AN ABSTRACT OF THE THESIS OF

Xiaoshan Liang for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on May 6th, 1993.

Title: Studies of Rainbow Trout Ki-*ras* Gene: Sequencing, Aflatoxin B1 Binding, and Chromatin Structure

Abstract Approved_____

Dr. Kensal E. van Holde

Characterization of the 5' flanking region of rainbow trout ki-*ras* gene was begun with the cloning and sequencing of this region by the inverse PCR technique and dideoxynucleotide chain termination method. In total, a nucleotide sequence of 1080 bp upstream from the first coding ATG was sequenced. Although this region showed certain promoter elements, it does not share common features with other mammalian *ras* promoters, which lack the TATA and contain multiple GC boxes with Sp1 binding activities. In contrast, this region in trout *ras* contains typical TATA and CCAAT boxes. This structural difference of the trout ki-*ras* promoter from that of other mammalian *ras* genes may suggest that different transcriptional regulation mechanisms of the *ras* gene are used at various levels in evolution.

The chromatin structure of the trout ki-*ras* gene was studied by probing *in vivo* for DNase I hypersensitive sites. To overcome the difficulties of using the traditional indirect end labeling method for a single-copy gene, the technique of ligation-mediated PCR was applied. No hypersensitive sites were observed at or near the codon 12 region of the gene, either in normal (protooncogene) or tumor (oncogene) tissue from the liver. This result suggests that the local chromatin structure of trout ki-*ras* gene may not be an important factor for codon 12 mutations induced by genotoxins, and that changes of chromatin structure are unlikely to be promoted after tumor formation. Studies by micrococcal nuclease demonstrate that this *ras* gene, in the region around 12, lacks ordered nucleosome positioning or may be even free of nucleosomes. Such an irregular organization of *ras* oncogenic chromatin would resemble that of many other "normal", highly active eukaryotic genes.

The intrinsic affinity of trout ki-*ras* gene for aflatoxin B_1 was determined by *in vitro* alkylation experiments. Exon 1 of the gene was synthesized and labeled at the 5' end of the coding strand by the PCR technique. Taking advantage of the selective cleavage of AFB1-DNA adducts by piperidine under alkali conditions, the frequency of AFB1 attack to each guanyl site was determined by densitometric scans after the cleaved fragments were electrophoresed on sequencing gels. The results demonstrated that two guanyl sites of codon 12 had differential affinity to AFB1, the more 5' G was relatively inaccessible but the more 3' G was accessible, indicating that the sequence selectivity of AFB1 may contribute to the preference of the initial adduction *in vivo*.

Studies of Rainbow Trout Ki-*ras* Gene:

Sequencing, Aflatoxin B1 Binding, and Chromatin Structure

by

Xiaoshan Liang

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed May 6, 1993

Commencement June 1993

APPROVED:

Redacted for Privacy

Professor of Biochemistry and Biophysics in charge of major

Redacted for Privacy

Head of Department of Biochemistry and Biophysics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 6, 1993

Typed by Xiaoshan Liang .

This thesis is dedicated to my mother in heaven.

Acknowledgment

I would like to give my acknowledgment to the following people:

Dr. Ken van Holde, my academic professor, for giving me guidelines, encouragement, and financial support; especially for providing me a chance to work on this project.

Dr. George Bailey, for his instruction and advice on the academic problems, and for being my committee member.

Dr. John Morris, for serving as my graduate representative.

Drs. W Curtis Johnson and Sonia Anderson, for their kindness and being my committee members.

People in the van Holde's and Bailey's laboratories, especially Kate Mathews who gave me many suggestions during the whole process of this thesis study; and Valiere Stanik for her help on the lab business..

I also want to take this chance to thank my wonderful husband- Guangwen, for his love, understanding, support as well as his unforgettable academic help; my lovely son- Albert who gives me enjoyment and fun. My thanks extend to my other family members- my father, parents-in-law, sister, and brother-in-law, who have always loved me and also helped me in any way they could during the period of my Ph. D. study.

I also would like to show my appreciation to my friends- Feng Dong and Xiaokui Jin, for their friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Studies of Rainbow Trout Ki-*ras* Gene:
Sequencing, Aflatoxin B_1 Binding and Chromatin Structure

I. INTRODUCTION

I.1. The *ras* genes and their products:

The members of the *ras* gene family were first identified as cellular homologies of the Harvey and Kirstein rat sarcoma virus oncogenes (Bishop, 1983; Varmus, 1984). Genes bearing a high degree of homology to the mammalian *ras* family have been identified in organisms from virtually every evolutionary group, including *Drosophila*, yeast, and slime mold. In mammals, the *ras* genes appear to be expressed in all cell types and at all developmental stages. The high degree of evolutionary conservation and constitutive expression of these genes in different cell types has led to the suggestion that they fulfill some essential cellular functions.

I.1.1. Characteristics of gene products:

The *ras* gene family consists three functional genes, Ha-*ras*, Ki-*ras* and N-*ras* (Hall, 1984), which encode a group of closely related proteins with 188 or 189 amino acid residues and molecular weights of 21 kilodalton (Barbacid, 1987; Downward, 1990). These proteins, all named p21, are located at the inner side of the plasma membrane where they are anchored into the lipid bilayer by a palmitic acid covalently attached at the carboxyl terminus. The proteins contain homologous regions to and share features in

common with a broader family of GTPase proteins in that they bind guanine nucleotides, and exhibit slow intrinsic GTPase activity (Barbacid, 1987). As in the case for the G proteins, the GTP-bound form of p21 is "active" while the GDP bound form is "inactive". The *ras* p21 requires the "active" form to fulfill its normal cellular functions. Therefore the two forms have to be dynamically equilibrated and well regulated in order to avoid the event of escape from normal cell growth. One cytoplasmic protein, GTPase-activating protein (GAP), has been found to regulate the transition of the two forms by accelerating the GTPase activity, and thus the hydrolysis of GTP to GDP, thereby inactivating the p21 (Trahey and McCormick, 1987; McCormick, 1989).

The three-dimensional structures of both the GTP- and GDP-forms of the protein to a resolution of 1.35Å have been determined (De Vos *et al.*, 1988; Milburn *et al.*, 1990). Combining this with mutagenesis studies (Schlichting *et al.*, 1990), several essential regions of the p21 have been identified, which include the phosphate binding sites located between residues 8 to 20 and 55 to 65, sites for the guanyl nucleotide base between residues 110 to 120 and 140 to 150, and the effector loop between residues 32 to 40. The conformational changes upon GTP hydrolysis have been observed at the effector domain, and at the region between residues 61 and 75.

The functional and structural resemblance of the *ras* proteins with the G proteins has led to the proposal that normal p21 are normally involved in the transduction of external stimuli, most likely induced by growth factors or factors involved in cell

differentiation (Levitzki, 1988). The current model is that the *ras* proteins become activated upon stimulation, transduce the signal to some effector molecules, and subsequently become inactivated (Barbacid, 1987; Shih *et al.*, 1984; Spiegel *et al.*, 1987; and Gilman *et al.*, 1987). Mutated *ras* proteins, however, have lost the ability to become inactivated and thus stimulate growth or differentiation autonomously.

I.1.2. Activation of *ras* genes:

The *ras* oncogene is distinguished from the normal cellular protooncogene by the presence of one or more point mutations (Barbacid, 1987). Such alterations could be correlated with increase in malignant behavior (Bishop, 1987). Around 30% of all human tumors examined have been found to contain the *ras* oncogenes activated by such point mutations (Barbacid, 1987; Bos, 1989). Most of the point mutations occur at or near sites of interaction of the protein with the β- and γ- phosphates of the nucleotide (residues 12, 13, 59, 61 & 63). Mutations can also occur at sites of interaction with the guanine base (116, 117, 119 & 146). Both types of mutation have the effect of causing an accumulation of the GTP-bound form of p21. Mutant p21 is resistant to inactivation by GAP, resulting in continuous signal transduction. However, not all mutations at different sites may be equally effective at inducing tumorigenic transformation (Fasano *et al.*, 1984). Another type of activation is the overexpression of normal p21 proteins, an effect which has been reported in 21-50% of human tumors. This may represent an alteration mechanism as yet unidentified for

deregulated signal transduction in cancer cells without obvious genetic alterations. These mechanisms may take the form of other cellular genes influencing the oncogene regulation, or the stability of the oncogene mRNA.

While the exact role of the oncogenic activation of *ras* (point mutations and overexpression) in the formation of tumors is not understood, it appears to be one of a member of steps that may occur in the progression of cells from a normal to a fully malignant phenotype. For example, recent investigations have shown that the activated *ras* p21 increases the turnover of intermediates in the phosphatidylinositol pathway (Fleischman *et al.*, 1986; Wolfman and Macara, 1987), and stimulates the activity of phospholipase A_2 and phospholipase C (PLC) (Bar-Sagi *et al.*, 1988; Bar-Sagi, 1989). Both events can affect the activity of protein kinase C (PKC) which could provide for the activation of new gene products that facilitate tumor progression. This cellular signaling pathway has been proposed as one of the *ras*-mediated mechanism for the tumorigenesis.

I.2. Chromatin structure:

Eukaryotic DNA is complexed in the cell with both histones and nonhistone proteins to the extent of approximately 2 grams of proteins per gram of DNA. This complex constitutes a special structure called chromatin (van Holde, 1989).

Chromatin structure possesses several levels of organization. Simplest is the nucleosomal core particle, the building block of chromatin, in which 146 bp of DNA are wrapped about a core of eight histones (2 H_2A, 2 H_2B, 2 H_3-H_4 dimers) for 1.75 turns in a

left-handed helical manner. These histone cores are spaced along the DNA at about every 200 bp to form the basic chromatin fiber, which has a diameter of about 10 nm, and is therefore called the 10 nm fiber. Next, the 10 nm fiber is itself folded back to form a coiled structure called the 30 nm fiber. This compaction is promoted by binding of lysine-rich histones (H_1 and its variants) to the linker DNA. Then the 30 nm fiber can be further compacted into highly coiled interphase heterochromatin and metaphasic chromosomes (Figure I.2.1).

I.2.1. Nuclease hypersensitive sites:

The nucleosome array, the 10 nm fiber, is punctuated at broad intervals by short region (on the order of 50-400 bp) that appear to be nucleosome free. These region are called DNase I hypersensitive sites because of the relative sensitivity of the DNA therein to cleavage by DNase I and other reagents such as micrococcal nuclease and MPE.Fe(II) (Eissenberg *et al.*, 1985; Elgin, 1988). Such sites were first detected in SV40 viral chromatin (Varshavsky *et al.*, 1978; Scott and Wigmore, 1978) and in *Drosophila* genomic chromatin (Wu *et al.*, 1979). The usual experimental procedure for localizing hypersensitive sites in chromatin has been the method termed indirect end labeling (Wu, 1980). Nuclei are briefly digested with the cleavage agent of choice, and the cuts in chromatin are mapped relative to known restriction sites by a blot/hybridization protocol.

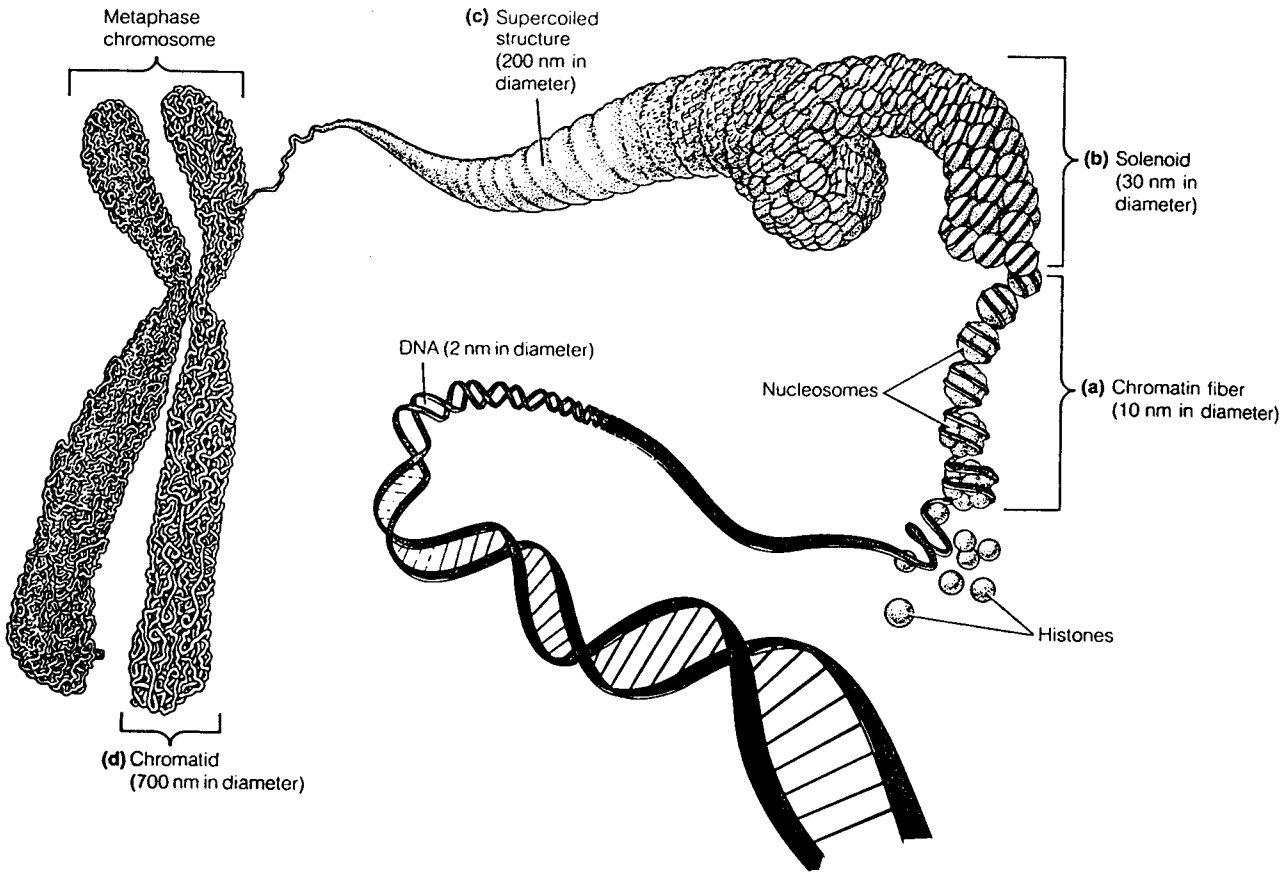


Figure I.2.1 A general picture of eukaryotic chromatin (from van Holde, 1989).

Specific cleavages in chromatin are assigned after careful comparison to protein-free DNA control digested to the same extent. While neither the structure nor function of chromatin hypersensitive sites are fully understood, many experiments have observed that these sites are frequently found at or near the 5' end of genes (see Reeves, 1984; Yaniv and Cereghini, 1986 for reviews), the region that a functionally important for the regulation of expression of transcriptionally active genes (Gross and Garrard, 1988). The formation of DNase I hypersensitive sites in some genes requires interaction of the DNA with specific proteins (Lowrey *et al.*, 1992). These results suggest that the hypersensitive sites may play a regulatory role for gene transcription. These sites may reflect the opening up of short regions of chromatin so that important DNA sequences, such as promoters, upstream activation sequences (UAS), terminators, enhancer elements of active or inducible genes, and origins of replication, recombination elements, and structural sites within or around telomeres and centromeres, become available for interactions with specific recognition molecules.

I.2.2. Nucleosome positioning:

Not all nucleosomes are arranged randomly on DNA. A number of researchers have demonstrated that at least nucleosomes can assemble on a specific fragment of DNA in a sequence-dependent manner (Simpson *et al.*, 1985; Eissenberg *et al.*, 1985). Thus, some nucleosomes do occupy defined positions within the chromatin, and certain homopolymers and specific DNA structure exclude nucleosome formation and thus can serve as boundaries for

nucleosome arrays (Kunkel *et al.*, 1981,). This phenomenon is called nucleosome "positioning". The exact mechanism is unknown, although factors such as the DNA sequence, local structure, or specific protein-DNA interactions may contribute to the nucleosome positioning (Chan *et al.*, 1984; Simpson and Stafford, 1983; Thoma and Simpson, 1985; Kleinschmidt and Steinbeisser, 1991; Hill *et al.*, 1991). Based on the fact that the spacer or linker DNA between nucleosomes is more accessible to certain agents including micrococcal nuclease than is the nucleosomal DNA, the nucleosome positioning can be determined by the indirect end labeling method (Wu, 1980). Chromatin with positioned nucleosomes, after digestion with micrococcal nuclease, gives rise to a nucleosomal ladder of DNA fragments from a given region at intervals of about every 200 bp (Keene and Elgin, 1981; Nedospasov and Georgiev, 1980). The absence of specific chromatin cleavage has usually been taken as evidence for random positioning of nucleosomes (assuming nucleosomes were shown by other analysis to be present on the sequence of interest).

I.2.3. Studies on the chromatin structure of oncogenes:

As an approach towards the understanding of possible relationship between chromatin organization and regulation or activation of the oncogenes, researchers have been paying special attention to the chromatin structure of oncogenes in the past several years. The most well studied so far is the c-*myc* gene. By mapping with DNase I, seven hypersensitive sites were found in the c-*myc* loci, primarily in the 5' flanking region (Kumar and Leffak, 1991).

The pattern of hypersensitive sites was changed after ALV-LTR integration (Schubach and Groudine, 1984), and translocation to Ig enhancer (Siebenlist *et al.*, 1984; Fahrlander *et al.*, 1985). The results also show that the chromatin structure can undergo rapid and reversible changes (folding and unfolding of nucleosomes) accompanying the activation, repression and superinduction of c-*myc* and c-*fos* (Chen and Allfrey, 1987; Dyson *et al.*, 1985), indicating that nucleosome structures may be dynamic. It was determined (Kumar and Leffak, 1991) that the local sequence information is a determinant in establishing the chromatin structure of the human c-*myc* upstream region. In addition, local curvature in the DNA structure also seemed to mediate the ordered nucleosome structure (Kumar and Leffak, 1989).

Another well characterized oncogene, the *ras* gene, has also been investigated for its chromatin structure features. Kasid and co-workers (1985) studied the human c-H-*ras* chromatin *in vivo* as probed by micrococcal nuclease digestion. Their result demonstrated that c-H-*ras* chromatin lacked typical nucleosomal packaging, a feature similar to that of "normal" actively transcribing eukaryotic genes. In addition, the chromatin structure of the mouse N-*ras* gene was studied by Paciucci and Pellicer (1991), who performed digestion of nuclei with DNase I. Their work revealed the presence of four hypersensitive sites at the 5' flanking region, two of which colocalized with the promoter and the negative regulatory elements respectively. These sites correlate with formation of specific DNA-interaction complexes *in vitro* and the pattern is conserved in different cell types. All these results suggest

that the organization of mammalian *ras* oncogenic sequence bears a relationship to the "normal" functional organization of eukaryotic DNA by nucleosomes.

I.3. Aflatoxin B_1 and carcinogenesis:

Aflatoxin B_1 is a mycotoxin produced by a number of strains of fungal genus *aspergillus*. These fungi are found as natural contaminants of the food chain. Experimentally, aflatoxin is a highly toxic, powerful mutagen and most potent liver carcinogen. In chemical form, it can be considered a highly substituted coumarin with a double bond at 8, 9 position (Figure I.3.1). This 8, 9 double bond can be oxidized to epoxide by a mixed functional oxidase complex *in vivo*, or by oxidants such as chloroperoxybenzoic acid *in vitro*. Such activation, in turn, allows aflatoxin to interact covalently with DNA and other macromolecules. Therefore this double bond and its oxidation are essential for the toxic, mutagenic and carcinogenic properties (Essigmann *et al.*, 1977; Lin *et al.*, 1977). Among the experimental animals studied, rainbow trout has been identified to be the species most sensitive to aflatoxin-induced carcinogenesis, as such, it can be considered as an excellent model for study of liver tumor induction by AFB1 (Bailey *et al.*, 1984).

I.3.1. Formation of aflatoxin-DNA adducts:

The target of aflatoxin B_1 is DNA. Therefore it is believed that the mutagenic and carcinogenic properties of AFB1 are mediated via the formation of DNA adducts (Miller and Miller, 1981). AFB1 interacts at the N-7 position of guanines in double strand DNA

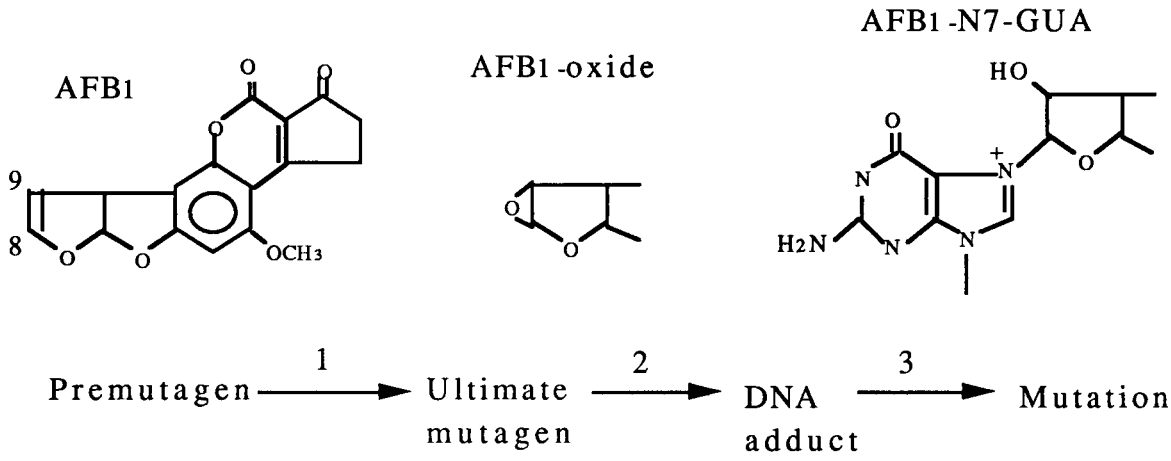


Figure I.3.1 Structure of aflatoxin B_1 and its DNA adduct.

(Figure I.3.1) forming a bulky N^7-guanyl AFB1 adduct (Groopman *et al.*, 1981). The binding of AFB1 to DNA has certain features. 1). The initial binding is possibly mediated by intercalation of activated AFB1 (Tullius and Lippard, 1982). Therefore it is not surprised to find that the frequency of adduction of single-stranded DNA is very low (Yu *et al.*, 1990). 2). Binding affinity is somewhat sequence-dependent (Muench *et al.*, 1983). Misra and co-workers (Misra *et al.*, 1983) provided a set of empirical rules to generalize the affinity of activated AFB1 for guanines within a known DNA sequence, based on nearest-neighbor nucleotides (Table 1). Benasutti and co-workers (1988) made an extensive study of various DNA sequences to ascertain a set of values determining activated AFB1 reaction intensities at particular guanines, again based on nearest neighbors. However, the results were not fully in agreement (Marien *et al.*, 1987 as an example), indicating a present lack in understanding of all the variables required to correctly predicted reaction affinities. Therefore, the reaction intensities of selected sites within any gene sequence have to be determined empirically.

Aflatoxin-DNA adducts can be removed by enzymatic hydrolysis of the bond, between C-8 of AFB1 and the N-7 of guanine, resulting in restoration of the DNA strand without damage. The hydrolysis can also occur at the N-9 glycosidic bond linking the adduct base to the deoxyribose of the DNA, resulting in the release of AFB1-N^7-G and depurination of DNA with the formation of an aguaninic site, or at the C-8 of guanine causing the opening of the imidazol ring and the formation of the persistent DNA-bond formadic-pyrimidine derivative (AFB1-FAPyr). The latter two

Table 1

The sequence context of aflatoxin B_1 binding to DNA (from Misra *et al.*, 1983).

---

(1) Single G residues flanked by A-T nucleotides [$(A/T)_n G(A/T)_m$ where n and m =1 or more] are poor targets.

(2) In GC bp clusters flanked by A-T nucleotides [e.g., $(A/T)_n GG(A/T)_m$], the following patterns are observed.

(a) In the dinucleotide GG the 5' G is an intermediate target while the 3' G is a strong site.

(b) In the dinucleotide CG the G residue is an intermediate site.

(c) In the dinucleotide GC the G is a poor site.

(d) In the trinucleotide CCG the G residue is a strong site, while in GCC it is a poor site.

(e) In the contiguous G clusters GGG and GGGG, the 5' and 3' residues are intermediate, and the middle residues are strong sites.

(f) The sequence-specific effects observed in GC bp clusters at the dimmer and trimmer level are essentially retained in longer GC bp clusters, except for certain additive effects. For example, the 5' G residue in the clusters GC, GCC, GCCC and GCGGC are all poor targets.

(3) The observed specificity is independent, within the limits of the experimental conditions used, of (a) methods of activation of AFB1, (b) source (prokaryotic or eukaryotic) of DNA, (c) dosage, (d) reaction time, and (e) ionic strength and thus appears to be an intrinsic feature of AFB1-DNA interactions.

---

reactions would require normal repair system to restore the DNA to its undamaged state. Failure of repair leads to mutation of the gene, and thus tumor formation.

I.3.2. *ras* gene mutations induced by AFB1:

The general mechanism of chemical carcinogenesis is still an unsolved problem. Recent studies have suggested that this process involves oncogene activation (Guerrero and Pellicer, 1987), which has been initiated by direct reaction between DNA and chemical intermediates (Zarbl *et al.*, 1985). Aflatoxins are one of the models for this kind of study.

Although there is no direct evidence for cause-effect correlation between N^7-G modification and mutagenesis, AFB1 modification does affect the template function of DNA, since depurination and strand scission are promoted and both of the events have biological consequences. The most significant consequence is the mutation of *ras* gene. McMahon and co-workers (McMahon *et al.*, 1986) have identified activated *ras* oncogenes in primary tumors derived by treatment of rats with AFB1, and the activation was found to be through point mutations at codon 12 of ki-*ras* gene (McMahon *et al.*, 1987) as well as codons 13 and 14 of the N-*ras* gene (McMahon *et al.*, 1990). Other workers have found similar mutations in AFB1-induced rat liver rumors (Sinha *et al.*, 1988). Recently, Chang *et al.* (1991) have identified point mutations at codon 12 and 13 of the ki-*ras* gene in rainbow trout liver tumors induced by AFB1. Table 2 summarizes these results. These reproducible activation of *ras* oncogenes in AFB1-induced tumors have made it possible to

Table 2

ras gene mutations induced by Aflatoxin B_1

ras gene	specie & tissue	mutations	reference
Ki-*ras*	trout liver	12 GGA-GTA	Yung-Jin Chang, *et al.*
		13 GGT-GTT	Mol. Carcinogenesis
		12 GGA-AGA	4:112,1991
Ki-*ras*	rat liver	12 GGT-GAT	Gerald McMahon, *et al.*
		12 GGT-TGT	PNAS 87:1104, 1990
		12 GGT-AGT	PNAS 84:4974, 1987
N-*ras*	rat liver	13 GGT-GTT	same as above
		14 GTT-ATT	
Ki-*ras*	rat liver	12 GGT-GAT	S. Sinha, *et al.*
			PNAS 85:3673, 1988

correlate their activating mutations with the known mutagenic effects of AFB1.

I.3.3. Chromatin-carcinogen interactions:

Chromatin-carcinogen interactions have been studied extensively in past several years (Boulikas, 1991). Experimental evidence indicates that genotoxin agents like AFB1 bind non-randomly in eukaryotic chromatin. By isolating nucleosomal core particles and oligonucleosomes from carcinogen-labeled chromatin as a direct approach to study the distribution of carcinogens, Bailey *et al.* (1980) found that the linker DNA is about 5-fold more susceptible to AFB1 than the histone-protected core DNA. Moyer *et al.* (1989) used nucleosomes that are homogenous in DNA sequence and hence in DNA-histone contact points, to study the site-specific aflatoxin B_1 adduction, and they found that the binding of AFB1 is suppressed about 2.4 fold at guanyl sites within nucleosomes compared with AFB1 affinity at the corresponding sites in naked DNA. All these results showed that AFB1 prefers to bind in linker region rather than nucleosome regions.

I.4. PCR and its applications:

PCR, the polymerase chain reaction, is an *in vitro* method for the enzymatic synthesis of specific DNA sequences, using two oligonucleotide primers that hybridize to opposite strands and flank the region of interest in the target DNA (Mullis *et al.*, 1986). Since it was invented in 1983 by Kary Mullis, the PCR technique has revolutionized molecular biology by virtue of its utility in

amplifying milligram amounts of a specific, targeted gene from as little as a single copy of starting template DNA. To data, a numerous research articles related to the PCR technique and applications have been published, and the rate of publication is believed to be consistently growing in a ridiculous speed. So it may be reasonable to say that today's molecular biology is the era of PCR. Because this technique was employed so extensively in this work, it is appropriate to review it briefly here.

