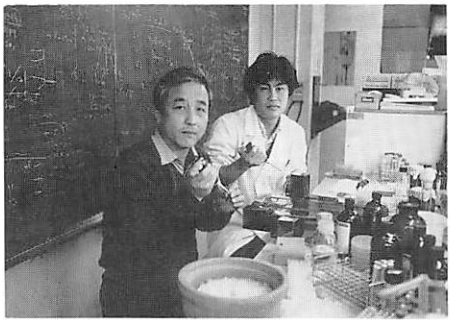
# SCIENTIFIC REPORTS

Unpublished data should not be cited without permission of the authors.

*Biophysics*






# DEPARTMENT OF BIOPHYSICS

## GENERAL SUMMARY

Dr. Hiroshi Yoshikawa, ex-professor of this department moved to Osaka University Medical School as professor of the Department of Genetics on January 1986. On April 1986 Dr. Fumio Harada, section head of the Virology Division, National Cancer Center Research Institute at that time, was installed as professor of the department. With this alternation, the major theme of this department changed from studies on the chromosomal replication of microorganisms to studies on the molecular mechanisms of oncogenesis by tumor viruses.

Recently several kinds of viruses have been watched as candidates for the cause of human cancers. These include hepatitis B virus (HBV) as a cause of liver cancer, HTLV-I as the cause of adult T-cell leukemia, human papilloma viruses as a cause of cervical cancer, and Epstein-Barr virus as a cause of Burkitt's lymphoma and nasopharyngeal carcinoma. However the molecular mechanisms of oncogenesis by these viruses have not been clarified yet.

Since 1983, we have continued molecular biological studies of HBV and woodchuck hepatitis virus, the most oncogenic hepadnavirus, in collaboration with a group in the First Department of Internal Medicine, School of Medicine, Kanazawa University. Patterns of viral DNA integrations in genomes of hepatocellular carcinomas (HCC), transcriptional activation of a host gene by viral DNA integration, enhancer activity of an integration site of viral DNA, correlation between expression of viral X gene product and HCC, and others have been investigated in the past three years.

The chromosomal DNA of vertebrates harbors many vertically transmitting endogenous retrovirus genomes. To clarify causal relationships of these endogenous retrovirus genomes to oncogenesis in human beings, a new research project was begun in 1986. Using a novel cloning technique, we obtained three kinds of endogenous retroviral sequences from a human genomic library. These elements contained sequences characteristic of the long terminal repeat structure of the retrovirus provirus.

Another new project is the search of human genes concerned with leukemogenesis by HTLV-I. HTLV-I does not contain oncogene in its genome. It is believed that virus encoded protein $p40^{x}$ activates transcription of specific host genes. To isolate these genes, we are searching genomic clones containing a homologous sequence with viral transcriptional enhancer from a human genomic library.

Further details of the studies described are presented in the following summaries.

(These studies were partly supported by a Grant-in-Aid from the Ministry of Health and Welfare for a Comprehensive 10-Year Strategy for Cancer Control, Japan and by a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture of Japan.)

## (1) Isolation of human endogenous retrovirus genomes.

F. Harada, N. Tsukada, Y. Hirose and N. Kato[1)]

RNA derected DNA synthesis by reverse transcriptase of retroviral genomic RNA is initiated from a specific primer tRNA. The primer binding site (PBS), which is complementary to the 3′-terminal 18 nucleotide sequence of primer tRNA, exists downstream of the 5′-long terminal repeat (LTR) of retrovirus provirus and all mammalian type C retroviruses whose sequences have been reported to possess a PBS corresponding to that of $tRNA^{Pro}$. Recently, several endogenous retrovirus genomes have been isolated from the human genome and their LTR sequences have been determined. These sequences were isolated by low stringency hybridization of various known retroviral sequences or were accidentally found in flanking regions of other genes. They have PBS sequences corresponding to $tRNA^{Glu}$, $tRNA^{Arg}$, $tRNA^{His}$, $tRNA^{Ile}$ or $tRNA^{Lys}$ but no sequence corresponding to $tRNA^{Pro}$ has yet been reported.

