THE UNIVERSITY of EDINBURGH

Title	Nucleosome positioning on the chicken β-globin genes
Author	Edgar, Melanie
Qualification	PhD
Year	2000

Thesis scanned from best copy available: may contain faint or blurred text, and/or cropped or missing pages.

Digitisation Notes:

- Pagination error; page 167 is omitted

Scanned as part of the PhD Thesis Digitisation project
http://libraryblogs.is.ed.ac.uk/phddigitisation

Nucleosome positioning on the chicken β-globin genes

by

Melanie Edgar

Ph.D. Thesis

University of Edinburgh

1999

Declaration of Originality

I declare that, unless otherwise stated, this thesis represents my own work and was composed by me.

Melanie Edgar

To Francis 'Vid' Voisey

and my parents and grandmother
in recognition of their
encouragement and support

Acknowledgements

To my supervisor, Dr. James Allan, many thanks for his constructive advice and infectious enthusiasm about chromatin. May his eyes always gleam with the light of a true believer.

To Dr. Colin Davey, many thanks for his omniscience on all technical matters, his thoughtful advice, and for his discussion and criticism of this manuscript.

Thanks to Alastair, Chang-Hui, George, Irina and Leo for their help and advice.

Thanks to Dr. Richard Meehan and Dr. Sari Pennings for their helpful suggestions during my studies.

In no particular order, I would like to thank Leo, Colin, Chang-Hui, Linda, Venkat, Dan, Jan, Jim, Anne-Marie, Sari, Richard, Alastair, Marietta, Simon, John, Cheng-Fu (Max), Jasmine, Carmel, George, Irina, Wipa and Ting-Kuan, whose friendship kept me sane and allowed me to go on typing. I would like to thank in particular (a) those of the above who took me into their homes, often at great personal risk (if you did, you will know what I mean); (b) those who helped me indulge my love of good food and (c) those who made the chromatin group such an hilarious place to work.

Thanks to Erich, Jeremy and Matthew who are the best 'big brothers' I could imagine.

Thanks to Christina for being so good a friend for so many years!

Thanks to mum and dad, because home is the place where, when you go there, they have to let you in.

Thanks to friends and family for their love and support: Phyllis and James Pick, Glenys Kidd, Jean and Derek Russel, Bev, Steve, Sam, Lou and Kathy Tyler, Marg and Annette Voisey, Sylvia, Rob, Danny, and Wendy Edgar, Milmo, Liz, Debbie, Bryony and Jean Kilminster.

Abstract

Chromatin structure plays a decisive role in controlling gene expression (reviewed in Wolffe et al., 1994 and Grunstein, 1990). The globin genes are regulated in a tissue-specific manner and the individual genes are expressed only at specific times during development. This temporal and tissue-specific regulation involves changes in chromatin folding, which encompass the entire globin domain, and alterations in local, gene specific chromatin structures. These long and short-range chromatin structures are dependent on the positioning of individual nucleosomes.

This project identified the positions occupied by nucleosomes *in vitro* and *in vivo* throughout a region containing the reciprocally-expressed chicken globin genes, epsilon and beta–adult. The placement of individual nucleosomes on their promoters and shared enhancer influences local chromatin architecture, which may in turn contribute to the regulation of these genes. In addition, periodicities detected in the long-range placement of nucleosomes may participate in the folding of long stretches of the chromatin fibre into an inactive higher order structure. Furthermore, as the in vitro map shows the distribution of nucleosomes as determined by the underlying DNA sequence, comparison of the *in vitro* and *in vivo* data would reveal the extent to which the DNA sequence of the globin genes determines their chromatin structure in active and inactive tissues.

The monomer extension technique was used *in vitro* in both original (Yenidunya et al., 1994) and modified form (this study), to map DNA sequence-directed positioning throughout a 3.8 kb region encompassing the epsilon-globin gene. This map, juxtaposed with that for the beta-adult globin region (Davey et al., 1995) reveals the precise location and relative strength of nucleosome positioning sequences throughout 8 kb of continuous sequence.

Indirect end-labelling was employed to map nucleosome positions *in vivo*, in chicken brain and adult chicken erythrocytes; in brain, the whole globin domain is inactive, and packaged into a repressed higher order structure; in adult erythrocytes, where the entire globin domain has adopted a more open, accessible chromatin structure, the beta-adult gene is in an active state and epsilon is silenced.

Many of the strong *in vitro* nucleosome positioning sites, which exhibited a marked periodicity, were occupied *in vivo* in brain. This suggests that the DNA sequence of the globin genes creates a long-range chromatin structure which may be involved in folding the domain into the inactive, higher order chromatin fibre. Nucleosomes on the inactive epsilon promoter *in vivo* occupy positions consistent with strong sites identified *in vitro*, a feature shared with the adult beta promoter. These data suggest positional information within the DNA sequence influences local chromatin architecture, which has the potential to regulate access of protein factors and the stability of promoter-enhancer interactions, an important factor in the competition of epsilon and beta-adult for their shared enhancer.

Table of Contents

Title Page	I
Declaration of originality	II
Dedication	III
Acknowledgements	IV
Abstract	V

Chapter 1 Introduction

1.1	**Overview**	**1**
1.2	**Chromatin Structure**	**2**
1.2.1	Histone proteins	2
1.2.2.1	The core particle and the nucleosome	3
1.2.2.2	The chromatosome	10
1.2.3	The 30 nm fibre	11
1.2.3.1	Helical models: the solenoid model	12
1.2.3.2	Helical models: The twisted ribbon models	15
1.2.3.3	Helical models: Cross-linker helical models	16
1.2.3.4	The supranucleosomal particle models	16
1.2.3.5	The 'irregular' model	17
1.2.3.6	Weighing the evidence for and against the various models	17
1.2.4	Further levels of higher order folding	20
1.3	**Gene activation is accompanied by changes in chromatin structure**	**21**
1.3.1	Nuclease sensitivity	23
1.3.1.1	DNase I hypersensitivity	24
1.3.2	Nucleosome spacing	27
1.3.3	Histone acetylation and gene activity	28
1.3.4	Linker histones and gene activity	30
1.3.5	Further characteristics associated with gene activity	32
1.3.6	Nucleosome Positioning	34
1.3.7	Determinants of nucleosome positioning	38
1.3.7.1	Boundary-directed nucleosome positioning	39
1.3.7.2	Role of higher order structure in determining nucleosome positioning	40
1.3.7.3	Role of linker histone in determining nucleosome positioning	40
1.3.7.4	DNA sequence-directed nucleosome positioning	40

1.3.8	Prediction of sequence-directed nucleosome positioning	43
1.3.8.1	Statistically based algorithms	43
1.3.8.2	Structurally based algorithms	44
1.4	**Mechanisms for activating genes**	**45**
1.4.1	Reprogramming during replication	45
1.4.2	Interactions between transcription factors and nucleosomes	46
1.4.3	Chromatin remodelling factors	47
1.5	**Regulation of the chicken β-globin genes**	**47**
1.5.1	Introduction	47
1.5.2	Tissue-specific expression of the globin genes	49
1.5.3	Developmental regulation of the globin genes	50
1.6	**Thesis perspective**	**56**

Chapter 2 Materials and Methods 58

2.1	**Materials and stock solutions**	**58**
2.2	**Protocols**	**63**
2.2.1	Manipulation of nucleic acids: cloning	63
2.2.2	Determination of DNA concentration	63
2.2.2.1	By spectrophotometry	63
2.2.2.2	By comparison with a set of DNA standards	63
2.2.3	Digestion of DNA with restriction enzymes	64
2.2.4	Purification of DNA from agarose gels	64
2.2.5	End-filling DNA	64
2.2.6	Ligation of DNA fragments	64
2.2.7	Preparation of competent *E. coli* cells	64
2.2.8	Transformation of competent *E. coli* cells	65
2.2.9	Purification of DNA	66
2.2.10	Precipitation of DNA	66
2.2.11	Small-scale preparation of double-stranded plasmid DNA (miniprep)	66
2.2.11.1	Qiagen miniprep kit	67
2.2.11.2	Alkaline lysis method	67
2.2.12	Large-scale preparation of double-stranded DNA	67
2.2.13	Preparation of genomic DNA	67
2.2.14	Dephosphorylation of DNA	68
2.2.15	End-labelling of DNA with ^{32}P	68
2.2.16	Sephadex G-25/G-50 column purification	68
2.2.17	Amplification of DNA fragments	69
2.2.18	Sequencing double-stranded DNA using radioactively end-labelled primers	70
2.2.19	Preparation of long-range sequencing ladders	71

2.2.20	Preparation of ^{32}P end-labelled DNA size markers	71
2.2.21	Denaturing polyacrylamide DNA sequencing gel	72
2.2.22	Agarose gel electrophoresis of DNA	72
Chapter 3	**Nucleosome positioning on the chicken epsilon globin genes in vitro**	**74**
3.1	**Introduction**	**74**
3.1.1	The monomer DNA extension technique	75
3.1.1.1	The original monomer DNA extension technique (MDE)	76
3.1.1.2	A new monomer DNA extension technique: oligonucleotide-based monomer DNA extension	78
3.2	**Materials and Methods**	**81**
3.2.1	Materials	81
3.2.2	Plasmid constructs	82
3.2.3	Preparation of single-stranded DNA	84
3.2.3.1	Preparation of helper phage	84
3.2.3.2	Titre helper phage	86
3.2.3.3	Preparation of DH11S phagemid SOB/glycerol stocks	86
3.2.3.4	Growing culture for production of single-stranded DNA	86
3.2.3.5	Isolation of phage particles from the medium	87
3.2.3.6	Purification of phage DNA from particles	87
3.2.3.6.1	Original method	87
3.2.3.6.2	Tri-reagent method	88
3.2.3.6.3	Qiagen method	88
3.2.3.7	Checking concentration and quality of single-stranded DNA	89
3.2.4	In vitro nucleosome reconstitution with linear plasmid DNA	89
3.2.5	Preparation of mononucleosomal DNA	90
3.2.6	5′ end-labelling of monomer DNA	90
3.2.7	Alkaline denaturation of monomer DNA	90
3.2.8	Monomer Extension for in vitro nucleosome mapping	91
3.2.8.1	Original method	91
3.2.8.1.1	Annealing monomer DNA to ssDNA	91
3.2.8.1.2	Extension of annealed monomer DNA	91
3.2.8.1.3	Analysis of the monomer DNA extension data obtained from the polyacrylamide gel	92
3.2.8.2	New oligonucleotide-based method of monomer DNA extension	92
3.2.8.2.1	Oligonucleotide annealing and restriction enzyme cutting of the ssDNA template	93
3.2.8.2.2	Annealing monomer DNA to ssDNA	93
3.2.8.2.3	Extension of annealed monomer DNA	93
3.2.9	Screening enzymes for their ability to cut ssDNA in the	94

	oligonucleotide-based monomer DNA extension technique (OMDE)	
3.3	**Results**	**95**
3.3.1	Development and Optimisation of Monomer DNA extension	95
3.3.1.1	Development of new oligonucleotide-based monomer DNA extension technique (OMDE)	95
3.3.1.2	Developing a strategy for mapping the entire epsilon globin gene	99
3.3.1.3	Development of an assay to check whether restriction enzymes are capable of cutting single-stranded DNA	100
3.3.1.4	Optimisation of the monomer extension technique	102
3.3.1.4.1	Making ssDNA by alternative protocols	103
3.3.1.4.2	Including single-stranded DNA binding protein in the extension reaction	103
3.3.1.4.3	Optimising extension reactions for Klenow polymerase	105
	Using an alternative polymerase in the extension reaction	105
3.3.1.5	Summary of the development and optimisation of monomer extension	105
3.3.2	Mapping nucleosome positions on the chicken ε-globin gene	109
3.3.2.2	Preparation of core particle (monomer) DNA	109
3.3.2.3	Results of monomer DNA extension mapping	111
3.3.2.4	Nucleosome positioning map for the globin genes	125
3.3.2.5	Features of the nucleosome positioning map for the β^A and ε genes	128
3.3.2.5.1	Distribution of strong positioning sites within the epsilon gene	128
3.3.2.5.2	Long-range periodicities of nucleosome positioning sites	129
3.3.2.5.3	Short-range periodicities shown by multiple overlapping sites	132
3.3.2.5.4	Positioning over the epsilon globin promoter	133
3.3.2.5.5	Occurrence of certain DNA sequence patterns within strong nucleosome positioning sites	133
3.3.2.5.6	Positioning at intron-exon boundaries	137
3.4	**Discussion**	**143**
3.4.1	Production of the nucleosome positioning map for the chicken ε-globin gene by monomer extension	143
3.4.2	Features of the ε-globin gene nucleosome positioning map: Comparison with the β^A map	143
3.4.3	Distribution of strong nucleosome positioning sites between introns and exons	144
3.4.4	Long-range periodicity of strong nucleosome positioning sites	146
3.4.5	Short-range periodicities relating multiple overlapping positioning sites	150
3.4.6	Nucleosome positioning over the epsilon globin promoter	152

3.4.7	DNA sequence patterns within strong nucleosome positioning sites	157
3.4.8	Positioning of nucleosomes at splice sites	160
Chapter 4	**Mapping nucleosome positions on the epsilon and beta-adult globin genes in vivo**	**161**
4.1	**Introduction**	**161**
4.2	**Materials and Methods**	**163**
4.2.1	Materials	163
4.2.2	Probes used for indirect end-labelling	166
4.2.3	Preparation of nuclei from chicken erythrocytes	166
4.2.4	Preparation of nuclei from brain tissue	168
4.2.5	Determining concentration of nuclei	169
4.2.6	Cleavage of chromatin in nuclei with DNA cleavage reagents	169
4.2.6.1	Cleavage of chromatin in nuclei by cuprous phenanthroline	169
4.2.6.2	Cleavage of chromatin in nuclei by methidium propyl EDTA (MPE)	170
4.2.6.3	Cleavage of chromatin in nuclei with micrococcal nuclease	171
4.2.7	Southern blotting	171
4.2.8	Filter hybridisation	172
4.2.8.1	Cross-linking the DNA to the filter	172
4.2.8.2	Pre-hybridisation of the filter	172
4.2.8.3	Preparation of ^{32}P-labelled probes by the OLB method	173
4.2.8.4	Hybridisation of the filter	173
4.2.9	Preparation of probes for Southern blotting	174
4.2.9.1	Making probes by PCR	174
4.2.9.2	Ligation of probe	174
4.3	**Results**	**176**
4.3.2	Strategy for mapping nucleosome placement on the ε and β^A genes	177
4.3.3.	Generation of a nucleosome placement map for the ε and β^A genes in brain and red blood cells	181
4.3.4	General features of nucleosome placement in brain and red blood cells on the ε and β gene region, and a comparison with the features of the in vitro nucleosome positioning map	202
4.3.5	Specific features of nucleosome placement in brain and red blood cells	204
4.3.5.1	Long-range patterns in the placement of nucleosomes in brain and red blood cells	204
4.3.5.2	Nucleosome positioning over the epsilon globin promoter	206

4.3.5.3	Distribution of nucleosomes between introns, exons and flanking regions	206
4.3.5.4	Nucleosome positioning at splice sites	206
4.4	**Discussion**	**210**
4.4.1	General features of nucleosome placement in brain and red blood cells	210
4.4.2	Specific features of nucleosome placement in brain and red blood cells	213
4.4.2.1	Long-range patterns in the placement of positioned nucleosomes	213
4.4.2.2	Nucleosome positioning over the epsilon promoter	216
4.4.3	Distribution of nucleosomes between introns, exons and flanking regions	219
4.4.4	Nucleosome placement over splice sites	220
	Thesis: Final Perspective	**222**
	References	**228**

Table of Figures

Figure 1.1a	The chromatosome	4
Figure 1.1b	Formation of heterodimers within the nucleosome is mediated by the histone fold motif	4
Figure 1.2	The nucleosome core particle structure at high resolution	6
Figure 1.3a	The nucleosome core particle structure: 73 bp half	8
Figure 1.3b	The H3-H4 histone-fold pair, showing the DNA binding structures	8
Figure 1.4	Various models for the 30 nm chromatin fibre	13
Figure 1.5	Schematic showing the structure of proposed chromatin loops when they are in active and inactive conformations	26
Figure 1.6a	Preferential placement of A/T and G/C nucleotides in nucleosomal DNA	35
Figure 1.7	Organisation of the chicken and human β-globin gene clusters	48
Figure 3.1	The original monomer DNA extension technique	77
Figure 3.2	Schematic diagram of the oligonucleotide-based monomer DNA extension technique (OMDE)	80
Figure 3.3	Clones used for monomer DNA extension mapping	85
Figure 3.4a	Schematic showing the mapping strategy for the monomer extension gel analysis shown in Figure 3.4b	96
Figure 3.4b	Demonstrating the reliability of the OMDE technique	98
Figure 3.5	Assay to check whether restriction enzymes are capable of cutting single-stranded DNA	101
Figure 3.6	Comparison of monomer DNA extension data from reactions using ssDNA prepared by different protocols	104
Figure 3.7a	The location of repeat elements within the chicken β-globin gene cluster	108
Figure 3.7b	Positions of CR1 repeat elements over the beta-adult and epsilon globin gene region	108
Figure 3.8a	Monomer DNA preparation and characterisation	110
Figure 3.8b	Integrity of the monomer DNA preparation	110
Figure 3.9a	Schematic representation of the monomer DNA extension analyses of clones PCBE1.9HE and PCBE1.9EH that are presented in Figure 3.9b	112
Figure 3.9b	Polyacrylamide gel analysis of monomer DNA extensions on PCBE1.9EH and PCBE1.9HE	113
Figure 3.9c	Size calibration of sequencing gels	114
Figure 3.9d	Nucleosome positioning map produced from the monomer DNA extension analysis on PCBE1.9HE with Xho I (Figure 3.9b)	115

Figure 3.10a	Schematic representation of the monomer DNA extension analyses on clones PCBEeco and PCBEsac presented in Figure 3.10b	117
Figure 3.10b	Polyacrylamide gel analysis of monomer extension on PCBEeco and PCBEsac clones	118
Figure 3.11a	Schematic representation of the oligonucleotide-based monomer DNA extension analyses on clones PCBE2.4Bam and PCBE2.4Bcl presented in Figure 3.11b	119
Figure 3.11b	6% polyacrylamide gel analysis of monomer extension reactions carried out by the new technique	120
Figure 3.12a	Schematic representation of the monomer DNA extension analyses carried out on clones PCBEacc and PCBEeag presented in Figure 3.12b	121
Figure 3.12b	6% polyacrylamide gel analysis of monomer DNA extension on PCBEeag and PCBEacc	122
Figure 3.13	Monomer extension provides quantitative data on the relative strength of nucleosome positioning sites	124
Figure 3.14a	Map of the histone octamer (nucleosome) positions on the epsilon globin gene region	126
Figure 3.14b	Combined map of the histone octamer (nucleosome) positions on the region containing the beta-adult and epsilon globin genes	129
Figure 3.15	Autocorrelation and Fourier analyses of monomer DNA extension data	130
Figure 3.16	Fourier analyses of monomer DNA extension data	131
Figure 3.17	Nucleosome positioning over the epsilon globin promoter in vitro	134
Figure 3.18a	Occurrence of the sequence motifs RRRNNRRR and MMMNNMMM within nucleosome positioning sites on the beta-adult and epsilon globin genes	136
Figure 3.18b	Table showing nucleosome positioning at putative DNA bend sites	138
Figure 3.18c	Table showing nucleosome positioning at CACACA motifs	139
Figure 3.19	Table showing nucleosome positioning at splice sites	141
Figure 4.1	Strategy for mapping nucleosome positions in vivo by indirect end-labelling	167
Figure 4.2	Time course of digestion of red blood cell nuclei by DNA cleavage reagents and micrococcal nuclease	178
Figure 4.3	A typical agarose gel prepared for mapping nucleosome placement by indirect end-labelling	182
Figure 4.4	Nucleosome mapping on brain nuclei using the Hind probe	183
Figure 4.5	Nucleosome mapping on brain nuclei using the Hind probe	185

Figure 4.6	Nucleosome mapping on red blood cell nuclei using the Hind probe	188
Figure 4.7	Nucleosome mapping on brain nuclei using the Eco probe	189
Figure 4.8	Nucleosome mapping on red blood cell nuclei using the probe	191
Figure 4.9	Nucleosome mapping on brain nuclei using the Dra probe	192
Figure 4.10	Nucleosome mapping on red blood cell nuclei using the Dra probe	194
Figure 4.11	Nucleosome mapping on brain nuclei using the Bam probe	195
Figure 4.12	Nucleosome mapping on red blood cell nuclei using the Bam probe	196
Figure 4.13	Nucleosome mapping on red blood cell nuclei using the Eag probe	198
Figure 4.14	Nucleosome mapping on brain nuclei using the Rsa probe	199
Figure 4.15	Nucleosome placement on the β^A and ε globin genes in brain and red blood cells	200
Figure 4.16	Map showing which in vitro core histone octamer positioning sites may be occupied in vivo on the beta-adult and epsilon globin genes	201
Figure 4.17	Nucleosome positioning over the ε globin promoter in brain and red blood cells	205
Figure 4.18	Nucleosome positioning over the exons of the beta-adult and epsilon globin genes	207
Figure 4.19	Nucleosome positioning at splice sites in brain and red blood cells	209

List of Abbreviations

Å	Angstrom (1Å=10^{-10} m)
ATP	adenosine triphosphate
β^A-globin/β^A globin	Chicken beta-adult globin gene of β-globin cluster
bp	base pair
BSA	bovine serum albumin
dATP	2-deoxyadenosine 5′-triphosphate
dCTP	2-deoxyctidine 5′-triphosphate
dGTP	2-deoxyguanosine 5′-triphosphate
dTTP	2-deoxythymidine 5′-triphosphate
dNTP	deoxynucleotide tri-phosphate
ddNTP	di-deoxynucleotide tri-phosphate
DNase I	deoxyribonuclease I
DTT	dithiothreitol
ε-globin/ε globin	chicken epsilon globin gene of β-globin cluster
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
g	gravity acceleration
Kbp	kilo base
MDE	Monomer DNA extension
OMDE	Oligonucleotide-based monomer

	DNA extension
MW	molecular weight
nt	nucleotide
OD	optical density
ssDNA	Single-stranded DNA

INTRODUCTION

1.1 Overview

In 1944, it was discovered that DNA was the store of information within the cell, the 'blueprint' encoding the characteristics of an organism, which could be passed from one generation to the next (Avery et al., 1944, reviewed in Lewin, 1994). With the elucidation of the structure of DNA (Watson and Crick, 1953), the challenge has been to understand how cells regulate the use of this information, so that each cell expresses the correct genes at the correct time.

Early studies implicated specific proteins (transcription factors) and the DNA elements they bound (promoters and enhancers) in the regulation of gene expression. With the discovery in the 1970's that DNA was packaged into a nucleoprotein complex, called chromatin, researchers realised that this could have a role in controlling gene expression. Further work showed that gene activation was accompanied by profound changes in chromatin structure which increased the accessibility of the DNA to the proteins involved in transcription (reviewed in Van Holde, 1989; Grunstein, 1990; Wolffe, 1994).

The β-globin gene cluster in human and chicken has been considered as a paradigm for the regulation of eukaryotic gene transcription. Studies have revealed much about the control of globin gene expression and the role played in this by chromatin structure. Despite this, the importance of chromatin structure and specific aspects of it such as nucleosome positioning in globin gene expression have not been fully evaluated. Further investigation of globin gene expression may well advance the understanding of eukaryotic gene expression in general.

1.2 Chromatin structure

The nuclei of eukaryotic cells contain ~2 metres of DNA. Therefore, the DNA must be folded up and tightly 'packed' for it to be accommodated in a nucleus only a few microns across (reviewed in van Holde, 1989). The DNA is folded into a higher order structure by association with histone proteins to form a nucleoprotein complex known as chromatin. The first level of packaging involves the wrapping of 146 bp of DNA (in 1.75 turns) around an octamer of histone proteins to form a core particle. This may associate with linker histone to form a chromatosome. The repeat unit of chromatin is the core particle or chromatosome and its linker DNA, which is referred to as a nucleosome. In the next level of packing, this 10 nm chromatin fibre is coiled into a 30 nm fibre, which in turn can be further coiled for an additional increase in the packing ratio.

1.2.1 Histone Proteins

The histones are fundamental in folding DNA into chromatin. The core histones, H2A, H2B, H3 and H4 form the core particle, and linker histones, such as H1 and H5, bind at the point where the DNA enters and exits the histone octamer, to form a chromatosome.

The histones are highly conserved; H4, which shows the slowest rate of change, differs by only two amino acid residues between pea and cow (Delange et al., 1969). The histones are basic proteins, rich in lysine and arginine (Van Holde, 1989). This contributes to the interactions between the histones and the DNA, which are largely electrostatic in nature.

Most histones have three structural domains: a central structured globular domain, a carboxy terminal flexible 'arm' and an amino terminal flexible 'tail'. The carboxy terminal end of the protein contains a 'histone-fold' domain; this is a long central helix,

bordered by a loop and a shorter helix on each side (Pruss et al., 1995). The histone fold motif is shared by many other proteins, including a number of transcription factors (Xie et al., 1996). In nucleosomes, it is involved in histone:histone and DNA:histone interactions. The amino terminal tails of the core histones are thought to be involved in modulating chromatin structure (Hayes and Wolffe, 1992a). The tails are critical to the formation of higher order structures and certain modifications, such as acetylation, may alter histone tail interactions, thereby influencing the structure of the chromatin fibre (reviewed in Wolffe and Hayes, 1999)

1.2.2.1 The core particle and the nucleosome

The existence of a repeated unit in chromatin was demonstrated by nuclease digestion (Williamson, 1970). Studies demonstrated that histones were a major component of chromatin, and that they could associate in specific combinations in solution (Kornberg and Thomas, 1974). This led to the proposal that the fundamental repeating unit in chromatin was an octamer of histone proteins, containing an H3/H4 tetramer and two H2A/H2B dimers (making the composition $(H3)_2(H4)_2(H2A)_2(H2B)_2$ around which ~ 200 bp DNA was wound (Kornberg, 1974).

It was further proposed that the chromatin fibre consisted of a continuous string of these particles. DNA (146 bp) wrapped around one core histone octamer to form a 'core particle', and then extended as a stretch of naked linker to the next. This repeating unit of chromatin - the core particle with its linker DNA - is referred to as a nucleosome (van Holde, 1989). The DNA would appear as a 'string', between 'beads'. This 'beads-on-a-string' structure can be visualised by electron microscopy when chromatin is unfolded in low salt conditions (Oudet et al., 1975). The 'beads-on-a-string' filament is often described as the 10 nm filament.

Extensive nuclease digestion of nucleosomes, which digests away the linker DNA, gives rise to a structure containing the octamer of histones and 146 bp of DNA; as

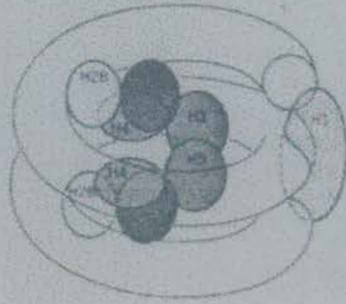


Figure 1.1a The chromatosome

The two turns of the DNA (168 bp) wraps around the octamer of histone proteins. Linker histone (in this case H1) binds at the site indicated, where the DNA enters and exits the core particle, to form the chromatosome.

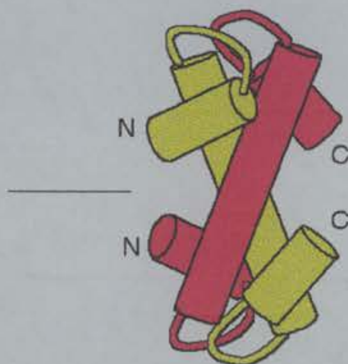


Figure 1.1b Formation of heterodimers within the nucleosome is mediated by the histone-fold motif

Each histone has a three helix, histone-fold motif. These domains pack through an interlocking 'handshake-type' interaction to form heterodimers.

mentioned above, this is termed a 'core particle' (Lutter, 1979). Figure 1.1a is a simple representation of the core particle, and it shows the position at which H1 binds to form the chromatosome. X-ray diffraction and electron microscopy studies of gradually improving resolution have revealed the structure of the core particle.

One early model, which became known as the MRC model, was based on electron microscopy data (Klug et al. 1980; Arents et al., 1991). It was proposed that the core histone octamer had a roughly cylindrical face, 70 Å in diameter and 55 Å in length, and that it had a two-fold axis of symmetry. The shape of the molecule suggested that it would accommodate two turns of a left-handed superhelix of DNA. The H3/H4 tetramer was thought to form the centre of this structure, associating with the central turn of DNA, whilst the two H2A/H2B dimers were located one on each side of the tetramer, each interacting with half a turn of DNA.

More information on core particle structure was provided by X-ray diffraction data at 7 Å resolution (Richmond et al., 1984). This confirmed that the shape of the octamer did provide a helical 'ramp' around which was coiled 1.7 turns of a left-handed DNA superhelix. It also revealed other structural features of the core particle: the path of the DNA on the octamer is not uniform as kinking and bending occurs 15 and 40 bp either side of the dyad; the H3/H4 tetramer and the central turn of DNA show symmetry about the dyad, but H2A/H2B dimers show some asymmetry.

The histone octamer was crystallised in the absence of DNA, and its structure determined to a resolution of 3.1 Å (Arents et al., 1991). The data provided by this study, considered in conjunction with previous studies where DNA was included in the structure, further refined the nucleosome core particle structure. It was revealed that the globular domains of the core histones were extended rather than compacted, leading to the proposal that each histone can make non-contiguous contacts with the DNA. Therefore, beginning where the DNA exits the core particle and working inwards to the dyad, the histones contact the surface of the histone core in the following order (where the superscript number indicates which of the two possible molecules is meant): $H2A^1$,

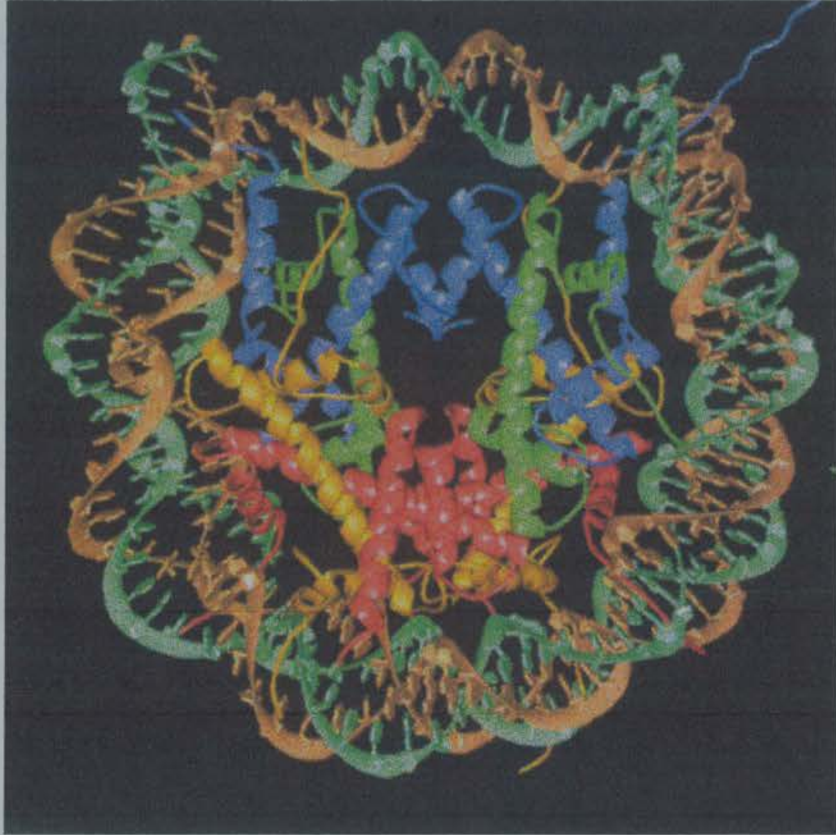


Figure 1.2 The nucleosome core particle structure at high resolution

The view is down the DNA superhelix axis. Eight histone protein main chains (H3:blue, H4:green, H2A:yellow, H2B:red) and 146 bp of DNA (brown and turquoise) are shown (taken from Luger et al., 1997).

$H3^2$, $H2A^1$, $H2B^1$, $H2A^1$, $H2B^1$, $H4^1$, $H3^1$, $H4^1$, $H3^1$ and $H3^2$. The octamer was thought to contact the DNA at 14 places on its surface (Arents et al., 1993).

The high resolution (2.8 Å) structure of the core particle (Luger et al., 1997) (Figure 1.2) has confirmed many of the structural features suggested by other studies and has also revealed other unexpected characteristics. The study was carried out on a core histone octamer formed on a 146 bp palindromic DNA of defined sequence. Despite the palindromic sequence the axis of dyad symmetry passes through a single base pair so that the octamer has 72 bp on one side of its dyad and 73 bp on the other. In the following description of the core particle, the terms SH0, SH1.5 (or 0 , 1.5) and so on are used. By convention, the dyad, where the major groove faces inward, is numbered as SH0; when the twist of the DNA helix next brings the major groove to face the octamer this is numbered as SH1, and so on to the final SH7. On the other side of the dyad, the numbering is SH-1 and so on (For example, see Figure 1.3a).

