Paul O.P. Ts'o, Ph.D.
363-34-4378

DOE/

DOE/ER/60636--T4

DE92 000706

Genes and Gene Expression:
Localization, Damage and Control - A Multilevel and Inter-Disciplinary Study

Progress Report

February 1, 1991 - January 31, 1992

Paul O.P. Ts'o, Ph.D.

The Johns Hopkins University
615 N. Wolfe Street
Baltimore, Maryland 21205-2179

September, 1990

Prepared for

The U.S. Department of Energy
Agreement #DE-FG02-88ER6036

Notice: This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed or represents that its use would not infringe privately-owned rights.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Overview

During the past year, significant progress has been made in all three projects. This overview will provide a coherent description of how the progress and directions in these three projects have merged together toward a common goal - the establishment of a three dimensional nuclear topography of genes and the establishment of procedures to mask or perturb a specific sequence of DNA within the living cells as well as method to monitor such an event. Six scientific papers supported in part or in full by the Department of Energy Contract have been published in 1900-1991, and additionally three papers are in press and three manuscripts have been submitted. All these materials are attached in the appendix.

Project I: Two significant advances has been made. First, we are now capable of detecting a single copy unique DNA sequence on a metaphase chromosome. This experiment was carried out using the human cervical cancer cell lines (SiHa cells) which contains one or two copies of integrated human papilloma virus (HPV) DNA, as well as with human Caski cell line which contains approximately one hundred copies of HPV DNA integrated within its genome. We are able to specifically locate and map the position of the integrated HPV DNA on the metaphase chromosome. Furthermore, the copy number of the integrated HPV DNA can be measured quantitatively⁽¹⁾. Second, we are able to simultaneously measure several three different fluorochromic centers in a single nucleus. Thus, we are able to provide a three dimensional description of a spatial relationship among these fluorochromic centers within a nucleus.

Utilizing this approach, we are able to locate the centromeres of three pairs of chromosomes, chromosome #7, #11, and #17 in eleven individual nuclei. This is done by the identification of specific satellite DNA sequences within the centromeres of each of these chromosomes in interphase nuclei. Through our computer-assisted microscopic imaging system (CMIS), we have optically sectioned the nucleus through the vertical or "Z" axis, and located each of the six centromeres using specific satellite DNA probes. In this report, we will show the preferential spatial location of the centromeres in these three sets of chromosomes with reference to each other. This preliminary data clearly indicates that the 3-D spatial relationship of the centromeres within these three chromosomes relative to each other is **not random** in an interphase nucleus. This data is the first available evidence to reach this conclusion.

We expect that this data will become much more defined as we complete our study on more nuclei. In addition, we expect to significantly improve these procedures for the cells on slides so that the nucleus structure will be less damaged.

Finally, it is important for us to greatly improve the fixation and the denaturation procedures. Current treatments for the denaturation and reannealing, which entail the heat and formamide treatment, could cause substantial damage and disorientation of the nuclear DNA topography. We would prefer to search and locate a specific sequence of DNA without resorting to denaturation and renaturation which is quite destructive. The only way we can do that is through the triplex formation which is currently being pursued in Project III. Preliminary success will be shown in the report of Project III. This approach will greatly improve the study of the nucleus in terms of

specific spatial orientation of DNA sequences with minimal perturbation to the original nuclear structure.

Project II: The main approach is in identifying restriction fragments with the highest resolution and sensitivity with the most convenient procedures and yielding maximal amounts of information. We have made two significant advances in this area.

First, we have now perfected the procedure for constructing maps for the restriction sites (in reality, two restriction maps), of DNA approximately 50-150kb in size. DNA fragments of this size usually come from cloning plasmids (Cosmids) or yeast artificial chromosomes (YACs). This facile and powerful procedure allows us to break-down the DNA of 50-150kb in size, specifically, into ten to twenty restriction fragments. We can construct a linear map for the restriction sites of such a large piece of DNA. The individual DNA restriction fragments arranged in a linear order in the plasmid or cosmid can now be sequenced to obtain nucleotide sequences. A manuscript to be submitted is attached in the appendix⁽²⁾. We have applied this technique and successfully mapped the restriction sites of a cosmid containing human sequence on chromosome #21.

Secondly, progress has been made in the study of DNA rearrangement. We can now study the DNA rearrangement with a genome size of approximately 5mb, such as that of bacteria and yeast with considerable assurance as shown in our paper⁽³⁾. Furthermore, this procedure has been written as a chapter in *Methods in Enzymology* as a procedure to study unknown DNA rearrangement in an organism with a genome size less than 10mb. Currently, we have used a human repetitive sequence (LINE-1 Family) as a probe to describe a human genome in a 2-D gel electrophoresis map.

Based upon our experience and achievement in the past five years, we are now trying to develop a procedure to map unknown DNA rearrangement and unknown insertion of DNA in mammalian genomes. The procedure is based upon the approach of locating one to several restriction fragment changes amidst two to three million restriction fragments. We hope this procedure, which is in the last stage of its development, will be ready in the next six months.

Project III consists of two sections, Section "A" is organic chemistry and Section "B" is physical chemistry. In the organic chemistry, there are two areas of research. The first area consists of the study of the change of the sugar portion of the backbone of the oligonucleotide analogue from deoxyribose to 2'-O-methylribose. Since the target nucleic acid is a RNA with ribose sugar backbone, there is an indication that the probe should also consist of ribose (2'-O-methylribose in this case) for higher stability. The affinity between the oligomer-probe and the target RNA in duplex formation may become higher with ribose/2'-O-methylribose backbone. Currently, all four methylphosphonates containing 2'-O-methylribose can now be synthesized in order to test this hypothesis. In the second area, the uptake of non-ionic oligonucleoside methylphosphonate is being investigated by the use of modified MatagenTM containing a fluorescent chromophore. Such a fluorescent probe of MatagenTM has been synthesized and the kinetics of the uptake of this fluorescent probe into living mouse L-cells is being followed, quantitatively, by the use of computer-assistant fluorescent microscopy. In the physical/chemical study, the major effort is on the investigation of triplex with considerable progress leading to five publications. We first

demonstrate the physico-chemical characteristics of a triplex with a model system of 16-mers, both with phosphodiester backbone and methylphosphonate backbone. Secondly, we show that with a pseudocytosine as an analogue in replacement of cytosine, a triplex of pIC•G•C motif can be formed with neutral pH and physiological condition. Finally, we are able to show the formation of a triplex with purine•purine•pyrimidine motif with the methylphosphonate backbone for the third purine strand. It appears that we may soon be able to study the formation of triplex between the Matagen™ probe and the dsDNA inside living cells.

In summary, all projects are working toward a goal for describing the three dimensional nuclear topography in terms of relative spatial relationships among genes (specific DNA sequence). Methods are now being perfected to detect these genes, quantitatively and spatially, to perturb these genes specifically, and to measure the perturbation in order to assure specificity. We are developing methods to assay, after perturbation of the target DNA within living cells, whether or not only the target sequence are attacked while other sequences remain unharmed. We are now at the stage to do chemical gene modification or masking within living cells in a strictly sequence-specific manner. Soon, we will be able to study the function and the physical location of each gene in living cells with exquisite specificity.