I.4.1. Method:

The PCR method consists of repetitive cycles of denaturation of the DNA, annealing of oligonucletides homologous to sequences flanking the segment of interest, and primer extension by a DNA polymerase (Figure I.4.1). Because the primer extension products synthesized in one cycle can serve as a template in the next, the number of target DNA copies approximately doubles at every cycle. Thus, the specific fragment defined by the positions of the 5' ends of the two primers is increasing in concentration exponentially, and theoratically 30 cycles of PCR yields about a billion fold (2^{30}) amplification.

PCR requires several key components, which include the target DNA, at least two synthetic primers, a DNA polymerase, dNTPs and a divalent metal ion (Mg^{2+}). No single set of conditions will guarantee success for all uses; consequently, each new PCR application is likely to require optimization. Nevertheless, a general guideline for the

Figure I.4.1 The polymerase chain reaction (PCR). The reaction involves repeated cycles of (a). heat denaturation of the DNA, (b). annealing of the primers (A and B), and (c). extension of the primers. In the step (a), the template is denatured by heating at 9 4^{o}C for 3-5 min and rapidly cooling on ice. Then in the step (b), two oligonucleotide primers (A and B) are annealed to the denatured template at their homologous regions, and this is followed by the step (c), in which, the primers are extended with DNA polymerase to synthesize the complementary strands of the template. By repeating these steps 30 to 40 cycles, the specific fragment defined by the positions of the 5' ends of the primers is increasing in concentration exponentially, yielding about a billion fold amplification (cont. on page 19).

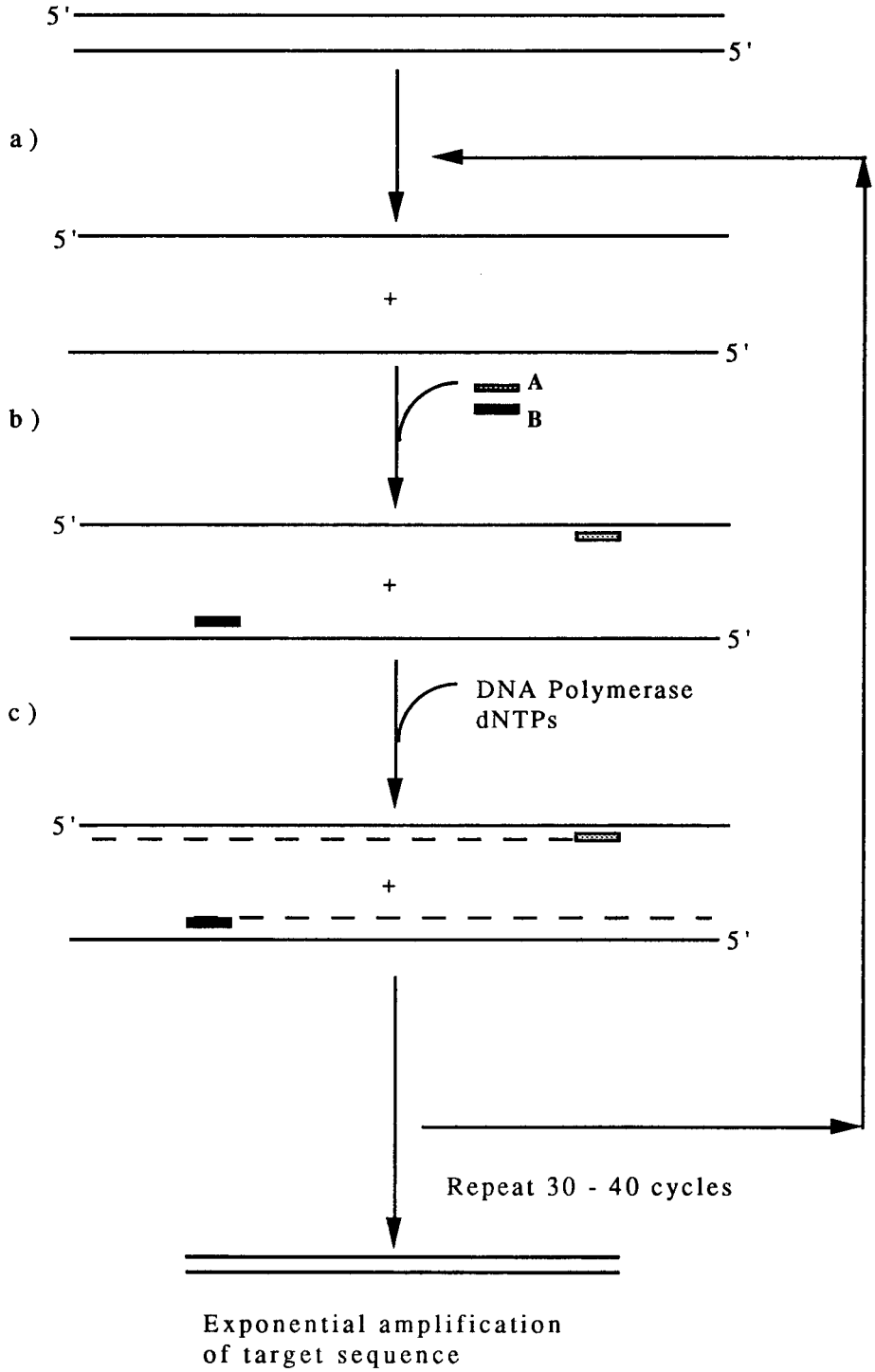


Figure I.4.1 The polymerase chain reaction (PCR).

consideration and optimization of these conditions has been provided (Erlich, 1989; Erlich *et al.*, 1989; Innis *et al.*, 1990).

One great breakthrough in PCR is the automation of the technique. This was made possible by an instrument with automatic temperature cycling and the substitution of a thermal stable DNA polymerase, purified from *Thermus aquaticus* (Taq) for the previously used *E. coli* DNA polymerase I, Klenow fragment (Saiki *et al*, 1985; Saiki *et al.* 1988). The advent of Taq DNA polymerase greatly simplifies the PCR procedure because it is no longer necessary to replenish the enzyme after each PCR cycle. Use of Taq DNA polymerase at high annealing and extension temperatures increases the specificity, yield and length of products that can be amplified and, thus increases the sensitivity of PCR for detecting rare target sequences.

I.4.2. Applications:

PCR has been applied to study a wide variety of problems in molecular biology (Bloch, 1991), such as genetic mapping (Welsh *et al.*, 1991; Parker *et al.*, 1991), genetic polymorphism (Weber and May, 1989), detection of mutations (Saiki *et al.*, 1985; Wong *et al.*, 1987; Saiki *et al.*, 1988), molecular virology (especially retrovirology) (Kwok *et al.*, 1989; Ehrlich *et al.*, 1989), and transcriptional splicing and regulation (including molecular immunology) (Rotenberg *et al.*, 1989). In addition, medical applications, especially diagnostic applications, are now a important area of PCR technology and this is based on the capacity of PCR to

amplify a target sequence from crude DNA preparations as well as from degraded DNA template.

I.4.3. Modifications:

A major limitation of conventional PCR is that DNA sequences situated immediately outside the primers are inaccessible. In many circumstances one wishes to characterize these kinds of unknown sequences which are contiguous to a known segment of DNA. Examples of these situations include characterization of 5' or 3' ends of cDNA, sequences surrounding insertional sequences, and translocation and rearrangements of antigen receptor genes. For such purposes, PCR can not be used without some modifications. Strategies to overcome this limitation have been developed and two of these appropriate to this thesis are described below.

1). Inverse PCR (Ochman *et al.*, 1988; Silver and Keerikatte, 1989; Triglia *et al.*, 1988):

This technique is based on the simple procedures of digestion of source DNA with restriction enzymes and circularization of cleavage products before amplification using primers synthesized in the opposite orientations to those normally employed for PCR. In general, inverse PCR permits the amplification of the upstream and/or downstream regions flanking a specified segment of DNA without resorting to conventional cloning procedures and can also be used to rapidly produce hybridization probes for identifying and orienting the adjacent or overlapping clones from a DNA library.

2). Ligation-mediated PCR (Mueller and Wold, 1991):

This technique involves the attachment of an unique DNA linker to a blunt-end by DNA ligase, by which one end defined by the linker primer is provided. Together with a gene specific primer, a specific segment of DNA can be amplified exponentially by conventional PCR. This method was originally developed for study *in vivo* protein-DNA interactions at single-copy genes in complex genomes (Mueller and Wold, 1989), but it is applicable to other PCR problems in which only one end of the region to be amplified is known. For example, it has been used for genomic sequencing (Pfeifer *et al.*, 1989), *in vivo* methylation analysis (Pfeifer *et al.*, 1989; Pfeifer *et al.*, 1990; Steigerwald *et al.*, 1990), and cloning of promoter elements (Fors *et al.*, 1990). Of the various single-sided PCR strategies, this technique is uniquely suited to the amplification of a genomic sequence ladder because it presents the single base resolution present in the starting material. In the next chapter of this thesis, the application of LM-PCR to the study of chromatin structure of a specific gene will be described.

I.5. Aims of this research:

It is known that in the trout system, extremely low levels of aflatoxins can result in the development of liver tumors (Bailey *et al.*, 1984). Studies have demonstrated that in many cases the genetic change has been a obviouse consequence, with evidence of that mutations in codons 12 and 13 of the *ras* gene are induced by aflatoxins (Chang *et al.*, 1991). The possibility can no longer be ignored that the chromatin structure of the *ras* gene in the liver

cells may participates in the extreme sensitivity to aflatoxins. There are several ways in which the local ordering of chromatin might promote or hinder access to critical sites.

1). DNase I hypersensitive sites represent nucleosome-free regions. If the critical sites like codons 12 and 13 of the *ras* gene are arranged in these regions, the accessibility to aflatoxins might be enhanced.

2). Even if nucleosomes are present in the region in question, they may be positioned so as to place a critical activation site within a linker region. We know from studies that the linker regions are more accessible to aflatoxin B_1 than are the nucleosomes themselves (Bailey *et al.*, 1980).

From these reasons, it seems that an important part of any thorough investigation of the *ras* oncogenes from trout must include a study of their chromatin structure, especially in the regions surrounding critical residues. This thesis represents an attempt to unveil such chromatin effect on aflatoxin B_1-mediated *ras* gene activation. The chromatin structure has been studied from two aspects: first, a search for DNase I hypersensitive sites and second, examination of possible nucleosome positioning by micrococcal nuclease. This study takes advantage of the amplification property of PCR and LM-PCR to increase the sensitivity and specificity, allowing such studies on single-copy genes on a large genome.

A number of studies have shown that aflatoxin B_1 binding to DNA shows some sequence specificity (Misra *et al.*, 1983; Marien *et al.*, 1987). Therefore it is of equal importance to determine whether or not the critical residues in the *ras* gene represent intrinsic "hot

spots" for AFB1 adduct formation. In this thesis, the study of chromatin structure in *ras* gene is accompanied by quantitative assessment of the relative aflatoxin affinity of "critical" guanine residues at the DNA level. This is carried out by *in vitro* alkylation of *ras* gene by activated aflatoxin B_1.

II. MATERIALS AND METHODS:

II.1. Enzymes, chemicals and buffers:

All restriction enzymes used in this study were purchased from New England Bio-Labs (NEBL) and Bio-Rad. T4 DNA ligase was from Promega. T4 polynucleotide kinase was from New England Bio-Labs. Taq DNA polymerase was from Perkin-Elmer. Sequenase Version 2.0 was from United States Biochemical (USB). DNase I and Micrococcal nuclease were from Worthington. The digestions using these enzymes were carried out under the conditions that the companies provided or suggested.

Dimethyl sulfate, bromine and piperidine were purchased from Aldrich, Baxter and Sigma respectively. They were used directly without further purification. Aflatoxin B_1 (AFB1) was purchased from Sigma. The AFB1 stock solution was made by dissolving of dried AFB1 in 95% ethanol to approximately 100 µg/ml. This stock solution was scanned from 220 to 400 nm on a Varian DMS UV-VIS spectrophotometer, and the concentration was determined using an extinction coefficient of 21,800 at 364 nm for AFB1 (Rodericks *et al.* 1970). The AFB1 was made to 10 µg/ml in dimethyl chloride and stored at -20°C until needed.

Buffers were prepared, autoclaved and stored at room temperature or -20°C. The components and their concentrations, as well as the applications of each buffer are listed in Table 3.

II.2. Purification of DNA by electroelution:

Most of the nucleic acids used in this thesis study were purified

Table 3 Buffers

Names	Components	Application
TAE	40 mM Tris-acetate, 2 mM EDTA,	Electrophoresis
TBE	89 mM Tris-borate,	Electrophoresis
	89 mM Boric acid, 10 mM EDTA.	
TE	10 mM Tris.HCl pH 8.0, 1 mM EDTA.	DNA buffer
GTEL	50 mM Glucose, 10 mM EDTA,	Plasmid prep.
	25 mM Tris-HCl pH 8.0,	
	4 mg/ml Lysozyme).	
SRB	40 mM Tris.HCl pH 7.5,	Sequenase reaction,
	20 mM $MgCl_2$, 50 mM NaCl.	
EDB	10 mM Tris.HCl pH 7.5,	Dilution for Sequenase
	5 mM DTT, 0.5 mg/ml BSA.	
H buffer	60 mM KCl, 15 mM NaCl,	Nuclei prep.
	0.15 mM Spermine,	
	0.5 mM Spermidine,	
	15 mM Tris.HCl pH 7.6,	
	15 mM β-MeOH, 0.2 mM PMSF.	
D buffer	10 mM Tris.HCl pH 7.6,	Nuclease digestion
	25 mM KCl, 5 mM $MgCl_2$,	
	0.5 mM $CaCl_2$, 0.35 M Sucrose.	

Table 3 Buffers (cont.)

Names	Components	Application
T4 kinase buffer	50 mM NaCl, 10 mM Tris.HCl pH 8.3, 3 mM $MgCl_2$, 5 mM DTT, 1 mM Spermidine.	End-labeling of DNA
T4 ligase buffer	3 mM Tris.HCl pH 7.8, 10 mM $MgCl_2$, 0.5 mM ATP, 10 mM DTT.	DNA ligation
BAP buffer	50 mM Tris.HCl pH 8.0, 50 mM NaCl.	Dephosphorylation
Taq A buffer	50 mM KCl, 10 mM Tris.HCl pH 8.0, 2 mM $MgCl_2$, 0.01% Gelatin.	PCR
Taq B buffer	40 mM NaCl, 10 mM Tris.HCl pH 8.9, 2 mM MgCl2, 0.01% Gelatin.	PCR
Loading buffer	0.01% Bromphenol blue, 0.01% Xylenecyanol, 20 mM EDTA, 95% Formamide.	Electrophoresis
LB	1% Bacto-tryptone, 0.25 M NaCl, 0.5% Yeast extract.	Growing of Bacteria
SOC	2% Bacto-tryptone, 10 mM $MgCl_2$, 0.5% Yeast extract, 25 mM KCl, 10 mM NaCl, 10 mM $MgSO_4$, 20 mM Glucose.	Growing of Bacteria

and recovered by electroelution. A specific apparatus called the S&S ELUTRAP from Schleicher & Schuell was used. This system consists of an Elutrap device and a standard electrophoresis chamber. The Elutrap device consists of a sample chamber limited at each end by a S&S BT1 membrane which does not pass molecules larger than 3-5 kilodaltons. A microporous membrane, S&S BT2, acts as a prefilter that prevents agarose and other particulate from entering the purified sample. Together, the BT2 and BT1 membranes form a "trap" into which the sample migrates. To purify DNA, gel slices containing desired nucleic acids were placed in the sample chamber of the elutrap device. Then the device was placed in the standard electrophoresis chamber. All chambers (sample chamber and electrophoresis chamber) were filled with 1x TAE buffer. The electrophoresis was performed for one hour to overnight (depending on the size of the molecules need to be purified) at 100-150 volts. When the electrophoresis was finished, the purified samples were collected from the "trap" and DNA was recovered by ethanol precipitation or SpeedVac lyophilization.

II.3. Oligonucleotides:

All oligonucleotides used in the sequencing, PCR and LM-PCR were synthesized on a 0.2 or 1 μM scale with an ABI Model 380B DNA synthesized in the Center for Gene Research and Biotechnology at Oregon State University. They were purified by HPLC or gel electroelution. First, the lyophilized oligonucleotides from the Center were resuspended in 300 μl of TE buffer, desalted through a

Table 4 Oligonucleotides

*The numbers in this column are the nucleotide positions in the 5' flanking region of the trout *ras* gene, referring +1 as the first nucleotide of exon 1. N/A represents the "not relative" sequence.

names	position*	sequences(5'-3')	use
xlprm2	-884 to -867	GGGGGAATTGTAAACAGC	sequencing
xlprm3	-198 to -215	CTCCCCTTCTCCATAGCC	sequencing
xlprm4	-631 to -615	CACATGGTTAGCAGATG	sequencing
xlprm5	-450 to -467	TTATGCCGCACACCATCG	sequencing
Prma1	-111 to -96	GGAGATCCCAGTTGAA	LM-PCR
Prma2	-105 to -81	CCCAGTTGAAGGATTGTTGTGTCGA	LM-PCR
Prma3	-105 to -70	CCCAGTTGAAGGATTGTTGTGTCGATAGTTAGTGTG	LM-PCR
Prmb1	+160 to +145	CATGCCTTGCTGAAGT	LM-PCR
Prmb2	+152 to +128	GCTGAAGTAAGAGACAAGGGTGTAC	LM-PCR
Prmb3	+152 to +119	GCTGAAGTAAGAGACAAGGGTGTACTGAAACTAC	LM-PCR
Prmc1	-1013 to -997	TACTGCCTGAGGCCAGA	LM-PCR
Prmc2	-1003 to -981	GGCCAGAACCCATTGTGTAGCTGTC	LM-PCR
Prmc3	-1003 to -971	GGCCAGAACCCATTGTGTAGCTGTCACTCGCAA	LM-PCR
Prmd1	-74 to -90	CTAACTATCGACACAAC	LM-PCR
Prmd2	-85 to -111	CACAACAATCCTTCAACTGGGATCTCC	LM-PCR
Prmd3	-85 to -120	CACAACAATCCTTCAACTGGGATCTCCAAAACCTGG	LM-PCR
Liprm1	N/A	GCGGTGACCCGGGAGATCTGAATTC	LM-PCR
Liprm2	N/A	GAATTCAGATC	LM-PCR
RHc37	+111 to +91	CTCGATGGTGGGGTCATATTC	PCR
RH38	+680 to +701	GACTCGTACAGGAAGCAGGTGG	PCR
RH01	+1 to +21	ATGACAGAATACAAGCTGGTG	PCR
5'-25	-25 to -5	GGGCTTTCTCTCCCACAGGTG	PCR
-40	N/A	GTTTTCCCAGTCACGAC	sequencing
reverse	N/A	AACAGCTATGACCATG	sequencing

15 ml Sephadex G-10 column and then purified by HPLC (Dr. Pui Shing Ho's lab) with the step gradient of a combination of solvent NH_4HAc from 90% to 60% and CH_3COOH/NH_4H from 10% to 40%. Alternatively, the lyophilized oligonucleotides were resuspended in 100 µl of TE buffer, mixed with an equal volume of 8 M urea/dye buffer and loaded into an 12% urea polyacrylamide gel with 8 cm wide lane followed by electrophoresis for about 4 hours at 350 volts in 1x TBE buffer. After electrophoresis was finished, the oligonucleotide band was located and excised from the gel under UV-shadowing. Then DNA was electroeluted from the gel slice in 1x TAE buffer at 150 volts for 2 hours as mentioned above. The oligos purified either by HPLC or by gel separation were recovered by ethanol precipitation followed by centrifugation at 10,000 rpm for 30 min in an Eppendorf microcentrifuge. The oligo pellets were dissolved in TE buffer and the DNA concentration was determined by UV absorbance (one OD_{260} equals 20 µg/ml). All the oligonucleotides were diluted to 60 pmol/µl in TE and stored at -20°C until needed. The names and sequences of oligonucleotides are listed in Table 4.

II.4. End-labeling of DNA by T4 polynucleotide kinase:

The oligonucleotide (10 to 60 pmole) was 5' end-labeled by T4 polynucleotide kinase (NEBL, 10 units/µl; 1 unit is required for incorporation of 1 nmole of acid insoluble ^{32}P in 30 min at 37°C) using gamma-^{32}P-ATP (5 µl at 6000 Curies/mmole, from NEN) in T4 kinase buffer (supplied with the enzyme by the company) in a total

volume of 20 µl. The mixture was incubated at 37°C for 30 min. The kinase was denatured by heating the tube to 65°C for 10 min. Free labels were removed by a spin-column (Quick-Spin™ G-25 Sephadex, from Boehringer Mannheim Corporation). The direct pass-through material, containing the labeled primer, was collected in a microcentrifuge tube and used as soon as possible. Usually the labeling efficiency was 10^5 cpm/µg.

II.5. Denaturing sequencing gel electrophoresis:

Two stock solutions were made. Stock solution 1 is called the "30% polyacrylamide solution", and contained 29% acrylamide, 1% Bis:acrylamide, 7 M urea and 1x TBE. Stock solution 2 is the "urea solution", and contained 7 M urea in 1x TBE. To make a 6% sequencing gel, 8 ml of polyacrylamide solution was mixed with 32 ml of urea solution. Then 0.4 ml of 10% ammonium persulfate and 50 µl of TEMED were added, mixed by swirling and the whole mixture was poured into a 0.4 x 30 x 50 ml glass plate sandwich. The polymerization was completed in about 15 min at room temperature. The gel was then prerun at constant current of 30 mA in 1x TBE for about 30 min or until the temperature of the gel reached to 40-45°C. Prior to loading, the samples were heated at 80°C for 3-5 min. The electrophoresis was performed at constant current of 30 mA until the blue dye ran at about 5 cm from the edge of the gel. The wet gel was transferred to a Whattman 3 MM paper, wrapped with plastic wrap and dried at 80°C under vacuum for 1 to 2 hours. The dried gel was exposed to a Kodak X-OAR film

(from Kodak) or Hyperfilm (from Amersham) for overnight to several days.

II.6. Isolation of nuclei and nuclease digestion:

Fresh rainbow trout liver tissue, either normal or tumor tissue, (from the Food Toxicology and Nutrition Laboratory) was perfused with H buffer, referred to as homogenization buffer (see Table 3), containing 0.32 M sucrose, using a 10 ml syringe with a 26 gauge needle. The tissue was then minced into four volumes of the same buffer with scissors, homogenized using a Polytron (Tissuemizer, from Teliman company) for 30 seconds, followed by passage through a motor-driven Teflon-glass homogenizer ten times. The homogenate was filtered through four layers of cheesecloth, diluted with eight volumes of H buffer containing 0.32 M sucrose and centrifuged at 4,000 rpm for 10 min. The pellet, which was referred to as "crude nuclear pellet", was resuspended in 5 ml of H buffer, and mixed with 20 ml of H buffer containing 2.3 M sucrose. The crude nuclear solution was then layered over 5 ml of 2.3 M sucrose in H buffer and centrifuged at 25,000 rpm in SW 28 rotor at 4°C for 90 min. After centrifugation, the nuclei formed a nice white tight pellet at the bottom of the centrifuge tube.

The nuclei were washed once with D buffer, referred to as digestion buffer, and resuspended in the same buffer at 2 mg/ml for nuclease digestion. An aliquot of nuclei was lysed in 0.1 M NaOH and the concentration of nuclei, in terms of DNA amount, was determined by UV absorbance at 260 nm, using a value that one OD

unit equals 50 μg/ml for double-stranded DNA (Sambrook *et al.*, 1989)

Nuclease Digestion: The nuclei were digested with DNase I or micrococcal nuclease to check the quality of the chromatin as well as for the following studies of the nucleosome structure and positioning of the *ras* gene.

DNase I digestion: About 0.2 mg of nuclei in D buffer were digested with different amounts of DNase I at 0, 0.25 u, 0.5 u, 1 u, 2 u, 4 u, 8 u and 16 u at 37°C water bath for 5 min. The reaction was stopped by adding SDS and EDTA to 1% and 20 mM respectively.

MNase digestion: 46 units of MNase were added to the 2 mg of nuclei suspension in D buffer and the mixture was incubated at 37°C in a water bath. Aliquots of 0.2 mg of digest were drawn from the mixture at different times and the digestion was terminated by adding SDS and EDTA to 1 % and 20 mM respectively.

II.7. Genomic DNA isolation:

Trout liver DNA, either from undigested nuclei or nuclease digested nuclei, was isolated by the following treatment. First, the suspension of nuclei was adjusted to 1 M NaCl, and the proteins were removed by several rounds of phenol, phenol/chloroform and chloroform extraction. Then RNA was removed from the nuclei by incubation with 50 μg/ml of ribonuclease A/T at 37°C for 3 hours. After the ribonuclease was destroyed by incubation of the mixture with Proteinase K at 100 μg/ml at 37°C for overnight, any residual proteins or peptides were removed by phenol, phenol/chloroform

and chloroform extraction. DNA was recovered by ethanol precipitation and centrifugation. The DNA concentration was determined by UV absorbance and made to 1 mg/ml.

II.8. Inverse polymerase chain reaction:

This technique was used to amplify the 5' flanking region of trout ki-*ras* gene, whose brief restriction map is shown in Fig. II.1.1. It involves two steps - rearrangement of the gene and the conventional polymerase chain reaction.

Rearrangement of the gene: The trout liver DNA was isolated as described above and digested with Pst I. The bulk of the Pst I digested DNA was loaded on a 1% preparative agarose gel and separated by electrophoresis. After the electrophoresis was finished, the gel was briefly stained with ethidium bromide at 50 µg/ml and DNA bands at a range of 1.5-3 Kb were located under UV light, excised and electroeluted from the gel in 1x TAE buffer at 150 volts for over 4 hours. using the S&S' Elutrap. DNA was then recovered by ethanol precipitation and centrifugation. The DNA pellet was resuspended in TE buffer and the concentration was determined by UV absorbance.

DNA isolated as above was circularized by ligation with 3 units of T4 DNA ligase in 1x ligation buffer (supplied with the enzyme from Promega) at 4^{o}C for overnight and then relinearized by digestion with Xba I. By doing so, the 5' flanking region of trout ki-*ras* gene was now located between its exon 2 and exon 1 (see results).

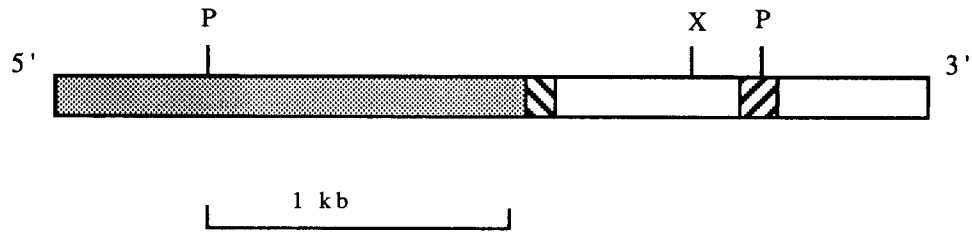


Figure II.1.1 Partial restriction map of trout ki-*ras* gene. P is indicated as Pst I site, X as Xba I site. Boxes with left or right diagonal lines represent exon 1 and exon 2 respectively. Open boxes depict introns, and dotted box depicts 5' flanking region. Digestion of Pst I generated 1.9 Kb fragment.

Polymerase Chain Reaction (PCR): In a 0.5 ml microfuge tube, about 1 µg of DNA from the above step was heat denatured at 94°C for 10 min and quickly chilled on ice. The amplification mix containing 1x Taq buffer A, 60 pmol of 5' primer RH38 and 3' primer RHc37, 0.2 mM each dNTPs and 5 units of Taq DNA polymerase was added. Total reaction volume was 100 µl. Then about 100 µl of mineral oil was layered on top of the reaction mixture. The tube was then placed in an automatic thermal cycler (Single Block EasyCycler, Ericomp, Inc.) and 40 cycles of amplification were performed, programmed for denaturation at 94°C for 2 min, annealing at 54°C for 2 min, followed by extension at 72°C for 3 min in each cycle. The PCR product was gel purified and recovered by electroelution as described above.

II.9. Cloning of PCR product into pUC 19:

Before cloning, the pure PCR product was phosphorylated by incubation of DNA with T4 polynucleotide kinase (20 units, from NEBL) in 50 µl of 1x kinase buffer containing 0.2 mM ATP at 37°C for 30 min. The reaction was terminated by addition of EDTA to 10 mM.