This crystal structure revealed the following features. The 146 bp of DNA are wrapped around the helix in 1.65 turns. The average twist of the DNA is 10.2 bp per turn, as compared with 10.6 bp per turn in solution. The twist does vary, from 9.4 to 10.9 bp per turn, over the length of the DNA superhelix. This variability could explain why longer DNA fragments can show histone contacts over 160 bps (Hayes et al., 1990). The average 10.2 bp pitch aligns the minor grooves to form a channel for histone N terminal tails, as will be discussed later. The path of the DNA is not uniform; as suggested by previous data, bending or kinking occurs at +/- SH1.5 and +/-SH 4.5. This is created by the way the octamer contacts the DNA.

This high resolution structure has revealed how the octamer is assembled. Each histone consists of a structured three helix domain called the histone fold (Figure 1.1b). This consists of an α-helix (α1) joined by a loop segment (L1) to a second helix (α2) which is joined by a loop segment (L2) to a third helix (α3). Outside the histone fold, there may be other motifs, like the α-helix at the H3 N-terminus (H3αN). The histone fold is crucial to the formation of the octamer. The histone fold domains, which are

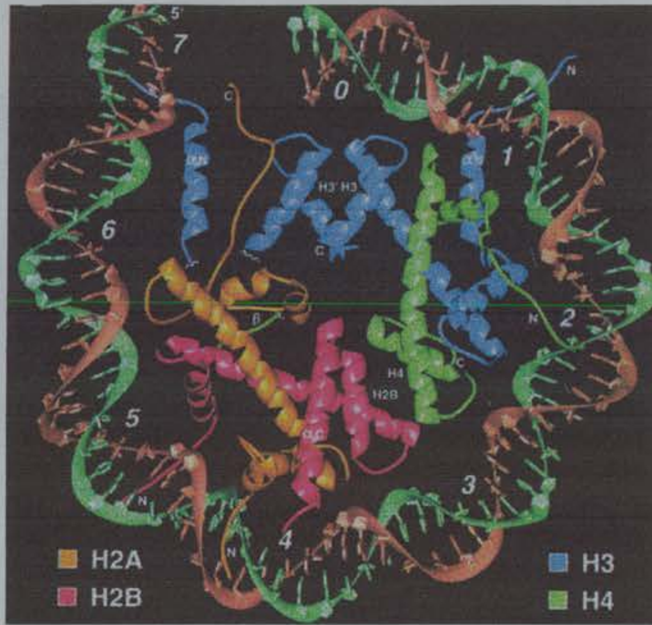


Figure 1.3a The nucleosome core particle structure: 73 bp half
The core particle binds the DNA with a single base pair at the dyad, dividing the DNA into 72 and 73 bp halves. The 73 bp half of the DNA is shown here. The view is down the DNA superhelix axis. The core particle dyad is indicated by '0', and other numbers indicate position as the number of DNA double helix turns from the dyad (see text). The 4-helix bundles mediating both the H4:H2B and H3:H3' interaction can be clearly seen (The copies of each histone pair are distinguished by being primed or unprimed). Histone fold extensions of H3 and H3' are labelled as αN and αN' respectively, and the H2B histone fold extension as αC.N and C-terminal tails are labelled as N and C (taken from Luger et al., 1997).

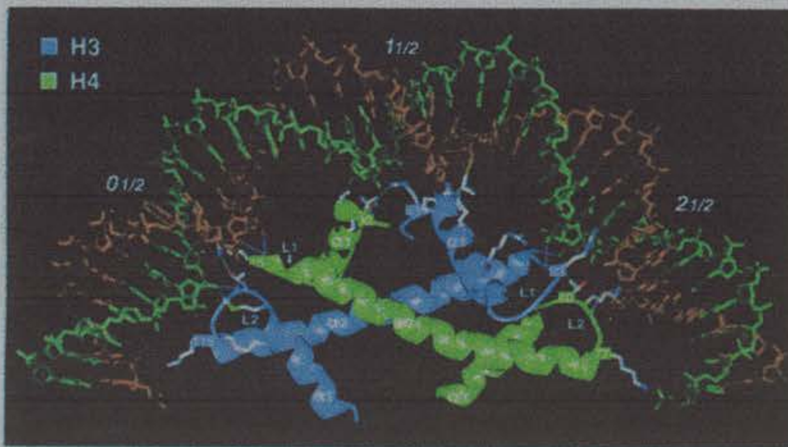


Figure 1.3b The H3-H4 histone fold pair, showing the DNA binding structures
The view is down the DNA superhelix axis. Examples of the two types of DNA binding site (L1L2 and $\alpha1\alpha1$) are indicated.

roughly U-shaped interlock in a handshake-type interaction to form the H2A/H2B and H3/H4 heterodimers (Figure 1.1b). The H3/H4 dimers interact with each other via a bundle of 4 helices; two of these belong to the one H3 molecule (H3), and the other two to the other H3 molecule (H3′). The H2A/H2B dimers interact with the tetramer through a bundle of four helices belonging to H4 and H2B; two of the four belong to H4, and two to H2B (Figure 1.3a).

The histone fold is also crucial to the interaction of the octamer with the DNA (Figure 1.3b). Out of the total 146 bp, 121 bp interacts with the histone-fold domains, with each of the heterodimers binding about 30 bp. The DNA is contacted at about 10 bp intervals as the minor groove faces the octamer. The H3/H4 tetramer binds SH-3 to SH+3 and the two H2A/H2B dimers bind SH-3 to SH-6 and SH3 to SH6. Each dimer creates two types of binding sites: the juxtaposition of the two α1 helices in the dimer creates an α1α1 binding site, and juxtaposition of loop segments from each of the histones in the dimer creates an L1L2 binding site (Figure 1.3b). For regions outside the histone fold, the N-terminal α-helix of H3 binds the terminal 13 bp each side of the dyad. The N-terminal tails of H2B and H3, which are ordered, but not structured, pass through channels created by adjacent minor grooves, so that there is a protruding histone tail every 20 bp. One H4 tail makes an internucleosome contact with the face of an H2A/H2B dimer. This raises the possibility that modifications like acetylation may alter interactions between core particles as opposed to, or in addition to altering the structure of an individual core particle.

The majority of the contacts between the histones and the DNA are non-specific: for example, main chain nitrogens of the histones make polar contacts with the phosphates of the DNA backbone; there are extensive non-polar contacts with the deoxyribose of the DNA backbone; the N-termini of α-helices make dipole interactions with the DNA backbone; there are extensive salt-links and H-bonds between basic and hydroxyl groups of the amino acid residues and the phosphate and oxygen groups of DNA, and arginine side-groups protrude into the 14 minor grooves. There are a very few specific

interactions: an H4 arginine (H4-R45) binds the oxygen of a thymidine group.

Therefore, it is not the DNA sequence *per se* which positions nucleosomes; instead aspects of DNA secondary structure, created by the sequence, make it energetically more or less favourable to bind the DNA around the nucleosome and to accommodate the kinks which occur at +/- SH1.5 and +/- SH4.5 thereby influencing nucleosome positioning.

1.2.2.2 THE CHROMATOSOME

The chromatosome is the next level of packaging of the DNA after the core particle (Figure 1.1a). A molecule of linker histone associates with the core particle, so that 168 bp of DNA is protected from nuclease digestion (Simpson, 1978). The incorporation of linker histone into chromatin stabilises nucleosomes, restricts nucleosome mobility and facilitates the packing of the DNA into a higher order structure, the 30 nm fibre (reviewed in Wolffe et al., 1997; Wolffe and Hayes, 1999).

Linker histones are a family of structurally related proteins including H1, H5 and $H1^0$. Typically, they have 3 structural domains: a central globular domain, a short basic N-terminal 'tail' and a long C-terminal 'tail'. The central globular domain is the most highly conserved (Hartman et al., 1977; Aviles et al., 1978). The NMR and crystal structures reveal that the globular domain of H5 has a structural core of 3 helices (Clore et al., 1987; Ramakrishnan et al., 1993).

The location of the linker histone globular domain within the chromatosome is uncertain. It was initially proposed to be located at the point where the DNA enters and exits the nucleosome, with symmetrical contacts protecting 10 bp of DNA on either side of the 146 bp nucleosome core (Allan et al., 1980). It has also been shown that the linker histone can be placed asymmetrically, protecting 15 and 5 bp of DNA at either side of the nucleosome core (Hayes et al., 1994). Recent work suggests both placements are possible for linker histone, but a survey of positioned chromatosomes showed that in

most of these the linker histone was symmetrically placed (Shen, 1997).

The globular domains are sufficient and necessary for the protection of 168 bp of DNA (Allan et al., 1980). They are also thought to be responsible for the correct positioning of the basic N and C-terminal tails so that they can facilitate the formation of a higher order structure (Allan et al., 1986).

The highly charged linker histones are important for chromatin compaction, because they can bind linker DNA and stabilise chromatin structure. The linker histone tails are sites of post-transcriptional modifications which may regulate the conformation of the linker histone tail, thereby altering its contribution to the structure of the chromatin fibre. These modifications include acetylation, phosphorylation and poly-ADP ribosylation. Phosphorylation during the cell cycle is thought to neutralise the charged tails somewhat and reduce their affinity for the DNA (Hall and Cole, 1986). Poly-ADP ribosylation of H1 is thought to decrease the stability of higher order structure (Aubin et al., 1983).

Linker histone may have a role as a repressor of transcription, because of its ability to stabilise higher order chromatin structure (van Holde, 1989; Wolffe et al., 1997). It has been proposed that the variety of linker histones occurring within a cell makes a contribution to the pattern of its gene expression. For example, levels of histone H5, which increase during the terminal differentiation in nucleated erythrocytes, may be involved in chromatin compaction and transcriptional repression. Furthermore, Histone $H1^0$ has been shown to be involved in the maintenance of the terminally differentiated state in some mammalian cells (reviewed in Wolffe, 1997).

1.2.3 THE 30 NM FIBRE

An unfolded chromatosome (or nucleosomal) fibre, with its 'beads-on-a-string' appearance, can be refolded, by altering the salt conditions, into a condensed higher order structure. Electron microscopy shows that compaction is a gradual process, with

the nucleosomes adopting a zigzag arrangement in the intermediate stages before the final highly condensed 30 nm (100 Å) fibre is achieved (Thoma and Koller, 1977; Thoma et al., 1979). The biophysical and biochemical properties of the 30 nm fibre have been extensively investigated: the mass per unit length of the fully compacted fibre, the dependence of the fibre diameter on linker length, the position of H1 in the higher order structure, the orientation of the nucleosomes with respect to the axis of the 30 nm fibre and whether the linker DNA is bent or straight. On the basis of these data, three classes of helical models have been proposed for the 30 nm fibre: the solenoid, the twisted ribbon, and the crossed-linker models. One non-helical model, the supranucleosomal particle model, and one irregular model have also been proposed. No single model fits all the data gathered, but the weight of evidence appears to favour the solenoid model.

1.2.3.1 Helical models: the solenoid model

It was proposed that the nucleosome fibre could be coiled into a helix with a pitch (rise per turn) of 11 nm and a diameter of ~30 nm, containing 4 to 10 nucleosomes per turn, with 6 or 7 nucleosomes per turn creating a particularly stable configuration. This model is sequential; that is, adjacent chromatosomes adopt adjacent positions in the 30 nm helix (Finch and Klug, 1976). A variation on this model was proposed to account for the zigzags observed in the fibre at low ionic strength. In the 10 nm ('beads-on-a-string') fibre, adjacent nucleosomes are proposed to adopt positions on alternate sides of the DNA strand. This is described as a contact helix, with two nucleosomes per turn, stabilised by H1-H1 interactions in the centre. This contact helix is coiled into a further helix with 6 to 8 nucleosomes per turn, a pitch of 11 nm and the nucleosomes in a radial orientation (Figure 1.4A). In this model, and the original model, the H1 binding sites are close together, so that H1 can polymerise in the centre of the superhelix and enhance solenoid stability. Therefore this model accommodates the proposed role of linker histone in catalysing or stabilising higher order structure (Maman et al., 1994).

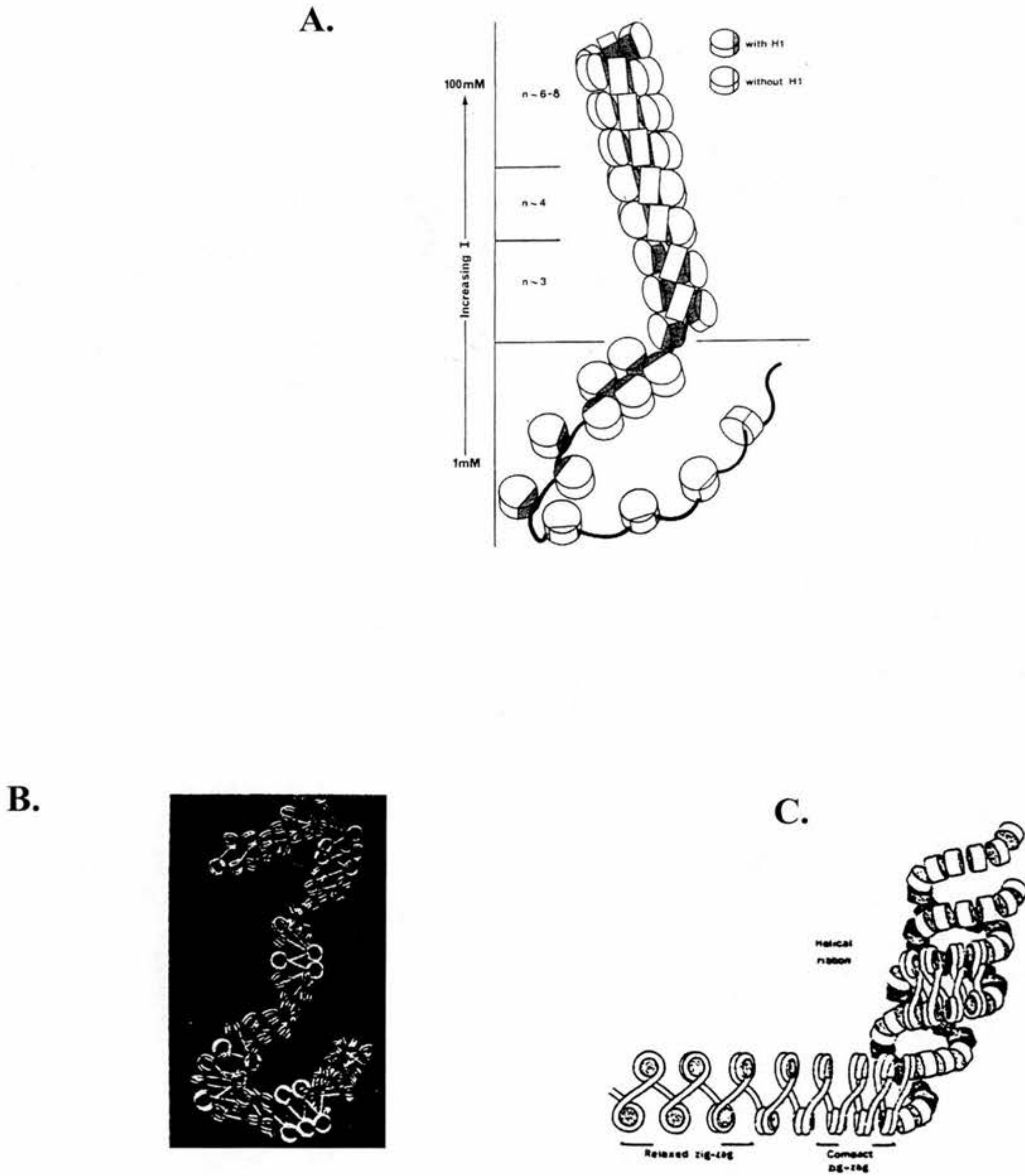


Figure 1.4 Various models for the 30 nm chromatin fibre.

A. The solenoid model (Finch and Klug, 1976). Increasing salt leads to increased compaction, with more nucleosomes per turn of the DNA Helix. B. The irregular model (Woodcock, 1994). C. The helical ribbon model (Woodcock et al., 1984).

The path which linker DNA adopts in solenoid models is uncertain. Butler (1984), proposed that the linker is internally located and can form reversed loops, kinks and sharp bends. In this structure, the dimensions of the solenoid would be largely independent of linker length. Certain linker lengths (163 to 176 bp) would preclude formation of a solenoid with this structure.

Electric dichroism data have been used to model the path taken by linker DNA in the solenoid (McGhee et al., 1983). A linker perpendicular to the solenoid axis did not fit their electric dichroism data and a linker parallel to the solenoid axis was considered unlikely as in this structure a greater rise in the pitch of the helix would occur with longer linkers. The following supercoiled linker model was proposed to fit the data. They suggested that the linker DNA formed a helix with 80 bp per turn and a diameter of 11 nm. They also proposed that the angle between the nucleosome face and the solenoid long axis (the tilt angle) had a maximum of 44° and that the tilt angle of the linker DNA to the solenoid axis was 20 to 25 degrees. This model has some major implications for the structure of the solenoid. The helical path of the linker DNA means that different linker lengths will complete a different fraction of a turn between one nucleosome and the next. Therefore the DNA entry and exit sites on the nucleosome could adopt different positions within the DNA solenoid. In sea urchin chromatin, for example, the 80 bp linker makes a full turn between nucleosomes, so that all chromatosomes adopt the same orientation within the solenoid. In chicken erythrocytes, by contrast, the ~40 bp linker makes half a turn, therefore entry and exit points of the DNA would be placed alternately on the inside and outside of the solenoid. This suggests that linker histone in some chromatins may adopt positions on both the inside and outside of the solenoid, but current data predict only an internal location for linker histone (Graziano et al., 1994). An alternative possibility is that the H1 varies its positions on the chromatosome, but this is currently a point of controversy.

1.2.3.2 Helical models: The twisted-ribbon models

In models of this type (Worcel et al., 1981; Woodcock et al., 1984), the chromatosomes form a 'ribbon' (indicated as relaxed zig-zag in Figure 1.4C); this ribbon is wrapped on the surface of a cylinder to form a helical structure. A *double* helix will be formed, where even-numbered chromatosomes are in one helix, and odd-numbered chromatosomes are in the other, with the linker DNA between adjacent nucleosomes 'zigzagging' back and forth between the two helices (Figure 1.4C).

Three variations on this model were proposed (Worcel et al., 1981), each with a different linkage number increment (ΔL). The increase in linkage number (ΔL) describes the number of turns introduced when the DNA is twisted about itself (Lewin, 1997). The differences in ΔL are created by the way in which the linker DNAs cross each other as they zigzag back and forth. In the model favoured by Worcel, $\Delta L = n$ (where n = number of nucleosomes), due to the contribution from the crossing of the linker DNA in the helical ribbon itself and from the 1.75 turns of DNA around the nucleosome. This fitted with the evidence from SV40 minichromosomes and circular DNA reconstituted into nucleosomes, where ΔL also equals n (McGhee and Felsenfeld, 1980; Wang, 1982). The model failed to take into account the contribution made to ΔL by the DNA helix which is overwound when bound to a nucleosome (average 10.2 bp per turn), compared with in solution (10.6 bp per turn). Therefore, Woodcock et al. (1984) proposed an alternative structure where $\Delta L = 2n$. In this model, the helix could be right or left-handed, with a pitch of 32 nm and 18 nucleosomes per turn when fully compacted. The very centre of the helix would be hollow and the linker histone would be located in the ribbon itself, not fully exposed on the surface. The diameter was proposed to be about 30 nm, though it has the potential to vary a great deal from this. Linker length would not influence the diameter of the helix, but it would affect the width of the ribbon, and therefore nucleosome spacing. However, studies show that nucleosome spacing in the

30 nm fibre is not influenced by linker length (reviewed in Ramakhrishnan, 1997).

1.2.3.3 Helical models: Cross-linker helical models

In these models, the linker DNA criss-crosses the fibre in a direction roughly perpendicular to its long axis. The nucleosomes are arranged non-sequentially; that is, nucleosomes which are adjacent in the 'beads-on-a-string' structure will not occupy neighbouring positions in the DNA helix. Instead, they will adopt positions opposite each other across the filament axis; hence the requirement for linker DNA to transversely cross the fibre axis. The linkers would be straight, or almost straight, and H1 would be located towards the centre of the helix.

The models are further subclassified according to the number of helical ramps in the structure. Staynov (1983) proposed a model with a single helix. Williams et al., (1986) proposed two variations on a model with two helices. Makarov et al. (1985) proposed a structure of 3 intertwining helices.

1.2.3.4 The supranucleosomal particle models

Zentgraf and Franke (1984) proposed the 'superbeads' model, on the basis of electron micrograph studies carried out on chromatin isolated in 100 mM salt. They observed that the chromatin fibre was composed of a series of globular structures which they dubbed 'superbeads'. If placed in low salt buffer, these structures could be seen 'unravelling', via helical intermediates, into the beads-on-a-string structure. Many criticisms of this model have been made; it has been suggested that the 'superbeads' are the result of partially unravelled helical structures, or the result of side-to-side packing of helices. These criticisms have been countered in turn by Zentgraf and Franke who propose that the different sizes of the 'superbead' from different species indicates that these structures are not produced by detached coils of solenoid.

1.2.3.5 The 'irregular' model

It has been suggested that as the physical properties of the chromatin fibre, such as linker length, can vary, a regular structure may not be achieved (Trifonov, 1990).

Therefore, a model was proposed where the nucleosomes formed a zigzag ribbon with straight linker DNA which was subsequently folded irregularly into a higher order structure (Horowitz, 1994; Woodcock, 1994). This is shown in Figure 1.4B.

This model would lead to an irregular placement of linker histone, whilst current evidence favours a location at the centre of the chromatin fibre (Graziano et al., 1994).

1.2.3.6 Weighing the evidence for and against the various models

The detailed structure of the 30 nm chromatin fibre remains a subject of controversy, and many different models have been proposed. Others have suggested that a 30 nm fibre does not exist at all (Horowitz et al., 1994; van Holde and Zlatanova, 1995). The models proposedhave a number of features which distinguish them, and these have been studied in order to define which structure is adopted *in vivo*. These features include the mass per unit length of the fully compacted fibre, the relationship between fibre diameter and linker length, the location of linker histone in the fibre and whether the linker DNA is straight or bent. A great deal of evidence has been collected on these properties, but much of it is contradictory or inconclusive.

The mass per unit length of the chromatin fibre was determined in solution, and it was suggested there were about 6 nucleosomes per turn of the 30 nm fibre) (Suau et al., 1979; ; McGhee et al., 1983; Butler and Thomas, 1990). This fits well with the value predicted for the solenoid models and the single helix crossed-linker model. Scanning transmission electron microscopy (STEM) suggests there are up to 12 nucleosomes per turn, and the helical ribbon models were proposed to account for this. Recent neutron scattering data suggest that 6 to 7 nucleosomes per turn may be the more accurate

estimate (Gerchmann and Ramakrishnan, 1987). Others have suggested however that this lower value is due to sub-maximal compaction of the chromatin fibre.

There has been some debate about the location of linker histone in the chromatin fibre. Much of the evidence, from antibody accessibility (Dimitrov et al., 1987), from radioimmunoassay and electron microscopy (Cattini et al., 1988) and restriction enzyme accessibility (Leuba et al., 1993), suggests an internal location for the linker histone. Neutron scattering studies (Graziano et al., 1994) locate linker histone at the inner face of the nucleosome, in the interior of the fibre. Other studies contradict this (Russanova et al., 1987). An internal location for linker histone fits the solenoid model (Finch and Klug, 1976), and the modification of the solenoid model suggested by Butler (1984); it also fits the crossed-linker model (Williams et al., 1986). It contradicts the 'irregular' model and other modification of the solenoid model (McGhee, 1983).

Studies to determine whether the linker DNA is straight or bent have been inconclusive. Photochemical dichroism studies have revealed certain constraints on the conformation of the linker DNA, but did not ascertain whether the linker is bent or straight (reviewed in Ramakrishnan, 1997). It has been suggested that dinucleosomes, which show a degree of compaction over a range of low ionic strengths, have bent or kinked linkers; though dinucleosomes obviously could not form a complete turn of a higher order structure, their organisation could have relevance to the formation of a higher order structure (Butler and Thomas, 1998). It has been suggested that the irregularity of the chromatin fibre may create problems in determining the linker conformation (van Holde and Zlatanova, 1995). Other evidence suggests that the linker may be straight; in electron micrographs the linker appears straight at low ionic strength, but it is possible that it could bend at higher ion concentrations. As the curvature of the DNA increases the rate of thymidine dimer formation changes. Studies on dimer formation in the linker suggest that the linker is straight, but it is possible that the presence of histone H1 may have affected the rate of dimer formation (Pehrson, 1989). Conclusive evidence for the conformation of the linker would eliminate certain models;

the solenoid has a bent linker, and the 'irregular', crossed-linker and helical ribbon models have straight linkers.

Early studies suggested that fibre diameter was independent of linker length; this favoured the solenoid and helical models. Other studies, using cryo-electron microscopy and small angle X-ray scattering suggest a linear relationship between linker length and fibre diameter (Williams et al., 1986). This led to the proposal of the crossed-linker model. However, in calf neuronal tissue where the linker length is close to zero, a normal ~30 nm fibre is still formed; this argues against the crossed-linker structure being formed, at least in this situation.

In summary, the current evidence favours the solenoid model, but the helical ribbon, crossed-linker and 'irregular' models cannot be ruled out and, in some situations at least, these models are a better fit to the data. If unequivocal data could be produced on properties of the fibre, in particular on the linker conformation and the location of linker histone, a precise model for the 30 nm fibre could be obtained.

For gene expression to occur, the higher order chromatin must undergo unfolding to allow transcription factor to access their recognition sites and to permit the formation of the transcription factor complex (reviewed in Felsenfeld, 1996). Therefore understanding the nature of the higher order chromatin structure is critical to understanding how different factors may influence its folding or unfolding, and thereby make a contribution to the regulation of gene expression. For example, transcription of the genes is thought to take place in the context of the 30 nm fibre with transient unfolding of the chromatin (Andersson et al., 1982; Ericsson et al., 1990). The ~ 30 kb domain which encompasses the β-globin gene has a more 'open' higher order structure in erythroid cells where the globins are active than in cells where the genes are inactive (Stalder et al., 1980). The actual conformation of this more 'open' 30 nm fibre is not understood. The 'open' chromatin structure is associated with modifications such as acetylation of the histones (Hebbes et al., 1994) and association of the chromatin with HMG's (Postnikov et al., 1991). In addition, it has been hypothesised that alterations in

nucleosome spacing on the chicken globin genes during development may alter chromatin higher order structure (Evans et al., 1990). As a precise model for higher order chromatin structure remains elusive, the affect of these modifications on chromatin structure is not resolved, so the means by which they may regulate gene expression await clarification.

1.2.4 Further Levels of Higher Order Folding

The 30 nm fibre must be further coiled so that it can be accommodated in the interphase nucleus; additionally, further folding is required to produce the metaphase chromosome. There is considerable evidence for further levels of chromatin folding in the interphase and metaphase nucleus (Mathog et al., 1984; Lichter et al., 1988; Yakoto et al., 1995).

It has been proposed that the next level of folding involves coiling the 30 nm filament into a 'supersolenoid' which is itself further coiled to create the mitotic chromosome (Butler, 1983). There is a great deal of evidence to suggest that the 30 nm fibre is organised into loops, ranging from 50 to 200 kb of DNA, in both interphase nuclei and in the metaphase chromosome. These loops can be further folded and condensed to create the metaphase chromosome (Callan, 1986 and references therein), possibly by the formation of an axial loop structure, and by the further folding of these loops. These loops are constrained at either end by a non-histone protein network, and appear to be attached to nuclear matrix material in interphase nuclei and to the chromosomal protein scaffold in metaphase chromosomes (Cook and Brazell, 1975; Paulson and Laemmli, 1977). The scaffold in interphase nuclei is thought to consist of topoisomerase II and Sc2 (Lewis and Laemmli, 1982). It was also proposed that the DNA loops might attach to a peripheral lamina, which might be further stabilised by attachment to an internal protein network (Lewis et al., 1984). DNA elements, called MAR's (matrix attachment regions) and SAR's (scaffold attachment regions) have been

detected which have an increased binding affinity for the nuclear matrix; it has been proposed that these are the attachment sites for the looped domains (reviewed in Bonifer et al., 1991).

As well as allowing greater packing of the DNA, the loops may serve other equally important functions. It was predicted that DNA might be compartmentalised into independent units to allow different regions to be structurally and functionally separate (Gasser and Laemmli 1987, and references therein). These loops might orchestrate this compartmentalisation. It has been proposed that these loops might allow topological stress to be applied over a particular region, and that MAR's and SAR's might define the boundaries of DNase sensitivity over regions containing active genes (Phi-Van and Stratling, 1988). It has also been suggested that replication origins are attached or closely located to the nucleoskeleton (Razin et al., 1985). Other workers favour a model where RNA polymerases, topoisomerases and transcription factors are attached to the nucleoskeleton, and that active transcription complexes are formed when these bind to enhancers/promoters on the DNA, forming active loops bound to the nucleoskeleton (Jackson et al., 1984; Cook, 1989). It is interesting to note in this context that MAR's have been shown to contain a high concentration of transcription factor binding sites (Boulikas, 1994).

1.3 GENE ACTIVATION IS ACCOMPANIED BY CHANGES IN CHROMATIN STRUCTURE

The packaging of DNA into nucleosomes and further folding into the 30 nm fibre and higher order structures beyond this, make a decisive contribution to the regulation of eukaryotic gene expression. Early studies, which compared the sedimentation of chromatin with its RNA polymerase binding efficacy, indicated that active regions of the genome had a more open, extended conformation (Gottesfeld et al., 1975). Since then, a wealth of evidence has shown that changes in chromatin structure play a decisive

role in both the activation and repression of transcription (reviewed in Felsenfeld, 1992; Wolffe, 1994; Wu, 1997; Wolffe and Hayes, 1999). For example, early studies led to the hypothesis that packaging of DNA into nucleosomes and a higher order structure was a general repressor of transcription. These included experiments which showed that nucleosomes assembled on a promoter *in vitro* inhibited initiation of transcription (Losa and Brown, 1987; Lorch et al., 1987). In addition, depletion of histone H4 in yeast led to the derepression of a number of genes, including PHO5 (Han, 1988). Other early experiments demonstrated that the exact chromatin architecture modulated the repressive affect of chromatin. For example, the location of an ARS (autonomously replicating sequence) within a nucleosome determined the extent to which replication was affected; if the ARS was in the central region of the nucleosome the copy number was lower than if it was at the edge of the nucleosome (Simpson, 1990). Further studies have shown that chromatin has more than just a simple occlusive function; it plays a more subtle role as a component and regulator of transcriptional mechanisms (reviewed in Wolffe, 1994). For example, a positioned nucleosome on the *Drosophila* hsp 26 gene promoter brings together two transcription factors which are far apart on the DNA, and may thus facilitate transcription by causing the activation domains of these proteins to cluster (Thomas and Elgin, 1988).

Genes can exist in three states with respect to transcription (van Holde, 1989). Firstly, the gene may be inactive. Secondly, it may be 'poised' for transcription. A poised gene is not actively transcribed, but it is capable of being transcribed in the tissue concerned, either at some point in the future, or in the past. When poised, a gene is more readily induced than when in its inactive form. Thirdly, a gene may be active; that is, actively transcribed. Inactive genes are packaged in a less accessible, higher order chromatin structure: most probably the 30 nm fibre and further levels of folding above this. When a gene becomes 'poised' or active, its chromatin becomes more 'open'. The opening of the chromatin structure involves loosening and (at least transient) unfolding of the 30 nm fibre, as well as rearrangement or loss of nucleosomes over regulatory

elements in the DNA which bind transcription factors (reviewed in Felsenfeld, 1996). These changes are thought to make DNA sequences more accessible to the proteins involved in transcription. Poised or active chromatin characteristically has the following features, which correlate with the more 'open' chromatin structure (reviewed in Wolffe et al., 1999): increased sensitivity to nucleases like DNase I, modifications of the nucleosomes like acetylation and ubiquitination, hypomethylation of the DNA, depletion of linker histone, increased association with proteins like HMG 14 and 17, changes in nucleosome spacing and changes in nucleosome positioning.

Some of these changes are localised; for example, DNase hypersensitivity is localised to short DNA elements like enhancers and promoters. Other changes may occur over the whole transcribed region and beyond. For example, when the 4 kb lysozyme gene is activated, 24 kb of DNA becomes DNase sensitive (Phi-Van and Stratling, 1988). In erythroid cells, where the globin genes are active, all 23 kb of the β-globin locus, transcribed and non-transcribed sequences alike, becomes hyperacetylated and DNase sensitive in erythrocyte cells (Stalder et al., 1980; Hebbes et al., 1994).

Some changes in chromatin structure may occur in the transition from inactive to poised; other changes may be specific for the transition from poised to actively transcribed.

1.3.1 Nuclease sensitivity

Nuclease sensitivity appears to be a global property of active chromatin. Genes are more sensitive to nucleases such as DNase I in tissues where they are poised for transcription or actively expressed than in tissues where they are inactive. The enhanced sensitivity extends beyond the borders of the transcribed region, implying that the whole domain containing an active gene may be more open and accessible (Cartwright and Elgin, 1986). These active domains may correspond to the loop domains which have been suggested to exist in chromatin (reviewed in Gasser and Laemmli, 1989).