Project I

Molecular Cytology

I. Introduction. The approach of Project I can be summarized in the following three aspects:

- A. The Development of instrumentation and detection technology. This phase of the work has been described in last year's Progress Report. Essentially, it is all completed in terms of the hardware. There will be a continuous improvement and upgrading of software as new demands appear and upgraded software packages become available. Currently, we have a very effective state of the art microscopy coupled with computer technology.
- B. Quantitative detection of a unique DNA sequence in metaphase chromosomes using a fluorescence microscope imaging system. We are able to detect single copy DNA within the nucleus as well as to quantitatively measure multiple copies of the same DNA with a narrow margin of numerical error. In order to proceed with such a study, we have utilized the integrated human papillomavirus Type 16 DNA in two human cervical carcinoma cell lines as the model system. Essentially, this study is complete as described below with a manuscript submitted for publication listed in the appendix.
- C. The Development of an approach and technology to study three dimensional genome topography within the nucleus, particularly within the interphase nuclei. After much hard work, we now have preliminary results on the location of the centromeres of three human chromosomes (chromosomes #17, #11, and #7) within human lymphocytes described by 3-D coordinates. As described below, eleven interphase nuclei has been studied and the location of the centromeres of these chromosomes within the interphase nucleus has been described in a 3-D system. The preliminary results strongly indicate that the location of these three

chromosomes relative to each other do have a definitive order within human lymphocytes. This is a very exciting discovery. On one hand, our results demonstrate the effectiveness of the instrumentation for such a study; and, on the other hand, the results strongly indicate that the direction of our research is fundamentally correct, i.e., that the chromosomes within the interphase nucleus have a definitive 3-D topography relative to each other. Apparently, what is required is a better procedure for fixing the cells on the slides in order to reduce artifacts in the preparation. Some very interesting photographs and illustrations are shown in the figures.

II. A more detailed description of our progress is outlined below.

A. Quantitation and Mapping of Integrated Human Papilloma Virus DNA in Human Metaphase Chromosome using a Fluorescence Microscope Imaging System⁽⁴⁾.

1. Integrated Human Papilloma Virus type 16 (HPV-16) DNA was directly visualized on metaphase chromosomes in the two human cervical carcinoma cells lines, SiHa and CaSki, by fluorescence *in situ* hybridization with a biotinylated DNA probe (7.9kb). A single fluorescein signal, usually about four pixels in size, was observed on one or both chromatids of a large acrocentric chromosome in metaphase chromosome spreads from SiHa cells. A similar event was also observed in CaSki metaphase spreads; however, about eight to ten of the other chromosomes had bright multipixel fluorescent signals. Bright, multi-pixel fluorescent signals were also observed in the interphase nuclei of CaSki cells.
2. The fluorescent intensities of hybridization signals from single copies and dispersed clusters of integrated HPV-16 DNA were quantified using a microscope equipped with a cooled-CCD camera that was interfaced to an image processor and host computer. Hybridization signals were localized on chromosomes using separate, registered images of 4',6-diamidino-2-phenylindole (DAPI) or propidium iodide stained metaphase chromosome spreads. In both SiHa and CaSki spreads, a single fluorescein signal was observed on one or both chromatids of chromosome #13 which was identified by simultaneous hybridization with a biotinylated centromere probe specific for chromosomes #13 and #21. Ratios of the distance from 13pter to the HPV-16 signals to the entire chromosome length were approximately 0.63 ± 0.05 in both SiHa and CaSki cells, indicating the possibility of a common integration domain on chromosome #13. In SiHa cells, no additional signals were observed on other chromosomes. This observation, taken together with literature reports that SiHa cells contain one to two copies of the HPV-16 genome in this region of chromosome #13 suggests that each fluorescein signal on chromosome #13 represents one equivalent of the HPV-16 genome.
3. The integrated intensities of 25 individual fluorescent signals on chromosome #13 of SiHa cells were calculated and compared to the integrated intensities of all fluorescent signals observed in twelve separate CaSki spreads. This data shows that the total integrated intensity in CaSki spreads was approximately two orders of magnitude greater than the integrated intensity of individual fluorescent signals on chromosome #13 of SiHa cells. The CaSki and SiHa cell lines are a good model system for validation of quantitative fluorescence

imaging because the relative amount of HPV-16 in the two cell lines has been determined and the integration site of a single copy of HPV-16 DNA has been mapped on SiHa cells to a well defined site on chromosome #13. Our data show that CaSki cells contain greater the 60 copies of the HPV-16 genome, which is in agreement with reports that this cell line contains 50 times more HPV-16 DNA than SiHa cells.

B. Genome Topography of the Interphase Nucleus.

1. Simultaneous *in situ* hybridization and detection of three differentially labelled DNA probes.

- a. DNA probes specific for the following human alpha satellite DNAs were purchased from ONCOR, Inc., biotinylated chromosome #11, chromosome #17, and chromosome #7. The chromosome #7 and #17 DNA probes were labelled by random primer extension with digoxigenin-11-dUTP and dinitrophenol-11-dUTP (DNP-11-dUTP), respectively. DNP-N-hydroxysuccinimide ester was prepared from DNP and N-hydroxysuccinimide and used to convert 5-(3-amino) allyl-dUTP to DNP-11-dUTP. *In situ* hybridization was carried out in 65% formamide-2x-SSC-10% dextran sulfate at 37° overnight using the instructions in a kit available from ONCOR, Inc. After the washes of appropriate stringency, the hybridized probes were detected with red, green and blue fluorophores as follows: digoxigenin with a rhodamine-conjugated sheep anti-digoxigenin antibody; DNP with a goat anti-DNP antibody followed by a fluorescein conjugated rabbit anti-goat antibody; biotin was detected with avidin cascade blue followed by amplification with a biotinylated anti-avidin antibody and a second layer of avidin cascade blue. Figure #1 shows the detection of three different centromere probes in the same human lymphocyte; chromosome #7 centromere homologs are red, the #11 homologs are blue and the #17 homologs are green.

2. Optical sectioning of interphase nuclei in fixed human lymphocytes.

- a. The methodology described in the above paragraph was used for the simultaneous *in situ* hybridization of fixed human lymphocytes with three differentially labelled human alpha satellite DNA probes. Eleven of the stained nuclei were optically sectioned using our Zeiss Axiovert 35 microscope equipped with a Z-axis stepping motor. Using a 100x objective, stacks of 32 images at 0.23 micron increments were obtained for blue, green and red fluorescence emission with a DAPI, FITC and rhodamine filter combination, respectively. At this magnification, the pixel dimensions are 0.23x0.23 microns. Subarrays (64x64) of the monochrome images were flat-fielded and registered in both the "X,Y" plane and the "Z" direction. Registration of the images was facilitated by the fact that the red and green fluorescent centromeres, viz #7 and #17, respectively, were faintly visible in the images taken with the DAPI filter set. Thus, due to the high sensitivity and dynamic range of the CCD camera and the excitation/emission characteristics of the blue filter combination, all six centromeres could be observed in images taken with the DAPI filter set. The registered images were overlaid to produce color composite images which were then assembled into a stack showing the 3-D orientation of the six different chromosomes. The registered, color composite stack of images was used to obtain the

following data: distance between the homologs, distance of the centromere from the center of the nucleus and the angle (counterclockwise from the x-axis) generated by a line from the center of the nucleus to its respective centromere. The normalized data (the distance from center and counterclockwise angle from the x-axis) for all eleven nuclei are presented in Figure #2. In this diagrammatic representation of the data, the chromosome #7 centromeres (red) were oriented to a common position with respect to the x-axis. In the 2-D projection, the red centromeres all fall in the right hemisphere, the blue centromeres all fall in the top hemisphere and the green centromeres lie in a path near the center of the nucleus along the y-axis. The data indicate that there exists a definite spatial organization of these three chromosomes in the interphase nuclei. It should be noted that optical sectioning was required to obtain this data since the six centromeres were in various optical planes through the nucleus. Two different views of six centromeres of one nucleus in 3-D space are presented in Figure #3A (top view) and Figure #3B (side view). At present, it is not possible to faithfully reproduce red, blue and green in the 3-D software package. One can readily see that centromere #11B is above the two chromosome #17 centromeres while #11A is below. The two chromosome #7 homologs appears to be in the same plane. It should be noted that the 3-D photos shown here were generated using unprocessed image stacks. The use of Agard's⁽⁵⁾ deconvolution software should greatly improve the resolution along the Z-axis. This software is currently processing our optical sections on the VAX station II/GPX.

