

SEX- AND TISSUE-SPECIFIC EXPRESSION OF DIFFERENT MEMBERS
OF THE MOUSE MAJOR URINARY PROTEIN MULTIGENE FAMILY.

IAIN McINTOSH

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

1988



To my parents.

I declare that this work is my own, except where otherwise stated,

Iain McIntosh.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. John Bishop, for giving me the opportunity to pursue the work presented in this thesis and for his patience, advice and encouragement during its course. Thanks are also due to George Newell, Denis Thoghill and Johnny Ireland in the workshop, Helen Gladstone in the library, and "Sir" Frank Johnston and Alan McEwan for their photographic expertise.

I am grateful to X. Ma and J.P. Whitaker for RNA samples, J. Lewis for pTZ vectors, and T. Spiegelberg and M. Richardson for plasmid constructs.

I would also like to express my gratitude to my friends and colleagues for their advice, help, encouragement, patience, support and tolerance (!) over the last three years, especially Raya, Ann, Pete, Cheryl, Sam, Joanne, Melville and Morag.

ABSTRACT

The major urinary proteins (MUPs) of the mouse are encoded by a large multigene family. The expression of the genes is under complex tissue-specific, developmental and hormonal control. The MUP genes have been characterised into four groups on the basis of sequence and restriction enzyme site homologies and their differential expression.

The genes show greater than 90% nucleotide sequence homology and analysis of gene expression between and within groups requires special methods. This thesis describes the development of such methods: the hybridisation of Northern blots with gene-specific oligonucleotide probes, with in vitro transcribed RNAs of known sequence as controls. This allows the discrimination of RNA species differing at a single nucleotide.

Evidence is presented that the expression of group 1 MUP genes is restricted to liver of both sexes and to the prelactational mammary gland. Although these genes have greater than 99% nucleotide sequence homology within exons, it has been possible to determine the relative abundance of the transcripts of three of them in male liver using gene-specific oligonucleotide probes. Two are also expressed at lower levels in female liver where all three are inducible by testosterone. The most abundantly expressed MUP gene in the liver is also expressed in the prelactational mammary gland.

Since the abundance of different group 1 gene transcripts can now be determined in vivo, the identification of cis- and trans-acting factors becomes possible.

CONTENTS

Chapter 1: INTRODUCTION	Page
The Control of Eukaryote Gene Expression	1
The Murine Major Urinary Protein Gene Family	
Tissue-specific and hormonally regulated expression	8
Gene structure and organisation	12
The evolution and possible function of MUPs and related proteins	21
Aims of Project	24
Chapter 2: A REVIEW OF FACTORS AFFECTING THE HYBRIDISATION OF OLIGONUCLEOTIDES TO FILTER-BOUND RNA	25
Duplex Stability	26
RNA.DNA Hybridisation	29
Special Considerations in Oligonucleotide Hybridisation	30
Chapter 3: GENERATING SPECIFIC RNA TRANSCRIPTS <u>IN VITRO</u>	
Introduction	35
Results	38
Chapter 4: THE EXPRESSION OF GROUP 1 AND GROUP 2 MUP GENES	
Introduction	44
Group 1- and Group 2-specific Probes	46
Tissue-specificity of Group 1 Gene Expression	51
Primer Extension Sequencing	65
Group 2 Pseudogenes Are Not Expressed	68
Conclusions	72

Chapter 5: SEXUALLY DIMORPHIC EXPRESSION OF INDIVIDUAL GROUP 1 GENES

Introduction	73
The Design of Gene-specific Oligonucleotide Probes	75
Hybridisation of the Gene-specific Probes	77
BL1 probe	78
BS1 probe	82
BS5 probe	86
BS6 probe	86
The Relative Expression of Group 1 MUP Genes	87
Conclusions	89

Chapter 6: DISCUSSION

DISCRIMINATION BETWEEN SIMILAR TRANSCRIPTS USING OLIGONUCLEOTIDE PROBES	91
The Stabilities of Mismatched Base-pairs	92
Estimating the T_m of a Short RNA.DNA Duplex	97
THE TISSUE-SPECIFIC AND SEXUALLY DIMORPHIC EXPRESSION OF GROUP 1 MUP GENES	
Tissue-specificity of Group 1 Gene Expression	101
Sexually-dimorphic Expression of Group 1 MUP genes	104

Chapter 7: MATERIALS AND METHODS

Animals	111
Enzymes	111
RNA Controls	111
Total RNA Samples	112
Gel Electrophoresis of RNA	112

Synthetic Oligonucleotides	113
Hybridisation of Oligonucleotide Probes	114
Labelling and Hybridisation of cDNA Probes	114
Primer Extension Sequencing	115
General Methods	115
Appendix: DETERMINATION OF THE RELATIVE AMOUNTS OF SPECIFIC MUP RNAs BY DENSITOMETRIC SCANNING	118
Abbreviations	121
References	124

Tables

	Page
1) T7 RNA Polymerase promoter sequence	37
2) T7 templates for generating RNA controls	40
3) Sequence of group 1- and group 2- specific oligonucleotide probes	47
4) The tissue distribution of group 1 MUP gene transcripts	55
5) Variation in MUP liver RNA levels between individual female mice	60
6) Sequence of gene specific oligonucleotide probes	76
7) The relative levels of group 1 MUP gene transcripts	84
8) Comparison of mismatch stability: DNA.DNA vs RNA.DNA duplexes	95
9) The estimated T_m and actual washing temperatures (T_w) of the oligonucleotide probes	98
10) Amino acid composition of group 1 MUP genes	103
11) Phylogenetic sub-groups of group 1 MUP genes	108

Figures

	Page
1) Exon structure of a MUP gene and positions of genomic sub-clones	16
2) The exonic sequences of cloned MUP genes	18
3) <u>In vitro</u> transcription with T7 RNA polymerase generates RNA of the expected size	43
4) Hybridisation of 14nt group 1- and group 2-specific oligonucleotide probes to homologous DNA and RNA	50
5) Tissue-specific expression of group 1 MUP genes	53

6) Variation in levels of MUP gene expression in the liver between individual female mice	58
7) Variation of different liver-specific mRNAs between individual female mice	63
8) Primer extension sequencing of total liver RNA	67
9) Expression of group 2 pseudogenes?	71
10) The expression of individual group 1 MUP genes	80
11) Putative association of urinary MUPs with individual clones	106

CHAPTER 1

INTRODUCTION

The Control of Eukaryote Gene Expression

A central problem in molecular biology is understanding the means by which organisms control the pattern of gene expression in different cell types throughout development. Gene expression can be regulated at various stages from transcription initiation to post-translational modification and protein transport. It appears, however, that transcriptional control is the most common (reviewed by Darnell, 1982). The correct and efficient initiation of transcription has been shown to involve interactions between cis-acting DNA sequence elements and trans-acting protein factors (Dyanan & Tjian, 1985; Serfling et al, 1985; Sassone-Corsi & Borrelli, 1986).

It has recently become apparent that the expression of genes encoding different types of protein is regulated in different ways. Those proteins which are present at low levels in most or all cell types e.g. enzymes like hypoxanthine phosphoribosyl transferase (HPRT) and dihydrofolate reductase (DHFR) appear to be the products of genes under one form of regulation, while those which are abundant in one or a few cell types (e.g. globins, insulin and albumin) are products of genes regulated in a different manner. Yet other genes which are expressed at low levels in most cell types but at high levels in one or two (e.g. metallothioneins) have some features common to both.

Regulation of transcription occurs at two stages: the frequency of transcription initiation and the site of initiation. In some systems both appear to be controlled by one set of sequence elements (GC-boxes) and a transcription factor (Sp1). This arrangement was first investigated in the SV40 early and late promoter region which consists of six GC-boxes arranged to form three 21bp repeats. Recent work has established that each GC-box in this promoter can be bound in vivo and in vitro by the cellular transcription factor Sp1 in a sequence-specific manner. Point mutations in different boxes have varying effects on transcription initiation (Gidoni et al, 1984, 1985; Barrara-Saldana et al, 1985). Steric constraints and differences in binding affinities at each site limit Sp1-DNA interactions to five out of six sites at any one time. In the late promoter, the GC-boxes alone regulate transcription initiation whereas in the early promoter they act via a TATA box (see below). GC-boxes are found upstream of a number of other genes, especially those which are expressed in most, if not all, cell types. One example is the mouse HPRT gene where the 49bp immediately 5' of the cap site appear to be sufficient to direct correct transcription (Melton et al, 1986). This region contains three GC-boxes and no TATA box. Furthermore, Dynan et al (1986) have shown that Sp1 binds to GC-boxes in, and stimulates transcription from, the promoter of the mouse DHFR gene.

Genes which are expressed only in certain cell types have different promoter structures with different elements involved in determining the site and frequency of initiation. These promoters usually lack GC-boxes and are generally more complex than those directing systemic

expression. The site of transcription initiation is defined by the position of a sequence element known as the TATA box (from its consensus sequence, TATAWAW; Breathnach & Chambon, 1981). This is found approximately 30bp upstream of the cap site of nearly all genes transcribed in a tissue-specific manner. Point mutations within the sequence often lead to heterogeneity in transcription initiation (Dierks et al, 1983; Myers et al, 1986), presumably through disruption of the preinitiation complex which forms at this site (Davison et al, 1983; Parker & Topol, 1984). The formation of the complex depends on a transcription factor known as TFIID (Sawadogo & Roeder, 1985) or BTF1 binding to the TATA box. This positions RNA polymerase II and ancillary initiation factors to direct transcription initiation approximately 30bp downstream (Zheng et al, 1987).

The frequency of initiation from such promoters is determined by elements positioned further upstream. These elements include the CCAAT box present in mammalian globin gene promoters, the chicken ovalbumin and conalbumin promoters and the herpesvirus thymidine kinase promoter (Benoist et al, 1980; Dierks et al, 1983; Jones et al, 1985; Myers et al, 1986). Mutations within this sequence affect the frequency of initiation in vivo and in vitro (Graves et al, 1986; Myers et al, 1986). The element appears to be bound by trans-acting factors present in HeLa cell and rat liver extracts (Jones et al, 1985; Graves et al, 1986; McKnight & Tjian, 1986).

Other elements present in the upstream regions of various genes confer inducibility by specific agents on heterologous promoters. For

example the glucocorticoid response elements (Yamamoto, 1985), the metal response elements of human and mouse metallothionein genes (Karin et al, 1984; Stuart et al, 1984), the heat shock elements of Drosophila (Pelham & Bienz, 1982), and the inducible elements of human α - and β -interferon genes (Ryals et al, 1985; Goodbournet al, 1985).

Not all upstream elements act in a positive manner. Dexamethasone, for example, has a negative effect on α -foetoprotein gene transcription (Huang et al, 1985) and the human β -interferon gene upstream element down-regulates transcription in the absence of an inducer (Goodbournet al, 1986). The clearest example of negative regulation of tissue-specific gene expression is provided by the phenomenon of extinction. Hybrid cells formed by fusing different cell types generally fail to express the tissue-specific products of either parent. There is evidence that specific genetic loci are involved in the trans-regulated repression of gene expression (Killary & Fournier, 1984; Chin & Fournier, 1987). These authors' data suggest that negative regulation may be a ubiquitous form of gene control.

Many upstream sequence elements share common features with enhancer elements first described in viruses of eukaryotic cells but subsequently found upstream, within and downstream of a number of eukaryotic genes (reviewed by Serfling et al, 1985; Maniatis et al, 1987). Enhancers increase the initiation of transcription from the cap site and function over long distances (> 1kb) independently of orientation. Those which have been examined in detail have been

shown to consist of separate domains bound by trans-acting factors which act independently (e.g. Wildeman et al, 1986; Zenke et al, 1986). The cell type and species specificity of viral enhancers varies but most enhancers associated with cellular genes act in a cell type-specific manner (Serfling et al, 1985; Maniatis et al, 1987).

The first ^{cellular} enhancer to be studied was found in an intron of the immunoglobulin heavy chain gene (Gillies et al, 1983) which functions only in lymphoid cells. Other tissue-specific enhancers have recently been identified, often a few kilobases upstream of the cap site. Examples include mouse albumin (10kb, Pinkert et al, 1987), chicken ovalbumin (7kb, Kaye et al, 1986) and chicken lysozyme (6kb, Theisen et al, 1986). Some have been shown to interact with trans-acting factors (e.g immunoglobulin κ -chain enhancer; Sen & Baltimore, 1986) and are associated with DNaseI hypersensitive sites (c.f. SV40, Jongstra et al, 1984). Erythroid-specific hypersensitive sites have been identified approximately 50kb 5' and 20kb 3' of the human β -globin gene. Sequences around these regions together specify complete tissue-specificity and developmental regulation of the human β -globin gene in transgenic mice (Grosveld et al, 1987). Other sequence elements within and 3' of the β -globin ^{gene} have also been shown to direct adult erythroid-specific expression in transgenic mice (Behringer et al, 1987). The cell type-specific elements of other gene promoters (which show some properties of enhancers) have been shown to be bound by as yet uncharacterised trans-acting factors e.g. insulin (Ohlsson & Edlund, 1986) and α_1 -antitrypsin (DeSimone et al, 1987).

A common feature of many upstream elements and enhancers is that they are present in two or more copies or consist of repeated sequences and domains, some of which are present in other promoters. For example, the mouse metallothionein I promoter contains five metal response elements of which a minimum of two are required to confer full metal inducibility on a heterologous promoter (Stuart et al, 1984, 1985; Searle et al, 1985). Further examples include the upstream repeat element in mammalian β -globin genes (Dierks et al, 1983; Myers et al, 1986), heat shock elements 5' to the Drosophila hsp70 gene (Dudler & Travers, 1984) and duplicated glucocorticoid receptor binding sites in the chicken lysozyme promoter and mouse mammary tumour virus long terminal repeat (Scheidereit et al, 1983; Renkawitz et al, 1984; Scheidereit & Beato, 1984).

In summary, the promoter regions of eukaryotic genes whose transcription is restricted to certain cell types or certain periods of development (rather than occurring in all cell types) consist of three basic cis-acting elements:- (1) the TATA box which mediates the site of initiation, (2) upstream elements which are involved in induction and/or tissue-specificity, and (3) enhancer elements which enable high frequencies of transcription initiation in a certain cell type. Some sequences have characteristics of both classes (2) and (3). Many elements have been shown to be bound by various trans-acting factors and are associated with alterations in chromatin structure (as measured by DNaseI hypersensitivity). It has been argued that the initiation of transcription requires interactions between DNA-bound proteins; and this has indeed been demonstrated in

prokaryotes and yeast (Ptashne, 1986; and references therein). Ptashne argues that enhancers are able to act at a distance because the DNA between the enhancer and proximal promoter elements is looped out allowing the interaction between proteins bound at each site.

It is sometimes possible to gain a greater understanding of the factors affecting gene expression by examining multigene families because subtle differences which lead to changes in regulation are not obscured. In certain cases the pattern of expression within a multigene family varies during development in different tissues and under the control of different inducers. Examples include the tissue-specific expression of different mouse kallikrein genes (Mason et al, 1983; van Leeuwen et al, 1986), varying patterns of globin synthesis during development, and the induction of metallothionein genes by glucocorticoids and heavy metals (Karin et al, 1984).

The gene families encoding the rodent urinary proteins appear to be regulated in all three ways. One particularly interesting aspect of their regulation is the differential induction of gene expression in the liver by testosterone and non-steroid hormones. The genes encoding the murine major urinary proteins (MUPs) have been studied in a number of laboratories over many years and it is an investigation of gene expression within this family which is presented in this thesis.

The Murine Major Urinary Protein Gene Family

Tissue-specific and hormonally regulated expression

The major urinary proteins (MUPs) of the mouse are a set of small (19,000 MW), acidic proteins (pI 4.0-4.6). They are synthesised mainly in the liver, but also in the lachrymal, submaxillary, parotid, sublingual and mammary glands and not in the kidney, testes, seminal vesicles or brain (Hastie et al, 1979; Shaw et al, 1983). MUPs excreted in the urine are synthesised in the liver, secreted into the serum and concentrated in the kidneys (Rumke & Thung, 1964; Finlayson et al, 1965). The in vitro translation products of hybrid-selected liver RNA have a pattern very similar to that of the urinary proteins (Clissold & Bishop, 1982; Knopf et al, 1983; Clissold et al, 1984). Expression in other tissues has been determined at the RNA level and not by the presence of proteins. In vitro translation products of hybrid-selected lachrymal gland RNA constitute a more basic set of proteins (Shaw et al, 1983; Shahan & Derman, 1984). MUP mRNA is the most abundant in the male liver (approximately 8%, Derman et al, 1981), where it is at least five times more abundant than in the female (Hastie et al, 1979; Derman, 1981). Female liver mRNA gives a simpler pattern of proteins after in vitro translation but a male-like pattern is observed, along with male-like levels of mRNA, in females treated with testosterone (Clissold et al, 1984).

The RNA and in vitro translation products have been examined from surgically altered (thyroidectomised and hypophysectomised females) and mutant (little and testicular feminisation) mice. These experiments showed that testosterone, growth hormone and thyroxine induce MUP mRNA synthesis in the liver, and that the action of testosterone in this tissue is at least partly dependent on growth hormone and thyroxine (Knopf et al, 1983). Different MUP species are induced to different extents by different combinations of hormones. Further evidence for the effect of growth hormone and thyroxine comes from experiments with cultured hepatocytes.

Primary hepatocytes show a pattern of gene expression similar to regenerating liver i.e. increased levels of transferrin RNA, the presence of α -foetoprotein RNA, and decreased transcription of adult liver RNA precursor species such as albumin and MUP. The rate of decay of MUP RNA transcription in hepatocytes can be slowed by culturing in the presence of growth hormone, thyroxine and insulin. The effects of growth hormone and thyroxine are synergistic (Spiegelberg & Bishop, 1988).

It has been proposed by Norstedt & Palmiter (1984) that testosterone acts on the pituitary to induce a discontinuous pattern of growth hormone production in males, compared with continuous production in females, and that this determines the sexual dimorphism of the liver. They further suggest that discontinuous stimulation of growth hormone receptors causes a discontinuous release of insulin-like growth factor I (IGF-1) which in turn induces MUP synthesis. Continuous stimulation of growth hormone receptors, on the other hand, up-regulates growth

hormone/prolactin receptors resulting in continuous IGF-1 release in turn down-regulating IGF-1 receptors and resulting in low MUP production.

Derman and co-workers have shown that the different MUP genes are expressed in different tissues by restriction analysis of uncloned MUP cDNA (Shahan & Derman, 1984) and by cloning and sequencing of MUP cDNAs from liver, lachrymal and salivary glands (Shahan et al, 1987a,b). Shaw et al (1983) had previously shown that the different MUP expressing tissues are under developmental control. Liver MUP mRNA is first detected at three weeks of age, but maximum levels are reached when animals are sexually mature at about six weeks. In mammary glands MUP mRNA is first detected during the first pregnancy. A cDNA cloned from a prelactational mammary gland library was found to be identical in sequence to one from a female liver library (Shahan et al, 1987b). Although the murine submaxillary gland is sexually dimorphic in the expression of some genes (e.g. kallikreins, van Leeuwen et al, 1987), this is not the case with MUP genes. MUP mRNA is first detected in this tissue at one week of age, peaks between 4 and 7 weeks, and then decreases. No hormonal regulation of MUP expression in the salivary glands has been demonstrated. MUP expression in the lachrymal gland is observed at two weeks of age when this gland can first be dissected. The level is about five times greater in males than in females, and testosterone treatment induces male-like levels in females. Expression in the lachrymal gland appears to be independent of both growth hormone and thyroxine (Shaw et al, 1983).

Recent experiments with transgenic mice have shown that 2kb of Mup 5' flanking sequence is sufficient to direct expression of the Herpes simplex virus thymidine kinase (HSV-tk) reporter gene in male liver and enables testosterone induction in female liver. There is no detectable expression in the submaxillary or lachrymal glands, nor in heart, muscle, brain or spleen (where no MUP gene expression is detectable, Hastie et al, 1979). There is, however, abundant expression in the male preputial gland and testes (Al-Shawi et al, 1988b). The rat analogue of MUP, α_{2u} -globulin, has been shown to be expressed at high levels in the preputial glands of male and female rats (Held & Gallagher, 1985; MacInnes et al, 1986) but not in the testes (R.A. Al-Shawi, unpublished results). It is thought that the high levels of thymidine kinase activity in the testes could be the cause of the sterility observed in male transgenic mice (Al-Shawi et al 1988b). The reason for the aberrant expression in these tissues is not known; it may be due to the removal of negative regulatory elements further upstream, within the gene itself or in the 3' flanking sequences. These possibilities are currently being investigated using alternative reporter genes and different lengths of 5' flanking sequence. Mice transgenic for an α_{2u} -globulin genomic clone including approximately 2kb of upstream sequence express α_{2u} -globulin in adult male liver and preputial gland. Expression can be induced by glucocorticoid treatment of adrenalectomised mice or testosterone treatment of ovariectomised mice (Da Costa Soares et al, 1987). The upstream sequences present in the transgene include the elements identified by Addison & Kurtz (1986) as being involved in glucocorticoid induction of α_{2u} -globulin gene expression in L cells. This suggests that these elements may also be involved in the

glucocorticoid regulation of $\alpha_2\mu$ -globulin in rat liver first demonstrated by Kurtz & Feigelson (1978).

Gene structure and organization

MUPs are encoded by a family of approximately 35 genes located on chromosome 4. Most MUP genes can be classified into two main groups (1 & 2) on the basis of hybridisation criteria (Bishop et al, 1982), and are linked pair-wise in a head-to-head manner about 15kb apart. With the 3' flanking regions each pair constitutes a 45kb imperfect palindrome and these are the units of duplication and evolution (Clark et al, 1984b; Bishop et al, 1985). It is estimated that there are thirteen to fifteen 45kb units and a further five to nine genes which do not fall into either group 1 or group 2. A group 1 gene (BS6) and a group 2 gene (BS2) have been sequenced throughout the entire transcription unit as well as 5' and 3' flanking regions. The data suggest that the group 2 gene is a pseudogene relative to the group 1 gene: it contains two in frame stop codons and a frame-shift mutation. There are also deletions upstream of the TATA box and in intron 3 and a base-change within the TATA box (Clark et al, 1985a). The TATA box of group 1 genes has been shown to direct transcription from a cap-site 30±1 bp downstream by primer extension and S1 analysis (Clark et al, 1985a) and by transformation of fibroblast cells (Held et al, 1987). Four group 2 genes have been sequenced to date and share a common base substitution in the first exon which generates a termination codon at the seventh position of the mature polypeptide (Ghazal et al, 1985). A synthetic oligonucleotide probe homologous to this region of

group 2 genes hybridises to 12 of 34 MUP genomic clones (Shahan et al, 1987a). It appears therefore, that most if not all group 2 genes are pseudogenes.

Of those genes outwith groups 1 and 2, some have been placed in two further groups (3 and 4) on the basis of nucleotide sequence homology (Al-Shawi et al, 1988a). Group 3 consists of two sequences, a liver cDNA clone (p199, Kuhn et al, 1984; => MUP15, Clark et al, 1985b; => cLIV4 => cLIV7, Shahan et al, 1987a) and a genomic clone (MUP16; Ma, 1987). Each differs from group 1 genes in exonic sequence by approximately 15% and from each other by 5%. Kuhn et al (1984) identified a potential glycosylation site in the sequence of the cDNA clone p199, and have shown that mRNA hybrid-selected with p199 codes for a group of four proteins which run more slowly in SDS-polyacrylamide gels than most MUPs. It was subsequently shown that mouse urine contains a minor MUP fraction which binds to Con-A-sepharose and is detected by a stain specific for glycosylated proteins (Clark et al, 1985b; Ma, 1987). MUP16 also encodes a protein with the same glycosylation site. It differs from all MUP genes cloned to date because it entirely lacks exon 7 and has a large 3' deletion.

