CHARACTERIZATION OF THE CHICKEN PHENOBARBITAL INDUCIBLE P450 GENE FAMILY

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE UNIVERSITY OF ADELAIDE

BY

LISA ANNE ELFERINK (nee Mattschoss) B. Sc. (Hons.)

DEPARTMENT OF BIOCHEMISTRY

FEBRUARY 1987

CONTENTS

BIBLIOGRAPHY.

LIST OF PUBLICATIONS.

THESIS SUMMARY

This thesis describes an investigation into the chicken phenobarbital (PB) inducible P450 gene family. The aims of this study fell into two broad categories:

(1) The construction and characterization of cDNA clones complementary to the chicken PB inducible P450s.
(2) The use of these clones as probes to isolate the corresponding genes and to determine the complexity of this P450 gene family by Southern hybridization analysis.

Such studies may ultimately lead to an understanding of the control of P450 gene expression.

Since a combination of the drugs 2-allyl-2-isopropylacetamide (AIA) and 3,5-diethoxycarbonyl-1,4-dihydrocollidine(DDC) has been shown in this laboratory to promote a greater induction of P450 mRNA in the chick embryo than PB alone, poly A^+ RNA was isolated from the livers of 18 day old chick embryos treated with these drugs. This mRNA was shown to direct the synthesis of a 50,000 molecular weight P450 protein in a cell free translation system, and was subsequently used in the construction of a cDNA library. Colony hybridization with a 720 bp cDNA insert from a previously isolated PB inducible chicken P450 clone detected the presence of ten homologous cDNA clones.

The cDNA insert of the largest recombinant, pCHP3, was fully sequenced and represents the first non-mammalian eukaryotic P450 sequence to be characterized. The reading frame encoded a predicted protein of 491 residues with the first fifteen amino acids confirmed by amino terminal sequencing of a major P450 protein purified from the livers of AIA treated chick embryos. Northern hybridization analysis of AIA and DDC induced hepatic chick embryo mRNA detected the presence of three mRNA species measuring 3.5, 2.8 and 2.2 Kb. All three mRNAs were inducible by each of the drugs AIA, DDC and PB and are not the consequence of polymorphisms within the chicken population. The 2.7 Kb cDNA insert of pCHP3 was shown by Northern hybridization analysis to be derived from the 3.5 Kb mRNA. Sequence analysis revealed that pCHP3 contained the complete 5' noncoding region of the 3.5 Kb mRNA but lacked a significant portion of the 3'noncoding region.

In an attempt to obtain the complete primary sequence of the 3.5 Kb mRNA, a second cDNA library was constructed using a procedure which preferentially clones the 3' end of mRNAs. Several cDNA clones were isolated and sequence analysis of one of these clones pCHPB15, extended the sequence of pCHP3 by 134 nucleotides in the 3' direction. Collectively however, the clones pCHP3 and pCHPB15, contain only 2846 nucleotides of sequence representing the 3.5 Kb mRNA.

The cDNA clones pCHP3 and pCHPB15 were used as probes to screen two chicken genomic libraries and several λ clones were isolated. Restriction mapping of these clones indicated that they represented two highly homologous genes termed A and B. Extensive Southern hybridization analysis of chicken genomic DNA revealed that these genes are nonallelic and exist as single copies per haploid genome. The presence of only two members in the chicken PB inducible P450 gene family, makes it the smallest PB inducible gene family characterized to

date. This is in direct contrast to the mammalian situation, where six to ten genes have been reported in different species.

Southern hybridization analysis using DNA probes specific for the 3.5 Kb and 2.2 Kb mRNAs established that both the chicken P450 genes are transcriptionally active. The 3.5 Kb mRNA is derived from the A gene whereas the 2.2 Kb mRNA is derived from the B gene. Since a third PB inducible P450 gene has not been detected in the chicken genome, the relationship of the 2.8 Kb mRNA to the two chicken P450 genes is presently unclear. It is possible that two of the P450 mRNAs are derived from a single gene either by alternative splicing and/or polyadenylation.

A λ clone containing the promoter for the chicken P450 A gene has been partially sequenced. The promoter contains a canonical TATA sequence and an inverted CCAAT sequence at the expected positions. The region sequenced also covers the first three exons of the gene, each exon/intron boundary conforming to the GT-AG rule. The sequence extends 1 Kb upstream from the site of transcription initiation. Computer aided comparison of the 5' flanking sequence of the P450 A gene with the sequences of PB inducible P450 genes from other species, did not detect the presence of potential regulatory sequences.

DECLARATION

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

In all of the experiments described in this thesis, the author was the principal worker. However, in Chapters 3 and 5, Dr. A. Hobbs and Prof. A. Steggles collaborated in the work due to the time consuming nature of the sequencing and screening techniques involved.

To the best of my knowledge and belief, this thesis contains no material previously published or written except where due reference is made in the text.

LISA ANNE ELFERINK (nee Mattschoss).

ACKNOWLEDGEMENTS

I wish to express my gratitude to the following people:

Prof. W. H. Elliott for an opportunity to work in the Department of Biochemistry.

Dr. B. K. May and Prof. W. H. Elliott for their support and guidance throughout the course of this work.

The past and present members of the group, for providing an enjoyable atmosphere in which to work. Parts of the work were undertaken in collaboration with Dr. A. Hobbs and Prof. A. W. Steggles, for which I am most grateful.

To those involved in the preparation of this thesis.

My family and friends for their loving support and continual interest.

My husband Kees, for his patience and encouragement.

During the course of this work, I was supported by a Commonwealth Postgraduate Research Award.

ABBREVIATIONS

AIA: 2-allyl-2-isopropylacetamide

A_n: Absorbance at wavelength, n

bp: base pairs

BSA: bovine serum albumin

cDNA: DNA complementary to mRNA

Ci: Curie

cpm: counts per minute

DDC: 3,5-diethoxycarbonyl-1,4-dihydrocollidine

DMSO: dimethyl sulphoxide

DNA: deoxyribonucleic acid

Kb: kilobases

PB: phenobarbital

RNA: ribonucleic acid — mRNA: mesenger RNA; rRNA: ribosomal RNA; tRNA: transfer RNA

rpm: revolutions per minute

S: Svedberg unit of sedimentation coefficient

SDS: sodium dodecyl sulphate

U.V.: ultraviolet

CHAPTER 1

INTRODUCTION.

1.1. GENERAL INTRODUCTION.

The tetrapyrrole heme serves as the prosthetic group for a large number of hemoproteins. These hemoproteins mediate essential physiological functions including the transport of oxygen (myoglobin and hemoglobin), the disposal of hydrogen peroxide (catalase and peroxidase), oxidative phosphorylation (mitochondrial cytochromes), the biosynthesis of steroid hormones (adrenal cytochromes) and the metabolism of steroids, fatty acids and xenobiotics (microsomal cytochromes). The mammalian liver synthesizes approximately 15% of the body's heme requirements, the remainder being accounted for by the erythropoietic cells of the bone marrow (Kappas et al., 1983).

In this laboratory we are investigating the regulation of hepatic heme biosynthesis. The synthesis of hepatic heme is controlled primarily by the activity of the first enzyme in the pathway, 5-amino-levulinate synthase (ALV-synthase). The level of this enzyme is normally low and regulated through negative feedback repression by the end product heme (De Matteis, 1973). However, ALV-synthase can be greatly induced in the livers of experimental animals by a wide variety of xenobiotics which also induce the synthesis of multisubstrate monooxygenases, the P450s.[1] As the synthesis of P450 constitutes one of the major requirements for hepatic heme, it has been proposed that the drug induced increases in ALV-synthase are a direct consequence of changes in the level of the intracellular free heme pool (May et al., 1986). In this proposal, the newly synthesized P450 apoprotein would conjugate the free heme, thereby removing the inhibitory effect of heme on ALV-synthase production. Studies in this group are therefore

1. The terms "cytochrome P450" and "P450" are used interchangeably to describe a hemoprotein whose ferrous-carbon monoxide complex exhibits a Soret absorption band at 450 nm and which functions as a multisubstrate monooxygenase or oxygen transfer agent.

directed towards understanding the relationship between the drug induced synthesis of both hepatic ALV-synthase and P450s, at the molecular level. The development of recombinant DNA technology has enabled the identification and isolation of the genes for these proteins. This will permit an investigation into the regulation of their expression. The research presented in this thesis focuses on the regulation of P450s. As this area of research is rapidly expanding and the literature is complex, the remainder of this chapter includes a condensed account of the current status of research on the P450s. Firstly, some general properties of the P450 superfamily are dealt with, followed by a discussion of specific members, with emphasis on the mechanism of drug induction and the control of gene expression. For a more detailed discussion of the P450s, the reader is directed towards a number of reviews which have recently appeared (Whitlock Jr., 1986; Adesnik and Atchison, 1985; Nebert and Gonzalez, 1985; Goldstein, 1984; Iverson et al., 1986; Schwab and Johnson, 1986).

1.2. THE CYTOCHROME P450 SUPERFAMILY.

Mammalian P450s constitute a diverse superfamily of membrane bound enzymes. They are involved in a large number of critical body functions, including the synthesis of steroid hormones and the metabolism of a wide variety of lipophilic compounds. These include exogenous compounds (collectively referred to as xenobiotics) and endogenous steroids and fatty acids. Many of the compounds metabolized by this enzyme system are converted to more polar derivatives, which are readily excreted from the organism. However, some compounds can be activated to reactive intermediates which are highly carcinogenic. While the liver is the principle site of P450s, lower concentrations are

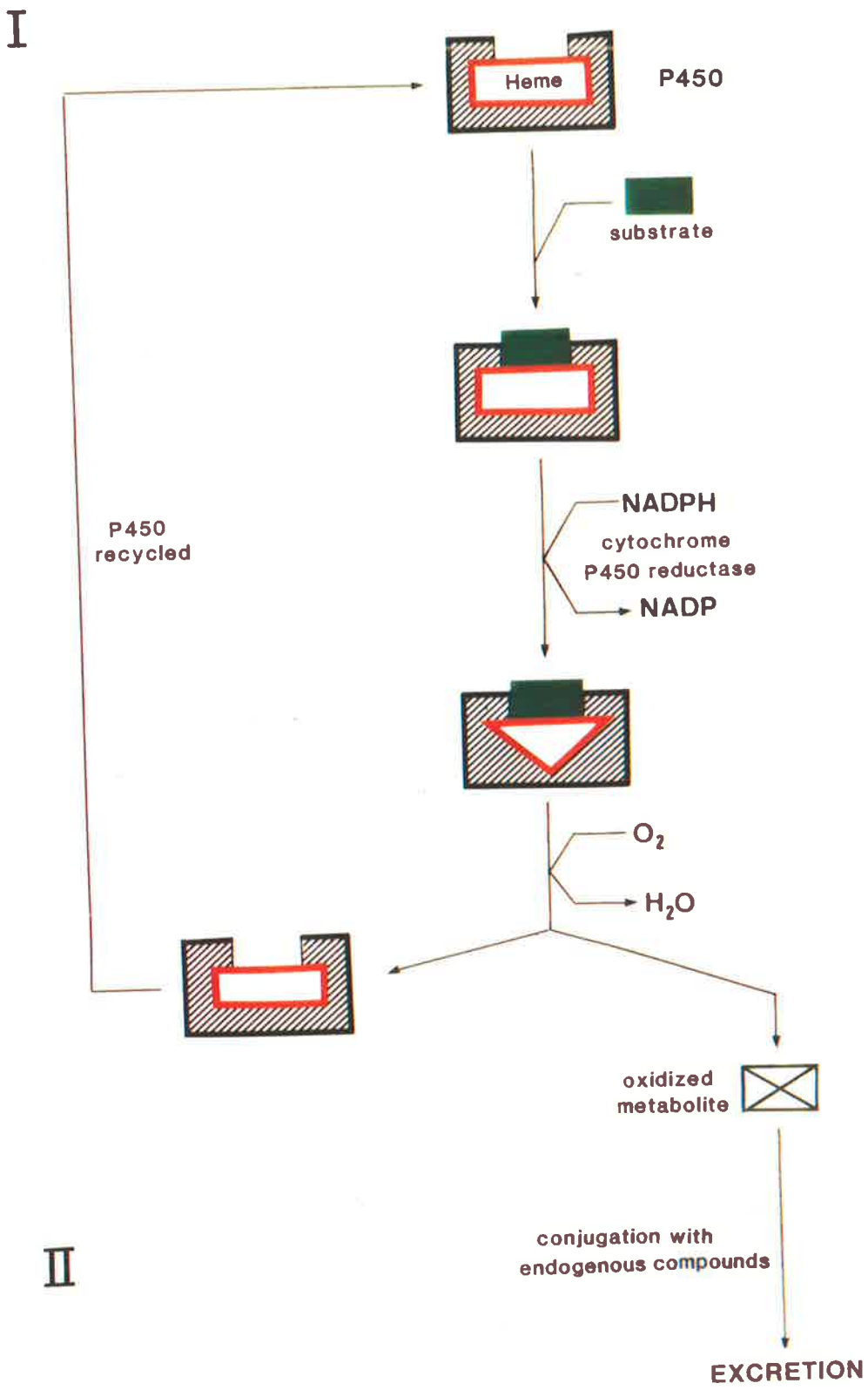
found in most tissues, with the exception of skeletal muscle and erythrocytes (Guengerich and Liebler, 1985).

The metabolism of a compound in the liver, occurs in a two phase process and is diagrammatically represented in Fig. 1.1. The oxidation of a compound by P450 during the phase I reactions, requires that the iron atom of the heme moiety be oxidized. The enzyme cytochrome P450 reductase utilises NADPH to reduce the iron to the ferrous form. It is in this state that the iron can bind a molecule of oxygen, thereby supplying one atom of oxygen to oxidize the P450 substrate. The other oxygen atom is used in the formation of water. Once oxygen is transferred to the substrate, the compound is susceptible to phase II reactions. These involve the conjugation of the polar compound by specific enzymes to endogenous molecules, rendering them more excretable from the organism. Many components of this enzyme system, including the phase I enzyme cytochrome P450 reductase (Gonzalez and Kasper, 1983) and the phase II enzymes epoxide hydrolase (Porter et al., 1986) and glutathione S-transferase (Suguoka et al., 1985), have recently been cloned and characterized. These enzymes are inducible by the same xenobiotics which induce P450.

The extremely broad substrate specificity of P450 is partly a result of the many different forms of the enzyme, encoded by a multiplicity of genes (Guengerich et al., 1982; Atchison and Adesnik, 1983; Gonzalez et al., 1986a; Gonzalez et al., 1985a). The categorization of a P450 isozyme into a particular family has been complicated not only by strain and colony variations between experimental animals, but also by the varied use of nomenclature between researchers. For example, rat P450h (Ryan et al., 1984) is catalytically and electrophoretically indistinguishable from rat P450 UTA, purified by Guengerich et al. (1982), and also equivalent to rat

Fig. 1.1. Schematic representation of substrate metabolism by P450.

The metabolism of P450 substrates has been schematically represented as a two phase process. Phase I involves the binding of a substrate to P450 while the iron of the heme moiety is in the ferric form (□). Cytochrome P450 reductase utilizes the coenzyme NADPH to reduce the iron of the heme to the ferrous form (▽). In this state, P450 can bind a molecule of oxygen, which is used to oxidize the substrate. In the process P450 reverts to its oxidized form. Phase II reactions involve further metabolism of the oxidized substrate towards ultimate excretion.


I
Heme
P450
substrate
NADPH
cytochrome
P450 reductase
NADP
P450
recycled
O_2
H_2O
oxidized
metabolite
II
conjugation with
endogenous compounds
EXCRETION

P4502c (Waxman et al., 1984). For the purpose of clarification, a table has been compiled which uses the characteristic regioselectivity[2] of particular P450 isozymes in the metabolism of the model substrate testosterone (Appendix 1). The table lists the source and inducibility of the respective isozymes and is intended to serve as a general guide for the comparison of these proteins.

Traditionally, P450s have been classified according to the type of xenobiotic by which they are induced. Initial classification of the P450 superfamily was on the basis of inducibility by either phenobarbital (PB) or 3-methylcholanthrene (3MC). The isolation of constitutive forms of P450 in hepatic and adrenal tissue and investigations with other inducers indicate that at least four families of P450 exist. Members of these separate families have been shown to be very different in their enzymic, structural and immunological properties and represent products of different genes (Atchison and Adesnik, 1985). The properties of the proteins which comprise these P450 families will now be discussed, together with specific reference to the structure and expression of the corresponding genes.

1.3. P450s INVOLVED IN STEROID BIOSYNTHESIS.

The adrenal cortex is the major site of mammalian steroidogenesis (Fevold, 1983). Of the five steps required to synthesize cortisol from cholesterol, four are catalyzed by P450s. The C-22,27 side chain of cholesterol is cleaved to form pregnenolone by P450ssc and the 3β hydroxyl is dehydrogenated yielding progesterone. Progesterone is successively hydroxylated by three different P450s at the 17α, 21 and 11β positions to yield cortisol. These steps occur in two subcellular compartments; the side chain cleavage (P450ssc) and 11β hydroxylation

2. Regioselectivity refers to the ability of P450 isozymes to catalyse the site specific hydroxylation of the substrate testosterone.

(P450-11β) steps occur in the mitochondria whereas the 17α (P450-17α) and 21-hydroxylation (P450c21) steps occur in the microsomes.

These P450 isozymes have been extensively studied in cultured bovine adrenocortical cells by Waterman and coworkers. The apparent lack of homology with the members of other P450 families indicates that they represent a distal member of the P450 gene superfamily. The expression of these P450 proteins has been shown to be tissue specific. For example, P450ssc mRNA is expressed in the adrenal cortex and ovaries, but not in the liver, kidney or heart (John et al., 1986). Similarly, P450c21 mRNA is expressed in the adrenal cortex and testis, but not in other steroidogenic tissue such as the ovaries or the placenta (Chung et al., 1986).

Unlike other P450 families which may contain up to 20 gene members (Adesnik and Atchison, 1983), the steroid metabolizing P450 family is unusually small. Southern hybridization analysis of bovine genomic DNA has revealed the presence of a single gene for P450ssc. Similarly, only two homologous copies of the gene for P450c21 have been detected in both human and bovine adrenal cells, although in the human one of these has been shown to be a pseudogene (White et al., 1986; Higashi et al., 1986).

In cultured bovine adrenocortical cells, the genes encoding these steroid synthesizing P450s are regulated by adrenocorticotrophic hormone (ACTH), via the action of cAMP (John et al., 1985; John et al., 1986; Zuber et al., 1986). John and coworkers (1985) have proposed that the action of cAMP is mediated through a labile trans-acting protein factor, tentatively named the steroid hydroxylase inducing protein (SHIP). cAMP is proposed to induce the synthesis of SHIP, which in turn interacts with a specific cis-acting genomic control element to stimulate P450 gene transcription.

1.4. GLUCOCORTICOID INDUCIBLE P450s.

Pregnenolone 16α-carbonitrile (PCN), is a synthetic steroid lacking hormonal activity which induces a unique protein profile of microsomal drug oxidizing activities in the rat liver (Schuetz et al., 1984). One form, P450-PCN has been purified from rat liver and differs from the major forms of P450 inducible by 3MC (Thomas et al., 1983) and PB (Waxman et al., 1983; Guengerich et al., 1982) both immunologically and catalytically. P450-PCN is also inducible by the glucocorticoid dexamethasone and to a lesser extent by PB (Schuetz et al., 1984).

Computer assisted comparison of the complete nucleotide and derived amino acid sequences of P450-PCN and a human inducible form (HLp), with those reported for other P450 families, supports the view that these isozymes are members of a distinct family. Rat P450-PCN shares 33% amino acid homology with both the major PB inducible form (P450b) and a 3MC inducible form (P450c), isolated from rat liver (Gonzalez et al., 1985b; Gonzalez et al., 1986b). Similar levels of homology were observed between human PCN, PB and 3MC inducible forms (Molowa et al., 1986). Southern hybridization analysis of genomic DNA suggests that there are three to five homologous genes in the glucocorticoid inducible family of each species (Hardwick et al., 1983a; Molowa et al., 1986).

The natural substrate for this glucocorticoid inducible form of P450 is not known. The higher constitutive level of expression in male rats suggests a possible role in androgen metabolism. However, the factors that control the sex dependent expression of P450-PCN are unclear and may be similar to those mechanisms controlling the sex dependent expression of other P450 genes (Section 1.5).

The mechanism by which PCN induces the accumulation of P450-PCN is not known. It has been established that glucocorticoids enter the nucleus as steroid-receptor complexes. These complexes interact with

specific areas of the chromatin, thereby regulating the coordinate expression of a large number of genes (Baxter and Rousseau, 1979). The induction of P450-PCN in cultured adult rat hepatocytes is stereospecific for glucocorticoids and PCN. The amount of induction is dependent on the concentration and the length of exposure to the steroids. Since these induction characteristics were reminiscent of the induction of many other glucocorticoid responsive proteins in hepatocytes, it was proposed that P450-PCN represented one of a group of genes whose expression is regulated by the cytoplasmic glucocorticoid receptor (Schuetz and Guzelian, 1984). However, the synthesis of P450-PCN in cultured hepatocytes treated with the glucocorticoid dexamethasone, failed to parallel the synthesis of a typical glucocorticoid-responsive protein such as tyrosine aminotransferase. Furthermore, two moderately potent inducers of P450-PCN either failed to induce tyrosine aminotransferase or actually antagonized induction of the protein by glucocorticoids. From such studies, Schuetz and Guzelian (1984) concluded that the induction of P450-PCN by glucocorticoids is mediated through a novel mechanism, readily distinguishable from the classic glucocorticoid receptor pathway. Whether such control involves a separate locus encoding a specific steroid receptor remains to be established.

1.5. POLYCYCLIC AROMATIC HYDROCARBON INDUCIBLE P450s.

Halogenated aromatic hydrocarbons such as 2,3,7,8-tetrachloro-dibenzo-ρ-dioxin (TCDD) and polycyclic aromatic hydrocarbons such as 3MC, are the standard drugs used in the study of this P450 family. The TCDD-induced family contains at least two members in the rat (Thomas et al., 1983), mouse (Gonzalez et al., 1984), rabbit (Norman et al., 1978) and human (Quattrochi et al., 1985; Quattrochi et al., 1986). cDNA

clones for these P450 isozymes have been isolated and the induction of these genes by TCDD and 3MC in the mouse and rat have been investigated extensively.

Mouse $P_1$450 and $P_3$450 exhibit 73% amino acid homology with each other (Kimura et al., 1984b) and are orthologous to rat P450c and P450d (Sogawa et al., 1984; Kawajiri et al., 1984) respectively, and to rabbit P450-6 and P450-4 respectively (Okino et al., 1985). The genes for $P_1$450 and $P_3$450 span 6.2 Kb and 6.7 Kb of genomic DNA respectively and are each divided into 7 exons (Gonzalez et al., 1985b). Analysis of rat P450c and P450d reveals similar exon/intron arrangements. Sequence comparison of the DNA in the 5' flanking regions of the rat P450c and P450d genes and the mouse $P_1$450 and $P_3$450 genes reveals little homology, despite the fact that each is inducible by TCDD and 3MC (Whitlock Jr., 1986). However, the genomic control elements involved in drug inducibility may be located further upstream of the regions which have been sequenced.

The mechanism of induction by 3MC and TCDD has been studied extensively. Whereas 3MC has been shown to increase the accumulation of mRNA specific for P450c and P450d in the rat, TCDD has been used to induce $P_1$450 and $P_3$450 in the mouse. Although the findings indicate that the increased levels of P450 isozymes result from the increased de novo synthesis of P450 apoprotein, this has not been investigated directly. Similarly it is not clear whether it is the drug itself or a metabolite which mediates the induction response.

Qualitative changes in P450 gene expression are observed during the development of the mouse embryo, such changes being tissue specific (Ikeda et al., 1983; Gonzalez et al., 1984; Tateja et al., 1985). Whereas $P_1$450 is substantially induced in the liver of the mouse embryo by polycyclic hydrocarbons, $P_3$450 induction is detectable only after

birth. In addition, $P_1$450 is inducible in the kidney by TCDD whereas $P_3$450 is not.

Recent evidence has suggested that the induction of mouse $P_1$450 by TCDD and other polycyclic aromatic hydrocarbons involves the coordination of both *trans*-acting proteins and *cis*-acting DNA elements (Jones *et al.*, 1986; Israel *et al.*, 1985).

The role of at least one *trans*-acting regulatory gene in the drug induction of P450s was discovered using mutant mice defective in P450 inducibility by 3MC (Thomas *et al.*, 1972; Whitlock Jr., 1986). The treatment of most wild and inbred strains of mice such as C57BL/6 with 3MC, induces aryl hydrocarbon hydroxylase (AHH) activity. The increase in AHH activity was shown to be due to increased transcription of the $P_1$450 and $P_3$450 genes (Israel and Whitlock Jr., 1984). In contrast, the mutant strain DBA/2 is "non responsive" to 3MC treatment, although these mice are not defective in the structural gene for $P_1$450. The unresponsiveness of these mice has been shown to be inherited as an autosomal dominant trait, suggesting that induction is controlled by a single locus, termed the Ah locus (Nebert and Jensen, 1979). This locus is postulated to encode a cytosolic receptor which binds the inducer. It is believed that the resultant inducer-receptor complex is translocated to the nucleus where it interacts with specific chromatin components to stimulate gene transcription.

Treatment of the 3MC inducible strain C57BL/6 with TCDD results in the induction of $P_1$450 gene transcription (Israel and Whitlock Jr., 1984). The level of induction observed with TCDD exceeds that observed with 3MC, indicating that TCDD is a more potent inducer. The enhanced potency of TCDD is further demonstrated by its ability to induce $P_1$450 in DBA/2 mice, *albeit* only slightly. Similar results were observed for the induction of mouse $P_3$450 by TCDD. As a consequence, it was proposed

that the C57BL/6 mice possess a receptor with a higher affinity than that in the DBA/2 mice (Jones et al., 1985). Although purification of the receptor has so far been unsuccessful, studies on partially purified cell extracts have indicated that the murine Ah receptor possesses many of the characteristics shared by the steroid hormone receptors (Denison et al., 1986a; Poland et al., 1986; Denison et al., 1986b).

Studies on mouse hepatoma cell lines varying in their responsiveness to TCDD, have supported the proposal that transcription of the $P_1$450 gene requires an inducer-receptor complex. Jones and coworkers (1986) have shown that the activity of the TCDD-receptor complex is mediated through a cis-acting control element, located upstream of the $P_1$450 gene. This control element spans 2.6 Kb of DNA and contains at least four functional domains: 1) a strong promoter; 2) an adjacent inhibitory domain which blocks promoter function; 3) two structurally distinct and functionally independent TCDD responsive domains, each of which have properties analogous to those of transcriptional enhancers; and 4) 5' sequences located 2 Kb upstream from the cap site of the transcript, which are proposed to augment the response to TCDD by increasing the binding of the TCDD-receptor complexes to their cognate regulatory domains. Similar studies by Sogawa and coworkers (1986) on the rat P450c gene have demonstrated the role of positive and negative regulatory DNA elements in the drug induced expression of this gene. Whether similar components control the transcription of mouse $P_3$450 and rat P450d genes, is yet to be confirmed. Marked differences among species and tissues in their biological responses to TCDD have been observed. Whether these occur as a consequence of the TCDD-receptor complex interacting with a specific DNA sequence, a distinct chromatin configuration or a combination of both, remains to be established.

Studies by Israel et al. (1985), imply that a second trans-acting element modulates the action of the TCDD-receptor complex. The inhibition of protein synthesis in mouse hepatoma cells by cycloheximide results in the superinduction of $P_1$450 gene transcription by TCDD. Since the inhibition of protein synthesis does not alter several properties of the TCDD-receptor complex, the existence of a labile repressor protein which acts via the inhibitory domain adjacent to the $P_1$450 promoter, has been proposed to inhibit transcription of the $P_1$450 gene. Therefore, the TCDD mediated increase in the transcription of the $P_1$450 gene involves the complex interaction between both positive and negative control elements.

1.6. PHENOBARBITAL INDUCIBLE P450s.

To date, the PB inducible P450s comprise the largest family and have been studied extensively in both the rat (Guengerich et al., 1982; Ryan et al., 1984) and the rabbit (Heinmann and Ozols, 1983; Leighton et al., 1984). Several isozymes with molecular weights ranging from 50,000 to 60,000 have been purified (Appendix 1). While the designation of the gene family as PB inducible is logical from a historical standpoint, it can be misleading. Members of this family are related by sequence homology and not necessarily by PB inducibility. For example, although the genes encoding rat P450f and rabbit P450-1 are approximately 50% homologous to PB responsive genes, both are constitutively expressed and unresponsive to treatment with PB (Gonzalez et al., 1986a; Tukey et al., 1985). As a consequence, this P450 family has been divided into two subfamilies (Adesnik and Atchison, 1985). These subfamilies will now be discussed separately.

1.6.A. The PB II Subfamily.

By definition, proteins within a family are greater than 50% homologous in their amino acid sequences (Dayhoff, 1976). Examination of nucleic and amino acid sequence homologies between rat and rabbit isozymes in this subfamily, strongly suggests that the rat f,g,h,i and j P450s are members of this subfamily and represent genes orthologous to rabbit P450-1 (Tukey et al., 1985) and the rabbit P450PBc series of genes (Govind et al., 1986; Adesnik and Atchison, 1985). These rat isozymes are expressed in hepatic tissue and appear unresponsive to treatment with PB. Some members of this subfamily are capable of metabolizing endogenous steroids, and in at least two cases expression of the gene is sex dependent (Waxman, 1984; Waxman et al., 1985; Ryan et al., 1984). The expression of rat P450h and P450i have been shown to be under strict hormonal control. The male specificity of rat P450h is reflected in its developmental induction by testosterone. Similarly, the female specificity of P450i is reflected by its induction in female rats during development (Waxman, 1984). Members of this P450 subfamily exhibit different patterns of tissue specific expression (Leighton and Kemper, 1981). The elevation of the mRNAs encoding these isozymes during development occurs primarily as a consequence of transcriptional activation of their respective genes (Gonzalez et al., 1986a; Song et al., 1986). Their role (if any) in the metabolism of exogenous compounds is not known. The nonresponsiveness of many of the isozymes in the PB II subfamily to previously identified inducers, suggests that these isozymes are constitutively expressed. Notable exceptions are P450h and P450i which are under hormonal regulation, P450j which is inducible by ethanol and P450-PB1 which is slightly inducible by PB.

On the basis of sequence homology, the rat isozyme P450-PB1 is also a member of this P450 subfamily (Gonzalez et al., 1986a). P450-PB1 is

75% homologous with P450f at the amino acid level and only 49% homologous with the major PB inducible isozyme P450b. However, in contrast to the other members of this subfamily, P450-PB1 is induced 4 fold by PB.

1.6.B. The PB I Subfamily.

Members of this subfamily are dramatically induced by PB. To date, only two major forms designated P450b and P450e have been purified (Ryan et al., 1982). These two isozymes were found to be immunochemically indistinguishable with polyclonal antisera (Ryan et al., 1982; Vlasuk et al., 1982). They are encoded by two distinct yet closely linked genetic loci, with at least 4 alleles at the P450b locus and 2 alleles at the P450e locus (Rampersaud and Walz Jr., 1983). Both amino acid and cDNA sequence analysis have demonstrated that these hemoproteins share 97% homology in their primary sequences (Yuan et al., 1983; Fujii-Kuriyama et al., 1982; Kumar et al., 1983). The genes for both P450s have been isolated and sequenced (Mizukami et al., 1983b; Suwa et al., 1985). The gene for P450b spans 23 Kb of DNA and is divided into 9 exons. The overall gene structure of P450b and P450e is very similar, with the exception that the first intron in P450b is 12 Kb long whereas in P450e it is only 3.2 Kb (Suwa et al., 1985). The 14 amino acid differences between these isozymes occur as a consequence of nucleic acid substitutions restricted to exons 6, 7, 8 and 9. The simple repeated sequence of $(CA)_5$ exists at an equivalent position in the 5'flanking region of both P450b and P450e, but has been extended to $(CA)_{19}$ in P450e (Suwa et al., 1985). Although as yet untested, this region has been proposed as a potential regulatory element to account for the difference between the levels of basal and PB inducible expression of P450b and P450e (Suwa et al., 1985). Southern hybridization analysis of rat

genomic DNA with a cDNA clone for P450e, indicates that 6-10 genes exist within this P450 subfamily (Kumar et al., 1983).

PB produces a 20-50 fold increase in P450b mRNA within 3-4 hours (Hardwick et al., 1983a). In vitro transcription experiments in hepatic nuclei isolated from PB treated rats indicate that the increase in P450b mRNA is due to transcriptional activation of the gene (Hardwick et al., 1983a; Pike et al., 1985). The possible involvement of mRNA stabilization during this response has not been determined and may contribute partly to the accumulation of mRNA. In PB treated adult rats, hepatic levels of P450b mRNA were 4-5 fold greater than those of P450e (Omiecinski et al., 1985).

The coordinate induction of P450b and P450e by PB is both tissue specific and age dependent (Omiecinski, 1986; Giachelli and Omiecinski, 1986). Whereas both isozymes are inducible in the liver, only P450b was found to be constitutively expressed in pulmonary and testicular tissue.

The mechanism by which PB causes an increase in gene transcription is not clear, and it is not known if a receptor mediated mechanism is involved in their induction. In contrast to compounds which are known to interact with receptors, inducers of the PB type are only effective at higher concentrations and exhibit no structural homology. These properties are not suggestive of a receptor mediated mechanism.

Heme has been implicated as a general positive regulator of PB inducible P450 gene expression (Ravishankar and Padmanaban, 1983), although evidence for this is limited. Ravishankar and Padmanaban (1985) have shown that inhibitors of heme synthesis such as cobalt chloride and 3-amino-1,2,4-triazole block the induction of rat P450b and P450e at the level of transcription. None of the treatments led to a significant change in the total transcription rates as measured in isolated nuclei. The effect was specific for P450 since no effect on

the transcription of the albumin gene was detectable (Satyabhama et al., 1986). However, it was not possible to counteract the effect of inhibitors by the administration of exogenous heme. Therefore, the exact role of heme in the regulation of P450 remains controversial and still requires detailed investigation.

1.7. AIMS OF THIS THESIS.

In this laboratory, we are investigating the drug induction of P450 and ALV-synthase in the chick embryo. The chick embryo was used since it is a convenient in vivo system in which both enzymes can be dramatically induced, particularly by the drugs 2-allyl-2-isopropyl-acetamide (AIA) and 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC), and to a lesser extent by PB. Although they appear to be coordinately induced, evidence for a relationship between the drug induction of P450 and ALV-synthase is at present limited.

At the commencement of this project, a small cDNA clone homologous to chick embryo hepatic P450 mRNA inducible by the drugs AIA and DDC had been isolated (Brooker and O'Connor, 1982). Rot analysis of hepatic mRNA from AIA and DDC treated chick embryos with this cDNA clone, revealed that each of these drugs induce P450 mRNA which is homologous to that inducible by PB (Brooker et al., 1983). The mechanism by which these structurally unrelated drugs induce homologous P450s is of immense interest.

The complete characterization of the chick embryo P450 gene(s) is a fundamental requirement for subsequent studies into understanding their drug inducibility at the molecular level. Therefore the main aim of the work presented in this thesis was to:

(1) characterize the P450 genes inducible by the drugs AIA, DDC and PB.

(2) determine the structure and organization of these genes.

Comparison of the nucleotide sequences between members of the P450 superfamily may reveal consensus sequnces in the control regions of the gene(s) which could be candidates for regulatory elements. Such studies may lead ultimately to an understanding of how these genes are controlled.

CHAPTER 2

MATERIALS AND METHODS.

2.1. MATERIALS.

2.1.A. Experimental Animals.

Fertilized White Leghorn eggs were obtained from the Department of Agriculture, Parafield Poultry Research Station, Adelaide and were incubated for 18 days prior to use in a humidified incubator at 37°C. Hatched chicks were maintained in a temperature controlled environment, being fed pellets and water ad libitum, in accordance with the guidelines specified by the NHMRC Animal Experimentation Ethics Committee (1984-1986).

2.1.B. Drugs and Chemicals.

2-Allyl-2-isopropylacetamide (AIA) was a generous gift from Roche, Australia. 3,5-Diethoxycarbonyl-1,4-dihydrocollidine (DDC) was purchased from Eastman Organic Chemicals and phenobarbitone (PB) from F.H. Faulding Pty. Ltd., Australia.

The following chemicals were purchased from Sigma Chemical Co.: Acrylamide and bisacrylamide (N,N'-methylene-bisacrylamide), SDS (sodium dodecyl sulphate), Trizma base, BCIG (5-bromo-4-chloro-3-indoyl-3-D-β-galactopyranoside), IPTG (isopropyl-thiogalactoside), salmon sperm DNA, deoxynucleotide triphosphates (dNTPs), ethidium bromide, EDTA (ethylenediaminetetraacetic acid), $CoCl_2$, bovine serum albumin (BSA) and ampicillin.

E. coli tRNA, agarose (type 1) and low melting point agarose were purchased from BRL. Phenol, glyoxal, PEG 6000 (polyethylene glycol) and Na cacodylate were purchased from BDH.

Oligo (dT) cellulose was obtained from Collaborative Research, TEMED from Eastern Kodak Co. and dideoxy DNA sequencing kits from Biotechnology Research Enterprises of South Australia (BRESA). BA45 paper and BA85 nitrocellulose were purchased from Schleicher and Schuell

Inc. Sephadex G50, CL-4B Sepharose oligo $(dT)_{12}$ and oligo $(dG)_{12}$ were from Pharmacia. Chloramphenicol and tetracycline were gifts from Parke-Davis and Upjoin Pty. Ltd. respectively. Rabbit reticulocyte lysate (code N150) was purchased from Amersham International (U.K.).

Other chemicals routinely purchased from Sigma Chemical Co., Ajax Chemicals Ltd and Pharmacia, were either analytical grade or of the highest available purity.

2.1.C. Enzymes.

The enzymes used during the course of this work were obtained from the sources listed below:

All restriction enzymes were initially purchased from New England Biolabs. EcoRI, HindIII, BamHI and T4 DNA polymerase were purchased from Integrated Sciences Pty. Ltd.

Calf intestinal phosphatase, proteinase K, E. coli DNA polymerase I (Klenow fragment), T4 polynucleotide kinase, S1 nuclease and T4 ligase were purchased from Boehringer Mannheim. Later, T4 DNA ligase, DNA polymerase I and Klenow fragment were obtained from BRESA.

Lysozyme, E. coli RNase A and E. coli DNase I were purchased from Sigma Chemical Co.

Terminal deoxynucleotidyl transferase was purchased from P.L. Biochemicals and AMV reverse transcriptase (avian myeloblastosis virus RNA-dependent DNA polymerase) was a generous gift from Dr. R. H. Symons.

2.1.D. Radiochemicals.

[α-^{32}P] dATP, [α-^{32}P] dCTP and [γ-^{32}P] ATP (>2000 Ci/mmol) were purchased from BRESA. [^{35}S] methionine (1200 Ci/mmol), [5-^{3}H] dCTP (25 Ci/mmol) and [8-^{3}H] dGTP (9.5 Ci/mmol) were purchased from Amersham

International (U.K.).

2.1.E. Synthetic DNA Oligonucleotides and Bacterial Vectors.

Oligonucleotides used as probes in Chapters 3 and 7 and as primers in Chapter 3 were manufactured by BRESA. pBR322, M13 mp8, mp18 and mp19 were purchased from BRESA.

2.1.F. Bacterial Strains.

E. coli MC1061: ara D139, Δ(ara, leu) 7697, Δlac x74, gal U^-, gal K^-, hsr^-, hsm^+, str A (Casadaban and Cohen, 1980). Host for pBR322 transformations.

E. coli JM101: lac, pro, F' traD36, sup E, pro AB, lac Iq, Z M15 (Messing, 1979). Host for M13 transformations.

E. coli LE392: F^- hsdR514, (r^-Km^-K), sup E44, sup F58 lac Y1. Host for bacteriophage λ propagation (Murray et al., 1977).

2.1.G. Media and Buffers.

Whereas media were prepared with monodistilled water, buffers and solutions were prepared using double distilled water. All media and buffers were sterilized by autoclaving.

1. Growth media for E.coli MC1061.

Luria broth (L broth) contained 1% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 1% (w/v) NaCl, adjusted to pH 7.0 with NaOH. Agar plates were prepared by adding 1.5% (w/v) Bacto-agar (Difco) to the L broth. Where appropriate, the media and plates were supplemented with either ampicillin (50 ug/ml) or tetracycline (20 ug/ml) after cooling to 50°C.