III. Future Studies.

- A. The experiments conducted thus far on genome topography of the interphase nucleus have been carried out with human lymphocytes that have been fixed and dropped on slides by conventional methods. These nuclei were rather flat due to drying. Additional heating of the slide was done so that optimal fluorescent signals could be obtained after *in situ* hybridization. We now plan to refine the fixation procedure in order to keep the nuclei as close to their normal shape as possible. *In situ* hybridization and subsequent detection of the hybridized probe will be carried with fixed cells in suspension in order to eliminate any drying which tends to flatten the cells. The fluorescently labelled cells will be suspended in 90% glycerol/antifade and then a 5 μ l of the suspension will be placed on a slide previously treated with a 0.1% gelatin of 0.01% chrom alum, before the gentle additional of a coverslip. It is hoped that this techniques will keep the cells in their normal shape and will sufficiently immobilize them on the slide so that they can be optically sectioned. When these new methods are in hand, we will use our fluorescence imaging system to determine the relative positions of the centromeres of chromosomes #7, #11, and #17 within the cell nucleus. A number of nuclei will be examined to determine if a particular 3-D pattern can be found in unstimulated human lymphocytes. If these cells exhibit a unique 3-D orientation with respect to chromosomes #7, #11, and #17, then a second set of these chromosome centromeres will be examined. If our prediction of a unique 3-D orientation for interphase chromatin holds when we have the second set of centromeres, then the effect of growth stimulation and differentiation on chromosome orientation will be examined.

Computer-generated composite image showing three probes in one interphase nucleus. Biotinylated chromosome 11 centromeres are stained blue, digoxigenin-labeled chromosome 7 centromeres are stained red and DNP-labeled chromosome 17 centromeres are stained green.

FIGURE #1

FIGURE #3

FIGURE #2

Project II

Genome Organization and DNA Arrangement - Investigator: Paul O.P. Ts'o, Ph.D.

- I. **Introduction.** Research in this project can be classified into two major areas. In the first area, we fully utilize the resolution power of the 2-D gel electrophoresis as well as the advantage of additional enzyme treatment of the restriction fragment in-gel. Thus, we are able to develop a very effective and powerful procedure for the construction of maps for restriction sites of DNA with an approximately size of 50-100kb. DNA of this size usually comes from cloning of sequence in plasmid, cosmid or in yeast artificial chromosomes (YACS). By appropriate choice of restriction fragments, we can break-down DNA of this size to 10-20 restriction fragment pieces and put them in a linear order for the construction of a restriction site map for this 50-100kb DNA. The individual restriction fragments can then be sequenced at the nucleotide level. Therefore, we consider this step very useful for human genome projects. The entire procedure has been worked out using λ -DNA as a model system. The procedure has now been written in a manuscript submitted to *Analytical Biochemistry*. In addition, the application of such a procedure for the construction of a restriction map for a cosmid containing human DNA from human chromosome #21 is now reported here as well. We consider this research from our group useful to the human genome project at Department of Energy. The next area involves the original studies for locating the unknown DNA rearrangement or unknown DNA insertion in a genome. Previous studies allow us to establish such an analysis for the genome of an organism with less than ten million base-pairs, a genome for bacteria or yeast. Based upon this experience, we are now in the process of constructing a procedure which will allow us to detect unknown DNA arrangement and unknown insertion in mammalian genome. These studies involve procedures for locating the changes of a few restriction fragments amidst two-three million pieces of restriction fragments. At the same time, we were able to utilize the mid-repetitive DNA sequences as probes to monitor the sequences neighboring to such a repetitive DNA family which serves as a detection post for change in sequences. Such a study has been carried out in the genome of Syrian hamsters with the repetitive IAP sequence for probes (Appendix I). Current studies involving the use of human repetitive DNA sequences of LINE-1 families to detect rearrangement of changes of human sequences adjacent to a LINE element. Preliminary results prove to be promising as indicated below. This can provide a meaningful and facile fingerprinting for the identification of individual DNA organization.
- II. A more detailed description of our progress is outlined below.
 - A. **"A Procedure for the Construction of Linear Restriction Fragment Maps of a 50-100kb DNA."**
 1. A simple and effective procedure for the construction of linear restriction fragment maps was developed. Using a two-enzyme digestion, two-dimensional (2-D) electrophoresis procedure, all the restriction fragments in a 50-100kb DNA can be individually resolved and displayed on a 2-D plane. This 2-D gel pattern, with appropriate markers, provides a fixed set of "x,y" coordinates for each fragment obtained from the single and double digestion as

well as the relationship between the two steps. A matrix is constructed from the 2-D pattern. The vertical column shows all the singly digested individual fragments and their sizes obtained from each restriction enzyme treatment, and the dividing horizontal row shows all the doubly digested DNA fragments and their sizes after treatment with two enzymes. The order of arrangement is always from the smallest to the largest fragments. Using this matrix, two linear DNA restriction maps for these two enzymes can be simultaneously constructed in a self-reconfirming manner. As examples for this procedure, we described the construction of two linear restriction fragment maps, a combination of *Eco RI* and *Bam HI* digestion, as well as a combination of *Eco RI* and *Hind III* digestion of λ phage DNA. A detailed description of this research is given in Appendix XIII.

B. Progress Report on Human Cosmid Restriction Mapping.

1. A cosmid consisting of approximately 40kb human genome sequence and 5.6kb vector sequence given by Stylianos E. Antonarakis, M.D., of the Johns Hopkins University, School of Medicine, was used as an example for constructing a previously unknown restriction fragment map. The cosmid was constructed by Murray, *et al.*⁽⁶⁾, with a vector which had been previously described⁽⁷⁾. Based upon the known vector sequence, the restriction sites for approximately 40 enzymes were determined for this vector. The restriction enzymes used in this 2-D restriction fragment analysis were chosen based upon the following criteria:
 - a. The position of the cutting sites in the vector DNA;
 - b. The suitable number of restriction fragments produced by the enzyme and the distribution of the sizes of these fragments;
 - c. The number and the distribution of the cutting sites for the second enzyme in the restriction fragments produced by the first restriction enzyme and conversely, the number and the distribution of the cutting sites for the first enzyme in the restriction fragments produced by the second enzyme;
 - d. The price of the enzyme since relatively large quantities of enzymes are required.
2. Based on the 1-D analysis of single and double digestion, *Hind III* and *Eco RI* were chosen in this experiment as the restriction enzyme pair.
3. In the first set of the experiment, 1.5 μ g of cosmid DNA was digested with *Hind III* in solution and the obtained *Hind III* fragments were electrophoresed in a one percent agarose gel slab. The DNA fragments containing portion of the gel was cut into a 16x0.8 strip which was then digested with 2,000 units of *Eco RV* *in situ*. After digestion, the gel strip was fused with another agarose gel slab. The second electrophoresis was performed at the dimension perpendicular to the first one. Visualization of the DNA fragments was achieved by ethidium bromide staining. In the second site of the experiment, the procedure was the same as described above except the two restriction enzymes were used in reversed order (Figure #4).