Group 4 comprises cDNA clones from lachrymal and submaxillary gland libraries (one each: cLAC1 and cSMX1, Shahan et al, 1987a) and two partial genomic clones (BL2 and CL12, Clark et al, 1982; Al-Shawi, 1985; Al-Shawi et al, 1988a). The cDNA clones differ in nucleotide sequence from group 1 genes by approximately 10%, from group 2 genes by 12%, and from each other by 7%. The protein encoded

by cLAC1 would be considerably more basic than any MUP encoded by a group 1 or group 3 gene (from a comparison of the amino acid sequences predicted by nucleotide sequences, see chapter 6). It is not known if group 3 or group 4 genes are located in 45kb palindromes.

The MUP transcription unit is 3.9kb in length and contains 7 exons (Clark et al, 1984a; Figure 1), the last of which is non-coding.

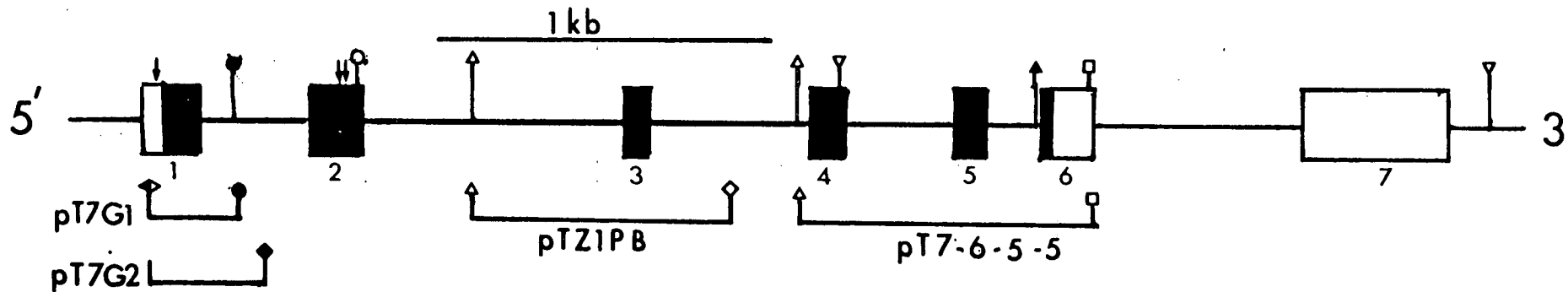
There are alternative splice sites within the non-translated part of exon 6 giving rise to three different splicing configurations. Two different lengths of MUP liver mRNA can be resolved by gel electrophoresis and the shorter, less abundant, of these (including all of exon 6 and entirely lacking exon 7) hybridises preferentially to a group 2 probe whereas the longer one (lacking all but the 5', coding, portion of exon 6 and including all of exon 7) hybridises preferentially to a group 1 probe (Al-Shawi, 1985; Clark et al, 1984a). The larger message is 10-20 fold more abundant in male mouse liver. From the available sequence data it was concluded that the shorter message is not the transcript of MUP16, a genomic clone lacking exon 7 (see above).

The exonic sequences of four genomic clones of group 1 genes have been determined (Clark et al, 1985b) as well as the sequence of a number of near full length liver cDNA clones from BALB/c mice (Kuhn et al, 1984; Clark et al, 1985b; Chave-Cox, 1986; Shahan et al, 1987a; compared in Figure 2). These clones encode proteins with an 18 amino acid signal peptide and a 162 amino acid mature protein with an expected molecular weight of approximately 19,000 Daltons. The nucleotide sequence homology between the different group

Figure 1

Exon structure of a Mup gene and positions of genomic sub-clones

The exon structure and restriction sites as determined by Clark et al (1984a) are shown. Exons are indicated by boxes, coding regions are shaded. The positions of genomic sub-clones used to generate the control transcripts described in Table 2 are indicated beneath the gene. The 5' ends of the cDNA clones used as templates are indicated by vertical arrows above the gene in exons 1 & 2 :- MUP15, MUP11, MUP8; from 5' to 3' (see Figure 2).



Key

- | | | | |
|---|--------|---|----------|
| ● | Bam HI | ◻ | Hind III |
| ◻ | Eco RI | ◆ | Bgl II |
| ∇ | Pst I | ◇ | Bst NI |
| ↑ | Pvu II | ⤴ | Sau 3A |
| ↑ | Kpn I | | |

Figure 2

The exonic sequences of cloned Mup genes

The exonic sequences of BALB/c genomic and cDNA clones were aligned and a consensus derived using the UWGCG programme, PRETTY. Sequences missing from various clones are shown by dots. The consensus for the 3' portion of exon 6 included in short message but not long (Clark et al, 1984a) is shown in lower case. The target sequences for gene or group specific probes (described in Tables 3 & 6) are underlined.

BL1 was isolated from a BALB/c genomic liver library (Clark et al, 1982, 1985b), BS1, BS5, BS6, MUP16 from a BALB/c genomic sperm library (Clark et al, 1982, 1985b; Ma, 1987), MUP8, MUP11, MUP15 and cLIV6, cLIV7 from separate BALB/c female liver libraries (Clark et al, 1985b; Shahan et al, 1987a), p1057 and cLIV1 from separate BALB/c male liver libraries (Kuhn et al, 1984; Shahan et al, 1987a), cLAC1 from a BALB/c lachrymal gland cDNA library (Shahan et al, 1987a) and cSMX1 from a BALB/c submaxillary gland cDNA library (Shahan et al, 1987a).

The termination point of MUP16 at the end of exon 6 is that predicted from the observation that the genomic clone lacks exon 7 (Ma, 1987).

1
 BLIRNA.
 cLIV6RNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI5RNA. g t c c CT6CT6CT6C
 NUPI6RNA g c a c t a ga t c t C
 cLACIRNA. g c c
 cSMXIRNA. g tg tc t tc a g t t c t a
 BS2RNA. a g ac c c t t ag ca CT6CT6 CT6CT6CT6C
 Consensus GGAGTGTAGC CACGATCACA AGAAAGACGT GGTCTGACA GACAGACAAAT CCTATICECI ACCAAAATGA AGATGCTGCT GCTG.....

101
 BLIRNA.
 cLIV6RNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.ag
 NUPI5RNA. TC a tt a t t a c g t t tt t
 NUPI6RNA. TG ta a g ta t c gc t t tt t
 cLACIRNA. a t c c a c a a a tt t
 cSMXIRNA. a a ga a ca aa tt t
 BS2RNA. TG a g t t a a a t c
 Consensus ..CT6T8TTT G6GACTGACC CTAGTCTGTG TCCATGACGA ABAAGCTAGT TCTAC666AA G6AACTTTAA T6TAGAAAAG ATTAATG666 AATG6CATAC

201
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA. gctgaa t tg a c g a g gc g aa c ac ct
 cLIV7RNA. ct
 NUPI6RNA. gctgaa tg a c g a g gc g aa c ac ct
 cLACIRNA. c a c g a g g g c c ct
 cSMXIRNA. c a c c a g g g c g ct
 BS2RNA. t a c g a g g c t ct
 Consensus TATTATCCTG GCTCTGACA AAAGAGAAA GATAGAAGAT AATG6CACT TTAGACTTTT TCTGGAGCAA ATCCATGTCT TGGAGAATC CTTAGTCTT

100
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

200
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

300
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

Consensus

301
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

401
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

501
 BLIRNA.
 cLIV6RNA.
 NUPBRNA.
 BSIRNA.
 cLIVIRNA.
 p1057RNA.
 BSSRNA.
 BS6RNA.
 NUPI1RNA.
 NUPI5RNA.
 cLIV7RNA.
 NUPI6RNA.
 cLACIRNA.
 cSMXIRNA.
 BS2RNA.
 Consensus

Consensus

1 genes is greater than 99% but they specify different proteins. Genomic clones from C57BL mice have been used to stably transform L-cells and the MUPs obtained co-migrate on two-dimensional gels with species of hybrid-selected, in vitro translated liver RNA (Held et al, 1987).

Comparison of the 5' flanking sequence of a number of MUP genes has identified a hypervariable A-rich region 15bp upstream from the TATA box (Held et al, 1987; Al-Shawi et al, 1988a). The length and base-composition of this region varies between different genes. The group 1 sequences can be divided into four sub-groups on the basis of the length and composition of this A-rich region and these coincide with those defined by parsimonies based on conservation of restriction sites (Al-Shawi et al, 1988a). Possible relationships between the expression pattern of different group 1 genes and their phylogenetic separation are discussed in chapter 6.

Approximately 10bp upstream of the A-rich region is a sequence, CCATAC, which shows limited homology to the consensus for the CAAT box (Ghazal, 1986; Held et al, 1987). The promoter regions of MUP genes contain a number of sequences showing homology to the published consensus for the glucocorticoid regulatory element (GRE), metal regulatory element (MRE), enhancer core sequences and nuclear factor 1 binding sites. Of these only that for the glucocorticoid regulatory element occurs more frequently than expected by chance allowing for sequence length and base composition (Ghazal, 1986). Two such consensus sequences are observed 10bp apart between 360bp and 380 bp upstream from the cap site of the group 1 genes analysed but it is not

known whether they are functional.

The evolution and possible function of MUPs and related proteins

Comparison of amino acid sequences relates the rodent urinary proteins, MUP and α_{2U} -globulin, to a number of mammalian secretory proteins:- human and rat retinol binding protein, human and rat α_1 -acid glycoprotein, human apolipoprotein D, human protein HC (high cysteine), a rat androgen-dependent epididymal secretory protein, ovine β -lactoglobulin and bovine odorant-binding protein (Ali & Clark, 1988; Sawyer, 1987; and references therein). Tobacco hornworm insecticyanin and frog Bowman's gland protein also share homology. Common three-dimensional protein structures have been demonstrated for some of these proteins and the arrangement of exons and introns is very similar in those cases where this information is available. It has been suggested that these proteins bind small molecules which are either insoluble at physiological pH, labile, or both (Sawyer, 1987). For example retinol binding protein binds retinol and transports it in the serum, and β -lactoglobulin has also been shown to bind retinol and may carry the vitamin in milk. Bovine olfactory binding protein has been shown to bind a variety of small hydrophobic odorant molecules as well as pyrazines (Pevsner et al, 1986) and α_{2U} -globulin is also reported to bind pyrazines in the nasal mucosa (A. Cavaggioni, unpublished results).

Male mouse urine contains androgen-related fractions which accelerate the onset of puberty in young females (Vandenbergh et al, 1975).

One of these is probably a protein of greater than 12,000 MW and the

second appears to be a mixture of oligopeptides of about 860 MW (Vandenbergh et al, 1976). Clark et al (1985a) have suggested that the protein is MUP and that the six N-terminal amino acids (N-Glu-Glu-Ala-Ser-Ser-Thr) are biologically active whilst the smaller, dialysable factor is the hexapeptide (N-Glu-Glu-Ala-Arg-Ser-Met) encoded by group 2 genes. Shaw et al (1983) have suggested that MUPs may act as carriers for active behavioural cues or pheromones. The high degree of conservation within the MUP and α_{2U} -globulin families could provide the variation and specificity required for binding different molecules.

The tissue distribution of α_{2U} -globulin synthesis is broadly similar to that of MUP, but differs in some details. Iso-electric focussing and Western blotting of S100 extracts from various tissues reveals a smaller number of proteins in male liver, a few more acidic proteins in submaxillary and lachrymal glands, and a large number (10-15) in both male and female preputial gland and in the mammary gland at 18 days gestation (MacInnes et al, 1986). α_{2U} -globulin RNA is undetectable in female liver (Gubits et al, 1984).

The 5' flanking sequences of α_{2U} -globulin genes lack a hypervariable A-rich region with a run of only seven A residues at the analogous position (Winderickx et al, 1987). The group 3 gene, MUP16, also has only seven A residues at this point suggesting that this arrangement is ancestral and the expansion of the array is specific to mouse (Al-Shawi et al, 1988a). There is, however, some preliminary evidence that α_{2U} -globulin genes are arranged in

a head-to-head manner similar to MUP genes (J.P. Whitaker, unpublished results). An $\alpha_2\text{U}$ -globulin pseudogene has been cloned and sequenced, but it does not have the same point mutation as group 2 MUP genes and has not been mapped relative to the other genes rat on chromosome 5 (Winderickx et al, 1987).

Aims of Project

The multigene family encoding the major urinary proteins of the mouse is an excellent model system in which to examine the factors affecting the regulation of gene expression in the mammalian liver. Testosterone, growth hormone and thyroxine are thought to affect the levels of expression of different genes within the family. In order to identify the cis-acting elements and trans-acting factors involved in this regulation a means of discriminating between the very similar transcripts is required.

The major aim of the work presented in this thesis was to develop a method by which the transcripts of individual group 1 MUP genes could be identified. The method chosen employed synthetic oligonucleotides as hybridisation probes to Northern blots using in vitro RNA transcripts of known sequence as controls. Once the levels of transcript from a given gene can be measured independent of others, it is possible to determine which genes are controlled by particular hormone regimes. This will enable the study of the putative regulatory regions in transgenic mice.

CHAPTER 2

A REVIEW OF FACTORS AFFECTING THE HYBRIDISATION OF OLIGONUCLEOTIDES TO FILTER-BOUND RNA

The experiments described in this thesis involve the hybridisation of group- or gene-specific oligonucleotide probes to filter-bound RNA. Such experiments have only recently become feasible due to advances in the chemistry of oligonucleotide synthesis which have made oligonucleotides readily available to the molecular biologist (reviewed by Itakura et al, 1984; Caruthers, 1985). Wallace et al (1979, 1981) have shown that it is possible to differentiate between DNA sequences differing at only one nucleotide using oligonucleotide probes and this, for example, allows one to identify the presence of mutant alleles for diagnostic purposes (Conner et al, 1983; Kidd et al, 1983; Studencki et al, 1985; DiLella et al, 1986).

Synthetic oligonucleotides have not been used extensively as probes to RNA targets. Szostak et al (1977) used a 15nt probe to yeast cytochrome c mRNA, and Capetanaki et al (1983) showed that a 14nt probe would hybridise to chicken vimentin mRNA. Synthetic oligonucleotides have also been used as primers for the generation of specific cDNAs (Houghton et al, 1980), direct sequencing of RNA using dideoxynucleotide triphosphates and reverse transcriptase (e.g. Epstein et al, 1986), and differentiating between similar transcripts (Field & Gross, 1985). During the course of the work presented in this thesis there have been a few reports on the use of

oligonucleotide probes to discriminate between similar transcripts. For example, Omiecinski et al (1985) used synthetic probes to examine differential induction of rat cytochrome P450s and van Leeuwen et al (1986, 1987) have used this approach to examine the tissue-specific expression of mouse kallikrein genes. In both cases the oligonucleotides were designed to hybridise to hypervariable regions within predicted RNA sequence containing 4 or 5 ^{mismatches} _^. Recently 15nt probes have been shown to discriminate between synthetic transcripts differing at a single nucleotide (Buvoli et al, 1987).

In this chapter I review aspects of nucleic acid hybridisation and duplex stability which have to be considered when designing and using oligonucleotide probes.

Duplex Stability

The stability of a nucleic acid duplex, either DNA.DNA, RNA.DNA or RNA.RNA, is usually quantified by its melting temperature (T_m). This is the temperature at which half the molecules are dissociated or denatured - the more stable a duplex, the higher its T_m . There are four factors which affect the T_m of a duplex:-

- 1) Salt concentration
- 2) Base composition
- 3) Fragment length
- 4) Degree of homology

1) Salt concentration

The two strands of a duplex repel each other in the absence of a cation, owing to the highly charged phosphate groups. The stability of a duplex, as measured by the T_m , increases linearly with the logarithm of monovalent cation concentration from 0.3mM to 0.5M (Dove & Davidson, 1962). At higher concentrations ($> 1M$) the logarithmic plot levels and the T_m is maximal between 1.0 and 2.0M M^+ (Schildkraut & Lifson, 1965). Divalent cations have an approximately 100 fold greater stabilising effect than Na^+ (Thomas, 1954); hence the inclusion of EDTA in hybridisation experiments to remove Mg^{2+} ions.

2) Base composition

Marmur & Doty (1959) established that a higher G+C content increases the thermal stability of a duplex - as might be expected from the greater number of hydrogen bonds. Subsequent work showed that T_m increased by $0.41^\circ C$ for each 1% rise in G+C content (Marmur & Doty, 1962). When combined with data on the effect of salt concentration, the Schildkraut & Lifson (1965) relation is obtained:-

$$T_m = 16.6 \log[M^+] + 0.41(\%G+C) + 81.5^\circ C \quad (1)$$

where M^+ = monovalent cation concentration.

3) Fragment length

This is a particularly important factor in the hybridisation of oligonucleotides. Since they form fewer hydrogen bonds, short DNA

molecules will form less stable duplexes. Thomas & Dancis (1973) have estimated that the T_m of a duplex is reduced by $(820/\lambda)^\circ\text{C}$, where λ = fragment length. These authors were thus able to modify the Schildkraut & Lifson relation to give:-

$$T_m = 16.6 \log[M^*] + 0.41(\%G+C) + 81.5 - (820/\lambda)^\circ\text{C} \quad (2)$$

4) Degree of homology

It is estimated that 1% mismatch reduces the T_m by about 1°C (Bonner et al, 1973). Adding this to equation (2) we get:-

$$T_m = 16.6 \log[M^*] + 0.41(\%G+C) + 81.5 - (820/\lambda) - (100-h)^\circ\text{C} \quad (3)$$

where h = % homology.

The rate of hybridisation is maximal $10-30^\circ\text{C}$ below the T_m (depending on the nucleic acid species involved). The same four factors are also involved in the rate of reaction although the relationships are not the same. The rate of hybridisation increases with salt concentration, is proportional to the square root of fragment length, is reduced by mismatching and is minimally dependent on base composition but is inversely proportional to the complexity (Wetmur & Davidson, 1968). These authors have shown that it is the initial formation of a short specifically base-paired region (nucleation) which limits the rate of hybridisation and that the formation of successive base-pairs thereafter (zippering) is rapid.

RNA.DNA Hybridisation

The points made in the previous section have been generalised to include both DNA.DNA and RNA.DNA hybridisation. Most early experiments studied the denaturation and renaturation of DNA duplexes (Marmur & Doty, 1961, 1962; Dove & Davidson, 1962; Schildkraut & Lifson, 1965). Subsequent work showed that RNA.DNA duplexes were affected in a similar, though not identical, manner (Nygaard & Hall, 1964).

Chamberlin (1965) demonstrated that duplexes of homopolymers have different thermal stabilities depending on whether one or other strand consists of ribo- or deoxyribonucleotides. For example, the stabilities in a G:C series are as follows:-



With A:U and A:T duplexes however, the order is



The low stability of the dA:rU base-pair may partly explain the observed instability of RNA.DNA hybrids relative to DNA duplexes.

The kinetics of RNA.DNA hybridisation are complex when both species are free in solution but can be greatly simplified if a large excess of DNA is employed, its renaturation is prevented by binding to a filter or if the DNA is single-stranded. Using this method it has been shown that the hybridisation of RNA to DNA is a slower reaction than DNA renaturation with an optimum temperature closer to the T_m , although the T_m is itself lower (Bishop, 1972).

Special Considerations in Oligonucleotide Hybridisation

Two major uses of synthetic oligonucleotide hybridisation probes are for the identification of cloned genes using a probe designed from a known amino acid sequence (reviewed by Lathe, 1985) and for discriminating between very similar sequences (e.g. Studencki et al, 1985). It is only possible to discriminate between similar sequences if hybridisation and washing conditions are sufficiently stringent. Hybridisation in high salt at a temperature well below T_m (i.e. conditions of low stringency) allows a probe to pair with similar but not identical sequences. This is useful when identifying related sequences in library screens or for site-directed mutagenesis. The stringency of hybridisation is most easily adjusted by changing the temperature. The work of Wallace and co-workers has shown that oligonucleotides can discriminate between sequences differing at a single nucleotide only when the hybridisation is sufficiently stringent (Wallace et al, 1981; Conner et al, 1983; Kidd et al, 1983). A less stringent hybridisation followed by a stringent wash allows some cross-hybridisation but a stringent hybridisation followed by a non-stringent wash permits discrimination between alleles (see Fig. 2 in Kidd et al, 1983). Note that such stringent hybridisations are at temperatures somewhat above those which would normally be employed (i.e. fairly close to T_m rather than 10-30°C below T_m). Although this lowers the rate of hybridisation, the rate is nevertheless very high due to the lack of any probe reassociation and the high mobility of the short probe.

The ability to discriminate between closely related sequences also relies on the increased specificity conferred by short fragment length. The stability of any duplex is partly dependent on the homology between strands (see above) and as the percentage of mismatched base-pairs in a duplex increases, the stability is lowered. If a single base-pair is mismatched, the percentage mismatch is obviously increased if the region of complementarity is short. In this case the length has a major effect on the T_m of a duplex (equation (3)). The length factor in this equation ($820/\lambda$) is only an average value (Thomas & Dancis, 1973) and there is no clear evidence that the relationship can be extrapolated to molecules 10-30 nucleotides in length, although some authors have assumed this (e.g. Lathe, 1985). Michelson & Monny (1967) have shown that a double reciprocal relationship exists between fragment length and T_m ($1/\lambda \propto 1/T_m$; T in Kelvin). This was determined using homopolymers and short oligonucleotides in low ionic strength buffers and may not apply in practice (random sequence and high ionic strength).

It should be noted that the Schildkraut & Lifson relation (as modified in equation (3)) can only be taken as an approximation for any hybridisation but this is especially true for oligonucleotide probes. As discussed in the preceding paragraph, the length effect has not been estimated under realistic experimental conditions. Oligonucleotide hybridisations are usually performed at high salt concentrations (e.g. $6 \times$ SSC, $\sim 1M$ Na^+) outwith the range of concentrations over which the Schildkraut & Lifson relation was determined. Errors introduced by assuming that the relation holds are, however, minimised by the logarithm. An empirical relationship

has been described for oligonucleotides of 11-20 bases in length in 1M

Na⁺ (Suggs et al, 1981):-

$$T_m = 2a + 4b \text{ } ^\circ\text{C} \quad (4)$$

where a = no. of A:T pairs

b = no. of G:C pairs.

Values obtained from this agree closely with those derived using equation (3) within the limits imposed. As mentioned above, the dA:rU base-pair is less stable than the dA:dT base-pair and this is believed to contribute to the reduced stability of RNA.DNA hybrids compared with the corresponding DNA duplex. With short probes this would reduce the T_m by more than is allowed for by either equation (3) or (4) when the G+C content is low (see chapters 4 & 6). In a study comparing synthetic oligonucleotide probes with nick- translated probes Berent et al (1985) used the simple equation of Suggs et al (equation (4)) to predict T_m and successfully identified the target transcript (rat 18S rRNA). A probe with a one base mismatch failed to hybridise under the same conditions. It should be noted however, that this 17 nucleotide probe was 59% G+C and included only 3 A residues. It is not clear if the relation would hold for a different base composition.