In the case of a cluster of genes with a related function which are expressed in the same tissue, a domain containing the entire cluster may become nuclease sensitive. For example, the entire ~ 30 kb domain containing the chicken β-globin genes becomes sensitive to DNase I in erythroid cells where the globin genes are active (Stalder et al., 1980; Hebbes et al., 1994). Though the individual globin genes are expressed at different stages in development, all the genes, poised and active alike, as well as the regions flanking the genes, show nuclease sensitivity (Benezra, 1986). These global changes in chromatin structure across a cluster of related genes may ensure that they exhibit co-ordinated, tissue-specific expression.

Poised and active genes show increased sensitivity to other nucleases, such as micrococcal nuclease. For example, micrococcal nuclease digestions of nuclei solubilise active chromatin more readily than bulk (Bloom and Anderson, 1978). The chicken β-globin genes show increased sensitivity to micrococcal nuclease in erythroid cells (Wood and Felsenfeld, 1992).

Enhanced DNase I and micrococcal nuclease sensitivity is a general property of poised and active genes, and the boundaries of the nuclease sensitivity often extend far beyond the transcribed region. Enhanced nuclease sensitivity precedes transcription, and appears necessary for it to occur. This is thought to indicate a more open chromatin structure which facilitates the access of transcription factors to the DNA. The exact mechanism which creates nuclease sensitivity and opens up the chromatin is uncertain, but it is thought to involve, for example, acetylation of the core histones, association with non-histone proteins like HMG's and alterations in the binding of linker histones (reviewed in Wolffe and Hayes, 1999). These will be discussed in more detail in sections 1.3 to 1.5.

1.3.1.1 DNase I hypersensitivity

Sites of DNase I hypersensitivity (DHS's) are generally located within larger regions

sensitive to DNase I. These very sensitive sites commonly occur 5′ to poised or active genes at sequences involved in gene regulation and generally cover 40 to 400 bp. Nucleosomes with an altered structure or nucleosome displacement may be responsible for their enhanced sensitivity (reviewed in Wolffe, 1992). DHS's occur primarily over promoter and enhancer regions, and probably help proteins involved in transcriptional regulation to gain access to their sites on the DNA (Gross and Garrard, 1988). A schematic (Figure 1.5) depicts how the unfolding of an entire chromatin domain, and nucleosome rearrangement over regulatory regions to create DHS's contributes to the activation of transcription.

Some DNase hypersensitive sites are constitutive; i.e. they appear when the gene becomes poised for transcription. Other sites are 'inducible'; i.e. they appear when the gene becomes active. One well-studied example of tissue-specific constitutive DHS's occurs upstream of the human β-globin genes. These 5′ hypersensitive sites are thought to open up the chromatin 'loop' domain containing the β-globin locus and to act as enhancers of transcription by interacting with the individual gene promoters (Grosveld et al., 1987).

The mechanisms which create DNase sensitivity and DNase hypersensitive sites are not fully understood. It is thought that DNA supercoiling may play a role in generating the open, DNase I sensitive domain. The looped domains which were proposed to form part of the higher order structure in eukaryotic nuclei may allow different degrees of supercoiling in different domains, and this could affect transcription (Cook, 1989). Other experiments also suggest a link between supercoiling, chromatin conformation and transcription. Blocking Topoisomerase II activity caused a decrease in DNase sensitivity over the chicken β^{A} globin gene, presumably by decreasing the degree of supercoiling (Reitman and Felsenfeld, 1990). DNA secondary structures may contribute to the formation of DHS's. Evidence for this comes from the observation that DHS's can sometimes also show sensitivity to nucleases specific for single-stranded DNA, including S1 nuclease. In many cases, the sites which are sensitive to S1 nuclease in

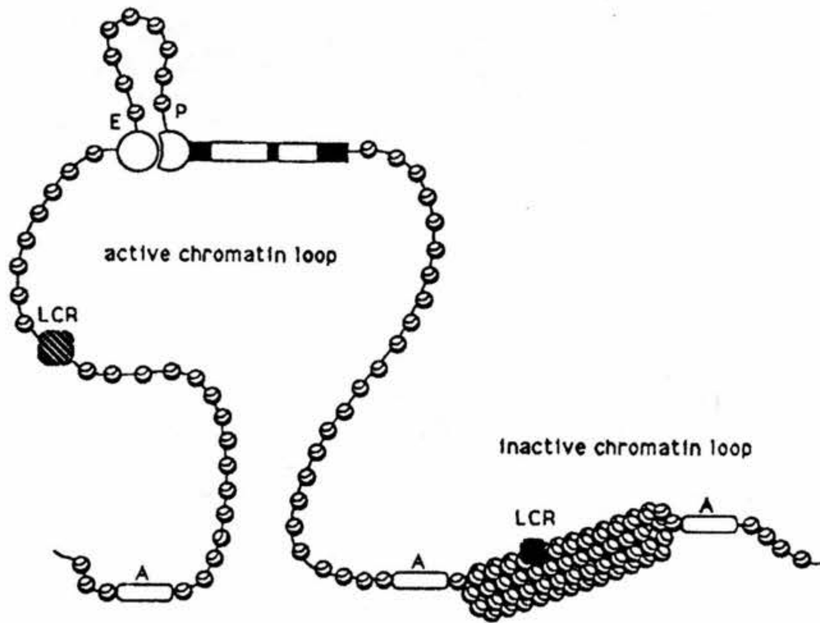


Fig. 1.5 Schematic showing the structure of proposed chromatin loops when they are in active and inactive configurations

The chromatin loop domain on the right is shown in a condensed, repressed state. Loosening of the chromatin structure within the domain, or unfolding of the chromatin domain as shown on the left, puts the gene contained within the loop into a 'poised' state. To allow gene transcription to occur, nucleosomes on the enhancer (E) and promoter (P) are rearranged or displaced, so that the promoter and enhancer can interact. The LCR may be important in regulating opening of the domain. A represent sites of attachment to the nuclear matrix, and potential domain boundaries (adapted from Lewin, 1997).

active chromatin are also sensitive in a supercoiled plasmid (the sensitivity is lost if the plasmid is linear so there is no supercoiling). This is observed in the region upstream of the chicken β^A globin gene (Larsen and Weintraub, 1982). It has been proposed that a subclass of hypersensitive sites might be associated with Z DNA (Rich et al., 1984), and that the presence of Z DNA may enhance transcription (Sinden, 1994).

1.3.2 Nucleosome spacing

Studies suggest increased transcriptional activity generally correlates with decreasing internucleosomal repeat distance (Browne and Sutcliffe, 1987), though a few reports contradict this (Smith et al., 1983).

Villeponteau et al, (1992) showed that nucleosomes on the actin, ovalbumin and globin genes are spaced up to 40 base pairs closer together when they are in an active or poised state than when they are in an inactive state. The compressed nucleosome spacing extends for several kilobases beyond the transcribed region, indicating that it is not just a result of mechanical disruption caused by transcription.

It has been suggested that nucleosome spacing could affect the folding of chromatin into a higher order structure. Butler (1984) proposed that if chromatin were to adopt a solenoid structure with a bent linker, certain linker lengths (163-176 bp) could not be folded into a higher order structure. Linker length would also have a profound affect on the stability of chromatin higher order structure if it adopted the structure proposed by McGhee et al. (1983). Therefore, alterations in linker length during development could significantly affect the packaging of chromatin into a higher order structure. It has also been suggested that the irregularity of linker length might preclude the formation of ordered structures like the solenoid, helical ribbon or crossed linker models (Trifonov, 1990). Therefore, it is energetically more favourable for regions which do exhibit a very regular nucleosome spacing to form a regular, stable higher order structure. Davey et al. (1995) suggested that the 200 bp periodicity of strong nucleosome positioning sites in

the β^A-gene seen *in vitro* might facilitate the packaging of chromatin into higher order structure in tissues where it is inactive.

Nucleosome spacing also has the potential to influence sequence directed nucleosome positioning. It was argued early on that the majority of nucleosomes in the genome could not be positioned with respect to the underlying DNA sequence because different tissues had different nucleosome repeat lengths. Therefore, altering nucleosome spacing within a cell modulates the recognition and interpretation of the positional information encoded within the DNA sequence.

The factors which determine and influence nucleosome spacing remain elusive. It has been proposed that changes in linker histone levels, leading to a change in the histone to DNA ratio, are responsible for the alteration in nucleosome spacing during development in red blood cells (Villeponteau et al., 1992). Periodic positioning signals within the DNA sequence may play a role in some situations (Davey et al., 1995; Liu et al., 1993; 1997), whilst others suggest that signals in the DNA sequence are subordinate *in vivo* to other determinants. Blank and Becker (1996), studied nucleosome spacing on DNA containing a strong periodic nucleosome positioning signal which had been assembled into chromatin in a *Drosophila* extract over a range of salt conditions. From their results they argued that electrostatic effects are more important than DNA sequence in determining nucleosome spacing. The extract, however, comes from *Drosophila* embryos, which must be able to replicate their DNA very rapidly, so the mechanism of chromatin assembly may not be representative of other cells.

There is some evidence that linker histone can affect nucleosome spacing but there is also some evidence against this (discussed in Wolffe, 1997).

1.3.3 Histone acetylation and gene activity

The amino-terminal tails of core histones contain lysine residues whose epsilon amino groups can be acetylated. Nascent histones are acetylated and then rapidly

deacetylated after incorporation into chromatin. Subsequent changes in core histone (H3, H4 and H2B) acetylation which accompany gene activation are a result of the differential activity of histone acetyltransferases and deacetylases. There is a great deal of evidence linking acetylation of core histones and gene activation (reviewed in Van Holde, 1988). It has been demonstrated that antibodies to acetylated lysine residues react preferentially with active alpha D globin, but not with the inactive ovalbumin gene (Hebbes et al., 1988). Conversely, gene repression is linked to hypoacetylation (Braunstein, 1993).

Acetylation is not just a property of actively transcribing genes; it appears to precede transcription. Indirect immunofluoresence shows H4 acetylation precedes transcription in *Tetrahymena* (Pfeffer et al., 1989). On active and poised genes, the hyperacetylated region extends beyond the transcribed region; in erythrocyte cells, for example, the entire β-globin locus, active genes, poised genes and flanking regions alike, is acetylated. The acetylated region appears to coincide with the region exhibiting generalised DNase I sensitivity (Hebbes et al., 1994).

The mechanism by which acetylation influences gene expression is not completely understood, and it may work in several ways. Acetylation may affect individual nucleosomes; by greatly reducing the affinity of the histone H4 tail for DNA (Hong et al., 1993) and causing a conformational change within the nucleosome itself. This enhances the accessibility of DNA bound by the nucleosome, which in turn can enhance transcription factor access to the DNA (reviewed in Wolffe et al., 1999). For example, acetylation has been shown to facilitate the binding of certain transcription factors such as GAL4 to the DNA (Vettese-Dadey et al., 1996). Therefore, histone acetylation may provide a molecular mechanism whereby DNA can be made more accessible to transcription factors whilst it is still bound by a nucleosome (Wade et al., 1997). It has also been suggested that acetylation might disrupt internucleosome contacts made by the H4 tails, and thereby destabilise chromatin higher order structure. Evidence supporting this comes from the treatment of cells with sodium butyrate, which leads to

hyperacetylation and apparent unfolding of the chromatin (reviewed in Wolffe et al., 1999). Some studies suggest, however, that that there is only a modest reduction in the wrapping of the DNA around the histone octamer, and a slightly less efficient packing of the nucleosome into arrays (Norton et al, 1989; Garcia-Ramirez et al., 1995).

Recently, there has been considerable progress in understanding the targeting of acetylation to particular regions of DNA, with the discovery of a number of transcription factors, including yeast Gcn5 (Grant et al., 1997) and human TAF_{II} 250 (Mizzen et al., 1996) which are also acetyltransferases. The recruitment of co-activators could lead to hyperacetylation which would cause the displacement/rearrangement of nucleosomes, allowing the formation of a transcriptional complex. There is also some evidence that deacetylase may associate with transcription factors which can act as repressors (reviewed in Wu, 1997). Acetylation is a balance of acetylation and deacetylation, the maintenance of gene activity might require the continued presence of the co-activators/acetyltransferases. Therefore the transcriptional machinery uses the packaging of the DNA as a way of continuously controlling and altering gene activity (Wade et al, 1997).

1.3.4 Linker histones and gene activity

Early experiments suggested that linker histones generally acted as repressors of gene expression (reviewed in Van Holde, 1989). This fitted well with data showing that some active genes were depleted in linker histone. It was proposed that nucleosomes played a role in organising higher order structure, so that depletion of linker histone might lead to chromatin unfolding (Thomas, 1984). The situation is complicated by the fact that some active genes are still associated with linker histones; other modifications may allow the chromatin to unfold. Studies have also demonstrated that transcription factors have decreased access to binding sites which occur within positioned chromatosomes containing linker histone (see section 1.3.6) (Richard-Foy and Hager,

1987; Pina et al., 1990). Studies have shown that binding of H1 can restrict the mobility of nucleosomes so that they can only adopt a subset of the possible alternate translational positions on the DNA (Ura et al., 1995; Pennings et al., 1994). On the MMTV promoter it was proposed that removal of linker histone H1 would enhance nucleosome mobility and facilitate transcription factor access (Bresnick et al., 1992). In agreement with this mechanism for H1 activity, Panetta et al. (1998) have shown *in vitro* that when H1 binds nucleosomes on the *Xenopus* 5S rRNA oocyte gene, repositioning of the nucleosomes is prevented, and TFIIIA binding is prohibited. In this way, H1 plays a direct dynamic role in regulating access of the transcriptional machinery.

More recent studies have indicated that linker histone has a more subtle role in regulating gene expression than just blanket repression. For example, H1 is required for the activated expression of CyP in starved *Tetrahymena* cells (reviewed in Wolffe et al., 1997). It has also been shown that the appearance of H1 in *Xenopus* is responsible for the inactivation the oocyte 5S rRNA gene, and the activation of the somatic 5S gene (Panetta et al., 1998).

Certain modifications of linker histones are known to occur: Poly-ADP ribosylation and phosphorylation. Poly-ADP ribosylation of H1 is thought to decrease the stability of higher order structure (Aubin et al., 1983).

It has been proposed that the linker histone variants occurring within a cell make a contribution to the pattern of its gene expression. It is also known that linker histone variants are sometimes localised in particular chromosome domains (Schulze et al., 1993). For example, replacement of linker histone B4 with H1 during embryogenesis in *Xenopus* leads to a switch in expression from the oocyte 5S rRNA gene to the somatic 5S gene (reviewed in Wolffe, 1997; Kandolf, 1994). The appearance of linker histone variant, H5, in chicken erythrocytes during terminal differentiation is thought to be responsible for chromatin compaction and transcriptional inactivation (reviewed in Wolffe, 1997).

1.3.5 Further characteristics associated with active chromatin

Other features which correlate with active chromatin include further modifications of core histones, such as H2B ubiquitination, association of chromatin with non-histone proteins like HMG's and hypomethylation of cytosine within the DNA.

In vertebrates, methylation occurs at the cytosine of the dinucleotide CpG. In general, hypomethylation of the DNA correlates with gene activation and hypermethylation with gene inactivation (Razin and Cedar, 1991; Bird, 1992). On the chicken ρ-globin gene for example, all CpG's in the promoter are methylated at day 14 *in vivo* when it is inactive; all CpGs are unmethylated at day 5 when it is active (Singal et al., 1997).

Various mechanisms have been proposed by which methylation could affect gene transcription. Firstly, methylation of transcription factor recognition sites could interfere with their binding. This is not necessarily the case as the methylation of some common transcription factor binding sites, such as that of Sp1, does not stop binding (Harrington et al., 1988). Experiments showing that methylation of the coding regions of a gene can lead to gene inactivation (Graessman et al., 1994) demonstrate that methylation can cause repression by other mechanisms.

Methylation may create changes in chromatin structure which could influence gene expression. This proposal is supported by experiments showing that repression of the HSV thymidine kinase gene by methylation required assembly of the gene into chromatin (Buschhausen et al., 1987). Other studies indicated that methylated DNA, when integrated into the mouse genome, adopted a chromatin conformation which had many of the features associated with a repressive chromatin structure (Keshet et al., 1986). Methylation could cause changes in chromatin structure directly, for example by altering nucleosome positioning, or by altering the interaction with linker histones. It

has been shown *in vitro* that methylation at a small number of sites is sufficient to displace a strongly positioned nucleosome upstream of the β^A-globin gene (Davey et al., 1997). It is thought that the methylation displaces this nucleosome by altering the DNA structure so that it can no longer accommodate nucleosome formation. Methylation can affect DNA bending, which will influence nucleosome formation, and promote the formation of unorthodox DNA structures such as Z-DNA which may exclude nucleosomes (Davey et al., 1997 and references therein). The effect of methylation on the binding of linker histone is still contentious (McArthur and Thomas, 1996; Campoy et al., 1995). There is also strong evidence that methylation has an indirect affect on chromatin structure, mediated by a set of non-histone proteins (dubbed MeCP's), which specifically bind methylated DNA (Meehan et al., 1992). MeCP2 has been shown to interact with histone deacetylase, so in some cases it may recruit the deacetylase complex and lead to gene inactivation (Nan et al., 1998).

Ubiquitin is attached to the ε amino groups of lysine side chains in H2A and H2B. Ubiquitination is increased on poised and active chromatin, and it may prevent close packing of the nucleosomes (Nickel et al., 1989; reviewed in Wolffe and Hayes, 1999).

The HMG proteins are a heterogeneous family of chromosomal non-histone proteins. They are divided into three classes: HMG 1/2, HMG 14/17 and HMG I/Y. The amount of HMG proteins in the nucleus is limited, and they are thought to be concentrated in particular regions of the genome. For example, HMG 14/17 are preferentially associated with regions containing transcribed genes; they occur on the β-globin gene cluster in erythrocyte cells (Postnikov et al., 1991). Although it is thought that they may contribute to the regulation of gene expression, more work is required before their role is understood. In particular, the mechanism by which HMG proteins may influence chromatin structure and gene expression remains uncertain. It has been suggested that HMG-1 can displace H1 to cause gene activation (Varga-Weisz et al., 1994). HMG 14/17, which bind nucleosomes with a higher affinity than DNA, are also thought to be capable of interfering with linker histone binding. They may also alter the

interaction of the core histone tails with the DNA, and thereby influence higher order structure (Alfonso et al., 1994). Other studies suggest that HMG 14/17 enhance elongation by transcription complexes, if incorporated into nucleosomes during chromatin assembly (Bustin, 1995). HMG I/Y may also be able to compete with H1 for binding, and may be able to alter the supercoiling of the DNA which could influence chromatin structure (Nissen, 1995).

1.3.6 Nucleosome Positioning

It has been argued that precise positioning of nucleosomes over a whole genome could not be possible, as nucleosome spacing varies in different cell types of an organism (Kornberg and Stryer, 1988). Early studies suggested that the majority of nucleosomes in *Drosophila* embryos and rat liver are not positioned (Prunell and Kornberg, 1978). A wealth of evidence has shown subsequently that a subset of nucleosomes in chromatin are positioned, and that they can play a decisive role in gene regulation (reviewed in Simpson, 1991; Wolffe, 1994; Lu et al., 1994).

It is necessary to define what is meant by nucleosome positioning. A positioned nucleosome occupies a precise location with respect to the underlying DNA sequence. The *translational* position of a nucleosome is defined by the entry and exit sites of the DNA wrapped around the nucleosome. The *rotational* position of the nucleosome describes which DNA sequences face the nucleosome, and which face the surrounding solution. The rotational and translational placement of a nucleosome are not, of course independent. At the dyad (centre) of the nucleosome, the minor groove faces outward, and the major groove inward. The orientation of the major groove is of interest as most sequence-specific contacts with DNA binding proteins (e.g. transcription factors) are made there. As the DNA which wraps around the nucleosome has an average ~10 bp per turn, every multiple of 10 bp from the dyad, the major groove faces inward. By contrast, every odd multiple of 5 bp from the dyad the major groove will face outward (Figure 1.6a).

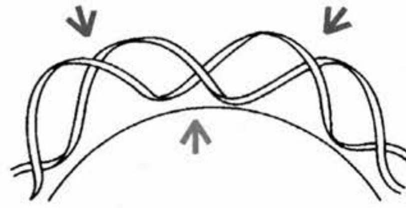


Figure 1.6a Preferential placement of A/T and G/C nucleotides in nucleosomal DNA

It has been observed that A/T dinucleotides occur preferentially at sites of minor groove compression (where the minor groove faces the histone core: red arrow), and that G/C nucleotide are favoured where the major groove faces inwards (blue arrows).

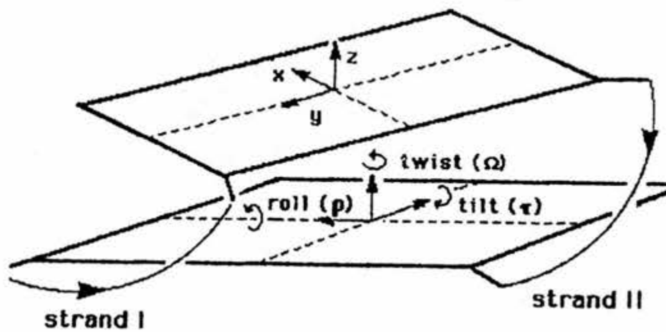


Figure 1.6b Nucleotide step orientational parameters

Further work has shown the mechanisms by which positioned nucleosomes regulate gene transcription. When nucleosomes are positioned on DNA regulatory elements such the formation of the transcriptional complex. There are four main ways in which positioned nucleosomes may regulate transcription factor access (reviewed in Thoma, 1992; Wolffe, 1994).

Firstly, nucleosomes can occlude transcription factor binding sites and reduce their accessibility. This explains the appearance of DNase hypersensitive sites in promoters and enhancers prior to gene transcription; the nucleosomes which occlude important sites must be rearranged or displaced to allow transcription complex formation. More recent experiments suggest that positioned nucleosomes can play a role which is more subtle than simple repression. Positioned nucleosomes can ensure that a recognition site is exposed in the linker where it is accessible for binding. Alternatively, positioned nucleosomes, by bringing together two distant protein binding sites on the surface of a nucleosome, may lead to enhanced transcription factor binding. Lastly, some transcription factors may require their recognition sites to be 'presented' on the surface of a nucleosome to enhance binding. Examples from the literature, covering these mechanisms for influencing transcription factor binding and gene expression, are given below.

Experiments in yeast showed how the removal of nucleosomes positioned over important transcription factor binding sites could relieve gene repression. For example, the loss of four positioned nucleosomes which occlude a TATA box and an upstream activating sequence are required for activation of the PHO5 gene (Fascher et al., 1990). Work on the MMTV LTR has shown that incorporation of a protein binding site into a nucleosome does not always abolish binding; the exact position of the recognition site within the nucleosome is critical in determining the accessibility of the site. These studies on MMTV *in vitro* (reviewed in Wolffe et al., 1994) showed that the rotational position of the glucocorticoid recognition element has a large effect on its accessibility to proteins; when the element faces towards the histone core binding is almost

completely abolished; when the element faces the solution binding affinity is reduced by only three to 11-fold compared with the affinity for free DNA. The translational position of the element within the nucleosome causes a variation in affinity, but over a narrower range. Therefore, repression by nucleosome placement is not simply 'yes' or 'no'; several levels of repression are possible (reviewed in Wolffe et al., 1994). Work on the accessibility of restriction enzyme sites bound by nucleosomes has also shown that a range of accessibility exists, dependent on position within the nucleosome (Polach and Widom, 1995).

A recent example indicates how the precise location of a positioned nucleosome relative to a transcription factor binding site can be critical. Differential nucleosome positioning on the *Xenopus* somatic and oocyte 5S rRNA genes has been shown to be responsible for the different expression patterns of these two genes (Panetta et al., 1998). Though the regulatory elements of these two genes show close sequence homology, there are big differences between the multiple *in vitro* nucleosome positioning sites. The predominant positions on the somatic gene occlude only part of the TFIIIA binding site, so that TFIIIA is able to bind the promoter. By contrast, most of the positions on the oocyte promoter completely occlude the site, and TFIIIA binding is only possible with extensive alteration of nucleosome positions, which is prevented if H1 is incorporated into the nucleosomes. Therefore nucleosome positioning plays a key role in regulating the transcription of the oocyte and somatic 5S rRNA genes.

It has also been shown that precise location of an ARS (autonomously replicating sequence) within a positioned nucleosome influences the extent to which replication is affected. If the ARS was located at the centre of the nucleosome the decrease in copy number was more marked than if it was at the edge (Simpson, 1990) The terminal 20 bp of the nucleosome is more accessible, and the histone DNA contacts are more easily broken (Polach and Widom, 1995).

Nucleosomes may be precisely positioned in order to expose important protein binding sites in the linker DNA. Nucleosomes on the *Xenopus* vitellogenin B1 genes

and *Drosophila* hsp 26 are positioned so that known transcription factor binding sites lie in the linker DNA (Schild et al., 1993; Thomas and Elgin, 1988).

Positioned nucleosomes can facilitate transcription on certain promoters by bringing together on their surface sites which are far apart on the DNA. The wrapping of the DNA around the positioned nucleosome on the *Drosophila* hsp26 promoter *in vivo* brings together two heat shock transcription factor sites in adjacent linkers; this may cluster the activation domains and initiate transcription (Thomas and Elgin, 1988). Positioning a nucleosome between the NF1 and oestrogen receptor binding sites is thought to bring these sites together, increasing the transcription level five to ten-fold (reviewed in Wolffe et al., 1994).

McPherson et al., (1993) proposed that activation of the serum albumin gene in mice requires the formation of three positioned nucleosomes over transcription factor binding sites in the enhancer. Therefore activation might require the recognition sites to be presented on the surface of the nucleosome for activation to occur. Certain rearrangements in nucleosome structure may facilitate factor binding.

In summary, positioned nucleosomes are nucleoprotein complexes which interact with transcription factors to allow dynamic, complex regulation of gene expression.

1.3.7 Determinants of nucleosome positioning

Five main mechanisms for creating positioned nucleosomes have been proposed. Firstly, nucleosomes could be positioned with respect to a boundary. This boundary could be created by a positioned non-histone protein, or a positioned nucleosome, or by a DNA sequence which excludes nucleosomes. Secondly, nucleosomes could be positioned relative to a replication origin, as nucleosomes form as soon as enough DNA has been replicated. There is as yet little direct evidence for this model and some evidence in yeast which suggests it does not occur (Thoma and Simpson, 1985). Thirdly, certain proteins might bind the nucleosome and influence its position. In

addition, the formation of higher order structures may influence nucleosome positioning. Finally, and perhaps most importantly, histone: DNA interactions can determine nucleosome positioning. DNA sequence and DNA secondary structure certainly play a major role in this.

1.3.7.1 Boundary-directed nucleosome positioning

It has been suggested that boundaries created by a protein, a positioned nucleosome or a DNA sequence which excludes nucleosomes, could direct nucleosome positioning. It has been proposed that next to the boundary, specific locations for the nucleosome would be strongly favoured, and the precision with which nucleosomes were positioned would decay further from the boundary. A longer linker would lead to a greater decay in precision with increasing distance (Kornberg and Stryer, 1988).

Several studies show that non-histone proteins could act as boundaries. The yeast proteins GRF2 and $\alpha 2$ have been shown to be capable of directing nucleosome positioning (Chasman et al., 1990; Shimizu et al., 1991). It has been proposed that boundaries could be created by regions containing certain DNA structures which exclude nucleosomes: for example, cruciform DNA, Z DNA or long stretches of poly(dA)•poly(dT) (Nickol et al., 1982; Simpson and Kunzler, 1979).

It has been proposed that two boundary elements, one created by a DNase sensitive site and the other by a positioned nucleosome, are important in directing nucleosome positioning on the ura3 gene in yeast (reviewed in Thoma, 1992). It has been proposed that boundary-directed nucleosome positioning is a common feature of yeast genes (Thoma, 1992).

Wang and Griffith (1996) showed that $(G/C)_3NN$ repeats exclude nucleosomes *in vitro*. Sequences like these are found *in vivo*; for example, they occur where the DNase I hypersensitive sites are located on the human dihydrofolate reductase genes (discussed in Wang and Griffith, 1996).

1.3.7.2 Role of higher order structure in determining nucleosome positioning

It has been suggested that folding a chromatin fibre into a higher order structure might influence nearest neighbour interactions, and thereby affect nucleosome positioning. This suggestion followed work by Thoma et al. (1979) which showed that manipulating the folding up of a yeast plasmid affected nucleosome positioning on the plasmid. This is still poorly understood as higher order chromatin structure is still poorly characterised.

1.3.7.3 Role of linker histone in determining nucleosome positioning

It has been suggested that linker histone binding to the nucleosome *per se*, or as part of a higher order structure could influence nucleosome positioning (Thoma and Zatchej, 1988). A comparison of nucleosome and chromatosome positions on 1.5 kb of the β-globin gene *in vitro*, in the absence of any higher order structure, suggested linker histone did not greatly alter nucleosome positioning (Shen, 1997). Other studies suggest that incorporation of linker histone into the nucleosome can restrict nucleosome mobility, and restrict it to one of several alternate positions (Pennings et al., 1994).

1.3.7.4 DNA sequence-directed nucleosome positioning

Many studies have indicated that histone:DNA interactions are capable of positioning nucleosomes *in vitro* (reviewed in Simpson, 1991; Thoma, 1992). Other studies suggest that DNA sequence also makes a contribution to nucleosome positioning *in vivo*; on a number of DNA sequences, it has been shown that positioning *in vitro*,

where only DNA and histones are present, is very similar to that seen *in vivo* (Butinelli et al., 1993, and recently, Adroer et al., 1998). Work by Tanaka et al. (1992) suggests that in some cases, the contribution of DNA sequence can be subordinate to other determinants of nucleosome positioning. This work showed that when a strong rotational positioning signal was introduced into yeast DNA *in vivo* it did not contribute to nucleosome positioning.

As nucleosomes package the entire genome, it is not surprising that nucleosome positioning signals are highly degenerate, and generally appear to be the additive result of many DNA-histone contacts within the core particle (Neubauer et al., 1986). To fold the DNA helix around a nucleosome, the DNA must be bent in a circular trajectory, with a radius of 44Å. The DNA helix becomes overwound and distorted at particular sites. Some DNA sequences will accommodate this distortion and bending better than others. There is now a great deal of evidence to suggest that both anisotropic bendability and the intrinsic curvature of the DNA, which is a function of its sequence, make a major contribution to rotational positioning of the nucleosome (reviewed in Thoma, 1992). Drew and Travers (1985) demonstrated that when DNA was circularised, AAA or TTT trinucleotides were positioned with the minor groove facing inwards, and trinucleotides like GGG, CCC and GGC were positioned so that their minor groove was on the outside of the DNA. By sequencing 200 nucleosome positioning sites, Satchwell et al. (1986) observed that AAA/TTT and AAT/ATT trinucleotides were positioned with a 10.2 bp periodicity, so that their minor groove faced the nucleosome core. GC sequences were also found to have a 10.2 bp periodicity, but this was out of phase with the AT periodicity so that GC bases occurred where the minor groove faced outwards. Put simply, A/T sequences were favoured at sites of minor groove compression, facing the histone octamer (Figure 1.6a). The tilt and roll (Figure 1.6b) characteristics of these A/T dinucleotide steps means adjacent nucleotides form a 'wedge' shape. Therefore the periodic occurrence of these nucleotides can introduce unidirectional bending into the DNA to facilitate nucleosome binding. There was concern that stripping H1 from the

chromatin influenced the results, because in a study of chromatosomes the periodicities were less marked. However, other studies have confirmed the importance of A/T and G/C periodicities and DNA bending. Ulyanov and Stormo (1995) demonstrated a significant periodic occurrence of the pattern AA/TT$(N)_3$CA/TG. Widlund et al., (1997) proposed that TATA tetranucleotides occurring with a 10 bp periodicity created strong nucleosome positioning sites. Guided by previous results, Schrader and Crothers (1989) designed DNA sequences containing two to twelve copies of an 'optimal' rotational positioning signal, $(A/T)_3NN(G/C)_3NN$, tandemly repeated. Competitive reconstitution experiments, which measure the binding affinity of sequences, were carried out on naturally occurring positioning sequences and these artificial sequences. They observed that nucleosome formation on four copies of the artificial sequence was as energetically favourable as on the best of the naturally occurring positioning signals. Hydroxy radical cleavage also suggested that the nucleosomes were bound to the DNA in the expected rotational setting.

