

TRANSCRIPTIONAL REGULATION OF  
HISTONE GENE EXPRESSION



*A thesis submitted to the University of Adelaide,  
for the Degree of Doctor of Philosophy*

by

**STEPHEN DALTON, B.Sc. (Hons.)**

Department of Biochemistry,  
University of Adelaide,  
Adelaide, South Australia

May, 1987

Awarded 25/8/87

## CONTENTS

	<u>Page No.</u>
STATEMENT	i
SUMMARY	ii
ACKNOWLEDGEMENTS	v
<b>CHAPTER 1. INTRODUCTION</b>	<b>1</b>
1.1 REGULATION OF TRANSCRIPTION	1
1.2 PROKARYOTIC AND EUKARYOTIC TRANSCRIPTIONAL CONTROL ELEMENTS	2
1.3 GENERAL REGULATORY ELEMENTS DIRECTING TRANSCRIPTION OF PROTEIN-CODING GENES	3
1.3.1 Specification of the Transcription Start-Site	4
1.3.2 General Upstream Promoter Elements	5
1.4 GENE-SPECIFIC MODULATOR ELEMENTS	6
1.4.1 Viral Enhancer Elements	6
1.4.2 Cellular Enhancer Elements	10
1.4.3 Tissue-Specific Regulatory Elements	11
1.4.4 Inducible Genes	14
1.4.5 Developmental Control of Gene Expression	15
1.4.6 Negative Regulatory Elements	16
1.5 STABLE TRANSCRIPTION INITIATION COMPLEXES	19
1.6 EUKARYOTIC TRANSCRIPTION COMPLEXES	20
1.6.1 The Transcription Factor SP1	21
1.6.2 DNA Binding and Activatory Domains of the Yeast Transcription Factors; GCN4 and GAL4	23
1.6.3 Metal-Binding Fingers of TFIIIA	25
1.6.4 Nuclear Protein Interactions with General Upstream Regulatory Elements	29
1.6.5 Gene-Specific Regulatory Elements and their Binding Factors	30

		<u>Page No.</u>
1.7	REGULATION OF TRANSCRIPTION BY PROTEINS ACTING AT A DISTANCE	32
1.7.1	Transcriptional Activation from a Distance: DNA Looping	33
1.7.2	Enhancers: Their Transcription Factors and Mechanism of Action	36
1.8	HISTONE PROTEINS	38
1.9	ORGANIZATION OF HISTONE GENES	40
1.10	HISTONE GENE ARCHITECTURE	42
1.10.1	General Promoter Elements	42
1.10.2	Gene-Specific Promoter Elements	43
1.10.3	3'-End Formation of Histone Transcripts	44
1.11	HISTONE GENE EXPRESSION	46
1.11.1	Developmental Switching of Histone Genes	46
1.11.2	Expression of Replication-Independent and Partially Replication-Dependent Histone Genes	48
1.11.3	Replication-Dependent (Cell-Cycle Regulated) Histone Genes	49
1.11.4	Multiple Levels of Histone Gene Regulation	49
1.11.5	<u>cis-</u> and <u>trans-</u> Regulatory Elements Required for Cell-Cycle Regulation	52
1.11.6	Histone mRNA 3' Ends and Differential Transcript Stability	53
1.12	THESIS AIMS	55
<b>CHAPTER 2.</b>	<b>MATERIALS AND METHODS</b>	
2.1	ABBREVIATIONS	56
2.2	MATERIALS	57
2.2.1	Chemicals and Reagents	57
2.2.2	Enzymes	58
2.2.3	Isotopically Labelled Compounds	58
2.2.4	Recombinant DNA Clones	58
2.2.5	Cloning Vectors	59
2.2.6	Synthetic Oligonucleotides	59
2.2.7	Bacterial Strains	61
2.2.8	Bacterial Media	62
2.2.9	Tissue Culture Cell-Lines and Media	62
2.3	METHODS	64
2.3.1	Isolation of Plasmid DNA	64
2.3.2	Preparation of M13-bacteriophage DNA	65
2.3.3	Restriction Enzyme Digestion	66
2.3.4	Gel Electrophoresis	67

	<u>Page No.</u>	
2.3.5	Autoradiography and Laser Densitometry	70
2.3.6	Subcloning of DNA Fragments into Plasmid and M13 Vectors	71
2.3.7	Transformation Procedure for Plasmid Recombinants	72
2.3.8	Transformation Procedure for M13 Recombinants	72
2.3.9	Containment Facilities	73
2.3.10	Preparation of <u>In Vitro</u> Labelled DNA	73
2.3.11	Manipulation and Harvesting Tissue Culture Cells	76
2.3.12	Isolation of RNA	77
2.3.13	Isolation of Nuclei from Tissue Culture Cells	78
2.3.14	Analysis of Nuclear Matrix-Bound DNA	79
2.3.15	Analysis of RNA	80
2.3.16	Dot Blot Filter Hybridization	82
2.3.17	Southern Blot Hybridization	84
2.3.18	Transfection and Selection of Clonal Cell-Lines	85
2.3.19	Oligonucleotide-Directed <u>In Vitro</u> Mutagenesis	85
2.3.20	DNA Sequencing by the Dideoxy-Chain Termination Method	87
2.3.21	Nuclear Extract Preparation	88
2.3.22	Chromatographic Enrichment of Nuclear Extract Components	88
2.3.23	Detection of DNA Binding Factors in Nuclear Extracts	90
<b>CHAPTER 3.</b>	<b>ANALYSIS OF H5 AND H2A<sub>F</sub> GENE EXPRESSION</b>	
3.1	INTRODUCTION	91
3.2	RESULTS	91
3.2.1	Histone Gene Expression in an Early- Erythroid Chicken Cell-Line	91
3.2.2	AEV <u>ts34</u> Cell Synchronization with Aphidicolin	93
3.2.3	Cell-Cycle Regulation of Chicken Histone Gene Expression	94
3.2.4	Transcription of H5 and H2A <sub>F</sub> Genes	96
3.2.5	5'-End Mapping of the H2A <sub>F</sub> Transcript	98
3.3	DISCUSSION	98
<b>CHAPTER 4.</b>	<b>TRANSCRIPTIONAL REGULATION OF HISTONE GENE EXPRESSION: A ROLE FOR H1 AND H2B GENE- SPECIFIC ELEMENTS</b>	
4.1	INTRODUCTION	104



		<u>Page No.</u>
4.2	RESULTS	106
4.2.1	The H1 Gene-Specific Element is a <u>cis-Acting</u> Transcriptional Regulator	106
4.2.2	Deletion and Substitution Analysis of the H2B Gene-Specific Element	112
4.3	DISCUSSION	113
4.3.1	An Intact H1-Specific Element is Required for Efficient Promoter Activity and Cell-Cycle Regulation	113
4.3.2	The H2B-Box is Required for Efficient Promoter Activity and Influences the Expression of an Adjacent H2A Gene	114
CHAPTER 5.	<b>IDENTIFICATION AND CHARACTERIZATION OF NUCLEAR FACTORS WHICH INTERACT WITH THE H1 AND H2B GENE-SPECIFIC PROMOTER ELEMENTS</b>	
5.1	INTRODUCTION	117
5.2	RESULTS	117
5.2.1	Detection of Nuclear Factors which Specifically Bind to the H1 and H2B Gene-Specific Elements	117
5.2.2	Sequence Requirements for Nuclear Factor Binding to H1 and H2B Oligonucleotide Duplexes	120
5.2.3	Enrichment of H1- and H2B-Specific Binding Factors	121
5.2.4	Levels of H1-Specific Nuclear Factor Fluctuate Throughout the Cell-Cycle	123
5.3	DISCUSSION	124
5.3.1	Sequence-Specific Nuclear Factors Bind to the H1 and H2B Gene-Specific Promoter Elements	124
5.3.2	Binding Activity of the H1-Specific Nuclear Factor Fluctuates Throughout the Cell-Cycle	126
CHAPTER 6.	<b>CHICKEN HISTONE GENES RETAIN NUCLEAR MATRIX ASSOCIATION THROUGHOUT THE CELL-CYCLE</b>	
6.1	INTRODUCTION	128
6.2	RESULTS	132
6.2.1	Enrichment of Core and H1 Histone Genes with the Nuclear Matrix is Independent of Their Transcriptional Activity	132

		<u>Page No.</u>
	6.2.2 Tissue-Specific Attachment of the Histone H5 Gene to the Nuclear Matrix	134
6.3	DISCUSSION	137
<b>CHAPTER 7.</b>	<b>FINAL DISCUSSION</b>	140
7.1	CELL-CYCLE CONTROL OF HISTONE GENE EXPRESSION	140
7.2	FUTURE WORK AND PERSPECTIVES	146
	 BIBLIOGRAPHY	

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge, it contains no material that has been previously published by any other person, except where due reference is made in the text.

STEPHEN DALTON

SUMMARY

The work presented and discussed in this thesis is primarily concerned with the periodic expression of chicken histone genes during S-phase of the cell-cycle and, in particular, their transcriptional activation.

1. Northern blot analysis revealed that regulation of steady-state mRNAs for the major core and H1 chicken histone genes is closely coupled to DNA replication.
2. Pulse-labelling experiments indicate that periodic transcriptional activation of these genes during S-phase plays a major role in this regard.
3. In contrast, transcription of two genes encoding the variant histones, H5 and H2A<sub>F</sub>, is uncoupled from DNA synthesis and accumulation of their respective mRNAs is not S-phase controlled.
4. The 5' cis-regulatory sequences required for transcriptional activation of a chicken histone H1 gene during S-phase of the cell-cycle have been identified. Specifically, a conserved heptanucleotide element (H1-box, 5'-AAACACA-3'), ubiquitous to all H1 genes, is essential for correct cell-cycle regulation. Introduction of point mutations or deletion of this element by site-directed mutagenesis reduces levels of gene-specific transcripts by 10-20 fold in vivo. Run-on transcription assays in isolated nuclei indicate that this reflects a loss of S-phase transcriptional control.

5. Similar to viral and cellular enhancers, the H1-box functions bi-directionally and from an upstream position in the H1-promoter.
6. Site-directed deletion and substitution mutagenesis was also performed to characterize the putative regulatory role for an H2B gene-specific motif (H2B-box, 5'-CTGATTTGCATAG<sub>C</sub>-3'). Deletion or disruption of this element decreased H2B steady-state mRNA levels by 10-20 fold in vivo. A concomitant decrease in transcripts generated from a paired and divergently arranged H2A gene was also detected. Pulse-labelling experiments indicate that while a decrease in absolute levels of H2B transcription accounts for this, a clear pattern of S-phase control was retained.
7. A sequence-specific cellular factor (H1-SF) has been partially purified from nuclear extracts and shown to interact selectively with the H1-specific motif as assayed by gel-mobility and DNAase protection assays.
8. Levels of this nuclear factor fluctuate in parallel with elevated and depressed rates of H1 gene transcription, an observation consistent with it being a trans-regulator of H1 gene expression during the cell-cycle. However, levels of a factor which specifically binds to the H2B-box are essentially invariant at corresponding stages.
9. The association between histone genes and the nuclear matrix (NM) during periods of high (S-phase) and low (non-S-phase) transcriptional activity has been investigated with synchronized cells. By DNAase I and restriction enzyme analysis, these studies

reveal that both core and linker histone genes (represented by H2A and H1 genes, respectively) are attached to the NM independent of their transcriptional activity during the cell-cycle.

10. The histone H5 gene, transcribed exclusively in cells of the erythroid lineage, is nuclear matrix-associated in AEV ts34 (early erythroid) cells but not in a T-lymphoid cell-line. DNA sequences necessary for NM-attachment of the H5 gene in AEV ts34 cells lie within a 780 bp region spanning 5' non-coding and coding sequences.

ACKNOWLEDGEMENTS

I would like to express my thanks to:

Professor W.H. Elliott for the opportunity of undertaking this work in the Department of Biochemistry.

Dr. Julian Wells for supervision during the course of this study.

Members of labs 16, 20 and 20A for their advice and assistance.

My wife, Dawn, and parents for encouragement and support throughout the years.

Julie Pocock for typing this thesis, Ruth Evans for diagrams and Leslie Crocker for photography.

During the course of this work I was supported by a Commonwealth Post-Graduate Research Award.

CHAPTER 1

INTRODUCTION:

TRANSCRIPTIONAL CONTROL OF GENE EXPRESSION





## CHAPTER 1 : INTRODUCTION

### TRANSCRIPTIONAL CONTROL OF GENE EXPRESSION

Eukaryotic cells possess the remarkable capacity to differentially express a small subset of their genetic complement in a precise and coordinated fashion. This is commonly orchestrated by selective activation and repression of gene transcription in a temporal, developmental or cell-type specific pattern. The molecular mechanisms directing transcriptional regulation of gene expression are therefore pivotal to our understanding of how differential gene expression controls aspects of cell specialization and homeostasis. Work presented in this thesis focuses on the regulation of transcription and the molecular events underlying this level of gene control are considered by focussing on the chicken histone multi-gene family as a model system. Particular emphasis is directed towards elucidating the cis-regulatory DNA elements and auxillary trans-acting molecules required for co-ordinate control of gene transcription throughout the cell-division cycle and to pronounce some of the more general experimental approaches employed. This chapter deals principally with two major areas which are directly relevant to the experimental work performed. They are; general aspects of eukaryotic gene transcriptional control and in particular, the organization, structure and expression of histone genes.

#### 1.1 REGULATION OF TRANSCRIPTION

Regulation of gene expression is controlled at several points along the pathway from gene to mature gene product. Apart from transcriptional regulation, differential RNA processing, nucleo-cytoplasmic RNA transport, RNA stability, translation and post-translational modifications are all potential stages at which gene expression can be modulated (reviewed by Nevins, 1983; Darnell, 1982). In addition to the

aforementioned levels of control other, less common, regulatory mechanisms have been identified. These include, gene activation by rearrangement of DNA sequences, gene amplification and DNA transposition (see Alberts et al., 1983). However, it is generally recognized that the most frequent mode of eukaryotic gene control lies at the level of transcription.

Both frequency and specificity of initiation are of prime importance in the control of transcription (Darnell, 1982). Obviously, apart from increasing the rate of initiation, which will generate higher levels of nascent transcripts, a key feature of transcriptional regulation is the ability to selectively express certain genes via a process involving promoter selection by RNA polymerase. This latter facet of transcriptional initiation is particularly important in the expression of tissue-specific, developmentally-regulated and inducible genes which are selectively activated in certain cell-types or in response to certain physiological stimuli.

## 1.2 PROKARYOTIC AND EUKARYOTIC TRANSCRIPTIONAL CONTROL ELEMENTS

Before discussing aspects of eukaryotic transcriptional control, reflection on prokaryotic cis-acting regulatory elements will be briefly considered.

Initiation of prokaryotic transcription is controlled by specific regions of 5'-flanking DNA called promoters. A large number of prokaryotic promoters have been defined and finely mapped by identifying mutations in potential regulatory sequence elements which either decrease or increase the level of transcription in a cis-dominant fashion. Two principle regions have been identified to participate in the initiation of transcription by RNA polymerase in E. coli. These are located approximately 35 bp and 10 bp upstream from the transcription start site (Reznikoff et al., 1985). The spatial relationship between

the DNA sequence elements located in -10 and -35 regions is important as insertion or deletion of a single base pair can dramatically alter transcription (Stefano and Gralla, 1982). Furthermore, these elements are generally believed to facilitate the initiation of transcription by directing binding of RNA polymerase. Other regulatory elements, located either further upstream or downstream of this region, interact with positive and negative regulatory proteins which modulate the efficiency of transcription initiation (see Section 1.7, Busby et al., 1982; Ogden et al., 1980; Irani et al., 1983; Majumdar and Adhya, 1984; Dunn et al., 1984).

Promoter regions of eukaryotic genes can be defined as the complement of DNA sequence elements required for accurate and efficient initiation of transcription. Eukaryotic transcription involves three distinctly separate RNA polymerase molecules which transcribe discrete classes of genes. All three polymerases are multisubunit enzymes and are distinguished biochemically by their differential sensitivity to the antibiotic  $\alpha$ -amanatin. One common feature of the transcriptional control elements associated with the three classes of genes is the presence of general and specific elements which control the frequency of initiation and/or transcriptional selectivity. For the purposes of this introduction, attention will focus predominantly on cis-acting elements regulating transcription of genes transcribed by RNA polymerase II (see Figure 1.1). The characteristics of these sequences and their interaction with the transcriptional machinery bear considerable similarity with their prokaryotic counterparts.

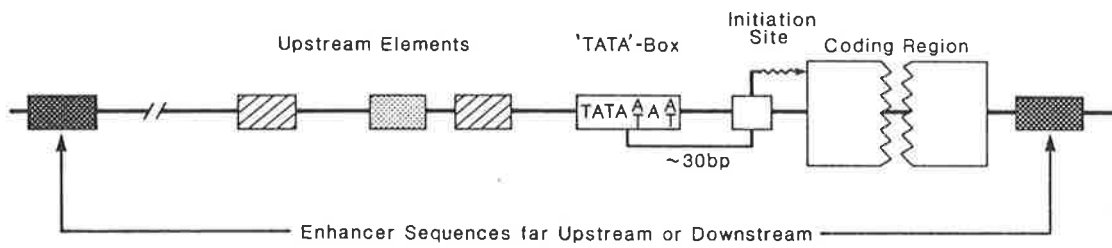
### 1.3 GENERAL REGULATORY ELEMENTS DIRECTING TRANSCRIPTION OF PROTEIN-CODING GENES

Transcription of RNA polymerase II genes culminates in the generation of transcripts ultimately translated into cellular protein

## FIGURE 1.1

### ORGANIZATION OF RNA POLYMERASE II-TRANSCRIBED GENES

The considerable variety in position of transcriptional control elements with respect to the initiation site in RNA polymerase II-transcribed genes is shown. In most cases, the 'TATA'-box specifies the transcription initiation site, while other elements located further upstream or downstream of this sequence generally play a role in modulating the frequency of transcription initiation and/or conferring promoter specificity. These include general motifs such as 'CCAAT'- and G/C-elements (Section 1.3) which may be represented in multiple copies or, enhancer sequences (Section 1.4). Adapted from Schaffner, 1985.



which determines cell fate and specialization in addition to maintenance of housekeeping functions. For this reason, special emphasis will be placed on several aspects of differential gene regulation displayed by protein coding genes and, the mechanism(s) underlying this control.

### 1.3.1 Specification of the Transcription Start-Site

The most striking sequence homology thus far identified in the promoter region of RNA polymerase II-transcribed genes is a highly conserved AT-rich region, with a consensus 5'-TATA<sup>A</sup><sub>T</sub><sup>A</sup>-3' (Corden et al., 1980), located 25-30 nucleotides upstream from the start-site of transcription in most eukaryotic protein-coding genes. In the few cases where it is absent, it may be replaced by a substitute 'TATA'-like element, which plays the same role (Brady et al., 1982). Mutations in (Wasylyk et al., 1980), or deletion of (Grosschedl and Birnstiel, 1980) the 'TATA'-box decreases the proportion of correctly initiated transcripts and generates multiple heterogeneous start-sites in vivo and in vitro. Deletion experiments have also shown the intervening spacing between this element and the in vivo start site to be critical in determining where initiation occurs (Breathnach and Chambon, 1981). The general role of the 'TATA'-box therefore appears in accurately specifying the transcription start-site which usually occurs at an adenine residue flanked by pyrimidine based nucleotides although no apparent consensus sequence exists in this region (Breathnach and Chambon, 1981). Moreover, its integrity is also required to achieve maximum RNA synthesis (McKnight and Kingsbury, 1982; Grosschedl and Birnstiel, 1980). A few notable exceptions of polymerase II-transcribed genes lacking any semblance of 'TATA'-homology in the -25 to -35 region have been identified. These include some Simian Virus 40 (SV40) genes, polyoma virus late genes, the E2 and IVa<sub>2</sub> genes of adenovirus (see Darnell, 1982) and several cellular housekeeping genes such as those for hypoxanthine phosphoribosyl transferase (Melton et al., 1984), 3-hydroxy-

3-methyl-glutaryl coenzyme A reductase (Reynolds et al., 1984) and adenosine deaminase (Valerio et al., 1985). In several of these cases, multiple transcription start-sites are generated in vivo.

### 1.3.2 General Upstream Promoter Elements

Several DNA sequence elements in the -40 to -110 region, upstream from the 'TATA'-box, are required for efficiency but not the accuracy of transcriptional initiation. These upstream promoter elements are polymorphic, and their number and position relative to the transcription start site is variable from one promoter to another. A prominent sequence motif in the immediate upstream region has the consensus sequence 5'-GGPyCAA<sup>A</sup>CT-3' ('CCAAT'-box), is located 70-90 bp from the mRNA start-site (Benoist et al., 1980) and is commonly found in a wide range of promoters. Mutations in the 'CCAAT'-element reduce levels of specific transcription from the rabbit  $\beta$ -globin (Dierks et al., 1983), Herpes Simplex Virus thymidine kinase (McKnight et al., 1984) and, adenovirus-2 EIIaE and major late promoters (Elkaim et al., 1983; Miyamoto et al., 1984). In several cases, this element has been shown to act in concert with other upstream regulatory elements (Grosveld et al., 1982; Bienz, 1986).

A second sequence (consensus: 5'-<sup>G</sup><sub>T</sub>GGGCGG<sup>GGG</sup><sub>AAAT</sub>-3'; Kadonaga et al., 1986), commonly known as the G/C-box, is a general polymerase II promoter element characteristic of many viral and cellular genes. Often present in multiple copies, this element is usually located in the -40 to -200 region and has been shown to play a critical role in directing efficient transcription from a select class of genes including the SV40 early (Benoist and Chambon, 1981), mammalian  $\beta$ -globin (Charnay et al., 1985), Herpes thymidine kinase (McKnight and Kingsbury, 1982) and glycoprotein-D (Everett, 1983) genes. Although this element lacks any discernable symmetry, it can function in either orientation (Gidoni et al., 1984; Dynan and Tjian, 1985).

## 1.4 GENE-SPECIFIC MODULATOR ELEMENTS

One striking feature of polymerase II promoter architecture is the presence of gene-specific regulatory elements in addition to the more general elements already outlined (Section 1.3). Identification of such elements, usually by deletion mapping or by detecting sequence homologies between commonly regulated sets of genes, has provided a significant insight into how differential gene expression is controlled. With few possible exceptions (see Kelly and Darlington, 1985), modulation of gene-specific transcription is controlled by a group of regulatory elements collectively known as enhancers. The remarkable ability of enhancer sequences to function upstream, within, or downstream from a structural gene and, in an orientation-independent manner, distinguishes them from classical promoter elements such as the 'TATA'-box (see Section 1.3), which are immediately 5' to the coding region. Generally, these elements function as activators and stimulate transcription from heterologous or homologous promoters, in some cases by over 1000-fold. The efficiency of enhancer-mediated transcriptional activation frequently decreases as the length of the interposed sequences is increased (Hen et al., 1983) and, proximal promoter elements are activated in preference to more distal ones (Wasylyk et al., 1983). In addition to activator sequences, a number of negative regulatory sequences ('silencers') have been identified which in some cases exhibit enhancer-like properties (see Section 1.4.6).

### 1.4.1 Viral Enhancer Elements

Enhancer elements were originally identified in the SV40 viral genome following deletion of an essential cis-acting region within 100 bp of the early gene transcription start site (Benoist and Chambon, 1981; Gruss et al., 1981), concomitantly reducing T-antigen expression by a factor of 100-fold and abolishing virus viability. It was subsequently established that sequences within this region activate early



viral gene expression, function in either orientation (Moreau et al., 1981) at positions distant from other necessary control sequences (Lee et al., 1981) and may activate genes from heterologous promoters such as  $\beta$ -globin (Capecchi, 1980; Banerji et al., 1981). Structurally the SV40 enhancer is composed of a tandem 72 bp repeat element, each duplicated sequence containing a 'core' motif (TGTGGAAAG; Weiher et al., 1983; Laimins et al., 1982) and an 8 bp purine/pyrimidine sequence, in addition to a second 8 bp purine/pyrimidine motif of different sequences that overlaps another core consensus sequences just upstream of the repeat element. Recent deletion studies have identified two functional domains in each repeat element, one overlapping the characteristic 'core' motif and a second harboring the purine/pyrimidine stretch of sequences (Zenke et al., 1986; Herr and Clarke, 1986). A third element required for transcriptional activation is upstream of the 72 bp repeats and encompasses a 'core' consensus-type sequence in addition to the upstream 8 bp pyrimidine/purine motif (Herr and Clarke, 1986).

No extensive sequence homologies between enhancer elements from different classes of viruses appear to exist, a feature illustrating the considerable heterogeneity between these elements. Nevertheless, a number of short and degenerate consensus sequences are shared by different sets of enhancers and in some instances harbor more than one short consensus element. Amongst the consensus sequences identified in viral enhancers include the 'core' motif  $\text{TGG}^{\text{AAA}}_{\text{TTT}}\text{G}$  (Weiher et al., 1983) which is common to SV40, polyoma, adenovirus, bovine papillomavirus and Molony murine sarcoma virus (Hen et al., 1983), in addition to stretches of alternating purines and pyrimidines (Nordheim and Rich, 1983) which may have the potential to form left-handed Z-DNA. Although enhancer sequences may not share extensive sequence homologies, they are often functionally equivalent as in the case of the Molony murine sarcoma virus enhancer which, in some instances, can substitute for the SV40 enhancer (Levinson et al., 1982).

Tandemly repeated elements of variable length are a common feature amongst viral enhancers, a factor believed to play a key role in their ability to potentiate transcriptional activation (Weber et al., 1984; de Villiers et al., 1984; Rogers and Saunders, 1985). Several exceptions to this exist, however (Hen et al., 1983). For instance, the polyoma enhancer is spread over a relatively large region of about 250 bp (de Villiers and Schaffner, 1981; Tyndall et al., 1981; Veldman et al., 1985) although like the SV40 enhancer, it contains at least two separate regions which can independently confer enhancer function (Herbomel et al., 1984; Veldman et al., 1985). While the human cytomegalovirus enhancer also does not possess classical enhancer architecture, it has retained some of the sequence motifs characteristic of other viral activator elements such as the SV40 'core' sequence (Boshart et al., 1986).

In general, viral activator sequences are located 5' to the transcription units on which they operate. Upstream enhancer elements have now been identified in the genomes of polyoma virus (de Villiers and Schaffner, 1981; de Villiers et al., 1982), within the long terminal repeats of retroviruses (Laimins et al., 1982; Jolly et al., 1983) and within other DNA virus genomes (Hearing and Shenk, 1983; Hen et al., 1983). Exceptions to the upstream positioning of these activators include the bovine papilloma virus enhancer which is located 3' to coding sequences (Lusky et al., 1983; Campo et al., 1983) and the visna virus enhancer which functions 3' to the transcription start-site (Hess et al., 1985).

Apart from anatomical and positional variations between viral enhancers, precedents exist for functional dissimilarities also. The Herpes simplex virus - type I enhancer is distinguished from other canonical viral enhancers by its inability to augment transcription from a downstream position (Preston and Tannahill, 1984), the adenovirus-2

enhancer functions uni-directionally (Hen et al., 1983) and the polyoma virus enhancer has a unique function as it contains sequences not only required for transcriptional potentiation but also for DNA replication (de Villiers et al., 1984).

One of the most remarkable properties of viral enhancers is their differential potential to stimulate transcription in various host cells/species which is probably indicative of the considerable sequence variation which exists between them. This property was initially recognized by Laimins and co-workers (1981) when insertion of a 72 bp repeat element from the Molony murine sarcoma virus promoter into the SV40 early promoter was shown to confer mouse cell-specific transcriptional activation in preference to human cells. The concept of enhancer host-cell specificity was later reinforced by de Villiers et al., (1984) who constructed hybrid recombinants by replacing late non-coding sequences of the polyoma virus with the SV40 enhancer which switched the polyoma virus host cell preference from murine to primate cells. Furthermore, the SV40 enhancer, by promoting T-antigen expression, can selectively direct tumour formation exclusively in the choroid plexus of transgenic mice (Palmiter et al., 1985). Removal of this element drastically reduces the specificity of SV40-related tumourigenesis (Messing et al., 1985). Apart from the primate and murine host cell preference of SV40 and Moloney sarcoma virus, respectively, human papovirus has a strong preferential activity in fetal glial cells (Rogers and Saunders, 1985), lymphotropic papovirus enhancer favours a hematopoietic cell host (Mosthaf et al., 1985) and the human hepatitis B enhancer has a preferred liver cell activity (Shaul et al., 1985). In contrast, several viral enhancers such as the human cytomegalovirus enhancer (see Boshart et al., 1986) display little cell-type or species preference.

#### 1.4.2 Cellular Enhancer Elements

Following the identification and characterization of viral activator sequences, analogous elements in cellular genes have been detected. Indeed, until recently, only a few cellular sequences with enhancer function had been identified. A general trend has emerged suggesting that these cellular cis-acting regulatory elements possess many of the inherent properties of viral enhancers and are now generally credited with conferring a level of selective gene expression including those associated with tissue-specificity, development, differentiation and cellular responses to induction. Gene-specific regulatory elements which confer selective transcriptional control by modulating the frequency of RNA polymerase initiation and differential promoter selection have been characterized by two main approaches. Transcription units mutagenized by systematic 5'- or 3'-end deletion, site-directed mutagenesis or by linker-scanning mutation analysis are routinely assayed in various expression systems including cell-free extracts, tissue culture cells or by introduction into the germ-line of usually Drosophila or mouse. Alternatively, hybrid genes are constructed to test if potential elements confer gene-specific differential regulation on a linked heterologous gene using the same expression systems. Collectively, these investigations have revealed that cellular cis-dominant activator elements can either be located in the upstream, downstream or intragenic region and generally display many of the properties exhibited by viral enhancers.

The first report of a cellular enhancer was made following the identification of sequences located downstream of the immunoglobulin heavy chain J locus which displayed properties both structurally and functionally homologous to some well characterized viral enhancers (Gillies et al., 1983; Banerji et al., 1983; Mercola et al., 1983). This element represented the first non-viral enhancer to which a

discrete function was assigned by accounting for the activation of rearranged variable genes when translocated adjacent to a constant-region gene during lymphoid differentiation.

The immunoglobulin heavy chain enhancer was shown to function in a position and orientation-independent manner and augmented transcriptional activation when fused to a heterologous gene (Gillies et al., 1983). One striking feature of this element is that it potentiates transcription in a strictly lymphoid-specific manner (Gillies et al., 1983; Banerji et al., 1983) although other additional intragenic and upstream regulatory regions have been implicated to contain tissue-specific regulatory information (Grosschedl and Baltimore, 1985; Foster et al., 1985). By itself however, in transgenic mice, the immunoglobulin heavy-chain enhancer confers lymphoid-specific expression on a rabbit  $\beta$ -globin gene when inserted into a linked conalbumin promoter (Gerlinger et al., 1986) and drives the *c-myc* gene to produce lymphoid-specific malignancy (Adams et al., 1985). Another notable feature of this element is its sequence homology to regions within the SV40 and other enhancers (Gillies et al., 1983). These include the GTGG<sup>AAA</sup><sub>TTT</sub>G 'core' motif characteristic of many viral enhancers (Weiher et al., 1983).

#### 1.4.3 Tissue-Specific Regulatory Elements

The majority of genes expressed in a cell-type specific manner are activated by cis-regulatory elements which have a corresponding cell-specific activity. In many cases, tissue-specific regulatory elements exhibit many characteristics common to viral transcriptional enhancers. These elements have generally been resolved by transfection of hybrid genes into tissue culture cells. Usually, these constructs consist of 5'-flanking regions of genes (expressed in a tissue-specific manner) which are fused to a 'reporter' gene. Through this type of analysis, tissue-specific cis-regulatory sequences in cellular genes have generally been located in the 5'-upstream region, usually within

400 bp of the transcription start site. One of the most striking examples of this type is the pituitary-specific activation of the bacterial chloramphenicol acetyltransferase (CAT) gene when fused to upstream regulatory elements from the rat prolactin and growth hormone genes (Nelson et al., 1986). Other examples of tissue-specific regulatory elements which fall into this category are described in Table 1.1.

Many studies involving the generation of transgenic mice have aimed at defining cis-regulatory elements required for differential gene expression. By introducing genes that are normally expressed in a tissue-specific manner, it has been possible to assess which cis-acting elements are involved in differential programming of gene expression. Producing transgenic mice with a series of mutant genes can then enable the precise sequences required for tissue-specific expression to be defined. This approach offers significant advantages over transfection of genes into cell-lines as testing the function of a potential cis-regulator throughout development and differentiation in a whole organism is a more stringent test of an elements tissue-specific role. Transgenesis has widely been used to define tissue-specific regulatory elements by microinjecting genes into the pronuclei of fertilized mouse eggs and also by P-element-mediated integration of DNA constructs into the germ-line of Drosophila (see Table 1.1). Perhaps the best characterized tissue-specific element identified by P-element-mediated transformation of Drosophila lies within a 125 bp region approximately 200 bp upstream of the yp1 yolk protein gene cap site (Garabedian et al., 1986). This element, when placed upstream of a heat-shock gene fused in frame to the E. coli  $\beta$ -galactosidase gene, directs tissue specific expression of the hybrid construct and behaves as a transcriptional enhancer. More impressively, the pancreas-specific regulator of the rat elastase-I gene has been delineated to a 134 bp region in the 5'-promoter region by analyzing expression in transgenic mice, and fulfils the definition of an enhancer when tested for position

and orientation independence (Palmiter and Brinster, 1986). Finally, Hanahan (1985) has elegantly shown cell-specific transcription signals of the rat insulin II gene lie within a 660 bp region 5' to the coding region by introducing an insulin promoter-large T-antigen coding sequence construct into transgenic mice which generated pancreatic-specific  $\beta$ -cell tumours.

While many regions have been identified which harbour cis-acting tissue-specific regulatory elements, it is becoming increasingly apparent that they are akin to immunoglobulin and viral enhancers. Furthermore, many tissue-specific activators contain the 'core' motif characteristic of their viral counterparts. Apart from this motif, these elements display considerable heterogeneity with respect to primary DNA sequence and in contrast to many viral enhancers are generally not composed of repeat elements.

As in the case of viral enhancers, activators of cellular genes are sometimes located downstream of the transcription initiation site. This point is illustrated by the murine heavy chain (Gillies et al., 1983; Banerji et al., 1983) and K-light chain (Emorine et al., 1983; Picard and Schaffner, 1984) immunoglobulin enhancers in the intron of rearranged structural genes. A precedent also exists for a tissue-specific enhancer element 3' to coding sequences. Recently, Choi and Engel (1986) have identified such an element within a 300 bp region, lying approximately 400 bp 3' to the polyadenylation signal in the chicken adult  $\beta$ -globin gene. While no other cellular gene is known to be associated with a 3' tissue-specific element, taken together with the downstream heavy and light chain immunoglobulin enhancers, it is clear that cell-specific control elements do not exclusively reside in the 5' upstream region.

## TABLE 1.1

### REGULATORY SEQUENCES REQUIRED FOR TISSUE-SPECIFIC EXPRESSION OF VARIOUS GENES

This table summarizes some of the general approaches used to identify and characterize cis-regulatory sequence elements required for differential gene control in a range of tissues. Similar approaches have been employed to delineate the DNA sequence requirements for developmentally regulated (Section 1.4.5) and inducible genes (Section 1.4.4). In many cases these regulatory elements exhibit characteristics similar to viral enhancers.



<u>Gene</u>	<u>Hybrid Construct</u>	<u>Expression System</u>
Rat: Insulin Chymotrypsin	5' promoter deletions fused to CAT gene	Hamster pancreatic cell-line rat exocrine cell-line
MHC E $\alpha$ <sup>d</sup> E $\beta$	Wild-type gene, 5' deletions	Transgenic mice not expressing the endogenous gene
Murine $\alpha$ -crystallin	-364 to +45 of 5' region fused to CAT gene	Transgenic Mice
Chicken $\alpha$ -crystallin	Wild type gene, 5' deletions	Mouse lens epithelial cell-line
Rat Insulin II	Insulin promoter - large T antigen coding sequence hybrid	Transgenic Mice
Chicken $\alpha$ -skeletal actin	5' flanking region fused to CAT gene	Primary Myoblast cell-line
Rat Prolactin Growth Hormone	5' promoter deletions fused to CAT gene	Pituitary cell-line
Human: Placental Lactogen	Fused 3' flanking region to CAT gene	Pituitary cell-line
Human antithrombin III	1.2 Kb of 5' flanking region fused to CAT gene	Myeloma cell-line
Murine Collagen I	-2000 to +54 of 5' region fused to CAT gene	Transgenic Mice
Rat Elastase I	5' promoter deletions, heterologous gene fusion	Transgenic Mice
Mouse alphafeto-protein	Alphafetoprotein-albumin gene fusion	Transgenic Mice
Mouse K-light chain $\mu$ -heavy chain	c-myc oncogene coupled to $\mu$ or K enhancer	Transgenic Mice
<u>yp1</u> Drosophila	Inserted upstream of heat-shock gene fused in frame to <u>E. coli</u> $\beta$ -galactosidase gene	<u>Drosophila</u> P-element transformation

<u>Tissue Where Expressed</u>	<u>Characteristics</u>	<u>Reference</u>
Endocrine $\beta$ -cells pancreas	Tissue-specific elements within 300 bp 5' of gene. Marked preference between pancreatic cell types.	Walker <u>et al.</u> , (1983)
Macrophages	Induced by $\gamma$ -interferon. Enhancer within 2 Kb of 5' flanking region.	Pinkert <u>et al.</u> , (1985); Gillies <u>et al.</u> , (1984)
Lens epithelia	5' Tissue-specific Regulatory Information.	Overbeek <u>et al.</u> , (1985)
Lens epithelia	Regulatory sequences in -93 to +58 region.	Hayashi <u>et al.</u> , (1985)
Pancreas	Generates pancreas-specific tumours. Required sequences within 660bp 5' to coding region.	Hanahan (1985)
Myoblasts	Activity in a 511 bp region 5' to gene.	Grichnik <u>et al.</u> , (1986)
Pituitary	Enhancer has 'core' SV40 enhancer motif.	Nelson <u>et al.</u> , (1986)
Various (eg.pituitary, placenta)	Operates in position and orientation-independent manner; 3' to gene. Enhancer not strictly cell-type specific.	Rogers <u>et al.</u> , (1986)
Lymphoid tissue	Enhancer exhibits tissue-specific activity and has marked homology to Ig enhancer sequences.	Prochownik (1985)
Connective Tissue	Sequences required for tissue-specific expression in 5' flanking region.	Khillan <u>et al.</u> , (1986)
Pancreas	Tissue-specific enhancer between -205 and -72.	Palmiter and Brinster (1986)
Visceral endoderm	Tissue-specific regulator 5' to coding region.	Krumlauf <u>et al.</u> , (1985)
B-lymphocytes	Classical enhancer-like properties. Drives <u>c-myc</u> to produce lymphoid-specific malignancies.	Adams <u>et al.</u> , (1985)
Ovarian follicle cells	5' sequences stimulate gene activity regardless of position and orientation.	Garabedian <u>et al.</u> , (1986)

#### 1.4.4 Inducible Genes

Another class of genes which have contributed to the understanding of transcriptional regulation are those modulated by various forms of induction. These stimuli include environmental effectors such as heat-shock, heavy metals and light or, physiological induction by steroid hormones and cyclic AMP.

Regulatory elements modulating inducible transcription are generally localized within 200 bp of the initiation site and typically consist of sequences between 10 and 30 bp in length. One notable exception to this is a glucocorticoid-responsive element in the first intron of the human growth hormone gene (Slater et al., 1984, 1985). It is not uncommon for induction-responsive elements to be reiterated in a dispersed fashion within the immediate upstream region. Heat-shock regulatory elements, inducible elements in human interferon genes, and heavy-metal responsive elements are amongst those repeated in the promoter region of inducible genes although it is not clear why multiple copies have been preserved throughout evolution and, if in fact they all have a regulatory role. The occurrence of multiple copies of small functional sequences is not exclusively confined to inducible elements but appears to be common amongst eukaryotic promoters and enhancers. Many sequences which modulate transcriptional activation of inducible genes can, by themselves, confer inducible gene control when inserted into or alongside the promoter of a heterologous gene and in most cases exhibit enhancer-like characteristics.

A feature of gene-specific regulatory elements is that they are often common to groups of co-regulated genes. For example, a well conserved sequence exists in cAMP-regulated genes including rat phosphoenolpyruvate carboxy kinase (Wynshaw-Boris et al., 1984), human proenkephalin (Comb et al., 1986), prolactin (Cooke and Baxter, 1982), bovine  $\alpha$ -subunit pituitary glycoprotein hormone (Goodwin et al., 1983) and the tyrosine aminotransferase (Shinomiya et al., 1984) genes. This

element is strikingly similar to the consensus element identified in all cAMP-responsive genes of prokaryotes (Ebright, 1982). Conserved consensus-type regulatory elements have also been detected between different heat-shock (Bienz and Pelham, 1986), interferon (Goodbourn et al., 1985) and metallothionein genes (Stuart et al., 1985). This homology is not limited to commonly regulated genes within a species but also is preserved across species barriers as in the case of metal-responsive elements in man, mouse, sheep and Drosophila (Stuart et al., 1984).

The diverse range of agents which selectively trigger gene activation is matched by their requirement for sequence-specific regulatory cis-elements in order to elicit a response. This is indicative of the control mechanisms which exist between different genes and reinforces the concept that gene-specific elements are integrally involved in specifying which genes are turned on and at what particular time.

#### 1.4.5 Developmental Control of Gene Expression

The molecular events controlling differential expression of genes responding to developmental cues are now beginning to be delineated. In several instances, cis-regulatory elements conferring stage-specific transcriptional activation have at least been partially resolved. Much work has focussed on the globin gene family where sequences responsible for correct regulation during erythroid maturation have been identified (Charnay et al., 1984; Wright et al., 1984). Interestingly, these sequences were found both 5' and 3' to the translation initiation site. More recently, an enhancer element has been identified 3' to the chicken  $\beta$ -globin coding region which modulates transcription in a stage-specific manner (Choi and Engel, 1986; Hesse et al., 1986).

In contrast to control sequences downstream of globin genes, most developmentally regulated genes appear to be activated by elements in the immediate 5' flanking region, usually within 200-300 bp of the

transcription initiation site. This has been clearly demonstrated for murine  $\alpha$ A-crystallin (Overbeek et al., 1985) and the  $\alpha_2$  type I collagen (Khillan et al., 1986) gene where 5' flanking sequences accurately confer stage-specific developmental regulation on a linked CAT gene in transgenic mice. Developmentally regulated genes are often also preferentially expressed in a tissue-specific manner. Remarkably, sequences directing stage-specific transcriptional control often coincide with those eliciting tissue-specific expression. Whether separate or common elements evoke these forms of differential control awaits to be demonstrated, although if separate elements exist some interaction is likely to be involved. Furthermore, while many control elements can confer developmental regulation on a heterologous gene, it is uncertain as to whether they also function in a downstream position or in either orientation. Several isolated examples, however, certainly do suggest that enhancers play a pivotal role in stage-specific gene expression. Gehring and co-workers have established, by P-element mediated germ-line transformation, that in the 5' flanking regions of fushi tarazu (ftz), (one of the Drosophila homeo-box containing genes required to establish segmental pattern formation) an enhancer-like element which when fused to the E. coli lacZ gene, directs expression of  $\beta$ -galactosidase in a characteristic 'zebra' pattern in transformed embryos (Hiromi et al., 1985). A more definitive set of experiments involved an element in the 5' flanking region of the developmentally regulated Xenopus GS17 gene. This 74 bp element functions in either orientation, in a downstream position and most significantly activates a heterologous gene in the mid-blastula stage about eight hours after fertilization, in parallel with the endogenous GS17 gene (Krieg and Melton, in press).

#### 1.4.6 Negative Regulatory Elements

It can be predicted that if elements can actively stimulate transcription, negative elements, complimentary to classical

viral/cellular enhancers, may also exist which suppress transcriptional activity, especially in view of analogous elements which have been well characterized in prokaryotes (reviewed by McClure, 1985). In fact, it is becoming increasingly clear that negative elements, commonly referred to as 'silencers', modulate the expression of a number of tissue-specific, inducible and developmentally-regulated genes.

Several examples of 'down' signals in the upstream region of yeast genes have now been identified. These 5' flanking elements, distinct from upstream activator sequences, often modulate transcription in response to catabolite repression such as the His3 (Struhl, 1986) and ADR11 (encoding alcohol dehydrogenase; Beier and Young, 1982) genes which are repressed by elevated glucose levels, and the CYC1 gene (encoding iso-1-cytochrome C) which is regulated by a set of physiological signals including O<sub>2</sub>, heme and catabolite repression (Guarente et al., 1984). The -100 to -150 region of the ADR11 gene has been shown to mediate in glucose catabolite repression by two criteria. First, if the element is deleted or displaced in an upstream direction, a glucose-insensitive, constitutive phenotype results (Beier and Young, 1982, Russel et al., 1983). Furthermore, if these sequences are inserted in the upstream region of a non-glucose regulated gene, such as ADC1 (coding for constitutive alcohol dehydrogenase), it renders the heterologous gene catabolite repressible. The ADR11 catabolite repressor element fails to operate in a position independent manner, in contrast to a corresponding element in the His3 gene. In addition, the heme regulated negative control region in CYC1 has an extensive run of alternating purine and pyrimidine residues which overlaps the upstream 'TATA' box (see Smith et al., 1979). Whether this is significant in relation to its potential to adopt a Z-DNA conformation is a point of conjecture.

Perhaps the best characterized negative regulatory element in yeast is located in the HMRE mating type locus. This element silences the

expression of one mating type allele present at the MAT locus (Brand et al., 1985). Even more remarkable is the ability of this transcriptional 'silencer' to act in a relatively position independent-manner and in either orientation, a feature reminiscent of classical viral enhancers. This system offers a defined mechanism whereby genes are selectively silenced or switched on in order to acquire an a or α mating type.

Transcriptional silencer elements, analogous to the yeast mating-type element, have been identified upstream of the rat insulin I gene (Laimins et al., 1986), the human T-cell lymphotropic virus type III long terminal repeat (Rosen et al., 1985) and the mouse c-myc gene (Remmers et al., 1986). The transcriptional silencer of the rat insulin I gene, located between 2 and 4 kb upstream of the start-site is typical of this class of elements in that it suppresses gene activity in an orientation and position-independent fashion and blocks transcription when linked to a heterologous gene (Laimins et al., 1986). Another class of repressor elements identified in higher eukaryotic genes which bear similarity to elements in yeast consist of regions corresponding to purine-pyrimidine rich tracts. Amongst those which fall into this class include elements 5' to the mouse  $\beta$ -major globin gene (Gilmour et al., 1984) and the cholesterol down-regulated 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG CoA reductase) gene (Osborne et al., 1985) which, when deleted, results in stimulation of transcription by an order of magnitude.

In retrospect it is not surprising that transcriptional silencers have been identified in the upstream region of eukaryotic genes. The examples cited suggest a common mechanism for repressing gene activity, just as there appears to be activator elements, which operate in a precise and co-ordinate fashion. It is conceivable that repressor elements located in insulin and globin genes operate to silence transcription in certain tissues while in others, they are inactive

resulting in differential expression. Undoubtedly, this mode of gene control will be found in a number of other differentially regulated genes. Likely candidates include genes which are developmentally regulated, where transcriptional silencing is vital up to a determined stage of cell/tissue maturation. An interplay between enhancers and silencers is likely, in at least some cases, to be responsible for the overall pattern of gene activity at various times during a cell-cycle or developmental programme.

#### 1.5 STABLE TRANSCRIPTION INITIATION COMPLEXES

The mechanisms of selective transcriptional initiation are not adequately explained alone by the involvement of cis-regulatory elements. Testimony to this is the preferential expression of viral and cellular genes in different cell types, implying that additional factors characteristic to a particular cell (apart from RNA polymerase) are required for control of transcriptional initiation. Indeed, in view of what is known regarding gene regulation in prokaryotes, it is conceivable that control of initiation is mediated by sequence-specific factors which act by facilitating or inhibiting RNA polymerase binding in the promoter region.

Our present understanding of the mechanisms governing transcriptional regulation of eukaryotic genes is limited. An important first step in resolving this problem involved the development of crude cell-free extracts that allowed accurate transcription of exogenous genes by RNA polymerase in vitro and provided the first direct indication that both DNA sequences and cellular (protein) factors are integrally involved in this process. Through this line of investigation it appears that the initiation of RNA biosynthesis is directed by a sequence-specific macromolecular assembly of transcription factors and RNA polymerase on the DNA template, which together, form an initiation complex capable of correctly specifying the first phosphodiester bond.



This has largely been established by the use and fractionation of cellular extracts which have been shown to enable accurate initiation of transcription of class I (Miesfeld and Arnheim, 1984; Cizewski and Sollner-Webb, 1983; Cizewski Culotta et al., 1985; Wandelt and Grummt, 1983), class II (Cizewski Culotta et al., 1985; Davison et al., 1983; Ackerman et al., 1983; Tolunay et al., 1984; Dignam et al., 1983; Manley et al., 1980) and class III (Gottesfeld and Bloomer, 1982; Bogenhagen et al., 1982; Cizewski Culotta et al., 1985; Schaak et al., 1983; Lassar et al., 1983; and reviewed by Brown, 1984) genes.

Collectively these studies indicate a common involvement of soluble factors (present in crude extracts) in transcriptional regulation of the three classes of eukaryotic genes in vitro. Furthermore, it has been generally shown that formation of highly stable pre-initiation complexes in each of the three classes of eukaryotic genes is an absolute prerequisite for transcriptional efficiency and fidelity in vitro (Davison et al., 1983; Miesfeld and Arnheim, 1984). Stable transcription complexes are typically composed of multiple soluble factors which are rapidly sequestered onto a DNA template following its addition to a cell-free extract (Lassar et al., 1983). Similar protein-DNA interactions are also thought to be generally required for control of gene expression in vivo.

## 1.6 EUKARYOTIC TRANSCRIPTION FACTORS

The enzymatic machinery which elicits RNA biosynthesis provides the cell with a means to adjust the patterns of transcription in response to environmental, physiological and developmental cues. In eukaryotes, this regulation is mediated in part by promoter-specific transcription factors (i.e. DNA binding proteins) with the inherent ability to discriminate between distinctive DNA sequence elements generally found in promoter regions of different genes. Of immediate importance then is an understanding of how transcription factors modulate transcription.

This has at least been partly achieved through the identification, purification and characterization of a small number of proteins required for promoter-specific expression. While a considerable number of nuclear factors have been described to interact with some of the gene-specific regulatory elements described (see Section 1.4), only a limited number have been shown to modulate transcription. While many of these auxiliary molecules exhibit the properties to be expected of trans-regulators, such as target selectivity, co-operative binding activity and cell-type specificity, an overwhelming proportion have yet to be confirmed as bona fide transcriptional modulators. Exceptions to this include the polymerase II transcription factors GAL4, GCN4 and SP1, and the polymerase III trans-acting protein, TFIIIA.

#### 1.6.1 The Transcription Factor SP1

The transcription factor SP1 is a sequence-specific DNA binding protein originally identified in HeLa cell extracts (reviewed by Kadonaga et al., 1986; Dynan and Tjian, 1985; Sassone-Corsi and Borelli, 1986) and more recently has been detected in a number of other cultured cell-lines (Briggs et al., 1986). It augments transcription of a select group of polymerase II genes that contain at least one appropriately positioned G/C-box element (see Section 1.3.2) within the promoter. A comparison of 36 different binding sites has revealed a range of binding affinities differing by at least 10 to 20-fold (Kadonaga and Tjian, 1986; Jones et al., 1986). A dodecanucleotide consensus sequence characteristic to all SP1 binding sites has been formulated, 5'- $\begin{array}{c} \text{GG} \\ \text{TA} \end{array} \text{GGCG} \begin{array}{c} \text{GGGC} \\ \text{TAAT} \end{array} \text{-3'}$ , and despite being markedly asymmetric, it can function in either orientation.

Recently, SP1 has been characterized and purified to more than 95% homogeneity by DNA-affinity chromatography (Briggs et al., 1986) revealing the presence of 105 and 95 Kd polypeptides which together are responsible for recognition and interaction with characteristic binding sites. Interestingly, a 50 Kd proteolytic subfragment protein isolated

from human placenta binds with the same target specificity as the 105 and 95 KD proteins characteristic of HeLa SP1, but retains only marginal transcriptional potentiation activity. This strongly implies that one domain of the protein is required for binding specificity and that binding alone is not sufficient to augment transcription. In all cases studied so far, the binding of SP1 to the promoter region can be correlated with in vitro transcriptional activation. More significantly, in many cases including activation of SV40, HSV TK and MTII<sub>A</sub> promoters, SP1 appears to act in conjunction with other cellular transcription factors (McKnight and Tjian, 1986). An important feature of this is that no common regulatory features have been identified among the diverse promoters to which SP1 binds, suggesting this factor is involved in some general transcriptional mechanism which may involve interactions with other, perhaps gene-specific, factors. This may also be true of other factors which bind to general regulatory elements, such as 'CCAAT' transcription factor (CTF) which binds to the 'CCAAT' motif common in many eukaryotic genes (reviewed by McKnight and Tjian, 1986, and see Jones et al., 1987). Once the protein sequence and X-ray crystal structure of SP1 has been resolved, the details of protein-DNA interactions will inevitably be uncovered.

The widespread involvement of general constitutive promoter elements, such as the G/C- and 'CCAAT'-boxes, and their interaction with discrete cellular factors may indicate such protein-DNA complexes play a general role in gene control, perhaps by accelerating a particular step in the initiation reaction. If this were true, the situation would follow established prokaryotic precedents as regulatory proteins in these systems often facilitate either the initial binding of RNA polymerase to a promoter, or the subsequent isomerization of RNA polymerase/promoter complex into an open configuration (see McClure, 1985). Recent lessons from prokaryotic systems such as the ara BAD operon where regulatory binding sites can modulate transcription at

variable positions and sometimes in an inverted orientation (Martin et al., 1986) suggests other similarities to G/C-box/SP1 mediated regulation in eukaryotes. However, it is quite conceivable that novel regulatory mechanisms may be involved as the chromatin template in eukaryotic cells may impose different constraints on the transcriptional apparatus and offer variant regulatory mechanisms, relative to their bacterial counterparts (reviewed by Weintraub, 1985). Also a distinct possibility is that general transcription factors such as SP1 and CTF are composed of multifunctional domains which may be involved in DNA recognition/binding, in catalytic activity or, may interact with other transcription complexes to influence regional DNA conformation.

#### 1.6.2 DNA Binding and Activatory Domains of the Yeast Transcription Factors; GCN4 and GAL4

One of the limiting factors in the understanding of trans-regulation of eukaryotic genes has been the absence of detailed information pertaining to the nature of DNA-binding proteins. Studies on two activator proteins in yeast have made significant contributions to knowledge regarding the nature of eukaryotic trans-regulatory molecules and how they may augment transcriptional modulation. The yeast DNA-binding protein, GCN4 interacts specifically with upstream activator regions in promoters of amino acid biosynthetic genes and coordinately induces their transcription (reviewed by Jones and Fink, 1982). GCN4 protein, synthesized from the cloned gene by transcription and translation in vitro, has been shown to bind specifically to a 12 bp promoter region of HIS3 and three other co-regulated genes critical for transcriptional induction in vivo (Hope and Struhl, 1985). Mutations that abolish his3 induction, without exception, also fail to bind GCN4 implying that binding of this protein to the HIS3 regulatory site directly mediates induction in vivo (Hill et al., 1986). Moreover, the palindromic nature of the optimal binding sequence, ATGACTCAT, suggest that GCN4 protein binds as a dimer to adjacent half sites (Hill et al.,

1986) similar to those recognized by prokaryotic regulatory proteins (Anderson et al., 1981; Oppenheim et al., 1980). A second yeast activator protein of 88 amino acids, GAL4, binds to the control regions upstream of GAL1 and GAL10 genes thereby mediating transcriptional activation in response to glucose induction (Guarente et al., 1982; Giniger et al., 1985).

The action of these regulatory proteins have been studied by similar approaches. By synthesizing a battery of N-terminally deleted GCN4 proteins, Hope and Struhl (1986) have shown the domain required for DNA binding resides in the 42 C-terminal amino acids. A fusion protein consisting of  $\beta$ -galactosidase coupled to end-deletions of the GCN4 protein also identified the DNA binding region to be in the basic C-terminal end, requiring only 74 terminal amino acids (Keegan et al., 1986). Significantly, N-terminally deleted proteins retaining binding activity and selectivity are incapable of inducing transcriptional activation in vivo in both cases. Conversely, Brent and Ptashne (1985) have shown that by fusing the first 74 amino acids of GAL4 to the DNA-binding domain of a bacterial repressor, LexA, the resulting fusion protein can activate a yeast gene placed downstream of a LexA operator. The observation that the hybrid protein binds DNA via the lexA repressor-operator interaction and stimulates transcription via activation by GAL4 implies that a common feature of regulatory proteins is the presence of separate binding and activatory domains within each protein. Functional dissection of GCN4 by constructing LexA-GCN4 hybrid proteins is consistent with this (Hope and Struhl, 1986).

Given that at least some eukaryotic transcription factors can be divided into protein domains according to their binding/recognition or activatory role, two fundamental questions remain, - how does the recognition process occur and, how does the protein elicit activation? The mechanistic aspects of protein-DNA recognition are at present unresolved and await x-ray crystallographic analysis of regulatory

complexes. At present the mechanism of activation by bound protein complexes is unclear but could involve allosteric activation of a component of the transcription apparatus, such as RNA polymerase, or interactions with histones to make local chromatin structure more permissive to other components of the transcriptional machinery. The small acidic region required for activation by GCN4 and GAL4 proteins argues against an enzymatic function such as in catalysis of DNA methylation, or nuclease and topoisomerase activity. Major conformational changes in DNA structure brought about by protein binding is also considered unlikely (Keegan et al., 1986). Evidence generated from the molecular mechanism of transcriptional activation in yeast by GCN4 and GAL4 proteins largely supports the precedents set by prokaryotic activators in that specific protein-protein contacts direct transcriptional modulation (Keegan et al., 1986; Brent and Ptashne, 1985; Hope and Struhl, 1986; Giniger et al., 1985).

### 1.6.3 Metal-Binding Fingers of TFIIIA

The protein TFIIIA acts as both a positive transcription factor for 5S RNA genes (Engelke et al., 1980) and a specific RNA-binding protein that complexes with 5S RNA in Xenopus oocytes to form the 7S ribonucleo-protein particle (Pelham and Brown, 1980). Accurate initiation of 5S ribosomal RNA gene transcription requires the interaction between this 40,000 dalton protein with a 50 nucleotide internal promoter region in the center of the gene for the formation of a stable pre-initiation complex and accounts, in part, for differential expression of oocyte and somatic genes during development (Brown and Schlissel, 1985). Proteolytic cleavage of the 7S particle divides the TFIIIA polypeptide into a 30 Kd amino-terminal domain which retains full binding activity and, a 10 Kd carboxy-terminus domain essential for transcriptional activation (Smith et al., 1984). The 30 Kd domain has been shown by analysis of the TFIIIA cDNA sequence to contain nine tandemly repeated units, each consisting of 30 amino-acid residues (Miller et al., 1985)

and containing two cysteine, and two histidine residues in conserved positions. Each repeated 30 residue motif has been found to bind zinc ions at the invariant cysteine and histidine residues (Diakun et al., 1986; and see Miller et al., 1985) and have subsequently been termed Zn-binding fingers.

Sequence comparisons with other eukaryotic regulatory proteins reveal a distinct similarity with TFIIIA with regard to the potential for Zn-binding fingers which may be directly involved in DNA binding. These include the yeast regulatory protein encoded by ADR1 (Hartshorne et al., 1986) and products of the developmental regulatory locus Kruppel (Rosenberg et al., 1986; Schuh et al., 1987) and Serendipity (Vincent et al., 1985) of Drosophila. This is perhaps indicative that TFIIIA represents a wider class of DNA-binding factors which interact with nucleotide recognition sequences via a mechanism involving Zn-binding fingers.

Rhodes and Klug (1986) have recently presented evidence to account for TFIIIA binding to the 50 bp internal regulatory element. Analysis of the binding region shows it to have an A-type conformation; similar to that required for binding to RNA. Within this region lies a repeated structure with a periodicity of half a helical turn. Remarkably, the nine half-helix repeats correspond precisely to the nine putative binding fingers in TFIIIA. McCall and co-workers (1986) have extended these observations by analyzing the crystal structure of the DNA-binding site and suggest the narrow and deep grooves in A-DNA harbours a polypeptide loop consisting of a twisted  $\beta$ -ribbon such as in a single finger of TFIIIA. While little is known regarding the mechanism of TFIIIA transcriptional activation, Kmiec and co-workers (1986) have recently proposed that factor-mediated DNA gyration could be involved and have proposed a model whereby this accounts for the differential activation of somatic and oocyte 5S RNA genes during development.