2. Growth media for E. coli JM101.

Minimal medium contained 1.05% K_2HPO_4, 0.45% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, 0.05% Na citrate, supplemented after autoclaving with 0.02% $MgSO_4$, 0.02% glucose and 0.0005% thiamine-HCl. M13 minimal medium plates were prepared with minimal medium containing 1.5% Bacto-agar.

2xYT broth contained 1.6% Bacto-tryptone, 1% yeast extract, 0.5% NaCl, adjusted to pH 7.0. Soft YT overlay contained 1xYT broth supplemented with 0.7% Bacto-agar.

3. Growth media for E. coli LE392.

NZCYM broth, for the propagation of λ bacteriophage,contained 1% NZ amine A, 0.5% yeast extract, 0.1% Casamino acids, 0.5% NaCl, 0.25% $MgSO_4.7H_2O$ and 0.2% maltose, with the pH adjusted to 7.5. NZCYM plates and soft overlay contained NZCYM medium supplemented with 1.5% and 0.7% Bacto-agar, respectively.

4. Buffers.

SM: 0.1 M NaCl, 8 mM $MgSO_4.7H_2O$, 0.01% gelatin and 50 mM Tris-HCl pH 8.0.

TE: 0.1 mM EDTA, 10 mM Tris-HCl pH 8.0.

NET: 0.1 M NaCl, 0.1 mM EDTA and 10 mM Tris-HCl pH 8.0.

TBE: 130 mM Tris base, 45 mM boric acid and 2.5 mM EDTA pH 8.0.

SSC: 150 mM NaCl, 15 mM Na citrate and 1 mM EDTA pH 7.0

5x loading buffer: 50% glycerol, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.2% bromophenol blue.

2.1.H. Preparation of Phenol.

Phenol was prepared by distillation directly into water and was stored at 4°C in a dark bottle until required. The phenol was buffer saturated by vigorous mixing with an equal volume of 1 M Tris-HCl pH 8.0

followed by three changes with an equal volume of TE. Once equilibrated the TE saturated phenol was stored at 4°C in a dark bottle for up to 5 weeks. Unless stated otherwise, TE saturated phenol was used for the extraction of DNA samples.

2.1.I. Miscellaneous.

All glassware and equipment was alkali washed (0.1 M NaOH) where necessary, to minimize RNase contamination and rinsed well with double distilled water or sterilized by dry heat or autoclaving.

Cellophane dialysis tubing (1 cm) and glass fibre (GF/A) filters (1 cm discs) were purchased from BDH chemicals and Whatman Ltd., respectively.

P450 rabbit antiserum was a gift from Dr. Srivastava and was raised against a partially purified form of AIA inducible chick embryo hepatic P450.

2.2. METHODS.

2.2.A. Drug Regime for P450 Induction in Chick Embryos.

17-18 day old chick embryos were administered either 3 mg AIA and 6 mg DDC, 6 mg DDC, 3 mg AIA or 4 mg of phenobarbitone. The drugs were dissolved in 0.1 ml DMSO and introduced into the fluid surrounding the embryo, through a small hole in the apex of the egg, using a 1.5 inch, 18 gauge needle. The hole was sealed with adhesive tape and the eggs incubated at 37°C for 16-18 hours.

2.2.B. Ethanol Precipitation of Nucleic Acids.

Unless stated otherwise, samples were adjusted to 0.2 M Na acetate pH 5.6 using a 4 M stock solution. Two and a half volumes of redistilled ethanol were added and after mixing thoroughly, the samples

chilled at either -20°C overnight or on dry ice for 60 minutes. Precipitates were collected from small volumes by centrifugation in an Eppendorf centrifuge at 12,000 rpm for 30 minutes at 4°C. Larger volumes were centrifuged at 10,000 rpm for 30 minutes at 4°C in a Sorvall SS34 rotor. The supernatant was removed, the tubes briefly respun and the residual ethanol aspirated off. The samples were dried *in vacuo* before resuspension in the appropriate solution.

2.2.C. RNA Preparation.

1. *Preparation of total hepatic RNA from chick embryos.*

6-10 chick embryos were killed by decapitation, their livers dissected out and the gall bladder removed. RNA was prepared using a phenol/SDS extraction procedure modified from Land *et al.* (1981), as detailed below.

The livers were homogenized in 10 ml/g of tissue of SDS buffer (0.5% SDS, 25 mM EDTA, 75 mM NaCl, 100 mM Tris-HCl pH 8.0) and 10 ml/g of tissue of phenol saturated with SDS buffer, in a Sorvall Omni Mix 17220. The resulting homogenate was centrifuged at 12,000xg for 30 minutes at 4°C and both the aqueous phase and interface retained. These were extracted with a half volume of SDS buffer saturated phenol and a half volume of chloroform/isoamyl alcohol (25:1, v/v). The aqueous phase was retained following centrifugation at 12,000xg for 20 minutes at 4°C. The aqueous phase was adjusted to 200 mM NaCl and the nucleic acids precipitated with two volumes of redistilled ethanol, on dry ice for 60 minutes. The nucleic acids were collected by centrifugation, the pellet dried *in vacuo* and resuspended in 0.5 ml of sterile water per liver. The suspension was adjusted to 0.5% SDS and treated with 25 ug/ml of proteinase K at 37°C for 45 minutes. The solution was extracted with half volumes of phenol and

chloroform/isoamyl alcohol (25:1, v/v). The aqueous phase was adjusted to 200 mM NaCl and precipitated with two and a half volumes of ethanol on dry ice for 60 minutes or at -20°C overnight. Following centrifugation, the pellet was dried *in vacuo* and resuspended in 2 ml of sterile water.

The RNA was precipitated from the mixture of nucleic acids by the addition of three volumes of 4 M Na acetate pH 5.6 and stored on ice for 30 minutes. The RNA was pelleted by centrifugation and the precipitation step repeated. The resulting pellet was dried *in vacuo*, resuspended in water, transferred to an Eppendorf tube and ethanol precipitated. The final RNA pellet was resuspended in TE to a final concentration of 2 mg/ml.

2. Isolation of poly A^+ RNA.

Chick embryo hepatic poly A^+ RNA was affinity purified from total RNA by two rounds of chromatography on oligo (dT) cellulose as described below.

The oligo (dT) beads were preswollen in water for at least 30 minutes, loaded into a Biorad Poly Prep column and allowed to settle. The column was washed in 0.1 M NaOH and 5 mM EDTA until the pH of the eluate was alkaline, to ensure that the column was free of RNA or ribonuclease contamination. The column was neutralized by washing with one column volume of 1 M Tris-HCl pH 7.5 and two column volumes of binding buffer (0.5 M KCl, 10mM Tris-HCl pH 7.5, 1 mM EDTA).

The RNA, resuspended in TE, was adjusted to 0.5% SDS and heated at 65°C for 5 minutes to dissociate RNA secondary structure. The RNA was immediately snap cooled on ice to prevent reaggregation and adjusted to 0.5 M KCl. Chromatography was performed at room temperature. The sample was slowly loaded onto the column equilibrated with binding buffer. The column was further washed with three to five column volumes

of binding buffer and the initial eluate containing poly A^- RNA was collected. The poly A^+ fraction was eluted from the column with elution buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) and readjusted to 0.5 M KCl. The RNA was again denatured by heating and passaged through the column once more. The poly A^+ and poly A^- RNA fractions were precipitated with ethanol, resuspended in TE and stored at -80^oC.

2.2.D. Analysis of Hepatic RNA.

All RNA preparations were analysed by either of two procedures to ensure that negligible degradation had occurred during the extraction procedure.

1. Cell free translation and analysis.

Poly A^+ RNA used in the construction of cDNA libraries was translated in a cell free reticulocyte lysate and the protein products immunoprecipitated with P450 polyclonal antiserum, according to the procedure of Brooker et al. (1983). The immunoprecipitates were analysed by SDS polyacrylamide gel electrophoresis (Laemmli, 1970) and fluorography (Bonner and Laskey, 1974).

2. Acid urea gel electrophoresis.

Subsequent RNA preparations were analysed by electrophoresis on a vertical 1.5% agarose gel (150 mm x 160 mm x 3 mm) prepared in 6.2 M urea and 25 mM Na citrate pH 3.5. The gels were pre-electrophoresed in 25 mM Na citrate pH 3.5 at 60 mA for 15 minutes at 4^oC. 10-20 ug of total or poly A^+ RNA was adjusted to a volume of 20 ul and mixed with 60 ul of loading buffer (20% sucrose, 8 M urea, 25 mM Na citrate pH 3.5 and 0.2% (w/v) bromophenol blue). The samples were loaded onto the gel and run into the gel at 30 mA. Electrophoresis was continued at 60 mA for 5 hours at 4^oC. The gel was stained with ethidium bromide and the

integrity of the RNA visualized by UV irradiation.

2.2.E. Preparation of [^{32}P]-Labelled DNA Probes.

1. Nick translations.

0.1-1.0 ug of DNA was labelled with [^{32}P] using E. coli DNA polymerase I and calibrated amounts of DNase I. DNA was nick translated in a 20 ul reaction containing 10 mM $MgCl_2$, 50 mM Tris-HCl pH 7.4, 0.1 mM DTT, 50 ug/ml BSA, 25 uM each of unlabelled dGTP and dTTP, 100 uCi each of [α-^{32}P] dATP and [α-^{32}P] dCTP (1700 Ci/mmol), a calibrated amount of DNase I and 5 units of DNA polymerase I. The reaction was incubated at 16^{o}C for 60 minutes, extracted with an equal volume of phenol and the unincorporated nucleotides removed by chromatography on a Sephadex G50 column equilibrated in TE.

DNase I was prepared as a 2 mg/ml stock solution in 10 mM HCl and stored in 10 ul aliquots at -20^{o}C. DNase I was calibrated by nick translating 1.0 ug of supercoiled pBR322 using serial dilutions of DNase I prepared in DNase I dilution buffer (10 mM $MgCl_2$, 50 mM Tris-HCl pH 7.4 and 50 ug/ml BSA), using only 2 uCi of [α-^{32}P] dATP. [^{32}P]-labelled DNA from each reaction was size fractionated on an alkaline agarose gel (Maniatis et al., 1982) using end-labelled HinfI cut pBR322 as size markers. 20-50 pg of DNase I was shown to generate fragments approximately 300-600 nucleotides in length, with a specific activity of 10^7-10^8 cpm/ug of DNA.

2. 5' end-labelling of synthetic DNA oligonucleotides.

The various synthetic DNA oligonucleotides used during the course of this work, were end-labelled at the 5' end by transfer of the [^{32}P] phosphate group from [γ-^{32}P] ATP, using T4 polynucleotide kinase. The reaction was performed in a 20 ul volume containing 10 mM $MgCl_2$,

50 mM Tris-HCl pH 7.4, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, 100 uCi of [γ-^{32}P] ATP and 2 units of T4 polynucleotide kinase at 37^{o}C for 60 minutes. The reaction was either extracted with an equal volume of phenol and ethanol precipitated, or 50 ug of tRNA was added and the mixture chromatographed on a G50 Sephadex column equilibrated in NET to remove unincorporated nucleotides.

3. End-filling of DNA restriction fragments.

Restriction enzyme generated 5' overhangs were end-filled by incubating 0.2-1.0 ug of DNA in a 20 ul reaction containing 10 mM Tris-HCl pH 7.4, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 25 uM of each dNTP and 3 units of Klenow. Alternatively, certain unlabelled dNTPs were replaced with 10 uCi of the appropriate [α-^{32}P] dNTP (1700 Ci/mmol) depending on the nature of the overhang. The reaction was performed at 37^{o}C for 20 minutes after which the DNA was extracted with an equal volume of phenol and recovered by ethanol precipitation.

4. Primer extended cDNA.

5 ng of [^{32}P]-labelled primer was hybridized with 5 ug of poly A^{+} RNA in a 10 ul volume in the presence of 0.1 M KCl, and annealed by heating to 65^{o}C for 10 minutes followed by gradual cooling to 25^{o}C. The reaction was adjusted to 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 20 mM DTT, 150 mM NaCl and 0.5 mM of all four dNTPs, and primer extension performed in a final volume of 20 ul at 37^{o}C for 60 minutes, using 4 units of AMV reverse transcriptase. The products were analysed by electrophoresis on a 8% polyacrylamide gel containing 7 M urea. The extension product was kindly purified and sequenced by Dr. A. Hobbs using the Maxam and Gilbert (1980) sequencing procedure.

2.2.F. Restriction Enzyme Digestions.

Restriction endonuclease digestion of DNA was performed using the conditions specified by the manufacturer for each enzyme. Reactions were permitted to proceed overnight and were terminated by the addition of either 5 mM EDTA or 1x load buffer (Section 2.1.G-4) as appropriate. Genomic DNA was routinely extracted with an equal volume of phenol following enzyme digestion. After removing 80% of the aqueous phase, the organic phase was re-extracted with an equal volume of TE and 80% of the aqueous phase combined with that obtained previously. The samples were then ethanol precipitated as outlined in Section 2.2.B.

2.2.G. Construction of cDNA Libraries.

During the course of the work two cDNA libraries were prepared using both the double tailing method of Land et al. (1981) and the loopback method of Maniatis et al. (1981). Modifications to the basic procedures employed during the construction of double stranded cDNA are detailed below.

1. Single stranded cDNA synthesis.

Oligo $(dT)_{12}$ primed synthesis of single stranded cDNA was performed on 10 ug of poly A^+ RNA at 45°C for 20 minutes in a 100 ul reaction containing 100 mM Tris-HCl pH 8.3, 10 mM $MgCl_2$, 20 mM DTT, 800 uM of dATP, dTTP and dGTP, 200 uM dCTP, 10 uCi of [α-^{32}P] dCTP, 4mM Na pyrophosphate, 1 ug of oligo $dT_{(12)}$ and 50 units of AMV reverse transcriptase. 1 ul of the reaction was removed to determine the percentage incorporation of labelled nucleotide into cDNA. The reaction was adjusted to 100 mM NaOH and heated to 90°C for 60 minutes to degrade the mRNA template. The reaction was neutralized with 100 mM Tris-HCl pH 8.0 and passaged through a Sephadex G50 column equilibrated in 200 mM ammonium bicarbonate. The DNA was lyophilized overnight.

2. Homopolymer tailing of cDNA and vector DNA.

Homopolymeric oligo (dC) tails were added to either single stranded cDNA or double stranded cDNA using calf thymus terminal deoxynucleotidyl transferase. Initially, enzyme activity was found to be extremely variable, however the following modifications routinely catalysed the addition of dCTP to the cDNA at a rate of 20 nucleotide residues in 4-8 minutes.

Na cacodylate buffer was a generous gift from Dr. A. Hobbs. BDH reagent grade Na cacodylate was recrystallized twice, dissolved in water and the pH adjusted to pH 6.9 with HCl. Homopolymeric tailing reactions were performed in a final volume of 100 ul, in a reaction containing 100 uCi [5-^{3}H] dCTP (25 Ci/mmol), 0.2 M Na cacodylate buffer pH 6.9, 2.8 mM 2-mercaptoethanol, 200 ug/ml BSA, 1 mM $CoCl_2$, 2-5 ug of single stranded or double stranded cDNA and 10 units of terminal deoxynucleotidyl transferase. The reaction was incubated at 30^{o}C until 20 nucleotides were added per end as determined by the conversion of labelled nucleotide into TCA precipitable material. The reaction was stopped by the addition of EDTA to 5 mM, phenol extracted and the aqueous phase chromatographed on Sepharose CL-4B in 200 mM ammonium bicarbonate. The fractions containing cDNA were lyophilized.

pBR322 was linearized with PstI and tailed with [8-^{3}H] dGTP (9.5 Ci/mmol) using the same conditions outlined for the cDNA until 20 nucleotides were added per end.

3. Synthesis of double stranded cDNA.

Double stranded cDNA was synthesized using either of two methods. In the first procedure, the synthesis of the second strand is primed by oligo $(dG)_{12}$ at a final concentration of 25 ug/ml. The oligo $(dG)_{12}$ is hybridized to the 3' oligo (dC) tail of single stranded cDNA in a 100 ul reaction containing 100 mM Tris-HCl pH 8.3, 10 mM $MgCl_2$, 50 mM KCl,

20 mM DTT, 800 uM dATP, dGTP and dTTP, 200 uM dCTP, 10 uCi [α-^{32}P] dCTP and 80 units of AMV reverse transcriptase. After incubation at 42^{o}C for 2 hours, the double stranded cDNA was phenol extracted and purified by chromatography on a CL-4B Sepharose column equilibrated in 200 mM ammonium bicarbonate.

Where appropriate, double stranded cDNA was synthesized using AMV reverse transcriptase by self priming of the 3' end of the single stranded cDNA. The single stranded cDNA was heated to 100^{o}C in the above conditions for 5 minutes and snap cooled on ice for 15 minutes. Synthesis of the double stranded cDNA was initiated by the addition of AMV reverse transcriptase.

4. Digestion with S1 nuclease.

Double stranded cDNA synthesized by the self priming of the 3' end of the single stranded cDNA, had to be modified prior to homopolymeric tailing. This was achieved by digestion of the single stranded hairpin loop of the double stranded cDNA with S1 nuclease, according to the procedure described by Maniatis et al. (1982).

5. Size fractionation of the double stranded cDNA and annealing to vector DNA.

Tailed double stranded cDNA generated by the double tailing method, was fractionated prior to insertion into pBR322 by electrophoresis on a 1.5% low melting point agarose gel. Double stranded cDNA greater than 1.5 Kb in length was purified as outlined in Section 2.2.J-1.

Equimolar amounts of dG-tailed pBR322 and dC-tailed double stranded cDNA were annealed in a 100 ul volume of annealing buffer (0.2 M NaCl and 10 mM Tris-HCl pH 8.3) at 65^{o}C for 5 minutes, followed by one hour at 42^{o}C and then allowed to cool slowly to 4^{o}C. 20 ng of vector DNA was

routinely used to transform competent E. coli MC1061 cells as outlined in Section 2.2.J-5.

6. Detection of recombinant colonies.

The detection of pBR322 plasmids containing cDNA inserts complementary to drug inducible P450s, was accomplished as follows. Colonies found to be Tet^R and Amp^S were toothpicked onto L-Tet plates and screened by the colony hybridization method of Grunstein and Hogness (1975). Nick translated inserts derived from existing cDNA clones were used as probes (Chapters 3 and 4) and positive recombinant clones were maintained as glycerol stocks at -80^oC.

2.2.H. Plasmid DNA Preparation.

The rapid and efficient alkaline lysis method of Birnboim and Doly (1979), was used for the analytical preparation of plasmid DNA. A final yield of 2-5 ug of DNA was routinely obtained from a 5 ml overnight culture and could be used directly for restriction enzyme analysis.

Large scale preparations of recombinant DNA plasmids were obtained by $CsCl_2$ gradient purification, according to the method of Yu-Lee and Rosen (1983). A final yield of 500-1000 ug of highly purified recombinant plasmid was routinely obtained for use in restriction enzyme analysis and subcloning.

2.2.I. Genomic DNA Preparation.

Genomic DNA was prepared from liver using a combination of the procedures described by Marshall et al. (1976) and Wallace et al. (1971), with the following modifications (Dr G. Cam, personal communication). The liver (2-5 gm) was removed and gently homogenized in 20 mls of buffer A (330 mM sucrose, 2 mM EDTA, 0.5 mM EGTA, 60 mM KCl, 15 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine and 15 mM Tris-HCl

pH 7.4). The homogenate was passed through muslin to remove fibrous material and the mixture poured over an equal volume of buffer B (2 M sucrose, 1 mM EDTA, 0.3 mM EGTA, 60 mM KCl, 15 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine and 15 mM Tris-HCl pH 7.4). The gradients were centrifuged in a swing out rotor at 16,000xg for 20 minutes at 4°C. The pelleted nuclei were gently resuspended in 5 ml of STE (10 mM NaCl, 10 mM Tris-HCl pH 8.0 and 10 mM EDTA) and the suspension very slowly added to 30 mls of STE supplemented with 0.5% SDS, with continual gentle stirring. Proteinase K was added to a final concentration of 0.1 mg/ml and the viscous solution incubated overnight at 37°C, with gentle mixing in an orbital shaking water bath. The mixture was gently extracted once with an equal volume of phenol, once with equal volumes of phenol and chloroform/isoamyl alcohol (25:1, v/v) and once with an equal volume of chloroform/isoamyl alcohol (25:1, v/v). The aqueous phase was carefully removed and dialysed against 4 l of TE overnight at 4°C. The dialysate was removed and treated with heat inactivated RNase A, added to 20 ug/ml, at 37°C for 4 hours. The mixture was phenol/chloroform extracted as described above and dialysed against 4 l of TE at 4°C. The dialysate was removed and the DNA stored as a diluted stock at 4°C.

2.2.J. Subcloning Restriction Fragments into DNA Vectors.

The following procedures were routinely used for the preparation of subclones in pBR322 and M13 vectors.

1. Preparation of Vectors.

pBR322 and M13 replicative form were linearized with the appropriate restriction enzyme(s) and electrophoresed on a 0.8%-1.4% low melting point agarose gel set up in a horizontal gel apparatus in TBE buffer. The DNA was visualized by staining with ethidium bromide and irradiation with a UV light. A gel slice containing the vector DNA was

excised from the gel, placed into a microfuge tube and an equal volume of NET buffer was added. The agarose was melted at 65°C for 10 minutes and the DNA recovered by one phenol extraction, one ether wash, followed by ethanol precipitation of the aqueous phase.

To remove the 5' terminal phosphate group from linearized vector DNA, 5 ug of the DNA was incubated with 0.1-0.2 units of calf intestinal phosphatase (CIP) at 37°C for 60 minutes in a reaction containing 10 mM Tris-HCl pH 9.0 and 0.1% SDS. The reaction was terminated with the addition of EDTA to 5 mM and phenol extracted. The DNA from the aqueous phase was ethanol precipitated and any contaminating uncut vector separated from the linearized vector by electrophoresis through a low melting point agarose gel as described above.

2. Preparation of DNA restriction fragments.

The cloned DNA was digested with the appropriate restriction enzyme(s) and depending on the size of the DNA fragment(s), electrophoresed on either a horizontal 0.8%-1.5% agarose gel or a vertical 6%-8% polyacrylamide gel (Maniatis et al. 1982). DNA fragments were visualized after staining with ethidium bromide by UV irradiation. Restriction fragments were isolated using either of three methods:

(i) DNA was recovered from a polyacrylamide gel slice by diffusion from the gel slice overnight in 500 ul of 0.1% SDS, 0.5 M NH_4 acetate, 10 mM Mg acetate and 1 mM EDTA pH 8.0. The DNA was recovered from the supernatant by ethanol precipitation.

(ii) DNA was recovered from low melting point agarose gels as previously outlined (Section 2.2.J-1). In situations where normal agarose was used, the DNA fragment was detected after staining with ethidium bromide by UV irradiation and a 10 mm x 5 mm piece of DEAE NA45 nitrocellulose paper was inserted into the gel, in front of the DNA fragment(s) of interest. The DNA was electrophoresed onto the paper and

transfer monitored using a hand held UV lamp. When transfer was completed, the filter was removed, rinsed gently in TE to remove adhering agarose and cut into smaller pieces enabling the DNA to be recovered in a microfuge tube. The paper was incubated in 200 ul of 1 M NaCl and 0.05 M arginine at 70°C for 60 minutes to elute the DNA. The eluate was centrifuged to pellet the paper, the supernatant extracted once with an equal volume of phenol/chloroform (1:1,v/v) and the DNA recovered by ethanol precipitation.

3. Quantitation of DNA.

A 1% agarose solution was prepared in TE and ethidium bromide added to a final concentration of 10 ug/ml. 10 ml of the solution was aliquoted into 36 mm petri dishes, the plates allowed to dry overnight and stored at 4°C. pBR322 DNA standards ranging in concentration from 500 ng/ul to 20 ng/ul were prepared in TE and 1 ul of each spotted directly onto the surface of a plate, with 1 ul of the test DNA. After the spots had dried (15-30 minutes) the petri dish was inverted onto a UV light source and the DNA visualized. Quantitation was achieved by the visual comparison of the intensity of fluorescence of the test DNA with the standards.

4. Ligation conditions.

The DNA fragment and appropriate vector were combined in a molar ratio of 3:1 respectively, in a 20-40 ul reaction containing 20 ng of vector DNA, 50 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP. All ligations (i.e. blunt or sticky ends) were incubated with 2.5 units of T4 DNA ligase for 4-18 hours at 14°C.

5. Transformation of E. coli strains.

E. coli MC1061 cells were made competent and transformed using a modification of the method described by Dagert and Ehrlich (1979). MC1061 cells were grown overnight in L broth at 37°C with constant aeration. The cells were subcultured 1/100 (v/v) into 100 ml of fresh L broth and grown at 37°C until an A_{600nm} of 0.6-0.7 was obtained. The cells were chilled on ice for 10 minutes and pelleted by centrifugation at 2000xg for 5 minutes at 4°C. The cells were gently resuspended in half the original volume of ice cold 10 mM $MgCl_2$ and placed on ice for 20 minutes. The cells were pelleted again by centrifugation and gently resuspended in 1/25 volume of ice cold 100 mM $CaCl_2$. The cells were kept on ice for at least 3 hours prior to use.

The annealed DNA was diluted to 100 ul with TE, mixed with 0.2 ml of competent MC1061 cells and stored on ice for 40 minutes with intermittent mixing. After heat shock for 2 minutes at 42°C, the cells were kept on ice for a further 20 minutes and then warmed to room temperature. 0.5 ml of L broth was added and the cells incubated at 37°C for 20 minutes. The transformed cells were either spread directly onto the surface of L plates containing the appropriate antibiotic or mixed with 3 ml of 0.8% L agar and overlayed onto the appropriate plates. The plates were incubated overnight at 37°C or until the colonies reached 1-2 mm in diameter.

Competent E. coli JM101 cells were prepared by diluting an overnight culture of JM101 grown in minimal medium, 1/100 (v/v) into fresh 2x YT broth. The culture was grown at 37°C with constant aeration until an A_{600nm} of 0.6-0.7 was obtained. The culture was chilled on ice for 10 minutes and the cells pelleted by centrifugation at 2000xg for 5 minutes at 4°C. The cells were gently resuspended in half of the original volume of ice cold 100 mM $CaCl_2$ and stored on ice for 20

minutes. The cells were pelleted by centrifugation and gently resuspended in 1/20 the volume of ice cold 100 mM $CaCl_2$.

0.2 ml of the competent cells was immediately combined with the ligation mix and stored on ice for 40 minutes with intermittent mixing. After heat shock at 42°C for 2 minutes, the cells were mixed with 3 ml of 0.7% YT agar containing 10 ul of 20 mg/ml IPTG, 20 ul of 20 mg/ml BCIG (in dimethylformamide) and 0.2 ml of exponentially growing JM101. The mixture was plated immediately onto minimal medium plates and incubated overnight at 37°C.

2.2.K. Southern Transfer and Hybridization Conditions.

Genomic DNA and cloned DNA digested with restriction enzymes were fractionated on horizontal 0.8%-1.5% agarose gels in 1x TBE and visualized after staining with ethidium bromide by UV irradiation. Genomic DNA was transferred to BA85 nitrocellulose using the method of Southern (1975) after treatment with 0.25 M HCl for 30 minutes. Bidirectional transfer of cloned DNA from agarose gels was performed according to the method of Smith and Summers (1980), in which the gel is neutralized and placed between two sheets of BA85 nitrocellulose.

Transfer of the DNA was allowed to proceed overnight before baking the filters at 80°C <u>in vacuo</u> for 2 hours. Prehybridization, hybridization and washing conditions varied according to the nature of the [^{32}P] probe used and are described below as well as in the figure legends.

Filters probed with nick translated restricted fragments of cDNA clones were prehybridized for 4 hours in 6x SSC, 0.1% (w/v) Ficoll, 0.1% (w/v) polyvinylpyrrolidone, 0.1% (w/v) BSA, 0.1% (w/v) SDS and 200 ug/ml of heat denatured salmon sperm DNA at 68°C. Hybridizations were normally performed with 50-150 ng of heat denatured [^{32}P] labelled

probe DNA for 16 hours with the above conditions. Filters were rinsed in 2x SSC, 0.1% SDS for 5 minutes at room temperature followed by two washes in 0.2x SSC, 0.1% SDS at 68°C for 30 minutes each, prior to autoradiography. Genomic Southerns were hybridized with 100-500 ng of [^{32}P] labelled probe DNA for 36-42 hours using the above conditions and washed in 1x SSC, 0.1% SDS at 68°C for 60 minutes. Low stringency hybridizations were performed under the same conditions except that the temperature was reduced to 52°C. The filters were subsequently washed at 52°C in 1x SSC, 0.1% SDS prior to autoradiography.

The hybridizations with labelled oligonucleotides were performed as follows. The filters were prehybridized in 6x SSC, 0.02% (w/v) Ficoll, 0.02% (w/v) BSA, 0.02% (w/v) polyvinylpyrrolidone, 0.05% Na pyrophosphate and 25 ug/ml heat denatured salmon sperm DNA at 42°C for 4 hours. Hybridizations were performed using these condition with 100 ng of [^{32}P] labelled oligonucleotide probe overnight at 42°C. Following hybridization, the filters were washed twice in 6x SSC, 0.05% Na pyrophosphate for 5 minutes each at room temperature, followed by a single 30 minute wash in 1x SSC, 0.05% Na pyrophosphate at either 42°C or 52°C.

For reuse, filters containing cloned DNA were washed at 90°C for 60 minutes in 10 mM Tris base to remove hybridized probe. Prior to prehybridization and hybridization as described above, the filter was rinsed in distilled water and autoradiographed to ensure that it was free of probe.

2.2.L. Northern Hybridization Analysis of RNA.

Deionized glyoxal was prepared according to the procedure of Carmichael and M^{c}Master (1980) and stored at -80°C. RNA samples were denatured by heating at 50°C for 60 minutes in a 20 ul reaction

containing 1.4 M glyoxal and 15 mM Na phosphate pH 6.5. After cooling on ice, an equal volume of 40% sucrose, 15 mM Na phosphate, 0.02% bromophenol blue and 0.02% (w/v) xylene cyanol was added and the RNA electrophoresed for up to 5 hours at 25 mA in a 1% agarose gel (150 mm x 160 mm x 3 mm), prepared in 0.1 M Na phosphate pH 6.5 . The electrode buffer was constantly recirculated to maintain a constant pH.

Denatured RNA was transferred unidirectionally to BA85 nitrocellulose using the procedure of Thomas (1980). The filters were baked for 2 hours _in vacuo_. Filters to be probed with [^{32}P] labelled cDNA clones were prehybridized in a solution containing 50% (v/v) deionized formamide, 0.9 M NaCl, 50 mM Tris-HCl pH 7.4, 5 mM EDTA, 0.5% SDS, 0.04% (w/v) Ficoll, 0.04% (w/v) polyvinylpyrrolidone, 0.04% (w/v) BSA and 250 ug/ml of heat denatured salmon sperm DNA at 42^{o}C for 3-4 hours. Hybridizations with 100-200 ng of denatured [^{32}P] labelled double stranded DNA were carried out overnight in the same conditions as above. The filters were then washed in two changes of 2x SSC, 0.1% SDS at room temperature, followed by two 15 minute washes in 0.2x SSC, 0.1% SDS at 50^{o}C prior to autoradiography.

Hybridizations involving [^{32}P] labelled oligonucleotides were performed in the same solution used for the double stranded probes with the exception that 50% deionized formamide was substituted by 0.05% Na pyrophosphate. Following hybridisation at 42^{o}C for 16-20 hours, the filters were washed in two changes of 6x SSC, 0.05% Na pyrophosphate at room temperature and once in 2x SSC, 0.05% Na pyrophosphate at 42^{o}C for 30 minutes prior to autoradiography.

2.2.M. Screening of the Genomic Libraries.

The chicken genomic library described in Chapter 5 was prepared by Dodgson _et al_. (1979) and kindly donated by Dr J. R. Wells. The library

was constructed by partial HaeIII and AluI digestion of chicken genomic DNA followed by the ligation of 15-21 Kb fragments to EcoR1 digested λ Charon 4A arms. The library had been amplified at least once and was screened with a nick translated cDNA clone.

The first round of screening was performed using 150 mm NZCYM plates containing 3 x 10^5 pfu plated out in E. coli LE392. Plating bacteria were prepared in 10 mM $MgSO_4$ as specified by Maniatis et al. (1982). A sterile 140 mm nitrocellulose filter was adsorbed to the plate for 10-20 minutes and orientation marks on the filter, were copied onto the base of the plate. After adsorption to the filter, the phage were denatured by soaking the filter in 0.5 M NaOH, 1.5 M NaCl for 2 minutes, neutralized with two 2 minute washes in 0.5 M Tris-HCl pH 7.4 and 1.5 M NaCl, air dried and baked in vacuo at 80°C for 2 hours. The conditions used for the hybridization of [^{32}P] labelled cDNA to these filters were as previously described (Section 2.2.K) except that the hybridization and washings were performed at 52°C. All potentially positive clones were picked into SM buffer and rescreened at lower density, giving plaque pure preparations after 4 successive rounds of screening. Three of the clones required an extra round of screening to achieve plaque purity.

The second genomic library described in Chapter 6, was kindly donated by Dr. P. Morris and was constructed by partial Sau3A digestion of chicken genomic DNA followed by ligation of 16-24 Kb fragments into SalI digested EMBL 3 arms. The library had been amplified at least once and was screened with a nick translated cDNA clone using the hybridization conditions described previously for high stringency screening (i.e. 68°C). Plaque pure preparations were achieved after 4 rounds of screening.

2.2.N. Titration and Preparation of λ Clones.

λ Phage titres were routinely determined using a plaque spot assay. 0.9 ml of E. coli LE392 plating bacteria was mixed with 9 ml of 0.7% NZCYM agar, spread onto a 150 mm NZCYM plate and allowed to set for 10 minutes at room temperature. Serial dilutions of the phage stock were prepared in SM buffer and 2 ul spotted directly onto the overlay. After drying, the plates were incubated at 37°C overnight and the phage titre determined by counting the number of plaques per 2 ul aliquot of the dilution.

The liquid culture method of Maniatis et al. (1982) was used to prepare phage DNA from recombinant phage.

2.2.O. Dideoxy Sequence Analysis of M13 Phage Recombinants.

Three different strategies were employed during the course of this work to generate a series of overlapping deletion clones for use in DNA sequencing.

1. DNase I digestion method.

M13 phage shotgun libraries of the cDNA clones of interest were constructed using the DNase I method of Anderson (1981). Briefly, entire plasmids containing the cDNA inserts were digested with a calibrated amount of DNase I in the presence of 1 mM Mn^{2+}, to decrease the sequence specificity of DNase I. Fragments between 500-1500 bp were isolated from a low melting point agarose gel and the ends repaired with Klenow. EcoRI linkers were ligated onto the ends, digested with EcoRI and the DNA fragments ligated into EcoRI cut M13 mp8 vector DNA (Section 2.2.J.4). After transformation of E. coli JM101 (Section 2.2.J.5), white plaques were spotted onto a lawn of E. coli, grown overnight and then subject to plaque hybridization (Benton and Davis, 1977) using nick translated insert to detect clones containing complementary DNA

sequences. Hybridization conditions were as previously described (Section 2.2.K.).

2. DNase I deletion subcloning method.

A sequential series of overlapping clones of the 5' genomic region of a λ recombinant were generated in mp18 using conditions described by Dr. A. Sivaprasad (personal communication). Essentially, 5 ug of the replicative form of the M13 recombinant was linearized by digestion with BamHI within the polylinker and the 5' phosphate removed by treatment with CIP (Section 2.2.J-1). Calibrated amounts of DNase I were used such that only 10-20% of the linearized DNA was digested, as judged by gel electrophoresis. The truncated insert molecules were blunt ended with Klenow and serial dilutions prepared to give final concentrations of 2 ug/ul, 0.2 ug/ul and 0.02 ug/ul. Following blunt end ligation and transformation into E. coli JM101, white plaques of varying sizes were toothpicked into a 98 well microtitre tray containing 100 ul of 2x YT broth innoculated 1/40 with an overnight of JM101. The recombinants were grown at 37°C for 6-8 hours with constant aeration and stored at 4°C overnight to settle the cells. 20 ul of the supernatant was incubated with 5 ul of phage lysis buffer (60% formamide, 3% SDS, 5 mM EDTA and 0.1% bromophenol blue) at 68°C for 30 minutes and the clones size fractionated on a 1% agarose gel using wild type M13 mp18 and the original recombinant in mp18 as size markers. Deletion clones were visualized by staining with ethidium bromide and UV irradiation. Clones differing in size by approximately 150-300 bp were chosen for subsequent sequence analysis.

3. Sequential single stranded cloning procedure.

The "band-aid" method of Dale et al. (1985) was the third technique utilized for generating a series of M13 clones containing overlapping

inserts. The procedure was performed essentially as described, except that 100 mM NaCl and 1 mM DTT were included in the initial EcoRI digestion of the recombinant without affecting the subsequent reactions. Recombinants were screened by size fractionation on 1% agarose gels and clones differing in size by 150-300 bp were prepared for subsequent sequence analysis.

4. Preparation of single stranded template DNA for sequence analysis.

An overnight culture of JM101 was diluted 1/100 into 2x YT and grown at 37°C for 60 minutes with aeration. 1.5 ml of the culture was aliquoted into sterile tubes, infected with either a recombinant M13 phage from a fresh plate or 5-50 ul of a recombinant phage stock. The cultures were grown for 6-8 hours at 37°C with constant aeration. The cells were pelleted by centrifugation at 12,000 rpm for 5 minutes in an Eppendorf centrifuge and 1 ml of the supernatant mixed with 0.3 ml of 2.5 M NaCl, 20% PEG. Phage particles were precipitated by incubating the tube at room temperature for 15 minutes. The remainder of the supernatant was stored at -20°C as a fresh phage stock. Phage pellets were collected by centrifugation for 15 minutes at 12,000xg in an Eppendorf centrifuge and the pellets resuspended in 100 ul of 0.5% SDS, 5 mM EDTA and 50 mM Tris-HCl pH 8.0. Phage DNA was recovered by extraction with an equal volume of phenol. 90 ul of the aqueous phase was ethanol precipitated overnight at -20°C. The DNA was recovered by centrifugation and resuspended in 30 ul of TE.

5. Dideoxy sequence analysis.

6-8 ul of the single stranded template DNA was annealed with 1 ul of the appropriate primer (2.5 ng), 1 ul of 10x TM (100 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$) in a final volume of 10 ul. Annealing was initiated

by heating at 90°C for 5 minutes followed by slow cooling to room temperature.

A BRESA DSK-A sequencing kit was used in all sequencing reactions. Prior to sequencing, 1 ul of the appropriate dNTP and ddNTP solutions were aliquoted into Eppendorf tubes and stored on ice. The annealed DNA and primer were mixed with 1 ul (10 uCi) of [α-^{32}P] dATP (1700 Ci/mmol) which had been dried down *in vacuo* and resuspended in water. 1 unit of Klenow was added to the reaction, vortexed gently and 2 ul added directly to the contents of each of the four reaction tubes. After 15 minutes at 37°C, the reactions were chased with 25 uM of each dNTP for a further 10 minutes at 37°C. The reactions were stopped by the addition of 4 ul formamide loading buffer, heated at 100°C for 3 minutes and immediately chilled on ice prior to gel electrophoresis.

6. Sequencing gels.

1 ul of each sequencing reaction was electrophoresed on a 6% polyacrylamide gel (200 mm x 400 mm x 0.2 mm) containing 7M urea in 1x TBE at 30-35 mA. The gels were routinely pre-electrophoresed for 30 minutes at 15 mA prior to sample loading.

After electrophoresis at 1800-2000 V, the gels were fixed in 10% (v/v) acetic acid to remove the urea and rinsed in several changes of 20% (v/v) ethanol. The gels were baked in a 100°C oven for 30-45 minutes and autoradiographed overnight at room temperature.

2.2.P. Miscellaneous.

RNA and DNA concentrations were measured in 1 cm quartz cuvettes in a Varian DMS 90 Spectrophotometer, assuming that one A_{260nm} unit was equal to 40 ug/ml of RNA and 50 ug/ml of DNA.

2.2.Q. Containment Facilities.