It is important to consider the complexity of the target sequences when deciding on what length of oligonucleotide probe to use. The shorter a probe, the greater the percentage mismatch to related but not identical sequences, thereby giving increased specificity. However, the probability of hybridising to unrelated sequences purely by chance also increases with decreasing size. Therefore the number

4 \times (λ = probe length) must be greater than the complexity of the sequences being studied (e.g. $\lambda > 13$ for yeast DNA, $\lambda > 17$ for mammalian genomic DNA).

When synthetic oligonucleotides designed from known amino acid sequences are used to identify cloned DNA sequences it is common to include a number of probes in a single hybridisation to allow for the degeneracy of the genetic code. This may result in alteration of the base composition. It has been reported that the use of 3M tetramethylammonium chloride (TEMAC) in filter washes (Wood et al, 1985; DiLella, et al, 1986) and hybridisation (Vries et al, 1986) can alleviate this problem. This compound has been shown to increase the stability of dA:dT base-pairs to that of dG:dC base-pairs (Melchoir & Von Hippell, 1973). It is not known if TEMAC has a similar effect on dA:rU base-pairs or if it stabilises mismatches.

Because oligonucleotide probes are short and single-stranded they can be used to probe DNA fixed in agarose gels. Although this avoids having to transfer the DNA to nitrocellulose or nylon membranes it may allow some cross-hybridisation due to inefficient washing as well as giving high levels of non-specific background (see Kidd et al, 1983; DiLella et al, 1986).

Most of the research into nucleic acid hybridisation and duplex stability has been based on experiments with both species free in solution. The work described in this thesis involves hybridising oligonucleotide probes in excess to filter bound RNA. The effects of immobilising one species on a filter have not been examined in detail.

It is known, however, that excess filter-bound DNA simplifies RNA-DNA hybridisation kinetics (see Bishop, 1972). It is believed that the major effect of using a filter-bound target (e.g. Southern or Northern transfers) is to allow excess double-stranded probe to reassociate in solution and reduce the rate of hybridisation (Anderson & Young, 1985, and references therein). Using single-stranded oligonucleotide probes there is no reassociation and so the factors affecting the rate of hybridisation are the probe concentration and the amount of target. The rate is proportional to both. There is no evidence that a filter-bound duplex is more or less stable than one in solution.

CHAPTER 3

GENERATING SPECIFIC RNA TRANSCRIPTS IN VITRO

Introduction

The major part of this thesis is concerned with investigating whether a particular RNA sequence is present in a family of sequences differing at only one or two nucleotides in a 900nt molecule. The approach taken involves hybridising synthetic oligodeoxyribonucleotides to filter bound RNA under conditions which allow one to distinguish between a perfect duplex and one containing a 1bp mismatch. The kinetics of RNA.DNA hybridisation and the stability of the resulting duplex are quite different from those pertaining to DNA.DNA reactions (Walker, 1969; Bishop, 1972; Smith, 1983; see chapter 2). In order to determine RNA.DNA hybridisation conditions of maximum specificity, RNA species of defined sequence are required as positive and negative controls for each oligonucleotide probe.

In vitro systems which generate relatively large amounts of specific RNA species have recently become available (Melton et al, 1984; Tabor & Richardson, 1985). These systems employ the promoter specificity of certain bacteriophage RNA polymerases to generate transcripts of DNA sequences cloned in plasmid vectors. RNAs generated by this method are biologically active and have been used to study the mechanisms of RNA splicing (Green et al, 1983) and translation (Krieg & Melton, 1984) and to generate mutated proteins (Kumar et al, 1986; Mead et al, 1986). The two commonly used

enzymes are those from the Bacillus typhimurium LT2 bacteriophage, SP6, and the E. coli bacteriophages T3 and T7. The T7 and SP6 RNA polymerases have been used in studies of polymerase-promoter interactions and shown to function in vitro without the addition of transcription factors (Chamberlin & Ring, 1973a; Butler & Chamberlin, 1982).

Since T7 RNA polymerase was used in the experiments described below it is the properties of this enzyme that are reviewed here. Initial studies showed that the enzyme requires only nucleoside triphosphates, a divalent metal ion, and a suitable template to generate RNA in vitro. A sulphhydryl reducing agent (β -mercaptoethanol or DTT) was found to be necessary for optimal transcription and salt concentrations greater than 50mM were found to inhibit the enzyme (Chamberlin & Ring, 1973a,b).

The seventeen T7 RNA polymerase promoter sequences have been determined and found to be similar. A 23bp consensus sequence exists from -17 to +6 of the transcription start site (Dunn & Studier, 1983; Table 1). The specificity of T7 RNA polymerase for its promoter is not absolute and the binding of the enzyme to its promoter is extremely weak. Chamberlin & Ring (1973a) report 50% efficient transcription of DNA from the related bacteriophage T3 and 30% from salmon sperm DNA but only minimal (less than 5%) amounts from bacterial and other bacteriophage DNAs. It is possible that the higher complexity of vertebrate DNA has included fortuitous promoter sites. Smeekens & Romano (1986) have found non-specific binding of plasmid DNA to the enzyme under conditions of low ionic strength.

This does not imply, however, that this binding necessarily leads to transcription. Evidence has been presented that a conformational change occurs after about 10 nucleotides have been incorporated resulting in the formation of a stable transcription complex (Ikeda & Richardson, 1986; Morris et al, 1987). Point mutations within the promoter sequence can greatly reduce the binding of T7 RNA polymerase and increase that of T3 RNA polymerase. In particular, the base at position -11 is critical in determining specificity (Morris et al, 1987).

Results

The T7 promoter vectors used in this work were pT7-1 and pT7-2 (US Biochemicals; similar to that named pT7-1 by Tabor & Richardson, 1985) and pTZ18R and pTZ19R (Mead et al, 1986). Both pairs include the strong class III promoter, $\phi 10$, upstream of a multiple cloning site in a pBR322 derived vector containing the β -lactamase gene for ampicillin resistance. The various MUP gene fragments were cloned into these vectors as outlined in Table 2 and described in Materials and Methods. The positions of the sub-clones relative to a schematised MUP gene are shown in Figure 1. pTZ18R and pTZ19R were used when they became available since they include part of the lacZ gene allowing blue/white colour selection of positive transformants. Plasmids pT761 and pT762 were prepared by M. Richardson, and pT7-6-5-5 by T. Speigelberg.

Template DNA was linearised downstream of the insert with a restriction enzyme that gives either a blunt end or one with a 5'

Table 2

T7 templates used to generate control RNAs

The table describes the plasmid constructs created as templates for the synthesis of specific RNA control transcripts. The regions of the MUP genes included in the larger genomic subclones are shown in Figure 1. The ends of the cDNAs used can be determined from their nucleotide sequences as shown in Figure 2 and the 5' ends are shown by arrows in Figure 1.

Notes

- ¹ Clark et al (1982)
- ² a Hong clone of BS2 (Clark et al, 1982;1985a)
- ³ Ma (1987)
- ⁴ Bishop et al (1982)
- ⁵ Clark et al (1985b)

Template	Vector	Source of Mup fragment	Restriction sites used	Linearised with	Transcript length
pT761	pT7-2	pBS6-2 ¹ from <u>Sau3A</u> (+14) to <u>BamHI</u>	<u>BamHI</u>	<u>SmaI</u>	301nt
pT762	pT7-2	DLIN50 ² from 5' end to <u>BglII</u>	<u>BamHI</u>	<u>SmaI</u>	362nt
pT7-MUP15 ³	pT7-1	pUC-MUP15 entire insert	<u>EcoRI/BamHI</u>	<u>BamHI</u>	892nt
pT7-6-5-5	pT7-1	pBS6-5-5 ⁴ entire insert	<u>PstI/HindIII</u>	<u>HindIII</u>	1011nt
pTZ1PB	pTZ19R	pBS1-A10 ¹ from <u>PstI</u> in intron 2 to <u>BstNI</u> in intron 3 (blunt)	<u>PstI/SmaI</u>	<u>EcoRI</u>	770nt
pTZ5BE	pTZ19R	pBS5-2 ¹ from <u>BamHI</u> to <u>EcoRI</u>	<u>BamHI/EcoRI</u>	<u>EcoRI</u>	363nt
pTZ8EB	pTZ18R	pUC-MUP8 ⁵ entire insert	<u>EcoRI/BamHI</u>	<u>BamHI</u>	730nt
pT7-MUP11 ³	pT7-1	pUC-MUP11 ⁵ entire insert	<u>EcoRI/BamHI</u>	<u>BamHI</u>	732nt

overhang. Schenbon & Mierendorf (1985) have shown that template molecules with 3' overhangs generate extraneous RNA sequences with both SP6 and T7 RNA polymerase. Transcription reactions were performed as described in Materials and Methods (chapter 7). Initial experiments used ^{32}P - α -UTP as the labelled nucleotide to enable visualisation of transcripts on polyacrylamide-urea gels (Figure 3). The slower moving bands in lanes 2 & 3 were not observed in all experiments. As purification by DNaseI digestion of template DNA followed by Sephadex G100 chromatography (Melton et al, 1984) gave a single band of the expected size on the autoradiograph, this purification procedure was used for all subsequent experiments. When generating RNA to be used as a target for a ^{32}P -labelled probe, [5- ^3H]-UTP was employed as a tracer to quantify yield without obscuring the signal of hybridised probe. Because the initial experiments with ^{32}P - α -UTP had given a single band on gel electrophoresis it was assumed that ^3H -labelled transcripts would behave similarly. When used as targets for oligonucleotide and cDNA probes, this was found to be the case (Figures 5,6,9 & 10).

Figure 3

In vitro transcription with T7 RNA polymerase generates RNA of the expected size

pT7G2 was linearised with SmaI, transcribed and prepared for electrophoresis as described in Materials and Methods. The figure shows an autoradiograph of a 4% polyacrylamide gel with RNA samples at various stages of purification. The large arrow marks the expected position of the in vitro product.

Lane 1) Negative control (no enzyme)

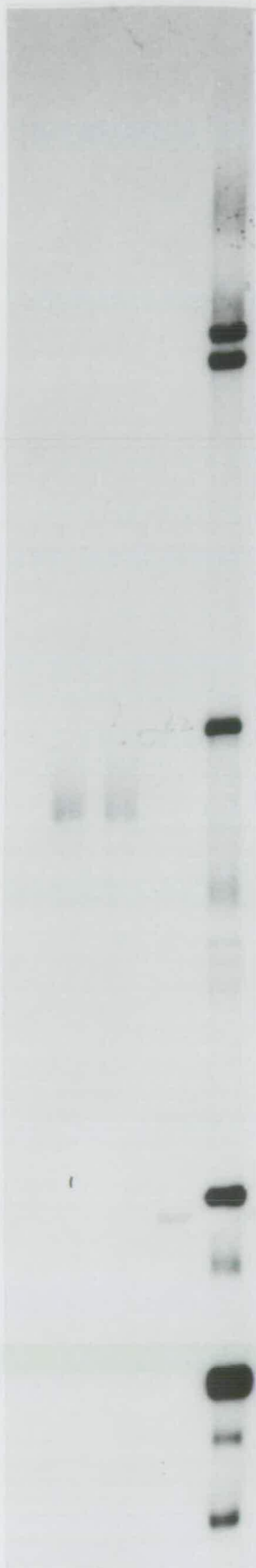
2) Crude reaction mix

3) After DNaseI digestion

4) After DNaseI digestion and G100 chromatography

M) Tag I digested pBR322 end-labelled with ^{32}P - α -dCTP and DNA polymerase I.

1 2 3 4 M



← 1450
← 1313

← 374

← 321
← 318

CHAPTER 4

THE EXPRESSION OF GROUP 1 AND GROUP 2 MUP GENES

Introduction

As described in the introduction, group 1 and group 2 MUP genes are arranged in a head-to-head manner within twelve to fifteen 45kb repeat units - one member of each group per repeat (Clark et al, 1984b; Bishop et al, 1985). This arrangement is believed to be the major unit of MUP gene duplication and evolution. There is closer similarity within groups than between them. The exonic sequences of group 1 genes share greater than 99% homology (Figure 2; Kuhn et al, 1984; Clark et al, 1985b; Shahan et al, 1987a); the exon 1 sequences of four group 2 genes differ by about 4% (Ghazal et al, 1985), and group 1 genes differ overall from group 2 genes by about 10%.

Group 1 genes are believed to generate the major proportion of MUP RNA in the liver for the following reasons:- i) all MUP liver cDNA clones isolated to date from six separate libraries in three laboratories are members of group 1, with the exception of the group 3 gene isolated as p199 (Kuhn et al, 1984) => MUP15 (Clark et al, 1985b) => cLIV4 => cLIV7 (Shahan et al, 1987a); ii) in vitro translation of liver RNA hybrid-selected with one of these group 1 clones gives an electrophoretic pattern very similar to that of the urinary MUPs synthesised in the liver (Knopf et al, 1983; Clissold et al, 1984). It is also known that submaxillary gland RNA hybrid selected

with the same cDNA clone generates a single species which comigrates with one of the liver products in a two-dimensional gel (Shaw et al, 1983). The in vitro translation products of hybrid-selected lachrymal gland RNA are much more basic than those generated by liver RNA and are thought to be encoded by different genes (Shaw et al, 1983; Shahan & Derman, 1984). Restriction enzyme analysis of uncloned cDNA prepared from submaxillary and lachrymal gland mRNA also suggests that the genes expressed in these tissues are different from those expressed in the liver (Shahan & Derman, 1984).

Prior to examination of the expression of individual group 1 genes it is necessary to determine the expression pattern of the group as a whole.

Before it was shown that group 2 genes are pseudogenes (Ghazal et al, 1985; Clark et al, 1985a) it was proposed that group 2 genes may be transcribed and spliced to generate the shorter form of MUP mRNA resulting from alternative splicing in the 3' untranslated portion of exon 6 (Clark et al, 1984a). This was deduced from the preferential hybridisation of a group 2 genomic sub-clone to the short message after stringent washing of Northern blots.

Since group 2 genes are now known to be pseudogenes, it is interesting to determine if they are transcribed or if the RNA hybridised by the group 2 probe was from a gene belonging to neither group 1 nor group 2 but showing greater similarity to group 2 genes in the region to which the probe hybridised. The transcription of unprocessed pseudogenes has not been widely reported in the literature; there is one example

of a human leucocyte interferon pseudogene isolated as a cDNA clone (Goeddel et al, 1981). Clark et al (1985a) have proposed that the truncated hexapeptide which could be encoded by the group 2 genes is the dialysable portion of the agent in male mouse urine which accelerates the onset of puberty when administered to young females (Vandenbergh et al, 1975; 1976; see chapter 1).

Group 1- and Group 2-specific Probes

The degree of homology is one factor determining the stability of a duplex and the shorter any duplex is, the greater the instability introduced by any one mismatch, i.e. a decrease in probe length should increase its specificity. Since cDNA and genomic sub-clone probes may not be specific enough to differentiate clearly between the transcripts of different but very similar genes, synthetic oligodeoxyribonucleotide probes were designed to be specific for group 1 and group 2 transcripts. The sequences of the probes are shown in Table 3 and the target sequences are marked in Figure 2. This region was chosen as a target because all sequenced group 2 genes are identical in this region (Ghazal et al, 1985) and are different at five positions from group 1 genes which are also identical to each other (Figure 2; Table 3). Both the group 1- and group 2- specific probes are different from group 3 gene sequences in this region. Initially, 14nt probes (corresponding to the first 14 nucleotides of the probes shown in Table 3) were chosen to maximise the percentage inhomology and thereby increase discrimination. These probes were, however, found to be inappropriate for the following reasons.

Table 3

Sequence of group 1- and group 2-specific oligonucleotide probes.

	60	43
6p1 consensus probe:	5' AGGGAATAGGATTGTCTG	3'
tT761:	3'... <u>u</u> cc <u>u</u> ua <u>u</u> cc <u>u</u> aacagac...5'	
tT762:	3'... <u>g</u> ac <u>u</u> uu <u>a</u> uc <u>u</u> uu <u>a</u> uagac...5'	
tT7MUP15:	3'... <u>u</u> cc <u>u</u> ua <u>u</u> cc <u>u</u> aacagga...5'	
	60	43
6p2 consensus probe:	5' CAGAAATAGAATTATCTG	3'
tT762:	3'... <u>g</u> ac <u>u</u> uu <u>a</u> uc <u>u</u> uu <u>a</u> uagac...5'	
tT761:	3'... <u>u</u> cc <u>u</u> ua <u>u</u> cc <u>u</u> aacagac...5'	
tT7MUP15:	3'... <u>u</u> cc <u>u</u> ua <u>u</u> cc <u>u</u> aacagga...5'	

The co-ordinates are as shown in Figure 2.

The sequences of the control RNAs are shown below each probe.

Mismatched bases are underlined.

Firstly, the group 1-specific probe hybridised to an RNA species intermediate in length between the long and short MUP mRNAs present in lachrymal, submaxillary and kidney, as well as liver RNA. Primer extension sequencing (see below) of kidney RNA using this primer gave an extension product longer than expected for MUP and which also showed no similarity to known MUP sequences. From tissue distribution, sexual dimorphism and the limited sequence data, the most likely candidate for this fortuitously hybridising transcript appears to be a member of the kallikrein multigene family (Mason et al, 1983). This matter was not investigated further.

Secondly, the group 2-specific probe failed to hybridise to the positive control transcripts or to cellular RNA even under relaxed conditions of hybridisation (1.2M M⁺, 12°C). This may be because the T_m of the DNA:RNA hybrid is much lower than expected because of 7/14 dA:rU base-pairs. Homopolymers of dA:rU have been shown to denature at ~20°C below the temperature at which homopolymers of dA:dT denature (Chamberlin, 1965). This probe gives a weak signal when hybridised to homologous DNA (Figure 4). Note that the group 1 probe gives comparable signals with both DNA and RNA targets while no signal is detectable from RNA challenged with the group 2 probe and the DNA signal is much weaker. This latter observation suggests that the weakness of the dA:rU base pair may not be the only problem here.

To overcome these difficulties the 18nt probes shown in Table 3 were synthesised. The longer probe is less likely to hybridise to transcripts of genes other than MUPs. Furthermore, any duplex formed

Figure 4

Hybridisation of 14nt group1- and group 2-specific oligonucleotide probes to homologous DNA and RNA.

The inserts were cut from the plasmids pT761 and pT762 by digestion with EcoRI and HindIII, electrophoresed in a 2%(w/v) agarose gel and transferred to nylon membranes by capillary blotting after depurination and denaturation. The corresponding synthetic RNAs (tT761 and tT762) were prepared and transferred as described in Materials and Methods. The expected position of tT762 is shown by the arrow.

The filters were hybridised and washed in 1.2M M⁺ at 12°C.

Lanes 1,2,5 & 6 were probed with the group 1-specific probe and lanes 3,4,7 & 8 were probed with the group 2-specific probe. Those probed with the group 1-specific oligonucleotide were subsequently washed at room temperature.

DNA Targets

Lanes 1,3) 20ng pT761 insert

2,4) 20ng pT762 insert

RNA Targets

Lanes 5,7) 10ng tT761

6,8) 10ng tT762

Probes

Gp1

Gp2

Gp1

Gp2

1 2

3 4

5 6

7 8



would be expected to be more stable due to the increase in probe length.

Tissue-specificity of Group 1 Gene Expression

Total cellular RNA was prepared from MUP expressing tissues, electrophoresed and transferred to nylon membranes. The group 1 cDNA clone MUP11 (Clark et al, 1985b) was labelled by the random primer method (Feinberg & Vogelstein, 1983) and used to probe one of the Northern blots. Figure 5 shows the relative levels of hybridisation with MUP RNA from male and female liver, submaxillary gland, lachrymal gland and in the prelactational mammary gland (Figure 5a). The filter was first hybridised with the group 1-specific 18nt probe. Group 1 MUP RNA was found only in the male and female liver and in the prelactational mammary gland (Figure 5b). Females treated with testosterone show male-like levels of MUP RNA in the liver (Figure 5; Knopf et al, 1983; Clissold et al, 1984). Figure 5b shows that group 1 gene expression in particular is induced to male-like levels in testosterone treated females.

These results were quantified by scanning densitometry of the autoradiographs (Table 4). Since the consensus sequence to which the probe hybridises was derived from MUP genes from each of the four phylogenetic sub-groups of group 1 (Al-Shawi et al, 1988a) and the RNA coding sequences of group 1 genes differ by <1%, it is expected that the probe would detect transcripts from all group 1 genes. Taking the figure of 0.45% as the percentage of total cellular RNA from male liver that is MUP mRNA and assuming that 3.6% of total



Figure 5

Tissue-specific expression of group 1 MUP genes

Total cellular RNA and RNA controls were prepared, electrophoresed and transferred as described in Materials and Methods. Panel (a) was probed with the group 1 cDNA clone MUP11 and exposed for 16hr; panel (b) was probed with the group 1 specific oligonucleotide and exposed for 6hr, except lane 9' was exposed for 3 days. This exposure did not reveal transcripts in the submaxillary, lachrymal or negative control lanes. The arrows indicate the expected positions of negative control RNA.