From the four eukaryotic transcription factors considered, a common feature to all is the involvement of multifunctional domains, each with a specialized function either in the binding/recognition process or in transcriptional activation. Binding itself is generally not sufficient for a transcription factor to exert an effect. The DNA conformation which TFIIIA recognizes and the mechanism by which this interaction occurs is markedly dissimilar from precedents established in prokaryotic systems. Representative of prokaryotic transcription factors is the bacteriophage 434 repressor (Wharton and Ptashne, 1985) which has an  $\alpha$ -helix-turn- $\alpha$ -helix supersecondary structure (reviewed by Wharton and Ptashne, 1986) which interacts with successive grooves of the major helix of right-handed B-type DNA. DNA-protein recognition is specified by the amino acid sequence of an  $\alpha$ -helical domain which forms strong bonds with DNA-phosphate groups along the DNA backbone (Anderson *et al.*, 1981; Pabo and Lewis, 1982; Bushman *et al.*, 1985). Similar binding mechanisms may also apply for the yeast MAT repressor (Shepherd *et al.*, 1984; Laughton and Scott, 1984) and in the homeo domains of several proteins in higher eukaryotes (McGinnis *et al.*, 1984; Gehring, 1985).

The identification and characterization of transcription factors such as SP1, GCN4, GAL4 and TFIIIA has strengthened the concept that trans-regulatory proteins contribute to the temporal and spatial control of gene expression. While these examples are precedents in providing a framework for what is becoming an established model for eukaryotic gene control, many other sequence-specific DNA-binding proteins have been identified in nuclear extracts which interact with important cis-regulatory elements of both viral and cellular genes. This has generally been achieved using DNAase protection ('footprint') assays (Galas and Schmitz, 1978), protein blotting (Miskimins *et al.*, 1985), nitrocellulose filter-binding assays (Jack *et al.*, 1981), gel mobility-shift assays (Piette *et al.*, 1985), exonuclease III protection assays (Wu, 1984) and light photo-footprinting (Becker and Wang, 1984). Once



identified, sequence-specific DNA-binding proteins are commonly partially purified and used to supplement in vitro transcription systems (Lee et al., 1987) to test for a biological role. In a limited number of cases, eukaryotic transcription factors have been purified to near homogeneity by DNA-affinity chromatography (Kadonaga and Tjian, 1986; Jones et al., 1987; Briggs et al., 1986). In the meantime much impetus is being placed on directly showing that the DNA-binding proteins thus far identified are bona fide transcription factors.

The requirement for a complex array for cis-control elements for efficient promoter function probably reflects a need for corresponding DNA-binding proteins. For example, nuclear factors have been shown to interact with the general (basal) and specialized (inducible) control elements required for induction of the human metallothionein II<sub>A</sub> (MTII<sub>A</sub>) gene by glucocorticoids and heavy metals (Lee et al., 1987). A eukaryotic promoter can therefore be viewed as being composed of modular arrangements of small sequence motifs which bind specific factors. Each motif-protein complex has a defined function such as conferring tissue-specificity, inducibility or in general transcriptional mechanics and together, combine to modulate polymerase initiation. Not only is there a general requirement for co-ordinate interaction between promoter-bound transcription factors (Jones et al., 1985) but also in many cases, co-operative interactions are commonly required for DNA-protein complex formation. This is illustrated by the binding of heat-shock transcription factor (HSTF) to dual recognition sites, each of which has different binding affinities for the same factor, where binding to the second, lower-affinity site, requires the first high-affinity site to be occupied (Topol et al., 1985). This bears many parallels to the binding of  $\lambda$ -repressor dimers to high and low affinity binding sites on the  $\lambda$ -chromosome (Ptashne, 1986a).

#### 1.6.4 Nuclear Protein Interactions with General Upstream Regulatory Elements

Components within crude and partially purified cellular extracts have now been shown to selectively bind to 'TATA'- (Davison et al., 1983; Parker and Topol, 1984; Sawadago and Roeder, 1985; Selleck and Majors, 1987), 'CCAAT'-(Cohen et al., 1986; Jones et al., 1985; Jones et al., 1987) and G/C (see Section 1.3.2) motifs in a number of unrelated promoters. Although SP1 has been purified and shown to transcriptionally activate a large number of RNA polymerase II-transcribed genes through selective binding to the G/C-box, little is known about the mechanism by which the above mentioned factors operate. Interactions between the 'TATA' factor and its cognate recognition sequence is absolutely required for accurate initiation in vitro (Davison et al., 1983) but little else is known regarding its modus operandi. Similarly, the requirements for the 'CCAAT'-box-factor complex are well recognized to be involved in general transcriptional mechanisms (Mellon et al., 1981), in tissue-specific expression by interaction with other factors (Myers et al., 1986; Bienz, 1986; Charnay et al., 1985) and in adenovirus DNA replication (Jones et al., 1987) by an unknown mechanism(s).

As G/C- and 'CCAAT'- motifs are commonly represented several times in a promoter, the reiteration and spatial arrangement of factor-DNA complexes may be of significance in their action. Perhaps an integral function of these complexes lies in their ability to stabilize promoter complexes (Brown, 1984) or in general RNA polymerase initiation processes. For example, they may possess enzymatic properties capable of modifying an incoming polymerase molecule (eg. phosphorylation). Alternatively, it is possible that these factors bind other nuclear proteins which otherwise lack sequence-specificity for promoters (see Learned et al., 1986). A general recurring feature, however, is that interactions with other bound nuclear factors is involved (McKnight

et al., 1985), sometimes positioned at distances far upstream (or downstream) from the initiation site (see Section 1.7).

#### 1.6.5 Gene-Specific Regulatory Elements and their Binding Factors

Nuclear factors have now been identified by conventional means (i.e. DNAase 'footprinting') to interact with gene-specific regulatory elements required for tissue-specific, developmental and inducible transcriptional control. While their function and mechanism of action is far from resolved, interactions between transcription complexes adjacent to one another, or separated by considerable distances are likely to be involved, especially in view of what is known about the nature of transcriptional enhancer elements. Although a specific interaction between a gene-specific element and a corresponding nuclear factor is relatively easy to demonstrate, their role as bona fide transcription factors is generally unproven. However, this possibility is supported by two related sets of observations which are well illustrated by using tissue-specific gene expression as an example. First, functional activity of a gene expressed in a tissue-specific manner can only be supported in extracts from cells in which it is transcribed in vivo (Bazett-Jones et al., 1985). This, combined with the fact that DNA binding proteins from these extracts selectively interact with tissue-specific activator elements is clearly consistent with them being transcriptional activators (Augereau and Chambon, 1986). Furthermore, in a number of cases, nuclear factor binding levels have been shown to correspond to the level of transcriptional activity of the gene to which it complexes. These include increased levels of factor binding to a specific regulatory element following poly(I)-poly(C) induction of the human  $\beta$ -interferon gene (Zinn and Maniatis, 1986), serum stimulation of the c-fos gene (Treisman, 1986), induction of the Drosophila heat-shock gene (Topol et al., 1985) and to the histone H1-box element during S-phase of the cell-cycle (Dalton et al., manuscript in preparation; and work presented in Chapter 5 of this thesis).

The mechanism of promoter recognition by specific DNA-binding proteins is unknown but quite likely involves some of the principles already outlined which apply to TFIIIA, SP1, GAL4 and GCN4 (see Section 1.6). It is tempting to speculate that transcriptional regulation is limited by the availability of specific trans-acting factor which confers some type of differential control. The two most obvious possibilities in this regard are de novo synthesis of the factor in response to some form of induction or developmental cue, or by modification of a pre-existing factor, for example by post-translational modification. Present knowledge implicates the latter possibility to be involved, at least in transcriptional regulation by heat-shock transcription factor (Topol et al., 1985) and activation of the immunoglobulin enhancer (Sen and Baltimore, 1986; Atchison and Perry, 1987).

The binding of nuclear proteins to cellular and viral enhancers can be a complex process which involves multiple factors interacting with discrete sequence domains (Weinberger et al., 1986; Sen and Baltimore, 1986; Davidson et al., 1986); the two most notable examples being the SV40 and immunoglobulin enhancers. Other enhancer elements, such as the heat-shock regulatory sequence (Bienz and Pelham, 1986), consist of a single protein recognition motif which binds one factor. In the case of multi-domain activator elements, an enhancer-specific factor is required for differential transcriptional activation (i.e. tissue-specific expression) in addition to more common factors for instance, which may also bind to the same recognition sequence in a totally unrelated enhancer which is subject to different control. This point is illustrated by the binding of a nuclear factor to an 'octamer' sequence which is common to the SV40 (Bochmann et al., 1987; Davidson et al., 1986) and K-immunoglobulin (Garcia et al., 1986; Singh et al., 1986) enhancers in addition to the promoter of U2 small nuclear RNA genes (Bochmann et al., 1987) and histone H2B genes (see Chapter 5, Sive and

Roeder, 1986). For the K-immunoglobulin enhancer, at least four separate factors are thought to bind within the activatory domain, three are ubiquitous to all tissues and one, which is B-cell specific, is thought to be involved in discriminant promoter activation (Sen and Baltimore, 1986; Weinberger et al., 1986; Atchison and Perry, 1987). When a full compliment of enhancer binding factors are assembled, it assumes full activatory potential.

### 1.7 REGULATION OF TRANSCRIPTION BY PROTEINS ACTING AT A DISTANCE

Already discussed, have been the involvement of DNA binding proteins which exert either a positive or negative level of control on transcription from positions far upstream (or downstream), relative to some of the more general regulatory motifs. By what mechanism(s) then, do distal transcription complexes modulate transcriptional initiation and, how do multiple and spatially separate regulatory factors interact to achieve this? In the formulation of a unifying hypothesis regarding eukaryotic gene control, two elementary principles must be recognized, (i) multiple cis-regulatory elements acting synergistically are required for correct transcriptional control/modulation and, (ii) DNA-binding proteins complexed to these sequence elements interact to determine promoter specificity and to regulate the frequency of RNA polymerase transcriptional initiation. Several models have been proposed to explain how each transcription complex interacts to form the assembled initiation apparatus. In any model describing trans-regulatory control, the role of all components should be viewed in the context of their interaction with other DNA-protein complexes.

This is quite simply explained in cases where proteins interact due to their binding sites being in juxtaposition, for example in the case of many prokaryotic systems (see McClure, 1985) and eukaryotic examples such as the Drosophila heat-shock promoter where a HSTF binding site and a CTF-site are adjacent (Parker and Topol, 1984). However, the

mechanism by which distal transcription complexes exert regulatory control at distances of kilobases or more, is not so obvious.

### 1.7.1 Transcriptional Activation from a Distance: DNA Looping

One mechanism in particular has been proposed to account for long-range trans-activation of transcription and is heavily supported by independent lines of experimental evidence. Briefly, this model entails bending of the intervening DNA between bound trans-acting factors so as to facilitate direct protein-protein contact which, would otherwise only be generally possible if adjacent binding sites were involved (eg. 'CCAAT' and G/C-elements in the HSV TK gene; Jones et al., 1985). This model of transcriptional regulation will initially be covered by considering the evidence from prokaryotic systems.

It is clear that proteins binding to adjacent regulatory sites proximal to prokaryotic promoter sequences interact to modulate transcriptional initiation (reviewed by McClure, 1985; Reznikoff et al., 1985). But, while it is easy to imagine such an arrangement of transcription complexes being involved in the direct facilitation of RNA polymerase initiation, it is more difficult to reconcile how similar complexes can exert an effect from positions further upstream or, downstream such as for the gal (Irani et al., 1983; Majumdar and Adhya, 1984) and araBAD operons of E. coli (Martin et al., 1986; Dunn et al., 1984). Early evidence in favour of the looping model followed the demonstration that repression of the araBAD operon, from an upstream operator site, could only be maintained if the repressor AraC was on the same DNA helical face as other downstream transcription complexes (Dunn et al., 1984, and see Schlieff and Lis, 1985; Schlieff, 1987). De-repression of the araBAD operon by AraC binding on the alternate helical face can be accounted for by its inability to interact with other regulatory elements. By in vivo footprinting, Martin and co-workers (1986) have more recently shown that binding of AraC to the upstream operator requires intact sequences downstream, at and around araI.

These observations not only are consistent with the looping model but indicate that co-operative binding involving other components of the ara system are required.

Perhaps the most direct demonstration of interactions between adjacent DNA-protein complexes comes from studies of the  $\lambda$ -phage repressor (see Ptashne, 1986a,b). Ptashne and co-workers have unambiguously shown that a co-operative binding mechanism operates where sequential binding of a repressor dimer to a high-affinity recognition sequence which is required for binding of a second dimer to an adjacent low-affinity operator site. Essentially, this co-operative binding between adjacent sites on the phage chromosome depends upon protein-protein (dimer-dimer) contacts, involving their carboxy-terminal ends, which increases the affinity of the second uncomplexed dimer for the weaker operator by an order of magnitude. Subsequently, the second repressor dimer complex contacts RNA polymerase which triggers initiation of transcription. A common theme observed here and with bacterial operons, is that interacting transcription complexes must lie on the same helical face of DNA. For instance, DNAase protection ('footprint') assays show that  $\lambda$  repressor dimers co-operatively interact when separated (by insertion of spacer DNA) by integral numbers of helical base turns (Hochschild and Ptashne, 1986). Insertion of intervening spacer DNA, resulting in operator sites being located on opposite helical faces, eliminates dimer-dimer contact and correct transcriptional modulation. In each case where repressor complexes made contact, the resulting DNAase digestion profile indicated a smooth DNA bend of the intervening DNA.

The theoretical and experimental evidence has now been supported by electron micrographs which has enabled loop-structures between  $\lambda$  operator sites to be visualized (Griffith et al., 1986). Remarkably, DNA looping was only apparent when repressor complexes were separated by integral turns of DNA, clearly confirming the structures predicted by

Hochschild and Ptashne (1986). Moreover, considerable evidence from studies of the lac (Mossing and Record, 1986), and deo (Dandanell and Hammer, 1985; Larsen and Hammer, 1987) operons are consistent with those mentioned implicating protein-protein contact and DNA looping as a general step in prokaryotic transcriptional control.

A prime question is, - will this model also generally apply to eukaryotic gene regulation, where DNA topology and arrangement is significantly more complex than in prokaryotes?

In eukaryotes, similar conceptual problems arise regarding the control of gene expression by transcription complexes separated by considerable distances from the RNA polymerase binding site. The situation can be initially simplified by considering interactions between adjacent protein binding sites and the requirements for transcriptional modulation. For example, the HSV TK promoter has a 'TATA'-motif and just upstream, an SP1 binding site either side of a CTF-binding site (see McKnight and Tjian, 1986). Significantly, these protein binding sites reside on the same helical face of DNA; by genetic analysis, are all required for correct transcriptional activation and, computer graphics predicts a direct protein-protein contact between the SP1 and CTF proteins. The idea that specific protein-protein contacts are required in the mechanism of transcriptional regulation by DNA-bound factors is consistent with the role of a specific activatory domain in auxillary proteins such as TFIIIA, GAL4 and GCN4 and, the well documented interaction between carboxy-terminal ends of  $\lambda$  repressor dimers (reviewed by Ptashne, 1986a). How then, can protein-DNA complexes influence transcription over large distances in eukaryotes? The simplest and most likely model, in view of the evidence available, involves looping out of intervening DNA between protein binding sites thereby facilitating protein-protein interactions, as appears to be the case in prokaryotes.



The idea originally put forward for the araBAD operon of E. coli has recently been extended and found to be consistent with the regulation of SV40 early genes. By inserting spacer DNA of different lengths (5-21 bp) between the SV40 enhancer and the 21 bp repeat elements (SP1 binding sites), or in the repeat-'TATA' intervening region; Takahashi and co-workers (1986) have shown that the respective protein binding sites are required on the same helical face in order to generate the normal transcriptional effect. Protein-protein contact by DNA bending is suggested as an explanation for this effect. This is supported by the recent observations of Shuey and Parker (1986) where binding of Drosophila heat-shock transcription factor was shown to induce a conformational change of the promoter in the form of a DNA bend. Furthermore, enhancer activation of immunoglobulin (Atchison and Perry, 1986) and SV40 (Wang and Calame, 1986) genes is consistent with a requirement for enhancer-factors interacting with downstream complexes.

#### 1.7.2 Enhancers: Their Transcription Factors and Mechanism of Action

While DNA looping at present, provides a useful model for how vastly separated transcription complexes interact in many cases, the net effect is not so obvious. Looping could foreseeably direct any one of several potential mechanisms. For instance, the 'chromatin structure model' predicts that enhancer-protein complexes modify chromatin conformation, which is supported by reports that enhancer regions become hypersensitive to nucleases during transcriptional activation (Cereghini et al., 1984). Another possibility is that enhancers serve to modulate superhelicity of linked DNA by providing sites for topoisomerase activity (Yang et al., 1985) where there is a requirement for the enhancer to be under superhelical strain to function, resulting in specific protein binding (Nordheim and Rich, 1983).

However, elegant work by Plon and Wang (1986) suggests that SV40 enhancer-driven transcription of a linked  $\beta$ -globin gene is independent of DNA superhelicity and, Petryniak and Lutter (1987) found that

expression of late SV40 genes does not require substantial topological tension, as has been suggested for Xenopus 5S genes (Ryoji and Worcel, 1984). Enhancers may alternatively act by targeting the transcriptional machinery to a nuclear location where transcription occurs. In light of the considerable evidence that the nuclear matrix is an attachment site for transcriptionally active genes and that attachment sites map to where enhancer activity lies in several genes (see Section 6.1) this seems a plausible model, in at least some cases. The idea that enhancers function as bidirectional entry sites for RNA polymerase also now seems unlikely following the observation that insertion of a heterologous promoter in between the K-enhancer and its natural promoter does not generate a dampening effect (Atchison and Perry, 1986). A possibility also exists where enhancer-binding factors interact with distal or proximal complexes which results in nucleoprotein particle condensation (Ptashne, 1986a). This most likely would entail DNA looping if distal enhancer complexes were involved, and in effect would stabilize the entire set of promoter complexes, as has been proposed by Sassone-Corsi et al., (1985) and Wang and Calame (1986) in their work on the SV40 enhancer. Evidence from the Drosophila heat-shock gene system suggests that interaction between HSTF and CTF aids binding of RNA polymerase to the TATA region by creating a single transcription complex which would involve looping of distal HSTF complexes (Bienz and Pelham, 1986).

In summary, it appears most likely that enhancer-protein complexes may be involved in the formation of stable transcription nucleoprotein structures by interacting with other promoter complexes which facilitates polymerase binding. Looping out of intervening DNA between spatially separate complexes is likely to be involved in assembly of the transcription machinery. This could also involve interactions with subnuclear structures such as the nuclear matrix which may be suitable for anchoring the transcriptional machinery.

## 1.8 HISTONE PROTEINS

The histone protein family consists of five classes of small basic proteins. Core histones (H2A, H2B, H3 and H4) are ubiquitous to all eukaryotic organisms (Isenberg, 1979; Von Holt et al., 1979) and are highly conserved throughout evolution. Together, core histones are involved in the formation of octameric nucleosome core particles which consist of equimolar amounts of H2A, H2B, H3 and H4 protein (reviewed in McGhee and Felsenfeld, 1980; Thomas, 1983; 1984). The fifth class of histone protein, the H1 or 'linker' histones, bind at the exit and entry points of DNA from the nucleosome core (Thoma et al., 1979; Belyavsky et al., 1980) and is thought to be required for the condensation of nucleosome chains into higher order chromatin structures (Thoma and Koller, 1981; reviewed in Reeves, 1984; Thomas, 1983, 1984).

Although generally well-conserved throughout evolution, each histone subtype with the possible exception of H4, is represented by several non-allelic primary sequence variants which display distinct patterns of expression during differentiation (Blankstein and Levy, 1976; Franklin and Zweidler, 1977; Grove and Zweidler, 1984; Lennox and Cohen, 1984; Wu et al., 1983; Zweidler, 1980, 1984), throughout the cell-cycle (Wu and Bonner, 1981; Wu et al., 1984; Wu et al., 1982), in a tissue-specific manner (Zweidler, 1984; Cole et al., 1984; Aviles et al., 1978) and, in response to hormonal antagonism (Gjerset et al., 1982). Furthermore, each histone subtype is subject to post-translational modification (Isenberg, 1979; Annunziato and Seale, 1983).

The distribution and relative abundance of variant H1, H2A, H2B and H3 subtypes have been examined in the chicken. Proportions of H2A, H2B and H3 histone proteins fluctuate independently throughout chicken embryonic development and in adult tissues where the relative quantity of each variant differs considerably (Urban and Zweidler, 1983). Similarly, the ratio of different chicken H1 subtypes fluctuate between tissues and during differentiation of particular cell-lines (Berdikov et al., 1975; Winter et al., 1985a).

Histone variants have been classified by Zweidler (1980, 1984) into three distinct classes on the basis of histone synthesis and the degree of linkage with DNA replication. The major class, replication-dependent variants, predominate in rapidly dividing tissues and their synthesis is tightly linked to DNA synthesis. Considerably less is known about the regulation of other classes of histones. In studies on regenerating mouse liver, partially replication-dependent histones have been identified as a group of variants which are induced during the initial stages of S-phase, but synthesis continues after the cessation of DNA replication and can even proceed up to a precise stage of differentiation. Expression of a third class of histones, the replication-independent variants, is completely uncoupled from DNA replication. Also known as 'replacement' variants, these histones gradually accumulate during cellular maturation, partially replacing replication-dependent histones in chromatin.

Apart from the histone subtypes already described, there is a small number of highly variant sub-types which are of particular interest as they may have the capacity to drastically modify chromatin structure in vivo. In chicken, two extremely variant histone proteins, H5 and H2A<sub>F</sub>, have been identified. H5 is a linker histone found in the nucleated erythroid cells of some non-mammalian vertebrates (Neelin et al., 1968; Miki and Neelin, 1977; Aviles et al., 1978), it shares sequence and structural similarities with histone H1 (Yaguchi et al., 1979; Von Holt et al., 1979; Aviles et al., 1978) and displays strong similarities with the mammalian linker histone variant, H1<sup>o</sup> (Smith et al., 1980; Cary et al., 1981). Like H5, H1<sup>o</sup> appears to accumulate during cellular maturation but its expression is not restricted to just one cell-type (Gjerset et al., 1982). Incorporation of H5 protein into chromatin increases during erythroid cell maturation, partially replacing H1, which is thought to generate pseudo-higher-order structures (Thomas et al., 1985). Thus, H5 may contribute to chromatin condensation observed in terminally differentiated erythroid cells.

An extremely variant H2A cDNA clone, designated H2A<sub>F</sub> (Harvey et al., 1983), encodes a protein which bears only 60% homology to the most abundant H2A in chicken erythrocyte chromatin. Interestingly, an H2A<sub>F</sub>-like protein detected in mammalian cells, H2A.Z, is enriched in active chromatin (Gabrielli et al., 1981; Allis et al., 1982). This protein is related to a Tetrahymena histone protein variant (hvl), found exclusively in the transcriptionally active macronucleus (Allis et al., 1980; Allis et al., 1986). It is conceivable that these related variants play some fundamental role in chromosomal organization of transcriptionally active DNA by altering the properties of nucleosome core particles (see Grove and Zweidler, 1984).

In summary, histones are a ubiquitous family of highly basic proteins, consisting of conserved non-allelic primary sequence variants within each subtype. The occurrence of precisely timed, stage-specific switches in synthesis of histone subtypes during development, suggests they play a key mediatory role in directing chromatin-associated cellular processes.

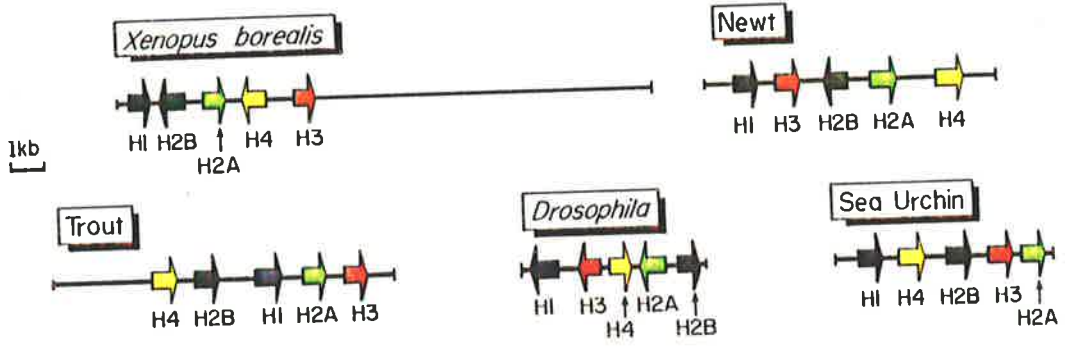
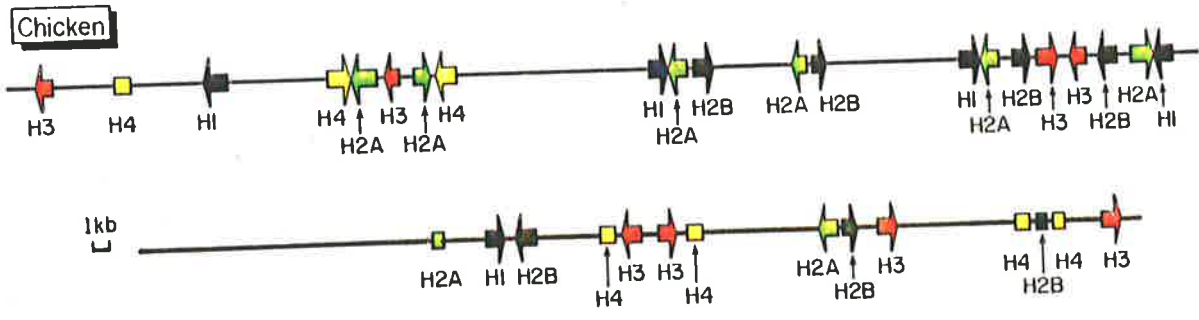
### 1.9 ORGANIZATION OF HISTONE GENES

Histone gene copy number and organization varies considerably between organisms (see Figure 1.2) and has been the subject of several recent reviews (Hentschel and Birnstiel, 1981; Maxson et al., 1983a; Old and Woodland, 1984; Stein et al., 1984). Studies on histone gene organization have shown that genes for the five principle classes of histone protein are usually clustered but their arrangement varies even between closely related organisms. Organization ranges from randomly arranged dispersed clusters of genes (such as in chicken) to highly regular, tandemly repeated quintets (such as sea urchin 'early' genes). Where histone gene copy number is high, the genes are typically organized into a highly reiterated quintet structure (see Figure 1.2). In sea urchin, this entails a head to tail quintet repeat spanning 6 to

## FIGURE 1.2

### DISPOSITION OF HISTONE GENE ARRANGEMENT IN SELECTED ORGANISMS

The data for chicken represents two separate chromosomal regions containing 35 of the 42 genes mapped in total (D'Andrea et al., 1985). Some 70% of Xenopus borealis histone genes occur in a quintet structure within a 15 kb region (Turner and Woodland, 1983). Each histone gene cluster in the newt, Notophthalmus, is separated by 50-100 kb satellite DNA spacer sequences (Stephenson et al., 1981a,b). Genes of the trout, Salmo gairdnerii, are clustered in a head to tail arrangement consisting of a quintet repeat unit within a 10 kb region (Connor et al., 1984), tandem repeats of Drosophila span approximately 5 kb (Lifton et al., 1977) and those of sea urchin 6-7 kb (Maxson et al., 1983a,b). Arrows indicate the relative direction of transcription where it is known. Note the scale for the chicken system is one-half of all other systems represented.



7 kb in length, each quintet unit arranged in the order  $H1-H4-H2B-H3-H2A$  (arrows indicate the direction of transcription) being reiterated several hundred times (see reviews by Kedes, 1979; Hentschel and Birnstiel, 1981; Maxson et al., 1983a,b). Reduced histone gene number correlates with a breakdown of such organization. This is most evident in the higher vertebrates such as birds (Harvey et al., 1981; Sugarman et al., 1983; Ruiz-Carrillo et al., 1983; D'Andrea et al., 1985) and mammals (Sittman et al., 1981; Marzluff and Graves, 1984; Zhong et al., 1983; Heintz et al., 1981; Sierra et al., 1982; Carozzi et al., 1984) in addition to lower eukaryotic organisms such as yeast (Hereford et al., 1979; Smith and Murray, 1983; Choe et al., 1985; Matsumoto and Yanagida, 1985), Neurospora crassa (Woudt et al., 1983) and Tetrahymena thermophila (Bannon et al., 1983, 1984). Histone genes in these organisms are, however, not completely disorganized. In the chicken, for example, preferred histone gene arrangements are observed between H2A/H2B and H3/H4 gene pairs (D'Andrea et al., 1985) although there is no apparent long-range order. Histone gene copy number appears to vary according to the requirements for histone gene transcription. This is best illustrated by the demand for large quantities of histone protein in sea urchin during the period of rapid cleavage between fertilization and late blastula (Maxson et al., 1983a,b; Section 1.11.1), hence, a high reiteration frequency of 'early' genes which are transcribed during embryogenesis (Old and Woodland, 1984).

Histone genes have been isolated from the chicken genome which are separated from the major core and H1 genes. Each of these genes code for a replication-independent protein variant (see Section 1.8). One of these codes for the H1-related, erythroid-specific H5 linker histone. This gene is present as a single copy, is not closely linked to any other histone gene (Krieg et al., 1983) and, is expressed exclusively in erythroid cells (Shannon et al., 1985; see Section 3.2). Another apparently solitary histone gene isolated is the H2A<sub>F</sub> gene (Harvey



et al., 1983; Robins et al., manuscript submitted for publication). It is likely that this gene codes for the replication-independent M1/H2A.Z, H2A protein variant described by Urban et al. (1979) (Section 1.8). The expression of these two highly variant histone genes is described in Chapter 3.

Two other solitary histone genes, H3.3A and H3.3B have been isolated (Engel et al., 1982; Brush et al., 1985) which code for the chicken replication-independent H3 variant, H3.3 (Urban et al., 1979; Urban and Zweidler, 1983). Unlike the major core and H1 genes, transcripts from the four chicken variant histone genes are polyadenylated and, in addition to the H2A<sub>F</sub> genes (Robins et al., manuscript submitted for publication), H3.3 genes contain introns (Brush et al., 1985)

## 1.10 HISTONE GENE ARCHITECTURE

The regulated expression of histone genes plays a pivotal role in programming of cellular differentiation and development. Towards an understanding of these processes an account of histone gene micro-structure will be given, specifically concentrating on DNA sequence elements such as the promoter region and conserved 3' elements required for mRNA termination, which are potentially involved in control of gene expression. An obvious feature of the majority of histone genes is the absence of intervening sequences and, unlike most RNA polymerase II-transcribed genes, are generally transcribed into mRNAs which are not polyadenylated (see Stein et al., 1984).

### 1.10.1 General Promoter Elements

The 5' termini (cap site) of sea urchin mRNAs map within a sequence with a consensus 5'- $\begin{matrix} \text{T} \\ \text{C} \end{matrix} \text{ATTTC} \begin{matrix} \text{G} \\ \text{A} \end{matrix} \text{-3'}$  (Hentschel et al., 1980; Sures et al., 1980), although this motif is not conserved at equivalent positions in histone genes of other species. At approximately 20-30 bp upstream from the transcription initiation site is a conserved 'TATA'-box element

which is also characteristic to most RNA polymerase II-transcribed genes (see Section 1.3.1). Between 40 and 50 bp further upstream, many histone genes have a 'CCAAT'-element (Section 1.3.2) which, in some cases such as in the intergene region of chicken H2A/H2B divergent gene pairs, is present in multiple copies (Wang et al., 1985; Sturm, 1985). Another general promoter element, usually found within 100 bp of the 'TATA'-motif in some histone genes is the G/C-rich element (Section 1.3.2), which again is often represented several times. Between 'TATA' and 'CCAAT' elements of sea urchin histone genes resides a sequence 5'-GATCC-3', which appears to be restricted to genes within this one specie (Hentschel and Birnstiel, 1981) but may be necessary for their expression (Etkin and Di Berardino, 1983; Etkin and Maxson, 1980).

As the majority of these elements are common to a wide range of unrelated genes transcribed by RNA polymerase II, it is likely they play a general role in transcriptional initiation mechanisms and not specifically in selective promoter recognition associated with developmental switching and co-ordinated expression throughout the cell-cycle. Insight into how individual histone genes are selectively activated has been advanced by the identification of gene-specific sequence elements.

#### 1.10.2 Gene-Specific Promoter Elements

Through extensive sequence analysis, numerous gene-specific elements have been identified (see Table 1.2), the majority of which may be important for transcriptional modulation of their cognate genes. In many cases this awaits experimental verification. These motifs are typically more distal from the cap site than the general promoter elements already described (see Section 1.10.1). Two gene-specific promoter elements are of particular concern to the work presented in this thesis (see Chapter 4). The first, was originally identified by Coles and Wells (1985) to be an H1 gene-specific element (5'-AAACACA-3') which is unique and ubiquitous among H1 genes. Secondly, the role of an

## TABLE 1.2

### HISTONE GENE-SPECIFIC UPSTREAM PROMOTER ELEMENTS

Conserved upstream promoter sequences of histone genes between different species. Notably, no common histone H3 gene-specific element has thus far been identified.

---

	<u>Reference</u>	<u>Distance of Upstream</u>
<u>H1 Genes</u>		
Chicken	Coles and Wells (1985) Coles <u>et al.</u> , manuscript submitted	49-79
Mouse	Cole (1984)	71
Rat	Cole <u>et al.</u> , (1986)	61
<u>Xenopus</u>	Turner <u>et al.</u> , (1983) Perry <u>et al.</u> , (1985)	75-115
Trout	Mezquita <u>et al.</u> , (1985)	69
Sea Urchin	Birnstiel <u>et al.</u> , (1978) Busslinger <u>et al.</u> , (1980)	69
Human	Cited in Perry <u>et al.</u> , (1985)	?
<u>H2A Genes</u>		
Sea Urchin	Grosschedl <u>et al.</u> , (1983)	~ 80
Chicken	Wang <u>et al.</u> , (1985)	125
<u>H2B Genes</u>		
<u>Xenopus</u>	Moorman <u>et al.</u> , (1982)	6
Mouse	Sittman <u>et al.</u> , (1983)	24
Human	Zhong <u>et al.</u> , (1983)	6
Sea Urchin	Sures <u>et al.</u> , (1978); Busslinger <u>et al.</u> , (1980)	25-31
Yeast	Wallis <u>et al.</u> , (1980)	?
<u>Drosophila</u>	Goldberg (1979)	12
Chicken	Sturm (1985)	6

Characteristics  
Element from 'TATA'-box (bp)

Consensus Amongst Six Chicken Genes  
: 5'-AAGAAACACAA<sup>A</sup>G-3'

Cross Species Consensus  
: 5'-<sup>AA</sup>TG<sup>AA</sup>NAAACACA<sup>AA</sup>GG-3'

Core Sequence  
: 5'-AAACACA-3'

...sea urchin: 5'-GCTGCT<sup>G</sup>TA<sup>C</sup> -----<sup>6 bp</sup> CACCAACAGATGG-3'  
...chicken: 5'-TGCTG<sup>TA</sup>GG ----- GCAGAGAGAGGG-3'

Other species have less obvious homology  
Sea urchin sequence has strong homology to Moloney murine sarcoma virus enhancer.

Cross Species Consensus Sequence  
: 5'-CTCATTTGCATNG-3'

Consensus sequence bears homology to the Drosophila heat-shock box (Pelham, 1982).

Sequence element also in promoter regions of small nuclear RNA genes (Mataj et al., 1985) and immunoglobulin genes (Falkner and Zachau, 1984; Parslow et al., 1984).

Chicken consensus sequence  
5'-CTGATTTGCATA<sup>G</sup>C-3'

CONTINUED OVERLEAF

<u>Reference</u>		<u>Distance of Upstream Element from 'TATA'-box (bp)</u>
<u>H4 Genes</u>		
Xenopus	Schaffner <u>et al.</u> , (1978) Clerc <u>et al.</u> , (1983)	12
Sea Urchin	Grosschedl and Birnstiel (1980) McKnight <u>et al.</u> , (1981)	10
<u>M. musculus</u>	Mous <u>et al.</u> , (1985)	9
Human	Struhl (1981)	11
<u>Drosophila</u>	Busslinger <u>et al.</u> , (1980)	8
Chicken	Wang <u>et al.</u> , (1985)	15

### Characteristics

Cross species consensus

5'-GTCA 1-5 bp GTCC-3'

element originally identified in the 5'-region of H2B genes (Harvey et al., 1982) has been investigated. Subsequently, this sequence has been found in immunoglobulin (Falkner and Zachau, 1984; Parslow et al., 1984) and small nuclear RNA genes (Mattaj et al., 1985) although as yet, has not been identified in other histone genes.

### 1.10.3 3'-End Formation of Histone Transcripts

With few exceptions (Molgaard et al., 1980; Krieg et al., 1982; Harvey et al., 1983; Bannon et al., 1983), histone mRNAs are generally not polyadenylated (Hentschel and Birnstiel, 1981; Old and Woodland, 1984; Birnstiel et al., 1985). A number of well conserved sequence motifs in the 3'-untranslated region are essential for efficient and accurate formation of histone 3'-ends (Birchmeier et al., 1982, 1983, 1984; Georgiev and Birnstiel, 1985; reviewed by Birnstiel et al., 1985). The most remarkable amongst these, due to its stringent conservation among diverse species, is the 12 bp hyphenated dyad symmetry element located at the 3'-end of mRNA and which has a consensus sequence 5'-AACGGC<sup>T</sup><sub>C</sub>CTTTTCAG<sup>A</sup><sub>G</sub>GCCACCA-3' (Hentschel and Birnstiel, 1981; Birchmeier et al., 1983;). This element is thought to generate a stable stem-loop structure in the unprocessed transcript which is involved in 3'-end formation (Birchmeier et al., 1983). Histone mRNA 3'-termini generally map at or near to the 3'-end of this motif (Hentschel and Birnstiel, 1981; Maxson et al., 1983b). A second sequence requirement, 5'-CAAGAAAAGA-3', is located 6 nucleotides downstream of the dyad symmetry element but is not well conserved in non-sea urchin species (Birnstiel et al., 1985). In addition to these motifs, an additional 50 to 80 nucleotides of spacer DNA is essential for faithful generation of correct 3'-ends (Birchmeier et al., 1983, 1984; Georgiev and Birnstiel, 1985).

Formation of correct histone mRNA 3'-ends involves the editing of a larger precursor (pre-mRNA) molecule (Price and Parker, 1984; Krieg and Melton, 1984; Birchmeier et al., 1984) rather than by sequence-specific transcriptional termination (see Salditt-Georgieff and Darnell, 1983) or sequential editing and addition of adenosines (i.e. polyadenylation; Darnell, 1982). A need for some form of post-transcriptional processing event was in fact indicated by experiments which showed that histone mRNAs are initially synthesized with heterogeneous end-points, terminating within the first 100-200 bp downstream of the final 3'-end (Birchmeier et al., 1984). The mechanism of histone mRNA 3'-processing was then partly deduced from a series of experiments in which in vitro synthesized histone mRNA precursors with 3'-extensions were accurately and efficiently processed when microinjected into the Xenopus oocyte (Birchmeier et al., 1984; Krieg and Melton, 1984). This is now known to involve endonucleolytic cleavage of the precursor RNA molecule (Gick et al., 1986).

Processing of sea urchin histone mRNA also requires a small nuclear RNA of approximately 60 nucleotides which is involved in formation of a small nuclear ribonucleoprotein (snRNP) (Birchmeier et al., 1984; Stunnenberg and Birnstiel, 1982; Galli et al., 1983; Strub et al., 1984; Birnstiel, 1984). This followed the observation that processing of synthetic sea urchin H3 RNA precursors in Xenopus oocytes requires the prior injection of a small nuclear RNA from sea urchin embryos (Birchmeier et al., 1984). Together with a protein component, this RNA is assembled into U7 snRNP particles which is attributed with supporting histone 3'-end formation in trans (Strub and Birnstiel, 1986). Sequence analysis of cDNA clones made to these U7 RNAs reveals that sequences at their 5'-ends could potentially base-pair within the conserved dyad and 5'-CAAGAAAGA-3' elements in a single-strand loop (Strub et al., 1984;



and see De Lorenzi et al., 1986). More recently, Schaufele et al. (1986) have shown this to be true. The mechanism of 3'-processing involves complimentary base-pairing between the conserved histone sequence and regions of snRNA in the U7 snRNP complex which involves RNA-RNA hybrid formation as a processing intermediate.

## 1.11 HISTONE GENE EXPRESSION

### 1.11.1 Developmental Switching of Histone Genes

Quantitative changes in histone gene expression are observed during development of some organisms. This is most dramatic in the sea urchin which contains two sets of genes transcribed during post-fertilization development (reviewed by Maxson et al., 1983b). The 'early' histone genes are organized in tandem repeating units (400 copies/haploid genome, see Section 1.9), while 'late' genes are represented as 10-20 copies/haploid genome and in a singular or paired arrangement (Childs et al., 1981).

In this organism, the mechanisms involved in regulation of the switch from 'early' to 'late' histone gene expression have been of central interest. 'Early' and 'late' variant mRNAs are expressed during partially overlapping stages of sea urchin embryogenesis. Accumulation of 'early' mRNAs is detectable after meiosis (Angerer et al., 1984, 1985) and thereafter accumulate in the haploid pronucleus of unfertilized eggs where they accumulate to high levels (Showman et al., 1982; DeLeon et al., 1983). During the period of rapid cell division, 'early' mRNA levels increase about 10-fold (Weinberg et al., 1983; Cox et al., 1984) as a result of increased transcriptional activity (Maxson and Wilt, 1982; Weinberg et al., 1983). At very early blastula stage, levels of early mRNAs decline due to decreased synthesis and increased turnover (Maxson and Wilt, 1982; Weinberg et al., 1983). Coincident

with this precipitous fall in 'early' message levels is a sharp rise in 'late' gene transcription at the early blastula stage. Prior to this, 'late' mRNAs do not represent a significant proportion of the histone transcript pool but represent approximately 90% of the total in the one-day-old mesenchyme blastula at which time late histone variants constitute a rising fraction of chromatin-incorporated histone (Maxson et al., 1983a,b). Thus, the histone mRNA pool of whole embryos undergoes a gradual shift from 'early' to 'late' messages, being about equal in concentration around the hatching blastula (18-20 hours after fertilization). The period of quantitative predominance of either 'early' or 'late' histone mRNAs correspond, respectively, to those of rapid cell division and the onset of cellular differentiation and morphogenesis. A similar developmental programme for histone gene expression has been noted during Drosophila oogenesis (Ambrosio and Schedl, 1985).

Unlike the developmental regulation of sea urchin genes, there is no differential expression of different sets of histone genes in mouse eggs and embryos (Graves et al., 1985). The same sets of histone genes are expressed throughout development. In the mouse egg, maternally-derived levels of histone mRNAs are high but following fertilization are rapidly degraded reaching a minimum level at the late two cell - early four cell stage. After this, histone transcript levels accumulate up until the embryo stage. Striking qualitative changes in the composition of histone mRNAs and protein are observed during the various developmental transitions resulting in a changing profile in the synthesis of histone variants.

### 1.11.2 Expression of Replication-Independent and Partially Replication-Dependent Histone Genes

The replication-independent (replacement) variant class of histones constitute a class of primary sequence variants which have a unique pattern of expression during development and differentiation in addition to their constitutive basal levels of synthesis during the cell-division cycle. Within this group are extreme variants, such as histone H5 which exhibits a developmentally distinct pattern of expression during avian erythrocyte maturation and, radical primary sequence variants such as H2A<sub>F</sub> (see Section 1.8).

Cellular differentiation and shut-down of DNA replication is accompanied by a severe drop in the accumulation of mRNAs transcribed from the major core and H1 genes (replication-dependent genes, see Section 1.11.3; Brown et al., 1985; Bird et al., 1985). Concomitant with this is the continued synthesis of some histone proteins and their corresponding mRNAs. For instance, levels of murine H3.3 transcripts are constitutively maintained throughout differentiation and remain unaffected by inhibitors of DNA or protein synthesis (Brown et al., 1985; Sittman et al., 1983). Furthermore, a mRNA transcript in rat myeloblasts, corresponding to a replication-independent H4 sub-species, is itself insensitive to changes in DNA replication and persists during differentiation (Bird et al., 1985). A pattern has therefore emerged suggesting that continued synthesis of replication-independent histone protein variants is a direct reflection of the translatable supply of respective transcripts.

Expression of genes coding for partially-replication independent histones have been studied, albeit to a limited degree. These proteins continue to be synthesized after the cessation of DNA replication but appear to be terminated at a programmed stage of differentiation.

Expression of some murine variant genes become transiently uncoupled from a reduced rate of DNA synthesis accompanying the initial stages of differentiation. However, synthesis of their mRNAs remain sensitive to replication inhibitors (Brown et al., 1985). Decreases in transcript levels corresponds to the decline in corresponding levels of histone protein prior to terminal differentiation.

### 1.11.3 Replication-Dependent (Cell-Cycle Regulated) Histone Genes

Synthesis of histone protein is generally coupled to DNA replication in the cell-division cycle. A close relationship between these two events is supported by the observation that histone protein biosynthesis is abruptly terminated when DNA replication is inhibited (Borun et al., 1967; Gallwitz and Mueller, 1969; Robbins and Borun, 1967). Subsequent studies have revealed that changes in levels of replication-dependent core and H1 histone proteins in eukaryotic cells is reflected by a differential accumulation of corresponding histone mRNAs during S-phase (or late G1) of the cell-cycle (Hereford and Osley, 1981; Sittman et al., 1983; Plumb et al., 1983, 1984; DeLisle et al., 1983; Heintz et al., 1983).

### 1.11.4 Multiple Levels of Histone Gene Regulation

Periodic fluctuations in the synthesis of replication-dependent histones are paralleled by similar changes in levels of corresponding mRNAs. The kinetics of histone mRNA accumulation during S-phase of the cell-cycle have been resolved by monitoring steady-state transcript levels in synchronized populations of cells. These studies have generally shown that there is a co-ordinate stoichiometric relationship between histone mRNA levels and the rate of DNA replication (Plumb et al., 1983; Heintz et al., 1983). Initial detection of histone transcripts generally coincides with initiation of S-phase (Heintz et al., 1983; Dalton et al., 1986a; DeLisle et al., 1983), although

events in late G1 may also be required for their periodic accumulation (Artishevsky et al., 1984; Hereford et al., 1982). Maximum levels of histone transcripts typically coincide with a peak in DNA synthesis and are transiently elevated by between 5-50 fold during this period (Alterman et al., 1984; Plumb et al., 1983; DeLisle et al., 1983; Dalton et al., 1986a, see Chapter 3)

The temporal elevation in histone transcript levels results from a combination of two interacting mechanisms (Heintz et al., 1983; Sive et al., 1984; DeLisle et al., 1983; Sittman et al., 1983). They are: (a) increased rates of transcription and (b) prolonged mRNA stability. Enhanced rates of histone gene transcription have generally been demonstrated by pulse-labelling experiments in either whole cells or isolated nuclei which indicate that levels of nascent transcripts increase significantly during S-phase (Heintz et al., 1983; Alterman et al., 1984; DeLisle et al., 1983; Sittman et al., 1983; Dalton et al., 1986a; see Chapter 3). In HeLa cells for instance, elevated steady-state histone mRNA levels are accompanied by at least a three-fold increase in the synthesis of nascent transcripts (Heintz et al., 1983). Similarly, experiments involving continuous or pulse-chase labelling of the cellular RNA pool has revealed that histone mRNAs rapidly decay outside of S-phase but are stabilized in S-phase (Heintz et al., 1983; Sittman et al., 1983; Alterman et al., 1984). In HeLa cells, for example, half-lives of histone transcripts increase from 8 minutes to 40 minutes in the absence and presence of DNA replication, respectively (Heintz et al., 1983). These reports are supported by the observation that in the presence of DNA synthesis inhibitors; such as hydroxyurea, fluorodeoxyuridine and methotrexate; histone gene transcription is depressed (Marzluff and Graves, 1984; Plumb et al., 1983; Sittman et al., 1983) and their mRNAs rapidly destabilized (DeLisle et al., 1983; Alterman et al., 1984; Heintz et al., 1983).

A considerable body of evidence suggests that transcriptional and post-transcriptional levels of control regulating histone mRNA levels are temporally, but not necessarily coupled functionally with DNA replication. This comes from evidence demonstrating that both levels of control can be uncoupled from DNA synthesis. For instance, some metabolic inhibitors which block DNA replication apparently have no detectable effect on either control of histone gene transcription or transcript stability (Marzluff and Graves, 1984). Notably, inhibitors which do not fall into this category reportedly interfere with deoxynucleotide metabolism. A functional link between dNTP levels and coupling of DNA synthesis to histone mRNA levels has therefore been postulated (see Marzluff and Graves, 1984). At this stage the identity of the putative transducing signal thought to be responsible for dNTP-coupling is unknown, or even if this is relevant to the control of histone mRNA levels during an unperturbed cell-cycle. One complicating feature of this work is the unknown specificity of DNA synthesis inhibitors employed in these studies. Protein synthesis inhibitors can counteract histone transcript destabilization and superinduce histone mRNAs by selectively increasing their stability, even in the absence of DNA replication (Stimac et al., 1984; Sive et al., 1984; Baumbach et al., 1984). The simplest interpretation of the results described is that deoxynucleotide metabolism is involved in the coupling of histone mRNA accumulation with DNA synthesis and that a labile regulatory protein is somehow involved in this process. It has also been proposed that histone biosynthesis is subject to autoregulation whereby levels of histone proteins regulate their own synthesis, perhaps by controlling levels of their corresponding mRNAs (see Stein and Stein, 1984). This seems unlikely, however, in view of evidence which suggests that in the absence of DNA and histone synthesis, protein synthesis inhibitors

superinduce histone mRNA levels (Marzluff and Graves, 1984). Moreover, factors affecting the rapid decay of histone transcripts detected in the absence of DNA synthesis are unclear since nuclear (Alterman et al., 1984) and cytoplasmic (Graves et al., 1987; see Section 1.11.5) determinants have been implicated.

#### 1.11.5 cis- and trans-Regulatory Elements Required for Cell-Cycle Regulation

The cis-regulatory elements required for differential transcription of histone mRNAs throughout the cell-cycle have been delineated by a number of approaches. Several reports indicate that histone genes introduced into cell-lines by DNA-mediated gene transfer require only several hundred base pairs of 5' and 3' flanking sequences for faithful cell-cycle regulation (Alterman et al., 1985; Green et al., 1986; Capasso and Heintz, 1985). Artishevsky and co-workers (1985) have shown that a DNA fragment derived from a hamster H3 gene which contains approximately 1 kb of 5' flanking sequence and encodes the first 20 amino acids of the protein, confers cell-cycle regulation on a bacterial neomycin gene. Sequences in the -70 to -110 region of a human histone H4 gene have also been shown to be essential for transcriptional regulation in an in vitro system (Hanly et al., 1985). Promoter sequences in yeast histone genes have also been identified to be necessary for periodic transcript accumulation during late G1-/early S-phase (Osley et al., 1986). These results are particularly interesting as two cell-cycle specific promoter functions have been identified by deletion analysis and by insertion of sequences into a heterologous promoter. One of these functions periodically activates transcription while the other, represses mRNA synthesis in a cell-cycle dependent manner. Furthermore, these authors suggest that negative regulation plays a predominant role in the transient fluctuation of histone

transcript levels as removal of the repressing function results in a constitutive pattern of expression.

A more comprehensive understanding of how histone mRNA levels are co-ordinately regulated throughout the cell-cycle at the transcriptional level of control is being facilitated by the characterization of interactions between cis-regulatory elements and sequence-specific DNA binding proteins (see Section 1.6). Several potentially important histone gene-specific conserved elements are now recognized to bind sequence-specific nuclear proteins in vitro (but see Mous et al., 1985). These include the H2B-box ('octamer') sequence (Sive and Roeder, 1986; and see Chapter 5), H1-box (Dalton et al., manuscript in preparation, see Chapter 5) and regions in the human histone H4 promoter (Dailey et al., 1986; van Wijnen et al., 1987). Several lines of evidence indicates these DNA-binding proteins to be important for efficient promoter activity, and perhaps in S-phase transcriptional activation. For example, H4 promoter sequences recognized by nuclear factors in vitro are also required for transcriptional efficiency (Hanly et al., 1985; Dailey et al., 1986; Capasso and Heintz, 1985). Furthermore, specific nuclear factors, titratable by promoter sequences, predominate in S-phase of the cell-cycle when transcriptional activity is optimal (Hanly et al., 1985). Work presented in this thesis explores the requirements for a sequence-specific trans-regulatory factor in the periodic activation of histone H1 gene transcription (see Chapters 4,5).

#### 1.11.6 Histone mRNA 3'-Ends and Differential Transcript Stability

An important function determining the steady-state levels of histone mRNAs is their rate of degradation. A structure characteristic of replication-dependent histone mRNAs is the 3'-end hairpin loop (see Section 1.10.3) which, is absent in replication-independent histone mRNAs (see Section 1.10.3) and appears to be required for differential



transcript stability. Several lines of evidence implicates this to be important in determining periodic changes in transcript stability. Fusion of the SV40 early promoter to a 463 bp fragment containing the 3'-terminal half of the mouse H4 gene, led to periodic accumulation of fusion transcripts during S-phase (Lüscher et al., 1985). Presumably, this is indicative that 3'-terminal H4 mRNA sequences contribute to differential transcript stabilization throughout the cell-cycle. Interestingly, a small proportion of SV40-initiated transcripts did not form authentic histone 3'-ends. Accumulation of these extended mRNAs was not apparently coupled to DNA replication, implying that transcript processing (see Section 1.10.3) or recognition of edited mRNAs by some degradative mechanism throughout the cell-cycle is important. More recently, the role of translation in selective degradation of histone mRNAs has been proposed and that this is dependent on a hairpin loop structure present at the 3'-terminus (Graves et al., 1987). Whether this form of degradation is the same as that responsible for differential histone mRNA stability during the cell-cycle is not clear, especially in view that nuclear mechanisms may also be involved (Alterman et al., 1984). Correct 3'-end formation is quite likely to be involved, an idea supported by the observation that replacing histone mRNA 3'-ends with a poly(A)<sup>+</sup> sequence enhances transcript stability (Alterman et al., 1985).

The 5' leader region of histone mRNAs has also been implicated to be important for transcript stability (Morris et al., 1986). Fusion of the first 20 nucleotides of a human H3 leader sequence to a leader-deleted  $\beta$ -globin gene, renders the stability of histone-globin fusion transcripts coupled to DNA replication.

### 1.12 THESIS AIMS

The aims of this project were as follows:

1. To determine if expression of two genes, encoding the extreme histone protein variants H5 and H2A<sub>F</sub>, was coupled to DNA replication.
2. Definition and characterization of the cis-regulatory sequence elements and trans-acting regulatory molecules required for appropriate regulated expression of histone genes in vivo.
3. To assess the role of a sub-nuclear structure, known as the nuclear matrix, in the mechanisms of histone gene regulation.

CHAPTER 2

MATERIALS AND METHODS

## 2.2 MATERIALS

### 2.2.1 Chemicals and Reagents

All reagents and chemicals were of analytical grade or of the highest available purity. Most chemicals and materials were obtained from a range of suppliers. Major sources of the more important chemicals and reagents are listed below.

Acrylamide, agarose, bisacrylamide, ampicillin, DTT, rNTP, dNTPs, NTPs,

BCIG, IPTG and SDS: Sigma Chemical Co., Missouri, U.S.A.

Aphidicolin was a generous gift from Dr. M. Suffness, National Cancer

Institute, Bethesda, Maryland, U.S.A.

Caesium Chloride (optical grade): Harshaw Chemical Co., U.S.A.

Chicken serum: Commonwealth Serum Laboratories, (CSL), Melbourne,

Australia

Chloramphenicol: gift from Parke-Davis and Co., Sydney, Australia.

DMEM and G418-sulfate (Geneticin): Gibco Laboratories, Grand Island, New York, U.S.A.

Fetal Calf Serum: Flow Laboratories, Ayrshire, Scotland; and CSL.

Gentamicin sulphate (Garamycin): Schering Corp., New Jersey, U.S.A.

Glyoxal, Nonidet P40, Formamide, PEG (6000 and 4000) and Bromocresol

Purple: BDH Ltd., Poole, England.

Low Gelling Temperature Agarose: Bethesda Research Laboratories (BRL)

Ltd., Gaithersburg, Maryland, U.S.A.

Mixed Bed Resin (AG 501-X8(D)): Bio-Rad, Richmond, California, U.S.A.

Oligo dT: Collaborative Research, Waltham, Massachusetts, U.S.A.

Tetracyclin: Upjohn Pty. Ltd., Sydney, Australia.

Tetramethyl ammonium chloride: Tokyo Kasei, Tokyo, Japan.

TEMED and Xylene Cyanol: Tokyo Kasei.

Tryptose phosphate broth: Difco Laboratories, Detroit, Michigan, U.S.A.

Versene solution: CSL.

### 2.2.2 Enzymes

The enzymes used in the course of this work were obtained from the sources listed below.

AMV reverse transcriptase: Molecular Genetic Resources, Tampa, Florida, U.S.A.

Calf intestinal phosphatase; lysozyme and RNAase A: Sigma.

E. coli DNA polymerase I, Klenow fragment and T4 DNA ligase: Boehringer Mannheim, Mannheim, Federal Republic of Germany; and Biotechnology Research Enterprises of South Australia (BRESA), Adelaide, Australia.

Proteinase K: Boehringer Mannheim.

Restriction Endonucleases: BRL and Boehringer Mannheim.

DNAase I: Worthington Biochemical Corp., New Jersey, U.S.A.; and Pharmacia, Uppsala, Sweden.

DNAase I (RNAase-free): Promega Biotec, Madison, Wisconsin, U.S.A.

T4 DNA polymerase: BRL and Pharmacia.

T4 Polynucleotide Kinase: Boehringer Mannheim; and U.S. Biochemical Co., Cleveland, Ohio, U.S.A.

Trypsin: Difco.

### 2.2.3 Isotopically Labelled Compounds

[ $\alpha$ -<sup>32</sup>P]dNTPs (1700 Ci/mmol), [ $\alpha$ -<sup>32</sup>P]rUTP (1500 Ci/mmol), [ $\gamma$ -<sup>32</sup>P]ATP (2000 Ci/mmol): BRESA.

[methyl-<sup>3</sup>H]thymidine (5 Ci/mmol): Amersham International, Buckinghamshire, England.

### 2.2.4 Recombinant DNA Clones

Except wherever specified, recombinant DNA clones and cloning vectors (Section 2.2.5) were obtained from the individuals indicated within the Department of Biochemistry, University of Adelaide. Details of constructs are given where appropriate in the text.

. Histone H2A, H2B, H3 and H4 genomic subclones spanning the entire coding region of the gene were inserted into pBR322, pBR325 or M13mp8 (Section 2.2.5) and have been described previously (Harvey, 1982; D'Andrea, 1985; Coles, 1986): Drs. R.A. D'Andrea and L.S. Coles.

. H2A<sub>F</sub> (Harvey et al., 1983) and H1 (Coles, 1986) cDNA clones were inserted into pBR322 or M13mp8: Drs. A.J. Robins and L.S. Coles

.  $\lambda$ CH5-01 (Histone H5 genomic  $\lambda$ -clone; Krieg et al., 1982):  
Dr. R.A. D'Andrea

. pSV2Neo (Southern and Berg, 1982): Dr. M.F. Shannon.

. pC $\beta$ A (chicken  $\beta$ -actin cDNA clone in pBR322, Cleveland et al., 1980): Dr. R. Harland, Fred Hutchinson Cancer Research Centre, Seattle, Washington, U.S.A.

. pC $\beta$ G (chicken  $\beta$ -globin genomic clone in pBR322, Colman et al., 1983): Dr. A.J. Robins.

. pC $\alpha$ G (chicken  $\alpha$ -globin genomic clone in pBR322, Colman et al., 1983): Dr. A.J. Robins.