4. One of the most important requirements of this procedure is to ensure that all of the DNA fragments are retained and detected. Higher concentration of the agarose gel (1.5%) and the increased amount of cosmid DNA (10 μ g) allowed the detection of small DNA fragments, such as 600bp in size. Polaroid film (#667) was used during the photography with one second exposure to avoid the bleaching of the ethidium bromide staining (Figure #5).
5. A matrix was constructed from the above two 2-D maps which provided the molecular sizes of all the DNA fragments and the relationship between the singly digested and doubly digested fragments in both the forward and reversed treatment order (Figure #6). The structure of the matrix and the basic procedure to trace-in the matrix to construct the restriction fragment maps had been described in Appendix XIII⁽⁸⁾.
6. The restriction fragment maps of the cosmid DNA for both *Hind III* and *Eco RV* were constructed by tracing throughout the matrix (Figure #7). The tracing in the matrix represented the order and the overlaps of the fragments. Because not all the restriction fragments obtained from the first restriction enzyme digestion contained cleavage sites for the second restriction enzyme and vice-versa, the order or the orientation of two fragments in *Hind III* map and the order of two fragments in *Eco RV* map could not be explicitly determined; however, the fragment order on the *Eco RV* map of the corresponding (or opposite) part of that two ambiguous *Hind III* fragments in the map remain to be unique. The two ambiguous fragments on *Eco RV* maps also had the corresponding (or opposite) part on the *Hind III* map uniquely determined. It is indicated that the success of this approach heavily depends on the choice of the enzyme pairs.
7. A sensitive *in situ* labelling method has been developed to decrease 20-100 fold the amount of DNA required by ethidium bromide staining approach as shown below. description of "In situ labelling of DNA fragments with sequenase."

C. *In-situ* Labelling of DNA Fragments with Sequenase.

1. To label DNA fragments *in situ* in the agarose gel with sequenase is a sensitive method which decreases the DNA amount required by ethidium bromide staining approach for restriction map construction for 20-100 times. Using λ -DNA 2-D pattern as an example, the procedure is as follows:
 - a. 50ng of λ -DNA was digested with *Hind III*. The λ /*Hind III* was electrophoresed in one percent agarose gel. The DNA containing portion of the gel was cut into a 16x0.8cm strip which was then digested by 2,000 units of *Eco RI* *in situ*. After digestion, the gel strip was put into a 26cm tube which contained 50uci of [α -³²P] dATP, 50 units of sequenase, Version 2.0 DNA polymerase (United States Biochemical) and sequenase reaction buffer (recommended by U.S.B.). The tube was gently rotated at 37° C overnight. The gel strip was washed with TPE electrophoresis buffer for a total of 20 hours with gentle shaking and with six times exchanging of the buffer. The gel strip was fused with another gel slab and the second electrophoresis was carried out at the dimension perpendicular to the first

one. After autoradiography, all of the expected 12 spots were revealed in 2-D pattern (Figure #8).

2. This approach can be used to label the DNA fragments having 5'-protruding end, such as those generated by *Eco RI* with high sensitivity. The DNA fragments with 3'-protruding end or blunt end can not be labelled by this approach. The classification of restriction enzymes that generate 5'-overhanging, staggered cuts and the radionucleosides triphosphates which can be used to label corresponding DNA termini had been described⁽⁹⁾.

D. LINE-1 Finger Printing of Human Genomic DNA.

1. We have used the 5'-end segment of a LINE-1 (L1) family sequence as a probe for finger printing of human genomic DNA organization. The following progress has been made.
2. DNA isolation from human blood has been done in accordance to a published procedure⁽¹⁰⁾. About 10ml of blood is sufficient for the preparation of 20 μ g of DNA. Later, for the purpose of developing the technique, commercially available human DNA from placenta or blood (Clonotech and Promega) were used.
3. Two dimensional electrophoresis and Southern transfer technique have been verified by the use of λ -DNA as a positive control. This step is to ensure the completeness of restriction enzyme hydrolysis, 2-D separation, and the transfer of DNA fragments from the gel to the membrane. Our experiments indicate that 700 units of restriction enzyme are needed for complete digestion of 10 μ g of λ -DNA in-gel. Both small 2-D electrophoresis (20cm x 25cm) and the large 2-D gel electrophoresis (40cm x 50cm) have been done for human DNA fragments.
4. Hybridization with a cloned (L1) fragment (208bp) has been investigated for varying degrees of stringency. Currently, the adopted procedure is as follows: Hybridized at 42° C for 20h in a solution containing 40% formamide, 0.6M NaCl, 0.06M Sodium Citrate, 0.1% SDS, 20mM Tris-Cl (pH 7.5), 40 μ g/ml Salmon Sperm DNA, 0.02% Ficoll, 0.02% BSA, 0.02% polyvinyl pyrrolidone, and radio-labelled L1 probe ($> 1.0 \times 10^7$ cpm/25ng) followed by washing with 2xSSC-0.1% SDS at 42° C for 15 minutes, twice and with 0.1xSSC-0.1% SDS at 68° C for ten minutes, twice. As shown in the duplicate x-ray film (a transparency created by photocopying - Figure #9), after exposure to the membrane obtained from the small 2-D gel electrophoresis, each spot was almost resolved in the gel. The enzyme treatment procedure is using one enzyme for the first digestion and two enzymes for the second digestion. Still, we need more resolution to identify and compare each spot from two or more different DNA sources; therefore, we attempted to use a bigger gel (40cm x 50cm). The hybridization with a membrane of large size has met with experimental difficulty in handling because of the size (35cmx40cm). The hybridization is much more convenient if the membrane is first cut into smaller pieces. Under such a situation, it can be difficult to maintain the original geometry/x-y coordinates of the spots on the membrane. We are experimenting with various approaches to solve this problem. Also, we are developing a method of hybridization of this large membrane without cutting it.

5. As shown in Figure #9, considerable fragments remain in the diagonal line. These fragments are the DNA fragments obtained after the first enzyme digestion (*Pst I*), which have no restriction sites for the second and third enzymes (*Eco RI* and *Hind III*). One way to improve the resolution is by the treatment of three enzymes (*Eco RI*, *Hind III* and another enzyme) in the *in situ* gel digestion, so that the number of fragments remaining in the diagonal line are substantially reduced. If the 2-D plane becomes too crowded because of the use of two or more enzymes in the second in-gel digestion, we can use only one enzyme in the second digestion; however, before the Southern blotting, we can cut up the gel area at the diagonal line containing considerable DNA fragments, and this gel strip can be treated again by the third (or even a fourth enzyme) followed by another electrophoretic separation. In other words, the first 2-D gel electrophoresis with two enzymes followed by the second 2-D gel electrophoresis with the diagonal gel strip treated by one or two more enzymes should be sufficient to resolve most of the fragments containing the LINE-1 element in human genomic DNA.

FIGURE #4

Hind III
↓

EcoRV
→

FIGURE #5

EcoRV
↓

Hind III
→

1c

1d

Eco R V
&
Hind III

Hind III

FIGURE #7

Cosmids

FIGURE #8

In-Situ Labelling of Restriction Enzyme

Hind III

→ EcoRI

Probe: L-1/H-X (fr)
Hybr: 43°C 21h. byproducts 76 Dext Sulf.
Ratner Bottle
wash: 2550cc-017505 43°C 15m 2, 0450cc-017505 43°C 15m 2, 45°C 15m 2

8/22 (W.C.) Pro. 10/73
Pst I (10)
Hind III (20)
Eco RI (20)
→

FIGURE #9

Project III

Antisense Oligonucleotide Analogues.