Lane 1) 3 μ g testosterone-induced total female liver RNA

2) -7) 3 μ g, 1 μ g, 0.3 μ g ... 0.01 μ g total male liver RNA

8) 30 μ g total female liver RNA

9) 30 μ g total mammary gland RNA

10) 30 μ g total male submaxillary gland RNA

11) 30 μ g total female submaxillarygland RNA

12) 30 μ g total male lachrymal gland RNA

13) 30 μ g total female lachrymal gland RNA

14) 3ng group 1 control RNA

15) 30ng group 1 control RNA

16) 30ng Mup15 control RNA (892nt)

17) 30ng group 2 control RNA (362nt)

Abbreviations:-

T = Testosterone-induced female liver RNA

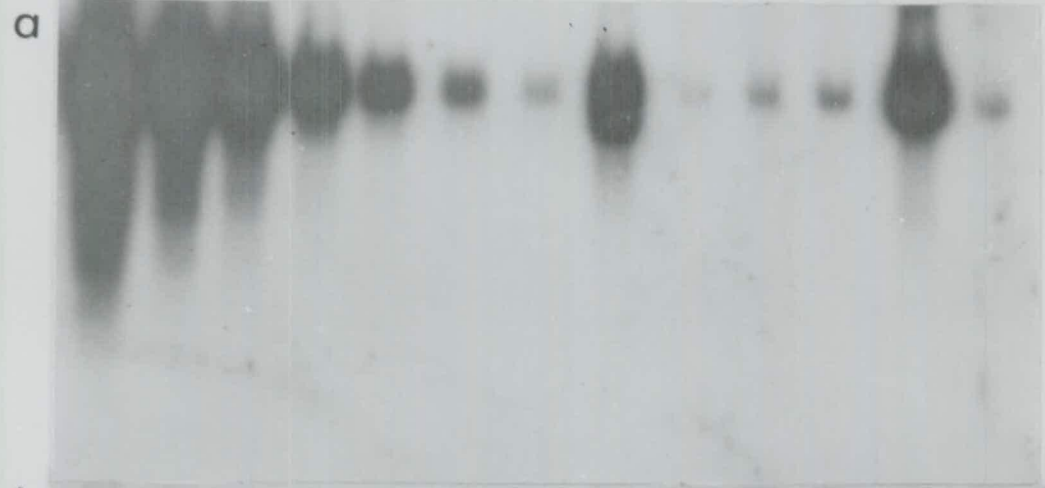
L = Liver

M = Prelactational mammary gland (14 days gestation)

Sax = Submaxillary gland

Lac = Lachrymal gland

T	♂ Liver						♀		Smx		Lac	
1	2	3	4	5	6	7	L	M	♂	♀	♂	♀
							8	9	10	11	12	13



	Controls			
	+		-	
M	14	15	16	17

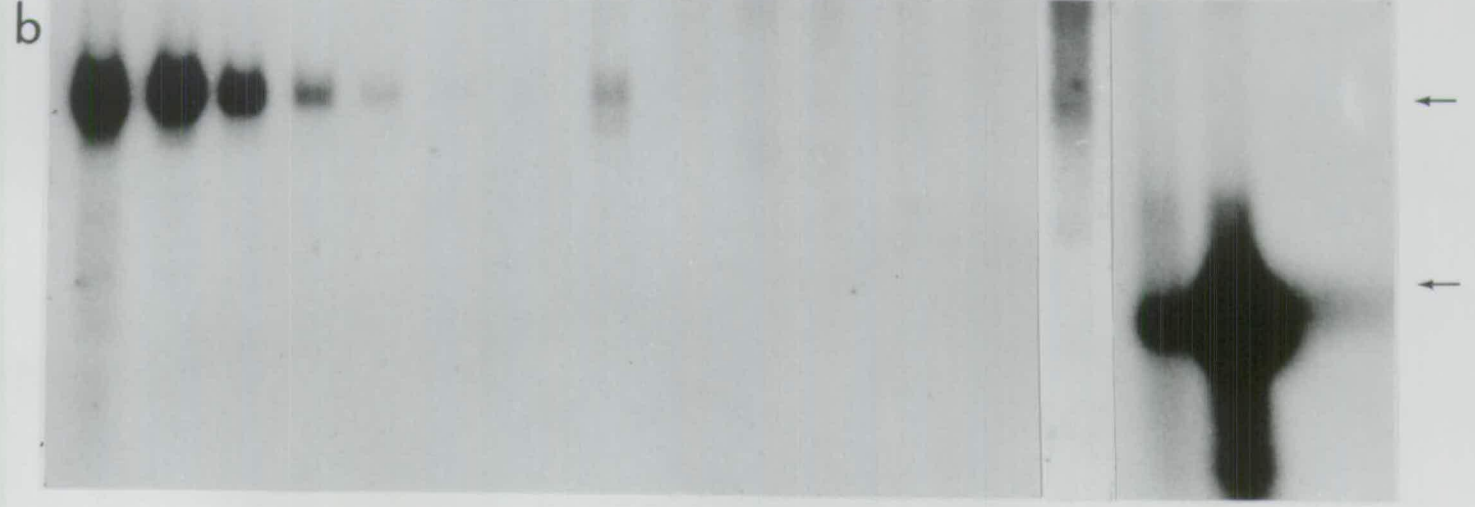


Table 4

The tissue distribution of group 1 MUP gene transcripts

The values were obtained by densitometric scanning with an integrator of the autoradiographs shown in Figure 5 and relating the signal from male liver samples to the positive controls allowing for different lengths of control transcript (as described in Appendix). Because it was not possible to determine the proportion of male liver RNA detected by the MUP11 cDNA probe, data in this Table is presented in percentage terms rather than in copies/cell (Table 7).

Abbreviations:-

M = male

T = testosterone-induced female

F = female

Mamm = prelactational mammary gland (14 days gestation)

Smx = submaxillary gland

Lac = lachrymal gland

* = no data from control transcript

Probe	M Liver RNA (% total liver)	T Liver	F Liver	Mann (% male liver RNA)	M Smx	F Smx	M Lac	F Lac
MUP11	*	103	1.2	0.01	0.05	0.05	1.0	0.05
Group 1	0.45	104	1.0	0.01	0	0	0	0

liver RNA is mRNA (Hastie & Bishop, 1976), then 12.5% of mRNA in BALB/c male liver is the product of group 1 MUP genes. This is in fairly good agreement with the level of 8% reported by Derman et al (1981) for NCS mice.

The level of MUP RNA in female liver is lower than expected. Hastie et al (1979) have shown by hybridisation kinetics that the male:female MUP liver RNA ratio is approximately 5:1 and similar results have been obtained using cDNAs and genomic sub-clones as probes to Northern blots (Shaw et al, 1983; Clissold et al, 1984; J.O. Bishop, unpublished results). This finding was investigated further by determining the relative amounts of MUP mRNA in the liver RNA of different females (Figure 6). Sample 1, from a 7 week old mouse, is the same preparation as the female liver sample in Figure 5; sample 2 is from a 6 week old mouse (gift of J.P. Whitaker) and samples 3 & 4 are from 8-10 week old mice (gift of X. Ma). The mice induced with testosterone were siblings of numbers 3 & 4. The variation between mice was quantified by scanning densitometry and the results are presented in Table 5a. The level of MUP mRNA in the sample from the 6 week old mouse is in fairly good agreement with previously published levels (1/7th of male levels c.f. 1/5th). The 8-10 week old mice have 2-3 fold less MUP mRNA. The amount of MUP mRNA from the 7 week old mouse is the same as determined previously (1% of male, Table 4).

It seems unlikely that a simple age effect underlies this variation since the highest level of MUP RNA is observed in the youngest mouse and intermediate levels in the oldest mice. Female mice of this age

Figure 6

Variation in levels of MUP gene expression in the liver between individual female mice.

Total cellular RNA and controls were prepared, electrophoresed and transferred as described in Materials and Methods. Panel (a) was probed with the group 1 cDNA clone MUP11 and exposed for 3hr, panel (b) was probed with the group 1-specific oligonucleotide and exposed for 16hr. The expected positions of group 2 and Mup15 negative control transcripts are shown by arrows.

- Lane 1) 3 μ g testosterone-induced total female liver RNA
- 2-4) 3 μ g, 1 μ g, 0.3 μ g total male liver RNA
- 5) 30 μ g total liver RNA from 7wk old female
- 6) 10 μ g total liver RNA from 7wk old female
- 7) 30 μ g total liver RNA from 6wk old female
- 8) 10 μ g total liver RNA from 6wk old female
- 9,11) 30 μ g total liver RNA from 8-10wk old females
- 10,12) 10 μ g total liver RNA from 8-10wk old females
- 13) 30 μ g total RNA from prelactational mammary gland
- 14) 3ng group 1 control RNA
- 15) 1ng group 1 control RNA
- 16) 30ng Mup15 control RNA (892nt)
- 17) 30ng group 2 control RNA (362nt)
- 18) 30 μ g total male kidney RNA

Abbreviations:-

T = Testosterone-induced female liver RNA

L = Liver

M = Prelactational mammary gland (14 days gestation)

Table 5

Variation in MUP liver RNA levels between individual female mice

The upper table (a) shows the variation in MUP liver RNA levels between individual female mice. The values were obtained by densitometric scanning, with an integrator, of the autoradiographs shown in Figure 6 and relating the signal from male liver samples to the positive controls allowing for different lengths of control transcript (as described in the Appendix). As in Table 4 the results are presented as percentages. The female liver samples are marked 1-4 as in Figure 6. Other abbreviations are as defined in Table 4.

The lower table (b) shows the variation between the individual females observed with different liver cDNA probes. The data was obtained by densitometric scanning, with an integrator, of the autoradiographs shown in Figures 6 & 7 and is normalised with respect to apolipoprotein A1.

a) Probe	M Liver (% total liver)	F Liver (% male liver)			
		1	2	3	4
MUP11	*	1.2	14.0	4.5	5.5
Group 1	0.42	1.0	13.9	7.8	7.3

b) Probe	Female liver sample				Male liver
	1	2	3	4	
Apolipoprotein A1	100	100	100	100	100
Albumin	27	24	42	40	7
Transferrin	524	197	105	85	188
MUP	38	287	261	173	3837

(8-10 weeks) have previously been shown to have about 20% of male liver MUP RNA levels (Clissold et al, 1984). The level of MUP liver RNA in female mice relative to that observed in males has been found to range from unity to less than 1:100 when MUP gene expression was examined in different species of mouse (Sampsell & Held, 1985). Sampsell & Held's data also show some variation in amounts of MUP RNA between mice of the same sex from the same stock.

To determine whether the variation was due to a general variation in RNA synthesis or was specific to MUP the filter was reprobed with apolipoprotein A1 (LVB6; Clissold & Bishop, 1981) and serum albumin (LVA304; Clissold & Bishop, 1981), transferrin (TRF32; Chave-Cox, 1986) and a Xenopus rDNA clone (pCM4305; Bishop, 1979). The results of these experiments are shown in Figure 7.

The results obtained with the rDNA probe show that the gel loadings were approximately correct. The relative levels of the other three species could be quantified by scanning densitometry with integration. The results are normalised with respect to apolipoprotein A1 (Table 5b). Since the probes were not labelled to the same specific activities it is not possible to determine the molar or mass ratios of the mRNAs. It is possible, however, to compare the relative levels of each RNA between different samples of female liver and male liver RNA. This comparison shows that there is little variation (< 2 fold) between females except that the 7 week old mouse has increased levels of transferrin mRNA and reduced levels of MUP mRNA. The male liver RNA shows the expected high level of MUP mRNA but a low level of serum albumin mRNA. The figure also shows that the prelactational mammary

Figure 7

Variation of different liver-specific mRNAs between individual female mice

The filter shown in Figure 6 was reprobbed with the cDNA clones homologous to apolipoprotein A1 (LVB6; Clissold & Bishop, 1981), serum albumin (LIV304; Clissold & Bishop, 1981) and transferrin (TRF32; Chave-Cox, 1986) to investigate further the variation observed with the MUP cDNA probe (Figure 6a). The filter was finally probed with a Xenopus laevis rDNA probe (pCM4305; Bishop, 1979) to verify equivalence of sample loading.

Abbreviations:-

ALB = Serum Albumin

APO = Apolipoprotein A1

TRF = Transferrin

T = Testosterone-induced female liver RNA

M = Prelactational mammary gland (14 days gestation)

K = Male kidney RNA

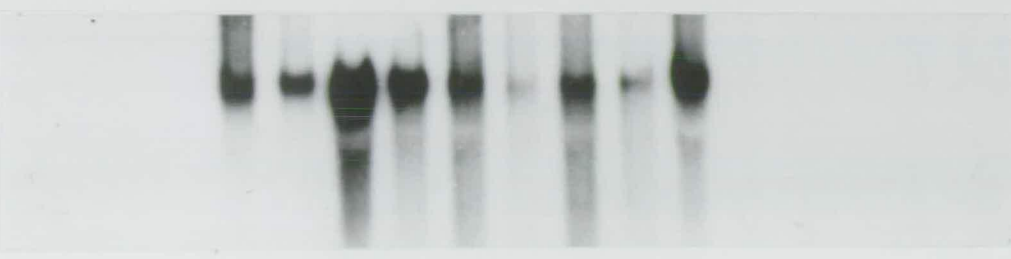
T ♂ Liver ♀ Liver M T7 RNA K

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18



ALB

APO



TRF



28S

rRNA

18S

gland expresses transferrin at levels similar to female liver (as previously reported by Chen & Bissell, 1987).

These results show that the original female liver RNA sample was deficient in MUP mRNA compared with other females. This may be related to the observed increase in the level of transferrin mRNA. It is possible that this is due to the general health of the animal. For example, transferrin gene expression is increased by iron deficiency and MUP gene expression may fall when animals are under physiological stress. It is also possible that hepatic MUP gene expression is influenced by the changing levels of the female sex hormones.

Clark et al (1984a) have shown that MUP gene transcripts can be spliced to generate two size classes of mRNA in the liver. The longer message lacks all but the 5', coding portion of exon 6 and includes exon 7 whereas the shorter message includes all of exon 6 and lacks exon 7. The longer species is seen to be more abundant using either a group 1 genomic sub-clone (Clark et al, 1984a), a cDNA clone or the group 1 specific oligonucleotide (Figure 5) as a probe.

The shorter MUP RNA species, however, appears to constitute a greater proportion of female liver RNA than male. This may reduce the apparent level of MUP mRNA in female liver relative to male as determined by densitometric scanning, due to the increased breadth of the band (compare Figure 6, lane 7 with Table 5a). Both female liver group 1 MUP cDNA clones isolated and sequenced to date (MUPs 8 and 11, Clark et al, 1985b; cLIV6=>MUP8, Shahan et al, 1987a) are reverse transcripts of the longer message (Figure 2). A comprehensive

analysis of this would require a series of S1 nuclease protection experiments with different 3' probes. A group 3 gene has been shown to encode transcripts spliced in a sexually dimorphic manner. The cDNA from a female library (MUP15, Clark et al, 1985b; cLIV7, Shahan et al, 1987a) includes the 31bp immediately 3' to the splice junction at the end of exon 6a (Figure 2), whereas an otherwise identical clone from a male library (cLIV4, Shahan et al, 1987a) is spliced in a similar manner to the longer group 1 cDNAs.

Primer Extension Sequencing

The group 1-specific oligonucleotide was used as a primer for primer extension sequencing (Epstein et al, 1986) to confirm the transcription initiation site previously proposed by Clark et al (1985a). Because RNA controls are available it is possible to show the specificity of the primer to the group 1 RNA template. Furthermore, any expressed group 1 gene which differs at any point in the 5' untranslated region from those cloned and sequenced to date could be identified by changes in the sequence ladder.

Figure 8a shows the primer extension sequence of male liver RNA. This is complementary to the consensus for group 1 RNA (Figure 8b). The termination site at the 5' end of the mRNA is visible as a doublet (lanes 1-6; arrow A). This is most likely to be due to the obstruction of the polymerase by the cap structure. The reverse transcript of group 1 control RNA (lane 8, arrow B) is longer than

Figure 8

Primer extension sequencing of total liver RNA

a) Determination of group 1 RNA 5' sequence

Annealing of primers and their extension was as described in Materials and Methods.

Lanes 1-5) 4µg total male liver RNA

6) 25µg total female liver RNA

7) 25µg total mammary gland RNA

8) 5ng group 1 control RNA

9) 25ng MUP15 control RNA

Reactions in lanes 2-5 contained 0.1mM ddGTP, ddATP, ddTTP, ddCTP respectively.

The termination sites for liver (A) and control (B) RNA are arrowed.

b) Sequence alignment of group 1 MUP mRNA consensus with that derived from primer extension sequencing

The sequence read from the gel is shown (lower line, upper case) against the consensus group 1 RNA sequence from the cap site to the end of the primer binding site (upper line, lower case). The sequence of the primer is shown in lower case. The first base after the primer, and the two terminal bases could not be unambiguously determined and are shown as X.

that from liver RNA because some polylinker sequence is included in the template. The smear visible in all tracks is probably due to end addition of the primer (Lee & Roeder, 1981). Note that, other than this smear, no extension products are present when MUP15 control RNA is used as a template. These results confirm the predicted transcription start site for group 1 genes and shows that any group 1 RNAs differing within the 5' untranslated region do not represent a significant portion of liver MUP RNA. A faint extension product is visible when female liver RNA is used as a template (Figure 8, lane 6). This RNA was from the preparation shown as sample 1 in Figure 5 and was subsequently shown to be deficient in MUP RNA.

Electrophoresis on a lower percentage gel did not reveal any initiation sites further upstream. No extension products were observed when this primer was used with submaxillary or lachrymal gland RNA as the template, nor when using the group 2 specific oligonucleotide as a primer. Shahan et al (1987a) have recently reported the isolation of a cDNA clone from a male liver library which includes 46bp of the immediate 5' flanking DNA. No upstream initiation sites have been observed in four separate analyses by primer extension (Figure 8; Clark et al, 1985a; Shahan et al, 1987a) and S1 nuclease protection (Clark et al, 1985a), and the origin of this clone remains unclear.

Group 2 Pseudogenes Are Not Expressed

The 18nt group 2-specific probe (Table 3) was used to probe Northern blots in an attempt to identify any transcripts of the group 2

pseudogenes. This probe proved extremely difficult to handle and only weak hybridisation to positive controls could be detected (30ng, equivalent to approximately 10^4 copies mRNA per cell). No signal was observed with 30µg total RNA from male or female liver, submaxillary gland or lachrymal gland, or 2µg of polyA⁺ RNA from male and female liver (Figure 9). A range of temperatures (20°C - 32°C) were used to try and improve the sensitivity but without success. The most likely explanation for the lack of sensitivity observed here is the base composition and possibly sequence of the probe. As mentioned above, homopolymers of dA:rU are much less stable than dA:dT but the target site for the probe is restricted to the regions over which group 1 and group 2 genes are identical to their respective consensus sequences but different from each other.

It is possible, therefore, that group 2 genes are expressed at levels below the detection limits of this assay. Shahan et al (1987b) have recently reported that an oligonucleotide probe homologous to the region containing a conserved point mutation (Ghazal et al, 1985) does not identify any transcripts either. The sequence and base composition (50% G+C) of this probe would not be expected to present the same difficulties as found in the experiments described here. The probe did, however, identify twelve different MUP genomic clones suggesting that the point mutation generating the stop codon occurred shortly after the original duplication event and prior to the amplification of the 45kb repeats (as postulated by Ghazal et al, 1985).

Since there is no evidence that group 2 genes are transcribed it is

Figure 9

Expression of group 2 pseudogenes?

Total cellular RNA and RNA controls were prepared, electrophoresed and transferred as described in Materials and Methods. Liver polyA⁺ RNA was prepared by the method of Wreschner & Herzberg (1984) using Hybond-mAP (Amersham) as directed by the supplier. The autoradiograph was exposed for 20hr. Longer exposures do not reveal any MUP-sized bands (upper arrow) without negative control showing. The lower arrow shows the expected position of negative control RNA.

- Lane 1) 30µg total male liver RNA
2) 30µg total female liver RNA
3) 30µg testosterone-induced female liver RNA
4) 2µg poly A⁺ male liver RNA
5) 2µg polyA⁺ female liver RNA
6) 30µg total mammary gland RNA
7) 30µg total male submaxillary gland RNA
8) 30µg total female submaxillary gland RNA
9) 30µg total male lachrymal gland RNA
10) 30µg total female lachrymal gland RNA
11) 30ng group 2 control RNA
12) 10ng group 2 control RNA
13) 30ng group 1 control RNA (301nt)
14) 30ng MUP15 control RNA (892nt)

Abbreviations:-

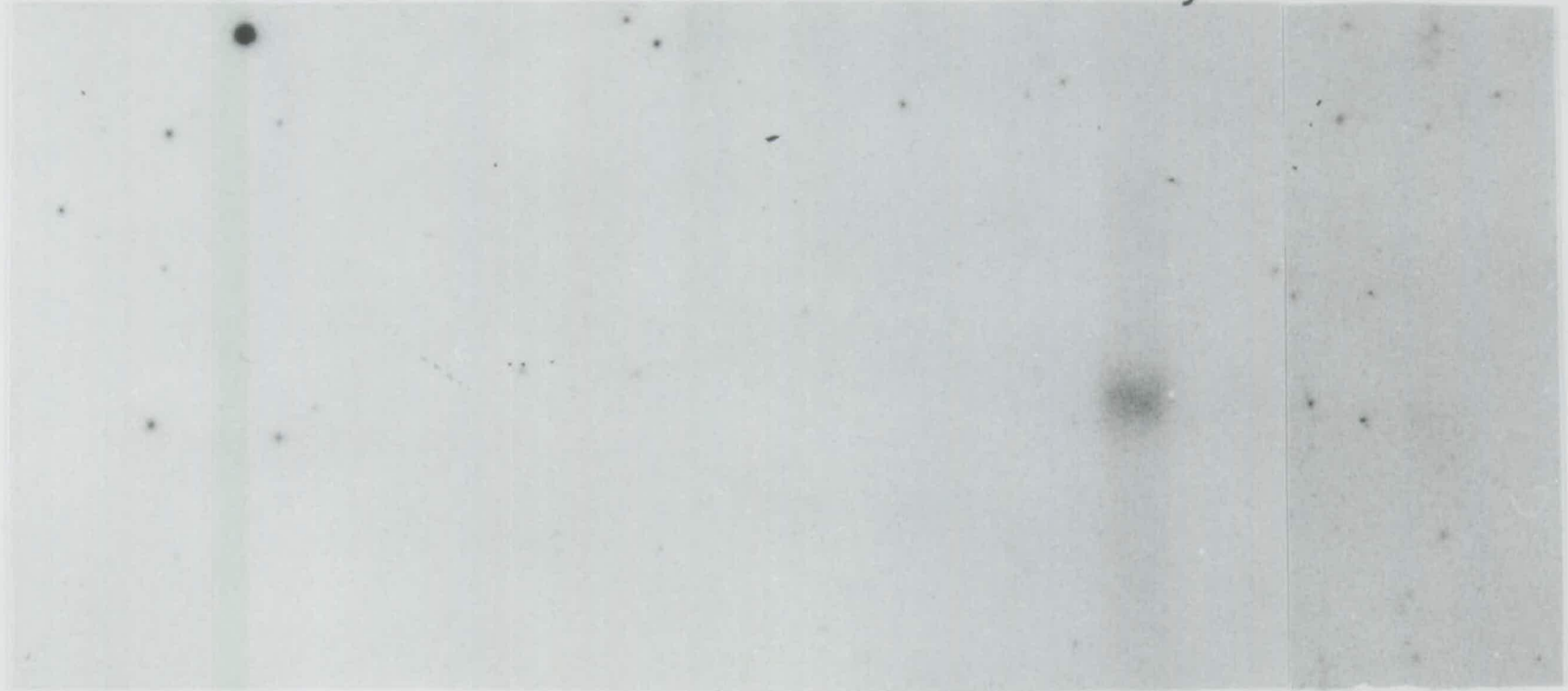
T = Testosterone-induced female liver RNA

M = Prelactational mammary gland (14 days gestation)

Smx = Submaxillary gland

Lac = Lachrymal gland

Liver						Smx	Lac	Controls					
♂	♀	T	♂	♀	M	♂	♀	♂	♀	+		-	
1	2	3	4	5	6	7	8	9	10	11	12	13	14



unlikely that the dialysable component of male urine is the hexapeptide which could be encoded by group 2 genes (Clark et al, 1985a). It is still possible, however, that the full length MUPs encoded by group 1 genes act as carriers for the active pheromone, or that they can be cleaved to give an active product, as originally postulated by Vandeburgh et al (1975, 1976).

Conclusions

From the results presented here it is concluded that group 1 MUP genes are expressed in the livers of male and female mice and in the prelactational mammary gland and represent the major portion of MUP RNA in these tissues. The group 2 pseudogenes are not transcribed at levels (10^4 copies per cell) that can be detected in the assay employed here.

As the tissue-specific expression of the group 1 genes has been established it is now possible to examine the expression of some of them individually. These experiments are described in the following chapter.

CHAPTER 5

SEXUALLY DIMORPHIC EXPRESSION OF INDIVIDUAL GROUP 1 GENES

Introduction

Knopf et al (1983) have shown by analysis of hybrid-selected and in vitro translated liver RNA that the relative levels of different MUPs are altered by various hormonal treatments. It is believed that the different MUP species are encoded by different group 1 genes. Held et al (1987) have recently shown that three different C57BL/6 genomic clones encode proteins which co-migrate with three different MUP species from in vitro translated liver RNA. However, it is possible that particular MUP species visualised on a two-dimensional gel could be the products of more than one gene. To show that a particular hormone treatment is affecting a given gene it is therefore necessary to measure levels of RNA transcribed from that gene in vivo i.e. a gene-specific probe is required.

If it can be shown that a certain gene (or genes) within the multigene family is regulated in a certain manner it then will be possible to determine whether cis-acting elements are involved in this regulation by comparing the sequences of flanking DNAs. Functional analysis of different promoter regions would involve the production of transgenic mice and determining the level of reporter gene expression. It has not been possible to analyse MUP gene promoter sequences in tissue culture cells because no cell line in which MUP genes are expressed has yet been found unless they are linked to the SV40

enhancer (J.O. Bishop, unpublished results).

