The nature of the rearrangements can then be investigated by restriction analysis and other methods. For example, information on deletions can be obtained by isolating the restriction fragments containing the end points and sequencing, to determine if, for example, there is sequence homology between the end points.

Extensive research with several cell lines in which very similar plasmid DNAs were incorporated in various places in the CHO genome has shown that the inserted pieces are usually very difficult to clone out. The problem is, essentially, a high frequency of rearrangement of the target DNA during cloning. This same problem has been found by others doing similar experiments, notably by Dr. Kenneth Tindall at the National Institutes of Environmental Health Sciences, North Carolina. It is supposed that the difficulty may arise from a tendency for exogeneous DNA to be selectively incorporated into highly repetitive DNA in the mammalian cell genome. Such repetitive DNAs are considered to be hard to clone, at least in part because of rearrangements. The difficulty of cloning DNA from plant cells, for example, is thought to be related to the large fraction of repetitive DNA in plants.

In last year's Progress Report, we suggested that we might use a lambda phage adapted for use in *E. coli recA* cells for cloning the DNA; such phage have been useful in cloning DNA from

repetitive sequences in plant cells, for example (see Loenen and Blattner, 1983). Although some work along these lines was done, we have adopted another procedure for the following reason. If the target DNA is located in highly repetitive DNA that is readily subject to rearrangement, the deletions, inversions, etc. measured in our experiments might well be types not representative of the gross rearrangements involving normally transcribed genes. Thus, while studies of rearrangements involving these hard-to-clone genes would undoubtedly be of interest, it was felt to be of greater importance to study first those rearrangements involving more normal genes.

To do this, the following stategy is being followed. A plasmid has been constructed which contains both the *gpt* and the *neo* genes, as well as several kilobases of cloned DNA from the CHO *aprt* gene, kindly given to us by Dr. Mark Meuth, of the Clare Hall Laboratories, Imperial Cancer Research Fund (London).

This plasmid will be used to transfect CHO-K1 cells, one line heterozygous for *aprt* (D422), the other hemizygous for *aprt* (D423), cells lines generously given us by Dr. E. C. Bradley, of the Institut du Cancer de Montreal (Canada). First, cell lines with stable inserts expressing the *gpt* and *neo* genes will be isolated by the same methods we have used before. From these lines we will isolate those that are also *aprt*$^-$ because the plasmid has incorporated into the genomic *aprt* locus by homologous recombination, as has been demonstrated by, for example, Thomas and Capecchi (1986). In these procedures, we are in consultation with Dr. John Sedivy, an Assistant Professor in this department, who is using the same basic scheme for targeting drug resistance markers to genes that are important in embryonic mouse cells.

We are reasonably confident that it will be possible to clone out various DNAs from the region of the *aprt* gene, because both Dr. Meuth at ICRF (see above) and Dr. Barry Glickman, of York University, Toronto, Canada, have been able to clone DNA from the *aprt* region. Dr. Meuth is also making available to us his extensive restriction map of the DNA surrounding the CHO *aprt* locus, as well as several probes for DNA sequences at various distances from the *aprt* gene. These probes can be used to ascertain the sizes of deletions involving this region of the chromosome.

The use of our construction, with a *gpt* and a *neo* gene inserted in the resident *aprt* gene,

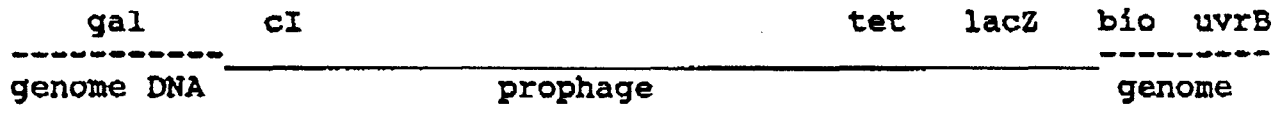
has several advantages. Most deletions can be quickly identified by assaying for both *gpt*$^-$ and *neo*$^-$. Gross DNA rearrangements with one end in the roughly two kilobases of DNA between the *gpt* and *neo* genes can be isolated quickly by assaying gamma-irradiated cells for *gpt*$^-$ in the presence of G418, so that only cells with active *neo* genes will grow. Those mutants that have altered Southern blots will have some kind of a rearrangement in the DNA between the *gpt* and *neo* genes, which can now be isolated for sequencing to determine in detail what has occured. (Mutants in which a point mutation in a restriction site causes a change in the Southern blot are expected to be rare.) The presence of an active *neo* gene in the genome should make it much easier to trace one end of the rearranged DNA and so determine what has taken place. The location of the insert in the *aprt* locus will make it possible to integrate the results with those published on rearrangements in the *aprt* locus by other workers (e.g., Breimar et al., 1986; Grosovsky et al., 1986). Finally, similar experiments can be done with the insert in the heterozygous and the hemizygous strains, to determine what effects adjacent DNA sequences present in the heterozygous strain, but not in the hemizygous strain, have on the observable spectrum of gross DNA rearrangements.