A recent study (Lowary and Widom, 1998) selected sequences with the highest affinity for the histone octamer from a pool of random sequences. They have discovered a number of sequences which enhance nucleosome formation. These include a 20 bp spacing of TA dinucleotides and CTA repeats at 10 and 20 bp intervals

The periodic repeats described above may provide a *rotational* signal for positioning nucleosomes. The sequences which contribute to the *translational* setting of a nucleosome have been more elusive. A disruption in the A/T and G/C periodicities was observed near the dyad (Satchwell et al., 1986). When the artificial positioning sequence (Schrader and Crothers, 1989) containing ten $(A/T)_3NN(G/C)_3NN$ repeats, was incorporated into yeast minichromosomes *in vivo*, nucleosomes centres were excluded by the sequence, and nucleosomes formed at the edge of the sequence (Tanaka et al., 1992). Again, this suggests that a discontinuity in the repeats may be required at the centre of the nucleosome, to accommodate the structural distortions which occur there. Other studies suggest that this discontinuity in the A/T and G/C periodicities, in

combination with an asymmetry between the two halves of the nucleosomal DNA, may act as a translational signal (Fitzgerald and Anderson, 1998). The high resolution structure of the core particle (Luger et al., 1997) reveals that there is 'kinking' of the DNA at ~15 and 45 bp either side of the dyad. Using a multi-alphabet consensus algorithm, Ulyanov and Stormo suggested that the patterns MMMNNMMM (where M=A or C) and RRRNNRRR (where R=A or G) are preferentially located 40 and 50 bp away from the dyad respectively. The authors do not address whether these motifs are particularly suited to accommodate the kinking that occurs at these sites.

1.3.8 Prediction of sequence-directed nucleosome positioning

The DNA sequence patterns which contribute to nucleosome positioning are not yet fully understood, but based on the current knowledge several algorithms have been produced to predict positioning. Some of these methods are based on looking for the periodic sequence patterns found to occur in nucleosomes, and some are based on theoretical models of how easily different sequences can accommodate bending around the nucleosome.

1.3.8.1 Statistically based algorithms

As discussed earlier, numerous studies have revealed a ~ 10 bp out of phase periodicity of A/T and G/C which causes unidirectional bending of the DNA, facilitating nucleosome formation (Satchwell, 1986; Muyldermans and Travers, 1994). This has been used as the basis of an algorithm for predicting nucleosome rotational positioning, which has proved quite successful (Calladine and Drew, 1986; Drew and Calladine, 1987).

In an attempt to discover other sequence patterns involved in rotational and translational positioning of the nucleosome, more advanced computational methods

have been applied. Multiple alignment algorithms have been used (Ioshikhes et al., 1996; Bolshoy et al., 1996). Known nucleosome positioning sequences, about 200 in number, were aligned and searched for similar patterns. Multi-alphabet consensus algorithms have also been used (Ulyanov and Stormo, 1995). Essentially, the occurrence of all possible sequence patterns, up to a given pattern length, is calculated. This is compared with the expected occurrence to see which patterns occur at a significantly high frequency. As this requires a huge amount of computation, the work involved can be reduced by searching for a degenerate alphabet rather than using the full alphabet. Hidden Markov models (Baldi et al., 1996) have also been applied. Though other features have been revealed, including the discontinuity of the A/T and G/C periodicity in the central ~30 bp, and certain sequences which appear to occur at set distances from the dyad, they do not yet allow successful prediction of nucleosome position.

1.3.8.2 Structurally based algorithms

The statistical studies which showed a 10 bp out of phase periodicity in A/T and G/C nucleotides could be explained satisfactorily by the sequence-directed curvature which these repeats would produce. Each pair of neighbouring nucleotides (known as a nucleotide step), has its own characteristic value for certain parameters, called roll and tilt angles (Figure 1.6b). The roll and tilt angles for certain dinucleotide steps is such that they produce a bend (higher roll and tilt angles can produce a 'wedge' shape: see Figure 1.6b). Therefore, a periodic repeat of these can lead to unidirectional bending.

Roll and tilt angles have been estimated for each of the possible 16 dinucleotide steps, allowing the overall curvature of a piece of DNA, and its suitability for nucleosome positioning to be investigated (De Santis et al., 1993). Their results at prediction have been encouraging, but their models need more refinement. These models describe static curvature. More recently, attempts have been made to incorporate dynamic elements into the algorithms. Using data quantifying the ease with which

different dinucleotide steps can bend upon interaction with DNase I, Bruckner et al. (1995) suggested that their model fitted some motifs, like TA, rather better. Much more work will be required before structural models can be reliable predictors of nucleosome position.

1.4 MECHANISMS FOR ACTIVATING GENES

Changes in chromatin structure play a decisive role in regulating gene expression (reviewed in Wolffe, 1994). For gene activation to occur, there must be unfolding of the chromatin fibre and rearrangement or loss of nucleosomes on regulatory elements so that the transcription complex can form. It has been proposed that there are three main mechanisms for altering the chromatin structure. Firstly, the chromatin structure can be 'reprogrammed' following replication. Secondly, some transcription factors may be capable of rearranging or displacing the nucleosome, so they can interact with their sites. Thirdly, 'remodelling complexes' alter chromatin structure by an ATP-dependent mechanism to allow transcription factor access.

1.4.1 Reprogramming during replication

In this model, transcription factors compete with nucleosomes for binding sites on the newly replicated DNA. Evidence supporting this comes from *in vitro* studies where addition of transcription factors before, or for some systems at the same time as chromatin assembly extracts allows the formation of a transcription complex (reviewed in Felsenfeld, 1996). For example, if TFIID and histones are present simultaneously in the assembly extract, a 'potentiated' chromatin template is produced which is capable of transcription when heat shock transcription factor is added. If the chicken β-globin gene locus is assembled into synthetic nuclei using a *Xenopus* extract, the chromatin structure can be 'reprogrammed' into an active state by replication in the presence of erythroid

transcription factors (Barton and Emerson, 1994).

1.4.2 Interactions between transcription factors and nucleosomes

Some transcription factors are capable of binding their recognition sites when they are positioned within a nucleosome. This may require or may lead to rearrangement or displacement of the nucleosome. For example, The PHO5 gene in yeast has four positioned nucleosomes in its promoter; one of these, nucleosome 2, covers the binding site for transcription factors PHO2 and PHO4. One PHO4 binding site occurs in the linker DNA next to nucleosome 2. It is thought that PHO4 transcription factor binds the site in the linker DNA, and co-operates with PHO2 to displace nucleosome 2 and expose the other PHO4 binding site (Almer et al., 1986).

In some cases, transcription factors may bind their sites in nucleosomal DNA without leading to the displacement of the nucleosome; in fact some proteins seem to require their sites to be 'presented' on the surfacce of a nucleosome. For example, activation of the serum albumin gene is accompanied by the appearance of positioned nucleosomes over transcription factor binding sites on its enhancer (McPherson et al., 1993).

There is a delicate energetic balance which decides whether a transcription factor binds its site in nucleosomal DNA; it depends on the levels of transcription factor and the rotational and translational position of the recognition site within the nucleosome (reviewed in Wolffe, 1994).

In summary, certain promoters can be activated without the need for reprogramming by replication. Nucleosomes do not play purely occlusive roles, and the interactions between nucleosome and a transcription factor are dynamic and are influenced by several factors, including the levels of transcription factor present and the precise placement of the transcription factor recognition site within the nucleosome.

1.4.3 Chromatin remodelling factors

Some transcription factors cannot access their sites when they are bound by nucleosomes. It has been shown that some genes have their chromatin structure remodelled by protein complexes to enhance transcription factor access. The SWI-SNF complex of yeast causes nucleosome disruption and facilitates transcription factor binding *in vitro* (Cairns et al., 1994). The SWI-SNF complex has been shown to be important for the transcription of a subset of yeast genes (Peterson and Heskowitz, 1992).

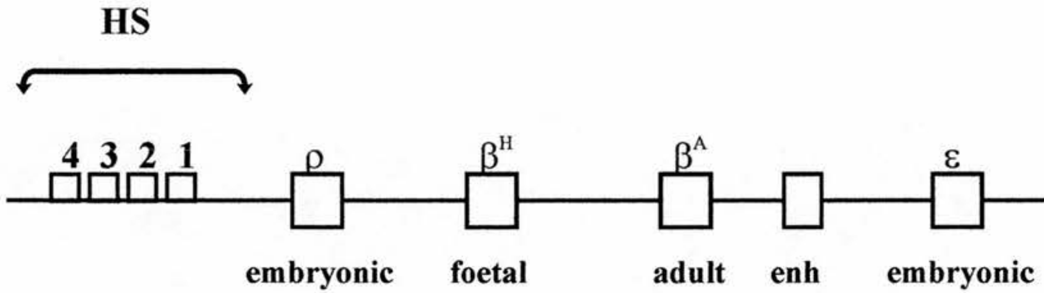
In Drosophila, the NURF protein complex and the GAGA factor cause disruption of the chromatin structure on the hsp 70 gene (Tsukiyama and Wu, 1995).

1.5 REGULATION OF THE CHICKEN β-GLOBIN GENES

1.5.1 Introduction

In vertebrates, haemoglobin, the oxygen-carrying molecule in the blood, is a tetramer of two α-globin and two β-globin polypeptide chains. Globin gene expression is tissue-specific, being restricted to erythroid cells. At the early embryonic stage in vertebrates, globin transcription first occurs in the yolk sac, then shifts to the foetal liver and finally to the bone marrow in adults. There are a variety of α- and β-globin genes, which are expressed at different times in development. In mammals and birds, the α-globin and β-globin genes occur in two distinct clusters on different chromosomes, so expression of the α- and β-globins must be carefully co-ordinated (Muller et al., 1988). Therefore, globin gene regulation is very complex; expression of the α- β-globin clusters must be co-ordinated, the individual globin gene clusters must be regulated in a

A



B

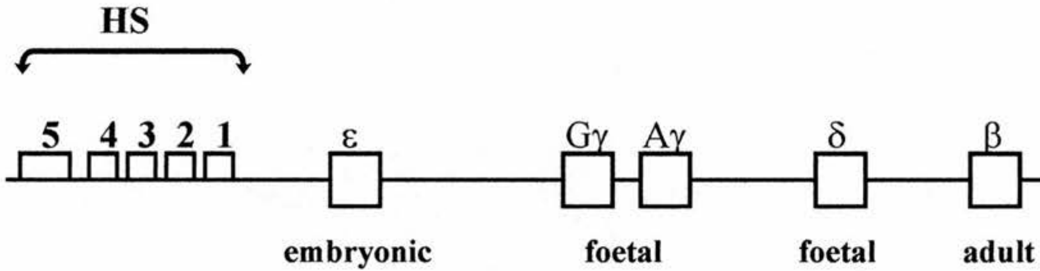


Figure 1.7 Organisation of the chicken and human β-globin gene clusters

This schematic depicts the individual globin genes and the important regulatory elements of the chicken (A) and human (B) β-globin gene clusters.The chicken cluster contains 4 genes: the ρ, β^A, β^H and ε globin genes. The enhancer (enh) between the and globin genes has LCR-like activity. There are 4 upstream hypersensitive sites (HS), and HS4 has insulator activity. In the human cluster, unlike the chicken cluster, the globin genes ε, Gγ, Aγ, δ and β occur in the order in which they are expressed. The five upstream hypersensitive sites (HS) have LCR activity, and HS5 is an insulator.

tissue-specific manner and the individual genes must be expressed only at the correct developmental stage. For these reasons, globin gene regulation has been intensively studied, with the aim of further understanding general mechanisms controlling temporal and tissue-specific eukaryotic gene expression.

The regulation of the globin genes has been investigated in mouse, human and chicken (the subject of this study). Certain parallels can be drawn between the different species; there are often strong homologies in the DNA elements and transcription factors involved (reviewed in Orkin, 1995). When the chicken β-globin gene cluster was introduced into mice, correct tissue-specific and temporal expression was seen (Mason et al., 1995). The chicken and human clusters evolved independently from a single β-globin gene in the common ancestor, so different gene regulatory mechanisms may have evolved (Czelusinak et al., 1982). Therefore, care must be taken in evaluating data from transgenic experiments, and in extrapolating ideas from one species to another.

1.5.2 Tissue-specific expression of the globin genes

The globin genes are expressed in erythroid tissues, and are silent in other tissues such as brain and oviduct. In erythroid cells, the chicken β-globin gene cluster (Figure 1.7) is contained within a ~30 kb domain that exhibits many of the characteristics associated with poised and active chromatin: hyperacetylation (Hebbes et al., 1994), hypomethylation, increased binding of HMG proteins (Postnikov et al., 1991), nuclease sensitivity (Stalder et al., 1980) and regions of nuclease hypersensitivity (reviewed in Wolffe, 1992). Some of these modifications, such as nuclease sensitivity, increased HMG binding and hyperacetylation extend across the entire globin domain. It was shown that the ~ 30 kb domain encompassing the entire cluster exhibited enhanced sensitivity to DNase I (Stalder et al., 1980; Hebbes et al., 1994). As discussed in section 1.3.1, the boundaries of the nuclease sensitivity extend far beyond the transcribed region and includes active genes, poised genes and non-transcribed flanking regions.

The possible role of acetylation, HMG proteins and linker histone in creating or maintaining an altered chromatin structure on poised or active genes has been discussed in previous sections (1.3.3 to 1.3.5). These modifications, which precede activation, may initiate or maintain a 'loosening' of the higher order structure, as indicated by the increased nuclease sensitivity. The poised state probably involves a 'loosening' rather than a complete unfolding of higher order chromatin structure, as it has been proposed that gene transcription takes place in the context of a 30 nm fibre, with only transient unfolding during transcription itself (Andersson et al., 1982; Ericsson et al., 1990). This loosening of the higher order structures increases the accessibility of the DNA to protein factors involved in transcription.

It has been shown that the nucleosome spacing is ~ 40 bp shorter across the chicken chicken β-globin cluster in active tissues (Villeponteau et al., 1992). It has been suggested that the change in nucleosome spacing on the cluster during development may create distinctive higher order chromatin structures, which could contribute to developmental regulation of the globin genes (Evans et al., 1990). The potential role played by nucleosome spacing in determining chromatin structure is discussed in greater detail in section 1.3.2.

Another modification in chromatin structure, DNase hypersensitivity, indicating a loss or rearrangement of nucleosomes, is localised to important regulatory elements in the DNA. One set of 4 four DNase hypersensitive sites, found upstream of the human β-globin gene cluster, and one downstream site, constitutes an LCR (locus control region). In transgenic systems, the LCR acts a powerful tissue specific enhancer, but perhaps most importantly it confers position-independent, copy number-dependent expression on linked genes (Grosveld et al., 1987). It has been suggested that in erythroid cells the LCR element defines the borders of the human β-globin gene cluster within a presumed active chromosomal loop domain. The LCR may be responsible for the creation of the 'open' chromatin structure across the whole β-globin gene cluster prior to transcription, and its maintenance afterwards (Tuan et al., 1985). Clustering of related genes such as

the human β-globins within a single domain allows domain-wide changes in chromatin structure to ensure that the genes are regulated in a co-ordinated, tissue-specific manner.

The chicken β-globin genes also occur within a cluster, and a ~ 400 bp enhancer, lying 3′ to the β^A-globin gene has been shown to have LCR-like activity, using copy number-dependent expression in transgenic mice as the assay for LCR activity (Reitman et al., 1990). The chicken enhancer is unable to open chromatin on its own (another definition for LCR activity). Interaction between the enhancer and other regulatory elements, like promoters, is required to open up chromatin in the domain containing the globin cluster (Reitman et al., 1993). It is argued that the upstream hypersensitive sites in chicken (discussed below) also contribute to LCR activity, and that neither the enhancer nor the upstream sites by themselves have 'true' LCR activity (Mason et al., 1995).

The enhancer is thought to confer tissue-specific expression, and to interact with the individual promoters to regulate developmental expression of the individual genes. Extensive studies have identified sequence elements in the enhancer, and the transcription factors which bind there (reviewed in Orkin et al., 1995). The GATA sites, bound by members of the GATA family, are thought to be particularly important in activating the globin cluster in erythroid cells (Fong and Emerson, 1992; Reitman and Felsenfeld, 1988; Orkin et al., 1995).

There are a set of four hypersensitive sites upstream of the chicken β-globin gene cluster (Figure 1.7). One of these, hypersensitive site 4 (HS4), acts as an insulator (Chung et al., 1993). An insulator is defined experimentally as a boundary element which does not have enhancer activity. The insulator may contribute to the formation and/ or maintenance of the putative active chromosomal loop domain in which the β-globin gene cluster is located. Three mechanisms of action have been proposed for insulators; they may act as a boundary to stop the spreading of heterochromatin; they may act as an anchor for the base of the proposed loop at the edge of the domain, and finally they may stop complexes tracking from nearby promoters and enhancers (Felsenfeld, 1996). The

other upstream hypersensitive sites have enhancer activity and are involved in regulating temporal expression of the globin genes. For example, enhancer activity of the 5 ′ HS 2 is mediated by its GATA sites (Abruzzo and Reitman, 1994).

1.5.3 Developmental regulation of the individual globin genes

The expression of individual globin genes is restricted to specific times during development. For expression of individual globin genes, which are all located within the 'open' chromatin domain in erythroid cells, further changes at regulatory elements are required. These alterations may include the binding of different transcription factors at enhancer elements, and the appearance of nuclease hypersensitive sites over individual promoters, accompanied by transcription factor binding (reviewed in Orkin et al., 1995). In chicken, the ρ and ε-globin (embryonic) genes are expressed in primitive erythroid cells up to and including day 5. From then onwards, the expression of ρ and ε-globin decreases, and the expression of β^H and β^A-globin (foetal and adult) genes starts, as definitive erythroid cells replace primitive erythroid cells. By about 12 days, the definitive lineage has completely replaced the primitive lineage, and ρ and ε-globin expression is undetectable (reviewed in Felsenfeld, 1993).

Some transcription factor binding sites are common to different globin promoters, and are involved in conferring erythroid specificity on the globin genes. Small differences in these common sites, and regions which differ between different globins (known as stage selector elements or SSE's) are responsible for the stage-specific expression of the globins. Transcription factor binding sites which are shared by the promoters of the chicken β-globin genes include a TATA box at ~ -30, a CCAAT element at ~ -70 to -90, and a CCACCC motif at a variable position from ~ 95 to -120 (relative the transcription start site). The TATA box binds TFIID. In the β^A, ε, and ρ-globins the TATA box is modified so that it can potentially bind members of the GATA family. It has been suggested that interaction of GATA-1 with the specialised TATA

box confers erythroid specificity on formation of the transcription complex at globin promoters (Fong and Emerson, 1992). Sp1-like proteins are known to bind the CCACCC sequence.

It is thought that the chicken ρ-globin gene is autonomously regulated. Activation of the gene is thought to occur primarily by interaction between the upstream hypersensitive sites in the cluster and the ρ-globin promoter. It has been suggested that decreases in the levels of Sp1 and GATA-1 in definitive cells contributes to the repression of ρ (Minie et al., 1992).

The silencing of the ρ-globin gene following expression in primitive cells is not fully understood. It has been proposed that silencing elements in 5′ and 3′ flanking regions and in the second intron of the ρ-globin gene, and the binding of YY1 protein to these silencing elements, contributes to gene silencing (Wandersee et al., 1996). Other studies suggest that methylation plays a role in regulating the stage-specific expression of the individual globin genes. In adult definitive cells, where ρ is inactive, all the CpG's in the promoter are methylated. In 5 day primitive cells, where ρ-globin is active all the CpG's are unmethylated (Singal et al., 1997).

There is strong evidence to suggest that the temporal control of the β^A and ε-globin genes is effected by a competition between these genes for interaction with their shared enhancer (Choi et al., 1988; Foley and Engel, 1992). Some studies, where the chicken globin gene cluster was introduced into mice, argue against the competition model (Mason et al., 1995). However, other researchers suggest that transgenic results must be considered carefully, because of the separate evolution of the clusters in mice and chicken.

Epsilon globin competes more effectively for the shared enhancer in 5 day (primitive) red blood cells, so that it is expressed whilst β^A-globin is silent; at about 8 days in development, a shift in the relative stabilities of the promoter-enhancer complexes occurs, and epsilon expression decreases whilst β^A-globin is upregulated, until in 12 day (definitive) red blood cells, epsilon is fully silenced. This mutually

exclusive expression is facilitated by the clustering of the genes.

The competition model for globin gene switching described above suggests transcription factors that bind the promoter and interact with proteins bound at the enhancer will decide the relative stability of the ε promoter-enhancer interaction and the β^A promoter-enhancer interaction, and thereby regulate temporal expression of these genes. Therefore, positioned nucleosomes which regulate the access of these transcription factors to the DNA could potentially play a role in globin gene switching. Researchers have suggested that the binding of the stage-specific factors NFE-4 and βCTF to the β^A promoter, and their interaction with proteins at the enhancer in definitive red-blood cells is the decisive factor in β^A/ε switching (Foley and Engel, 1992). The binding of GATA (which is not stage-specific, though its levels change during development) at the enhancer and β^A promoter may also contribute to the promoter-enhancer interaction. It was proposed that by binding to itself, or other transcription factors, it could add to the stability of a loop formed between the promoter and enhancer. Foley and Engel (1992) suggested that chromatin structure, which was not considered in their study, could also contribute to switching. In agreement with this, it was proposed that the precisely positioned nucleosome on the β^A promoter which occludes most of the important transcription factor binding sites prior to the β^A/ε switch, could be displaced by the binding of NFE-4 (Buckle et al., 1991).

A decrease in the stability of the ε promoter-enhancer interaction may also contribute in the switch from ε to β^A-globin expression; enhancement of ε-globin expression in a construct containing the enhancer and ε-globin only is 10-fold less in definitive cells than primitive cells (Mason et al., 1996). No stage-specific factors have been identified on the ε-globin promoter, but changes in the levels of non stage-specific factors may influence switching, as GATA 1 and Sp1 levels are higher in primitive than in definitive cells (Minie et al., 1992) whilst GATA 3 is higher in definitive cells (Leonard et al., 1993). It has been proposed that the presence of GATA 1 at the enhancer increases the stability of its interaction with the ε-globin promoter, whilst

GATA 3 at the enhancer stabilises its interaction with the β^A-globin promoter. The decrease in GATA 1 levels could also affect the occupancy of the GATA site on the ε-globin promoter. Therefore, there are transcription factor binding sites on the ε-globin promoter which could be important to the stability of the ε promoter-enhancer interaction; if a nucleosome were positioned on the ε-globin promoter it might have the potential to influence β^A/ε switching. The majority of experiments on switching have not been carried out in a chromatin context because transient transfections have been used (for example Foley and Engel, 1992). As such, a thorough investigation of the chromatin architecture of the ε-globin promoter may deepen our understanding of the switching mechanism.

In summary, the activation of the chicken globin genes in erythroid cells involves the formation of a domain around the whole cluster with a more 'open' chromatin structure, as indicated by the increased nuclease sensitivity of the cluster in erythroid tissues. The β^A/ε enhancer and the upstream hypersensitive sites are thought to be important in creating this open chromatin domain. Chromatin structure thus plays a decisive role in regulating the tissue-specific expression of the globin genes. The stage-specific regulation of the genes is thought to be regulated by the interaction of enhancer elements with individual globin promoters. The upstream hypersensitive sites are thought to be more important in the regulation of ρ, and the β^A/ε enhancer is thought to be more important in the temporal expression of the β^A and ε genes. Competition for the enhancer, as influenced by the relative stabilities of the promoter-enhancer complexes is thought to play the decisive role in controlling the switch in expression from ε to β^A. The promoter-enhancer stability, and as a result the switch, is influenced by the appearance of stage-specific transcription factors such as NFE-4, but it is also thought that subtle alterations in the levels of non-stage specific factors may play a role. There is a positioned nucleosome on the β^A-globin promoter which has the potential to influence the switch. The chromatin architecture of the ε-globin promoter before and after the switch is unknown, but if a nucleosome were positioned there it would have the

potential to modulate the stability of the ε promoter:enhancer interaction. Whilst much is known about the contribution of chromatin structure to tissue specificity, most of the work done on developmental regulation of the individual genes has been carried out in a non-chromatin context. It is interesting, therefore, that expression of the entire chicken β-globin cluster in mice (Mason et al., 1995) did not support the competition model for β^{A}/ε switching. This may be due to differences in regulation between mice and chicken, but it could also suggest that the switch is influenced by the chromatin structure of the globin genes.

1.6 THESIS PERSPECTIVE

The work presented in the following chapters investigates nucleosome positioning on the chicken β-globin genes *in vitro* and *in vivo*.

Sequence-dependent nucleosome positioning on the ε-globin gene was determined *in vitro*. The *in vitro* nucleosome positioning data obtained for this gene could be compared with those already obtained for the β^{A}-globin gene (Davey et al., 1995). A comparison of the two sets of data was a primary aim of this study; because the two genes share some features, but also exhibit some important differences. For example, the genes are adjacent to each other, and occur within the same domain; they are evolutionarily related; and they are both expressed in erythroid cells, but at different times in development. In addition, the two sets of data are contiguous, allowing a single map to be produced. This allows long-range features in nucleosome positioning to be studied.

A large database of nucleosome positioning sequences can be compiled from the two sets of data. This is far superior to others in the literature (Ioshikhes and Trifonov, 1993) because it contains more positioning sequences, mapped at better resolution, and with a weighting which denotes their relative affinities.

The techniques used for mapping nucleosome positioning *in vitro* were monomer

DNA extension (Davey et al., 1995) and oligonucleotide-based monomer DNA extension (this study). These methods can map nucleosome placement at high resolution, on long stretches of DNA, and they provide quantitative data on the relative strength of nucleosome positioning sequences.

The other important aim of the study was to investigate nucleosome placement *in vivo* on the β^A and ε-globin gene regions mapped *in vitro*. A comparison of the *in vivo* and *in vitro* nucleosome positioning data would reveal the extent to which DNA sequence directed nucleosome positioning *in vivo*. Two tissues, brain and adult red blood cells were studied, to compare how the sequence-encoded positional information was recognised and interpreted in different tissues. There is the potential for nucleosome positioning to contribute to the tissue-specific and temporal regulation of the globin genes, and it was hoped that the data produced in this study could deepen our understanding of its role in regulating globin gene expression.

CHAPTER 2: MATERIALS AND METHODS

This chapter contains details of the general methods used throughout this work, and most are adapted from Sambrook et al., (1989) unless otherwise stated. Chemicals were obtained from BDH unless otherwise stated. Methods specific to particular sections are described later in the relevant sections.

2.1 Materials and stock solutions

Acrylamide: A 20% (w/v) stock solution in 8M urea/1 X TBE was prepared from a 40% 19:1 acrylamide: N,N'-methybisacrylamide solution (Appligene), and diluted to 6% polyacrylamide in 8 M urea and 1 x TBE for denaturing DNA sequencing gel electrophoresis.

Agarose: Flogen (FMC) or Seakem GTG agarose was used for separation of DNA.

Agarose Gel Loading buffer (5x): This contains 0.25% bromophenol blue and 40% sucrose in 5x TBE gel buffer.

Ammonium acetate: A 0.3 M solution, pH 5.5 was used for neutralisation of DNA in Southern blotting.

Ammonium persulphate: 25% (w/v) fresh solutions of ammonium persulphate were made in water and used for catalysing polymerisation of polyacrylamide gels.

Antibiotics: Solutions of ampicillin (Sigma) (100 mg/ml, in distilled water) and kanamycin (Sigma) (50 mg/ml in distilled water) were sterilised by filtration and stored at -20°C.

BSA: For restriction enzyme digestions and monomer extensions, stock solutions at 10mg/ml were obtained from New England Biolabs. For southern blots, a 10% solution was prepared by dissolving 5g BSA (Sigma) in 50 ml water.

Dextran Sulphate: 9% (w/v) dextran sulphate was used in filter hybridisation as a molecular crowding agent.

Dialysis Tubing: Tubing was boiled in a solution containing 2% sodium hydrogen carbonate and 1 mM EDTA. It was washed three times in distilled water and autoclaved in 1mM EDTA, before storage at 4°C.

DNA Polymerase I large (Klenow) fragment: This was obtained from NBL for use in monomer extension experiments.

DTT: Dithiothreitol was dissolved at a concentration of 1M in 10 mM sodium acetate (pH5.5), sterilised by filtration and stored at -20°C.

EDTA: A 0.5 M solution was made by dissolving disodium ethylene diamine tetra-acetate in double-distilled water and adjusting the solution to pH 8.0 with sodium hydroxide.

Ethidium Bromide (EtBr): A 10 mg/ml solution was prepared by dissolving 1g of EtBr (BCL) in 100 ml distilled water. The solution was stored in a light-proof bottle at room temperature.

Heparin: Heparin (Sigma) was dissolved in double-distilled water at a concentration of 25 mg/ml and stored at -20°C.

5′ Kinase Buffer (10x): This contains 0.5 M Tris-Cl pH 7.5, 20 mM $MgCl_2$, 50 mM DTT, 1 mM EDTA and 1 mM spermidine in double-distilled water. It was filter-sterilised and stored at -20°C.

LB: 10 g/l of bactotryptone (Difco), 5g/l of yeast extract (Difco) and 10g/l of NaCl was dissolved in double-distilled water and autoclaved.

LB agar: 15 g/l of agar was added to LB medium and autoclaved as above. The appropriate antibiotic was added immediately before pouring the plates.

Micrococcal nuclease (MNase): from Worthington was used for preparation of core particles and digestion of nuclei.

dNTPs: Deoxyribonucleotide triphosphates (dATP, dCTP, dTTP and dGTP) (Pharmacia Biotech) were diluted to 10 mM each in distilled water and stored at -20°C.

Buffered phenol (pH 8.0): 250 g of solid phenol (Fluka 77610) was dissolved overnight at room temperature in 127 ml of 2 M Tris-Cl (pH 7.5). Following mixing, the phenol was centrifuged at 2000 x g for 1 minute, and the aqueous bottom layer was discarded. Then, 55 ml of 2 M Tris (pH 8.0), 13.75 ml of m-cresol and 550 μl of β-mercaptoethanol and 275 mg of 8-hydroxyquinone was added. Following mixing the mixture was divided into aliquots which were centrifuged to separate the buffer and phenol layers prior to storage at 4°C in the dark.

PMSF: Phenylmethyl sulphonyl fluoride (Sigma) was dissolved in isopropanol at a concentration of 250mM and stored at -20°C. Prior to use, it was warmed to 37°C to redissolve PMSF.

Proteinase K: Proteinase K, obtained from NBL at 50 mg/ml, was stored at -20 °C.

RNase A: Bovine pancreatic ribonuclease A (BCL) was dissolved at 10 mg/ml in 0.01 M sodium acetate pH 5.2 and boiled for 15 minutes to destroy any contaminating DNase I activity. After the RNase had cooled slowly to room temperature 0.1 volumes of 1 M Tris-Cl (pH 7.4) were added.

Salmon sperm DNA: Salmon sperm DNA (Sigma) was dissolved at 10 mg/ml in TE on a roller overnight at 4°C. Following sonication to shear the DNA, the solution was stored at - 20°C.

Sephadex G-50 or G-25: Sephadex G-50 (coarse) (Pharmacia) or G-25 (fine) (Sigma) was added to TE and autoclaved. Then 0.05% sodium azide was added.

SDS (sodium dodecyl sulphate): A 10% solution was prepared by dissolving 100 g/l SDS in double-distilled water.

Sodium acetate: A 3 M solution was prepared for DNA precipitation by dissolving 408.1 g/l sodium acetate in distilled water, and adjusting the pH to 5.2 with glacial acetic acid.

SSC (20x): the solution used for southern transfer of DNA was prepared as a 20 x stock by dissolving 175.3g/l NaCl, 88.2g/l sodium citrate in 800 ml of distilled water. The pH was adjusted to 7.0 with NaOH.

SSM (Sequencing Stop Mix): SSM contains 90 % formamide, 1x TBE, (an extra) 9 mM EDTA, 0.04 % bromophenol blue and 0.04 % xylene cyanol.

TBE (20X): TBE gel running buffer (pH 7.5) was prepared as a 20x stock by dissolving 121 g/l Tris base, 62.5 g/l boric acid and 7.4g /l EDTA in distilled water.

TPE (10 X): TPE gel running buffer was made as a 10 x stock, containing 0.3 M Tris-base, 0.36 M Sodium dihydrogen phosphate and 1 mM EDTA. If necessary, the pH was adjusted to 7.5.

TBG: To prepare TBG medium 2 g/l bactotryptone (Difco), 24 g/l yeast extract (Difco), 4 g/l glycerol, 17 mM KH_2PO_4, 55 mM K_2HPO_4 was dissolved in water and autoclaved. Before use glucose was added to a final concentration of 20 mM.

SOB: 20 g/l bactotryptone, 5 g/l yeast extract, 2.5 mM KCl, 10 mM NaCl, 10 mM $MgCl_2$ and 10 mM $MgSO_4$ was dissolved in distilled water. The solution was sterilised by autoclaving.

SOB:Glycerol: 60 ml SOB was mixed with 40 ml of glycerol and autoclaved.

TE (50 X): For a 50 X stock, a solution containing 500 mM Tris-Cl pH 8.0 and 5 mM EDTA pH 8.0 was prepared in distilled water. The solution was autoclaved.

TEP: **TEP_{16mM}** 1 X TE was supplemented with 16 mM NaCl; **$TEP_{0.4M}$** 1 X TE was supplemented with 0.4 M NaCl and **TEP_{2M}** 1 X TE was supplemented with 2 M NaCl

Tri Reagent: Tri reagent (Sigma), a mixture of guanidine thiocyanate and phenol in a mono-phase solution, was used in the isolation of single-stranded and genomic DNA.