Group 1 MUP genes differ by less than 1% in their exonic sequences. Short synthetic oligonucleotides are necessary to discriminate between the transcripts of individual group 1 genes. From the sequence comparison presented in Figure 2 it can be seen that group 1 MUP genes differ from each other at point mutations throughout exons 1-7. Each shall require a unique positive control RNA, and hybridisation and washing conditions must be determined for individual probes so that transcripts from the target genes can be identified unambiguously.

Wallace and co-workers have shown that oligonucleotide probes can discriminate between alleles of a gene which differ at only a single nucleotide (Conner et al, 1983; Kidd et al, 1983; Studencki et al, 1985; DiLella et al, 1986), for example to diagnose disease traits. An important difference between these examples of DNA-DNA hybridisation and the proposed use of oligonucleotide probes to discriminate between mRNA species is that a small degree of cross-hybridisation (up to 10%) is acceptable in the former case because it does not affect the conclusion:- a positive/negative ratio of 10 implies one allele is homozygous. This is not the case when one is attempting to measure the amount of a single mRNA species in a mixture.

In this chapter the design of the gene-specific oligonucleotides is described and their use to probe Northern blots is presented.

The Design of Gene-specific Oligonucleotide Probes

The sequences of the gene-specific oligonucleotides are shown in Table 6 and their target sequences are underlined in Figure 2. In choosing appropriate target sequences, the 3' portion of exon 6 omitted from the long form of MUP mRNA (Clark et al, 1984a) and exon 7 were not considered because it is not known how the transcript of each individual group 1 MUP gene is spliced. The expression of genes cloned only as cDNAs (p499, p1057 and MUP11) was not investigated for the following reasons. Firstly, it is the regulation of different MUP genes that is of primary interest since these clones may contain upstream regulatory elements involved in the control of their expression. Secondly, p499 was cloned from a C57BL liver library and its relationship to BALB/c genes is unknown, and finally, MUP11 has no unique point mutations and p1057 was not available as a cDNA to generate a control transcript.

The chosen target sequences include the greatest number of differences between the gene of interest and other MUP genes, although there is only ever one known mismatch between group 1 genes in any given region. For example, the BL1-specific probe was targeted to exon 5 rather than to the EcoRI site in exon 2 because there are more differences between the BL1-specific probe and other MUP gene sequences (Figure 2). Furthermore, the C57BL cDNA clone p499 shares the same T→A mutation in exon 2 and the homologous gene may be present and expressed in BALB/c mice. Similarly, the BS1 probe was targeted to exon 3 rather than exon 5 because of greater variation at this

Table 6

Sequence of gene specific oligonucleotide probes.

		553		536
BL1 specific probe:	5'	CACATAGTTTTGCAAACC	3'	
tTZ8EB:	3'	...guguauc ^u aaaacg ^u uu ^g g...5'		
tT7-MUP11:	3'	...guguauc <u>a</u> ac ^c g ^u uu ^g g...5'		
		341		324
BS1 specific probe:	5'	GATAATCCGAGCACTCT	3'	
tTZ1PB:	3'	...cuauuaag ^c c ^u cg ^u g ^a ga...5'		
tT7-MUP11:	3'	...cuauuaag <u>c</u> c ^u cg ^u g ^a ga...5'		
		262		245
BS5 specific probe:	5'	GAAAAAGTTTAAAGTTGC	3'	
tTZ5BE:	3'	...cuuuuuca ^a auuuca ^a cg...5'		
tT7-MUP11:	3'	...cuuuuu <u>c</u> ag ^a uuu ^a ca ^a cg...5'		
		499		482
BS6 specific probe:	5'	CATAGAGACCCATCAGCT	3'	
tT7-6-5-5:	3'	...g <u>a</u> u <u>c</u> u <u>c</u> g <u>g</u> g <u>u</u> ag <u>u</u> cg <u>a</u> ...5'		
tT7-MUP11:	3'	...g <u>a</u> u <u>c</u> u <u>c</u> g <u>g</u> g <u>u</u> ag <u>u</u> cg <u>a</u> ...5'		

The co-ordinates are as shown in Figure 2.

The sequences of positive and negative control RNAs are shown below each probe. Mismatched bases are underlined.

point. Hybridisation conditions which allow discrimination between group 1 genes differing at a single nucleotide will presumably also discriminate between the group 1 gene of interest and other MUP genes. It is therefore not necessary to include control RNA derived from group 2 and group 3 genes. The probes for BS5 and BS6 are targeted for these genes' unique point mutations within the coding region. In each case the single mismatch was positioned near the middle of the probe as this has the greatest destabilising effect (Smith, 1983).

Control templates were prepared and RNA transcribed from them as described in chapter 3 (Figure 1, Table 2). The cDNA clone MUP8 was used as a positive control for BL1 because it is identical to the exonic gene sequence (Figure 2). MUP11 was used as the negative control in each case because at each target site its sequence corresponds to the consensus group 1 sequence (Figure 2).

Hybridisation of the Gene-specific Probes

Northern blots of RNA from tissues that express group 1 genes (male and female liver and prelactational mammary gland) together with the appropriate control RNA species were prepared as before (chapters 3 & 4). A large amount of the negative control RNA was loaded so as to minimise the possibility of overestimating the abundance of a particular species.

It should be noted that MUP genes may exist which contain the same point mutation as those being investigated, but which have not been cloned and sequenced. If these are transcribed, erroneous conclusions

might be drawn. It may be possible to show how many genes share a given point mutation by hybridising each oligonucleotide probe to Southern blots of genomic DNA digested with a number of restriction enzymes but this could only be accurate \pm one copy. Of the genes whose expression is investigated here, it has been estimated by comparing the intensity of hybridisation signal between cloned and genomic DNA, that BL1 and BS1 are each present in 1-2 copies per haploid genome but that there are 2-3 genes homologous to BS6 (Al-Shawi, 1985; Al-Shawi et al, 1988a). It is not known how many of the BS6-like genes share the same point mutation.

BL1 probe

Using the Schildkraut & Lifson relation (see chapter 2, equation 3) the T_m of the BL1 probe and its target sequence was estimated in the final wash solution at 48°C. Hybridisations and washes were performed as described in Materials and Methods on a rotary shaking platform in an incubator at 43°C. These conditions were found to allow the BL1 probe to pair with the complementary positive control sequence but not with the negative control. Under these conditions the probe hybridised to RNA from both male and female liver (Figure 10a). The individual variation between samples of female liver RNA observed using the group 1 consensus oligonucleotide probe (chapter 4) is also seen here. A faint band is visible in mammary gland RNA after a longer exposure.

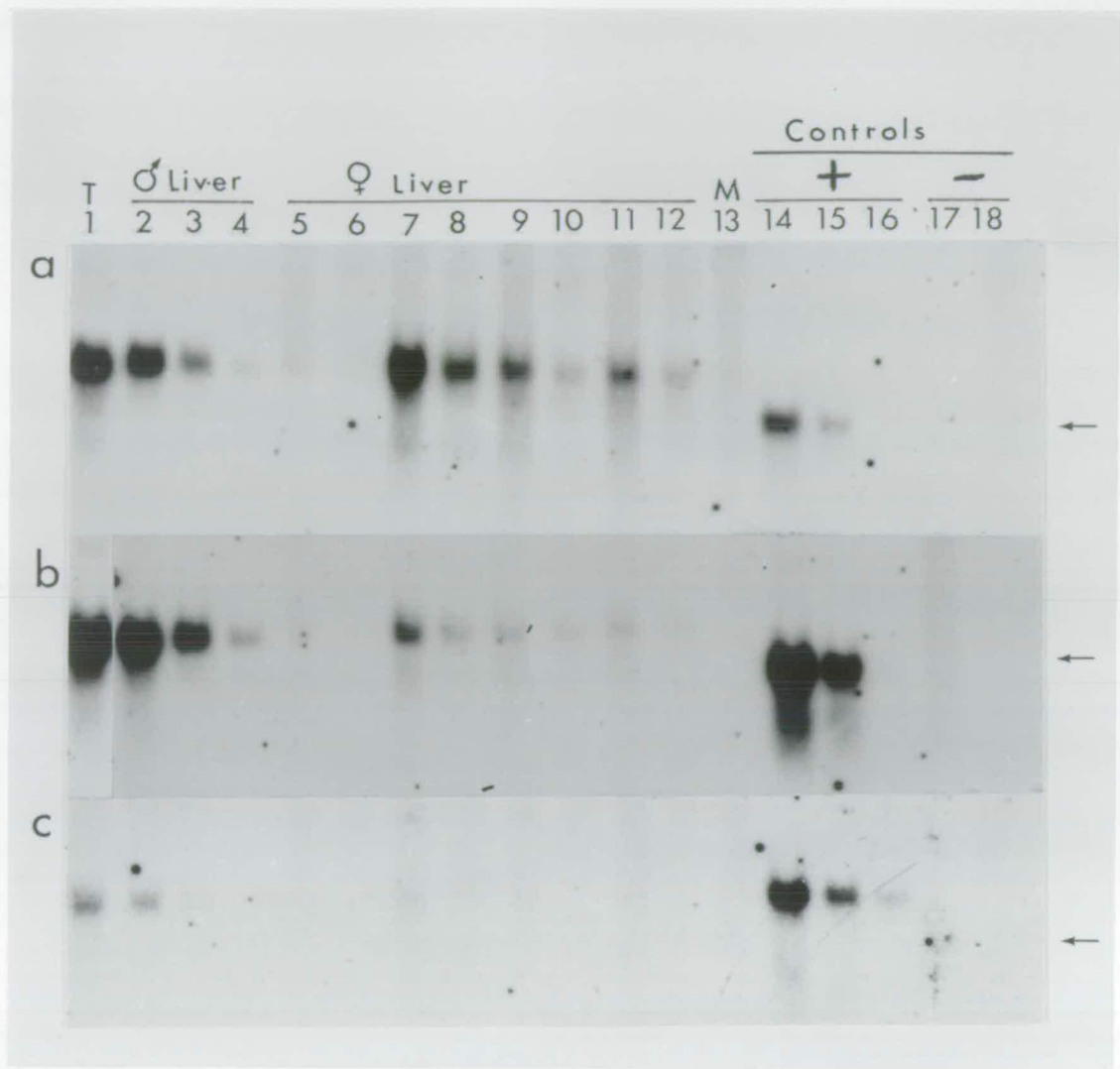
The expression of the BL1 gene has coincidentally been investigated in other laboratories. Shahan et al (1987 a,b) have isolated a cDNA

Figure 10

The expression of individual group 1 MUP genes

Filters were prepared with appropriate RNA controls as before (Figure 5) and probed as follows:- panel (a) BL1-specific probe, 3 day exposure; panel (b) BS1-specific probe, 3 day exposure; panel (c) BS6-specific probe, 10day exposure. Arrows show the expected positions of negative control RNA.

- Lane
- 1) 3 μ g testosterone-induced female liver RNA
 - 2-4) 3 μ g, 1 μ g, 0.3 μ g total male liver RNA
 - 5) 30 μ g total liver RNA from 7wk old female
 - 6) 10 μ g total liver RNA from 7wk old female
 - 7) 30 μ g total liver RNA from 6wk old female
 - 8) 10 μ g total liver RNA from 6wk old female
 - 9, 11) 30 μ g total liver RNA from 8-10wk old females
 - 10, 12) 10 μ g total liver RNA from 8-10wk old females
 - 13) 30 μ g total RNA from prelactational mammary gland
 - 18) 30 μ g total male kidney RNA
- Panel a, lane
- 14) 3ng tT28EB
 - 15) 1ng tT28EB
 - 16) 0.03ng tT28EB
 - 17) 30ng tT7-MUP11
- Panel b, lane
- 14) 3ng tT21PB
 - 15) 1ng tT21PB
 - 16) 0.03ng tT21PB
 - 17) 30ng tT7-MUP11
- Panel c, lane
- 14) 30ng tT7-6-5-5
 - 15) 10ng tT7-6-5-5
 - 16) 3ng tT7-6-5-5
- / Continued p.81



17) 30ng tT7-MUP11

The BL1, BS1 and BS6 positive control RNAs are 730nt, 770nt, and 1011nt in length, respectively (Table 2).

Abbreviations:-

T = Testosterone-induced female liver RNA

M = Prelactational mammary gland RNA (14 days gestation)

clone (cLIV6) from a female liver library and from a library prepared from prelactational mammary gland which is identical to the exonic sequence of BL1. Eight out of nine MUP cDNAs isolated from this liver library were identical in sequence (Shahan et al, 1987a). Of five group 1 clones from the mammary gland library, two were sequenced and found to be identical to the group 1 clones from the liver library (cLIV6, Shahan et al, 1987b). On this basis these authors propose that this is the most abundantly expressed group 1 gene in female liver and in the prelactational mammary gland.

A BALB/c genomic clone (BJ-31H; Held et al, 1987) which is identical in its restriction map and very similar in its 5' flanking sequence to BL1, has been used to transform L cells. The protein synthesised in these cells co-migrates with a MUP species present in both male and female BALB/c liver in vitro translation products. Because they have identical coding sequences BJ-31H and BL1 may be clones of the same gene or, if they are clones of different genes, then they probably resulted from a recent duplication event and would encode electrophoretically indistinguishable proteins.

BS1 probe

The BS1-specific probe has a higher G+C content than the BL1 probe and a correspondingly higher estimated T_m (50°C). Hybridisation and washing in a shaking water bath at this temperature did not discriminate between positive and negative control transcripts. If hybridisation and washing were carried out at 53°C then it was possible to discriminate clearly between the controls. Under these

conditions the probe identified transcripts in both male and female liver, although the levels in female liver were lower relative to male liver than is the case for the BL1 probe (compare Figures 10a & 10b; Table 7). Examination of the original autoradiograph shows that the BS1 probe is sensitive enough to detect 0.03ng of positive control RNA (corresponding to approximately 50 copies per cell of mRNA) but this is too faint to quantify by scanning densitometry. The signal is of approximately the same intensity as that seen for 30µg of liver RNA from the female with the lowest level of MUP expression. At this level of sensitivity no transcripts are detectable in prelactational mammary gland RNA.

The exonic sequence of BS1 is identical to the male liver cDNA clone cLIV1 (Shahan et al, 1987a). This is one of four cDNA clones, identical in coding sequence, which appears to be the reverse transcript of an RNA molecule ^{5'}initiated upstream of the usual cap site since it includes some 5' flanking sequence up to and slightly beyond the TATA box. There is a single nucleotide difference in this region (Shahan et al, 1987a). As described in chapter 4, primer extension and S1 nuclease protection experiments provide no evidence for an alternative transcription initiation site and so the origin of this clone remains unclear. There may be a very weak upstream promoter which occasionally directs transcription from an alternative cap site. It is possible that cLIV1 is the cDNA corresponding to a gene which is related to BS1 in much the same way that BL1 is related to BJ-31H, i.e. that in each case the two genes are the result of a recent duplication event and have not diverged.

cont. on p 86

Table 7

The relative levels of group 1 Mup gene transcripts

Probe	Male Liver RNA (copies/cell)	Female Liver RNA (copies/cell)				
		T	1	2	3	4
Group 1	58,000	60,300	580	8,200	4,600	4,300
BL1	39,000	63,000	230	9,100	2,900	1,700
BS1	14,000	15,000	<14	150	48	25
BS6	15,000	13,000	0	0	0	0

The values were obtained by densitometric scanning of the autoradiographs shown in Figures 6 & 10, with an integrator, and relating the signal from male liver samples to the positive controls allowing for different lengths of control transcripts (as described in Appendix). Results are expressed as copies/cell, correct to two significant figures.

The female liver RNA samples are marked 1-4 as in Figure 6.

T = testosterone-induced female liver RNA.

cont from p 83

BS5 probe

The G+C content of the BS5-specific probe is very low and the large proportion of dA:rU base-pairs (8 of 18) in the DNA:RNA duplex suggests that this oligonucleotide would have quite a low T_m when paired with complementary RNA (see group 2 consensus probe, chapter 4). Furthermore, the mismatched base-pair is rG:dT which has previously been shown to be quite stable (see Smith, 1983; chapter 3). Unfortunately, this is the only position at which BS5 is known to differ from other group 1 genes. These two factors suggest a priori that the BS5-specific probe may not give a clear result. This was indeed the case: it was not possible to find a temperature at which the positive signal was any greater than the negative and so the expression of the gene BS5 could not be examined.

BS6 probe

The G+C content of the BS6-specific probe is 50% and the estimated T_m is 52°C. Hybridisation at this temperature did not allow discrimination between positive and negative control RNAs. With hybridisation and washing at 57°C the negative control signal was a trace of the positive control (Figure 10c). This is not ideal but as the positive signal itself was markedly lower after hybridisation at this higher temperature it seems likely that the hybridisation reaction is near equilibrium. Consequently more stringent washing conditions were not attempted. Again a clear signal is obtained from male and testosterone-induced female liver RNA. The signal from uninduced female liver RNA was similar to that from the negative

control; after a shorter exposure of the autoradiograph, neither was visible. No transcripts were detected in prelactational mammary gland RNA.

5' flanking DNA from the BS6 gene has been used to direct expression of an HSV-tk reporter gene in transgenic mice (Al-Shawi et al, 1988b). Hybrid gene mRNA and enzyme activity could be detected in the liver of male transgenic mice and in the livers of testosterone-induced females. Large amounts of transgene RNA were also present in the testes and male preputial glands (see chapter 1).

The Relative Expression of Group 1 MUP Genes

The relative proportion of each group 1 transcript can be estimated by densitometric scanning and integration (Table 7). Each track was scanned at least three times and the figures given in the Table are the averages of at least two determinations from different concentrations of control RNA (see Appendix for an example). The Table shows that transcripts of BL1 are the most abundant MUP RNA species in both male and female liver. Furthermore, BS1 and BS6 are also expressed in male liver and BS1 is expressed in untreated female liver, although at a much lower level than BL1. The expression of each gene in females is induced to male-like levels by testosterone treatment.

The variation in MUP mRNA levels between individual female mice noted using MUP11 and group 1 probes (chapter 4) is observed with each of the gene-specific probes. This implies that the variation is in MUP

gene expression as a whole rather than any particular group 1 gene.

Four out of five male liver MUP cDNA clones were identical to cLIV1 and of the nine female liver MUP cDNA clones, eight were identical to cLIV6 (Shahan et al, 1987a). Shahan et al propose that these clones represent the major MUP mRNAs in the livers of male and female mice. Hybridisation with oligonucleotide probes was also used to investigate the liver RNA of both sexes (Shahan et al, 1987b).

The probes used do not discriminate between cLIV1 and cLIV6. Consequently, as they note, the hybridisation results show that each may be present in the liver RNA of the opposite sex and that related genes may be expressed in the liver. The data presented above clearly show that this is the case. Expression of both BL1 and BS1, identical in exonic sequence to cLIV6 and cLIV1 respectively, is detected in both male and female liver. BS6, a gene for which no cDNA has yet been identified, is also expressed in male liver and furthermore, the group 1 cDNA clone, MUP11, was isolated from a female liver library (Clark et al, 1985b). Thus three different group 1 genes have been shown to be expressed in male and female liver. Possible relationships between genes and the MUP products observed on iso-electric focussing gels are discussed in chapter 6.

Longer exposure of the autoradiograph shown in Figure 10a revealed a faint band of the expected size in prelactational mammary gland RNA. This confirms the report of Shahan et al (1987b) that the MUP component of mammary gland RNA is the same as the most abundant MUP component of female liver. Any BS1 transcripts in the mammary gland are present at levels below the sensitivity of this assay (50 copies

per cell). Because longer exposures reveal cross-hybridisation of the BS6 probe to negative control RNA, it is not possible to rule out expression of this gene at low levels in female liver and the prelactational mammary gland.

No mRNAs identified with the gene-specific probes clearly corresponded to the shorter message seen with the group 1 consensus probe or the MUP11 cDNA probe (Figures 6 & 10). This suggests that other group 1 genes are transcribed and spliced to generate this shorter RNA species. This supports one of the two hypotheses to explain the relationship of long and short mRNAs (Clark et al, 1984a) since expression of group 2 genes could not be detected by specific oligonucleotides (Chapter 4; Shahan et al, 1987b). It is also possible that as yet unidentified MUP genes, sharing greater homology to group 2 than group 1 genes in the 3' region of exon 6, are transcribed and spliced to give the short mRNA. The group 3 gene, MUP16, entirely lacks exon 7 (Ma, 1987) and would presumably encode a short mRNA. Comparison of exon 6 sequences does not suggest, however, that it would hybridise more strongly to a group 2 than a group 1 sub-clone probe and therefore is unlikely to represent a major portion of non-group 1 short mRNA.

Conclusions

The data presented in this chapter show that BL1 mRNA is the most abundant MUP mRNA in male liver. Gene BL1 is also transcribed in female liver. Gene BS1 is also transcribed in both male and female liver, although at lower levels, and the gene BS6 is transcribed in

male liver at a similar level to BS1. No RNA homologous to the BS6-specific probe can be detected in female liver under conditions which prevent hybridisation to negative control RNA. The expression of all three genes in female liver can be induced to male-like levels by testosterone treatment.

These experiments have shown that it is possible to discriminate between RNA transcripts differing at a single nucleotide using oligonucleotide probes to Northern blots. This not only enables one to investigate the influence of hormones on MUP gene expression but also the differential expression of members of any other multigene family where sequence data is available.

CHAPTER 6

DISCUSSION

DISCRIMINATION BETWEEN SIMILAR TRANSCRIPTS USING OLIGONUCLEOTIDE PROBES

Short synthetic oligonucleotide probes have been shown to discriminate between DNA sequences differing at a single nucleotide (Wallace et al, 1979; Conner, et al, 1983; Kidd et al, 1983; Studencki et al, 1985; DiLella et al, 1986) but have not been widely used as hybridisation probes to RNA sequences (see chapter 2). In some cases oligonucleotide probes cross-hybridise with DNA sequences differing at a single nucleotide (e.g. Wallace et al, 1979; Kidd et al, 1983; DiLella et al, 1986). This may be because hybridisation or washing were performed at too low a stringency. Alternatively, the mismatched bases may be only moderately destabilising (see below) or the particular experimental conditions (hybridising DNA in a dried gel rather than transferred to a filter, or using tetramethylammonium chloride for washes; DiLella et al, 1986) may prevent perfect discrimination. Such cross-hybridisation can usually be tolerated in diagnoses since a positive/negative signal of approximately 10 implies homozygosity for one allele. This is not the case when the objective is to estimate the levels of expression of closely related genes. Consequently particular attention must be paid to hybridisation and washing conditions in the latter case. The results presented in chapter 5 show that in some cases oligonucleotide probes can discriminate between RNA species which differ at only one

nucleotide to a high enough degree. In these cases it is possible to examine the relative expression of genes which share greater than 99% homology.

The Stabilities of Mismatched Base-pairs

Oligonucleotide probes of the same length and similar base composition discriminate between RNA species differing at a single nucleotide with varying efficiencies (Figure 10). The probable reason for this is that different mismatched base-pairs destabilise duplexes to different extents. The stabilities of mismatched base-pairs in duplexes between oligodeoxyribonucleotides and RNA remains to be examined systematically but there is pertinent data from RNaseA cleavage experiments and homopolymer duplex stability determinations (see below). These can be compared with the results obtained from DNA-DNA hybridisation.