III. Introduction.

- A. Project III consists of two sections, Section "A" - Organic Chemistry and Section "B" - Physical Chemistry. In Section "A," the original objective described last year, is to increase the effectiveness in the binding of Matagen™, the oligonucleoside methylphosphonate, to the target RNA in a sequence specific manner. Since the target RNA consists of a ribose backbone and there exists a hypothesis that ribose - ribose duplex would be more stable than ribose - deoxyribose duplex. For the reason of chemical stability 2'-O-methylribose is used instead of ribose for the methylphosphonate linkage; therefore, current efforts are now being made to prepare methylphosphonate monomer containing 2'-O-methylribose for the synthesis of the oligomer, Matagen™, replacing the deoxy sugar with the 2'-O-methylribose sugar. The synthetic objective has now been achieved. In the upcoming months, data will show whether this hypothesis is correct. If so, then the affinity between the 2'-O-methylribose Matagen™ to the ribose target may be found to be higher than the affinity for binding of the 2'-deoxyribose Matagen™ with the RNA target.
- B. In the second section of section "A," a fluorescently labelled Matagen™ is being synthesized. This project was simulated by the availability and the power of the computer-assisted microscopy mentioned in Project I. Currently, this research is a collaborative effort between Project III and Project I. The objective is to observe the properties of the Matagen™ labelled with a fluorescent chromophore, particularly regarding its entrance into or exit from a living cell as well as accumulation within the cell. Again, the synthetic objective has now been achieved and the current study of the uptake and the fate of the fluorescent Matagen™ inside the mouse L-cell is in progress. In addition, work is ongoing to link a psoralen to the fluorescent Matagen™ as well, thereby a photochemically-induced crosslinking of fluorescent labelled Matagen™ to the cellular RNA target can be achieved. These studies are very interested, related to probing the function of nucleic acid within living cells as observed directly under the microscope.
- C. In Section "B," Physical Chemistry, the main objective in this section is to study the properties of triplex formation using oligonucleotides as models. We have chosen a 16-mer of homopurine sequence such as d(AG)₈ as well as 16-mer of homopyrimidine sequence such as d(CT)₈ as model compounds consisting of phosphodiester backbones or methylphosphonate backbones. At first, we established the physical/chemical characteristics of triplex both by optical studies and by NMR studies. Subsequently, the triplex formation was studied with the emphasis on the formation of triplex in neutral pH with high stability in physiological condition or inside living cells. In the first case, a specially designed analogues, 2'-O-methyl-pseudoisocytidine, was built to replace 2'-deoxy-cytidine in the homopyrimidine oligomer with 2'-O-methylribose backbone. The data clearly indicates a pyrimidine•purine•pyrimidine motif can be found in a stable triplex at neutral pH. Subsequently, our investigation has lead to the study of triplex formation of

purine•purine•pyrimidine motif utilizing d(AG)₈ of methylphosphonate backbone as the third strand. Again, a stable triplex at neutral pH was found solely with the third strand of d(AG)₈ with a methylphosphonate backbone and not with d(AG)₈ with phosphodiester backbone. Current progress encourages the possibility that with methylphosphonates together with analogues such as pseudoisocytidine, we will soon be able to probe, to mask, and perhaps to perturb in a sequence specific manner with a segment of DNA duplex inside living cells

D. Organic Chemistry - Paul S. Miller, Ph.D.

Scheme #1

Oligonucleoside methylphosphonates which are complementary to sequences in the initiation codon region or splice junction regions of mRNA or precursor mRNA can be used to control gene expression in mammalian cells and virus-infected cells in culture^(11,12,13,14). Methods are currently being sought to improve the effectiveness with which these nuclease resistant analogues bind to their targeted RNAs. Two strategies are being tested. One approach involves the synthesis of oligomers which contain

2' - O - methylribonucleosides in place of the 2'-deoxyribonucleoside units⁽¹⁵⁾. The other approach involves modifying the oligomer with a photoreactive group, psoralen, which allows the oligomer to cross-link to its mRNA target when irradiated with long wavelength (365nm) ultraviolet light^(16,17,18,19,20,21).

2. Oligo-2'-O-methylribonucleoside Methylphosphonates. We have previously reported on the synthesis of a pentamer r-U^mpA^mA^mU^mC, which contains 2'-O-methylribonucleosides, N^m, and methylphosphonate linkages as indicated by the underline⁽²²⁾. This pentamer forms a duplex with a complementary ribopentamer, r-GAUC_A, which is more stable than the duplex formed by d-TpGATC and the r-GAUC_A. We are extending our studies to a dodecamer, r-A^mpU^mA^mG^mG^mA^mU^mU^mU^mG^mU^mC, whose sequence is complementary to the coding region of the M protein mRNA of vesicular stomatitis virus (VSV). This oligomer

contains 2'-O-methylguanosine, a nucleoside which can form three hydrogen bonds with C, in place of the previously used 2'-O-methylinosine, which only forms two hydrogen bonds with C.

3. 2'-O-methylguanosine is unusually difficult to synthesize. It has been prepared in a six step procedure starting from the 3',5'-O-(tetraisopropylidisiloxane-1,3-diyl)-2-amino-6-chloropurine riboside using methyl iodide and the sterically hindered base, 2-*tert*-butylimino-2-diethylamino-1,3-dimethylperhydro-1.3.2-diazaphosphorin⁽²³⁾. Alternatively 2'-O-methylguanosine can be obtained directly by methylation with diazomethane at low temperature in dimethylformamide in the presence of the Lewis acid catalyst, stannous chloride⁽²⁴⁾. Although the latter method is more direct, it employs the hazardous and potentially explosive reagent, diazomethane.
4. We have found that trimethylsilyldiazomethane (TMSCHN₂) is a suitable substitute for diazomethane for the methylation of N²-isobutyryl-5'-O-(4-methoxytrityl)-guanosine (**1a**) and N²-isobutyryl-5'-O-(4,4'-dimethoxytrityl)guanosine (**1b**) (see Scheme #1). Monomers **1a** or **1b** are methylated in dichloromethane in the presence of stannous chloride and TMSCHN₂. At the end of the reaction a mixture of products corresponding to the 2'-O-methyl, the 3'-O-methyl and dimethylated nucleosides is obtained. The desired 2'-O-methylated nucleoside (**2a** or **2b**) is isolated and purified by silica gel column chromatography in 20% overall yield.
5. The protected nucleoside, **2b** is then converted to the methylphosphoramidite synthon **3b**, which will be used for the preparation of r-A^mpU^mA^mG^mG^mA^mU^mU^mG^mU^mC.

E.

Scheme #3

Fluorecently-Labelled Psoralen-Derivatized Oligonucleoside Methylphosphonates. In an effort to better understand the uptake and mechanism of action of antisense oligonucleoside methylphosphonates, we are preparing an oligonucleoside methylphosphonate derivatized with aminoethyltrimethylpsoralen at its 5'-end and with a rhodamine based fluorescent label at its 3'-end. The oligomer, whose sequence is complementary to VSV M protein mRNA and whose structure is shown schematically in Scheme #3, was synthesized on a controlled pore glass support derivatized with an aminopropyl linker arm as shown

in Scheme #2. This support was used instead of commercially available linker arm supports, when it was found that methylphosphonate oligomers could not be removed from these supports without considerable degradation. The support was synthesized in an eight step procedure starting from *bis*(1-hydroxyethyl)amine and the phthalimide derivative of 1-amino-3-bromopropane. The loading of the support was 33 μ mole/gram based on the amount of dimethoxytrityl group cleaved from the support.