. pCFK (chicken feather keratin cDNA clone, Molloy et al., 1982):  
Dr. K. Gregg.

#### 2.2.5 Cloning Vectors

. pBR322 (Sutcliffe, 1978): Dr. R.J. D'Andrea

. pBR325 (Bolivar, 1978): Dr. A.J. Robins

. M13mp8 (Messing and Vieira, 1982): Dr. R.A. Sturm

. M13mp18 and M13mp19 (Messing, 1983): Dr. A.J. Robins

. pAT153 (Twigg and Sherratt, 1980): Dr. R.A. Sturm

. pJL4 (Gough et al., 1985): Dr. M.F. Shannon

#### 2.2.6 Synthetic Oligonucleotides

Synthetic DNA oligonucleotides were synthesized with an Applied Biosystems 380B DNA Synthesizer at the University of Adelaide Centre of Gene Technology by Dr. D. Skingle and S. Rogers.

(a) Oligonucleotides used for Primer Extension Analysis

The gene-specific oligonucleotides used from primer-extension analysis (Section 2.3.15b) were as follows:

H1 primer:

5' dGGCGGGAGCGGTCTCGGACATCGCGG 3'

H2A primer:

5' dAGCGACTGAACACTCAGAGAGCAAAC 3'

H2A<sub>F</sub> primer:

5' dCCAGCCTTCCCACCTGCCATGGTGCCGC 3'

H2B primer:

5' dGGCTCGGGCATAGTGGCACAACGCGC 3'

(b) Gel-Mobility Shift Assay Oligonucleotides

Synthetic oligonucleotides used in gel mobility-shift assays (Section 2.3.23a) were synthesized as a series of complimentary DNA fragments (40-mers). The oligonucleotide homologous to the coding strand of DNA is designated '+' and its complimentary sequence '-'. Complimentary oligomers were designed so that when annealed, an EcoRI and BamHI overhang was generated at the +5' and -5' end, respectively. Further descriptions are given in Chapter 5.

H1-box oligomers: Parental/Wild Type

+ 5'dAATTCTTTTGTAGTCCAAAGAAACACAAATCGAGCACAG3'  
- 3'GAAAACAATCAGGTTTCTTTGTGTTTAGCTCGTGCCTAGd5'

H2B-box oligomers: Parental/Wild Type

+ 5'dAATTCGAAAGGAATGCTTCTCATTTGCATAGAGGGGCTAG3'  
- 3'GCTTTCCTTACGAAGAGTAAACGTATCTCCCCGATCCTAGd5'

ΔH1-box (4): 4 base substitutions

+ 5'dAATTCTTTTGTAGTCCAAAGGACTACGAATCGAGCACAG3'  
- 3'GAAAACAATCAGGTTTCTGATGCTTAGCTCGTGCCTAGd5'  
\* \*\* \*

ΔH2B-box (3): 3 base substitutions

\* \* \*

+ 5'dAATTCGAAAGGAATGCTTCTCAGTAGCCTAGAGGGGCTAG3'  
- 3'GCTTTCCTTACGAAGAGTCATCGGATCTCCCCGATCCTAGd5'  
\* \* \*

(c) Site-Directed Mutagenesis

Oligonucleotides designated as (-) ΔH1-box(4), (-) ΔH2B-box(3), H1-box UI and H1-box R0 (shown below) were used for oligonucleotide site-directed mutagenesis of a 5 Kb BamHI fragment (see Chapter 4) cloned into M13mp19. Mutagenesis was confirmed by dideoxy chain termination sequencing (Section 2.3.20) using the H1 and H2B primers (Section 2.2.6a) to sequence mutated regions in the respective genes in all cases except pCH.H1/UI (see Chapter 4) where the synthetic oligomer, 5' TTTTGAAGAGGAGGAAGT 3', was used to sequence the upstream insertion.

H1-Box Upstream Insertion (UI)

5'dCTATTGCACTTATTATAGAGCTCGATTTGTGTTTCTTTGGACTGTAAACAATTCGTTAATG 3'

H1-box Reverse Orientation (R0)

5'dCCCCGGGAGCCCTTCGGTGTGAGGTTTCTTTGTGTGTTTAGCTCCTAACAAAAGTGATTTTCT 3'

(d) Universal Sequencing Primer (17-mer)

This synthetic oligonucleotide was used for synthesis of M13 probes (Section 2.3.10c), primes 20 nucleotides back from the Hind III insertion site of M13mp8 (Section 2.2.5) and has the sequence:

5'dGTAAAACGACGGCCAGT3'

2.2.7 Bacterial Strains

E. coli MC1061: araD139, Δ(ara,leu) 7697, ΔlacX74, galU<sup>-</sup>, galK<sup>-</sup>, hsr<sup>-</sup>, hsm<sup>+</sup>, strA; (Casadaban and Cohen, 1980).

E. coli JM101: Δ(lac-pro), F' lacI<sup>q</sup> Z Δ M15, traD1; (Messing, 1979).

Both bacterial strains were gifts from Dr. A.J. Robins.



### 2.2.8 Bacterial Media

L-broth (Luria broth): 1% (w/v) bacto-tryptone (Difco), 0.5% (w/v) yeast extract (Difco), 1% (w/v) NaCl, pH 7.0. Where appropriate, L-broth was supplemented with ampicillin (50 µg/ml) or tetracycline (20 µg/ml).

L-agar plates: L-broth containing 1.5% (w/v) bacto-agar (Difco).

Minimal medium: 2.1% (w/v)  $K_2HPO_4$ , 0.9% (w/v)  $KH_2PO_4$ , 0.2% (w/v)  $(NH_4)_2SO_4$ , 0.1% (w/v) tri-sodium citrate.

Minimal plus glucose plates: minimal medium, containing 0.4% (w/v) glucose, 0.0001% (w/v) thiamine and 1.5% (w/v) bacto-agar.

2 x YT broth: 1.6% (w/v) bacto-tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl, pH 7.0.

JM101 was grown in minimal medium, 2 x YT broth and on minimal plus glucose plates whereas MC1061 was propagated in L-broth or on L-agar plates. All media listed in Sections 2.2.7 and 2.2.8 were sterilized by autoclaving, except heat-labile reagents, which were filter sterilized. Media were prepared with bi-distilled water.

### 2.2.9 Tissue Culture Cell-Lines and Media

#### (a) Cell-lines

##### AEV ts34

Designated ts34 A6 L1 LSCC HD2, but abbreviated in this thesis as AEV ts34, the temperature-sensitive clonal cell-line consists of early chicken erythroblasts transformed by the avian erythroblastosis virus, AEV-RAV 2 (Beug et al., 1979; Graf et al., 1976). These cells were kindly provided by Dr. T. Graf, Institute of Virology, German Cancer Research Center, Heidelberg, Federal Republic of Germany.

##### HeLa

Originally derived from a human cervical carcinoma, these cells were a gift from Dr. A. Dunn, Ludwig Cancer Research Institute, Melbourne, Australia.

### T-cells

This cell-line consists of a Marek-virus-transformed chicken T-lymphoid clonal line (Powell et al., 1974) provided by Dr. T. Graf.

### Chick Embryo Fibroblasts

Primary chick embryo fibroblast cultures were prepared as described by Rein and Rubin (1968) and provided by Dr. A. Koltunow, Department of Biochemistry, University of Adelaide.

#### (b) Cell Culture Growth Conditions

Tissue culture cells were maintained at 37°C in an atmosphere of 5% CO<sub>2</sub> by a Forma-Scientific water-jacketed incubator. Unless otherwise stated, cells were routinely grown in Nunc-T75 disposable flasks (80 cm<sup>2</sup>/260 ml; Kamstrup, Denmark) except in situations involving scale-up (Section 2.3.21) when larger 150 cm<sup>2</sup> (650 ml) flasks were used (Costar, Cambridge, Massachusetts, U.S.A.).

The following media (+ supplements) was used for cell culture:

AEV ts34 and T-Cells: DMEM, 10% fetal calf serum plus 2% chicken serum.

HeLa: DMEM plus 10% fetal calf serum.

Chick Embryo Fibroblasts: DMEM, 5% fetal calf serum plus 10% tryptose phosphate broth.

DMEM and tryptose phosphate broth were prepared according to the manufacturers specifications except that DMEM was supplemented with 7 mg/ml extra glucose, 9.52 mg/ml Hepes, 123 µM 2-mercaptoethanol and 4.8 mg/ml NaHCO<sub>3</sub>. Tissue culture media was prepared using water filtered through a Millipore Milli-Q Water System and filter sterilized before use. After filtration, gentamicin was added to a final concentration of 50 mg/ml. A solution of Dulbecco's PBS (Dulbecco and Vogt, 1954) plus 0.001% phenol red (sodium salt) was routinely used to wash cells (Section 2.3.11).

## 2.3 METHODS

### 2.3.1 Isolation of plasmid DNA

#### (a) Growth and amplification of plasmid DNA

A loopful of E. coli glycerol stock transformed with the plasmid to be isolated, was streaked on an L-agar plate (usually supplemented with an appropriate antibiotic to maintain selective pressure for the retention of the plasmid; 30 µg/ml) and incubated overnight at 37°C. A single colony was used to infect 5 ml of L-broth (supplemented with antibiotic), which was then grown overnight at 37°C with vigorous shaking (if amplification of the plasmid was intended). If amplification was not required, a 100 ml solution of L-broth was infected with a single colony and the broth incubated overnight with shaking.

To amplify plasmid, the 5 ml overnight culture was diluted 100-fold into 500 ml of fresh broth (without antibiotic) and grown with aeration at 37°C to an  $A_{600}$  of 1.0, at which time chloramphenicol was added to a final concentration of 150 µg/ml. Incubation continued overnight.

#### (b) Large scale isolation

Cells from the 100 ml overnight or the 500 ml amplified culture were harvested by centrifugation (JA-10 rotor, 6000 rpm for 10 minutes at 4°C). Plasmid DNA was isolated by a modified procedure of the alkaline extraction method of Birnboim and Doly (1979).

Cell pellets were resuspended in 4 ml of 15% (w/v) sucrose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA, 2 mg/ml lysozyme and incubated on ice for 20 minutes. Freshly made 0.2 M NaOH, 1% (w/v) SDS (8 ml) was added and gently mixed until the suspension became almost clear and slightly viscous. The solution was then left on ice for 10 minutes and 5 ml of 3 M sodium acetate pH 4.6 was added and gently mixed by inversion for a few seconds. The tube was maintained on ice for 30 minutes to allow protein, high molecular weight RNA and chromosomal DNA to precipitate,

then centrifuged for 25 minutes at 15,000 rpm (JA-20 rotor) to remove the clot. The supernatant containing supercoiled plasmid DNA was carefully aspirated, avoiding lumps of the precipitate, treated with RNAase A (heat-treated at 80°C for 20 minutes to inactivate deoxyribonucleases; 50  $\mu$ l of a 10 mg/ml stock) for 20 minutes at 37°C, phenol-chloroform extracted, then ethanol precipitated (see Section 2.3.3).

Precipitated nucleic acids were resuspended in an aqueous solution of CsCl (1 mg/ml), ethidium bromide (0.082 mg/ml) and adjusted to a refractive index of 1.390-1.396 by adding CsCl or H<sub>2</sub>O if required. Following centrifugation at 42,000 rpm (Ti 50 rotor, 20°C) for 48 hours, the low band in the density gradient was recovered (supercoiled, circular plasmid DNA), extracted three times in equal volumes of butan-1-ol and the aqueous phase ethanol precipitated. Precipitated DNA was washed in 70% aqueous ethanol, resuspended in the appropriate volume of H<sub>2</sub>O and stored at -20°C.

(c) Miniscreen procedure

Colonies were grown overnight in 2 ml of L-broth plus an appropriate antibiotic at 37°C with continual shaking. The cells were then pelleted by centrifugation for 5 minutes in an Eppendorf microfuge. Plasmid DNA was extracted by the same, but scaled down, procedure as described (Section 2.3.1b).

2.3.2 Preparation of M13-bacteriophage DNA

(a) Isolation of M13 replicative form (R<sub>F</sub>)

A single plaque was toothpicked into 100 ml of 2 x YT broth containing 10 ml of fresh JM101 overnight culture, grown from a single colony picked from a minimal plus glucose plate. This culture was incubated at 37°C with vigorous aeration for 6 hours. The method of isolating M13 replicative form from this culture was the same as that for the isolation of plasmid DNA (Section 2.3.1b). The yield of DNA prepared in this way was approximately 100  $\mu$ g.

(b) Preparation of single-stranded M13 recombinant DNA

Recombinant plaques were toothpicked into 1 ml of 2 x YT broth containing 25  $\mu$ l of an overnight culture of JM101 (grown in minimal medium; Section 2.2.8) and grown with shaking for 5 hours at 37°C. Cells were pelleted by centrifugation in an Eppendorf centrifuge for 5 minutes. To each supernatant, 0.2 ml of 2.5 M NaCl, 20% PEG (6000) was added and, after leaving at room temperature for 15 minutes, the phage pellet was collected by centrifugation. After removing the supernatant, the pellet was resuspended in 0.1 ml of 10 mM Tris-HCl pH 8.0, 0.1 mM EDTA and extracted with an equal volume of buffer (100 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.001% 2-mercaptoethanol)-saturated phenol. The aqueous phase was re-extracted with 0.5 ml of diethyl ether and ethanol precipitated. Phage DNA was collected by centrifugation, resuspended in 25  $\mu$ l of 10 mM Tris-HCl pH 8.0, 0.1 mM EDTA and stored at -20°C.

Phage supernatants, generated after removal of cells from 1 ml cultures, were kept at 4°C as phage stocks. 20  $\mu$ l of these stocks were often used to re-innoculate 1 ml cultures for single-stranded DNA preparation.

2.3.3 Restriction Enzyme Digestion

Restriction endonuclease digestion of DNA was performed using the conditions of each enzyme described by the supplier.

Reactions were terminated by adding EDTA to a final concentration of 5 mM. The reaction mix was then extracted with an equal volume of phenol/chloroform (v/v;1/1). The aqueous phase was adjusted to 0.3 M-sodium acetate and the DNA precipitated by the addition of 2.5 volumes of nuclease-free ethanol. DNA was pelleted by centrifugation at 4°C for 10 minutes in an Eppendorf microfuge or a JA-20 rotor (10,000 rpm). The DNA pellet was washed with 70% nuclease-free ethanol and dried in vacuo before being redissolved in an appropriate volume of water.

Alternatively, digestion was terminated by the addition of a one-third-volume of 3 x urea load buffer (4 M urea, 50mM EDTA, 0.5% (w/v) bromocresol purple, 50% (w/v) sucrose).

#### 2.3.4 Gel Electrophoresis

##### (a) Agarose Gel Electrophoresis

###### (i) Analytical DNA gels

Agarose was dissolved in 1 x TEA buffer (40 mM Tris-acetate, 20 mM Na-acetate, 1 mM EDTA, pH 8.2) and cast either in 14 cm x 14 cm x 0.3 cm vertical slab gel templates or, for horizontal gels, in a Hoeffer HE99 gel template (Hoeffer Scientific Instruments, San Francisco, U.S.A.) or onto 7.5 cm x 5 cm microscope slides. Vertical gels were electrophoresed between tanks containing 1 x TEA at 65 mA for approximately 3 hours while horizontal gels were run submerged in 1 x TEA at 80-100 mA for the same amount of time. DNA samples were loaded in 1 x urea load buffer (Section 2.3.3) and visualized by staining with 10 µg/ml ethidium bromide for 5 minutes followed by destaining with deionized water and examination under ultraviolet light.

###### (ii) Analytical RNA Gels

Agarose was dissolved in 10 mM sodium phosphate buffer pH 6.5 and cast in either vertical slab templates, a Hoeffer-HE99 horizontal gel template or on microscope slides (Section 2.3.4a(i)). Vertical gels were generally electrophoresed between tanks containing 10 mM sodium phosphate buffer pH 6.5 at 65 mA for approximately 3 hours while horizontal gels were run submerged in the same buffer at 65 mA (20 minutes) or, 100 mA (2-3 hours) for mini-gels and large gels (on Hoeffer apparatus), respectively. Electrophoresis buffer was constantly recirculated between tanks to maintain the pH below 8.0, so that glyoxal did not dissociate from RNA. Samples were loaded in 1 x agarose load buffer (4 x agarose load = 50% glycerol, 10 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.5 and 0.5% bromocresol purple) and where appropriate, stained with ethidium bromide and visualized under ultraviolet light.

(iii) Preparative DNA Gels

Low gelling temperature (LGT) agarose was dissolved in 1 x TEA and cast either into vertical templates or on to horizontal slides, as described above. Electrophoresis was carried out at 4°C. DNA was loaded in 1 x urea load.

DNA was detected by brief ethidium bromide staining and the desired bands excised from the gel with a scalpel. Two volumes of 10 mM Tris-HCl pH 8.0, 0.1 mM EDTA were added to the slice, and the agarose melted at 65°C for 15 minutes. An equal volume of buffer-saturated phenol (at 37°C) was added, the phases rapidly mixed and then immediately separated by centrifugation. The aqueous phase was re-extracted with phenol, adjusted to 0.3 M sodium acetate pH 5.5 and the DNA ethanol precipitated. Recovery of DNA from gel slices was usually in excess of 60%.

(b) Polyacrylamide gel electrophoresis

(i) Analytical

Electrophoresis of DNA fragments less than 1 kilobase (kb) in length was carried out in vertical 14 cm x 14 cm x 0.5 mm gels containing 5-20% acrylamide/bisacrylamide (30/1; w/w) polymerized in 1 x TBE buffer (90 mM Tris-borate, 2.5 mM EDTA, pH 8.3) by the addition of 250 µl 10% (w/v) ammonium persulphate and 12.5 µl TEMED. Electrophoresis was performed at 250 volts for approximately 90 minutes. DNA samples were loaded onto gels in either 1 x urea load or 1 x acrylamide load buffer (1 mM EDTA, 10% (w/v) sucrose, 50 mM Tris-HCl pH 7.4, 0.02% (w/v) bromocresol purple, 0.02% (w/v) xylene cyanol). DNA was visualized under UV light following ethidium bromide staining.

(ii) Preparative

DNA fragments that had been fractionated preparatively were excised from the gel and the DNA eluted into 200-500 µl of 10 mM Tris-HCl pH 8.0, 0.1 mM EDTA at 37°C for between 1 and 16 hours. The eluate was

adjusted to 0.3 M-sodium acetate and the DNA ethanol precipitated. If [<sup>32</sup>P]-labelled DNA fragments (Section 2.3.10) were to be eluted from gels, such fragments were first located by exposure to X-ray film at room temperature for a few minutes (see Section 2.3.5). Eluted fragments were not ethanol precipitated. Recovery of DNA from gel slices was between 50% to 90% depending on the size of the DNA fragments.

#### (iii) Denaturing Analytical Gels

Products of primer extension (Section 2.3.15b), DNAase I protection (Section 2.3.23b) assays and dideoxy-chain termination sequencing (Section 2.3.20) were separated by electrophoresis on polyacrylamide gels which included 7 M urea as a denaturant. The gels used were 40 cm x 40 cm x 0.35 mm. Typically, 6% polyacrylamide gels were used for these purposes and were made by mixing together an 81 ml mixture of acrylamide monomer (20:1, acrylamide to bisacrylamide) in 1 x TBE buffer containing 7 M urea, 800  $\mu$ l of 10% (w/v) ammonium persulfate and 65  $\mu$ l of TEMED. The final mixture was poured into a gel template and allowed to polymerize.

Reservoir tanks contained approximately 4 litres of TBE buffer and the gels were pre-electrophoresed for 45 minutes at 25 mA. Debris and urea were removed from sample wells prior to loading by flushing with buffer from a syringe. All gels were run at 20-30 mA and kept at high temperature during electrophoresis to facilitate DNA denaturation.

Gels were fixed with 200 ml of 10% (v/v) acetic acid and washed with 4 litres of 20% (w/v) aqueous ethanol. After drying, gels were autoradiographed for the desired time (Section 2.3.5).

#### (iv) Electrophoresis and Isolation of Histones

Histone protein was isolated from AEV ts34 cells and mature erythrocytes as described by Dalton et al. (1986a). Protein samples were electrophoresed on Triton X-100-acetic acid-urea-polyacrylamide



gels (6 mM Triton X-100, 5% acetic acid, 6 M urea, 15% acrylamide) essentially as described by Zweidler (1978) followed by staining with amido black.

### 2.3.5 Autoradiography and Laser Densitometry

[<sup>32</sup>P]-labelled DNA which had been electrophoresed on polyacrylamide gels was visualized by autoradiography. Preparative gels (Section 2.3.4b(ii)) were covered with a thin sheet of plastic-wrap and a sheet of Fuji X-ray film was placed over the gel enclosed in an Ilford autoradiography cassette and exposed at room temperature for the required amount of time. For detection of low level radiation, autoradiography was carried out with a tungsten intensifying screen at -80°C. After fixing, analytical polyacrylamide gels (Section 2.3.4b(i)) were autoradiographed as described, usually at -80°C with an intensifying screen. Detection of [<sup>32</sup>P]-labelled hybrids immobilized on filter membranes was by autoradiography at -80°C with a tungsten intensifying screen following hybridization, washing and drying (Section 2.3.17).

When DNA was to be eluted from a gel, one of the glass plates of the gel template was removed and plastic wrap placed over the gel. Two strips of tape were adhered to the plastic wrap each side of the tracks concerned and spotted with radioactive marker dots. A sheet of X-ray film was placed on the gel and autoradiographed for the required amount of time. After developing and fixing the film, radioactive marker dots were lined up with the X-ray film and bands excised from the gel for elution (Section 2.3.4b(ii)).

Quantitation of radioactivity displayed on autoradiograms was performed on a Zeineh Soft Laser Densitometer (Model SL-504-XL). Densitometry was carried out in duplicate, usually on two exposures of the autoradiogram.

### 2.3.6 Subcloning of DNA fragments into plasmid and M13 vectors

#### (a) End-filling of DNA fragments and blunt-end ligations

Vector DNA was linearized with a suitable restriction enzyme then dephosphorylated with calf-intestinal phosphatase in a 50  $\mu$ l reaction mix containing 50 mM Tris-HCl pH 9.0, 1 mM MgCl<sub>2</sub>, 0.1 mM ZnSO<sub>4</sub> and 0.18 units of enzyme. After a 1 hour incubation at 37°C, the enzyme was heat inactivated at 70°C for 15 minutes in the presence of 0.5% SDS and the protein removed by phenol/chloroform extraction. The linearized dephosphorylated vector was purified from uncut vector by passaging the DNA through a low melting point agarose gel (Section 2.3.4a(iii)).

Restriction fragments to be subcloned were preparatively isolated from either agarose or polyacrylamide gels. When the DNA fragment had protruding 5' or 3' termini, it was treated with DNA polymerase I Klenow fragment to end-fill the single-strand regions to blunt-ends. This was done, before purification of the fragment by gel electrophoresis, in a 20  $\mu$ l reaction mix containing 10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 30  $\mu$ M of each dNTP and 1 unit of Klenow fragment.

Ligation of insert into vector was performed in a 10  $\mu$ l volume containing 50 mM Tris-HCl pH 7.4, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 0.5 mM ATP and 0.5 units of T4 DNA ligase at 4°C for 16 hours. Sufficient insert to give a three fold molar excess over vector was generally used. 50 ng of plasmid or 10 ng of M13 vector was normally included in a ligation mix. Recombinant molecules were transformed into bacteria as described below (Sections 2.3.7, 2.3.8).

#### (b) Sticky-end ligations

These were performed as described (Section 2.3.6a) except 0.1 units of ligase was used in the reaction and incubation carried out at 14°C for 4-16 hours.

### 2.3.7 Transformation procedure for plasmid recombinants

A single colony of E. coli strain MC1061 was used to infect 5 ml of L-broth which was then grown overnight at 37°C with aeration. One ml of the overnight culture was diluted 50 fold in 50 ml L-broth and grown with shaking to an  $A_{600}$  of 0.8. Cells were then rapidly chilled on ice for 30 minutes, pelleted by gentle centrifugation (HB-4 rotor, 5000 rpm for 2 minutes), washed in a half volume of ice cold 0.1 M  $MgCl_2$ , then resuspended in a one-twentieth volume of ice cold 0.1 M  $CaCl_2$  and left on ice for at least 1 hour. 200  $\mu$ l of this cell suspension was mixed with 100  $\mu$ l of ligation mix diluted in 100 mM Tris-HCl pH 7.4 and left on ice for 30 minutes with occasional mixing. The transformation mix was heated to 42°C for 2 minutes then returned to ice for 30 minutes. After slowly warming to room temperature, 0.5 ml of L-broth was added and the cells incubated at 37°C for 30 minutes. After this time, 3 ml of L-broth containing 0.7% agar was added to the transformation mix and poured onto an L-agar plate containing an appropriate antibiotic, depending on the resistance carried by the plasmid. The plate was incubated overnight at 37°C.

### 2.3.8 Transformation procedure for M13 recombinants

A loopful of E. coli strain JM101, from a minimal plus glucose plate, was used to infect 5 ml of minimal medium which was then grown overnight at 37°C with aeration. The overnight culture was diluted 50 fold into 50 ml of 2 x YT and grown with shaking to an  $A_{600}$  of 0.4. Cells were pelleted by gentle centrifugation (HB-4 rotor, 5000 rpm for 2 minutes), then resuspended in a one-twentieth volume of ice cold 50 mM  $CaCl_2$  and left on ice for at least 1 hour. 200  $\mu$ l of this cell suspension was mixed with a sample of the ligation mix and left on ice for 40 minutes. The transformation mix was then heated to 42°C for 2 minutes. L-broth (3 ml) containing 0.7% agar, 20  $\mu$ l of BCIG (20 mg/ml in dimethylformamide); 20  $\mu$ l of IPTG, (24 mg/ml in water) and 0.2 ml of

a JM101 overnight culture (diluted 1:5 in 2 x YT broth) were then added and the mixture plated directly onto a minimal plus glucose plate. Incubation was overnight at 37°C.

### 2.3.9 Containment Facilities

All manipulations involving recombinant DNA were carried out in accordance with the regulations and approval of the Australian Academy of Science Committee On Recombinant DNA and, the University Council of the University of Adelaide.

### 2.3.10 Preparation of In Vitro Labelled DNA

#### (a) Nick-translation (Maniatis et al., 1975)

100 ng of DNA was labelled in a 20  $\mu$ l reaction mix containing 50 mM Tris-HCl pH 7.8, 5 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 50  $\mu$ g/ml bovine serum albumin, 25  $\mu$ Ci each of [ $\alpha$ -<sup>32</sup>P]dCTP and [ $\alpha$ -<sup>32</sup>P]dATP, 25  $\mu$ M each of unlabelled dGTP and dTTP, 20 pg of E. coli DNAase I and 5 units of E. coli DNA polymerase I. The solution was incubated at 15°C for 90 minutes, phenol/chloroform extracted and the un-incorporated nucleotides removed by chromatography on a Sephadex G-50 column. For use as a hybridization probe, DNA strands were separated by boiling for 2 minutes followed by snap-chilling on ice.

#### (b) End-fill labelling

DNA fragments with 5' overhangs were incubated in a 10  $\mu$ l reaction mix with E. coli DNA polymerase I, Klenow fragment, as previously described (Section 2.3.10(a)) except that certain unlabelled dNTPs were replaced with 25  $\mu$ Ci of labelled [ $\alpha$ -<sup>32</sup>P]dNTPs depending on the nature of the overhang. The reaction was terminated by phenol/chloroform extraction and DNA recovered by ethanol precipitation.

Alternatively, DNA templates with a recessed or blunt 3' terminus were end-labelled with T4 DNA polymerase. Labelling a DNA fragment with this enzyme basically involved a two-step replacement synthesis reaction. In the absence of dNTPs, the enzyme's 3' exonuclease activity

digests DNA which is repaired following addition of dNTPs (and [ $^{32}\text{P}$ - $\alpha$ ]-dATP) through extensive labelling of the digested strand which serves as a primer template for polymerase activity (see Maniatis et al., 1982). The 3' exonuclease step involved incubating the DNA fragment in a 10  $\mu\text{l}$  reaction mix with E. coli T4 polymerase (1 unit), 1  $\mu\text{l}$  10 x medium restriction enzyme buffer (see Section 2.3.3) plus the appropriate volume of  $\text{H}_2\text{O}$  at 37°C. The incubation time was selected according to the extent of labelling required. This reaction mix was added to 1  $\mu\text{l}$  each of 1 mM-dCTP, dGTP, dTTP, 1  $\mu\text{l}$  of 10 x medium restriction enzyme buffer plus 25  $\mu\text{Ci}$  lyophilized [ $^{32}\text{P}$ - $\alpha$ ]dATP and brought to a volume of 20  $\mu\text{l}$  with  $\text{H}_2\text{O}$ . Incubation was for a further 20 minutes at 37°C at which time 1  $\mu\text{l}$  of 1 mM dATP was added and incubated for a further 10 minutes.

(c) M13 Probes

2.5 ng of universal sequencing primer (Section 2.2.6d) was annealed to 500 ng of recombinant M13 single-stranded DNA in a 10  $\mu\text{l}$  hybridization reaction (primer, M13 DNA, 1  $\mu\text{l}$  medium restriction enzyme buffer (Section 2.3.3), adjusted to 10  $\mu\text{l}$  with  $\text{H}_2\text{O}$ ), heated to 70°C for 3 minutes and then at 37°C for 30 minutes. Hybridization mix was then added to 10  $\mu\text{l}$  of 1 mM-DTT, 0.15 mM each of dTTP and dGTP, 25  $\mu\text{Ci}$  each of [ $^{32}\text{P}$ - $\alpha$ ]dATP and [ $^{32}\text{P}$ - $\alpha$ ]dCTP, 1 unit of E. coli DNA polymerase I Klenow fragment in 1 x medium restriction enzyme buffer at 37°C, for synthesis of the complimentary [ $^{32}\text{P}$ ]-labelled DNA strand. After incubation for 15 minutes, the mix was adjusted to 0.25 mM dATP, 0.25 mM-dCTP and incubated for a further 15 minutes. Labelled DNA was then digested with the appropriate restriction enzyme to release double-stranded insert DNA which was then isolated by electrophoresis through a 6% polyacrylamide gel (Section 2.3.4b). For use as a hybridization probe, (Section 2.3.15c) DNA was heat-denatured before use.

(d) 'Oligo-labelling' DNA Fragments (Feinberg and Vogelstein, 1983).

Random oligonucleotides (8-12 mers) generated from nuclease-digested calf thymus DNA (gift from Dr. S. Wilton; BRESA), were annealed to the heat-denatured, single-strand DNA fragment to be radiolabelled. The Klenow fragment of E. coli DNA polymerase I was then used to add nucleotide residues to the 3'-OH terminus of these oligonucleotides which randomly prime along the single-strand template facilitating synthesis of a second, labelled strand. In this way, it is possible to prepare [<sup>32</sup>P]-labelled DNA of high specific activity which can be used as probes in hybridization experiments (Sections 2.3.15a; 2.3.17).

Linearized DNA (50-100 ng) was suspended in 10  $\mu$ l of H<sub>2</sub>O, heat denatured, chilled on ice and added to 12.5  $\mu$ l nucleotide/buffer cocktail (40  $\mu$ M-dATP, dGTP, dTTP; 100 mM Tris-HCl pH 7.6; 100 mM NaCl; 20 mM MgCl<sub>2</sub>; 200  $\mu$ g/ml gelatin and 500  $\mu$ g/ml random primer) plus 25  $\mu$ Ci of lyophilized [ $\alpha$ -<sup>32</sup>P]dCTP. The solution was mixed, 2.5  $\mu$ l enzyme mix (E. coli DNA polymerase I Klenow fragment, 5.0 BRESA units/ $\mu$ l in 50 mM potassium phosphate pH 7.0; 0.25 mM DTT and 50% glycerol (v/v)) was added and then incubated at 40°C for 20 minutes. Reactions were essentially terminated and labelled fragments prepared for hybridization experiments as described (Section 2.3.10a).

(e) 5' End-Labeling Synthetic Oligonucleotides

Synthetic oligonucleotides/primers were 5' end-labelled with T4 polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP. Normally, 50-100 ng of synthetic primer was kinased in a 10  $\mu$ l reaction mix consisting of 50 mM Tris-HCl pH 7.4, 10 mM MgCl<sub>2</sub>, 5 mM DTT, 7  $\mu$ l of lyophilized [ $\gamma$ -<sup>32</sup>P]ATP (~ 35  $\mu$ Ci) and 1 unit of polynucleotide kinase. The reaction was incubated for 45 minutes at 37°C, after which 5  $\mu$ l of formamide load buffer (Section 2.3.15b) was added, and the mixture loaded onto a 20% polyacrylamide gel for purification of labelled DNA (Section 2.3.4b(ii)).

### 2.3.11 Manipulation and Harvesting Tissue Culture Cells

Approximately 16 hours prior to experimental manipulation, cells were seeded in fresh media at  $\approx 1.10^6$ /ml for suspension cultures and 30-40% confluence for attached cells in order to ensure they were highly mitotic and in an exponential phase of growth.

Attached cells (HeLa and chick embryo fibroblasts) were collected by initially washing in 15-20 ml of PBS and after decantation, cells were detached by incubation at 37°C for 5 minutes with an overlay ( $\approx 1$  ml) of trypsin (1 mg/ml) dissolved in 1x versene solution (Section 2.2.1) and resuspended in 200 volumes of chilled PBS. Once resuspended, cells were routinely harvested by sequential steps involving centrifugation (1200 x g, 5 minutes), decantation of the supernatant and resuspension (200 volumes of chilled PBS) a total of three times before being finally pelleted. Alternatively, attached cells were detached by scraping off with a teflon coated policeman and washed as described. Suspension cultures were thoroughly resuspended in media before being harvested.

#### (a) Synchronization of Tissue Culture Cells

##### (i) AEV ts34

AEV ts34 cells were synchronized by addition of 5  $\mu$ g/ml aphidicolin (5 mg/ml stock in dimethylsulphoxide) for 20 hours in order to allow accumulation at the G1/S-phase cell-cycle border. Cells were released from the cell-cycle block by being washed three times in 200 volumes of warm DMEM (minus serum supplements). After being resuspended in DMEM, cells were centrifuged each time at 1200 x g for 5 minutes. Finally, cells were resuspended in warm DMEM plus serum supplements and seeded at a density of  $1.5-2.0 \times 10^6$ /ml and incubated as described (Section 2.2.9b).

(ii) HeLa

Synchronization of HeLa cells was by the method of Heintz et al., (1983). This procedure essentially involved the steps described in Section 2.3.11a(i) but, in addition, a thymidine block was incorporated into the procedure prior to the aphidicolin block.

(b) Kinetic Analysis of DNA Synthesis

The rate of DNA synthesis was monitored by pulse-labelling cells with [<sup>3</sup>H]thymidine (1  $\mu$ Ci/ml) for 15 minutes. Amounts of radiolabel incorporated into nascent DNA was determined by pelleting (centrifugation at  $\sim 12000 \times g$ , 5 minutes) the 10% trichloroacetic acid (TCA) insoluble cellular material 3 times in 50 volumes of chilled 10% TCA followed by 2 washes in chilled 100% ethanol. After lyophilization and resuspension in 50  $\mu$ l of 0.1N NaOH, TCA insoluble material was dissolved in 1 ml of toluene scintillation fluid and assayed on a Beckman Tri-Carb 3255 Scintillation Spectrometer (10 minute counts).

Homogeneity of the cell population with respect to cell-synchronization was determined by pulse-labelling cells with [<sup>3</sup>H]thymidine, centrifuging them onto gelatin-coated microscope slides (gift from Dr. E. Gowans, IMVS, Adelaide) until flattened and coating them with Ilford liquid nuclear emulsion (Ilford Ltd., Basildon, Essex, England; Type L4). Autoradiography and development was as described by Rogers (1979). The proportion of cells synthesizing DNA was determined visually by counting the number of labelled nuclei.

2.3.12 Isolation of RNA

(a) Tissue Culture Cells

After harvesting and washing in PBS (Section 2.3.11), cells were resuspended in 50 volumes of chilled Nonidet P40 (NP40) lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM KCl and 0.5% (v/v) NP40) and left on ice for 10 minutes with occasional vortex mixing. Lysates were centrifuged (10 minutes, 5000  $\times g$ , 4°C) and the recovered supernatant was extracted



with an equal volume of phenol/chloroform (1/1; v/v). Following separation of aqueous and organic solvent phases by centrifugation (4000 x g, 10 minutes, 4°C) the aqueous phase was recovered and re-extracted with an equal volume of chloroform. Nucleic acids in the aqueous phase were ethanol precipitated, resuspended in 0.2-1.0 ml of 40 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.5 mM MgCl<sub>2</sub> and digested with 10 units of RNAase-free DNAase at 37°C for 30 minutes. Following extraction with an equal volume of phenol/chloroform, RNA was stored at -20°C as an ethanol precipitate.

(b) Fractionation of poly(A)<sup>+</sup> and poly(A)<sup>-</sup> RNA

(Aviv and Leder, 1972)

1.8 g of oligo dT-cellulose was packed into a 10 ml disposable column (Bio-rad, EconoColumn) and contaminating nuclease activity was inactivated by flushing through 10 ml of wash buffer (0.1 M NaOH, 5 mM EDTA). The column was neutralized by equilibration with 1 M Tris-HCl pH 7.6, followed by equilibration in high-salt binding buffer (1 mM EDTA, 10 mM Tris-HCl pH 7.6, 0.5% SDS and 0.5 M NaCl). After heat denaturation (65°C for 3 minutes) in 4 ml binding buffer minus NaCl, RNA was loaded on the column after being adjusted to 0.5 M NaCl. A further 10 ml of binding buffer was applied to the column and eluted, unbound, RNA was detected from the A<sub>254</sub> elution profile using a Uvicord column monitor and reapplied to the column. Bound RNA was eluted by flushing the column with 5-10 ml of 1 mM EDTA, 10 mM Tris-HCl pH 7.6 and 0.1% SDS. In general, chromatography was performed at 20-25°C. All RNA fractions were recovered from the elution buffer by ethanol precipitation and stored at -20°C as an ethanol suspension.

2.3.13 Isolation of Nuclei from Tissue Culture Cells

Cells were harvested, washed (Section 2.3.11), resuspended in chilled NP40 lysis buffer (Section 2.3.12(a)) and homogenized with 10-15 strokes of the B (tight) pestle in a Dounce homogenizer. Nuclei were

pelleted by overlaying the cell lysate on 4 volumes of lysis buffer containing 30% sucrose and centrifuged at 800 g for 5 minutes (4°C). The nuclear pellet was washed in 50 volumes of lysis buffer and then in nuclear wash buffer (10 mM Tris-HCl pH 7.4, 5 mM MgCl<sub>2</sub>) plus 0.1% NP40. Nuclei were stored for up to one week at 80°C by resuspending at  $\approx 3 \cdot 10^8$ /ml in storage buffer (50 mM Tris-HCl pH 8.0, 25% glycerol, 0.1% EDTA, 10 mM 2-mercaptoethanol and 5 mM MgCl<sub>2</sub>) before being snap frozen in liquid nitrogen.

#### 2.3.14 Analysis of Nuclear Matrix-bound DNA

##### (a) DNAase Digestion and Nuclear Matrix Preparation

Nuclear matrices were prepared from cells pre-labelled for 24 hours with [<sup>3</sup>H]thymidine (50 nCi/ml) by the addition of an equal volume of 4M NaCl, 10 mM MgCl<sub>2</sub> to nuclei ( $3 \cdot 10^6$ /ml; in nuclear wash buffer plus 0.1% NP40) and gently mixed at 0°C for 15 minutes. DNAase I was added to a final concentration of 25 µg/ml and incubated at 37°C with occasional gentle mixing. At various times, aliquots of the digestion mix were taken and reactions terminated by addition of 25 mM EDTA at 0°C, followed by centrifugation at 15,000 x g for 5-10 minutes at 4°C. The supernatant was again centrifuged and resulting pellets were pooled and treated with 0.5% SDS, 500 µg/ml proteinase K and incubated at 37°C for 4 hours. Digestion of DNA associated with matrix structures was monitored by determining the amount of radioactivity in aliquots from pellets and supernatants prepared at various times.

Low salt extraction of AEV ts34 nuclei used for preparation of nuclear matrices was by the method of Mirkovitch et al., (1984).

##### (b) Restriction Enzyme Digestion of Matrix Bound DNA

Isolated nuclei, were washed twice in 300 µl of restriction enzyme buffer (Section 2.3.3) and digested with the appropriate restriction enzyme (250 units) for 2 hours at 37°C with frequent gentle mixing. Nuclei were then extracted by addition of an equal volume of 4 M NaCl,

10 mM MgCl<sub>2</sub>. The high-salt extracted nuclei were collected by centrifugation at 15,000 x g for 15 minutes. Pellets were washed with high-salt solution also containing 0.5% NP40 and Tris-HCl pH 7.4 followed by a wash in restriction enzyme buffer. Matrix structures were again resuspended in restriction enzyme buffer plus 0.1% NP40 and digested for 3-4 hours at 37°C with the appropriate restriction enzyme (250 units). Digestions with a second restriction enzyme requiring a different buffer were carried out by adjusting the NaCl and Tris concentrations in the reaction mix followed by addition of the enzyme (250 units) and incubation for a further 3-4 hours at 37°C. Reaction mixtures were centrifuged (15,000 x g, 10 min.) and pellets digested in 1% SDS, 500 µg/ml proteinase K at 37°C for 4-5 hours. Supernatant fractions were also SDS/proteinase K treated.

#### (c) DNA Purification and Analysis

Following digestion of DNA with DNAase I or restriction enzymes and proteinase K treatment, DNA from supernatants and pellets (NM) was purified by phenol/chloroform extraction and RNAase A treatment followed by a further phenol/chloroform extraction and ethanol precipitation.

#### 2.3.15 Analysis of RNA

##### (a) Electrophoresis, Transfer to Filter Membranes and Hybridization

Ethanol precipitated RNA (up to 20 µg) was washed in 70% aqueous ethanol, dried in vacuo, re-suspended in 10 µl formamide (deionized in mixed bed resin), incubated at 80°C for 3 minutes and snap chilled on ice. Resuspended RNA was adjusted to 10 mM NaH<sub>2</sub>PO<sub>4</sub> (sodium phosphate buffer) pH 6.5, 0.86 M glyoxal (deionized) in a final volume of 21 µl and incubated at 50°C for 20 minutes. RNA samples were loaded onto 1.5% agarose gels in 1 x agarose load buffer and electrophoresed on vertical or horizontal slab gels (Section 2.3.4a(ii)). The integrity of RNA was routinely checked by electrophoresis on mini-gels (Section 2.3.4a(ii)) with the RNA (≈ 3 µg) prepared as described, but in an appropriately

scaled-down volume followed by ethidium bromide staining and visualization under ultraviolet light (Section 2.3.4a(i)).

Following electrophoresis, RNA was transferred to GeneScreen filter membranes by capillary action, as recommended by the manufacturer (New England Nuclear Corporation, Boston, Massachusetts, U.S.A.) using 25 mM  $\text{NaH}_2\text{PO}_4$  pH 6.5 as the transfer buffer. Transfer of RNA onto filter membranes was usually complete within 18 hours. Filters were then air-dried, baked at 80°C for 2-3 hours, immersed in 20 mM Tris-HCl pH 8.0 at 100°C and allowed to cool (to remove residual glyoxal).

Filters were pre-hybridized for 16-20 hours at 42°C in sealed plastic bags with 0.5 ml/cm<sup>2</sup> of pre-hybridization buffer (50% deionized formamide, 0.2% polyvinyl pyrrolidone, 0.2% bovine serum albumin (Pentax, fraction V), 0.2% ficoll, 50 mM Tris-HCl pH 7.5, 1 M NaCl, 0.1% sodium pyrophosphate, 1% SDS, 10% dextran sulphate, 200 µg/ml denatured salmon sperm DNA). Filters were hybridized for 16-24 hours at 42°C in 0.2 ml/cm<sup>2</sup> of buffer essentially the same as pre-hybridization buffer except that 10 ng/ml of heat-denatured [<sup>32</sup>P]-labelled hybridization probe (nick-translated or 'oligo'-labelled, see 2.3.10a,d), was included. Once hybridization was complete, filters were washed twice in 100 ml of 2 x SSC (1 x SSC = 0.15 M NaCl, 0.015 M sodium citrate) at room temperature, twice in 100 ml of 2 x SSC plus 1% SDS at 65°C, for 30 minutes and finally, twice in 100 ml of 0.1 x SSC at 65°C for 30 minutes.

Filters required for re-probing were washed twice in 250 ml of 5 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.002% polyvinyl pyrrolidone, 0.002% ficoll and 0.05% sodium pyrophosphate with agitation for one hour at 65°C to remove hybridized probe. Alternatively, filters were treated with 50% deionized formamide, 1 x SSC at 65°C for 30-45 minutes. Detection of DNA:RNA filter immobilized hybrids was as described (Section 2.3.5).

(b) Primer Extension Analysis of RNA (McKnight et al., 1981)

Typically, 1.5 ng of each 5'-end labelled primer (Section 2.3.10e) was added to RNA (usually 10-20  $\mu$ g) and ethanol precipitated. The pellet was washed in 70% aqueous ethanol, re-suspended in 10  $\mu$ l of hybridization mix (200 mM NaCl, 10 mM Tris-HCl pH 8.3), heated to 70° for 3 minutes and then primers were allowed to anneal with RNA by incubating at 42°C for 1-3 hours. Following hybridization, samples were incubated with reverse transcriptase (8 units) in a 34  $\mu$ l volume containing 60 mM NaCl, 10 mM Tris-HCl pH 8.3, 10 mM DTT, 500  $\mu$ M dNTPs, 10 mM MgCl<sub>2</sub> for 1 hour at 42°C. Samples were then usually incubated for a further 30 minutes at 37°C with DNAase-free RNAase A (250  $\mu$ g/ml).

Extension products were phenol/chloroform extracted, ethanol precipitated, washed in 70% aqueous ethanol and dried in vacuo. After resuspension in 5  $\mu$ l formamide load buffer (100% formamide, 0.1% (w/v) bromocresol purple, 0.1% (w/v) xylene cyanol, 20 mM EDTA), primer extension products were heat denatured at 100°C for 3 minutes and electrophoresed on 6% denaturing urea-polyacrylamide gels (Section 2.3.4b) and detected by exposure to X-ray film (Section 2.3.5).

(c) In situ Hybridization of Cell Sections

AEV ts34 cells were wax embedded, sectioned and hybridized in situ to gene specific probes (Section 2.3.10c) by the method of Jilbert et al., 1986.

2.3.16 Dot Blot Filter Hybridization

(a) Preparation of Filter Membranes for Dot Blot Analysis

Single-stranded M13 DNA or double-stranded DNA recombinant insert was resuspended in H<sub>2</sub>O, heat-denatured, snap-chilled on ice and adjusted to 6 x SSC so that the final volume of each sample was 200  $\mu$ l. Samples were loaded onto GeneScreen filter membranes, mounted in a Schleicher and Schu'll (New Hampshire, U.S.A.) dot blot manifold. Aliquots of DNA were filtered under gentle vacuum after the filters had been initially

washed in 2 x SSC before being mounted in the manifold. Filters were then washed twice under gentle vacuum with 6 x SSC, removed from the manifold and baked at 80°C for 2-3 hours.

(b) Detection of Nascent Transcripts in Pulse-Labelled Nuclei

Nuclear run-on transcription assays were performed as described by Marzluff (1978). Briefly, in vitro labelling of nuclear RNA was carried out by incubating nuclei ( $1.3 \cdot 10^8$ /ml) in a 150  $\mu$ l volume which contained 50 mM Tris-HCl pH 7.9, 1 mM ATP, 0.5 mM CTP, 0.5 mM GTP, 0.15 M KCl and 100  $\mu$ Ci of [ $\alpha$ - $^{32}$ P]UTP for 15 minutes at 25°C. The reaction was terminated by the addition of SDS to 0.5% (w/v) and adjusted to 100 mM sodium acetate pH 5.2, 1 mM EDTA. Following phenol/chloroform extraction, the aqueous phase was adjusted to 0.2 M sodium acetate pH 5.2 and the RNA was ethanol precipitated. Incorporation of radiolabel into RNA was determined by the number of counts present in 10% trichloroacetic acid insoluble material. All reactions were carried out in duplicate.

DNA dots were applied to GeneScreen nylon membranes as described (Section 2.3.16a). Prehybridization of filters was carried out for 16-20 hours at 52°C in 50% formamide, 5 x SSC, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% SDS, 10  $\mu$ g denatured E. coli tRNA per ml, 0.4% polyvinyl pyrrolidone and 0.4% ficoll. Membrane filters were then hybridized in the same buffer with the addition of  $3 \cdot 10^6$ - $3 \cdot 10^7$  cpm of [ $^{32}$ P]-labelled RNA for 36 hours at 52°C. Filters were washed twice at room temperature for 10 minutes in 2 x SSC, 0.1% SDS followed by a wash twice at 65°C in 0.5 x SSC, 0.1% SDS for 2 hours. Hybridization of labelled DNA was determined by scintillation spectrometry using a toluene-based scintillant after exposure to X-ray film (Section 2.3.5).

(c) Analysis of Nuclease-Digested DNA Dot Blot Analysis

Preparation of GeneScreen filter membranes, with varying amounts of heat-denatured DNA per dot (1-5  $\mu$ g) were prepared as described (Section

2.3.16a). Pre-hybridization, hybridization and filter washing conditions are detailed in Section 2.3.17.

### 2.3.17 Southern Blot Hybridization

Analysis of DNA digested by restriction enzymes was carried out by electrophoresis through 0.8-1.0% agarose gels (Section 2.3.4a(i)) and transfer to nitrocellulose filter membranes (Schleicher and Schull, BA85) by the method of Southern (1975). Nitrocellulose filters were baked for 2-3 hours at 80°C in vacuo and pre-hybridized overnight at 42°C in 50% formamide, 5 x SSC, 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.5, 100 µg/ml denatured salmon sperm DNA and 1 x Denhardt's solution (0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin and 0.02% ficoll). Hybridization was at 42°C for approximately 18 hours under the same conditions except that 2 x SSC was included. Heat-denatured, [<sup>32</sup>P]-labelled nick-translated or oligo-labelled probes ( $\approx 1 \times 10^8$ - $10^9$  cpm/µg DNA) were added to hybridization solutions for probing dot-blot and Southern blots were prepared as described previously (Sections 2.3.10a,d). Recombinant DNA clones (gene inserts) used in the preparation of [<sup>32</sup>P]-labelled probes have been described (Section 2.2.4). All filters were washed twice in 2 x SSC, 0.5% SDS at room temperature for 15 minutes and then twice in 0.2 x SSC, 0.1% SDS at 65°C for 45 minutes.

An H5 gene-specific synthetic oligonucleotide (26 mer, 'H5 primer'; Section 2.2.6a) was 5' end-labelled with [ $\gamma$ -<sup>32</sup>P]ATP (Section 2.3.10e). This 26-mer is complimentary to the H5 gene sequence from -24 to +2 (with respect to the numbering of the gene by Krieg et al., 1983). Filters probed with [ $\gamma$ -<sup>32</sup>P]-kinased oligonucleotides were treated by a modification to the method of Noyes et al., (1979). Briefly, pre-hybridization was carried out at 42°C overnight in 6 x NET (0.9 M NaCl, 90 mM Tris-HCl pH 7.6, 6 mM EDTA), 0.5% NP40, 5 x Denhardt's solution and 100 µg/ml denatured salmon sperm DNA. Hybridization was carried out

under the same conditions except that 1 x Denhardt's solution was included in the mix. Following hybridization, filters were washed twice for 1 hour at 65°C in 4 x SSC and DNA:DNA hybrids detected by autoradiography (Section 2.3.5).

### 2.3.18 Transfection and Selection of Clonal Cell-Lines

Approximately 18 hours prior to transfection  $1.0-1.5 \cdot 10^6$  cells ( $\approx 10^5$  cells/ml) in exponential growth phase were seeded in 57 cm<sup>2</sup> petri dishes. Transfection of DNA into HeLa cells was by a modification to the calcium-phosphate precipitation procedure of Graham and van der Eb (1973). A 1.25 ml transfection cocktail containing 8.4 µg/ml of Eco RI-linearized 'test' DNA, 4 µg/ml Bam HI-linearized pSV2Neo (Section 2.2.4), 25 µg/ml sheared chicken DNA, 1 x HEPES buffered saline (137 mM NaCl, 5 mM KCl, 6 mM dextrose, 20 mM HEPES pH 6.92, 0.7 mM Na<sub>2</sub>HPO<sub>4</sub>) and 100 mM CaCl<sub>2</sub> pH 5.55 was made to volume with H<sub>2</sub>O, incubated at room temperature with occasional mixing and added to a 10 ml cell culture. Media was replaced after 18 hours, supplemented with 750 µg/ml G418 followed by colony selection for up to 14 days. During the selection period, media was replaced every 3 days each time carefully aspirating dead-sloughed cells from the dish. Discrete, G418-resistant colonies visible after the selection period were isolated from other colonies with glass cloning rings. Isolated colonies were transferred following trypsinization (Section 2.3.11) into 1.7 x 1.6 cm multi-well plates (Flow Laboratories) and after colony amplification to larger tissue culture vessels, G418-selected clonal HeLa cell-lines were maintained as described (Section 2.2.9) but under continued selective pressure.

### 2.3.19 Oligonucleotide Directed In Vitro Mutagenesis (Zoller and Smith, 1983)

DNA on which site-directed mutagenesis was to be performed was cloned into an M13 vector. The mutagenesis reaction, in a volume of



15  $\mu$ l, consisted of approximately 100 ng ssM13, 4 ng phosphorylated deletion primer (Section 2.2.6c), 2.5 ng phosphorylated universal sequencing primer (Section 2.2.6d), 200 mM NaCl, 100 mM Tris-HCl pH 8.0, 10 mM MgCl<sub>2</sub> and was heated at 65°C for 5 minutes before allowing annealing at room temperature for 10 minutes. Following heteroduplex formation, 32  $\mu$ l of extension mix (0.1 mM each of dTTP, dGTP, dCTP and dATP, 1.6 mM ATP, 1.6 mM DTT), 2 units of Klenow DNA polymerase and 1 unit of T4 DNA ligase was added. This reaction was left at room temperature for 4 hours and then various dilutions transformed into E. coli JM101 (Section 2.3.8).

Screening for mutant phage was performed by a modified procedure of the in situ plaque hybridization method of Benton & Davis (1977). Unwashed nitrocellulose discs were marked assymmetrically and placed onto the transformed plate for 2 minutes (until uniformly wet) and replaced by a second filter, which after 4 minutes was peeled off the plate. Filter discs were dried at room temperature (plaques facing up), sandwiched between two sheets of Whatman 3MM paper, and baked at 80°C in vacuo for 2 hours.

Filters were pre-hybridized in 50% formamide, 5 x SSC, 5 x Denhardt's solution, 50 mM sodium phosphate pH 6.5 and 10  $\mu$ g/ml sonicated salmon sperm DNA for 2 hours at 42°C. Approximately 100 ng of [<sup>32</sup>P]-kinased mutagenesis primer was added and incubation of filters continued for 16-18 hours at 42°C. Filters were briefly washed in 6 x SSC at room temperature and then in 3M tetramethyl ammonium chloride (Wood et al., 1985) at 65°C for at least one hour before being wrapped in gladwrap and autoradiographed overnight. 'Positive' plaques which appeared as intense spots on the autoradiogram were picked off the original plate, expanded and sequenced (Section 2.3.20) to confirm the mutagenesis.

2.3.20 DNA Sequencing by the Dideoxy-Chain Termination Method (Sanger et al., 1977)

(a) Sequencing Reactions

Four separate reactions, each specific for one of the bases in DNA, were used to analyze the insert sequence in the M13 single stranded template. In each of the sequencing reactions, the sequencing primer (Section 2.2.6c,d) was extended in the presence of a different ddNTP such that there was a partial incorporation of each, which resulted in termination of synthesis. The method described below is for the sequencing of one M13 clone but is readily expanded to allow the concurrent sequencing of 8 to 16 clones.

(b) Hybridization

2.5 ng of sequencing primer (Section 2.2.6c,d) was annealed to 3  $\mu$ l of M13 single-stranded template (prepared as described, Section 2.3.2b) in a 10  $\mu$ l volume containing 10 mM Tris-HCl pH 8.0, 50 mM NaCl and 10 mM MgCl<sub>2</sub> by heating the solution at 70°C for 3 minutes and then incubating at room temperature for 45 minutes.

(c) Polymerization

1  $\mu$ l of [ $\alpha$ -<sup>32</sup>P]dATP was lyophilized, the hybridization mix was added, vortex mixed to resuspend the labelled dATP, and then 1  $\mu$ l of 10 mM DTT added. 1.5  $\mu$ l of each appropriate zero mix (T° for ddTTP: 10  $\mu$ M dTTP, 200  $\mu$ M dCTP, 200  $\mu$ M dGTP, 5 mM Tris-HCl pH 8.0, 0.1 mM EDTA; C° for ddCTP: 200  $\mu$ M dTTP, 10  $\mu$ M dCTP, 200  $\mu$ M dGTP, 5 mM Tris-HCl pH 8.0, 0.1 mM EDTA; G° for ddGTP: 200  $\mu$ M dTTP, 200  $\mu$ M dCTP, 10  $\mu$ M dGTP, 5 mM Tris-HCl pH 8.0, 0.1 mM EDTA; A° for ddATP: 200  $\mu$ M of dTTP, dCTP and dGTP, 5 mM Tris-HCl pH 8.0, 0.1 mM EDTA) and ddNTP solutions (0.1 mM for ddCTP and ddATP, 0.5 mM for ddTTP and ddGTP, each in water) were added together. 2  $\mu$ l of the zero-ddNTP mixtures were added separately to four Eppendorf 'reaction tubes'.

0.5  $\mu$ l of DNA polymerase I Klenow fragment (1 unit/ $\mu$ l) was added to the hybridization mixture-label-DTT solution. 2  $\mu$ l of this was added to each of the four reaction tubes and the solutions were mixed by centrifugation for 1 minute. After 10 minutes incubation at 37°C, 1  $\mu$ l of dATP chase (500  $\mu$ M dATP in 5 mM Tris-HCl pH 8.0, 0.1 mM EDTA) was added to each of the four tubes, mixed by a 1 minute centrifugation and incubated for a further 15 minutes at 37°C.

3  $\mu$ l of formamide loading buffer was added to stop the reactions and the solutions were mixed by a brief centrifugation. Samples were heated at 100°C for 3 minutes and then loaded onto a sequencing gel (Section 2.3.4b(iii)).

#### 2.3.21 Nuclear Extract Preparation

Cells were washed in PBS (Section 2.3.11) and re-suspended in lysis buffer (10 mM Hepes pH 8.0, 50 mM NaCl, 0.5 M sucrose, 0.1 mM EDTA, 0.5% Triton X-100, 1 mM DTT and 5 mM MgCl<sub>2</sub>) at  $4 \cdot 10^7$  cells/ml. Nuclei were pelleted by centrifugation of the cell lysate at 500 x g, re-suspended at  $5 \cdot 10^7$  nuclei/ml in lysis buffer followed by addition of spermidine and NaCl to 5 mM and 0.5 M, respectively. The suspension was incubated on ice for 30-60 minutes and centrifuged at 4°C for 10 minutes in a microcentrifuge ( $\approx 15000$  x g). Protein in supernatant fractions was precipitated in ammonium sulphate (0.33 g/ml), re-suspended in 1 ml Tris-Magnesium (TM) buffer (50 mM Tris-HCl pH 7.9, 12.5 mM MgCl<sub>2</sub>, 1 mM EDTA, 1 mM DTT, 20% glycerol) plus 0.1 M KCl, and dialyzed extensively against the same buffer at 4°C in dialysis bags. Protein concentration in nuclear extracts was routinely determined by the method of Bradford (1976) using bovine serum albumin as a standard.

#### 2.3.22 Chromatographic Enrichment of Nuclear Extract

##### Components

##### (a) Gel Filtration

Crude nuclear extract (Section 2.3.21; 500-750 mg in 1 ml) was

loaded on a 480 ml (2.8 cm diameter) Sephacryl S-300 (Pharmacia) gel filtration column, pre-equilibrated in TM buffer (Section 2.3.21) plus 0.1 M KCl, and calibrated using standard molecular weight protein markers. Gel filtration was according to Pharmacia's specifications at 4°C with a flow rate of 1 ml/minute and TM plus 0.1 M KCl as the chromatography buffer. Specific nuclear factor components in fractionated extract samples were detected using a DNA-binding assay (gel mobility shift assay, Section 2.3.23a). Those fractions displaying specific binding activity were pooled for further analysis.