Chamberlin and colleagues partly defined the different thermal stabilities of DNA, RNA and RNA-DNA homopolymers (Chamberlin, 1965). The relative instability of the dA:rU base-pair was discovered (see chapter 2). The instability of this base-pair appears to be the most likely explanation for the apparently low T_m of the duplex formed by the group 2 specific oligonucleotide and the group 2 RNA control (chapter 4). This duplex would have seven dA:rU base-pairs and five G:C base-pairs. By comparison, the duplex formed by the group 1 specific oligonucleotide and the group 1 control RNA contains five dA:rU and eight G:C pairs. This explanation is supported by data reported by Buvoli et al (1987) from preliminary studies on two

isoforms of hnRNP protein A1. These authors show that 15nt probes can discriminate between synthetic transcripts differing at a single nucleotide. A duplex with a dA:rU base-pair in the middle is less stable than an otherwise identical sequence with dT:rA in the middle. In RNA duplexes AA:UU is less stable than AU:UA in a nearest neighbour analysis (Tinoco et al, 1973) but it is not known if dAdA:rUrU is less stable than dAdT:rUrA, although one might expect this from experiments with homoduplexes.

The effect of mismatching on short DNA duplexes has been investigated by replacing a single dA residue in pdA_n with either dG or dT and hybridising this to oligo-dT cellulose. The results show that the mismatches significantly reduce the T_m of the duplex and that the nearer to the middle of a duplex a mismatch is, the greater the destabilising effect (reviewed by Smith, 1983). Studies of RNA and DNA helices which showed that duplex stability could be accurately predicted from the base-sequence using nearest neighbour analysis (Tinoco et al, 1971, 1973; Breslaur et al, 1986). Further work by Tinoco and co-workers on the thermodynamics of DNA helix formation using mismatched oligonucleotides has shown that mismatches to cytosine are generally more destabilising than those to guanine, and pyrimidine:pyrimidine mismatches are the least stable (Aboul-ela et al, 1985). Similar experiments with RNA.DNA duplexes have not been reported.

The stability of mismatches in RNA.DNA can be analysed using RNaseA digestion. Cleavage of the RNA strand of an RNA.DNA duplex at mismatched base-pairs by RNaseA has been developed as a method of

detecting new mutations in genes (e.g. Myers & Maniatis, 1986). Not all mismatches are cleaved to the same extent, and some are not cleaved at all (Myers et al, 1985; Myers & Maniatis, 1986). A mismatch within a duplex induces more frequent "breathing" of the duplex. It is presumably the temporary separation of the duplex strands which allows RNaseA to act. Consequently more destabilising mismatches will induce more frequent strand separation and hence are more likely to be digested by RNaseA. As an example, single rC:dA, rC:dC, and rC:dT mismatches are always cleaved by RNaseA but rU:dC, rG:dT, rG:dG, and rA:dA are not. This is partly due to the specificity of RNaseA which cleaves at pyrimidine residues. However, not only is rU:dC completely resistant to RNaseA, but rU:dG and rU:dT are only digested in a proportion of cases. This implies that rU:dC is a stable base-pair and that rU:dG and rU:dT are moderately stable.

The sequence context surrounding the mismatch is important - a given mismatch may be cleaved in one context but not in another - but there is no overall sequence pattern correlated with cleavage (Myers et al, 1985).

A comparison of the data of Aboul-ela et al (1985) and Myers et al (1985) shows that the stability of mismatched base-pairs is partly dependent on whether one or both of the strands in a helix is DNA (Table 8). This emphasises the importance of using RNA controls to determine hybridisation and washing conditions for oligonucleotide probes. It is not possible to use the RNaseA cleavage data when designing oligonucleotide probes to the various group 1 MUP genes because each has only one or two distinguishing point mutations. It

Table 8

Comparison of mismatch stability: DNA.DNA vs RNA.DNA duplexes

	DNA.DNA	RNA.DNA
Stable	G:T 0.0-0.7	rG:dT 0/14
	G:G 0.2	rG:dG 0/6
	G:A 0.3-1.2	rA:dA 0/6
	C:T 1.2-1.5	rU:dC 0/3
	A:A 1.5	rG:dA 1/7
	T:T 1.5	rU:dT 1/4
	A:C 1.9-2.0	rU:dG 3/11
	C:C 2.1	rA:dC 4/12
		rA:dG 2/3
		rC:dC 4/4
Unstable		rC:dT 9/9
		rC:dA 22/22

The data for DNA duplexes is from Aboul-ela et al (1985); the figures represent ΔG° (25°C, kcalmol⁻¹).

RNA.DNA mismatch stabilities are from the RNaseA cleavage data of Myers et al (1985) as summarised by Myers & Maniatis (1986).

The fractions give the number of times more than 50% of duplexes containing these mismatches were digested by RNaseA.

is interesting, however, to compare the results obtained (chapter 5) with what would be predicted by RNaseA cleavage experiments taking the efficiency of RNaseA cleavage as an indication of mismatch instability.

Two of the four group 1 MUP gene-specific oligonucleotide probes give a clear result i.e. no cross-hybridisation to negative control RNA (BL1 and BS1, Figures 10a & 10b). Both mismatches (rC:dT and rC:dC) are efficiently cleaved by RNaseA in most or all sequence contexts (Myers et al, 1985). The BS6-specific oligonucleotide has a rG:dA mismatch when duplexed to negative control RNA and under the conditions employed generates a weak signal (Figure 10c). The RNaseA cleavage data predicts that a rG:dA mismatch would be cleaved only in particular contexts of RNA sequence. This is not due to the pyrimidine mismatches. This may be partly due to the flanking sequence. For example, the BS6-specific oligonucleotide is 50% G+C which may partly stabilise the mismatched duplex.

The only unique mutation in the determined BS5 exonic sequence is a dG→dA transition in exon 2 (Figure 2). This generates an rG:dT mismatch in the heteroduplex between the BS5-specific oligonucleotide and the negative control RNA. It was not possible to establish hybridisation and washing conditions under which this probe was discriminatory (chapter 5). This might be predicted by RNaseA cleavage data which shows that the mismatch is stable and resistant to RNase attack (Myers et al, 1985).

These results would not be predicted by extrapolating analysis of

mismatches in DNA duplexes. This data would predict the unstable C:C mismatch but in certain contexts dC:dT and dG:dA are quite stable (Table 8, Aboul-ela et al, 1985). A dG:dT mismatch is less stable than rG:dT or rG:rU, presumably because of the required fidelity of base-pairing required in DNA replication (Smith, 1983). Of the more stable DNA.DNA mismatches, G:T and G:A are known to be paired by hydrogen bonds (Aboul-ela et al, 1985). These authors also propose configurations for G:G, A:A and A:C mismatches but any hydrogen bonds in C:C, T:T or T:C mismatches would require geometrics quite different from Watson-Crick base-pairs.

The results of Buvoli et al (1987) do not agree with the predicted stabilities of Myers et al shown in Table 8. These authors show that dA:rA and dT:rU mismatches destabilise duplexes with 15nt probes at a temperature 1-2°C above the estimated T_m . It should be noted, however, that in this example the mismatched base has an rU residue on either side: such a sequence context was not investigated by Myers et al.

Estimating the T_m of a Short RNA.DNA Duplex

As mentioned in chapter 2, equations which estimate the T_m of a duplex are not directly applicable to all duplexes formed between oligodeoxyribonucleotides and RNA. Table 9 compares the T_m values predicted by the modified Schildkraut & Lifson relation (chapter 2, equation 3) and Suggs' empirical relation with the temperatures at which hybridisations and washes were carried out to discriminate between sequences differing at a single nucleotide (chapter 5). A

Table 9

The estimated T_m and actual washing temperatures (T_w) of the oligonucleotide probes

Probe	%G+C	dA:rU	Estimated T_m		T_w (°C)
			A °C	B	
Gp1	44	5	50	52	43
Gp2	28	8	43	46	25
BL1	39	6	48	50	43
BS1	44	5	50	52	53
BS6	50	6	52	54	57

Estimated T_m values A:- from the modified Schildkraut & Lifson relation (chapter 2, equation (3)) with $[Na^+] = 0.549$.

B:- from the empirical relation of Suggs et al (1981).

T_w :- the temperature at which hybridisation and washing were performed

stringent wash would normally be performed approximately 2°C below the estimated T_m . It can be seen from the comparison in Table 9 that in some cases this would be too stringent, in others, too relaxed. All probes are 18 nucleotides in length and each was used in the same hybridisation and washing solutions, so the only variables affecting T_m are base composition and homology. Because the group 1 consensus probe differs from other sequences by at least 3/18 nucleotides, it does not require such stringent conditions. Therefore although the group 1 consensus and the BS1-specific probes are both 44% G+C, the difference in homology allows the former to be used at a lower temperature.

The consequence of changing the base composition is highlighted by comparing results with the BL1- and BS1-specific oligonucleotides. Although both probes have quite destabilising mismatches (see above), the former discriminates at a temperature approximately 5°C below the estimated T_m , while the latter, apparently differing only in base composition, does not discriminate until about 3°C above the estimated T_m . The BS6-specific probe (50% G+C) gives good but not absolute discrimination at a temperature 5°C above the estimated T_m . Two 15nt probes (33% G+C) have been shown to discriminate between transcripts differing at a single nucleotide after a final wash at a temperature 1-2°C above the estimated T_m (Buvoli et al, 1987).

The data presented in chapter 5 and discussed above also suggest that even the most destabilising mismatches (e.g. rC:dC) do not reduce the T_m by the predicted amount. Experiments with larger fragments

have estimated that a 1% mismatch would reduce the T_m by 1°C (Bonner et al, 1973). This predicts a reduction in T_m of approximately 5°C for a duplex with 1/18 mismatched base-pairs. The results obtained with the BS1- and BS6-specific probes show that this is not the case. It may be that the destabilising effect of a single mismatch is reduced or increased by the base composition, and/or the sequence, of the duplex.

With short probe length, therefore, the base composition has a greater effect on the stability of RNA.DNA duplexes. This may be partly due to the weakness of the dA:rU base-pair or different sequence contexts. The best method of predicting the stability of different duplexes is by nearest neighbour analysis of each sequence (Tinoco et al, 1971, 1973; Breslaur et al, 1986), but thermodynamic data for RNA.DNA base-pairs is not available.

There is insufficient experimental data to adjust the Schildkraut & Lifson relation to make it applicable to RNA.DNA duplexes. Increasing G+C content probably stabilises an RNA.DNA duplex by more than the predicted amount because rG:dC and dG:rC are both more stable than dG:dC (Chamberlin, 1965), while duplexes with a low G+C content may be less stable than expected since dA:rU and rA:dT are both less stable homoduplexes than dA:dT.

THE TISSUE-SPECIFIC AND SEXUALLY DIMORPHIC EXPRESSION OF GROUP 1 MUP GENES

Tissue-specificity of Group 1 Gene Expression

The MUP multigene family is a useful system in which to study eukaryotic gene expression. The expression of different genes within the family is modulated by various hormones. Regulation is primarily at the level of transcription (Derman, 1981) and varies in a developmental and tissue-specific manner (Shaw et al, 1983). The classification of MUP genes into phylogenetic groupings on the basis of restriction site and nucleic acid sequence homology (Al-Shawi et al, 1988a) is supported by the tissue-specific expression and hormonal regulation of genes from each group.

Group 1 genes are expressed in the liver under the influence of testosterone, growth hormone and thyroxine and also in the prelactational mammary gland, probably under the influence of the female sex hormones (Knopf et al, 1983; Shaw et al, 1983).

The influence of oestrogen and progesterone on MUP gene expression in female liver outwith and during pregnancy has not been reported. Oestrogen administration is known to reduce the levels of $\alpha_2\mu$ -globulin mRNA in male rat liver (Kurtz et al, 1976). This may simply be due to antagonism of testosterone action since the hepatic levels of dihydrotestosterone receptor are reduced. No group 1 transcripts are detectable in the subaxillary or lachrymal glands (Figure 5).

Members of group 3 are also expressed in the liver under the influence of testosterone and growth hormone but are not affected by thyroxine administration or withdrawal (Kuhn et al, 1984; Ma, 1987). It has been shown that group 3 gene products are glycosylated and their transcripts represent a relatively greater proportion of female liver RNA than male (Clark et al, 1985b; Ma, 1987). The expression of group 3 genes in other MUP expressing tissues has not been reported. MUP gene expression in the submaxillary gland does not appear to be under hormonal control while that in the lachrymal gland expression is influenced by testosterone (Shaw et al, 1983).

MUP cDNA clones have recently been isolated from submaxillary and lachrymal libraries and found to be quite different from group 1 genes (~10% diverged; Shahan et al, 1987a). Analysis of the amino acid composition of their respective translation products explains the earlier observations: the lachrymal protein would be considerably more basic than any other MUP whose amino acid composition can be predicted and the submaxillary gland protein would have a similar charge to one of the more acidic group 1 proteins (Table 10). These genes have been classified as belonging to group 4 (Al-Shawi et al, 1988a) since they differ markedly from those in groups 1, 2 and 3 at the nucleotide sequence level as well as in tissue-specific expression. It is interesting that iso-electric focusing of hybrid-selected submaxillary and lachrymal gland mRNA in vitro translation products resolves a number of products (Shaw et al, 1983; Shahan & Derman, 1984) and yet only one submaxillary and one lachrymal cDNA were isolated (19 and 11 times each).

Table 10

Amino Acid Composition of Group 1 MUP Genes

Clone	Acidic		(H	Basic		R)	Net Charge
	(D	E)		K			
(BL1	11	21	5	11	9	-7	
(BS1	11	20	5	10	9	-7	
Gp1 (p1057	11	20	5	10	9	-7	
(MUP11	11	19	5	9	8	-8	
(BS6	11	21	5	9	9	-9	
Gp3 (MUP15	7	24	4	10	7	-10	
(MUP16	8	24	3	9	8	-12	
Gp4 (cLAC1	10	20	6	17	5	-2	
(cSMX1	11	23	5	13	7	-9	

The table shows the number of each acidic and basic amino acid present in the mature protein product of each of the BALB/c clones as derived from the nucleotide sequences (Figure 2).

Abbreviations:- D - Aspartic acid

E - Glutamic acid

H - Histidine

K - Lysine

R - Asparagine

No expression of group 2 pseudogenes has been detected using two different oligonucleotide probes (Figure 9; Shahan et al, 1987b).

Sexually-dimorphic Expression of Group 1 MUP Genes

The relative levels of different group 1 gene transcripts appear to be influenced by various hormones in the liver (Knopf et al, 1983). To study the regulation of individual MUP genes in vivo, hybridisation probes specific for each gene are required. The work presented in this thesis has shown that synthetic oligonucleotide probes can discriminate between RNA sequences differing at only one nucleotide. This has enabled determination of the relative levels of expression of three group 1 MUP genes (BL1, BS1 and BS6) in male liver. It has been shown that two of these (BL1 and BS1) are also expressed in female liver at quite different levels, and all three are induced to male-like levels in the livers of testosterone-treated females. RNA homologous to the BL1-specific probe is the most abundant MUP RNA in both male and female liver.

Iso-electric focusing of the urine from males and females of BALB/c and C57BL strains shows that the more abundant MUPs in BALB/c urine are more basic than the abundant C57BL proteins. It is also clear that the abundant female product is more basic in BALB/c mice (Figure 11; Clissold et al, 1984). Analysis of the coding regions of the MUP genes and cDNAs sequenced to date allows them to be placed in sub-groups on the basis of predicted net charge (Table 10). It has not yet been possible to correlate a cloned gene to its encoded protein because two or more different genes could encode proteins with the

Figure 11

Putative association of urinary MUPs with individual clones

The figure showing isoelectric focusing of mouse urine is adapted from Clissold et al (1984). The anode is at the top.

M = male urine

T = testosterone-induced female urine

F = female urine

The bands marked 1-5 in C57BL male urine correspond to spots 1-5 on the two-dimensional gels of Knopf et al (1983) as described in the text. Bands a-f in BALB/c male urine may be the products of the following MUP sequences:-

a	p1057
b,c	BS1, BL1
d	MUP11
e or f	BS6

These associations are based on the presence or absence of homologous RNA in female liver RNA (Figure 10; MUP11 was isolated from a female liver cDNA library) and the predicted net charge of the translation products (Table 10). The alternative association of BS6 with bands e or f is because of the possible low level expression of this gene in female liver.

C57BL

BALB/c

M

T

F

T

F

M

1

2

3

4

5

a

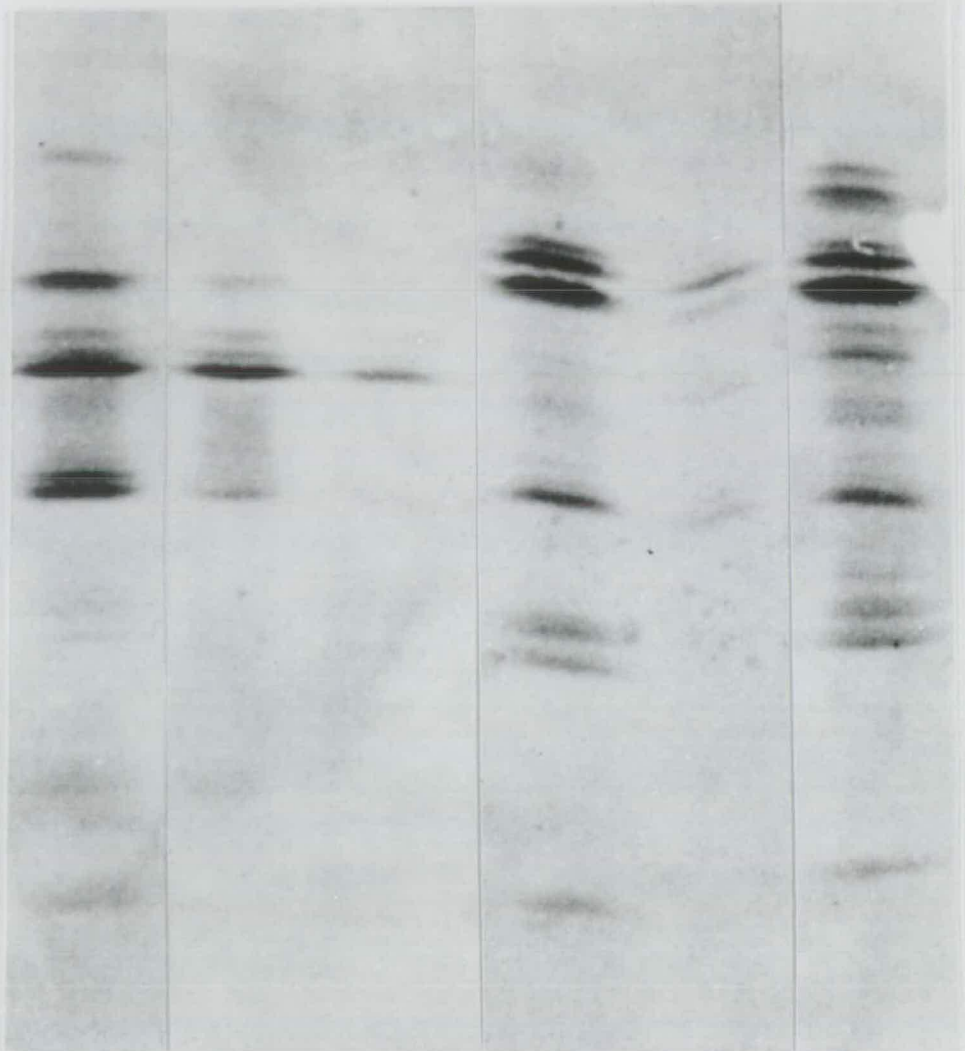
b

c

d

e

f



same electrophoretic mobilities. The bands in C57BL male urine can be associated with the spots seen on a two-dimensional gel (Figure 11) by comparing their modes of expression and hormonal regulation. MUP3 is present in female urine and MUP 1 and MUP 5 are not present in the liver in vitro translation products of testosterone-induced females (Knopf et al, 1983).

The data of Held et al (1987) link a number of genes cloned from a C57BL library with the products of mRNA translated in vitro: - MUP2 appears to be encoded by BL6-51, MUP3 by BL6-3 and BL6-11, and MUP4 by BL6-42. Some of these genes share a phylogenetic sub-group with clones from BALB/c mice (Table 11) but they would not encode proteins with similar iso-electric focusing points (as predicted by amino acid composition, Table 10). From the predicted amino acid composition, p1057, BS1 and BL1 would encode the more basic proteins. The data presented in chapter 5 show that the expression of BS1, BS6 and BL1 in female liver is induced by testosterone and BL1 and BS1 are expressed in uninduced females. Furthermore, MUP11 is a cDNA clone from a female liver library (Clark et al, 1985b).

From these observations, the tentative association shown in Figure 11 is proposed. Note that the relative levels of MUP liver RNA do not agree with the abundance of individual MUPs in the urine. The BL1-specific oligonucleotide probe detects a greater proportion of liver RNA, from both sexes, than the BS1-specific probe, but the most abundant band in iso-electric focussed male urine is not the most abundant in the urine of uninduced females. This suggests some degree of post-transcriptional regulation, probably at the level of

Table 11

Phylogenetic Sub-groups of Group 1 MUP Genes

Sub-group	BALB/c Clones	C57BL clones
1a	BS5, BS6	BL6-3
1b	BS1	BL6-42
1c	BL1, BJ-31H	BL6-11
1d	-	BL6-51

From Al-Shawi et al (1988a).

translation or protein secretion or transport. Indeed, Berger & Szoka (1981) have shown that the testosterone induction of MUP levels in the liver parallels that for RNA, rather than the 30- fold difference observed in urinary MUP levels (Szoka & Paigen, 1978). These authors suggest that this post-translational control affects the excretion of all MUPs, but as their assay could not discriminate between the individual proteins there may be qualitative control at this level.

It is possible that BS6 is expressed in female liver but at levels below the detection and discrimination limits of the specific oligonucleotide probe. It is for this reason that two alternative associations are proposed for the more acidic proteins (Figure 11). The lowest, most acidic protein is probably the glycosylated product of a group 3 gene. The questions raised by these comparisons are currently being investigated by transcribing MUP cDNA's in vitro, translating the RNA and examining the protein products on two-dimensional gels.

Equating the sexually dimorphic expression of group 1 MUP genes with their phylogenetic classification is not straightforward. Genes expressed in both male and female BALB/c liver are present in sub-groups 1b and 1c (BS1 and BL1, respectively; Table 11). The gene (or genes) encoding the predominant MUP in C57BL female urine could be a member of either sub-group 1a or 1c (BL6-3 or BL6-11, respectively), as determined by co-migration of in vitro translation products (Held et al, 1987). It may be that the genetic differences affecting the sexually dimorphic expression of individual group 1 genes are as subtle as those determining electrophoretic mobilities of

the gene products i.e. a small number of base substitutions. Such differences may also be responsible for the different responses of MUP genes to hormones such as thyroxine. MUP gene sequences are highly conserved in the 5' flanking region as well as the coding region (98% homology; Ghazal, 1986; Held et al, 1987).

The work presented in this thesis together with results from other laboratories discussed above has identified certain members of the MUP multigene family which are expressed in a sex- and tissue-specific manner. The transcripts of individual group 1 genes can be detected by gene-specific oligonucleotides and it is now possible to identify the genes whose expression is influenced by hormones other than testosterone e.g. thyroxine and glucocorticoids. Once these genes have been identified it is possible to determine the sequence elements involved in the hormonal regulation.

CHAPTER 7

MATERIALS AND METHODS

Animals

All mice used in this work were of the BALB/c strain maintained in the Department of Genetics, University of Edinburgh.