We used the 3′-fragment of $tRNA^{Pro}$ 3′-end labeled with [5′-$^{32}$p] pCp as a probe for cloning. This probe hybridized to not only the PBS of endogenous retrovirus genomes, but also with the genomes of $tRNA^{Pro}$ However, the genome of $tRNA^{Pro}$ does not have the 3′-CCA sequence of the mature tRNA molecule, whereas PBS contains this sequence. Therefore, if the hybrid is treated with RNase A, the radioactivity of the probe hybridized to the $tRNA^{Pro}$ genome should be eliminated, and only the endogenous retroviral sequence should be selected. By this method we isolated three kinds of human endogenous retrovirus genomes (HuERS-P1, P2 and P3) from a HeLa cell genomic library.

These elements contained sequences characteristic of the LTR structure of the retrovirus provirus as follows. The PBS sequences of these clones shared 17 of 18 nucleotides with the complementary sequence of the 3′-terminus of $tRNA^{Pro}$. These elements were flanked by cellular direct repeats of 4 base pairs, and each LTR started with the dinucleotide TG and ended with CA. Polypurine tracts were observed on the immediate 5′-side of the 3′-LTRs. LTR sequences contained a CAT box, a TATA box and a polyadenylation signal.

The LTR sequences of these three elements do not resemble each other or those of any retroviruses or endogenous retroviruses whose sequences have been reported. From the Southern hybridization experiment there may be 10 to 20 copies of HuERS-Pl, 20 to 30 copies of HuERS-P2 and 30 to 40 copies of HuERS-P3 in the human genome. These elements exist not only in the human genome but also in the simian genome. However, mouse DNA does not contain these elements. Probably germ cells of the common ancestor of humans and monkeys were infected by retroviruses, and these sequences still exist in the genomes of both species.

1) Virology Division, National Cancer Center Research Institute.

## (2) The search of genes concerned with leukemogenesis by HTLV-I.

F. Harada, N. Tsukada and Y. Hirose

Human T-cell leukemia virus I (HTLV-I), the causative virus of adult T-cell leukemia (ATL), does not contain oncogene in its genome. HTLV-I possesses a characteristic gene called pX which encodes a protein $p40^X$. $p40^X$ functions as a trans-acting factor to increase viral transcription through the enhancer sequence in the LTR. Conceivably, $p40^X$ also activates transcription of specific cellular genes and causes ATL. Assuming that the transcriptions of these cellular genes are activated through the sequences homologous to the viral enhancer, the following experiments have been carried out.

In the viral enhancer, a 21 nucleotide-long sequence is repeated three times (Fig. 1), and these sequences are very important for enhancer activity. Very similar 21 nucleotide repeats are also observed in the LTR of HTLV-II, another human retrovirus which causes hairy cell leukemia (Fig. 1). We chemically synthesized this sequence and used it as a probe for cloning DNA regions containing this sequence from a HeLa cell genomic library. About $3X10^5$ phage plaques were screened, and 14 different clones were isolated. Seven well hybridized clones were selected, and the nucleotide sequences of the region hybridized to the probe and the flanking sequences were determined. Each clone contained a sequence similar to the probe (Fig. 1), but the repeat of this sequence was not observed in any clones. In clones 5 and 21, this sequence existed in open reading frames consisting of more than 130 amino acids.

Specific expression of these sequences in the cells infected by HTLV-I or HTLV-II is under investigation.

```
HTLV-I     AAGGCTCTGACGTCTCCCCCC
           T....C.......G......T
           C....GT.....A.AA....T

HTLV-II    ....................T  (probe)
           .C...C........C.T....
           ............AT.A.....

Clone 2    G..............TTTTA
Clone 5    TG.A....TC..........T
Clone 9    TG......A.....A.....T
Clone 13   TC..............TA.T
Clone 21   CCT..............T.T
Clone 23   .GC..............TAT
Clone 24   ..A......C.........TC
```

Fig. 1. 21 Nucleotide repeats in the LTRs of HTLV-I and HTLV-II, and homologous sequences existing in the human genomic clones. Dots indicate the same nucleotides as in the first sequence of HTLV-I and letters indicate different nucleotides from those in the first sequence of HTLV-I.

## (3) Stable maintenance of integrated woodchuck hepatitis virus (WHV) DNA during establishment of a cell line of woodchuck hepatoma.