To increase the cloning efficiency, after the digestion of vector pUC 19 with Sma I, the plasmid was dephosphorylated by treatment with 0.6 unit of bacterial alkaline phosphatase (BAP, 0.29 u/µl; from IBI) in 1x BAP buffer at 68°C for 90 min. BAP was then denatured by digestion with Proteinase K at 100 µg/ml at 55°C for 30 min followed by two cycles of phenol/chloroform extraction. After the above treatment, the DNA molecules (both insert and vector) were

recovered by ethanol precipitation and centrifugation.

Next, insert and vector were ligated at the ratio of vector to insert of 5 to 1, using T4 DNA ligase in 20 µl of 1x ligation buffer (supplied with the enzyme from Promega) at 4°C for overnight. The next day, the ligation mixture was diluted 5 fold with TE buffer. 5 µl of diluate containing about 2.5 ng DNA was added to 20 µl of competent cell DH5α (library efficiency, from BRL), mixed by gentle pipetting and incubated on ice for 30 min. Then the mixture was heat shocked at 42°C for 50 sec and placed back on ice for an additional 2 min. Then 0.98 ml of room temperature SOC medium (Table 3) was added and the cells were put in a shaker for constant shaking at 300 rpm for 1 hour. 100 µl of cells were then plated on a LB/amp$^+$/x-gal plate and the plate was incubated at 37°C for overnight.

II.10. Plasmid minipreparation:

Cells containing the plasmid were grown at 37°C overnight in 5 ml of LB containing the appropriate antibiotic. Next day, the overnight culture was centrifuged at 2,500 rpm for 15 min in a table centrifuge. The supernatant was removed, and the cell pellet was resuspended in 200 µl of GTEL buffer (Table 3) and kept at room temperature for 5 min. Then 400 µl of freshly made 0.2 M NaOH/1% SDS were added, mixed by inversion several times and kept on ice for 5 min. 200 µl of ice cold 10 M ammonium acetate were then added and the mixture was vortexed for 5 sec. and incubated again on ice for 10 min. Cell debris was removed by centrifugation in a microcentrifuge at room temperature for 5 min.

The supernatant containing the plasmid was precipitated by incubating with 0.6 volume of isopropanol for 10 min at room temperature. The plasmid was then recovered by centrifugation for 15 min at room temperature in a microcentrifuge and redissolved in 100 µl of TE buffer.

For restriction enzyme analysis, 5 µl were used and digested with the desired restriction enzyme in the presence of 100 µg/ml of RNsae A. For sequencing analysis, the DNA was denatured in 0.2 M NaOH/ 20 mM EDTA at 37°C for 30 min, then neutralized with 0.1 volume of 3 M sodium acetate pH 5.2 and precipitated with 3 volumes of 100% ethanol. After washing with 70% ethanol, the DNA pellet was redissolved in TE buffer.

II.11. Sequencing of double-stranded DNA template:

Sequenase, version 2.0 (USB, 13 unit/µl; 1 unit catalyzes the incorporation of 1 nmole of dNTPs into acid insoluble form in 1 min at 37°C using 5 µg of primered M13 mp18 DNA as template) was used in the sequencing. The protocol provided by USB was followed. Briefly, in a centrifuge tube, approximately 1 µg of plasmid DNA to be sequenced was combined with 0.5 pmole of sequencing primer in 1x SRB buffer (Table 3) in total volume of 10 µl. The mixture was warmed to 65°C for 2 min, then cooled slowly to room temperature over a period of 30 min or more to anneal the template and primer. To the annealed template-primer, DTT (to final concentration of 10 mM), 1x labeling mix (1.5 µM dGTP, 1.5 µM dCTP, 1.5 µM dTTP) and 0.5 µl of alpha-^{32}P dATP (at 10 µCi/µl , 10 mM; from NEN) were added. The labeling reaction was initiated by adding 2 µl of diluted

Sequenase version 2.0 (diluted 1:8 in ice-cold 1x EDB buffer) and incubated for 2-5 min at room temperature. When the labeling incubation was complete, 3.5 µl of labeled mix was removed and transferred to each of the 4 tubes which contained 2.5 µl of each of the ddXTP termination mix (80 µM ddNTPs of 4 kinds, plus 8 µM ddXTP in 50 mM NaCl, X=G, A, T or C) and prewarmed to 37°C for at least 1 min. The tubes were mixed, centrifuged and incubated at 37°C for 3-5 min. Then 4 µl of stop solution (95% Formamide, 20 mM EDTA, 0.05% Bromophenol Blue and 0.05% Xylene Cyanol FF) was added to each termination reaction, and the tubes were kept on ice. The samples were heated to 75-80°C for 2 min and loaded immediately on sequencing gel, which had been pre-run; 2-3 µl was used in each lane.

II.12. Ligation-mediated polymerase chain reaction:

Ligation-mediated polymerase chain reaction (abbreviated as LM-PCR) was used in attempt to map DNase I hypersensitive sites. Briefly:

1. Nuclei were cleaved with varying amounts of endonuclease.
2. Genomic DNA was isolated and purified from the cleaved nuclei.
3. A gene-specific primer was used to initiate synthesis of a complementary strand producing duplex DNA extending from the gene-specific primer to the cleavage point.
4. A synthetic-linker duplex was ligated to the blunt end of each newly synthesized duplex.
5. These duplexes were amplified by PCR using two primers;

one was the long strand of the linker, and the second was the gene-specific sequence.

6. The fragments were visualized by hybridization or by primer extension.

Linker preparation: Linkers were prepared by annealing in 250 mM Tris-HCl pH 7.7, a 25-mer (5'-GCGGTGACCCGGGAGATCTGAATTC) to an 11-mer (5'-GAATTCAGATC) by heating to 95°C for 3 min and gradually cooling to 4°C over a time period of 3 hours. Linkers at 20 pmole/µl were stored at -20°C and thawed on ice when needed.

The LM-PCR technique contained three basic steps: Sequenase reaction, ligation and PCR. The final products were visualized by the primer extension method:

Sequenase reaction: The first step was to anneal a gene-specific primer (primer 1) to cleaved genomic DNA (0.5-3 µg) in Sequenase buffer (40 mM Tris-HCl pH 7.7 and 50 mM NaCl) in 15 µl final volume. The DNA was denatured at 95°C for 5 min, annealed at 4 7°C for 30 min, and the sample was quick-cooled on ice to prevent formation of nonhomologous duplexes. The Mg^{2+}, DTT and dNTPs necessary for primer extension were added as a cocktail (7.5 µl) to a final concentration of 20 mM, 20 mM and 0.25 mM respectively. Sequenase version 2.0 was diluted 1:4 in ice cold TE pH 8 and 1.5 µl was added to the reaction mix. The mixture was incubated at 47°C for 15 min and cooled on ice. Then 6 µl of ice cold 0.3 M Tris-HCl pH 7.7 was added, the reaction was stopped by heating to 67°C for 15 min and returned to ice.

Ligation reaction: After heat denaturation, 45 µl of freshly made ligation mix containing 13.3 mM $MgCl_2$, 30 mM DTT, 1.66 mM ATP, 83.3 µg/ml BSA, 100 pmole linker and 3 units of T4 DNA ligase (3 units/µl; Promega) as final concentration was added. The reaction was incubated overnight at 17°C and stopped by heating at 70°C for 10 min. DNA was then precipitated by adding 9.4 µl of precipitation mix (8.4 µl 3 M NaOAc pH 5.2 and 1 µl Yeast tRNA at 10 mg/ml) and 220 µl of 100% ethanol and incubating on dry ice for 20 min. The DNA was then recovered by centrifugation at 4°C for 15 min in a microcentrifuge. The DNA pellet was washed once with 75% ethanol and dried in a Speedvac. The DNA was then redissolved in 50 µl of water.

Polymerase chain reaction: To the 50 µl of ligated sample, 50 µl of freshly prepared Taq DNA polymerase mix was added and mixed by pipetting. The final reaction condition was: 1x Taq Buffer A (Table 3), 10 mM each dNTPs, 10 pmole each of linker primer and a second gene-specific primer (primer 2) and 3 units of Taq polymerase (AmpliTaq, 3 u/µl; from Perkin-Elmer). The PCR reaction mix was overlaid by 90 µl of sterile mineral oil to prevent evaporation and the PCR was performed immediately in a thermal cycler. Twenty cycles were performed as follows: (1). 95°C for 1 min to denature the template; (2). 66°C for 2 min to anneal the primers; and (3). 7 6°C for 3 min to extend the primers. To assure complete extension of all primers, after the 20th cycle a final extension was performed at 76°C for 10 min.

Labeling of LM-PCR products by primer extension: A gene specific primer was end-labeled with T4 polynucleotide kinase as described in section II.4. After use of a spin-column to remove free labels, the labeled primer was diluted into buffer containing 1x Taq A buffer, 2.5 nmole each dNTPs and 0.2 unit of Taq polymerase in a total volume of 60 µl. This mixture is termed the "labeling mix". Then the LM-PCR reaction mixture was mixed with 10 µl of labeling mix and another 3 cycles of PCR were performed at 95°C for 1 min, 66°C for 2 min and 76-80°C for 3 min. The labeled samples were precipitated in 2.5 volumes of ethanol and recovered by centrifugation as before.

The labeled samples were redissolved in 5 µl of denaturing loading buffer, heated at 95°C for 3 min and cooled on ice. The denatured samples were then loaded onto a denaturing polyacrylamide gel (25 x 40 x 0.04 cm^3, 6 or 8% polyacrylamide, 7 M urea in 1x TBE buffer) and the gel was run at 25 mA for 2 hours. Following electrophoresis, the gel was dried under vacuum, exposed with or without intensifying screen to Hyperfilm (Amersham) for from overnight to 4 days, and developed.

II.13. DNA alkylation:

Preparation of aflatoxin B1 8, 9-dibromide Aflatoxin B_1 (10 µg/ml in dichloromethene) was mixed with equi molar concentration of bromine (in dichloromethene) at a variety of different levels (o.1 µg to 5 µg). The mixture was incubated for 5 min at room temperature and then dried completely under nitrogen in fume hood. AFB1-dibromide (AFB1-Br_2) was used immediately for DNA alkylation.

Preparation of the template: A 136 bp fragment, including 111 bp of exon 1 and adjacent 25 bp of 5' flanking sequence, from trout ki-*ras* gene was used for alkylation. This fragment was prepared by normal PCR synthesis. Briefly, 5-10 ng of plasmid, pTras-1 containing the trout ki-*ras* gene, was mixed with 100 µM of each dNTPs, 60 pmoles of the specific primers R01 and RHc37 and the 5 units of Taq polymerase (AmpliTaq, Perkin-Elmer) in the 1x Taq A buffer. PCR was performed for 40 cycles of 94°C for 1 min, 60°C for 30 sec and 76°C for 1 min. The 136 bp fragment was purified by gel purification. To end-label one strand, one of the two PCR primers (01 or RHc37) was kinased before it was applied to PCR synthesis. The purified labeled template was redissolved in 50 mM Tris.HCl pH 7.4 and 10 mM EDTA.

DNA alkylation: The labeled DNA template was added directly to freshly prepared AFB1-Br_2. Equivalent amounts of DNA were treated in parallel with dimethyl sulfate (DMS) or AFB1-Br_2 . DMS was added to a final concentration of 1-10 mM, and the reaction quenched after 30 min by addition of β-mecaptoethanol to 200 mM. Dried AFB1-Br_2 was dissolved in the labeled template in 50 µl of 50 mM Tris.HCl pH 7.4 and 10 mM EDTA by vigorous shaking. The final concentration of AFB1-Br_2 was at 2 µg/50 µl. After 30 min at room temperature, the reaction was quenched with β-mecaptoethanol as above. All samples (either DMS or AFB1-Br_2 treated) were extracted once with chloroform, precipitated with ethanol, recovered by centrifugation and then redissolved in 1 M freshly diluted piperidine. Samples were placed in a 90°C heating block for 30 min,

then dried in the SpeedVac. Each pellet was redissolved in 4 μl water. Then 4 μl of formamide dye loading buffer was added. Samples were heated at 80°C for about 2 min before loaded on a 6% sequencing gel. Roughly equivalent radioactive counts were loaded for DMS and AFB1-Br_2 cleaved samples. The gel was run, dried, and autoradiogrammed, and film was developed.

III. RESULTS:

III.1. Isolation of nuclei and nuclease digestion:

Electrophoretic analysis of DNA fragments generated by the digestion of nuclei with exogenous nuclease has provided important information on the structural organization of chromatin; therefore this technique has been used widely to study the chromatin structure of various particular DNA sequences. The requirement for nuclear preparations of good quality is very important for the accurate interpretation of the chromatin structure *in vivo*. There are generally two concerns when nuclei are isolated from a tissue-physical instability of the nuclei and the presence of endogenous nuclease activity. Stability of the nuclei is often markedly influenced by the buffer conditions. Lysis of nuclei is a common problem when non-optimal buffer conditions are used. The second problem is that an endogenous nuclease activity is sometimes copurified with the nuclei (Vanderbilt *et al.*, 1982). This activity can result in rapid digestion of DNA during nuclear preparation or incubation.

In this study, several methods have been tried in attempts to isolate nuclei of good quality from trout liver tissue. The first two methods were found to be inappropriate, for the lysis of nuclei occurred during the preparation. The third method was successfully used.

III.1.1. Isolation of nuclei:

Method I was from Bailey *et al.* (1980). It had been originally

used to purify trout liver nuclei for studies of distribution of aflatoxins in chromatin. Briefly, fresh trout liver was excised, perfused and homogenized with homogenization buffer (60 mM KCl, 15 mM NaCl, 0.15 mM Spermidine, 15 mM β-MeOH, 15 mM Tris-HCl pH 7.4, 0.2 mM PMSF and 50 mM $NaHSO_2$ pH 7.2). The liver homogenate was then filtered through cheese cloth followed by centrifugation at 4,000 rpm for 10 min. The pelleted nuclei were resuspended in homogenization buffer and washed with the same buffer. This was followed by washing with homogenization buffer containing 0.15% Non-idet NP-40 to break the outer nuclear membrane. The nuclei were then pelleted by centrifugation. Unfortunately, I found that when the detergent was added, the nuclei were lysed immediately, yielding a very sticky and insoluble mass of chromatin. Some modifications of the method were tried: for example, before pelleting the nuclei and breaking the nuclear membrane, the homogenate was centrifuged at low speed to clear cell debris, or the concentration of detergent was changed. But the preparation was still unsuccessful.

I then attempted Method II, as described by Blobel & Potter (1966), who successfully modified the procedure that generates almost quantitative separation of clean nuclei from a homogenate of rat liver. This method involved a different homogenization buffer - TKM (10 mM Tris-HCl pH 7.5, 5 mM KCl and 1 mM $MgCl_2$) and a step sucrose gradient with a cushion of 1.6 M/2.3 M (tissue homogenate in 1.6 M sucrose laid on top of the 2.3 M sucrose buffer) to remove cytoplasmic contamination. In addition, a different detergent, Triton X-100 at concentration of 0.05% to 2%, was also used to break the

outer nuclear membrane and release the attached ribosomes. By following this procedure, the trout liver nuclei was isolated by 1). ultracentrifugation of homogenate with 1.6 M/2.3 M sucrose cushion, and 2). washing the nuclear pellet with 0.5% Triton X-100. Again after addition of the detergent, the nuclei were also lysed immediately, although less completely. Nevertheless, the yield was dramatically decreased.

Since attempts at preparation of nuclei by above two methods were unsuccessful, probably due to the use of detergent, a third method was tried, which was successful for this preparation in terms of the stability of nuclei. This procedure, which is that of Bres & Eales (1988), is described in the schematic diagram shown in Figure III.1.1. As shown in the diagram, the procedure is very simple, including only two rounds of centrifugation. Also, it utilizes a similar way to remove cytoplasmic contamination as does method II, by using a sucrose gradient cushion. However, no detergent is used in this procedure. The original paper did not mentioned the reason why the detergent was not used, I presume that it was because trout liver nuclei seem very labile to any kinds of detergent. So it may be concluded that the preparation should avoid the use of detergent. It is clear that the purity of the nuclei could be poorer without detergent treatment, since chromosomal proteins might be contaminated by nuclear membrane proteins and lipids. However, the integrity of nuclei is more important for studies such as those of chromatin structure because the maintenance of physical structure *in vivo* during the preparation is absolutely required. In this study, by using method III, nuclei of reasonably high purity

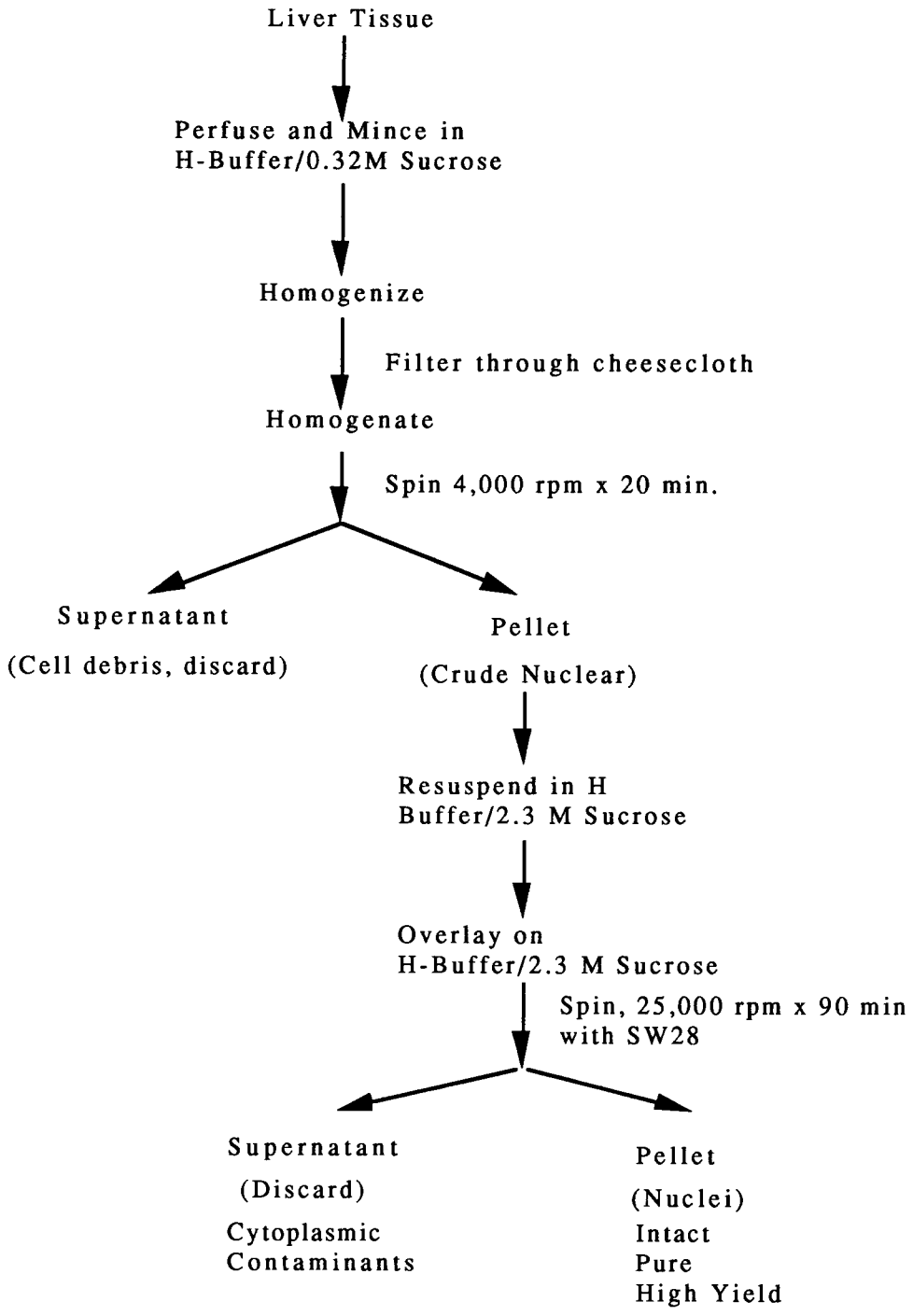


Figure III.1.1 Schematic diagram of the final procedure for the preparation of the trout liver nuclei.

were isolated from trout liver (normal or tumor tissue), without nuclear disintegration. Between 5 and 10 mg of nuclei could be isolated from about 10 g of fresh liver tissue.

III.1.2. Endonuclease digestion:

Although the quality of the nuclei was satisfaction as judged by direct visualization, it had also to be determined at the molecular level. This was accomplished by digesting the nuclei with exogenous endonuclease and examining the DNA products on gel electrophoresis. The nuclei were digested with MNase at 46 u/mg/ml for different time periods. The digested DNA was isolated from the chromatin and resolved on 1.8% agarose gel electrophoresis. The result is shown in Figure III.1.2. Lane M shows the DNA markers. Lanes 1 to 10 contain the DNA from MNase treated nuclei digested for times of 0, 30", 1', 2', 3', 5', 7', 10', 15' and 20' respectively. As shown in the figure, at 0 time and very short length of digestion, most DNA was in the high molecular range. This indicates that little endogenous nuclease activity was present in the preparation. As the digestion went on, chromatin was digested down to the oligonucleosome and mononucleosome level. This result indicated that chromatin structure was intact in these nuclei.

The quality of the nuclei preparation was also checked by histone gels, as shown in Figure III.1.3. Aliquots of a suspension of nuclei were boiled for 5 min in SDS/gel loading buffer and loaded on a histone gel (12% separating gel, 5% stacking gel). The electrophoresis was stopped after 1.5 hours at 100 volts. Lane M contained the protein markers. Lanes 1 to 4 contained the histones

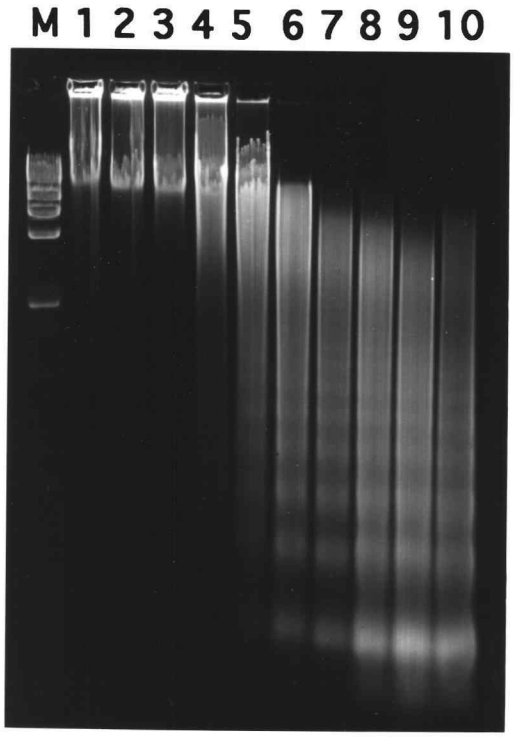


Figure III.1.2 Micrococcal nuclease digestion pattern of trout liver nuclei. Freshly prepared trout liver nuclei at 1 mg/ml were digested with micrococcal nuclease at 46 u/mg/ml at 37°C. An aliquot of digest was withdrawn at different times: 30 sec (lane 2), 1 min (lane 3), 2 min (lane 4), 3 min (lane 5), 5 min (lane 6), 7 min (lane 7), 10 min (lane 8), 15 min (lane 9) and 20 min (lane 10). DNA was then isolated and electrophoresed on 1.8% agarose gel. Lane 1 exhibits the undigested control. M shows the DNA markers: lDNA digested by Hind III.

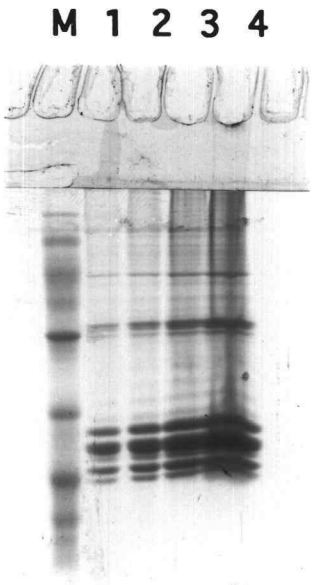


Figure III.1.3 Histones from trout liver nuclei. 20 µl of freshly made trout liver nuclei at 1 mg/ml was boiled for 5 min in 20 µl of 2x SDS loading buffer. Then, 2 µl (lane 1), 5 µl (lane 2), 10 µl (lane 3) or 20 µl (lane 4) was loaded on a mini SDS-PAGE gel followed by electrophoresis at 100 volts for 1.5 hours. The gel was coommassie stained. M shows the protein markers.

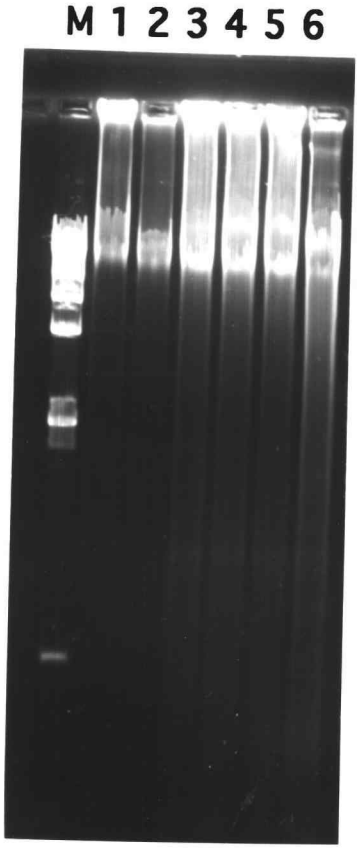


Figure III.1.4 Autodigestion of trout liver nuclei. Trout liver nuclei in nuclear digestion buffer were incubated at 37°C for different lengths of time: zero (lane 1), 0.5 h (lane 2), 1 h (lane 3), 2 h (lane 4), 3 h (lane 5) and 4 h (lane 6). The DNA was isolated and resolved on 1.8% agarose gel electrophoresis. M depicts DNA markers.

from the same nuclei suspension at different loading of proteins. All histones were present in good quality, with very little or no evidence for degradation.

III.1.3. Test for endogenous nuclease activity:

Although preliminary tests had indicated little, I wished to test more rigorously if there was any endogenous nuclease activity present in the trout liver nuclei. Freshly prepared nuclei were resuspended in nuclease digestion buffer and incubated at 37°C for different time periods, ranging from 0.5 hour to 4 hours. The DNA was isolated and resolved on 1.8% agarose gel electrophoresis. The result is shown in Figure III.1.4. Lane M contains the DNA markers. Lanes 1 to 4 were from autodigestion at 0.5 h, 1 h, 2 h, 3 h and 4 h respectively. Virtually no low molecular weight DNA was observed, indicating that little or no endogenous nuclease activity was present in this nuclear preparation.

In summary, trout liver nuclei were successfully isolated by the method III. The results of MNase digestion, histone gels, and tests for autodigestion of the nuclei indicated that the chromatin structure was maintained during the preparation.

III.2. Characteristics of the 5'-flanking region of the trout ki-*ras* gene:

The ki-*ras* gene has been identified to be present in rainbow trout, and the sequence homology with human ki-*ras* gene found to be about 80% (Mangold *et al.*, 1991). In addition, this gene was also found to be responsible for AFB1 induction of liver tumor in trout (Chang *et al.*, 1991). We became interested, therefore, in the features of the sequence of this gene at its 5' flanking region, as compared to other *ras* genes, and in chromatin structure in this region. In particular, we were concerned with how this structure might influence *ras* mutation by carcinogens. In order to be able to examine this sequence, as well as the region surrounding codon 12, the regulatory and hypersensitive sites, it was necessary to sequence the 5' flanking region of the trout *ras* gene.

This section first describes the strategy by which the trout ki-*ras* gene was cloned and sequenced by PCR technique; the results of the sequencing are then presented.