All manipulations involving recombinant DNA were performed in accordance with the regulations and approval of the Australian Academy of Science Committee on Recombinant DNA and the University Council of the University of Adelaide.

CHAPTER 3

CONSTRUCTION AND CHARACTERIZATION OF cDNA CLONES FOR PB INDUCIBLE CHICKEN P450s.

3.1. INTRODUCTION.

Information concerning the structure and organization of the chicken P450 genes and their respective regulatory sequences must be obtained in order to understand the control of P450 synthesis. Previously Brooker and O'Connor (1982) showed that the drugs AIA, DDC and PB cause a rapid induction of P450 in 18 day old chick embryo livers. Using a polyclonal antiserum raised against the AIA inducible form of P450, it was established that in each case, the P450 induced by AIA, DDC and PB respectively was a protein of 50,000 molecular weight which shared common antigenic determinants.

A cDNA clone (p2H10) for chicken P450 had been previously obtained in our laboratory using hepatic mRNA isolated from chick embryo livers induced with AIA and DDC (Brooker et al., 1983). Hybridization analysis and thermal denaturation of mRNA:cDNA hybrids indicated that AIA, DDC and PB all induce homologous mRNA species, possibly due to enhanced expression of the same gene. Northern hybridization analysis of AIA and DDC induced hepatic chick embryo RNA detected the presence of a 3.3-3.5 Kb mRNA species homologous to the cDNA clone p2H10. However, p2H10 was found to contain only a 720 bp insert excisable by the enzyme PstI. Therefore, a full length cDNA clone was required for determination of the protein sequence and the subsequent isolation and characterization of the corresponding gene. The construction of a chick embryo hepatic cDNA library and the isolation and sequence analysis of cDNA clones for P450 are described in this chapter.

3.2. RESULTS.

3.2.A. Preparation and Analysis of Hepatic RNA.

It has been routinely found that a combination of the drugs AIA and DDC promote a greater induction of P450 mRNA in the livers of chick

embryos than PB (Brooker et al., 1982). For this reason, total hepatic RNA from 18 day old chick embryos induced with a combination of AIA and DDC was routinely prepared using the rapid phenol/SDS extraction procedure described in Section 2.2.C. Poly A^+ RNA was affinity purified by two passages through oligo (dT) cellulose to yield 50-100 ug of poly A^+ RNA per gram of liver. Total RNA and poly A^+ RNA integrity were analysed by electrophoresis in an acid-urea gel system (Fig. 3.1-A). Although poly A^+ RNA contained some residual 28S rRNA, both samples appeared to be undegraded and were free of contaminating chromosomal DNA (Fig 3.1-A).

Prior to generating a cDNA library, the poly A^+ RNA was examined for the presence of mRNA species which encoded P450 proteins. Hepatic poly A^+ RNA isolated from drug treated and untreated chick embryo livers was translated in vitro in a cell free reticulocyte lysate (Section 2.2.D.). The [^{35}S] methionine labelled protein products were immunoprecipitated with rabbit antiserum raised against a 50,000 molecular weight P450 protein, isolated from the livers of AIA treated chick embryos (Section 2.2.D). The immunoprecipitates were analysed by SDS polyacrylamide gel electrophoresis and fluorography. The results (Fig. 3.1-B) clearly demonstrate that the poly A^+ RNA isolated from drug treated chick embryos codes for a major immunoprecipitable protein product of molecular weight 50,000. The protein is drug inducible, since in untreated embryos, this immunoprecipitation product was not observed. A protein of molecular weight 74,000 was also precipitated by the antiserum, but the identity of this protein is unknown. Although the purified P450 protein used to generate the antiserum was shown by SDS PAGE and Coomassie Blue staining to be homogeneous (Dr. G. Srivastava, personal communication), the possibility of an undetectable contaminant with high antigenicity in the preparation cannot be

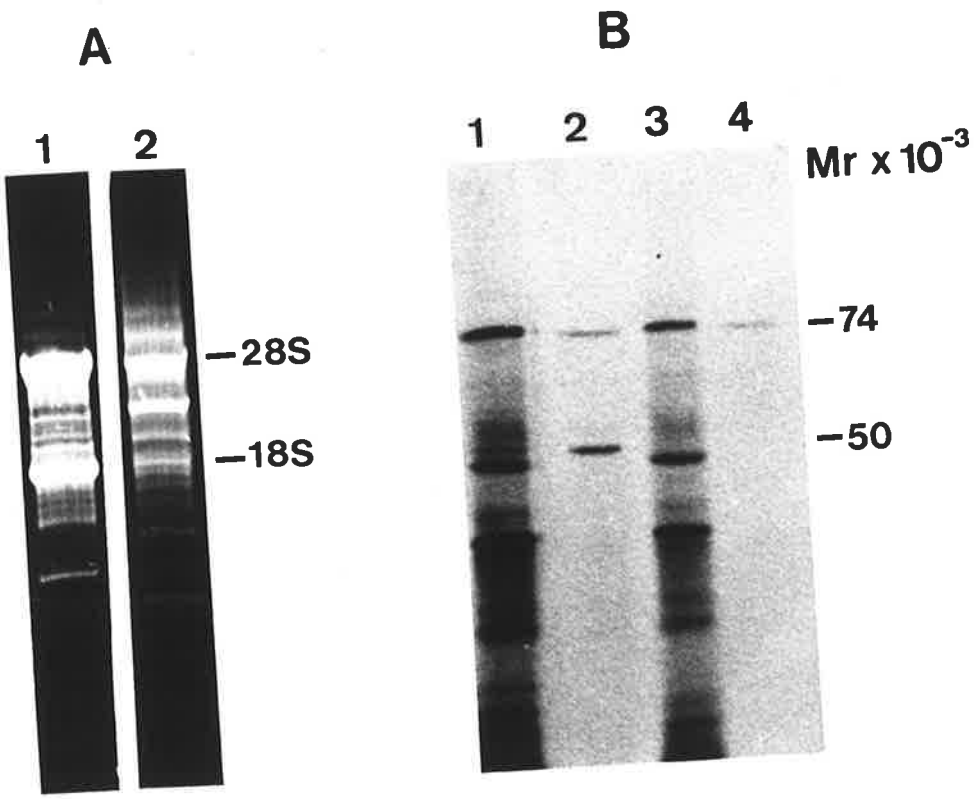
Fig. 3.1. Analysis of chicken hepatic RNA integrity.

Total RNA and poly A^+ RNA was isolated from the livers of untreated, and AIA and DDC treated chick embryos as described in section 2.2.D. RNA integrity was analysed by acid urea gel electrophoresis (panel A) and cell free translation in a reticulocyte lysate (panel B).

Panel A: 10 ug of total RNA (lane 1) and 10 ug of poly A^+ RNA (lane 2) was analysed by electrophoresis on a 1.5% acid urea agarose gel (Section 2.2.D-2). The gel was stained with ethidium bromide and RNA integrity visualized by U.V. irradiation. The 28S and 18S rRNA species are indicated.

Panel B: 2 ug of poly A^+ RNA isolated from AIA and DDC treated (lanes 1 and 2) and untreated (lanes 3 and 4) chick embryo livers, was translated in a 100 ul cell free reticulocyte lysate. A 20 ul aliquot was removed for analysis of total protein synthesis (lanes 1 and 3) and the remainder immunoprecipitated using antiserum to the AIA induced form of P450 (lanes 2 and 4). Total protein and immunoprecipitated protein were analyzed by electrophoresis on a 13% SDS-polyacrylamide gel and fluorography.

The additional protein band of Mr 74,000 may represent albumin, which often contaminates antibody preparations.


A
1
2
28S
18S
B
1
2
3
4
Mr x 10-3
74
50

excluded. The presence of this 74,000 molecular weight protein in both drug treated and control samples, and the observation that all the P450s characterized to date exhibit a molecular weight of approximately 50,000, further argues against it being a P450. Therefore a detailed characterization of the additional 74,000 molecular weight protein was not performed.

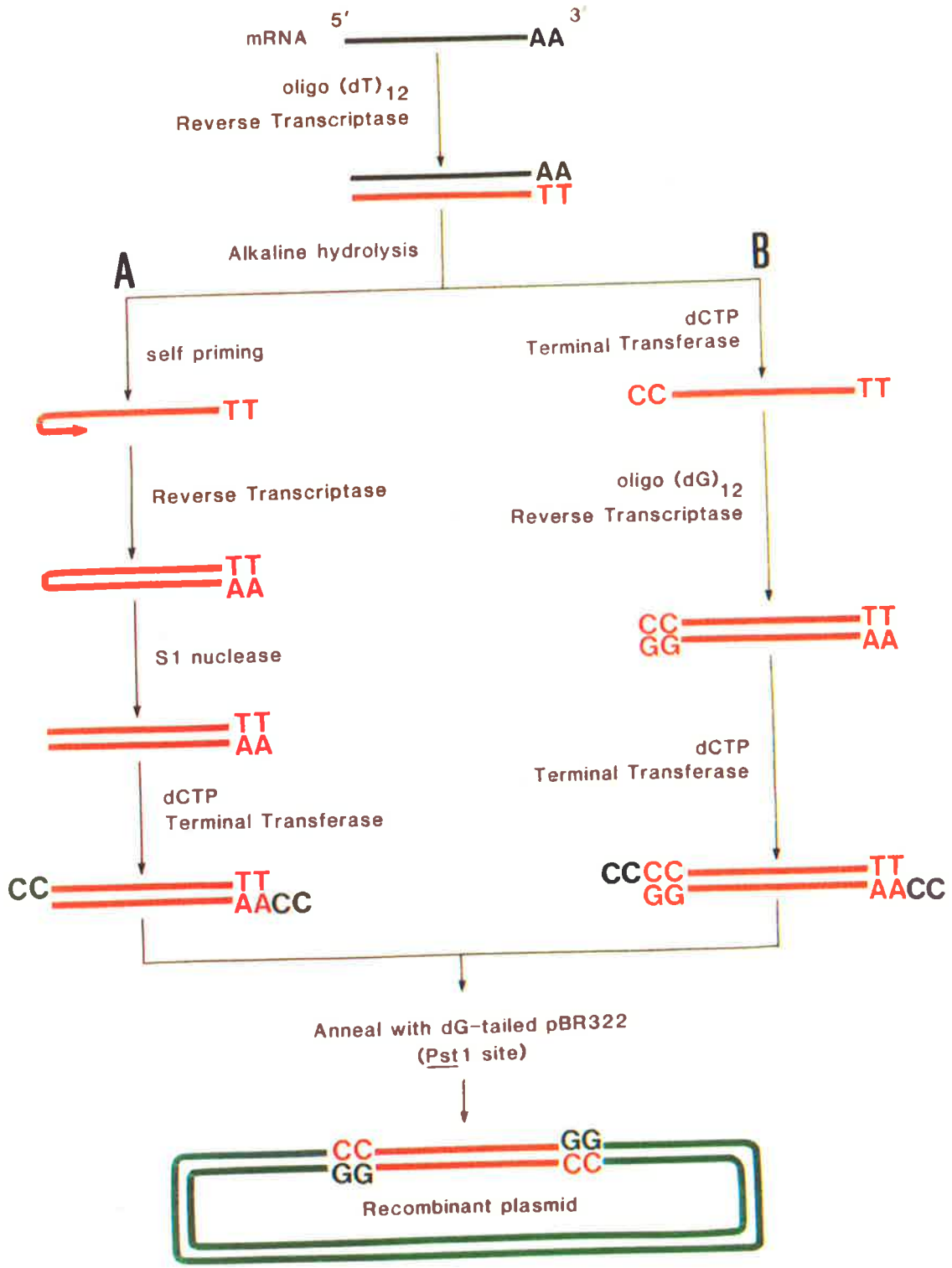
Since the mRNA isolated from the livers of chick embryos induced with AIA and DDC directed the synthesis of a 50,000 molecular weight protein, immunoprecipitable with an antiserum to P450, it was concluded that intact P450 mRNA was present. Therefore, this mRNA was used in the construction of a cDNA library.

3.2.B. Synthesis of a cDNA Library.

Recent advances in recombinant DNA technology have enabled the refinement of several experimental approaches for the synthesis of double stranded cDNA (Gubler and Hoffman, 1983; Okayama and Berg, 1982). However, at the time at which the library described in this chapter was synthesized, only two cloning procedures were available, the loopback method of Maniatis et al. (1982), and the double tailing method of Land et al. (1981). Both of these cloning strategies are detailed in Chapter 2, Section 2.2.G. and summarized diagrammatically in Fig. 3.2. Briefly, single stranded cDNA is synthesized from mRNA using AMV reverse transcriptase, by priming off the 3' poly A tail with oligo $(dT)_{12}$. The mRNA template is removed by alkaline hydrolysis and double stranded cDNA synthesized by either of the two methods. In the loopback method (Fig. 3.2., pathway A), the hairpin loop formed by the 3' end of the single stranded cDNA, is utilized to prime the synthesis of double stranded cDNA using reverse transcriptase. The hairpin loop which joins the two strands of cDNA is cleaved by S1 nuclease and the double stranded cDNA

Fig. 3.2. Comparison of the cDNA cloning strategies.

Comparison of the loopback method (pathway A) described by Maniatis et al. (1982) and the double tailing method (pathway B) described by Land et al. (1981) for the synthesis of double stranded cDNA is schematically represented. A detailed description of the respective procedures is contained in the text (Section 3.2.B).


mRNA
5'
3'
AA
oligo (dT)12
Reverse Transcriptase
AA
TT
Alkaline hydrolysis
A
B
self priming
TT
Reverse Transcriptase
TT
AA
S1 nuclease
TT
AA
dCTP
Terminal Transferase
CC
TT
AACC
dCTP
Terminal Transferase
CC
TT
oligo (dG)12
Reverse Transcriptase
CC
GG
TT
AA
dCTP
Terminal Transferase
CCCC
GG
TT
AACC
Anneal with dG-tailed pBR322
(Pst1 site)
CC
GG
GG
CC
Recombinant plasmid

subsequently tailed with dC residues. In the double tailing method (Fig. 3.2., pathway B), a homopolymeric oligo (dC) tail is added to the 3' end of the single stranded cDNA using terminal transferase. Synthesis of the second strand by reverse transcriptase is primed by annealing oligo $(dG)_{12}$ to the dC-tailed single stranded cDNA. When synthesis of the second strand is complete, the double stranded cDNA is tailed with dC residues. The dC-tailed double stranded cDNA is subsequently annealed with dG-tailed vector DNA under conditions which stabilize the dC-dG base pairing of the complementary homopolymeric tails. The result is the formation of hybrid molecules capable of transforming E. coli.

A major disadvantage associated with the loopback method is the use of the 3' terminus of the single stranded cDNA to prime second strand synthesis and the subsequent removal of the hairpin loop structure by S1 nuclease. This results in the unavoidable loss of cDNA sequences complementary to the 5' end of the mRNA. Therefore, by the very nature of the technique, full length cDNA clones cannot be generated. The double tailing method permits the cloning of the complete 5' ends of mRNAs, by avoiding the limitations associated with the loopback method. The addition of a homopolymeric dC tail to the 3' terminus of the single stranded cDNA acts as a template for oligo $(dG)_{12}$ to bind and prime the synthesis of the second strand.

Therefore, to enable the generation of full length P450 cDNA clones, the double tailing method of Land et al. (1981) was used to construct a chicken liver cDNA library. Previous studies in our laboratory by Brooker et al. (1983) with chick embryo livers suggested that the levels of P450 mRNA induced by AIA and DDC represented approximately 1% of the total mRNA population. Due to the enrichment of

P450 mRNA following drug induction, a small library of approximately 1500 colonies was constructed from poly A^+ RNA.

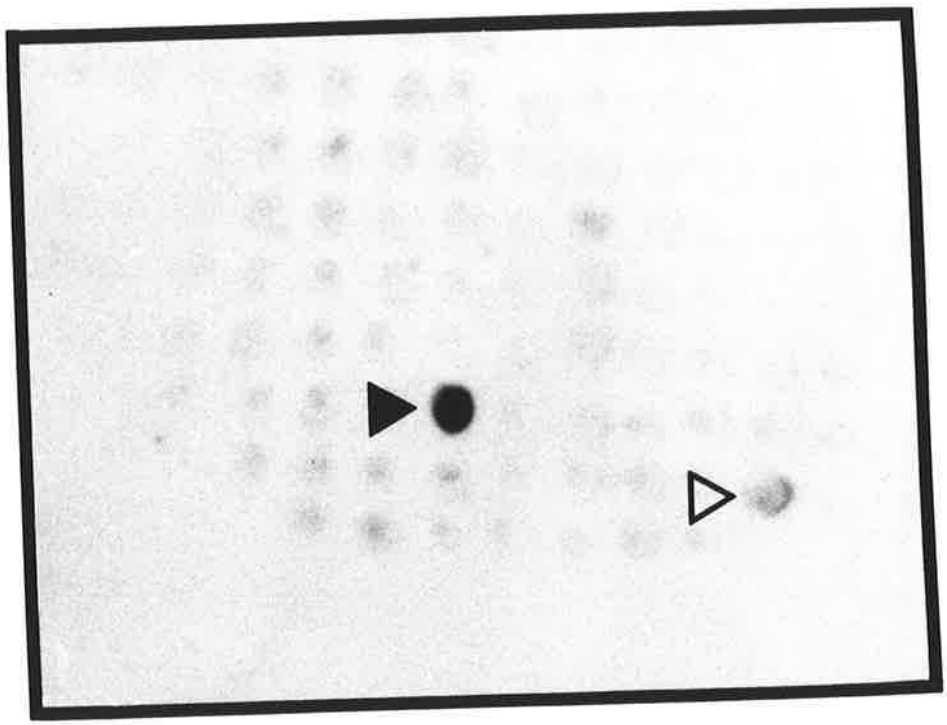
Prior to cloning into E. coli, the tailed double stranded cDNA was size fractionated by electrophoresis on a 1% low melting point agarose gel and fragments larger than the 1500 bp marker were extracted for cloning. Equimolar amounts of the dC-tailed double stranded cDNA and dG-tailed pBR322 DNA were annealed and used to transform E. coli MC1061. A transformation efficiency of $1x10^5$ per ug of pBR322 was obtained with a background level of $2x10^2$ per ug of PstI cleaved dG-tailed pBR322. Of the 1500 colonies obtained, approximately 1200 colonies were tetracycline resistant (tet^R) and ampicillin sensitive (amp^S). Since insertion of DNA into the PstI site results in inactivation of the gene encoding ampicillin resistance, this indicates that approximately 80% of the colonies obtained were recombinants.

3.2.C. Detection of cDNA Clones Containing P450 Sequences.

Recombinants containing P450 sequences were identified by colony hybridization using the previously isolated chicken P450 cDNA clone, p2H10, as a specific probe (Brooker and O'Connor, 1982). Nitrocellulose filters containing tet^R and amp^S recombinants were prepared by the method of Grunstein and Hogness (1975) and probed with [^{32}P]-labelled 720 bp PstI insert derived from p2H10. pBR322 and p2H10 were included as negative and positive controls respectively. Of the 1200 colonies screened, 10 colonies hybridized strongly to the probe. Fig. 3.3 shows the result of screening a portion of this library, illustrating that positive clones were readily identifiable. To confirm that these 10 clones were true positives they were toothpicked onto a nitrocellulose filter and subjected to a second round of screening with the probe. Whereas pBR322 showed negligible cross-reactivity with the probe, each

Fig. 3.3. Detection of P450 cDNA clones by colony hybridization.

Nitrocellulose filters containing tet^R and amp^S colonies were prepared according to the method of Grunstein and Hogness (1975). The filters were probed with the [^{32}P]-labelled 720 bp PstI insert of p2H10 using the hybridization and washing conditions described in Section 2.2.K. Positive clones were readily identifiable although the signal intensities varied from strongly hybridizing clones (▶) to weakly hybridizing clones (▷). p2H10 and pBR322 were included as positive and negative controls respectively, but are not shown.



of the 10 recombinants hybridized strongly, confirming the presence of P450 coding sequences within the cDNA inserts. These cDNA clones were designated pCHP1 through to pCHP10 (plasmid chicken P450 clone 1-10), in order of isolation.

3.2.D. Restriction Analysis of P450 cDNA Clones.

To determine the size of the cDNA inserts, plasmid DNA was prepared (Section 2.2.H) from pCHP1 through to pCHP10. Since the procedure used for the construction of these recombinants regenerated the PstI sites flanking the cDNA inserts, the size of the inserts was easily determined by digestion with PstI and analysis on a 2% agarose gel (Fig. 3.4). Although p2H10 contains a 720 bp PstI fragment, a residual 200 bp of the cDNA insert could not be excised indicating the loss of a PstI site. The inability to regenerate either of the flanking PstI sites during the cloning procedure is not uncommon and some workers have reported as many as 60% of the DNA sequences inserted in this way were irretrievable by digestion with PstI (Villa-Komaroff et al., 1978). The entire cDNA inserts of the ten clones isolated from the library described here, were excisable by digestion with PstI and ranged in size from 1200-2800 bp. The largest of these cDNA clones, pCHP3, contained four PstI fragments of estimated sizes 1350, 700, 650 and 75 bp (Fig. 3.4, lane 4). The fragments were designated A, B, C and D in order of decreasing length. This clone was further characterized by sequencing.

3.2.E. Sequence Analysis of pCHP3.

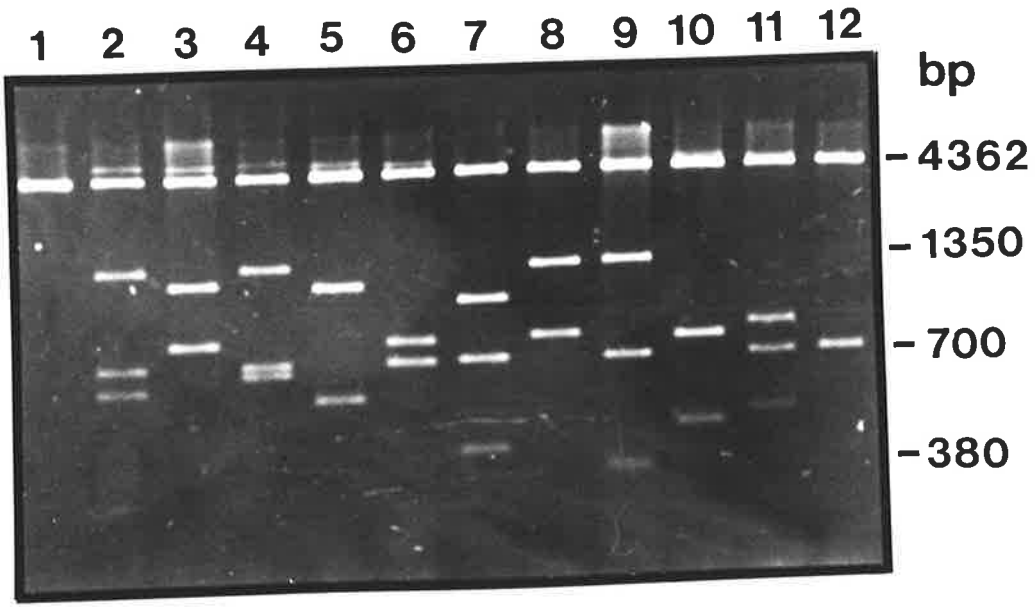
M13 phage shotgun libraries of the cDNA clone pCHP3, were prepared using the DNaseI digestion method outlined previously in Section 2.2.P-1. The M13 phage recombinants were sequenced by the dideoxy sequencing method of Sanger et al. (1980) according to the strategy

Fig. 3.4. Restriction analysis of the ten P450 cDNA clones.

Plasmid DNA was prepared (Section 2.2.H.) from the ten P450 cDNA clones, pCHP1-pCHP10, and digested with PstI. The digestion products were electrophoresed on a 2% agarose gel along with PstI digests of pBR322 and p2H10. The gel was stained with ethidium bromide and the fragments were visualized by U.V. irradiation.

Lanes:

1. pBR322
2. pCHP1
3. pCHP2
4. pCHP3
5. pCHP4
6. pCHP5
7. pCHP6
8. pCHP7
9. pCHP8
10. pCHP9
11. pCHP10
12. p2H10


1
2
3
4
5
6
7
8
9
10
11
12
bp
4362
1350
700
380

shown in Fig. 3.5-A. Sequence analysis of pCHP3 revealed an insert of 2712 bp (Fig. 3.5-B), flanked by dG and dC tails (not shown). This sequence contained an open reading frame starting at the first methionine codon 40 nucleotides from the 5' terminus, giving a predicted protein of 491 amino acids in length. The predicted sequence of the first 15 amino acid residues was identical to that determined by direct amino terminal sequencing of the P450 purified from the livers of chick embryos induced with AIA (Dr G. Srivastava, personal communication). This unambiguously established that pCHP3 was derived from a P450 mRNA and confirmed the identity of the initiation codon.

In order to determine the size and sequence of the complete 5' noncoding region, a [^{32}P]-labelled 17 nucleotide oligomer complementary to the sequence underlined in Fig. 3.5-B, was used in primer extension analysis of poly A^+ RNA isolated from the livers of chick embryos induced with AIA and DDC. The extension products were electrophoresed on a sequencing gel and showed a major extension product of 57 nucleotides in length (Fig. 3.6-A). The dideoxy sequencing reaction used to size the extension product is shown in Fig. 3.6-B. Direct sequencing of the extension product by chemical degradation showed that the sequence was identical to the first 57 nucleotides of pCHP3, with both sequences starting at the same 5' nucleotide. Therefore, the 39 nucleotides of the 5' noncoding sequence in pCHP3 represented the complete 5' noncoding region of the mRNA.

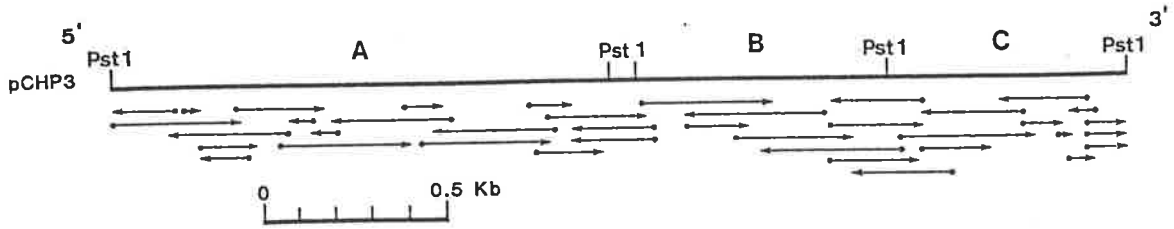
The 3' noncoding region of pCHP3 extended for 1197 nucleotides after the termination codon, TAG. The sequence contained a characteristic oligo $(A)_{14}$ region starting at nucleotide 2399, the significance of which is as yet unclear. The sequence lacked the consensus polyadenylation signal AAUAAA or one of its usual variants (Birnstiel et al., 1985), as well as a poly A tail at its 3' end,

Fig. 3.5. Nucleotide and derived amino acid sequence of pCHP3.

The PstI restriction map for pCHP3 is shown in panel A, with the three largest PstI fragments labelled A, B and C, in order of decreasing length. The sequencing strategy is represented below the restriction map. The direction and extent of sequencing are indicated by arrows.

Panel B presents the nucleotide and derived amino acid sequence of pCHP3. The sequence complementary to the synthetic primer used in primer extension analysis (Fig. 3.6) is underlined. The termination codon TAG is denoted by an asterisk and the oligo $(dA)_{14}$ region starting at nucleotide 2399 is indicated by an arrow.

A

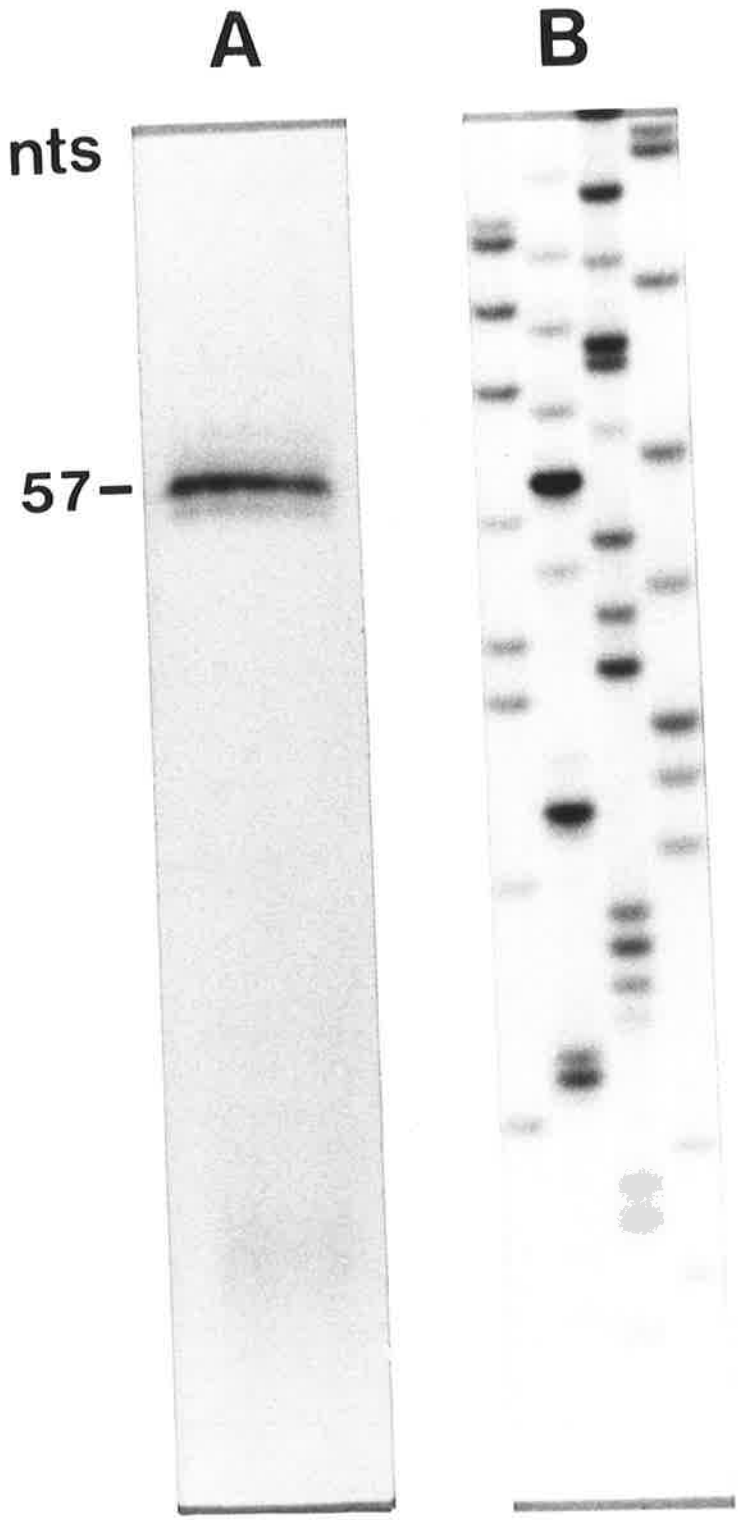


B

```
                                                                M  D  F  L  G  L  P  T  I  L  L  L  V  C  I  S  C  L  L  I  A  A  W  R  S  T  S
AGACTGTTCTGACACTTGACATCTCTTCCTCTGCCCACCATGGACTTCCTGGGATTGCCCACAATCCTCTTGCTGGTCTGTATCTCATCGCTTCTCATTGCTGCATGGAGGAGTACATCG
                                                                                                                                        120

 Q  R  G  K  E  P  P  G  P  T  P  I  P  I  I  G  N  V  F  Q  L  N  P  W  D  L  M  G  S  F  K  E  L  S  K  K  Y  G  P  I
CAAAGAGGGAAGGAGCCTCCTGGTCCCACGCCAATCCCCATCATTGGAAATGTTTTTCAGCTGAACCCGTGGGAGTTGATGGGAAGCTTTAAGGAGCTCAGCAAAAAGTACGGTCCTATC
                                                                                                                                        240

 F  T  I  H  L  G  P  K  K  I  V  V  L  Y  G  Y  D  I  V  K  E  A  L  I  D  N  G  E  A  F  S  G  R  G  I  L  P  L  I  E
TTCACAATACATTTAGGCCCCAAAAAGATTGTGGTGCTGTATGGCTACGACATTGTGAAAGAAGCCCTAATTGATAACGGAGAAGCCTTCAGTGGGAGAGGAATACTGCCTCTGATTGAA
                                                                                                                                        360

 K  L  F  K  G  T  G  I  V  T  S  N  G  E  T  W  R  Q  L  R  R  F  A  L  T  T  L  R  D  F  G  M  G  K  K  G  I  E  E  R
AAGCTGTTCAAGGGCACAGGCATTGTGACCAGCAATGGGGAGACCTGGAGGCAACTGCGACGATTTGCCCTCACCACCCTGCGTGATTTTGGAATGGGGAAGAAGGGCATTGAGGAGCGA
                                                                                                                                        480

 I  Q  E  E  A  H  F  L  V  E  R  I  R  K  T  H  E  E  P  F  N  P  G  K  F  L  I  H  A  V  A  N  I  I  C  S  I  V  F  G
ATCCAGGAGGAAGCTCATTTTCTGGTGGAGAGGATCAGGAAGACACACGAGGAACCTTTCAACCCTGGGAAGTTCTTAATCCATGCTGTTGCCAACATCATCTGCTCCATTGTATTTGGG
                                                                                                                                        600

 D  R  F  D  Y  E  D  K  K  F  L  D  L  I  E  M  L  E  E  N  N  K  Y  Q  N  R  I  Q  T  L  L  Y  N  F  F  P  T  I  L  D
GATCGGTTTGACTACGAGGACAAGAAATTTCTGGATTTAATTGAGATGCTGGAAGAAAATAACAAATACCAGAACAGAATACAAACACTGCTCTACAATTTCTTCCCAACCATCTTAGAT
                                                                                                                                        720

 S  L  P  G  P  H  K  T  L  I  K  N  T  E  T  V  D  D  F  I  K  E  I  V  I  A  H  Q  E  S  F  D  A  S  C  P  R  D  F  I
TCTTTGCCTGGGCCTCATAAAACACTAATAAAAAACACTGAAACGGTTGATGACTTCATTAAAGAAATCGTAATAGCACACCAGGAATCGTTCGATGCCAGGTGCCCTCGAGATTTTATT
                                                                                                                                        840

 D  A  F  I  N  K  M  E  Q  E  K  E  N  S  Y  F  T  V  E  S  L  T  R  T  T  L  D  L  F  L  A  G  T  G  T  T  S  T  T  L
GATGCTTTCATTAACAAAATGCAACAGGAGAAAGAGAACTCTTATTTCACTGTTGAGTCCTTGACCAGAACCACACTTGACTTGTTCGTTGCGGGAACGGGGACAACCAGCACCACACTG
                                                                                                                                        960

 R  Y  G  L  L  I  L  L  K  H  P  E  I  E  E  K  M  H  K  E  I  D  R  V  V  G  R  D  R  S  P  C  M  A  D  R  S  Q  L  P
AGATATGGACTTCTGATTCTCCTGAAACACCCAGAGATTGAAGAGAAAATGCACAAGGAGATTGACCGTGTGGTTGGCCGAGACCGAAGCCCCTGCATGGCAGACAGGAGCCAGTTGCCA
                                                                                                                                       1080

 Y  T  D  A  V  I  H  E  I  Q  R  F  I  D  F  L  P  L  N  V  P  H  A  V  I  K  D  T  K  L  R  D  Y  F  I  P  K  D  T  M
TACACGGATGCCGTCATCCACGAAATCCAAAGATTCATTGATTTCCTTCCCCTTAATGTTCCACATGCTGTGATCAAGGACACCAAGCTCAGAGACTATTTTATCCCCAAGGACACCATG
                                                                                                                                       1200

 I  F  P  L  L  S  P  I  L  Q  D  C  K  E  E  P  N  P  E  K  F  D  P  G  H  F  L  N  A  N  G  T  F  R  R  S  D  Y  F  M
ATATTCCCTTTGCTGTCTCCCATCCTGCAAGACTGCAAGGAATTTCCAAACCCAGAAAAATTTGACCCAGGACACTTCCTGAATGCAAACGGCACGTTCAGAAGGAGTGACTACTTCATG
                                                                                                                                       1320

 P  F  S  A  G  K  R  I  C  A  G  E  G  L  A  R  M  E  I  F  L  F  L  T  S  I  L  Q  N  F  S  L  K  P  V  K  D  R  K  D
CCATTCTCTGCAGGAAAACGCATCTGTGCAGGAGAAGGGCTGGCCCGGATGGAGATATTCGTGTTCTTAACATCCATCCTGCAGAACTTTTCTCTGAAGCCTGTTAAAGATCGCAAGGAC
                                                                                                                                       1440

 I  D  I  S  P  I  I  T  S  L  A  N  M  P  R  P  Y  E  V  S  F  I  P  R  *
ATTGACATTTCTCCAATAATCACCTCTCTGGCAAATATGCCCCGACGTTATGAGGTCTCATTTATTCCACGTTAGACAAGTGAAGACAGGTGCCACCTCCCAAATTCATGAGACGTCTTC
                                                                                                                                       1560

TTGAATGATAGTCAATGCCCAGCCTGATAACACTGAGACACTCAGATAATTTCGTAGATTCTTTAGGGACAGAAGACTTACGTTGTTCTATGTGATGCCAGAGGACATGATGCAGCAGAA
                                                                                                                                       1680

AACAATTCTCCACAAAGCTTAGGCTCCTACACACTGCCCAAAATAGGTACATATCTTGACTGTTTGTTTTCCCTCATTTATAAAAAATGCATTATGGCTTTAAGACACTATCTGGAAACC
                                                                                                                                       1800

CTCCAGTCTGTCTAACAACATGGAAGATTGGTGGACTGTACTTACTGGACCTATCAGAATTAAAAGTGCAATTACAGCACTGCTCATTGAGACAAAACAGAGTGAAATAATTGTGGGATC
                                                                                                                                       1920

TCCACGATGCCTTTTAGAAGATGAATGTCAGCTTGAATCTTCCCATTACTTTTATGAAATACGTAAAAATTGCTACAGAAATTGTACAAACTTCAGGTAAAGGGTTTCTTTGAGCTAACA
                                                                                                                                       2040

AAGACAGATGGATGTAGCATTGGCATTTCCCAGATTGACCGCTTTCCCTGCAGTGTATAGCTTACTGTGATTGCAGTGACAACACAGCTGAATCCATGCTACGTGGCATAATACAGCACT
                                                                                                                                       2160

TGCAATGTCACCACAGTGGCATTAATTGCTTAGGTAAGCAGCCAAATGTTCATTTCTTTCACTGTGTTAATAGTACTGATATACCATGGGTTATTTGTTTTGTGTTATACAATAAGTTCA
                                                                                                                                       2280

AAGCATAAAATATCACAGAATGACAGAAATACAAATGAAAGAAAATACATGAAAATATTGCCTACCCTATCGATATTCGTAGCCTAGACCAACTGTTAGAAGTCTTCTTGACAGAAAGAA
                                                                                                                                       2400

AAAAAAAAAAAAGAAAGAAGGAGAAAGCTCACTGTGCTTCCTTCGACTTTGCCATGGCTAAGTAAATGGCTTGAAAGGCAACCTACGTTCCAGACACCTTCCTCCATACTGAGCAGACAA
                                                                                                                                       2520

GGCAGCAGCTTGAGCAAAGCAGCAAGAAAACAGCTGTAGCTCCCAGAGACTTCACCATCATATGGAGCAGGAATCTGAGATCAGAGAAGCGCTAAAGCTGTGCCCCGCAAGTCTCAACTT
                                                                                                                                       2640

TGTGTCTCACTGTGCTGTGGTACAAGCAGCAGTGCAATAGTTACCAATGGCTGCTTGCACTTCCTCCAAGTTC
                                                                                                                                       2712
```

Fig. 3.6. Primer extension analysis of chicken hepatic poly A^+ RNA.

Extended cDNA products were synthesized using a 5' [^{32}P]-labelled 17 nucleotide primer on poly A^+ RNA, isolated from the livers of chick embryos treated with AIA and DDC (Section 2.2.E-4). The products were analyzed on an 8% sequencing gel (panel A) with a dideoxy sequence of M13 mp18 DNA as size standards (panel B). The major extension product of 57 nucleotides (nts) is indicated.


A
B
nts
57-

suggesting that pCHP3 did not contain the entire 3' noncoding region of the homologous P450 mRNA. As discussed in a later section of this chapter, Northern hybridization analysis supported this contention.