**S. Kaneko[1], K. Kodama, H. Yoshikawa, K. Kobayashi[1], N. Hattori[1] and S. Murakami.**

The fate of integrated WHV DNA was systematically analyzed in DNA samples from primary hepatoma of woodchuck, solid tumors transplanted in athymic mice derived from a woodchuck primary hepatoma, and an established cell line of tissue culture, WH257GE10, originating from the transplanted tumor. In four out of five woodchuck hepatomas, WHV DNA integration was demonstrated in addition to various amount of extrachromosomal replicating intermediate WHV DNA. The integration patterns of the primary hepatoma do not indicate existence of a common integration site on the host DNA. The integration pattern in WH257GE10 is identical to that in the transplanted tumor and similar but slightly different from that of the primary hepatoma. No extrachromosomal or replicative intermediates of WHV DNA were detected in the transplanted tumors or in WH257GE 10.

Three integration sites on the chromosomes of WH257GE10 were deduced, and all of three integrated WHV DNA are a part of the genome, subgenomic integration. A small portion corresponding to the cohesive region of the genome was not detected in all of these integrated WHV DNA. A positive role of WHV DNA integration on generation of hepatoma in woodchuck is suggested by the high incidence of WHV DNA integration in primary hepatoma and the stable maintenance of a certain mode of WHV DNA integration in the hepatoma-derived cell populations during passages of transplantation or serial growth of tissue culture.

1) The 1st. Int. Med., Med. School Kanazawa Univ.

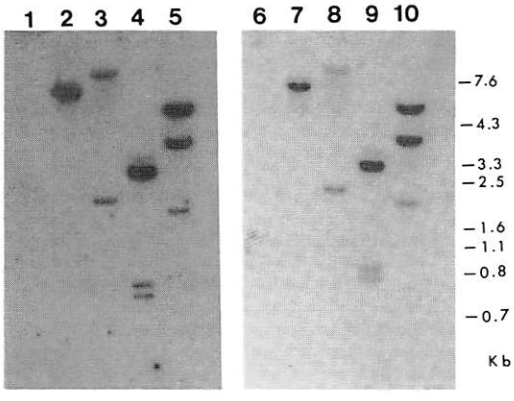



Fig. WHV DNA in the transplanted cell line and WH257GE10. DNA samples of the transplanted cell line in the 8th passage (*lanes 1* to *5*) and WH257GE10 in the 100th passage (*lanes 6* to *10*) were applied onto 1.0% horizontal agarose gels without digestion (*lanes 1* and *6*) or after digestion with *Kpn*I (*lanes 2* and *7*), with *pvu*II (*lanes 3* and *8*), with *Eco*RI (*lanes 4* and *9*), or with *Hind*III (*lanes 5* and *10*). Molecular markers are shown on the right. *Kb*, kilobase.

## (4) WHV DNA integration in multicentric primary hepatocellular carcinomas of woodchucks.

A. Shimoda[1], K. Kuroki, M. Uchijima, S. Kaneko[1], M. Unoura[1], K. Kobayashi[1], N. Hattori[1] and S. Murakami.

Woodchuck hepatoma related to WHV infection is manifested by its multicentric property in backgrounds of chronic hepatitis without cirrhosis. A high frequency of integration was observed in woodchuck primary hepatoma.

We have extended our study to the analysis of pattern of WHV DNA integration pattern in each tumor of multi-tumor bearing woodchucks. Results are 1) High frequency of WHV DNA integration was again confirmed, namely 14 out of 24 tumor samples or 5 out of 6 woodchuck liver samples. 2) The same integration pattern among DNA samples of different tumors in a certain individual is rare, especially among large tumor samples. However, there could be detected bands having the same size in *EcoRI* or *Hind III* digests among hepatoma samples in the same individual, suggesting that they have been derived from the same origin. 3) Among 10 minute tumors in a woodchuck liver, 4 samples seem to have the same pattern of WHV DNA integration. The pattern is similar to that of one of two major tumors but different from that of another one in the same individual.

Multicentric properties of hepatoma in woodchuck can be interpreted as the result of multi-origin tumorigenesis, or alternatively, as the result of multi-step progressions of disseminated single-origin tumor. Our results may support the latter hypothesis.