(b) Ion-Exchange Chromatography

Further enrichment of specific nuclear factors (as detected by the DNA binding assay, Section 2.3.23a) was achieved by applying pooled fractions (see Section 2.3.22a) to a 10 ml DEAE-sepharose CL-6B (Pharmacia) ion-exchange column. Chromatography was according to Pharmacia's recommendations at 4°C with a flow rate of 1 ml/minute using TM plus 0.1 M KCl as the application buffer. After loading the sample ( $\approx$  5 ml), 25 x 1 ml fractions were collected after the initial void column volume had flowed through. Protein was eluted from the column with TM buffer plus 1 M KCl. Flow-through and column-bound fractions were assayed for specific binding activity (Section 2.3.23a) and pooled where appropriate.

(c) Affinity Chromatography

DNA-protein affinity chromatography was essentially by the method of Kadonaga and Tjian (1986) except that synthetic 40-mer H1-box oligonucleotides (Section 2.2.6b) were ligated together for coupling to the Sepharose CL-2B gel matrix. Binding and elution buffer consisted of 25 mM HEPES pH 7.5, 1 mM EDTA, 5 mM DTT, 10% glycerol, 0.5 mM PMSF plus 0.15 mM or 1 M NaCl, respectively.

CHAPTER 3

ANALYSIS OF H5 AND H2A<sub>F</sub> GENE EXPRESSION

## CHAPTER 3 : ANALYSIS OF H5 AND H2A<sub>F</sub> GENE EXPRESSION

### 3.1 INTRODUCTION

Histone protein synthesis during S-phase of the cell-cycle is chiefly regulated by the availability of corresponding mRNAs (see Sections 1.11.3 and 1.11.4). This is reflected by the well documented coupling between DNA synthesis and accumulation of histone mRNAs in the cytoplasm. In this chapter, the parameters involved in the cell-cycle regulation of major core and H1 genes were investigated in addition to those encoding for the histone variants H5 and H2A<sub>F</sub>. This essentially entailed analysis of gene transcription and steady-state mRNA levels throughout the cell-cycle. These experiments were specifically performed to assign the H2A<sub>F</sub> gene to the replication-dependent/replication-independent (basal) category and, to identify features of H5 gene expression which accounts for continued synthesis of its gene product during erythroid differentiation. Much of this work has been previously published (Dalton et al., 1986a) or has been submitted for publication (Robins et al., submitted for publication).

### 3.2 RESULTS

#### 3.2.1 Histone Gene Expression in an Early-Erythroid Chicken Cell-Line

The AEV ts34 cell-line used in these studies consisted of early chicken erythroblasts transformed by an avian erythroblastosis virus (AEV-RAV2, Section 2.2.9). Specifically, the HD2 subline was chosen because it is incapable of differentiating to mature erythrocytes at elevated temperatures (42°C), even though other AEV ts34 clonal-lines can be induced to do so under special culture conditions (Beug et al., 1982). Extensive studies of cell-specific markers such as surface antigens, heme, hemoglobin, carbonic anhydrase and histone H5 have characterized the clonal cell-line as early-erythroid. This was further

vindicated by the exceptionally low levels of  $\alpha$ -globin mRNA detected by Northern-blot analysis in AEV ts34 cytoplasmic RNA extracts compared to that in mature erythrocytes (Figure 3.1b). Furthermore, a comparison between the profiles of chromatin-extracted H1 and H5 protein derived from these two sources was made by gel electrophoresis (Figure 3.1a). Protein loads were selected so that the amount of H1 in each sample was approximately equal, thereby allowing comparisons to be made. Clearly, the ratio of H5 to H1 in erythrocytes was substantially greater than in dividing AEV ts34 cells. The factors which contribute to elevated H5 levels in erythrocytes are infact likely to be associated with the rate of H5 gene transcription and stability of its mRNA. Results presented in this chapter address these factors.

Authentic expression of H5 and H2A<sub>F</sub> genes in AEV ts34 cells was verified by confirming that both transcripts were polyadenylated and expressed in an appropriate manner with respect to parameters such as cell-type specificity. Cytoplasmic RNA was separated into poly(A)<sup>+</sup>/poly(A)<sup>-</sup> fractions by oligo dT cellulose chromatography (Section 2.3.12b), electrophoresed through an agarose slab gel, blotted onto a filter membrane and hybridized to a [<sup>32</sup>P]-labelled homologous chicken histone gene probe (Northern blot analysis, see sections 2.2.4; 2.3.15a). Both the 930 base-H5 and 820 base-H2A<sub>F</sub> transcripts were significantly enriched in the poly(A)<sup>+</sup> fraction (Figure 3.2a), consistent with previous findings (Molgaard et al., 1980; Krieg et al., 1982). Major core (represented by H2A mRNA), H1 (data not shown) and  $\beta$ -actin transcripts were specifically enriched in the poly(A)<sup>-</sup> fraction.

The cell-type specificity of H5 and H2A<sub>F</sub> gene expression in cultured cell-lines was determined by probing filter-immobilized cytoplasmic RNA extracted from three chicken cell types; AEV ts34 cells (early erythroid), a Marek-virus transformed T-lymphoid cell-line and a primary fibroblast culture (Section 2.2.9). Core histone transcripts

### FIGURE 3.1

#### CHARACTERIZATION OF AEV ts34 CELLS AS EARLY-ERYTHROID

- A. Analysis of H5 (lower band) and H1 (upper bands) levels in dividing AEV ts34 cells and in mature erythrocytes by electrophoresis through a Triton X-100-acetic acid-urea-polyacrylamide gel (Section 2.3.4b(iv)). Levels of protein loaded were chosen such that the amount of H1 histone in each track was equivalent. H5 is barely detectable in the preparation from AEV ts34 cells.
- B. Cytoplasmic RNA (10  $\mu$ g) from AEV ts34 cells and chicken erythrocytes (gift from Dr. A.J. Robins) was electrophoresed through a 1.5% agarose gel and analyzed by Northern blot hybridization (Section 2.3.15a) using a [ $^{32}$ P]-labelled nick-translated genomic chicken  $\alpha$ -globin probe (Section 2.2.4). Transcript lengths were confirmed by comparison against HinF I-digested pBR322 markers.



A.

Pre-Erythroblast  
(AEV ts34)

Erythrocyte



H1

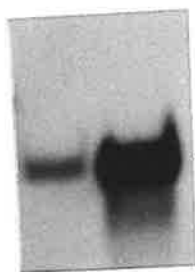


H5



B.

AEV ts34  
Erythrocyte



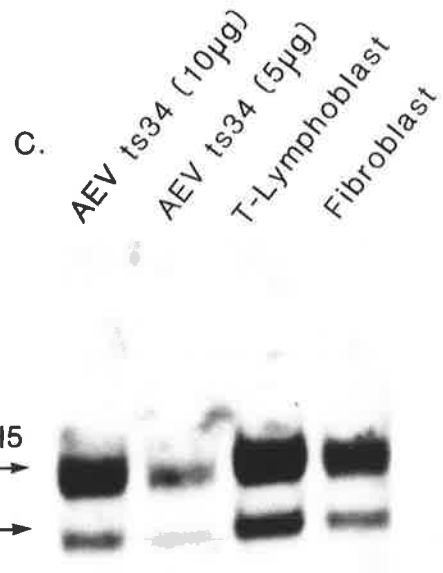
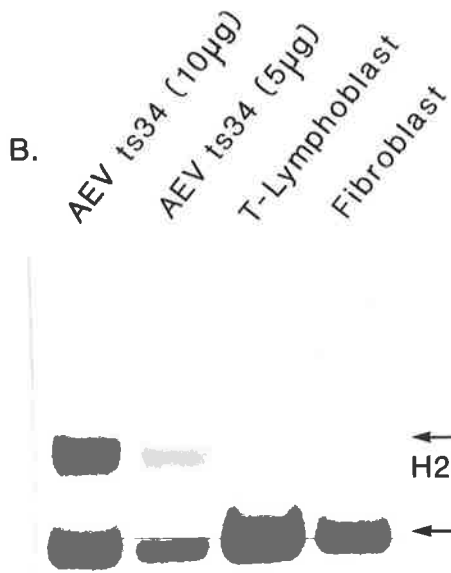
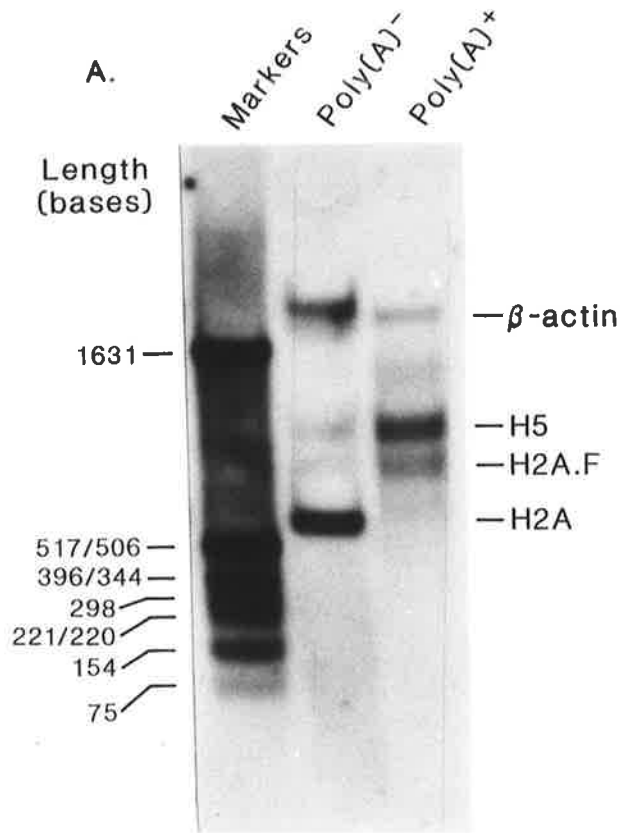
Length

— 680b

### FIGURE 3.2

#### POLYADENYLATION AND TISSUE-DISTRIBUTION OF H5 AND H2A<sub>F</sub> TRANSCRIPTS

- A. Cytoplasmic RNA from AEV ts34 cells was separated into poly(A)<sup>+</sup> and poly(A)<sup>-</sup> fractions by oligo dT chromatography (Section 2.3.12c). 250 ng and 10 µg of RNA from poly(A)<sup>+</sup> and poly(A)<sup>-</sup> fractions, respectively, was electrophoresed through a 1.5% agarose gel, blotted onto a nylon membrane (Section 2.3.15a) and probed with gene-specific nick-translated [<sup>32</sup>P]-cDNA or genomic histone gene fragments (Section 2.2.4). Bands representing β-actin, H5, H2A<sub>F</sub> and H2A are shown relative to HinF I-digested pBR322 markers.
- B. Northern analysis (Section 2.3.15a) was performed, as described in the legend to Figure 3.2a, on total cytoplasmic RNA (10 µg, unless otherwise indicated) extracted from AEV ts34 cells, a chicken T-lymphoblast cell-line and a chicken fibroblast primary culture (Section 2.2.9) using [<sup>32</sup>P] oligo-labelled H2B and H5 histone gene fragments (Section 2.2.4). Bands corresponding to H5 and H2B transcripts are indicated.
- C. After removing hybridized [<sup>32</sup>P]-labelled probe, the same filter as in Figure 3.2b was reprobed with H2B and H2A<sub>F</sub> [<sup>32</sup>P] oligo-labelled histone gene fragments, respectively (Section 2.2.4). Bands corresponding to each specific transcript are indicated.



(in this case represented by H2B), were ubiquitous to all three cell-types (Figure 3.2b). H5 transcripts were detected exclusively in the AEV ts34 early-erythroid cell-line (Figure 3.2b) whereas an 820b-H2A<sub>F</sub> transcript was present in extracts from all three sources (Figure 3.2c).

An important feature of histone gene expression is its reported coupling to DNA replication (see Section 1.11.3; 1.11.4). The regulated expression of major core and H1 chicken histone genes during the cell-cycle was therefore investigated with special attention focussing on the highly variant H5 and H2A<sub>F</sub> genes. This entailed obtaining homologous populations of AEV ts34 cells with respect to cell synchronization.

### 3.2.2 AEV ts34 Cell Synchronization with Aphidicolin

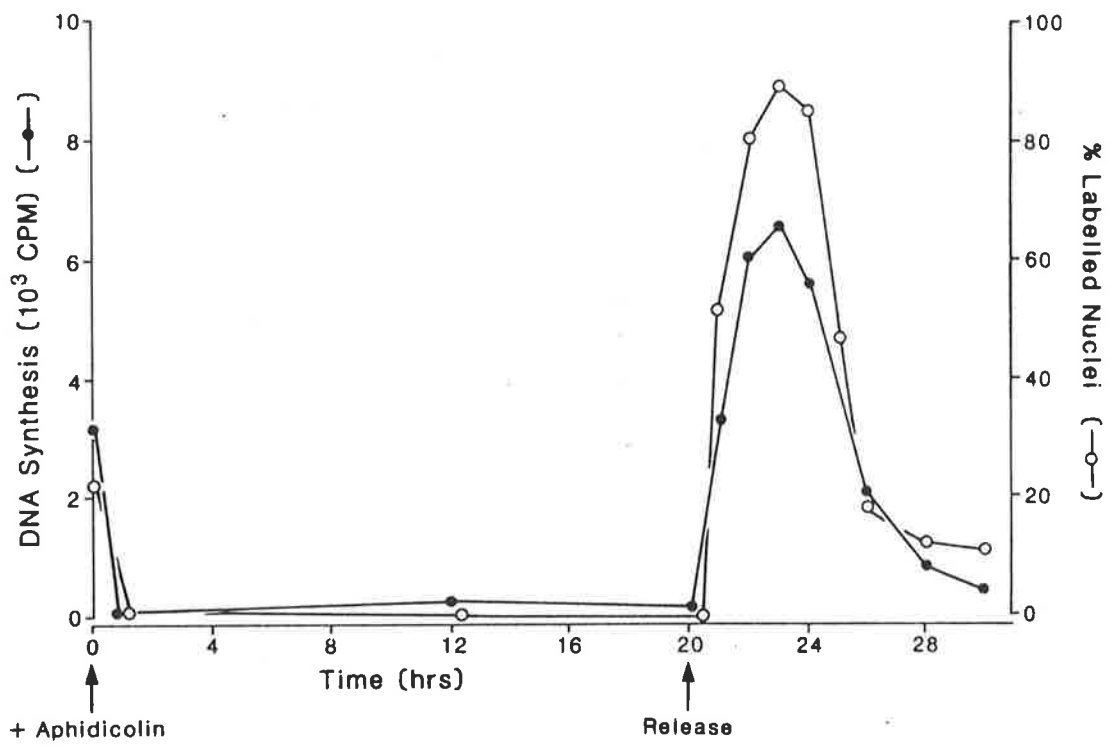
Initial attempts to synchronize AEV ts34 cells with hydroxyurea and thymidine or, by isoleucine starvation, were unsuccessful. The successful application of the antibiotic aphidicolin in synchronizing HeLa cells (Heintz et al., 1983; Pedrali-Noy et al., 1980) suggested this approach for AEV ts34 cells. This fungal-derived antibiotic inhibits DNA replication by reversibly binding to DNA polymerase  $\alpha$  (reviewed by Huberman, 1981), thereby inhibiting replication and facilitating cells to accumulate at the G1/S cell-cycle border.

After addition of aphidicolin (5  $\mu$ g/ml) to unsynchronized AEV ts34 cells in exponential growth phase, incorporation of [<sup>3</sup>H]thymidine into DNA (trichloroacetic acid insoluble material) by pulse-labelling rapidly declined to a minimal rate within 1 hour (Figure 3.3). Throughout the period of aphidicolin exposure (20 hours), the basal rate of DNA synthesis remained very low and relatively constant. The effects of the antibiotic on DNA synthesis were shown to be clearly reversible after the washing procedure (see Section 2.3.11) by a sharp increase in [<sup>3</sup>H]thymidine incorporation. Pulse-labelling data indicates that the S-phase proceeded for 6 to 7 hours after which the rate of DNA synthesis returned to approximately the pre-release rate. In addition, the

### FIGURE 3.3

#### KINETICS OF DNA SYNTHESIS IN AEV ts34 CELLS

Where indicated, aphidicolin (5  $\mu\text{g/ml}$ ) was added to randomly growing cultures of cells in exponential phase and washed in fresh media after 20 hours (Section 2.3.11a). Cells were pulse-labelled with [ $^3\text{H}$ ]thymidine for 15 minutes. The rate of radioactivity incorporation into trichloroacetic acid-insoluble material, indicating the rate of DNA synthesis, and the percentage of labelled nuclei, indicating the percentage of cells in S-phase, are shown (see Section 2.3.11b).



increased rate of DNA replication was paralleled by an elevation in the percentage of [<sup>3</sup>H]thymidine-labelled nuclei, indicating that a high degree of cell synchronization was achieved. With a single application and subsequent release of AEV ts34 cells from the aphidicolin block, in excess of 90% entered S-phase in synchrony (Figure 3.3) indicating that the cell population was essentially homogenous, with respect to synchronization and, therefore suitable for cell-cycle studies.

### 3.2.3 Cell-Cycle Regulation of Chicken Histone Gene Expression

To monitor the steady-state levels of specific histone mRNA species during the cell-cycle, cytoplasmic RNA was prepared from cells before, during and after treatment with aphidicolin. RNA (10 µg) was resolved electrophoretically and analysis of individual histone mRNAs was performed by Northern blot hybridization. In each experiment, full-length nick-translated or oligo-labelled histone gene (cDNA or genomic) fragments spanning all or most of the coding region were used (see Section 2.2.4). Amounts of histone transcripts were quantitated by densitometric scanning (see Section 2.3.5) and results expressed as relative increases in hybridization signals over pre-release (20 hours) levels (Figures 3.4, 3.5).

It is apparent that steady-state levels of major core (H2A, H2B, H3, H4) and H1 histone mRNAs dramatically increased as the cells progressed through S-phase (Figures 3.4, 3.5). Levels of these transcripts generally plateau at approximately 3 to 4 hours post-release. Accumulation of H2A mRNA during this period is typical of the other histone mRNA sub-species. During mid-S-phase, there was approximately 15-fold more H2A mRNA present than prior to release from the cell-cycle block or, after completion of S-phase (Figure 3.4). Following the peak in steady-state transcript accumulation, levels declined sharply in parallel with the rate of DNA replication, as monitored by the kinetics of [<sup>3</sup>H]thymidine incorporation (see Figure

3.3), reaching a minimum level at approximately 6 to 7 hours post-release. Increases in the steady-state levels of H3 (Figure 3.4) and H1, H2B, H4 (Figure 3.5) transcripts during S-phase had a similar kinetic profile to that observed for H2A mRNA, indicating that levels of mRNAs transcribed from the major chicken histone gene types are generally co-ordinately coupled to DNA replication.

Identical RNA samples were probed to detect transcripts encoding the variant H5 and H2A<sub>F</sub> histones (Figures 3.4; 3.5). In contrast to the major core and H1 mRNAs, levels of H5 and H2A<sub>F</sub> transcripts were essentially invariant at all cell-cycle time points and certainly displayed no coupling to DNA replication. These results illustrate the striking difference between the general pattern of histone gene S-phase regulation (eg. H1, H2A, H2B, H3, H4) and the distinctive absence of linkage between DNA replication and steady-state levels of the two highly variant histone mRNAs examined. When synchronized AEV ts34 cells were allowed to pass through another cell-cycle after aphidicolin release, levels of both transcripts remained constant in the secondary G<sub>1</sub>-, S- and G<sub>2</sub>-phases; whereas H2A mRNA levels transiently increased exclusively during S-phase (data not shown). Therefore, H5 and H2A<sub>F</sub> gene expression, as measured here, is independent of DNA synthesis (and probably other phases of the cell-cycle).

To test the possibility that H5 and H2A<sub>F</sub> transcript levels observed were not due to dividing cells, but a small population of spontaneously differentiating cells, in situ hybridization (see Section 2.3.15c) was performed on paraffin-embedded AEV ts34 cell sections. For this part of the study, the assistance of Alison Jilbert (Institute of Medical and Veterinary Science, Adelaide, South Australia) is gratefully acknowledged. Greater than 95% of cell nuclei were specifically and intensely labelled by hybridization with a H5 [<sup>32</sup>P]-cDNA insert (Section 2.2.4; labelled by second-strand synthesis of an M13 clone, Section

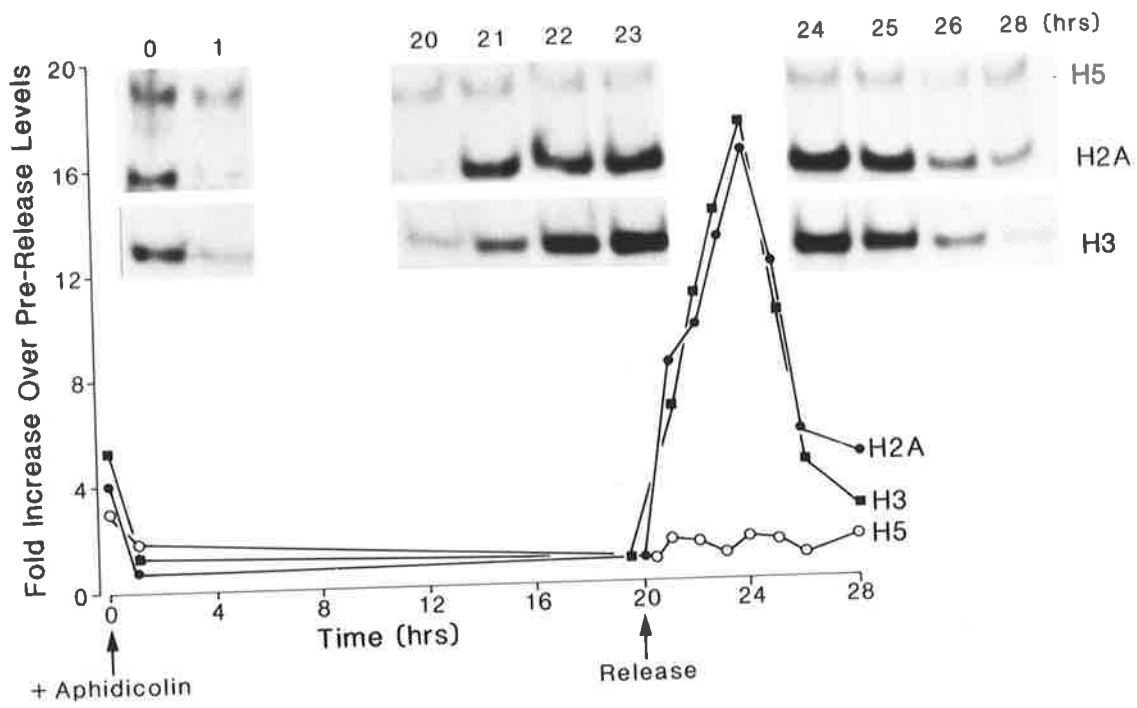


## FIGURE 3.4

### HISTONE mRNA LEVELS DURING THE CELL-CYCLE

Steady-state levels of individual histone mRNAs during the cell-cycle are shown. Cytoplasmic RNA samples (10  $\mu$ g) prepared at the times indicated were resolved on 1.5% agarose gels after denaturation and levels of different histone mRNAs were detected by Northern blot hybridization (Section 2.3.15a) with specific gene probes (Section 2.2.4). Densitometric quantitation of the autoradiograms shown in each panel is also represented.

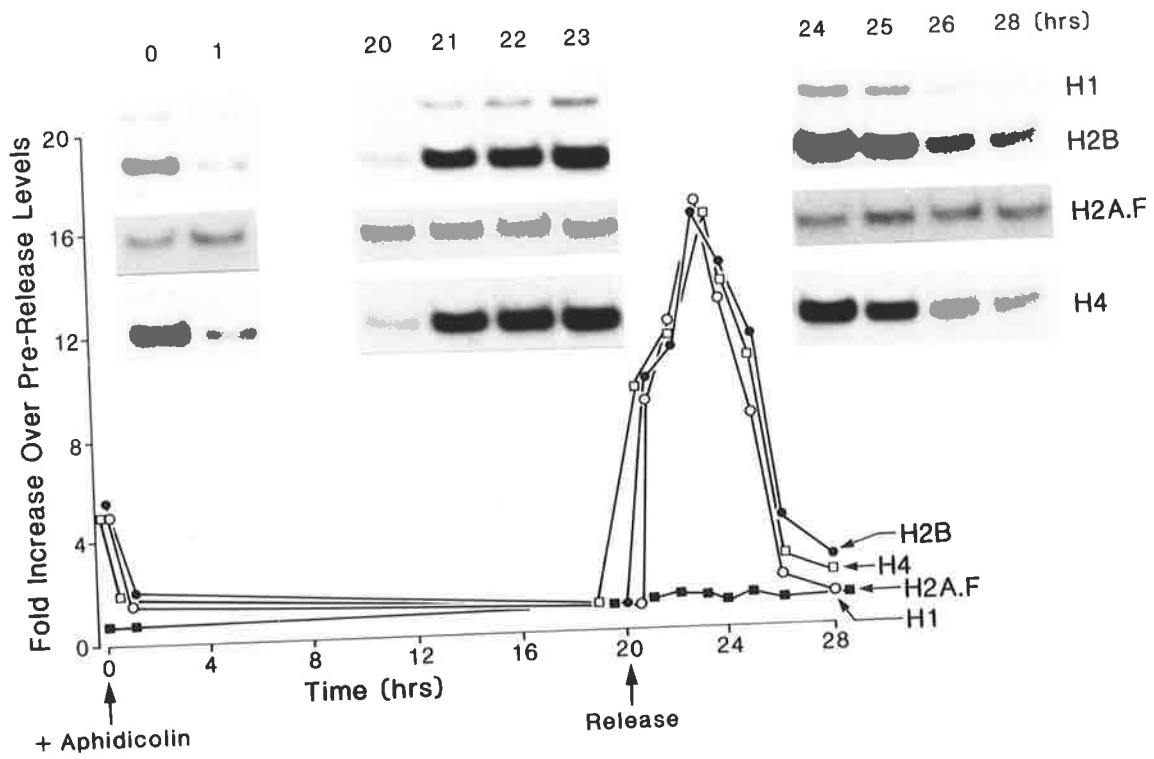
Symbols: ● , H2A mRNA; ○ , H5 mRNA; ■ , H3 mRNA. Graphical data is representative of the mean values from two independent experiments.



### FIGURE 3.5

#### HISTONE mRNA LEVELS DURING THE CELL-CYCLE

Steady-state levels of individual histone mRNAs during the cell-cycle are shown (see legend to Figure 3.4). Symbols:  $\circ$  , H1 mRNA;  $\bullet$  , H2B mRNA;  $\square$  , H4 mRNA;  $\blacksquare$  , H2A<sub>F</sub> mRNA. Graphical data is representative of the mean values from two independent experiments.



2.3.10c) complimentary to H5 mRNA (Table 3.1). A complimentary H5 coding fragment, which would not anneal to H5 mRNA, was used as an indicator of non-specific background hybridization. Similar observations apply to the widespread presence of H2A<sub>F</sub> mRNA in these cell sections (see Table 3.1). As a further control, chicken  $\beta$ -globin and feather keratin probes failed to detect significant levels of corresponding transcripts and were only marginally above background levels in intensity. Thus, the invariance of H2A<sub>F</sub> and H5 mRNA levels throughout the cell-cycle is accounted for by the accumulation of transcripts in the majority of dividing cells as opposed to a small (possible) subset of differentiating AEV ts34 cells. These results are represented by cell-sections in which H5 mRNA was detected in about 93% of the cell population (Table 3.1, inset C), in addition to positive ( $[^3\text{H}]$ poly(dT) probe; Table 3.1, inset B) and negative (no probe - Table 3.1, inset A; - H5 probe, Table 3.1 inset D) control experiments.

#### 3.2.4 Transcription of H5 and H2A.F Genes

Although results presented thus far indicate that steady-state levels of H5 and H2A<sub>F</sub> mRNA do not fluctuate throughout the cell-cycle, the possibility remains that both genes may be transcriptionally activated by a stage-specific mechanism. This would also mean that some other level of control would need to operate (ie. differential transcript stability or translation) so as to account for equivalent H5 and H2A<sub>F</sub> levels already reported (Section 3.2.3). To test this possibility, isolated nuclei were prepared from randomly growing, aphidicolin-treated and aphidicolin-released (non-S-phase and S-phase) AEV ts34 cells and pulse-labelled for 15 minutes with  $[^{32}\text{P}]$ UTP. During this period, the rate of  $[^{32}\text{P}]$ label incorporation into the cellular RNA pool was linear and remained so for at least a further 20 minutes (data not shown). Labelled, nascent nuclear transcripts were used to probe excess quantities of filter-immobilized ssM13 with a plus (non-coding strand, complimentary to RNA) or minus (coding strand, homologous to

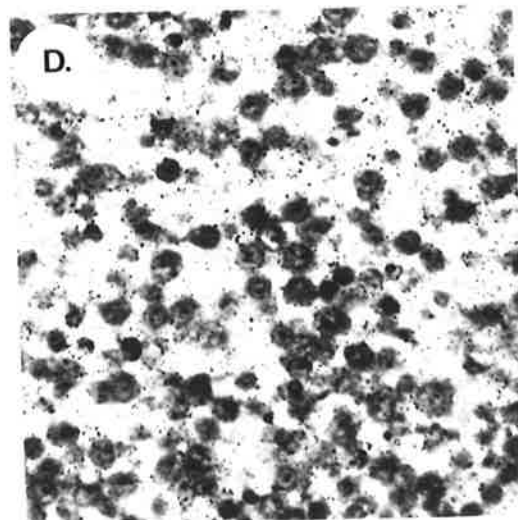
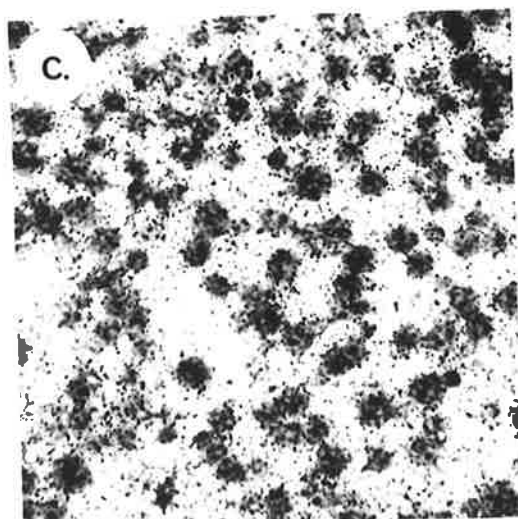
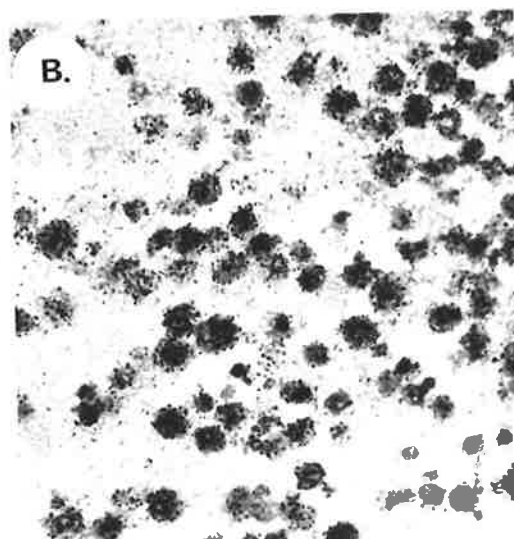
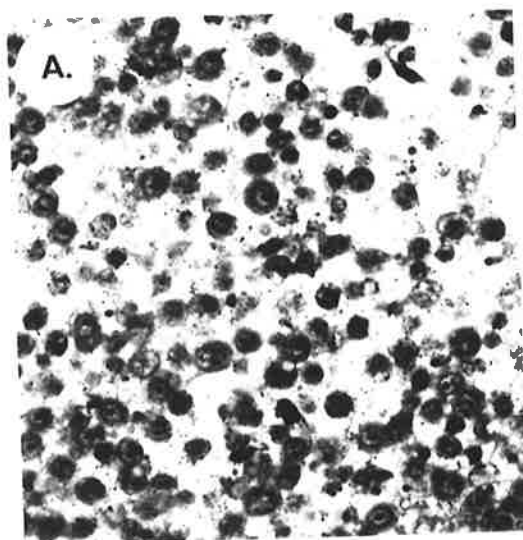
TABLE 3.1

DETECTION OF HISTONE TRANSCRIPTS BY IN SITU HYBRIDIZATION

AEV ts34 cells were fixed, sectioned and prepared for in situ hybridization as described in Section 2.3.15c. Probes were synthesized by the M13 second strand synthesis method (Section 2.3.10c) and separated away from vector DNA on polyacrylamide gels, following restriction enzyme digestion. Data are expressed as the combined results of two independent experiments, each carried out on duplicate cell-sections. Labelling intensity varied considerably, even when probes of equal specific activity were used (i.e.  $10^9$  cpm/ $\mu$ g DNA). This is represented empirically ( $\checkmark$ ). In each section >95% of cells were positively labelled with a [ $^3$ H]-labelled poly(dT) probe (Jilbert et al., 1986) which was an indicator of general poly(A)<sup>+</sup> mRNA content in each case. (+) designates the labelled strand which hybridizes to mRNA, while (-) represents the complimentary labelled probe (see Section 2.2.4). Cell sections, probed as indicated, were stained with eosin and hematoxylin as described by Jilbert et al., (1986). FK = feather keratin.

- A. Unprobed cell section.
- B. [ $^3$ H]-labelled poly(dT) probe.
- C. [ $^{32}$ P]H5 probe (+).
- D. [ $^{32}$ P]H5 probe (-).

	% Labelled Nuclei	Labelling Intensity
H5 (+)	93	✓✓✓✓
H5 (-)	5	✓
H2A <sub>F</sub> (+)	95	✓✓✓✓
H2A <sub>F</sub> (-)	3	✓
FK (+)	3	✓
FK (-)	5	✓
β-Globin (+)	9	✓ <sup>1/2</sup>
β-Globin (-)	4	✓
M13 (no insert)	4	✓



RNA) H5, H2A<sub>F</sub>, H2A, H2B or H1 histone gene insert. As enhanced rates of transcription during S-phase makes a major contribution to the accumulation of major core and H1 histone mRNAs (Graves and Marzluff, 1984; Heintz et al., 1983; Sittman et al., 1983), it was of particular interest to see if the same was true for the control of H5 and H2A<sub>F</sub> genes.

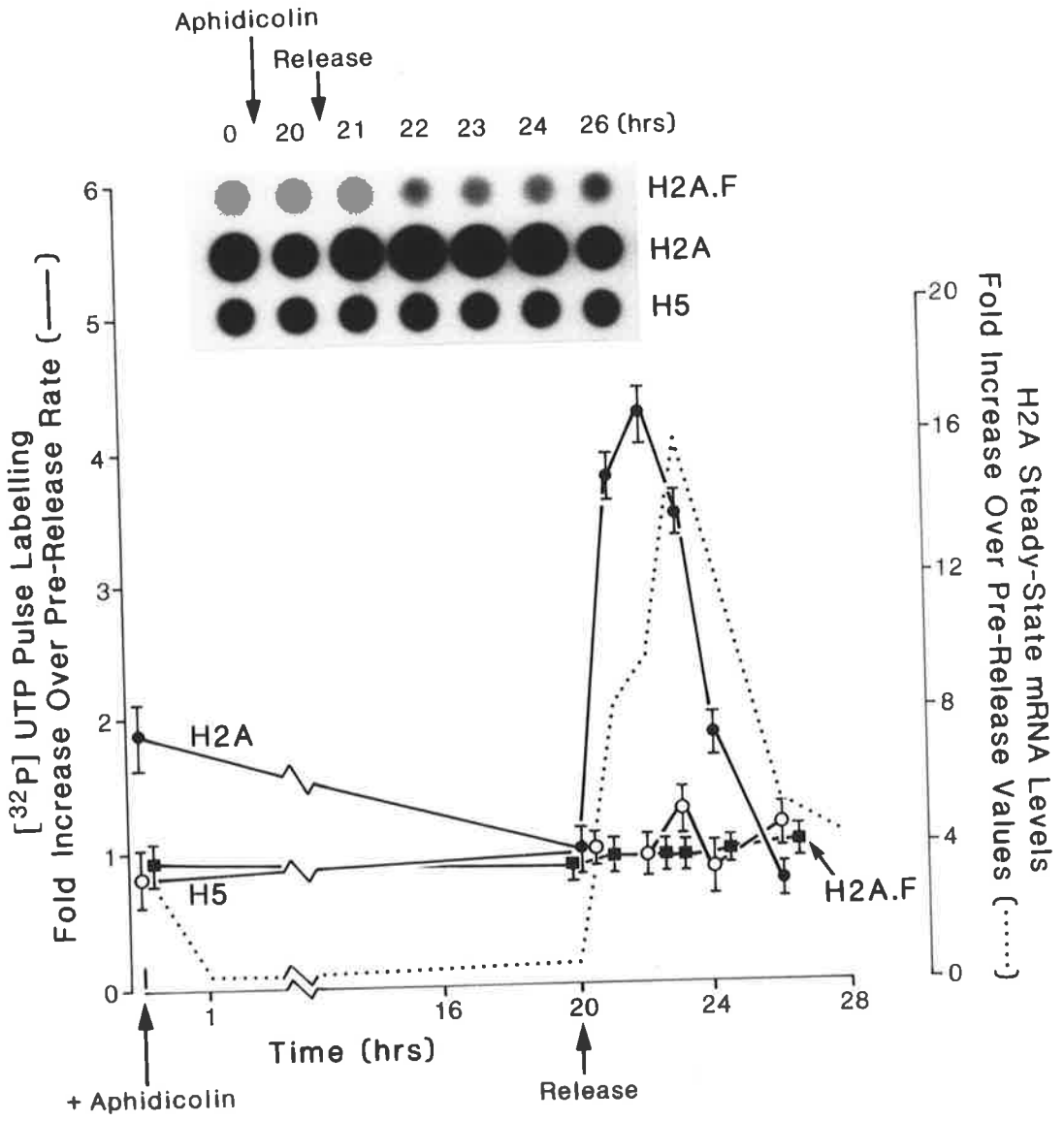
Dot-blot hybridization of labelled nuclear RNA to filter-bound ssM13 constructs was detected by autoradiography and quantitated by scintillation spectrometry of individual dots. From the autoradiograph, (Figure 3.6; 3.7), a comparatively low, but nevertheless obvious, level of H1, H2A and H2B transcription is evident during the period of aphidicolin exposure. After reversing the inhibition of DNA replication, a sharp increase in levels of nascent transcripts was detected. Maximum rates of H1, H2A and H2B transcription occurred between 1 and 2 hours post-release and declined thereafter for the duration of the cell-cycle monitored. Significantly, the maximal rate of transcription occurred before the peak accumulation of their respective steady-state mRNAs in the cytoplasm (Figure 3.6; 3.7). This is consistent with previous observations involving transcription of major histone genes in HeLa cells (Heintz et al., 1983), mouse lymphoma cells (Sittman et al., 1983) and hamster fibroblasts (Artishevsky et al., 1984). Non-specific hybridization (background) was determined by probing ssM13 with inserts non-complimentary to histone mRNAs and was subtracted to give specific hybridization values (data not shown). Analysis of hybridization between [<sup>32</sup>P]-labelled RNA and ssM13-H5 and -H2A<sub>F</sub> gene constructs indicated that transcription of these genes was invariant throughout the cell-cycle (Figure 3.6). This pattern of constitutive transcription contrasts that of H1, H2A and H2B genes and indicates that H2A<sub>F</sub> and H5 genes are not transcriptionally activated in a temporal manner during S-phase of the cell-cycle.



## FIGURE 3.6

### TRANSCRIPTIONAL ACTIVITY OF HISTONE GENES THROUGHOUT THE CELL-CYCLE

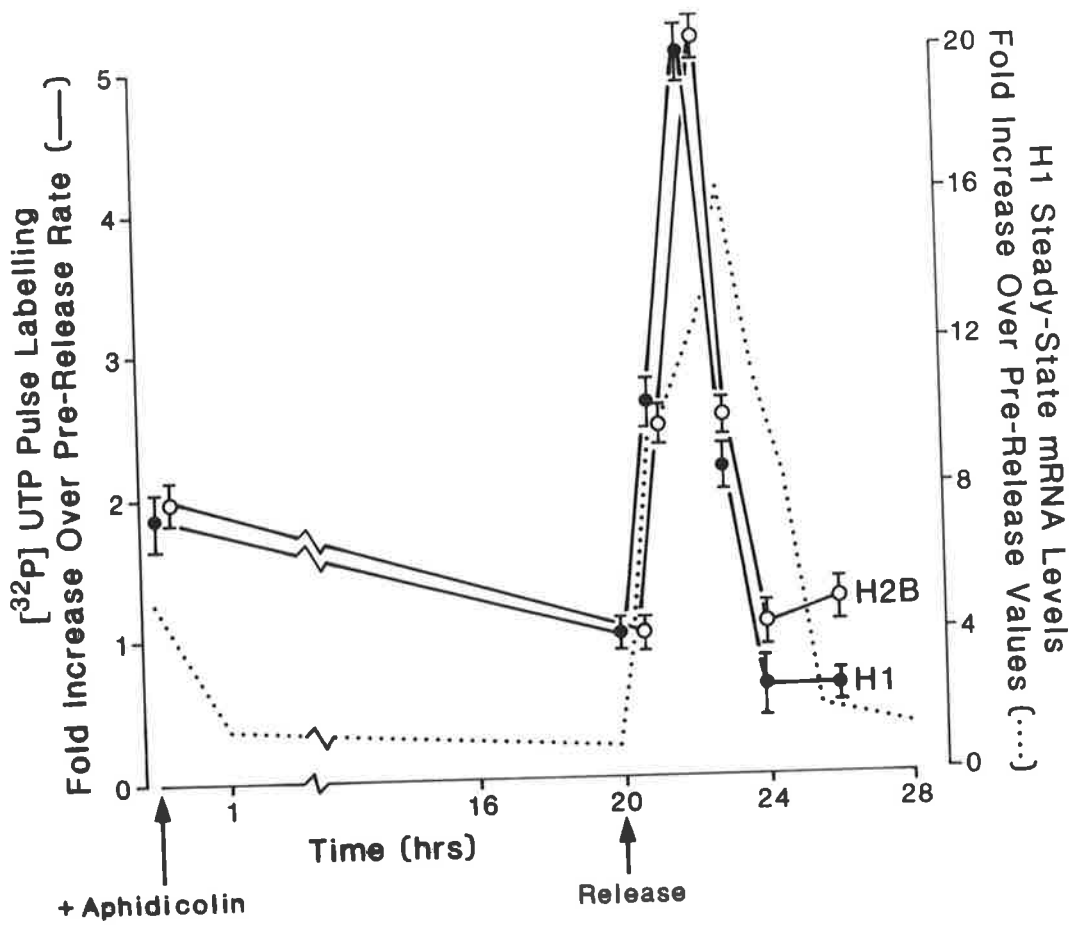
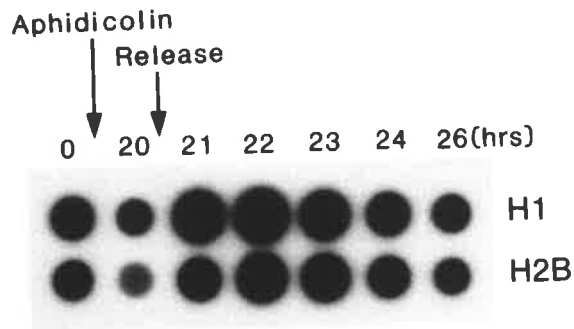
Detection of nascent H2A, H5 and H2A<sub>F</sub> mRNA transcripts from pulse-labelled nuclei. Nuclei were isolated from unsynchronized, aphidicolin-blocked and aphidicolin-released AEV ts34 cells and incubated under the conditions described (Section 2.3.16b). Labelled RNA obtained from nuclei isolated at various stages of the cell-cycle was used to probe filter-immobilized DNA dots at the times indicated (see Section 2.3.16b). Hybridization of RNA to ssM13 DNA dots with H5, H2A or H2A<sub>F</sub> full-length gene inserts (see Section 2.2.4) was determined by autoradiography and quantitated by scintillation spectrometry after levels of non-specific hybridization with non-complimentary ssM13/insert DNA had been subtracted (see text for description, and Section 3.2.4). Each bar represents  $\pm$  the standard error of the mean from three separate experiments. Each hybridization was carried out in duplicate. Symbols: ●, H2A mRNA; ○, H5 mRNA; ■, H2A<sub>F</sub> mRNA. The broken line represents the profile of steady-state H2A mRNA levels (see Figure 3.4) plotted against the right-hand-side vertical axis.



### FIGURE 3.7

#### TRANSCRIPTIONAL ACTIVITY OF HISTONE GENES THROUGHOUT THE CELL-CYCLE

Detection of nascent H1 and H2B transcripts from pulse-labelled nuclei. Experimental conditions were as described (see legend to Figure 3.6) except that ssM13 constructs with H1 and H2B histone gene inserts (see Section 2.2.4) were filter immobilized for probing with [<sup>32</sup>P]-labelled RNA. Symbols: ●, H1 mRNA; ○, H2B mRNA. The broken line represents the profile of steady-state H1 mRNA levels (see Figure 3.5) plotted against the right-hand-side vertical axis. Each bar represents  $\pm$  the standard error of the mean from three separate experiments.



### 3.2.5 5'-End Mapping of the H2A<sub>F</sub> Transcript

To determine the 5'-terminal end of chicken H2A<sub>F</sub> mRNA, primer extension analysis (Section 2.3.15b) was carried out on poly(A)<sup>+</sup> cytoplasmic RNA isolated from AEV ts34 cells with a H2A<sub>F</sub>-specific synthetic oligonucleotide (28-mer; Section 2.2.6). This yielded an extension product of 144 bases (Figure 3.8) corresponding to the designated 'cap' site located at an A residue, 124 bp upstream from the 'ATG' codon (Figure 3.9).

### 3.3 DISCUSSION

Work presented in this chapter has reinforced the concept that histone gene expression is tightly linked to DNA replication in dividing cells (see Section 1.11.3) by analyzing the periodic accumulation of mRNAs and rates of gene transcription throughout the cell-cycle. This analysis has been extended by focussing on parameters governing the expression of two genes encoding the extreme histone variants, H5 and H2A<sub>F</sub>.

Cytoplasmic concentrations of steady-state histone mRNAs during the AEV ts34 S-phase increased by approximately 15-fold. A general trend emerged suggesting that accumulation of each major core and H1 histone mRNA was stoichiometrically in proportion to the rate of DNA synthesis (Figures 3.3, 3.4 and 3.5). Cells not actively engaged in DNA replication (G1/S border, G<sub>1</sub>, G<sub>2</sub>) all contain low, but detectable, levels of H1, H2A, H3 and H4 mRNA. There are several possible explanations for this phenomenon, including an incomplete block of DNA synthesis in the presence of aphidicolin, which is indicated by a small percentage of labelled nuclei at the G1/S cell-cycle border (20 hours + aphidicolin; see Figure 3.3). Apart from this, it is likely that there is a sub-population of replication-independent chicken histone gene (major core and H1) subtype variants which do not respond to transient

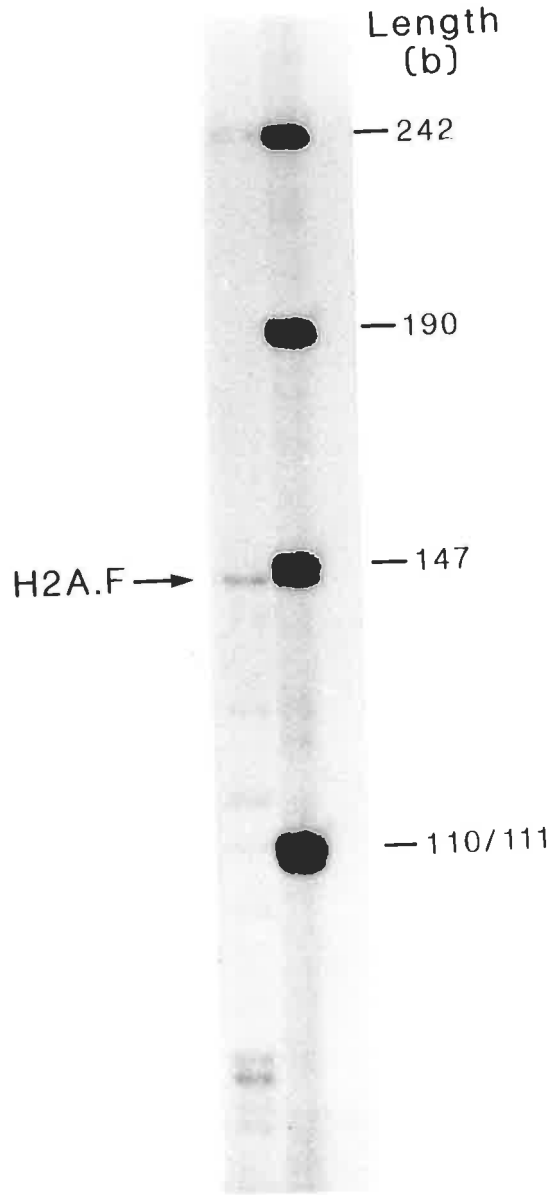
## FIGURE 3.8

### 5'-END MAPPING OF H2A<sub>F</sub> TRANSCRIPTS IN

#### AEV ts34 CELLS

5'-end mapping of the H2A<sub>F</sub> transcript. Primer extension analysis (Section 2.3.15b) of AEV ts34 poly(A)<sup>+</sup> RNA (300 ng) was performed using a synthetic oligonucleotide (Section 2.2.6a), complimentary to H2A<sub>F</sub> mRNA (28-mer, Figure 3.9) which was extended using reverse transcriptase in the 3' to 5' direction. The predicted cap-site is indicated in Figure 3.9.

AEV ts34  
Markers



### FIGURE 3.9

#### MAPPING 5' TERMINI OF THE H2A<sub>F</sub> TRANSCRIPT

Mapping the cap-site of H2A<sub>F</sub> mRNA using data derived from primer extension analysis in Figure 3.8. Including the length of the gene-specific synthetic primer (28-mer, indicated by the horizontal line), which hybridizes to the transcript corresponding to the positions indicated on the gene, an extension product of 144 b was generated, therefore allowing the cap-site to be assigned to an 'A' residue (indicated by a vertical arrow).



GATCCGCAGGCCACGCCTC **CCGCCC** CCCCCACCCCGCCTTCAGCCTC TCGTGCCGTGTGCA

AAGAAAAACAGCCGTTAAACGGGCCGCG **GGGCGG** GGCCAGCGCGGGCCCAACAAGCGGC

CGGGCGATTGGCGGCAGCGGGGCG **TTCAAAT** CAGTACGGGGCGCTCCGATTGGCTCGGGGTGC

CCGCGGCAGGCCGGGCTCCCTCCTCGCGCTCGGGCGGGTATTGTCTGGCGGCTCCTGAGGCGGTT

CGAGCGCGGATTTCGGGGACCGGGGCTCGGGCGGGCGGCACC ATG **1459 bp**

**intron** GCA GGT GGG AAG GCT GGG AAG GAC AGC GGG  
ala gly gly lys ala gly lys asp ser gly

AAG GCC AAG GCG AAG GCG GTG TCT CGC TCG CAG AGA GCC GGA TTG CAG  
lys ala lys ala lys ala val ser arg ser gln arg ala gly leu gln

GTAACGAGCGGGCGCGCAGTGCTCCCCATCTGTACCACGCTAATTGCTGTAAGCACCGCTCCCG  
CTGCTGACACAGTGCCCCGCGACGCCTTTTCCCCCAATCCAAGGTGAAGTTGGCTGCTGTGCG  
TTGTCCTTCGGTCTCGCTCACCCCTCTGGCTGTGCCTGCAGGGCAGGGCAGGGCCGGATGCTCCCG

fluctuations in the rate of DNA synthesis (see Section 1.11.2). These genes foreseeably would be transcribed constitutively, thus accounting for the 'basal' levels of steady-state mRNA and transcription outside of S-phase. The chicken H3.3 gene (Brush et al., 1985) is such an example. Furthermore, differential coupling of chicken H1 proteins (Winter et al., 1985a) and mRNAs (Winter et al., 1985b) to DNA replication has been observed during myogenesis. Finally, the possibility exists that transcription of S-phase regulated genes is not completely 'silenced' in the absence of DNA synthesis.

To determine the extent which transcription is responsible for histone mRNA accumulation during S-phase, rates of transcript synthesis were measured, during and in the absence of DNA synthesis. Quantitation of nascent transcripts synthesized over a 15 minute pulse-labelling interval in isolated nuclei indicates that the rate of core and H1 histone mRNA synthesis increases by approximately 5-fold during S-phase, in comparison to pre-release rates (see Figure 3.6 and 3.7). Significantly, the kinetics of transcript synthesis precedes the steady-state accumulation of corresponding mRNAs and the kinetics of DNA replication by approximately 1 hour (Figure 3.6 and 3.7). Elevated rates of histone gene transcription per se clearly do not account for the total increase in cytoplasmic mRNAs detected. In light of considerable evidence (Heintz et al., 1983; Graves and Marzluff, 1984), it is likely that prolonged mRNA stability is an important factor in regulating chicken histone transcript levels in S-phase.

Transcription of genes encoding the two highly variant histone subtypes, H5 and H2A<sub>F</sub>, is constitutive and constant throughout the cell-cycle (Figure 3.6). As steady-state levels of H5 and H2A<sub>F</sub> mRNAs also are invariant (Figure 3.4 and 3.5), it seems unlikely that these transcripts exhibit differential stability throughout the cell-cycle. Both H5 and H2A<sub>F</sub> genes can therefore be categorized into the

replication-independent class, in marked contrast to the expression of major core and H1 chicken histone genes which, in general, are S-phase regulated. The AEV ts34 cell-line specifically chosen for these studies (see Section 2.2.9) does not have the capacity to differentiate into non-dividing erythrocytes, even under suitable conditions and is essentially a non-producer of globin (Beug et al., 1982). Furthermore, the profile of H1 and H5 proteins and low abundance of  $\alpha$ -globin mRNA in these cells (Figure 3.1) is characteristically early-erythroid. Taken together with the in situ hybridization studies presented (Table 3.1), these facts render it extremely unlikely that H5 and H2A<sub>F</sub> mRNA detected in these experiments arose from a subpopulation of cells which was not under cell-cycle control.

Replication-uncoupled expression of the H5 gene is clearly consistent with a requirement for continued synthesis of H5 protein during erythroid differentiation, when cell-division has ceased and DNA-replication terminated. H5 protein has the capacity to displace H1 molecules from internucleosomal DNA and to condense chromatin (Thoma et al., 1979). If H5 (and its mammalian counterpart H1<sup>o</sup>, Gjerset et al., 1982) appeared only post-mitotically in differentiating cell-lines, a direct causal relationship between the presence of this histone and the cessation of DNA synthesis may be implied. However, H5 certainly appears in early erythroblasts which undergo a number of cell divisions before maturing into erythrocytes (Figure 3.1a; Beug et al., 1982). This unusual histone accumulates to higher levels in nondividing cells, and its major biological role may be to condense chromatin in maturing cell-lines thereby precluding further transcription. If this is the case, H5 (or H1<sup>o</sup>) synthesis is not required to maintain the normal chromatin structure of dividing cells, and thus synthesis need not be coupled to S-phase. Instead, stable H5 mRNA is constitutively synthesized, facilitating the accumulation of H5 protein.



Considerably less is known about regulation of H2A<sub>F</sub> levels in dividing and non-dividing tissues, or even if it is incorporated into chromatin. Evidence presented that the H2A<sub>F</sub> gene is not under cell-cycle control is probably indicative that synthesis of H2A<sub>F</sub> protein is also unlinked to DNA replication, similar to H5, or is partially replication-independent (see Section 1.11.2). A clue as to the role of this protein could lie in what is known about related histone subtypes identified in other species. For instance, synthesis of an apparently similar histone protein (M1) to H2A<sub>F</sub>, identified in mouse erythro-leukemia cells, is completely independent of DNA replication and is not thought to participate in 'normal' nucleosome formation (Grove and Zweidler, 1984). Hence it may have a specialized role in dividing and postmitotic cells. Another related histone variant, H2A.Z, is enriched in the nucleoli of mammalian cells and its homologue, hvl, is found exclusively in the transcriptionally active macronucleus of Tetrahymena (Allis et al., 1982; see Grove and Zweidler, 1984; Allis et al., 1986). Nevertheless, the biological role of H2A<sub>F</sub> (if any) awaits experimental verification.

It has been suggested that the structural features which distinguish 'basal' histone genes from the more abundant core and H1 replication-dependent genes is an important feature in their regulation (Brush et al., 1985; Seiler-Tuyns and Paterson, 1986). This has been implied due to the correlation between levels of 'basal' histone mRNAs which are polyadenylated (i.e. have no 3' stem-loop structure) and, contain introns in their primary transcripts. However, the requirement for introns is not absolute as intervening sequences are absent in the chicken H5 gene (Krieg et al., 1982). The absence of typical histone 3' termination signals in this class of genes is a recurring feature which, in part at least, is likely to account for their constitutive pattern of expression during the cell-cycle. This is consistent with the

observation that replacement of histone 3' termination sequences with globin mRNA polyadenylation signals greatly enhances transcript stability (Graves et al., 1987). In fact, the downstream sequences required for efficient and accurate processing of replication-dependent histone transcripts coincide with those required to confer cell-cycle regulation on a heterologous gene (Lüscher et al., 1985; Schaufele et al., 1986), which is consistent with the synthesis of polyadenylated histone mRNAs being replication-independent. A simple model accounting for this could involve differential processing of primary transcripts at different cell-cycle stages resulting in modified stability or, by degradative mechanisms operating on identical transcripts but preferentially active outside of S-phase. The former possibility is consistent with observations that histone mRNA degradation occurs by an exonucleolytic event in the 3' to 5' direction (Ross and Kobs, 1986). Incorrect formation of 3' ends outside of S-phase, as a result of differential processing, conceivably could act as a preferential substrate for this activity which would enhance mRNA degradation. This situation would not apply to polyadenylated 'basal' histone transcripts, such as H5 and H2A<sub>F</sub>, which appear to have constant half-lives throughout the cell-cycle and are not subject to the same 3' processing events.

A requirement for poly(A)<sup>-</sup> mRNA is clearly not applicable to cell-cycle regulated mRNAs in all eukaryotes. Yeast (Hereford and Osley, 1981; Hereford et al., 1982) and Tetrahymena (reference cited within Pederson et al., 1986; and see Bannon et al., 1983, 1984) histone transcripts are generally polyadenylated but also their regulation is typically linked to DNA replication. Perhaps 3' processing of histone transcripts in lower eukaryotes is of minor importance and transcriptional activation is the predominant feature in their periodic regulation. Several lines of evidence suggest that polyadenylation of 'basal' histone mRNAs is not exclusively responsible for their

replication-independent accumulation. Transcription of H5 and H2A<sub>F</sub> genes for instance, is uncoupled from DNA replication (Figure 3.6), implying that 5' sequences may also be involved. Furthermore, in contrast to work presented by Seiler-Tuyns and Paterson (1986), evidence suggests that 5' flanking histone sequences can confer cell-cycle regulation on hybrid poly(A)<sup>+</sup> mRNAs (Artishevsky et al., 1985).

Although the H5 gene has some recognizable remnants of H1 histone gene structure (Coles and Wells, 1985; Coles et al., submitted for publication), a 5' element which is ubiquitous to all H1 genes (H1-box) is absent from the H5 gene. The lack of this sequence near the H5 gene may explain the absence of S-phase transcriptional control. Evolution of the H5 gene from the H1 gene(s) may have resulted in the loss of putative elements involved in S-phase activation and the concomitant gain of sequences responsible for erythroid-specific expression. The role of the H1-gene-specific element will be addressed in Chapter 4.

C H A P T E R 4

TRANSCRIPTIONAL REGULATION OF HISTONE

GENE EXPRESSION:

A ROLE FOR H1 AND H2B GENE-SPECIFIC ELEMENTS

## CHAPTER 4

### TRANSCRIPTIONAL REGULATION OF HISTONE GENE EXPRESSION :

#### A ROLE FOR H1 AND H2B GENE-SPECIFIC ELEMENTS

##### 4.1 INTRODUCTION

Through extensive DNA sequence analysis conducted in this laboratory a number of highly conserved promoter elements have been identified in each of the major core and H1 chicken histone genes. In particular, gene-specific elements identified in the 5'-flanking regions of H1 (Coles and Wells, 1985) and H2B (Harvey et al., 1982) genes are of special interest as they may play a pivotal role in transcriptional regulation of their cognate genes.