Tris-Cl: A 1 M solution was prepared by dissolving 121.1 g/l of Tris base in distilled water and adjusting the pH to the required value with concentrated hydrochloric acid.

2.2 Protocols

2.2.1 Manipulation of nucleic acids: cloning

Cloning was carried out using the protocols described in 'Molecular Cloning, a laboratory manual' by Sambrook et al., (1989), or in Ausubel et al., (1997). 'Current Protocols in Molecular Biology', volumes 1-3. Recombinant DNA was prepared using the following techniques: digestion of DNA with restriction enzymes, purification of fragments from agarose gels, dephosphorylation of DNA, end-filling DNA, ligation of DNA and transformation of recombinant plasmids into competent cells.

2.2.2 Determination of DNA Concentration

2.2.2.1 By spectrophotometry

DNA concentration was determined by spectrophotometry, assuming that an OD of 1 at 260 nm is equivalent to 50 μg/ml of double-stranded DNA.

2.2.2.2 By comparison with a set of DNA standards

DNA concentration was also determined by comparison with a set of DNA standards of known concentration dotted onto Whatman DE81 paper. A dilution series of the DNA was made and 2 μl dotted onto DE81 paper. A set of DNA samples of known concentrations(2μl of each) were dotted next to these. The filter was stained with ethidium bromide by shaking in 1x TBE containing 5 μg/ml EtBr for 20 minutes. The DE81 was destained 5 washes, 10 minutes each, in 1x TBE. The filter was then exposed to UV light and the concentration of the DNA estimated by comparing the intensity of

fluorescence of the sample with the DNA standards.

2.2.3 Digestion of DNA with Restriction Enzymes

DNA was digested with restriction enzymes for the preparation of markers for gels, for preparing recombinant DNA and for diagnostic digests on recombinant DNA. DNA was digested using restriction enzymes from several manufacturers (New England Biolabs, Fermentas and Promega), according to their instructions. All digests included 0.1mg/ml BSA and 1 mM spermidine.

2.2.4 Purification of DNA from agarose gels

Restriction enzyme-treated DNA was separated by agarose gel electrophoresis and the desired DNA fragments were collected by electroelution. Briefly, the required band was cut out from the gel and transferred to a dialysis tubing bag and subjected to electrophoresis in 0.5 x TBE at 80V for 25 minutes. The DNA recovered from the dialysis bag was purified as described in section 2.9.

2.2.5 End-filling DNA

A 25 μl reaction containing 1x Klenow reaction buffer, 50 μM dNTP's, 0.1 to 4 μg of DNA and 1 to 4U of Klenow fragment polymerase was incubated at 37°C for 30 minutes. The reaction was stopped by heating to 75°C for 10 minutes. Following purification (section 2.9), the DNA was ethanol precipitated and resuspended in TE.

2.2.6 Ligation of DNA fragments

Ligation reactions of 10 to 100 μl containing 1 x ligation buffer (50 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP and 25 μg/ml BSA), 1 to 2U of T4 DNA ligase (NBL), and DNA at a concentration of approximately 5 ng/μl, were incubated overnight at 12°C. If the fragments were blunt-ended a higher DNA concentration (10 ng/μl), a lower ligation temperature (4°C) and 4 to 8U of T4 DNA ligase were used. If the ligation was intramolecular, a lower DNA concentration (2 ng/μl) was used.

2.2.7. Preparation of competent *E. coli* cells

E. coli strains JM101 and DH11S (Lin et al., 1992) were made competent for transformation using the protocol described by Nishimura et al. (1990). 50 ml of medium A (LB supplemented with 10 mM $MgSO_4$ and 0.2% glucose and filter sterilised) was inoculated with 0.5 ml of an overnight culture and incubated at 37°C in a shaking incubator until the culture reached mid-log phase. The cells were incubated on ice for 10 minutes and pelleted at 750 x g for 10 minutes at 4°C. The pelleted cells were resuspended gently in 0.5 ml of ice-cold medium A. Then 2.5 ml of storage solution B (LB supplemented with 36% glycerol, 2% PEG 7500 and 12 mM $MgSO_4$ and filter sterilised) was added. The competent cells were, for preference, used immediately or stored at -70°C in 0.1 ml aliquots.

2.2.8 Transformation of competent *E. coli* cells

Competent *E. coli* cells were transformed with plasmids according to Hanahan (1983). 30 to 50 ng of plasmid was mixed with 200 μl of competent cells and incubated

on ice for 20 minutes. The cells were then heat shocked by incubation at 37°C for 30 seconds, followed by incubation on ice for 5 minutes. The cells were supplemented with 300 μl of LB and incubated at 37°C for 60 minutes with occasional mixing. The cells were plated out onto LA agar plates containing the appropriate antibiotic and incubated overnight at 37°C.

2.2.9 Purification of DNA

An equal volume of buffered phenol/chloroform (1:1 v/v) was added and mixed by vortexing (or by inversion, if it was necessary to avoid shearing, as in the case of large DNA fragments) and left on ice for 5 minutes, with occasional mixing. The aqueous phase, containing the DNA, was separated by spinning at 12000 x g for 5 minutes in a centrifuge. The phenol/chloroform extraction was repeated until there was no material at the interface between the phenol and aqueous layers. Then 1 volume of chloroform/isoamyl alcohol (24:1 v/v) was added to the aqueous layer, mixed and left on ice for 5 minutes. The aqueous layer was separated from the chloroform by centrifugation as above. Following purification, DNAs were always alcohol precipitated.

2.2.10 Precipitation of DNA

Sodium acetate (pH 5.5) was added to the DNA, to a final concentration of 0.3 M, and then either an equal volume of propan-1-ol or 2.5 volumes of ethanol was added. If precipitating with ethanol, the sample was precipitated for 1 hour at -70°C or at -20°C overnight. If precipitating with propan-1-ol, the sample was precipitated or at -20°C overnight.

2.2.11 Small-scale preparation of double-stranded plasmid DNA (miniprep)

Single colonies, or 40 µl of a solution of bacteria (in SOB/glycerol, see section 3.2.3.3) containing the appropriate plasmid , were inoculated into 5 ml of LB medium containing appropriate antibiotics and shaken overnight at 37°C. The cells in 3 ml of culture were pelleted by spinning at 15000 x g for 30 seconds. The supernatant was removed and one of the following methods was used for preparation of plasmid DNA:-

2.2.11.1 Qiagen miniprep kit

The pellet was resuspended and treated according to manufacturer's instructions.

2.2.11.2 Alkaline lysis method

The pellet was resuspended in 200 µl of TEG (25 mM Tris-Cl pH 8.0, 10 mM EDTA and 50 mM glucose). 400 µl of a freshly made solution containing 0.2 M NaOH and 1% SDS was added, and mixed by inversion. 300 µl of ice-cold 3 M potassium 5 M acetate was added, mixed by inversion and incubated on ice for 5 minutes. The samples were centrifuged at 15000 x g in a bench top centrifuge for 5 minutes, and the supernatants transferred to fresh tubes and deproteinated by phenol/chloroform extraction (section 2.9). After propan-2-ol precipitation, the plasmid DNA was resuspended in TE supplemented with RNase to a final concentration of 100 µg/ml, and incubated for 30 minutes at 37°C. After adding 0.6 volumes of 20% PEG 6000 in 2.5 M NaCl, the DNA was incubated on ice for 60 minutes, pelleted by a spin at 13000 rpm, washed with 70% ethanol and resuspended in TE.

2.2.12 Large-scale preparation of double-stranded DNA

Large scale preparations were carried out, essentially as a scaled-up version of the small-scale method, as described in Sambrook et al., (1989).

2.2.13 Preparation of genomic DNA

Chicken genomic DNA was prepared from frozen chicken red blood cell nuclei using Tri-reagent (Sigma) according to the manufacturer's instructions.

2.2.14 Dephosphorylation of DNA

5 µg of DNA was ethanol precipitated (section 2.10) and resuspended in 50 µl of 1x calf intestinal alkaline phosphatase buffer. After adding 10 U of calf intestinal alkaline phosphatase (CIAP), the reaction mix was incubated at 37°C for 20 minutes. A further 10 U of CIAP was added and the reaction incubated for a further 20 minutes at 37°C (or at 50°C for 20 minutes if the DNA had 5′ recessed or blunt ends). The reaction was quenched by adding EDTA to a final concentration of 5 mM and incubating at 75°C for 10 minutes. The DNA was deproteinated by phenol/chloroform extraction, and then ethanol precipitated.

2.2.15 End-labelling of DNA with ^{32}P

200 to 500 ng of dephosphorylated DNA was labelled in a 10 µl reaction containing 1x 5' kinase buffer, 2.5 µl of γ^{32}PATP, at 3000 Ci/mmol (Amersham) and 10U of T4 polynucleotide kinase. The reaction was incubated at 37°C for 1 hour then boosted to

50 μl with TE and phenol/chloroform extracted. Unincorporated label was removed by spinning through a 1ml Sephadex G50 column equilibrated with distilled water or TE. The labelled DNA was ethanol precipitated and resupended in TE or double-distilled water to the required volume.

2.2.16 Sephadex G-25/G-50 column purification

Sephadex spin columns were employed to purify DNA from unincorporated radioactive label; G-50 was used for core particle (monomer) DNA, size markers and southern probes, and G-25 for oligonucleotides and primers. A 1 ml syringe was filled in Sephadex G-25 or G-50 and spun at 2000 x g for 1 minute to pack the column. Columns were then equilibriated with TE or H_2O. G-50 columns were washed by spinning through the wash volumes three times at 2000 x g for 1 minute. G-25 columns were washed by pipetting TE or H_2O onto the column and allowing it to drip through Liquid was excluded from the columns prior to loading the sample by spinning at 1.6 K for 6 minutes (5 minutes for G-25). The sample was loaded on the top of the column and the column was spun at 2000 x g for 5 minutes. The eluted sample was collected, and purified by phenol extraction (section 2.9), and ethanol precipitation (2.10).

2.2.17 Amplification of DNA fragments by polymerase chain reaction (PCR)

PCR is the *in vitro* enzymatic amplification of a specific DNA segment by repeated cycles of extension of two oligonucleotide primers defined as the upstream and downstream ends of the desired PCR product.

Linearised DNA (20 ng) was added to a 25 μl reaction containing 2.5μl of 10 x Taq buffer (0.5M KCl, 100 mM Tris-HCl (pH 7.5)), 2.5 μl of 2 mM dNTP's, 2.5 μl 5% Tween, 2.5 μl of each appropriate primer at 10 μM, 1 to 2 mM $MgCl_2$, Taq enzyme and

double-distilled water. The sample was placed in a thermal cycler (Hybaid) and heated to 95°C for 5 minutes, followed by 1 minute at the annealing temperature and 1 minute at the extension temperature (74°C). The sample was then heated at 95°C for 1 minute, followed by 1 minute at the annealing temperature and 1 minute at 74°C, and this was repeated 25 times.

The annealing temperature was calculated by using the following formula:

$$T_{ann} = 61.2 + 0.41(\%GC) - 500/L$$

Where %GC is the percentage of G+C bases in the oligonucleotide primer and L is the length in base pairs of the oligonucleotide. As two oligonucleotide primers are used in each PCR, the annealing temperature used is a compromise between the two T_{ann}.

PCR does not require purified DNA as a template; PCR can also be carried out directly on a bacterial colony. A colony is picked from the LB agar plate on which it was grown using a toothpick, and put in the PCR premix.

2.2.18 Sequencing double-stranded plasmid DNA using radioactively end-labelled primers

DNA sequencing of double-stranded plasmid was performed using the dideoxy chain-termination method (Sanger et al., 1977).

0.5 pmole of plasmid DNA was denatured in 0.2 M NaOH for 10 minutes at room temperature. The solution was neutralised by adding a 1/5 volume of 5 M ammonium acetate pH 7.5, and the DNA precipitated with 2.5 volumes of ethanol. After two washes with 70% ethanol, the dry pellet was dissolved in distilled water to a concentration 0.05 pmole/μl and used as soon as possible.

Sequencing reactions were performed in two stages. In the first step, primers end-labelled with ^{32}P were annealed to the template: in a 10 μl solution, containing 1.25 μl of

T7 promoter DNA primer (0.25 pmole), 5 μl of denatured DNA template (0.25 pmole), 2 μl of 5X sequencing reaction buffer (200 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, and 250 mM NaCl), and 1.75 μl H_2O. The samples were incubated at 65 °C for 4 minutes and then cooled down to 30 °C at a rate of 75 seconds/ °C. The next step was extension and termination: 1 μl of 0.1 M DTT, 4.7 μl of 10 mM Tris pH 7.5, and 0.3 μl of T7 DNA polymerase (at 8U/μl) (Amersham) were added to the annealing mix. The mix was split into four 3.5 μl aliquots and placed in 4 PCR tubes, each of which contained 2.5 μl of one of the four dNTP/ddNTP mixes: dNTP/ddGTP, dNTP/ddCTP, dNTP/ddATP, or dNTP/ddTTP (each mix contained 50 mM NaCl, 80 μM dNTPs and 8 μM of on of the four ddNTPs). These samples were incubated at 37 °C for 5 minutes and then for 15 minutes at 50 °C. The reaction was stopped by adding 1 volume of sequencing stop mix (SSM). The samples were heated to 95 °C for 3 minutes prior to loading onto a denaturing polyacrylamide sequencing gel.

2.2.19 Preparation of long-range sequencing ladders

Long C/T sequencing reactions were prepared as size markers for running on polyacrylamide gels with monomer extension reactions.

1 pmole of M13-20 primer was mixed with 1 pmole of M13mp18 ssDNA (USB) in 1 x sequencing buffer (40 mM Tris pH 7.5, 20 mM $MgCl_2$, and 50 mM NaCl). The mixture was incubated at 65°C for 4 minutes, ramped to 30°C at 75 sec/°C. After this annealing step, a labelling premix containing 4 μl of 10 mM Tris pH7.5, 2 μl of 0.1 M DTT, 2 μl of x 5 labelling dNTP mix for αdATP (7.5 μM dCTP, dGTP and dTTP), 3 μl of α^{35}SdATP (600 Ci/mmol), and 1 μl of T7 DNA polymerase (8 units/μl) was added to the annealing reaction and incubated at room temperature for 10 minutes.

The labelling mix was split into 2 new PCR tubes which contained 10 μl of C term

or T term (50 mM NaCl, 80 μM dNTPs, 8 μM ddC or ddT), respectively, and the extension reaction was carried out at 37°C for 5 minutes and 50°C for 15 minutes. The samples were extracted, ethanol precipitated, and the pellet resuspended in SSM.

2.2.20 Preparation of ^{32}P end-labelled DNA size markers

These were prepared as size markers for running on polyacrylamide gels with monomer extension reactions.

Phage lambda DNA was cleaved by HinFI or DdeI restriction enzyme, purified by phenol extraction and ethanol precipitation and then 5′ end-labelled with T4 polynuclease kinase as follows. 300 ng of HinFI or DdeI digested DNA were mixed with 1 X 5′ kinase buffer, 3 μl of γ^{32}P-ATP, and 10 units of T4 polynuclease kinase (Amersham), and incubated at 37°C for 60 minutes. The markers were phenol extracted (2.9), purified on G-50 spun columns and ethanol precipitated (2.10). Following a wash in 70% ethanol, the dry pellet was resuspended in SSM and stored at –20°C.

2.2.21 Denaturing polyacrylamide DNA sequencing gel

DNA sequencing reactions were analysed on 48 x 20 cm 6% polyacrylamide gels (19:1 acrylamide to N, N′-methylene bisacrylamide) containing 8 M urea and 1 X TBE. Immediately following the addition of the catalyst (0.04% fresh ammonium persulphate), and the polymerising agent (0.12% TEMED), the gel was mixed and poured between the assembled cleaned glass plates. The gel was pre-run for at least 40 minutes so that it reached 50°C before loading the samples. Sequencing reaction samples were heated to 95 °C for 5 minutes before loading and the gel was run in 1 X TBE buffer at a constant power setting of 50 W for a period of time depending on the size of DNA fragments. The required time was judged using the bromophenol blue and xylene cyanol dye fronts. At the end of the electrophoresis run, the gel was fixed in 10% glacial acetic acid and 12%

methanol for 10 minutes, washed with distilled water and transferred to 17 chromatography paper (Whatman) and dried at 80 °C under vacuum for 1 hour. The dry gel was exposed to an X-ray film (Fuji) at -70 °C with Kodak intensifying screens.

2.2.22 Agarose gel electrophoresis of DNA

DNA restriction digests, amplified DNA prepared by PCR, single-stranded DNA, and double-stranded DNA preparations were analysed by horizontal electrophoresis. Agarose was dissolved at the desired percentage by boiling in 1 X TBE containing 0.5 μg/ml of EtBr. The gel was set with an appropriate comb inserted and allowed to solidify. DNA samples dissolved in 1 X TBE loading buffer were loaded into the wells. The agarose gel was run in 1 X TBE containing 0.5 μg/ml EtBr at a constant voltage of 50 V until the desired resolution was achieved. DNA was detected on an UV transilluminator.

CHAPTER THREE: NUCLEOSOME POSITIONING ON THE CHICKEN EPSILON GLOBIN GENE *IN VITRO*

3.1 INTRODUCTION

Chromatin structure has an important role to play in controlling gene expression (reviewed in Wolffe et al., 1994; Grunstein, 1990). The globin genes are regulated in a tissue-specific manner, and the individual genes are expressed only at specific times during development. Changes in higher-order chromatin folding which encompass the entire globin domain, and alterations in local, gene-specific chromatin structures are thought to contribute to this temporal and tissue-specific regulation (Larsen and Weintraub, 1982; Buckle et al., 1991; Felsenfeld, 1993; Orkin et al., 1995). The position of individual nucleosomes on globin DNA sequences may be a major determinant of long and short-range chromatin structures.

In this investigation, nucleosome positions were to be mapped *in vitro*, where the influence of DNA sequence on nucleosome positioning can be investigated independently of the other factors which can act *in vivo*, over a 3.8 kb DNA region containing the chicken epsilon (ε) globin gene. This map, when juxtaposed with that for the beta-adult (β^A) globin gene region (Davey et al., 1995), would show the precise location and relative strength of all nucleosome positioning sites throughout 8 kb of continuous sequence. This would reveal patterns in the long-range positioning of nucleosomes, as well as the placement of individual nucleosomes over regulatory regions such as the ε-globin promoter. The long-range placement of nucleosomes may have implications for the folding of the chromatin fibre into a higher order structure (Davey et al., 1995), and the presence of a positioned nucleosome on DNA regulatory regions (promoters and enhancers) can affect the accessibility of transcription factors, thereby contributing to the regulation of gene expression (reviewed in Wolffe et al.,

1994). The *in vitro* nucleosome positioning data from the β^A-globin gene has revealed some interesting long-range and short-range patterns in nucleosome placement. A comparison of these data with that for ε-globin gene would be of interest because the two genes evolved from a common ancestral gene, occur within the same chromatin domain (Stalder et al., 1980; Reitman et al., 1993), and are both expressed in erythroid cells, but at different times during development (reviewed in Felsenfeld, 1993).

The monomer DNA extension technique (Yenidunya et al., 1994) was chosen for mapping nucleosome positions because it has several advantages over other commonly used methods. Firstly, it is quantitative, so that the relative strengths of different positioning sites can be compared. Secondly, it can map nucleosome positions with base pair resolution. Thirdly, nucleosomes can be mapped over a long distance (up to ~1 kb) in a single reaction, so the production of a long-range map is less labour-intensive. Finally, the region of DNA being mapped is contained within a much longer DNA fragment so the positions mapped are not influenced by *end effects* (non-specific interactions of core histones with the ends of (short) DNA fragments).

Both the original monomer DNA extension technique (Yenidunya et al., 1994), and an adaptation of it (developed as part of this project) have been used in the mapping experiments described in this study.

3.1.1 The monomer DNA extension technique

Two versions of the monomer DNA extension technique have been used for mapping: the original method (Yenidunya et al., 1994) and a new method. The original monomer DNA extension technique will hereafter be referred to as 'MDE' and the new method, oligonucleotide-based monomer DNA extension, will be hereafter referred to as OMDE.

3.1.1.1 The original monomer extension technique (MDE)

A schematic illustration of the original monomer extension technique is shown in Figure 3.1. The technique is described in detail in section 3.2.8.1. A 4.3 kb DNA fragment containing the entire chicken ε-globin gene region was cloned into a phagemid vector (pBluescript KS II minus), creating the clone PCBE4.3 (Figure 4.3). PCBE4.3 was used to prepare the mononucleosomes. Linearised double-stranded DNA from PCBE4.3 is reconstituted with core histones by salt-gradient dialysis, and mononucleosomes are prepared from the reconstitute by digestion with micrococcal nuclease. Monomer DNA (~146 bp in length) is purified from an agarose gel.

After the monomer DNA has been 5′ end-labelled with ^{32}P and alkali-denatured, it is annealed to ssDNA, made from subclones of PCBE4.3 (Figure 3.3), containing the particular region of ε-globin to be mapped. The monomer DNA is then extended on the ssDNA with the Klenow fragment of DNA polymerase in the presence of an appropriate restriction enzyme which cuts uniquely within the phagemid or insert. The enzyme cleaves the nascent double-stranded DNA formed by the extension of the monomer DNA to produce labelled extension fragments, which are subsequently sized on a denaturing polyacrylamide gel.

The length of any particular fragment corresponds to the distance from the 5′ end of an annealed monomer DNA to the restriction enzyme site, thus defining one boundary of a nucleosome. The other boundary of the same nucleosome can be determined by annealing and extending the same monomer DNA (the complementary strand) on the complementary ssDNA, using another restriction enzyme site as the reference point. The relative intensity of the fragments denotes the relative binding strengths of the different positioning sites.

As a routine control to ensure the extension fragments were restriction enzyme-dependent, monomer DNA is extended on the ssDNA in the absence of any restriction enzyme.

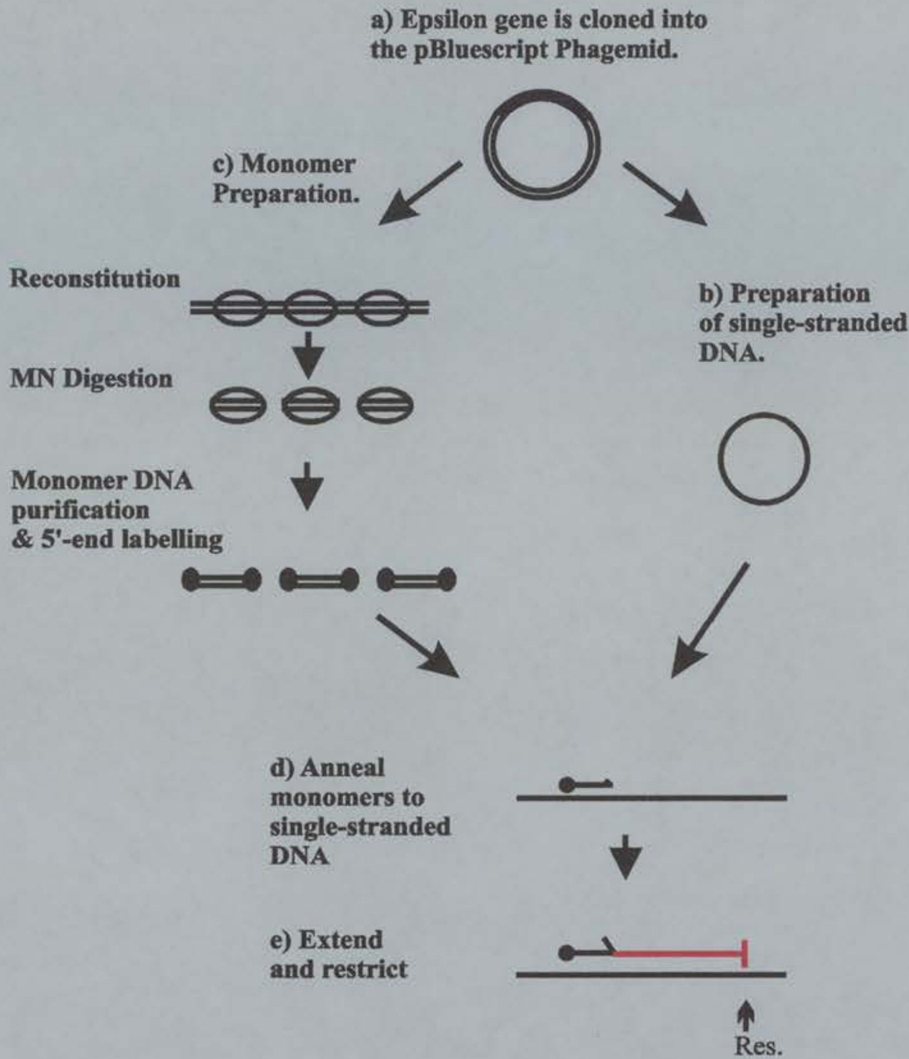


Figure 3.1 The original monomer DNA extension technique

a) The epsilon globin gene is cloned into the pBluescript phagemid.
b) Single-stranded DNA is prepared from this full-length phagemid construct, or from subclones of it.
c) Monomer DNA is prepared by micrococcal nuclease (MN) digestion of the histone-reconstituted phagemid construct. The monomer DNA is purified and end-labelled with ^{32}P.
d) The radiolabelled monomer DNA is annealed to the single-stranded DNA. (For simplicity the single-stranded DNA is shown as linear; it is, of course, circular).
e) The annealed monomer DNA is extended by DNA polymerase in the presence of a restriction enzyme which cuts uniquely within the phagemid construct. The resulting fragments are sized on a denaturing polyacrylamide gel.

3.1.1.2 A new monomer DNA extension technique: oligonucleotide-based monomer DNA extension (OMDE)

The original monomer extension technique requires a restriction enzyme that has a unique recognition site within the plasmid from which ssDNA is prepared. Therefore, mapping a long sequence may require a large number of subclones to ensure that this criterion can be met. For this reason, a new technique was developed to dispense with the requirement that a restriction enzyme should have only one recognition site within the mapping plasmid. The restriction enzyme-cutting is 'targeted' to one of several potential sites by employing an oligonucleotide designed to anneal specifically at the desired restriction site.

The new, oligonucleotide-based technique is depicted in Figure 3.2. It is described in detail in section 3.2.8.2. After preparing monomer DNA and single-stranded DNA as described for the original technique, a complementary oligonucleotide is annealed to the restriction enzyme site on the ssDNA and incubated with restriction enzyme. The restriction enzyme can only cut at the unique double-stranded DNA created by the annealed oligonucleotide; at the other potential restriction enzyme sites, the DNA remains single-stranded and the enzyme cannot cut. Monomer DNA is then annealed to the pre-cut ssDNA and extended, but this time in the absence of restriction enzyme; Klenow polymerase extends the monomer DNA to the end created by restriction enzyme cutting, where the extension terminates. As with the original method, the resulting fragments are subsequently sized on a denaturing polyacrylamide gel, and the size of each fragment gives the distance of a nucleosome boundary from the restriction enzyme site.

Two control reactions are also included in each analysis. In one, the ssDNA is incubated with restriction enzyme in the absence of oligonucleotide to check whether the enzyme is capable of cutting ssDNA. In the second, annealed oligonucleotide is incubated in the absence of restriction enzyme to verify that the bands seen in the monomer extension are restriction enzyme-dependent.

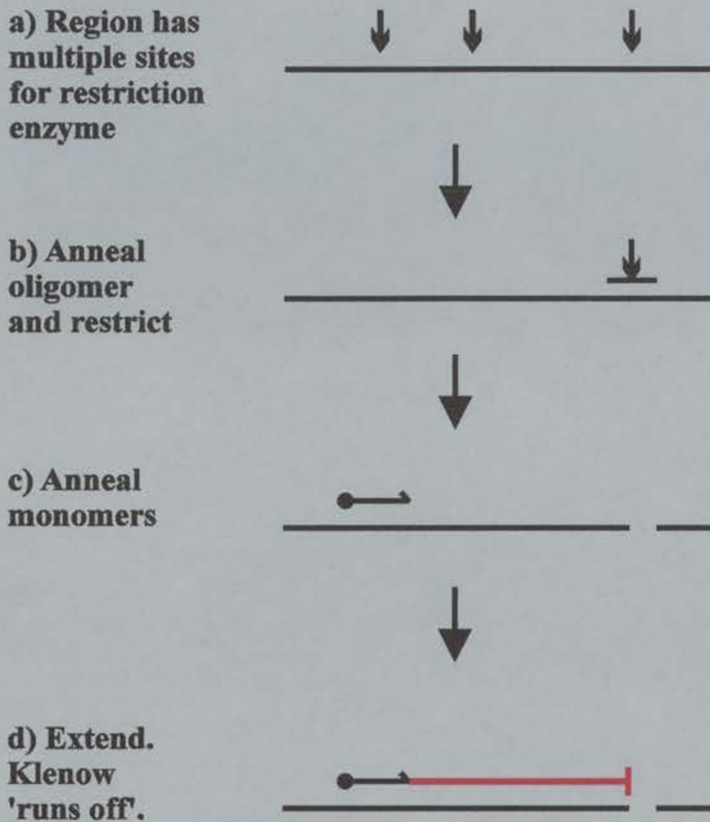


Figure 3.2 Schematic diagram of the oligonucleotide-based monomer DNA extension technique (OMDE)

The epsilon globin gene region to be mapped is cloned into the pBluescript plasmid, and single-stranded DNA and monomer DNA prepared as described for the original method (Figure 3.1).

a) The clone containing the region to be mapped may have several recognition sites for the restriction enzyme.

b) An oligonucleotide, designed to be complementary to only one of these sites, is annealed to create a uniquely double-stranded region which can be cut by the restriction enzyme.

c) Radiolabelled monomer DNA is annealed to this pre-cut single-stranded DNA.

d)The monomer DNA is then extended as described for the original method, except that restriction enzyme is not included in the reaction.

The DNA polymerase extends the annealed monomer DNA to the restricted end, and terminates at this defined point. The resulting fragments are sized on a denaturing polyacrylamide gel.

3.2 MATERIALS AND METHODS

3.2.1 Materials

M9 salt: made as a 5 X stock containing 0.24 M Na_2HPO_4 , 0.11 M KH_2PO_4, 0.43 M NaCl and 0.093 M NH_4Cl.

M9 minimal medium: (Sambrook et al., 1989) One litre of M9 minimal medium was made by dissolving 15 g of agar in 770 ml of ddH_2O and autoclaving it. After cooling the autoclaved agar to 55°C, 200 ml of 5 X M9 salt (at 55°C), 2 ml of 1 M $MgSO_4$, 20 ml of 20 % glucose, 100µl of 1 M $CaCl_2$, 1 ml of 1 mg/ml thiamine and the appropriate antibiotic was added. The plates were then poured and dried at 37°C for 2 hours before use.

TBG: (Lin et al., 1992) To prepare TBG medium for growing a ssDNA culture, 2 g/l bactotryptone (DIFCO), 24 g/l yeast extract (DIFCO), 4g/l glycerol, 17 mM KH_2PO_4 (BDH), 55 mM K_2HPO_4 (BDH) were dissolved in water and autoclaved. Before use glucose was added to a final concentration of 20 mM.

SOB: (Lin et al., 1992) This contains 20 g/l bactotryptone, 5 g/l yeast extract, 2.5 mM KCl, 10 mM NaCl, 10 mM $MgCl_2$ and 10 mM $MgSO_4$ dissolved in distilled water. The solution was sterilised by autoclaving.

SOB:Glycerol: 60 ml SOB was mixed with 40 ml of glycerol and autoclaved.

TEP_{YM}: YM NaCl in TEP. For example, $TEP_{0.4M}$ is 0.4 M NaCl in TEP (10 mM Tris, pH 8, 0.1 mM EDTA and 0.1 mM PMSF).

TYP medium: (per litre) 16 g bactotryptone, 16 g yeast extract, 5 g NaCl, 2.5 g K_2HPO_4.

3.2.2 Plasmid constructs

Twelve overlapping clones were generated to cover the entire epsilon globin gene (Figure 3.3). To create clones 1-11 described below, fragments of the chicken ε globin gene were cloned from λCβG1 (Dolan et al., 1981) into the polylinker of the pBluescript II KS minus phagemid vector (Stratagene). The clones were designed so that both the 5′ and 3′ nucleosome boundaries could be mapped. Clone 12 (PCBEoverlap) was created by cloning part of the globin insert from pCBG17 (Villeponteau et al., 1982) into pBluescript KS minus. The epsilon globin insert in this clone covers 2853 to 3833 relative to the β^A cap site, so it can be used to map nucleosome positions to be mapped upstream of the 4.3 kb region cloned into PCBE4.3. This allows the positioning map for the ε globin gene to be joined to that for the β^A globin gene (Davey et al., 1995). To map nucleosome placement on this region, monomer DNA was prepared from pCBA6.2, a clone covering part of the β^A and ε globin genes, from -1054 to 5114 relative to the β^A cap site. This monomer DNA was extended on ssDNA prepared from PCBEoverlap.

All numbers below indicate the position of the cloned globin fragments relative to the β^A gene cap site. The chicken β-globin gene cluster was sequenced by Reitman et al. (1993), gen bank ID gghbbre.