1) The 1st. Int. Med. School Kanazawa Univ.

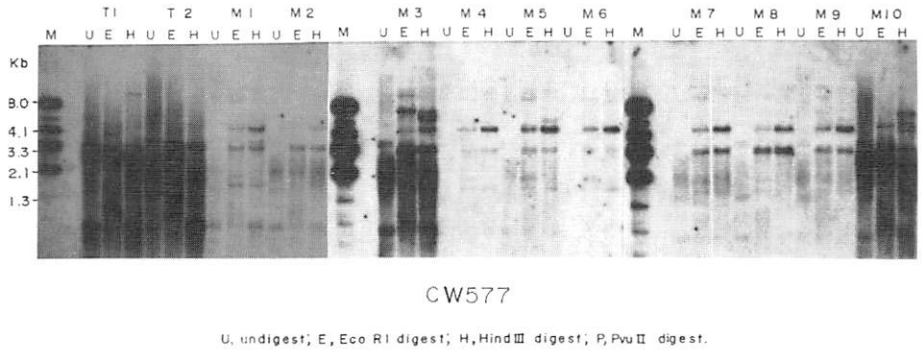



Figure. Integration patterns of WHV DNA in multicentric tumors in the same liver.

Tumor samples were collected from a WHV-infected woodchuck CW577. Tl and T2 are major tumors, and M1 to M10 are minute disseminated tumors. U, E, H, P are undigested, digested with *EcoRI, HindIII, PvuII* respectively. M is mixture of molecular markers. Whole genome of WHV2 is used as the probe of Southern hybridization. Smear bands below 3.3 kb are extrachromosomal replicative intermediates of WHV, and the 3.3 kb bands in E and H, are the linear form of the unit-length WHV. *PvuII* has no recognition site on WHV genome.

**(5) A novel family of short interspersed repetitive sequence found in the vicinity of an integration site of WHV DNA in a hepatoma-derived cell line of woodchuck.**

**S. Aoyama[1], S. Kaneko[1], K. Kuroki, K. Kodama, A. Shimoda[1], and S. Murakami.**

In a woodchuck hepatoma, WH257GE10, three independent integrations of WHV subgenome have been stably maintained. During the course of analysis of one cloned region, region A, repetitive sequences were found close to the host-viral junction points. To understand the process of the WHV DNA integration and control of transcription of the region, the character of the sequence (s) was analyzed. A 16 times AC repeat was found in the upstream of the WHV DNA, belonging to one of the purine-pyrimidine alternating repeats which has enhancer activity per se as reported by Hamada et al. (Hamada, H, et al., Mol. Cell Biol., *4*, 2622-2630 (1984)). Fragments having sequence homology to the repetitive sequence located downstream of integrated WHV DNA were cloned from a woodchuck genomic library of a normal liver tissue. Three of them were sequenced and compared to that in region A. As a result, a family of repetitive sequences was characterized.

The size of the consensus sequence of the repetitive sequence is about 0.45 kb. Since each cloned region (ranging 8 kb to 15 kb) has only one copy of the sequence and has no homology to the other fragments in region A, the repetitive sequence might not be a tandem repetitive, and not a part of a family of long interspersed repetitive sequence (LINE). Thus the repetitive sequence is considered to be a family of short interspersed repetitive sequence (SINE). The SINE tentatively designated the 0.45 kb repetitive sequence. The copy number of the repetitive sequence is around $10^4$ per haploid genome of woodchuck deduced from the cloning frequency. The SINE has no strong homology to the SINEs or LINEs of rodents or primates reported previously.

There are homologous sequences in other mammalian cells, even in nematodes, but not in baker's yeasts. Thus the 0.45 kb SINE may have homologous counterparts in multicellular genomes. Diversity of DNA sequences among the family is about 20%, and there is a non-conserved region in the middle of the repetitive sequence. The 26 times AT repeat in region A is located at the region and unique among the sequences of the family isolated so far, although a short stretch of AT alternatives is found in the region of the consensus sequence.

Preliminary results of CAT assays suggest that not only the 0.75 kb *EcoRI-HindIII* fragment in region A, but also fragments having other sequences of the SINE family exhibit enhancer activity.