Sequence analysis of the full complement of six chicken H1 histone genes has revealed a number of common promoter elements and a characteristic promoter architecture (see Figure 4.1). Four discrete promoter elements have been localized within a region between -120 and -32 base pairs (bp), relative to the cap site and include the classical 'TATA' and 'CCAAT'-boxes in addition to a G/C-rich upstream region and a characteristic A-rich motif. The G/C-rich hexanucleotide sequence (5'-GGGCGG-3') at approximately -55 has been previously described in the 5'-flanking regions of several other RNA polymerase II-transcribed genes (Section 1.3.2). This motif is identical to the consensus core recognition sequence required for binding of the transcription factor SP1 (Kadonaga et al., 1986). Furthermore, an intact G/C element is required for efficient H1 promoter activity in Xenopus oocytes and HeLa cells (Younghusband et al., 1986), presumably by functioning as a recognition binding sequence for SP1. The A-rich heptanucleotide sequence 5'-AAACACA-3', referred to as the H1-box, is located at approximately position -120 and is ubiquitous to all H1 genes sequenced



in species ranging from sea urchin to human (Coles et al., submitted for publication). An extended homology 5'-AGGAAACACACAA-3' exists amongst the six chicken H1 genes. Notably, this element is absent in the extreme H1 variant, histone H5 gene, which in chicken is expressed exclusively in the erythroid cell lineage (Shannon et al., 1985; Section 3.2.1) and is not under cell-cycle control (Chapter 3). The extreme evolutionary conservation with respect to nucleotide sequence and relative position of this element within its cognate promoter implies a distinct regulatory function.

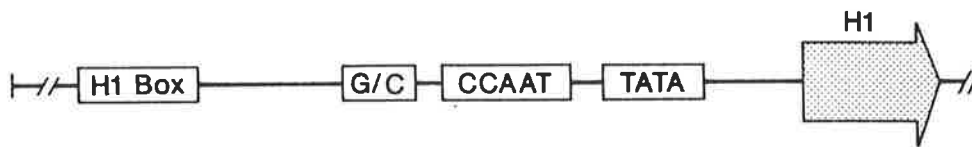
Analysis of the nine H2B genes in chicken at the DNA sequence level has revealed a highly structured and conserved promoter architecture (see Figure 4.2). Each H2B gene possesses a 'TATA'-box, three 'CCAAT'-elements and a gene-specific motif (H2B-box; 5'-CTCGATTTGCATA(G/C)-3'), originally described by Harvey et al. (1982). Six of the chicken H2B genes are paired with an H2A gene and ordered in a divergent arrangement (Sturm, 1985). In each case, the relative order and DNA sequence of H2B and H2A elements in the intragenic region are conserved to a high degree (Figure 4.2). Furthermore, this region may constitute an overlapping promoter whereby DNA sequence elements in the intergene region may influence expression of both genes (Sturm, 1985). While apparently absent in other histone genes, an element closely resembling the H2B-box has been identified in the upstream regions of immunoglobulin (Falkner and Zachau, 1984; Parslow et al., 1984) and small nuclear RNA genes (Mattaj et al., 1985). These studies show this element to be required for efficient promoter activity.

Work presented in this chapter focusses on the regulatory role of the H1 and H2B gene-specific elements in control of histone gene expression, particularly in the periodic transcriptional activation of their cognate genes during S-phase of the cell-cycle.

## FIGURE 4.1

### PROMOTER ARCHITECTURE OF CHICKEN H1 GENES

The promoter regions of six chicken H1 genes have been sequenced (Coles, 1986; Coles et al., manuscript submitted) and conserved regions have been aligned for comparison. Positions of these elements are indicated, relative to the predicted cap-sites, in addition to the inter-element spacing. Histone gene nomenclature has been described (D'Andrea et al., 1985; Coles, 1986). Protein subtype variants (\*) have recently been assigned to each of the H1 genes listed (Shannon and Wells, manuscript submitted for publication). Adapted from Coles (1986).



<u>Gene</u>	<u>Subtype* Variant</u>	<u>H1 Specific</u>	<u>G-Rich</u>	<u>CCAAT</u>	<u>TATA</u>
01 H1	H1-c'	-118 AAGAAACACAAA...32...GCGGGGCGGGCT...7...GCACCAATCA...15...CTAAAAATA	-74	-55	-32
02 H1	H1-d	-120 AAGAAACACAAA...32...GCGGGGCGGGCT...7...GCACCAATCA...15...CTATAAATA			
.01 H1	H1-c	-118 AAGAAACACAGA...32...GTTAGGCGGGCT...7...GCACCAATCA...15...CTATAAATT			
11L H1	H1-a	-105 AAGAAACACAAG...27...GCGGGGAGGGCT...7...GCACCAATCA...15...CTATAAATA			
11R H1	H1-a'	-113 AAGAAACACAAG...31...GCGGGGCGGGC...5...GCACCAATCA...14...CTATAAATG			
03 H1	H1-b	-89 AAGAAACACAAC...17...GCGGGGCGA-----ACCAATCA...15...CTATAAAGG			

## FIGURE 4.2

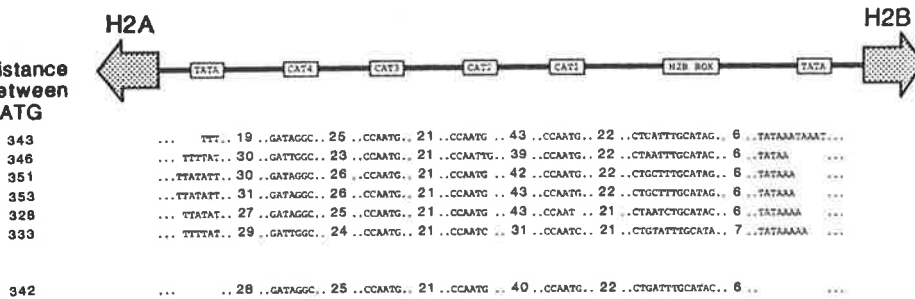
### PROMOTER ARCHITECTURE OF CHICKEN H2B GENES

The promoter regions of nine chicken H2B genes have been sequenced (Sturm, 1985) and conserved regions have been aligned for comparison. Positions of these elements are indicated, relative to the predicted cap-sites, in addition to interelement spacing. Six of the H2B genes are paired with an H2A gene. Adapted from Sturm (1985).

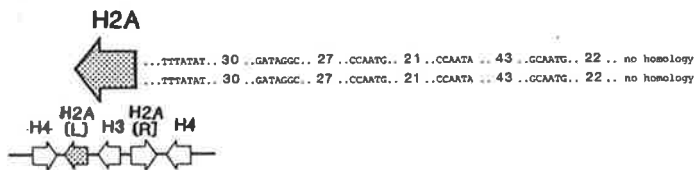
**Subclone**

**Distance  
Between  
ATG**

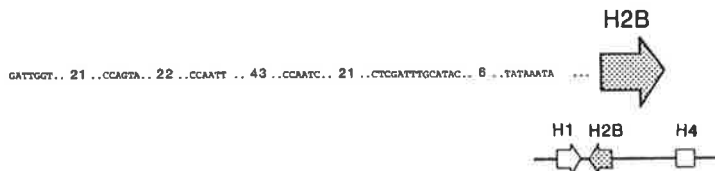
pCh1.3E/pTAT  
pCh2.0W  
pCh1.5K  
pCh1.0E L  
pCh1.0E R  
λCh1.10



pCh0.4E L  
pCh0.4E R



λCh1.02



## 4.2 RESULTS

The basic construct (pCH5B.P) used in these studies consisted of a 5 kb Bam HI fragment subcloned from a lambda genomic clone,  $\lambda$ CH01 (D'Andrea et al., 1985), and has been described in detail elsewhere (Sturm, 1985; Younghusband et al., 1986). This 5 kb fragment, consisting of an H1 gene with 5'- and 3'-flanking DNA in addition to a linked and divergently transcribed H2A/H2B gene pair (see Figure 4.4b), was subcloned into the vector pAT153 (Section 2.2.5) for expression studies in cultured cells.

### 4.2.1 The H1 Gene-Specific Element is a cis-Acting Transcriptional Regulator

The regulatory function of the highly conserved H1-specific promoter element (5'-AAACACA-3') was initially investigated by transfecting the linearized pCH5B.P construct into HeLa cells with either the element intact or deleted by site-directed mutagenesis using synthetic oligonucleotides (Section 2.3.19).

Following deletion mutagenesis of the H1-box, the 5 kb BamHI chicken histone gene fragment was sub-cloned in the vector pAT153, generating the deletion construct pCH. $\Delta$ H1 (Figure 4.3 and 4.4). Constructs were linearized at the Eco RI site and co-transfected with the selectable marker plasmid, pSV2Neo (Southern and Berg, 1982) into HeLa cells by the calcium phosphate precipitation method (Section 2.3.18). Following selection of transformed, G418-resistant clonal cell-lines, levels of transcripts were determined by primer extension analysis using a gene-specific [<sup>32</sup>P]-labelled synthetic primer (26-mer, Section 2.3.10e). Similarly, transcripts generated from the linked chicken histone H2A and H2B genes were used as an internal control. Products generated by primer extension analysis on RNA using specific synthetic oligonucleotides (H1, H2A and H2B primers, 26-mer, Section 2.2.6a) have previously been characterized (Sturm, 1985). Typically,

major bands of 49, 58 and 68 bases are observed for H2A, H1 and H2B extension products, respectively. Three clonal cell-lines transfected with pCH. $\Delta$ H1 and three pCH5B.P ('parental')-transfected cell-lines were selected for analysis of transcript levels. Amounts of RNA from each cell-line, on which primer extension analysis was performed, generally varied so that intensities of extension products on autoradiograms were approximately equal.

Initially, relative steady-state transcript levels for H1, H2A and H2B genes were determined by primer extension analysis using gene-specific oligonucleotides (Section 2.2.6a). Between each of the six clonal cell-lines studied, relative H2B and H2A transcript levels were essentially invariant (Figure 4.4). In contrast, a striking reduction in H1 mRNA levels was detected in pCH. $\Delta$ H1-transfected HeLa cells. To determine the degree which H1 transcript levels drop in response to deletion of the H1 gene-specific element, relative amounts of H1 and H2B extension products ( $\alpha$  mRNA abundance) were determined by densitometric scanning of autoradiograms (Section 2.3.5) shown in Figure 4.4. Densitometric scans revealed a marked decline in H1 mRNA levels relative to H2B transcripts in pCH. $\Delta$ H1-transfected cell-lines compared to pCH5B.P-transfected ('parental') cells (Table 4.1). These results represent a general decrease of approximately 20 fold in H1 transcript levels as a consequence of H1-box deletion.

Cell-cycle regulated expression of CH5B.P histone genes was assayed in one of the G418-selected 'parental' cell-lines (Figure 4.4; lane 1), synchronized by a sequential thymidine/aphidicolin block (Section 2.3.11b; see Figure 4.5). Primer extension analysis of RNA isolated from cells at various stages of the cell-cycle indicated that H1 and H2B genes within the pCH5B.P construct are regulated in co-ordination with DNA replication (see Figure 4.5). Sequences required for cell-cycle regulation of these genes are therefore present within the 5 kb BamHI

### FIGURE 4.3

#### H1 AND H2B GENE PROMOTER MANIPULATION BY SITE-DIRECTED MUTAGENESIS

Site-directed mutagenesis of the wild-type/parental 5 kb Bam H1 fragment (see Figure 4.4) in M13 was performed using synthetic oligonucleotides (Section 2.2.6c), generating a series of mutant constructs in the H1 or H2B promoter region. DNA sequence data for H1 and H2B genes in CH5B.P was taken from Coles (1986) and Sturm (1985), respectively. Sequences indicated display the relevant region of the resulting Bam H1 fragment which was cloned into the vector pAT153 (Section 2.2.5) for expression studies in cultured cell-lines. Positioning of the H1 and H2B gene-specific elements, in addition to an upstream insertion site in the CH. $\Delta$ H1/4 H1 promoter, are indicated relative to the respective transcription start sites. Base substitutions, where applicable, are indicated (\*). Sequences represented are as follows:

CH5B.P/H1: Sequences flanking and including the H1-box in the 'parental' ( $\lambda$ CH01) H1 promoter.

CH. $\Delta$ H1: Deletion of the core H1-box sequence (from Younghusband et al., 1986).

CH. $\Delta$ H1/4: Base substitutions within the H1 box.

CH.H1/R0: H1-specific element replaced in the reverse orientation.

CH.H1/US.P: Sequences in the upstream region of the CH. $\Delta$ H1/4 H1 promoter where insertional mutagenesis of the H1-box was performed.

CH.H1/UI: Upstream insertion of the H1-box into the pCH. $\Delta$ H1/4 construct.

CH5B.P/H2B: Sequences flanking and including the H2B-box in the 'parental' H2B promoter.

CH5B. $\Delta$ H2B: Deletion of the H2B-box element (Sturm, 1985)

CH5B. $\Delta$ H2B/3: Base substitutions within the H2B-box.

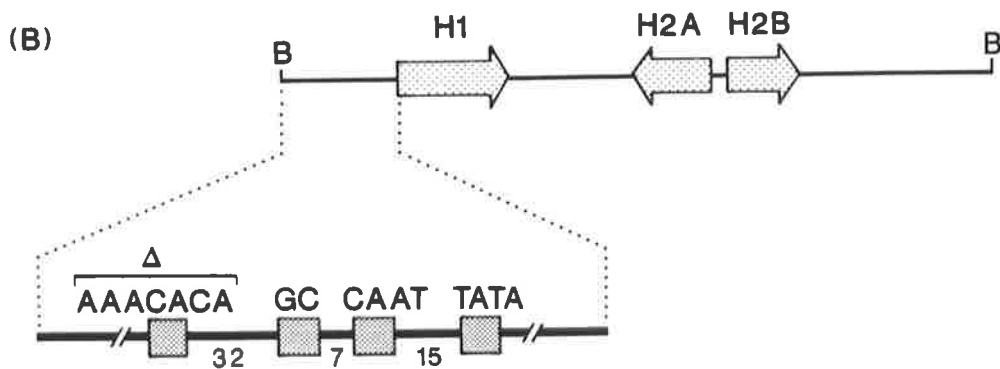
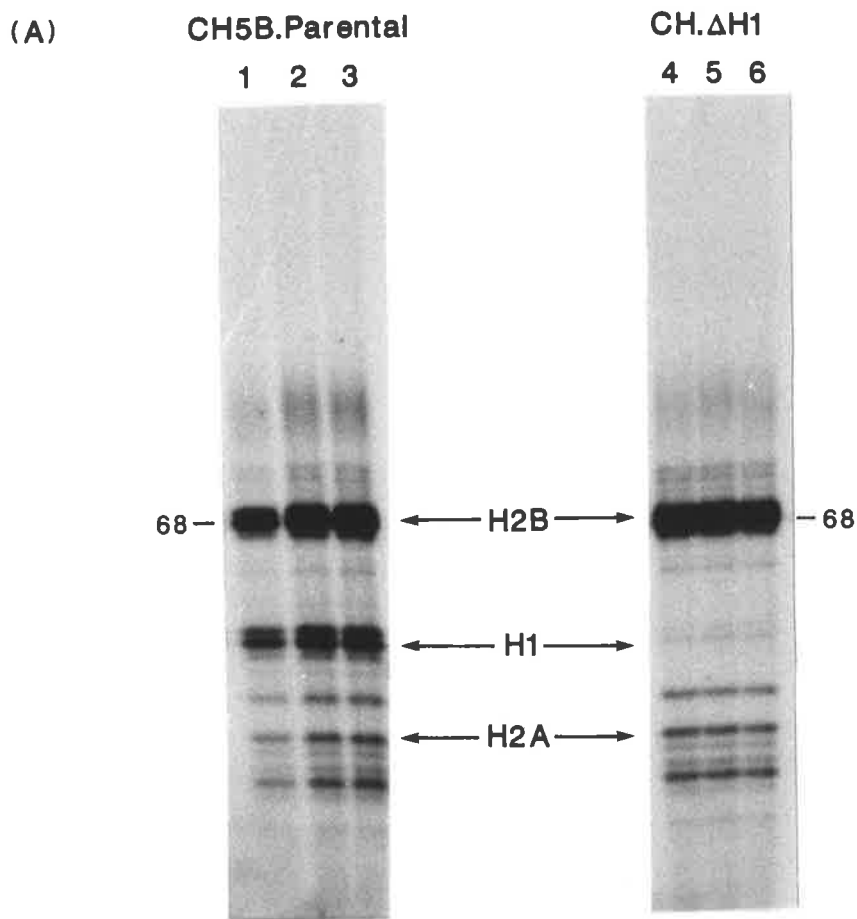




## FIGURE 4.4

### DELETION OF THE H1 GENE-SPECIFIC ELEMENT REDUCES H1 TRANSCRIPT LEVELS

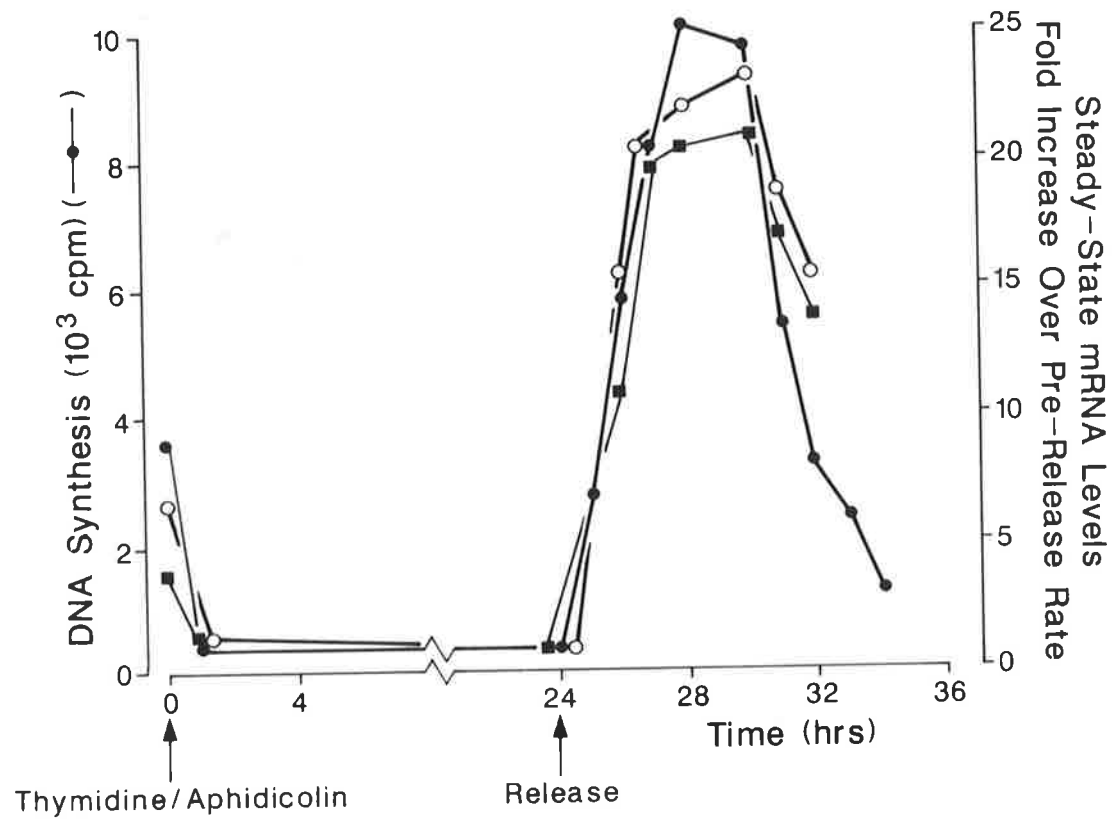
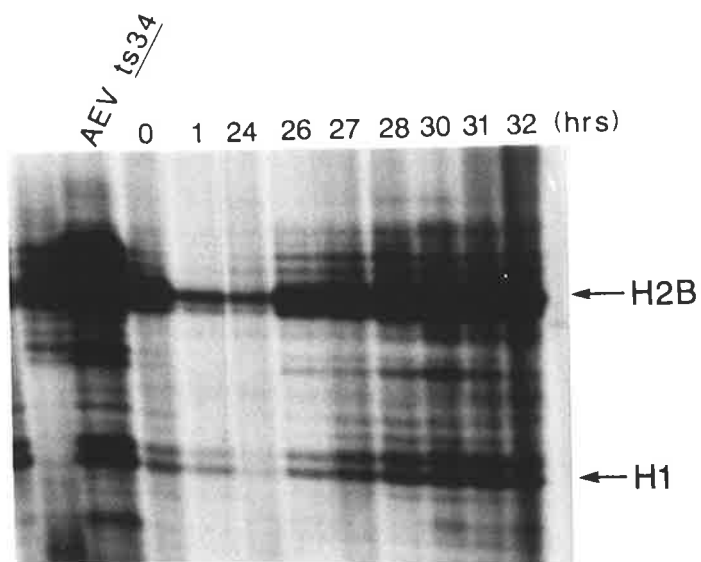
- A. HeLa cells in exponential growth phase were cotransfected with pCH5B.P ('parental' construct, see Figures 4.4b, 4.3; lanes 1-3) or pCH.ΔH1 (H1-box deleted, see Figure 4.3; lanes 4-6) and pSV2Neo. Following selection with G418, three clonal cell-lines, for each construct transfected, were selected for analysis. Steady-state H1, H2A and H2B mRNA, transcribed from exogenous-transfected genes, were detected in RNA samples from unsynchronized clonal cell-lines by primer extension analysis (Section 2.3.15b). Amounts of RNA from each cell-line used for primer extension analysis were normalized so that signal intensities were approximately equal. Relative levels of transcripts in each cell-line are presented in Table 4.1. Extension products are shown relative to a pBR322 Hpa II-digested (68b) marker.
- B. The 5 kb Bam H1 fragment shown containing H1, H2A and H2B chicken histone genes was subcloned into pAT153, generating the construct, pCH5B.P. Positions of conserved elements in the H1 promoter are shown, including the H1-box on which deletion or substitution mutagenesis (see Figure 4.3) was performed. The 'AAACACA' motif represents the H1-box. In addition, a G/C-box, 'CCAAT'-box, and a 'TATA'-box are represented in addition to the length of intervening DNA (see Figure 4.1)



## FIGURE 4.5

### CELL-CYCLE REGULATION OF EXOGENOUS H1 AND H2B GENES IN A HELA CELL-LINE

HeLa cells in exponential growth phase were cotransfected with pCH5B.P ('parental' construct; see Figure 4.4) and pSV2Neo. Following selection with G418 a clonal cell-line (as in Figure 4.4a; lane 1) was selected for cell-cycle analysis. Cells were synchronized by sequential treatment with thymidine and then aphidicolin (Section 2.3.11a(ii)), cytoplasmic RNA extracts were prepared at various time points, and analyzed by primer extension analysis (Section 2.3.15b). Levels of H1 (○) and H2B (■) mRNA were determined by laser densitometry of the autoradiogram (Section 2.3.5) and plotted as a function of time relative to pre-release (24 hr. + aphidicolin) levels. Progression of cells through S-phase and into G<sub>2</sub> was monitored by pulse-labelling with [<sup>3</sup>H]- thymidine (●) over 15 minutes intervals (Section 2.3.11b).



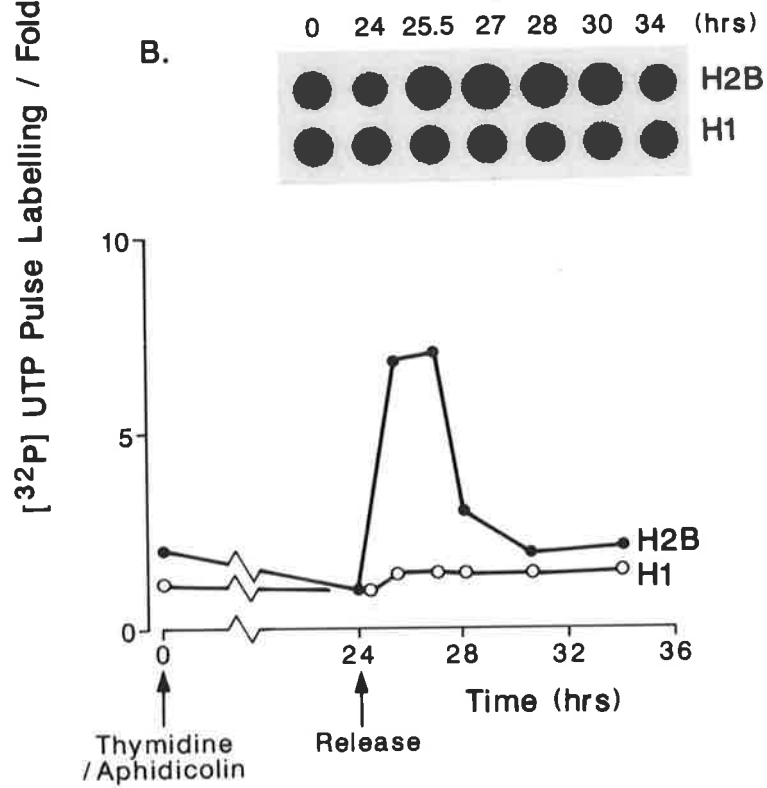
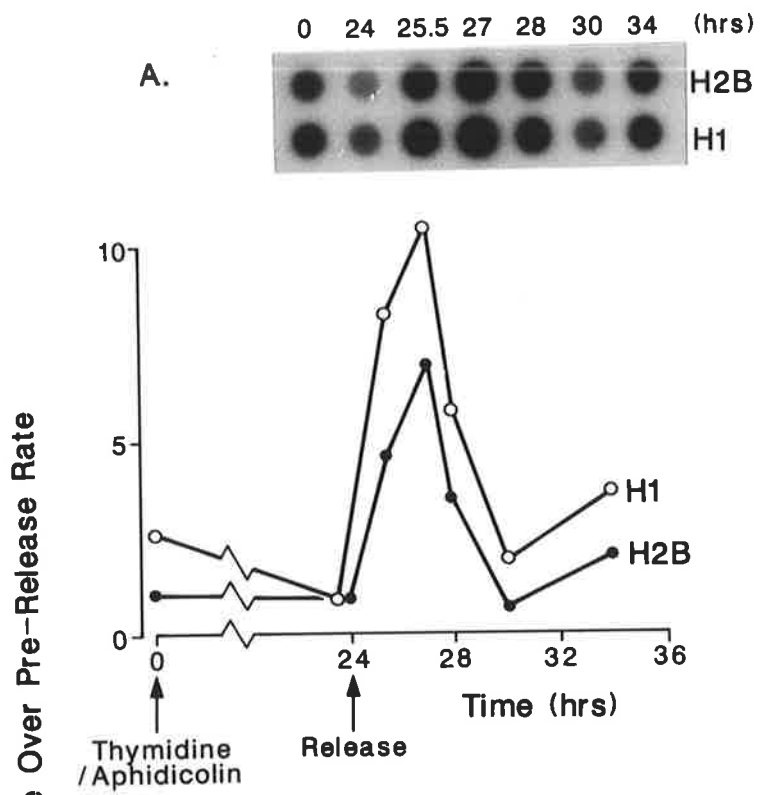
restriction fragment. Although deletion of the H1-specific element clearly reduced levels of steady-state H1 mRNA levels in vivo (see Figure 4.4), the level at which this element controls transcript levels was not defined by this analysis. Next, the possibility was examined that the H1-box functions as a transcriptional regulator and, if this element is required for cell-cycle regulation. Nuclei were isolated from pCH5B.P and pCH. $\Delta$ H1 transformed cell-lines following cell synchronization (during, before and after S-phase), pulse-labelled with [ $^{32}$ P]rUTP and nascent [ $^{32}$ P]-transcript levels quantitated by dot-blot hybridization so as to determine, (a) if the H1-box regulates the rate of H1 gene transcription and, (b) if this transcriptional control is cell-cycle specific. This method allows transcriptional activity to be monitored in isolated nuclei which directly reflects the rate of transcription in cells from which they were derived.

In nuclei derived from pCH5B.P ('parental') and pCH. $\Delta$ H1-transfected cell-lines, levels of H2B gene-specific nascent transcripts increased in the order of 7-fold over pre-release (aphidicolin blocked) levels (Figure 4.6), displaying a transcriptional profile typical of S-phase regulated genes (see Figure 3.7). Similarly, levels of H1 nascent transcripts in 'parental' nuclei were regulated in a cell-cycle-dependent manner displaying a 10-fold increase over pre-release levels, 3 hours into S-phase. In contrast, however, [ $^{32}$ P]H1 transcript levels from ' $\Delta$ H1' nuclei were constant throughout the cell-cycle. Furthermore, relative rates of maximal transcription in 'parental' nuclei (3 hours in S-phase) were approximately 7-fold greater than in ' $\Delta$ H1' nuclei at the same time point (Table 4.1). Deletion of the H1-specific element not only generally reduced H1 transcriptional efficiency from its cognate promoter, but also resulted in a loss of S-phase regulatory control. Hybridization to a negative control set of M13 DNA dots was always less than 20% of the corresponding control signal with histone gene insert alone.

## FIGURE 4.6

### DELETION OF THE H1-BOX ELIMINATES CELL-CYCLE CONTROL OF H1 GENE TRANSCRIPTION

HeLa cells in exponential growth phase were transfected with, (a) the 'parental' pCH5B.P construct or, (b) pCH.ΔH1 (H1-box deleted; see Figure 4.3 and 4.4); in addition to pSV2Neo. Following selection with G418, nuclei were isolated from a pCH5B.P (as in figure 4.5) and pCH.ΔH1-transfected (Figure 4.4; lane 5) clonal cell-lines at various stages of cell-synchronization (before, during and after addition of aphidicolin). Nuclei were pulse-labelled with [<sup>32</sup>P]UTP (Section 2.3.16b) and levels of H1 (○) and H2B (●) nascent transcripts were monitored by dot-blot hybridization with 2.5 μg of filter-immobilized H1 or H2B chicken histone gene insert (Section 2.2.4). Quantitation of nascent transcript levels was determined by scintillation spectroscopy of individual DNA-RNA hybrid dots. Non-specific hybridization of labelled RNA (to 2.5 μg of M13 DNA) was subtracted, and fluctuations in H1 and H2B transcript levels plotted as a function of time. Results are presented as the mean value (background subtracted) of two independent experiments.





While deletion of the H1-box clearly had a profound effect on steady-state H1 mRNA levels, in addition to the rate of gene transcription and its coupling to DNA replication, it was possible that this effect was a consequence of general promoter disruption rather than being element-specific in nature. It was therefore imperative to demonstrate that the H1-box had an important regulatory role by maintaining the general spatial arrangement of sequences within the promoter. For this purpose, four base substitutions were introduced into the H1-box core region (5'-AAACACA-3') by site-directed mutagenesis, generating the mutant sequence 5'-GACTACG-3' and the construct pCH. $\Delta$ H1/4, thereby conserving spatial architecture of the promoter (see Figure 4.3).

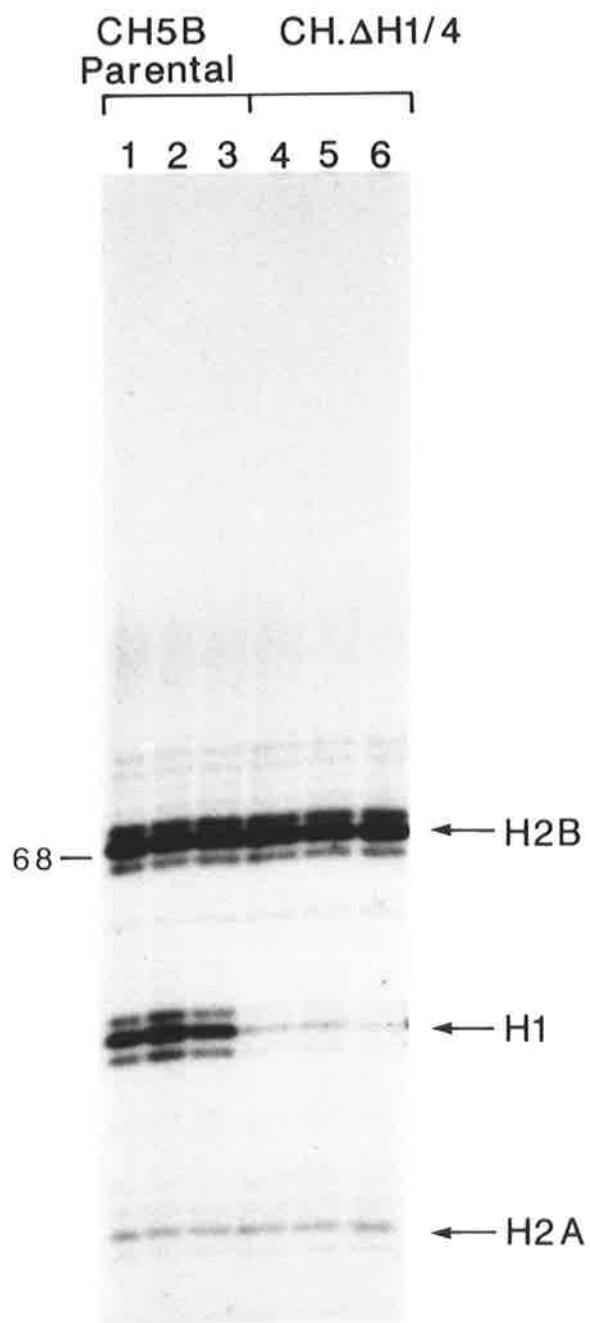
Following the general procedure described for pCH5B.P/pCH. $\Delta$ H1, three G418-resistant clonal cell-lines transfected with pCH. $\Delta$ H1/4 were selected and transcript levels for H1, H2A and H2B genes were analyzed in comparison to levels from pCH5B.P ('parental') cell-lines. H1 transcript levels in cells transfected with a base-substituted H1-box were clearly lower than in 'parental' cell-lines (Figure 4.7). This is indicated by a relative decrease of approximately 20-fold in H1 transcript levels (Table 4.1). The general decrease in steady-state mRNA levels due to H1-box base substitution was also reflected by a marked decline in the rate of gene transcription which was accompanied by a loss of S-phase regulatory control (Figure 4.8). Significantly, H2B gene transcription retained a tight link with DNA replication and kinetically the rate of transcription closely resembled that in the 'parental' cell-line.

The role of the H1-specific element in transcriptional regulation was further investigated by determining its ability to function in either orientation as well as its relative dependence on position within the promoter. Orientation-independent regulation by the H1-box was

## FIGURE 4.7

### BASE SUBSTITUTION WITHIN THE H1-BOX REDUCES H1 TRANSCRIPT LEVELS

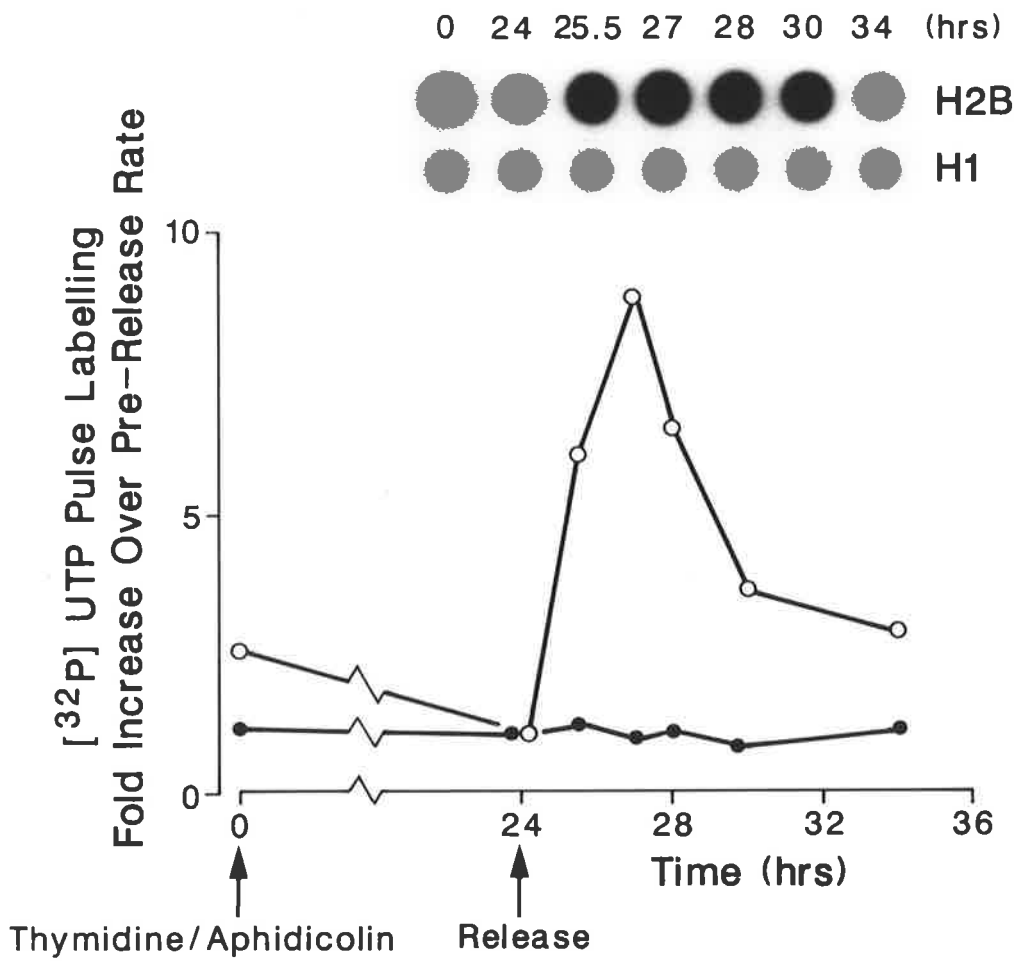
HeLa cells in exponential growth phase were cotransfected with pCH5B.P (lanes 1-3) or pCH.ΔH1/4 (lanes 4-6) and, pSV2Neo. Following selection with G418, three clonal cell-lines for each construct transfected were selected for analysis. Steady-state H1, H2A and H2B mRNA, transcribed from exogenous-transfected genes, was detected in RNA samples extracted from unsynchronized clonal cell-lines by primer extension analysis (Section 2.3.15b). Amounts of RNA from each cell-line used for primer extension analysis were normalized so that signal intensities were approximately equal. Relative levels of transcripts in each cell-line are presented in Table 4.1. Extension products are shown relative to a pBR322 Hpa II-digested marker (68b).



## FIGURE 4.8

### BASE SUBSTITUTION WITHIN THE H1-BOX ELIMINATES CELL-CYCLE CONTROL OF H1 GENE TRANSCRIPTION

HeLa cells in exponential growth phase were cotransfected with pCH. $\Delta$ H1/4 (4 base substitutions within the H1-box; see Figure 4.3 and 4.4) and pSV2Neo. Following selection with G418, nuclei were isolated from a selected clonal cell-line (See Figure 4.7; lane 6) at various stages of cell synchronization (before, during and after addition of aphidicolin). Nuclei were pulse-labelled with [ $^{32}$ P]UTP (Section 2.3.16b) and rates of H1 ( $\bullet$ ) and H2B ( $\circ$ ) nascent transcripts monitored by dot-blot hybridization with 2.5  $\mu$ g of filter-immobilized H1 or H2B gene insert (Section 2.2.4). Quantitation and presentation of data is as described in the legend to Figure 4.6. Results are presented as the mean value (M13 hybridization background subtracted) of two independent experiments.



tested by reversing its polarity in the CH5B.P-H1 promoter using oligonucleotide-directed mutagenesis (see Figure 4.3). The 5 kb Bam H1 fragment with an H1-box in reverse orientation was excised from the M13 vector and subcloned into pAT153, generating the construct pCH.H1/R0 (Figure 4.3), for co-transfection into HeLa cells with pSV2Neo. Three G418-resistant clonal cell-lines were selected for analysis and levels of H1, H2A and H2B transcripts compared with those in 'parental' (pCH5B.P-transfected) cell-lines.

Unlike deletion or base substitution of this element, reversing the orientation of the H1-box had only a marginal effect on H1 transcript levels (Figure 4.9). Analysis of H1/H2B transcripts in 'parental' and pCH.H1/R0-transfected cell-lines indicate that H1 mRNA levels only decreased by about 10-15% as a consequence of reversing H1-box polarity (see Table 4.1). Cell-cycle regulated transcription of the H1 gene in one of the pCH.H1/R0-transfected cell-lines (Figure 4.9; lane 6) was tested by pulse-labelling nuclei isolated from cells at various stages of synchronization. Not only did the H1 gene display periodic transcriptional activation in S-phase, but increases in rates of transcription relative to pre-release (+ aphidicolin) rates were comparable to those in parental-lines (Figure 4.10; represented by broken lines). Only a 20% relative decrease in the peak rate of H1 gene transcription was detected as a consequence of reversing H1-box polarity (Table 4.1).

To determine if the H1-box could function from an upstream position, it was inserted approximately 500bp upstream from its natural site, within the pCH. $\Delta$ H1/4 H1 promoter (which has a mutated H1-box, see Figure 4.3). This insertion consisted of the heptanucleotide 'core' sequence (5'-AAACACA-3') in addition to 7 bp of 3'- and 5'-flanking sequence (see Figure 4.3). Insertion into this construct was selected in an attempt to see if an intact H1-specific element could reactivate

cell-cycle transcriptional control as the H1 gene in pCH. $\Delta$ H1/4 is constitutively expressed at low levels throughout the cell-cycle (see Figure 4.8). The insertion site and the insertion itself are shown in Figure 4.3. The site of insertion was chosen so that the upstream H1-box would be on the same helical face as the G/C-box (see Section 1.7). The 5 kb Bam HI fragment was subcloned into pAT153 generating the construct pCH.H1/UI, co-transfected with pSV2Neo into HeLa cells, and G418-resistant colonies selected for analysis of steady-state mRNA levels (Figure 4.9). In each case, levels of H1 transcripts were approximately 50% lower in these cell-lines compared to pCH5B.P-transfected cells (Table 4.1). This, however, represents a substantial increase over H1 mRNA levels in pCH. $\Delta$ H1/4-transfected cells. While H1 transcript levels did not seem to be fully restored to 'parental' levels, insertion of the H1-box in an upstream position resulted in a significant increase in H1 mRNA levels compared to the pCH. $\Delta$ H1/4 control containing a crippled H1-box (i.e. pCH. $\Delta$ H1/4). Pulse-labelling experiments indicate that the increase in steady-state mRNA levels (relative to CH. $\Delta$ H1/4 cell-lines) is reflected by an increase in the rate of H1 gene transcription and a concomitant reactivation of periodic control throughout the cell-cycle (Figure 4.10). Profiles of H1 and H2B gene transcription in the pCH.H1/UI-transfected cell-line are shown in Figure 4.10b. These results and those presented in Table 4.1, show that there is a considerable increase in the rate of H1 transcription during S-phase. Although an obvious pattern of S-phase regulatory control is evident, relative increases over 'basal' (+ aphidicolin) levels are reduced by about 50% when compared to the 'parental' H1 promoter in pCH5B.P (represented by the broken line). However, this represents a substantial increase (~ 8-fold) in the rate of transcription, at the time where gene activity is maximum, when compared to the same promoter (in pCH. $\Delta$ H1/4) without an H1-box upstream insertion (Table 4.1). Not

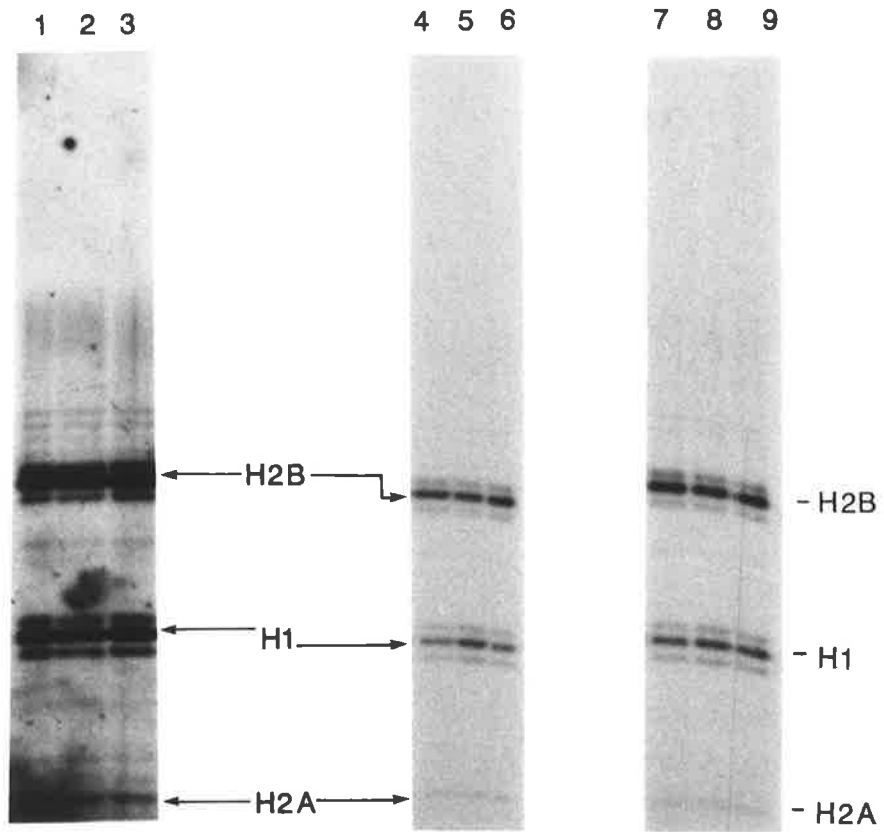
## FIGURE 4.9

### EFFECTS OF H1-BOX POLARITY REVERSAL AND UPSTREAM INSERTION ON STEADY-STATE H1 mRNA LEVELS

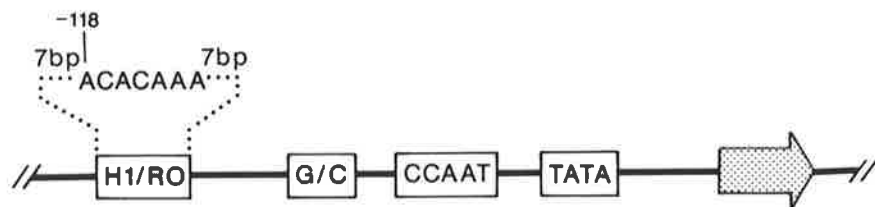
- A. HeLa cells in exponential growth phase were cotransfected with either pCH5B.P (lanes 1-3); pCH.H1/RO (H1-box in reverse orientation, Figure 4.3; lanes 4-6) or pCH.H1/UI (H1-box inserted upstream in the pCH. $\Delta$ H1/4-H1 promoter, Figure 4.3; lanes 7-9) and pSV2Neo. Following selection with G418, three clonal cell-lines, for each construct transfected, were selected for analysis. Steady-state H1, H2B and H2A mRNA, transcribed from exogenous-transfected genes, were detected by primer extension analysis (Section 2.3.15b). Amounts of RNA from each cell-line used for primer extension analysis were normalized so that signal intensities were approximately equal. Relative levels of transcripts in each cell-line are presented in Table 4.1
- B. Architecture of the H1 promoter with an H1-box in, (i) reverse orientation (pCH.H1/RO) and, (ii) inserted in an upstream position of the pCH. $\Delta$ H1/4-H1 promoter (pCH.H1/UI).



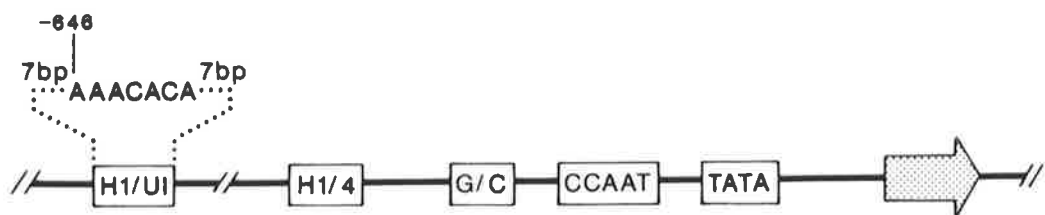
A.



B. (i)



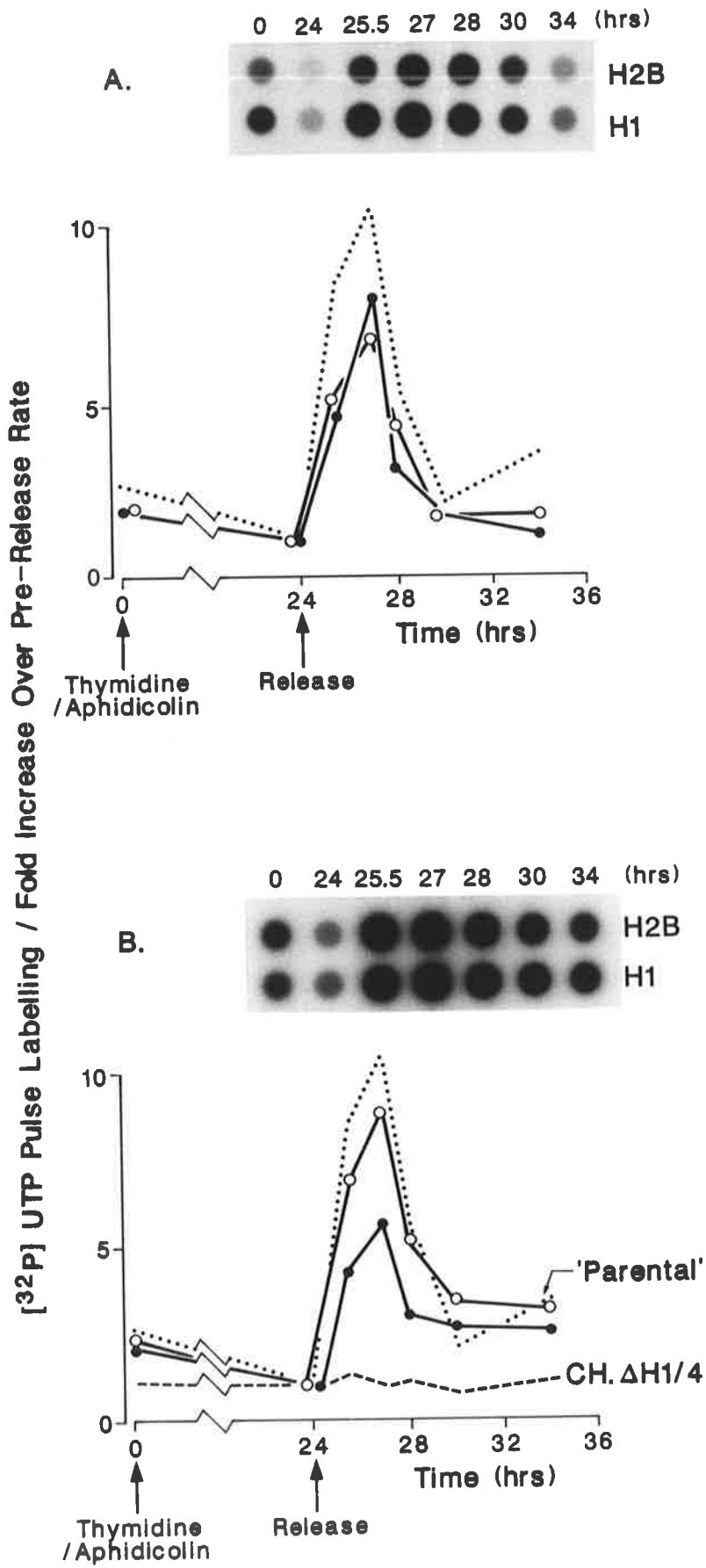
(ii)



## FIGURE 4.10

### EFFECTS OF H1-BOX POLARITY REVERSAL AND UPSTREAM INSERTION ON H1 GENE TRANSCRIPTION

HeLa cells in exponential growth phase were cotransfected with, (a) pCH.H1/RO (pCH5B.P 'parental' construct with H1-box in the reverse orientation, Figure 4.3 and 4.4) or, (b) pCH.H1/UI (H1-box inserted in an upstream position within pCH. $\Delta$ H1/4, see Figure 4.3 and 4.4) and pSV2Neo. Selection of clonal cell-lines, analysis of transcription rates, and presentation of data are described in the legend to Figure 4.7. Cell-lines used for transcription analysis of pCH.H1/RO and pCH.H1/UI are described in Figure 4.9 (lanes 6 and 7, respectively). Results are presented as the mean value (background subtracted) of two independent experiments. Broken lines represent transcription rates of, (a) 'parental' H1 gene (data from Figure 4.6), (b) 'parental' H1 gene and pCH. $\Delta$ H1/4-H1 gene (data from Figures 4.6, 4.8). Symbols: H1 transcripts (●), H2B transcripts (○).



only is periodic activation of transcription restored by insertion of a 'native' H1-box, but the kinetic profile of gene activity roughly parallels that of the 'parental' construct.

#### 4.2.2 Deletion and Substitution Analysis of the H2B Gene-Specific Element

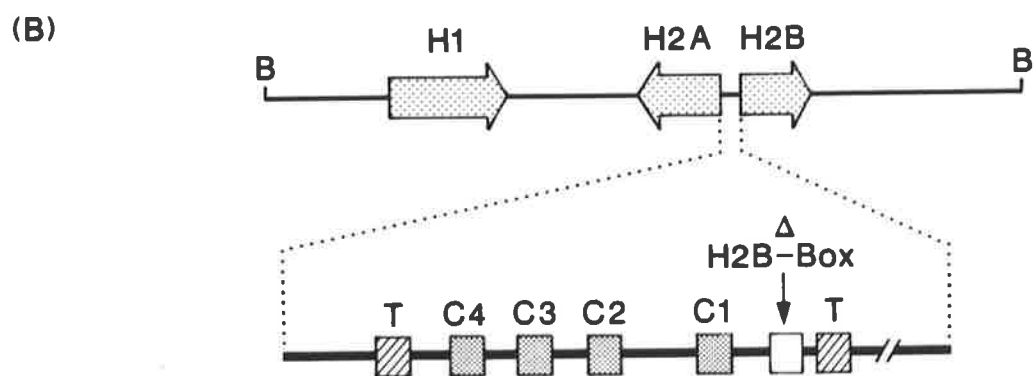
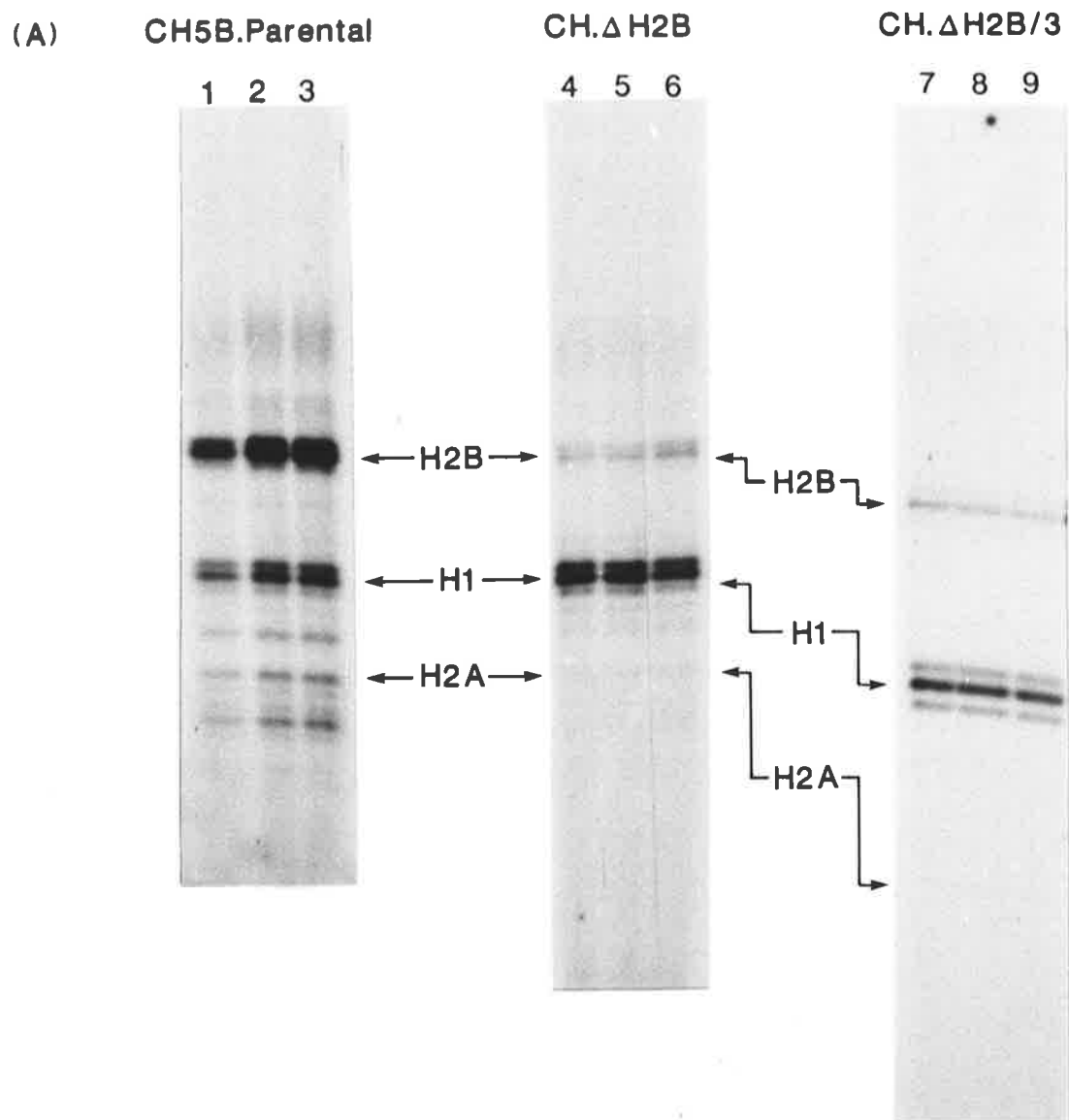
A similar approach employed to delineate the regulatory function of the H1-box was applied in an analysis of the putative transcriptional control mediated by the H2B-box. Apart from deletion of the element itself (pCH. $\Delta$ H2B), three base substitutions were introduced into a central, highly-conserved core region of the H2B-box (5'-CTCATTGTCAT-3') in the pCH5B.P construct, generating the new motif 5'-CTCAGTAGCCT-3' and consequently a new construct, pCH. $\Delta$ H2B/3 (Figure 4.3, 4.11). Deletion of the H2B-specific motif and introduction of point mutations resulted in a distinct decline in H2B steady-state mRNA levels for the three pCH. $\Delta$ H2B and three pCH. $\Delta$ H2B/3 clonal cell-lines assayed (Figure 4.11). Compared to relative transcript levels in 'parental' cell-lines, this represented a decrease ranging from approximately 8 to 15-fold (Table 4.1). Remarkably, transcript levels of the divergently transcribed H2A gene were also reduced, generally by a factor of 4-6 fold (see Figure 4.11). In contrast, levels of H1 mRNA were apparently invariant.

Transcription of H1 and H2B genes throughout the cell-cycle in a pCH. $\Delta$ H2B/3-transfected cell-line was analyzed by pulse labelling isolated nuclei. H1 gene transcription was elevated by approximately 9-fold during S-phase and retained a similar transcriptional profile as in previous experiments (see Figure 4.6a). In contrast, levels of H2B gene transcription were depressed by a factor of 3-4 fold, although a periodic pattern of transcriptional activation was retained (Figure 4.12).