III.2.1. Experimental strategy: Inverse PCR technique:

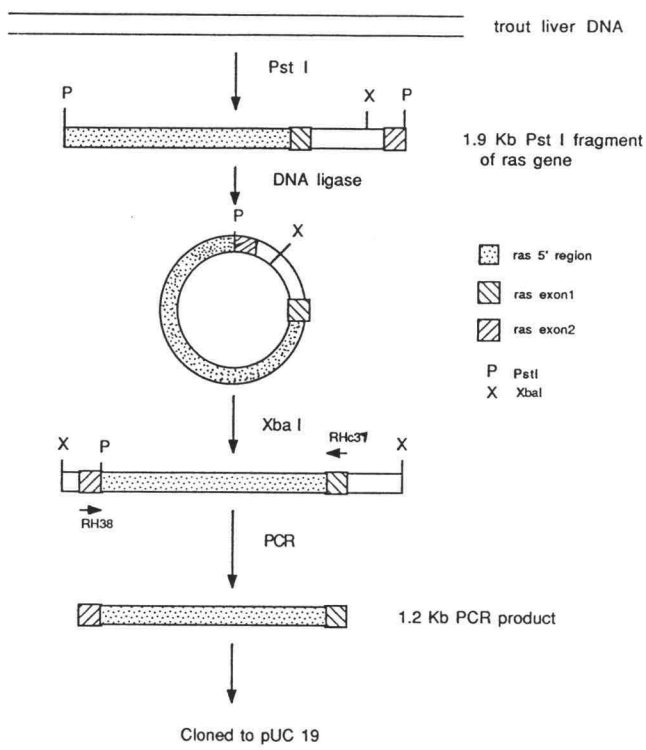
PCR, as reviewed in the first chapter, is an *in vitro* method for nucleic acid synthesis by which a particular segment of DNA can be specifically replicated and greatly amplified. It requires two oligonucleotide primers that flank the DNA fragment to be amplified. However, this requirement has greatly limited its use only to targets whose sequence is known *a priori*, whereas regions that lie outside the boundaries of known sequences cannot be

amplified in this way. This problem has been overcome by several modifications of the method. One modification is called "inverse PCR". This procedure involves cutting the DNA with restriction endonucleases and ligating the resulting free ends under conditions that favor circularization of the DNA. As a result of this manipulation, oligonucleotides complementary to the 5' and 3' ends of the known sequences can now be used to amplify the unknown sequence of interest located at the 5' and/or 3' ends. In other words, this technique first rearranges the gene in such a way that the region of interest is positioned between two known sequences. These can then be used to define primers for normal PCR (See Figure III.2.1).

In this study, we are interested in the 5' flanking region and exon 1 of trout ki-*ras* gene. But only the sequence of exon 1, the first intron and part of exon 2 were known at the beginning of this research. In order to determine the sequence around the 5' region, the ki-*ras* gene containing this region needs to first be pulled out from the total genome and then be amplified. The technique of inverse PCR seemed appropriate for this study. The scheme of the technique is shown in panel A of Figure III.2.2. The region defined by two arrows above the trout liver genome represent the ki-*ras* gene. The enlarged segment below is bounded by Pst I sites (P), and contains exon 1 and part of exon 2 plus a 5' flanking segment. The dotted portion represents the 5' flanking sequence. Two boxes with diagonal lines represents exon 1 and exon 2. The open box represents the intron. In the first step, the trout liver DNA was digested with Pst I to generate the 1.9 kb fragment shown. The

Figure III.2.1 The synthesis of trout ki-*ras* gene by inverse polymerase chain reaction. In panel A, regions defined by the arrows on top of the genome represent the ki-*ras* gene (not to scale). The detailed procedure is described in the text. Panel B represents the amplified 1.2 kb product indicated by arrow (cont. on page 57).

A.



B.

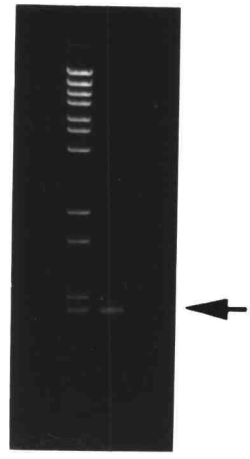


Figure III.2.1 The synthesis of trout ki-*ras* gene by inverse polymerase chain reaction.

total digest was electrophoresed on a preparative agarose gel, the gel was briefly stained with ethidium bromide, and the positions of the DNA and markers were revealed by UV illumination. The DNA in the 1.5 to 3 kb size range was excised from the gel and DNA was recovered by electroelution. DNA was then ligated using T4 DNA ligase, by which means the 1.9 kb fragment of the trout ki-*ras* gene was circularized. To rearrange the gene, the circular form of the gene was digested with Xba I (Note that there is one Xba I site located within the intron). After the digestion, the gene was now rearranged so that its 5' flanking region was positioned between its exon 2 (in the upstream site) and its exon 1 (in the downstream site). In the second step, the rearranged gene was amplified by polymerase chain reaction with Taq polymerase using two primers RH38 and RHc37. The sequence of primer RH38 was from the exon 2 between codon 38 to 44, primer RHc37 from the complementary strand of exon 1 from codon 37 to 31 (Table 4). These two primers defined the ends of the PCR product, one was the 5' end of the exon 2 (defined by RH38), the other was the 3' end of the exon 1 (defined by RHc37). The regions outside of these boundaries (the intron in particular) were not involved in the PCR process. Therefore after 40 cycles of PCR, a 1.2 kb fragment (not 1.9 kb) of the rearranged gene was produced. The product was gel-purified, recovered by electroelution and the purity was checked on agarose gel, shown in panel B of Figure III.2.1.

III.2.2. Cloning of PCR synthesized trout ki-*ras* gene into pUC19:

To sequence the 5' flanking region, the 1.2 kb product could have been directly sequenced in the PCR reaction with Taq polymerase (Ruano and Kidd, 1991; Innis *et al.*, 1988; Engelke *et al.*, 1988). Alternatively, the product could be cloned into a plasmid and then sequenced with T7 DNA polymerase using the standard dideoxynucleotide chain termination method (Scharf *et al.*, 1986). There are several reasons for preferring the cloning approach. First, once the fragment is cloned into a plasmid, the target fragment for sequencing can be easily made in good quality and purity at one time through large plasmid preparation. But direct sequencing needs PCR synthesis for every experiment, and is therefore tedious and more expensive. Second, direct sequencing of templates by PCR is occasionally interfered with by impurities in the template, leading to results which could be misinterpreted. On the other hand, sequencing of the plasmid by standard method will generate a very clean, unambiguous pattern. Third, universal M13 sequencing primers can be used if the sequence has been cloned into pUC system. Based on above consideration, the 1.2 kb PCR product from inverse PCR was cloned into the plasmid pUC 19 and then transformed to DH5α cells. The plasmid containing the insert was screened by plating the transformant on LB plates with inducer IPTG and indicator X-gal. The blue colonies, which were expected to have the insert, were analyzed by restriction enzyme digestion. One colony, called cp 42, contained the plasmid with the insert of the

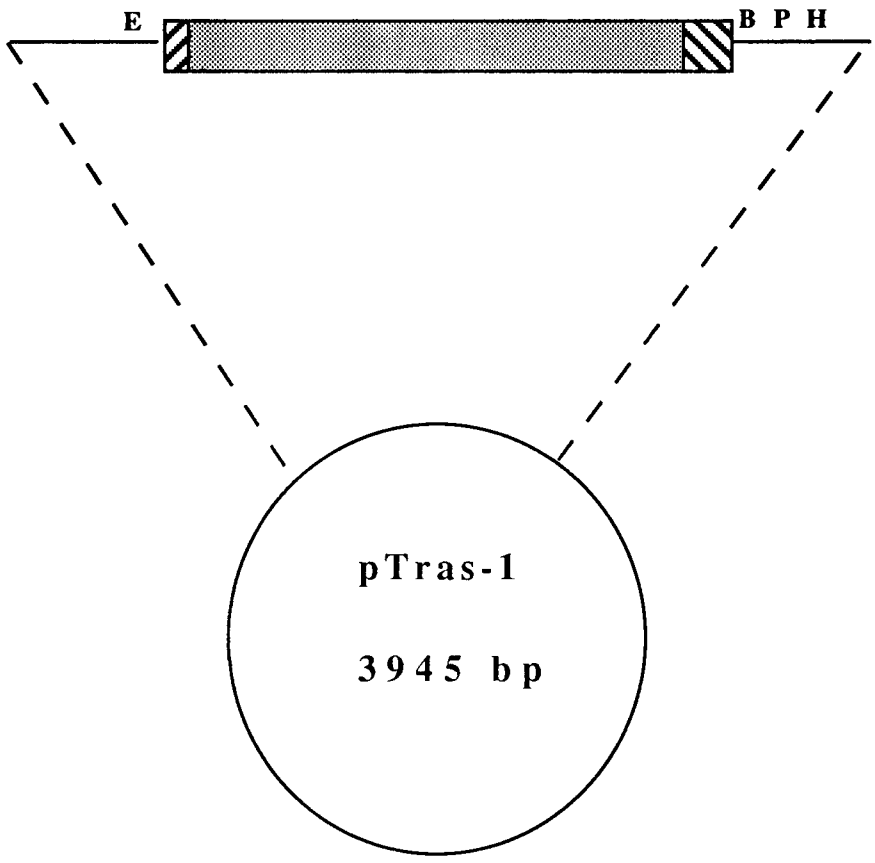


Figure III.2.2 The structure of pTras-1 construct. The 1.2 kb insert was cloned into the Sma I site of pUC 19. The 5' end of the insert is flanked by an EcoR I site, the 3' end by BamH I, Hind III and Pst I sites. The total size of the plasmid is 3,945 bp.

right size. This plasmid was called pTras-1, and its structure is shown in Figure III.2.2.

III.2.3. Sequencing of the 5' portion of the trout ki-*ras* gene:

The enzymatic chain-termination sequencing technique allows rapid sequencing from a single-stranded DNA template like M13 (Sanger *et al.*, 1977). However, sequencing using single-stranded vectors can only be performed in one direction. Furthermore, DNA inserted into these single-stranded vectors is subject to recombinational and excision events more frequently than if it is inserted into plasmids (Wang and Sodja, 1991). Therefore, in this study, the 1.2 Kb fragment from inverse PCR amplification was cloned into plasmid pUC 19 instead of M13 vector. After ligation, the insert DNA is flanked by T7 polymerase promoters enabling sequencing of both strands from a single clone, thus saving time.

One problem with the double-stranded plasmid sequencing, however, is that the readable range of sequence is less than in the case of single-strand sequencing. In our gel system, only about 250 to 300 bp could be read in one gel if double loading is used. In order to get a maximum length of sequence information, deletion mutation of the fragment could be made. But this method is very inefficient, tedious and time-consuming. Since now the oligonucleotide synthesis is both easy and cheap, an alternative way was to have new primers made, based on the previous sequencing results. These could then be used for sequencing of the next range. The schematic of this sequencing strategy is shown in Figure III.2.3.

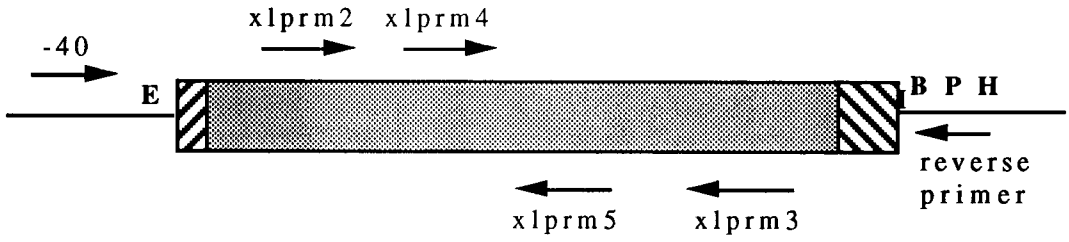


Figure III.2.3 The schematic strategy of sequencing of pTras-1. The 1.2 kb insert and its flanking polylinker sites are shown. Reverse primer and -40 are universal M13 sequencing primers. Primers xlprm2-xlprm5 are trout ki-*ras* specific primers.

First, two universal M13 sequencing primers were used to begin the sequencing of plasmid pTras-1 from both directions. The sequence information initially obtained included the exon 2, exon 1 and some of the 5' flanking region (as well as some of the polylinker sequence from the vector). Figure III.2.4 shows the sequence of both exons represented by the underlined region. In 1991, Mangold *et al* (1991) reported for the first time that ki-*ras* gene was expressed in normal trout liver. By using the PCR technique, they synthesized a 800 bp fragment derived from a portion of the genomic sequence, and identified that this fragment is the ki-*ras* gene. The 800 bp product includes 111 bp of exon 1, 568 bp of intron, and 126 bp of partial of exon 2. So the exon sequence results from this study were compared with results of Mangold *et al.* and it was found that the homology in the exon sequences was almost 100% except for three base differences present in the positions of +6, +12 and +18 of the exon 1, which includes G, T, and T respectively in Mangold *et al.*'s sequence, but A, C and G in our sequence. One explanation could be that our sequence represents another allele or another gene copy. When the difference was compared at the amino acid level, it was found that these base differences correspond to degenerate codons, so the corresponding amino acids they code are not changed. Another simple explanation is that Mangold et al.' sequence was generated with PCR using a rat 5' primer covering this region, suggesting that these base differences may actually represent the difference in sequence between rat and trout.

To continue the sequencing in the 5' flanking region (Figure

III.2.3), two new primers, called xlprm2 and xlprm3, were designed from the previous sequencing result, and used to sequence more of the 5' flanking region, again from both directions. This generated a new set of sequences. By repeating this procedure twice, the whole 5' flanking region was sequenced. In total, 4 specific primers, xlprm2 to xlprm5 (Table 4), were used. The summary of the sequence of trout ki-*ras* gene, expressed in the normal order rather than in the order of plasmid construct, is shown in Figure III.2.4. As shown in the figure, +1 represented the first codon ATG, upstream from this point is designed minus and downstream plus. The underlined regions represent the exon sequences. This figure includes 111 bp of exon 1 sequence, as labeled from +1 to +111; 568 bp of intron as labeled from +112 to +679 (this sequence was cited from the result of Mangold *et al.*, for the purpose of integrity of the sequence); and 126 bp of exon 2 sequence labeled from +680 to +806. The region between -1080 to -1 represents the 5' flanking region immediately close to exon 1. Analysis for functional elements in the 5' flanking region have revealed some interesting features which will be described in the Discussion section.

III.2.4. The target *ras* gene represents the expressed gene:

Given the ancestral tetraploid origin and ongoing diploidization of salmonids, it is believed that trout will exhibit a correspondingly large number of duplicated and diverging *ras* genes and psuedogenes, both cryptic and expressed. It is important for us to know that the ki-*ras* gene we studied represented the expressed one, because the expressed gene is the one for mediating the normal

Figure III.2.4 The sequence of the 5' flanking region and 5' portion of the trout ki-*ras* gene. The underlined regions represent the exon sequences (all of exon 1 and part of exon 2). The point labels +1 represent the start of the first coding ATG. Upstream from this point is labeled as minus, downstream from it is labeled as plus. In total 1080 bp in the 5' flanking region have been sequenced by the dideoxynucleotide chain termination method (cont. on page 66).

-1080 GAGCAGTGTG TGCTTCCTAG TTGTTTGTCA ATTTTGGGAC CATTATGTAA CATCATAGAT
-1020 CTGTGTCTAC TGCCTGAGGC CAGAACCCAT TGTGTAGCTG TCACTCGCAA CTATCTTTTG
-0960 AACGTGTGTT GTTACAGTTA GAGAGACATT TCAGTTGGAG GATCATTTAG TAGCAACATT
-0900 TGTGGGTATG ATATTTGGGG GAATTGTAAA CAGCAATCTT CTGTTCAAAC AATGTTAAAT
-0840 AATGATGACA TGTCATTATT TAATCAGGGG GAACAGTGAT TGTGCCAATT CAAAGTAATA
-0780 TAATAGTCAA CAGACCTGAT AGGGCAGGCA GCTTATATAT ATCTTTCCAG AATTCCAGTC
-0720 TCTTTCATAG TTTTCCATAT CTTCTCCTTG TTGATAATCC CATTTAAGAA TGATAGTAGA
-0660 GTATAATATC AAATGTTATT GGTCACATAC ACATGGTTAG CAGATGTTAA TGCGAGTGTA
-0600 GTGAAATGCT TGTGCTTCTA GTTCCGACTA TGCCGTAATA TCTAGCAAGT AATCTAACAT
-0540 TCACAACAAC TACCTTATAC ACAGAAATGT AAAGGGCCGA ATAAGAATAG GTACATATAA
-0480 ATATATCGGT GAGCGATGGT GTGCGGCATA ATATAATAAT ATATACACTC ATCCCATTTT
-0420 CTGTAAAATT GCAGGATGAA ATTGGATAGA ATGTAAACAT TCTTGTATCC ACAGAGATCT
-0360 GTTATTGTAT ACATTCCCTA CAGCTCTGGA GAACATAGGA GAGAAATTGT GCAAATTAAA
-0300 TGCAATCTGG AGACATTCGG GACAAAATAC AATTTTGTAT TTGCACTTAA GTTTAGTTTA
-0240 AGGAATTAAT ACAACTTGAC CAGATGGCTA TGGAGAAGGG GAGAAATTCT TCATATCCTC
-0180 AGCTTCAAAC TCACTATGTG TGTGTAGGGT TGGAACATTT CAGAAACTTT CCAGAAATGT
-0120 CCAGGTTTTG GAGATCCCAG TTGAAGGATT GTTGTGTCGA TAGTTAGTGT GTAGCTCACA
-0060 AAAGGTTTTA TATATTCACT GACCTTTATT TCTGTGTGCT TTCTCTCCCA CAGGTGAAGG
+1 ATGACAGAAT ACAAGCTGGT GGTGGTGGGG GCAGGAGGTG TGGGCAAGAG CGCGCTCACC
+61 ATCCAGCTCA TTCAGAACCA CTTTGTGGAT GAATATGACC CCACCATCGA GGTAATAAGT
+121 AGTTTCAGTA CACCCTTGTC TCTTACTTCA GCAAGGCATG CCTGGGAATC GGGAGACAAA
+181 AATCTTTTAG AAATAAAGTT GTGTAAACAG GTTTCGATCA TTTCTGATCA TTTTCTGTTA
+241 TTTGAATGCA TGAATATTAT AATATCATAA ACATGGTGTC TTTCAGTCAG TTTAATTTTG
+301 TTGGTTGATG AAAGATTACT ATGTGATTAA ATGCTTTAGT AGTTTCACAT GTCAGATATT
+361 ATTACTCAAT TATACTTTCT GAAAGCTATC AATTTTATGA CTGCTTGAAG ATTGAGTGCA
+421 TTTTTCCCTT GGTGCACATA CAGTCGCTCT TAGTCATCCC ATCCATTCTC TCTGATTCTA
+481 ATATTATTGA ATTTCCTCAG CGACTTTATT AGATGAGCAT TCTATGACCA TATAGTTAGG
+541 TCTAGACATT CTGATTCCTA GTCAGGTAGA CTGGACACCT ATGAATATGT GATCATTCAT
+601 AAGTGTCCAG TCTACCTGAT TACCGCTAAG GTGCCCTCCT GACCCAACCC ATCCCTCTCT
+661 TTCTCTCTCC CTCCTTCAGG ACTCGTACAG GAAGCAGGTG GTGATTGATG GGGAGACATG
+721 TCTGCTGGAC ATCCTGGACA CTGCAGGTCA GGAGGAGTAC AGCGCCATGA GGGACCAATA
+781 CATGAGGACA GGGGAGGGCT TCCTC

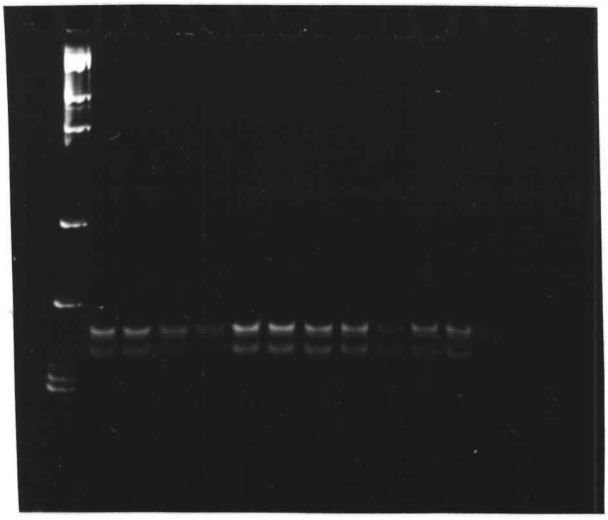
Figure III.2.4 The sequence of the 5' flanking region and part of the 5' portion of the trout ki-*ras* gene.

cellular function, and therefore the one responsible for the abnormality if mutated.

This question was approached by a technique called "mismatch" PCR in the following way: an oligo primer from the 5' flanking region (e.g. primer a2, see Table 4) was used as a 5' primer, in conjunction with a 3'-primer RHc37, to synthesize DNA from trout liver genome from normal tissues, as well as tumor tissues, which had been identified by sequencing to contain a G to T transversion at codon 12 of the ki-*ras* gene. This would generate a specific product at size of 216 bp (defined by primer a2 and RHc37). Then the first round PCR products were reamplified with the same 3' primer RHc37 and a 5' "mismatch" primer, 12(G) or 12(T). The primer 12(G) contains G at the codon 12 region so it will pick up the normal *ras* sequence (*ras* gene from normal liver DNA or the normal allele from tumor DNA). The primer 12(T) contains a mutated base T at the codon 12 region so it will pick up only the mutated *ras* sequence from tumor DNA. To avoid cross-priming (faulty priming of the normal allele sequence with 12(T)), the PCR conditions used for reamplification were very stringent, e.g. very low concentration of dNTPs to increase the accuracy of incorporation; fast timing and high temperature of PCR. If a specific product at size of 97 bp could be reamplified with 12(T) and RHc37 only from the tumor DNA, it would be a strong indication that the primer used for PCR amplification from tumor tissue (a2) represented the expressed sequence, and therefore strongly support the idea that the sequence from which primer a2 was derived did represent the expressed one. The data is shown in Figure III.2.5. Panel A represents the

Figure III.2.5 Reamplification of trout liver DNA with 12 mismatch primer. The trout liver DNA either from normal or tumor tissues were amplified by PCR with 5' primer a2 and RHc37. Then 1 ul from first round PCR product with 1 to 50 dilution was reamplified with RHc37.and 12 mismatch primer 12(G) (panel A), or 12(T) (panel B). Lane 1 is the negative control called L2. Lanes 2 to 6 are the positive controls, L7, L23, L47, YJ11, and YJ5 respectively. Lanes 7 to 11 are other negative controls obtained from normal DNA (cont. on page 69).

A. M 1 2 3 4 5 6 7 8 9 10 11



B. M 1 2 3 4 5 6 7 8 9 10 11

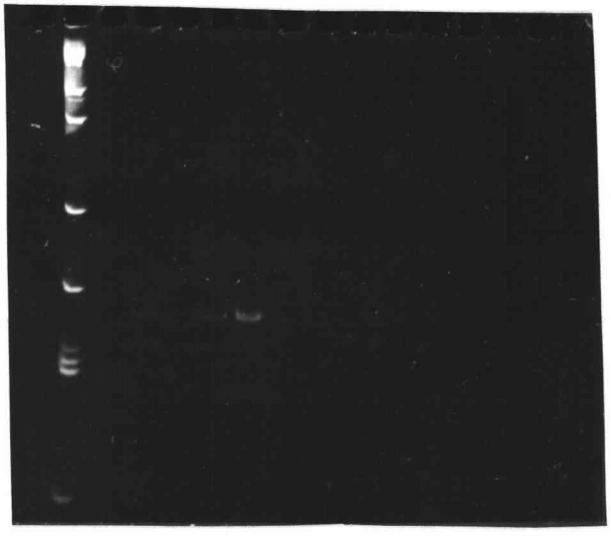


Figure III.2.5 Reamplification of trout liver DNA with 12 mismatch primer.

reamplification products with primer 12(G) and RHc37, and the panel B reamplification products with 12(T) and RHc37. Lane 1 is the negative control called L2. Lanes 2 to 6 are the positive controls, L7, L23, L47, YJ11, and YJ5 respectively. All these controls were verified in sequence. Lanes 7 to 11 are other negative controls which were not verified by sequencing. It is clearly shown that the reamplification products were obtained in every cases with 12(G) and RHc37 as expected (panel A). However, the specific products were observed only in the positive controls when 12(T) and RHc37 were used (panel B). This experiment strongly suggested that the ki-*ras* gene we studied was the expressed gene, although the last and absolute proof would be through cloning and sequencing of these specific PCR products..

To summarize this section: a 1.2 Kb fragment of surrounding the 5' end of the trout ki-*ras* gene was synthesized by inverse PCR, cloned to pUC 19 and sequenced by the dideoxynucleotide chain termination method. In total 1080 bp of sequence in the 5' flanking region was determined. By "mismatch" PCR technique, the sequences obtained in these studies have been proven to be the expressed gene.

III.3. Analysis of the chromatin structure in the 5' region of trout ki-*ras* gene:

DNase I hypersensitive sites usually correspond to defined gaps in the nucleosomal structure of genes; they may be found in 5' flanking regions, introns, and occasionally in exons (Eissenberg *et al.*, 1985; Reeves, 1984; Yaniv *et al.*, 1986). To study the hypersensitive sites of a gene in chromatin, nuclei are isolated from tissue and digested very lightly with DNase I. After digestion, genomic DNA is purified from proteins, and digested with a restriction enzyme which is located outside of the regions to be studied. A classical way to visualize the digestion pattern is called the indirect-end labeling method, developed by Carl Wu in 1980 (Wu, 1980), in which a DNA probe abutting the restriction site is used to probe a single strand gel of the digest, using Southern blotting. However, for a single-copy gene, it is very difficult to detect DNase I sensitive sites by this method, and resolution is generally poor, since the signal is so low compared to background.

In the past several years, the PCR technique has been widely used in molecular biology. Many modifications of this technique have been developed to study different questions, or to solve particular problems. The most significant modification, I think, is ligation-mediated PCR (Mueller and Wold, 1991). This technique provides the opportunity to study certain aspects of genome structure *in vivo* much more easily, such as direct genomic DNA sequencing (Pfeifer *et al.*, 1990), and footprinting (Mueller and Wold, 1991; Pfeifer and Riggs, 1991), even for single-copy genes. In

this work I have utilized an extension of this method to search *in vivo* for DNase I hypersensitive sites in or around the codon 12 region of the trout ki-*ras* gene.

III.3.1. The ligation-mediated PCR:

PCR, as described above, can be used to exponentially amplify any segment of DNA located between two specified primer hybridization sites. However, in its original form, it cannot be used to amplify certain kinds of fragments produced by nuclease cleavage of a genome. For example, the members of a sequence ladder, which have one fixed end defined by a restriction site and variable opposite ends generated by chemical cleavage or enzyme digestion, cannot be amplified by this method, for only one end is known. This problem has now been overcome by the "ligation-mediated polymerase chain reaction", called LM-PCR. In this technique, a gene-specific 5' primer is used in an extension reaction to synthesize the sequence of interest on a single-strand template, then a linker is attached at the newly created 3' blunt end and amplification is performed through 15-18 cycles of conventional PCR using a linker primer and a second gene specific 5' primer. A schematic summary of LM-PCR is presented in Figure III.3.1. As shown in the figure, LM-PCR is composed of three basic steps: (a). first strand synthesis by primer extension, (b). ligation reaction, and (c). PCR. In the first step, the starting gnomic DNA is randomly cleaved by a chemical reagent or a non-specific enzyme. It is then denatured by heat and a gene-specific primer (primer 1) is used that will anneal somewhere to the 5' side of the region of interest.

Figure III.3.1 The ligation-mediated polymerase chain reaction (cited from Mueller and Wold, 1991. LM-PCR is composed of three basic steps: (a). first strand synthesis by primer extension, (b). ligation reaction, and (c). PCR. The products are visualized by labeled primer extension. Details are described in the text (cont. on page 74).