1) The 1st. Int. Med., Med. School Kanazawa Univ.

## (6) Transcriptional activity of a chromosome region where a WHV subgenome integrated in an established cell line of woodchuck hepatoma.

K. Kuroki, S. Kaneko[1], K. Kodama, and S. Murakami.

In three integration sites of a cell line of woodchuck hepatoma, WH-257GE10, region A contains only one fourth of a WHV genome encompassing the section from the end of S to the middle of X of the genome. The nucleotide sequence of an integrated WHV subgenome, WHV3, and its flanking host DNA revealed that the integrated WHV DNA has an enhancer region of hepadnaviruses and that WHV3 is close to characteristic host DNA sequences including a 28 poly T just upstream to WHV3, two independent purine-pyrimidine alternating repeats located in its upstream and in its downstream.

The possibility of insertion activation by the integrated WHV DNA in this region was tested by measuring transcriptional activity of the region. Results indicate 1) WHV DNA sequence in the region has a 33 bp long consensus sequence tightly conserved among mammalian hepadnaviruses. DNA fragments of WHV3 covering the consensus sequence exhibit enhancer activity detected by CAT assay. The enhancer activity of WHV3 is comparable to those of WHV and HBV enhancer regions cloned from virions. In addition to that, both flanking host DNA sequences also exhibit enhancer activity. 2) A 2.5 kb poly $(A)^+$ mRNA transcribed from a region about 2 kb away from WHV3 was detected in the cell line, and its polarity is away from WHV3 detected by RNA probes. The amount of the transcript in WH257GE10 is at least 10 fold higher than that in the normal liver mRNA. Therefore transctiption of the region is actually activated in WH257GE10. 3) Transcripts from the WHV3 sequence could not be detected in WH257GE10.

These results tend to suggest that integration of the WHV enhancer element activates transcription of a certain gene located nearby in the hepatoma-derived cell line.

1) The 1st. Int. Med., Med. School Kanazawa Univ.

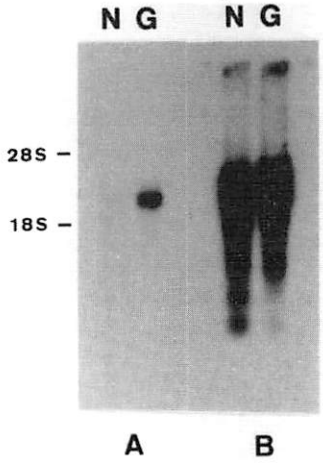



Figure. An activated transcript from a region near to the integration site in WH257GE10 cells.

Northern hybridization was carried out by using a RNA probe of fragment V which is 1.6 kb apart from the integrated WHV3 in region A (A) or a RNA probe of fragment IV having the 0.45 kb repetitive sequence which is transcribable and used as a positive control (B). The polarity of RNA detected by the probes is away from WHV3. Poly $(A)^+$ RNA fraction of WH257GE10 (lane G) or of an uninfected normal liver (lane N) were used.

## (7) Expression of the product of X ORF of woodchuck hepatitis virus (WHV) genome in *Eschericha coli.*

**K. Kodama, H. Yoshikawa, and S. Murakami.**

The fourth long open reading frame in hepadnavirus genomes, X gene, is well conserved among mammalian hepadnaviruses although its biological function in relation to virus production or oncogenic properties of hepadnaviruses remains to be elucidated. We have constructed recombinant plasmids which express a full-size product of X ORF of WHV2 clone. A 0.7 kb *NcoI-HindIII* fragment which starts exactly at the first initiation codon of X ORF of WHV2 and ends in the middle of C ORF was prepared and inserted between SD sequence downstream of *Tac* promoter and the *Rrn* terminator of an expression vector, pkk223-3. By pulse labeling with $^{35}$S-methionine in the maxicell system, two bands of 17k and 11k were detected as the result of insertion. The former band was deduced to be the X product since the band was eliminated by frame-shifting in X-ORF. A plasmid, pKX10, with 10bp distance SD and the initiation codon of X ORF can produce the band but a plasmid, pKX6, with 6 bp distance can not produce it at a detectable level. The 11k band is deduced to be a fused protein having the N-terminal half of C ORF translated from a polycistronic mRNA transcribed from *Tac* promotor, since deletion of *Tac* promotor eliminated both bands and frame-shifting in C ORF eliminate the 11k band, although frame-shifting in X ORF does not affect the amount of the 11k band.