3.2.F. Analysis of the Derived Amino Acid Sequence of pCHP3.

The molecular weight of the protein encoded by pCHP3 was calculated to be 56,196. This is significantly higher than the molecular weight of 50,000 estimated for the protein by SDS polyacrylamide gel electrophoresis, although similar discrepancies have been observed for other drug inducible and constitutively expressed P450s (Heinmann and Ozols, 1983; Kimura et al., 1984; Ozols et al., 1981). The amino terminal region of the predicted protein sequence contains a hydrophobic stretch of 21 amino acids flanked by charged residues, typical of the signal sequence of other P450s. Examination of the protein sequence revealed two potential glycosylation sites at asparagine residues 417 and 456 (Fig. 3.5-B). However, previous work on P450s from other species has found no evidence of post-translational glycosylation (Armstrong et al., 1983; Haugen and Coon, 1976). More significantly, chicken P450s synthesized in vivo migrated to the same position on SDS polyacrylamide gels as chicken P450 sythesized in a cell free wheat germ translation system (Brooker et al., 1983), a system which does not glycosylate proteins. Therefore, it was concluded that the chicken P450 is not post-translationally modified.

3.2.G. Comparison of pCHP3 with other P450s.

Sequence comparisons between pCHP3 and other P450s were performed at the nucleic acid and derived amino acid levels for two reasons. Firstly, it would enable the classification of the protein encoded for by pCHP3 and secondly, permit the identification of conserved regions of

homology between members of different families. Dayhoff (1977) proposed for the purpose of classification, that proteins within a family usually exhibit greater than 50% homology at the amino acid level. To determine which family pCHP3 represented, the DNA coding and derived amino acid sequences were compared with representatives from every other gene family of mammalian P450s identified to date. A second cDNA clone pCHP7, isolated from the cDNA library described here, was recently sequenced by A. Hansen (this laboratory) and made available for comparison. As discussed in a later section of this chapter, pCHP7 corresponds to a 2.2 Kb mRNA species.

The percentage homologies obtained from the best alignment for any two sequences are presented in Table 3.1. The chicken cDNA clones pCHP3 and pCHP7, share strong homology (92%) at the amino acid level. Eight of the thirty seven amino acid substitutions are nonconservative. When comparisons were extended to mammalian members of the PB inducible P450 family, pCHP3 and pCHP7 shared 49-56% homology at the amino acid level. The chicken P450s shared 56% homology with rat P450-PB1, a poorly PB inducible isozyme, and with rabbit P450-1, a constitutive form. 49% homology was observed between the major rat PB inducible form, P450b and the chicken P450s. However, 30% or less homology was observed when the derived amino acid sequences of pCHP3 and pCHP7 were compared to members of other P450 families such as the steroid synthesizing P450s of the adrenals or the TCDD and PCN inducible hepatic isozymes. When similar comparisons were conducted on the DNA coding sequences (Table 3.1), the chicken sequences could be classified as members of the PB inducible family of P450s.

Since all P450s perform a common function, it would be expected that these isozymes contain common domains. When amino acid comparisons were performed between the chicken sequences and other members of the PB

Table 3.1. Percentage sequence homology of pCHP3 and representative members of other P450 gene families.

A protein is grouped to a specific family of P450s if the amino acid sequence homology is greater than 50% (Dayhoff, 1978). Values above the diagonal refer to amino acid homology. Values below the diagonal refer to nucleic acid sequence homology in the coding regions. The references presented below are numbered 1-10 and relate to the P450s compared in the table.

REFERENCES:

1. Hobbs et al. (1986)
2. A. Hansen (personal communication)
3. Gonzalez et al. (1986a)
4. Tukey et al. (1985)
5. Fujii-Kuriyama et al. (1982)
6. Jaiswel et al. (1984)
7. Molowa et al. (1986)
8. White et al. (1986)
9. Morohashi et al. (1984)
10. Zuber et al. (1986)

		1	2	3	4	5	6	7	8	9	10
chicken pCHP3	1		95	63	64	58	48	47	45	43	47
chicken pCHP7	2	92		62	64	58	47	47	45	43	47
rat P450-PB1	3	56	56		77	55	47	38	43	42	45
rabbit P450-1	4	56	56	73		57	49	39	47	42	45
rat P450b	5	49	49	51	51		49	42	45	43	47
human P_1450	6	30	30	30	29	31		42	49	41	46
human HLp	7	23	23	24	23	25	22		40	39	41
human P450c21	8	29	29	27	28	28	25	21		44	51
bovine P450ssc	9	18	18	16	17	18	13	22	17		43
bovine P450-17α	10	27	27	26	26	26	25	21	29	18	

inducible P450 family, extensive regions of homology were observed (Fig. 3.7). When these comparisons are extended to members of other P450 families, such as the human PCN inducible form of P450 (HLp), three distinct highly conserved regions of homology are observed between the P450s (underlined in Fig. 3.7). These are the amino-terminal and carboxyl-terminal conserved cysteine containing peptides, and the "analogous" peptide originally identified from comparisons of rabbit P450-3b and P450-2 (Ozols et al., 1981). Significantly, the amino-terminal conserved region specified by pCHP3 (region 1) has lost the cysteine residue which has been conserved in most other species of P450 and was originally thought to be involved in the thiolate bond providing the 5th ligand to the iron atom of the heme moiety (Tarr et al., 1983). The loss of this residue and the conservation of the cysteine in the carboxyl-terminal region (region 3), strongly indicate that the carboxyl-terminal cysteinyl peptide is involved in heme binding (Gonzalez et al., 1984; Zuber et al., 1986). The "analogous" peptide (region 2) is highly conserved between members of the PB inducible family, but shows poor homology with the corresponding region of P450 HLp, suggesting it plays a role in substrate binding. A detailed comparison of the presumed heme binding site and the "analogous" peptide between 15 members of different P450 families from 5 different species has recently appeared (Zuber et al., 1986) and strongly supports the above findings. Both these regions display extensive sequence homology (57-100%) between orthologous members of the same gene family, but homology is less pronounced between P450 families. For example, rat P450b and rabbit P450-2 display 90% homology in the sequence surrounding the proposed heme binding site and 83% homology in the "analogous" peptide. When the sequence comparisons are extended to rat P450b and rat P450-PCN, the isozymes display only 33% and 26% homology in the

Fig. 3.7. Comparison of the derived amino acid sequences of six members of the PB inducible P450 gene family.

The complete derived amino acid sequence of pCHP3[1] was compared with those of pCHP7[2], rabbit P450-1[3], rat P450-PB1[4], rat P450b[5] and human HLp[6]. The complete sequence of pCHP3 is given. Where amino acid residues are conserved in the other P450 sequences they are indicated as dashes. The boxes represent spaces created for maximal alignment of the sequences. The conserved amino-terminal cysteinyl peptide (region 1), the analogous peptide (region 2) and the carboxyl-terminal conserved cysteinyl peptide (region 3) are underlined.

REFERENCES:

1. Hobbs et al. (1986)
2. A. Hansen (personal communication)
3. Tukey et al. (1985)
4. Gonzalez et al. (1986a)
5. Fujii-Kuriyama et al. (1982)
6. Molowa et al. (1986)


PCHP3
PCHP7
RABBIT 1
RAT PB1
RAT b
HLP
1
2
3

sequence surrounding the proposed heme binding site and the "analogous" peptide respectively. The structural and functional role of these peptides in vivo remains to be determined.

3.2.H. Northern Hybridization Analysis of Chicken P450 mRNAs Homologous to pCHP3.

To characterize the mRNA coding for chicken P450, Northern hybridization analysis was performed on poly A^+ RNA isolated from the livers of chick embryos induced with AIA and DDC. The RNA was denatured by glyoxylation and fractionated on a 1% agarose gel. After transfer to nitrocellulose filters, the RNA was probed with nick translated pCHP3. The resultant autoradiogram (Fig. 3.8) shows a strong band at 3.5 Kb with a weaker band at 2.8 Kb. With longer exposures, a minor 2.2 Kb mRNA can also be observed. Each mRNA has the potential to code for a PB inducible P450, since all three are larger than the 1400 nucleotides required to encode a 50,000 molecular weight protein. Earlier comparison of the sequences of pCHP7 and pCHP3 (Table 3.1; Fig. 3.7) showed that the cDNA clones represent mRNAs which encode separate P450 isozymes, with 92% homology at the amino acid level. Northern hybridization analysis of hepatic poly A^+ RNA isolated from noninduced chick embryos showed no detectable hybridization signal (Fig. 3.8, lane 2), indicating that all three mRNAs are inducible with AIA and DDC.

Since the mRNA fraction used in the previous Northern hybridization analysis was prepared from chick embryos induced with AIA and DDC, the relationship between the individual drugs and the different mRNA species was uncertain. In addition, previous studies by Brooker et al. (1983) indicated that PB also induces the accumulation of chicken P450 mRNAs homologous to those induced by AIA and DDC. Therefore three chick embryos were induced with either AIA, DDC or PB for 16 hours. Northern

Fig. 3.8. Size estimation of chicken P450 mRNAs by Northern hybridization analysis.

5 ug of poly A^+ RNA, isolated from the livers of AIA and DDC treated (lane 1) and untreated (lane 2) chick embryos, was denatured by glyoxylation and electrophoresed on a 1% agarose gel as described in Section 2.2.L. Following transfer to nitrocellulose, the filter was probed with nick translated pCHP3 using the hybridization and washing conditions described in Section 2.2.L.

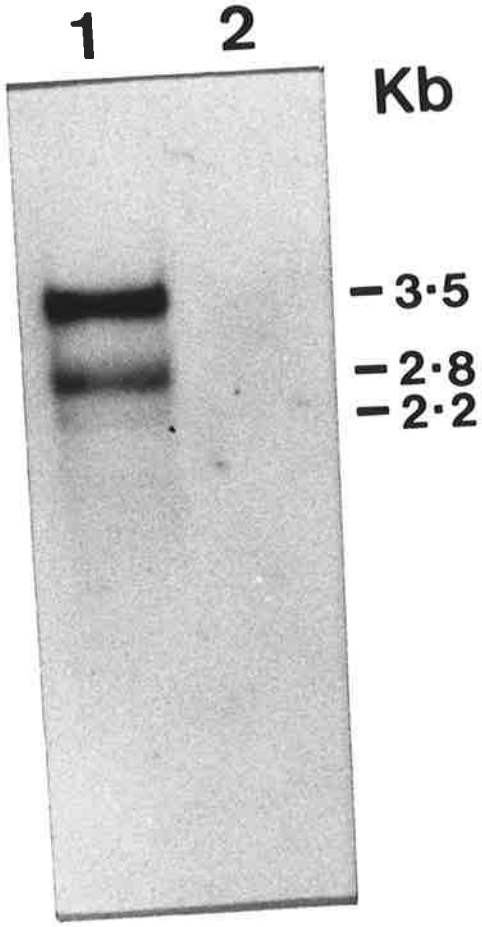
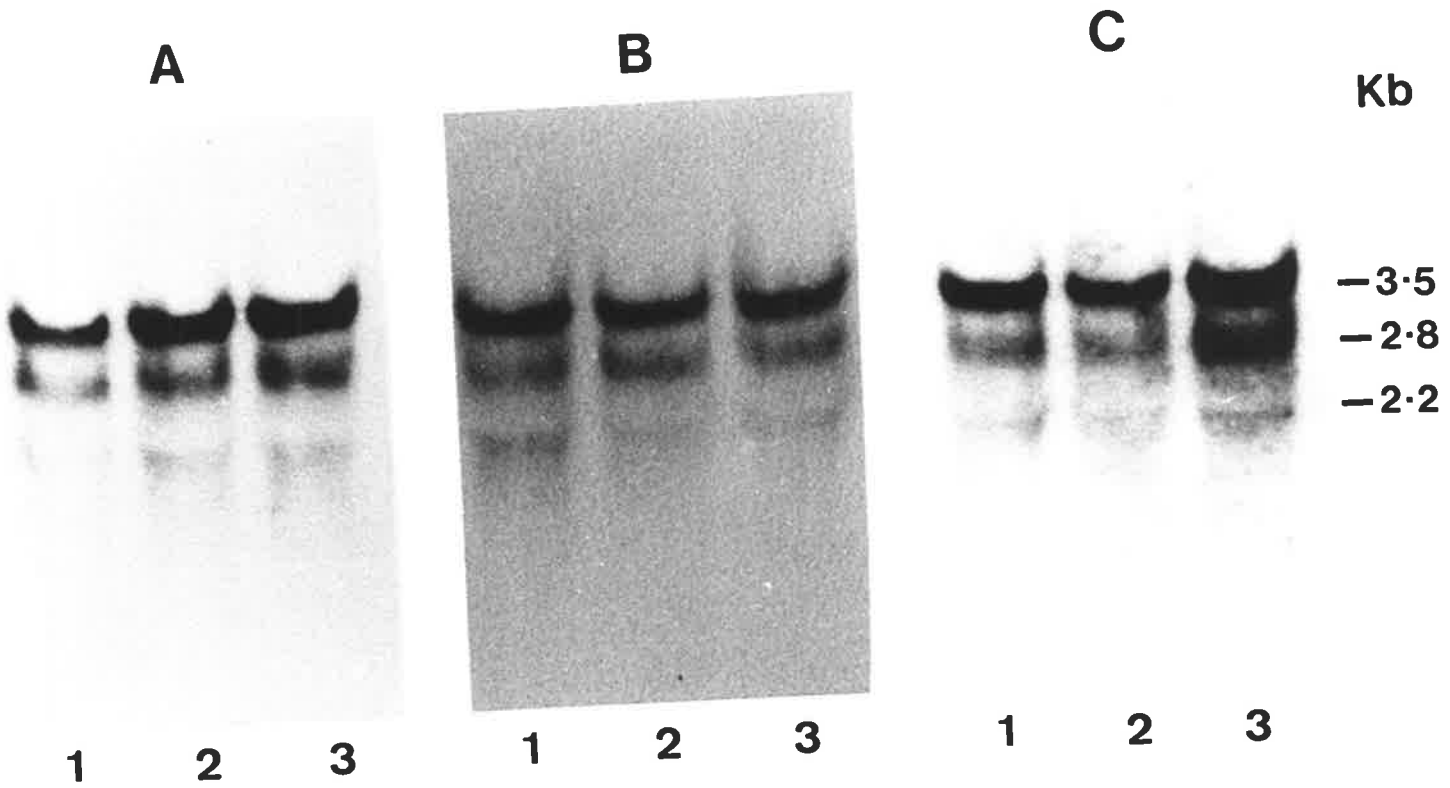

1
2
Kb
3·5
2·8
2·2

Fig. 3.9. Characterization of the three chicken P450 mRNAs in individual embryos by Northern hybridization analysis.

20 ug of total RNA isolated from the livers of chick embryos induced for 18 hours with either AIA (panel A), DDC (panel B) or PB (panel C), was denatured by glyoxylation and electrophoresed on a 1% agarose gel. The RNA was transferred to nitrocellulose, probed with nick translated pCHP3 as described in Section 2.2.L and autoradiographed at -70^{o}C. The exposure times for panels A, B and C were 7, 13 and 93 hours respectively. In each panel, lanes 1, 2 and 3 refer to separate RNA samples isolated from the livers of individual chick embryos following treatment with the appropriate drug.


A
B
C
Kb
3·5
2·8
2·2
1
2
3
1
2
3
1
2
3

hybridization analysis was performed on total RNA isolated separately from the liver of each individual embryo. The RNA was denatured by glyoxylation and fractionated on a 1% agarose gel. Following transfer to nitrocellulose, the filters were probed with nick translated pCHP3. As shown in Fig. 3.9, AIA, DDC and PB independently induce all three mRNA species in each individual embryo tested. The relative intensity of the hybridization signals obtained with the various drug treatments indicates that AIA was the strongest inducer of these P450 mRNAs, at the time point measured.

3.2.I. The Relationship between pCHP3 and the Three Homologous mRNA Species.

Although the cDNA insert of pCHP3 shows strong homology with the 3.5 Kb and 2.8 Kb mRNA species, the nature of cDNA cloning specifies that pCHP3 can only be derived from one of these mRNAs. The 2.7 Kb insert of pCHP3 must be derived from either the 3.5 Kb or 2.8 Kb mRNA. To investigate the relationship between pCHP3 and the three mRNAs more directly, Northern hybridization analysis was performed on poly A^+ RNA isolated from the livers of chick embryos induced with AIA and DDC, using DNA probes complementary to the coding and noncoding regions of pCHP3. For the purpose of synthesizing oligonucleotide probes capable of distinguishing between the coding regions of the three chicken P450 mRNAs, the derived amino acid sequence of pCHP3 was compared with that of rat P450b, a major PB inducible isozyme. As indicated in Fig. 3.10-A, regions A and B denote two peptides whose corresponding DNA sequence was used to synthesize oligonucleotide probes. Region A represents a highly divergent region between the two P450s, whereas region B represents a highly conserved region, containing the cysteinyl residue assumed to be involved in the active site of all P450s. Northern

Fig. 3.10. The relationship between pCHP3 and the three homologous chicken P450 mRNAs.

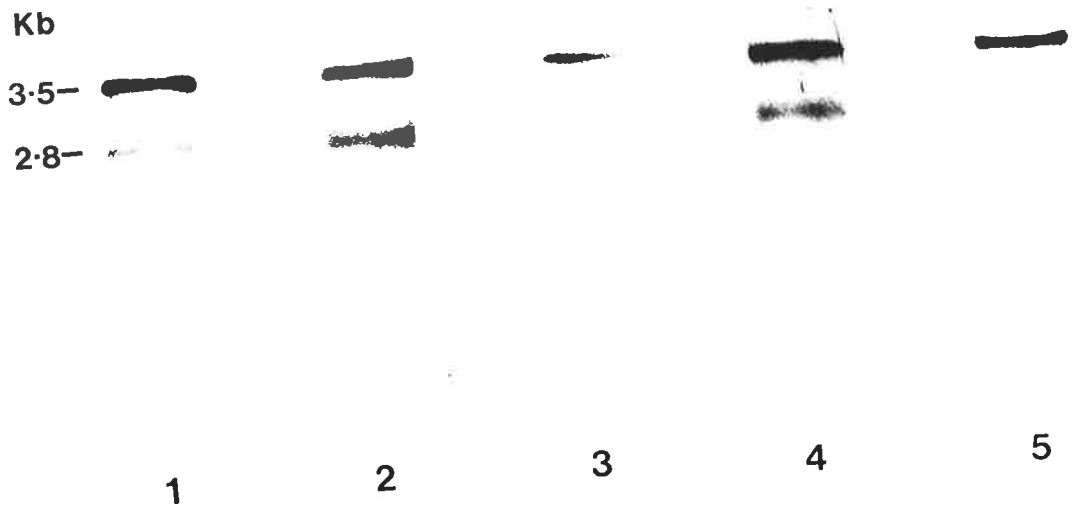
Panel A: A comparison of the derived amino acid sequence of pCHP3 and rat P450b is shown. Dashes indicate positions of conserved amino acids between the two sequences. As discussed previously in the text (Section 3.2-G), the underlined sequences are the previously identified amino-terminal conserved region (region 1), the analogous peptide (region 2), and the carboxyl-terminal conserved region (region 3). The filled circles indicate two peptides (regions A and B), the corresponding DNA sequences of which were used to synthesize oligonucleotides.

Panel B: Poly A^+ RNA (5ug) was electrophoresed on an agarose gel following treatment with glyoxal. After transfer to nitrocellulose, the filters were hybridized with [^{32}P]-labelled probes, specific for conserved and nonconserved regions of the mRNA. The filters were probed with pCHP3 (lane 1), an oligonucleotide complementary to the 30 nucleotides at the 5' end of pCHP3 (lane 2), an oligonucleotide complementary to the 24 nucleotides encoding peptide A (see above) (lane 3), an oligonucleotide complementary to the 30 nucleotides encoding peptide B (see above) (lane 4), or the 650 bp PstI fragment (fragment C) at the 3' end of pCHP3 (lane 5).

A

```
                                                                    50
CHICKEN  MDFLGLPTILLLVCISCLLIAAWRSTSQRGKEPPGPTPIPIIGNVFQLNPWDLMGSFKELSKKYGPIFTIH
RAT      -EPSI-LLLA---G□F--LLVRGHPKS--NF----R-L-LL--LL--DRGG-LN--MG-RED--DV--V-

                                100
CHICKEN  LGPKKIVVLYGYDIVKEALIDNGEAFSGRGILPLIEKLFKGTGIVTSNGETWRQLRRFALTTLRDFGMGKK
RAT      ---RPV-M-C-T-TI----VGQA-D-----TIAV--PI--EY-VIFA---R-KA----S-A-M-------R

                                                        200 A
                150
CHICKEN  GIEERIQEEAHFLVERIRKTHEEPFNPGKFLIHAVANIICSIVFGDRFDYEDKKFLDLIEMLEENNKYQNR
RAT      SV---V----QC---EL--SQGAPLD-TFLFQCIT-----T----E----T-RQ--R-L-LFYRTFSLLSS
              1

                                        250
CHICKEN  IQTLLYNFFPTILDSLPGPHKTLIKNTETVDDFIKEIVIAHQESFDASCPRDFIDAFINKMEQEKEN□SY
RAT      FSEQVFE--SGF-KYF--A-RQISL-LQEIL-YIGH--EK-RATL-PNA------TYLLR--K--S-HHTE

                                                                350
                        300
CHICKEN  FTVESLTRTTLDLFLAGTGTTSTTLRYGLLILLKHPEIEEKMHKEIDRVVGRDRSPCMADRSQLPYTDAVI
RAT      -HH-N-MISL-S--F---E-S-------F-LM--Y-HVA--VQ----Q-I-SH-L-TLD---KM--------
                                                                          2

                                                400
CHICKEN  HEIQRFIDFLPLNVPHAVIKDTKLRDYFIPKDTMIFPLLSPILQDCKEFPNPEKFDPGHFLNANGTFRRSD
RAT      ------S-LV-IG---R-T---MF-G-LL--N-EVY-IR-SA-HYPQY-DH-DS-N-E---D---ALKK-E

                    B                   450
CHICKEN  YFMPFSAGKRICAGEGLARMEIFLFLTSILQNFSLKPVKDRKDIDISPIITSLANMPRPYEVSFIPR
RAT      A-----T-----L---I--N-L---F-T------VSSHLAP----LT-KESGIGKI-PT-QIC-SA-
              3
```

B



hybridization analysis of induced chicken poly A^+ RNA using these and other probes specific for the 5' and 3' noncoding regions of pCHP3 is shown in Fig. 3.10-B.

When pCHP3 was used as a probe, all three mRNAs were detected. However, the faint hybridization signal obtained with the 2.2 Kb mRNA in the exposure shown (Fig. 3.10-B, lane 1), was not reproducible by photography. An oligonucleotide complementary to the 30 nucleotides at the extreme 5' end of pCHP3 hybridized strongly to both the 3.5 Kb and 2.8 Kb mRNA species (lane 2). Similarly, an oligonucleotide complementary to the highly conserved region (region B), also hybridized to both mRNAs (lane 4). However, both an oligonucleotide complementary to a sequence encoding a highly divergent region (region A), and the 650 bp 3' PstI fragment of pCHP3 (Fig. 3.4-A, fragment C), hybridized only to the larger mRNA (lanes 3 and 5). This result unequivocally establishes that pCHP3 is derived from the 3.5 Kb mRNA. Since the 3.5 Kb and the 2.8 Kb mRNA species differ both in their coding and noncoding regions, these mRNAs are not colinear with differences attributable to different polyadenylation sites.

As mentioned previously, the chicken cDNA clone pCHP7, has recently been shown in this laboratory to be full length for the 2.2 Kb mRNA (A. Hansen, personal communication). No observable hybridization of the 3' PstI fragment of pCHP3 (fragment C) to the 2.2 Kb mRNA occurred. Therefore, it was tentatively concluded that the 3.5 Kb and 2.2 Kb mRNA species were not homologous in this region. The subsequent sequence comparison of pCHP3 and pCHP7 in their respective 3' noncoding regions confirmed this conclusion.

When Northern hybridization analysis was performed using oligonucleotide probes corresponding to sequences present in pCHP3 and pCHP7, no detectable hybridization of these probes to the 2.2 Kb mRNA

was observed. The lower abundance of the 2.2 Kb mRNA, rather than a lack of sequence homology, was considered responsible for the inability to detect this mRNA by Northern hybridization using oligonucleotide probes.

3.3. DISCUSSION.

This chapter described the preparation and characterization of cDNA clones complementary to mRNAs encoding drug inducible chicken P450s. mRNA was extracted from the livers of AIA and DDC treated chick embryos and used to generate a library of cDNA clones. The double tailing method of Land et al. (1981) was used in an attempt to generate full length cDNA clones. Colony hybridization analysis with a 720 bp insert from a previously isolated chicken P450 cDNA clone, detected the presence of ten homologous cDNA clones. The 2.7 Kb cDNA insert of the largest recombinant, pCHP3, was fully sequenced and represented the first non-mammalian eukaryotic P450 sequence to be characterized.

pCHP3 contains a complete 5' noncoding region which is similar in length to those of the rat PB inducible P450b and P450e mRNAs, but shows no sequence homology with these regions (Suwa et al., 1985). A similar lack of homology has also been reported amongst the 5' noncoding regions of mouse 3MC inducible P450 mRNAs (Kimura et al., 1984). The initiation codon in the 3.5 Kb mRNA is the first AUG triplet downstream from the 5' end, which is consistent with Kozak's modified model for the initiation of translation (Kozak, 1981). The reading frame encoded a predicted protein of 491 residues and the first fifteen amino acids were confirmed by amino-terminal amino acid sequencing of a major P450 protein purified from the livers of chick embryos after induction with AIA. Since these sequences were identical it was concluded that the signal peptide is not removed during insertion of the protein into the microsomal membrane.

This is consistent with other mammalian hepatic P450s (Ozols et al., 1981; Ozols et al., 1983).

Northern hybridization analysis of AIA and DDC induced chicken mRNA established unequivocally that pCHP3 was derived from the 3.5 Kb mRNA. Since the 2712 bp cDNA insert contains the complete coding and 5' noncoding regions, the absence of a poly A tail or consensus polyadenylation signal confirms the observation that pCHP3 is not full length. The increased size of the 3.5 Kb chicken mRNA compared to the 2.0-2.5 Kb size generally found for mammalian P450s, is due to the unusually long 3' noncoding region. Further details on the 3' noncoding region will be dealt with in Chapter 4.

Northern hybridization analysis detected the presence of two additional drug inducible chicken P450 mRNAs measuring 2.8 Kb and 2.2 Kb in length, which were highly homologous to pCHP3. Further studies showed that all three mRNAs were induced by each of the drugs AIA, DDC and PB in the livers of individual chick embryos, indicating that they are not due to polymorphisms within the chicken population. Therefore, it seems reasonable to conclude that these three mRNAs are derived from genes which are members of the PB inducible P450 family.

The failure to detect noninduced levels of the chicken mRNAs by Northern hybridization analysis using a glyoxal gel system, made quantitation of the induction response impossible. This may be due to incomplete transfer and/or binding of the mRNA to the nitrocellulose. More recent studies in this laboratory using formaldehyde gel systems, have shown that AIA produces a 70-100 fold increase in the hepatic level of all three chicken mRNAs within 3-6 hours (A. Hansen, personal communication). This level of induction is higher than that previously observed for mammalian members of the PB inducible P450 family. For example, in the livers of rats, PB produces a 20-50 fold increase in the

level of P450b mRNA, a major PB inducible isozyme, within 3-4 hours (Omiecinski, 1986). Therefore, the greater level of P450 mRNA induction observed in the livers of chick embryos treated with AIA, may indicate an increased potency of this drug to induce members of the PB inducible P450 family. Although the results of Northern hybridization analysis of total RNA isolated from the livers of AIA and PB induced chick embryos supports this proposal (Fig. 3.9), the level of P450 mRNA induction was examined at only one time point. Therefore, further studies on the time course of P450 mRNA induction by AIA and PB need to be performed to establish this generality.

The relationship between the 3.5 Kb mRNA and the two additional 2.8 Kb and 2.2 Kb mRNAs was investigated using DNA probes specific for defined regions of pCHP3. Northern hybridization analysis indicated that the 3.5 Kb and 2.8 Kb mRNAs were highly homologous in their respective 5' non-coding regions, but differed in their coding and 3' non-coding regions. During these studies, no detectable hybridization of the oligonucleotide probes to the 2.2 Kb mRNA was observed. It is probable that this was due to a combination of the low abundance of the 2.2 Kb mRNA and the sensitivity problems often associated with the use of oligonucleotide probes. Therefore the lack of hybridization was probably not a true indication of the degree of homology between the 3.5 Kb and 2.2 Kb mRNAs.

Although the analysis of the nucleotide sequences for each of the three mRNAs would permit a more detailed comparison, the studies presented here indicate that in at least two cases, the chicken P450 mRNAs are not simply colinear, with differences attributable to differences in polyadenylation sites. The relationship of the mRNAs to the corresponding chicken P450 genes will be dealt with in Chapter 7.

During the preparation of this thesis, studies performed in this laboratory indicated that four of the original cDNA isolates described earlier in this chapter, corresponded to either the 2.8 Kb or 2.2 Kb mRNAs. The DNA sequence of the largest of these cDNA clones, pCHP7, was recently determined and shown to be full length for the 2.2 Kb mRNA (A. Hansen, personal communication). Comparison of the nucleic acid and derived protein sequences of pCHP3 and pCHP7, showed that the 3.5 Kb and 2.2 Kb mRNAs encoded P450 isozymes containing 491 amino acid residues, which share 92% homology in their primary sequences. The 5' noncoding regions of both mRNAs are identical and the 670 nucleotide 3' noncoding region of pCHP7 shares 56% homology with the corresponding 3' noncoding region of pCHP3. Studies are currently underway to determine the sequence of the cDNA clones corresponding to the 2.8 Kb mRNA.

As discussed earlier in Section 1.6 of Chapter 1, the PB inducible P450 family is comprised of two subfamilies named PB I and PB II. Members of these subfamilies are related by sequence homology and not necessarily by PB inducibility. Therefore, to enable the classification of the P450 proteins encoded by the 3.5 Kb and 2.2 Kb chicken P450 mRNAs, comparison of the derived amino acid sequences of pCHP3 and pCHP7 with those of mammalian P450s was performed. The chicken P450s shared the greatest extent of homology (56%) with rat P450-PB1, a poorly PB inducible isozyme and rabbit P450-1, a constitutive form. 49% homology was observed between pCHP3 and pCHP7, and the major rat PB inducible form, P450b. Since rat P450-PB1 and rabbit P450-1 are members of the PB II subfamily (Govind et al., 1986; Adesnik and Atchison, 1985), on the basis of sequence homology it would seem reasonable to classify pCHP3 and pCHP7 as members of this subfamily. Previous Northern hybridization analysis has shown that the three chicken P450 mRNAs are

slightly inducible by PB, but to a greater extent by AIA. Whether AIA induces the mammalian members of the PB II subfamily is not known.

CHAPTER 4

cDNA CLONING OF THE 3' END OF THE 3.5 Kb CHICKEN mRNA.

4.1. INTRODUCTION.

In the previous chapter, the cDNA clone pCHP3 was shown to contain a 2.7 Kb insert complementary to the 3.5 Kb chicken P450 mRNA. Sequence and Northern hybridization analysis revealed that pCHP3 was not full length and lacked a significant portion of the 3' noncoding region adjacent to the poly A tail. Since pCHP3 contained the longest cDNA insert but still lacked sequences complementary to the 3' end of the mRNA, it was considered unlikely that the other cDNA clones isolated from the library would cover this region. For this reason a second cDNA library was generated using a technique which preferentially clones the 3' end of mRNAs. This chapter discusses the construction of a second chicken cDNA library using the loopback method of Maniatis et al. (1982) and the isolation and characterization of recombinants specific for the 3' noncoding region of the 3.5 Kb P450 mRNA.

4.2. RESULTS.

4.2.A. Construction and Screening of a Second cDNA Library.

A second cDNA library was constructed using the loopback method of Maniatis et al. (1982) which results in the preferential cloning of the 3' ends of mRNAs. The cloning strategy is diagrammatically outlined in Fig. 3.3 of Chapter 3 and the experimental conditions detailed in Section 2.2.G. Briefly, single stranded cDNA was synthesized from 10 ug of poly A^+ RNA, isolated from the livers of chick embryos induced with a combination of AIA and DDC. Synthesis of the second strand was achieved using AMV reverse transcriptase, by self priming of the 3' end of the single stranded cDNA. The hairpin loop was removed by S1 nuclease digestion and the double stranded cDNA subsequently tailed with dC residues. The double stranded dC-tailed cDNA was not size fractionated prior to cloning, in an attempt to preserve small P450 cDNAs. Equimolar

amounts of dG-tailed pBR322 and dC-tailed double stranded cDNA were annealed and the mixture used to transform E. coli MC1061, generating $5x10^4$ transformants per ug of double stranded cDNA. A small library of only 1500 colonies was generated due to the enrichment of P450 mRNAs following drug induction with AIA and DDC. Approximately 1200 colonies, which represented recombinants, were shown to be tet^R, amp^S.

Although the primary objective of the work was to obtain recombinants complementary to the 3' end of the P450 mRNA, it was decided to use the entire cDNA insert of pCHP3 as an initial probe, in order to detect all P450 cDNA clones present in the library. Since both the new cDNA library and pCHP3 were cloned into pBR322, it was necessary to isolate the cDNA insert of pCHP3 before it could be used as a probe. As shown in Fig. 3.3 of Chapter 3, digestion of pCHP3 with PstI generates four cDNA fragments termed A, B, C and D. Since fragments A, B and C were readily obtainable and effectively covered the entire length of the cDNA insert, these fragments were collectively chosen to screen the library. Positive clones were subsequently screened with the 3' PstI fragment of pCHP3 (fragment C), for the purpose of identifying cDNA clones containing sequences complementary to the 3' end of the 3.5 Kb mRNA.

Strong hybridization signals were observed with twenty eight colonies following two successive rounds of screening with fragments A, B and C (results not shown). These colonies were picked for further analysis of their cDNA inserts.

4.2.B. Characterization of the P450 cDNA Clones by Southern Hybridization Analysis.

Plasmid DNA was prepared from the P450 clones (Section 2.2.H) and designated pCHPB1 through to pCHPB28 (plasmid chicken P450 library B

isolates 1-28). The size of the inserts was determined by digestion with PstI and analysis by electrophoresis on a 2% agarose gel. A representative sample of ten clones is shown in Fig. 4.1. All the isolated recombinants contained inserts excisable with PstI, ranging from 390-1580 bp in size.

Previous sequence analysis and Northern hybridization studies showed that fragment C of pCHP3 could be used as a specific probe for the 3' noncoding region of the 3.5 Kb mRNA (Chapter 3). Therefore, to determine which of the cDNA inserts contained sequences complementary to the 3.5 Kb mRNA, Southern hybridization analysis was performed. The DNA fragments were transferred bidirectionally to nitrocellulose, according to the procedure of Smith and Summers (1980). Southern hybridization analysis of the top filter with nick translated fragment C, revealed that twenty of the twenty eight clones contained sequences homologous to the 3' noncoding region of the 3.5 Kb mRNA. As shown in Fig. 4.2-A, the largest PstI fragments which hybridized to fragment C, measured 840 bp and represented the cDNA inserts of pCHPB15, pCHPB16 and pCHPB28 (lanes 5, 6 and 10 respectively). Southern hybridization analysis of the duplicate filter revealed that a small 200 bp PstI fragment of pCHPB28, which did not hybridize to fragment C, hybridized strongly with fragment B of pCHP3 (Fig. 4.2-B, lane 10) and therefore contained sequences which extended in the 5' direction. Similarly restricton fragments of other cDNA clones which did not cross react with fragment C, hybridized strongly with fragment B (Fig. 4.2-B, lanes 3, 8 and 9). One exception was the 550 bp PstI fragment of pCHPB8 (Fig 4.1, lane 3) which did not hybridize to fragments B and C (Fig. 4.2-B and -C, lanes 3), and therefore was expected to hybridize to fragment A of pCHP3. Since the 840 bp fragments of pCHPB15, pCHPB16 and pCHPB28 did not hybridize to fragment B, it was considered that these fragments must

Fig. 4.1. Restriction analysis of chicken P450 cDNA clones.

Plasmid DNA was prepared from the twenty eight cDNA clones (Section 2.2.H), digested with PstI and electrophoresed on 2% agarose gels. The gels were stained with ethidium bromide and the DNA visualized by U.V. irradiation. A representative sample of ten clones along with PstI digested pCHP3 is shown.

Lanes:

1. pCHPB1
2. pCHPB6
3. pCHPB8
4. pCHPB11
5. pCHPB15
6. pCHPB16
7. pCHPB20
8. pCHPB21
9. pCHPB22
10. pCHPB28
11. pCHP3

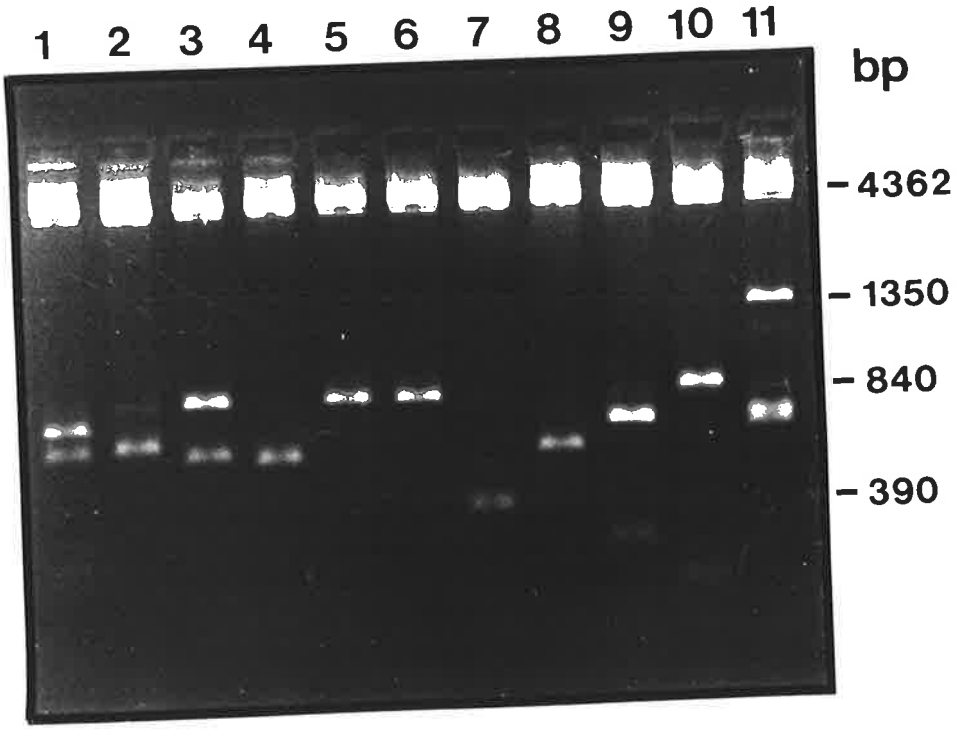

1
2
3
4
5
6
7
8
9
10
11
bp
4362
1350
840
390

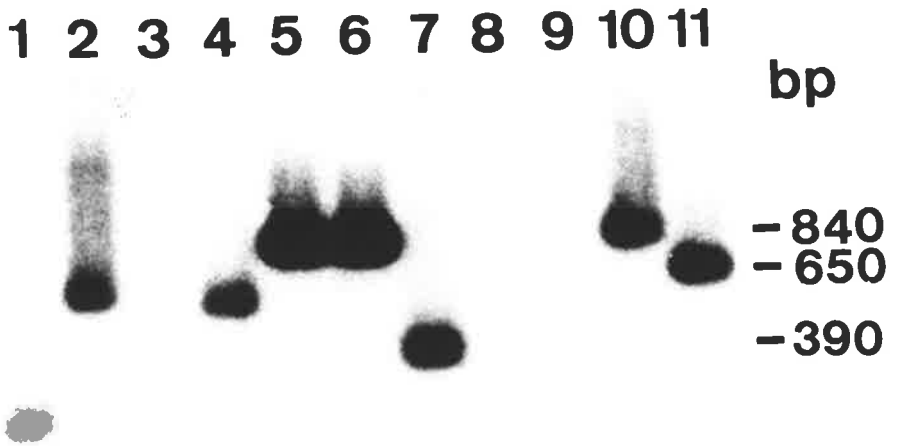
Fig. 4.2. Southern hybridization analysis of chicken P450 cDNA clones.