1. PCBE4.3: This BamHI-Hind III fragment, covering positions 3351 to 7635, was cloned into the BamHI and Hind III sites in the pBluescript polylinker.

2. PCBE2.4Bam: This BamHI-Bcl I fragment, covering 3351 to 5749, was cloned into the BamHI site of the pBluescript polylinker.

3. PCBE2.4Bcl: This contains the same fragment as the PCBE2.4Bam clone, but it is cloned into the BamHI site of the pBluescript polylinker in the opposite orientation.

4. PCBEacc: This AccI-Bcl I fragment, covering 4548 to 5749, was cloned into the BamHI and Acc I sites in the polylinker.

5. PCBEeag: This BamHI-Eag I fragment, covering 3351 to 4828, was cloned into the BamHI and Eag I sites in the pBluescript polylinker.

6. PCBEeco: This BamHI-EcoRI fragment, covering 3351 to 5114, was cloned into the BamHI and EcoRI sites in the polylinker.

7. PCBEsac: This Sac I-Bcl I fragment, covering 4655 to 5749 was cloned into the SacI and BamHI sites in the pBluescript polylinker.

PCBEacc and PCBEeag map 5′ nucleosome boundaries; PCBEeco and PCBEsac map the 3′ boundaries over the same region.

8. PCBE1.9EH: This EcoRI-EcoNI fragment, covering 5114 to 7005, was cloned into the EcoRI and EcoNI sites in the pBluescript polylinker.

9. PCBE1.9HE: This contains the same fragment as the PCBE1.9EH clone, but it is cloned in the opposite orientation into the pBluescript polylinker.

10. PCBE2.5EH: This EcoRI-Hind III fragment, covering 5114 to 7635, was cloned into the EcoRI-Hind III sites in the pBluescript polylinker.

11. PCBE2.5HE: This contains the same fragment as the PCBE2.5EH clone, but it is cloned in the opposite orientation into the pBluescript polylinker.

12. PCBEoverlap: This Sac I-EcoRV fragment, covering 2853 to 3833, was cloned into the Sac I and EcoRV sites in the polylinker of pBluescript KS minus.

All phagemids were transformed into *E. coli* strain DH11S (Lin et al., 1992) for the preparation of single-stranded DNA.

3.2.3 Preparation of single-stranded DNA

Single-stranded DNA was produced by superinfection of bacteria containing the pBluescript phagemid with M13 KO7 helper phage (Promega).

3.2.3.1 Preparation of helper phage

Dilutions of helper phage stock were made over the range of 1x 10^2pfu/ml to x 1 x 10^{12} pfu/ml. 0.1 ml of each phage dilution was added to 0.2 ml of log-phase DH11S (grown in TYP medium for 6 to 8 hrs). The bacteria-phage mixture was incubated at room temperature for 5 minutes. 50 μl of each bacteria-phage mixture was plated on TYP plates (TYP + 1.5% agar) containing 70 μg/ml kanamycin and incubated overnight at 37°C. Single, well-isolated colonies were picked with a small Gilson pipette tip into 5 ml TYP medium containing 70 μg/ml kanamycin and grown for 8 to 12 hrs at 37°C. Cells were pelleted by centrifugation at 5500 x g for 15 minutes. The supernatant, which contains the helper phage, was spun again at 5500 x g for a further 15 minutes to remove any remaining bacteria. The helper phage was sterilised by heating at 65°C for 20

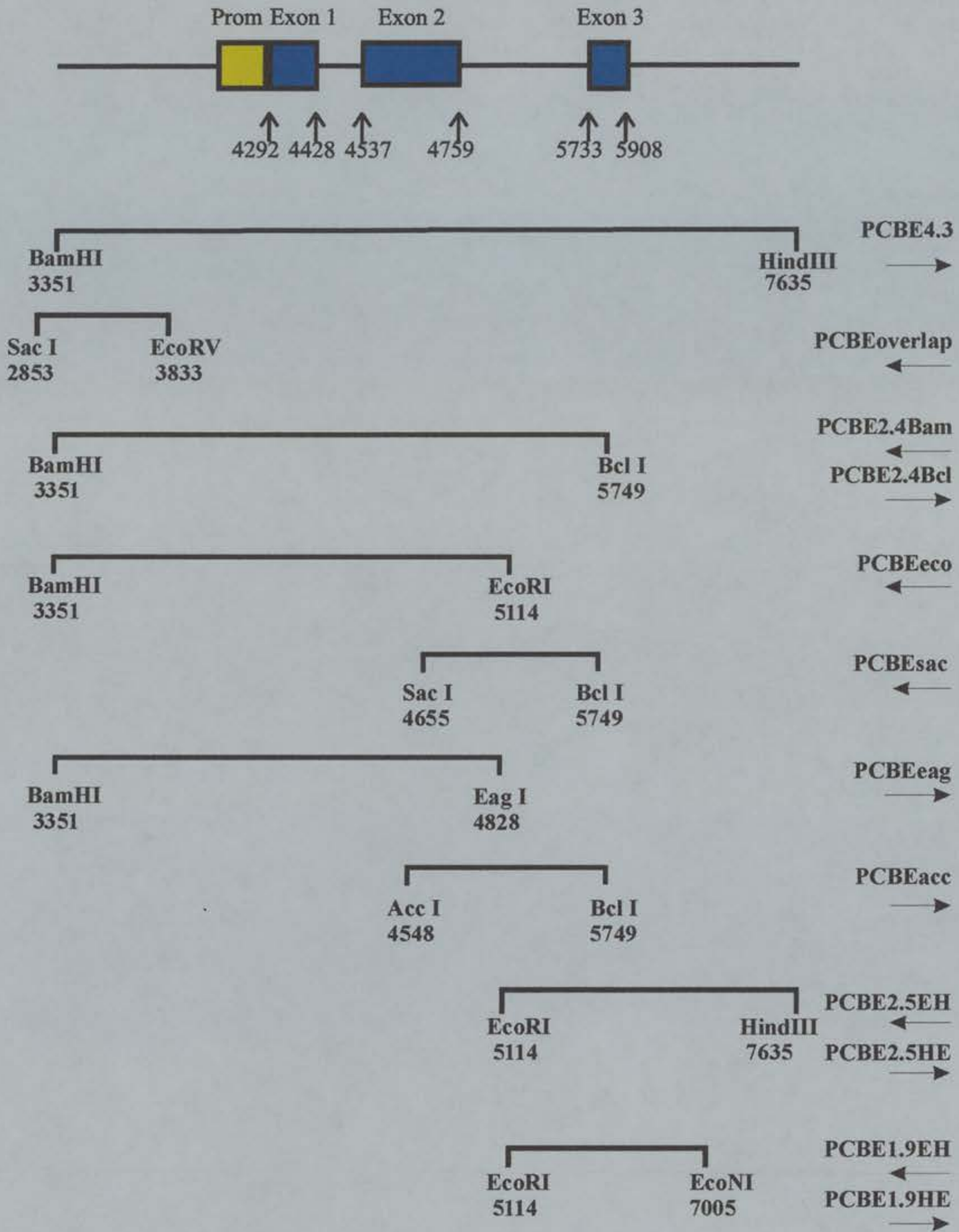


Figure 3.3 Plasmids used for monomer DNA extension mapping
The fragments shown above were cloned into plasmids for use in MDE mapping. The numbering indicates position of the fragments with respect to the β^Acap site. Arrows below the name of each clone show the direction in whch monomer DNA is extended during monomer DNA extension. Arrows in the upper part of the diagram indicate the start and end points of the exons, relative to the β^Acap site. (Not to scale).

minutes. The helper phage was titred as below.

3.2.3.2 Titre helper phage

Ten-fold serial dilutions of the phage were made in B broth (10% bactotryptone, 8% NaCl and 0.2% glucose) over the range $1x10^9$ to $1x10^2$ pfu/ml (assuming initial concentration approximately $1x10^{11}$). 0.1 ml of each phage dilution was added to 0.2 ml of log-phase DH11S. The phage-bacteria mixture was incubated at room temperature for 5 minutes and added to 0.7% B broth agar at 52 °C. This was poured onto B broth plates. The plates were incubated overnight at 37°C. The number of plaques was counted next day, to estimate the concentration of the helper phage stock in pfu/ml. The helper phage stock solution was then diluted with B broth to 10^{10} pfu/ml.

3.2.3.3 Preparation of DH11S phagemid SOB/glycerol stocks

All phagemids were transformed into *E. coli* strain DH11S (Lin et al., 1992; available from Gibco BRL) and 60:40 SOB/glycerol stocks made by mixing the culture with an equal volume of SOB:glycerol. These phagemid stocks were spread on LB plates containing 100 μg/ml ampicillin and incubated at 37°C overnight. Colonies were picked and streaked on M9 minimal medium plates containing ampicillin, to select for *E.coli* containing the F′ plasmid required for phage superinfection. After 48 hours growth at 37°C, colonies were picked into SOB and grown for 12-14 hours. An aliquot of this culture was mixed with an equal volume of 60:40 SOB/glycerol and stored at -70°C.

3.2.3.4 Growing culture for production of single-stranded DNA

100 ml of TBG was inoculated with 300 μl DH11S pBluescript plasmid SOB:glycerol (produced from a colony grown on an M9 plate-see above). Then 1 ml of

M13KO7 helper phage at approximately 10^{10} pfu/ml was added. This culture was incubated in a rotational shaker at 37°C, for 90 minutes. Ampicillin was added to a final concentration of 100 ug/ml and kanamycin to 70 ug/ml. The culture was incubated for a further 14 hours at 37°C.

3.2.3.5 Isolation of phage particles from the medium

The culture was spun at 7500 x g for 20 minutes at 4°C to pellet the bacteria. The supernatant, which contains the single-stranded DNA packaged in phage particles, was spun again to remove any remaining bacteria. The supernatant was added to 0.25 volumes of 40 % PEG 6000 in 2.5 M NaCl and incubated on ice for at least 60 minutes to precipitate the phage particles.

The phage particle precipitate was collected by centrifugation at 7500 x g for 30 minutes at 4°C. The pellet was resuspended in TE.

3.2.3.6 Purification of phage DNA from particles

Three different methods were used for phage purification.

3.2.3.6.1 Original Method

The phage particles, resuspended in TE, were frozen and thawed to help disrupt them. NaAc (pH 5.2) was added to a final concentration of 0.3 M, and the phage resuspension was extracted on ice; once with buffered phenol, twice with buffered phenol:chloroform and once with chloroform:IAA. After ethanol precipitation, the ssDNA was resuspended in TE.

SDS was added to a final concentration of 0.5%, EDTA pH 8 to a final concentration of 3 mM and NaAc (pH5.2) to 0.3 M. Extraction was carried out on ice repeatedly with

buffered phenol:chloroform (typically x 4) and once with chloroform:IAA. After ethanol precipitation, the ssDNA was resuspended in TE.

After addition of 0.6 volumes of 20 % PEG 6000 in 2.5 M NaCl, the ssDNA was incubated on ice for at least 60 min. Following a spin at 15000 x g for 15 minutes, residual PEG/NaCl was removed and the pellet washed with 70 % ethanol. The pellet was resuspended in TE. EDTA (pH 8) was added to a final concentration of 3 mM, SDS to a final concentration of 0.5%, and NaAc (pH 5.2) to 0.3 M. The ssDNA was extracted on ice repeatedly with buffered phenol:chloroform and once with chloroform:IAA. After ethanol precipitation, the pellet was resuspended in TE.

3.2.3.6.2 Tri-reagent method

After the phage particle pellet was resuspended in TE, 5 volumes of Tri-reagent (Sigma) were added. The samples were incubated for 15 minutes at room temperature. 0.1 ml of choloroform was added, and incubated at room temperature for a further 15 minutes, with occasional vortexing. The resulting mixture was centrifuged at 15000 x g for 15 minutes. The resulting aqueous layer was discarded. To precipitate the DNA from the interface and organic phase 0.15 ml of ethanol was added, and the sample incubated at room temperature for 5 minutes. To pellet the DNA the sample was centrifuged at 2000g for 5 minutes.

The resulting pellet was washed twice with 0.1 M sodium citrate, 10% ethanol for 30 minutes, and then washed twice with 75% ethanol for 10 minutes. The pellet was dried under vacuum and resuspended in TE. If the ssDNA was still contaminated with protein it was further cleaned by extraction with buffered phenol-chloroform in the presence of 0.3 M NaAc pH 5.5 and 0.5% SDS, followed by ethanol precipitation and resuspension in TE.

3.2.3.6.3 Qiagen method

The phage particle pellet was resuspended in 2.1 ml of TE and loaded onto columns from the Qiagen M13 ssDNA kit. The purification of the ssDNA was carried out as described in the manufacturers handbook. The purified DNA was eluted from the column in TE. If the ssDNA was still contaminated with protein it was further cleaned by extraction with buffered phenol-chloroform followed by ethanol precipitation and resuspension in TE.

3.2.3.7 Checking concentration and quality of ssDNA

Samples of the ssDNA, before and after heating at 95°C for 4 minutes, were run on a 1.6% agarose gel. Single-stranded DNA of known concentration was run on the gel for estimation of DNA concentration. Protein contamination of the ssDNA was indicated by a smeared band in the unheated, which typically resolved into a single clean band when heated. (A smeared band was considered to indicate protein contamination as steps taken to remove protein contamination were found empirically to reduce smearing in the unheated sample).

3.2.4 Preparation of chicken core histones

Nuclei prepared from adult chicken red blood cells as described in section 4.2.3 were washed three times in micrococcal nuclease buffer (section 4.2.1) (all centrifugations were carried out at 1200 x g). The final pellet was resuspended in 13 ml of micrococcal nuclease buffer and, after the addition of $CaCl_2$ to a concentration of 1 mM, digested with 50 U/ml micrococcal nuclease for 10 minutes. The digestion was stopped with 5 mM EDTA, and the sample was centrifuged at 12000 x g for 15 minutes. The supernatant was collected, supplemented with NaCl to a final concentration of 0.65 M and loaded onto a CL4B Sepharose column, equilibrated with TE and 0.65 M NaCl.

Fractions were collected, and their OD was read at 280 nm to identify those that contained chromatin. A sample of each fraction was assayed for linker histone contamination by SDS polyacrylamide gel electrophoresis. Fractions not contaminated by linker histone were pooled and loaded on a hydroxyapatite column, equilibrated with 0.65 M NaCl, 10 mM phosphate buffer, pH 6.8 (Sambrook et al., 1989), 0.2 mM PMSF and 0.1 mM EDTA. The column was washed with 0.65 M NaCl/ 10 mM phosphate buffer, pH 6.8 and core histones were eluted with 2M NaCl/ 10 mM phosphate buffer, pH 6.8. An aliquot of each fraction was assayed by SDS polyacryamide gel electrophoresis. The fractions were pooled and the concentration determined by the measuring the optical density at 280 nm (1 mg/ml of core histones has an OD of 0.36 at 280 nm). The preparation was divided into 1ml aliquots and stored at -70°C.

3.2.5 *In vitro* nucleosome reconstitution with linear plasmid DNA

The method of reconstitution involved dialysis from high salt and urea as described by Camerini-Otero et al. (1976) and Clark and Felsenfeld (1991).The pBluescript-ε construct, containing the region to be mapped was linearised by restriction with Bsa I, which cuts at ~1.4 kb from the insert. Following purification, ethanol precipitation and resuspension in TE, the linearised DNA was reconstituted with chicken erythrocyte core histones. Chromatin reconstitution was performed in 250 μl of TEP_{2M} (2M NaCl in TEP), containing 12.5 μg of core histones and 25 μg of linearised plasmid DNA (0.5:1 w/w ratio). The reconstitute was incubated at room temperature for 10-20 minutes and dialysed against a linear TEP_{2M}-$TEP_{0.4M}$ gradient at 4°C for approximately 5 hours. The reconstitutes were further dialysed against TEP_{16mM} overnight.

3.2.6 Preparation of mononucleosomal DNA

After reconstitution, chromatin samples were digested in TEP_{16} supplemented with

1 mM $CaCl_2$, with 10 U/ml micrococcal nuclease for 35 minutes at 0°C , followed by 5 minutes at 37°C. DNA protected from nuclease digestion by the presence of a core particle was purified by phenol/chloroform extraction, and ethanol precipitated. These DNAs were electrophoresed in a 1.6% agarose gel and the monomer band, hence monomer DNA, (~146 bp long) was cut from the gel, electroeluted, and purified by phenol/chloroform extraction, followed by ethanol precipitation.

3.2.7 Treatment of monomer DNA prior to use in monomer DNA extension

3.2.7.1 5′ end-labelling of monomer DNA

Monomer DNA (200 to 500 ng) was labelled in 50 mM Tris pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, 3.5 μl of [γ-^{32}P]ATP (3000 Ci/mmol) and 10 U of T4 polynucleotide kinase (Amersham). The reactions were incubated at 37°C for 60 minutes, and purified by phenol/chloroform extraction. The DNA was separated from unincorporated label by spinning through a 1 ml Sephadex G50 (Pharmacia) column equilibrated with water. The volume of the eluent was adjusted to 45 μl with water.

3.2.7 2Alkaline denaturation of monomer DNA

Monomer DNA was denatured by the addition of fresh NaOH to 0.2 M and incubation at room temperature for 10 minutes. To neutralise the reaction, ammonium acetate pH 7.5 was added to a 1M final concentration. The monomer DNA was ethanol precipitated, washed thoroughly, and resuspended in TE.

An aliquot of the monomer DNA was run on a denaturing 6% polyacrylamide gel to

check monomer DNA size and to assess the extent of single-strand nicking.

3.2.8 Monomer Extension for *in vitro* nucleosome mapping

Two monomer DNA extension techniques were used: the original method (MDE), and a new, oligonucleotide-based method (OMDE).

3.2.8.1 Original method

3.2.8.1.1 Annealing monomer DNA to ssDNA

A 3-fold molar excess of ssDNA (~ 1 to 2 μg) was mixed with 30 to 50 ng of the labelled, denatured monomer DNA in 100 mM NaCl, 2 mM DTT, 20 mM Tris-Cl pH 7.5, 20 mM $MgCl_2$ in a final volume of 25 μl. In a thermal cycler, the samples were denatured for 3 min at 95°C, incubated at 80°C for 30 seconds then the temperature was gradually reduced at a rate of 1°C every 100 seconds until a temperature of 55°C was reached. The samples were removed to ice.

3.2.8.1.2 Extension of annealed monomer DNA

These annealed samples were diluted to 50 μl by the addition of distilled water, dNTP's to a final concentration of 10 μM for each of the four dNTPs, BSA to a final concentration of 0.1 mg/ml, 4 U of DNA polymerase (Klenow fragment), and 20 U of the appropriate restriction enzyme. This reaction mix was incubated at 37°C for 60 minutes. The extension products were purified by buffered phenol/chloroform extraction and ethanol precipitation.

The extension products, resuspended in polyacrylamide gel loading buffer (SSM) and heat-denatured, were run on a denaturing 6 % polyacrylamide (19:1 polyacrylamide:bis)

gel. Size standards of C and T ^{35}S sequencing reactions of M13mp18 ssDNA (USB) made using the -20 Universal forward primer. ^{32}P end-labelled HinFI and DdeI digests of phage lambda were also run.

3.2.8.1.3 Analysis of the monomer DNA extension data obtained from the polyacrylamide gel

Quantitative densitometer scans were made for each gel after PhosphorImager (Molecular Dynamics) analysis of the dried gels. Using the co-ordinate position and size of marker DNA fragments on the gel, a six-order polynomial was generated in Sigma Plot, which could be used to convert the co-ordinates of fragments on the gel into DNA size. After the densitometry data had been transferred to Microsoft Excel, this equation was employed to accurately size the extended monomer DNA fragments. As the length of each fragment represents the distance from a nucleosome boundary to the restriction enzyme site used in the mapping, the location of the positioning site boundary relative to the beta-adult cap site can be determined. Densitometry traces of reactions carried out in the absence of restriction enzyme were considered as background for quantitative adjustment of traces of reactions including restriction enzyme.

Normalisation of traces from different reactions was achieved by employing common nucleosome positioning sites within the vector or globin sequence. In this way, a map of the whole ε globin gene could be generated, showing the location of each nucleosome positioning site with respect to the DNA sequence and the relative affinity of these binding sites.

3.2.8.2 New oligonucleotide-based method of monomer DNA extension

This differed from the original method, in that the ssDNA template is cut with a restriction enzyme prior to annealing of the monomer DNA.

3.2.8.2.1 Oligonucleotide annealing and restriction enzyme cutting of the ssDNA template

The conditions for the annealing reaction were adapted from the conditions used for annealing primers in the polymerase chain reaction (PCR); 0.9 pmoles of ssDNA were mixed with 4.5 pmoles of the appropriate oligonucleotide (obtained from Oswel) in 20 µl reactions containing 50 mM NaCl, 1.5 mM $MgCl_2$. and 10mM Tris-Cl pH 7.5. In a thermal cycler, the samples were heated to 95°C for 3 minutes, incubated at 80°C for 30 seconds, then the temperature was decreased by 1°C every 90 seconds until a temperature was reached which was 10°C below the annealing temperature of the oligonucleotide. The samples were put on ice.

To make the conditions more suitable for digestion by restriction enzyme, these samples were boosted in volume to 70 µl with distilled water, BSA (final concentration 0.1 mg/ml), 5 µl of the appropriate 10 X restriction buffer and 10-20 U of the appropriate restriction enzyme, and incubated at 37°C for 60 minutes. Following purification by phenol/chloroform extraction, and precipitation with ethanol, the cut ssDNA was resuspended in TE at approximately 200 ng/µl.

3.2.8.2.2 Annealing monomer DNA to ssDNA

The monomer DNA was annealed to the ssDNA as described in 3.2.8.1.1

3.2.8.2.3 Extension of annealed monomer DNA

Extension was carried out as described in 3.2.8.1.2, except that the ssDNA was incubated with Klenow polymerase in the absence of restriction enzyme. The samples were run on a polyacrylamide gel and analysed by densitometry as described for the

original technique in section 3.2.8.1.3.

3.2.9 Screening enzymes for ability to cut ssDNA in the oligonucleotide-based monomer DNA extension technique (OMDE)

Restriction enzymes used in OMDE must not be able to cut ssDNA. If an enzyme were able to cut at more than one site during OMDE, it would not be possible to define the position of nucleosome boundaries with respect to a single restriction site.

Therefore it was necessary to develop an assay to check whether restriction enzymes were capable of cutting ssDNA. Three OMDE-style annealing reactions were set up as described in section 3.2.8.2, except for the amendments mentioned; in (a), a positive control, the reaction was set up as usual; oligonucleotide was annealed to the ssDNA, then digested with the enzyme to be tested. In (b), the oligonucleotide was omitted from the annealing reaction, and then the ssDNA digested as usual; in (c), the oligonucleotide was annealed and the digestion reaction carried out in the absence of restriction enzyme. The three ssDNA samples were then dephosphorylated, purified, precipitated and 5′ end labelled with ^{32}P as described in sections 2.13 and 2.14. The labelled DNA was separated from unincorporated label by spinning through a Sephadex column (section 2.15). Following precipitation, the samples, resuspended in SSM, were run with size standards on a denaturing 6 % polyacrylamide gel to check for restriction products.

3.3 RESULTS

3.3.1 Development and Optimisation of Monomer DNA Extension

The new, oligonucleotide-based method of monomer DNA extension was adapted from the original technique (Yenidunya et al., 1994), with the intention of reducing the amount of subcloning required for the mapping. The strategy of the original and adapted methods is given in section 3.1.

3.3.1.1 Development of new oligonucleotide-based monomer DNA extension technique (OMDE)

In this section, evidence is presented which shows that the adapted, oligonucleotide-based technique can give reliable nucleosome positioning data.

The reliability of the new technique was tested by using the original and adapted methods to map nucleosome positions on the same piece of DNA. The reliability and validity of the original technique has already been demonstrated, as the nucleosome positioning data it produces are consistent with the results given by various other approaches (Yenidunya et al., 1994).

For both the original and adapted monomer extension reactions, single-stranded DNA was prepared from pCBA4.4 (Davey et al., 1995). This clone is a pBluescript KS minus phagemid containing a 4.4 kb EcoRI-Bam HI fragment of the chicken β^{A}-globin gene. 5′ end-labelled monomers were prepared from reconstituted pCBA64, a pCBA4.4 subclone (p64, Davey et al., 1995). Sac I was used as the restriction enzyme for mapping (Figure 3.4a).

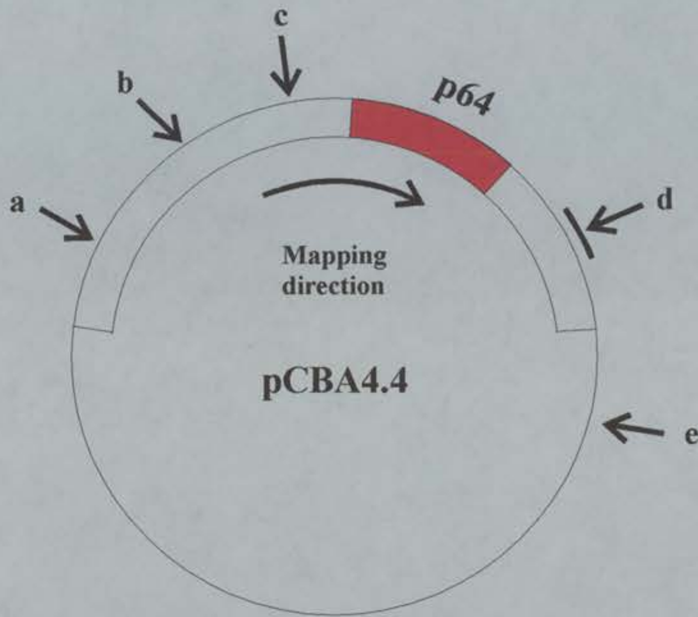


Figure 3.4a Schematic showing the mapping strategy for the monomer extension gel analysis shown in Figure 3.4b

This schematic shows the mapping strategy for the monomer DNA extension carried out on pCBA4.4 ssDNA with monomer DNA prepared from p64. The open box represents the 4.4 kb EcoR1-BamHI insert encompassing the β^A gene (from -1052 to +3369 relative to the β^A cap site). The red box indicates the position of the globin insert in p64, a pCBA4.4 subclone which extends from +2092 to +2780. The recognition sites for Sac I, the enzyme used in the mapping, are indicated by arrows labelled a to e. The oligonucleotide used in the oligonucleotide-based monomer DNA extension reaction is shown by the line at Sac I site d. In the oligonucleotide-based monomer extension reaction, monomer DNA prepared from p64 extends in the direction shown by the curved arrow on ssDNA prepared from pCBA4.4 which has been precut at Sac I site d only. In the original monomer DNA extension reaction the monomer DNA extends in the same direction to both Sac site d and e, before being cut. The diagram is not to scale.

The results of this experiment are shown in Figure 3.4b. The monomer extension reaction, carried out using the original method (lane 5), shows the features of a typical monomer extension mapping analysis. Bands up to ~ 250 bp in length are produced by self-extension of the monomer DNA (arrowed). Template-length and high molecular weight (HMW) DNA at the top of the gel (also arrowed) are produced by monomer DNA which extends on the single-stranded DNA and escapes restriction. The fragments occurring in the region (labelled as region Z) between the self-extended monomers and the HMW DNA are produced by extended monomer DNAs which are cut by restriction enzyme. Therefore, the size of each fragment in region Z represents the distance from one boundary of a nucleosome to the site of the restriction enzyme used in the mapping. In this case, the position of the 5′ boundary of the nucleosomes is mapped, as the minus strand of the ssDNA was used (see Figure 3.4a).

By comparing lane 5 (original method of monomer extension), with lane 1 (oligonucleotide-based method of monomer extension), it can be seen that the original and oligonucleotide-based methods give equivalent results. Therefore, the OMDE technique can produce good, reliable nucleosome positioning data.

Careful inspection reveals some differences between lanes 1 and 5 in the pattern of fragments longer than ~670 bp; these were predicted. In the original monomer extension reaction, Sac I cutting occurs at 5 sites (a, b, c, d and e in Figure 3.4a) within the plasmid, whilst in the adapted method, cutting is targeted to one site (d) where the oligonucleotide anneals. In the original monomer DNA extension reaction, pCBA64 monomer DNA can extend to both Sac site d and Sac site e where it is cut, whilst in the oligonucleotide-based method, monomer DNA can only extend to the pre-cut site d. Therefore, there will be extra fragments in the original monomer DNA extension reactions, produced by extended monomer DNA being cut at site e. The positions of the p64 clone and Sac I site (e) mean that these extra fragments are all greater in size than ~670 bp.

Controls for oligonucleotide-based monomer DNA extension are shown in lanes 2, 3

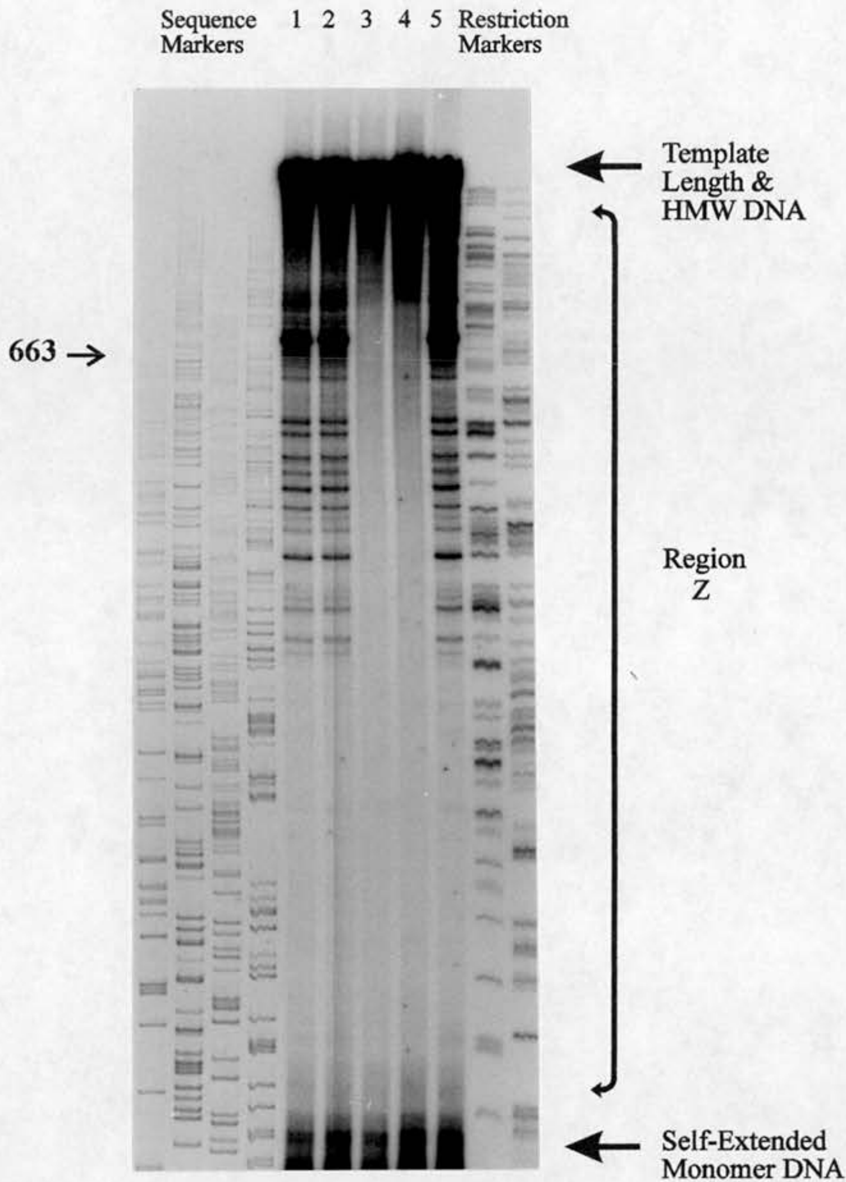


Figure 3.4b Demonstrating the reliability of the OMDE technique
6% denaturing polyacrylamide gel analysis of monomer DNA (prepared from p64) extended on ssDNA prepared from pCBA4.4, and restricted with Sac I. Lane 1 shows the result for the new (OMDE) technique, and lane 5 the original (MDE) technique. The fragments in region Z represent the positions of nucleosome boundaries from the Sac I cut site. Lanes 2, 3 and 4 are controls for the new technique; in 2, oligonucleotide was removed following restriction; in 3, oligonucleotide annealing and digestion steps were carried out in the absence of restriction enzyme, and in 4, oligonucleotide annealing and restriction were carried out in the absence of oligonucleotide. The sequence markers are A, G, C and T sequencing reactions of M13mp18 phage DNA, and the restriction markers are Dde I and Hinf I digest of lambda DNA.

and 4. Comparison of lane 2 with the standard OMDE reaction (lane 1) demonstrates that removal of the oligonucleotide prior to annealing of monomer DNA is not necessary, as results are identical whether it is removed or not. Lane 3 (digestion carried out in the absence of annealed oligonucleotide) demonstrates, by the absence of bands in region Z, that the Sac I restriction enzyme is unable to cut the ssDNA in the absence of annealed oligonucleotide. Lane 4 (oligonucleotide digestion carried out in the absence of restriction enzyme) demonstrates that the fragments in region Z in the standard OMDE reaction (lane 1) are restriction-dependent. Therefore, taken together, controls 2, 3 and 4 confirm that fragments in region Z in the standard OMDE reaction are restriction-dependent, specific DNA products which represent the distance from a nucleosome boundary to the site of the restriction enzyme (in this case Sac I).