Detection of anti-X antibody was carried out by immunoprecipitation using $^{35}$S-labeled maxicell lysate. Anti-X antibody could be detected in five out of ten serum samples which were sAg or sAb positive. Among them, four out of five sera of hepatoma-bearing woodchucks are anti-X positive. We started to establish an *in vitro* system to prepare X product sufficient for biochemical analysis. X-RNA transcribed with SP6 RNA polymerase *in vitro*, was translated with avian reticulocyte lysates. A 17k product was efficiently synthesized in the system. Biochemical analysis of the product is on going.

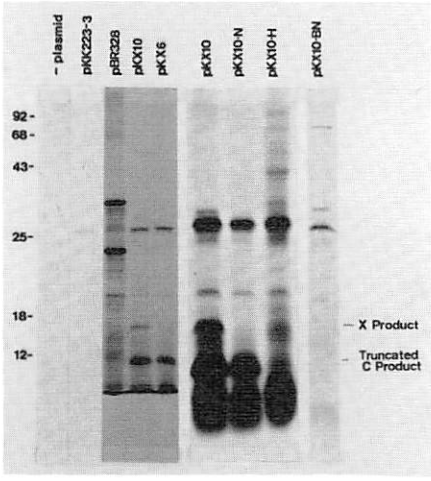



Figure. Expression of X protein in maxicells detected by SDS PAGE.

Two bands, 17k and 11k, were detected in pKX110. The 17k band was deduced as X protein because of its absence in pKX110-N having a frame-shifting in X ORF. The 11k band is a fused protein with N-terminal of C ORF since pKX10-H with a frame-shifting in C ORF could not produce it. Since pKX110-BN lacking *Tac* promoter could not synthesize 17k and 11k, both products seem to be translated from the same polycistronic mRNA.

## (8) Expression of PreC-C and C of human hepatitis B virus (HBV) and woodchuck hepatitis virus (WHV) in HeLa cells.

A. Shimoda[1], K. Kuroki, M. Unoura[1], K. Kobayashi[1], N. Hattori[1] and S. Murakami.

The preC-C region of hepadnaviruses codes two polypeptides, preC-C and C, having different roles on viral infection cycles. We have decided to establish mammalian cell clones stably expressing each of them. A DNA fragment having the PreC-C or C region of cloned WHV2 or HBVadr4 was inserted into the polylinker site downstream to MMTV LTR of pMSG plasmid which has *Ecogpt* gene as the selective marker. Constructed plasmid DNA was introduced into HeLa cells by transfection with calcium phosphate co-precipitation. Numbers of *Ecogpt* transformed colonies were in the same range per microgram donor DNA of each of 4 different constructs and the parental pMSG, suggesting that expression of PreC-C or C is not harmful for HeLa cells. Tandem integration of donor DNA shown by Southern blotting in all clones tested so far, including 6 WHV PreC-C, 5 WHV C, 3HBV PreC-C and 5 HBV C clones.

Expression of PreC-C or C region detected by Northern blotting was varied considerably among these clones. The level of the amount of the mRNA increased by 10 folds by addition of dexamethasone. The existence and localization of core-related proteins was examined immunohistochemically by using human anti-HBc IgG as the first antibody, biotinylated anti-human IgG as the second, and the avidin conjugated alkaline phosphatase H as the third. Cells having HBV PreC-C region were stained positive in the cytoplasmic region, whereas the nuclear region was stained in cells harboring HBV C region. The same tendencies were observed in cells with WHV DNA fragment, although extents of staining are rather weak.

1) : The 1st. Int. Med., Med. School Kanazawa Univ.