Enzymes

All restriction and modifying enzymes and radioisotopes were purchased from Amersham International and used with the buffers supplied or as directed by the supplier (unless otherwise stated).

RNA Controls

Template DNA (2 μ g), restricted with the appropriate enzyme (Table 1) was transcribed in a 50 μ l mix of 40mM Tris-Cl pH8.0, 15mM MgCl₂, 5mM DTT, 500 μ g/ml BSA (nuclease free), 1mM each rNTP, 2 μ Ci [5-³H]-UTP or ³²- α -UTP, 50 units RNase inhibitor (Promega or Pharmacia) and 10-30 units T7 RNA polymerase (US Biochemicals or Boehringer). Incubation was at 37°C for 45 min. RNA yield was determined from the TCA precipitable counts in a 2 μ l aliquot. RNA was purified by digesting the template DNA with RNase-free DNaseI (Worthington; 4 units, 37°C, 10min), phenol/chloroform extraction and Sephadex 6100 chromatography (in 10mM Tris-Cl pH7.0, 0.1% SLS) and precipitated from 0.3M NaOAc pH6.5 and ethanol (Melton et al,

1984). Pellets were resuspended in sterile H₂O, final yield determined by TCA precipitation, and the RNA stored frozen at -20°C. Samples (2µl) for polyacrylamide gel electrophoresis were denatured by boiling in an equal volume of formamide dye mix for 5mins. 4% polyacrylamide (29:1 acrylamide:bis-acrylamide), 7.6M urea, 1xTBE pH8.8 gels were pre-run at 30W for 1 hour before samples were loaded. Electrophoresis was at 30W for 2 hours.

Total RNA Samples

Total cellular RNA was prepared by the method of Chirgwin et al (1979) including a CsCl cushion. Total RNA from testosterone-induced female liver was a gift from X. Ma. The mice were induced with subcutaneous pellets (30mg) for 2 weeks. Total mammary gland RNA was prepared from mice at 14 days gestation, liver RNA from 7-8 week old mice and submaxillary and lachrymal gland RNA from 8-9 week old mice.

Gel Electrophoresis of RNA

RNA samples were denatured in 24% formamide, 5% formaldehyde, 10mM NaHPO₄ pH7.0 at 60°C for 5min and electrophoresed in 1.8% agarose, 6% formaldehyde, 10mM NaHPO₄ pH7.0 gels at 0.17 V/cm/hr. The RNA was then transferred to Hybond-N membranes (Amersham) by capillary blotting from 20 x SSC and covalently cross-linked by U.V. irradiation (Church & Gilbert, 1984).

Synthetic Oligonucleotides

The synthetic oligonucleotides were purchased from the Oswel DNA Service, Chemistry Dept, Edinburgh University, except for the group 2-specific probe which was a gift from Dr. B.J. Hammill (Cruachem Ltd.). Concentrations of oligonucleotide in water were determined by measuring the A_{260} of the solution and taking the extinction coefficients of each base to be: - dA, 15.4; dG, 11.7; dC 7.3; dT, 8.8 $\text{cm}^2/\mu\text{mol}$; and the molecular weights of the monophosphates to be 331.2, 347.2, 307.2 and 322.2, respectively. Stocks and 5 $\mu\text{g}/\text{ml}$ working solutions were stored at -20°C . Labelling with ^{32}P - γ -ATP was performed largely as described by Maxam & Gilbert (1980). Equimolar amounts of ^{32}P - γ -ATP (3,000 Ci/mmol) and oligonucleotide were mixed in 50mM Tris-Cl pH7.5, 10mM MgCl_2 , 5mM DTT, 0.1mM spermidine, 0.1 mM EDTA, with 10 units of T4 polynucleotide kinase (final volume, 50 μl) and incubated at 37°C for 30 min. Efficiency of incorporation was determined by spotting an aliquot onto 3MM paper discs, washing in four changes of cold 5% TCA and drying with ethanol and ether. Specific activities of between 3×10^8 and 2×10^9 dpm/ μg were routinely obtained by this method. Labelled DNA was phenol/chloroform extracted and ethanol precipitated by adding 4 vols of 2.5M NH_4OAc , 2 μg Dextran 2000 as carrier, and 3 vols ethanol. Tubes were kept at -20°C for at least 1hr and spun in an SS34 rotor at 15,000rpm, -20°C for 30min.

Hybridisation of Oligonucleotide Probes.

Filters were prehybridised in 0.5M NaHPO₄ pH7.2, 7% SLS, 2mM EDTA for 1hr and then hybridised overnight in the same solution plus labelled probe. Washes were as follows:- 2 x 10min in 0.5M NaHPO₄ pH7.2, 5% SLS, 1mM EDTA; 2 x 20min in 0.5M NaHPO₄ pH7.2, 1% SLS, 1mM EDTA. Hybridisations and washes were performed in either a shaking water-bath or on a rotating table in an oven at the temperatures given in the text.

Labelling and Hybridisation of cDNA Probes

The double-stranded DNA probes described in the text were labelled by a variation of the random primer method (Feinberg & Vogelstein, 1983) as follows:- 0.2µg of linearised plasmid was boiled in water with 120ng of p(dN)₆ (Pharmacia) for 3min and quenched in ice. Labelling was with 30-50µCi ³²P-α-dCTP and 2.5 units *E. coli* DNA polymerase I at 25°C for 1hr in 66mM Tris-Cl pH7.4, 6mM MgCl₂, 2.5mM DTT, 30µM each dATP, dGTP, dTTP (final volume 100µl). The efficiency of incorporation was determined by TCA precipitation of an aliquot. Specific activities in excess of 10⁸dpm/µg were obtained. Labelled DNA was phenol/chloroform extracted and ethanol precipitated with 25µg of sonicated salmon sperm DNA as a carrier. Prehybridisation and hybridisation were in the solutions described above at 68°C. Washes were 2 x 10min 40mM NaHPO₄ pH7.2, 5% SLS, 1mM EDTA; 2 x 20min 40mM NaHPO₄ pH7.2, 1% SLS, 1mM EDTA at 68°C.

Primer Extension Sequencing

Primer extensions were carried out according to Epstein et al (1986). Either 5ng of control RNA or 4µg of total male liver RNA was annealed with 2.5ng of ³²P-labelled oligonucleotide probe (or 25µg of total RNA with 15ng of primer) in 250mM KCl, 1mM EDTA, 10mM Tris-Cl pH8.0 (final volume 15µl) at 65°C for 90min and allowed to cool to 30°C. The primers were extended in a 50µl reaction after adjusting to 80mM KCl, 0.3mM EDTA, 18mM Tris-Cl pH8.3, 10mM MgCl₂, 5mM DTT, 0.2mM each dATP, dGTP, dCTP, dTTP, with either 2 units or 5 units of AMV reverse transcriptase (Pharmacia; depending on the amount of RNA template) at 37°C for 1hr. Separate reactions were performed with each dideoxynucleotide at 0.1mM. Extension products were precipitated by adding 5µl 3M NaOAc, pH6.5, and 165µl ethanol, leaving to stand at -70°C for 1hr, and spinning in a microfuge at 10,000rpm, 4°C for 15min. Pellets were dried under vacuum and resuspended in 2µl water and heated to 50°C for 5min. An equal volume of formamide dye mix was added and the samples were boiled (6min) prior to electrophoresis on a 10% polyacrylamide (29:1 acrylamide:bis-acrylamide), 7M urea gel, 1 x TBE pH8.8 for 3hr. Gels were fixed for 10min in 10% methanol, 10% acetic acid and exposed directly to X-ray film at room temperature.

General Methods

Phenol/chloroform extraction

Enzyme reactions were phenol/chloroform extracted, after the addition

of sufficient EDTA to chelate Mg^{2+} , by adding 0.5 vol of water-saturated phenol and incubating at 37°C for 5min; 0.5 vol chloroform was added and the mixture centrifuged for 3min. The aqueous layer was extracted twice more with 1 vol of chloroform and the organic layers were back-extracted with 0.5 vol of 0.1M Tris-Cl pH7.5 and this was added to the aqueous layer. DNA was recovered by precipitation from ethanol.

DNA gel electrophoresis and fragment purification

Restriction digests were assayed by electrophoresis of an aliquot on an agarose gel of the appropriate density (0.5% to 2.0% w/v) containing 1µg/ml ethidium bromide in 1 x TA pH7.9 at 100 - 200V. DNA fragments were purified by electrophoresis in agarose gels run in 1 x TBE pH8.3 at 0.17V/cm/hr and electro-eluted onto dialysis membrane, purified by ion-exchange chromatography using a Scheicher & Schuell Elutip-d column and ethanol precipitated.

Ligation, transformation and plasmid DNA preparation

Ligations were performed in 20µl of 66mM Tris-Cl pH7.6, 10mM $MgCl_2$, 40mM NaCl, 1mM EDTA, 5mM DTT, 0.25mM ATP with 1 unit T4 DNA ligase (2 units for a blunt end) at 14°C for 2hr with a 3x, 5x, or 10x ratio of insert:vector at 5ng vector/plate. The ligation mixes were used to transform competent E. coli strains HB101 (pT7 plasmids) or NM522 (pTZ phagemids) using X-gal (50µl of 20mg/ml) and IPTG (20µl of 24mg/ml) in the top agar as required. Transformants were selected by ampicillin resistance. Plasmid DNA was prepared as

described (Bishop, 1979) except that only one CsCl gradient step was performed, dialysis was against 10mM Tris-Cl pH7.5 and contaminating bacterial DNA was removed by Sepharose 2B chromatography with 50mM Tris-Cl pH7.5, 0.3M NaCl.

HB101: F⁻, hsdS20 (r⁻,m⁻),recA13, ara-14, proA2,
lacY1, galK2, cpdL20(Sm^r), xy1-5, mtl-1, supE44

NM522: hsdΔ5, Δ(lac, pro), F' lacZM15, lacI^q, pro⁺

Solutions

1 x SSC	is 0.15M NaCl, 15mM Na ₃ citrate
1 x TA pH7.9	is 50mM Tris-base, 20mM NaOAc, 2mM EDTA, 10mM NaCl adjusted to pH7.9 with glacial acetic acid
1 x TBE pH8.3	is 89mM Tris-base, 89mM boric acid, 2mM EDTA
1 x TBE pH8.8	is 0.13M Tris-base, 89mM boric acid, 2mM EDTA
NaHPO ₄ pH7.0	is 83g/l Na ₂ HPO ₄ , 62g/l NaH ₂ PO ₄ .2H ₂ O
NaHPO ₄ pH7.2	is 70.6 g/l Na ₂ HPO ₄ , 1.2% v/v H ₃ PO ₄
Formamide dye mix	is 7.5mg/ml EDTA, 0.3mg/ml bromophenol blue, 0.3mg/ml xylene cyanol FF, in double-distilled formamide

APPENDIX

DETERMINATION OF THE RELATIVE AMOUNTS OF SPECIFIC MUP RNAs BY DENSITOMETRIC SCANNING

The relative levels of MUP gene expression shown in Tables 4,5 & 7 were determined by scanning densitometry with an electronic integrator. As an example, the raw data obtained from the autoradiograph probed with the BL1 specific probe (Figure 9a) are presented here (Table A1). The figures in the bottom two rows of the table were obtained using a fainter wedge. Only results obtained from the same autoradiograph with the same wedge are directly comparable.

To determine the proportion of male liver RNA that is homologous to this probe the values from the male liver tracks and the positive control tracks were compared. It is necessary to allow for the different lengths of each control transcript (Table 2). Each track was scanned three times and the readings (less background) were averaged. Furthermore, readings from tracks containing different amounts of the same RNA sample were averaged after allowing for the amount of RNA.

From Table A1,

3 μ g of male liver RNA gives a value of 139

1 μ g of male liver RNA gives a value of 33

3ng of positive control RNA gives a value of 52.3

Using these figures

$$\frac{139 \times 900}{52.3 \times 730} = 3.3\text{ng}/\mu\text{g} \text{ or } 0.33\%$$

and

$$\frac{33 \times 3 \times 900}{52.3 \times 730} = 2.3\text{ng}/\mu\text{g} \text{ or } 0.23\%$$

an average value of 0.28% is obtained.

The values for each sample of female liver RNA relative to male levels were determined in a similar manner allowing for the different amounts of RNA in each track but not, of course, for transcript length. For example, for female no. 2:-

30µg of female no. 2 RNA gives a value of 284.3

10µg of female no. 2 RNA gives a value of 86.3

Male values are as above.

Using these figures

$$\frac{284.3}{139 \times 10} \\ \Rightarrow 20.4\%$$

$$\frac{86.3}{33 \times 10} \\ \Rightarrow 26\%$$

$$\frac{86.3 \times 3}{33 \times 10} \\ \Rightarrow 18.6\%$$

and

$$\frac{284.3}{33 \times 30} \\ \Rightarrow 28.7\%$$

an average value of 23.4% is obtained.

It was not possible to quantify the weak signal from mammary gland RNA. These figures can be converted to copies/cell using the following approximations (Hastie & Bishop, 1976). There are approximately 5×10^5 mRNA molecules per liver cell and these constitute 3.6% of cytoplasmic RNA. This gives a figure of 1.4×10^7 RNA molecules per liver cell. Using these figures, it was found that at 0.28% of total RNA in male liver, BL1 transcripts are present at approximately 39,000 copies per cell. At 23.4% of the male level, female no. 2 therefore has about 9,100 copies per cell of BL1 transcripts. Similar calculations were performed on the data obtained

scanning the autoradiographs generated with the other probes to give the results shown in Tables 4,5 & 7.

Table A1

Sample	Reading on integrator			Background reading			Average less background
	1	2	3	1	2	3	
3µg T	230	229	229	13	13	13	216.3
3µg M	149	150	149	11	10	10	139
1µg M	39	39	39	6	6	6	33
3ng control	59	60	59	7	7	7	52.3
30µg F2	288	290	290	5	5	5	284.3
10µg F2	89	89	88	3	2	2	86.3
30µg F3	107	107	107	21	21	21	86
30µg F4	72	72	72	20	21	21	51.3
0.3µg M	50	50	50	2	2	2	48
30µg F1	79	82	81	54	52	51	28.3

M = male liver RNA

F1-F4 = female liver RNA samples 1-4 (as described in chapter 4 and Figure 5)

T = testosterone-induced female liver RNA

Other details in text.

ABBREVIATIONS USED IN TEXT

A	adenine
AMV	avian myeloblastosis virus
ATP	adenosine 5' triphosphate
bp	base-pair
BSA	bovine serum albumin
C	cytidine
cDNA	DNA copy of RNA
d	deoxyribo-
DHFR	dihydrofolate reductase
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dpm	depositions per minute
DTT	dithiothreitol
EDTA	diaminoethanetetra-acetic acid
G	guanidine
h	% homology
hnRNP	heterologous nuclear ribonucleoprotein
HPRT	hypoxanthine phosphoribosyl transferase
hsp	heat shock protein
HSV	herpes simplex virus
IEF	isoelectric focusing
IGF-I	insulin-like growth factor I
IPTG	isopropyl- β -D-thiogalactoside
kb	kilobase pair
λ	probe length
LTR	long terminal repeat

M ⁺	monovalent cation
MMTV	mouse mammary tumor virus
mRNA	messenger RNA
MUP	major urinary protein
MW	molecular weight
nt	nucleotide
OAc	acetate
pH	-log [H ⁺]
poly(A) RNA	polyadenylated RNA
r	ribo-
rNTP	ribonucleotide triphosphate(s)
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
S1	single strand specific nuclease
SDS	sodium dodecyl sulphate
SLS	sodium lauryl sulphate
Sp1	specific transcription factor 1
SV40	simian virus 40
T	thymidine
TCA	trichloroacetic acid
T _m	melting temperature of a nucleic acid duplex
Tris	tris-[hydroxymethyl]-aminomethane
U	uridine
UTP	uridine triphosphate
U.V.	ultra-violet
W	A or T
w/v	weight per volume

X-gal

5-Bromo-4-chloro-3-indolyl- β -D-galactopyranoside

REFERENCES

- Aboul-ela, F., Koh, D., Tinoco, I., Jr., & Martin, F.H. (1985). Base- base mismatches. Thermodynamics of double helix formation for $dCA_3XA_3G + dCT_3YT_3G$ ($X,Y = A,C,G,T$). Nucl. Acid Res. 13: 4811-4824
- Addison, W.R., & Kurtz, D.T. (1986). Nucleotide sequences required for the regulation of a rat α_2u -globulin gene by glucocorticoids. Mol. Cell. Biol. 6: 2334-2346
- Ali, S., & Clark, A.J. (1988). Characterisation of the gene encoding ovine beta-lactoglobulin: similarity to the genes for retinol binding protein and other secretory proteins. J. Mol. Biol. 199: 415-426
- Al-Shawi, R.A. (1985). PhD. Thesis, University of Edinburgh.
- Al-Shawi, R., Ghazal, P., Clark, A.J., & Bishop, J.O. (1988a). Intraspecies evolution of a multigene family. J. Mol. Evol. manuscript submitted.
- Al-Shawi, R., Burke, J., Jones, C.T., Simons, J.P., & Bishop, J.O. (1988b). A MUP promoter-thymidine kinase reporter gene shows relaxed tissue-specific expression and confers male sterility upon transgenic mice. Mol. Cell. Biol. in press.
- Anderson, M.L.M., & Young, B.D. (1985). Quantitative filter

hybridisation. in Hames, B.D., & Higgins, S.J. (eds) Nucleic acid hybridisation. A practical approach. pp 73-111 IRL Press Oxford

Barrera-Saldana, H., Takahashi, K., Vigneron, M., Wildeman, A., Davidson, I., & Chambon, P. (1985). All six GC-motifs of the SV40 early upstream element contribute to promoter activity in vivo and in vitro. EMBO J. 4: 3839-3849

Behringer, R.R., Hammer, R.E., Brinster, R.L., Palmiter, R.D., & Townes, T.M. (1987). Two 3' sequences direct adult erythroid-specific expression of human β -globin genes in transgenic mice. Proc. Natl. Acad. Sci. 84: 7056-7060

Benoist, C., O'Hare, K., Breathnach, R., & Chambon, P. (1980). The ovalbumin gene-sequence of putative control regions. Nucl. Acid Res. 8: 127-142

Berent, S.L., Mahmoudi, M., Torczynski, R.M., Bragg, P.W., & Bollon, A.P. (1985). Comparison of oligonucleotide and long DNA fragments as probes in DNA and RNA dot, Southern, Northern, colony and plaque hybridisations. BioTechniques 3: 208-220

Berger, F.G., & Szoka, P. (1981). Biosynthesis of the major urinary proteins in mouse liver: A biochemical genetic study. Biochem Genet. 19: 1261-1273

Bishop, J.O. (1972). Molecular hybridisation of ribonucleic acid with a large excess of deoxyribonucleic acid. Biochem. J. 126:

171-185

Bishop, J.O. (1979). A DNA sequence cleaved by restriction endonuclease R.EcoRI in only one strand. J. Mol. Biol. 128: 545-559

Bishop, J.O., Clark, A.J., Clissold, P.M., Hainey, S., & Francke, U. (1982). Two main groups of mouse major urinary protein genes, both largely located on chromosome 4. EMBO J. 1: 615-620

Bishop, J.O., Selman, G.G., Hickman, J., Black, L., Saunders, R.D.P., & Clark, A.J. (1985). The 45-kb unit of major urinary protein gene organization is a gigantic imperfect palindrome. Mol. Cell. Biol. 5: 1591-1600

Bonner, T.I., Brenner, D.J., Neufeld, B.R., & Britten, R.J. (1973). Reduction in the rate of DNA reassociation by sequence divergence. J. Mol. Biol. 81: 123-135

Breslaur, K.J., Frank, R., Blocker, H., & Marky, L.A. (1986). Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. 83: 3746-3750

Breathnach, R., & Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. Ann. Rev. Biochem. 50: 349-383

Butler, E.T., & Chamberlin, M.J. (1982). Bacteriophage SP6-specific

RNA polymerase I. Isolation and characterization of the enzyme.
J. Biol. Chem. 257: 5772-5778

Buvoli, M., Biamonti, G., Riva, S., & Morandi, C. (1987).
Hybridization of oligodeoxynucleotide probes to RNA molecules:
specificity and stability of duplexes. Nucl. Acid Res. 15: 9091

Capetanaki, Y.S., Ngai, J., Flytzanis, C.N., & Lazarides, E. (1983).
Tissue-specific expression of two mRNA species transcribed from a
single vimentin gene. Cell 35: 411-420

Caruthers, M.H. (1985). Gene synthesis machines: DNA chemistry and
its uses. Science 230: 281-285

Chamberlin, M., & Ring, J. (1973a). Characterization of T7-specific
ribonucleic acid polymerase I. General properties of the enzymatic
reaction and the template specificity of the enzyme. J. Biol.
Chem. 248: 2235-2244

Chamberlin, M. & Ring, J. (1973b). Characterization of T7-specific
ribonucleic acid polymerase II. Inhibitors of the enzyme and their
application to the study of the enzymatic reaction. J. Biol.
Chem. 248: 2245-2250

Chamberlin, M.J. (1965). Comparative properties of DNA, RNA, and
hybrid homopolymer pairs. Fed. Proc. 24: 1446-1457

Chave-Cox, A.J. (1986) PhD. Thesis, University of Edinburgh.