Scheme #2

1. The oligomer was synthesized using standard protected nucleoside-3'-N,N-bis(diisopropyl)aminomethylphosphonamidite and protected nucleoside-3'-N,N-bis(diisopropyl)amino- β -cyanoethylphosphoramidite synthons in a Biosearch DNA synthesizer. The oligomer was deprotected by sequential treatment with hydrazine hydrate followed by ethylenediamine, and was purified by reversed phase HPLC. The deprotected oligomer ($R = OH$, $R' = H$) was derivatized with the fluorescent label by treatment with the commercially available isothiocyanate derivative of the dye, TRITC. This oligomer was phosphorylated using polynucleotide kinase and ATP to give derivatized oligomer ($R = OH$, $R' = \text{rhodamine label}$) which was then purified by reversed phase HPLC. This oligomer will be further derivatized with aminoethyltrimethylpsoralen using standard procedures⁽²⁵⁾.
2. The stability and uptake of the fluorescently tagged oligomer, ($R = OH$, $R' = \text{rhodamine dye}$) have been examined in preliminary studies. The phosphodiester group which links the rhodamine dye to the 3'-end of the oligomer is resistant to nuclease hydrolysis. Thus incubation of the oligomer in cell culture medium containing 10% fetal bovine serum for 24hrs at 37° C resulted in no change of the oligomer as assayed by reversed phase HPLC. This stability may be due to the presence of the bulky rhodamine dye and the positively charged tertiary amino group of the linker arm. The latter may form a zwitterion with the negatively charged phosphodiester group.
3. The oligomer, ($R = OH$, $R' = \text{rhodamine dye}$) appears to be taken up by mouse L cells (fibroblasts) growing in monolayer as determined by fluorescence microscopy. The cells were treated with media containing 1 μ M oligomer and then washed several times with

phosphate buffer saline followed by replacement with medium containing no oligomer. The cells were then immediately examined. Fluorescence was detected in distinct regions surrounding the nuclear membrane in the cells after the cells had been incubated with the oligomer for 45 minutes. We are currently attempting to quantitate the amount of oligomer which enters the cells, the rate of uptake and the distribution of the oligomer in the cells as a function of time. We will also examine uptake of the oligomer in VSV-infected cells to determine if the presence of target sequences influences the amount of oligomer which is taken up. Similar experiments will be carried out with oligomer which is derivatized with the aminoethyltrimethylpsoralen before and after irradiation at 365nm.

F. Physical Chemistry - Lou-Sing Kan, Ph.D.

1. The physical chemistry section of this project has resulted in the publication of four papers and one abstract which are listed in the appendix.

a. *"Proton NMR and Optical Spectroscopic Studies on the DNA Triplex Formed by d-A-(A-G)₇-G and d-C-(T-C)₇"*

- (1) Triplex and duplex formation of two deoxyribohexadecamers d-A-(G-A)-G (a) and d-C-(T-C)-T (b) have been studied by U.V., C.D. fluorescence, and proton NMR spectroscopy. Optical studies of (a) and (b) at dilute concentrations (μM range) yielded results similar to those seen for polymers of the same sequence, indicating that these hexadecamers have properties similar to the polymers in regard to triplex formation. The CD spectra of concentrated NMR samples (mM range) are similar to those observed at optical concentrations at both low and high pH, making possible a correlation between CD and NMR studies. In NMR spectra, two imido NH-N hydrogen bonded resonance envelopes at 12.6 and 13.7 ppm indicate that only the duplex conformation is present at pH > 7.7. Four new NH-N hydrogen-bonded resonance envelopes at 12.7, 13.5, 14.2 and 14.9 ppm are observed under acidic conditions (pH 5.6) and the two original NH-N resonances gradually disappear as the pH is lowered. Assignment of these four peaks to Watson-Crick G•C, Hoogsteen T•A, Watson-Crick A•Y and Hoogsteen C•G hydrogen-bonded imidos, respectively, confirm the formation of triple stranded DNA. NMR results also show that triplex is more stable than duplex at the same salt condition and that triplex melts to single strands directly without going through a duplex intermediate. However, in the melting studies, a structural change within the triple-stranded complex is evident at temperatures significantly below the major helix-to-coil transition. These studies demonstrate the feasibility of using NMR spectroscopy and oligonucleotide model compounds (a) and (b) for the study of DNA triplex formation.

b. *"Comparative Circular Dichroism and Fluorescence Studies of Oligodeoxyribonucleotide and Oligodeoxyribonucleoside Methylphosphonate Pyrimidine Strands in Duplex and Triplex Formation."*

- (1) An analogue of the homopyrimidine oligodeoxyribonucleotide $d(CT)_8$ has been synthesized. This analogue, $d(\underline{CT})_8$, contains non-ionic methylphosphonate internucleoside linkages. The pH-dependent conformational transitions of $d(\underline{CT})_8$ have been studied and its ability to form duplexes and triplexes with the normal homopurine oligonucleotide $d(AG)_8$ has also been investigated as a function of pH. Circular dichroism spectroscopy and ethidium bromide fluorescence enhancement have been used to monitor pH-dependent conformational transitions driven by the protonation of cytosine residues, and the difference behavior of $d(\underline{CT})_8$ and $d(CT)_8$ has been compared. It was possible to form self-associated complexes by using either $d(CT)_8$ or $d(\underline{CT})_8$ and both compounds combined with $d(AG)_8$ to form duplex or triplex DNA. At neutral pH, the CD spectrum of $d(AG)_8 \cdot d(\underline{CT})_8$ duplex was quite different from the CD spectrum of $d(AG)_8 \cdot d(CT)_8$ duplex, reflecting most likely a difference in conformation. The duplex to triplex transition characteristic of this DNA sequence occurred at a lower pH when $d(\underline{CT})_8$ was substituted for $d(CT)_8$; however, at pH 4.2, triplex containing $d(\underline{CT})_8$ was similar in conformation to triplex containing $d(CT)_8$. Several of these observations can be related to the alterations in electrostatic and steric interactions that occur when the negatively charged phosphodiester backbone of $d(CT)_8$ is replaced with a non-ionic methylphosphonate backbone.

c. *"Triplex Formation of Oligonucleotides Containing 2'-O-Methylpseudoisocytidine in Substitution for 2'-Deoxycytidine."*

- (1) Recently, studies of sequence-specific triplex formation of short synthetic oligonucleotides (or their analogues) have been of prime interest. In this report, a synthesis of new oligonucleotide analogue containing 2-amino-5-(2-O-methyl-B-D-ribofuranosyl)-4(1H)-pyrimidinone (2'-O-methylpseudoisocytidine or 1) in substitution for 2'-deoxycytidine and the triplex formation of this oligonucleotide analogue with a target duplex at neutral pH are reported.
- (2) A triplex can be formed as a triad consisting of a homopyrimidine strand, a homopurine strand, and a homopyrimidine strand. The third strand (i.e., the second pyrimidine strand) is located in the major groove of a duplex with Watson-Crick base pairing. Thymine or cytosine in the third strand form Hoogsteen type hydrogen bondings with adenine or guanine in the homopurine strand, respectively. Since protonation of cytosine bases in acidic conditions is essential in order to provide the second hydrogen bonding between the protonated cytosine and guanine in the Hoogsteen pair of the triad (Figure #10, upper left panel), this C-G-C⁺ triad is unstable in physiological conditions. This requirement limits the formation of triplex in living cells since the cellular pH is usually above pH 7.0.

Figure #10

- (3) To overcome this limitation, we designed and synthesized an oligonucleotide containing 1 which may form Hoogsteen type base pairings through hydrogen bondings with guanine in neutral and basic conditions (Figure #10, upper right panel). As indicated in the figure, this nucleoside 1 already contains one hydrogen at the N-3 position for hydrogen bonding with the guanine in the Hoogsteen pair of the triad. Also, methoxy substitution at 2'-position of pyrimidine nucleosides in a third strand stabilized the triplex formation.

- (4) A scheme for the synthesis of 1^{21} and its amidite synthon 5 is shown in the lower panel of Figure #10. An octamer 5'-(TT1TT1TT)3' (a) was synthesized on a DNA synthesizer. After deblocking and purification the oligomer a of this preparation showed one sharp peak by HPLC analysis. After hydrolysis of the oligomers by snake venom phosphodiesterase and alkaline phosphatase to nucleosides, the nucleoside composition for each oligomer was confirmed by HPLC. We studied a triplex formation of the oligomer a with an undecameric target duplex d u p l e x 5'-d(AAGAAGAAGAA)3'-5'-d-(TTCTTCTTCTT)3' (d). Both 5'-d(TTCTTCTT)3' (b) and 5'-(TTCmTTCmTT)3' (c), Cm = 2'-O-methylcytidine²²) were used as controls for the third strand.