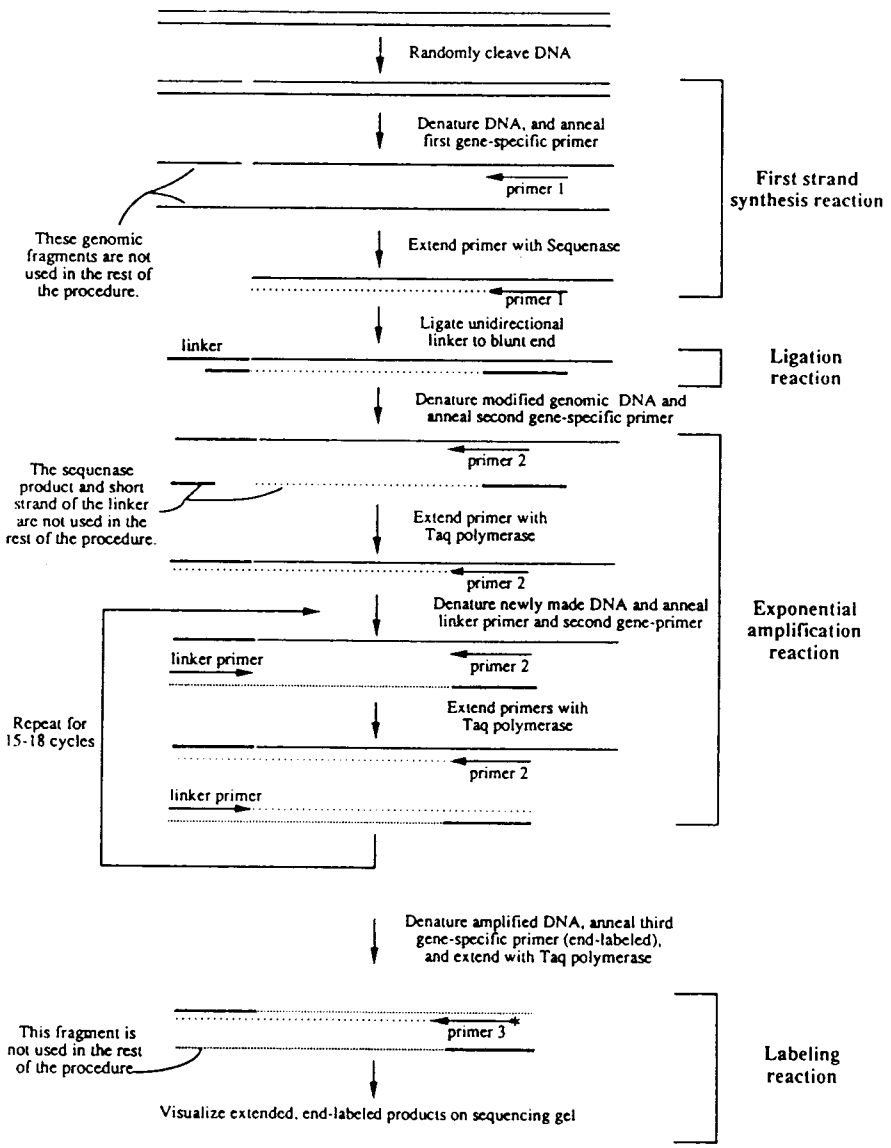


Figure III.3.1 The ligation-mediated polymerase chain reaction.

Extension of this primer creates a blunt double-stranded end at the variable cleavage site of each member of the "ladder" of sequences. During the second step, T4 ligase catalyzes the addition of an unidirectional blunt-end linker to each of these sites. Through this reaction, the 5'-phosphate of the genomic DNA is ligated to the 3'-hydroxyl of the longer strand of the staggered linker. Because the shorter strand of the linker lacks a 5'-phosphate, it cannot be ligated to the extension product and, following ligation, does not participate in the rest of the procedure. In the third step, the ligation product is denatured and a second gene-specific primer (primer 2) is annealed to the genomic DNA and extension of the primer proceeds past the cleavage sites and through the attached linker sequence. This set of modified sequences are all potential PCR substrates and after 15-18 cycles of PCR can be amplified about 10^4 fold.

LM-PCR products can be visualized by Southern hybridization (Pfiefer *et al.*, 1989) or by labeled primer extension (Mueller & Wold, 1989). The appearance of the ladder will be that of the original ladder of fragment defined by the 5' primer and the set of variable sites, except that each member will be uniformly longer, due to the attached linker.

Nested primers:

As mentioned above, PCR is a very efficient way to synthesize DNA *in vitro*. This efficiency has the consequence that very often, molecules with sequences similar to, but not necessarily identical to the primer can also be significantly amplified during the reaction. This will generate a background and reduce the synthesis of the

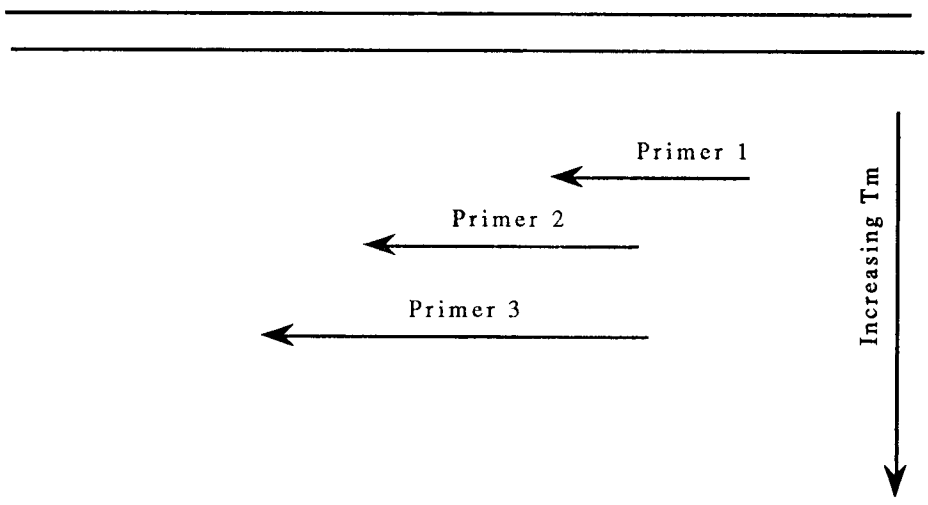


Figure III.3.2 The arrangement for the nested gene-specific primers. Primer 1 is used in the first-strand synthesis reaction, primer 2 is used in the amplification reaction, and an end-labeled primer 3 is used in the labeling reaction. As shown on the right, the T_m should increase in the order 1<2<3.

desired specific product, especially when the amount of starting template desired to be amplified is very low (for example, a single-copy gene in a large genome). This "primer mispriming" phenomenon can be a critical limitation to the whole procedure. To reduce the incidence of this to the lowest possible level, a nested set of gene-specific primers can be used (Mueller and Wold, 1991). Figure III.3.2 illustrates the primer arrangement. Primer 1 is used for the Sequenase reaction in the first step, to synthesize a blunt-ended duplex which is the substrate for the ligation in the second step. Primer 2 is used for PCR amplification during the third step, used in conjunction with the linker primer. To visualize the LM-PCR product by primer extension method, primer 3 is used. It is kinased first at its 5'-end and applied to the PCR reaction. These primers should be positioned so that the extending end of each is 3' or internal to that of previous one, and the T_m should increase in the order of 1<2<3 (either by changing the GC% or increasing the length of the primers). Using multiple primers allows optimal temperature to be used for each step of the procedure and will introduce a level of specificity that cannot be obtained with a single primer (Engelke *et al.*, 1988).

Primer selection:

The approach to the selection of efficient and specific primers is somewhat empirical. There is no set of rules that will ensure the extension of an effective primer pair. However, the majority of primers can be made to work if the following factors are considered (Erlich, 1989): (1). The sequence composition of the primer must

avoid obviously unusual structures such as long homopurine or homopyrimidine tracts or regions of self-complementary, for these might generate unusual DNA structures upon denaturation. (2). Balanced melting temperatures, (T_ms) should be used for a given primer pair. The length and GC contents of the primers can be used to approximately calculate the T_m. For shorter primers, T_m is estimated from the following rule-of-thumb calculation: 2°C for each A or T and 4°C for each G or C. For long primer, T_m can be calculated from the formula:

$$T_m = 81.5 + 16.6 (\log M) + 0.41 (\%GC) - 500/n$$

where M is the molarity of the monovalent cations and n is the length of the oligomer. (3). The spatial relationship of one primer to the other must be considered. Avoiding primers with 3' overlaps is most important and will reduce the incidence of an artifact termed "primer-dimmer".

Based on above considerations, several sets of primers were designed for my studies. The positions of the primers in the trout liver genome are shown in Figure III.3.3. To seek for hypersensitive sites in the trout ki-*ras* gene, primer sets a and b were used to study the codon 12 region from both directions, and the primer sets c and d were designed to study the 5' flanking region. Primer 1 was used for the initial extension, primer 2 used for PCR, and primer 3 used to detect the amplified fragments by labeled primer extension.

Practical test of the primers:

Primers were carefully designed, based on the considerations mentioned above, and synthesized on the DNA autosynthesizer in

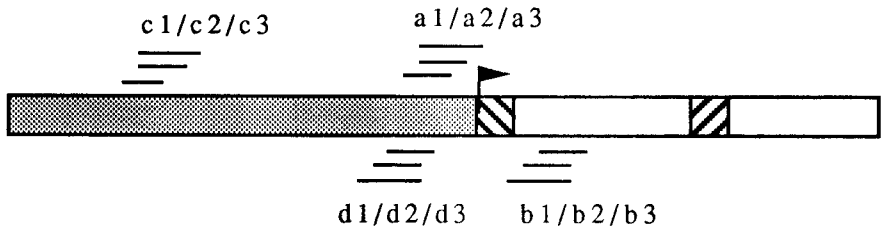


Figure III.3.3 Oligonucleotide primer sets used for LMPCR-aided analysis of the 5' region of trout ki-*ras* gene. Dotted box represents the 5' flanking region. Two boxes with diagonal lines depict the exon regions. The arrow marks the translation start site. The lines represent the positions of the nested primers. Sequences of these primers are listed in Table 4.

the Center for Gene Research at Oregon State University. Since it is the primer more than anything else that determines the success or failure of an amplification reaction, the actualy usability of primers had to be tested before they could be applied to the real experiment. This can be established by performing conventional PCR using these primers to see if they can lead to the synthesis of specific products from the trout liver genome. In this experiment, Primers a2, b2 and RHc37 were used. Primer RHc37 is identical to the sequence of the complementary strand of ki-*ras* gene, from codon 37 to 31 in exon 1, and is used here as a 3' primer. Primer a2 and b2 were used as 5' primers. Synthesis with the a2/RHc37 pair should generates a 216 bp fragment, and a 257 bp fragment should be obtained using the a2/b2 pair. In the experiment, about 1 μg of trout liver genomic DNA, from normal tissue purified as described in section II.7, was mixed with 0.2 mM of each dNTP and 60 pmole each of 5' and 3' primers in 1x Taq A buffer (50 mM KCl, 10 mM Tris pH 8.3, 2 mM $MgCl_2$ and 0.01% gelatin), and the reaction was initiated by addition of 2.5 unit of Taq polymerase. After 40 cycles of denaturation, annealing and extension steps, the DNA was precipitated with ethanol and was redissolved in TE buffer. Equal aliquots of products from each set of primer pairs were loaded on a 2% NuSieve/ 1% agarose gel, followed by electrophoresis. The result is shown in Figure III.3.4. Lanes 1 and 2 are from primer pair a2/RHc37, lanes 3 and 4 from primer pair a2/b2. Arrows at the right side of the picture indicated the two expected products in the different reactions- (a) 216 bp for lanes 1 and 2, (b) 257 for lanes 3 and 4. The bands were very sharp and free of contamination,

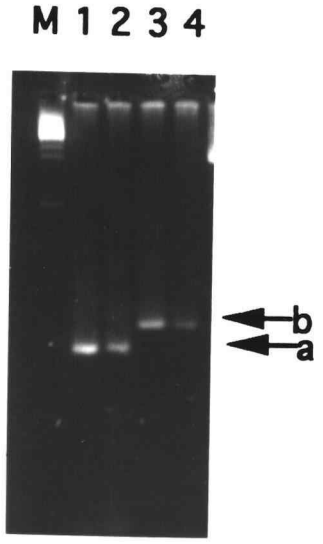


Figure III.3.4 Synthesis of trout ki-*ras* specific fragments by PCR and the optimization of Mg^{++} concentration. Lanes 1 and 2 are from primer pair a2/RHc37, which generates fragment of 216 bp (arrow a). Lanes 3 and 4 are from a2/b2, which generates fragment of 257 bp (arrow b). Lanes 1 and 3 represent the PCR reaction in 2 mM Mg^{++}, and lanes 2 and 4 the PCR reaction in 4 mM Mg^{++}. The products were electrophoresed on 2% NeuSieve/1% agarose gel.

suggesting that those primers could be used to make specific products. There is no evidence for mismatching or the spurious products.

III.3.2. Optimization of LM-PCR:

In addition to the primer factor, there are several other factors which are also important for the efficiency of LM-PCR. Each of these factors is discussed below and all were tested one by one to make sure that the interference from each has been eliminated to the lowest possible level.

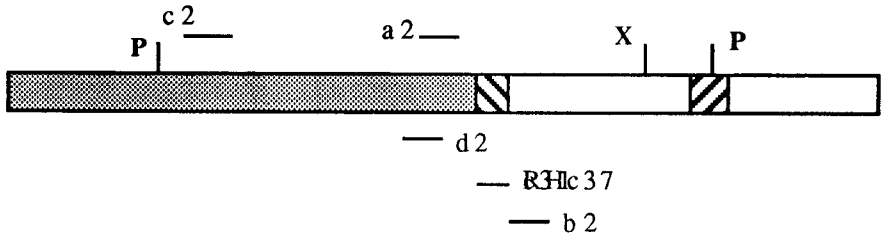
Magnesium concentration: It is believed that the magnesium concentration is an important factor for PCR efficiency. It can affect the PCR efficiency in different ways, including primer annealing, strand dissociation, formation of primer-dimmer artifact, enzyme activity and fidelity, etc, (Mueller and Wold, 1991). The magnesium concentration used in PCR can vary from low as 0.5 mM to high as 20 mM, depending on the properties and complexity of the template. The optimal magnesium concentration for most PCR work is usually in the range of 0.5 to 2 mM. To determine the Mg^{++} concentration for trout ki-*ras* gene synthesis with the specific primers, the PCR was carried out at five different Mg^{++} concentrations (0.5, 1, 1.5, 2, 3 and 4 mM) for each set of primer pairs. In Figure III.3.4, data for two concentrations (2 mM and 4 mM) are shown. Lanes 1 and 3 are for 2 mM, lanes 2 and for 4 mM. From the figure we can see that PCR is much more efficient in 2 mM than in 4 mM Mg^{++}, the difference being at least 5 fold. This

difference is seen in both primer pairs indicating that 2 mM magnesium is optimal for this particular template, and so 2 mM was used for all following experiments. The PCR efficiency at lower Mg^{++} concentrations range from 0.5 mM to 2 mM did not shown significant difference (data not shown).

The PCR buffer: Although there are many different PCR buffer conditions successfully used by researchers all over the world, the overall conditions can be simply classified into two basic kinds: one is the KCl-based buffer system (recommended by Cetus, this contains 50 mM KCl, 10 mM Tris pH 8.3, 2 mM $MgCl_2$ and 0.01% gelatin) and the other is the NaCl-based buffer system (developed by Muller and Wold (1989), this contains 40 mM NaCl, 10 mM Tris pH 8.9, 2 mM $MgCl_2$ and 0.01% gelatin). The main difference, as shown above, is the nature of the monovalent cation. In many reported experimental results, the NaCl based buffer showed better performance in PCR (Mueller and Wold, 1991), probably because it helps Taq polymerase pass through polyGC rich regions better than does KCl. In our system, the relative efficiencies of these two buffer conditions were determined experimentally. PCR synthesis from trout ki-*ras* was carried out in parallel in these two buffers. Figure III.3.5 shows the strategy and results. In addition to a2, b2 and RHc37, primer c2 and d2 were also tested in this experiment (panel A). Specific fragments were expected to be amplified with different combinations of primer pairs, and their sizes are listed in panel B. As in the previous experiment, 1 μg of trout genomic DNA was mixed with 60 pmoles of each primer pair and the synthesis with Taq polymerase was performed in NaCl buffer and KCl buffer. Panel

Figure III.3.5 Optimization of Taq polymerase buffer. Panel A depicts the strategy of PCR. Panel B lists the expected sizes of products for each corresponding primer pair. Panel C shows the ethidium stained gel electrophoresed on 2% NeuSieve/1% agarose gel. In the panel C, I depicts NaCl buffer, II KCl buffer. Lanes 1 to 5 represent the results from each primer pair as in panel B. Gels a to d represent 4 independent experiments (cont. on pages 85 & 86).

A.



B.

	5' primer	3' primer	product(bp)
1.	a 2	RHc37	215
2.	a 2	b 2	256
3.	c 2	d 2	918
4.	c 2	RHc37	1113
5.	c 2	b 2	1154

Figure III.3.5 Optimization of Taq polymerase buffer.

C.

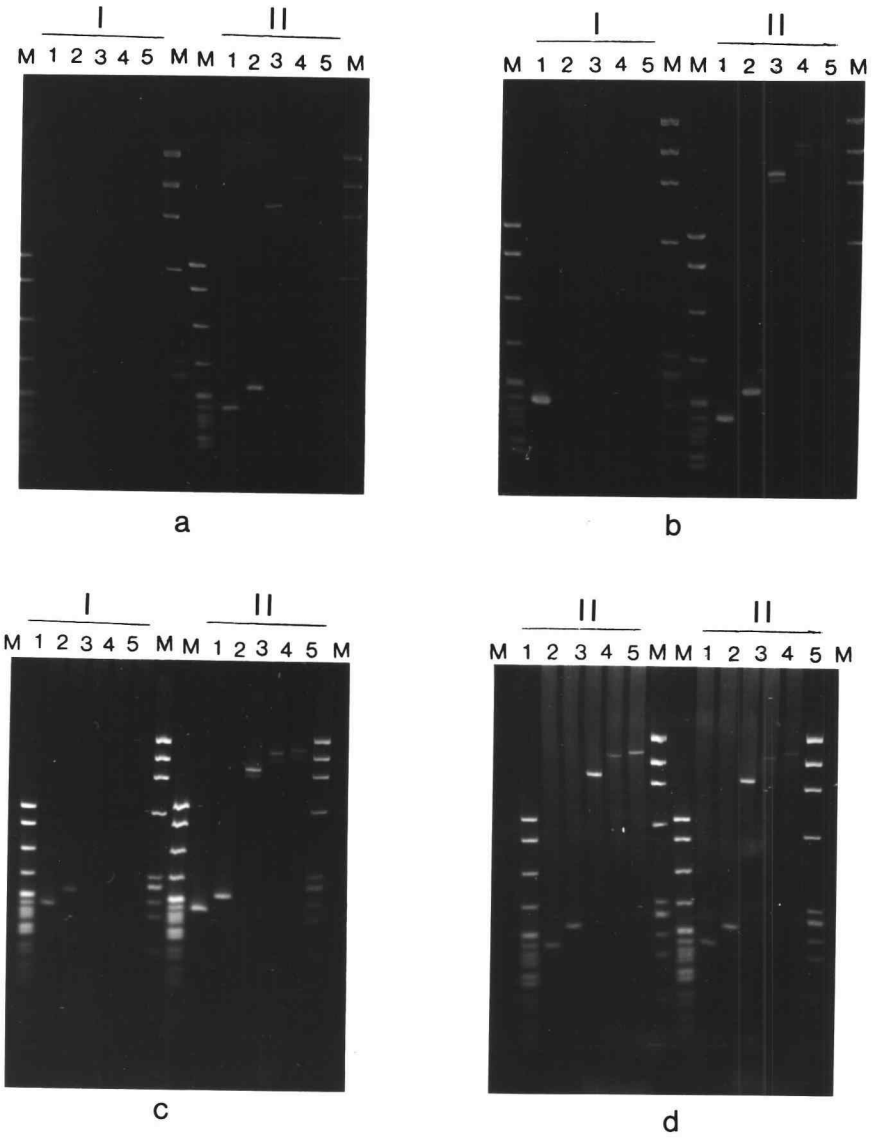


Figure III.3.5 Optimization of Taq polymerase buffer.

C, in each case, I depicts the NaCl buffer, II the KCl buffer. Lane 1 is from the primer pair a2/RHc37, lane 2 from a2/b2, lane 3 from c2/d2, lane 4 from c2/RHc37 and lane 5 from c2/d2. The result clearly showed that the synthesis was more efficient in KCl buffer than in NaCl buffer. The experiment was repeated several times and the result was consistent (a to c) and the result was reproducible also in a repetition of the same experiment (d, tested for KCl buffer). The specificity of the products obtained in the KCl buffer system was examined by Southern blots, using a restriction fragment PX from pasmid pTras-1, that covers the 1 kb of the 5' flanking region and 111 bp of exon 1 of the trout ki-*ras* gene, the result is shown in Figure III.3.6. As the same result from ethidium bromide stained gel in Figure III.3.5, the *ras* specific products with correct expected sizes from different primer pairs were observed in KCl buffer (panel C), not in NaCl buffer (data not shown). This result suggested that the specific template, trout ki-*ras* gene, needs KCl-mediated buffer system for its maximum synthesis by PCR with these specific primers.

Ligation reaction: In the ligation reaction, a synthetic linker is attached to the blunt-ended duplexes which are formed by extension of the gene-specific primer. In this reaction, the 5'-phosphate of the genomic DNA is ligated to the 3'-hydroxyl of the longer strand of the staggered linker, whose sequence is shown below:

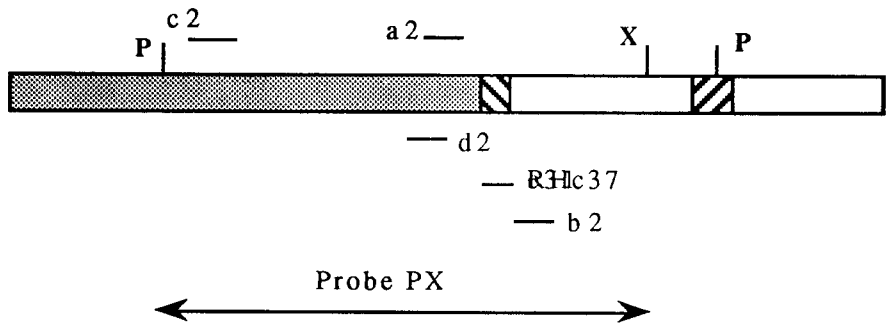
5'-OH-GCGGTGACCCGGGAGATCTGAATTC-OH-3' 25-mer

CTAGACTTAAG-OH-5' 11-mer

T4 DNA ligase catalyzes this reaction, and it requires that both the

Figure III.3.6 Specificity of trout ki-*ras* fragments from PCR synthesis by Southern blot. The gel from the experiment d in Figure III.3.5 was Southern transferred to a nylon membrane, blotted and probed with probe PX that covers the 1 kb of the 5' flanking region and the exon 1 of the trout ki-*ras* gene (cont. on pages 89 & 90).

A.



B.

	5' primer	3' primer	product(bp)
1.	a 2	RHc37	215
2.	a 2	b 2	256
3.	c 2	d 2	918
4.	c 2	RHc37	1113
5.	c 2	b 2	1154

Figure III.3.6 Specificity of trout ki-*ras* fragments from PCR synthesis by Southern blot.

C.

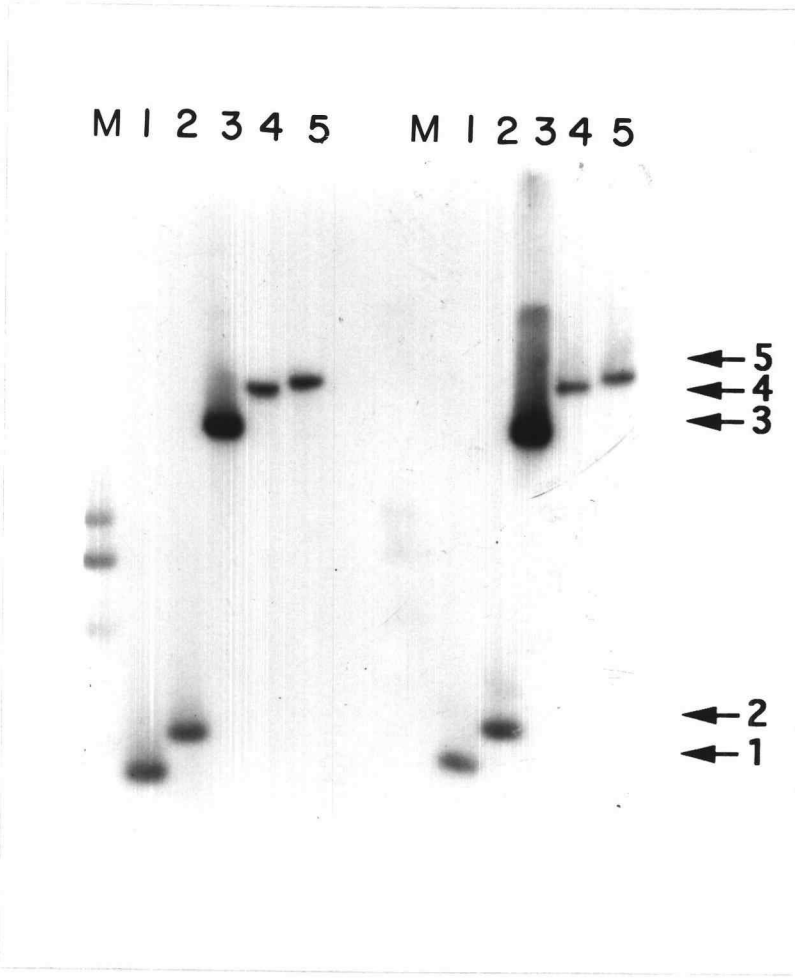


Figure III.3.6 Specificity of trout ki-*ras* fragments from PCR synthesis by Southern blot.

linker and the genomic substrates be double-stranded. The staggered structure of the linker combined with its low GC content in the duplexed region ensures that the hybridized linker will not be stable under Taq polymerase conditions. This permits PCR amplification to be performed without removing the unused linker, as the shorter oligonucleotide cannot prime under these conditions. The efficiency of this blunt-end ligation is very important for the LM-PCR because the substrates for following PCR amplification are formed in this step. However, the efficiency of blunt-end ligation is usually very low and is dependent upon buffer conditions. To determine the conditions to optimize this reaction, I planned to directly use trout liver DNA that had been digested with a blunt-end restriction enzyme, for example, Rsa I in the ki-*ras* gene, to do the ligation reaction with ^{32}P end-labeled linker, and then look for the shift of the radioactive band on an autoradiogram after electrophoresis. But it is impossible to see this shift for a single-copy gene without amplification. Therefore I decided to carry out this experiment *in vitro*. After the ligation of the template with linker, the whole ligation mixture is then modified to Taq buffer conditions and PCR is carried out with Taq polymerase. If the ligation works and is efficient enough, then suitable substrates with attached linker should be formed and the amplification of the template should be observed. Based on this hypothesis, pTras-1 was used for this study (panel A in the Figure III.3.7). There are four Rsa I sites in this plasmid, three of which are located in the vector, and one is in the insert. Digestion of the plasmid generated four blunt-ended linear fragments: a through d, with sizes of 2262,

Figure III.3.7 Gel analysis of ligation reaction. Lane M contains the marker pBR322/msp I. Lanes 1 to 7 demonstrate the same reaction within different reaction tubes. Arrows on the left side represent the expected sizes of amplified fragments, corresponding to each of the Rsa I fragments (cont. on page 93).