The gel presented in Fig. 4.1 was used in Southern hybridization analysis. The DNA was transferred bidirectionally to nitrocellulose according to the method of Smith and Summer (1980). The duplicate filters were probed with either nick translated fragment C of pCHP3 (panel A) or nick translated fragment B of pCHP3 (panel B)

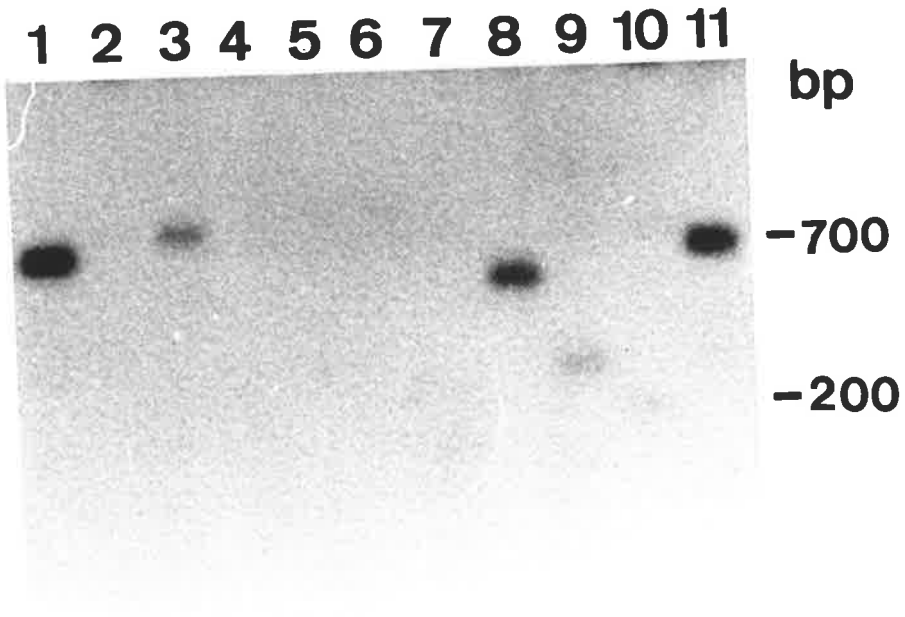
Lanes:

1. pCHPB1
2. pCHPB6
3. pCHPB8
4. pCHPB11
5. pCHPB15
6. pCHPB16
7. pCHPB20
8. pCHPB21
9. pCHPB22
10. pCHPB28
11. pCHP3

A.



B.



extend the sequence of pCHP3 in the 3' direction. This presumption is based on the observation that the cDNA fragments are longer than fragment C of pCHP3 and do not extend 5' into regions homologous to fragment B.

Since the restriction patterns of pCHPB15, pCHPB16 and pCHPB28 were identical, with the exception of an additional PstI fragment in the cDNA insert of pCHPB28, these plasmids were considered to contain the same cDNA sequence (results not shown). Therefore, only pCHPB15 was chosen for sequence analysis. Since several of the cDNA clones contained a single 390 bp PstI fragment which hybridized strongly to fragment C, one representative clone, pCHPB20, was also sequenced for the purpose of comparison with pCHPB15 and pCHP3.

4.2.C. Sequence Analysis of pCHPB15 and pCHPB20.

An M13 phage shotgun library of pCHPB15 was constructed using the DNase I digestion method previously described in Section 2.2.O-1. The 390 bp PstI fragment of pCHPB20 was cloned in both orientations into PstI cut M13mp18 vector DNA. The recombinants were sequenced using the dideoxy chain termination method of Sanger (1980).

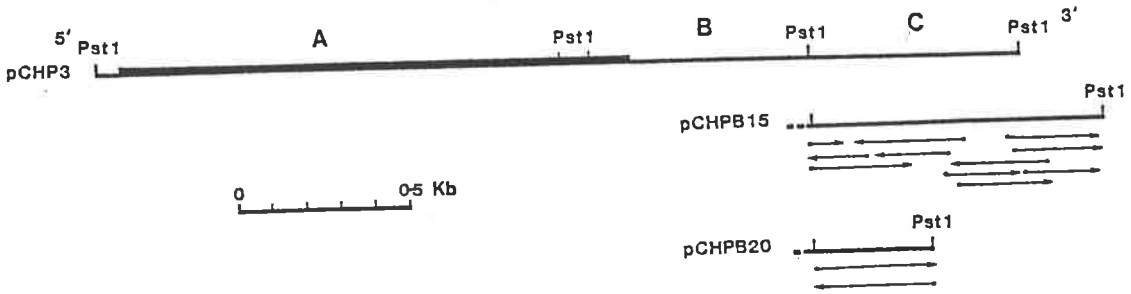
pCHPB15 was sequenced as indicated in Fig. 4.3-A. The 840 bp fragment contained a region of 628 nucleotides identical to pCHP3, with the exception of an additional guanosine residue at position 2399 (Fig. 4.3-B). The sequence of pCHPB15 extended the sequence of pCHP3 a further 134 nucleotides in the 3' direction and terminated with a poly A tract followed by the oligo (dC) tails (not shown). However no consensus polyadenylation signal (Birnstiel et al., 1985) was found close to this poly A tail. Sequence analysis of pCHPB20 revealed sequences identical to the 5' end of pCHPB15, with a 3' end extending to nucleotide 2398 of pCHP3 (arrow 1), followed by an oligo (dA) tract and

Fig. 4.3. Combined nucleotide sequence analysis of pCHPB15 and pCHPB20.

The PstI restriction maps for pCHP3, pCHPB15 and pCHPB20 are shown in panel A, with the three largest PstI fragments of pCHP3 labelled A, B and C. The sequencing strategy for both pCHPB15 and pCHPB20 is represented below the restriction maps by a series of arrows which indicate the direction and extent of sequencing. The 5' ends of pCHPB15 and pCHPB20, indicated by the dashed lines, were not sequenced.

Panel B presents the combined nucleotide sequence data from pCHPB15 and pCHPB20. Arrow 1 refers to the 3' end of pCHPB20 and in addition, refers to the extra guanosine residue in pCHPB15. Arrow 2 indicates the 3' end of pCHP3. The sequence is numbered with respect to pCHP3.

A



B

CTGCAGTGTATAG
2100

CTTACTGTGATTGCAGTGACAACACAGCTGAATCCATGCTACGTGGCATAATACAGCACT
2160

TGCAATGTCACCACAGTGGCATTAATTGCTTAGGTAAGCAGCCAAATGTTCATTTCTTTC
2220

ACTGTGTTAATAGTACTGATATACCATGGGTTATTTGTTTTGTGTTATACAATAAGTTCA
2280

AAGCATAAAATATCACAGAATGACAGAAATACAAATGAAAGAAAATACATGAAAATATTG
2340

↓1

CCTACCCTATCGATATTCGTAGCCTAGACCAACTGTTAGAAGTCTTCTTGACAGAAAGAA
2400

AAAAAAAAAAAAGAAAGAAGGAGAAAGCTCACTGTGCTTCCTTCGACTTTGCCATGGCTA
2460

AGTAAATGGCTTGAAAGGCAACCTACGTTCCAGACACCTTCCTCCATACTGAGCAGACAA
2520

GGCAGCAGCTTGAGCAAAGCAGCAAGAAAACAGCTGTAGCTCCCAGAGACTTCACCATCA
2580

TATGGAGCAGGAATCTGAGATCAGAGAAGCGCTAAAGCTGTGCCCCGCAAGTCTCAACTT
2640

TGTGTCTCACTGTGCTGTGGTACAAGCAGCAGTGCAATAGTTACCAATGGCTGCTTGCAC
2700

↓2

TTCCTCCAAGTTCACCCGATCGAACAAAAGTGGTATTAGCACTTCCTGGGAGATTAAGGA
2760

TCAGACGTTTAAAAGGGAAGAAGACTTAACGAGCAGGCTTTATTTCATCAGTCTTTGTTG
2820

CCTGCTAAACTGGAAAACATTAGTTAAAAAAAAAAAAAAA
2820

oligo (dC) tails due to the cloning procedure. From the sequence analysis it appeared that the insert of pCHPB20 was generated by priming off the oligo (dT) primers which had annealed, during the original cDNA synthesis, to the oligo (dA) region following nucleotide 2398. The absence of oligo (dC) tails at the 5' ends of both the 840bp and 390 bp cDNA fragments of pCHPB15 and pCHPB20 respectively, suggested that the cDNA inserts of these clones were longer than the sequenced fragments. This implies that these cloned inserts contained an extra internal PstI site.

4.3. DISCUSSION.

In an attempt to obtain cDNA clones complementary to the extreme 3' end of the 3.5 Kb chicken mRNA, a second cDNA library was generated using the loopback method of Maniatis et al. (1982). As discussed previously in Chapter 3, this technique preferentially clones the 3' ends of mRNAs. Colony hybridization analysis using three PstI fragments of the previously characterized chicken clone pCHP3, detected twenty eight homologous clones. The cDNA fragments of two clones were shown by hybridization and sequence analysis to be complementary to the 3' PstI fragment of pCHP3 (fragment C). Sequence analysis of the cDNA clone pCHPB20 revealed that the 390 bp fragment was identical to the sequence of pCHP3 from nucleotide 1977 to nucleotide 2367 and offered no new sequence. However, the 840 bp cDNA fragment of pCHPB15 was found to extend the sequence of pCHP3 by 134 nucleotides in the 3' direction, terminating in a poly A tract of 15 nucleotides.

In Chapter 3, Northern hybridization analysis of AIA and DDC induced chick embryo hepatic mRNA revealed that the 2712 nucleotide cDNA insert of pCHP3 was derived from the 3.5 Kb mRNA. However, together the two cDNA clones pCHP3 and pCHPB15 contain a sequence of 2846

nucleotides, not including the poly A tail. Therefore, assuming the 3.5 Kb mRNA contains a poly A tail measuring 200-300 nucleotides long, the length of the cDNA sequence presented here is still 350-400 nucleotides short of the expected size of the mRNA. In addition, the absence of a hexanucleotide similar to the polyadenylation signal AAUAAA or one of its usual variants (Birnstiel et al., 1985) close to the poly A tail of pCHPB15, confirms the observation that pCHPB15 lacked a significant portion of the 3' noncoding region. Therefore it was concluded that the poly A tail at the 3' terminus of pCHPB15, did not represent the actual poly A tail of the 3.5 Kb mRNA.

The presence of two poly A tracts within the 3' noncoding region of the 3.5 Kb chicken mRNA may explain the inability to generate cDNA clones corresponding to the extreme 3' end. During the synthesis of the single stranded cDNA, efficient annealing of the oligo $(dT)_{12}$ primer to these internal poly A tracts, would impede the synthesis of cDNA primed from the true poly A tail and hence none of the cDNA clones would contain this 3' mRNA region. The presence of poly A tracts at the extreme 3' ends of pCHPB15 and pCHPB20 would support this explanation.

It is possible that the two poly A tracts observed in the 3' noncoding region of the cDNA clones pCHP3 and pCHPB15 may be a consequence of retroviral insertions into the chicken gene, resulting in the incorporation of heterologous mRNA sequences and their associated poly A tails. Examination of pCHP3 and pCHPB15 indicated that the first poly A stretch could have contained up to 36 bases (Fig. 4.3-B, nucleotides 2391-2426), some of which have since mutated. Poly A tracts have also been observed in the 3' noncoding regions of the mRNAs encoding the 3MC inducible isozymes, human P450-4 (Quattrochi et al., 1986) and mouse P_3450 (Kimura et al., 1984a). Their significance at present is unclear.

A striking feature of the 3' noncoding regions of rat P450-PB1 (Gonzalez et al., 1986a) and human P450-4 (Quattrochi et al., 1986) is the presence of repetitive R.dre.1 and Alu sequences respectively. The role of these sequences is not known. Computer aided analysis of the 3' noncoding sequences of pCHP3 and pCHPB15 did not detect the presence of such repetitive sequences.

Since the 840 bp cDNA insert of pCHPB15 was the largest DNA fragment complementary to the 3' noncoding region of the 3.5 Kb chicken mRNA, it was expected that none of the other cDNA clones isolated from the library described here would extend the nucleotide sequence further in the 3' direction. Therefore, rather than sequence additional cDNA clones, it was decided to use the characterized cDNA clone pCHP3 to probe a chicken genomic library to isolate the corresponding P450 gene. In this manner, the complete primary sequence of the 3.5 Kb mRNA could be determined from the gene.

CHAPTER 5

CHARACTERIZATION OF THE CHICKEN PB INDUCIBLE P450 GENE FAMILY.

5.1. INTRODUCTION.

In mammals, the PB inducible P450 family is complex, with many highly homologous genes. Four isozymes of rabbit PB inducible P450 have been characterized, either from sequencing of the isolated protein (Heinemann and Ozols, 1983) or from cDNA clones (Leighton et al., 1984). Similarly, several isozymes have been isolated from rat liver, each of which displays characteristic catalytic activities (Appendix 1). Southern hybridization analysis of rat genomic DNA with cDNA clones specific for PB inducible rat P450s reveals a complex pattern of bands, suggesting the presence of at least six genes homologous to PB inducible P450b in the rat genome (Mizukami et al., 1983a). Similar analysis of human genomic DNA by Southern hybridization indicates that the human PB inducible P450 family may contain up to four highly homologous genes (Phillips et al., 1985).

In view of the multiplicity of homologous P450 genes in mammals it was considered necessary to first determine the number of genes present in the chicken genome before work on the control of gene expression could be initiated. Therefore, this chapter describes the isolation and characterization of λ phage genomic clones representing the family of chicken PB inducible P450s. Both quantitative and qualitative techniques strongly suggest that this family contains only two closely related genes.

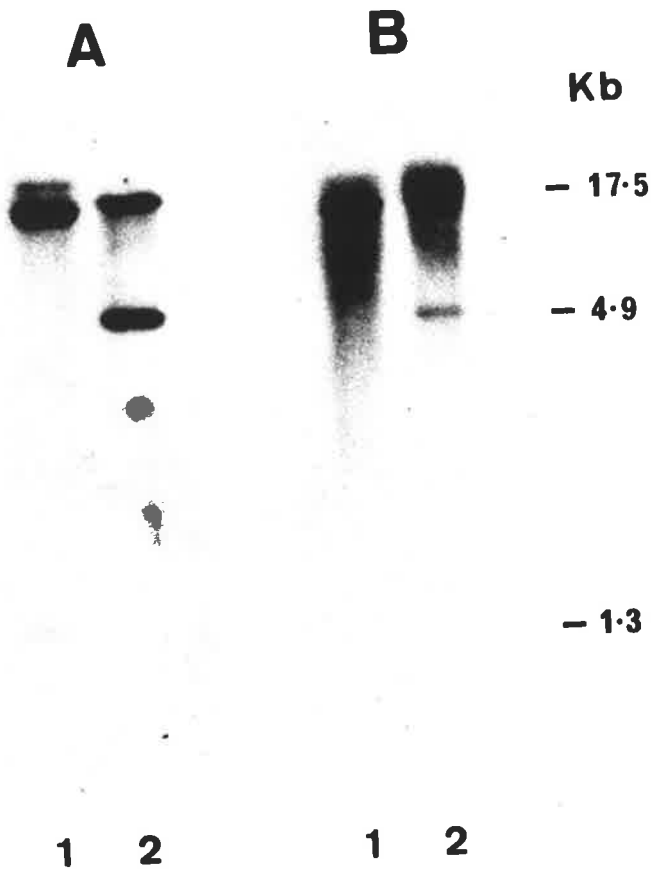
5.2. RESULTS.

5.2.A. Analysis of Total Genomic Chicken DNA.

Prior to isolation of the genomic clones, Southern hybridization analysis was performed on total genomic DNA in order to qualitatively determine the number of PB inducible P450 genes in the chicken genome. Hepatic chicken genomic DNA was digested with EcoRI or BamHI,

Fig. 5.1. Hybridization analysis of total chicken DNA.

Chicken DNA (10ug) was digested with EcoRI (lanes 1) or BamHl (lanes 2) and electrophoresed on a 0.8% agarose gel. After transfer to nitrocellulose, the filters were probed with nick translated pCHP3 at 68°C (panel A) or 52°C (panel B) as described in Section 2.2.K.


A
B
Kb
17·5
4·9
1·3
1
2
1
2

electrophoresed on an agarose gel and transferred to nitrocellulose. The filters were probed with nick translated pCHP3 under conditions of high (Fig. 1-A) or low (Fig. 1-B) stringency. DNA digested with EcoRI (lane 1), generated a strongly hybridizing 15 Kb band and two weakly hybridizing bands at 20 Kb and 1.3 Kb when hybridized at high stringency ($68^{o}C$). Following digestion of the DNA with BamHI (lane 2), two bands at 17 Kb and 4.9 Kb were visible. Hybridization at low stringency ($52^{o}C$) resulted in an increased background signal but did not reveal any additional bands. Since the hybridizations at both temperatures were carried out for the same duration, the lower hybridization rate at $52^{o}C$ resulted in a reduced signal, as observed by others (Beltz et al., 1983). Significantly, the failure to detect additional bands under low stringency conditions, indicates that sequences with significant partial homology to pCHP3 do not exist in the chicken genome. This indicates that the chicken PB inducible P450 gene family consists of only a few members. Quantitation of the gene copy number of P450 genes homologous to pCHP3 is described elsewhere in this chapter.

5.2.B. Isolation and Characterization of Genomic Clones.

P450 genomic clones were isolated from a λ Charon 4A/chicken genomic library. The library was constructed by partial HaeIII/AluI digestion of chicken genomic DNA, followed by the ligation of 15-21 Kb DNA fragments to EcoRI digested λ Charon 4A arms (Dodgson et al., 1979). Approximately $1x10^{6}$ pfu, which represented thirteen genome equivalents, were screened with pCHP3 according to the procedure outlined in Section 2.2.M. Twenty nine recombinant phage were detected and plaque purified after four rounds of screening. A representative portion of the screening is shown in Fig. 5.2 and clearly illustrates that positive clones were easily identifiable. The clones were designated λCHP1-29

Fig. 5.2. Detection of recombinants homologous to pCHP3 in a λ Charon 4A genomic library.

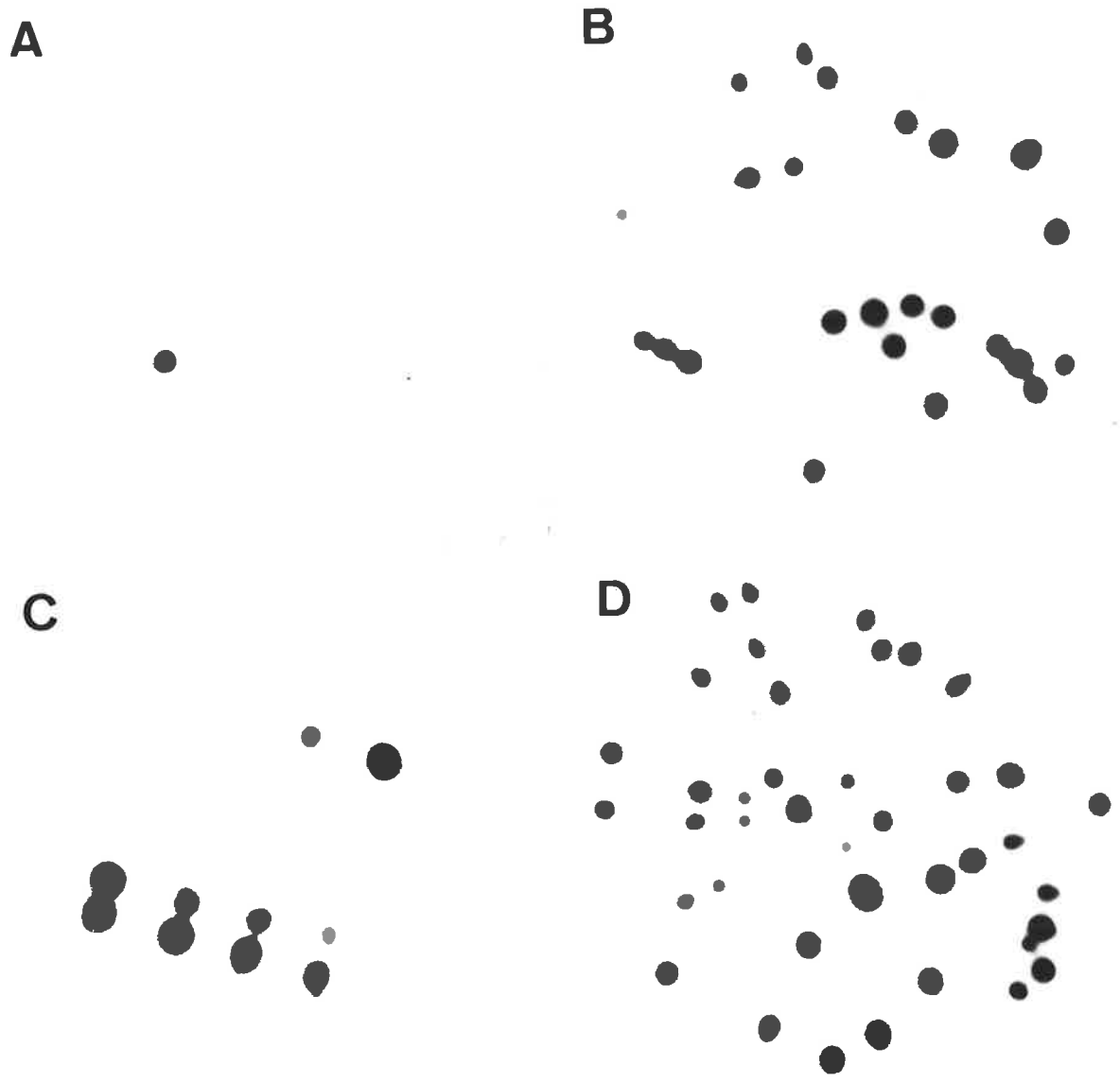
Approximately $1x10^6$ pfu, which represented thirteen genome equivalents, were screened according to the procedure outlined in Section 2.2.M, using nick translated pCHP3 as a probe. A representative portion of four successive rounds of screening (panels A to D) is shown:

First screen **(panel A)**: 30,000 pfu, 1 positive.

Second screen **(panel B)**: 4,000 pfu, 24 positives.

Third screen: **(panel C)**: 30 pfu, 10 positives.

Fourth screen **(panel D)**: 40 pfu, 40 positives.


A
B
C
D

(λ, chicken, P450, isolates 1-29) in order of their isolation. The number of clones obtained from this genomic library supported the previous observation that the chicken PB inducible P450 family contained only a few gene members.

Restriction mapping revealed that these genomic clones were derived from two non-overlapping groups. One of these, group A, comprised nineteen clones (Fig. 5.3-A) which contained a characteristic 3.8 Kb BamHI/EcoRI fragment, and 1.3 Kb and 9.4 Kb EcoRI fragments, each of which hybridized to pCHP3. Southern hybridization analysis on total genomic DNA with pCHP3 did not reveal a 9.4 Kb EcoRI fragment (Fig. 5.1). Therefore, it was concluded that this fragment was generated as a consequence of the introduction of an artificial EcoRI site at the ends of λCHP23, λCHP22, λCHP1828, λCHP15, λCHP18 and λCHP27 (Fig. 5.3), during the construction of the library. The second group, group B (Fig. 5.3-B), contained a characteristic 1.3 Kb EcoRI/BamHI fragment and 3.4 Kb and 8.4 Kb BamHI/EcoRI restriction fragments, each of which hybridized to pCHP3. Since the EcoRI sites flanking λCHP20, λCHP24 and λCHP25 were also generated during cloning, the EcoRI 8.4 Kb fragment present in these clones, represented a fragment of intermediate size. Clones λCHP4, λCHP8 and λCHP10 contained an additional BamHI site absent in the other group A clones which was considered to represent a polymorphism. The observation that the 29 genomic recombinants could be categorized into two groups suggested that the PB inducible P450 gene family contains only two members.

5.2.C. Restriction Mapping of the Genomic Clones.

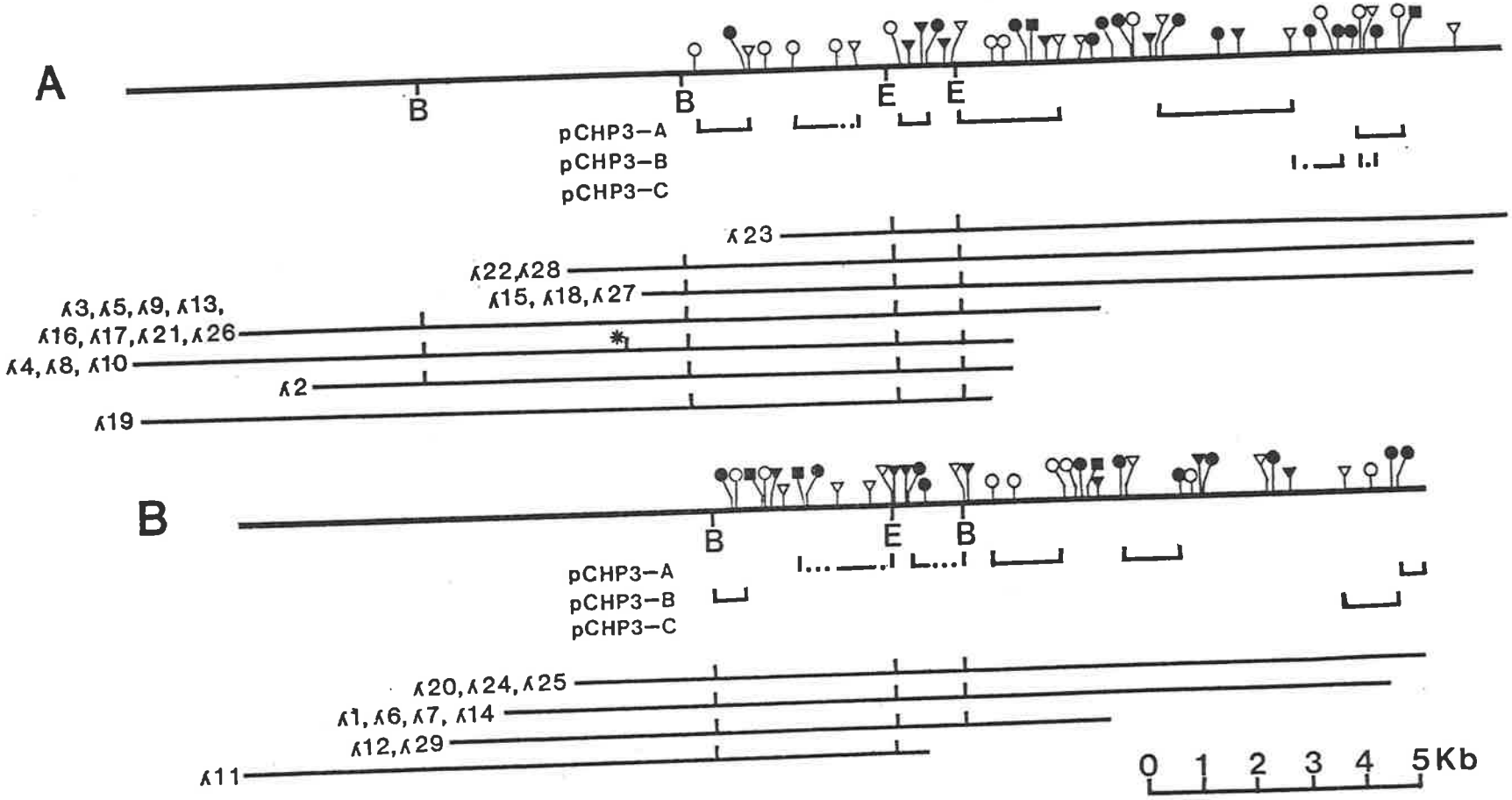
To investigate whether these two groups represented the non-overlapping ends of the same gene, the 3 largest PstI fragments of pCHP3 (see Chapter 3, Fig. 3.5) were subcloned into pBR322. pCHP3-A,

Fig. 5.3. Restriction maps of group A and group B genes.

Two overlapping groups of genomic clones denoted group A (Panel A) and group B (Panel B), were distinguished after digestion with EcoRI and BamHI and Southern analysis using pCHP3 as probe. Restriction fragments from the genomic clones specified in the text were subcloned into pBR322 and mapped in more detail using the indicated restriction enzymes. Southern hybridization analysis of the restriction digests was performed using the cDNA subclones pCHP3-A, pCHP3-B and pCHP3-C. Solid and broken lines indicate strongly and weakly hybridizing bands respectively. The restriction site marked with an asterisk is a BamHI site found only in the indicated clones. We suspect this represents an allelic variation.

B = BamHI	○ = PstI	▽ = HindIII
E = EcoRI	● = PvuII	▼ = AvaI
	■ = SacI	

It has been shown that the 2.2 Kb P450 mRNA is derived from the chicken B gene (Fig. 7.2). Recent work using specific 5' and 3' specific probes for this mRNA, has established that the orientation of gene B is reversed to that shown opposite.

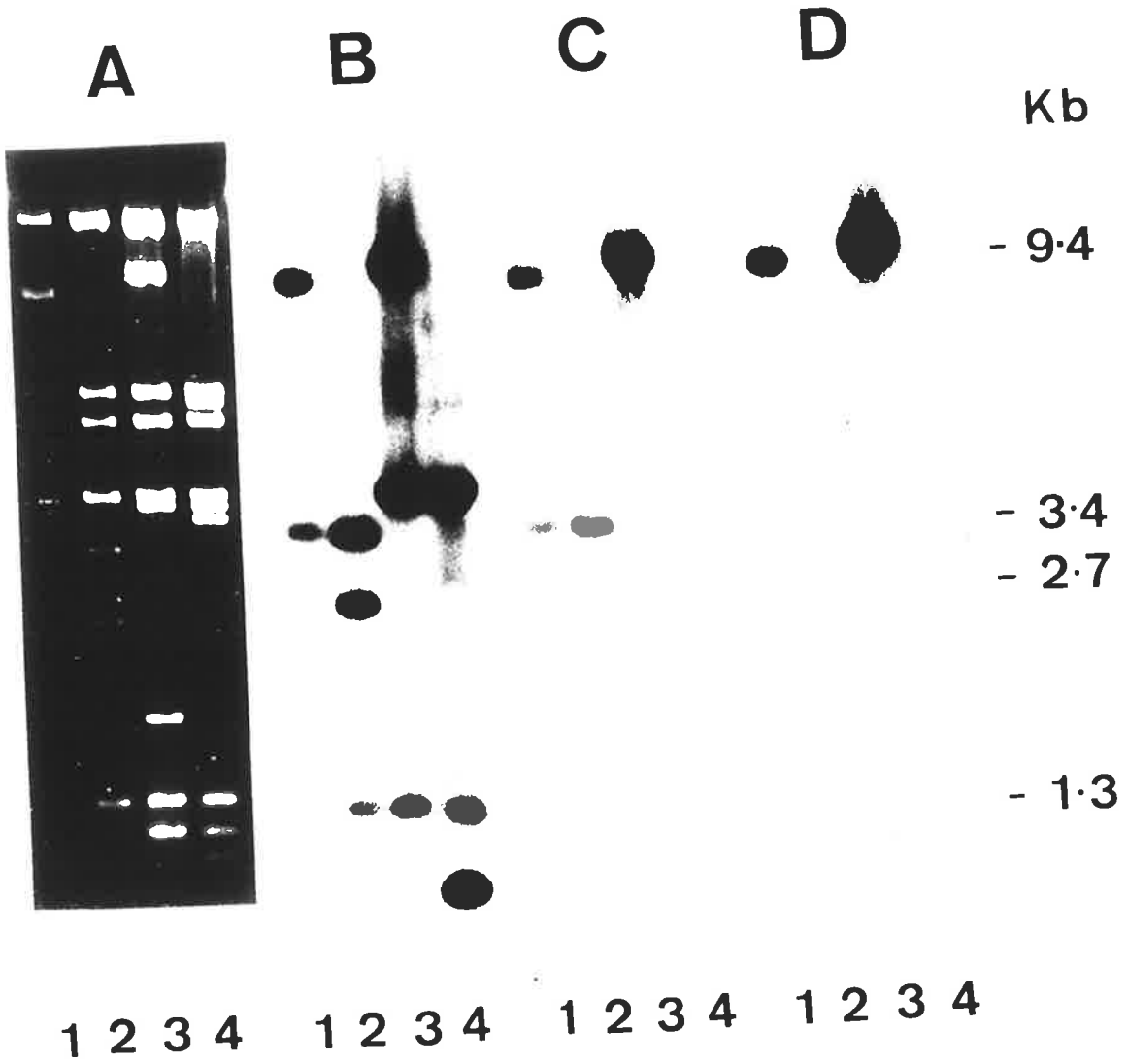

A
B
B
E
E
pCHP3–A
pCHP3–B
pCHP3–C
λ23
λ22,λ28
λ15, λ18,λ27
λ3,λ5,λ9, λ13,
λ16, λ17,λ21,λ26
λ4,λ8, λ10
λ2
λ19
B
B
E
B
pCHP3–A
pCHP3–B
pCHP3–C
λ20,λ24,λ25
λ1,λ6,λ7, λ14
λ12,λ29
λ11
0 1 2 3 4 5Kb

contained a 1.35 Kb PstI fragment which represented a small part of the 5' noncoding region as well as most of the coding region of the mRNA. pCHP3-B contained a 700 bp insert covering a small portion of the carboxyl terminal half of the protein and part of the 3' noncoding region. pCHP3-C, contained only 3' noncoding region. These three subclones were used as probes in Southern hybridization analysis of EcoRI/BamHI digests of four genomic clones, which covered the entire length of the cloned genomic DNA from groups A and B (Fig. 5.4). The results show that the subclone pCHP3-A hybridized to the same restriction fragments which hybridized to pCHP3. In group A, pCHP3-A hybridized to the 3.8 Kb BamHI/EcoRI fragment and the 1.3 Kb and 9.4 Kb EcoRI fragments (Fig. 5.4-B, lanes 3 and 4). In group B, pCHP3-A hybridized to the 1.3 Kb EcoRI/BamHI fragment, and the 3.4 Kb and 8.4 Kb BamHI/EcoRI fragments (Fig. 5.4-B, lanes 1 and 2). pCHP3-A also hybridized to a 0.9 Kb EcoRI fragment from the group A clone λCHP10 (Fig. 5.4-B, lane 4) and to a 2.7 Kb EcoRI fragment from the group B clone λCHP12 (Fig. 5.4-B, lane 2). However, these fragments were a consequence of artificial EcoRI sites produced during the cloning procedure. pCHP3-C hybridized to both the 9.4 Kb EcoRI fragment of group A and the 8.4 Kb BamHI/EcoRI fragment of group B (Fig. 5.4-D, lanes 3 and 1 respectively). Since the extreme 3' and 5' PstI fragments of pCHP3 were represented in both group A and B genes, it was concluded that these groups must represent separate genes rather than two non-overlapping ends of a single P450 gene. Subclone pCHP3-B also hybridized to the 9.4 Kb EcoRI and the 8.4 Kb BamHI/EcoRI fragments, as well as the 3.4 Kb BamHI/EcoRI fragment of group B (Fig. 5.4-C). The reason for this later anomaly is unknown and should be resolved by sequence analysis.

To permit a more detailed comparison of the two genes, selected

Fig. 5.4. All three regions of the cDNA clone, pCHP3, hybridize to the chicken P450 group A and B genomic clones.

DNA from clones λCHP20 (lanes 1), λCHP12 (lanes 2), λCHP22 (lanes 3), and λCHP10 (lanes 4) was digested with EcoRI and BamHI and electrophoresed on a 0.8% agarose gel. Panel A shows the ethidium bromide stained gel. After transfer to nitrocellulose, the filters were probed with nick translated pCHP3-A (panel B), pCHP3-B (panel C), or pCHP3-C (panel D).


A
B
C
D
Kb
- 9·4
- 3·4
- 2·7
- 1·3
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4

restriction fragments were subcloned into pBR322 and detailed restriction analysis performed. The 3.8 Kb BamHI/EcoRI fragment and the 1.3 Kb EcoRI fragments of λCHP3 and the 9.4 Kb EcoRI fragment of λCHP18 were isolated as representatives of group A and subcloned into pBR322. Similarly, the 8.4 Kb BamHI/EcoRI fragment of λCHP20 and the 3.4 Kb, 1.3 Kb and 2.7 Kb BamHI/EcoRI fragments of λCHP12 were subcloned into pBR322 to represent group B. Each subclone was subjected to extensive restriction analysis and the restriction fragments were probed with the three subclones of pCHP3 (Fig. 5.3). The fine maps indicated that both genes were at least 12 Kb in length. pCHP3-A hybridized to restriction fragments which covered 10 Kb and 7 Kb of genes A and B respectively while pCHP3-B and pCHP3-C hybridized to a small region close to the ends of the gene maps. The maps also indicated that the coding regions of both genes were interrupted, indicative of introns and that the distribution of exons appeared different for both genes.

Previous sequence analysis of pCHP3 revealed a 5' to 3' order for the PstI fragments of pCHP3-A, pCHP3-B and pCHP3-C respectively. However, in both the group A and B genes, pCHP3-C hybridized to restriction fragments located between regions which hybridized strongly to pCHP3-A and pCHP3-B. The identity of the subclones was confirmed by sequence analysis (Chapter 3) and the restriction maps were rigorously checked. The reason for this anomaly is presently unclear. Furthermore, the previously observed anomaly with pCHP3-B was limited to a short segment near the 5' end of the group B gene.

5.2.D. Comparison of Restriction Digests of Total Genomic DNA and the Genomic Clones.

In order to determine whether any additional genes in the chicken genome homologous to pCHP3 had escaped isolation from the genomic

library, Southern hybridization analysis of the genomic clones and total genomic DNA was performed. Total chicken genomic DNA along with DNA from four clones representing group A (λCHP18 and λCHP3) or group B (λCHP12 and λCHP20) were digested with EcoRI/BamHI (Fig. 5.5-A), BamHI/SacI (Fig. 5.5-B) and HindIII (Fig. 5.5-C), electrophoresed on an agarose gel, transferred to nitrocellulose and probed with pCHP3. Digestion of chicken genomic DNA with EcoRI/BamHI, generated four bands which hybridized to pCHP3 (Fig. 5.5-A, lane 1). The 1.3 Kb band was accounted for by 1.3 Kb fragments in clones from both group A and B genes. The 3.8 Kb and 3.4 Kb fragments were accounted for by fragments in group A (lanes 2 and 4) and group B (lanes 3 and 5) genes respectively. The cloned 9.4 Kb and 8.4 Kb fragments (Panel A; lanes 2 and 3) corresponded to the 15 Kb genomic fragment with the reduced sizes accounted for by the omission of part of the fragments during cloning. The same explanation is offered to account for the presence of the 2.4 Kb and 2.7 Kb bands in lanes 4 and 5 respectively. Digests of the group B clone λCHP20 (lane 3) were contaminated by the group A clone λCHP3, as is indicated by the 2.4 Kb and 3.8 Kb bands (Panel A; lane 3).

The BamHI/SacI (Panel B) and HindIII (Panel C) digests were similarly analysed. In every case, bands in the digested genomic DNA could be accounted for by bands from one or more of the genomic clones. Some of the bands in the digested genomic clones were larger than could be accounted for by the insert alone. Therefore, these fragments must be situated at the ends of the respective insert and thus contained parts of the λ arms. These results provide very strong evidence that there are only two genes (A and B) in the chicken genome which hybridize to pCHP3.

5.2.E. Southern Hybridization Analysis of Individual Chicken Genomic DNAs.

Although the detailed restriction maps (Fig. 5.3) revealed a large number of differences between the two genes, it was possible that these two genes represented different alleles from the same locus. To investigate this possibility, genomic DNA was isolated from seven individual chicken livers and after digestion with EcoRI and EcoRI/BamHI, was analysed by Southern hybridization (Fig. 5.6). All seven chickens contained the 1.3 Kb EcoRI fragment from group A (Fig. 5.6, A-G, lane 1) and the 3.8 Kb and 3.4 Kb BamHI/EcoRI restriction fragments characteristic of group A and B genes respectively (Fig. 5.6, A-G, lane 2). Assuming that the genes are alleles which occur with equal frequency in the population, the probability of selecting at random, seven heterozygotes for a single gene is 1 in 128. If either allele were present at a lower frequency, then the probability would be even smaller. Therefore, it was considered that the group A and B genes were probably not allelic polymorphisms, but occur at separate loci in the chicken genome.