Figure #11

An octamer a or b or c was mixed with the duplex d in a buffer, and the thermally induced transitions of the helices in each mixture was studied by measurement of UV absorption at 260nm at pH 7.0 (Figure #11 upper panel). Both the duplex d alone and the mixtures consisting of b and d or of c and d showed only one transition ($T_m = 42^\circ\text{C}$, Figure #11, lower panel) which was attributable to the melting of the duplex d itself. On the other hand, the mixture of a and d showed two transitions (Figure #12, lower panel). One transition ($T_m = 42^\circ\text{C}$) was identical with the melting of d and the second transition ($T_m = 12^\circ\text{C}$) was surmised to be the dissociation of the third strand a from the duplex d as illustrated in the upper panel of Figure #12. This mixture of a and d showed similar transition profiles between pH 7.0 and pH 8.7 (data not shown).

Also, we studied the circular dichroism (CD) spectra of mixtures of oligomers. Spectra of a mixture of a and the duplex d at both 3°C (dotted line) and 30°C (solid line) in neutral condition are shown in Figure #12A. As a control, the sum of a spectrum of a and a spectrum of d in the same buffer at 3°C is shown in the same figure (broken line). The control spectrum and the spectrum of the mixture at 30°C showed similar patterns. In contrast, the spectrum of the mixture at 3°C showed a different pattern, especially in the shorter wavelength region. A positive band at 215nm in the control spectrum changed to a negative band in the spectrum of the mixture at 3°C . This

Figure #12

difference indicated the association of the third strand with the duplex. Furthermore, a CD mixing study showed that the association has a 1:1 stoichiometry. These results indicated the triplex formation of a with d. This results in reported in Appendix VII.

d. "Triplex formation of the oligonucleotide containing 2'-O-methylpseudoisocytidine with a DNA duplex at neutral pH."

- (1) The synthesis of the hexadecanucleotide 5'TTTT1TTTT11111T3' (1-16 mer) containing 2-amino-5-(2'-O-methyl-B-D-ribofuranosyl)-4(1H)-pyrimidinone (2'-O-methylpseudoisocytidine or 1) is described. Triplex formation of 1 - 16 mer with a deoxyribonucleotide duplex: 5'd(ACCAAAGAAAAGGGGGGGACCA)3' - 5'd(TGGTCCCCCCTTTTCTTTGGT)3'

(duplex-22) which contains the "polypurine tract" found in the genome of human T-cell leukemia (lymphotropic) virus (HTLV-III) was studied by thermal denaturation and circular dichroism (CD) spectra in aqueous solution at neutral pH. The "polypurine tract" contains a homoguanine cluster consisting of six deoxyguanine residues. The results indicate that 1-16 mer and duplex-22 formed at triplex at neutral condition (0.01M Na cacodylate, 0.5M NaCl, 5mM MgCl₂, pH 7.2). In contrast, a hexadecadeoxynucleotide, 5'TTTTMTTTTMMMMMT3' (M-16 mer), containing 5-methyl-2'-deoxycytidine (M) did not form a stable triplex with duplex-22 at the same condition (Tm < 0° C). The CD mixing titration indicated that the triplex was formed with 1:1 (duplex:third strand) molecular stoichiometry.

e. "Formation of a purine•purine•pyrimidine triplex with purine oligomers having non-ionic methylphosphonate linkages."

- (1) Oligonucleoside methylphosphonates [(RO)₂POCH₃] are non-ionic backbone analogues of the naturally occurring negatively charged nucleic acids. The study of these analogues is of interest as they may be able to form stable, sequence specific complexes with charged nucleic acid target sequences. The absence of charge should make these interactions favorable; however, questions concerning the chirality of the phosphate methyl and its possible unfavorable steric effects remain.

Using the oligomers $d(AG)_8$ and $d(CT)_8$ as a model system, we have investigated the complexes which form at moderate pH and salt conditions (pH 8.0, 0.1M NaCl) when the homopurine strand is replaced with a fully modified methylphosphonate strand, $d(\underline{AG})_8$. Using several physical methods, we find the non-ionic analogue is more stable than the parent phosphodiester oligomer in regard to complex formation. Mixing studies show that $d(\underline{AG})_8$ forms both 1:1 and 2:1 (pur:pyr) complexes with $d(CT)_8$, whereas this is the case for the parent phosphodiester system. NMR studies at 300 MHz of the 2:1 $d(\underline{AG})_8:d(CT)_8$ system shows the presence of an additional hydrogen-bonded resonance envelope in the imido proton region which is not observed in the 1:1 mixture. Also, the Watson-Crick G•C and A•T NH-N resonances show different chemical shifts upon triplex formation. The thermal stabilities of the new triplex at both optical ($10^{-6}M$, $T_m = 51/36^\circ C$) and NMR ($10^{-3}M$, $T_m \approx 60^\circ C$) concentrations indicates that this model system complex melts directly to single strands without going through a duplex intermediate. These physico-chemical characterizations show for the first time the formation of a purine•purine•pyrimidine type triplex in a model system having purine strands of mixed adenine guanine residues. Also, the ability to form this triplex appears to be possible by reduction of charge-charge repulsive interactions through the use of non-ionic methylphosphonate purine oligomers.

f. *"DNA Triplex Formation of Oligonucleotide Analogues Consisting of Linker Groups and Octamer Segments which have opposite Sugar-phosphate Backbone Polarities."*

- (1) The DNA oligomer analogues $3'(\text{CTTTCTTT})5'-(\text{TTCTTCTT})3'$ (IV), $5'(\text{TTTCTTTC})3'-P2-3'(\text{CTTTCTTT})5'$ (V), and $5'd(\text{TTTCTTTC})3'-P2-3'(\text{CTTTCTTT})5'-P4-5'(\text{TTCTTCTT})3'$ (VI) ($P2 = P^*P$, $P4 = P^*P^*P^*P^*$, when $P = \text{phosphate}$, $^* = 1,3\text{-propandiol}$) have been synthesized. These oligomers consist of a linker group or groups and homopyrimidine oligonucleotides which have opposite sugar-phosphate backbone polarities. These oligomer analogues are designed to form triplexes with a duplex, $5'd(\text{AAAGAAAGCCCTTTCTTTAAGAAGAA})3':5'D(\text{TTCTTCTTAAAGAAAGGGCTTTCTTT})3'$ (I), which contains small homopurine clusters alternately located in both strands. The length of the linker groups, P2 and P4, was based upon a computer modeling analysis. Triplex formation by the unlinked octamers $5'd(\text{TTCTTCTT})3'$ (II) and $5'd(\text{TTTCTTTC})3'$ (III) and the linked oligomers analogues, IV, V, and VI with the target duplex was studied by thermal denaturation at pH 5.2. The order of stabilities of triplex formation by these oligomers were $I-V \gg I-IV > I-(II, III)$. The mixture of I and VI showed two transitions corresponding to the dissociation of the third strand. The higher transition corresponded to the dissociation of 3'-3' linked octamer segments and the lower one corresponded to the dissociation of 5'-5' linked octamer segments. The T_m of the latter transition was higher than that of the I-IV triplex, thus the triplex formed by the 5'-5' linked octamer segment was stabilized by the triplex formed by the 3'-3' linked octamer segments in I-VI triplex. Triplex formation of this system was also studied in the presence of ethidium bromide. Only the I-IV triplex and the part of triplex formed by the 5'-5' linked

oligomers segments in the I-IV triplex were stabilized in the presence of ethidium bromide, whereas the I-V triplex or the I-(II, III) triplex were not. In the presence of ethidium bromide, the three octamer segments in VI dissociated from the duplex I in a synergistic manner. Negative bands at 230 to 210nm were detected in the CD spectra of I-(II, III), and I-IV, I-V or I-VI at pH 5.2, indicating triplex formation. Addition of ethidium bromide to the mixtures of oligomers changed CD spectra only slightly. These results are useful when considering the design of oligonucleotide analogues which can bind as third strands to DNA duplex of higher complexity.