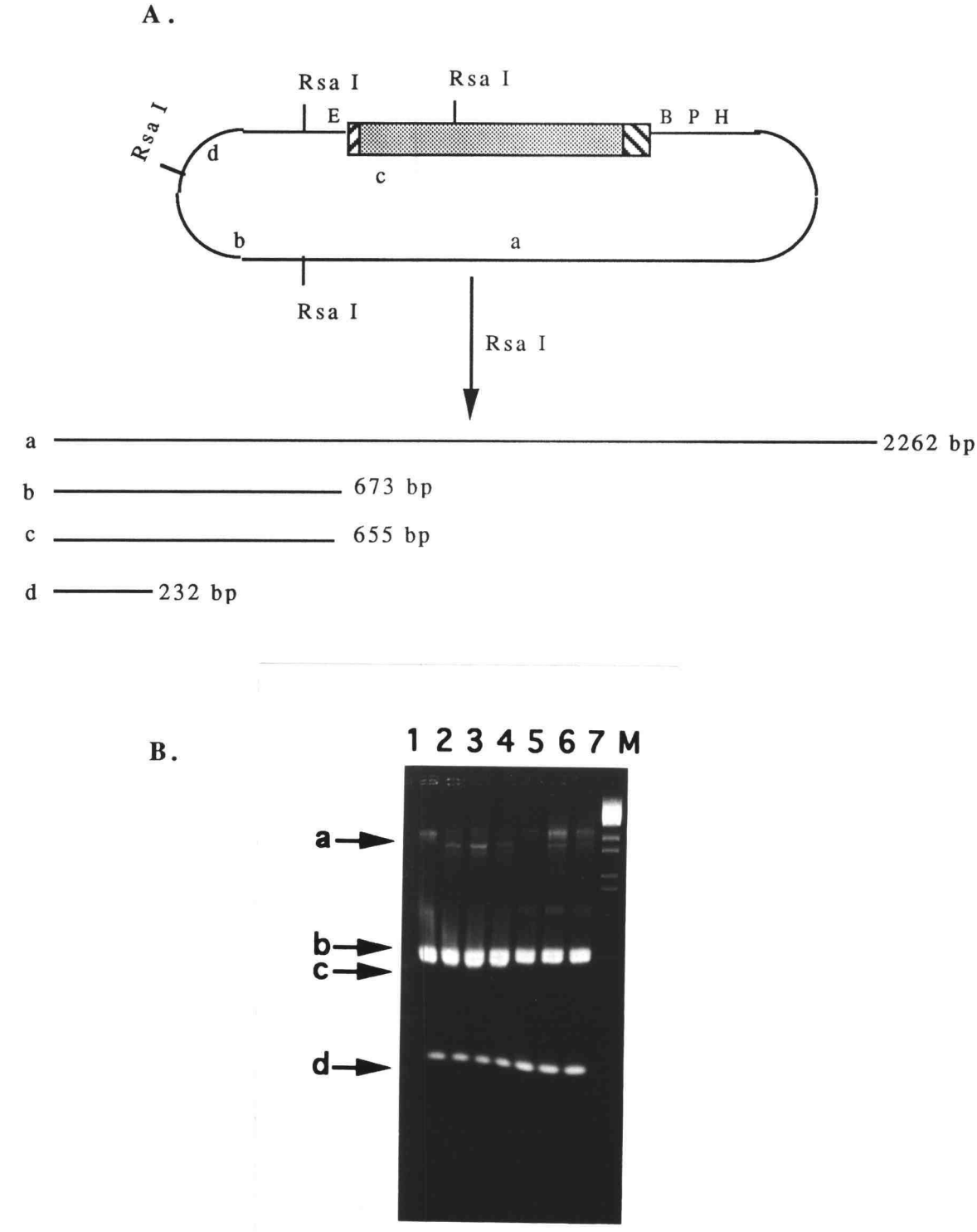


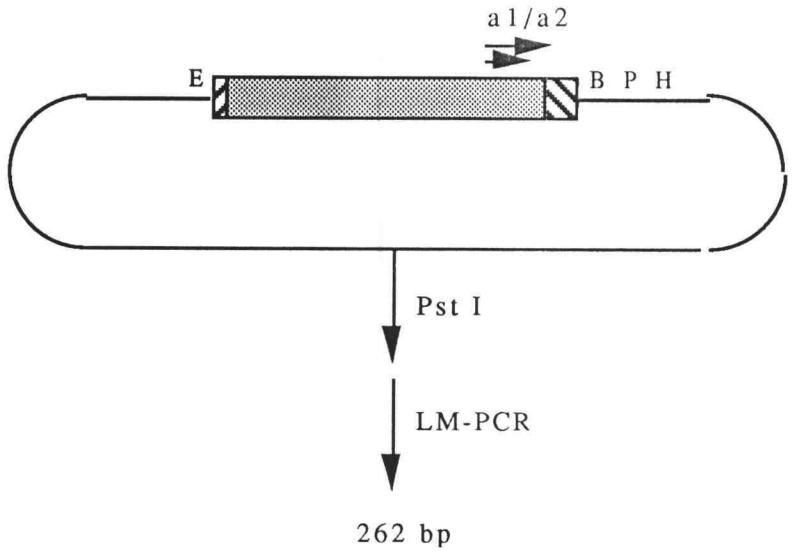
Figure III.3.7 Gel analysis of ligation reaction.

673, 655 and 232 bp respectively. In the reaction, 5 ng of Rsa I digested pTras-1 was ligated with 20 pmoles of linker in the ligation buffer (Pfeifer *et al*, 1989) for overnight at 17°C. Then the mixture was modified to Taq buffer (as mentioned before) and carried through 10 cycles of PCR in the presence of 2.5 units of Taq polymerase. The DNA was then chloroform extracted and ethanol precipitated. The resulting product was analyzed on agrose gel shown in panel B of Figure III.3.7. Lane M shows the markers (pBR322/Msp I). Lanes 1 to 7 demonstrate the same reaction within different reaction tubes. Arrows on the left side represent the sizes of the expected amplified fragments, corresponding to each of the Rsa I fragments. As shown in the figure, all of the Rsa I fragments were significantly amplified after ligation and PCR reaction. The largest fragment was less amplified due to the low efficiency of PCR amplification for longer DNA. This result showed that the blunt-end ligation worked well and was efficient for following PCR amplification.

Sequenase reaction: From the studies described above, it is clear that the PCR amplification and ligation work in the test conditions. Since the Sequenase reaction is the first step in LM-PCR, its efficiency will also affect the overall efficiency of LM-PCR. Therefore this reaction was also tested empirically. The plasmid pTras-1 was used again in this study. Instead of using Rsa I, Pst I was chosen to digest the plasmid (see panel A in Figure III.3.8). Digestion by Pst I generated a sticky end, instead of the necessary blunt-end. Therefore a specific primer (primer a1 in this study)

Figure III.3.8 Test of the Sequenase reaction. LM-PCR of pTras-1 digested with Pst I was performed with primer set a. M was the DNA marker. Lane 1 was the no-template control. Lanes 2 to 6 show production of the reaction using different amount of pTras-1/Pst I, from 0.2 ng (lane 2), 1 ng (lane 3), 3 ng (lane 4), 5 ng (lane 5) to 10 ng (lane 6). The arrow on the right side indicates the specific amplification product (cont. on page 96).

A.



B.

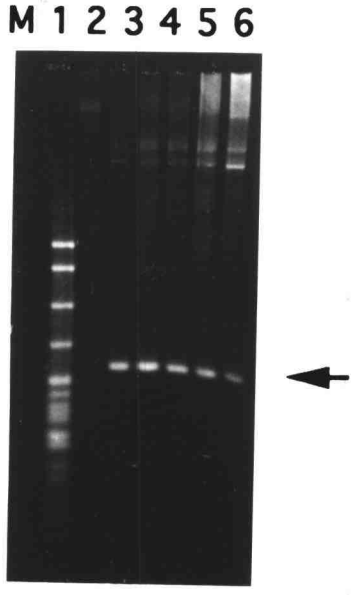


Figure III.3.8 Test of the Sequenase reaction.

was required to generate blunt-ends in order to allow the following two steps of LM-PCR to work. The final product should be 262 bp, and this includes 237 bp from the plasmid and 25 bp from the linker. To do the experiment, 5 ng of Pst I digested pTras-1 was mixed with 0.6 pmole of a trout ki-*ras* specific primer, (called prm a1) in 1x Sequenase buffer. The mixture was heat-denatured, and then annealed at 47°C. Extension of the primer was initiated by addition of Sequenase and continued for 15 min at 47°C. After the enzyme was heat-denatured, the reaction mixture was modified to ligation conditions and the remainder of reactions of LM-PCR were followed. The product was analyzed by agrose gel electrophoresis; the results are shown in panel B of Figure III.3.8. M designates the DNA marker. Lane 1 shows the no template control, lanes 2 to 6 demonstrate the reaction using different amount of pTras-1/Pst I. The arrow on the right side indicates the specific amplification fragment of 262 bp as expected. This result showed that the Sequenase reaction worked well under this test condition and that templates with blunt-ends were efficiently formed by extension of the gene-specific primer with Sequenase.

To summarize this section: Primers were designed based on the considerations important to this work and were tested for their usability in the PCR reactions. LM-PCR procedures were carefully tested in a step by step fashion in order to optimize conditions. The importance of this section is that it demonstrates that well defined, amplified products could be obtained from single-copy sequences in the trout genome, using the LM-PCR method. This section should be

thought of as a series of very critical controls, showing that the methods were capable of detecting specific, single-copy fragments under the conditions used.

III.3.3. Analysis of the chromatin structure of the trout ki-*ras* gene:

After the LM-PCR technique was carefully tested in the step-by-step fashion to optimize the reaction conditions, the chromatin structure of the trout ki-*ras* gene was examined by this technique. Two aspects of the chromatin structure were studied: a search for the DNase I nuclease hypersensitive sites was carried out, and tests for the nucleosome positioning were conducted by micrococcal nuclease digestion. In both studies, the LM-PCR technique described above were employed.

III.3.3.1. Search for DNase I hypersensitive sites:

Figure III.3.3 shows the strategy of the LMPCR-aided DNase I mapping experiments. The names of the primers, a to d, were used to locate the positions of the primers along the genome. The numbers by each name represent the use of the primer in LM-PCR procedure. Primer set a and b were used to map the hypersensitive sites around codon 12 region, from both directions. Primer set c and d were designed for possible use in future studies of the 5' flanking region..

Study of the noncoding strand with primer set a:

The importance for this study was the hypothesis that the codon 12 region, which represents the activation site for aflatoxin-

induced carcinogenesis, might be especially accessible in a carcinogen-sensitive tissue like trout liver. As mentioned in the last section, a set of nested primers was used. Primer a1 was used for the Sequenase reaction, primer a2 for PCR amplification and primer a3 for the labeling reaction to visualize the products. Their sequences were derived from the coding strand upstream from codon 12 so the results would reflect the nuclease cutting sites, were there any, in the noncoding strand around the codon 12 region. In other words, if the codon 12 region is especially exposed in trout liver chromatin hypothesized, these studies should demonstrate this.

Nuclei were isolated from trout liver and digested with different amounts of DNase I. Genomic DNA was then purified, free of all proteins, and digested with restriction enzyme Rsa I. One Rsa I site is located in the intron, close to the exon 1, so the digestion with Rsa I will limit the full length of LM-PCR to fragments no longer than the fragment defined by the Rsa I site. As a starting material, 3 μg of Rsa I digested DNA was used and the complete process of LM-PCR was followed, under the optimized conditions described above. The final result was shown in Figure III.3.9. The full length of amplification product should extend from the 5' end of the primer a3 to the Rsa I site, plus an additional 25 due to attached linker, a total of 259 bp. If there were any hypersensitive sites, well defined smaller fragments should be seen on the gel in DNase I digested template and the intensity of these bands should increase as the DNase I concentration increases. As shown in the figure, the full length of 259 bp product was observed in every lane, and the intensity of the band decreased as the DNase I concentration

Figure III.3.9 Search for DNase I hypersensitive site around codon 12 region of trout ki-*ras* gene by LM-PCR. Genomic DNA was isolated from trout liver nuclei which were previously digested with different amounts of DNase I at 0 (lane 1) to 32 u/mg/ml (lane 6). DNA was then digested with Rsa I followed by the performance of LM-PCR, using primer set a. Each LM-PCR reaction involved 3 μg of Rsa I digested DNA as starting material (cont. on page 101).

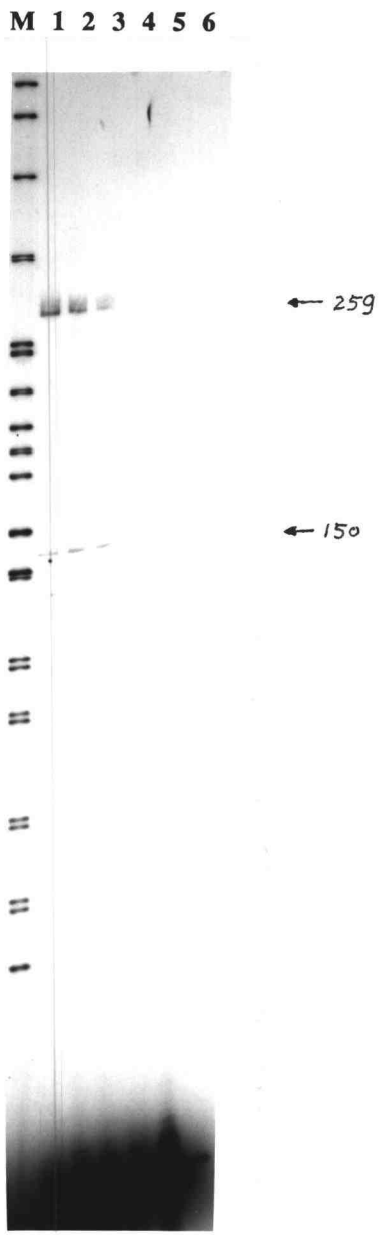


Figure III.3.9 Search for DNase I hypersensitive site around codon 12 region of trout ki-*ras* gene by LM-PCR.

increased. This is expected, as a consequence of random degradation of the template by DNase I. However, no specific bands that correspond to the DNase I digestion could be observed. This result suggested that there are no hypersensitive sites in the region we were studying in this experiment, which extends from residues -100 to +148, including codon 12 (Fig. III.2.4).

An artifactual band at 150 bp:

Although no specific band corresponding to DNase I digestion could be observed, there was indeed a small fragment at about 150 bp range seen in every lane (Fig. III.3.9). What is this fragment? It cannot correspond to a hypersensitive site, because it was also observed in the control lane (no DNase I). It is also unlikely that this was due to endogenous nuclease cleavage because it was demonstrated in section III.1 that there was no significant endogenous nuclease activity present in the preparation. It seemed to me that one possibility was the presence of resistant secondary structure which could cause the premature termination of the PCR synthesis. To seek evidence for this, the sequence around this 150 bp region was carefully checked and the sequences noted below were detected between codon 6 to codon 15 (150 bp fragment referred to the stop at codon 8):

6 7 8 9 10 11 12 13 14 15

CTG GTG GTG GTG GGG GCA GGA GGT GTG GGC

The melting temperature for this region was calculated and it turned out to be over 76°C. This happened to be the temperature used for extension during PCR amplification. So the result could be

Figure III.3.10 Test for the artifactual product during LM-PCR. The experiment mentioned in Figure III.3.9 was repeated under the same conditions, except the extension temperature was increase from 76°C to 80°C (cont. on page 104).

Figure III.3.10 Test for the artifactual product during LM-PCR.

explained by the hypothesis that due to the very high T_m around this region, the template was not totally melted during extension at 76°C, therefore causing the premature termination of the synthesis. This hypothesis was proved by repeating the experiment as before, but raising the extension temperature from 76°C to 80°C. The result was shown in Figure III.3.10. Again the full length product was observed, but the 150 bp band has disappeared or dramatically decreased in intensity. So it was concluded that the 150 bp fragment was an artifactual product.

Effect of the extension temperature for PCR:

By comparing two figures shown above, it was immediately noticed that at 76°C extension temperature, the gel was cleaner, in terms of the background (except the 150 bp artifactual band) than at 80°C extension. At 80°C, the artifactual band was gone, but the contamination by smaller fragments was increased. What is the explanation for this difference? The film of Figure III.3.10 was scanned and each contamination band was aligned to the corresponding sequence. The result was shown in Figure III.3.11. Panel A shows the corresponding sequence, panel B the scan result. The solid line in panel B depicts the pBR322/Msp I as Marker, the dotted line is from lane 3 in Figure III.3.10. It was observed that most of these peaks aligned very well to high A/T region, as in the sequence. So my interpretation is that when the extension temperature is raised too high, the synthesis is not stable in high A/T region, especially for smaller products. Therefore these products easily fall off the template as the polymerase passes. The

Figure III.3.11 Effect of extension temperature for LM-PCR. The densitometric scan of the LM-PCR products from the lane 3 of Figure III.3.10 was carried out and the result is shown as dotted line in the panel B. Each peak represents the premature termination product of the LM-PCR. The corresponding sequence to this scan profile is shown on the top of the scan in the panel A. Solid line represents the scan from the DNA marker (cont. on page 107).

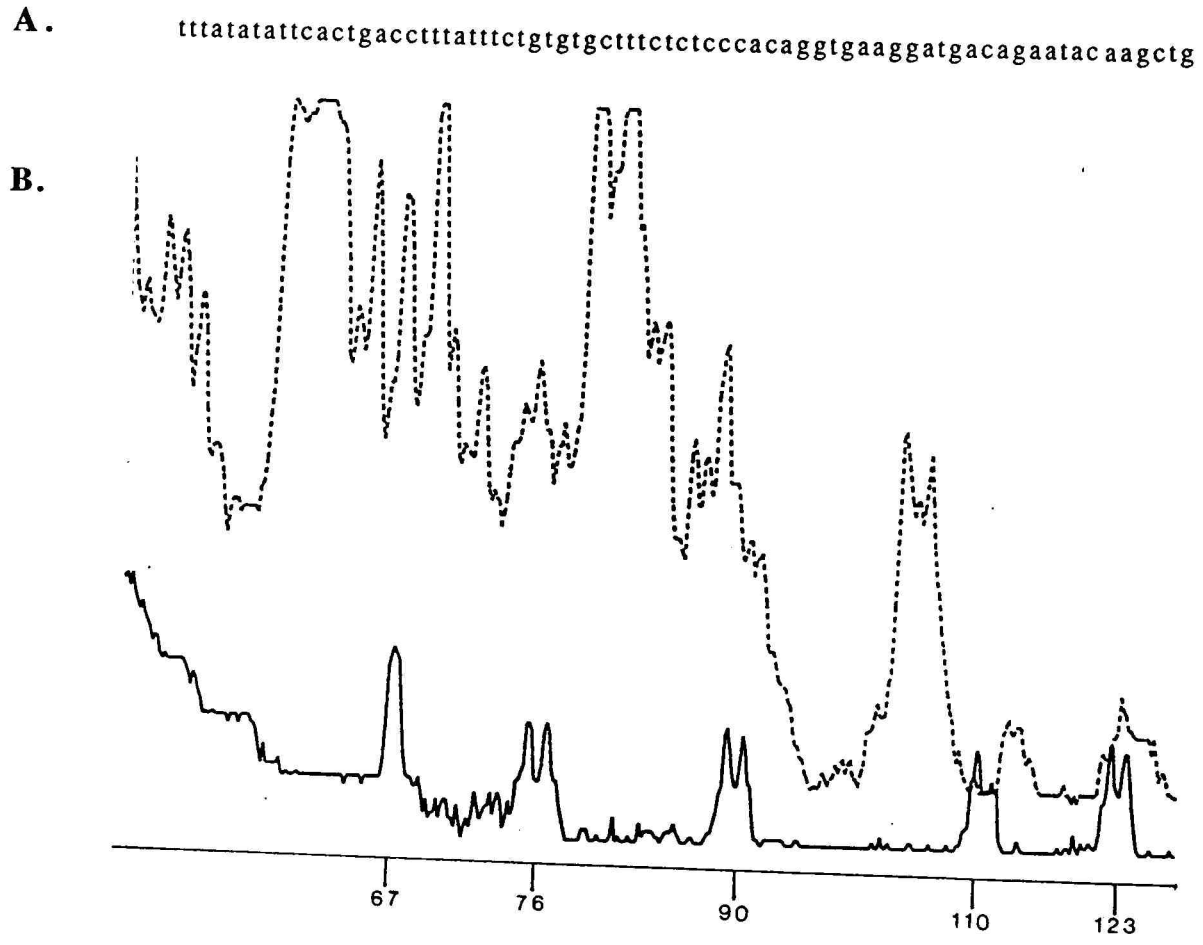


Figure III.3.11 Effect of extension temperature for LM-PCR.

above results indicate that the extension temperature during PCR is very critical for the efficiency of the amplification. It is also important for the interpretation of the result, particularly in differentiating the true results from artifacts or background. For each set of primer pairs and each template, the effect of temperature should be examined.

Tumor studies:

It is equally important to test if the carcinogen AFB1 can induce a change of chromatin structure after it binds to chromatin and damages the DNA. The same set of experiments was carried out using trout liver tumor tissue, to see if there will be any difference or changes in the chromatin structure following tumor formation. The result is shown in Figure III.3.12, and is very similar to the results from normal tissue study as shown in Figure III.3.9. No hypersensitive sites could be found in the DNase I digested sample, suggesting that the chromatin structure is the same for both tissues and changes in the chromatin structure are probably not important in the tumor formation.

Study on the coding strand with primer set b:

The attempt to seek the DNase I hypersensitive sites has been extended to the study of the coding strand of the *ras* gene using the same technique. To study the codon 12 region, the down stream gene specific primer set b was used. The procedure and conditions followed were exactly the same as mentioned before, except BsrI enzyme was used here to restrict the full length fragment to 289 bp. Figure III.3.13 shows the result. Basically, we could not see any

Figure III.3.12 Search for DNase I hypersensitive sites in trout ki-*ras* chromatin from tumor tissue. Genomic DNA was purified from trout liver nuclei of tumor tissue, which were previously digested with different amount of DNase I at 0 (lane 1) to 64 u/mg/ml (lane 8). DNA was then digested with Rsa I and followed by performance of LM-PCR using primer set a (cont. on page 110).

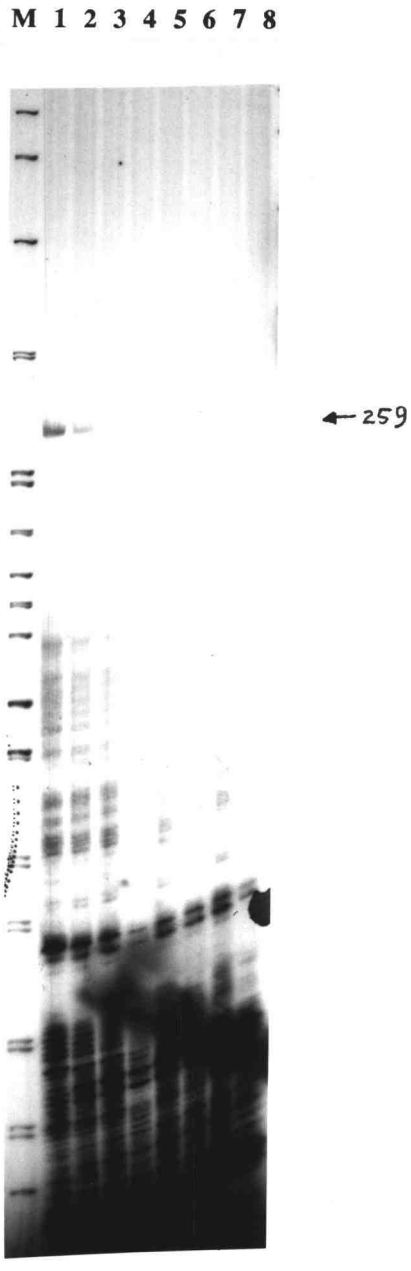


Figure III.3.12 Search for DNase I hypersensitive sites in trout ki-*ras* chromatin from tumor tissue.

product which corresponds to DNase I cleavage. The overall efficiency with primer set b was lower, probably because of Bsr I as poor cutter. BsrI did not cut genomic DNA very well, therefore reducing the amount of template available for LM-PCR, causing poor synthesis of the full-length fragment.

In summary, by using LM-PCR method, the chromatin structure of trout ki-*ras* gene was studied by DNase I mapping. Apparently there are no hypersensitive sites around codon 12 region, either in normal or tumor tissue, suggesting the chromatin structure may not play a role in *ras* activation involved in tumor formation and that there are no changes upon tumor induction.

III.3.3.2. Search for nucleosome positioning in trout ki-*ras* gene around codon 12 by MNase.

In some, but by no mean all cases, there exists a regular nucleosomal arrangement in the 5' flanking region and /or the coding region of genes (van Holde, 1989). Such "positioning" can be detected by genomic analysis using micrococcal nuclease. Micrococcal nuclease prefers to cut in the linker regions between nucleosomes, so after such digestion, the chromatin will release a mixture of oligonucleosomes, resulting in a regularly spaced "ladder" on a DNA gel. So MNase digestion has been used to study the nucleosome positions in a particular sequence. To seek evidence for regularly positioned nucleosomes in the region around codon 12 of the *ras* gene, I have used MNase digestion coupled with LM-PCR. Unlike DNase I, MNase digestion generates a free 5'-OH which cannot

Figure III.3.13 Study on the coding strand by LM-PCR. Genomic DNA was purified from nuclei of normal tissue, which was previously digested with different amount of DNase I at 0 (lane 1) to 32 u/mg/ml (lane 6), and then digested with Bsr I, followed by carrying out of LM-PCR, using primer set b (cont. on page 113).

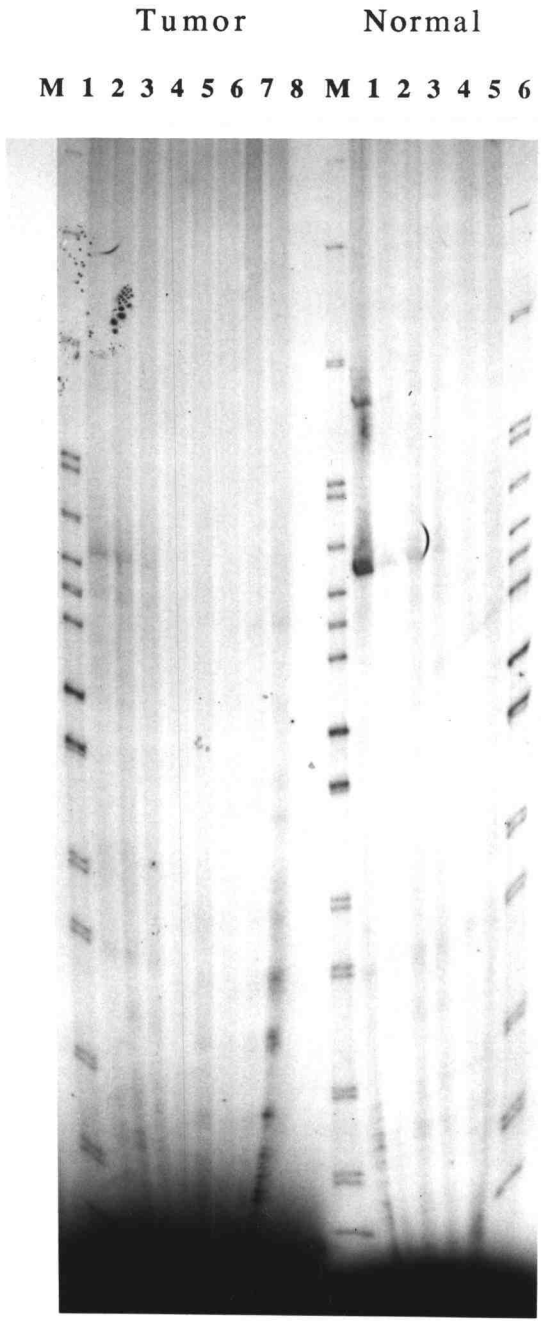


Figure III.3.13 Study on the coding strand by LM-PCR.

be a substrate for linker attachment during LM-PCR. Therefore the MNase digested template has to be phosphorylated before LM-PCR.

As in the DNase I studies, the trout liver nuclei were digested with different amounts of MNase. Then the genomic DNA was purified, digested with Rsa I and phosphorylated with T4 polynucleotide kinase. The phosphorylated templates were used for the LM-PCR experiment, using primer set a. The result is shown in Figure III.3.14. Again, the full length 259 bp product was observed in each lane. No specific bands, corresponding to strongly preferred sites of MNase cutting, were observed, although the 150 bp artifactual band was also seen in this gel. This result can be explained in two ways. Either the nucleosomes around the codon 12 region are randomly positioned, or there are no nucleosomes in this region. These two possibilities can not be differentiated in this study.

Figure III.3.14 Search for nucleosome positioning in trout ki-*ras* chromatin by micrococcal nuclease study coupled with LM-PCR. Nuclei from normal tissue were digested with different amount of MNase at 0 (lane 1) to 46 u/mg/ml (lane 6). DNA was isolated, digested with Rsa I, followed by performance of LM-PCR using primer set a (cont. on page 116).

Figure III.3.14 Search for nucleosome positioning in trout ki-*ras* chromatin by micrococcal nuclease study coupled with LM-PCR.

III.4. Sequence specificity of aflatoxin B_1 binding to trout ki-*ras* gene:

In the previous two sections, results have been presented from studies of rainbow trout ki-*ras* gene, from sequencing to chromatin structure, with the aim of testing the possible effect of chromatin structure on AFB1 induction of liver tumors. The results, however, were negative. In this section, I have turned efforts to test the second hypothesis- that intrinsic accessibility of guanines at codon 12 to AFB1 is higher than at other nearby guanine sites.

Virtually all of the aflatoxins bind to guanine sites, but the relative affinity varies greatly between different sites. This affinity depends upon still unrecognized features of DNA sequence or structure, for the simple nearest-neighbor rules (Misra *et al.*, 1983) could not account for all the variation observed, and therefore could not be used to predict the attack frequency of sites in a particular sequence. The intrinsic affinity of each guanine site has to be determined empirically.

Previous workers (Marien *et al.*, 1989) have studied the sequence selectivity of human c-Ha-*ras* gene to AFB1. In this section, the same question was studied using trout ki-*ras* gene, with focus on the codon 12 region.