## FIGURE 4.11

### DELETION AND BASE SUBSTITUTION OF THE H2B-BOX REDUCES LEVELS OF H2B AND H2A TRANSCRIPTS

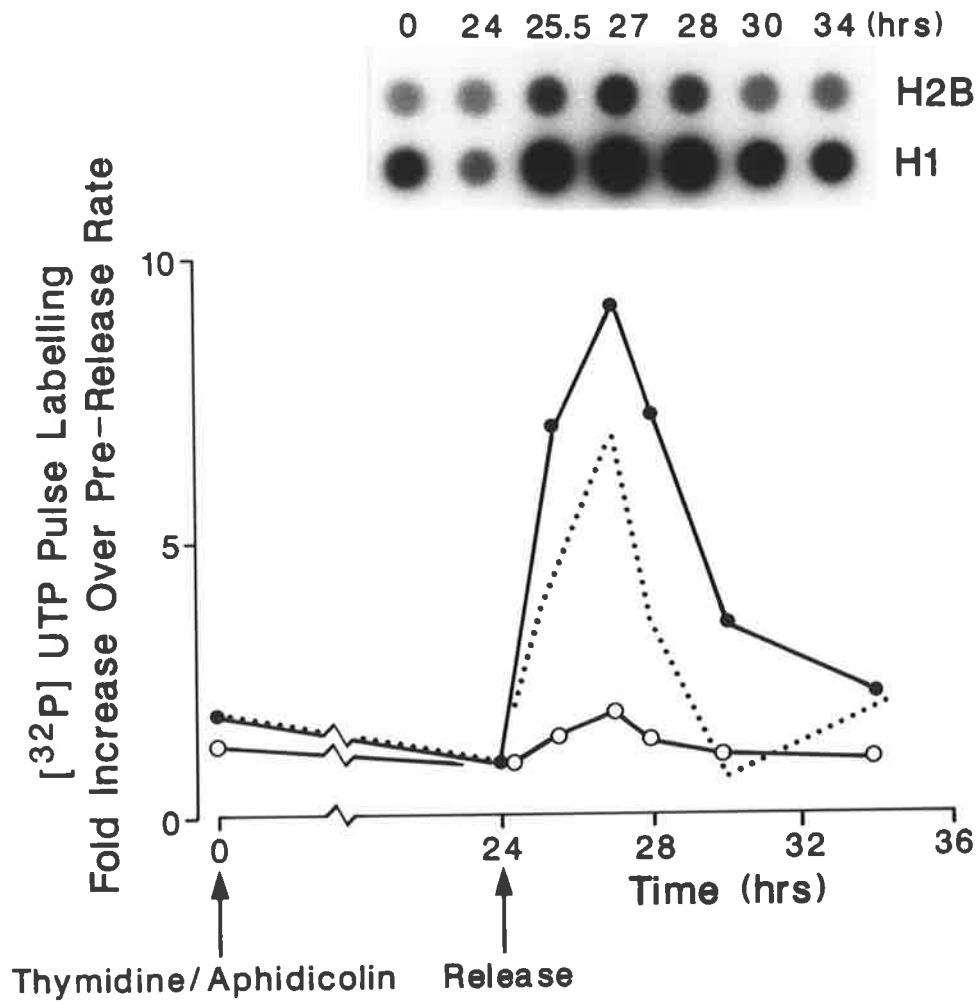
- A. HeLa cells in exponential growth phase were cotransfected with pCH5B.P (lanes 1-3), pCH. $\Delta$ H2B (H2B-box deleted, see Figure 4.3, lanes 4-6) or pCH. $\Delta$ H2B/3 (base substitutions in H2B box, see Figure 4.3, lanes 7-9) and, pSV2Neo. Following selection with G418, three clonal cell-lines, for each construct transfected, were selected for analysis. Steady-state H1, H2A and H2B mRNA, transcribed from exogenous-transfected genes, were detected in RNA samples from unsynchronized clonal cell-lines by primer extension analysis (Section 2.3.15b). Amounts of RNA from each cell-line used for primer extension analysis were normalized so that signal intensities were approximately equal. Relative levels of transcripts in each cell-line are presented in Table 4.1.
- B. The 5 kb Bam H1 fragment shown containing H1, H2A and H2B chicken histone genes was subcloned into pAT153, generating the construct, pCH5B.P. Positions of conserved elements in the H2A/H2B intergene region are shown, including the H2B-box on which deletion or substitution mutagenesis was performed. C1, C2, C3 and C4 represent 'CCAAT'-box motifs, and T represents the respective 'TATA'-box elements (see Figure 4.2).



## FIGURE 4.12

### BASE SUBSTITUTION WITHIN THE H2B-BOX REDUCES H2B GENE TRANSCRIPTION

HeLa cells in exponential growth phase were co-transfected with pCH. $\Delta$ H2B/3 (3 point mutations in H2B-box, Figure 4.3 and 4.11) and pSV2Neo. Selection of clonal cell-lines, analysis of transcription rates and presentation of data are described in the legend to Figure 4.8. Results are presented as the mean value (M13 background hybridization subtracted) of two independent experiments. Broken lines represent transcription rates of the 'parental' H2B gene in pCH5B.P (data from Figure 4.6). H1 transcripts ( $\bullet$ ), H2B transcripts ( $\circ$ ).





**TABLE 4.1**

LEVELS OF HISTONE mRNA LEVELS AND RATES OF GENE  
TRANSCRIPTION IN CELL-LINES TRANSFECTED WITH VARIOUS CONSTRUCTS

- A. Autoradiograms shown in Figures 4.4, 4.7, 4.9, 4.11 were scanned with a Zeineh laser densitometer (see Section 2.3.5) and relative amounts of H1/H2B, or H1/H2A transcripts determined for each cell-line transfected with different constructs.
- B. For each cell-line, rates of transcription (i.e. fold-increase over pre-release rates, see Figures 4.6, 4.8, 4.10, 4.12) at T=27 hrs (3 hours into S-phase) for H1 and H2B genes were determined and expressed as a ratio (H1/H2B). This ratio is then expressed relative to the same ratio in the 'parental' cell-line to indicate the net change in rates of transcription due to mutagenesis of H1 or H2B gene-specific promoter elements.

A. RELATIVE LEVELS OF STEADY STATE mRNA LEVELS

<u>Cell-line</u>	<u>H1/H2B</u>			<u>H1/H2A</u>		
	1	2	3	1	2	3
CH5B.P	0.20	0.32	0.29			
CH. ΔH1	0.01	0.01	0.02			
CH5B.P	0.32	0.51	0.30			
CH. ΔH1/4	0.02	0.02	0.01			
CH5B.P	0.58	0.72	0.45			
CH.H1/RO	0.49	0.62	0.38			
CH.H1/UI	0.30	0.39	0.47			
CH5B.P	0.20	0.28	0.29	4.11	6.32	7.74
CH. ΔH2B	3.46	2.87	2.68	21.71	17.59	14.73
CH. ΔH2B/3	2.57	2.05	1.73	23.6	24.72	25.50

B. RELATIVE RATES OF TRANSCRIPTION (T=27 hrs)

<u>Cell-Line</u>	<u>H1/H2B</u>	<u>Δ Relative to Parental Ratio</u>
CH5B.P	1.50	
CH. ΔH1	0.23	+ 6.5- fold
CH. ΔH1/4	0.10	+ 15-fold
CH.H1/RO	1.21	+ 1.2-fold
CH.H1/UI	0.78	+ 1.9-fold
		+ 7.8-fold (relative to CH. ΔH1/4)
CH/ ΔH2B/3	5.07	+ 3.4-fold

### 4.3 DISCUSSION

#### 4.3.1 An Intact H1-Specific Element is Required for Efficient Promoter Activity and Cell-Cycle Regulation

Experiments described in this chapter were designed to delineate the role of a highly conserved heptanucleotide element (5'-AAACACA-3'), ubiquitous to histone H1 genes, in directing transcription from the CH5B.P gene promoter. Evidence is presented from quantitative primer extension analysis which shows that deletion of this 7 bp core sequence element by site-directed mutagenesis significantly reduces steady-state levels of H1 gene-specific transcripts. In contrast, levels of mRNA transcribed from a closely linked H2B gene, which essentially served as an internal control, remained unaltered following deletion of this element.

Pulse-labelling experiments involving nuclei isolated from pCH5B.P permanently transformed cell-lines revealed that deletion or disruption of the H1-specific element not only reduces promoter activity but also is accompanied by a concomitant loss of cell-cycle regulation. In contrast, an H2B gene closely linked to the H1 gene retains a periodic pattern of transcriptional regulation. Although the evidence presented indicates that deletion of the H1-box eliminates cell-cycle regulated transcriptional activity of the CH5B.P ( $\lambda$ CH01)-H1 gene, it is possible other regulatory elements not defined by these experiments may also be involved in periodic activation of H1 genes during S-phase.

Transcriptional control of histone genes in yeast is believed to be a function of multiple control elements, either having positive or negative regulatory roles (Osley et al., 1986). The H1-box may in fact be analogous to the yeast promoter element proposed to activate histone gene transcription in a cell-cycle-dependent manner. Other experiments indicate that the H1-box can function in either orientation and from an upstream position. This latter result is particularly important as an

intact H1-box was shown to reactivate cell-cycle control, albeit at slightly reduced levels, on an H1 gene with a non-functional (mutated) H1-specific element. Together, these findings are consistent with the possibility that this heptanucleotide element is a cell-cycle specific transcriptional enhancer. Breeden and Nasmyth (1987) have recently reported such an element upstream of the yeast homothallism (HO) gene which can confer cell-cycle regulation when inserted upstream of a heterologous gene. Experiments are currently in progress to determine if the H1-box can confer cell-cycle regulatory control on a metallothionein gene in transfected cell-lines.

Previously, it had been shown that deletion of the H1 gene-specific element (AAACACA) apparently had no effect on H1 gene expression (Younghusband et al., 1986), whereas evidence presented here shows that deletion or mutation of this conserved sequence drastically decreases H1 gene expression in stably-transformed cell-lines. This discrepancy could be due to two reasons, (1) the transient assay system previously employed is possibly inappropriate for studying some aspects of gene control or, (2) the vector employed in these studies, which contained an SV40 enhancer element, may have 'masked' the deletion effect.

#### 4.3.2 The H2B-Box is Required for Efficient Promoter Activity and Influences the Expression of An Adjacent H2A Gene

Similar experiments to those already described have been performed for the H2B-box to determine if, (a) it is required for H2B promoter activity and, (b) to see if disruption of this element in any way influences the expression of a divergently transcribed H2A gene. Levels of steady-state H2B mRNA were reduced by about 10-fold following deletion and base substitution mutagenesis of the H2B-box. Remarkably, H2A mRNA levels decreased by approximately 3-4 fold indicating that the H2B-specific motif influences the expression of both genes. Pulse-labelling experiments in isolated nuclei indicate that initiation of

transcription is a major factor contributing to the fall in H2B mRNA levels following H2B-box mutagenesis. While the levels of H2B gene transcription were considerably reduced, a detectable transient increase in the rate of transcription was detected during S-phase. This is taken to indicate that promoter efficiency is severely reduced as a result of mutagenesis within the H2B-box even though cell-cycle specific mechanisms were still operative. The H2B-box presumably also operates by influencing H2A gene transcription. This would account for the decrease in H2A steady-state mRNA levels when the H2B-box is deleted. Perhaps this element is a general transcriptional enhancer which is operative in a number of unrelated genes and has no specific cell-cycle regulatory function. This infact is quite likely in view of evidence from studies on small nuclear RNA genes (Mattaj et al., 1985) and immunoglobulin genes (Falkner and Zachau, 1984) which also contain this element.

Sive and co-workers (1986) have also recently carried out work on the H2B-box using in vitro transcription studies. These investigations are consistent with those presented in this chapter as deletion and point mutagenesis of the H2B-box significantly reduced transcriptional activity of H2B gene templates in vitro. Studies presented here extend these observations by analysis of transcription throughout the cell-cycle in vivo and the effects on a divergently transcribed H2A gene. The observation that cell-cycle specific transcriptional control of the H2B gene is retained even when the H2B-box is mutagenized, further reinforces the concept that the H1-box is likely to be an S-phase specific element as mutagenesis of this sequence completely eliminates all remnants of periodic regulation. Both elements therefore apparently serve different functions.

The regulatory role of the H2B-box in H2A and H2B gene expression is compatable with there being a requirement for a highly conserved

intergene region and divergent arrangement. Strict evolutionary pressures may have maintained this arrangement due to a requirement for co-ordinate expression of H2A and H2B genes by the H2B-box. This overlapping/dual promoter concept (see Sturm, 1985) could also be extended to involve other spatially conserved sequence elements in the intergene region, such as multiple 'CCAAT' boxes (see Figure 4.2).

A recurring feature of experiments described in this chapter is the relative consistency obtained with expression of H1, H2A and H2B genes in different clonal cell-lines. This was noted for such parameters as cap site usage and relative levels of expression. In several cell-lines, however, genes were expressed at low levels and in such instances were generally excluded from further analysis. Overall, this implies that the integration site of linearized templates into the genome is not a critical factor in analysis of gene expression, so long as the appropriate cis-regulatory information is present in the integrating DNA. The same is true in most transgenic animal experiments (Palmiter and Brinster, 1986).

CHAPTER 5

IDENTIFICATION AND CHARACTERIZATION OF NUCLEAR

FACTORS WHICH INTERACT WITH THE H1 AND H2B

GENE-SPECIFIC PROMOTER ELEMENTS

## CHAPTER 5: IDENTIFICATION AND CHARACTERIZATION OF NUCLEAR FACTORS WHICH INTERACT WITH THE H1 AND H2B GENE-SPECIFIC PROMOTER ELEMENTS

### 5.1 INTRODUCTION

Having established that the H1 and H2B gene specific elements are essential for efficient activity of their cognate promoters (see Chapter 4) a series of experiments was performed aimed at identifying and characterizing sequence-specific nuclear factors which bind to these regulatory motifs. The principle object of work presented in this chapter, therefore, was to explore the possibility that putative trans-acting molecules interacting with these elements actually play a role in modulating histone gene transcription.

### 5.2 RESULTS

#### 5.2.1 Detection of Nuclear Factors which Specifically Bind to the H1 and H2B Gene-Specific Elements

Initially, an electrophoretic (gel mobility-shift) assay that exploits the high-affinity binding properties of site-specific DNA binding proteins which results in decreased electrophoretic mobility of DNA-protein complexes was employed (Section 2.3.23a). Two complimentary oligonucleotides (40-mers) were synthesized such that they span the H1-box and additional 5'- and 3'-sequences flanking this element in the genomic fragment CH5B.P (see Figure 4.4; Section 2.2.6b). Complimentary, [<sup>32</sup>P] end-labelled H1 oligonucleotides (designated H1-box '+' and '-', Section 2.2.6b) were annealed and the double-stranded form incubated with crude nuclear extract (2 µg) prepared from AEV ts34 cells, under the conditions described (in Section 2.3.23a), for 20 minutes at ambient temperature with, or without, varying amounts of non-specific, unlabelled polydI.polydC competitor DNA. Protein-duplex DNA complexes were separated from free DNA by electrophoresis through a



high-ionic strength polyacrylamide gel and visualized by autoradiography.

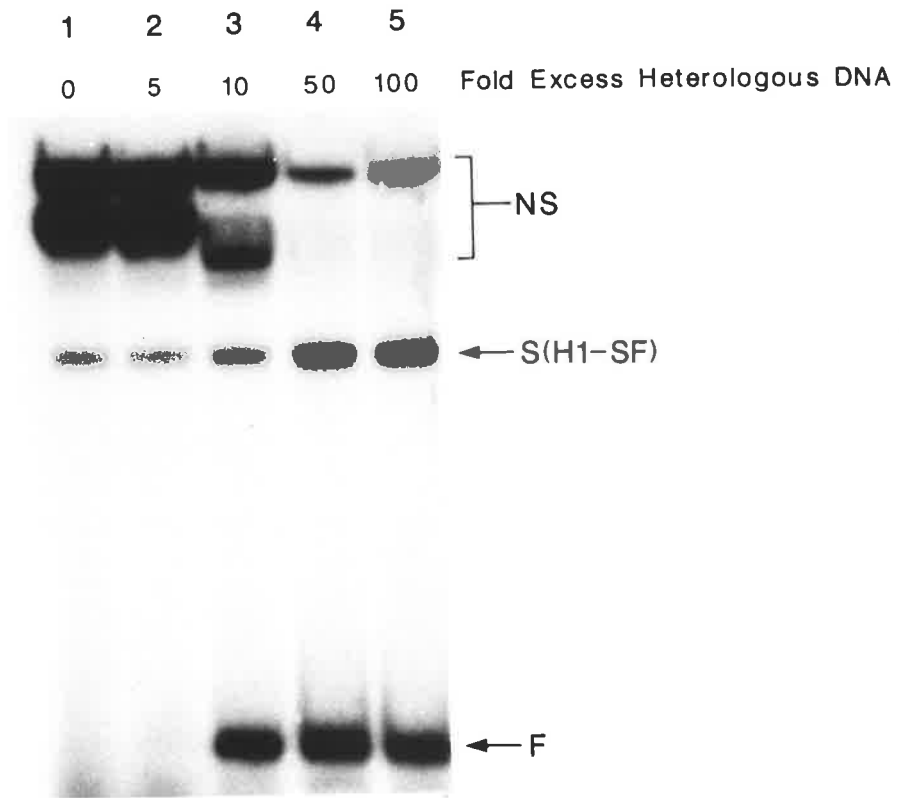
This assay allowed several slow-migrating DNA-protein complexes to be detected resulting from binding of factors present in the crude extract to [<sup>32</sup>P]-end-labelled H1 duplex oligonucleotides (Figure 5.1a; lane 1). Several very slow migrating bands detected in the absence of competitor DNA co-migrate with complexes formed when the H1 [<sup>32</sup>P]ds-oligonucleotides were mixed with purified histone protein (data not shown) and, are therefore likely to represent non-specific interactions involving histones and possibly some other nuclear components. The specificity of protein-DNA complexes formed in the presence of competitor DNA was stringently tested by incubating [<sup>32</sup>P]ds-oligonucleotides with crude extract in the presence of increasing amounts of unlabelled ('cold') heterologous or homologous DNA. Prior to this, the concentration of NaCl and proportions of other components included in the incubation mix were optimized (see legend to Figure 5.1) to maximize factor binding. Addition of increasing amounts of homologous DNA (unlabelled, H1 duplex oligonucleotide, 0-100 ng) effectively decreased the intensity of 'non-specific' upper bands and putative H1-specific complexes (Figure 5.1b). In contrast, similar assays using heterologous DNA (0-100 ng, polydI.polydC) abolished the upper 'non-specific' series of bands but had no apparent effect on the faster migrating 'H1-specific' complex (Figure 5.1a). These experiments were repeated with heterologous duplex-oligonucleotide and size fractionated/sonicated E. coli DNA with similar results (data not shown). These competition experiments revealed that a nuclear factor(s) present in the AEV ts34 crude nuclear extract displays a high degree of binding specificity and that resulting complex formation can not be titrated out using excess quantities (0-100 fold excess) of non-specific heterologous DNA. The major specific extract-DNA complex described will be referred to as H1-SF.

## FIGURE 5.1

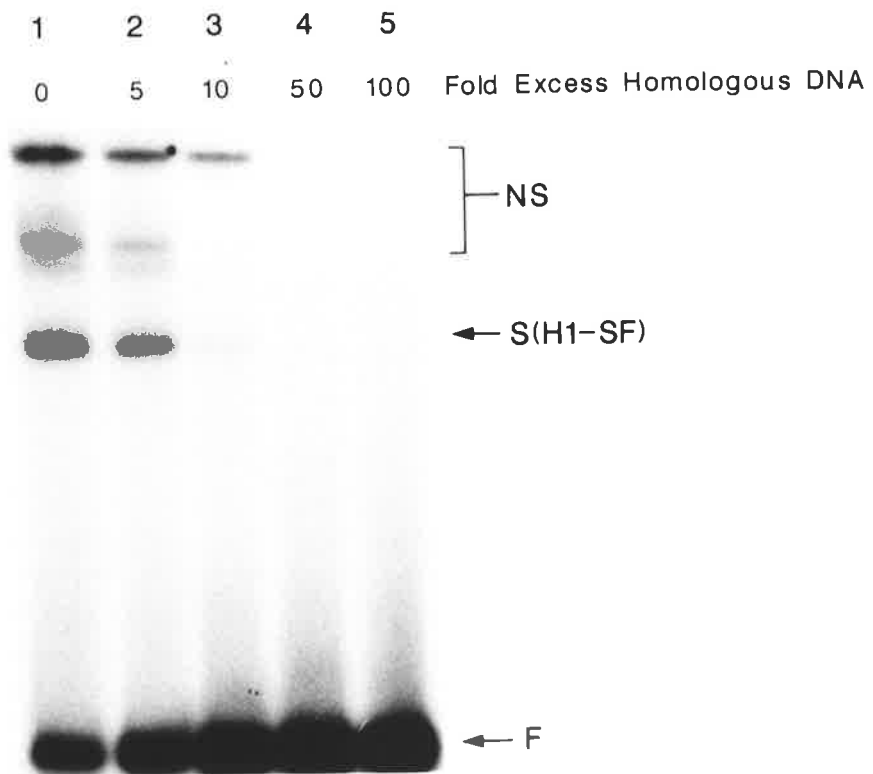
### DETECTION OF AN H1-BOX SPECIFIC NUCLEAR BINDING FACTOR BY COMPETITION ASSAY

- A. Crude nuclear extract (5  $\mu$ g) was incubated with 100 pg of annealed [ $^{32}$ P]-labelled H1-box oligomer (+/-; Section 2.2.6b) and various fold-excesses of unlabelled heterologous (polydI.polydC) DNA for 15 minutes in a 20  $\mu$ l cocktail consisting of 25mM Hepes pH 7.4, 25mM EDTA pH 7.4, 25mM DTT, 10% glycerol, 150mM NaCl, 0.5mM PMSF at room temperature and electrophoresed on a polyacrylamide gel (Section 2.3.12a). Unlabelled, heterologous DNA was omitted (lane 1), but added to binding reactions in 5, 10, 50 and 100-fold excess (lanes 2-5, respectively). Non-specific (NS) and specific (S, H1-SF = H1 specific factor) complexes formed between nuclear extract components and labelled DNA in addition to free (F), uncomplexed [ $^{32}$ P]duplex H1-oligomer were detected by autoradiography and are indicated. Similar results were obtained by including excess quantities of unlabelled heterologous ds-oligomer or size fractionated E. coli DNA instead of polydI.polydC (data not shown).
- B. 100 pg of annealed [ $^{32}$ P]-labelled H1-box oligomer-duplex (+/-; Section 2.2.6b) was incubated under the conditions described (in (A) above) except that various amounts of unlabelled homologous [ $^{32}$ P] duplex-oligomer competitor DNA was included instead of heterologous DNA. After incubation for 15 minutes, binding reactions were electrophoresed through a polyacrylamide gel (Section 2.3.12a). Non-specific (NS) and specific (S, H1-SF = H1 specific factor) complexes are indicated, in addition to free (F), uncomplexed DNA. Mobility-shift assays of the binding reactions are shown where a 0,5,10,50 and 100-fold excess of heterologous competitor DNA was included (lanes 1-5, respectively).

A.



B.



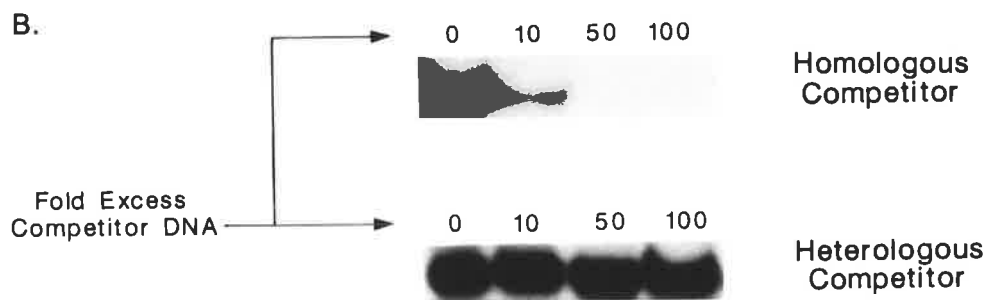
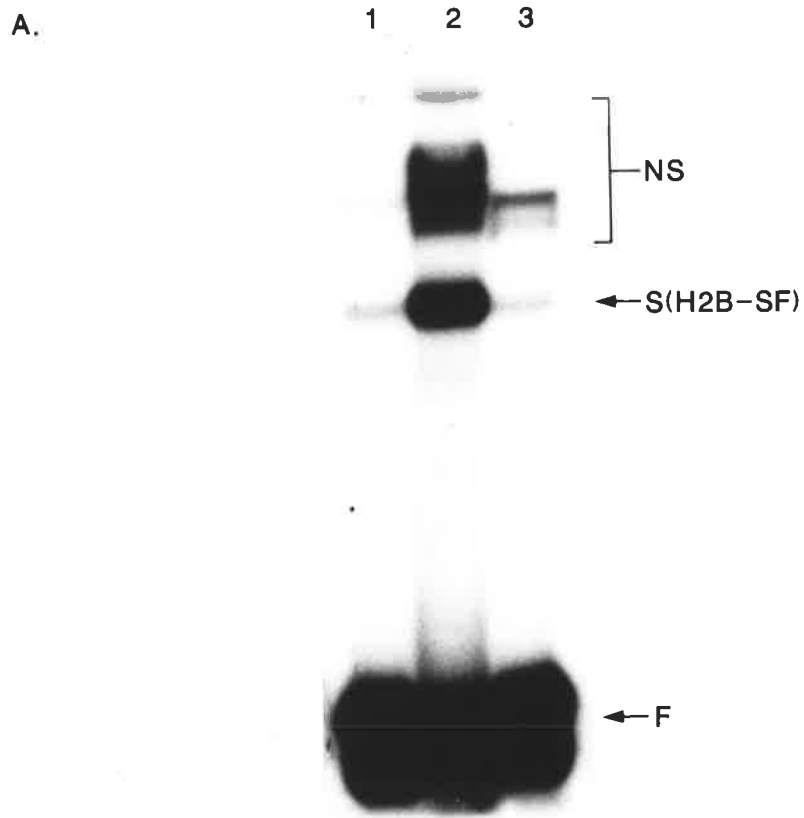
Similar competition assays were also performed using annealed complimentary oligonucleotides spanning sequences across the H2B-box in CH5B.P (see Figure 4.11). These oligonucleotides (40-mers), designated H2B-box '+' or '-', are represented in Section 2.2.6b. Binding of a nuclear component in AEV ts34 extracts, indicated by the presence of a slow-migrating (retarded) complex, to annealed [<sup>32</sup>P]H2B 40-mers was unaffected by the presence of excess heterologous oligonucleotide duplexes, sonicated E. coli DNA or polydI.polydC but, was substantially reduced by competition with increasing amounts of excess, unlabelled homologous synthetic 40-mer duplexes (Figure 5.2). Formation of a series of low-mobility complexes was abolished by inclusion of non-specific competitor DNA in the incubation mixture (data not shown). The H2B-specific nuclear factor (designated here as H2B-SF) was not the same as the factor which bound to ds-H1-box oligonucleotides as incubation with a 100-fold excess of unlabelled H1-box duplexes failed to reduce binding to [<sup>32</sup>P]-labelled H2B duplexes (result not shown). In addition, both nuclear factors required different NaCl concentrations for maximal binding. Similarly, H2B oligonucleotide failed to reduce H1-box-nuclear factor binding, as indicated by the gel mobility-shift assay (data not shown).

In order to determine the nature of the H1 and H2B-specific factors, a series of gel mobility-shift assays were carried out with nuclear extracts which were either RNAase A digested, proteinase K treated or heated to 70°C for 5 minutes before incubation with [<sup>32</sup>P]-labelled DNA. Results indicate that the H1 specific factor present in nuclear extracts was sensitive to proteinase K and heat treatment but insensitive to RNAase A activity (Figure 5.3), indicating that the binding factor is a protein(s) or requires a proteinaceous component. Similar results were obtained for the binding of H2B-SF following treatment of extracts by these procedures (data not shown).

## FIGURE 5.2

### DETECTION OF AN H2B-BOX SPECIFIC NUCLEAR BINDING FACTOR BY COMPETITION ASSAY

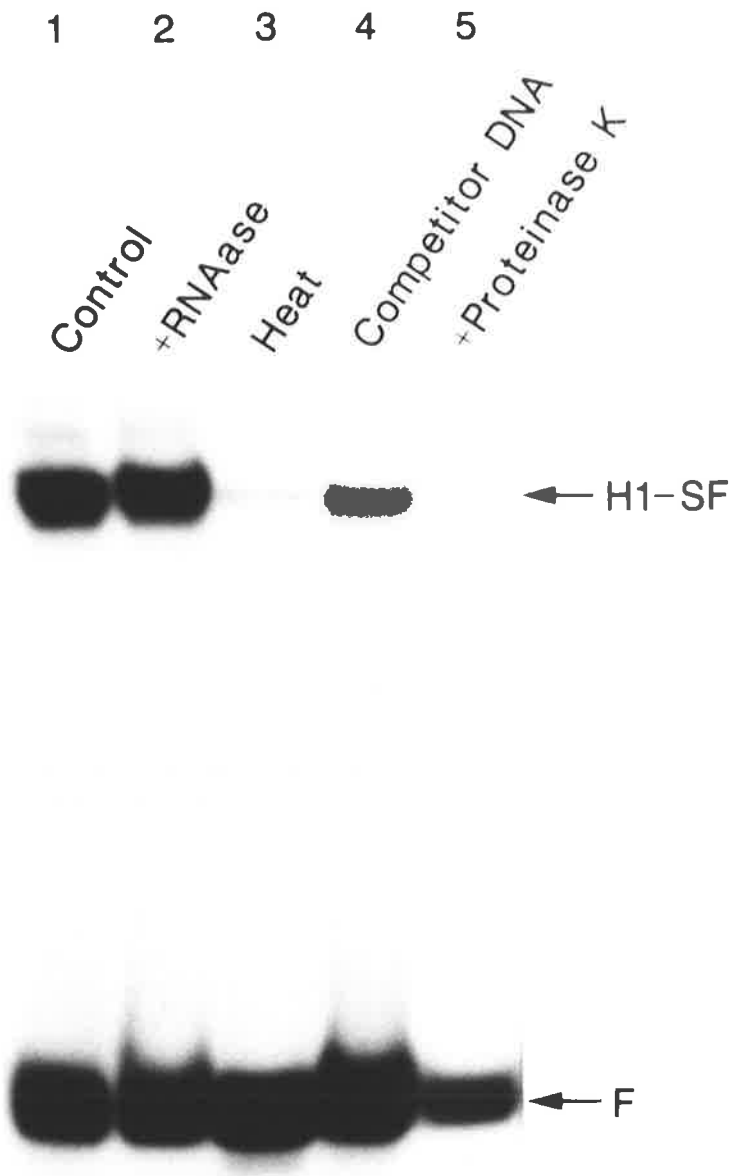
- A. Each incubation mix consisted of 25mM Hepes pH 7.4, 25mM EDTA pH 7.4, 25mM DTT, 10% glycerol, 200mM NaCl, 0.5M PMSF in a 20  $\mu$ l volume. In addition, various amounts of annealed [ $^{32}$ P]-labelled H2B-box oligomer (+/-; Section 2.2.6b), unlabelled homologous DNA and crude nuclear extract were added as follows: lane 1, 0.1  $\mu$ g extract, 300 pg [ $^{32}$ P]DNA, 10 fold excess (3 ng) unlabelled H2B duplex oligonucleotide; lane 2, 5  $\mu$ g extract, 300 pg [ $^{32}$ P]DNA; lane 3, 0.1  $\mu$ g extract, 300 pg [ $^{32}$ P]DNA. Incubation was for 15 minutes at room temperature. DNA-nuclear extract complexes were detected by electrophoresing the binding reaction through a polyacrylamide gel (Section 2.3.12a) and autoradiographic exposure. NS = non-specific complexes; S (H2B-SF) = H2B specific factor and; F = free label.
- B. Gel mobility-shift competition assays were performed with various amounts (0-100 fold excess) of homologous (upper panel) or heterologous polydI.polydC DNA (lower panel) under the conditions described above (Figure legend, part A) with 5  $\mu$ g extract, 300 pg [ $^{32}$ P] DNA. Similar experiments employing heterologous oligonucleotide duplex and *E. coli* DNA were also performed giving similar results to when polydI.polydC was included (data not shown). The DNA-protein complex shown is the H2B-SF retarded complex shown in part A.



### FIGURE 5.3

#### BIOCHEMICAL CHARACTERIZATION OF THE H1-SF

Nuclear extract (5  $\mu$ g) was incubated under the conditions described in the legend to Figure 5.1 in the presence of a 100-fold excess (10 ng) of polydI.polydC (lane 1). Before initiation of incubation reactions, the following procedures were involved: lane 2, extract was treated with 100  $\mu$ g/ml RNAase A at 37°C for 15 minutes; lane 3, extract was heated at 70°C for 5 minutes; lane 4, extract pre-incubated with 1 ng unlabelled homologous DNA duplex for 5 minutes before addition of labelled DNA; lane 5, extract treated with 200  $\mu$ g/ml proteinase K at 42°C for 15 minutes. H1-SF represents binding of the H1-specific factor and F represents free label.





## 5.2.2 Sequence Requirements for Nuclear Factor Binding to H1 and H2B

### Oligonucleotide Duplexes

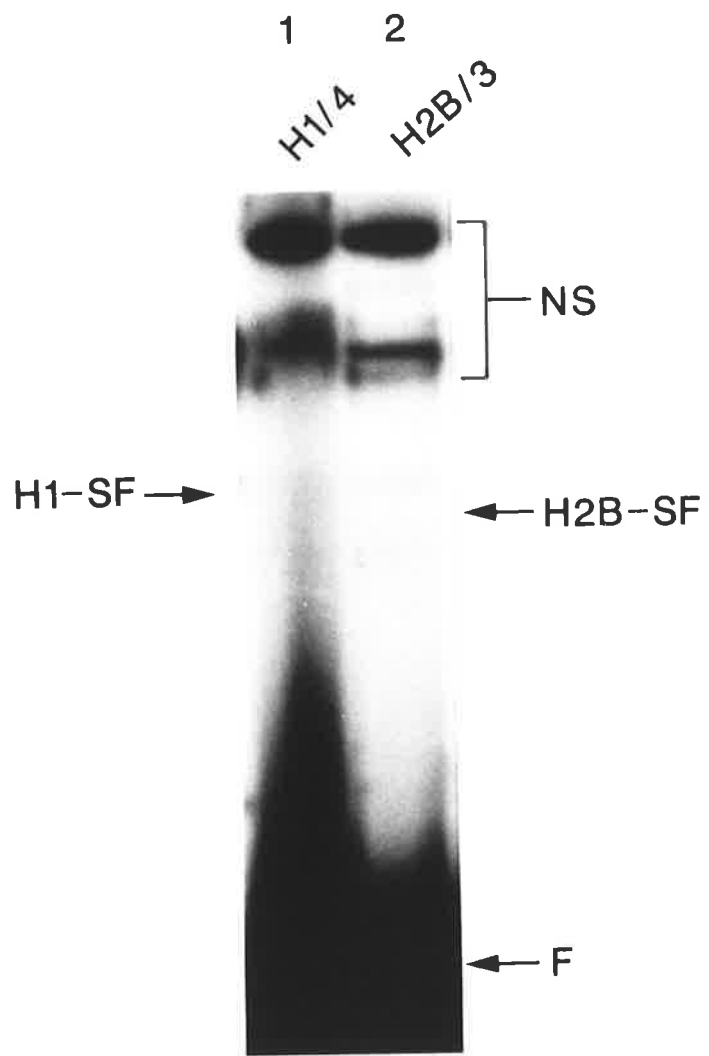
The requirements for an intact, conserved H1-box element in binding the sequence-specific nuclear factor (see 5.2.1) was tested, again using gel mobility-shift assays to monitor complex formation. Oligonucleotide duplexes containing four base substitutions in the highly conserved 5'-AAACACA-3' motif (see 2.2.6b), corresponding to the same regions and base substitutions made in the construct pCH. $\Delta$ H1/4 (Section 4.2; Figure 4.3) used for expression studies, were tested for their ability to bind the H1-specific nuclear factor, H1-SF, when annealed. Equivalent base substitutions in this expression construct markedly reduced transcription of the H1 gene in vivo (see Chapter 4). Effectively, this series of base changes resulted in loss of factor binding (Figure 5.4), presumably due to the sequence-specific target recognition site being destroyed.

To confirm that the H1-box was the binding site for the H1-specific nuclear factor, DNAase I protection assays ('footprints') were performed. This assay has the advantage over mobility-shift experiments as it enables accurate mapping of the DNA sequences involved in factor binding by resolving DNAase I-protected zones along an end-labelled restriction fragment. A DNA restriction fragment of approximately 600 bp was used, spanning the entire CH.01 promoter region (Coles, 1985) which contains sequences required for efficient and accurate expression of H1 genes in tissue culture cells and Xenopus oocytes (Youngusband et al., 1986). This fragment was [<sup>32</sup>P]end-labelled at the Hind III terminus (Figure 5.5), incubated with heterologous polydI.polydC competitor DNA plus varying amounts of nuclear extract and digested with DNAase I. Digested products were separated on a 6% polyacrylamide urea sequencing (analytical) gel and visualized by autoradiography. One obvious protected region was detected approximately 90 bp from the

## FIGURE 5.4

### BASE SUBSTITUTIONS WITHIN THE H1 AND H2B-BOX ELIMINATES NUCLEAR FACTOR BINDING

Nuclear extract (5  $\mu$ g) was incubated with 100 pg of, (1) base substituted [ $^{32}$ P]H1-box duplex oligomer, H1/4 (under the conditions described in the legend to Figure 5.1) or, (2) base substituted [ $^{32}$ P]H2B-box duplex oligomer, H2B/3 (under the conditions described in the legend to Figure 5.2). Oligonucleotides used in these experiments are shown in Section 2.2.6b. Non-specific binding activity (NS) and free duplex oligonucleotide (F) are indicated in addition to the expected positions of H1-SF and H2B-SF. The autoradiogram represents a longer exposure than normally required to easily detect H1-SF (H1-specific factor) /H2B-SF (H2B specific factor) complexes with DNA so as to ensure that binding activity to mutated H1 and H2B-box sequences was absent. Identical nuclear extracts used for experiments shown in Figure 5.1 and 5.2 were also used for these binding experiments.



Hind III terminus of the fragment (Figure 5.5). This protected region corresponds to a 15 bp region encompassing the H1-specific element and some flanking sequences. Increasing amounts of extract (0-1  $\mu$ g) led to increased DNAase protection (Figure 5.5). Furthermore, inclusion of a 100-fold molar excess of 'cold' homologous H1-promoter restriction fragment, reversed the extract-dependent DNAase protection in the specified region, consistent with a titration effect resulting in nuclear factor being competed away from labelled H1 promoter DNA. Other protected zones were detected (one of these corresponds to the G/C-box) although their exact significance is not known. Prior to the addition of [<sup>32</sup>P]-labelled promoter fragment, extract was pre-incubated with polydI.polydC competitor DNA to ensure binding specificity.

The requirement for a conserved H2B-box motif in binding the nuclear factor which forms specific complexes with 'H2B-duplexes' (H2B-SF; see Section 5.2.1) was addressed by introducing 3 point mutations into the highly conserved H2B-box core region. Oligonucleotides used in this experiment were designated  $\Delta$ H2B-box(3) '+' or '-' and the sequences of these oligomers are represented in Section 2.2.6b. Analogous to the H1-box base substitution binding experiments, sequences of these synthetic oligomers are identical to the corresponding region of the H2B promoter in the mutated expression construct, pCH. $\Delta$ H2B/3 (described in Chapter 4). Substitution of bases in this region which has already been shown to be essential for efficient promoter activity (in Chapter 4), eliminates nuclear factor binding to mutated H2B 'promoter' sequences completely (Figure 5.4).

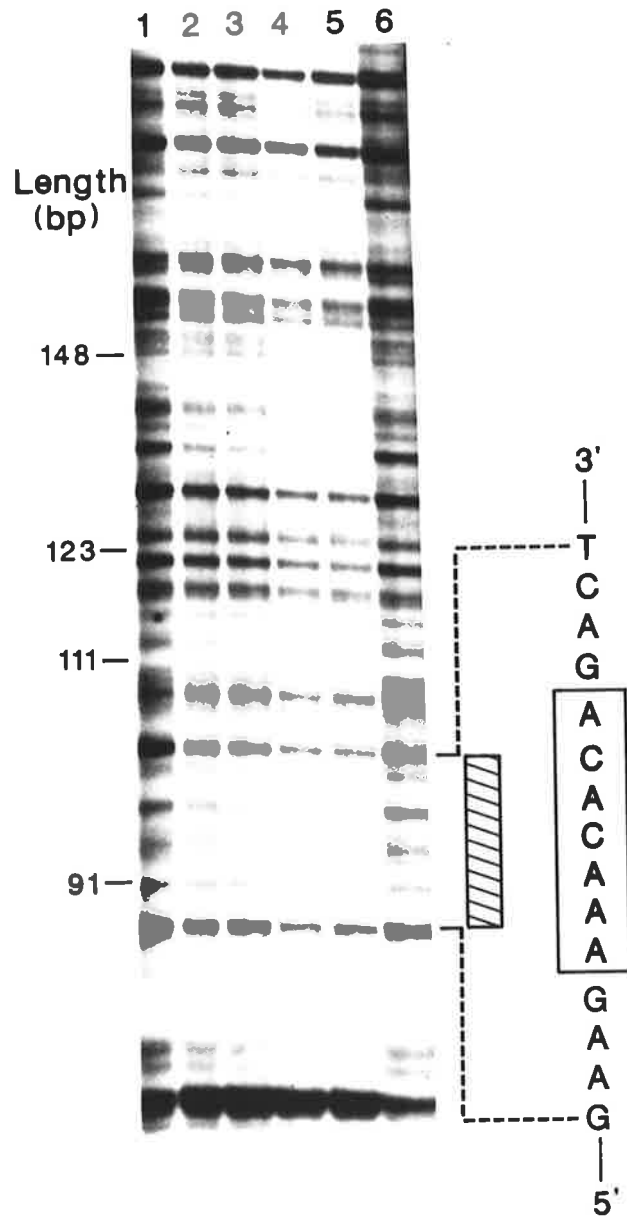
### 5.2.3 Enrichment of H1- and H2B-Specific Binding Factors

Characterization of the nuclear factor(s) which displays specific binding activity to the region of DNA encompassing the H1- and H2B-specific elements, involved two enrichment steps. Crude extract (approximately 750 mg) was initially passed through a Sephacryl S300 gel

## FIGURE 5.5

### PROTECTION OF THE H1-BOX FROM DNAase DIGESTION BY A NUCLEAR FACTOR

A 600bp H1 Hind III/Bam H1 promoter fragment was end-labelled at the Hind III terminus using T4 DNA polymerase (Section 2.3.10b). 2 ng of this fragment ( $\sim 10^7$  Cerenkov cpm/ $\mu$ g DNA) was incubated under the conditions described (Section 2.3.10b) with various amounts of nuclear extract as follows:- lane 1, no extract; lane 2, 0.25  $\mu$ g; lane 3, 0.5  $\mu$ g; lane 4, 0.75  $\mu$ g; lane 5, 1  $\mu$ g and, lane 6, 1  $\mu$ g plus a 50-fold excess of unlabelled promoter fragment. Following DNAase digestion (see Section 2.3.12b), digested products were electrophoresed on a polyacrylamide sequencing gel (Section 2.3.4b(iii)) and visualized by autoradiography. The protected zone and corresponding promoter sequences along the restriction fragment were determined by running Maxam and Gilbert sequencing reactions (Section 2.3.12b) alongside the tracks (not shown). Positions of Hpa II-digested pBR322 markers are indicated in addition to the position of the H1-box in the promoter fragment relative to the [ $^{32}$ P]-labelled Hind III terminus.



filtration column. Fractions eluting off the column (see Figure 5.6) were collected and assayed for binding activity using the mobility-shift assay (Figure 5.7). By calibrating the column with molecular weight standards, it is estimated that the H1-SF elutes at a molecular weight of ~ 80,000-100,000 under non-denaturing conditions while H2B-SF elutes at between 65,000-75,000 (Figure 5.6). Fractions containing binding activity were pooled and DNA-target specificity tested by competition assay with excess amounts of homologous and heterologous DNA (data not shown).

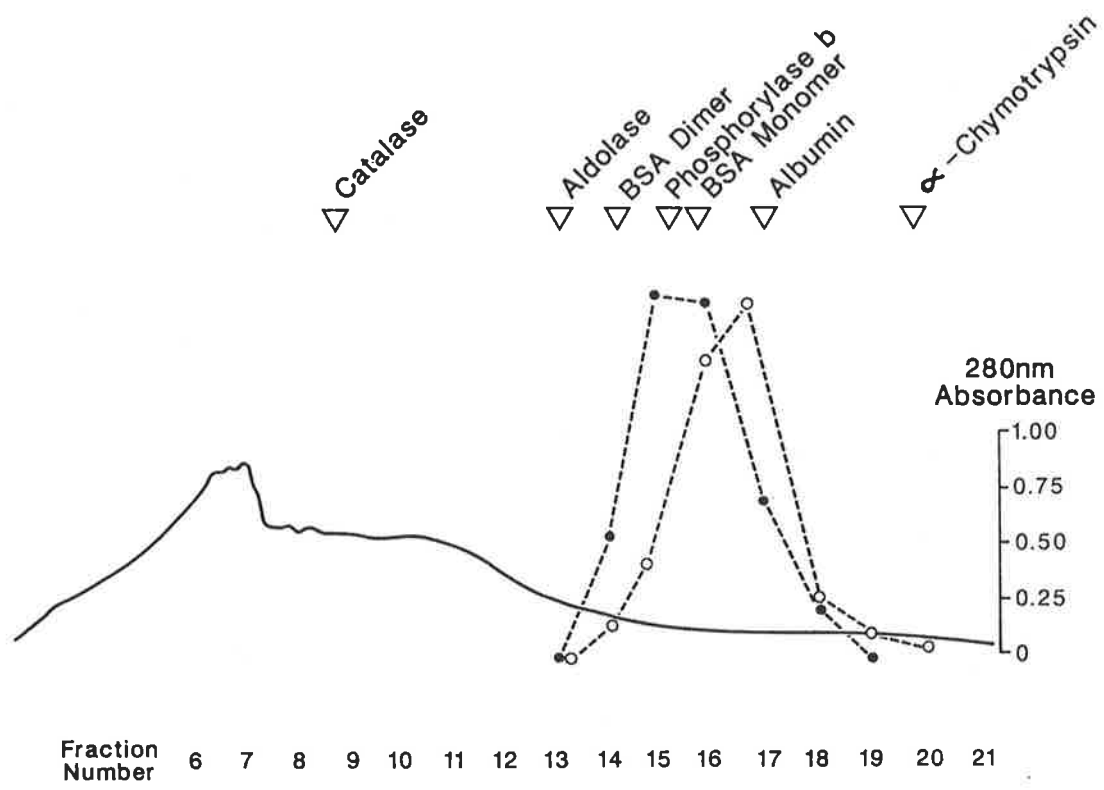
Following confirmation of nuclear factor-DNA binding specificity in each case by competition assay, pooled Sephacryl fractions were chromatographed through a DEAE-cellulose anion exchange column. Flow-through fractions, in addition to fractions eluting at 250 mM and 1 M-KCl were assayed for binding activity. A number of consecutive flow-through fractions contained major binding-activity while no such activity was detected in protein fractions eluting under higher ionic strengths (data not shown). Following anion exchange chromatography, only a single retarded band was detected in gel mobility-shift assays, even in the absence of heterologous competitor DNA, indicating that the H1 and H2B-specific factors had been purified from most of the non-specific DNA binding protein (Figure 5.8). Competition assays indicated that these retarded complexes were the result of specific interactions with H1 and H2B-oligonucleotide duplexes, respectively (data not shown). Scintillation counting of specific DNA-protein bands excised from gels showed that binding activity of the H1-specific nuclear protein had increased by a factor of 15 and 16-fold following Sephacryl and ion-exchange purification steps, respectively (Table 5.1 and see Figure 5.8). Similar enrichment of the H2B factor was achieved following these purification steps. Gel mobility-shift binding assays were carried out in each case using an excess of specific duplex [<sup>32</sup>P]-

## FIGURE 5.6

### ENRICHMENT OF SEQUENCE-SPECIFIC NUCLEAR FACTORS BY GEL FILTRATION

Crude nuclear extract (~ 750 mg protein) was loaded on a Sephacryl S-300 gel filtration column (Section 2.3.22a) and 5 ml fractions were collected. A typical  $A_{280}$  protein elution profile is shown (solid line, —) with respect to fraction number and standard molecular weight protein markers (Sigma) which were used to calibrate the column. They were as follows:  $\alpha$ -chymotrypsin (MW 25,000), albumin (MW 45,000), bovine serum albumin monomer (MW 66,000), phosphorylase b (MW 97,000), bovine serum albumin dimer (MW 130,000), aldolase (MW 158,000), catalase (MW 232,000). Superimposed on the protein elution profile are the relative levels of the H1-SF (--●--) and H2B-SF (--○--) collected from the corresponding fractions, as determined by autoradiographic densitometry (Section 2.3.5) of the gel mobility-shift assays shown in Figure 5.7.



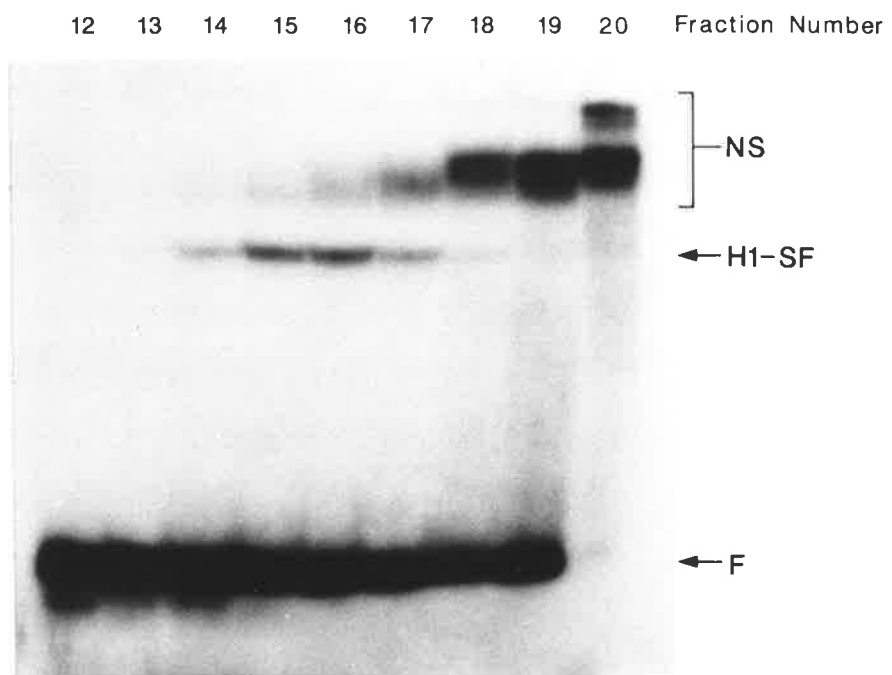


## FIGURE 5.7

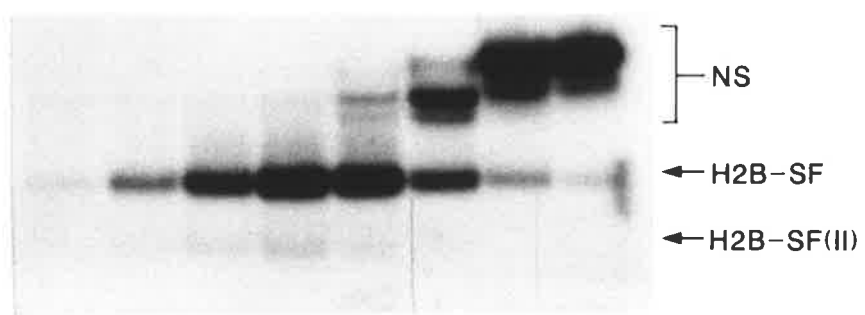
### MONITORING ELUTION OF H1-SF AND H2B-SF FOLLOWING GEL FILTRATION BY GEL MOBILITY-SHIFT ASSAY

- A. Gel mobility-shift assays were performed under conditions described in the legend to Figure 5.1 for detection of H1-SF except that 0.5  $\mu$ g extract protein was used. Lanes are numbered corresponding to the extract fraction used in the corresponding binding assays (see Figure 5.6). Positions of non-specific (NS), free-oligomer (F) and H1-SF-DNA complex (H1-SF) are indicated.
  
- B. Identical protein fractions (as in (A)) were used to monitor H2B-SF elution from the Sephacryl column. Binding reactions were as described (Figure legend 5.2). Other details are described above (legend part (A)). Fraction numbers in part A also correspond to the underlying tracks in part B. A second, low-intensity complex (HSB-SFII) was also detected.

A.



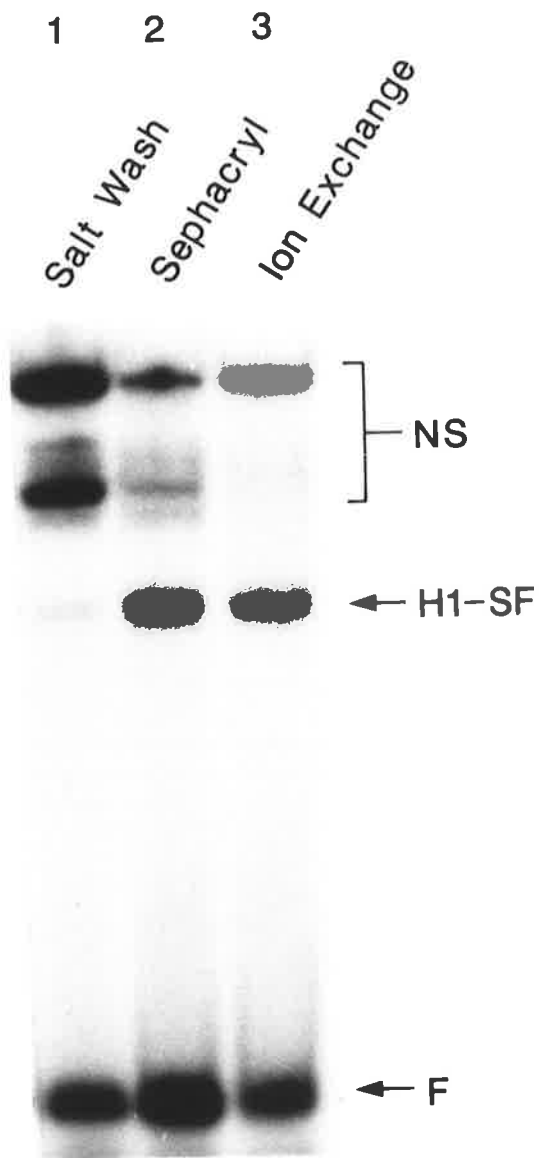
B.



## FIGURE 5.8

### INCREASE IN H1-SF BINDING ACTIVITY FOLLOWING GEL FILTRATION AND ION EXCHANGE CHROMATOGRAPHY

Nuclear extract from, (1) chromatin salt washes (crude, 5  $\mu$ g), (2) Sephacryl fractions (see Figure 5.6, 5.7; 3.5  $\mu$ g) or, (3) ion exchange fractions (0.2  $\mu$ g) were incubated with [ $^{32}$ P]H1-oligonucleotide duplex under the conditions described in the legend to Figure 5.1. Positions of specific (H1-SF) and non-specific (NS) complexes are shown in addition to free DNA (F). Similar quantitative experiments were performed in the presence of heterologous polydI.polydC (see Figure 5.3) to eliminate non-specific binding. Results for this (specific-activity of binding) are shown in Table 5.1. This figure emphasizes both the enrichment of H1-SF after each step and, separation of this binding activity from non-specific components. The autoradiogram represents a short exposure so as to highlight relative intensities of each band.



oligonucleotide to ensure a quantitative assessment of complex formation. In semi-purified extract preparations, a second low-intensity complex (H2B-SFII) interacting with H2B-oligonucleotides was sometimes detected (see Figure 5.7b). This interaction was sequence-specific as judged by competition assay (data not shown).

A final enrichment step attempted was to purify the H1-SF to near homogeneity by employing DNA-affinity column chromatography (Section 2.3.22c). This essentially involved coupling size-fractionated concatamerized H1-box oligonucleotide duplexes (~ 0.5-1 kb average length) to a cyanogen bromide-activated Sepharose gel matrix and performing chromatography under conditions described for binding assays (see legend to Figure 5.1). Protein binding to the column was eluted with binding buffer plus 1M NaCl. Eluting protein fractions were collected and analyzed for enrichment of H1-SF by gel mobility-shift assays. Very little H1-SF was detected in the 'bound' fraction but instead, eluted preferentially in the flow-through fractions directly after being applied to the column (Figure 5.9). Therefore, the affinity column failed to bind H1-SF which was applied in ion-exchange-purified fractions.

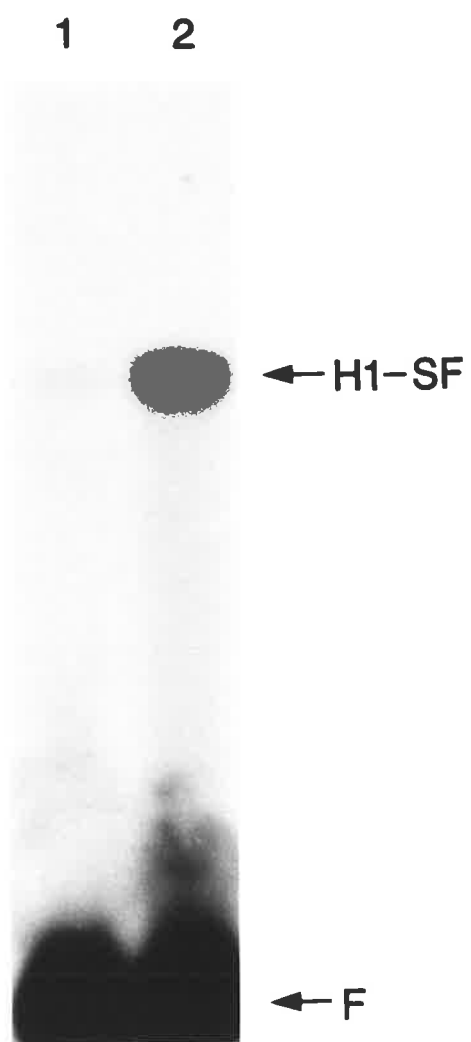
#### 5.2.4 Levels of H1-Specific Nuclear Factor Fluctuate Throughout the Cell-Cycle

As it is well established that transcription of histone H1 genes is transiently elevated during S-phase of the cell-cycle and that the H1-specific hexanucleotide sequence is required for this periodic regulation, it was of interest to determine if levels of the H1-specific factor also fluctuate. Levels of this H1-box sequence-specific factor varied considerably in crude and partially purified nuclear extracts prepared from an AEV ts34 chicken early erythroid cell-line at different stages of the cell-cycle. Analysis of H1-SF levels by the gel mobility-shift assay indicate an increase of 10-15 fold over levels in

## FIGURE 5.9

### GEL MOBILITY-SHIFT ANALYSIS OF PROTEIN FRACTIONS ELUTING OFF A DNA/H1-BOX AFFINITY COLUMN

A DNA affinity column consisting of cyanogen bromide-activated Sepharose CL-2B coupled to concatamerized H1-box duplex oligonucleotides was prepared and used for chromatography as described (Section 2.3.22c). Pooled fractions containing specific binding activity (6 mg total protein) following ion exchange chromatography were pre-incubated with polydI.polydC and applied to the DNA affinity column in ~ 20 ml of binding buffer (25 mM Hepes pH 7.5, 1 mM EDTA, 5 mM DTT, 10% glycerol, 0.5 mM PMSF, 150 mM NaCl). After flow-through fractions were collected, column-bound protein was collected by applying 10 ml of elution buffer (same as binding buffer except that 1 M NaCl was included) and dialyzed extensively against application buffer. 0.1 µg of pooled flow-through and 1 M NaCl-eluting protein fractions were incubated with [<sup>32</sup>P]H1-box duplex oligonucleotides (as described in the legend to Figure 5.1), electrophoretically resolved on a polyacrylamide gel (Section 2.3.12a) and complexes detected by autoradiography. Lane 1, 1 M NaCl eluting fractions; Lane 2, flow-through eluting fractions. The H1-SF binding complex (H1-SF) and free/uncomplexed DNA are indicated.





the absence of DNA replication (+ aphidicolin, G<sub>2</sub>-phase; Figure 5.10). Fluctuating levels of this factor during the cell-cycle therefore appear to reflect the transcriptional activity of the CH5B.P-H1 gene which is most active in early S-phase (Chapter 4). In contrast, levels of the H2B-SF nuclear factor which binds to a histone H2B promoter element (5'-CTGATTTGCATAG-3') was found to be invariant at corresponding times of the cell-cycle. Presumably, levels of H1-SF detected in gel assays, bears a true representation of H1 factor binding levels at corresponding stages of the cell-cycle in vivo. The possibility that fluctuating levels of H1-SF is due to some general artifact is unlikely as H2B-SF levels in the same extracts are essentially invariant.