In some of the EcoRI digests, a 4.8 Kb band was detectable, the presence of which correlated with the disappearance of a 20 Kb band. Therefore this 4.8 Kb band was thought to represent allelic variation.

5.2.F. Quantitative Southern Hybridization Analysis of the A and B Genes.

To determine the copy number of the two PB inducible P450 genes in the chicken genome, quantitative Southern hybridization analysis was performed using the 3.8 Kb and 3.4 Kb BamHI/EcoRI restriction fragments of the genomic clones as characteristic markers for group A and B genes respectively (Fig. 5.3). Genomic DNA isolated from a single chicken

Fig. 5.5. The P450 A and B genomic clones can account for all the restriction fragments in digests of chicken genomic DNA.

The restriction enzymes EcoRI/BamHI (panel A), BamHI/SacI (panel B), and HindIII (panel C) were used to digest 10 ug of total chicken liver DNA (lanes 1) and DNA from the genomic clones λCHP18 (gene A) (lanes 2), λCHP20 (gene B) (lanes 3), λCHP3 (gene A) (lanes 4), and λCHP12 (gene B) (lanes 5). After electrophoresis on a 0.8% agarose gel the DNA was transferred to nitrocellulose and the filters probed with nick translated pCHP3.

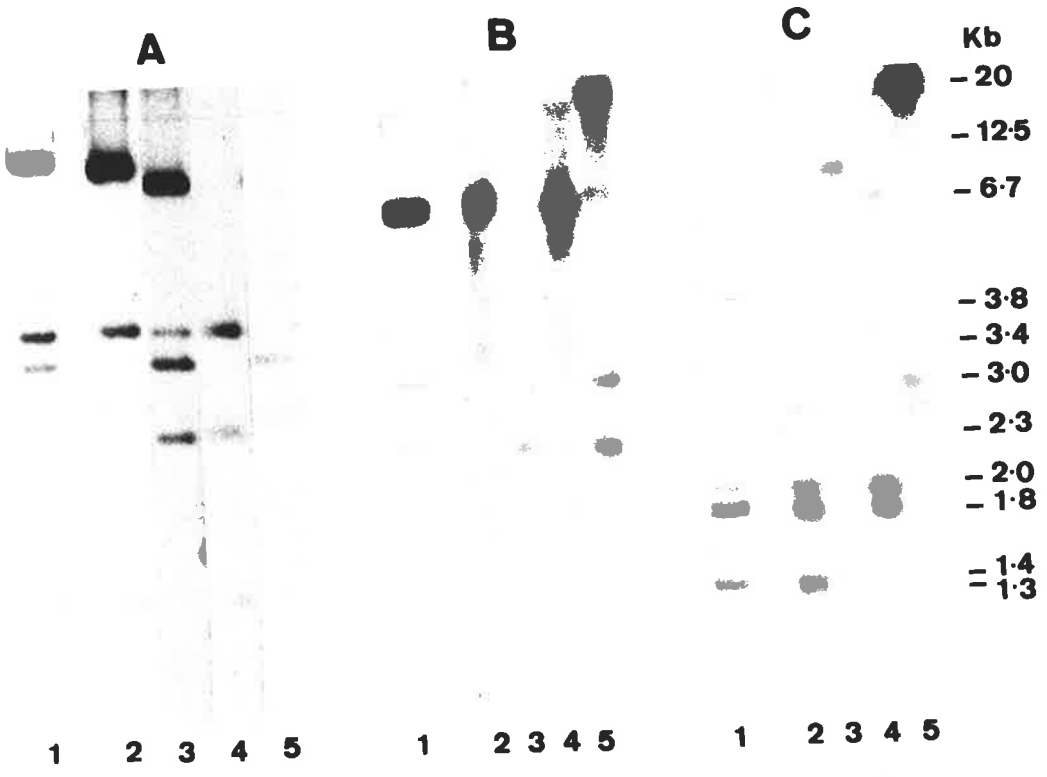

A
B
C
Kb
– 20
– 12·5
– 6·7
– 3·8
– 3·4
– 3·0
– 2·3
– 2·0
– 1·8
– 1·4
– 1·3
1 2 3 4 5
1 2 3 4 5
1 2 3 4 5

Fig. 5.6. Both A and B genes are detected in all individual chicken genomes by Southern hybridization analysis.

DNA (10 ug) from seven individual chicken livers (A-G) was digested with EcoRI (lanes 1) or EcoRI/BamHI (lanes 2) and electrophoresed on a 0.8% agarose gel. After transfer to nitrocellulose, the filters were probed with nick translated pCHP3 (Section 2.2.K.).

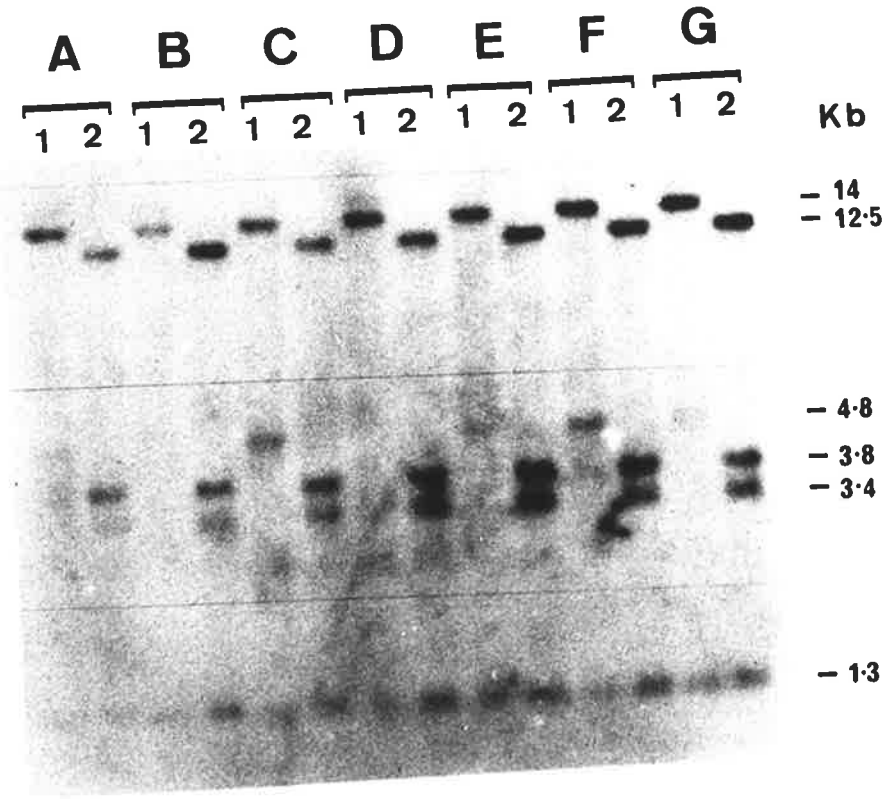
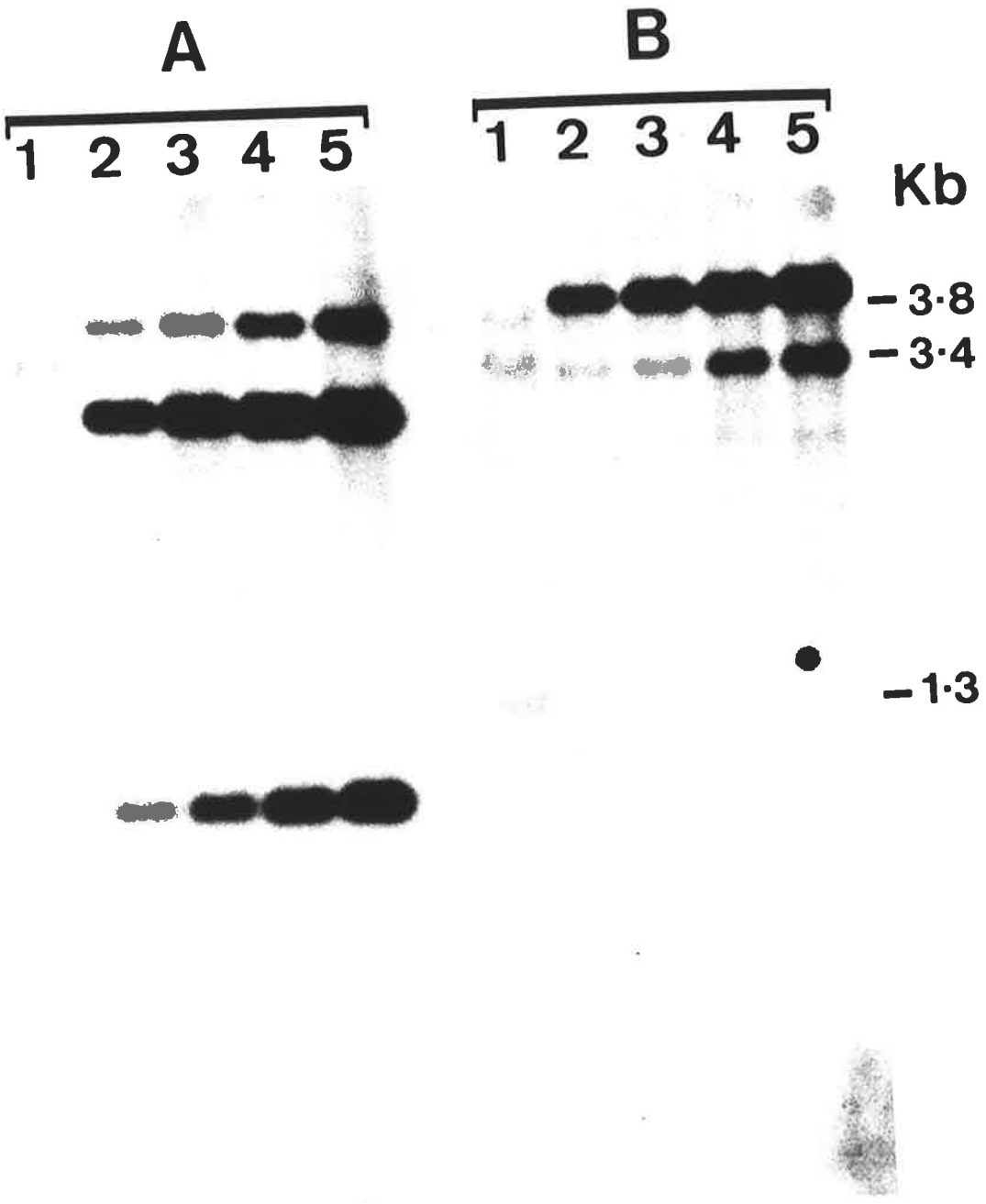

A B C D E F G
1 2 1 2 1 2 1 2 1 2 1 2 1 2
Kb
14
12·5
4·8
3·8
3·4
1·3

Fig. 5.7. Quantitative Southern hybridization analysis of the A and B genes.

Total chicken DNA and subclones containing the 3.8 Kb BamHI/EcoRI fragment from gene A (panel A) and the 3.4 Kb BamHI/EcoRI fragment from gene B (panel B) were digested with EcoRI/BamHI, electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose, and probed with nick translated pCHP3. The lanes contained 10 ug of total chicken liver DNA (lanes 1) or 0.5 (lanes 2), 1.0 (lanes 3), 2.0 (lanes 4), or 4.0 (lanes 5) haploid genome equivalents of the subclone based on a haploid genome size of 2.0×10^9 base pairs (Rosen et al., 1973).


A
B
1 2 3 4 5
1 2 3 4 5
Kb
– 3·8
– 3·4
– 1·3

liver, was digested with EcoRI and BamHI and fractionated on an agarose gel (Fig. 5.7-A and -B, lane 1). The 3.8 Kb and 3.4 Kb subclones described in Section 5.2.C. were digested with the same enzymes and amounts corresponding to 0.5, 1.0, 2.0 and 4.0 haploid genome equivalents were coelectrophoresed in adjacent lanes (Fig. 5.7-A and -B, lanes 2-5 respectively). After transfer to nitrocellulose, the filter was hybridized with nick translated pCHP3. Direct comparison of the hybridization signals obtained from the genomic DNA with the relative copy number of each subclone indicated that there was no more than one copy of either the 3.8 Kb or 3.4 Kb restriction fragments per haploid genome. Although the hybridization signal of the 3.8 Kb band in the genomic DNA was significantly less than the signal corresponding to one genome equivalent of plasmid DNA (Fig. 5.7-A, lane 3), hybridization to this band was found to be variable between experiments. However, in every experiment the results consistently indicated no more than one copy. Therefore, it was concluded that there was only one group A and one group B gene per haploid genome.

5.3. DISCUSSION.

The comprehensive study presented in this chapter clearly indicates that two independent PB inducible P450 genes exist in the chicken genome. Southern hybridization analysis of genomic DNA under conditions of low and high stringency, suggested that only a limited number of genes homologous to pCHP3 exist in the chicken genome. When pCHP3 was used to screen a chicken genomic library, twenty nine recombinant phage were detected and isolated. Restriction mapping and Southern hybridization analysis indicated that these clones could be organized into two non-overlapping groups, A and B. Comparison of restriction digests of chicken genomic DNA and cloned DNAs revealed that each of the

genomic bands homologous to pCHP3 could be accounted for by restriction fragments from one or more of the genomic clones.

Fine mapping of the chicken genomic clones indicated that both genes were over 12 Kb in length and that both contained at least five introns. The size estimate and organization of the chicken genes are in keeping with the values of 14 Kb and eight introns found for the rat PB inducible P450e gene (Suwa et al., 1985). The rat P450b gene is much longer with a length of 20 Kb, but the size difference is due almost entirely to the first intron, which is 12.5 Kb in length compared to 3.2 Kb in the rat P450e gene (Mizukami et al., 1983b).

Restriction mapping revealed (Fig. 5.3) a skewed distribution of the chicken genomic clones with respect to both genes A and B. Whereas many of the clones extended up to 10 Kb past the 5' ends of the genes, it is possible that a small section of the 3' end of the genes may not be represented. Since the clones were isolated from a genomic library which had been amplified at least twice, it is possible that the absence of sequences representing the 3' ends of these genes, was due to selective losses after amplification. Furthermore, the order of the restriction fragments which hybridized to pCHP3-B and pCHP3-C was reversed to that expected. Therefore, it was also possible that some recombination events had occurred during the isolation of the clones. Sequence analysis of these clones and additional clones isolated from a second genomic library (Chapter 6) will clarify these uncertainties.

Quantitative Southern hybridization analysis of chicken genomic DNA indicated that both genes A and B occur as single copies per haploid genome. The detection of restriction fragments characteristic of the A and B genes in the genomes of seven individual chickens, strongly implies that these genes are nonallelic and exist at separate genetic loci.

The detection of only two genes in the chicken PB inducible family makes it the smallest PB inducible P450 family characterized to date. In contrast to the situation in chickens, the mammalian PB inducible family is complex and contains many highly homologous members. For example, in the rat, 6-10 highly homologous PB inducible P450 genes have been detected using Southern hybridization analysis (Simmons et al., 1985), quantitative hybridization analysis (Mizukami et al., 1983a) and characterization of isolated genomic clones (Atchison and Adesnik, 1983). However, many of these genes are due to allelic polymorphism within the rat population (Vlasuk et al., 1982). Analysis of the phenotypes from F1 progenies of inter-strain crosses indicates that at least six isozymes are encoded by two closely linked genetic loci, with at least 4 alleles at the P450b locus and 2 alleles at the P450e locus (Rampersaud and Walz Jr., 1983).

It should be noted that the classification of the chicken A and B genes as PB inducible has been determined by homology with a cDNA clone derived from a drug inducible mRNA. Further details on the inducibility of these genes by the drugs PB and AIA, will be discussed later in Chapter 7.

CHAPTER 6

CHARACTERIZATION OF THE GENOMIC CLONES REPRESENTING THE 3' END OF THE CHICKEN P450 A GENE.

6.1. INTRODUCTION.

The previous chapter described the isolation and characterization of λ phage genomic clones representing two chicken PB inducible genes (A and B). Extensive Southern hybridization analysis of the genomic clones with subclones of pCHP3 indicated that: (1) a small section at the 3' end of the genes may not be represented and (2) the order of the restriction fragments which hybridized to pCHP3-B and pCHP3-C was reversed to that expected. During the preparation of this library, the chicken genomic DNA was partially digested with HaeIII and AluI and size fractionated prior to cloning into λ Charon 4A. It is possible that the 3' ends of the chicken A and B genes are rich in restriction sites for HaeIII and/or AluI. Consequently, DNA fragments corresponding to the 3' ends of the genes would not be represented in this library. Alternatively, since the genomic library was amplified at least twice, it is possible that the absence of sequences representing the extreme 3' ends of these genes was due to selective losses during amplification.

Because of the apparent absence of sequences corresponding to the extreme 3' ends of the A and B genes it was necessary to screen a second library in an attempt to obtain genomic clones covering these regions. This chapter describes the isolation and characterization of recombinant phage containing sequences complementary to the 3' ends of the chicken P450 A gene.

6.2. RESULTS.

6.2.A. Isolation of Chicken P450 Genomic Clones.

Previous Southern hybridization analysis of phage genomic clones corresponding to the chicken P450 A and B genes, indicated that the 3' PstI fragment of pCHP3, fragment C, hybridized strongly to the 9.4 Kb EcoRI and 8.4 Kb BamHI/EcoRI fragments situated at the extreme ends of

group A and B clones respectively (refer to Fig. 5.4-D, lanes 3 and 1). However, the 840 bp cDNA insert of the previously characterized chicken P450 clone, pCHPB15, was shown by sequence analysis to extend the sequence of pCHP3 by 134 nucleotides in the 3' direction (refer to Chapter 4). Therefore, to selectively isolate genomic clones which contained sequences complementary to the 3' ends of the chicken P450 genes, it was considered appropriate to use pCHPB15 as a probe to screen the second genomic library.

The second genomic library was kindly donated by Dr. P. Morris (this department) and was constructed by partial Sau3A digestion of chicken genomic DNA. DNA fragments ranging in size from 16-24 Kb, were ligated into SalI digested EMBL 3 phage arms (Frischauf et al., 1983) and amplified only once. $1x10^6$ pfu from this library, representing thirteen chicken genome equivalents (Rosen et al., 1973), were screened with nick translated pCHPB15, as described previously in Section 2.2.M. Thirty four recombinant phage were detected and plaque purified following four rounds of screening. A representative portion of the library screening is shown in Fig. 6.1 and clearly demonstrates that genomic clones containing P450 sequences were easily identified. The clones were designated λCHPB1-34 (λ, chicken, P450, library B, isolates 1-34), in order of their isolation.

6.2.B. Southern Hybridization Analysis of the P450 Genomic Clones.

To determine the relationship between these genomic clones and the chicken PB inducible P450 A and B genes, restriction mapping was performed. As illustrated in Fig. 6.2-B, each of genomic clones contained a 1.3 Kb EcoRI fragment. Since restriction analysis of genomic clones isolated from the first library indicated that the 1.3 Kb EcoRI fragment was characteristic of the chicken P450 A gene, it was

Fig. 6.1. Detection of recombinants homologous to pCHPB15 in an EMBL 3 genomic library.

Approximately $1x10^6$ pfu, representing thirteen chicken genome equivalents were screened according to the method outlined in Section 2.2.M, using nick translated pCHPB15 as a probe. A representative portion of four successive rounds of screening (panel A to D) is presented:

First screen **(panel A)**: 30,000 pfu, 1 positive.

Second screen **(panel B)**: 5,000 pfu, 24 positives.

Third screen **(panel C)**: 30 pfu, 4 positives.

Fourth screen **(panel D)**: 47 pfu, 47 positives.

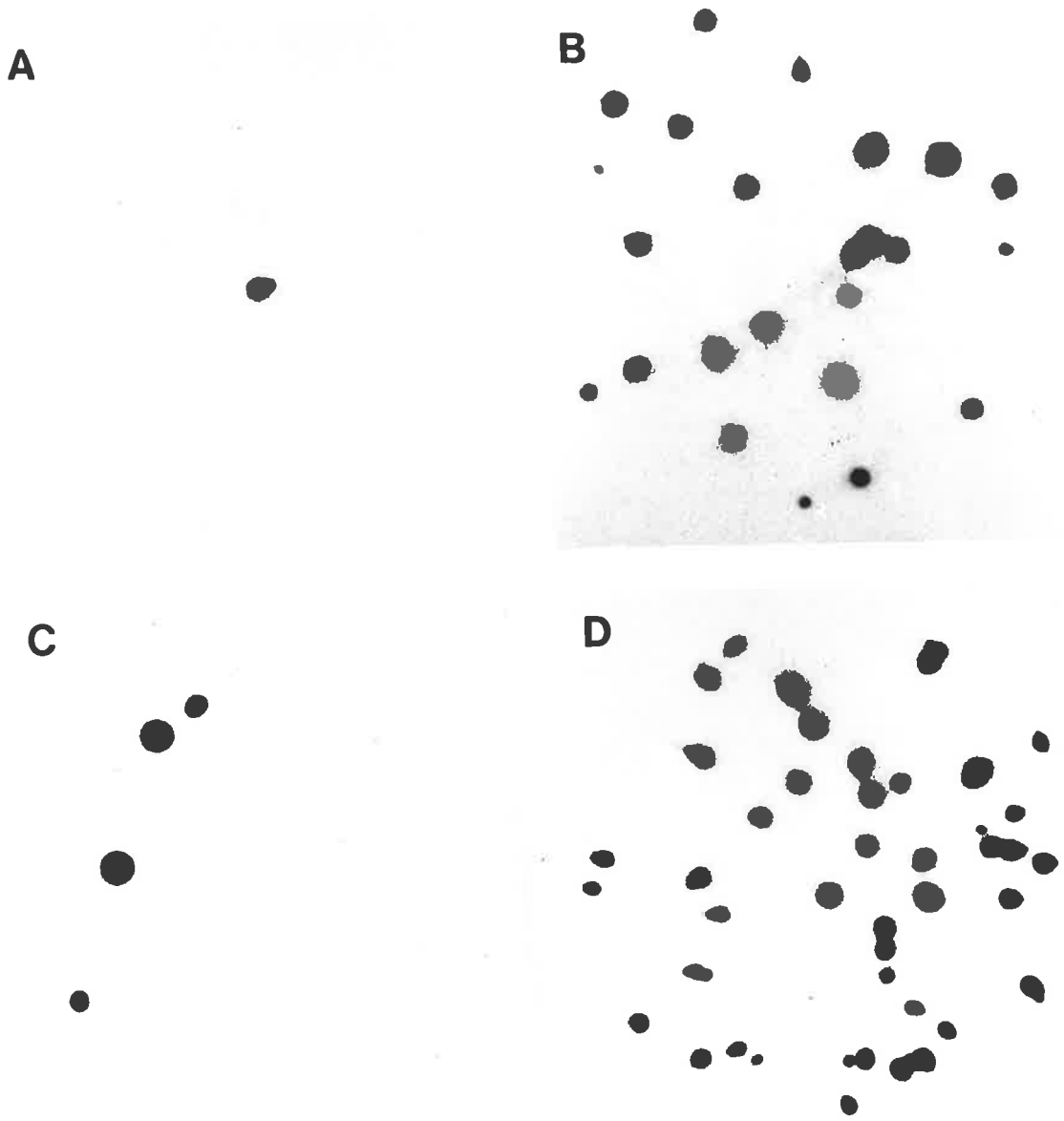
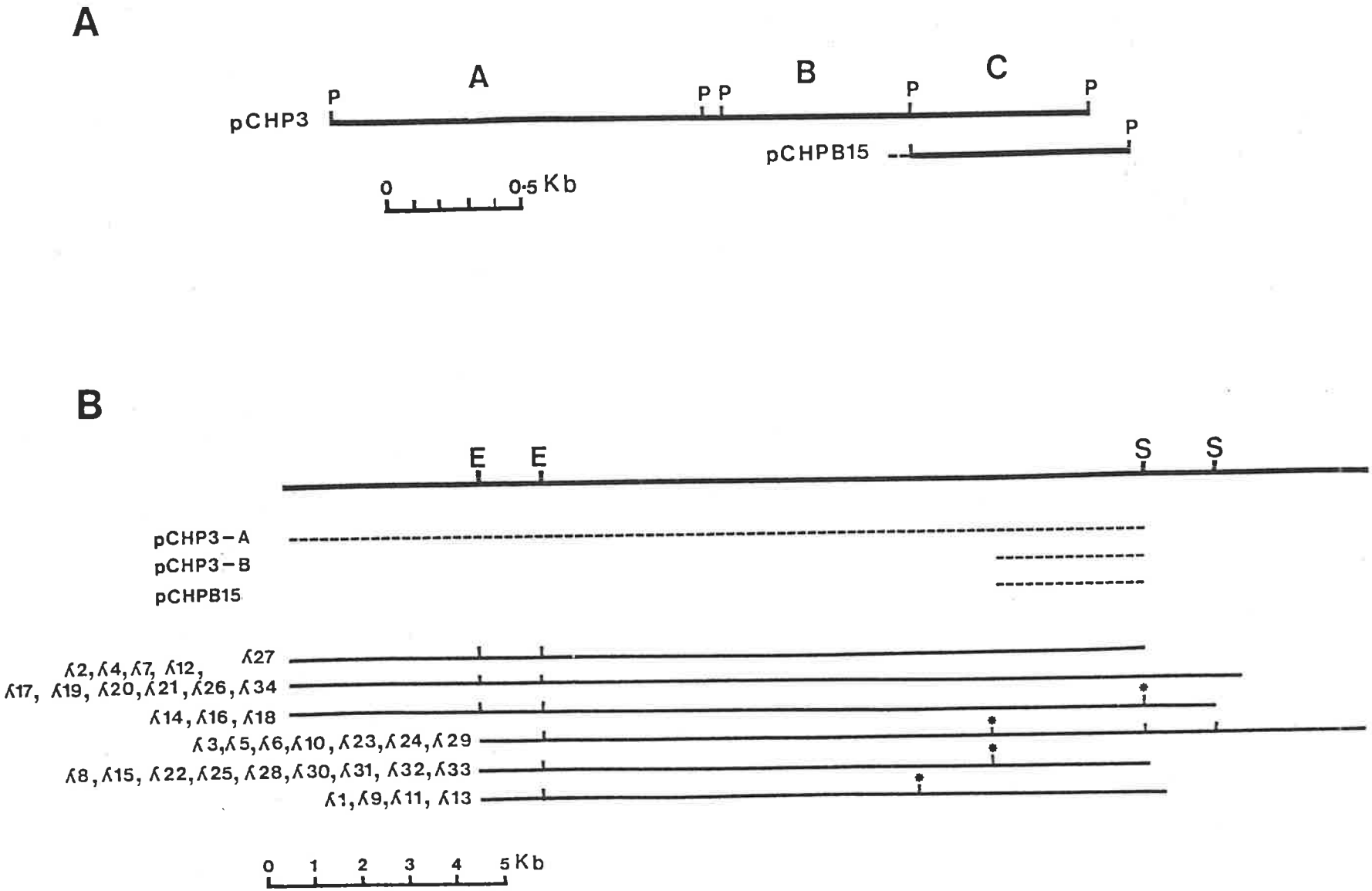

A
B
C
D

Fig. 6.2. Restriction maps of P450 genomic clones isolated from the EMBL 3 library.

The relationship of the PstI restriction maps of pCHP3 and pCHPB15 is shown in panel A. The three largest PstI fragments of pCHP3 are labelled A, B and C.

The EcoRI/SalI restriction maps of the thirty four genomic clones are presented in panel B. Southern hybridization analysis of the restricted DNA was performed using nick translated pCHPB15 and the cDNA subclones pCHP3-A and pCHP3-B. The broken lines indicate the relative hybridization patterns of these three probes to the genomic clones. The restriction sites marked with an asterisk are EcoRI star (*) sites found in the indicated clones.

P = PstI E = EcoRI S = SalI


A
A
B
C
P
P P
P
P
pCHP3
P
pCHPB15
0
0·5 Kb
B
E
E
S
S
pCHP3–A
pCHP3–B
pCHPB15
λ27
λ2, λ4, λ7, λ12,
λ17, λ19, λ20, λ21, λ26, λ34
λ14, λ16, λ18
λ3, λ5, λ6, λ10, λ23, λ24, λ29
λ8, λ15, λ22, λ25, λ28, λ30, λ31, λ32, λ33
λ1, λ9, λ11, λ13
0 1 2 3 4 5 Kb

concluded that all the chicken genomic clones isolated from this second library corresponded to the A gene. One clonal type (e.g. λCHPB3 and λCHPB29), contained two internal SalI sites not present in other overlapping clones and probably represented a polymorphism. Twenty three of the clones (e.g. λCHPB1, λCHPB3, λCHPB8 and λCHPB14) contained an internal EcoRI site, denoted by an asterisk in Fig. 6.2-B. Several preparations of EcoRI have been shown to contain secondary endonuclease activity or star activity (*), which under certain conditions recognized the DNA sequence $^{5'}$AATT$^{3'}$ (Mayer, 1978). Since the additional EcoRI sites present in these genomic clones were recognized only in the presence of SalI, these sites were designated EcoRI*.

To establish the transcriptional orientation of the genomic clones isolated in this chapter, Southern hybridization analysis was performed using pCHP3-A, pCHP3-B and pCHPB15 as specific probes. The alignment of these probes with the cDNA clone pCHP3 is shown in Panel A of Fig. 6.2 and the results of the hybridization studies are presented schematically in Fig. 6.2-B. pCHP3-A hybridized strongly to a 4.0 Kb SalI/ EcoRI fragment, the 1.3 Kb EcoRI fragment and the large 13 Kb EcoRI/SalI fragment. Hybridization of pCHP3-B and pCHPB15 was localized to one end of the 13 Kb EcoRI/SalI fragment. Since three of the clonal types extended up to 5 Kb beyond pCHPB15 in the 3' direction, it was considered highly probable that the genomic clones contain the entire 3' end of the chicken P450 A gene, not previously represented by the cDNA and genomic clones.

6.3. DISCUSSION.

In an attempt to obtain genomic clones which contained sequences complementary to the extreme 3' end of the chicken P450 A and B genes, a second genomic library was screened using pCHPB15 as a probe. Detailed

restriction mapping indicated that each of the thirty four recombinant phage isolated, contained a 1.3 Kb EcoRI fragment which hybridized strongly to pCHP3-A. These results indicated that the genomic clones characterized here all represented the chicken P450 A gene.

Restriction mapping and Southern hybridization analysis revealed that the chicken P450 A gene was at least 18 Kb in length. Furthermore, the genomic clones extended the previous clones up to 5 Kb in the 3' direction. Therefore, the overlapping clones obtained from the two libraries described here and in Chapter 5, are considered to encompass the entire chicken P450 A gene and its flanking sequences. Although it is preferable to obtain the complete sequence of a gene from one contiguous DNA fragment, the maximum amount of foreign DNA that can be packaged into λ phage heads (approximately 23 Kb) renders it impossible for large genes, such as the chicken P450 A gene discussed here, to be isolated intact (Frischauf et al., 1983). However, the complete DNA sequences of several large eukaryotic genes, including the 15 Kb rat α-casein gene (Lu and Rosen, 1983), the 21 Kb Xenopus laevis vitellogenin A1 gene (Wahli et al., 1981) and the 37 Kb chicken α2(I) collagen gene (Ohkubo et al., 1980), have been successfully determined from overlapping λ clones.

Previous sequence analysis and Northern hybridization studies (Chapters 3 and 4) revealed that the cDNA clone pCHPB15 extended the sequence of pCHP3 in the 3' direction. Since only genomic clones corresponding to gene A were isolated from this genomic library when pCHPB15 was used as a probe, it is probable that the 3.5 Kb mRNA is derived from gene A. As discussed in the next chapter, Southern hybridization analysis using oligonucleotide probes specific for the 3.5 Kb chicken mRNA, confirms this view.

The inability to detect genomic clones for gene B using pCHPB15 as a probe was completely unexpected. Previous mapping results indicated that both genes A and B contained sequences complementary to pCHPB15 (refer to Chapter 5). However, in light of the studies presented in this chapter, it is possible that the previously characterized group B clones have undergone a recombination event either during the construction or screening of the library. Sequence analysis of these clones and additional clones from a chicken cosmid library currently being screened in this laboratory (A. Hansen, personal communication), should clarify the situation. The use of cosmid vectors should overcome the limitations of λ phage vectors, by permitting the cloning of the entire chicken PB inducible P450 genes into single recombinants.

CHAPTER 7

SEQUENCE ANALYSIS OF THE CHICKEN P450 A GENE.

7.1. INTRODUCTION.

Determination of the nucleotide sequences for the 3.5 Kb, and more recently the 2.2 Kb chicken mRNAs, has permitted an investigation into the relationship between these two mRNAs and the genomic clones representing genes A and B. As described in this chapter, by using DNA probes which distinguish between these two mRNAs, both the chicken P450 genes (A and B) were shown to be transcriptionally active.

The nucleotide sequence of the 5' flanking region and the first three exons of the chicken P450 A gene is also presented and compared with the sequences of other eukaryotic genes, including in particular the mammalian PB inducible P450 genes. Such comparative studies are useful as a first indication of consensus sequences in the control regions of genes which may be strong candidates as regulatory elements.

7.2. RESULTS.

7.2.A. Hybridization Analysis of the A and B Genes using DNA Probes which Distinguish between the 3.5 Kb and 2.2 Kb mRNAs.

Previous Northern hybridization analysis of poly A^+ RNA from AIA and DDC induced chick embryo livers, indicated the presence of three homologous P450 mRNA species. These were shown to differ in their respective coding and 3' noncoding regions (Chapter 3). A synthetic oligonucleotide probe complementary to a highly divergent region of pCHP3, was previously shown by Northern hybridization analysis to be specific for the 3.5 Kb chicken mRNA (refer to Chapter 3). This oligonucleotide was used as a probe to examine the relationship between the 3.5 Kb mRNA and the two chicken P450 genes, A and B. Genomic clones λCHP18 and λCHP3, representing the A gene and λCHP20 and λCHP12, representing the B gene, were used in these studies since they have been characterized extensively by fine mapping as reported in Chapter 5. The

DNA from these four clones was digested with HindIII, electrophoresed on a 0.8% agarose gel and the DNA fragments transferred bidirectionally to nitrocellulose, according to the procedure of Smith and Summers (1980). One filter was probed with the oligonucleotide probe specific for the 3.5 Kb mRNA (oligonucleotide A) and the second filter was probed with an oligonucleotide probe complementary to the region encoding the highly conserved putative heme binding site (oligonucleotide B). The position of these oligonucleotide probes with respect to the protein sequence specified by pCHP3, is illustrated in Fig. 3.10-A. The results of the Southern hybridization analysis are shown in Fig. 7.1. At 42^{o}C, oligonucleotide B bound to clones for both the A and B genes (panel A). Raising the temperature of washing to 52^{o}C, made little difference to the intensity of the hybridization signal. In contrast, oligonucleotide A bound strongly to the clones specific for gene A but poorly to the clones representing gene B at 42^{o}C (panel B). Raising the temperature of washing to 52^{o}C had no effect on the binding to the group A clones but completely eliminated the hybridization signal observed with the group B clones. This result clearly indicates that the chicken P450 A gene is transcriptionally active, giving rise to the 3.5 Kb mRNA.

During the preparation of this thesis, studies performed in this laboratory indicated that a 228 bp BamHI/AvaII fragment derived from the 3' noncoding region of the cDNA clone pCHP7, was specific for the 2.2 Kb chicken P450 mRNA (A. Hansen, personal communication). To determine the relationship between the 2.2 Kb mRNA and the chicken A and B genes, Southern hybridization analysis was performed in collaboration with A. Hansen, using the 228 bp BamHI/AvaII fragment as a specific probe. DNA from two overlapping clones representing the A gene (λCHP18 and λCHP3) and two clones representing gene B (λCHP20 and λCHP12), were digested with EcoRI and BamHI and electrophoresed on a 0.8% agarose gel

Fig. 7.1. Hybridization analysis of genomic clones for the A and B genes using oligonucleotide probes.

DNA from clones λCHP18 (lanes 1) and λCHP3 (lanes 2) representing the A gene and λCHP12 (lanes 3) λCHP20 (lanes 4) representing the B gene was digested with HindIII and electrophoresed on a 0.8% agarose gel. After transfer to nitrocellulose the filters were hybridized to [^{32}P]-labelled oligonucleotide B derived from the conserved putative heme-binding site (A) or [^{32}P]-labelled oligonucleotide A derived from the highly divergent region (B) previously described in Fig. 3.10. The filters were then washed at 42^{o}C or 52^{o}C after hybridization. The faint bands above the 1.8 Kb band in lanes 1 are due to partial digestion of this clone.