III.4.1. Preparation of the template:

A fragment of 136 bp, including 111 bp of exon 1 and 25 bp adjacent 5' flanking sequence, was used for AFB1 binding experiments. It was prepared by PCR amplification from the

```
                                    1   2         3   4 5 6 7 8 9 101112
GTGCTTTCTCTCCCACAGGTGAAGGATGACAGAATACAAGCTGGTGGTGGTGGG

1314151617181920212 2324 2526
GGCAGGAGGTGTGGGCAAGAGCGCGCTCACCATCCAGCTCATTCAGAACCACTT

TGTGGATGAATATGACCCCACCATCGAG
```

Figure III.4.1 Sequence of 5' region of exon 1 region used as template for aflatoxin B_1 alkylation studies. The template was synthesized by PCR with primers 5'-25 and RHc37. Numbers on the top of the sequence are assigned to each guanine as a reference number, starting from the 5' terminus of the exon 1. Underlined nucleotide bases represent the codon 12.

plasmid pTras-1, using two specific primers 5'-25 and RHc37, and its sequence shown previously in Figure III.2.4, is shown here again in Figure III.4.1 for more clarity. The numbers are assigned as references to each guanine site, starting from the 5' terminus of exon 1. Resolution in the region close to the primer is usually low, therefor the "enlarged" fragment (instead of exon 1 only) was used so that the guanines at the first several codons could be examined. To end label one strand, PCR was carried out with these two primers, in which one primer was kinased. Theoretically, this method should result in very clean template with only one strand labeled. However, in practice, the purity of the template was not so perfect, presumably due to very high efficiency of PCR and high sensitivity of ^{32}P labeling. Therefore, after PCR synthesis, the product was purified by PAGE gel electrophoresis. The desired fragment was located by brief exposure of the wet gel to x-ray film (Hyperfilm, from Amersham) and then excised from the gel. This was followed by several rounds of elution in TE buffer. The labeled template was then ethanol precipitated and dissolved in 50 mM Tris.HCl pH 7.4 and 10 mM EDTA.

III.4.2. Activation of AFB1:

AFB1 requires activation before it can significantly react with biological macromolecules. The activation *in vivo* is believed to proceed via epoxidation of the 8, 9 double bond (Essignmann *et al.*, 1977; Lin *et al.*, 1977). Usually, for *in vitro* experiments, activated derivatives of AFB1 are used. These include naturally produced AFB1-epoxide, chemically synthesized AFB1-Cl_2 (Fahmy *et al.*,

1978) and AFB1-Br_2 (Marien *et al.*, 1989). One feature of these derivatives is that they are very unstable in an aqueous environment, which makes the experiment difficult to handle. Therefore, all the experiments should be done with special care. For example, one must work in a clean and dry area in a hood, and use a flow of dry nitrogen to dry AFB1-Br_2. Special precaution is also required because of the extreme carcinogenety of the activated aflatoxin.

Several studies have demonstrated that the sequence selectivities are the same between these analogues (Fahmy *et al.*, 1978; Misra *et al.*, 1983; Muench *et al.*, 1983; Refolo *et al.*, 1985; Marien *et al.*, 1989). Therefore the chemically synthesized analogs can be as substitutes of naturally produced epoxide compound to study the AFB1 binding affinities in a particular sequence. In this study, AFB1-dibromide was used because this analog is relatively more stable than others, is more reactive, and easier to synthesize (Marien *et al.*, 1989). The structure of this compound is shown in Figure III.4.2.

III.4.3. Determination of adduction frequency:

It is important that the adduction frequency be limited to 0.5 hit/strand or less, since multiple cuts on a given strand will bias the result toward observation of cutting sites close to the labeled 5' end. The cutting frequency, expressed as the average number of cuts/strand, can be calculated using the Poisson Distribution function:

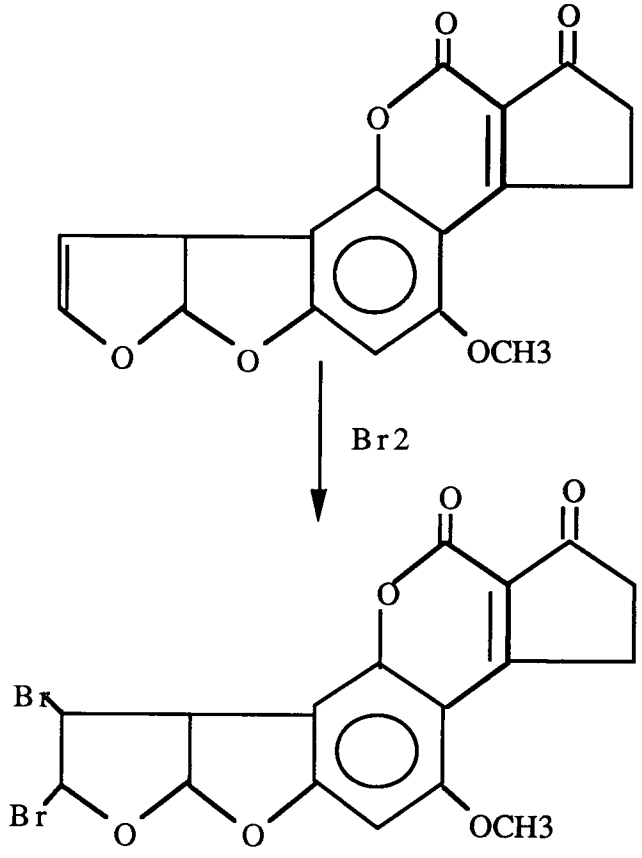


Figure III.4.2 Synthesis of aflatoxin B_1 dibromide.

$$P = \exp(-Y)$$

in which P means the probability that no cuts occur in the treated samples, and Y means the average number of cuts/strand. At less than 0.5 cleavage/strand, less than 10% of the fragments will be cut twice or more, and such low levels are not visible from the background on an autoradiogram. Practically, this frequency can be readily established by the following method. Fixed amounts of labeled template are mixed with various amounts of AFB1-Br_2. At the end of alkylation, the reaction is stopped and DNA is ethanol precipitated. The pellet is redissolved in freshly prepared piperidine to allow cleavage of adducts to occur. Then the cleavage pattern is resolved on a sequencing gel and the film is scanned by densitometry. The intensity of the uncut DNA band from alkylated samples was compared with that of control sample. Then the intensity difference between these two gave the probability that any strand was not cut. Observing a 50% decrease in the uncut band intensity demonstrates represents 0.5 hit/strand.

III.4.4. Intrinsic accessibility of guanines of codon 12 to AFB1:

To examine the sequence interaction selectivities of AFB1 in trout ki-*ras* sequences, freshly prepared AFB1-Br_2 was dissolved in a solution containing end-labeled template. After about 30 min exposure, N-7 guanyl adducts in the template fragment were cleaved quantitatively by alkaline treatment. Radiolabeled fragments were then separated on a sequencing gel. Figure III.4.3 is an autoradiograph of a typical sequencing gel, with DNA ^{32}P labeled on the 5' end of the coding strand. Lane 6 represents DMS-

Figure III.4.3 DMS and AFB1-Br_2 cleavage of exon 1 of trout ki-*ras* gene. The DNA fragment (coding strand) was prepared by PCR synthesis with two specific primers, ^{32}P labeled 5'-25 as 5' primer and RHc37 for 3' primer. The labeled fragment was treated with 10 mM DMS (lane 6) or 2 μg of AFB_1-Br_2 (lanes 2-5) for 30 min at room temperature. Control (lane 1) was carried through all steps except alkylation. The alkylated pattern was visualized by piperidine cleavage followed by electrophoresis on sequencing gel. Lane M was the DNA marker, pBR322/msp I (cont. on page 124).

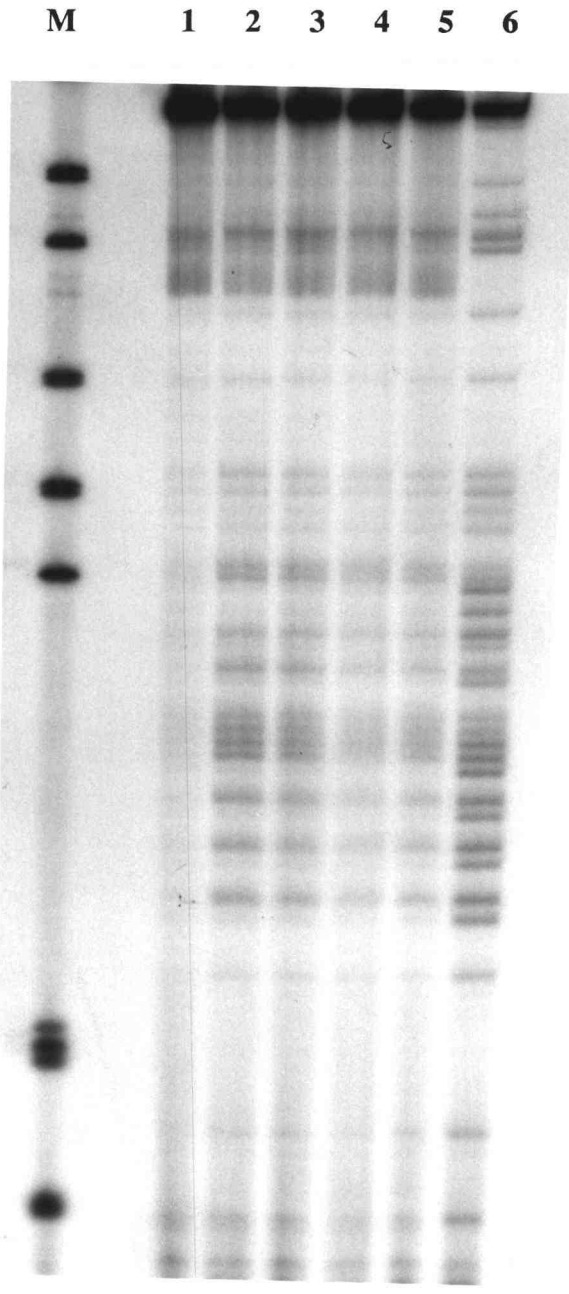


Figure III.4.3 DMS and AFB1-Br_2 cleavage of exon 1 of trout ki-*ras* gene.

DNA binding and the subsequent cleavage pattern. Lanes 2 to 5 represent the AFB1-DNA binding and cleavage pattern. Lane 1 is the control. To determine the background cleavage levels which were from sources other than alkylation, DNA was exposed to all steps, except alkylation with DMS or AFB1-Br_2 (lane 1). From this length of film exposure, some faint bands were seen in the control lane, indicating some non-specific cleavage by piperidine. To quantitate the cleavage of guanines due to alkylation, the intensities of the background level were subtracted from each peak. Representative densitometric scans of the alkylation pattern produced by DMS or AFB1-Br_2, represented by the relative intensity versis the position of guanine, are shown in Figure III.4.4. Cleavage was observed only at guanine nucleotides as seen in the DMS cleavage pattern, no significant peaks were seen at positions of other bases. In the AFB1-Br_2 cleavage pattern, the reaction intensities at various guanines were quite different, suggesting that AFB1 does show sequence specificity in binding to *ras* gene. To provide a clearer picture of the results from the scan, the frequency spectrum of adduct formation among the various guanine sites for AFB1 is illustrated in Table 5. The assigned reference number of each guaninyl site was shown in Figure III.4.1. Based on the observed relative intensity of adduction, each site susceptibility was categorized into four classes: - as weak site (relative intensity < 0.4), + as intermediate site (0.4-0.6), ++ as strong site (0.6-0.8), and +++ as strongest site (0.8-1.0). For trout ki-*ras* gene, the guanines of codon 12, reference numbers of 15 & 16, showed differential accessibility, with the 5' G being an intermediate or weak site, the 3 G a strong

Figure III.4.4 Densitometric scans of guanine N-7 alkylation pattern produced by DMS and AFB1-Br_2. The intensity of each peak was normalized in each scan, and is shown as relative intensity. Based on the relative intensity, the susceptibility of each guanine site is categorized into four classes: - as weak site (< 0.4), + as intermediate site (0.4-0.6), ++ as strong site (0.6-0.8), and +++ as strongest site (0.8-1.0), which are listed in Table 5. Numbers above peaks indicate the position of guanines. Codon 12 is indicated by the arrow which corresponds to guanine 15 and 16 in the figure (cont. on page 127).

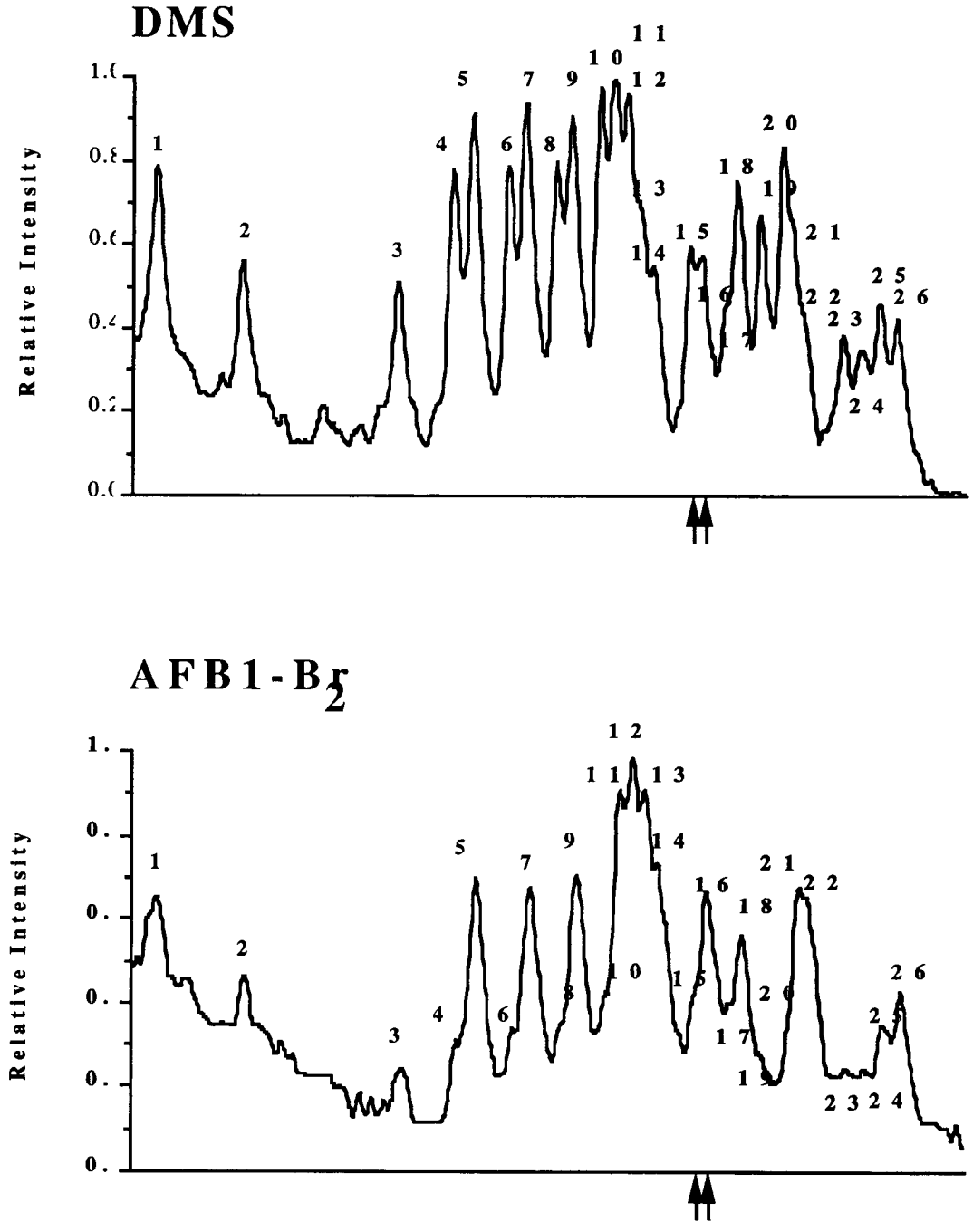


Figure III.4.4 Densitometric scans of guanine N-7 alkylation pattern produced by DMS and AFB1-Br2.

site. This result of differential accessibility of guanines in codon 12 suggests that the more accessible guanine in this codon may experience more frequent mutation than the less accessible one. This is supported by the result of Chang *et al.* (1991). By using PCR and the oligonucleotide hybridization method, they found from 14 AFB1-induced trout liver tumors that 10 out of 14 bear activating point mutations in the ki-*ras* gene. Of the 10 mutant *ras* genotypes, seven were codon 12 GGA-GTA transversions, two were codon 13 GGT-GTT transversions, and one was a codon 12 GGA-AGA transition. This higher incidence of point mutation at 3' G than the 5' G of codon 12 does agree the result that the accessibility of this site to AFB1 is also significantly higher.

However, the strongest sites were assigned to guanines 11, 12 and 13, which correspond to codons 10 and 11. These two codons have never been found to be mutated in human tumors studied so far. This result indicated that the predominance of point mutations detected *in vivo* at codon 12 does not reflect sequence-mediated preferential susceptibility of these sites to the initial DNA adduction.

In summary, the sequence interaction specificities of AFB1 in the codon 12 region of the trout ki-*ras* gene was examined by *invitro* alkylation experiment. Among the adducts produced, those at N-7-guanyl sites were alkali-labile and were identified on sequencing gel. The data indicated that the middle guanine of codon 12 has above average potential for adduct formation by AFB1 but the first guanine was less accessible. Neither represents an extremely accessible site.

Table 5

Site reactivity of AFB1-Br_2 to trout ki-*ras* gene. For each guanine position the reaction intensities were determined and then categorized with -, +, ++ and +++. - represents the weak site, + intermediate site, ++ strong site, and +++ strongest site. The trinucleotide sequence (Tri-N) is given for each position. The guanines of codon 12 are referenced number 15 and 16 respectively.

Guanine Ref. No.	Tri-N seq.	Reaction Intensity
1	TGA	++
2	AGA	+
3	AGC	-
4	TGG	-
5	GGT	++
6	TGG	-
7	GGT	++
8	TGG	-
9	GGT	
10	TGG	-
11	GGG	+++
12	GGG	+++
13	GGG	+++
14	GGC	++
15	AGG	+/-
16	GGA	++
17	AGG	-
18	GGT	+
19	TGT	-
20	TGG	-
21	GGG	++
22	GGC	++
23	AGC	-
24	CGC	-
25	CGC	+
26	AGC	+

IV. DISCUSSION:

IV 1. Instability of trout liver nuclei in the presence of detergent:

Isolation of nuclei from tissues is essential for research in many fields, including studies of chromatin structure of genes. The quality, and sometimes the quantity, of the nuclei from this initial step is very important, and may determine the success or failure of the research. By the term quality, I refer to the purity and stability of the nuclei. The purity is disturbed by the presence of cytoplasmic organelles such as mitochondria and other contaminants, including endoplasmic reticulum, cytoplasmic RNA *etc*,. By differential centrifugation in isotonic medium, most of the these contaminants can be removed from the nuclei. However, there is still a significant amount of endoplasmic reticulum, accompanied by ribosomes, attached to the outer nuclear membrane. To remove the outer nuclear membrane, detergent has traditionally been introduced into the procedure. Penman (1966) used a mixture of deoxycholate and Tween 40 to successfully clean Hela cell nuclei. But it was found that rat nuclei could only tolerate Triton X-100, not Tween-40 (Blobel & Potter, 1966). These results suggested that the nuclei from different tissues or species show different sensitivity to detergent. In the trout system, in particular, I have found the nuclei seem to be very sensitive to detergent treatment, using either NP-40 or Triton X-100. Although I did not test all kinds of detergent, it seems that trout liver nuclei are less tolerant of detergent than are those from other tissues or species.

The reason for this is unknown, but it must be considered in any study of trout liver chromatin.

IV.2. Properties of the 5' flanking sequences of trout ki-*ras* gene:

In these studies, a fragment surrounding the 5' end of the trout ki-*ras* gene was synthesized by inverse PCR technique, cloned into pUC 19 and sequenced by the dideoxynucleotide chain termination method. In addition to the sequences of exon 1 and part of exon 2, 1080 bp of 5' flanking sequence was also determined. To seek the evidence that the amplified gene was indeed the *ras* gene, the exon sequences were compared with that from Mangold *et al.*'s study (1991), and the result showed that two sequences matched very well, except only three degenerate bases in the exon 1, indicating that this PCP amplified sequence does represent a ki-*ras* gene..

IV.2.1. Conservation of the *ras* gene:

Sequence homology in the region of exon 1 and exon 2 was compared between the trout ki-*ras* gene and the human *ras* genes by Mangold *et al.* (1991), who found that the homology was 76.8-79% to the c-ki-*ras*, 77.6% to the N-*ras*, and about 83.4% to the c-Ha-*ras* gene. However, if the comparison of homology was conducted at the amino acid level, it was observed that the conservation of the sequence was as high as 99% to the c-ki-*ras* p21. Therefore the *ras* gene that Mangold *et al.* identified in trout was classified as a ki-*ras* gene as should be the one studied here. This feature of high

sequence homology supports the general conservation of the *ras* gene through evolution (Barbacid, 1987).

IV.2.2. The target *ras* gene in this study represents the expressed gene:

The original aims of these thesis studies were to characterize the 5' flanking region, and to study the chromatin structure around codon 12 region of the trout ki-*ras* gene. Since trout exhibits a large number of duplicated and diverging *ras* genes and psuedogenes, it is important to know that the ki-*ras* gene we studied represented the expressed one, because the expressed gene is the one for mediating the normal cellular function, and therefore the one responsible for the abnormality if mutated. So the final results will have significance to answer the questions we asked in this thesis.

This question was approached by the technique called "mismatch" PCR, in which a mismatch primer 12(G/T) was used to amplify the specific fragment from normal or tumor DNA of trout liver tissue, under the conditions that only allow the synthesis from the primer with the exact match of the sequence to the genomic DNA. The results (Figure III.2.4) strongly supported that the ki-*ras* gene we studied was the expressed gene.

IV.2.3. Properties of *ras* promoter elements:

The 5' flanking region is very important to a gene, not only in that this region contains promoter sequences for the initiation of the gene transcription, but in that it also contains many protein binding sites and regulatory sequences which are responsible for the

regulation of transcription. There are two kinds of sequences that are characteristics of most eukaryotic promoters: the TATA-box and the GC box. Studies of *ras* gene promoters from a number of species, mainly mammalian, have revealed that *ras* genes are in general TATA-less genes, sharing certain properties with other TATA-less genes, such as housekeeping genes (Osborne *et al.*, 1985; Valerio *et al.*, 1985; Masters and Attardi, 1985; McGrogan *et al.*, 1985; Melton *et al.*, 1984). A brief summary of these properties is listed in Table 6. However, when the 1080 bp of the 5' flanking sequence of this ki-*ras* gene from trout was carefully examined, it was found that this gene was somewhat unusual as compared to mammalian *ras* genes. There is a very low GC content (only about 32%) and there is no GC box (which is a binding site for Sp 1 factor). Interestingly, there are a nearly perfect TATA box and CAAT box located at -745 and -795 respectively, shown as the dotted region in Figure III.2.4. These are typical promoter elements for many eukaryotic genes. However, it has not been established that these sequences actually represent promoter elements for the trout ki-*ras* gene, since the transcriptional start sites have not been determined yet. It is not unreasonable that the trout ki-*ras* gene may have a different promoter structure than other higher eukaryotic *ras* genes, using a different regulation mechanism. Curiously, the only other *ras* gene known to have TATA and CCAAT boxes is the *ras* gene from *Drosophila* (Table 6), an organism evolutionarily very distant from trout.

Table 6

Characteristics of *ras* promoters

Species	Features	References
Mammalian	high GC contents	Ishii *et al.*, 1985; 1986
	multiple GC boxes	Hoffman *et al.*, 1987; 1990
	no TATA or CAAT	Nagase *et al.*, 1990
	Sp 1 binding sites	Lowndes *et al.*, 1989
	multiple transcriptional	Paciucci & Pellicer, 1991
	sites	Pintzas & Spandidos, 1991
	bi-directional	Honkawa *et al.*, 1987
Drosophila	TATA box, and CAAT box	Bishop III and Corces, 1988
	bi-directional	Lev *et al.*, 1988
Trout	(TAAT) at -745	This study
	(CAAT) at -795	

IV 3. Chromatin structure in the 5' flanking region of trout ki-*ras* gene.

IV.3.1. Chromatin structure:

In this study, we proposed to seek any chromatin structure of *ras* gene in the codon 12 region that might provide an explanation for the high level of chemical carcinogenesis induced by AFB1 in trout livers. Our working hypothesis was that codon 12 of the *ras* gene might be positioned in a nucleosome-free region (linker region or hypersensitive region), which could make guanines of this codon have more chance to interact with AFB1 in this tissue. Therefore in this study, the chromatin structure of trout ki-*ras* gene, in the range from -105 bp to +129 bp which includes the codon 12 region, was delineated by mapping of sites accessible to nuclease attack, using ligation-mediated PCR technique. Apparently no DNase I hypersensitive sites were located in this region. In addition, the result from studies with micrococcal nuclease digestion indicated that no ordered nucleosome structure was present in this range. This result was similar to the study by Kasid *et al.* (1984) who demonstrated that human c-Ha-*ras* chromatin lacked typical nucleosomal packaging. Although these two results cannot be compared directly since the regions studied on the genes was different, the overall tendency might be the same.

IV.3.2. LM-PCR technique:

To explain the above negative result, it might be argued that

the LM-PCR technique was not sufficiently sensitive to detect such structure in a single-copy gene in a large genome. Therefore the results might not represent the real structure *in vivo*. However, by doing a series of experiments as controls to test the reaction conditions in a step-by-step fashion, I have clearly demonstrated the needed sensitivity. Therefore, there is no reason to believe that this negative result was not correct.

In fact, the usability of LM-PCR technique is not really questioned at all. Pfeifer *et al.* (1989, 1990) have been able to study the *in vivo* footprinting and the methylation patterns of CpG islands of human x-chromosome-linked phosphoglycerate kinase 1 (PGK1) by LM-PCR aided genomic sequencing. They have also been able to map *in vivo* adductions of photoproducts in PGK1 by the same technique (1991). Mueller *et al.* (1989, 1991) have also successfully used this technique to study the footprints of a muscle specific enhancer. All of these reports showed very positively that LM-PCR is a powerful and sensitive technique. It can be used to study specific sites in a complicated genome much more easily than before. It makes direct genomic sequencing and footprinting not only possible but also practical, at the single base resolution.

IV.4. Sequence interaction specificities of AFB1 binding to trout ki-*ras* gene in the codon 12 region

A main thrust of this work was to examine the frequency of adduct formation at the two guanines of codon 12 of the *ras* gene, to see if the intrinsic affinity of these guanines to AFB1 is higher than that of other guanines in the sequence. This was determined in this

study by *in vitro* alkylation of exon 1 of the *ras* gene by AFB1. The result demonstrated that two guanines of codon 12 have different affinities to AFB1; the 3' guanine was an accessible site with strong affinity, but the 5' guanine was relatively less accessible (intermediate or weak site). Other guanines in exon 1 showed differential affinities, some of them accessible, some of them not accessible. The extremely accessible sites were guanines 11, 12, and 13, corresponding to the codons 10 and 11.

A similar study has been carried out by Marien *et al.* (1989) on the human c-Ha-*ras* gene. The guanines of codon 12, reference numbers of 14 & 15 in its sequence, were all strong targets, though were not the strongest. The similarity between our work and theirs was that the guanine(s) of codon 12 were accessible to AFB1, which may suggest that codon 12 could easily experience adduction *in vivo* with AFB1, especially at the 3' guanine site. However, codon 12 was not the strongest site for aflatoxins, indicating that the initial adduction frequency at guanine sites cannot account alone for pronounced tendency of aflatoxin to induce tumors in trout liver by codon 12 mutation. Some other still unknown mechanisms, e.g. an *in vivo* selection process, must be involved in these events.

The difference between these two studies was that in human c-Ha-*ras* gene, both guanines of codon 12 were strong targets, but in trout ki-*ras* gene, these two sites had differential accessibilities. This difference could be explained by the sequence effect of the nearest-neighbor-rule (Misra *et al.*, 1983). In the trout ki-*ras* gene, guanines of codon 12 (GGA) are flanked by adenines at both ends. By the nearest-neighbor-rule, the 5' G is a poor or intermediate site,

while the 3' G is a strong site. In the human c-Ha-*ras* gene, guanines of codon 12 (CGG) are flanked by cytidines at both sides. By the nearest-neighbor-rule, both Gs should be strong sites. These results seem to fit this rule quite well in both cases.

BIBLIOGRAPHY

Bailey, G. S. Nixon, J. E., Hendricks, J. D., Sinnhuber, R. O. and van Holde, K. E. (1980) "Carcinogen Aflatoxin B_1 is Located Preferentially in Internucleosomal Deoxyribonucleic Acid Following Exposure *in vivo* in Rainbow Trout." Biochem. 19:5836-5842.

Bailey, G. S., Hendricks, J. D., Nixon, J. E. and Pawlowski, N. E. (1984) "The Sensitivity of Rainbow Trout and Other Fish to Carcinogens." Drug Metabolism Rev. 15:725-750.

Barbacid, M. (1987) "*ras* Genes". Ann. Rev. Biochem. 56:779-827.

Bar-Sagi, D., Suhan, J. P., McCormick, F. and Feramisco, J. R. (1988) "Localization of Phospholipase A_2 in Normal and *ras*-transformed Cells." J. Cell. Biol. 106:1649-1658.