### 5.3 DISCUSSION

#### 5.3.1 Sequence-Specific Nuclear Factors Bind to the H1 and H2B Gene-Specific Promoter Elements

The work presented in this chapter specifically focussed on the identification and characterization of nuclear factors within cellular extracts which bind to H1 and H2B gene specific promoter elements with high affinity and sequence selectivity. Both of these sequence motifs have already been shown to be essential for promoter efficiency and, in the case of the H1-box, for periodic activation of transcription during S-phase (see Chapter 4).

Initial studies employing a gel retardation assay detected several retarded complexes between duplex-oligonucleotide, spanning the H1-box plus flanking promoter sequences, and crude extract components. Competition assays with excess amounts of unlabelled, heterologous or homologous DNA indicated that one of these retarded complexes involved a nuclear component binding with sequence-specificity. Low-mobility complexes detected in the absence of competitor DNA probably involved histone protein and perhaps some other nuclear components. Size

## FIGURE 5.10

### LEVELS OF H1-SF BINDING TO THE H1-BOX ARE COUPLED TO S-PHASE

Crude nuclear extract was initially prepared from AEV ts34 cells at various stages of synchronization (section 2.3.11a(i)). Specific complex formation in each of the cell-cycle extracts was determined by electrophoretic mobility shift assay (see insets), excising bands from the gel and scintillation counting. Binding conditions are described in legends to Figures 5.1 and 5.2 and included a 100-fold excess of polydI.polydC. Relative complex formation is expressed as the fold change over pre-release (20 hrs + aphidicolin) levels. Superimposed is the profile of H1 gene transcriptional activity at corresponding stages of the cell-cycle (from Figure 3.7; ---). T = 42 hrs represents the time at which cells have progressed through the cell-cycle to a second S-phase (see Chapter 3). Symbols; H1-SF = top row of bands (●); H2B-SF = bottom row of bands (○). Similar results were also obtained when partially purified (Sephacryl column fractionated) extracts were used (data not shown).

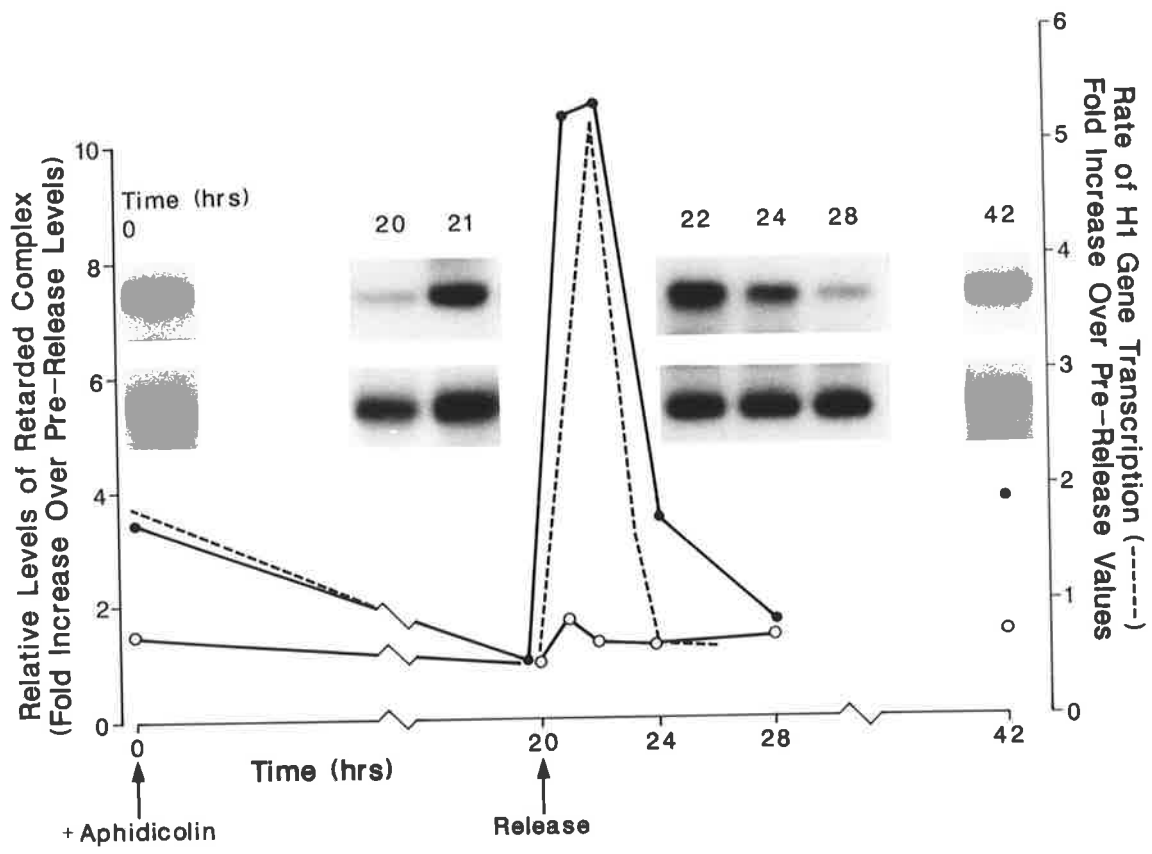


TABLE 5.1

PROTOCOL FOR THE ENRICHMENT OF NUCLEAR FACTORS FROM  
NUCLEAR EXTRACTS

Enrichment of salt-extracted nuclear factors involved gel filtration through Sephacryl S-300 and then ion exchange chromatography using DEAE-Sephrose (CL-6B). The amount of protein after each purification step is indicated in addition to sample volumes. Binding reactions were performed with protein samples from each stage of the enrichment scheme (Section 2.3.22) using the conditions previously described (legends to Figures 5.1 and 5.2). After binding, bound and unbound DNA was separated by polyacrylamide gel electrophoresis, visualized by autoradiography and appropriate bands excised from the gel and counted by liquid scintillation spectroscopy. The counts per minute of input DNA ( $[^{32}\text{P}]$  duplex oligonucleotide) in the bound complex was used as a measure of the binding activity in that fraction. Specific activity of factor binding was calculated as the counts per minute of total input DNA resolved as a bound complex per microgram of input protein. Excess quantities of labelled DNA were included to ensure quantitative results. Data presented applies to the enrichment of H1-SF and H2B-SF but specific binding activity of H2B-SF is indicated in the parenthesis. Nuclear extract was prepared from  $\sim 2 \cdot 10^{10}$  AEV ts34 cells (see Section 2.3.21).

<u>Fraction</u>	<u>Vol. (ml)</u>	<u>Amount of Protein</u>			<u>Specific Binding Activity</u>
		mg/ml	total mg	% total	
Salt Extract	7.8	10	78	100	1 (1)
Sephacryl S-300	25	0.7	17.5	22	15 (12)
DEAE-Sephrose	10	0.15	1.5	1.9	237 (182)

fractionation of extract components revealed that in the native state, the specific binding factor (H1-SF) had an apparent molecular weight of 80,000-100,000.

To demonstrate that a conserved H1-box was an absolute requirement for specific binding activity, gel retardation assays were performed using duplex-oligonucleotides with multiple base substitutions in the conserved 'core' (5'-AAACACA-3') heptanucleotide region. This set of point mutations were previously shown to result in a loss of H1 gene cell-cycle transcriptional regulation (see Chapter 4). Mutations which abolished the conserved H1-box motif also completely eliminated H1-SF binding. This is consistent with this factor being important for transcriptional regulation as mutagenesis of its target binding site also is accompanied by a loss of periodic transcriptional regulation of the H1 gene in vivo. Specific binding of a nuclear factor to the H1-box was also confirmed by DNAase footprint experiments.

A sequence-specific nuclear factor has also been identified to specifically interact with the H2B-box by a similar approach used in the characterization of H1-SF binding to the H1-box. This factor is likely to be similar as that described by Sive and Roeder (1986) which interacts with the H2B-box of human histone genes and also to similar sequences in the human U2 small nuclear RNA gene promoter, mouse immunoglobulin heavy chain enhancer and, a mouse light chain promoter. A factor has also been reported to bind a similar sequence in the SV40 early promoter (Bohmann et al., 1987). Together, these observations make it likely that a common or closely related factor contributes to the regulation of these and other genes which utilize this sequence for transcriptional regulation. Unlike H1-SF, the H2B-box specific factor uniformly binds to its target recognition sequence in extracts prepared from cells at different stages of the cell-cycle. Constant and constitutive binding of this factor may be indicative that it is not

required for S-phase regulation per se, although it may be necessary for absolute promoter efficiency. This is consistent with results presented in Chapter 4 which show that deletion of the H2B-box does not completely eliminate periodic activation of transcription but significantly reduces the general rate of transcriptional initiation during S-phase.

### 5.3.2 Binding Activity of the H1-Specific Nuclear Factor Fluctuates Throughout the Cell-Cycle

It has been established that transcription of H1 histone genes is transiently elevated during S-phase and that a gene-specific element is required for this to occur. Therefore, it was of interest to determine if levels of nuclear factor binding to this element also fluctuated in a cell-cycle dependent manner. Levels of H1-SF complex formation were monitored by gel mobility-shift assays using nuclear extracts prepared from AEV ts34 cells at various stages of synchronization. These studies revealed that factor binding to the H1-box was transiently elevated in extracts prepared from S-phase cells and depressed at other stages of the cell-cycle. The observation that factor binding correlated with the kinetics of transcriptional activation during S-phase implies that it may have a specific role in cell-cycle control. Presumably, levels of complex formation between H1-SF and its target binding sequence bears a true representation of binding activity at corresponding stages of the cell-cycle in vivo. The possibility that fluctuating levels of H1-SF factor binding is due to some cell-cycle related artefact is unlikely as H2B-SF binding levels were essentially invariant in the same set of extracts.

Elevated H1-SF levels in S-phase, together with the observation that levels of this factor are depressed when gene activity is low (+ aphidicolin, G<sub>2</sub>-phase), implies that increased availability of H1-factor could account for elevated gene transcription. Increased levels of H1-SF could be due to a periodic increase in synthesis, prolonged

stability (i.e. half-life) or an increase in the free pool of factor available for binding. Apart from modulating the absolute levels of this factor, alternatively, factor levels may remain constant throughout the cell-cycle but selective binding ability/affinity could be controlled by post-translationally modifying the protein.

CHAPTER 6

CHICKEN HISTONE GENES RETAIN NUCLEAR

MATRIX ASSOCIATION THROUGHOUT THE CELL-CYCLE



## CHAPTER 6

### CHICKEN HISTONE GENES RETAIN NUCLEAR MATRIX ASSOCIATION THROUGHOUT THE CELL-CYCLE

#### 6.1 INTRODUCTION

The three dimensional conformation of DNA sequences appears to play an important role in the control of gene expression in vivo. Several differences exist between the conformation of actively transcribed and inactive genes. For instance, transcribed DNA sequences exhibit increased nuclease sensitivity as compared to untranscribed sequences (Garel and Axel, 1976; Weintraub and Groudine, 1976; Bellard et al., 1977; Bloom and Anderson 1978). Furthermore, transcribed genes are predominantly enriched in euchromatin as opposed to heterochromatin and are often complexed with specific proteins associated with their activation (Weisbrod, 1982; Cartwright et al., 1982).

Chromatin fibres appear to be arranged into topologically constrained domains or loops which are anchored by a sub-nuclear residual framework in both metaphase (Paulson and Laemmli, 1977; Adolph et al., 1977) and interphase (Benyajati and Worcel, 1976; Cook and Brazell, 1976; Igo-Kemenes and Zachau, 1978; Lebkowski and Laemmli, 1982) chromosomes. This framework resembles a 'scaffold' structure and is thought to be intimately involved in the organization of interphase chromatin into topologically distinct loops. The nuclear scaffold structure is composed of three elements; a peripheral lamina, an internal protein network and, residual nucleolar structures. A model predicting the organization of chromatin in a series of loops anchored at specific attachment sites to scaffold components has been proposed (Laemmli et al., 1978; Marsden and Laemmli, 1979). These attachment sites have been postulated to stabilize topological domains along the

chromatin fibre in order to facilitate a number of processes which occur proximal to the scaffold structure.

In order to study the organization of interphase chromatin and attachment of genes to the nuclear scaffold, a corresponding structure known as the nuclear matrix (NM) has been employed for in vitro studies at the molecular level. The NM is a residual framework to which supercoiled loops of DNA are anchored and is commonly prepared by extraction of isolated nuclei with high-salt, resulting in removal of histones and other proteins. Electron microscopy reveals that following high-salt extraction of nuclei, NM structures appear as a proteinaceous scaffold-like network anchoring a halo of large supercoiled loops of DNA estimated to be approximately 30-90kb in length (Benyajati and Worcel, 1976; Vogelstein et al., 1980; Paulson and Laemmli, 1977). The association of actively transcribed genes with the nuclear scaffold and the way in which this influences the dynamic conformation of these DNA sequences has been implicated in gene activation. In addition, the scaffold structure is also thought to be involved in other nuclear processes such as DNA synthesis (Berezney and Bucholtz, 1981; Pardoll et al., 1980) and RNA processing (Ciejek et al., 1982; Herman et al., 1978; van Eekelen and van Venrooij, 1981).

Due to the nature of native chromatin, which aggregates under isotonic conditions, it is usually necessary to employ hypotonic or hypertonic salt conditions when analyzing chromatin structure in vitro. Consequently, it is sometimes difficult to interpret results generated from experiments involving chromatin prepared under these different conditions. Micrococcal nuclease digestion of chromatin, isolated under hypotonic conditions, indicates that actively transcribed genes partition into an insoluble nuclear fraction in contrast to untranscribed genes which are solubilized (Cohen and Sheffery, 1985; Stratling et al., 1986; Xu et al., 1986; Ryoji and Worcel, 1985). This

phenomenon can be explained by actively transcribed genes coalescing into large complexes, or alternatively, by an association with a nuclear structure such as the matrix. A large body of evidence indicates that transcriptionally active genes selectively partition into an insoluble, high-salt extracted, chromatin fraction following nuclease digestion through selective attachment to the nuclear matrix. Actively transcribed ovalbumin (Robinson et al., 1982), vitellogenin 2 (Jost and Seldran, 1984), heat-shock (Small et al., 1985; Mirkovitch et al., 1984), actin (Small et al., 1985), globin (Robinson et al., 1982; Hentzen et al., 1984) and ribosomal RNA (Pardoll and Vogelstein, 1980) genes are amongst those found to be preferentially associated with the insoluble NM. Furthermore, essentially all rapidly-labelled nascent HnRNA is associated with the NM (Pardoll, et al., 1980; Miller et al., 1978), suggesting that a sub-nuclear structure plays a fundamental role in the mechanisms governing gene transcription.

Convincing evidence for transcribed genes being insoluble due to NM-association and the possible role of this structure in influencing gene transcription comes from two independent lines of investigation. Firstly, it has been shown that in intact nuclei, under physiological salt conditions, nascent transcripts, active RNA polymerase and transcribed genes are closely associated with a large sub-nuclear structure in preference to untranscribed genes (Jackson and Cook, 1985). Secondly, in a series of elegant mapping experiments, Cockerill and Garrard (1986) have identified specific DNA sequence elements necessary for NM-attachment of  $\kappa$ -immunoglobulin genes which have since been shown to co-map with enhancer sequences (Cockerill et al., in press). Cohabitation of scaffold binding regions with the upstream enhancer regions of alcohol dehydrogenase, fushi tarazu and glue protein genes have also been identified in Drosophila (Gasser and Laemmli, 1986). Therefore, anchorage of active genes to the NM appears, at least

in some cases, to be integrally related to the transcription apparatus itself.

Work previously described in this thesis focussed principally on the cis- and trans- regulatory factors involved in the control of histone gene transcription. This chapter is directed towards analyzing another major determinant of gene regulation, - DNA topology. Specifically, this involved studying the relationship between histone genes and the nuclear matrix.

As both core and H1 chicken histone genes have been shown to be periodically activated with respect to transcription during S-phase of the cell-cycle (Chapter 3), it was of interest to see if corresponding changes in nuclear matrix attachment of these genes could be detected. It is well established that active genes are closely associated with the nuclear matrix and that anchorage via proximal DNA sequences is required for efficient transcriptional activity. In contrast, little is known regarding attachment of 'silent' genes or those transcribed at low levels. To test the possibility that a direct correlation exists between the transcriptional state of a gene and its conformation relative to the nuclear matrix, the temporally activated histone gene system has been analyzed. As histone gene transcription transiently fluctuates through the cell-cycle, this provides an ideal situation where interactions between genes and the nuclear matrix can be explored during periods of high and low transcriptional activity. Apart from the major core and H1 genes, the AEV ts34 cell-line expresses a tissue-specific linker histone gene, H5. This presents a unique situation where attachment of a histone gene (H5) to the NM can be studied in cells in which it is 'silent' and where in others it is transcribed. Most of the work presented has been published (Dalton et al., 1986b).

## 6.2 RESULTS

### 6.2.1 Enrichment of Core and H1 Histone Genes with the Nuclear Matrix is Independent of Their Transcriptional Activity

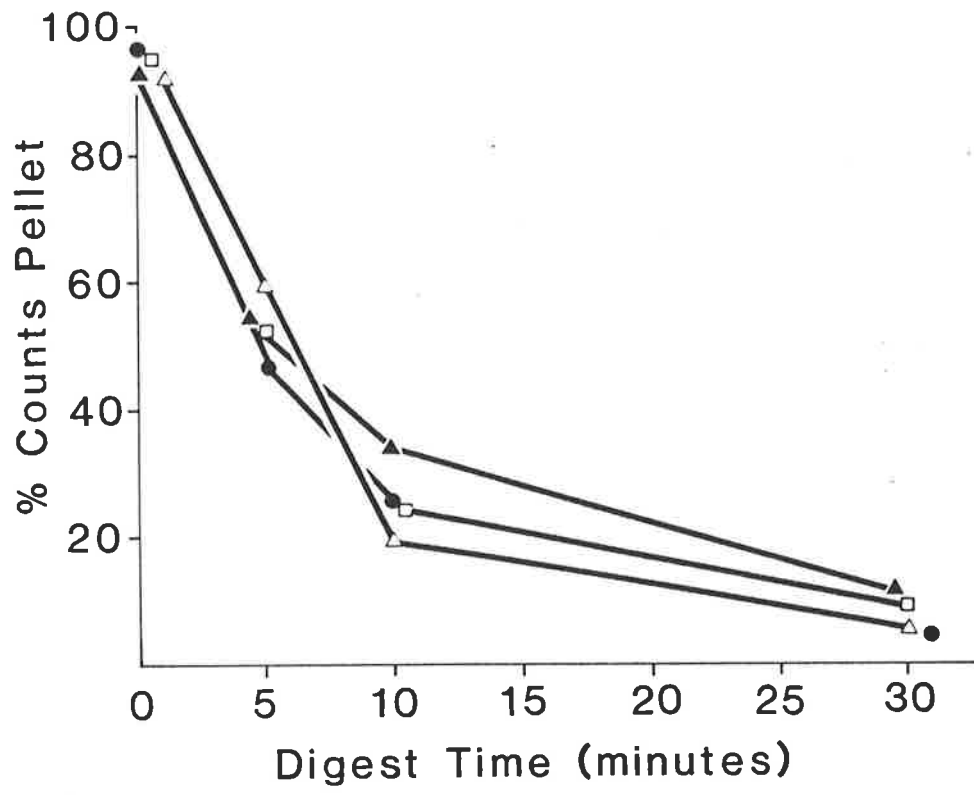
Following high-salt extraction of nuclei isolated at different stages of cell-synchronization, intact genomic DNA remains arranged in large supercoiled loops anchored to the NM (Pardoll et al., 1980). These loops of DNA can be treated with nucleases such as DNAase I or restriction endonucleases, blotted onto filter membranes and probed so that sequences remaining associated with the matrix (located in the pellet after brief centrifugation) and supernatant can be determined. This approach was employed in an attempt to determine if histone genes are preferentially anchored to the NM and, if so, whether this is strictly dependent on them being transcriptionally active during S-phase of the cell-cycle.

Initially, digestion of nuclear matrices prepared from [<sup>3</sup>H]-labelled populations of AEV ts34 cells before, during and after S-phase were digested with DNAase I (see Chapter 3 for details on cell-synchronization). DNA sequences selectively attached to the NM are preferentially retained by the insoluble (pellet) fraction and, increasing digestion results in the enrichment of such sequences which can be detected by hybridization to a specific probe following DNA purification. In order to monitor if histone genes are anchored to the NM (i.e. enriched following DNAase digestion), matrices were digested for up to 30 minutes, and the quantity of DNA released/retained was detected by [<sup>3</sup>H]thymidine label present in pellet and S/N fractions (see Figure 6.1). This procedure was performed using matrices prepared from synchronized AEV ts34 cells at various stages of cell synchronization (see Chapter 3). Equal amounts of DNA from pellet (NM) fractions were purified following digestion from 0-30 minutes, dotted onto nitrocellulose filter membranes and then probed (Section 2.3.16). An

## FIGURE 6.1

### RELEASE OF NUCLEAR MATRIX BOUND DNA FOLLOWING DNAase DIGESTION

Release of [ $^3\text{H}$ ]-thymidine labelled DNA following DNAase I (25  $\mu\text{g}/\text{ml}$ ) digestion of nuclear matrices isolated from pre-labelled cells (by the high-salt method, Section 2.3.14) at various stages of cell-synchronization (see Chapter 3). Data is plotted as the percentage of residual [ $^3\text{H}$ ]-label (see Chapter 3) remaining in the insoluble nuclear matrix fraction (pellet) in relation to total input radioactivity (pellet plus supernatant fractions) versus the time of digestion. Matrices from unsynchronized ( $\Delta$ ); aphidicolin blocked ( $\blacktriangle$ ); 2 hrs. post-release (S-phase,  $\bullet$ ) and  $G_2$ -cells ( $\square$ ) were digested from 0-30 min., pelleted, SDS and proteinase K treated (Section 2.3.14) and assayed on a scintillation spectrometer (10 minutes counts). Results are expressed as the mean of two independent experiments.



H2A histone gene probe (Section 2.2.4) was selected as an example of a core histone gene, previously shown to be transcribed at elevated levels during S-phase (see Chapter 3). Analysis of residual DNA remaining associated with the NM following DNAase digestion, by probing dot blots, revealed that H2A gene sequences were enriched in matrices isolated from all cell-cycle time points (Figure 6.2). This enrichment (approximately 6-8 fold) was essentially the same during, before and after S-phase (Table 6.1), indicating that core histone genes are anchored to the NM, independent of their transcriptional state. Similar results were also obtained when these experiments were repeated but with 0.5 and 1  $\mu$ g of purified DNA applied per dot (Table 6.1). A positive (pJL4 vector + H2A gene insert) and negative (pJL4 vector minus insert; Section 2.2.5) control set of dots were included to determine non-specific and specific levels of hybridization (Figure 6.2)

Although DNAase treatment allowed enrichment of particular sequences with the NM to be detected, a more useful approach was to digest supercoiled DNA with restriction endonucleases so that DNA products of discrete sizes in the matrix and/or S/N were obtained. Nuclear matrices (prepared as in Figure 6.1) were pre-digested and then digested to completion with EcoRI or EcoRI/HindIII, resulting in 17-23% and 9-16% of total [<sup>3</sup>H]-thymidine labelled DNA remaining associated with the nuclear matrix following digestion, respectively. The amount of residual DNA remaining anchored following EcoRI digestion was slightly higher than previously reported for avian NM but is lower than that for Drosophila matrices (Buongiorno-Nardelli et al., 1982). This variation is likely to be a consequence of differences in supercoiled DNA loop size and arrangement. In cases where matrix-associated DNA was digested with different restriction enzymes, variations are also likely to be caused by the presence of satellite DNA which remains undigested due to the absence of cleavage recognition sequences. Following restriction



## FIGURE 6.2

### ENRICHMENT OF H2A GENE SEQUENCES WITH THE NUCLEAR MATRIX BY DNAase DIGESTION

Enrichment of H2A hybridizing sequences associated with unsynchronized (U), aphidicolin blocked (+A), S- and G<sub>2</sub>-phase high-salt prepared (Section 2.3.14) NM following DNAase I digestion (see Figure 6.1). DNA remaining associated with pellet fractions was purified by SDS, proteinase K and RNAase treatment (Section 2.3.14). Following phenol/chloroform extraction and ethanol precipitation, 2.5 µg DNA was analyzed by dot-blot hybridization (Section 2.3.16) with an H2A gene-specific probe (gene insert; Section 2.2.4), for each digestion time point (0,5,10,30 min.). Similar results were obtained when 0.5 and 1 µg of DNA was used (see Table 6.1). Hybridization to a +ve set of control DNA dots (10 ng pJL4 + H2A insert) or -ve control set of DNA dots (10 ng pJL4 - H2A insert) are shown.

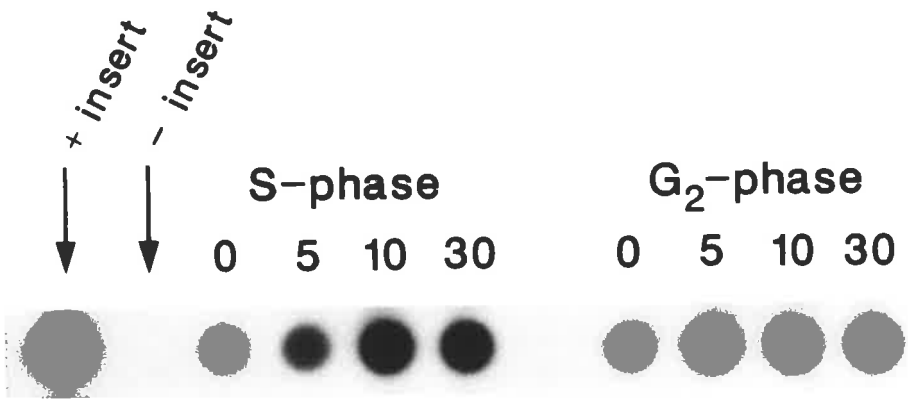
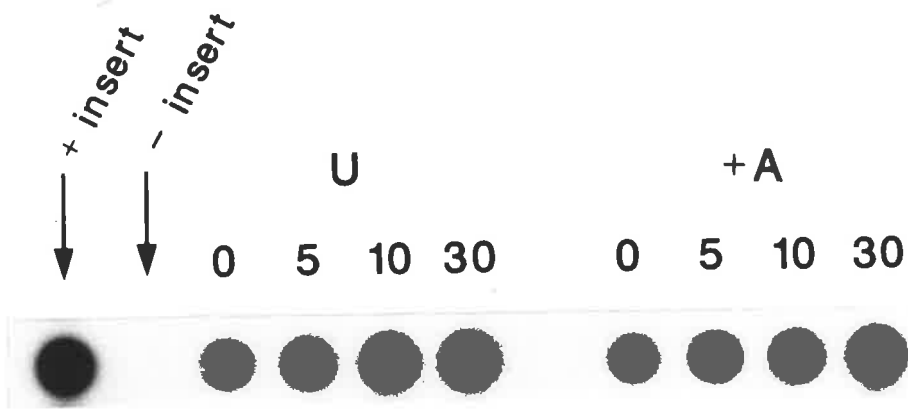


TABLE 6.1

ENRICHMENT OF H2A GENE SEQUENCES ATTACHED TO THE NUCLEAR  
MATRIX THROUGHOUT THE CELL-CYCLE

Analysis of H2A hybridizing sequences present in pellet (NM) fractions following DNAase digestion of matrices (Section 2.3.14) obtained from unsynchronized (U), aphidicolin-blocked (+A), S-phase (S) and G<sub>2</sub>-phase (G<sub>2</sub>) cells. Data is expressed as the ratio of the hybridization signals (after subtraction of non-specific hybridization to pJL4 vector in each case) between matrix-bound DNA (0.5, 1 and 2.5 μg) following digestion for 30 min. (T30) and undigested matrix-bound DNA (T0), as determined by scintillation spectroscopy (see Figure 6.1).

<u>Quantity of DNA (μg)</u>	<u>T0/T30 RATIO</u>			
	<u>Stage of Cell Synchronization</u>			
	U	+A	S	G <sub>2</sub>
0.5	6.5	7.9	7.0	6.9
1.0	6.2	7.6	7.1	6.8
2.5	8.8	7.5	8.1	8.2

enzyme digestion, DNA from matrices and supernatants was purified and fractionated by agarose gel electrophoresis. Purified matrix and S/N DNA samples (10 µg) prepared from matrices isolated at different stages of cell synchronization were electrophoresed, blotted onto nitrocellulose filters and hybridized with a [<sup>32</sup>P]-labelled probe for the linker histone gene, H1 (see Figure 6.3). For all cell-cycle stages, the intensity of bands in supernatant tracks was substantially lower than in matrix tracks, for EcoRI/HindIII digests. Bands in NM tracks ranging from approximately 11-2.1kb indicate that all six histone H1 genes are preferentially associated with the scaffold and lie within 11kb of this structure. The presence of a 2.1kb band in the matrix track indicates an anchorage site relatively close (within 2.1kb) to at least one H1 gene in the presence or absence of high transcription levels (S- and non-S phase, respectively). It is recognized that in some cases there appears to be more DNA in supernatant than matrix-fractions (see Figure 6.3). However, this only serves to confirm the general observation that histone H1 genes are preferentially retained by the nuclear matrix throughout the cell-cycle. Southern analysis of DNA isolated from EcoRI digested matrices revealed a similar distribution of H1 sequences between pellet (NM) and S/N fractions throughout the cell-cycle. A similar experiment using EcoRI digested nuclear matrices from different populations of synchronized AEV ts34 cells indicates that H2A hybridizing sequences are also present predominantly in matrix fractions (data not shown). The pattern of H1 and H2A hybridizing bands closely resembled that of total genomic DNA when digested with the same restriction enzymes (see Ruiz-Carillo et al., 1983).

#### 6.2.2 Tissue-Specific Attachment of the Histone H5 Gene to the Nuclear Matrix

The tissue-specific linker histone H5 gene (see Krieg et al., 1983) was also examined to see if it is matrix bound in either AEV ts34

### FIGURE 6.3

#### CELL-CYCLE INDEPENDENT ENRICHMENT OF H1 GENE SEQUENCES WITH THE NUCLEAR MATRIX FOLLOWING RESTRICTION ENZYME DIGESTION

Restriction enzyme analysis of H1 sequence association with NM (Section 2.3.14) at various stages of cell synchronization (see Chapter 3). Matrix structures (Section 2.3.14) from aphidicolin blocked (+A), S-phase (2 hrs after release from the aphidicolin block) and G<sub>2</sub>-cells (10 hrs post-release) (see Figure 3.3), were digested with Eco RI/Hind III and the purified DNA resolved on a 0.8% agarose gel. DNA (10 µg) present in pellet (P) and supernatant (S) fractions was analyzed by hybridization (Section 2.3.17) with a [<sup>32</sup>P]-nick translated H1 gene-specific probe (Section 2.2.4) and detected by autoradiography. Lower panels;- Ethidium bromide stained agarose gel showing digested DNA samples corresponding to the autoradiogram in the upper panels. Sizes of detected bands are shown relative to Eco RI-cut SPP1-DNA markers.

+A

S-phase

G2

P S

P S

P S



kb

- 7.8
- 7.0
- 5.9
- 4.7
- 3.4
- 2.7
- 1.9
- 1.8
- 1.5
- 1.3

P S

P S

P S



kb

- 7.8
- 7.0
- 5.0
- 4.7
- 3.4
- 2.7
- 1.9
- 1.8
- 1.5
- 1.3

cells or in a chicken T-cell line. This gene is unusual among the histone genes in two respects. Firstly, it is only expressed in cells of erythroid-lineage and, unlike the majority of other histone genes, H5 transcription is not temporally coupled to DNA replication (see Chapter 3). BamHI/SacI digestion of total matrix DNA produces a characteristic band of approximately 1.7kb in the matrix-associated fraction of AEV ts34 cells but exclusively in the S/N fraction of the T-cell line (Figure 6.4). This band corresponds to the H5 coding region in addition to some 5'- and 3'-non-coding sequences (see Krieg et al., 1983). Nuclear matrix association of transcriptionally active genes in the T-cell-line was demonstrated by the preferential distribution of  $\beta$ -actin gene sequences with the insoluble chromatin (pellet) fraction (data not shown). Digestion of AEV ts34 or T-cell matrices with Bam HI/Sac I resulted in 92% and 87% of total nuclear [<sup>3</sup>H]-labelled DNA being released into the supernatant, respectively.

To define the region of DNA required for tissue-specific attachment of the H5 gene to the NM, further mapping experiments were performed. A triple digestion of AEV ts34 cell matrices with Sac I, Bam HI and Nae I (resulting in 93% of total [<sup>3</sup>H]-nuclear DNA being released), generated two distinct hybridizing bands when probed with a 1.7 Kb genomic fragment containing the H5 gene. Digestion of matrix DNA with these three enzymes resulted in hybridizing fragments of approximately 780 and 920 bp being detected in the pellet (NM) and S/N fraction, respectively (Figure 6.5). This indicates that further digestion of the 1.7 Kb Sac I/Bam HI fragment (as seen in Figure 6.4) with Nae I, results in dissection of the H5 gene and release of some H5 coding and 3'-untranslated regions, while part of the coding and 5'-flanking regions are retained by the matrix (refer to Figure 6.6). The DNA sequences necessary for H5 gene attachment to the NM are therefore confined within a 780 bp fragment spanning part of the coding and 5'-

## FIGURE 6.4

### TISSUE-SPECIFIC ATTACHMENT OF THE H5 GENE TO THE NUCLEAR MATRIX

Association of the tissue specific H5 gene with pellet and supernatant fractions in AEV ts34 cells and T-cells. Matrices from unsynchronized cells were prepared by the high-salt method (Section 2.3.14) and then digested with Bam HI/Sac I. DNA was purified, 10  $\mu$ g was electrophoresed through a 1% agarose gel, transferred to a nitrocellulose filter and probed with an H5-specific [ $^{32}$ P]-kinased synthetic oligonucleotide (26-mer; Section 2.2.5) in order to determine the distribution of H5 sequences between pellet (P) and supernatant (S) fractions in both cell-lines (see Section 2.3.17). Sizes of detected bands are shown relative to Eco RI-cut SPP1-DNA markers.

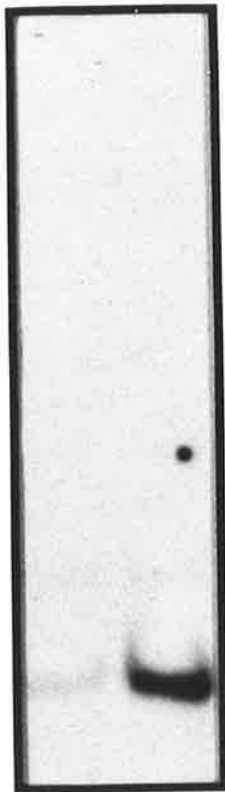


ts34

T-cell

S P

S P



kb

7.8

7.0

5.9

4.7

3.4

2.7

1.9

1.8

1.5

1.3

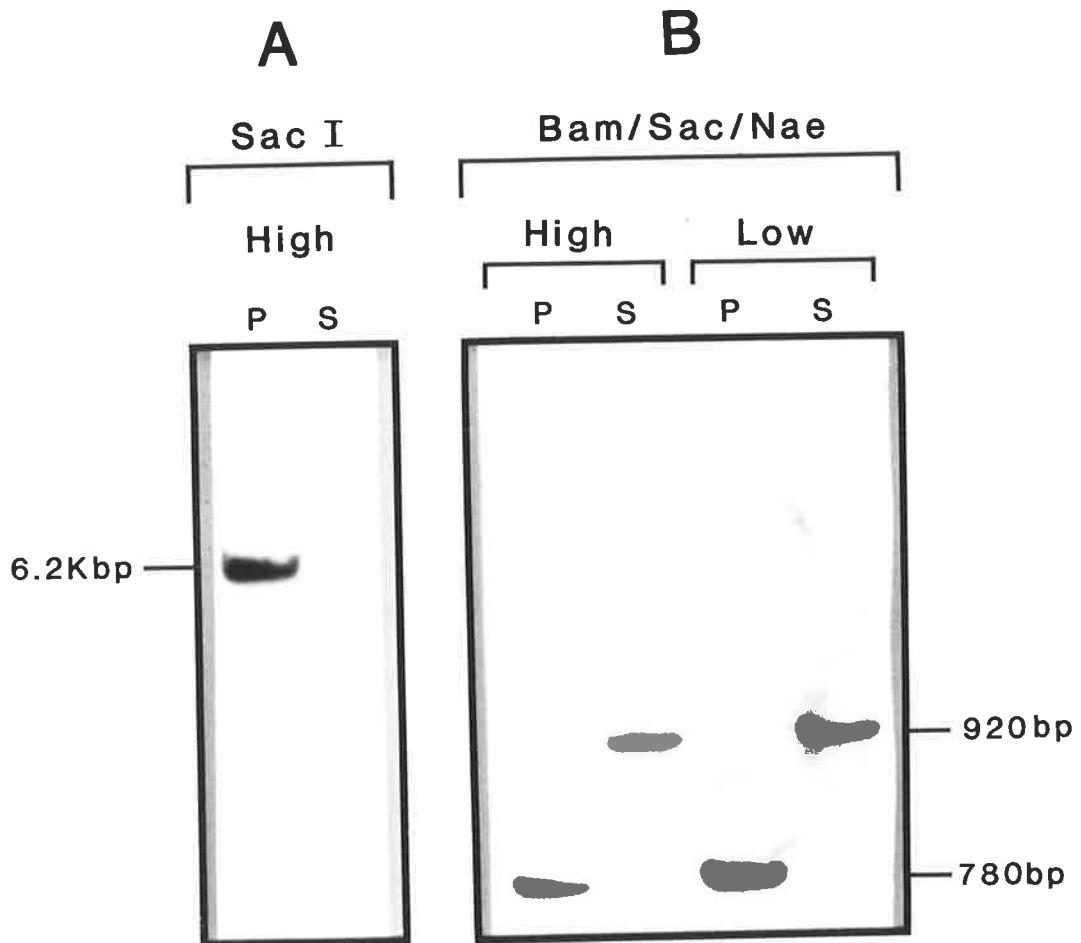
1.1

## FIGURE 6.5

### IDENTIFICATION OF THE H5 GENE MAR IN

#### AEV ts34 CELLS

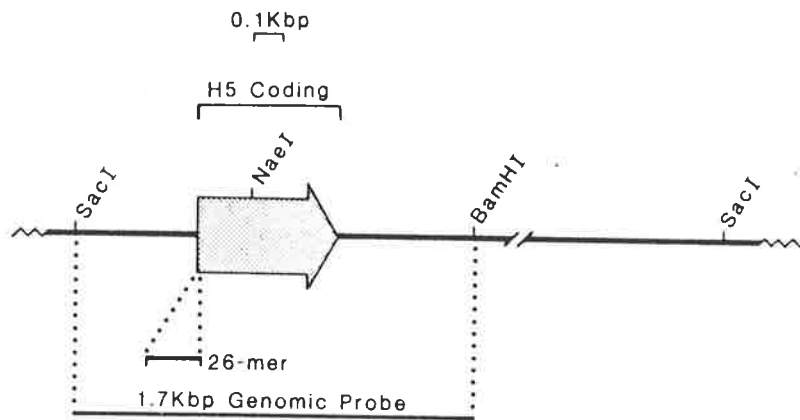
Mapping of the H5 gene nuclear matrix anchorage region (MAR) in AEV ts34 matrices. Panel A: Nuclear matrices were extracted by the high-salt method (Section 2.3.14), digested with Sac I followed by electrophoresis of DNA (10  $\mu$ g) from pellet (P) and supernatant (S) fractions through a 1% agarose gel. Panel B: Matrices were isolated in high or low-salt buffer (Section 2.3.13) and digested with Bam HI, Sac I and Nae I; followed by electrophoresis of pellet (P) and supernatant (S) fractions (1.5 and 15  $\mu$ g DNA, respectively) through a 1.5% agarose gel. Gels were then blotted and probed (Section 2.3.17) with a 1.7 Kb Bam HI/Sac I [ $^{32}$ P]-nick translated H5 genomic insert (see Figure 6.6). Hybridizing DNA fragments (sizes indicated) were detected by autoradiography.



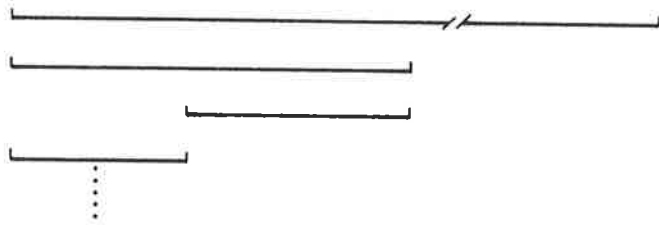
## FIGURE 6.6

### STRATEGY USED TO IDENTIFY THE H5 GENE MAR IN AEV ts34 CELLS

This figure summarizes the strategy used to identify the matrix anchorage region (MAR) of the H5 gene with AEV ts34 matrices. Restriction enzyme sites (Krieg et al., 1983) relevant to the mapping of the MAR are indicated relative to the H5 coding region. Probes (synthetic 28-mer; Section 2.2.6a; and 1.7 Kb genomic fragment) used to identify H5 gene sequences in supernatant and pellet (NM) fractions (see Figures 6.4 and 6.5) are indicated relative to the gene.



NM Attachment	Length (Kbp)
Yes	6.2
Yes	1.7
No	0.92
Yes	0.78



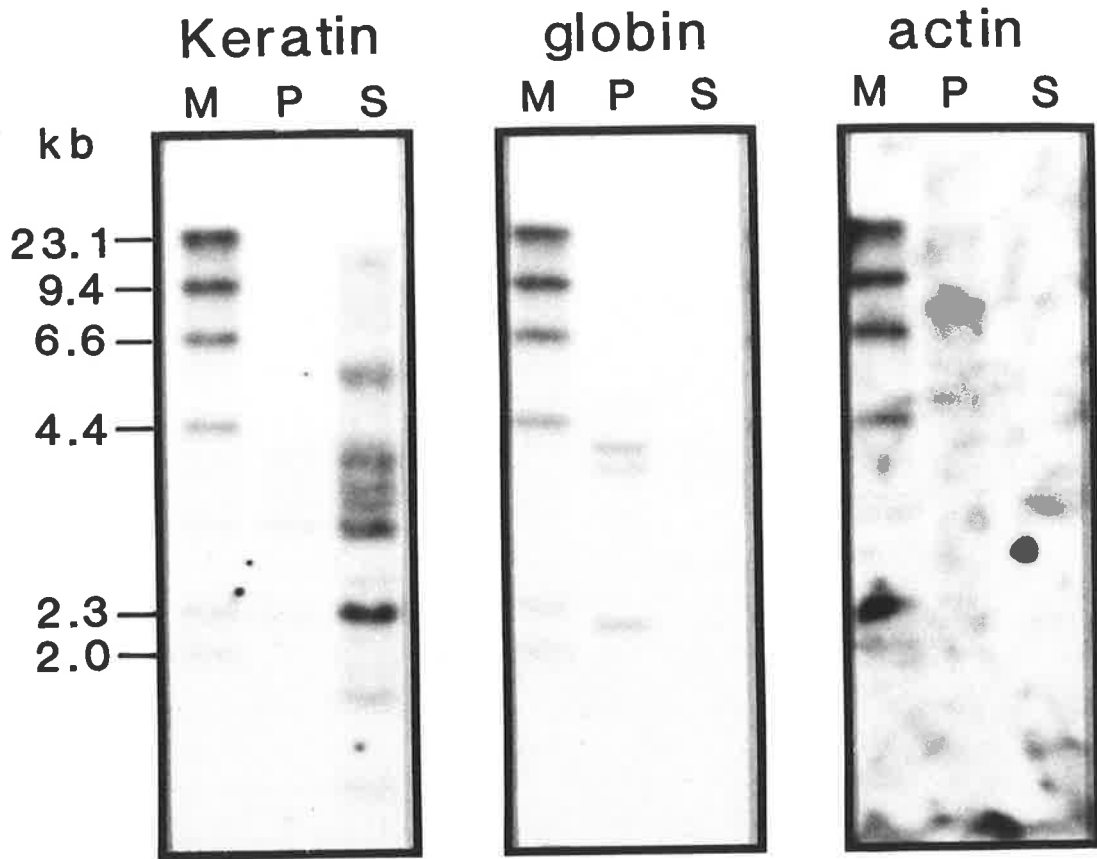
untranscribed region. This experiment was repeated using matrices prepared by the low-salt method of Mirkovitch et al. (1984) with similar results (Figure 6.5). Other experiments using this genomic probe revealed that the H5 gene, located within a 6.2 Kb Sac I fragment, is selectively associated with the matrix following digestion.

So as to ensure that hybridizing gene sequences were not anchored to the NM of ts34 cells in an obligatory manner, the distribution of  $\beta$ -actin,  $\beta$ -globin and feather keratin genes between matrix and S/N fractions were characterized. In the chicken AEV ts34 early-erythroid cell-line, it has been shown that mRNA for  $\beta$ -actin, but not feather keratin is present as determined by Northern analysis (see Section 3.2.1; and unpublished observations). Globin, a classical marker used for determination of erythroid differentiation has been shown by Beug and co-workers (1982) to be essentially absent in this cell-line by radioimmunoassay.  $\alpha$ -globin (see Chapter 3) and  $\beta$ -globin (result not shown) mRNA is only present at very low levels in AEV ts34 cells whereas the level of  $\beta$ -actin mRNA is relatively high (see Figure 3.2). Consistent with previous findings on the disposition of  $\beta$ -actin gene sequences in Drosophila matrices (Small et al., 1985), the  $\beta$ -actin hybridizing bands of approximately 8.7 and 5.0 kb are observed in the matrix lane while essentially absent in the S/N following Eco RI/Hind III digestion. Similar experiments examining the feather keratin genes revealed the presence of several hybridizing bands ranging from 14.6-1.64 kb exclusively in the supernatant but absent in the pellet fraction (Figure 6.7). However, when digested with Eco RI alone, keratin hybridizing sequences >14 Kb in length were found predominantly in the insoluble pellet fraction (result not shown). The probe used to detect feather keratin genes (Molloy et al., 1982) also cross-hybridizes with scale genes (R. Presland, personal communication), therefore implying that scale genes (not transcribed in this cell-line) are also unattached

## FIGURE 6.7

### PARTITIONING OF FEATHER KERATIN, $\beta$ -GLOBIN AND $\beta$ -ACTIN SEQUENCES BETWEEN NUCLEAR MATRIX AND PELLET FRACTIONS

Partitioning of feather keratin,  $\beta$ -globin and  $\beta$ -actin sequences between pellet (P) and supernatant (S) fractions following Eco RI/Hind III digestion of AEV ts34 NM DNA (Section 2.3.14). 10  $\mu$ g of digested DNA was electrophoresed through a 1% agarose gel, blotted and probed with [ $^{32}$ P]-labelled nick translated gene insert (Sections 2.3.10a; 2.2.4). Sizes of detected bands are shown relative to Hind III cut lambda-DNA markers (M).





to the NM. Unexpectedly, Southern analysis indicates that  $\beta$ -globin sequences are anchored to the nuclear scaffold as indicated by hybridizing bands of 4.2, 4.0 and 2.0kb which are all preferentially present in NM tracks. This is not consistent with the proposition that gene transcription and NM anchorage are directly coupled, although it is possible that attachment of this gene to the NM is purely fortuitous. The generation of a 2.1kb  $\beta$ -globin hybridizing band by EcoRI/HindIII digestion suggests a close physical link between the globin gene locus and the NM in this cell-line.

### 6.3. DISCUSSION

Core and H1 histone genes remain NM-associated in both maximal and minimal states of transcription. That is, the same quantitative profile of NM-associated histone genes is observed between cell populations in which greater than 90% are in S-phase, with associated maximal histone gene transcription, and those in which less than 1% are in S-phase (see Chapter 3). This result reinforces the concept that NM-attachment is necessary although not sufficient for transcription of specific genes, in agreement with results of Small et al (1985) for  $\beta$ -actin and heat shock genes.

There is some uncertainty about the mechanism of gene attachment to the NM, whether it occurs in vivo and, if so, what its biological significance may be. Because high salt conditions are commonly used in NM preparation, there is clearly the possibility of artifactual associations occurring in vitro. For instance, Mirkovitch et al. (1984) have suggested that under some conditions, 2M-NaCl can induce the precipitation of transcriptional complexes. However, Keppel (1986) has recently shown transcriptional-dependent anchorage of human ribosomal RNA genes to the NM using both high and low salt conditions (also see Razin et al., 1986) and, similar results have been presented under both

sets of conditions in mapping the domain responsible for H5 gene attachment (see Figure 6.5). A biological role for the anchorage of transcriptionally active genes to the nuclear scaffold in vivo is suggested by the identification of evolutionarily conserved sequence elements consisting of several hundred base pairs of A/T-rich DNA also containing the topoisomerase 2 consensus sequence which anchor such genes as the mouse  $\kappa$ -immunoglobulin gene to the NM (Cockerill and Garrard, 1986). This, together with anchorage sites co-mapping with enhancer regions of several genes (Gasser and Laemmli, 1986; Cockerill and Garrard, 1986; Cockerill et al., in press), make it tempting to speculate that the matrix is integrally involved in assembly of transcription complexes. If, as it seems, attachment is a pre-requisite for transcription in vivo, then other factors are likely to be necessary for expression of these genes. The attachment process may be triggered by specific factors as with steroid hormone inducible genes (Taniguchi et al., 1984; Swaneck and Alvarez, 1985; see Section 1.7.2) or it may result from DNA-specific interaction with existing matrix proteins. For histone genes, ori sequences could be involved since DNA synthesis is a NM-associated phenomena and with a few exceptions (Dalton et al., 1986a), is also coupled to histone gene transcription.

Notably, in the two cell-lines tested, a correlation was shown between NM attachment and transcription of the H5 gene. While little is known regarding the mechanism of NM attachment of transcriptionally active genes in vivo, in light of these results, it is conceivable that DNA-protein complex (i.e. stable transcription pre-initiation complexes) formation at the scaffold is involved. For the H5 gene, this may involve a tissue-specific cellular component (i.e. DNA-binding protein) which may be required for erythroid-specific transcriptional activity.

There is no difficulty in demonstrating selective NM-association of specific genes experimentally. The H5 gene for example, is active

and NM-attached in erythroid cells but inactive and unattached in T-cells (Figure 6.4, and see Chapter 3). Similarly, there is a correlation between the absence of close feather keratin gene anchorage to the NM, and the inactive state of these genes in erythroid cells. It is reasonable to equate gene attachment to the NM with DNAase sensitivity as markers of genes with potential for expression within cells. A good example is the DNAase hypersensitivity of globin 5'-sequences prior to expression of the gene during red cell development (Weintraub et al., 1981). This, together with the finding that  $\beta$ -globin genes are NM-associated in AEV ts34 cells suggests that such attachment may be necessary to 'prime' the gene ready for transcription during erythroid differentiation. It is interesting to note, however, that prior to the onset of transcription, potentially active  $\alpha$  and  $\beta$ -major globin genes can be enriched in an insoluble nuclease-digested chromatin fraction (Cohen and Sheffery, 1985). Similarly, Chrysogelos et al (1985) have shown that human histone genes exhibit an intermediate level of regulation whereby they are never in a completely inactive conformation throughout the cell-cycle and can be considered as being 'primed' for transcriptional activation at the onset of S-phase. This is consistent with the observations presented that core and H1 histone genes are anchored to the NM, independent of their transcriptional activity. While clear and repeatable patterns of DNA attachment for given genes in specific cell-types can be seen, it will be necessary to further clarify the DNA sequence elements involved in such associations and the identification of putative NM proteins involved in these interactions.

CHAPTER 7

FINAL DISCUSSION

## CHAPTER 7

### FINAL DISCUSSION

#### 7.1 CELL-CYCLE CONTROL OF HISTONE GENE EXPRESSION

While it is clear that strict regulatory controls are required for the co-ordinate synthesis of histone proteins during S-phase, the underlying mechanisms facilitating this are not fully understood. At least two levels of control are involved and these act at the levels of histone gene transcription and, differential histone mRNA stability. In the case of mRNA half-life, the conserved 3' stem loop structure of major core and H1 histone transcripts has been implicated to be involved (see Section 1.11.6), although a mechanism has not been clearly elucidated. Work presented in this thesis has focussed on the other major determinant of S-phase control, the temporal control of histone gene transcription. It is not surprising that transcriptional control is a prime factor in this regard as it is the primary step in gene control and offers the simplest/most direct and energy efficient mode of co-ordinately expressing a family of genes.

Attempts have been made previously to characterize factors influencing the coupling of histone gene expression to DNA replication primarily by focussing on parameters which effect histone mRNA levels. One common approach has been to employ metabolic inhibitors which interfere with DNA or protein synthesis. Several putative control mechanisms have been proposed as a result of this approach. For instance, there is evidence that histone transcript levels are coupled to dNTP metabolism. Furthermore, feedback regulation of histones themselves or labile regulatory proteins have been suggested as regulators of histone biosynthesis via transcriptional regulation. Results from such experiments are difficult to interpret as the specificity of inhibitors has not been analyzed in detail.

Two fundamental questions have yet to be fully answered in vertebrate systems. When precisely is histone gene transcription initiated and, to what extent are S-phase controlled genes 'silenced' in the absence of DNA replication? Studies in yeast suggest that activation of histone gene transcription occurs in late G1-phase and essentially terminates in early S-phase (Hereford et al., 1982). The cell-cycle controlled yeast H0 gene also appears to be activated in late G1 also (Nasmyth, 1983). Similar results have been obtained using a temperature-sensitive hamster cell-line although transcription does not appear to be complete until mid-late S-phase (Artishevsky et al., 1984). Most other time-course experiments have either used synchronization procedures at the G1/S boundary, which are not really permissible to answer this question, or have not focussed on detailed analysis at the G1/S cell-cycle transition. In all probability then, initiation of histone gene transcription in vertebrates is likely to precede the onset of DNA replication just as it does in lower eukaryotes. The question of histone gene silencing in the absence of DNA synthesis is more difficult to answer in vertebrates. It is quite likely, however, that only a small 'basal' level of transcription actually exists in vivo for truly S-phase regulated genes. There are several reasons for this. First, gene-specific probes have not been generally employed in pulse-labelling experiments, making it impossible to determine if detectable transcription in the absence of DNA replication is due to a population of constitutively (non-S-phase controlled) expressed genes. For example, an H2A.1 probe in chicken may detect basal level activity of another H2A gene. Second, a potential source contributing to detectable histone gene transcription outside of S-phase is the small population of cells not correctly synchronized. For instance, it is estimated that approximately 5% of AEV ts34 cells do not become effectively synchronized at the G1/S border after treatment

with aphidicolin (see Chapter 3). Finally, a trend emerged in the course of experiments reported here which indicate that the relative level of transcription from exogenous chicken histone genes in the absence of DNA replication in transformed HeLa cells (see Chapter 4) is distinctly lower than endogenous levels of histone mRNAs, at corresponding times in AEV ts34 cells (see Chapter 3). The most likely explanation for this is that in HeLa cell-lines, expression of only one gene at a time was being detected (i.e. no replication-independent gene expression would be detected if the transfected gene itself was S-phase regulated). In support of this idea is that transcription of yeast histone genes occurs exclusively during late-G1/S-phase (Hereford et al., 1982). Therefore, rates of overall transcription in the absence of DNA replication may not necessarily be a true indication of the activity of replication-dependent genes.

Transcription in general is recognized to be generally influenced by regional chromatin structure. This phenomenon together with the fact that a close temporal relationship exists between histone gene expression and DNA replication implies that some regional chromosomal changes synonymous with replication may also coincide with those required for activation of histone gene transcription. One prominent idea relating to this has been the possibility that putative origins of DNA replication (autonomously replicating sequences, ars) associated with yeast histone genes are actually also required for their periodic expression (see Osley and Hereford, 1982) or, influence absolute transcript levels (Osley et al., 1986). A remarkably high incidence of ars activity has also been identified with histone genes of other species (Wilhelm et al., 1986; Dalton et al., manuscript in preparation) although several studies have failed to identify a regulatory role (Dalton et al., manuscript in preparation). The significance of putative origins of replication being closely associated with histone

and other cell-cycle regulated genes (see Miller and Nasmyth, 1984), is therefore unresolved.

A general feature of transcriptionally active genes is that they are proximally anchored to a sub-nuclear structure known as the nuclear matrix (see Chapter 6). These anchorage sites are thought to punctuate DNA loops into transcriptionally active domains, perhaps by stabilizing transcription complexes associated with cis-regulatory control elements (i.e. enhancers; see Cockerill et al., 1986; Gasser and Laemmli, 1985). Furthermore, fixation of DNA loops may also be required to present DNA templates in the correct conformation/supercoiled state. With respect to regulation of histone gene expression, it is likely that their attachment is an absolute requirement for their activity. This is supported by observations that there is an overwhelming coincidence of core and H1 histone gene-NM anchorage. Furthermore, this is highlighted by selective attachment of the H5 gene exclusively in the cell-line in which it is expressed. Fine mapping of these matrix anchorage regions (MAR) are likely to reveal that they coincide with important cis-regulatory elements. Already, MAR regions associated with immunoglobulin (Cockerill et al., in press) and developmentally regulated Drosophila genes (Gasser and Laemmli, 1985) have been shown to co-map with corresponding enhancer elements.

A possibility exists that transcriptional activation of histone genes in S-phase involves common elements which facilitate their coordinate expression. For instance, one control element may simultaneously activate several genes within a gene cluster. However, this seems unlikely, at least in vertebrate systems, as it appears that each histone gene has its own complement of cis-elements required for S-phase transcriptional control. This seems like a cumbersome way of controlling a clustered set of genes, especially in view of the situation in yeast where a core element 5'-GCGAAA-3' is found in the



upstream regions of H2A, H2B, H3 and H4 genes. However, it is clear that other cell-cycle controlled genes in yeast such as CDC9 and HO, do not contain the histone 'core' element (see Osley et al., 1986). Of the limited sequence data available on other vertebrate non-histone S-phase controlled genes such as, thymidine kinase (Kwoh and Engler, 1984), no universal sequence elements have been found which are potential (cell-cycle) control elements. The likelihood of S-phase specific transcriptional cis-regulators being gene-specific, or at least limited to a small number of genes, is reinforced by the observation that the H1-box is a cell-cycle specific control element not found in other histone genes.

Apart from the probability of there being multiple S-phase regulatory control elements in vertebrate histone genes, it is plausible that a complement of DNA sequence elements are in fact required. The complexities of S-phase regulatory control are well illustrated by the genetic approach used by Breeden and Nasmyth (1987) in delineating control mechanisms of the yeast HO gene. These studies have shown that repeats of the element CACGA<sub>4</sub> present in the 5'-upstream region are capable of conferring cell-cycle control on a heterologous gene. At least seven other genes have been identified by genetic complementation analysis to be involved in the overall regulation of HO and that two of these are required for CACGA<sub>4</sub>-mediated activation of transcription. Negative ('silencer') control elements in yeast histone genes also appear to act in concert with cell-cycle specific upstream activator elements (Osley et al., 1986) and in the control of HO (Sternberg et al., 1987).

By a biochemical approach, work presented here shows that the H1-specific promoter element is essential for periodic transcriptional activation of its cognate gene in vivo. The ability of this element to function in either orientation and from an upstream position is

reminiscent of classical viral and cellular enhancer elements. Whether this element can by itself confer cell-cycle regulation on a heterologous gene is currently being tested. It is not known whether non-histone S-phase controlled genes also possess similar control elements to the H1-box. Certainly, no histone-like promoter elements have been universally found in the 5'-upstream regions of other cell-cycle controlled genes. However, the H1-box 'core' sequence, 'AAACACA', is present in the promoter region of a cell-cycle controlled murine dihydrofolate reductase (DHFR) gene (see Farnham and Schimke, 1986) and may play a similar regulatory role in controlling transcription of this gene as it does for the H1 gene.

Sequences composing the H1-box bind a specific nuclear factor and mutations in this element which abolish S-phase control of transcription also preclude factor binding. It is reasonable to speculate that this nuclear factor (H1-SF) is prevented from complexing to the H1 promoter region in vivo in cells transformed with the gene containing the mutated element, thus accounting for loss of S-phase regulatory control. Base substitutions in the H2B-box which abolish binding of its sequence-specific factor, H2B-SF, do not entirely eliminate S-phase control indicating that the H1 and H2B-specific elements (and their respective binding factors) serve different functions. It is proposed that the H1-box/H1-SF complex forms an important step in the mechanism of H1 gene transcriptional activation during S-phase, particularly in view that binding levels fluctuate in parallel with rates of transcription.