As mentioned in last year's Progress Report, I have become interested in the induction of gross DNA rearrangements in prokaryotes. In particular, I have hypothesized that the low level of observed rearrangements, such as deletions, induced by ionizing radiation in bacteria may be a consequence of the organization of the bacterial genome. From data in mammalian cells, it is reasonable to guess that the mean size of a deletion induced by gamma rays is the order of tens of kilobases. Since most mammalian genes are larger than this (because of the introns), such a deletion will usually affect only a single gene. However, if gamma rays should induce deletions of similar size in an *E. coli* genome, the deletions would in general involve a number of genes. If any one of these were essential, the deletion would be a lethal mutation, the cell would not survive, and the deletion would not be observed.

As mentioned in last year's Report, we are testing this idea by constucting a lambda phage with several assayable genes, and which can be incorporated as a prophage in the genome of the *E. coli* host cell. Since the prophage DNA is not essential to cell survival, about 50 kilobases of

DNA can be removed without, presumably, affecting the ability of the cell to survive.

The phage reported in the Progress Report for 1987-88 turned out to have properties that made it unsuitable for the proposed experiments. Therefore, a new lambda phage has been constructed. When this phage is incorporated in the genome, assayable genes are located as follows.



In practice, the activity of the *lacZ* gene is assayed by the usual method with X-gal, selecting white or light blue colonies. The activities of all the other genes, cI and *tet* on the prophage, and *gal*, *bio* and *uvrB* on the genomic DNA, can then readily be assayed.

In preliminary experiments, *lacZ* mutants induced by gamma rays have tended to have lack of *tet* and cI activity, whereas *lacZ* mutants induced by ultraviolet light have, in nearly all cases, have had active *tet* and *uvrB* genes. An advantage of the prophage structure is that it will be relatively simple to get data on the sizes of the induced deletions. The base sequence of the lambda prophage is known. Therefore, the sizes of various deletions can be determined by synthesizing oligonucleotide probes with the base sequences of a number of sites along the prophage, and using these in colony hybridization experiments.

These experiments are now continuing, but at a relatively low level of urgency, since we more interested in some experimental data on rearrangements in the mammalian cells as soon as possible.

A theory of the formation of deletions has been derived on the basis of concepts introduced by Szostak et al. (1983). It is assumed that deletions are formed by the two ends of a DNA double strand break recombining with a part of the same DNA molecule having at least some homology with the sequence at the site of the break, in the manner suggested by Szostak et al. If the break is filled in and the resulting Holliday junction is resolved in one direction, the

end result is a repair of the double strand break. If the Holliday junction is resolved in the other direction, the result can be a deletion.

The partial homology between the base sequence at the break and that at some other part of the molecule is assumed to arise more or less at random. The probability of a suitable sequence in a short length dx of DNA is taken as $q\,dx$. It is then possible to obtain analytical expressions for the probability of a deletion of length D involving a gene of length L, where the nearest essential genes are at distances of A and B to the left and the right of the gene. (If a deletion removes an essential gene, the cell will not live and the deletion will not be observed.)

The analytical expressions obtained show various obvious features: a large deletion is more likely to be observed than a small one, because it has a greater chance of removing part of the gene in which mutations are being observed; the number of observed deletions involving a given gene will be low if the gene is closely flanked by essential genes; etc. It is hoped to obtain soon some data on deletions which can be used to determine the usefulness of the theoretical expressions for deletion frequencies and relative sizes.

PUBLICATIONS SUPPORTED BY THIS GRANT

Hutchinson, F., in *Mechanisms and Consequences of DNA Damage Processing*, edited by E.C. Friedberg and P.C. Hanawalt, Alan Liss, New York, 1988, pp. 505-509. Interpreting the DNA sequence changes induced by mutagenesis of mammalian cells.

Tindall, K. R., J. Stein, and F. Hutchinson, Genetics 118: 551-560 (1988). Changes in DNA base sequence induced by gamma-ray mutagenesis of lambda phage and prophage.

Hutchinson, F., Mutat. Res., in press (1989). Use of data from bacteria to interpret data on DNA damage processing in mammalian cells.

REFERENCES

Breimer, L. H., J. Nalbantoglu and M. Meuth, J. Mol. Biol. 192: 669-674 (1986).

Grosovsky, A.J., P. J. Drobetsky, P. J. deJong and B. W. Glickman, Genetics, 113: 405-415 (1986).

W.A. Loenen and F. R. Blattner, *Gene* 26: 173, 1983.

Szostak, J.W., T.L. Orr-Weaver, and R.J. Rothstein, *Cell* 33:25-35, 1983.

K.R. Thomas and M. R. Capecchi, Cold Spring Harbor Symp. Quant. Biol. 51: 1101, 1986.