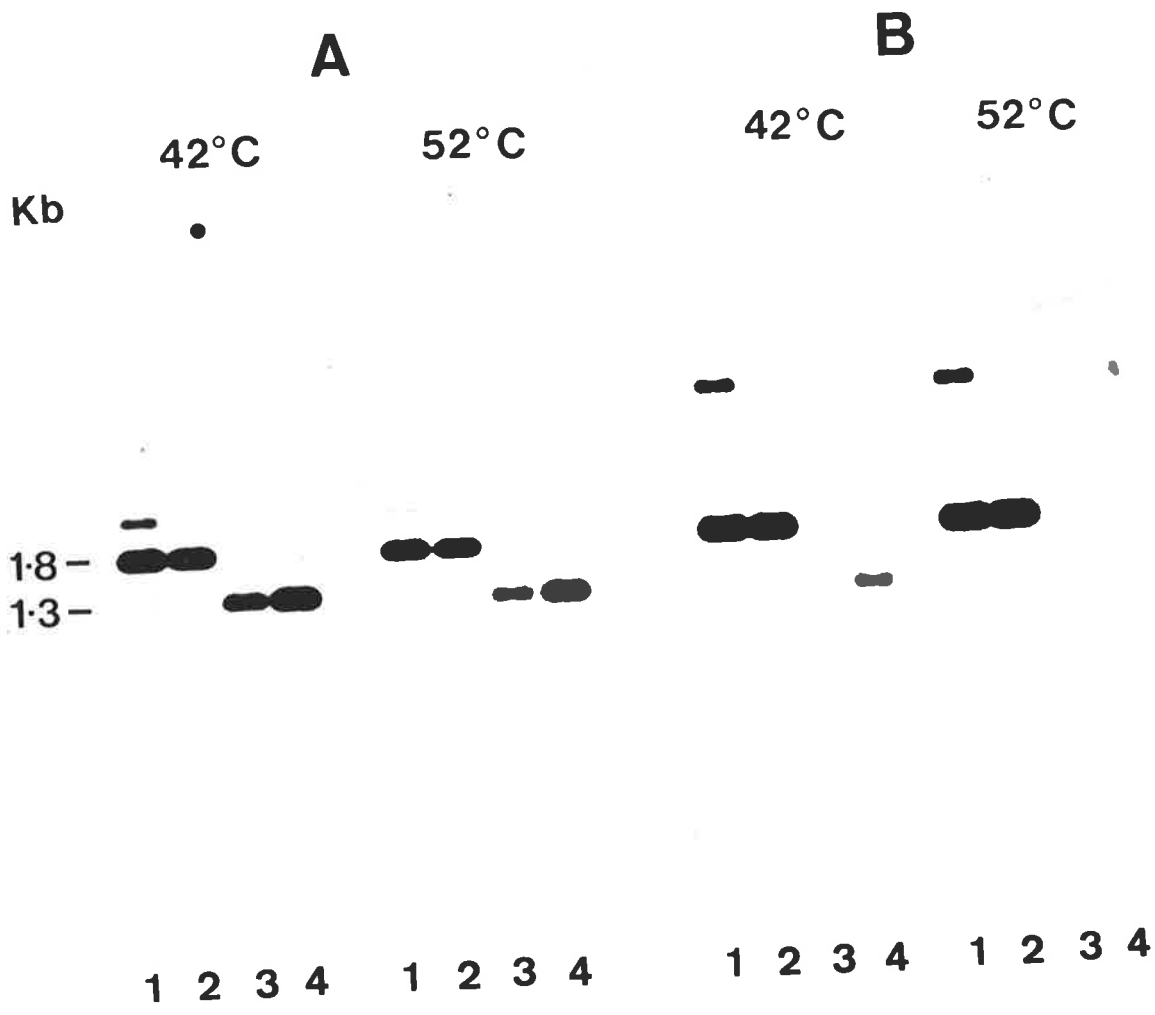
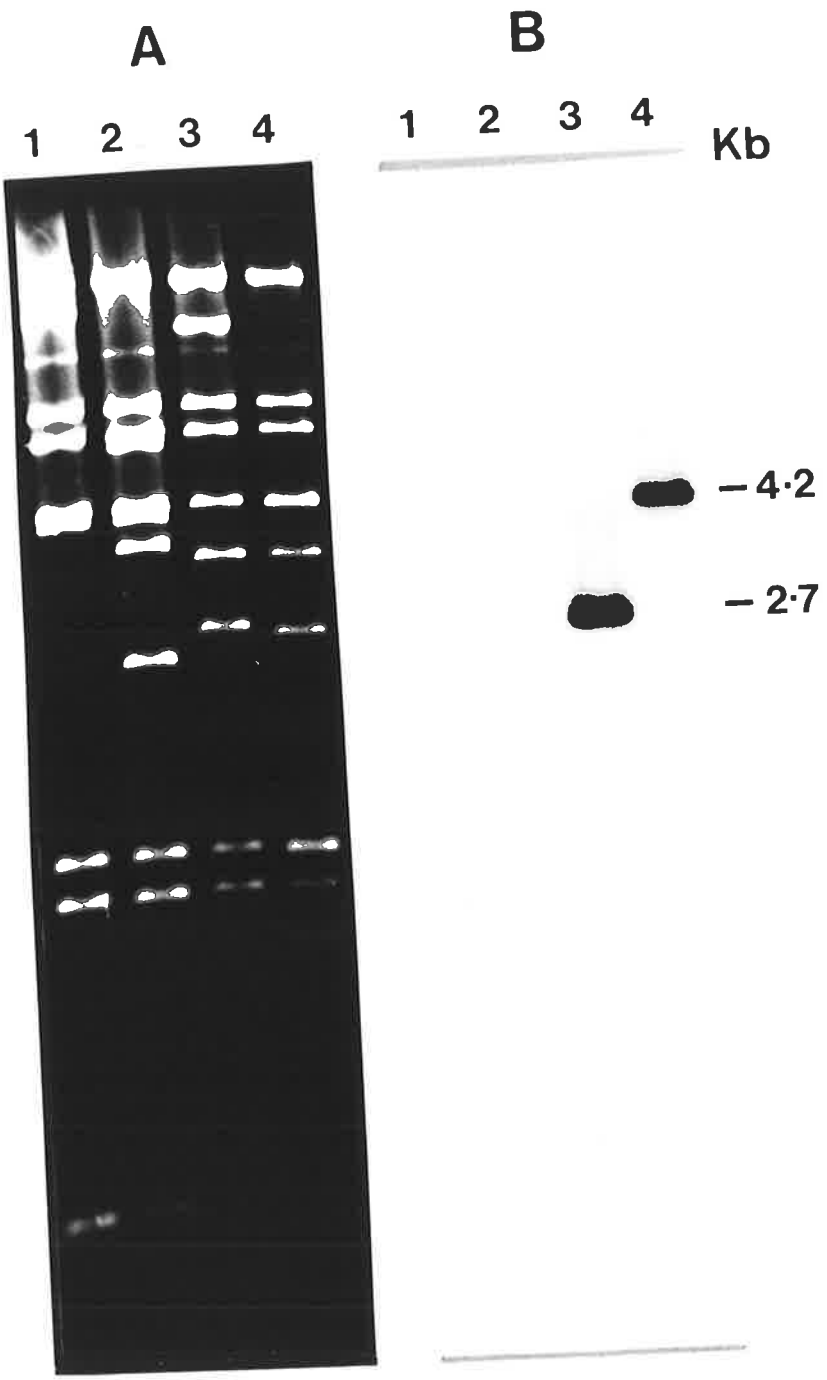

A
B
42°C
52°C
42°C
52°C
Kb
1·8 –
1·3 –
1 2 3 4 1 2 3 4
1 2 3 4 1 2 3 4

Fig. 7.2. Southern hybridization analysis of genomic clones for the A and B genes using a probe specific for the 2.2 Kb mRNA.

DNA from clones λCHP18 (lanes 1), λCHP3 (lanes 2), λCHP20 (lanes 3) and λCHP12 (lanes 4) was digested with EcoRI/BamHI and electrophoresed on a 0.8% agarose gel. The ethidium bromide stained gel is shown in panel A. After transfer to nitrocellulose, the filter was probed with the nick translated 228 bp BamHI/AvaII fragment of pCHP7 (panel B), according to the hybridization and washing conditions described in Section 2.2.K.


A
B
1 2 3 4
1 2 3 4
Kb
– 4·2
– 2·7

(Fig. 7.2-A). Following transfer to nitrocellulose, the filter was probed with the nick translated BamHI/AvaII fragment of pCHP7. As shown in Fig. 7.2-B, the probe bound strongly to the clones specific for gene B (lanes 3 and 4), but did not hybridize to the clones representing gene A. This result proves unequivocally that the 2.2 Kb mRNA is derived from the B gene. Therefore, in the chicken, both the group A and B genes are transcriptionally active.

7.2.B. Partial Sequence Analysis of the Chicken P450 A Gene.

When the work described in this section was performed, it had been established that the 3.5 Kb mRNA was derived from the A gene. However, at the time it was not certain if gene B was transcriptionally active. Therefore, it was considered appropriate to focus further attention on the genomic clones representing gene A. Although two overlapping clones were required to encompass the entire P450 A gene and its flanking sequences, fine mapping studies with subclones of pCHP3 indicated that the 5' end of the gene terminated in the 3.8 Kb BamHI/EcoRI fragment present in clone λCHP3 (refer to Fig. 5.3-A). Therefore, this fragment was chosen for sequence analysis in an effort to deduce the promoter sequence for gene A.

The 3.8 Kb BamHI/EcoRI fragment was cloned directly into M13 mp18 and mp19 phage DNA and a sequential series of overlapping deletion clones generated for use in DNA sequencing (Section 2.2.P-2,-3). The 3.8 Kb BamHI/EcoRI fragment was sequenced in its entirety as outlined schematically in Fig. 7.3, and was expected to contain up to 1 Kb of 5' flanking DNA which could contain sequences important for the control of gene A expression. The complete nucleotide sequence of 3766 bp is presented in Fig. 7.4.

Fig. 7.3. Sequencing strategy for the genomic 3.8 Kb BamHI/EcoRI fragment.

A diagrammatic representation of the structural features of the 3.8 Kb genomic fragment is shown. The positions of the inverted CCAAT and canonical TATA sequences are shown and the first three exons (solid boxes) of the A gene labelled 1, 2 and 3. The extent and direction of sequencing are indicated by arrows.

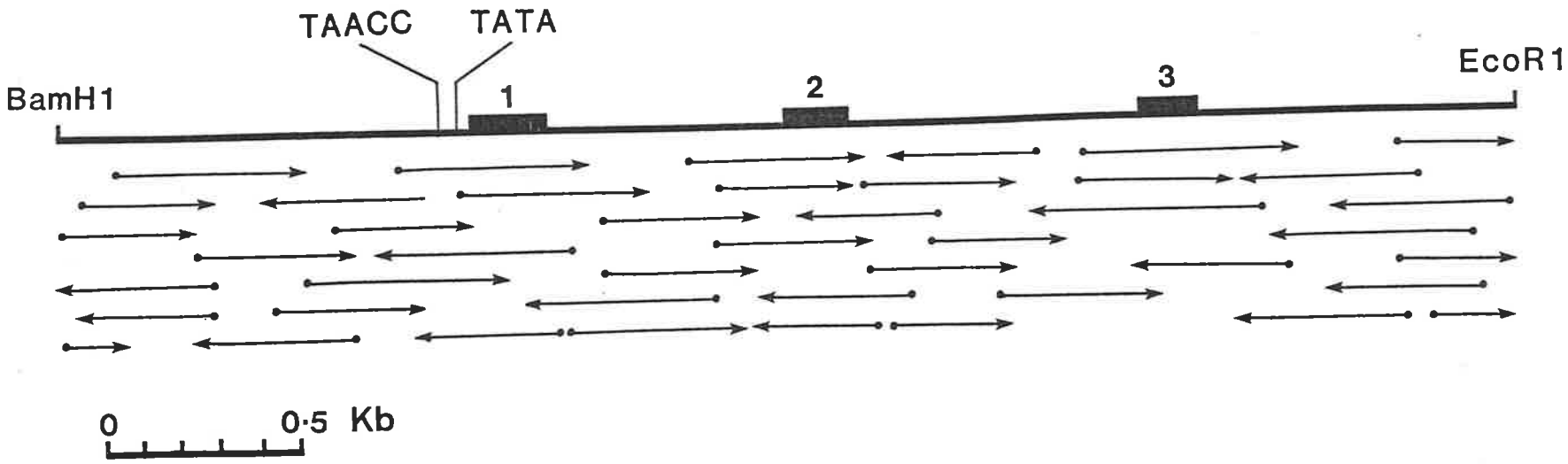
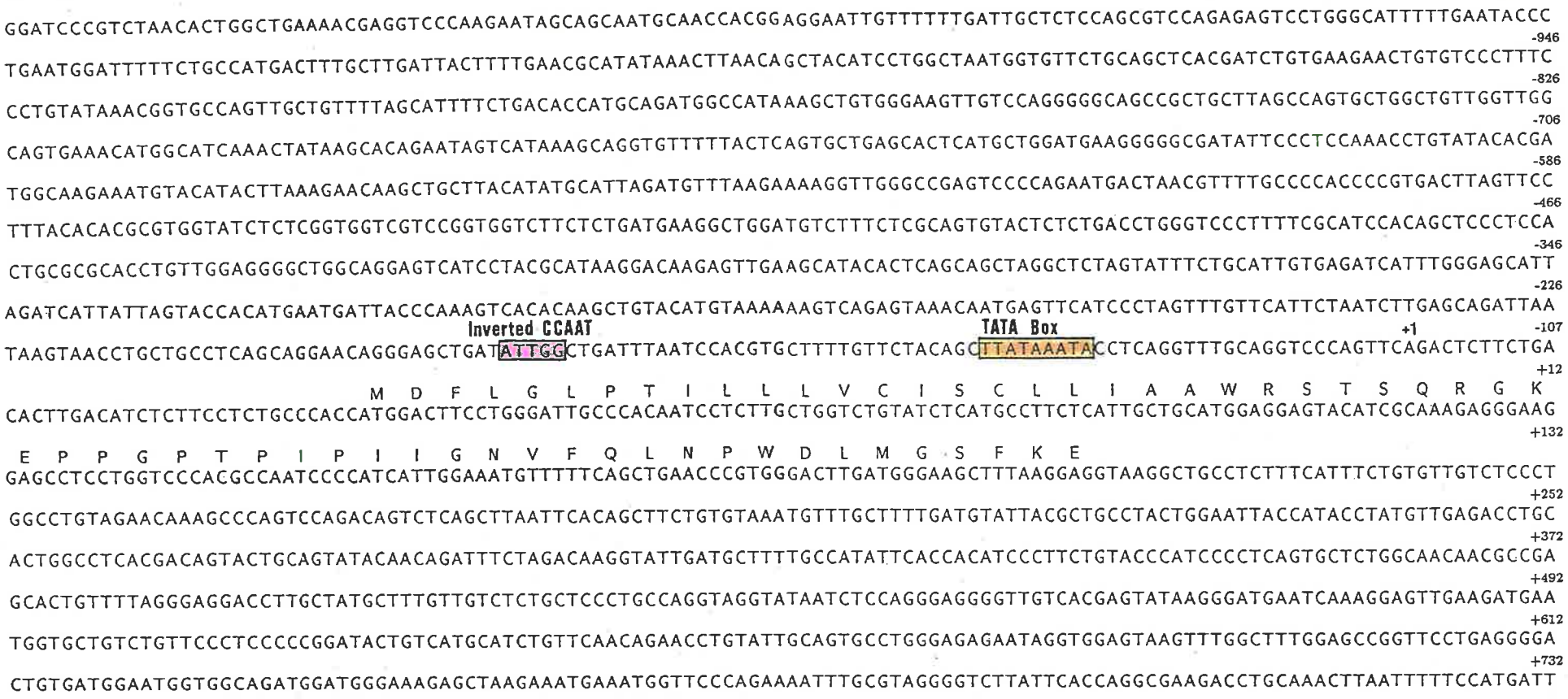

TAACC
TATA
BamH1
1
2
3
EcoR1
0
0·5 Kb

Fig. 7.4. The nucleotide sequence of the P450 A gene promoter region.

The nucleotide sequence of the 3.8 Kb BamHI/EcoRI genomic fragment was determined (Section 2.2.O.). The inverted CCAAT and canonical TATA sequences are highlighted. The sequence is numbered in base pairs (bp) with the mRNA cap site at +1. The amino acid sequences encoded by the first three exons are indicated.


GGATCCCGTCTAACACTGGCTGAAAACGAGGTCCCAAGAATAGCAGCAATGCAACCACGGAGGAATTGTTTTTTGATTGCTCTCCAGCGTCCAGAGAGTCCTGGGCATTTTTGAATACCC
-946
TGAATGGATTTTTCTGCCATGACTTTGCTTGATTACTTTTGAACGCATATAAACTTAACAGCTACATCCTGGCTAATGGTGTTCTGCAGCTCACGATCTGTGAAGAACTGTGTCCCTTTC
-826
CCTGTATAAACGGTGCCAGTTGCTGTTTTAGCATTTTCTGACACCATGCAGATGGCCATAAAGCTGTGGGAAGTTGTCCAGGGGGCAGCCGCTGCTTAGCCAGTGCTGGCTGTTGGTTGG
-706
CAGTGAAACATGGCATCAAACTATAAGCACAGAATAGTCATAAAGCAGGTGTTTTTACTCAGTGCTGAGCACTCATGCTGGATGAAGGGGGCGATATTCCCTCCAAACCTGTATACACGA
-586
TGGCAAGAAATGTACATACTTAAAGAACAAGCTGCTTACATATGCATTAGATGTTTAAGAAAAGGTTGGGCCGAGTCCCCAGAATGACTAACGTTTTGCCCCACCCCGTGACTTAGTTCC
-466
TTTACACACGCGTGGTATCTCTCGGTGGTCGTCCGGTGGTCTTCTCTGATGAAGGCTGGATGTCTTTCTCGCAGTGTACTCTCTGACCTGGGTCCCTTTTCGCATCCACAGCTCCCTCCA
-346
CTGCGCGCACCTGTTGGAGGGGCTGGCAGGAGTCATCCTACGCATAAGGACAAGAGTTGAAGCATACACTCAGCAGCTAGGCTCTAGTATTTCTGCATTGTGAGATCATTTGGGAGCATT
-226
AGATCATTATTAGTACCACATGAATGATTACCCAAAGTCACACAAGCTGTACATGTAAAAAAGTCAGAGTAAACAATGAGTTCATCCCTAGTTTGTTCATTCTAATCTTGAGCAGATTAA
Inverted CCAAT
TATA Box
+1
-107
TAAGTAACCTGCTGCCTCAGCAGGAACAGGGAGCTGATATTGGCTGATTTAATCCACGTGCTTTTGTTCTACAGCTTATAAATACCTCAGGTTTGCAGGTCCCAGTTCAGACTCTTCTGA
+12
M D F L G L P T I L L L V C I S C L L I A A W R S T S Q R G K
CACTTGACATCTCTTCCTCTGCCCACCATGGACTTCCTGGGATTGCCCACAATCCTCTTGCTGGTCTGTATCTCATGCCTTCTCATTGCTGCATGGAGGAGTACATCGCAAAGAGGGAAG
+132
E P P G P T P I P I I G N V F Q L N P W D L M G S F K E
GAGCCTCCTGGTCCCACGCCAATCCCCATCATTGGAAATGTTTTTCAGCTGAACCCGTGGGACTTGATGGGAAGCTTTAAGGAGGTAAGGCTGCCTCTTTCATTTCTGTGTTGTCTCCCT
+252
GGCCTGTAGAACAAAGCCCAGTCCAGACAGTCTCAGCTTAATTCACAGCTTCTGTGTAAATGTTTGCTTTTGATGTATTACGCTGCCTACTGGAATTACCATACCTATGTTGAGACCTGC
+372
ACTGGCCTCACGACAGTACTGCAGTATACAACAGATTTCTAGACAAGGTATTGATGCTTTTGCCATATTCACCACATCCCTTCTGTACCCATCCCCTCAGTGCTCTGGCAACAACGCCGA
+492
GCACTGTTTTAGGGAGGACCTTGCTATGCTTTGTTGTCTCTGCTCCCTGCCAGGTAGGTATAATCTCCAGGGAGGGGTTGTCACGAGTATAAGGGATGAATCAAAGGAGTTGAAGATGAA
+612
TGGTGCTGTCTGTTCCCTCCCCCGGATACTGTCATGCATCTGTTCAACAGAACCTGTATTGCAGTGCCTGGGAGAGAATAGGTGGAGTAAGTTTGGCTTTGGAGCCGGTTCCTGAGGGGA
+732
CTGTGATGGAATGGTGGCAGATGGATGGGAAAGAGCTAAGAAATGAAATGGTTCCCAGAAAATTTGCGTAGGGGTCTTATTCACCAGGCGAAGACCTGCAAACTTAATTTTTCCATGATT

+852

L S K K Y G P I F T I H L G P K K I V V L Y G Y D I

AACATCCAGATGTGTTGTTCCTTGTGTTCTTCCTTCCTGCAGCTCAGCAAAAAGTACGGTCCTATCTTCACAATACATTTAGGCCCCAAAAAGATTGTGGTGCTGTATGGCTACGACATT

+972

V K E A L I D N G E A F S G R G I L P L I E K L F K G T

GTGAAAGAAGCCCTAATTGATAACGGAGAAGCCTTCAGTGGGAGAGGAATACTGCCTCTGATTGAAAAGCTCTTCAAGGGCACAGGTGCGTCCCAGGTAACACTGCCATACAGAACTCGA

+1092

TCGTGTCATTGGAACAGGAATCCACCTGTGAGGGCAGTGGAGGAGTGACTGGAGGGTAATACACTACATTTCTACATAAAAATTATGGAATAAAGCATTAAAAGCCCAACCCATTACTTG

+1212

GGCCTAGCAGTAGGTTGAAGCCCTGGAGAACTGTTTAAGGTCCTAGGTTGCCTTCAGTGAGTTCCTTTCAATTTTCAGCACCATTCATATTCTTTCTACAGTTTCGTTCCCAAAAGGTGC

+1332

CCCTAAGATCTTACTGTTATTCTTCCATTTCTGAAAATACTGGCAAAGTCCCTACAGCTCAGGAGCCATCAACTGAACACAACACTGACAAGCAAATACAGGCTGTGCTAGTTAAGGGTC

+1452

CTACTCAAGATTAGGCCCTGGGCTCATTCACCATGACAGTCCCCAGCAGAGCATTTATTCAAGTAAATGCTTGTCCCGTGGTATTTTCAGTAATGATACGATGTGTGAAGGTGCATCACT

+1572

GCACATCTGGGAATTTGTCTTAGGCACCTGCTTGTGGCCAAGTTGACTTTAGCACCAGTGGTCCCTTTCTATTGACAGCCCATTGTCACTTCTCTTTTTCTGGAAATAAACGGTGGTGCC

+1692

TTGATCAGATTGATCTATTTGTCACCTCCATAACTGCAGGAAGATTAGGGTTTGGGGCATATCCCAATTGTGGTAGATATCTATTCCAGAACGAGGTTGGGGGGCAGAAAATTTACTTAC

+1812

G I V T S N G E T W R Q L R R F A L T T L R D F G M G K K

TCATGGAATCTTTGTTGTTCCTTCTTGTGACAGGCATTGTGACCAGCAATGGGGAGACCTGGAGGCAACTGCGACGATTTGCCCTCACCACCCTGCGTGATTTTGGAATGGGGAAGAAGG

+1932

G I E E R I Q E E A H F L V E R I R K T H

GCATTGAGGAGCGAATCCAGGAGGAAGCTCATTTTCTGGTGGAGAGGATCAGGAAGACACACGGTAGGACACTGACGTGTGTCTGCTGGGTAACACTGCGTATCTGCAATGAGACTGTAA

+2052

GAAAATTCAGTCTGGATTCACAAATAACAATCCCTGCTACCAATCTGAGAAAATGTGCTCATCTCTACTGGTCTTTTCATTTCACTGATTCTGAATTGTATATGGCCTCACTGTAAGCTT

+2172

CCAAGTCTCGGTCGCTTTGTTACGCTGTATACTACATTCTATGCTTTACTATGTACTAAAAGGCTTCACAAGTAACTGAAAAGTCCAGAAGCGTTTTTATACTTAATGCTAACAATAAAT

+2292

AGATAGGACTACTAACCTACTAGACTTACTTACGCTGAGCATTACTATGCTTACCACTTCTGACAAGATATCCGATGAGGTTTTGCTGGCGAGGAGTCGCGAGCAAGGCAGAGAAGGGCA

+2412

GGCTCCAGTGCTGCACAGAAGAGGCTGATCCTTAACGTTACCATTTTCTCTTCCATCTTTTTTCCTTTAGAATTTTATAACTTTCCATCTTGATGATTCTCATCTTCCATACTTGAGATA

+2532

TACCTTACCGCGCTCTCCAGGTTCAACTTTCGTGTCTTTCATGGATCTAATCTCATTTTTTCTCACGCTTCCATAGTTATTTACTTAGAAAATTCTGTAGCCGTGCTGTCAAAACTCTAT

+2652

TTTTGATCACGGGTATCTGAAGTTTCATTGGGTTATGGTGGAATTC

The nucleotide sequences encoding the 3.5 Kb mRNA were identified by comparing the sequence of pCHP3 with the sequence of the 3.8 Kb genomic fragment. The genomic sequence covered the first three exons of the gene and contains 1068 bp of the 5' flanking region. The P450 exonic regions in the 3.8 Kb BamHI/EcoRI fragment were perfectly homologous with the corresponding region of pCHP3. Exons 1, 2 and 3 measure 206, 163 and 150 bp respectively. The sequences at the exon/intron boundaries are in agreement with the consensus sequences noted for such splice junctions in several eukaryotic genes (Breathnach and Chambon, 1981), whereby the invariant GT of the donor and AG of the acceptor sequence define the 5' and 3' terminal dinucleotides of each intron respectively.

Previous primer extension analysis of poly A^+ RNA isolated from the livers of chick embryos induced with AIA and DDC, indicated that the complete 5' noncoding region of the 3.5 Kb mRNA was represented in pCHP3 (refer to Chapter 3). By comparing the 39 nucleotides of the 5' noncoding sequence of pCHP3 with the genomic sequence, the A residue denoted +1 in Fig. 7.4 was identified as the site of transcription initiation or the mRNA cap site. It has been observed that most, but not all eukaryotic gene transcripts are initiated with an A residue flanked by pyrimidines (Breathnach and Chambon, 1981; Sadler et al., 1983). The mRNA cap site of this gene is an A residue flanked by a pyrimidine and a purine and therefore is in keeping with these observations.

Analysis of the sequences immediately upstream from the mRNA cap site revealed the presence of the sequence $^{5'}$TTATAAATA$^{3'}$ at positions -33 to -25, bounded on both sides by GC rich sequences. This region is similar in sequence and position to the TATA box described by Goldberg (1979). The function of the TATA box appears to be in specifying the

site of transcription initiation by RNA polymerase II (Grosveld et al., 1982).

A second consensus sequence best described by the sequence $5'(G/A)G(C/T)CAA(T/A)^{3'}$ has been identified as a moderately conserved element found approximately 70 to 80 bp upstream from the mRNA cap site of many eukaryotic genes (Benoist et al., 1980; Breathnach and Chambon, 1981). Although the presence of this sequence is not essential (Govind et al., 1986; Suwa et al., 1985), several in vitro mutagenesis experiments using globin promoters (Grosveld et al., 1982; Charnay et al., 1984) and the herpes simplex virus thymidine kinase promoter (Graves et al., 1986) have shown that the integrity of the sequence plays an important role in modulating the rate of gene transcription. Examination of the equivalent region of the chicken P450 A gene revealed the sequence $^{5'}CCAAT^{3'}$ at positions -70 to -66 on the noncoding strand, in the reverse orientation with respect to the gene.

A computer aided search of the 3.8 Kb genomic sequence failed to detect any additional regulatory sequences within the 5' flanking and intronic regions of the chicken P450 A gene. Previous sequences which have been identified as regulators of gene expression include enhancer elements (Khoury and Gruss, 1983; Banerji et al., 1983; Goodburn et al., 1985; Choi and Engel, 1986), G-C boxes (Kadonaga et al., 1986), binding sites for hormone receptor proteins (Moore et al., 1985) and left-handed or Z-DNA (Rich et al., 1984).

7.3. DISCUSSION.

The study presented in this chapter clearly indicated that the chicken P450 A and B genes are both transcriptionally active. Southern hybridization analysis of genomic clones representing each of the genes revealed that the 3.5 Kb mRNA is derived from gene A and the minor

2.2 Kb mRNA is derived from gene B. The classification of the chicken P450 A and B genes as PB inducible has been previously determined by homology with a cDNA clone derived from a PB inducible mRNA. Previous studies by Brooker and coworkers (1983) established that AIA, DDC and PB caused an increase in the level of P450 mRNA in the livers of chick embryos. Future *in vitro* transcription experiments in hepatic nuclei will determine whether this increase in P450 mRNA is due to transcriptional activation of the chicken A and B genes. Similar studies with isolated rat liver nuclei indicate that only one out of six genes homologous to a cDNA clone for P450e, is markedly inducible by PB (Atchison and Adesnik, 1983).

The relationship of the 2.8 Kb mRNA species to the two genes is at present unclear. Previous studies indicated that all three chicken P450 mRNAs differed in their respective coding and 3' noncoding regions. However, the comprehensive study presented in Chapter 5, argues against the existence of a third gene in the chicken PB inducible P450 gene family. The possibility exists that two of the PB inducible P450 mRNAs detectable in the livers of drug treated chick embryos, originate from a single gene. In such an event, the selective use of multiple polyadenylation sites and/or alternative RNA processing could be involved in generating different mRNAs from a single gene. Examples of this include mouse α-amylase (Tosi *et al.*, 1981; Young *et al.*, 1981), chicken vimentin (Zehner *et al.*, 1983) and chicken smooth muscle α-actin (Carroll *et al.*, 1986) gene expression. Currently there is no unequivocal evidence for two or more mRNAs encoding different P450 isozymes, originating from the same gene. Three transcripts have been reported for bovine P450-11β (John *et al.*, 1985), each of which directs the synthesis of P450-11β in an *in vitro* translation system. However, the number of genes encoding P45011β and the mechanism of production of

the three mRNA species is not known. Sequence analysis of cDNA clones for the 2.8 Kb mRNA and the two chicken P450 A and B genes should clarify these uncertainties.

The sequence of a 3.8 Kb BamHI/EcoRI fragment encompassing 1 Kb of the 5' flanking region and the first three exons of the chicken P450 A gene was determined. The nucleotide sequence of the three exons is identical to the corresponding region of pCHP3, each exon-intron boundary conforming to the GT-AG rule (Breathnach and Chambon, 1981). Examination of the sequences immediately upstream from the mRNA cap site of the gene revealed that this gene contained a canonical TATA box in the expected position and a potential CCAAT box in the expected position, albeit inverted relative to the orientation in which these elements are more commonly found. This is in contrast to the other PB inducible P450 genes analyzed to date, none of which contain a consensus CCAAT box or one of its usual variants (Govind et al., 1986; Suwa et al., 1985; Mizukami et al., 1983b). Inverted CCAAT sequences have been reported for several genes including the chicken smooth muscle α-actin gene (Carroll et al., 1986), the chicken adipocyte serine protease (Phillips et al., 1986) and herpes simplex virus thymidine kinase (Graves et al., 1986). In vitro mutagenesis and DNase footprinting experiments by Graves and coworkers (1986) have shown that the inverted CCAAT box of the herpes simplex virus thymidine kinase gene is necessary for optimal transcriptional efficiency and that it interacts with a CCAAT binding transcriptional factor. The significance of the inverted CCAAT box in the chicken P450 A gene is presently unclear.

Comparison of the partial sequence of the chicken P450 A gene with other eukaryotic gene sequences failed to detect any additional consensus sequences previously associated with the transcriptional regulation of gene expression (Khoury and Gruss 1983; Kadonaga et al.,

1986; Moore et al., 1985; Rich et al., 1984). Negligible homology was observed between the chicken P450 A gene and the PB inducible rabbit P450Bc2 (Govind et al., 1986), rat P450b and P450e genes (Suwa et al., 1985), in their respective 5' flanking regions. Stretches of poly $(CA)_n$ present in the 5' flanking regions of rat P450b and P450e genes (Suwa et al., 1985) are absent in the 5' flanking region of the chicken A gene. It is possible that the cis-acting genomic elements involved in the PB inducibility of these P450 genes may be located further upstream of the transcriptional unit. The techniques of DNase footprinting, deletion studies and in vitro mutatagenesis will enable the identification of cis-acting genomic elements and their role in the transcriptional regulation of the chicken P450 A gene by the drugs AIA, DDC and PB.

CHAPTER 8

SUMMARY AND CONCLUDING DISCUSSION.

8.1. INTRODUCTION.

The mechanism by which structurally unrelated drugs induce homologous P450s is of immense interest. A fundamental requirement for understanding drug inducibility at the molecular level involves the complete characterization of the P450 genes and their regulatory sequences. The work described in this thesis dealt with an investigation into the chicken PB inducible gene family. The aim of this study fell into two main categories:

(1) the characterization of the chicken P450 genes inducible by the drugs AIA, DDC and PB.

(2) the analysis of the structure and multiplicity of these P450 genes.

Although several questions have arisen as a consequence of this work, the initial aims were achieved in this study. The general conclusions derived from these studies are discussed in this chapter.

8.2. SUMMARY AND CONCLUDING DISCUSSION.

At the commencement of this project, information concerning the chicken PB inducible P450 gene family was limited. A small cDNA clone (p2H10) homologous to P450 mRNA induced by the drugs AIA and DDC in the livers of chick embryos, had been isolated in this laboratory (Brooker and O'Connor, 1982). Additional studies by Brooker and coworkers (1983) established that each of the drugs AIA, DDC and PB induced homologous forms of hepatic P450 mRNA. In order to study the structure and expression of the chicken PB inducible gene family, the construction of a full length cDNA clone was an essential prerequisite. As described in Chapter 3, hepatic poly A^+ RNA isolated from the livers of chick embryos induced with AIA and DDC was used to generate a cDNA library. The double tailing method of Land et al. (1981) was used in an attempt to

generate full length cDNA clones. Ten P450 cDNA clones were isolated from this library, using the previously characterized cDNA clone p2H10 as a specific probe.

The largest recombinant, pCHP3, was sequenced. The 2712 bp cDNA insert of pCHP3 contained the complete 5' noncoding region and the entire coding region of the corresponding mRNA. The first fifteen amino acid residues of the predicted protein were confirmed by amino-terminal sequencing of the major P450 protein purified from the livers of chick embryos induced with AIA. The 3' noncoding region of pCHP3 lacked a poly A tail and consensus polyadenylation signal, which suggested that the clone was not full length for the corresponding P450 mRNA. Northern hybridization analysis of AIA and DDC induced chicken mRNA, detected the presence of three mRNA species homologous to pCHP3, measuring 3.5 Kb, 2.8 Kb and 2.2 Kb. The presence of all three mRNAs in the livers of individual chick embryos induced with either AIA, DDC or PB indicated that they were all inducible by each drug and were not the consequence of polymorphic variation within the chicken population. The results of further Northern hybridization studies established unequivocally, that pCHP3 represented the 3.5 Kb chicken mRNA. In addition, the 3.5 Kb and 2.8 Kb mRNAs were shown to be highly homologous in their respective 5' noncoding regions, but differed in their coding and 3' noncoding regions. It was concluded that these two mRNAs encode different isozymes. The DNA sequence of a second cDNA clone pCHP7, isolated from the library described in Chapter 3, was recently determined and shown to be full length for the 2.2 Kb mRNA (A. Hansen, personal communication). Comparisons of the nucleic acid and derived amino acid sequences of pCHP3 and pCHP7, revealed that the 3.5 Kb and 2.2 Kb mRNAs encoded P450 isozymes, containing 491 residues, which shared 92% homology in their primary sequences. Whereas the 5' noncoding regions of the mRNAs were

identical, they differed substantially in the length and sequence of their respective 3' noncoding regions. However, the preliminary study presented in this thesis did not determine whether the protein encoded by the 2.8 Kb mRNA is different from the isozyme encoded by the 2.2 Kb mRNA. The ongoing analysis of additional clones complementary to the 2.8 Kb mRNA should enable a more detailed comparison of the mRNAs and the proteins they encode.

As mentioned previously, pCHP3 was not full length for the 3.5 Kb mRNA; it lacked up to 800 nucleotides of the 3' noncoding region. In an attempt to obtain cDNA clones complementary to this region of the 3.5 Kb mRNA, a second library was constructed using the loopback method of Maniatis et al. (1982), which preferentially clones the 3' ends of mRNAs. Twenty P450 cDNA clones were shown by hybridization analysis to contain sequences corresponding to the extreme 3' PstI fragment of pCHP3, fragment C. The largest hybridizing fragment was 840 bp in length and represented the cDNA insert of pCHPB15. Sequence analysis of this 840 bp fragment extended the sequence of pCHP3 by 134 nucleotides in the 3' direction. However, the two clones pCHP3 and pCHPB15 collectively contained only 2846 nucleotides of sequence representing the 3.5 Kb chicken mRNA. The absence of a consensus polyadenylation signal or one of its usual variants (Birnstiel et al., 1985) confirmed the observation that pCHPB15 did not represent the complete 3' noncoding region of the 3.5 Kb mRNA. Therefore, the cDNA sequence presented in Chapter 3 is still approximately 400 nucleotides short of the expected size of the mRNA, excluding the poly A tail. The presence of two oligo (dA) tracts within the 3' noncoding region of the 3.5 Kb mRNA may explain the failure to obtain clones corresponding to the extreme 3' end. Efficient annealing of the oligo $(dT)_{12}$ primer to these internal oligo (dA) tracts would interfere with the synthesis of cDNA from the

true poly A tail. Therefore, the likelihood of the P450 cDNA clones isolated from either of the two libraries containing sequences complementary to this 3' region of the 3.5 Kb mRNA would be low.

The determination of sequences for the chicken P450 isozymes encoded by the 3.5 Kb and 2.2 Kb mRNAs, enabled the classification of these proteins. Comparison of the derived amino acid sequences of pCHP3 and pCHP7 with those of the mammalian P450s, revealed that the greatest extent of homology (56%) was observed with the poorly PB inducible isozyme, rat P450-PB1, and the constitutive form, rabbit P450-1. Since rat P450-PB1 and rabbit P450-1 are members of the PB II subfamily, by definition, pCHP3 and pCHP7 have been classified as members of this subfamily.

During the preparation of this discussion, an alternative nomenclature for the P450 gene family was proposed. Based on the alignment of amino acid sequences, Nebert and coworkers (1987) have proposed the existence of at least five subfamilies within the PB inducible P450 family (i.e. P450IIA-E). By definition, any two proteins within the same subfamily are 70% or more homologous in their amino acid sequences. Within a gene family in any given species, protein members of different subfamilies are 40-65% homologous in their amino acid sequences. According to this criterion, the derived amino acid sequence of pCHP3 has been designated P450IIC10 (Nebert et al., 1987).

An analysis of the derived amino acid sequences of pCHP3 based on an evolutionary comparison has recently appeared (Hobbs et al., 1986). Since this was not a primary aim of the project, the conclusions drawn from these studies have not been included in this discussion.

In view of the complexity of the mammalian PB inducible P450 gene family (Atchison and Adesnik, 1983), it was considered necessary to first determine the number of P450 genes homologous to pCHP3 in the

chicken genome, prior to isolating genomic clones. Southern hybridization analysis of genomic DNA under conditions of low and high stringency, suggested that only a small number of P450 genes homologous to pCHP3 existed in the chicken genome. When pCHP3 was used as a probe to screen a genomic library, twenty nine recombinants were isolated and mapped into two nonoverlapping groups, A and B. Both groups contained restriction fragments which hybridized to both the 5' and 3' fragments of pCHP3. Therefore, it was concluded that groups A and B represented two separate genes. Comparison of restriction digests of the cloned DNAs and genomic DNA discounted the possibility that other closely related P450 genes existed in the chicken genome. The results of Southern hybridization studies presented in Chapter 5, strongly indicated that both the A and B genes exist at separate genetic loci, as single copies per haploid genome. The detection of only two genes within the chicken genome makes it the smallest PB inducible family characterized to date. This is in direct contrast to the mammalian families where 6-10 members have been reported (Adesnik and Atchison, 1983; Phillips et al., 1985).

Extensive restriction mapping with subclones of pCHP3 indicated that a small section of the 3' ends of the A and B genes may not be represented by the existing genomic clones. In an attempt to isolate clones covering this region, a second genomic library was screened. pCHPB15 was used a probe since this clone contained the extreme 3' sequences, which had been previously shown to hybridize to genomic clones representing both the A and B genes. Thirty four recombinants were detected and shown by mapping studies to represent the chicken P450 A gene. These genomic clones extended the previous clones up to 5 Kb in the 3' direction. Therefore, the overlapping clones obtained from the

two libraries were considered likely to encompass the entire chicken P450 A gene and its immediate flanking sequences.

The failure to detect genomic clones for gene B using pCHPB15 as a probe was completely unexpected. Previous mapping results indicated that the sequences contained within the probe were common to both genes. As discussed in Chapter 5, the genomic clones isolated for the B gene may have undergone a recombination event. Since a rec A^- strain was used for phage propagation, it seems unlikely that such an event would have occurred during screening of the library. Therefore, a recombination event may have occurred during either construction or subsequent amplification of the library. A chicken cosmid library is currently being screened for the B gene and should help to clarify the situation.

Determination of the nucleotide sequences for the 3.5 Kb, and more recently the 2.2 Kb chicken P450 mRNAs, permitted an investigation into the relationship between these two mRNAs and the genomic clones representing the genes A and B. The results presented in Chapter 7, clearly indicate that both genes are transcriptionally active. Southern hybridization studies indicated that the 3.5 Kb mRNA was derived from the A gene and the 2.2 Kb mRNA from the B gene. It should be noted that the classification of the chicken P450 A and B genes as PB inducible was determined through their homology with a cDNA clone derived from a PB inducible mRNA. However, it is possible that the accumulation of P450 mRNA following drug treatment is due to enhanced mRNA stability rather than an increase in the rate of P450 gene transcription. Extensive studies have established that several members of the steroid hydroxylase (John et al., 1986), the TCDD inducible (Israel and Whitlock, 1984; Kimura et al., 1986) and the PB inducible (Hardwick et al., 1983a) P450 gene families are transcriptionally activated by their respective

inducing agents. The rapid increase in rat P450-PCN by PCN is also probably due to transcriptional regulation (Hardwick et al., 1983b). However, the induction of rabbit P450-3c enzyme by the macrolide antibiotic triacetyloleandomycin (TAO), appears to be a consequence of either enhanced mRNA or protein stability (Dalet et al., 1986). A similar post-transcriptional mechanism has been proposed for the induction of rat P450j protein by acetone, pyrazole and 4-methylpyrazole (Song et al., 1986). Although it seems reasonable to assume that the drug induced increase in chicken P450 mRNA is due to transcriptional regulation, the potential role of enhanced mRNA stability cannot be ignored. Therefore, similar "run on" experiments in isolated chicken hepatic nuclei are essential to confirm if AIA, DDC and PB increase the rate of chicken P450 A and B gene transcription.