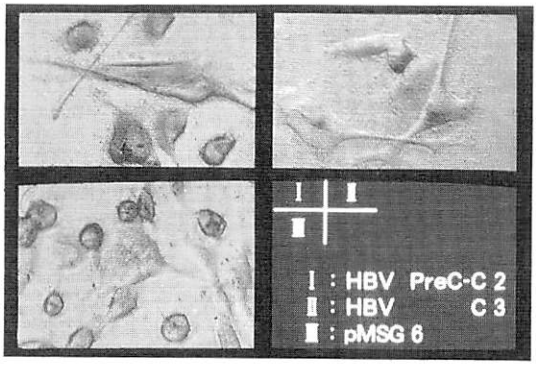



Figure. Intracellular localization of C related proteins in HeLa cells having HBV PreC-C expression plasmid.

Fixed cells were incubated with a 100 folds diluted human anti-HBc antibody, then incubated with biotinylated anti-human IgG sera of goats. Finally cells were treated with avidin-conjugated alkaline phosphatase H and stained. Original color of stained regions is light red.

## (9) Binding of nuclear factors to the WHV enhancer.

S. Murakami, S. Kaneko[1], K. Kodama, F. Harada and K. Kuroki.

A DNA region having enhancer activity has been demonstrated in the HBV genome, and there is a 33 bp long conserved sequence located in the region of mammalian hepadnavirus genomes. To obtain further information about hepadnavirus enhancer, we have decided to search nuclear factor (s) which can interact with the region *in vitro* detected by a gel retardation method and DNase I foot printing. The results are as follows. 1) Nuclear factor which could bind a 93 bp long *HpaII-HaeIII* DNA fragment including the consensus sequence was detected by gel retardation assay in nuclear extracts of HeLa or WH257GE10 cells. Retarded bands could be observed in the presence of excess M13 fragments or pBR322 fragments. Such retarded bands of the WHV enhancer element can not be formed in the presence of excess DNA of other hepadnavirus enhancer regions or SV40 enhancer region. 2) All bands in the 33bp region were protected and clusterings of hyper-sensitive bands surrounding the region were observed. 3) The binding activity could increase by four to five fold in HeLa cells pretreated with a phorbor ester, TPA. In parallel to this, activity of WHV enhancer increases in TPA treated HeLa cells.

Specific binding of nuclear factors to WHV enhancer was demonstrated, and at least one of binding factors in HeLa cells is a TPA-inducible positive trans-acting factor. There must be conditions which induce hepadnavirus enhancer activity in hepatic cells.

1) : The 1st. Int. Med., Med. School Kanazawa Univ.

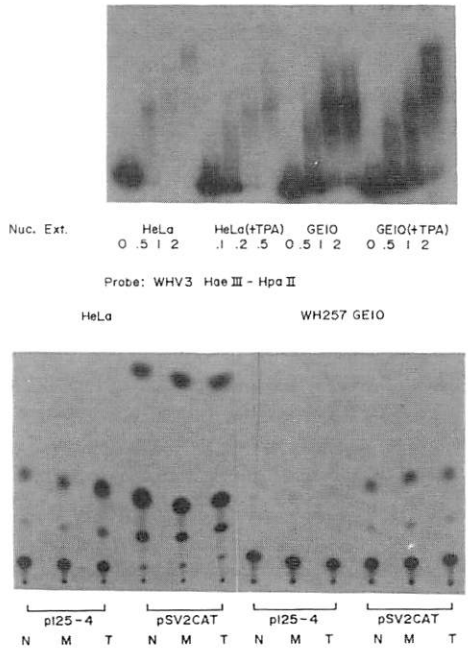



Figure. Activity of WHV enhancer region and binding activity of nuclear factors are TPA inducible in HeLa cells.

(Upper figure) Gel retardation assay. The $^{32}$p-labeled *HaeIII-HpaII* fragment of WHV3 enhancer region (lng) was incubated in the presence of various amount of nuclear extracts (in $\mu$1) prepared from HeLa or WH257GE10 cells with or without TPA pretreatment for 24hr. DNA bands were separated by electrophoresis in a 4% agarose gel.

(Lower Figure) CAT assay. Plasmid DNA of pSV2CAT or p125-4 having WHV3 enhancer region in the downstream of CAT gene of pSV1CAT was introduced to HeLa or WH257GE10 cells by electroporation. Crude extracts were prepared from cells with or without treatment of TPA or mitomycin C for one day, harvested at 48 hr after electroporation.