References

1. Callahan, D.E., Karim, A., Zheng, G., Ts'o, P.O.P., Lesko, S.A. "*Quantitation and Mapping of Integrated Human Papillomavirus on Human Metaphase Chromosomes using a Fluorescence Microscope Imaging System.*" (In Press).
2. N. Ichikawa, M. Yi, P.O.P. Ts'o. "*A Simple and Effective Procedure for the Construction of Linear Restriction Fragment Maps of a 50-250kb DNA.*" Analytical Biochemistry, 1991 (Submitted).
3. Yi, M., Au, L.C., Ichikawa, N., Ts'o, P.O.P.: *Enhanced Resolution of DNA Restriction Fragments: A Procedure by Two-Dimensional Electrophoresis and Double-Labeling.* Proc. Natl. Acad. Sci., 87(10):3919-3923, 1990.
4. Callahan, D.E., *op. cit.*, (Submitted).
5. Ann. Rev. Biophys. Bioeng., 13:191, 1984.
6. Murray, *et al.*: Mol. Gen. Genet., 150:53-61, 1977.
7. Cross, S.H., Little, P.F.R.: Gene, 49:9-22, 1988.
8. LeBlond, G., Ts'o, P.O.P.: Anal. Biochem., 177:413-418, 1989.
9. Ichikawa, N., Yi, M., Paul P.O.P.: Analytical Biochemistry, 1991 (Submitted).
10. Kunkel, L.M., *et al.*, Proc. Natl. Acad. Sci., 74:1245-1249, 1977.
11. Miller, P.S. & Ts'o, P.O.P.: "*Oligonucleotide Inhibitors of Gene Expression in Living Cells: New Opportunities in Drug Design.*" Ann. Rep. Med. Chem., 23:295-304, 1988.
12. Miller, P.S.: "*Antisense nucleic acid analogues as potential antiviral agents,*" in Herpesviruses, the Immune System and AIDS. L. Aurelian (ed.), Kluwer Academic Publishers Boston, 1990, pp. 343-360.
13. Miller, P.S.: BioTechnology, 9:358-362, 1991.
14. Miller, P.S.: "*Antisense oligonucleoside methylphosphonates,*" in Antisense RNA and DNA. J.A.H. Murray (ed.), J. Wiley and Sons, New York (In press).
15. Miller, P.S., Bhan, P., Cushman, C.D., Kean, J.M. & Levis, J.T.: Nucleosides and Nucleotides, 10:37-46, 1991.

16. Lee, B.L., Murakami, A., Blake, K.R., Lin, S.-B. and Miller, P.S.: *Biochemistry*, 27:3197-3203, 1988.
17. Kean, J.M., Murakami, A., Blake, K.R., Cushman, C.D. and Miller, P.S.: *Biochemistry*, 27:9113-9121, 1988.
18. Lee, B.L., Blake, K.R. and Miller, P.S.: *Nucleic Acids Res.*, 16:10681-10697, 1988.
19. Bhan, P. and Miller, P.S.: *Bioconjugate Chemistry*, 1:82-88, 1990.
20. Chang, E.H., Miller, P.S., Cushman, C.D., Devadas, K., Pirollo, K.F., Ts'o, P.O.P., Yu, Z.P.: *Biochemistry*, 30, 1991 (In press).
21. Miller, P.S.: "*Psoralen-Derivatized Antisense Oligonucleoside Methylphosphonates*," in *Gene Regulation: Biology of Antisense RNA and DNA*. R.P. Erickson & J.G. Izant (eds.), Raven Press, New York, 1992 (In press).
22. Miller, *op. cit.*, *Nucleotides and Nucleotides*, 10:37-46, 1991.
23. B.S. Sproat, B. Beijer, A. Iribarren.: *Nucl. Acids Res.*, 18:41-49, 1990.
24. M.J. Robbins, S.R. Naik, A.S.K. Lee.: *J. Org. Chem.*, 39:1891, 1974.
25. Lee, *op. cit.*, *Biochemistry*, 27:3197-3203, 1988.

September 12, 1991

Appendix List *

- I. Lo-Chun Au, Paul O.P. Ts'o, Ming Yi: "Monitoring Mammalian Genome Rearrangement with Mid-Repetitive Sequences as Probes." *Analytical Biochemistry*, 90:326-330, 1990.
- II. L.S. Kan, D.E. Callahan, T.L. Trapane, P.S. Miller, P.O.P. Ts'o, D.H. Huang. "Proton NMR and Optical Spectroscopic Studies on the DNA Triplex Formed by d-A-(A-G)₇G and d-C-(T-C)₇T." *J. Biomolecular Structure & Dynamics*, 8(5):911-933, 1991.
- III. D.E. Callahan, T.L. Trapane, P.S. Miller, P.O.P. Ts'o and L.S. Kan. "Comparative Circular Dichroism and Fluorescence Studies of Oligodeoxyribonucleotide and Oligodeoxyribonucleoside Methylphosphonate Pyrimidine Strands in Duplex and Triplex Formation." *Biochemistry*, 30:1650-1655, 1991.
- IV. N. Ichikawa, M. Yi, P.O.P. Ts'o. "A Simple and Effective Procedure for the Construction of Linear Restriction Fragment Maps of a 50-250kb DNA." *Analytical Biochemistry*, 1991 (Submitted).
- V. T. Trapane, L.S. Kan, M. Reynolds, R. Hogrefe, P.O.P. Ts'o. "Formation of a purine•purine•pyrimidine triplex with purine oligomers having non-ionic methylphosphonate linkages." *Seventh Conversation in Biomolecular Stereodynamics*, 8(6):a229, 1991 (Abstract).
- VI. A. Ono, C.N. Chen, L.S. Kan. "DNA Triplex Formation of Oligonucleotide Analogues Consisting of Linker Groups and Octamer Segments which have opposite Sugar-phosphate Backbone Polarities." *Biochemistry*, 1991 (Accepted).
- VII. A. Ono, P.O.P. Ts'o, L.S. Kan. "Triplex Formation of Oligonucleotides Containing 2'-O-Methylpseudoisocytidine in Substitution for 2'-Deoxycytidine." *J. Am. Chem. Soc.*, 113:4032-4033, 1991.
- VIII. Callahan, D.E., Karim, A., Zheng, G., Ts'o, P.O.P., Lesko, S.A. "Quantitation and Mapping of Integrated Human Papillomavirus on Human Metaphase Chromosomes using a Fluorescence Microscope Imaging System." (In Press).
- IX. Miller, P.S., Bhan, P., Cushman, C.D., Kean, J.M. & Levis, J.T. "Antisense Oligonucleoside Methylphosphonates and Their Derivatives." *Nucleosides and Nucleotides*, 10:37-46, 1991.
- X. Miller, P.S. "Oligonucleoside Methylphosphonates as Antisense Reagents." *BioTechnology*, 9:358-362, 1991.
- XI. Leonard, T.E., Bhan, P., Miller, P.S. "Preparation of Protected 2'-O-Methylguanosine." *Tetrahedron Letters*, 1991 (Submitted).
- XII. Miller, P.S. "Psoralen-Derivatized Antisense Oligonucleoside Methylphosphonates," in *Gene Regulation: Biology of Antisense RNA and DNA*. R.P. Erickson & J.G. Izant (eds.), Raven Press, New York, 1992 (In press).
- XIII. Ichikawa, N., Yi, M., Ts'o, Paul P.O.P.: "A Procedure for the Construction of Linear Restriction Fragment Maps of A 50-100kb DNA." *Analytical Biochemistry*, 1991 (Submitted).

* all reprints and preprints removed as & enclosed separately—

END

**DATE
FILMED**

11 106191

1