The relationship of the 2.8 Kb mRNA to the two chicken P450 genes is presently unclear. The comprehensive study presented in Chapter 5 discounts the possibility that a third gene exists in the chicken PB inducible gene family. The favoured view is that two of the PB inducible P450 mRNAs detectable in the livers of drug treated chick embryos originate from a single gene, by either the selective use of multiple polyadenylation sites or alternative splicing. As described in Chapter 3, Northern hybridization studies established that the 3.5 Kb and the 2.8 Kb mRNAs are not colinear, but differ in their respective coding and 3' noncoding regions. If both the 3.5 Kb and 2.8 Kb mRNAs are derived from the A gene, the differences in their respective coding regions must be due to alternative splicing of the primary transcript. However, the differential use of polyadenylation sites may also contribute to the differences observed in their respective 3' noncoding regions. Recent Northern hybridization studies in this laboratory have established that the 2.8 Kb and 2.2 Kb mRNAs differ in their respective

3' noncoding regions (A. Hansen, personal communication). However, these studies were unable to determine if the 2.8 Kb and 2.2 Kb mRNAs encode different P450 isozymes. Therefore, if the two mRNAs are derived from the B gene, it is not possible from the existing information to distinguish between the involvement of either alternative splicing or differential polyadenylation for generating these transcripts. The complete sequence analysis of the A and B genes and cDNA clones corresponding to the 2.8 Kb mRNA will provide the answer to this question. Multiple mRNA transcripts have been reported for the bovine P450-11β (John et al., 1986) and the human 21-hydroxylase, P450-1 (Tukey et al., 1986). Although Tukey and coworkers proposed that the two human P450-1 transcripts may be generated by alternative splicing, the studies did not determine if these transcripts were functional in vivo. However, the three mRNA transcripts observed for the bovine P450-11β, were all shown to direct the synthesis of P450-11β in an in vitro translation system. Therefore, the precedent exists for the potential involvement of alternative splicing for the generation of different P450 transcripts from the same gene.

To understand the control of chicken P450 gene expression, an initial prerequisite is to sequence the promoter and 5' flanking regions. Attention was focused on the chicken P450 A gene since at the time it was clear that this gene was transcriptionally active and that overlapping genomic clones representing the entire gene were available. Fine mapping studies localized the promoter to a 3.8 Kb BamHI/EcoRI fragment of λCHP3. The nucleotide sequence of this genomic fragment encompassed 1 Kb of the 5' flanking region and the first three exons of the gene. The nucleotide sequence of the three exons was identical to the corresponding region of pCHP3, each splice junction conforming to the GT-AG rule (Breathnach and Chambon, 1981). The sequence upstream of

the mRNA cap site contained a canonical TATA and an inverted CCAAT box in the expected positions.

The 5' flanking sequence of the PB inducible rabbit P450 PBc2, and the rat P450b and P450e genes have been reported (Suwa et al., 1985; Govind et al., 1986). Comparison of these sequences detected the presence of alternating purine/pyrimidine residues at positions -254 of both the rat P450b and P450e genes. The sequence has the potential to form Z-DNA and is much longer in the P450e gene than in the P450b gene. Z-DNA has been proposed to play a role in the transcriptional regulation of gene expression (Nordheim and Rich, 1983). The difference in the length of the Z-DNA between the rat P450b and P450e genes has been postulated to play a role in regulating the difference in the basal and PB induced levels of these two P450 transcripts. To date, this hypothesis remains untested. The absence of Z-DNA in the 5' flanking sequences of the chicken P450 A gene and the rabbit PB inducible P450Bc2 gene (Govind et al., 1986) suggest that its presence in the rat P450b and P450e genes is coincidental. Therefore, the role of Z-DNA in the inducibilty of P450 by PB remains tenuous.

Additional computer aided comparison of the chicken genomic sequence with the sequences of other eukaryotic genes, including the mammalian PB inducible P450 genes (Suwa et al., 1985; Govind et al., 1986), did not detect the presence of consensus sequences which have been previously identified as regulatory elements of gene expression. Such sequences include transcriptional enhancers (Khoury and Gruss, 1983), G-C boxes (Kadonaga et al., 1986) and binding sites for hormone receptor proteins (Moore et al., 1985). The absence of such consensus sequences does not necessarily indicate the absence of regulatory regions in the chicken P450 gene promoter. Extensive deletion studies in mouse hepatoma cells have established that induction of P_1450 by TCDD

involves a balance between positive and negative cis-acting control elements, which span 2.6 Kb of the immediate 5' flanking region of the gene. Similar studies by Sogawa et al. (1986) on the rat P450c gene have established that the DNA sequences responsible for mediating TCDD inducibility are located up to 3.6 Kb 5' from the transcription start site. Therefore, it is possible that the sequences involved in the PB inducibility of the chicken P450 genes may be located further upstream of the 1 Kb flanking region sequenced. The ultimate determination of regulatory elements within the chicken P450 A gene, will require direct functional studies such as deletion analysis, DNase footprinting studies and in vitro mutagenesis.

The isolation of the chicken PB inducible P450 A and B genes, is a crucial step towards investigating the mechanism by which structurally unrelated drugs induce homologous forms of P450. The molecular mechanism of P450 induction by PB is poorly understood. Several studies have established that the increase in P450 mRNA by PB is due to transcriptional activation of the corresponding gene (Hardwick et al., 1983a; Pike et al., 1985; Gonzalez et al., 1986a). In addition, the expression and responsiveness of these genes to PB is tissue specific (Leighton and Kemper, 1984; Omiecinski, 1986). However, it is not clear if a receptor mediated mechanism analogous to that established for P450s inducible by polycyclic hydrocarbons (Whitlock Jr., 1986) is involved in P450 induction by PB. The complexity of the mammalian families studied and the absence of suitable PB responsive cell lines has hindered progress in this area. Given the simplicity of the chicken PB inducible P450 gene family, the chick embryo should prove valuable as an experimental system for pursuing such studies. As a first approach to understanding this mechanism of drug induction, future experiments are aimed at identifying the DNA sequences involved in the control of P450

gene expression. Recent studies in this laboratory have established that expression of the chicken P450 A and B genes is restricted to the liver and kidney (A. Hansen, personal communication). Since it is likely that expression of the P450 genes in the liver and kidney is mediated by specific trans-regulatory proteins acting at the level of transcription (Dynan and Tjian, 1985; Whitlock Jr., 1986; Veres et al., 1986), future experiments on the chicken P450 A and B genes will be limited to these tissues. Due to the lack of suitable hepatoma cell lines capable of P450 induction by PB, primary chicken hepatocytes will be used as an alternative system for studying the regulation of P450 gene expression.

APPENDIX

P450 NOMENCLATURE.

ISOZYME	SOURCE	INDUCERS (where applicable)
RAT P450a	HEPATIC MICROSOMES	AROCLOR 1254, PB[1], 3MC[2]
RAT P450b	HEPATIC MICROSOMES	AROCLOR 1254, PB
RAT P450c	HEPATIC MICROSOMES	AROCLOR 1254, 3MC, TCDD[3], βNF[4]
RAT P450d	HEPATIC MICROSOMES	AROCLOR 1254, 3MC, TCDD, βNF, ISOSAFROLE
RAT P450e	HEPATIC MICROSOMES	AROCLOR 1254, PB
RAT P450f	HEPATIC MICROSOMES	CONSTITUTIVE
RAT P450g	HEPATIC MICROSOMES	CONSTITUTIVE
RAT P450h	MALE HEPATIC MICROSOMES	CONSTITUTIVE
RAT P450i	FEMALE HEPATIC MICROSOMES	CONSTITUTIVE
RAT P450j	HEPATIC MICROSOMES	ACETONE, ETHANOL, PYRAZOLE
RAT P450-PB1	HEPATIC MICROSOMES	PB
RAT P450-PCN	HEPATIC MICROSOMES	PCN[5], PB, TAO[6]
MOUSE P450b	HEPATIC MICROSOMES	PB
MOUSE $P_1$450	HEPATIC MICROSOMES	TCDD, 3MC
MOUSE $P_2$450	HEPATIC MICROSOMES	ISOSAFROLE
MOUSE $P_3$450	HEPATIC MICROSOMES	TCDD, 3MC

REFERENCE	ALTERNATIVE NOMENCLATURE
Ryan et al. (1979)	P450 PB3; Waxman et al. (1983) P450 UTF; Guengerich et al. (1982)
Ryan et al. (1979)	P450 PB4; Waxman et al. (1982) P450 PB-B; Dannan et al. (1983)
Ryan et al. (1979)	P450 βNFB; Guengerich et al. (1982)
Ryan et al. (1980)	P450 βNF/IFG; Guengerich et al. (1982)
Vlasuk et al. (1982)	P450 PB5; Waxman et al. (1982) P450 PB-D; Guengerich et al. (1982)
Ryan et al. (1984)	—
Ryan et al. (1984)	—
Ryan et al. (1984)	P450 RLM5; Cheng & Schenkman (1982) P450 UT-A; Guengerich et al. (1982) P450-male; Kamataki et al. (1983) P4502c; Waxman (1984)
Ryan et al. (1984)	P4502d; Waxman (1984) P450-female; Kamataki et al. (1983)
Ryan et al. (1985)	P450et; Tu and Yang (1985)
Waxman & Walsh (1983)	P450 PB-C; Guengerich et al. (1982)
Elshourbagy et al. (1980)	P450 PB/PCN-E; Guengerich et al. (1982) P450p; Wrighton et al. (1985)
Stupans et al. (1984)	—
Negishi et al. (1979)	—
Ohyama et al. (1984)	—
Gonzalez et al. (1984)	—

ISOZYME	SOURCE	INDUCERS (where applicable)
RABBIT P450-1	HEPATIC MICROSOMES	CONSTITUTIVE
RABBIT P450-2	HEPATIC MICROSOMES	PB
	PULMONARY MICROSOMES	CONSTITUTIVE
RABBIT P450-3a	HEPATIC MICROSOMES	ETHANOL, BENZENE, IMIDAZOLE
RABBIT P450-3b	HEPATIC MICROSOMES	CONSTITUTIVE
RABBIT P450-3c	HEPATIC MICROSOMES	RIFAMPICIN, PB, TAO, DEXAMETHASONE
RABBIT P450-4	HEPATIC MICROSOMES	ISOSAFROLE, 3MC, TCDD, βNF
RABBIT P450-5	HEPATIC MICROSOMES	PB, AROCLOR 1254
	PULMONARY MICROSOMES	CONSTITUTIVE
RABBIT P450-6	HEPATIC MICROSOMES	TCDD, βNF, AROCLOR 1254
	PULMONARY MICROSOMES	AROCLOR 1254, TCDD, 3MC
Pseudomonas putida P450cam	P. putida	CONSTITUTIVE
CHICKEN P450-A	HEPATIC MICROSOMES	PB, AIA[7], DDC[8]
CHICKEN P450-B	HEPATIC MICROSOMES	PB, AIA, DDC
BOVINE P450ssc	ADRENAL MITOCHONDRIA	ACTH, cAMP
BOVINE P450c21	ADRENAL MICROSOMES	ACTH, cAMP
BOVINE P450-17α	ADRENAL MICROSOMES	ACTH, cAMP
BOVINE P450-11β	ADRENAL MITOCHONDRIA	ACTH, cAMP

REFERENCE	ALTERNATIVE NOMENCLATURE
Dieter et al. (1982)	—
van der Hoeven et al.(1974)	LM2; Black et al. (1982)
Slaughter et al. (1981)	—
Koop et al. (1982)	LMeb; Ingelman-Sundberg et al.(1984)
Johnson (1980)	—
Schwab & Johnson (1986)	LM3; Ingelman-Sundberg et al.(1980)
Johnson et al. (1977)	LM4; Haugen & Coon (1976) P448; Kawalek et al. (1975)
Robertson et al. (1983)	—
Wolf et al. (1979)	P450 II; Wolf et al. (1979)
Norman et al. (1978)	P450b; Johnson et al. (1977)
Serabjit-Singh et al. (1983)	P450 III; Ueng & Alvares (1982)
Katagiri et al. (1986)	—
Hobbs et al. (1986)	—
Hobbs et al. (1986)	—
DuBoie et al. (1980)	—
Kominami et al. (1980)	—
Zuber et al. (1985)	—
Watanuki et al. (1977)	—

ISOZYME	SOURCE	INDUCERS (where applicable)
PORCINE P450-17α	ADRENAL MICROSOMES	ACTH, cAMP
	TESTICULAR MICROSOMES	ACTH, cAMP
HUMAN P450bufI	HEPATIC MICROSOMES	UNKNOWN
HUMAN P450bufII	HEPATIC MICROSOMES	UNKNOWN
HUMAN P450-HLp	HEPATIC MICROSOMES	PCN, PB, TAO, DEXAMETHASONE

1. Phenobarbital.
2. 3-methylcholanthrene.
3. 2,3,7,8-tetrachlorodibenzo-ρ-dioxin.
4. β-Napthoflavone.
5. Pregnenolone-16α-carbonitrile.
6. Triacetyloleandomycin.
7. 3,5-diethoxycarbonyl-1,4-dihydrocollidine.
8. 2-allyl-2-isopropylacetamide.

REFERENCE	ALTERNATIVE NOMENCLATURE
Nakajin *et al.* (1984)	—
Watanuki *et al.* (1978)	—
Gut *et al.* (1986)	—
Gut *et al.* (1986)	—
Watkins *et al.* (1985)	—

BIBLIOGRAPHY.

Adesnik, M., and Atchison, M. (1983) **J. Biol. Chem.** 258, 11285-11295.

Adesnik, M., and Atchison, M. (1985) **CRC Crit. Rev. Biochem.** 19, 247-305.

Anderson, S. (1981) **Nucleic Acids Res.** 9, 3015-3027.

Armstrong, R. N., Pinto-Coelho, C., Ryan, D. E., Thomas, P. E., and Levin, W. (1983) **J. Biol. Chem.** 258, 2106-2108.

Banerji, J., Olson, L., and Schaffner, W. (1983) **Cell** 33, 729-740.

Baxter, J. D., and Rousseau, G. G. (1979) in **Glucocorticoid Hormone Action** (Baxter, J. D., and Rousseau, G. G., eds), pp. 1-24, Springer-Verlag, New York.

Beltz, G. A., Jacobs, K. A., Eickbush, T. H., Cherbas, P. T., and Kafatos, F. C. (1983) **Methods Enzymol.** 100, 266-285.

Benoist, C., and Chambon, P. (1981) **Nature** 290, 304-309.

Benton, W. D., and Davis, R. W. (1977) **Science** 196, 180-182.

Birnboim, H. C., and Doly, J.(1979) **Nucleic Acids Res.** 7, 1513-1523.

Birnstiel, M. L., Busslinger, M., and Strub, K. (1985) **Cell** 41, 349-359.

Black, S. D., Tarr, G. E., and Coon, M. J. (1982) **J. Biol. Chem.** 257, 14616-14619.

Bonner, W. M., and Laskey, R. A. (1974) **Eur. J. Biochem.** 46, 83-88.

Breathnach, R., and Chambon, P. (1981) **Annu. Rev. Biochem.** 50, 349-383.

Brooker, J. D., and O' Connor, R. (1982) **Eur. J. Biochem.** 129, 333-353.

Brooker, J. D., Srivastava, G., Borthwick, I. A., May, B. K., and Elliott, W. H. (1983) **Eur. J. Biochem.** 136, 327-332.

Carmichael, G., and M^{C}Masters, G. (1980) **Proc. Natl. Acad. Sci. U.S.A.** 74, 4835-4838.

Carroll, S. L., Bergsma, D. J., and Schartz, R. J. (1986) **J. Biol. Chem.** 261, 8965-8976.

Casadaban, M. J., and Cohen, S. N. (1980) **J. Mol. Biol.** 138, 179-207.

Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R., and Maniatis, T. (1984) **Cell** 38, 251-263.

Cheng, K. C., and Schenkman, J. B. (1982) **J. Biol. Chem.** 257, 2378-2385.

Choi, O., and Engel, J. D. (1986) **Nature** 323, 731-734.

Chung, B., Matteson, K. J., and Miller, W. L. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 4243-4247.

Dagert, M., and Ehrlich, S. D. (1979) **Gene** 6, 23-28.

Dale, R. M. K., McClure, B. A., and Houchins, J. P. (1985) **Plasmid** 13, 31-40.

Dalet, C., Blanchard, J. M., Guzelian, P., Barwick, J., Hartle, H., and Maurel, P. (1986) **Nucleic Acids Res.** 14, 5999-6015.

Dannon, G. A., Guengerich, F. P., Kaminsky, L. S., and Aust, S. T. (1983) **J. Biol. Chem.** 258, 1282-1288.

Dayhoff, M. O. (1976) in **Atlas of Protein Sequence and Structure** (Dayhoff, M. O., ed.) Vol 5, suppl. 2, National Biomedical Research Foundation, Md.

De Matteis, F. (1973) **Drug Metab. Disp.** 1, 267-272.

Denison, M. L., Vella, L. M., and Okey, A. B. (1986a) **J. Biol. Chem.** 261, 3987-3995.

Denison, M. S., Harper, P. A., and Okey, A. B. (1986b) **Eur. J. Biochem.** 155, 223-229.

Dieter, H., Muller-Eberhard, U., and Johnson, E.F. (1982) **Biochem. Biophys. Res. Commun.** 105, 515-520.

Dodgson, J. D., Strommer, J., and Engel., D. J. (1979) **Cell** 17, 879-887.

Dubois, R. N., Simpson, E. R., Hall, P. F., and Waterman, M. R. (1981) **J. Biol. Chem.** 256, 7000-7005.

Dynan, W. S., and Tjian, R. (1985) **Nature** 316, 774-777.

Elshourbagy, N. A., and Guzelian, P. S. (1980) **J. Biol. Chem.** 255, 1279-1285.

Fevold, H. R. (1983) **Annu. Rev. Physiol.** 45, 19-36.

Finkelstein, M., and Shaefer, J. M. (1979) **Physiol. Rev.** 59, 353-406.

Frischauf, A., Lehrach, H., Poustka, A., and Murray, N. (1983) **J. Mol. Biol.** 170, 827-842.

Fujii-Kuriyama, Y., Mizukami, Y., Kawajiri, K., Sogawa, K., and Muramatsu, M. (1982) **Proc. Natl. Acad. Sci. U.S.A.** 79, 2793-2797.

Fujisawa-Sehara, A., Sogawa, K., Nishi, C., and Fujii-Kuriyama, Y. (1986) **Nucleic Acids Res.** 14, 1465-1477.

Giachelli, C. M., and Omiecinski, C. J. (1986) **J. Biol. Chem.** 261, 1359-1363.

Goldstein, J. A. (1984) **TIPS REV.** 9, 290-292.

Gonzalez, F. J., MacKenzie, P. I., Kimura, S., and Nebert, D. W. (1984) **Gene** 29, 281-292.

Gonzalez, F. J., and Kasper, C. B. (1983) **J. Biol. Chem.** 258, 1363-1368.

Gonzalez, F. J., Nebert, D. W., Hardwick, J. P., and Kasper, C. B. (1985a) **J. Biol. Chem.** 260, 7435-7441.

Gonzalez, F. J., Kimura, S., and Nebert, D. W. (1985b) **J. Biol. Chem.** 260, 5040-5049.

Gonzalez, F. J., Kimura, S., Song, B. J., Pastewka, J., Gelboin, H. V., and Hardwick, J. P. (1986a) **J. Biol. Chem.** 261, 10667-10672.

Gonzalez, F. J., Song, B., and Hardwick, J. P. (1986b) **Molec. Cell. Biol.** 6, 2969-2976.

Govind, S., Bell, P. A., and Kemper, B. (1986) **DNA** 5, 371-382.

Graves, B. J., Johnson, P. F., and McKnight, S. L. (1986) **Cell** 44, 565-576.

Grosveld, G. C., Rosenthal, A., and Flavell, R.A. (1982) **Nucleic Acids Res.** 10, 4951-4971.

Grunstein, M., and Hogness, D. (1975) **Proc. Natl. Acad. Sci. U.S.A.** 72, 3961-3965.

Gubler, U., and Hoffman, B. J. (1982) **Gene** 25, 263-269.

Guengerich, F. P. (1979) **Mol. Pharmacol.** 13, 911-923.

Guengerich, F. P., Dannan, G. A., Wright, S. T., Martin, M. V., and Kaminsky, L. S. (1982) **Biochemistry** 21, 789-798.

Guengerich, F. P., and Liebler, D. C. (1985) **CRC Crit. Rev. Toxicol.** 14, 259-307.

Haugen, D. D., and Coon, M. J. (1976) **J. Biol. Chem.** 251, 7929-7939.

Hardwick, J. P., Gonzalez, F. J., and Kasper, C. B. (1983a) **J. Biol. Chem.** 258, 8081-8085.

Hardwick, J. P., Gonzalez, F. J., and Kasper, C. B. (1983b) **J. Biol. Chem.** 258, 10182-10186.

Heinemann, F. S., and Ozols, J. (1983) **J. Biol. Chem.** 258, 4195-4201.

Higashi, Y., Yoshioka, H., Yamane, M., Gotoh, O., and Fujii-Kuriyama, Y. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 2841-2845.

Hobbs, A. A., Mattschoss, L. A., May, B. K., Williams, K. E., and Elliott, W. H. (1986) **J. Biol. Chem.** 261, 9444-9449.

Ikeda, T., Altieri, M., Cheny, T., Nakamura, M., Tukey, R. H., Nebert, D. W., and Negishi, M. (1983) **Eur. J. Biochem.** 134, 13-18.

Ingelman-Sundberg M., and Glaumann, H. (1980) **Biochim. Biophys. Acta** 599, 417-431.

Ingelman-Sundberg, M., and Jornvall, H. (1984) **Biochem. Biophys. Res. Commun.** 124, 375-382.

Israel, D. I., and Whitlock Jr., J. P. (1984) **J. Biol. Chem.** 259, 5400-5402.

Israel, D. I., Estalono, M. G., Galeazzi, D. R., and Whitlock Jr., J. P. (1985) **J. Biol. Chem.** 260, 5648-5653.

Iverson, P. L., Hines, R. N., and Bresnik, E. (1986) **Bioessays** 4, 15-19.

Jaiswel, A. K., Gonzalez, F. J., and Nebert, D. W. (1984) **Science** 228, 80-82.

John, M. E., John, M. C., Simpson, E. R., and Waterman, M. R. (1985) **J. Biol. Chem.** 260, 5760-5767.

John, M. E., John, M. C., Boggaram, V., Simpson, E. R., and Waterman, M. R. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 4715-4719.

Johnson, E. F., and Muller-Eberhard, U. (1977) **Biochem. Biophys. Res. Commun.** 76, 644-651.

Jones, P. B. C., Galeazzi, D. R., Fisher, J. M., and Whitlock Jr., J. P. (1985) **Science** 227, 1499-1502.

Jones, P. B. C., Durrin, L. K., Galeazzi, D. R., and Whitlock Jr., J. P. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 2802-2806.

Kadonaga, J. T., Jones, K. A., and Tjian, R. (1986) **TIBS REV.** 11, 20-23.

Kappas, A., and Alvares, A. P. (1975) **Scientific American** 232, 22-31.

Kappas, A., Sassa, S., and Anderson, K. E. (1983) in **The Metabolic Basis of Inherited Diseases** (Stanbury, J. B., Wyngaarden, J. B., Frederickson, D. S., Goldstein, J. C., and Brown, M. S., eds) 5th Ed., pp. 1301-1384, McGraw-Hill, New York.

Kamataki, T., Maeda, K., Yamazoe, Y., Nagai, T., and Kato, R. (1983) **Arch. Biochem. Biophys.** 225, 758-770.

Katagiri, M., Ganguli, B. N., and Gunsalus, I. C. (1968) **J. Biol. Chem.** 243, 3543-3546.

Kawajiri, K., Gotoh, O., Sogawa, K., Tagashira, Y., Muramatsu, M., and Fujii-Kuriyama, Y. (1984) **Proc. Natl. Acad. Sci. U.S.A.** 81, 1649-1653.

Kawalek, J. C., Levin, W., Ryan, D., Thomas, P. E., and Lu, A. H. Y. (1975) **Mol. Pharmacol.** 11, 874-878.

Khoury, G., and Gruss, P. (1983) **Cell** 33, 313-314.

Kimura, S., Gonzalez, F. J., and Nebert, D. W. (1984a) **Nucleic Acids Res.** 12, 2917-2927.

Kimura, S., Gonzalez, F. J., and Nebert, D. W. (1984b) **J. Biol. Chem.** 259, 10705-10713.

Kimura, S., Gonzalez, F. J., and Nebert, D. W. (1986) **Mol. Cell. Biol.** 6, 1471-1477.

Kominami, S., Ochi, H., Kobayashi, Y., and Takemori, S. (1980) **J. Biol. Chem.** 255, 3386-3394.

Koop, D. R., Morgan, E. T., Tarr, G. E., and Coon, M. D. (1982) **J. Biol. Chem.** 256, 10704-10711.

Kozak, M. (1981) **Nucleic Acids. Res.** 9, 5233-5252.

Kumar, A., Raphael, C., and Adesnik, M. (1983) **J. Biol. Chem.** 258, 11280-11284.

Laemmli, U. K. (1970) **Nature** 227, 680-685.

Land, H., Grez, M., Hauser, H., Lindermaier, W., and Schutz, G. (1981) **Nucleic Acids Res.** 9, 2251-2266.

Leighton, J. K., and Kemper, B. (1982) **Arch. Biochem. Biophys.** 216, 272-288.

Leighton, J. K., De Brunner-Vossbrick, B. A., and Kemper, B. (1984) **Biochemistry** 23, 204-210.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) **Molecular Cloning: A Laboratory Manual**, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Marshall, A. J., and Burgoyne, L. A. (1976) **Nucleic Acids Res.** 3, 1101-1110.

May, B. K., Borthwick, I. A., Srivastava, G., Pirola, B. A., and Elliott, W. H. (1986) **Current Topics Cell. Regulation** 28, 233-261.

Mayer, H. (1978) **FEBS Lett.** 90, 341-348.

Maxam, A. M., and Gilbert, W. (1980) **Methods Enzymol.** 65, 499-560.

Messing, J. (1979) **Recombinant DNA Technology Bulletin** 2, 43-48.

Miki, N., Sugiyama, T., Yamano, T., and Miyake, Y. (1981) **Biochem. Int.** 3, 217-223.

Mizukami, Y., Fujii-Kuriyama, Y., and Muramatsu, M. (1983a) **Biochemistry** 23, 204-210.

Mizukami, Y., Sogawa, K., Suwa, Y., Muramatsu, M., and Fujii-Kuriyama, Y. (1983b) **Proc. Natl. Acad. Sci. U.S.A.** 80, 3958-3962.

Molowa, D. J., Schuetz, E. G., Wrighton, S. A., Watkins, P. B., Kremers, P., Mendel-Picon, G., Parker, G. A., and Guzelian, P. S. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 5311-5315.

Moore, D. D., Marks, A. R., Buckley, D. I., Kapler, G., Payvar, F., and Goodman, H. M. (1985) **Proc. Natl. Acad. Sci. U.S.A.** 82, 699-702.

Morahashi, K., Fujii-Kuriyama, Y., Sogawa, O. K., Inayama, H. S., and Omura, T. (1984) **Proc. Natl. Acad. Sci. U.S.A.** 81, 4647-4651.

Murray, N. E., Brammar, W. J., and Murray, K. (1977) **Mol. Gen. Genet.** 150, 53-61.

Nakajin, S., Shinoda, M., Hanui, M., Shively, J. E., and Hall, P. F. (1984) **J. Biol. Chem.** 259, 3971-3976.

Nebert, D. W., and Jenson, N. M. (1979) **CRC Crit. Rev. Biochem.** 6, 401-437.

Nebert, D. W., and Gonzalez, F. J. (1985) **TIPS REV.** 6, 160-164.

Nebert, D. W., Adesnik, M., Coon, M. J., Estabrook, R. W., Gonzalez, F. J., Guengerich, F. P., Gunsalus, I. C., Johnson, E. F., Kemper, B., Levin, W., Phillips, I. R., Sato, R., and Waterman, M. R. (1987) **DNA** 6, 1-11.

Negishi, M., and Nebert, D. W. (1979) **J. Biol. Chem.** 254, 11015-11023.

Norman, R. L., Johnson, E. F., and Muller-Eberhard, U. (1978) **J. Biol. Chem.** 253, 8640-8647.

Ohkubo, H., Vogeli, G., Mudryj, M., Avvedimento, V. B., Sullivan, M., Pastan, I., and de Crombrugghe, B. (1980) **Proc. Natl. Acad. Sci. U.S.A.** 77, 7059-7063.

Ohyama, T., Nebert, D., and Negishi, M. (1984) **J. Biol. Chem.** 259, 2675-2682.

Okayama, H., and Berg, P. (1982) **Mol. Cell. Biol.** 2, 161-170.

Okino, S. T., Quattrochi, L. C., Barnes, H. J., Osanto, S., Griffin, K. J., Johnson, E. F., and Tukey, R. H. (1985) **Proc. Natl. Acad. Sci. U.S.A.** 82, 5310-5314.

Omiecinski, C. J., Walz Jr., F. G., and Vlasuk, G. P. (1985) **J. Biol. Chem.** 260, 3247-3250.

Omiecinski, C. J. (1986) **Nucleic Acids Res.** 14, 1525-1539.

Ozols, J., Heinemann, F. S., and Johnson, E. F. (1981) **J. Biol. Chem.** 256, 11405-11408.

Phillips, I. R., Shephard, E. A., Ashworth, A., and Rabin, B. R. (1985) **Proc. Natl. Acad. Sci. U.S.A.** 82, 983-987.

Phillips, M., Djian, P., and Green, H. (1986) **J. Biol. Chem.** 261, 10821-10827.

Pike, S. F., Shephard, E. A., Rabin, B. R., and Phillips, I. R. (1985) **Biochem. Pharmacol.** 34, 2489-2494.

Poland, A., Glover, E., Hal Ebitino, F., and Kende, A. S. (1986) **J. Biol. Chem.** 261, 6352-6265.

Porter, T. D., Beck, T. W. and Kasper, C. B. (1986) **Arch. Biochem. Biophys.** 248, 121-129.

Quattrochi, L. C., Okino, S. T., Pendurthi, U. R., and Tukey, R. H. (1985) **DNA** 4, 395-400.

Quattrochi, L. C., Pendurthi, U. R., Okino, S. T., Potenza, C., and Tukey, R. H. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 6731-6735.

Rampersaud, A., and Walz Jr., F. G. (1983) **Proc. Natl. Acad. Sci. U.S.A.** 80, 6542-6546.

Ravishankar, N., and Padmanaban, G. (1983) **Arch. Biochem. Biophys.** 225, 16-24.

Ravishankar, N., and Padmanaban, G. (1985) **J. Biol. Chem.** 260, 1588-1592.

Rich, A., Nordheim, A., and Wang, A. H. J. (1984) **Annu. Rev. Biochem.** 53, 791-846.

Robertson, I. G. C., Serabjit-Singh, C., Croft, J. E., and Philpot, R. M. (1983) **Mol. Pharmacol.** 24, 156-162.

Rosen, J. M., Liarokos, C. D., and O'Malley, B. W. (1973) **Biochemistry** 12, 2803-2809.

Rosen, J. M., Woo, S. L. C., Holder, J. W., Means, A. R., and O'Malley, B. W. (1975) **Biochemistry** 14, 69-78.

Ryan, D. E., Thomas, P. E., Koreniowski, D., and Levin, W. (1979) **J. Biol. Chem.** 254, 1365-1374.

Ryan, D. E., Thomas, P. E., and Levin, W. (1980) **J. Biol. Chem.** 255, 7941-7955.

Ryan, D. E., Thomas, P. E., Levin, W., and Shively, J. E. (1982) **Arch. Biochem. Biophys.** 216, 272-288.

Ryan, D. E., Shinji, I., Wood, A. W., Thomas, P. E., Lieber, C. S., and Levin, W. (1984) **J. Biol. Chem.** 259, 1239-1250.

Ryan, D. E., Ramanathan, L., Shinji, I., Thomas, P. E., Hanui, M., Shively, J. E., Lieber, C. S., and Levin, W. (1985) **J. Biol. Chem.** 260, 6385-6393.

Sadler, J. R., Waterman, M. S., and Smith, T. F. (1983) **Nucleic Acids Res.** 3, 863-877.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A. (1980) **J. Mol. Biol.** 143, 161-178.

Satayabhama, S., Sathiagnana, S. R., and Padmanaban, G. (1986) **Biochemistry** 25, 4508-4512.

Schuetz, E. G., and Guzelian, P. S. (1984) **J. Biol. Chem.** 259, 2007-2012.

Schuetz, E. G., Wrighton, S. A., Barwick, J. L. and Guzelian, P. S. (1984) **J. Biol. Chem.** 259, 1999-2006.

Schwab, G. E., and Johnson, E. F. (1986) in **Mammalian Cytochrome P450s** (Guengerich, F., ed.) CRC Press, New York, in press.

Serabjit-Singh, C. T., Albro, P. W., Robertson, I. G. C., and Philpot, R. M. (1983) **J. Biol. Chem.** 258, 12827-12834.

Simmons, D. L., Lalley, P. A. and Kasper, C. B. (1985) **J. Biol. Chem.** 260, 515-521.

Slaughter, S. R., Wolf, C. R., Marciniszyn, J. P., and Philpot, R. M. (1981) **J. Biol. Chem.** 256, 2499-2503.

Smith, G. E, and Summers, M. D. (1980) **Anal. Biochem.** 109, 123-129.

Sogawa, K., Gotoh, O., Kawajiri, K. I., and Fujii-Kuriyama, Y. (1984) **Proc. Natl. Acad. Sci. U.S.A.** 81, 5066-5070.

Sogawa, K., Fujisawa-Sehara, A., Yamane, M., and Fujii-Kuriyama, Y. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 8044-8048.

Song, B. J., Gelboin, H. V., Park, S. S., Yang, C. S., and Gonzalez, F. J. (1986) **J. Biol. Chem.** 261, 16689-16697.

Southern, E. M. (1975) **J. Mol. Biol.** 98, 503-517.

Stupans, I., Ikeda, T., Kessler, D., and Nebert, D. W. (1984) **DNA** 3, 129-137.

Suguoka, Y., Kano, T., Okuda, A., Sakai, M., Kitagawa, T., and Murumatsu, M. (1985) **Nucleic Acids Res.** 13, 6049-6057.

Suwa, Y., Mizukami, Y., Sogawa, K., and Fujii-Kuriyama, Y. (1985) **J. Biol. Chem.** 260, 7980-7984.

Tarr, G. F., Black, S. D., Fujita, V. S., and Coon, M. J. (1983) **Proc. Natl. Acad. Sci. U.S.A.** 80, 807-817.

Thomas, P. E., Kouri, R. E., and Hutton, J. J. (1972) **Biochem. Genet.** 6, 157-168.

Thomas, P. (1980) **Proc. Natl. Acad. Sci. U.S.A.** 77, 5201-5202.

Thomas, P. E., Reik, L. M., Ryan, D. E., and Levin, W. (1983) **J. Biol. Chem.** 258, 4590-4598.

Tosi, M., Young, R. A., Hagenbuchle, O., and Schibler, U. (1981) **Nucleic Acids Res.** 9, 2313-2323.

Tu, Y. Y., and Yang, C. S. (1985) **Arch. Biochem. Biophys.** 242, 32-40.

Tukey, R. H., Okino, S., Barnes, H., Griffin, K. J., and Johnson, E. F. (1985) **J. Biol. Chem.** 260, 13347-13354.

Tukey, R. H., Quattrochi, L. C., Pendurthi, U. R., and Okino, S. T. (1986) **Fed. Proc.** 45, 1663.

Tuteja, N., Gonzalez, F. J., and Nebert, D. W. (1985) **Dev. Biol.** 112, 177-184.

Ueng, T. H., and Alvares, A. P. (1982) **Mol. Pharmacol.** 22, 221-228.

van der Hoeven, T. A., Haugen, D. A., and Coon, M. J. (1974) **Biochem. Biophys. Res. Commun.** 60, 569-575.

Veres, G., Craigen, W. J., and Caskey, C. T. (1986) **J. Biol. Chem.** 261, 7588-7591.

Villa-Karmaroff, L., Efstratiadis, S., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L., and Gilbert, W. (1978) **Proc. Natl. Acad. Sci. U.S.A.** 75, 3727-3731.

Vlasuk, G. P., Ghrayeb, J., Ryan, D., Reik, L., Thomas, P. E., Levin, W., and Walz Jr., F. G. (1982) **Biochemistry** 21, 789-798.

Wahli, W., Dawid, I. B., Ryffel, G. U., and Weber, R. (1981) **Science** 212, 298-304.

Wallace, P. G., Hewish, D. R., Venning, M. M., and Burgoyne, L. A. (1971) **Biochem. J.** 125, 47-54.

Watanuki, M., Tilley, B. E., Hall, P. F. (1977) **Biochim. Biophys. Acta** 483, 236-247.

Watanuki, M., Granger, G. A., and Hall, P. F. (1978) **J. Biol. Chem.** 253, 2927-2931.

Watkins, P. B., Wrighton, S. A., Maurel, P., Schuetz, E. G., Mendez-Picon, G., Parker, G. A., and Guzelian, P. S. (1985) **Proc. Natl. Acad. Sci. U.S.A.** 82, 6310-6314.

Waxman, D. J., and Walsh, C. (1982) **J. Biol. Chem.** 257, 10446-10457.

Waxman, D. J., Ko, A., and Walsh, C. (1983) **J. Biol. Chem.** 258, 11937-11947.

Waxman, D. J., and Walsh, C. (1983) **Biochemistry** 22, 4846-4855.

Waxman, D. J. (1984) **J. Biol. Chem.** 259, 15480-15490.

Waxman, D. J., Dannan, G. A., and Guengerich, F. P. (1985) **Biochemistry** 24, 4409-4417.

White, P. C., New, M. I., and Dupont, B. (1984) **Proc. Natl. Acad. Sci. U.S.A.** 81, 1986-1990.

White, P. C., New, M. I., and Dupont, B. (1986) **Proc. Natl. Acad. Sci. U.S.A.** 83, 5111-5115.

Whitlock Jr., J. P. (1986) **Annu. Rev. Pharmacol. Toxicol.** 26, 333-369.

Wolf, C. R., Szutowski, M. M., Ball, L. M., and Philpot, R. M. (1978) **Chem. Biol. Interact.** 21, 29-33.

Wolf, C. R., Smith, B. R., Ball, L. M., Serabjit-Singh, C., Bend, J. R., and Philpot, R. M. (1979) **J. Biol. Chem.** 254, 3658-3663.

Wrighton, S. A., Maurel, P., Schuetz, E. G., Watkins, P. B., Whitley, B., and Guzelian, P. S. (1985) **Biochemistry** 24, 753-760.

Young, R. A., Hagenbuchle, O., and Schibler, U. (1981) **Cell** 23, 451-458.

Yu-Lee, L., and Rosen, J. M. (1983) **J. Biol. Chem.** 258, 10794-10804.

Yuan, P., Ryan, D. E., Levin, W., and Shively, J. E. (1983) **Proc. Natl. Acad. Sci. U.S.A.** 80, 1169-1173.

Zehner, Z. E., and Paterson, B. M. (1983) **Proc. Natl. Acad. Sci. U.S.A.** 80, 911-915.

Zuber, M. X., Simpson, E. R., Hall, P. F., and Waterman, M. R. (1985) **J. Biol. Chem.** 260, 1842-1848.

Zuber, M. X., John, M. E., Okamura, T., Simpson, E. R., and Waterman, M. R. (1986) **J. Biol. Chem.** 261, 2475-2482.

LIST OF PUBLICATIONS.

Isolation of two chicken cytochrome P450 genes homologous to mRNA induced by phenobarbital, 2-allyl-2-isopropylacetamide and 3,5-diethoxycarbonyl-1,4-dihydrocollidine. Mattschoss, L. A., Hobbs, A. A., Steggles, A. W., May, B. K., and Elliott, W. H. (1984), in **Microsomes and Drug Oxidations** (Boobis, A. R., Caldwell, J., De Matteis, F., and Elcombe, C. R., eds), pp. 157-164, Taylor and Francis Ltd., London.

Control of synthesis of hepatic 5-aminolevulinic acid synthase and cytochrome P450: Relationship to hepatic porphyrias. Borthwick, I. A., Srivastava, G., Hobbs, A. A., Mattschoss, L. A., Steggles, A. W., May, B. K., and Elliott, W. H. (1985) in **Cellular Regulation and Malignant Growth** (Ebashi, S., ed.), pp. 144-151, Japan Sci. Soc. Press, Tokyo/Springer-Verlag, Berlin.

Isolation and characterization of genomic clones for two chicken phenobarbital inducible cytochrome P450 genes. Mattschoss, L. A., Hobbs, A. A., Steggles, A. W., May, B. K., and Elliott, W. H. (1986) **J. Biol. Chem.** 261, 9438-9443.

The cDNA and protein sequence of a phenobarbital-induced chicken cytochrome P450. Hobbs, A. A., Mattschoss, L. A., Brian, B. K., Williams, K. E., and Elliott, W. H. (1986) **J. Biol. Chem.** 261, 9444-9449.