In summary, Figure 3.4b shows that the protocol for the new technique can give good, reliable nucleosome positioning data.

3.3.1.2 Developing a strategy for mapping the entire epsilon globin gene

In a single monomer extension reaction, nucleosomes can be mapped at high resolution over a distance of up to 1 kb (Yenidunya et al., 1994; Davey et al., 1995). Therefore, restriction enzymes chosen for mapping needed to be separated by a maximum distance of about 800 bp, so that the positioning data from individual reactions would overlap. This would allow the data to be assembled into a single map, showing nucleosome positions on the entire gene region.

After choosing the restriction enzyme sites from the globin sequence (for which the EMBO gene bank identification is gghbbre), oligonucleotides were designed to anneal at these sites. The oligonucleotides were designed so that the recognition site of the restriction enzyme was in the centre, flanked by at least 6 bp on either side. It was necessary to check that the oligonucleotide annealed to the correct site only. The

oligonucleotide was end-labelled with ^{32}P, and then used as a primer in a sequencing reaction as described in chapter 2. The samples were run on a denaturing polyacrylamide gel, and the sequence was read to ensure the oligonucleotide was annealing at the correct site.

The restriction enzymes chosen for mapping had to fulfil certain criteria. Firstly, these enzymes should be able to cut near the ends of DNA molecules, as they would have to cut oligonucleotides ~ 20 bp in length. Secondly, the enzymes must not require an incubation temperature higher than the annealing temperature of the oligonucleotide. Thirdly, they must not be able to cut the single-stranded DNA used in the monomer DNA extension at any place other than the double-stranded region created by the annealed oligonucleotide. The data from a monomer extension in which a restriction enzyme cut more than once cannot be analysed, as nucleosome positions could not be defined relative to a single, unique restriction site.

3.3.1.3 Development of an assay to check whether restriction enzymes are capable of cutting single-stranded DNA

It was observed that a number of restriction enzymes were capable of cutting ssDNA because, in these cases, the pattern of fragments in the standard OMDE reaction was identical to that in the control reaction where restriction was carried out in the absence of annealed oligonucleotide. Therefore, it was necessary to develop an assay to check whether restriction enzymes were capable of cutting ssDNA before they were actually used in a monomer DNA extension reaction.

The protocol for this assay is given in section 3.2.9. The results of an assay, carried out on ssDNA obtained from PCBE2.5HE (Figure 3.3), cut with either Sau3AI or Pst I, are shown in Figure 3.5. Oligonucleotide-annealling reactions were carried out and the DNA digested, with the amendments described below, and then DNA was purified, ^{32}P labelled, and run on a polyacrylamide gel. Reactions without restriction enzyme were

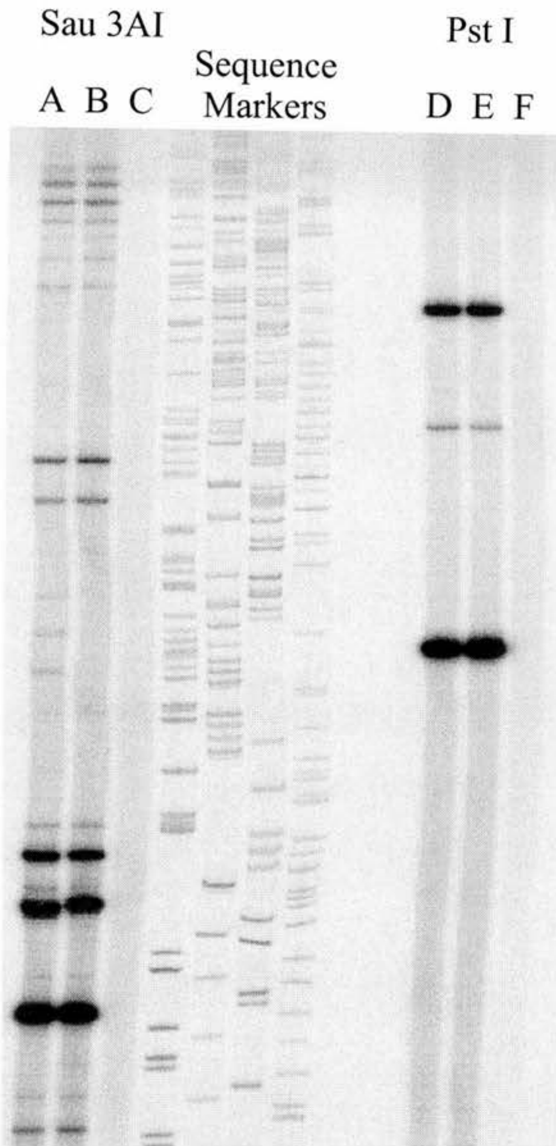


Figure 3.5 Assay to check whether restriction enzymes are capable of cutting single-stranded DNA.

This assay was carried out on PCBE2.5HE with either Sau3AI (lanes A, B and C) or Pst I enzyme (D, E and F). Lanes A and D show the fragments created by restriction in the presence of both annealed oligomer and restriction enzyme. Lanes B and E show the results of digestion with restriction enzyme in the absence of annealed oligonucleotide. Lanes C and F show the results of digestion in the presence of annealed oligonucleotide, in the absence of restriction enzyme. The sequence markers are A, G, C and T sequencing reactions of M13mp18 phage DNA.

included to show whether fragments on the gel were restriction-dependent (lanes C and F). When the restriction enzyme digestion was carried out in the absence of annealed oligonucleotide (lanes B and E), it gave an identical pattern of fragments to restriction enzyme digestion with annealed oligonucleotide (lanes A and D), indicating that Sau3AI and Pst I can cut the ssDNA whether oligonucleotide is annealed or not. Sizing of these fragments demonstrated that Sau3AI and Pst I were cutting at all possible sites within the PCBE2.5HE plasmid.

This assay was routinely used to check the suitability of enzymes for use in the new monomer extension technique.

3.3.1.4 Optimisation of the monomer extension technique

Experiments were carried out to find a more rapid and reliable method of making single-stranded DNA for monomer extension. The original protocol (adapted from Lin et al., 1992 and Sambrook, 1989) was very time consuming. In addition, the full-length clone, PCBE4.3 (Figure 3.3), consistently produced a lower yield of single-stranded DNA, and inferior quality monomer extension data compared with the shorter PCBE2.4Bam/PCBE2.4Bcl and PCBE1.9EH/PCBE1.9HE clones (Figure 3.3) which together covered the same region.

It was thought that the quality of ssDNA produced from PCBE4.3 might not be as high, and that this was affecting the processivity of the polymerase. Therefore, as well as trying to improve single-stranded DNA quality, other ways of improving the processivity of the polymerase were investigated. These included (i) adding single-stranded DNA-binding protein in the Klenow extension reaction, (ii) optimising the extension conditions for Klenow and (iii) using Taq DNA polymerase in the extension reaction instead of Klenow polymerase.

3.3.1.4.1 Making ssDNA by alternative protocols

The original protocol for ssDNA preparation, involving multiple rounds of phenol extraction, can take two days from the harvesting of the phage particles, whilst the Tri-reagent method takes only one day, and the Qiagen method less than an hour. The three protocols are described in sections 3.2.3.6.

The three preparations of single-stranded DNA produced data of equivalent quality when used in monomer extension reactions (Compare lanes A, B and C in Figure 3.6). Therefore, the more rapid Qiagen method was adopted for future experiments.

The two alternative methods of preparing single-stranded DNA did not give better monomer extension results on the PCBE4.3 clone than the original method, so the following experiments were carried out to try to optimise the processivity of the polymerase.

3.3.1.4.2 Including single-stranded DNA binding protein in the extension reaction

It was considered possible that the ssDNA was forming secondary structures which were impeding the progress of the polymerase. 1 or 2 μg of single-stranded DNA binding protein (Promega), which has been used in sequencing reactions to try to prevent the DNA from forming secondary structures (Ausubel et al., 1997), was added in the monomer DNA annealing step of a monomer DNA extension reaction.

It was found that including single-stranded DNA binding protein in the reaction did not improve the quality of monomer extensions on the PCBE4.3, PCBE2.4Bam or PCBE2.4Bcl clones (data not shown).

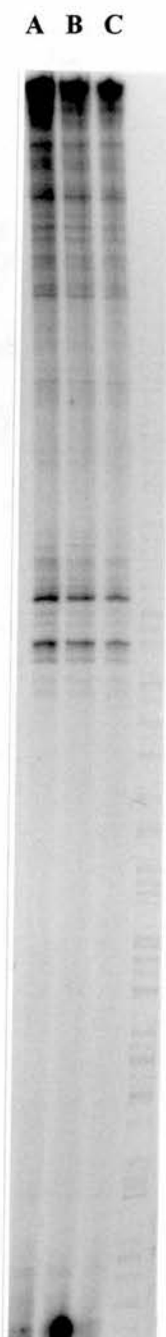


Figure 3.6 Comparison of monomer DNA extension data from reactions using ssDNA prepared by different protocols

6 % polyacrylamide gel analysis of monomer extensions on PCBE2.4Bcl single-stranded DNA prepared by three different protocols: the Tri-reagent method (lane A), the Qiagen method (lane B) and the original method (lane C).

3.3.1.4.3 Optimising extension reactions for Klenow polymerase

If the Klenow polymerase was falling off the DNA, its processivity might be improved by optimising the temperature for the extension reaction; the effect of temperature was investigated over the range 15°C to 50°C. The salt conditions were also altered so that they were as close as possible to the optimal buffer for Klenow polymerase.

The quality of the monomer extension data obtained from PCBE4.3 (Figure 3.3) clone was not improved by altering the temperature or salt conditions of the extension reaction (data not shown).

3.3.1.4.4 Using an alternative polymerase in the extension reaction

Taq DNA polymerase was tested in the extension reaction because it has a higher processivity than Klenow (Ausubel et al., 1997). The OMDE reaction was carried out as normal, except that in the monomer DNA extension step the Klenow polymerase was replaced with Taq polymerase, the reaction buffer was as close as possible to the standard Taq buffer as recommended by the manufacturer, and the temperature for extension was kept as close as possible to the optimum for Taq polymerase. Polyacrylamide gel analysis of the products of the reaction showed that the Taq polymerase only produced short (< 50 bp) fragments (data not shown), so it is not suitable for use in monomer DNA extension.

3.3.1.5 Summary of the development and optimisation of monomer extension

The new monomer extension technique was shown to give reliable nucleosome positioning data (Figure 3.4b). A more rapid method of producing single stranded DNA

was employed.

Changes in salt conditions, temperature and polymerase did not improve the quality of data produced by monomer DNA extensions on the full-length PCBE4.3 clone. However, this was not a problem for the project, as overlapping subclones (PCBE2.4Bam/PCBE2.4Bcl and PCBE1.9EH/PCBE1.9HE, Figure 3.3) of PCBE4.3 (Figure 3.3) gave good results when used in monomer extensions. PCBE4.3 was initially subcloned as PCBE2.4Bam/ PCBE2.4Bcl and PCBE2.5EH/ PCBE2.5HE (Figure 3.3); i.e. the 4.3 kb ε globin fragment from PCBE4.3 was split into two overlapping 2.4 and 2.5 kb fragments, and recloned in both orientations. The clones containing the 2.4 kb fragment (PCBE2.4Bam and PCBE2.4Bcl, Figure 3.3) produced good quality monomer DNA extension data, but the clones containing the 2.5 kb fragment (PCBE2.5EH and PCBE2.5HE, Figure 3.3) proved problematic. The 3′ end of the 2.5 kb fragment (from clones PCBE2.5EH and PCBE2.5HE) was removed to create subclones containing a 1.9 kb ε globin fragment (PCBE1.9EH and PCBE1.9HE, Figure 3.3), which produced good monomer DNA extension data.

Why the clones containing the 2.4 kb and 1.9 kb fragments produced better monomer DNA extension data than the clones containing the 4.3 kb or 2.5 kb fragments is uncertain. The length of the single-stranded DNA *per se* cannot be the most important factor, as the PCBE2.4Bam/ PCBE2.4Bcl clones, containing a 2.4 kb fragment, gave much better results than the PCBE2.5EH/ PCBE2.5HE clones, containing their similarly sized 2.5 kb fragment.

It is possible that repetitive elements which occur at high frequency in the β-globin cluster, and in particular around the chicken ε-globin gene (Villeponteau et al., 1982; Reitman et al., 1993) could play a role. The location of repetitive DNA within the cluster is shown in Figures 3.7a and 3.7b. The presence of these elements might adversely affect either the production of the ssDNA, or the extension reaction with Klenow polymerase. This proposal is supported by the fact that the PCBE1.9 clones produced better monomer extension data than the PCBE2.5 clones from which they were derived by the removal of

a region containing a CR1 repeat element at position 7248 relative to the β^A cap site (Figure 3.7b).

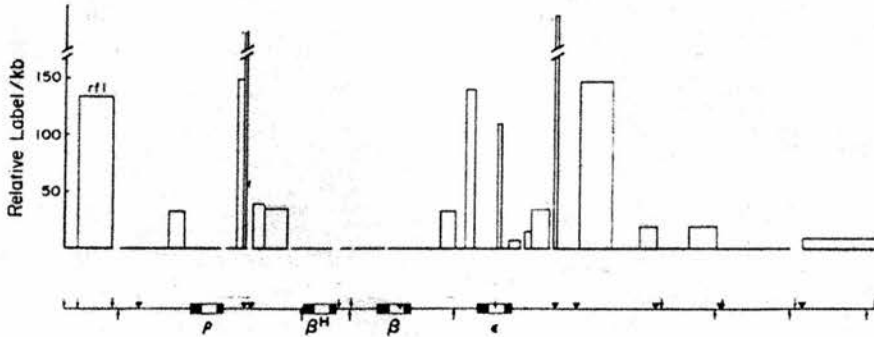


Figure 3.7a The location of repeat elements within the chicken β-globin gene cluster

Villeponteau et al. (1982) demonstrated by blot hybridisation that there are many repeat elements around the epsilon globin gene. The relative labelling (y axis) is related to the frequency of the repeat, the length of the repeat, and the homology to the probe (taken from Villeponteau et al., 1982).

Position of CR1 repeats relative to the β^A cap site	Length (bp)
2782 - 2819	38
3120 - 3361	222
3813 – 4066	254
5202 – 5358	157
6193 - 6369	177
6583 – 6756	174
7248 - 7640	213

Figure 3.7b Positions of CR1 repeats over the beta-adult and epsilon globin gene region

Reitman et al. (1993) sequenced the cluster, and identified many repeat elements, including CR1 repeats which occur at a frequency 35 times higher than in the bulk of the chicken genome. The CR1 repeat at 7248 occurs in the region which was removed (7005 to 7635 relative to the β^A cap site) from the PCBE2.5 clones to create the PCBE1.9 subclones (Figure 3.3) which produced much better quality monomer extension data.

3.3.2 Mapping Nucleosome Positions On The Chicken ε-Globin Gene

3.3.2.1 Introduction

Mapping of nucleosome positions was carried out using both the original and oligonucleotide-based monomer DNA extension methods (described in section 3.1). To achieve this, core particle DNAs, which are nucleosome positioning sequences, were prepared from the epsilon region. The core particle ('monomer') DNA was then extended on ssDNA produced from several subclones of the epsilon region. A variety of restriction enzymes were used in the monomer DNA extension reactions. The products formed during these reactions, which define the positions of nucleosome boundaries, were run on denaturing polyacrylamide gels. From densitometry analysis of these gels, a map was produced which revealed the placement and relative binding strength of nucleosome positioning sites on a 3.8 kb region containing the epsilon globin gene.

3.3.2.2 Preparation of core particle (monomer) DNA

Core histones were reconstituted onto DNA, and the resulting reconstituted chromatin was digested with micrococcal nuclease to produce core particles (as described in sections 3.2.4 and 3.2.5). The results of a typical reconstitution/digestion experiment are presented in Figure 3.8a.

After purification by excision and elution from an agarose gel, 5′ end-labelled monomer DNA was routinely run on a 6% denaturing polyacrylamide gel to assess its spread in the size and to check whether it had acquired significant levels of internal single-stranded nicking. A typical result is shown in Figure 3.8b.

All the monomer DNA extensions described in the next section, unless otherwise

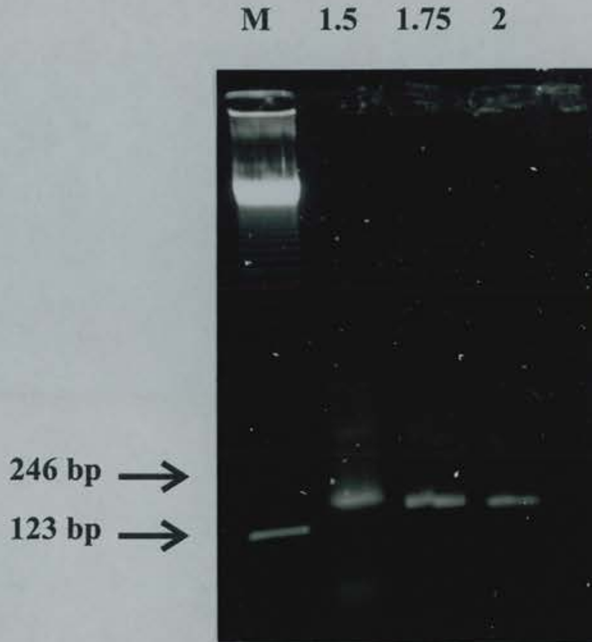


Figure 3.8a Monomer DNA preparation and characterisation

Reconstituted chromatin was digested with micrococcal nuclease at 0°C for 30 minutes, followed by 37 °C for 1.5, 1.75 and 2 minutes. Purified DNA fragments were separated on a 1.6% agarose gel. The markers (lane M) are a 123 bp ladder.

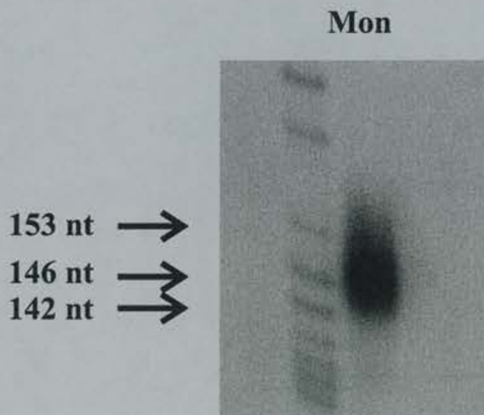


Figure 3.8b Integrity of the monomer DNA preparation

6% denaturing polyacrylamide gel showing purified, 5' end-labelled monomer DNA (lane labelled mon). Only a small region of the gel is shown, but on the whole gel fragments > 15 bp can be seen. Gels such as this were routinely used to confirm that the monomer DNA did not exhibit internal nicking.

stated, used monomer DNA prepared from PCBE4.3 (Figure 3.3), which encompasses the entire region to be mapped.

3.3.2.3 Results of monomer DNA extension mapping

A large number of monomer extension reactions, on a number of clones (Figure 3.3), were carried out to create a complete nucleosome positioning map for the epsilon gene region. A typical example of these monomer DNA extension reactions, run on a 6 % polyacrylamide gel, is shown in Figure 3.9b. In this example, monomer DNA was extended on single-stranded DNA prepared from PCBE1.9HE (Figure 3.3), cutting with Sph I, BstEII and Xho I enzymes. The schematic which accompanies this gel (Figure 3.9a) shows the region of the ε gene which was mapped, the cutting sites of the mapping enzymes and the direction in which monomer DNA was extended.

The monomer extension reactions in this gel were analysed as follows to produce a nucleosome positioning map for the 3′ region of the ε-globin gene. A quantitative densitometry scan was made of gel 3.9b and the data from the monomer extension in each lane were transferred to Microsoft Excel. Using the co-ordinate position and size in nucleotides of the marker fragments on the gel, a six order polynomial was generated in Sigma Plot to convert the co-ordinate position on the gel into DNA size (Figure 3.9c). This equation was employed to accurately size the extended monomer DNA fragments. The length of any one of these fragments, as discussed in section 3.3.1, defines the distance of one nucleosome boundary from the restriction enzyme site used in the mapping, and the relative intensity of the fragments reveals the relative binding strength of the positioning sites. Therefore, for each monomer extension in the gel, a densitometry trace can be produced which shows the relative strength of nucleosome positioning sites and their location with respect to the underlying globin sequence. For example, the densitometry trace for the monomer DNA extension on PCBE1.9HE where Xho I was the mapping enzyme (Figure 3.9b, lane a) is shown in Figure 3.9d. This map

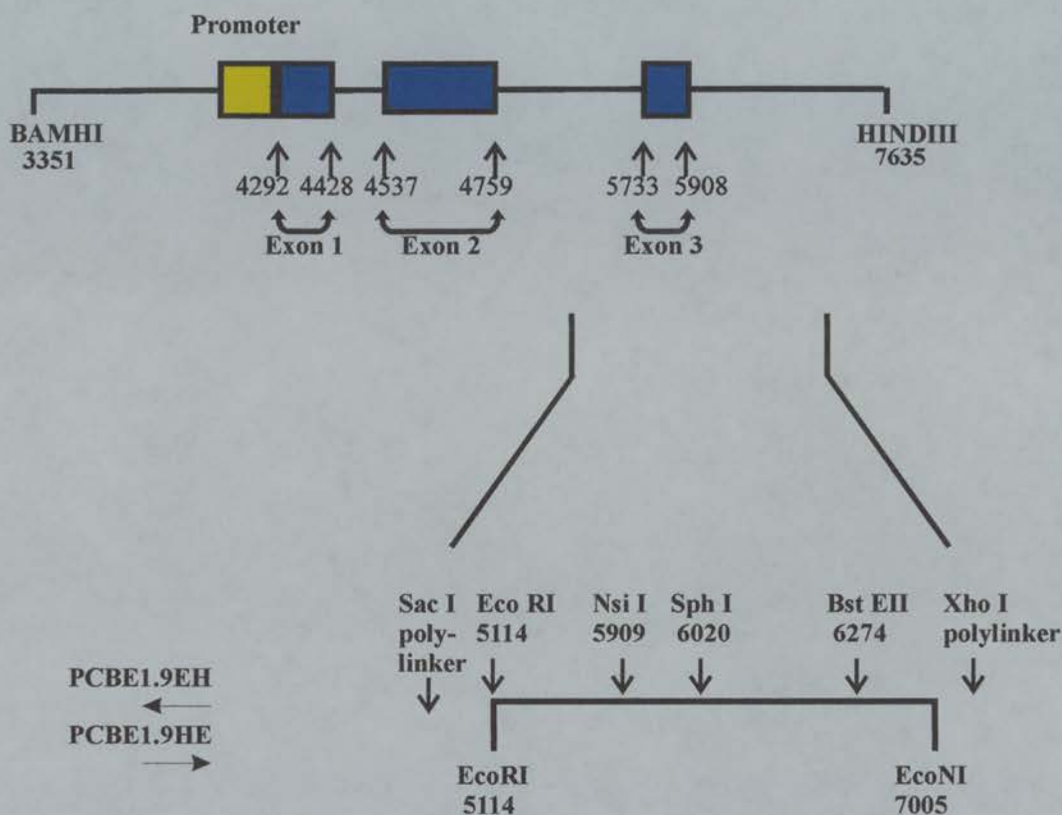


Figure 3.9a Schematic representation of the monomer DNA extension analyses of clones PCBE1.9HE and PCBE1.9EH that are presented in Figure 3.9b

The upper part of the figure shows the region of the ε–globin gene on which nucleosome positions were mapped in this analysis. The lower part of the figure shows the ε–globin gene fragment in the clones PCBE1.9EH and PCBE1.9HE which were used for the mapping. Monomer DNA was extended on ssDNA produced from these clones in the direction shown by the horizontal arrows, and restricted with one of the restriction enzymes indicated by the vertical arrows. All positions are relative to the β^{A}cap site.

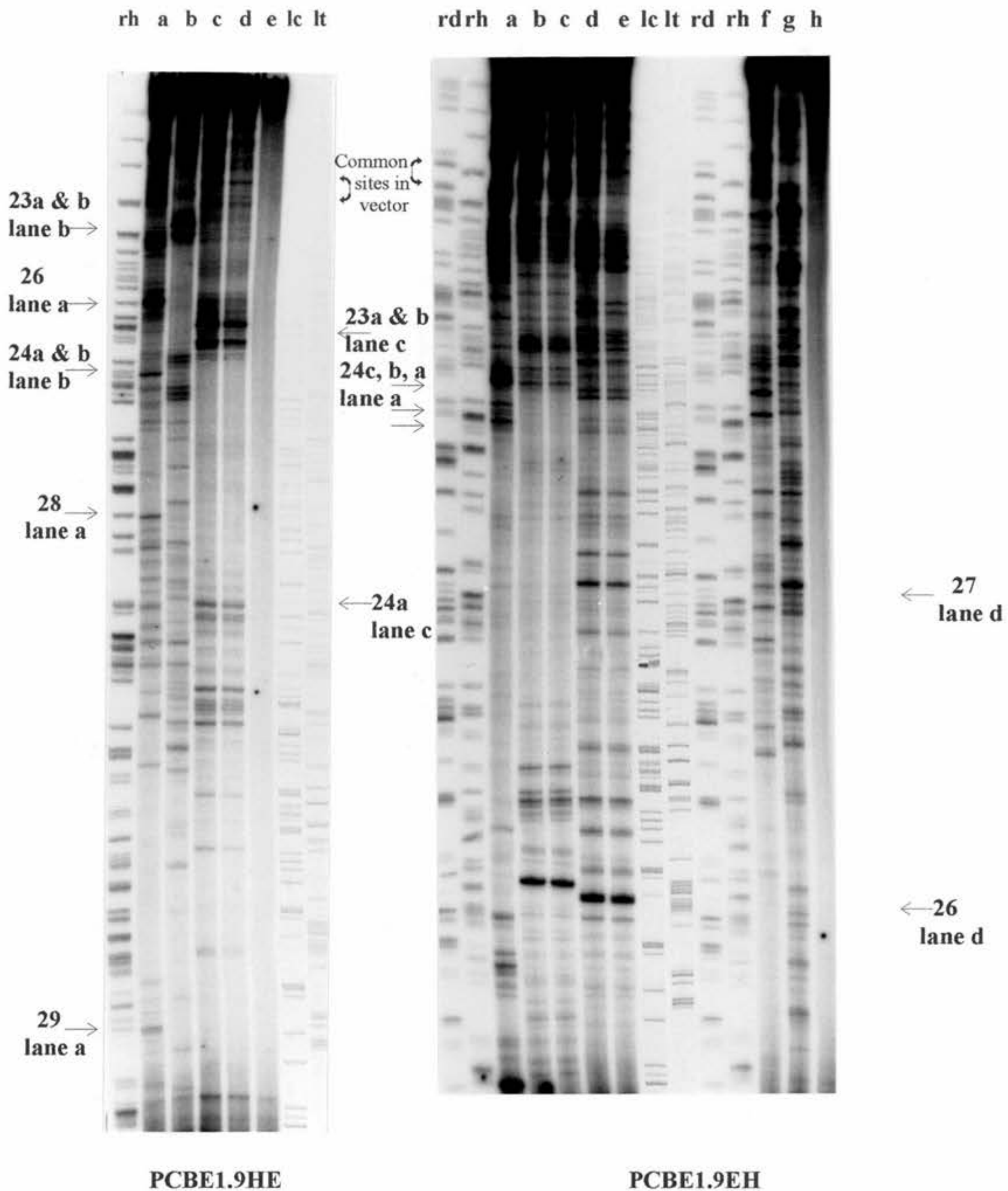


Figure 3.9b Polyacrylamide gel analysis of monomer DNA extensions on PCBE 1.9EH and PCBE1.9HE

Monomers were extended on single-stranded DNA prepared from PCBE1.9HE, and restricted with Xho I (lane a), Bst EII (lane b) and Sph I (lanes c and d). Lane e is a control extension without restriction enzyme cutting.
The right hand gel shows monomer extensions on PCBE1.9EH which have been restricted with Eco RI (lane a), Sac I (lanes b and c), Sph I (lanes d and e), Bst EII (lane f), PflMI (lane g; this lane was not analysed). Lane h is a control without restriction enzyme. Restriction markers (rd and rh) are lambda DNA cut with Dde I and HinFI, respectively. C and T sequencing reactions (lc and lt respectively) on M13 phage DNA are used as additional size markers. The arrows refer to strong positioning sites, which are numbered as shown in the final nucleosome positioning map (Figure 3.14a, and in Figure 3.9d). The curved brackets indicate the position of common histone octamer positoning sites in the vector which were used for normalising some of the lanes.

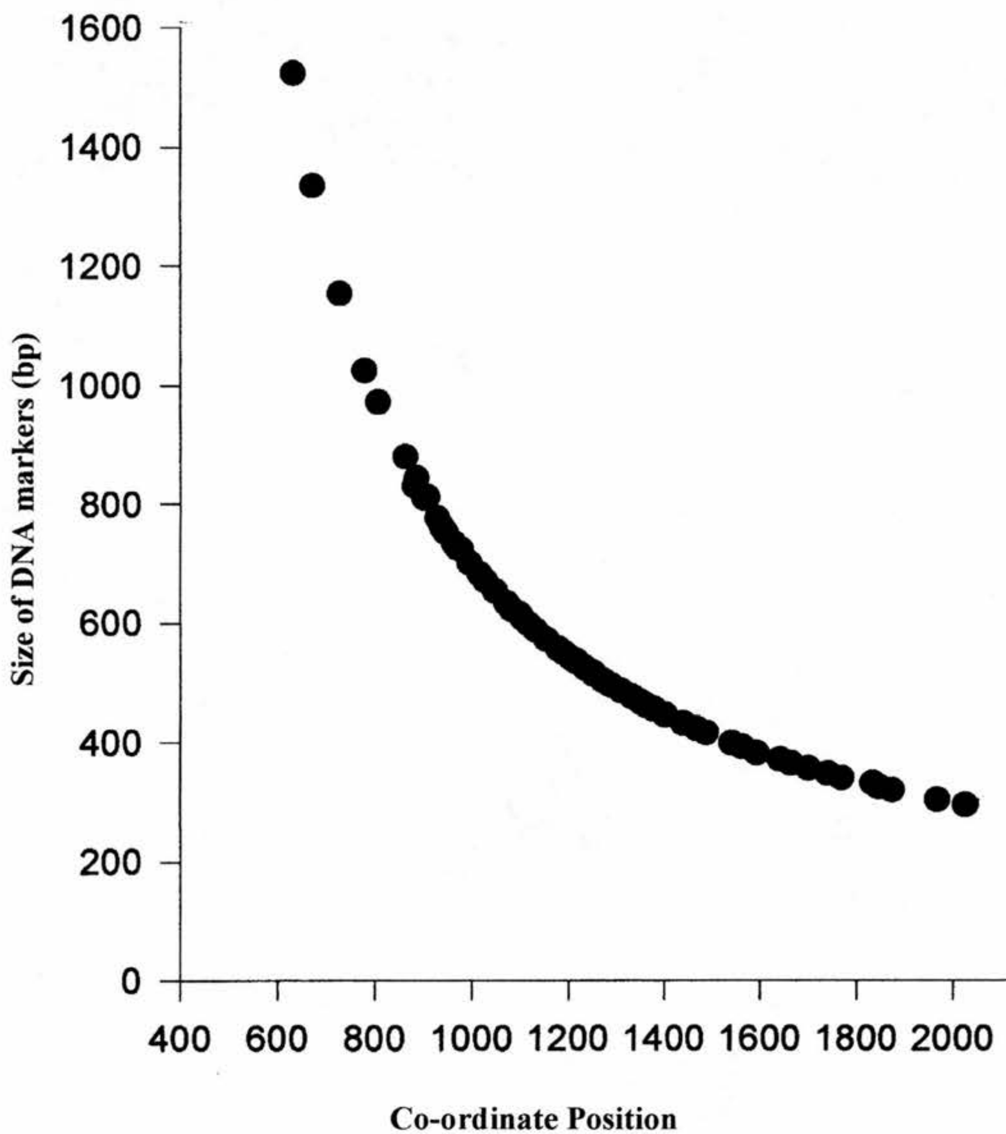


Fig 3.9c Size calibration of sequencing gels

The size (in nucleotides) of marker bands from the PCBE1.9HE polyacrylamide gel (Figure 3.9b) was plotted against their co-ordinate position to generate the curve shown above. A six order polynomial was generated in Sigma Plot from these data and the equation was employed to accurately size extended monomer DNA fragments on this gel.