A prominent feature of chicken H1 promoter architecture is the presence of highly conserved DNA sequence elements which are also conserved relative to their spatial arrangement. In particular, the spacer regions between the H1-box and G/C-box may be of significance as they are generally conserved to a high degree. In general, they also lie on the same helical face of DNA. This would facilitate direct

interactions between bound complexes at these two sites, a feature thought to be important for transcriptional regulation in many cases (see Section 1.7) and, which may be crucial for S-phase control of the H1 gene. Interestingly, the 'AAACAC' (H1-Box) element in the DHFR gene is also separated from a G/C-box by the same distance (generally) as in chicken H1 promoters and therefore lies on the same helical face of DNA, thus reinforcing the possibility of an interaction between these two elements and further establishing a link with S-phase transcriptional control. These elements are not on the same helical DNA face in all H1 genes, but these are exceptions, and can perhaps be accounted for by the fact that not all chicken H1 genes are S-phase regulated (Winter et al., 1985b).

## 7.2 FUTURE WORK AND PERSPECTIVES

One of the most perplexing questions facing molecular biologists at present is the mechanism by which trans-acting factors modulate transcription in a temporal, cell-specific or inducible fashion. In most cases, the parameters influencing factor binding to target sequences are unresolved. For instance, are levels of trans-regulatory proteins generally increased by de novo synthesis in order to effect gene activation or, are relatively stable proteins post-translationally modified to exert an effect? Studies to be performed on the nature of fluctuating H1-SF binding to the H1-box at different stages of the cell-cycle should contribute to an understanding of the underlying mechanisms controlling this aspect of gene regulation.

On a more general note, although the overall hierarchy of gene control is extremely complex, a finite number of control points must exist. For instance, one gene product (i.e. a trans-regulator) may modulate transcription of several co-regulated genes. This idea has recently been supported and extended by reports in yeast that a common

trans-regulator (HAP1 protein) binds to entirely different promoter elements in separate genes (Pfeifer et al., 1987). Conceivably, this could apply to a wide group of regulatory protein molecules which utilize different functional domains and are therefore multivalent with respect to DNA binding site target specificity.

Future work is now being directed towards purifying H1-SF to homogeneity by repeating DNA-affinity chromatography experiments. This will enable several avenues of experimental approach to be pursued. These include testing the transcriptional activatory role of this protein in vitro, determining the regulatory nature of the protein (i.e. binding, activatory domains), and ultimately cloning the gene so that in turn we will become one step closer to understanding the overall mechanism by which genes are periodically expressed throughout the cell-division cycle.

B I B L I O G R A P H Y

## BIBLIOGRAPHY

- Ackerman, S., Bunick, D., Zandomeni, R. and Weinmann, R. (1983) Nucl. Acids Res. 11: 6041-6064.
- Adams, J.M., Harris, H.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D. and Brinster, R.L. (1985) Nature 318: 533-538.
- Adolph, K.W., Cheng, S.M. and Laemmli, U.K. (1977) Cell 12: 805-816.
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J.D. (1983) in "Molecular Biology of the Cell". Garland Publishing Co., New York, U.S.A.: 456-467.
- Allis, C.D., Glover, C.V.C., Bowen, J.K. and Gorovsky, M.A. (1980) Cell 20: 609-617.
- Allis, C.D., Ziegler, Y.S., Gorovsky, M.A. and Olmsted, J.B. (1982) Cell 31: 131-136.
- Allis, C.D., Richman, R., Gorovsky, M.A., Ziegler, Y.S., Touchstone, B., Bradley, W.A. and Cook, R.G. (1986) J. Biol. Chem. 261: 1941-1948.
- Alterman, R.-B., Ganguly, S., Schulze, D.H., Marzluff, W.F. and Skoultchi, A.I. (1984) Mol. Cell. Biol. 4: 123-132.
- Alterman, R.-B., Sprecher, C., Graves, R., Marzluff, W.F. and Skoultchi, A.I. (1985) Mol. Cell. Biol. 5: 2316-2324.
- Ambrosio, L. and Schedl, P. (1985) Dev. Biol. 111: 220-231.
- Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981) Nature 290: 754-758.
- Angerer, L.M., DeLeon, D.V., Angerer, R.C., Showman, R.M., Wells, D.E. and Raff, R.A. (1984) Dev. Biol. 101: 477-484.
- Angerer, L., DeLeon, D., Cox, K., Maxson, R., Kedes, L., Kaumeyer, J., Weinberg, E. and Angerer, R. (1985) Dev. Biol. 112: 157-166.
- Annunziato, A.T. and Seale, R.L. (1983) J. Biol. Chem. 258: 12675-12684.

- Artishevsky, A., Ganguly, S., Schulze, D.H., Marzluff, W.F.,  
Schildkraut, C.L. and Skoultchi, A.I. (1984) *Mol. Cell. Biol.* 4:  
123-132.
- Artishevsky, A., Grafsky, A. and Lee, A.S. (1985) *Science* 230: 1061-  
1063.
- Atchison, M.L. and Perry, R.P. (1986) *Cell* 46: 253-262.
- Atchison, M.L. and Perry, R.P. (1987) *Cell* 48: 121-128.
- Augereau, P. and Chambon, P. (1986) *EMBO J.* 5: 1791-1797.
- Aviles, F.J., Chapman, G.E., Kneale, G.G., Crane-Robinson, C. and  
Bradbury, E.M. (1978) *Eur. J. Biochem.* 88: 363-371.
- Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* 69: 1408-  
1412.
- Banerji, J., Rusconi, S. and Schaffner, W. (1981) *Cell* 27: 299-308.
- Banerji, J., Olson, L. and Schaffner, W. (1983) *Cell* 33: 729-740.
- Bannon, G.A., Calzone, F.J., Bowen, J.K., Allis, C.D. and Gorovsky, M.A.  
(1983) *Nucl. Acids Res.* 11: 3903-3917.
- Bannon, G.A., Bowen, J.K., Yao, M.C. and Gorovsky, M.A. (1984) *Nucl.*  
*Acids Res.* 12: 1961-1975.
- Baumbach, L.L., Marashi, F., Plumb, M., Stein, G. and Stein, J. (1984)  
*Biochemistry* 23: 1618-1625.
- Bazett-Jones, D.P., Yeckel, M. and Gottesfeld, J.M. (1985) *Nature* 317:  
824-828.
- Becker, M.M. and Wang, J.C. (1984) *Nature* 309: 682-687.
- Beier, D.R. and Young, E.T. (1982) *Nature* 300: 724-728.
- Bellard, M., Gannon, F. and Chambon, M. (1977) *Cold Spring Harb. Symp.*  
*Quant. Biol.* 42: 779-791.
- Belyavsky, A.V., Bavykin, S.G., Goduadze, E.G. and Mirzabekov, A.D.  
(1980) *J. Mol. Biol.* 139: 519-536.
- Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) *Nucleic*  
*Acids Res.* 8: 127-142.

- Benoist, C. and Chambon, P. (1981) *Nature* 290: 304-310.
- Benton, W.D. and Davis, R.W. (1977) *Science* 196: 180-182.
- Benyajati, C. and Worcel, A. (1976) *Cell* 9: 393-407.
- Berdikov, V.A., Gorel, C., Argutinskaya, S.V., Cherepanova, V.A. and Kileva, E.V. (1975) *Mol. Biologiya* 10: 887-896.
- Berezney, R. and Bucholtz, L.A. (1981) *Exp. Cell. Res.* 132: 1-13.
- Beug, H., von Kirchbach, A., Doderlein, G., Couscience, J.-F. and Graf, T. (1979) *Cell* 18: 375-390.
- Beug, H., Doederlein, G., Freudenstein, C. and Graf, T. (1982) *J. Cell. Physiol.* 1: 195-207.
- Bienz, M. (1986) *Cell* 46: 1037-1042.
- Bienz, M. and Pelham, H.R.B. (1986) *Cell* 45: 753-760.
- Birchmeier, C., Grosschedl, R. and Birnstiel, M.L. (1982) *Cell* 28: 739-745.
- Birchmeier, C., Folk, W. and Birnstiel, M.L. (1983) *Cell* 35: 433-440.
- Birchmeier, C., Schümperli, D., Sconzo, G. and Birnstiel, M.L. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 1057-1061.
- Bird, R.C., Jacobs, F.A., Stein, G., Stein, J. and Sells, B.H. (1985) *Biochim. Biophys. Acta* 824: 209-217.
- Birnboim, H.C. and Doly, J. (1979) *Nucl. Acids Res.* 7: 1513-1523.
- Birnstiel, M.L., Portmann, R., Busslinger, M., Schaffner, W., Probst, E. and Kressman, A. (1978) *Alfred Benzon Symp.* 13: 117-129.
- Birnstiel, M.L. (1984) *Philos. Trans. R. Soc. Lond. B.* 307: 293-295.
- Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) *Cell* 41: 349-359.
- Blankstein, L.A. and Levy, S.B. (1976) *Nature* 260: 638-641.
- Bloom, K.S. and Anderson, J.N. (1978) *Cell* 15: 141-150.
- Bochmann, D., Keller, W., Dale, T., Schöler, H.R., Tebb, G. and Mattaj, I.W. (1987) *Nature* 325: 268-272.
- Bogenghagen, D.F., Wormington, W.M. and Brown, D.D. (1982) *Cell* 28: 413-421.



- Bolivar, F. (1978) *Gene* 4: 121-126.
- Borun, T.W., Scharff, M.D. and Robbins, E. (1967) *Proc. Natl. Acad. Sci. U.S.A.* 58: 1977-1983.
- Boshart, M., Weber, F., Dorsch-Hasler, J.G., Fleckenstin, B. and Schaffner, W. (1986) *Cell* 41: 521-530.
- Bradford, M.M. (1976) *Analyt. Biochem.* 72: 248-254.
- Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J. and Salzman, N.P. (1982) *Cell* 31: 625-633.
- Brand, A.H., Breeden, L., Abraham, J., Sternglanz, R. and Nasmyth, K. (1985) *Cell* 41: 41-48.
- Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50: 349-383.
- Breeden, L. and Nasmyth, K. (1987) *Cell* 48: 389-397.
- Brent, R. and Ptashne, M. (1985) *Cell* 43: 729-736.
- Briggs, M.R., Kadonaga, J.T., Bell, S.P. and Tjian, R. (1986) *Science* 234: 47-52.
- Brown, D.D. (1984) *Cell* 37: 359-365.
- Brown, D.D. and Schlissel, M.S. (1985) *Cell* 42: 759-767.
- Brown, D.T., Wellman, S.E. and Sittman, D.B. (1985) *Mol. Cell. Biol.* 5: 2879-2886.
- Brush, D., Dodgson, J.B., Choi, O.K.R., Stevens, P.W. and Engel, J.D. (1985) *Mol. Cell. Biol.* 5: 1307-1317.
- Buongiorno-Nardelli, M., Michele, G., Carrie, M.T. and Marilley, M. (1982) *Nature* 298: 100-103.
- Busby, S., Irani, M. and deCrombrughe, B. (1982) *J. Mol. Biol.* 154: 197-209.
- Bushman, F.D., Anderson, F.E., Harrison, S.C. and Ptashne, M. (1985) *Nature* 316: 651-653.
- Busslinger, M., Portmann, R., Irminger, J.C. and Birnstiel, M.L. (1980) *Nucl. Acids. Res.* 8: 957-977.

- Campo, M.S., Spandidos, D.A., Lang, J. and Wilkie, N.M. (1983) *Nature* 303: 77-80.
- Capasso, O. and Heintz, N. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82: 5622-5626.
- Capecchi, M.R. (1980) *Cell* 22: 479-488.
- Carozzi, N., Marashi, F., Plumb, M., Zimmerman, S., Zimmerman, A., Coles, L.S., Wells, J.R.E., Stein, G. and Stein, J. (1984) *Science* 224: 1115-1117.
- Cartwright, I.L., Abmayr, S.M., Fielschmann, G.M., Lowenhaupt, K., Elgin, S.C.R., Keene, M.A. and Howard, G.C. (1982) *CRC Crit. Rev. Biochem.* 13: 1-85
- Cary, P.D., Hines, M.L., Bradbury, E.M., Smith, B.J. and Johns, E.W. (1981) *Eur. J. Biochem.* 120: 371-377.
- Casadaban, M.J. and Cohen, S.N. (1980) *J. Mol. Biol.* 13: 179-207.
- Cereghini, S., Sargosti, S., Yaniv, M. and Hamer, D.W. (1984) *Eur. J. Biochem.* 144: 545-553.
- Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984) *Cell* 38: 251-263.
- Charnay, P., Mellon, P. and Maniatis, T. (1985) *Mol. Cell. Biol.* 5: 1498-1511.
- Childs, G., Maxson, R., Cohn, R.H. and Kedes, L.H. (1981) *Cell* 23: 651-663.
- Choe, J., Schuster, T., and Grunstein, M. (1985) *Mol. Cell. Biol.* 5: 3261-3269.
- Choi, O.-R. and Engel, J.D. (1986) *Nature* 323: 731-734.
- Chrysogelos, S., Riley, D.E., Stein, G., and Stein, J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82: 7535-7539.
- Ciejek, E.M., Nordstrom, J.M., Tsai, M.-J. and O'Malley, B.W. (1982) *Biochemistry* 21: 4945-4953.

- Cizewski, V. and Sollner-Webb, B. (1983) Nucl. Acids. Res. 11: 7043-7056.
- Cizewski Culotta, V., Wides, R.J. and Sollner-Webb, B. (1985) Mol. Cell. Biol. 5: 1582-1590.
- Clerc, R.G., Buchner, P., Strub, K. and Birnstiel, M.L. (1983) Nucl. Acids. Res. 11: 8641-8657.
- Cleveland, D.W., Lopata, M.A., MacDonald, R.J., Cowan, N.J., Rutter, W.J. and Kirchner, M.W. (1980) Cell 20: 95-105.
- Cockerill, P.N. and Garrard, W.T. (1986) Cell 44: 273-282.
- Cockerill, P.N., Yuen, M.-H. and Garrard, W.T. J. Biol. Chem. (In press).
- Cohen, R.B. and Sheffery, M. (1985) J. Mol. Biol. 182: 109-129.
- Cohen, R.B., Sheffery, M. and Kim, C.G. (1986) Mol. Cell. Biol. 6: 821-832.
- Cole, R.D. (1984) Anal. Biochem. 136: 24-30.
- Cole, R.D., York, R.G. and Kistler, W.S. (1986) Biochem. Biophys. Acta. 869: 223-229.
- Coles, L.S and Wells, J.R.E. (1985) Nucl. Acids Res. 13: 585-594.
- Coles, L.S. (1986) Ph.D. Thesis, University of Adelaide, Adelaide, South Australia.
- Coles, L.S., Robins, A.J., Madley, L.K. and Wells, J.R.E., Submitted for publication.
- Colman, A., Krieg, P. and Wells, J.R.E. (1983) Eur. J. Biochem. 131: 551-554.
- Comb, M., Birnberg, N.C., Seasholtz, A., Herbert, E. and Goodman, H.M. (1986) Nature 323: 353-356.
- Conner, W., Meqzuita, J., Winkfein, R.J., Staes, J.C. and Dixon, G.H. (1984) J. Mol. Evol. 20: 227-235.
- Cook, P.R. and Brazell, I.A. (1976) J. Cell. Sci. 22: 287-302.
- Cooke, N.E. and Baxter, J.D. (1982) Nature 292: 603-606.

- Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980) *Science* 209: 1406-1414.
- Cox, K.H., DeLeon, D.V., Angerer, L.M. and Angerer, R.C. (1984) *Dev. Biol.* 101: 485-502.
- Dailey, L., Hanly, S.M., Roeder, R.G and Heintz, N. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 7241-7245.
- Dalton, S., Coleman, J.R. and Wells, J.R.E. (1986a) *Mol. Cell. Biol.* 6: 601-606.
- Dalton, S., Younghusband, H.B. and Wells, J.R.E. (1986b) *Nucl. Acids Res.* 14: 6507-6523.
- Dalton, S., Shannon, M.F and Wells, J.R.E., Manuscript in Preparation.
- Dandanell, G. and Hammer, K. (1985) *EMBO J.* 4: 3333-3338.
- D'Andrea, R.J. (1985) Ph.D. Thesis, University of Adelaide, Adelaide, South Australia.
- D'Andrea, R.J., Coles, L.S., Lesnikowski, C., Tabe, L. and Wells, J.R.E. (1985) *Mol. Cell. Biol.* 5: 3108-3115.
- Darnell, J.E. (1982) *Nature* 297: 365-371.
- Davidson, I., Fromental, C., Augereau, P., Wildeman, A., Zenke, M. and Chambon, P. (1986) *Nature* 323: 544-548.
- Davison, B.L., Egly, J.-M., Mulvihill, E.R. and Chambon, P. (1983) *Nature* 301: 680-686.
- DeLeon, D.V., Cox, K.H., Angerer, L.M. and Angerer, R.C. (1983) *Dev. Biol.* 100: 197-200.
- DeLisle, A.J., Graves, R.A., Marzluff, W.F. and Johnson, L.F. (1983) *Mol. Cell. Biol.* 3: 1920-1929.
- DeLorenzi, M., Rohrer, U. and Birnstiel, M.L. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 3243-3247.
- de Villiers, J. and Schaffner, W. (1981) *Nucl. Acids Res.* 9: 6251-6264.
- de Villiers, J., Olson, L., Tyndall, C. and Schaffner, W. (1982) *Nucl. Acids Res.* 20: 7965-7976.

- de Villiers, J., Schaffner, W., Tyndall, C., Lupton, S. and Kamen, R.  
(1984) *Nature* 312: 242-246.
- Diakun, G.P., Fairall, L. and Klug, A. (1986) *Nature* 324: 698-699.
- Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and  
Weissman, C. (1983) *Cell* 32: 695-706.
- Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucl. Acids. Res.*  
11: 1475-1489.
- Dulbecco, R. and Vogt, M. (1954) *J. Exptl. Med.* 99: 167-172.
- Dunn, T., Hahn, S., Ogden, S. and Schlieff, R. (1984) *Proc. Natl. Acad.*  
*Sci. U.S.A.* 81: 5017-5020.
- Dynan, W.S. and Tjian, R. (1985) *Nature* 316: 774-778.
- Ebright, R.H. (1982) in "Molecular Structure and Biological Function".  
Griffin, J. and Duax, W., editors, Elsevier/North Holland, New  
York: 91-102.
- Elkaim, R., Goding, C. and Kedinger, C. (1983) *Nucl. Acids. Res.* 11:  
7105-7117.
- Emorine, L., Kuehl, M., Weir, L., Leder, P. and Max, E.E. (1983) *Nature*  
304: 447-449.
- Engel, J.D., Sugarman, B.J. and Dodgson, J.B. (1982) *Nature* 297: 434-  
436.
- Engelke, D.R., Ng, S.-Y., Shastry, B.S. and Roeder, R.B. (1980) *Cell* 19:  
717-728.
- Etkin, L.D. and Maxson, R.E (1980) *Dev. Biol.* 75: 13-25.
- Etkin, L.D. and Di Berardino, M.A. (1983) in "Eukaryotic Genes, their  
Structure, Activity and Regulation". MacLean, N., Gregory, S.P.  
and Flavell, R.A., editors, Cambridge, U.K.: 127-156.
- Everett, R. (1983) *Nucl. Acids Res.* 11: 6647-6666.
- Falkner, F.G. and Zachau, H.G. (1984) *Nature* 310: 71-74.
- Farnham, P.H. and Schimke, R.T. (1986) *Mol. Cell. Biol.* 6: 2392-2401.
- Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* 132: 6-13.

- Foster, J., Stafford, J. and Queen, C. (1985) *Nature* 315: 423-425.
- Franklin, S.G. and Zweidler, A. (1977) *Nature* 266: 273-275.
- Gabrielli, F., Hancock, R. and Faber, A.J. (1981) *Eur. J. Biochem.* 120: 363-369.
- Galas, D. and Schmitz, A. (1978) *Nucl. Acids Res.* 5: 3157-3170.
- Galli, G., Hofstetter, H., Stunnenberg, H.G. and Birnstiel, M.L. (1983) *Cell* 34: 823-828.
- Gallwitz, D. and Mueller, G.C. (1969) *J. Biol. Chem.* 244: 5947-5952.
- Garabedian, M.J., Shepherd, B.M. and Wensink, P.C. (1986) *Cell* 45: 859-867.
- Garcia, J.V., Le thi Bich-Thuy., Stafford, J. and Queen, C. (1986) *Nature* 322: 383-385.
- Garel, A. and Axel, R. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73: 3966-3970
- Gasser, S.M. and Laemmlí, U.K. (1986) *Cell* 46: 521-530.
- Gehring, W.J. (1985) *Cold Spring Harbor Symp. Quant. Biol.* 50: 243-252.
- Georgiev, O. and Birnstiel, M.L. (1985) *EMBO J.* 4: 481-489.
- Gerlinger, P., Le Meur, M., Irrmann, C., Renard, P., Wasylyk, C. and Wasylyk, B. (1986) *Nucl. Acids Res.* 14: 6565-6577.
- Gick, O., Kramer, A., Keller, W. and Birnstiel, M.L. (1986) *EMBO J.* 5: 1319-1326.
- Gidoni, D., Dynan, W.S. and Tjian, R. (1984) *Nature* 312: 409-413.
- Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) *Cell* 33: 717-728.
- Gillies, S.D., Folson, V. and Tonegawa, S. (1984) *Nature* 310: 594-597.
- Gilmour, R.S., Spandidos, D.A., Vass, J.K., Gow, J.W. and Paul, J. (1984) *EMBO J.* 3: 1263-1272.
- Giniger, E., Varnum, S.M. and Ptashne, M. (1985) *Cell* 40: 767-774.
- Gjerset, R., Goraka, C., Hasthorpe, S., Lawrence, J.J. and Eisen, H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 403-427.

- Goldberg, M.L. (1979) Ph.D. Thesis, Stanford University, Palo Alto, California, U.S.A.
- Goodbourn, S., Zinn, K. and Maniatis, T. (1985) *Cell* 41: 509-520.
- Goodwin, R.G., Moncman, C.L., Rottman, F.M. and Nilson, J.H. (1983) *Nucl. Acids. Res.* 11: 6873-6882.
- Gottesfeld, J. and Bloomer, L.S. (1982) *Cell* 28: 781-791.
- Gough, M.N., Metcalf, D., Gough, J., Grail, D. and Dunn, A.R. (1985) *EMBO J.* 4: 645-653.
- Graf, T., Royer-Pakora, B., Schubert, G.E. and Beug, H. (1976) *Virology* 71: 423-433.
- Graham, F.L. and van der Eb, A.J. (1973) *Virology* 52: 456-467.
- Graves, R.A. and Marzluff, W.F. (1984) *Mol. Cell. Biol.* 4: 351-357.
- Graves, R.A., Marzluff, W.F., Giebelhaus, D.M. and Schultz, G.A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82: 5685-5689.
- Graves, R.A., Pandey, N.B., Chodchoy, N. and Marzluff, W.F. (1987) *Cell* 48: 615-626.
- Green, L., Schlaffer, I., Wright, K., Moreno, M.L., Berand, D., Hager, G., Stein, J. and Stein, G. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 2315-2319.
- Grichnik, J.M., Bergsma, D.J. and Schwartz, R.J. (1986) *Nature* 14: 1683-1761.
- Griffith, J., Hochschild, A., and Ptashne, M. (1986) *Nature* 322: 750-752.
- Grosschedl, R. and Birnstiel, M.L. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77: 1432-1436.
- Grosschedl, R., Mächler, M., Rohrer, V. and Birnstiel, M. (1983) *Nucl. Acids. Res.* 11: 8123-8136.
- Grosschedl, R. and Baltimore, D. (1985) *Cell* 41: 885-897.
- Grosveld, G.C., de Boer, E., Shewmaker, C.K. and Flavell, R.A. (1982) *Nature* 295: 120-126.

- Grove, G.W. and Zweidler, A. (1984) *Biochemistry* 23: 4436-4443.
- Gruss, P., Dhar, R. and Khoury, G. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78: 943-947.
- Guarente, L., Nye, J.S., Hochschild, A. and Ptashne, M. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 2236-2240.
- Guarente, L., Lalonde, B., Gifford, P. and Alani, E. (1984) *Cell* 36: 503-511.
- Hanahan, D. (1985) *Nature* 315: 115-122.
- Hanly, S.M., Bleecker, G.C. and Heintz, N. (1985) *Mol. Cell. Biol.* 5: 380-389.
- Hartshorne, T.A., Blumberg, H. and Young, E.T. (1986) *Nature* 320: 283-287.
- Harvey, R.P., Krieg, P.A., Robins, A.J., Coles, L.S. and Wells, J.R.E. (1981) *Nature* 294: 49-53.
- Harvey, R.P. (1982) Ph.D. Thesis, University of Adelaide, Adelaide, South Australia.
- Harvey, R.P., Robins, A.J. and Wells, J.R.E. (1982) *Nucl. Acids Res.* 10: 7851-7863.
- Harvey, R.P., Whiting, J.A., Coles, L.S., Krieg, P.A. and Wells, J.R.E. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80: 2819-2823.
- Hayashi, S., Kondoh, H., Yasuda, K., Soma, G., Ikawa, Y. and Okada, T.S. (1985) *EMBO J.* 4: 2201-2207.
- Hearing, P. and Shenk, T. (1983) *Cell* 33: 695-703.
- Heintz, N., Zernick, M. and Roeder, R.G. (1981) *Cell* 24: 661-668.
- Heintz, N., Sive, H.L. and Roeder, R.G. (1983) *Mol. Cell. Biol.* 3: 539-550.
- Hen, R., Borrelli, E., Sassone-Corsi, P. and Chambon, P. (1983) *Nucl. Acids. Res.* 11: 8747-8760.
- Hentschel, C., Irminger, J.-C., Bucher, P. and Birnstiel, M.L. (1980) *Nature* 285: 147-151.



- Hentschel, C.C. and Birnstiel, M.L. (1981) *Cell* 25: 301-313.
- Hentzen, P.C., Rho, J.H. and Bekhor, I. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 304-307.
- Hereford, L., Fahrner, K., Woolford, J., Rosbach, M. and Kaback, D.B. (1979) *Cell* 18: 1261-1271.
- Hereford, L.M. and Osley, M.A. (1981) *Cell* 24: 367-375.
- Hereford, L.M., Bromley, S. and Osley, M.A. (1982) *Cell* 30: 305-310.
- Herman, R., Weymouth, L. and Penman, S. (1978) *J. Cell Biol.* 78: 663-674.
- Hermbomel, P., Bourachot, B. and Yaniv, M. (1984) *Cell* 39: 653-662.
- Herr, W. and Clarke, J. (1986) *Cell* 45: 461-470.
- Hess, J.L., Clements, J.E. and Narayan, O. (1985) *Science* 229: 482-485.
- Hesse, J.E., Nickol, J.M., Lieber, M.R. and Felsenfeld, G. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 4312-4316.
- Hill, D.E., Hope, I.A., Macke, J.P., and Struhl, K. (1986) *Science* 234: 451-457.
- Hiromi, Y., Kuriowa, A. and Gehring, W.J. (1985) *Cell* 43: 603-613.
- Hochschild, A. and Ptashne, M. (1986) *Cell* 44: 681-687.
- Hope, I.A. and Struhl, K. (1985) *Cell* 43: 177-188.
- Hope, I.A. and Struhl, K. (1986) *Cell* 46: 885-894.
- Huberman, J.A. (1981) *Cell* 23: 647-648.
- Igo-Kemenes, T. and Zachau, J. (1978). *Cold Spring Harb. Symp. Quant. Biol.* 42: 109-118.
- 'Instructions to Authors' (1978) *Biochem. J.* 169: 1-27.
- Irani, M.H., Orosz, L. and Adhya, S. (1983) *Cell* 32: 783-788.
- Isenberg, I. (1979) *Ann. Rev. Biochem.* 48: 159-191.
- Jack, R.S., Gehring, W.J. and Brack, C. (1981) *Cell* 24: 321-331.
- Jackson, D.A. and Cook, P.R. (1985) *EMBO J.* 4: 919-925.
- Jilbert, A.R., Burrell, C.J., Gowans, E.J. and Rowland, R. (1986) *Histochem.* 85: 505-514.

- Jolly, D.J., Esty, A.C., Subramani, S., Friedmann, T. and Verma, I.M.  
(1983) Nucl. Acids Res. 11: 1855-1872.
- Jones, E.W. and Fink, G.R. (1982) In: 'The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression'. Strathern, E.W., Jones, E.W. and Broach, J.R., eds. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory: 181-299.
- Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985) Cell 42: 559-572.
- Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R. (1986) Science 232: 755-759.
- Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. and Tjian, R. (1987) Cell 48: 79-89.
- Jost, J.-P. and Seldran, M. (1984) EMBO J. 3: 2005-2008.
- Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) Trends in Biochem. Sci. 11: 20-23.
- Kadonaga, J.T. and Tjian, R. (1986) Proc. Natl. Acad. Sci. U.S.A. 83: 5889-5893.
- Kedes, L.H. (1979) Ann. Rev. Biochem. 48: 837-870.
- Keegan, L., Gill, G. and Ptashne, M. (1986) Science 231: 699-703.
- Kelly, J.H. and Darlington, G.J. (1985) Ann. Rev. Genet. 19: 273-296.
- Keppel, F. (1986) J. Mol. Biol. 187: 15-21.
- Khillan, J.S., Schmidt, A., Overbeek, P.A., de Crombrughe, B. and Westphal, H. (1986) Proc. Natl. Acad. Sci. U.S.A. 83: 725-729.
- Kmieciak, E.B., Razvi, F. and Worcel, A. (1986) Cell 45: 209-218.
- Krieg, P.A., Robins, A.J., Colman, A. and Wells, J.R.E. (1982) Nucl. Acids. Res. 10: 6778-6785.
- Krieg, P.A., Robins, A.J., D'Andrea, R. and Wells, J.R.E. (1983) Nucl. Acids. Res. 11: 619-627.
- Krieg, P.A. and Melton, D.A. (1984) Nature 308: 203-206.
- Krieg, P.A. and Melton, D.A. (1987) Proc. Natl. Acad. Sci. U.S.A., in press.

- Krumlauf, R., Hammer, R.E., Tilghman, S.M. and Brinster, R.L. (1985)  
Mol. Cell. Biol. 5: 1639-1648.
- Kwoh, T.J. and Engler, J.A. (1984) Nucl. Acids Res. 12: 3959-3971.
- Laemmli, U.K., Cheng, S.M., Adolph, K.W., Paulson, J.R., Brown, J.A. and  
Baumbach, W.R. (1978) Cold Spring Harb. Symp. Quant. Biol. 42:  
109-118.
- Laimins, L.A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982)  
Proc. Natl. Acad. Sci. U.S.A. 79: 6453-6457.
- Laimins, L., Holmgren-König, M. and Khoury, G. (1986) Proc. Natl. Acad.  
Sci. U.S.A. 83: 3151-3155.
- Larsen, J.E.L. and Hammer, K. (1987) Nature 325: 823-826.
- Lassar, A.B., Martin, P.L. and Roeder, R.G. (1983) Science 222: 740-748.
- Laughton, A. and Scott, M.P. (1984) Nature 310: 25-31.
- Learned, R.M., Learned, T.K., Haltiner, M.M. and Tjian, R.T. (1986) Cell  
45: 847-857.
- Lebkowski, J.S. and Laemmli, U.K. (1982) J. Mol. Biol. 156: 309-324.
- Lee, F., Mulligan, R., Berg, P. and Ringold, G. (1981) Nature 294: 228-  
232.
- Lee, W., Haslinger, A., Karin, M. and Tjian, R. (1987) Nature 325: 368-  
372.
- Lennox, R.W. and Cohen, L.H. (1984) in 'Histone Genes, Structure,  
Organization and Regulation'; Stein, G.S., Stein, J.L. and  
Marzluff, W.F., eds., J. Wiley and Sons, New York: 373-395.
- Levinson, B., Khoury, G., Van de Woude, G. and Gruss, P. (1982) Nature  
295: 568-572.
- Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1977) Cold  
Spring Harbour Symp. Quant. Biol. 42: 1047-1051.
- Lüscher, B., Stauber, C., Schindler, R. and Schumperli, D. (1985) Proc.  
Natl. Acad. Sci. U.S.A. 82: 4389-4393.

- Lusky, M., Berg, L., Weiher, H., and Botchan, M. (1983) *Mol. Cell. Biol.* 3: 1108-1122.
- Majumdar, A. and Adhya, S. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 6100-6104.
- Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72: 1184-1188.
- Maniatis, T., Fritsch, E.F. and Sambrook J. (1982) "Molecular Cloning (a Laboratory Manual)", Cold Spring Harbour Laboratory, New York.
- Manley, J.L., Fire, A., Cano, A., Sharp, P.A. and Gefter, M.L. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77: 3855-3859.
- Marsden, M. and Laemmli, U.K. (1979) *Cell* 17: 849-858.
- Martin, K., Hvo, L. and Schlieff, R.F. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 3654-3658.
- Marzluff, W.F. (1978) *Methods Cell Biol.* 19: 317-331.
- Marzluff, W.F. and Graves, R.A. (1984) in "Histone Genes, Structure, Organization and Regulation"; Stein, G.S., Stein, J.L. and Marzluff, W.F. eds., J. Wiley and Sons, New York: 281-315.
- Matsumoto, S. and Yanagida, M. (1985) *EMBO J.* 4: 3531-3538.
- Mattaj, I.W., Lienhard, S., Jirieny, J. and DeRobertis, E.M. (1985) *Nature* 316: 163-167.
- Maxam, A.M. and Gilbert, W. (1980) *Methods in Enzymol.* 65: 499-560.
- Maxson, R.E. and Wilt, F.H. (1982) *Dev. Biol.* 94: 435-440.
- Maxson, R., Cohn, R. and Kedes, L. (1983a) *Ann. Rev. Genet.* 17: 239-279
- Maxson, R., Mohun, T., Gormezano, G., Childs, G. and Kedes, L. (1983b) *Nature* 301: 120-125.
- McCall, M., Brown, T., Hunger, W.N. and Kennard, O. (1986) *Nature* 322: 661-664.
- McClure, W.R. (1985) *Ann. Rev. Biochem.* 54: 171-205.
- McGhee, J.D. and Felsenfeld (1980) *Ann. Rev. Biochem.* 49: 1115-1156.

- McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. and Gehring, W.J.  
(1984) *Cell* 37: 403-408.
- McKnight, S.L., Gavis, E.R., Kingsbury, R. and Axel, R. (1981) *Cell* 25:  
385-398.
- McKnight, S.L. and Kingsbury, R. (1982) *Science* 217: 316-324.
- McKnight, S.L., Kingsbury, R., Spence, A. and Smith, M. (1984) *Cell* 37:  
253-262.
- McKnight, S., Kingsbury, R., Eisenberg, S., Tufaro, F., Weinheimer, S.,  
Triezenberg, S., Johnson, P. and Graves, B. (1985) in "Eukaryotic  
Transcription" Cold Spring Harbor Laboratory: 116-122.
- McKnight, S.L. and Tjian, R. (1986) *Cell* 46: 795-805.
- Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981) *Cell* 27:  
279-288.
- Melton, D.W., Konecki, D.S., Brennand, J. and Caskey, C.T. (1984) *Proc.*  
*Natl. Acad. Sci. U.S.A.* 81: 2147-2151.
- Mercola, M., Wang, X.-F., Olsen, J. and Calame, K. (1983) *Science* 221:  
663-665.
- Messing, J. (1979) *Recomb. DNA Tech. Bull.* 2: 43-48.
- Messing, J. and Vieira, J. (1982) *Gene* 19: 269-276.
- Messing, J. (1983) in "Methods in Enzymology" 101 (part C). Wu, R.,  
Grossman, L. and Moldave, K. eds., Academic Press, New York.
- Messing, A., Chen, H.Y., Palmiter, R.D. and Brinster, R.L. (1985) *Nature*  
316: 461-463.
- Mezquita, J., Connor, W., Winkfein, R.J. and Dixon, G.H. (1985) *J. Mol.*  
*Evol.* 21: 209-219.
- Miesfeld, R. and Arnheim, N. (1984) *Mol. Cell. Biol.* 4: 221-227.
- Miki, B.L.A. and Neelin, J.M. (1977) *Can. J. Biochem.* 55: 1220-1227.
- Miller, T.E., Huang, C.-Y. and Pogo, A.O. (1978) *J. Cell Biol.* 76: 675-  
691.
- Miller, A.M. and Nasmyth, K.A. (1984) *Nature* 312: 247-251.

- Miller, J., MacLachlan, A.D. and Klug, A. (1985) *EMBO J.* 4: 1609-1614.
- Mirkovitch, J., Mirault, M.-E. and Laemmli, U.K. (1984) *Cell* 39: 223-232.
- Miskimins, W.K., Roberts, M.P., McClelland, A. and Ruddle, F.H. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82: 6741-6744.
- Miyamoto, N.G., Moncollin, V., Wintzerith, M., Hen, R., Egly, J.M. and Chambon, P. (1984) *Nucl. Acids. Res.* 12: 8779-8799.
- Molgaard, H.V., Perucho, M. and Ruiz-Carillo, A. (1980) *Nature* 283: 502-504.
- Molloy, P.L., Powell, B.C., Gregg, K., Barone, E.D. and Rogers, G.E. (1982) *Nucl. Acids Res.* 10: 6007-6021.
- Moorman, A.F.M., de Boer, P.A.J., de Laaf, R.T.M. and Destree, O.H.J. (1982) *FEBS Letts.* 144: 235-241.
- Moreau, P., Hen, R., Everett, R., Gaub, M.P. and Chambon, P. (1981) *Nucl. Acids. Res.* 9: 6047-6068.
- Morris, T., Marashi, F., Weber, L., Hickey, E., Greenspan, D., Bonner, J., Stein, J. and Stein, G. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83: 981-985.
- Mossing, M.C. and Record, M.T. (1986) *Science* 233: 889-892.
- Mosthaf, L., Pawlita, M. and Gruss, P. (1985) *Nature* 315: 597-600.
- Mous, J., Stunnenberg, H., Georgiev, O. and Birnstiel, M.L. (1985) *Mol. Cell. Biol.* 5: 2764-2769.
- Myers, R.M., Tilly, K. and Maniatis, T. (1986) *Science* 232: 613-618.
- Nasmyth, K.A. (1983) *Nature* 302: 670-676.
- Neelin, J.M. (1968) *Can. J. Biochem.* 46: 241-247.
- Nelson, C., Crenshaw, E.B., Franco, R., Lira, S.A., Albert, V.R., Evans, R.M. and Rosefeld, M.G. (1986) *Nature* 322: 557-562.
- Nevins, J.R. (1983) *Ann. Rev. Biochem.* 52: 441-446.
- Nordheim, A. and Rich, A. (1983) *Nature* 303: 674-679.

- Noyes, B.A., Maverech, M., Stein, R. and Agarwal, K.L. (1979) Proc. Natl. Acad. Sci. U.S.A. 76: 1770-1774.
- Ogden, S., Haggerty, D., Stoner, C., Kolodrubetz, D. and Schlieff, R. (1980) Proc. Natl. Acad. Sci. U.S.A. 77: 3346-3350.
- Old, R.W. and Woodland, H.R. (1984) Cell 38: 624-626.
- Oppenheim, D.S., Bennett, G.N. and Yanofsky, C. (1980) J. Mol. Biol. 144:133-142.
- Osborne, T.F., Goldstein, J.L. and Brown, M.S. (1985) Cell 42: 203-212.
- Osley, M.A. and Hereford, L. (1982) Proc. Natl. Acad. Sci. U.S.A. 79: 7689-7693.
- Osley, M.A., Gould J., Kim, S., Kane, M. and Hereford, L. (1986) Cell 45: 537-544.
- Overbeek, P.A., Chepelinsky, A.B., Khillan, J.S., Piatigorsky, J. and Westphal, H. (1985) Proc. Natl. Acad. Sci. U.S.A. 82: 7815-7819.
- Pabo, C.O. and Lewis, M. (1982) Nature 298: 443-447.
- Palmiter, R.D., Chen, H.Y., Messing, A. and Brinster, R.L. (1985) Nature 316: 457-460.
- Palmiter, R.D. and Brinster, R.L. (1986) Ann. Rev. Genet. 20: 465-499.
- Pardoll, D.M. and Vogelstein, B. (1980) Exp. Cell Res. 128: 446-470.
- Pardoll, D.M., Vogelstein, B. and Coffey, D.S. (1980) Cell 19: 527-536.
- Parker, C.S. and Topol, J. (1984) Cell 36: 357-369.
- Parslow, T.G., Blair, D.L., Murphy, W.J. and Granner, D.K. (1984) Proc. Natl. Acad. Sci. U.S.A. 81: 2650-2654.
- Paulson, J.R. and Laemmli, U.K. (1977) Cell 12: 817-828.
- Pederson, D.S., Shupe, K., Bannon, G.A. and Gorovsky, M.A. (1986) Mol. Cell. Biol. 6: 3014-3017.
- Pedrali-Noy, G., Spadari, S., Miller-Faures, A., Miller, A.D.A., Kruppa, J. and Koch, G. (1980) Nucl. Acids. Res. 8: 377-387.
- Pelham, H.R.B. and Brown, D.D. (1980) Proc. Natl. Acad. Sci. U.S.A. 77: 4170-4174.

- Pelham, H.R.B. (1982) *Cell* 30: 517-528.
- Perry, M., Thomsen, G.M. and Roeder, R.G. (1985) *J. Mol. Biol.* 185: 479-499.
- Petryniak, B. and Lutter, L.C. (1987) *Cell* 48: 289-295.
- Pfeifer, K., Prezant, T. and Guarente, L. (1987) *Cell* 49: 19-27.
- Picard, D. and Schaffner, W. (1984) *Nature* 307: 80-82.
- Piette, J., Kryszke, M.-H. and Yaniv, M. (1985) *EMBO J.* 4: 2675-2685.
- Pinkert, C.A., Widera, G., Cowing, C., Heber-Katz, E., Palmiter, R.D., Flavell, R.A. and Brinster, R.L. (1985) *EMBO J.* 4: 2225-2230.
- Plon, S. and Wang, J. (1986) *Cell* 45: 575-580.
- Plumb, M., Stein, J. and Stein, G. (1983) *Nucl. Acids. Res.* 11: 7927-7945.
- Plumb, M., Marashi, F., Green, L., Zimmerman, A., Zimmerman, S., Stein, J. and Stein, G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 434-438.
- Powell, P.C., Payne, L.N., Frazier, J.A., Rennie, M. (1974) *Nature* 251: 79-80.
- Preston, C.M. and Tannahill, D. (1984) *Virology* 317: 439-444.
- Price, D.H. and Parker, C.S. (1984) *Cell* 38: 423-429.
- Prochownik, E.V. (1985) *Nature* 316: 845-848.
- Ptashne, M. (1986a) 'A Genetic Switch'; Cell and Blackwell Scientific Press, Cambridge and Palo Alto.
- Ptashne, M. (1986b) *Nature* 322: 697-701.
- Razin, S.V., Kekelidze, M.G., Lukanidin, E.M., Scherrer, K. and Georgiev, G.P. (1986) *Nucl. Acids Res.* 14: 8189-8207.
- Reeves, R. (1984) *Biochim. Biophys. Acta* 782: 343-393.
- Rein, A. and Rubin, H. (1968) *Exp. Cell. Res.* 49: 666-678.
- Remmers, E.F., Yang, J.-Q. and Marcu, K.B. (1986) *EMBO J.* 5: 899-904.



- Reynolds, G.A., Basu, S.K., Osborne, T.F., Chin, D.J., Gill, G.,  
Brown, M.S., Goldstein, J.L. and Luskey, K.L. (1984) *Cell* 38: 275-  
285.
- Reznikoff, W.S., Siegele, D.A., Cowing, D.W. and Gross, C.A. (1985) *Ann.*  
*Rev. Genet.* 19: 355-387.
- Rhodes, D. and Klug, A. (1986) *Cell* 46: 123-132.
- Robbins, E. and Borun, T.W. (1967) *Proc. Natl. Acad. Sci. U.S.A.* 57:  
409-416.
- Robins, A.J., Dalton, S., Sturm, R.S., Lesnikowski, L.S. and  
Wells, J.R.E.: Manuscript submitted for publication.
- Robinson, S.I., Nelkin, B. and Vogelstein, B. (1982) *Cell* 28: 99-106.
- Rogers, A.W. (1979) "Techniques of Autoradiography"; Elsevier Publishing  
Co., New York.
- Rogers, B.L. and Saunders, G.F. (1985) *BioEssays* 4: 62-65.
- Rogers, B.L., Sobnosky, M.G. and Saunders, G.F. (1986) *Nucl. Acids. Res.*  
14: 7647-7659.
- Rosen, C.A., Sodroski, J.G. and Haseltine, W.A. (1985) 41: 813-823.
- Rosenberg, U.B., Schröder, C., Preiss, A., Kienlin, A., Cote, S.,  
Riede, I. and Jackle, H. (1986) *Nature* 319: 336-339.
- Ross, J. and Kobs, G. (1986) *J. Mol. Biol.* 188: 579-593.
- Ruiz-Carrillo, A., Affolter, M. and Renaud, J. (1983) *J. Mol. Biol.* 170:  
843-859.
- Russell, D.W., Smith, M., Cox, D., Williamson, V.M. and Young, E.T.  
(1983) *Nature* 304: 652-654.
- Ryoji, M. and Worcel, A. (1984) *Cell* 37: 21-32.
- Ryoji, M. and Worcel, A. (1985) *Cell* 40: 923-932.
- Salditt-Georgieff, M. and Darnell, J.E. Jr. (1983) *Proc. Natl. Acad.*  
*Sci. U.S.A.* 80: 4694-4698.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci.*  
*U.S.A.* 74: 5463-5467.

- Sassone-Corsi, P., Wildeman, A. and Chambon, P. (1985) *Nature* 313: 458-463.
- Sassone-Corsi, P. and Borrelli, E. (1986) *Trends in Genet.* 2: 215-219.
- Sawadago, M. and Roeder, R.G. (1985) *Cell* 43: 165-175.
- Schaak, J., Sharp, S., Dingermann, T. and Söll, D. (1983) *J. Biol. Chem.* 258: 2447-2453.
- Schaffner, W., Kunz, G., Daetwyler, H., Telford, J., Smith H.O. and Birnstiel, M.L. (1978) *Cell* 14: 655-671.
- Schuaffner, W. (1985) In: 'The Role of Cis-and Trans- Acting Elements in Initiation. Gluzman, Y., Editor. Cold Spring Harbour Laboratory, New York.
- Schaufele, F., Gilmartin, G.M., Bannwarth, W. and Birnstiel, M.L. (1986) *Nature* 323: 777-781.
- Schlieff, R. and Lis, J. (1985) *J. Mol. Biol.* 95: 417-431.
- Schlieff, R. (1987) *Nature* 325: 14-15.
- Schneider, R., Gander, I., Müller, V., Mertz, R. and Winnacker, E.L. (1986) *Nucl. Acids Res.* 14: 1303-1317.
- Schuh, R., Aicher, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R. and Jäckle, H. (1987) *Cell* 47: 1025-1032.
- Seiler-Tuyns, A. and Paterson, B.M. (1986) *Nucl. Acids. Res.* 14: 8845-8862.
- Selleck, S.B. and Majors, J. (1987) *Nature* 325: 173-177.
- Sen, R. and Baltimore, D. (1986) *Cell* 46: 705-716.
- Shannon, M.F., Wigley, P.L. and Wells, J.R.E. (1985) *FEBS Letts.* 186: 180-186.
- Shannon, M.F. and Wells, J.R.E. Manuscript Submitted for Publication.
- Shaul, Y., Rutter, W.J. and Laub, O. (1985) *EMBO J.* 4: 427-430.
- Shepherd, J.C.W., McGinnis, W., Carrasco, A.E., DeRobertis, E.M. and Gehring, W.J. (1984) *Nature* 310: 70-71.

- Shinomiya, T., Scherer, G., Schmid, W., Zentgraf, H. and Schutz, G.  
(1984) Proc. Natl. Acad. Sci. U.S.A. 81: 1346-1350.
- Showman, R.M., Wells, D.E., Anstrom, J., Hurse, D.A. and Angerer, R.A.  
(1982) Proc. Natl. Acad. Sci. U.S.A. 79: 5943-5947.
- Shuey, D.J. and Parker, C.S. (1986) J. Biol. Chem. 261: 7934-7940.
- Sierra, F., Lichtler, A., Marashi, F., Rickles, R., Van Dyke, T.,  
Clark, S., Wells, J., Stein, G. and Stein, J. (1982) Proc. Natl.  
Acad. Sci. U.S.A. 79: 1795-1799.
- Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986) Nature 319:  
154-158.
- Sittman, D.B., Chiv, I.-M., Pan, C.-J., Cohn, R.H., Kedes, L.H. and  
Marzluff, W.F. (1981) Proc. Natl. Acad. Sci. U.S.A. 78: 4078-4082.
- Sittman, D.B., Graves, R.A. and Marzluff, W.F. (1983) Nucl. Acids. Res.  
11: 6679-6697.
- Sittman, D.B., Graves, R.A. and Marzluff, W.F. (1983) Proc. Natl. Acad.  
Sci. U.S.A. 80: 1849-1853.
- Sive, H.L., Heintz, N. and Roeder, R.G. (1984) Mol. Cell. Biol. 4: 2723-  
2734.
- Sive, H.L. and Roeder, R.G. (1986) Proc. Natl. Acad. Sci. U.S.A. 83:  
6382-6386.
- Sive, H.L., Heintz, N. and Roeder, R.G. (1986) Mol. Cell Biol. 6: 3329-  
3340.
- Slater, E.P., Searle, P.K., Chen, H.Y., Brinster, R.L. and  
Palmiter, R.D. (1984) Proc. Natl. Acad. Sci. U.S.A. 81: 7318-7322.
- Slater, E.P., Rabenau, O., Karin, M., Baxter, J.D. and Beato, M. (1985)  
Mol. Cell. Biol. 5: 2984-2992.
- Small, D., Nelkin, B. and Vogelstein, B. (1985) Nucl. Acids Res. 13:  
2413-2431.
- Smith, B.J., Walker, J.M. and Johns, E.W. (1980) FEBS Lett. 12: 42-44.
- Smith, D.R., Jackson, I.J. and Brown, D.D. (1984) Cell 37: 645-652.

- Smith, M., Leung, D.W., Gillam, S. and Astell, C.R. (1979) *Cell* 16: 753-761.
- Smith, M.M. and Murray, K. (1983) *J. Mol. Biol.* 169: 641-661.
- Southern, E.M. (1975) *J. Mol. Biol.* 98: 503-517.
- Southern, P.J. and Berg, P. (1982) *J. Mol. Appl. Genet.* 1: 327-341.
- Stefano, J.E. and Gralla, J.D. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 1069-1072.
- Stein, G.S., Stein, J.L. and Marzluff, W.F., eds. (1984) 'Histone Genes, Structure, Organization and Regulation', J. Wiley and Sons, New York.
- Stephenson, E.C., Erba, H.P. and Gall, J.G. (1981a) *Cell* 24: 639-647.
- Stephenson, E.C., Erba, H.P. and Gall, J.G. (1981b) *Nucl. Acids. Res.* 9: 2281-2295.
- Sternberg, P.W., Stern, M.J., Clark, I. and Herskowitz, I. (1987) *Cell* 48: 567-577.
- Stimac, E., Groppi, V.E. and Coffino, P. (1984) *Mol. Cell. Biol.* 4: 2082-2090.
- Stratling, W.H., Dolle, A. and Sippel, A.E. (1986) *Biochemistry* 25: 495-502.
- Strub, K., Galli, G., Busslinger, M. and Birnstiel, M.L. (1984) *EMBO J.* 3: 2801-2807.
- Strub, K. and Birnstiel, M.L. (1986) *EMBO J.* 5: 1675-1682.
- Struhl, K. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78: 4461-4465.
- Struhl, K. (1986) *Nature* 317: 822-824.
- Stuart, G.W., Searle, P.K., Chen, H.Y., Brinster, R.L. and Palmiter, R.D. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 7318-7322.
- Stunnenberg, H.G. and Birnstiel, M.L. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79: 6201-6204.
- Sturm, R.A. (1985) Ph.D. Thesis, University of Adelaide, Adelaide, South Australia.

- Sugarman, B.J., Dodgson, J.B. and Engel, J.D. (1983) *J. Biol. Chem.* 258: 9005-9016.
- Sures, I., Lowry, J. and Kedes, L.H. (1978) *Cell* 15: 1033-1044.
- Sures, I., Levy, S. and Kedes, L.H. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77: 1265-1269
- Sutcliffe, J.G. (1978) *Cold Spring Harbour Symp. Quant. Biol.* 43: 77-91.
- Swaneck, G.E. and Alvarez, J.M. (1985) *Biochem. Biophys. Res. Comm.* 128: 1381-1387.
- Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. and Chambon, P. (1986) *Nature* 319: 121-126.
- Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. and Hamuro, J. (1984) *Nature* 302: 305-309.
- Thoma, F., Koller, T. and Klug, A. (1979) *J. Cell. Biol.* 83: 403-427.
- Thoma, F. and Koller, T.L. (1981) *J. Mol. Biol.* 149: 709-733.
- Thomas, J.O. (1983) in "Eukaryotic Genes, their Structure, Activity and Regulation"; MacLean, N., Gregory, S.P. and Flavell, R.A. eds., Butterworths and Co., Cambridge U.K.: 9-30.
- Thomas, J.O. (1984) *J. Cell. Sci. Suppl* 1: 1-20.
- Thomas, J.O., Rees, C. and Pearson, E.C. (1985) *Eur. J. Biochem.* 147: 143-151.
- Tolunay, H.E., Yang, L., Anderson, W.F. and Safer, B. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81: 5916-5920.
- Topol, J., Ruden, D.M. and Parker, C.S. (1985) *Cell* 42: 527-537.
- Treisman, R. (1986) *Cell* 46: 567-574.
- Turner, P. and Woodland, H.R. (1983) *Nucl. Acids Res.* 11: 971-896.
- Turner, P.C., Aldridge, T.C., Woodland, H.R. and Old, R.W. (1983) *Nucl. Acids Res.* 11: 4093-4107.
- Twigg, A.J. and Sherratt, D. (1980) *Nature* 283: 216-218.
- Tyndall, C., LaMantia, G., Tacker, C.M., Favaloro, J. and Kamen, R. (1981) *Nucl. Acids Res.* 9: 6231-6250.

- Urban, M.K., Franklin, S.G. and Zweidler, A. (1979) *Biochemistry* 18: 3952-3960.
- Urban, M.K. and Zweidler, A. (1983) *Dev. Biol.* 85: 421-428.
- Valerio, D., Duyvesteyn, M.G.C., Dekker, B.M.M., Weeda, G., Berkevens, Th.M., Van der Voorn, L., van Ormondt, H. and van der Eb, A.J. (1985) *EMBO J.* 4: 437-443.
- van Eekelen, C.A.G. and van Venrooij, W.J. (1981) *J. Cell Biol.* 88: 554-563.
- van Wijnen, A.J., Stein, J.L. and Stein, G.S. (1987) *Nucl. Acids. Res.* 15: 1679-1698.
- Veldman, G.M., Lupton, S. and Kamen, R. (1985) *Mol. Cell. Biol.* 5: 649-658.
- Vincent, A., Colot, H.V. and Rosbash, M.J. (1985) *J. Mol. Biol.* 186: 149-166.
- Vogelstein, B., Pardoll, D.M. and Coffey, D.S. (1980) *Cell* 20: 79-85.
- Von Holt, C., Strickland, W.N., Brandt, W.F. and Strickland, M.L. (1979) *FEBS Lett.* 100: 201-218.
- Walker, M.D., Edlund, T., Boulet, A.M. and Rutter, W.J. (1983) *Nature* 306: 557-561.
- Wallis, J.W., Hereford, L. and Grunstein, M. (1980) *Cell* 22: 799-805.
- Wandelt, C. and Grummt, I. (1983) *Nucl. Acids. Res.* 11: 3795-3809.
- Wang, S.-W., Robins, A.J., D'Andrea, R. and Wells, J.R.E. (1985) *Nucl. Acids. Res.* 13: 1369-1387.
- Wang, X.-F. and Calame, K. (1986) *Cell* 47: 241-247.
- Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Teoule, R. and Chambon, P. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77: 7024-7028.
- Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) *Cell* 32: 503-514.
- Weber, F., deVilliers, J. and Schaffner, W. (1984) *Cell* 36: 983-992.

- Weiber, H., Konig, M. and Gruss, P. (1983) *Science* 219: 626-631.
- Weinberg, E.W., Hendricks, M., Hemmink, D., Kuwabara, L. and Farrelly, L.A. (1983) *Dev. Biol.* 98: 117-129.
- Weinberger, J., Baltimore, D., and Sharp, P.A. (1986) *Nature* 322: 846-848.
- Weintraub, H. (1985) *Cell* 42: 705-711.
- Weintraub, H. and Groudine, M. (1976) *Science* 193: 848-856.
- Weintraub, H., Larsen, A. and Groudine, M. (1981) *Cell* 24: 333-344.
- Weisbrod, S. (1982) *Nature* 297: 289-285.
- Wharton, R.P. and Ptashne, M. (1985) *Nature* 316: 601-605.
- Wharton, R.P. and Ptashne, M. (1986) *Trends in Biochem. Sci.* 11: 71-73.
- Wilhelm, M.L., Demouveau, M.F. and Wilhelm, F.X. (1986) *Curr. Genet.* 10: 459-462.
- Winter, E., Levy, D. and Gordon, J.S. (1985a) *J. Cell. Biol.* 101: 167-174.
- Winter, E., Palatnik, C.M., Williams, D.L., Coles, L.S., Wells, J.R.E. and Gordon, J.S. (1985b) *J. Cell. Biol.* 101: 175-181.
- Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82: 1585-1588.
- Woudt, L.P., Pastink, A., Kempers-Veenstra, A.E., Jansen, A.E.M., Hager, W.H. and Planta, R.J. (1983) *Nucl. Acids. Res.* 11: 5347-5360.
- Wright, S., Rosenthal, A., Flavell, R. and Grosveld, F. (1984) *Cell* 38: 265-273.
- Wu, C. (1984) *Nature* 309: 229-234.
- Wu, R.S. and Bonner, W.M. (1981) *Cell* 27: 321-330.
- Wu, R.S., Nishioka, D. and Bonner, W.M. (1982) *J. Cell. Biol.* 93: 426-431.
- Wu, R.S., Tsai, S. and Bonner, W.M. (1983) *Biochemistry* 22: 3868-3873.

- Wu, R.S., West, W.H.P., Bonner, W.M. (1984) in "Histone Genes, Structure, Organization and Regulation", Stein, G.S., Stein, J.L., and Marzluff, W.F. eds.; J. Wiley and Sons, New York: 457-483.
- Wynshaw-Boris, A., Lugo, T.G., Short, J.M., Fournier, R.E.K. and Hanson, R.W. (1984) *J. Biol. Chem.* 19: 12161-12169.
- Xu, M., Barnard, M.B., Rose, S.M., Cockerill, P.N., Huang, S-Y. and Garrard, W.T. (1986) *J. Biol. Chem.* 8: 3838-3845.
- Yaguchi, M., Roy, C., Dove, M. and Seligy, V. (1979) *Biochem. Biophys. Res. Comm.* 76: 100-160.
- Yang, L., Rowe, T.C., Nelson, E.M. and Liu, L.F. (1985) *Cell* 41: 127-132.
- Younghusband, H., Sturm, R.A. and Wells, J.R.E. (1986) *Nucl. Acids Res.* 14: 635-644.
- Zenke, M., Grundström, T., Matthes, H., Wintzerith, M., Wildeman, A. and Chambon, P. (1986) *EMBO J.* 5: 387-397.
- Zhong, R., Roeder, R.G. and Heintz, N. (1983) *Nucl. Acids. Res.* 11: 7409-7425.
- Zinn, K. and Maniatis, T. (1986) *Cell* 45: 611-618.
- Zoller, M.J. and Smith, M. (1983) *Methods Enzymol.* 100: 468-500.
- Zweidler, A. (1978) *Meth. Cell Biol.* 17: 223-233.
- Zweidler, A. (1980) *Dev. Biochem.* 15: 47-56.
- Zweidler, A. (1984) in "Histone Genes, Structure, Organization and Regulation". Stein, G.S., Stein, J.L. and Marzluff, W.F. eds., J. Wiley and Sons, New York: 339-375.