Bar-Sagi, D. (1989) "*Ras* Proteins: Biological Effects and Biochemical Targets." Anticancer Res. 9:1427-1437.

Benasutti, M., Ejadi, S., Whitlow, M. D. and Loechler, E. L. (1988) "Mapping the Binding Site of Aflatoxin B_1 in DNA: Systematic Analysis of the Reactivity of Aflatoxin B_1 with Guanines in Different DNA Sequences." Biochem. 27:472-481.

Bishop, J. M. (1983) "Cellular Oncogenes and Retroviruses." Ann Rev. Biochem. 52:301-354.

Bishop, J. M. (1987) "The Molecular Genetics of Cancer." Science. 235:305-311.

Bishop, J. G. 3d, and Corces, V. G. (1988) "Expression of an Activated *ras* Gene Causes Developmental Abnormalities in Transgenic *Drosophila melanogaster.*" Genes & Dev. 2:567-577.

Blobel, G. and Potter, V. R. (1966) "Nuclei from Rat Liver: Isolation Method that Combines Purity with High Yield." Science. 154:1662-1665.

Bloch, W. (1991) "A Biochemical Perspective of the Polymerase Chain Reaction." Biochemistry. 30:2735-2747.

Bos, J. L. (1989) "*ras* Oncogenes in Human Cancer: A Review." Cancer Res. 49: 4682-4689.

Boulikas, T. (1991) "Relation between Carcinogenesis, Chromatin Structure and Poly(ADP-ribosylation)." Anticancer Res. 11:489-527.

Bres, O. and Eales, J. G. (1988) "High-affinity, Limited-capacity Triindothyronine-binding Sites in Nuclei from Various Tissues of the Rainbow Trout (*Salmo gairdneri*)." General & Comparative Endocrinology. 69:71-79.

Chang, Y-J., Mathews, C., Mangold, K., Marien, K., Hendricks, J. and Bailey, G. (1991) "Analysis of *ras* Gene Mutations in Rainbow Trout Liver Tumors Initiated by Aflatoxin B_1." Mol. Carcinogenesis. 4:112-119.

Chan, D. C., Biard-Roche, J., Gorka, C., Girardet, J. L., Lawrence, J. J. and Piette, L. I. (1984) "The Role of the Central Globular Domain of Histone H5 in Chromatin Structure." J. Biomole. Struc. and Dynam. 2:319-332.

Chen, T. A. and Allfrey, V. G. (1987) "Rapid and Reversible Changes in Nucleosome Structure Accompany the Activation, Repression, and Superinduction of Murine Fibroblast Protooncogenes c-*fos* and c-*myc*." Proc. Natl. Acad. Sci. USA. 84: 5252-5256.

De Vos, A. M., Tong, L., Milburn, M. V., Matias, P. M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E. and Kim, S. H. (1988) "Three-Dimensional Structure of an Oncogene Protein: Catalytic Domain of Human c-Ha-*ras* p21." Science. 239:888-893.

Downward, J. (1990) "The *ras* Superfamily of Small GTP-binding Proteins." Trends in Biochem. Sci. 15:469-472.

Dyson, P. J., Littlewood, T. D., Forster, A. and Rabbitts, T. H. (1985) "Chromatin Structure of Transcriptionally Active and Inactive Human c-*myc* Alleles." EMBO J. 4:2885-2891.

Eckert, K. A. and Kunkel, T. A. (1991) "DNA Polymerase Fidelity and the Polymerase Chain Reaction." PCR Methods and Applications. 1:17-24.

Ehrlich, G. D., Glaser, J. B., LaVigne, K., Quan, D., Mildvan, D., Sninsky, J. J., Kwok, S., Papsidero, L. and Poiesz, B. J. (1989) "Prevalence of Human T-cell Leukemia/Lymphoma Virus (HTLV) Type II Infection Among High-Risk Individuals: Type-Specific Identification of HTLVs by Polymerase Chain Reaction." Blood. 74:1658-1664.

Eissenberg, J. C., Cartwright, I. L., Thomas, G. H. and Elgin, S. C. (1985) "Selected Topics in Chromatin Structure." Ann. Rev. Genet. 19:485-536.

Elgin, S. C. (1988) "The Formation and Function of DNase I Hypersensitive Sites in the Process of Gene Activation." J. Biol. Chem. 263:19259-19262.

Engelke, D. R., Hoener, P. A. and Collins, F. S. (1988) "Direct Sequencing of Enzymatically Amplified Human Gemonic DNA." Proc. Natl. Acad. Sci. USA. 85:544-548.

Erlich, H.A. ed. (1989) PCR Technology: Principles and Applications for DNA Amplification." Stockton Press, New York.

Erlich, H.A., Gibbs, R. and Kazazian, H. H. Jr. eds. (1989) "Current Communications in Molecular Biology: Polymerase Chain Reaction." Cold Spring Harbor Laboratory Press.

Essigmann, J. M., Croy, R. G., Nadzan, A. M., Busby, W. F. Jr., Reinhold, V. N., Buchi, G. and Wogan, G. N. (1977) "Structural Identification of the Major DNA Adduct formed by Aflatoxin B_1 *in vitro*." Proc. Natl. Acad. Sci. USA. 74:1870-1874.

Fahmy, M. J., Fahmy, O. G. and Swenson, D. H. (1978) "Aflatoxin B_1-2, 3-dichloride as a Model of the Active Metabolite of Aflatoxin B_1 in Mutagenesis and Carcinogenesis." Cancer Res. 38:2608-2616.

Fahrlander, P. D., Piechaczyk, M. and Marcu, K. B. (1985) "Chromatin Structure of the Murine c-*myc* Locus: Implications for the Regulation of Normal and Chromosomally Translocated Genes." EMBO J. 4:3195-3202.

Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M. and Wigler, M. (1984) "Analysis of the Transforming Potential of the Human H-*ras* Gene by Random Mutagenesis." Proc. Natl. Acad. Sci. USA. 81:4008-4012.

Fleischman, L. F., Chahwala, S. B. and Cantley, L. (1986) "*Ras*-Transformed Cells: Altered Levels of Phosphatidylinositol-4, 5-Bisphosphate and Catabolites." Science. 231:407-410.

Fors, L., Saavedra, R. A. and Hool, L. (1990) "Cloning of the Shark P_O Promoter Using a Genomic Walking Technique Based on the Polymerase Chain Reaction." Nucleic Acids Res. 18:2793-2799.

Frohman, M. A., Dush, M. K. and Martin, G. R. (1988) "Rapid Production of Full-length cDNAs from rare Transcripts: Amplification using a Single Gene-specific Oligonucleotide Primer." Proc. Natl. Acad. Sci. USA. 85:8998-9002.

Gilman, A. G. (1984) "G Proteins and Dual Control of Adenylate Cyclase." Cell. 36:577-579.

Gilman, A. G. (1987) "G Proteins: Transducers of Receptor Generated Signals." Ann. Rev. Biochem. 56:615-649.

Groopman, J. D., Croy, R. G. and Wogan, G. N. (1981) "*In Vitro* Reactions of Aflatoxin B1-Adducted DNA." Proc. Natl. Acad. Sci. USA. 78:5445-5449.

Gross, D. S. and Garrard, W. T. (1988) "Nuclease Hypersensitive Sites in Chromatin." Ann. Rev. Biochem. 57: 159-197.

Guerrero, I. and Pellicer, A. (1987) "Mutational Activation of Oncogenes in Animal Model Systems of Carcinogenesis." Mutation Res. 185:293-308.

Gyllensten, U. B. and Erlich, H. A. (1988) "Generation of Single-Stranded DNA by the Polymerase Chain Reaction and its Application to Direct Sequencing of the HLA-DQA Locus." Proc. Natl. Acad. Sci. USA. 85:7652-7656.

Hall, A. (1984)" The *ras* Gene Family." Oxford Surveys on Eukaryotic Genes. 1:111-144.

Hill, C. S., Rimmer, J. M., Green, B. N., Finch, J. T. and Thomas, J. O. (1991) "Histone-DNA Interactions and Their Modulation by Phosphorylation of -Ser-Pro-X-Lys/Arg- Motifs." EMBO J. 10:1939-1948.

Hoffman, E. K., Trusko, S. P., Freeman, N. and George, D. L. (1987) "Structural and Functional Characterization of the Promoter Region of the Mouse c-ki-*ras* Gene." Mol. Cell. Biol. 7:2592-2596.

Hoffman, E. K., Trusko, S. P., Murphy, M. and George, D. L. (1990) "An S1 Nuclease-sensitive Homopurine/Homopyrimidine Domain in the c-ki-*ras* Promoter Interacts with a Nuclear Factor." Proc. Natl. Acad. Sci. USA. 87:2705-2709.

Holtzman, E., Smith, I. and Penman, S. (1966) "Electron Microscopic Studies of Detergent-treated HeLa Cell Nuclei." J. Mol. Biol. 17:131-135.

Honkawa, H., Masahashi, W., Hashimoto, S. and Hashimoto-Gotoh, T. (1987) "Identification of the Principal Promoter Sequence of the c-H-*ras* Transforming Oncogene: Deletion Analysis of the 5'-Flanking Region by Focus Formation Assay." Mol. Cell. Biol. 7:2933-2940.

Innis, M. A., Myambo, K. B., Gelfand, D. H. and Brow, M. A. (1988) "DNA Sequencing with *Thermus aquaticus* DNA Polymerase and Direct Sequencing of Polymerase Chain Reaction-Amplified DNA." Proc. Natl. Acad. Sci. USA. 85:9436-9440.

Innis, M. A., Gelfaud, D. H., Sninsky, J. J. and White, T. J. eds. (1990) "PCR Protocols: A Guide to Methods and Applications." Academic Press, Inc., New York.

Ishii, S., Merlino, G. T. and Pastan, I. (1985) "Promoter Region of the Human Harvey *ras* Proto-oncogene: Similarity to the EGF Receptor Proto-oncogene Promoter." Science. 230:1378-1381.

Ishii, S., Kadonaga, J. T., Tjain, R., Brady, J. N., Merlino, G. T. and Pastan, I. (1986) "Binding of the Sp1 Transcription Factor by the Human Harvey *ras*1 Proto-oncogene Promoter." Science. 232:1410-1413.

Kasid, U. N., Hough, C., Thraves, P., Dritschilo, A. and Smulson, M. (1985) "The Association of Human c-Ha-*ras* Sequences with Chromatin and Nuclear Proteins." Biochem. Biophy. Res. Commu. 128:226-232.

Keene, M. A. and Elgin, S. C. (1981) "Micrococcal Nuclease as a Probe of DNA Sequence Organization and Chromatin Structure." Cell. 27:57-64.

Kleinschmidt, J. A. and Steinbeisser, H. (1991) "DNA-dependent Phosphorylation of Histone H2A.X During Nucleosome Assembly in *Xenopus Laevis* Oocytes: Involvement of Protein Phosphorylation in Nucleosome Spacing." EMBO J. 10:3043-3050.

Kumar, S. and Leffak, M. (1989) "DNA Topology of the Ordered Chromatin Domain 5' to the Human c-*myc* Gene." Nucleic Acids Res. 7:2819-2833.

Kumar, S. and Leffak, M. (1991) "Conserved Chromatin Structure in c-*myc* 5' flanking DNA after Viral Transduction." J. Mol. Biol. 222:45-57.

Kunkel, G. R. and Martinson, H. G. (1981) "Nucleosomes Will not Form on Double-stranded RNA or Poly(dA).Poly(dT) Tracts in Recombinant DNA." Nucleic Acids. Res. 9:6869-6888.

Kwok, S., Mack, D. H., Kellogg, D. N., Mckinney, N., Faloona, F. and Sninsky, J. J. (1989) "Application of the Polymerase Chain Reaction to the Detection of Human Retroviruses." in Current Communications in Molecular Biology (Erlich *et al.* eds.). Cold Spring Harbor Laboratory Press. pp151-158.

Lev, Z., Segev, O., Cohen, N., Salzberg, A. and Shemer, R. (1988) "Structure of the *Drosophila ras*2 Bidirectional Promoter." in "*ras* Oncogenes." edited by Demetrios Spandidos, NATO ASI Series, Series A: Life Science, Vol. 170:75-81.

Levitzki, A. (1988) "Transmembrane Signalling to Adenylate Cyclase in Mammalian Cells and in *Saccharomyces Cerevisiae.*" Trends in Biochem. 13:298-301.

Lin, J., Miller, J. A. and Miller, E. C. (1977) "2, 3-Dihydro-2-(guan-7-yl)-3-hydroxy-aflatoxin B_1 a Major Acid Hydrolysis Product of Aflatoxin B_1-DNA or -ribosomal RNA Adducts Formed in Hepatic Microsome-mediated Reactions and in Rat Liver *in vivo.*" Cancer Res. 37:4430-4438.

Lowndes, N. F., Paul, J., Wu, J. and Allan, M. (1989) "c-Ha-*ras* Gene Bidirectional Promoter Expressed *in vitro*: Location and Regulation." Mol. Cell. Biol. 9:3758-3770.

Lowrey, C. H., Bodine, D. M. and Nienhuis, A. W. (1992) "Mechanism of DNase I Hypersensitive Site Formation within the Human Globin Locus Control Region." Proc. Natl. Acad. Sci. USA. 89:1143-1147.

Mangold, K., Chang, Y-J., Mathews, C., Marien, K., Hendricks, J. and Bailey, G. (1991) "Expression of *ras* Genes in Rainbow Trout Liver." Mol. Carcinogenesis. 4:97-102.

Marien, K., Moyer, R., Loveland, P., van Holde, K. E. and Bailey, G. (1987) "Comparative Binding and Sequence Interaction Specificities of Aflatoxin B_1, Aflatoxicol, Aflatoxin M_1, and AFlatoxicol M_1 with Purified DNA." J. Biol. Chem. 262:7455-7462.

Marien, K., Mathews, K., van Holde, K. and Bailey, G. (1989) "Replication Blocks and Sequence Interaction Specificities in the Codon 12 Region of the c-Ha-*ras* Proto-oncogene Induced by Four Carcinogenes *in vitro*." J. Biol. Chem. 264:13226-13232.

Masters, J. N. and Attardi, G. (1985) "Discrete Human Dihydrofolate Reductase Gene Transcripts Present in Polysomal RNA map with their 5' Ends Several Hundred Nucleotides Upstream of the Main mRNA Start Sites." Mol. Cell. Biol. 5:493-500.

McCormick, F. (1989) "*ras* GTPase Activating Proteins: Signal Transmitter and Signal Terminator." Cell. 56:5-8.

McGrogan, J. P., Simonsen, C. C., Smouse, D. T., Farnham, P. J. and Schimke, R. T. (1985) "Heterogeneity at the 5' Termini of Mouse Dihydrofolate Reductase mRNAs: Evidence for Multiple Promoter Regions." J. Biol. Chem. 260:2307-2314.

McMahon, G., Hanson, L., Lee, J. J. and Wogan, G. N. (1986) "Identification of an Activated c-Ki-*ras* Oncogene in Rat Liver Tumors Induced by Aflatoxin B_1." Proc. Natl. Acad. Sci. USA. 83:9418-9422.

Mcmahon, G., Davis, E. and Wogan, G. N. (1987) "Characterization of c-Ki-*ras* Oncogene Alleles by Direct Sequencing of Enzymatically Amplified DNA from Carcinogene-Induced Tumors." Proc. Natl. Acad. Sci. USA. 84:4974-4978.

Mcmahon, G., Davis, E. F., Huber, L. J., Kim, Y. and Wogan, G. N. (1990) "Characterization of c-Ki-*ras* and N-*ras* Oncogenes in Aflatoxin B_1-Induced Rat Liver Tumors." Proc. Natl. Acad. Sci. USA. 87:1104-1108.

Melton, D. W., Konecki, D. S., Brennand, J. and Caskey, C. T. (1984) "Structure, Expression, and Mutation of the Hypoxanthine Phosphoribosyltransferase Gene." Proc. Natl. Acad. Sci. USA. 81:2147-2151.

Milburn, M. V., Tong, L., deVos, A. M., Brunger, A., Yamaizumi, Z., Nishimura, S. and Kim, S. H. (1990) "Molecular Switch for Signal Transduction: Structural Differences between Active and Inactive Forms of Protooncogenic *ras* Proteins" Science. 247:939-945.

Miller, E. C. and Miller, J. A. (1981) "Searches for Ultimate Chemical Carcinogens and their Reactions with Cellular Macromolecules." Cancer. 47:2327-2345.

Misra, R. P. Muench, K. F. and Humayun, M. Z. (1983) "Covalent and Noncovalent Interactions of Aflatoxin with Defined Deoxyribonucleic Acid Sequences." Biochem. 22:3351-3359.

Moyer, R., Marien, K., van Holde, K. E. and Bailey, G. (1989) "Site-specific Aflatoxin B_1 adduction of Sequence-positioned Nucleosome Core Particles." J. Biol. Chem. 264:12226-12231.

Mueller, P. R. and Wold, B. (1989) "*In vivo* Footprinting of a Muscle Specific Enhancer by Ligation Mediated PCR." Science. 246:780-786.

Mueller, P. R. and Wold, B. (1991) "Ligation-Mediated PCR: Applications to Genomic Footprinting." Methods: A Companion to Methods in Enzymology. 2:20-31.

Muench, K. F., Misra, R. P. and Humayun, M. Z. (1983) "Sequence Specificity in Aflatoxin B_1-DNA Interactions." Proc. Natl. Acad. Sci. USA. 80:6-10.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) "Specific Enzymatic Amplification of DNA *in vitro*: The Polymerase Chain Reaction." Cold Spring Harbor Symposia on Quantitative Biology. 51:263-273.

Mullis, K. B. (1991) "The Polymerase Chain Reaction in an Anemic Mode: How to Avoid Cold Oligodeoxyribonuclear Fusion." PCR Methods and Applications. 1:1-4.

Nagase, T., Ueno, Y. and Ishii, S. (1990) "Transcriptional Control of the Human Harvey *ras* Proto-oncogene: Role of Multiple Elements in the Promoter Region." Gene. 94:249-253.

Nedospasov, S. A. and Georgiev, G. P. (1980) "Non-Random Cleavage of SV40 DNA in the Compact Minichromosome and Free in Solution by Micrococcal Nuclease." Biochem. Biophys. Res. Commu. 92:532-539.

Ochman, H., Gerber, A. S. and Hartl, D. L. (1988) "Genetic Applications of an inverse polymerase Chain Reaction." Genetics. 120:621-623.

Ohara, O., Dorit, R. L. and Gilbert, W. (1989) "One-Sided Polymerase Chain Reaction: The Amplification of cDNA." Proc. Natl. Acad. Sci. USA. 86:5673-5677.

Osborne, T. F., Goldstein, J. L. and Brown, M. S. (1985) "5' End of HMG CoA Reductase Gene Contains Sequences Responsible for Cholesterol-mediated Inhibition of Transcription." Cell. 42:203-212.

Paciucci, R. and Pellicer, A. (1991) "Dissection of the Mouse N-*ras* Gene Upstream Regulatory Sequences and Identifiation of the Promoter and a Negative Regulatory Element." Mol. Cell. Biol. 11:1334-1343.

Parker, J. D., Rabinovitch, P. S. and Burmer, G. C. (1991) "Targeted Gene Walking Polymerase Chain Reaction." Nucleic Acids Res. 19:3055-3060.

Pfeifer, G. P., Steigerwald, S. D., Mueller, P. R., Wold, B. and Riggs, A. D. (1989) "Genomic Sequencing and Methylation Analysis by Ligation Mediated PCR." Science. 246:810-813.

Pfeifer, G. P., Tanguay, R. L., Steigerwald, S. D. and Riggs, A. D. (1990) "*In vivo* Footprint and Methylation Analysis by PCR-aided Genomic Sequencing: Comparison of Active and Inactive X Chromasomal DNA at the CpG Island and Promoter of Human PGK-1." Genes & Dev. 4:1277-1287.

Pfeifer, G. P., Steigerwald, S. D., Hansen, R. S., Gartler, S. M. and Riggs, A. D. (1990) "Polymerase Chain Reaction-aided Genomic Sequencing of an X Chromosome-Linked CpG Island: Methylation Patterns Suggest Clonal Inheritance, CpG Site Autonomy and an Explanation of Actvity State Stability." Proc. Natl. Acad. Sci. USA. 87:8252-8256.

Pfeifer, G. P. and Riggs, A. D. (1991) "Chromatin Difference Between Active and Inactive X Chromosomes Revealed by Genomic Footprinting of Premeabilized Cells using DNase I and Ligation-Mediated PCR." Genes & Dev. 5:1102-1113.

Pintzas, A. and Spandidos, D. A. (1991) "Sp1 Specific Binding Sites Within the Human H-*ras* Promoter: Potential Role of the 6 bp Deletion Sequence in the T24 H-*ras*1 Gene." Anticancer Res. 11:2067-2070.

Reeves, R. (1984) "Transcriptionally Active Chromatin." Biochimica et Biophysica Acta. 782:343-393.

Refolo, L. M., Conley, M. P, Sambamurti, K., Jacobsen, J. S. and Humayun, M. Z. (1985) "Sequence Context Effects in DNA Replication Blocks Induced by Aflatoxin B_1." Proc. Natl. Acad. Sci. USA. 82:3096-3100.

Rodericks, J. V., Stolofft, L., Pons, W. A., Robertson, J. A. and Goldblatt, L. A. (1970) "Molar Absorptivity Values for Aflatoxins and Justification for Their Use as Criteria of Purity of Analytical Standards." JAOAC. 53:96-101.

Rotenberg, M. O., Chow, L. T. and Broker, T. R. (1989) "Characterization of Rare Human Papillomavirus Type 11 mRNAs Coding for regulatory and Structural Proteins, Using the Polymerase Chain Reaction." Virology. 172:489-497.

Ruano, G. and Kidd, K. K. (1991) "Coupled Amplification and Sequencing of Genomic DNA." Proc. Natl. Acad. Sci. USA. 88:2815-2819.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N. (1985) "Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science. 230:1350-1354.

Saiki, R. K., Gelfaud, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988) "Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA polymerase." Science. 239:487-491.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. (second edition). page C.1.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) "DNA Sequencing with Chain-terminating inhibitors." Proc. Natl. Acad. Sci. USA. 74:5463-5467.

Scharf, S. J., Horn, G. T. and Erlich, H. A. (1986) "Direct Cloning and Sequencing Analysis of Enzymatically Amplified Genomic Sequences." Science. 233:1076-1078.

Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F.and Petsko, G. A. (1990) "Time-resolved X-Ray Crystallographic Study of the Conformational Change in Ha-*Ras* p21 Protein on GTP Hydrolysis." Nature. 345:309-315.

Schubach, W. and Groudine, M. (1984) "Alteration of c-*myc* Chromatin Structure by Avian Leukosis Virus Integration." Nature. 307:702-708.

Scott, W. A. and Wigmore, D. J. (1978) "Sites in Simian Virus 40 Chromatin which are Preferentially Cleaved by Endonucleases." Cell. 15:1511-1518.

Shih, T. Y. and Weeks, M. O. (1984) "Oncogenes and Cancer: The p21 *ras* Genes." Cancer Invest. 2:109-123.

Siebenlist, U., Hennighausen, L., Battey, J. and Leder, P. (1984) "Chromatin Structure and Protein Binding in the Putative Regulatory Region of the c-*myc* Gene in Burkitt Lymphoma." Cell. 37:381-391.

Silver, J. and Keerikatte, V. (1989) "Novel Use of Polymerase Chain Reaction to Amplify Cellular DNA Adjacent to an Integrated Provirus." J Viol. 63:1924-1928.

Simpson, R. T. and Stafford, D. W. (1983) "Structural Features of a Phased Nucleosome Core Particles." Proc. Natl. Acad. Sci. USA. 80:51-55.

Simpson, R. T., Thoma, F. and Brubaker, J. M. (1985) "Chromatin Reconstituted from Tandomly Repeated Cloned DNA Fragments and Core Histones: A Model System for Study of Higher Order Structure." Cell. 42:799-808.

Sinha, S., Webber, C., Marshall, C. J., Knowles, M. A., Proctor, A., Barrass, N. C. and Neal, G. E. (1988) "Activation of *ras* Oncogene in Aflatoxin-Induced Rat Liver Carcinogenesis." Proc. Natl. Acad. Sci. USA. 85:3673-3677.

Spiegel, A. M. (1987) "Signal Tranduction by Guanine Nucleotide binding Proteins." Mol. Cell. Endocrinol. 49:1-16.

Steigerwald, S. D., Pfeifer, G. P. and Riggs, A. D. (1990) "Ligation-Mediated PCR Improves the Sensitivity of Methylation Analysis by Restriction Enzymes and Detection of Specific DNA Strand Breaks." Nucleic Acids Res. 18:1435-1439.

Thoma, F., Bergman, L. W. and Simpson, R. T. (1984) "Nuclease Digestion of Circular TRP1ARS1 Chromatin Reveals Positional Nucleosomes Separated by Nuclease-Sensitive Regions." J. Mol. Biol. 177:715-733.

Thoma, F. and Simpson, R. T. (1985) "Local Protein-DNA Interactions May Determine Nucleosome Positions on yeast Plasmids." Nature. 315:250-252.

Trahey, M. and McCormick, F. (1987) "A Cytoplasmic Protein Stimulates Normal N-*ras* p21 GTPase, but does not Affect Oncogenic Mutants." Science. 238:542-545.

Triglia, T., Peterson, M. G. and Kemp, D. J. (1988) "A Procedure for *in vitro* Amplification of DNA Segments that Lie outside the Boundaries of Known Sequences." Nucleic Acids Res. 16:8186.

Tullius, T. D. and Lippard, S. J. (1982) "Ethidium Bromide Changes the Nuclease-sensitive DNA Binding Sites of the Antitumor Drug cis-diamminedichloroplatinum (II)." Proc. Natl. Acad. Sci. USA. 79:3489-3492.

Valerio, D., Duyvesteyn, M. G., Dekker, B. M., Weeda, G., Berkvens, T. M., van der Voorn, L., van Ormondt, H. and van der Eb, A. J. (1985) "Adenosine Deaminase: Characterization and Expression of a Gene with a Remarkable Promoter." EMBO J. 4:437-443.

van Holde, K. E. (1989) "Chromatin." Springer-Verlag, New York.

Vanderbilt, J. N., Bloom, K. S. and Anderson J. N. (1982) "Endogenous Nuclease: Properties and Effects on Transcribed Genes in Chromatin." J. Biol. Chem. 257:13009-13017.

Varmus, H. E. (1984) "The Molecular Genetics of Cellular Oncogenes" Ann. Rev. Genet. 18:553-612.

Varshavsky, A. J., Sundin, O. H. and Bohn, M. J. (1978) "SV40 Viral Minichromosome: Preferential Exposure of the Origin of Replication as probed by Restriction Endonucleases." Nucleic Acids Res. 5:3469-3477.

Wang, B. and Sodja, A. (1991) "Alternate Approach to Sequencing Double Stranded Template DNAs." BioTechniques 10:198-201.

Weber, J. L. and May, P. E.(1989) "Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction." American J. Human Genet. 44:388-396.

Welsh, J., Petersen, C. and McClelland, M. (1991) "Polymorphisms Generated by Arbitrarily Primed PCR in the Mouse: Application to Strain Identification and Genetic Mapping." Nucleic Acids Res. 19:303-306.

Wolfman, A. and Macara, I. G. (1987) "Elevated Levels of Diacylgylcerol and Decreased Phorbol ester Sensitivity in *ras* Transformed Fibroblasts." Nature. 325:359-361.

Wong, C., Dowling, C. E., Saiki, R. K., Higuchi, R. G., Erlich, H. A. and Kazazian, H. H. Jr. (1987) "Characterization of beta-Thalassaemia Mutations Using Direct Genomic Sequencing of Amplified Single Copy DNA." Nature. 330:384-386.

Wu, C. (1980) "The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I." Nature 286:854-860.

Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R. and Elgin, S. C. R. (1979) "The Chromatin Structure of Specfic Genes: I. Evidence for High Order Domains of Defined DNA Sequence." Cell. 16:797-806.

Yaniv, M. and Cereghini, S. (1986) "Structure of Transcriptionally Active Gene." CRC Critical Rev. in Biochem. 21:1-26.

Yu, F., L., Bender, W. and Geronimo, I. H. (1990) "Base and Sequence Specificities of Aflatoxin B_1 Binding to Single- and Double-stranded DNAs." Carcinogenesis. 11:475-478.

Zarbl, H., Sukumar, S., Arthur, A. V., Martin-Zanca, D. and Barbacid, M. (1985) "Direct Mutagenesis of Ha-*ras*-1 Oncogenes by N-nitroso-N-methylurea During Initiation of Mammary Carcinogenesis in Rats." Nature. 315:382-385.