Chen, L.-H., & Bissell, M.J. (1987). Transferrin mRNA level in the mouse mammary gland is regulated by pregnancy and extracellular matrix. *J. Biol. Chem.* 262: 17247-17250

Chin, A.C., & Fournier, R.E.K. (1987). A genetic analysis of extinction: Trans-regulation of 16 liver-specific genes in hepatoma-fibroblast hybrid cells. *Proc. Natl. Acad. Sci.* 84: 1614-1618

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., & Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18: 5294-5299

Church, G.M., & Gilbert, W. (1984). Genomic sequencing. *Proc. Natl. Acad. Sci.* 81: 1991-1995

Clark, A.J., Clissold, P.M., & Bishop, J.O. (1982). Variation between mouse major urinary protein genes isolated from a single inbred line. *Gene* 18: 221-230

Clark, A.J., Clissod, P.M., Al-Shawi, R., Beattie, P., & Bishop, J.O. (1984a). Structure of mouse major urinary protein genes: different splicing configurations in the 3' non-coding region. *EMBO J.* 3: 1045- 1052

Clark, A.J., Hickman, J., & Bishop, J.O. (1984b). A 45-kb DNA domain with two divergently orientated genes is the unit of organisation of the murine major urinary protein genes. *EMBO J.* 3: 2055-2064

Clark, A.J., Ghazal, P., Bingham, R.W., Barrett, D. & Bishop, J.O. (1985a). Sequence structure of a mouse major urinary protein gene and pseudogene compared. *EMBO J.* 4: 3159-3165

Clark, A.J., Chave-Cox, A., Ma, X., & Bishop J.O. (1985b). Analysis of mouse major urinary protein genes: variation between the exonic sequences of group 1 genes and a comparison with an active gene outwith group 1 both suggest that gene conversion has occurred between MUP genes. *EMBO J.* 4 : 3167-3171

Clissold, P.M., & Bishop, J.O. (1981). Molecular cloning of cDNA sequences transcribed from mouse liver endoplasmic reticulum poly(A)mRNA. *Gene* 15: 225-235

Clissold, P.M., & Bishop, J.O. (1982). Variation in mouse major urinary protein (MUP) genes and the MUP gene products within and between inbred lines. *Gene* 18: 211-220

Clissold, P.M., Hainey, S., & Bishop, J.O. (1984) Messenger RNAs coding for mouse major urinary proteins are differentially induced by testosterone. *Biochem. Genet.* 22: 2232-2240

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L., & Wallace, R.B. (1983). Detection of sickle cell β^s -globin allele by hybridisation with synthetic oligonucleotides. *Proc. Natl. Acad. Sci.* 80: 278-282

Da Costa Soares, V., Gubits, R.M., Feigelson, P., & Constantini, F. (1987). Tissue-specific and hormonally regulated expression of a rat α_{2u} globulin gene in transgenic mice. *Mol. Cell. Biol.* 7: 3749-3758

Darnell, J.E., Jr. (1982). Variety in the level of gene control in eukaryotic cells. *Nature* 297: 365-371

Davison, B.L., Egly, J.-M., Mulvihill, E.R., & Chambon, P. (1983). Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. *Nature* 301: 680-686

Derman, E. (1981). Isolation of a cDNA clone for mouse urinary proteins: Age- and sex-related expression of mouse urinary protein genes is transcriptionally controlled. *Proc. Natl. Acad. Sci.* 78: 5425-5429

Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M., & Darnell, J.E., Jr. (1981). Transcriptional control in the production of liver-specific mRNAs. *Cell* 23: 731-739

De Simone, V., Ciliberto, G., Hardon, E., Paonessa, G., Palla, F., Lundberg, L., & Cortese, R. (1987). Cis- and trans-acting elements responsible for the cell-specific expression of the human α_1 -antitrypsin gene. *EMBO J.* 6: 2759-2766

Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J., &

Weissmann, C. (1983). Three regions upstream from the cap site are required for efficient and accurate transcription of rabbit β -globin gene in mouse 3T6 cells. *Cell* 32: 695-706

DiLella, A.G., Marvit, J., Lidsky, A.S., Guttler, F., & Woo, S.L.C. (1986). Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. *Nature* 322: 799-803

Dove, W.F., & Davidson, N. (1962). Cation effects on the denaturation of DNA. *J. Mol. Biol.* 5: 467-478

Dudler, R., & Travers, A.A. (1986). Upstream elements necessary for optimal function of the hsp70 promoter in transformed flies. *Cell* 38: 391-398

Dunn, J.J., & Studier, F.W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. *J. Mol. Biol.* 166: 477-535

Dynan, W.S., Sazer, S., Tjian, R., & Schimke, R.T. (1986). Transcription factor Sp1 recognizes a DNA sequence in the mouse dihydrofolate reductase promoter. *Nature* 319: 246-248

Dynan, W.S., & Tjian, R. (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. *Nature* 316: 774-778

Epstein, L.M., Mahon, K.A., & Gall, J.G. (1986). Transcription of

satellite DNA in the newt J. Cell. Biol. 103: 1137-1144

Feinberg, A.P., & Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-13

Field, L.J., & Gross, K.W. (1985). Ren-1 and Ren-2 loci are expressed in mouse kidney. Proc. Natl. Acad. Sci. 82: 6196-6200

Finlayson, J.S., Asofsky, R., Potter, M., & Runner, C.C. (1965). Major urinary protein complex of normal mice: origin. Science 149: 981-982

Ghazal, P., Clark, A.J., & Bishop, J.O. (1985) Evolutionary amplification of a pseudogene. Proc. Natl. Acad. Sci. 82: 4182- 4185

Ghazal, P. (1986) PhD. Thesis, University of Edinburgh.

Gidoni, D., Dynan, W.S., & Tjian, R. (1984). Multiple specific contacts between a mammalian transcription factor and its cognate promoter. Nature 312: 409-413

Gidoni, D., Kadonaga, J.T., Barrera-Saldana, H., Takahashi, K., Chaabon, P., & Tjian, R. (1985). Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science 230: 511-517

Gillies, S.D., Morrison, S.L., Oi, V.T., & Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33: 717-728

Goeddel, D.V., Leung, D.W., Dull, T.J., Gross, M., Lawn, R.M., McCandliss, R., Seeburg, P.H., Ullrich, A., Yelverton, E., & Gray, P.W. (1981). The structure of eight distinct cloned human leukocyte interferon cDNAs. *Nature* 290: 20-26

Goodboun, S., Burstein, H., & Maniatis, T. (1986). The human β -interferon gene enhancer is under negative control. *Cell* 45: 601-610

Goodboun, S., Zinn, K., & Maniatis, T. (1985). Human β -interferon gene expression is regulated by an inducible enhancer element. *Cell* 41: 509-520

Graves, B.J., Johnson, P.F., & McKnight, S.L. (1986). Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* 44: 565-576

Green, M.R., Maniatis, T., & Melton, D.A. (1983). Human β -globin pre-mRNA synthesised in vitro is accurately spliced in *Xenopus* oocyte nuclei. *Cell* 32: 681-694

Grosveld, F., van Assendelft, G.B., Greaves, D.R., & Kollias, G. (1987). Position-independent, high-level expression of the human

B-globin gene in transgenic mice. *Cell* 51: 975-985

Gubits, R.M., Lynch, K.R., Kulkarni, A.B., Dolan, K.P., Gresik, E.W., & Feigelson, P. (1984) Differential regulation of α_{2U} -globulin gene expression in liver, lachrymal gland, and salivary gland. *J. Biol. Chem.* 259: 12803-12809

Hastie, N.D., & Bishop, J.O. (1976). The expression of three abundance classes of messenger RNA in mouse tissues. *Cell* 9: 761-774

Hastie, N.D., Held, W.A., & Toole, J.J. (1979). Multiple genes coding for the androgen-regulated major urinary proteins of the mouse. *Cell* 17: 449-457

Held, W.A., & Gallagher, J.F. (1985). Rat α_{2U} -globulin mRNA expression in the preputial gland. *Biochem. Genet.* 23: 281-290

Held, W.A., Gallagher, J.F., Holman, C.M., Kuhn, N.J., Sampsell, B.M., & Hughes Jr., R.G. (1987) Identification and characterization of functional genes encoding the mouse major urinary proteins. *Mol. Cell. Biol.* 7: 3705-3712

Houghton, M., Stewart, A.G., Doel, S.M., Entage, J.S., Eaton, M.A.W., Smith, J.C., Patel, T.P., Lewis, H.M., Porter, A.G., Birch, J.R., Cartwright, T., & Carey, N.H. (1980). The amino-terminal sequence of human fibroblast interferon as deduced from reverse transcripts obtained using synthetic oligonucleotide primers. *Nucl. Acid Res.*

8: 1913-1931

Huang, D.P., Cote, G.J., Massari, R.J., & Chiu, J.F. (1985). Dexamethasone inhibits α -fetoprotein gene transcription in neonatal rat liver and isolated nuclei. *Nucl. Acid Res.* 13: 3873-3890

Ikeda, R.A., & Richardson, C.C. (1986). Interactions of the RNA polymerase of bacteriophage T7 with its promoter during binding and initiation of transcription. *Proc. Natl. Acad. Sci.* 83: 3614-3618

Itakura, K., Rossi, J.J., & Wallace, R.B. (1984). Synthesis and use of synthetic oligonucleotides. *Ann. Rev. Biochem.* 53: 323-356

Jones, K.A., Yamamoto, K.R., & Tjian, R. (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. *Cell* 42: 559-572

Jongstra, J., Reudelhuber, T.L., Oudet, P., Benoist, C., Chae, C.-B., Jeltsch, J.-M., Mathis, D.J., & Chambon, P. (1984). Induction of altered chromatin structures by simian virus 40 enhancer and promoter elements. *Nature* 307: 708-714

Karin, M., Haslinger, A., Holtgrev, H., Richards, R.I., Krauter, P., Westphal, H.M., & Beato, M. (1984). Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-II_A gene. *Nature* 308: 513-519

Kaye, J.S., Pratt-Kaye, S., Bellard, M., Dretzen, G., Bellard, F., & Chambon, P. (1986). Steroid hormone dependence of four DNaseI-hypersensitive regions located within the 7000bp 5'-flanking segment of the ovalbumin gene. *EMBO J.* 5: 277-285

Kidd, V.J., Wallace, R.B., Itakura, K. & Woo, S.L.C. (1983). α_1 -Antitrypsin deficiency detection by direct analysis of the mutation in the gene. *Nature* 304: 23-234

Killary, A.M., & Fournier, R.E.K. (1984). A genetic analysis of extinction: Trans-dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. *Cell* 38: 523-534

Knopf, J.L., Gallagher, J.F., & Held, W.A. (1983) Differential, multihormonal regulation of the mouse major urinary protein gene family in the liver. *Mol. Cell. Biol.* 3: 2232-2240

Krieg, P.A., & Melton, D.A., (1984). Functional messenger RNAs are produced by SP6 in vitro transcription of cloned cDNAs. *Nucl. Acid Res.* 12: 7057-7070

Kuhn, N.J., Woodworth-Gutai, M., Gross, K.W., & Held, W.A. (1984). Subfamilies of the mouse major urinary protein (MUP) multi-gene family: sequence analysis of cDNA clones and differential regulation in the liver. *Nucl. Acid Res.* 12: 6073-6090

Kumar, V., Green, S., Staub, A., & Chambon, P. (1986). Localisation of the oestradiol-binding and putative DNA-binding domains of the

human oestrogen receptor. *EMBO J.* 5: 2231-2236

Kurtz, D.T., & Feigelson, P. (1978). Multihormonal control of the messenger RNA for the hepatic protein α_{2u} -globulin. in Litwack, G. (ed) *Biochemical Actions of Hormones Vol. 5* pp433-455. Academic Press, New York.

Kurtz, D.T., Sippel, A.E., Ansah-Yiadom, R., & Feigelson, P. (1976). Effects of sex hormones on the level of the messenger RNA for the rat hepatic α_{2u} globulin protein. *J. Biol. Chem.* 251: 3594-3598

Lathe, R. (1985). Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* 183: 1-12

Lee, D.C., & Roeder, R.G. (1981). Transcription of adenovirus type 2 genes in a cell-free system: Apparent heterogeneity of initiation at some promoters. *Mol. Cell. Biol.* 1: 635-651

Ma, X. (1987). PhD. Thesis, University of Edinburgh

MacInnes, J.I., Nozik, E.S., & Kurtz, D.T. (1986). Tissue-specific expression of the rat α_{2u} globulin gene family. *Mol. Cell. Biol.* 6: 3563-3567

McKnight, S., & Tjian, R. (1986). Transcriptional selectivity of viral genes in mammalian cells. *Cell* 46: 795-805

Maniatis, T., Goodbourn, S., & Fischer, J.A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* 236: 1237-1245

Marmur, J., & Doty, P. (1959). Heterogeneity in deoxyribonucleic acids I. Dependence on composition of the configurational stability of deoxyribonucleic acids. *Nature* 183: 1427-1429

Marmur, J., & Doty, P. (1961). Thermal renaturation of deoxyribonucleic acids. *J. Mol. Biol.* 3: 585-594

Marmur, J., & Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.* 5: 109-118

Mason, A.J., Evans, B.A., Cox, D.R., Shine, J., & Richards, R.I. (1983). Structure of mouse kallikrein gene family suggests a role in specific processing of biologically active peptide. *Nature* 303: 300-307

Maxam, A.M., & Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Meths. Enzymol* 65: 499-560

Mead, D.A., Szczesna-Skorupa, E., & Kemper, B. (1986). Single-stranded DNA 'blue' T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Prot. Engg.* 1: 67-74

Melchoir, W.B., & Von Hippell, P. (1973). Alteration of the relative stability of dA.dT and dG.dC base pairs in DNA. Proc. Natl. Acad. Sci. 70: 298-302

Melton, D.A., Krieg, P.A., Rebegliati, M.R., Maniatis, T., Zinn, K. & Green, M.R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acid Res. 12: 7035-7056

Melton, D.W., McEwan, C., McKie, A.B., & Reid, A.M. (1986). Expression of the mouse HPRT gene: Deletional analysis of the promoter region of an X-chromosome linked housekeeping gene. Cell 44: 319-328

Michelson, A.M., & Monny, C. (1967). Polynucleotides X. Oligonucleotides and their association with polynucleotides. Biochim. Biophys. Acta 149; 107-126

Morris, C.E., McGraw, N.J., Joho, K., Brown, J.E., Klement, J.F., Ling, M.L., & McAllister, W.T. (1987). Mechanisms of promoter recognition by the bacteriophage T3 and T7 RNA polymerases. in Reznikoff, W.S., Burgess, R.R., Dahlberg, J.E., Gross, C.A., Record, M.T., Jr., & Wickens, M.B. (eds). RNA polymerase and the regulation of transcription. pp 47-58 Elsevier New York.

Myers, R.M., Larin, Z., & Maniatis, T. (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA

duplexes. *Science* 230: 1242-1246

Myers, R.M., & Maniatis, T. (1986). Recent advances in the development of methods for detecting single-base substitutions associated with human genetic diseases. *Cold Spring Harb. Symp. Quant. Biol.* 51: 275-284

Myers, R.M., Tilly, K., & Maniatis, T. (1986). Fine structure genetic analysis of a β -globin promoter. *Science* 232: 613-618

Norstedt, G. & Palmiter, R.D. (1984). Secretory rhythm of growth hormone regulates sexual differentiation of mouse liver. *Cell* 36: 805-812

Nygaard, A.P., & Hall, B.D. (1964). Formation and properties of RNA-DNA complexes. *J. Mol. Biol.* 9: 125-142

Ohlsson, H., & Edlund, T. (1986). Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell* 45: 35-44

Omiecinski, C.J., Walz, F.G., JR., & Vlasuk, G.P. (1985). Phenobarbital induction of rat liver cytochrome P-450b and P-450e. Quantitation of specific RNAs by hybridisation to synthetic oligodeoxyribonucleotide probes. *J. Biol. Chem.* 260: 3247-3250

Parker, C.S., & Topol, J. (1984). A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* 36: 357-369

Pelham, H.R.B., & Bienz, M. (1982). A synthetic heat-shock promoter element confers heat-inducibility of the herpes simplex thymidine kinase gene. *EMBO J.* 1: 1473-1477

Pevsner, J., Sklar, P.B., & Snyder, S.H. (1986). Odorant-binding protein: Localization to nasal glands and secretions. *Proc. Natl. Acad. Sci.* 83: 4942-4946

Pinkert, C.A., Ornitz, D.M., Brinster, R.L., & Palmiter, R.D. (1987). An albumin enhancer located 10kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes Devel.* 1: 268-276

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* 322: 697-701

Renkawitz, R., Schutz, G., van der Ahe, D., & Beato, M. (1984). Sequences in the promoter region of the chicken lysozyme gene required for steroid regulation and receptor binding. *Cell* 37: 503-510

Rumke, P.H., & Thung, P.J. (1964). Immunological studies on the sex-dependent prealbumin in mouse urine and on its occurrence in the serum. *Acta Endocrinologia* 47: 156-164

Ryals, J., Dierks, D., Ragg, H., & Weissmann, C. (1985). A 46-nucleotide promoter segment from an IFN- α gene renders an unrelated promoter inducible by virus. *Cell* 41: 497-507

Sampsel, B.M., & Held, W.A. (1985). Variation in the major urinary protein multigene family in wild-derived mice. *Genetics* 109: 549-568

Sassone-Corsi, P., & Borrelli, E. (1986). Transcriptional regulation by trans-acting factors. *Trends Genet.* 2: 215-219

Sawadogo, M., & Roeder, R.G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43: 163-175

Sawyer, L. (1987). One fold among many. *Nature* 327: 659

Scheidereit, C., Geisse, S., Westphal, H.M., & Beato, M. (1983). The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumour virus. *Nature* 304: 749-752

Scheidereit, C., & Beato, M. (1984). Contacts between hormone receptor and DNA double helix within a glucocorticoid regulatory element of mouse mammary tumour virus. *Proc. Natl. Acad. Sci.* 81: 3029-3033

Schenborn, E.T., & Mierendorf, R.C., Jr. (1985). A novel transcription property of SP6 and T7 RNA polymerases: dependence on template structure. *Nucl. Acid Res.* 13: 6223-6236

Schildkraut, C., & Lifson, S. (1965). Dependence of the melting

temperature of DNA on salt concentration. *Biolpolymers* 3: 195-208

Searle, P.F., Stuart, G.W., & Palmiter, R.D. (1985). Building a metal-responsive promoter with synthetic regulatory elements. *Mol. Cell. Biol.* 5: 1480-1489

Sen, R., & Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* 46: 705-716

Serfling, E., Jasin, M., & Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. *Trends Genet.* 1: 224-230

Shahan, K., & Derman, E. (1984). Tissue-specific expression of major urinary protein (MUP) genes in mice: characterization of MUP mRNAs by restriction mapping of cDNA and by in vitro translation. *Mol. Cell. Biol.* 4: 2259-2265

Shahan, K., Gilmartin, M. & Derman, E. (1987a). Nucleotide sequences of liver, lachrymal, and submaxillary gland mouse major urinary protein mRNAs: mosaic structure and construction of gene-specific synthetic oligonucleotide probes. *Mol. Cell. Biol.* 7: 1938-1946

Shahan, K., Denaro, M., Gilmartin, M., Shi, Y., & Derman, E. (1987b). Expression of six mouse major urinary protein genes in the mammary, parotid, sublingual, submaxillary, and lachrymal glands and in the liver. *Mol. Cell. Biol.* 7: 1947-1954

Shaw, P.H., Held, W.A., & Hastie, N.D. (1983). The gene family for major urinary proteins: expression in several secretory tissues of the mouse. *Cell* 32: 755-761

Søeekens, S.P., & Romano, L.J. (1986). Promoter and nonspecific DNA binding by the T7 RNA polymerase. *Nucl. Acid Res.* 14: 2811-2827

Smith, M. (1983). Synthetic oligodeoxyribonucleotides as probes for nucleic acids and as primers in sequence determination. in Weissman, S.M. (ed). *Methods of DNA and RNA sequencing*. Praeger. New York.

Spiegelberg, T., & Bishop, J.O. (1988). Tissue-specific gene expression in mouse hepatocytes cultured in growth-restricting medium. *Mol. Cell. Biol.* 8: 3338-3344

Stuart, G.W., Searle, P.F., Chen, H.Y., Brinster, R.L., & Palmiter, R.D. (1984). A 12-base-pair DNA motif that is repeated several times in metallothionein gene promoters confers metal regulation to a heterologous gene. *Proc. Natl. Acad. Sci.* 81: 7318-7322

Stuart, G.W., Searle, P.F., & Palmiter, R.D. (1985). Identification of multiple regulatory elements in mouse metallothionein-I promoter by assaying synthetic sequences. *Nature* 317: 828-831

Studencki, A.B., Conner, B.J., Impraim, C.C., Teplitz, R.L., & Wallace, R.B. Discrimination among the human β^A , β^B , and β^C -globin genes using allele-specific oligonucleotide hybridization

probes. Am. J. Hum. Genet. 137: 42-51 (1985)

Suggs, S.V., Hirose, T., Miyake, T., Kamashima, E.H., Johnson, M.J., Itakura, K., & Wallace, R.B. (1981). Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. in Brown, D.D., & Fox, D.F. (eds) Developmental biology using purified genes. ICN-UCLA Symp. Mol. Cell. Biol. Vol. 23 pp 683- 693. academic Press New York.

Szoka, P.R., & Paigen, K. (1978). Regulation of mouse major urinary production by the MUP-A gene. Genetics 90: 597-612

Szostak, J.W., Stiles, J.I., Bahl, C.P., Wu, R. (1977). Specific binding of a synthetic oligodeoxyribonucleotide to yeast cytochrome c mRNA. Nature 265: 61-63

Tabor, S. & Richardson, C.C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. 82: 1074-1078

Theisen, M., Steif, A., & Sippel, A.E. (1986). The lysozyme enhancer: cell- specific activation of the chicken lysozyme gene by a far-upstream DNA element. EMBO J. 5: 719-724

Thomas, C.A., Jr., & Dancis, B.M. (1973). Ring stability. J. Mol. Biol. 77: 43-55

Thomas, R. (1954). Recherches sur la denaturation des acides

desoxyribonucleiques. *Biochim. Biophys. Acta* 14: 231-240

Tinoco, I., Jr., Uhlenbeck, O.C., & Levine, M.D. (1971). Estimation of secondary structure in ribonucleic acids. *Nature* 230: 362-367

Tinoco, I., Jr., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M., & Gralla, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.* 246: 40-41

Vandenbergh, J.G., Whitsett, J.M., & Lombardi, J.R. (1975). Partial isolation of a pheromone accelerating puberty in female mice. *J. Reprod. Fert.* 43: 515-523

Vandenbergh, J.G., Finlayson, J.S., Dobrogosz, W.J., Dills, S.S., & Kost, T.A. (1976). Chromatographic separation of puberty accelerating pheromone from male mouse urine. *Biol. Reprod.* 15: 260-265

van Leeuwen, B.H., Evans, B.A., Tregear, G.W., & Richards, R.I. (1986). Mouse glandular kallikrein genes. Identification, structure, and expression of the renal kallikrein gene. *J. Biol. Chem.* 261: 5529-5535

van Leeuwen, B.H., Penschow, J.D., Coghlan, J.P., & Richards, R.I. (1987). Cellular basis for the differential response of mouse kallikrein genes to hormonal induction. *EMBO J.* 6: 1705-1713

Vries, M.V., Bogaard, M.E., van den Elst, H., van Boom, J.H., van der Eb, A.J., & Bos, J.L. (1986). A dot-blot screening procedure for mutated ras oncogenes using synthetic oligodeoxynucleotides. *Gene* 50: 313-320

Walker, P.M.B. (1969). The specificity of molecular hybridisation in relation to studies on higher organisms. *Prog. Nucl. Acid Res.* 9: 301-326

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., & Itakura, K. (1979). Hybridisation of synthetic oligodeoxyribonucleotides to ϕ x174 DNA: the effect of single base pair mismatch. *Nucl. Acid Res.* 6: 3543-3557

Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H., & Itakura, K. (1981). The use of synthetic oligonucleotides as hybridisation probes. II. Hybridisation of oligonucleotides of mixed sequence to rabbit β -globin DNA. *Nucl. Acid Res.* 9: 879-894

Wetmur, J.G., & Davidson, N. (1968). Kinetics of renaturation of DNA. *J. Mol. Biol.* 31: 349-370

Wildeman, A.G., Zenke, M., Schatz, C., Wintzerith, M., Grundstrom, T., Matthes, H., Takahashi, K., & Chambon, P. (1986). Specific protein binding to the Simian Virus 40 enhancer in vitro. *Mol. Cell. Biol.* 6: 2098-2105

Winderickx, J., van Dijck, P., Dirckx, L., Volckaert, G., Rombauts,

W., Heyns, W., & Verhoeven, G. (1987). Comparison of the 5' upstream putative regulatory sequences of three members of the $\alpha_2\mu$ -globulin gene family. *Eur. J. Biochem.* 165: 521-529

Wood, W.I., Gitschier, J., Lasky, L.A., & Lawn, R.M. (1985). Base composition-independent hybridisation in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci.* 82: 1585-1588

Wreschner, D.H., & Herzberg, M. (1984). A new blotting medium for the simple isolation and identification of highly resolved messenger RNA. *Nucl. Acid Res.* 12: 1349-1359

Yamamoto, K.R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Ann. Rev. Genet.* 19: 209-252

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A., & Chambon, P. (1986). Multiple sequence specific motifs are involved in SV40 enhancer function. *EMBO J.* 5: 387-397

Zheng, X.-M., Moncollin, V., Egly, J.-M., & Chambon, P. (1987). A general transcription factor forms a stable complex with RNA polymerase B (II). *Cell* 50: 361-368