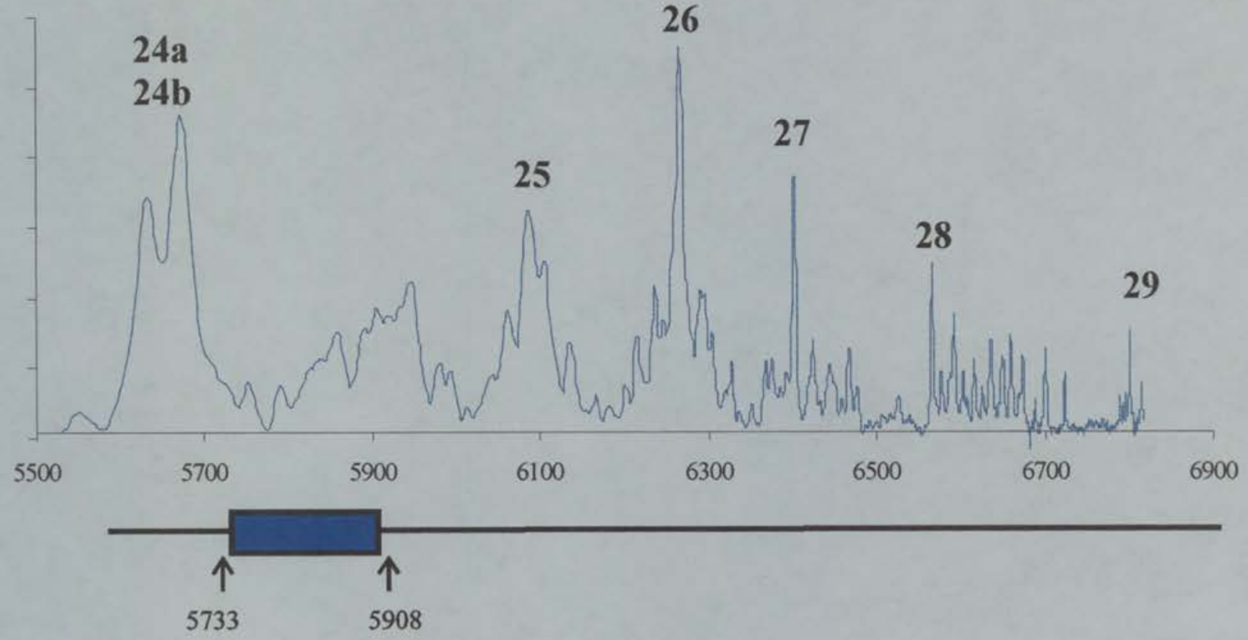


Figure 3.9d Nucleosome positioning map produced from the monomer DNA extension analysis carried out on PCBE1.9HE with Xho I (Figure 3.9b)

The centre of each positioning site is shown relative to the underlying ε-globin gene. The sequence (x axis) is numbered with respect to the cap site of the β^{A}-globin gene. The height of the peaks indicates the relative strength of the positioning sites. The figure below the trace shows the position of exon 3 of the ε-globin gene.

has been quantitatively adjusted for background by using densitometry data from a reaction carried out in the absence of restriction enzyme (Figure 3.9b, PCBE1.9HE, lane e).

A large number of monomer extension reactions were analysed by the method described above to produce densitometry traces (nucleosome positioning maps) covering different parts of the ε globin gene. A selection of the monomer extension gels analysed are shown in Figures 3.9b, 3.10b, 3.11b and 3.12b. There are schematics accompanying these (Figures 3.9a, 3.10a, 3.11a and 3.12a) which show the region of ε-globin mapped, the direction of mapping, and the restriction sites of the enzymes used for mapping. Nucleosome mapping was carried out by the original technique (Figures 3.9b, 3.10b and 3.12b) and by the oligonucleotide-based method (Figure 3.11b).

The monomer DNA extension reactions required for mapping the entire ε-globin gene region represent a small proportion of the total which were actually carried out. There were several reasons for this.

Firstly, each monomer extension was carried out with two different batches of monomer DNA (prepared from PCBE4.3) to demonstrate that there were no significant differences between the results of monomer extensions carried out with different core particle preparations.

Secondly, both 5′ and 3′ nucleosome boundaries were mapped. This allows the precise location of each nucleosome positioning site to be mapped, and then confirmed by the second mapping. To achieve this, pBluescript clones were prepared which would produce the minus strand ssDNA for mapping 5′ nucleosome boundaries over the entire ε globin gene region (Figure 3.3, for example, PCBE2.4Bcl, PCBEeag, PCBEacc and PCBE1.9HE, etc.). Other clones were generated which would produce plus strand ssDNA for mapping 3′ nucleosome boundaries over the same region (Figure 3.3 PCBE2.4Bam, PCBEeco, PCBEsac and PCBE1.9EH)

Mapping both nucleosome boundaries has also allowed another important control to be carried out. It can be used, as in Yenidunya et al. (1992) to demonstrate that monomer

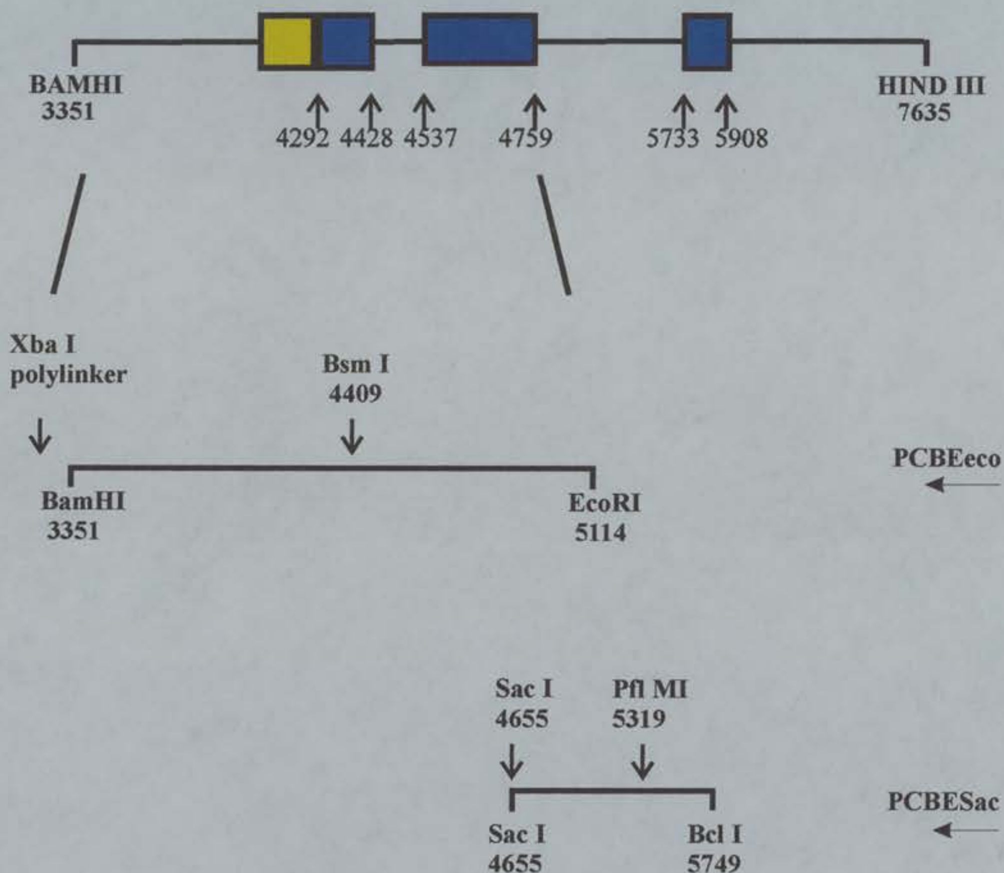


Figure 3.10a Schematic representation of the monomer DNA extension analyses on clones PCBEeco and PCBEsac presented in Figure 3.10b

The upper part of the figure indicates the region of the ε-globin gene on which nucleosome positions were mapped in this analysis. The lower part of the figure shows the ε-globin gene fragment in the clones (PCBEeco and PCBEsac) which were used for the mapping. Monomer DNA was extended on ssDNA produced from these clones in the direction shown by the horizontal arrow next to the name of the clone, and restricted at one of the sites indicated by a vertical arrow. All positions are numbered relative to the β^{A}cap site.

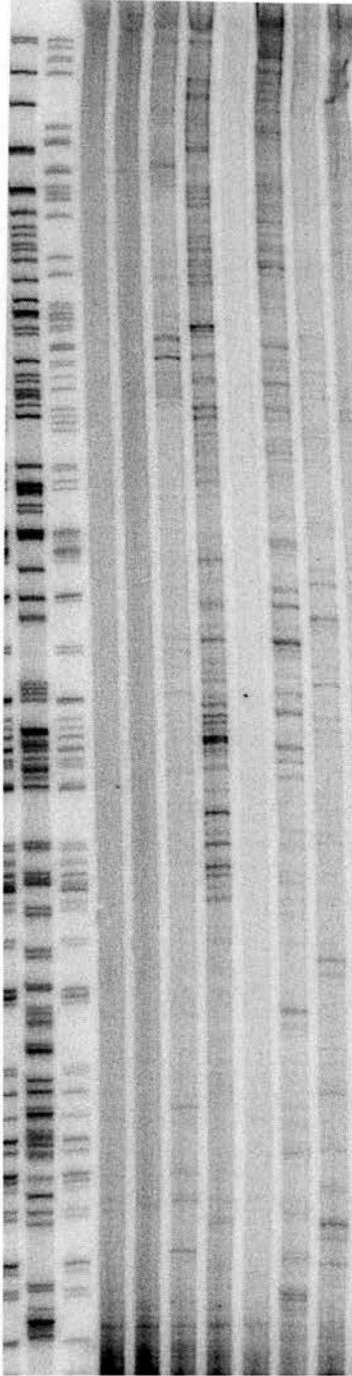


Figure 3.10b Polyacrylamide gel analysis of monomer extension reactions on PCBEeco and PCBEsac clones

Monomers were extended on ssDNA prepared from PCBEsac (lanes a to d), and from PCBEeco (lanes e to g), and cut with Sac I (lane c), PflMI (lane d),Xba I (lane f) and Sac II (lane g). Lanes a, b, and e are controls without restriction enzyme. The Sac II and PflMI reactions were not analysed for mapping nucleosome placement. The size markers are the same as those described in Fig. 3.9b.

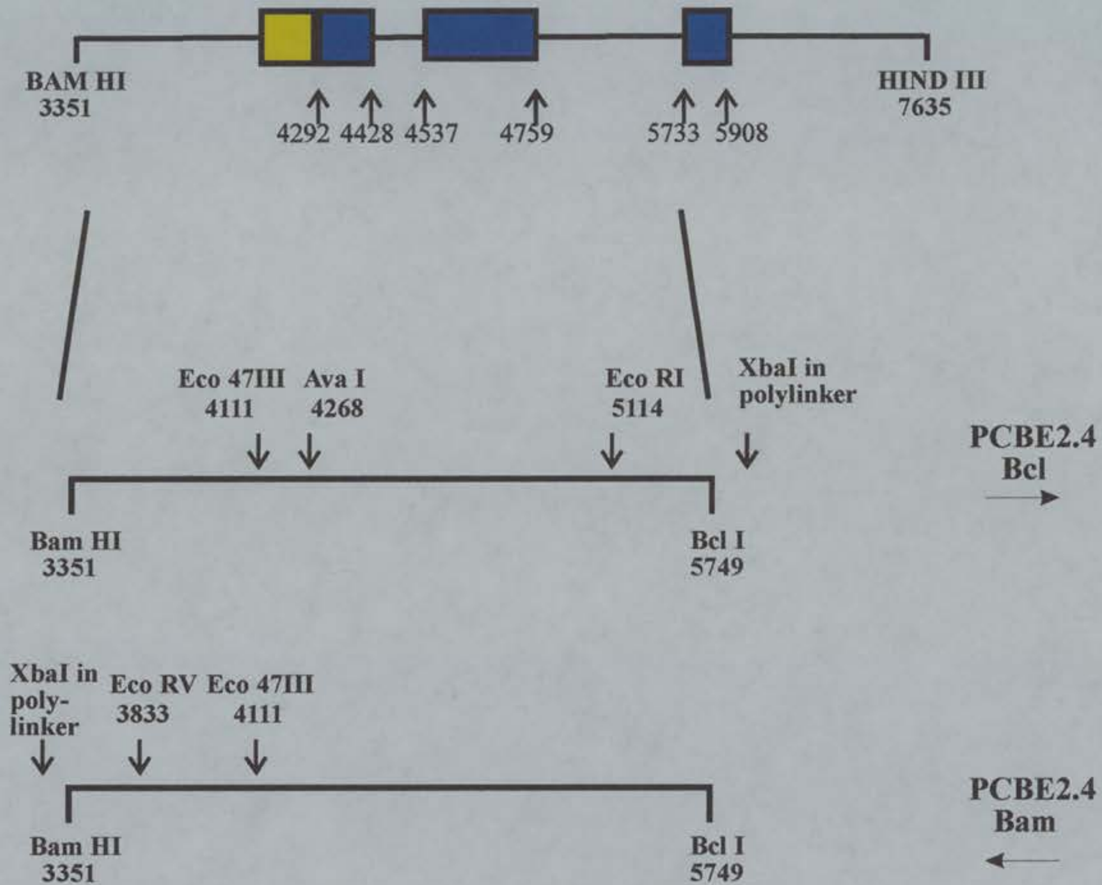


Figure 3.11a Schematic representation of the oligonucleotide-based monomer DNA extension analyses on clones PCBE2.4Bam and PCBE2.4Bcl presented in Figure 3.11b

The upper part of the figure indicates the region of the ε-globin gene on which nucleosome positions were mapped in this analysis. The lower part of the figure shows the ε-globin gene fragment in the clones (PCBE2.4Bam and PCBE2.4Bcl) which were used for the mapping. Monomer DNA was extended on ssDNA produced from these clones in the direction shown by the horizontal arrows, and restricted at one of the sites indicated by the vertical arrows. All positions are numbered relative to the beta-adult cap site.

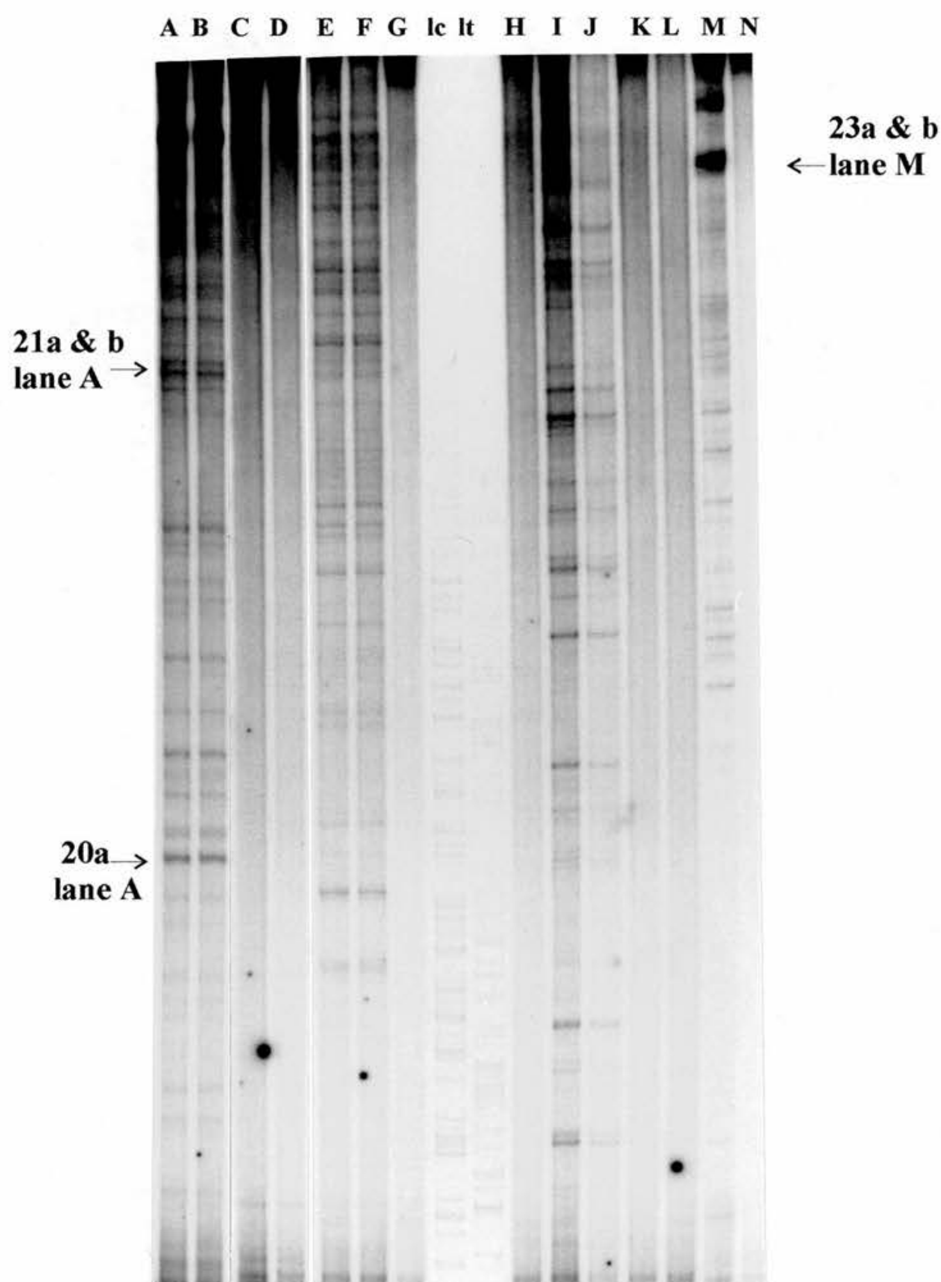


Figure 3.11b 6% polyacrylamide gel analysis of monomer extension reactions carried out by the new technique

The OMDE method of monomer extension was carried out using ssDNA prepared from PCBE24.Bcl (lanes E to L) and PCBE24.Bam (lanes A to D), restricting at an annealed oligomer with Eco RV (A and B), Ava I (lanes E and F) and Eco RI (I and J). Control reactions, carried out either without oligonucleotide, or without restriction enzyme are shown in lanes C and D (control for lanes A and B), G and H (for lanes E and F), and K and L (for lanes H and I). An analysis using the original method, on ssDNA produced from PCBE2.4Bam, restricted with Eco47III is presented (lane M), along with its control without restriction enzyme (lane N) . Size markers are as described in Figure 3.9b.

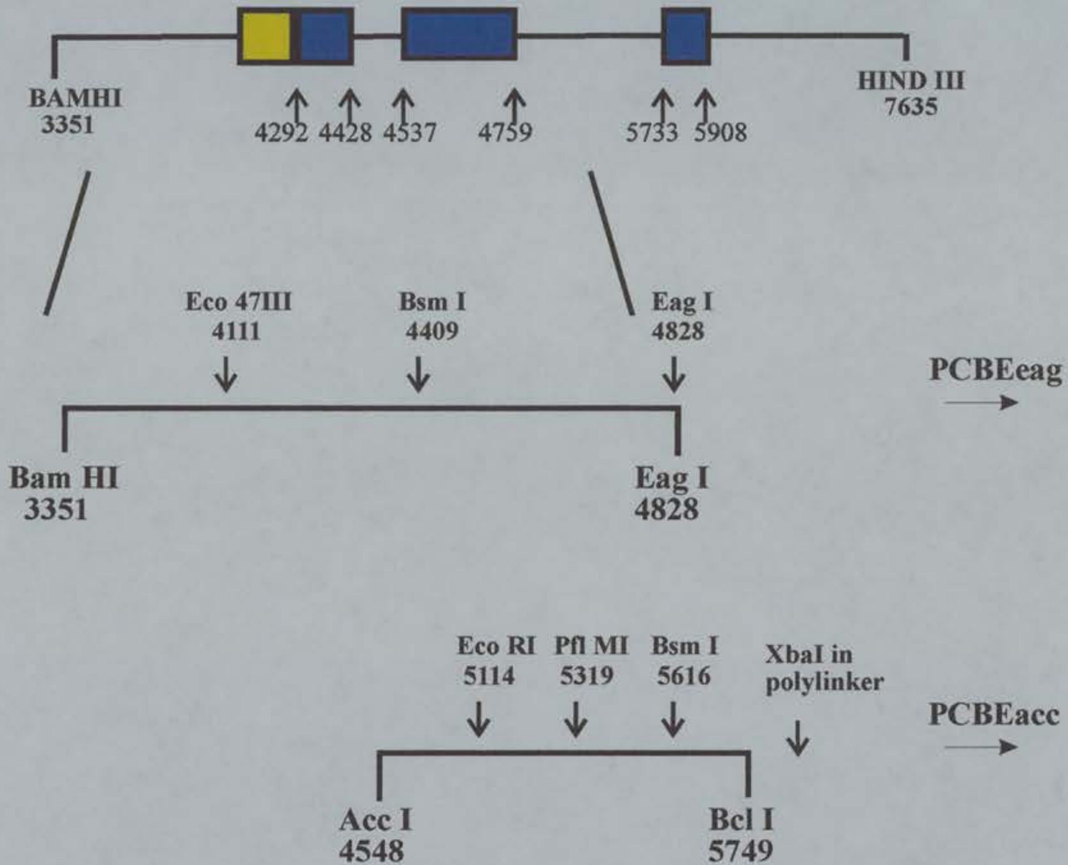


Figure 3.12a Schematic representation of the monomer DNA extension analyses on clones PCBEacc and PCBEeag presented in Figure 3.12b

The upper part of this figure indicates the region of the ε-globin gene on which nucleosome positions were mapped in this analysis. The lower part of the figure shows the ε-globin fragment in the clones (PCBEeag and PCBEacc) which were used for the mapping. Monomer DNA was extended on ssDNA produced from these clones in the direction shown by the horizontal arrows, and restricted at one of the sites indicated by the vertical arrows. All positions are numbered with respect to the beta-adult cap site.

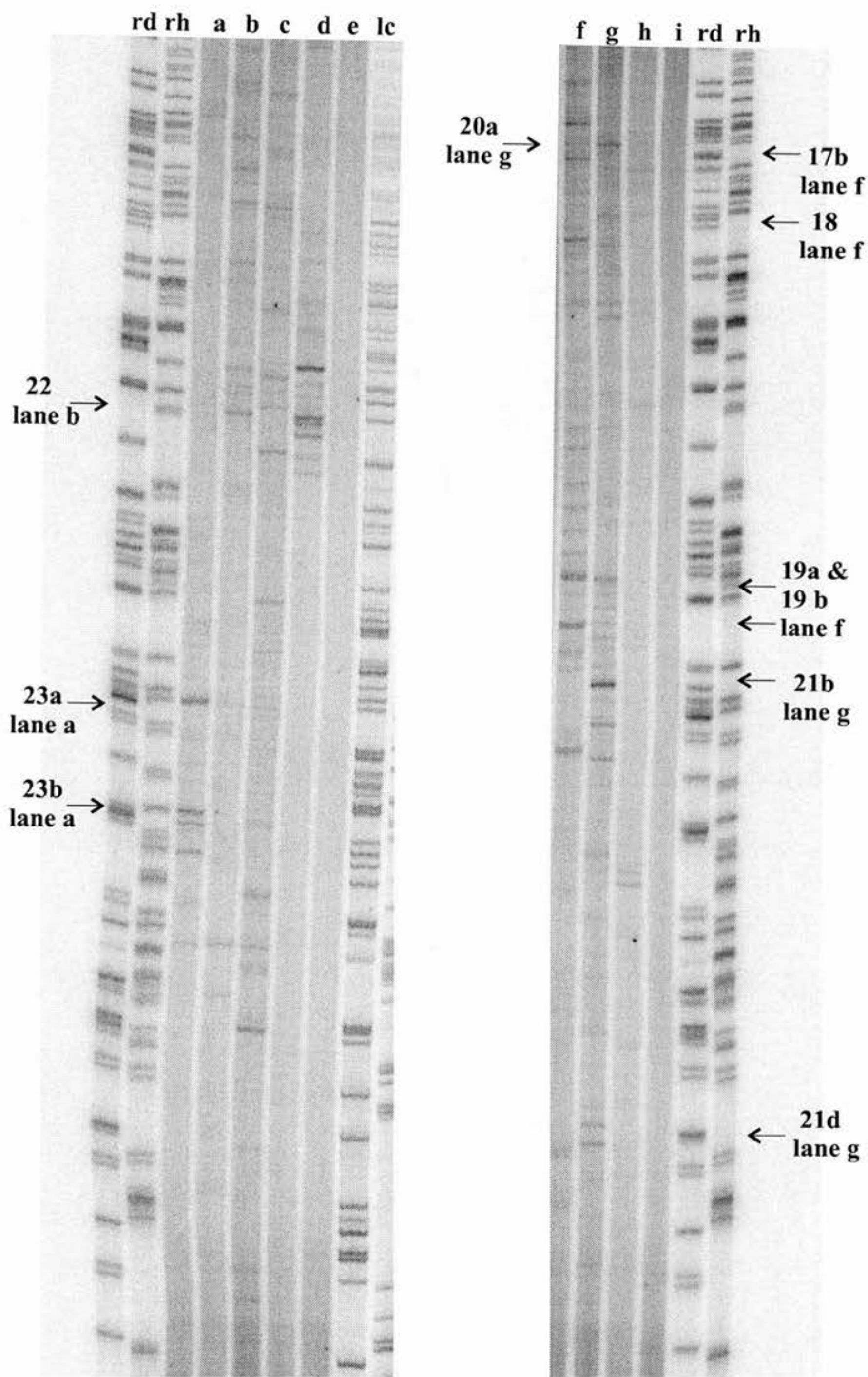


Fig. 3.12b 6% polyacrylamide gel analysis of monomer extension on PCBEeag and PCBEacc

Monomer extension reactions were carried out on PCBEacc (lanes a to e) and PCBEeag (lanes f to i). Restriction was carried out with Bsm I (lane a), PflMI (b), Eco RI (c), Xba I (d), Eco 47III (f), Eag I (g) and Bsm I (h). Controls without restriction enzyme are presented in lanes e and i. Size markers are the same as those described for Figure 3.9b.

DNA extension provides quantitative information about the relative strength of nucleosome positioning sites. It was used in this way to confirm that the nucleosome positioning study on the ε-globin gene was quantitative. When both 5′ and 3′ nucleosome boundaries were mapped on the ε-globin gene, and the data compared, no significant qualitative or quantitative differences were observed (for an example, see Figure 3.13). This confirms that the labelling efficiency of the plus and minus strands of a monomer DNA is equivalent. Therefore the intensity of extended monomer DNA fragments can be taken to reflect their relative abundance, rather than differences in end-labelling efficiency, so monomer DNA extension on the epsilon globin gene does provide quantitative information about the relative strength of nucleosome positioning sites.

Thirdly, it has been possible to compare the data produced by the original and oligonucleotide-based monomer extension methods. Carrying out oligonucleotide-based monomer DNA extension on ssDNA from PCBE2.4Bam and PCBE2.4Bcl (Figure 3.3) would have been sufficient to map nucleosome placement across the 2.4 kb region contained within these clones. However, this region was subcloned into PCBEsac, PCBEeco, PCBEeag and PCBEacc (Figure 3.3), so that nucleosome positioning on this region could also be mapped by the original monomer DNA extension method. The purpose of this was not to test the validity of the adapted method, which has already been demonstrated (section 3.3.1.1). This was done to compare the *quality* (for, example, the amount of background) of the nucleosome positioning data produced by the original and adapted techniques over a longer stretch of DNA. The quality of the data produced by the original and adapted methods of monomer DNA extension was equivalent (compare Figures 3.11b and 3.12b).

The positioning maps covering different parts of the epsilon region were linked together to create a single nucleosome positioning map covering the entire epsilon globin gene region. To achieve this, densitometry traces from different monomer extension reactions were normalised with respect to each other by employing common nucleosome

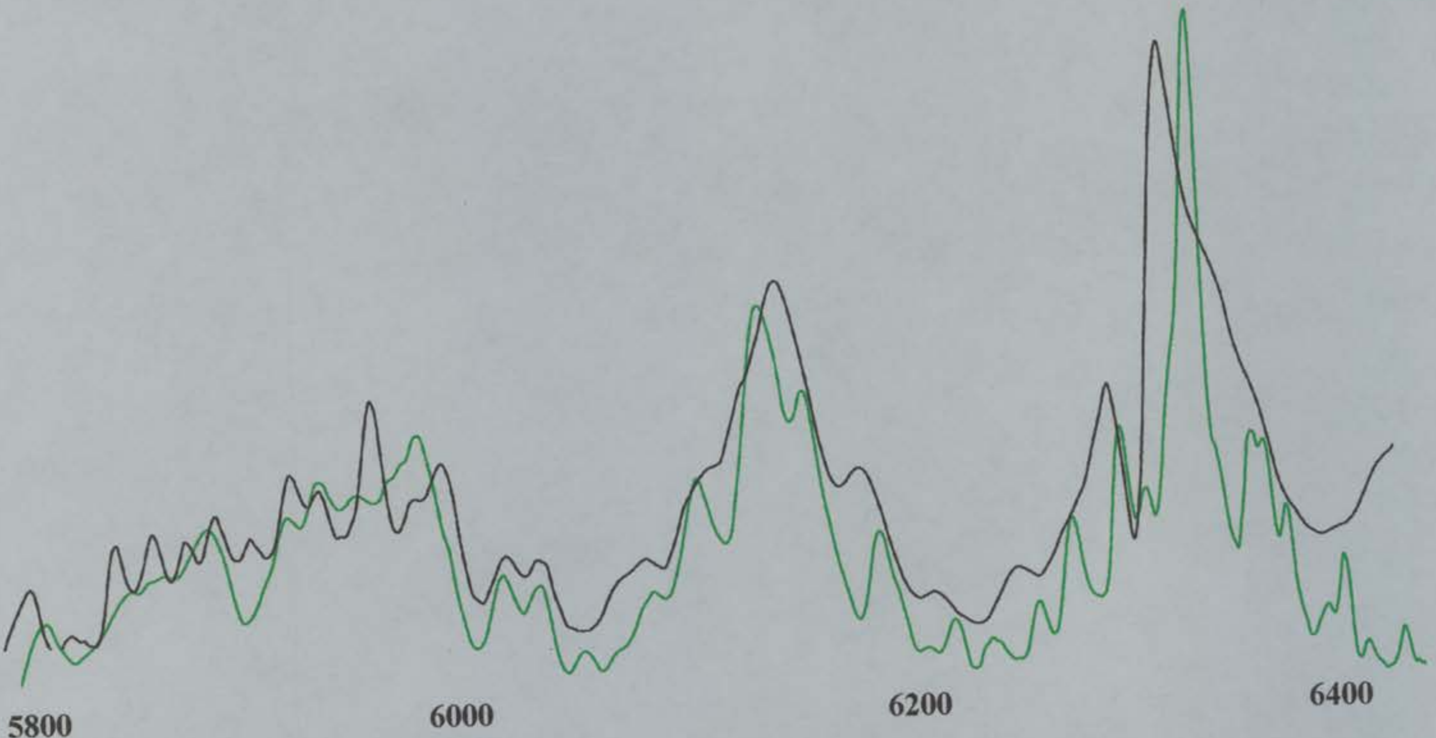


Figure 3.13 Monomer extension provides quantitative data on the relative strength of nucleosome positioning sites
The plus strand (green trace) and the minus strand (black trace) of monomer DNA were extended on complementary ssDNAs to map nucleosome positions over the same region of ε-globin, but in opposite orientations. The position of the nucleosome dyads (which occur at the midpoint between the two boundaries) are shown. From comparison of the overlaid traces, it is evident that nucleosomes, which are mapped to virtually identical positions, show effectively the same relative intensities in both orientations. Analysis of the smallest extension products when mapping in one direction will reveal positioning on a region which is covered by the largest extension products when mapping in the other direction, leading to the difference in resolution seen between the two traces.

positioning sites within the vector or globin sequence. For example, common nucleosome positioning sites in the vector which were used to normalise data from monomer DNA extension on the PCBE1.9EH and PCBE1.9HE clones are indicated by asterisks in Figure 3.9b. The resulting ~ 3.8 kb map of the entire globin gene region shows the location and relative strength of nucleosome positioning sites with respect to the DNA sequence (Figure 3.14a). The major nucleosome positioning sites in this map are numbered for the purpose of discussion, and these main sites are also indicated in the monomer extension gels shown in Figures. 3.9b, 3.10b 3.11b and 3.12b. The numbering begins at 17 so that it follows on from the numbered positions in the contiguous β^A map (Davey et al., 1995).

3.3.2.4 Nucleosome positioning map for the globin genes

The map showing nucleosome positioning sites over ~ 3.8 kb of the ε globin gene (Figure 3.14a) adjoins that for the β^A-globin gene region (Davey et al., 1995). The data for the ε globin genes were normalised with respect to that for β^A-globin using common core histone positioning sites within pBluescript vector sequence, to produce a map covering 7.2 kb of continuous sequence containing the β^A and ε-globin genes and the shared enhancer that lies between them. This is shown in Figure 3.14b.

Comparison of the β^A and ε-globin regions of the map reveals some strong similarities. In both cases, many of the possible positioning sites are occupied at low (essentially background) frequencies, whilst a subset of sites are occupied at a much higher frequency. Integration of the area under the strongest nucleosome positioning site (23b in Figure 3.14b) reveals that its intensity is ~1000 times greater than background. By comparison, the strongest positioning site in β^A has an intensity ~ 300 times greater than background. This is a measure of the range in the relative strength of all potential core histone octamer positioning sites on the ε and β^A globin genes.

The long and short-range features of the nucleosome positioning map shows some

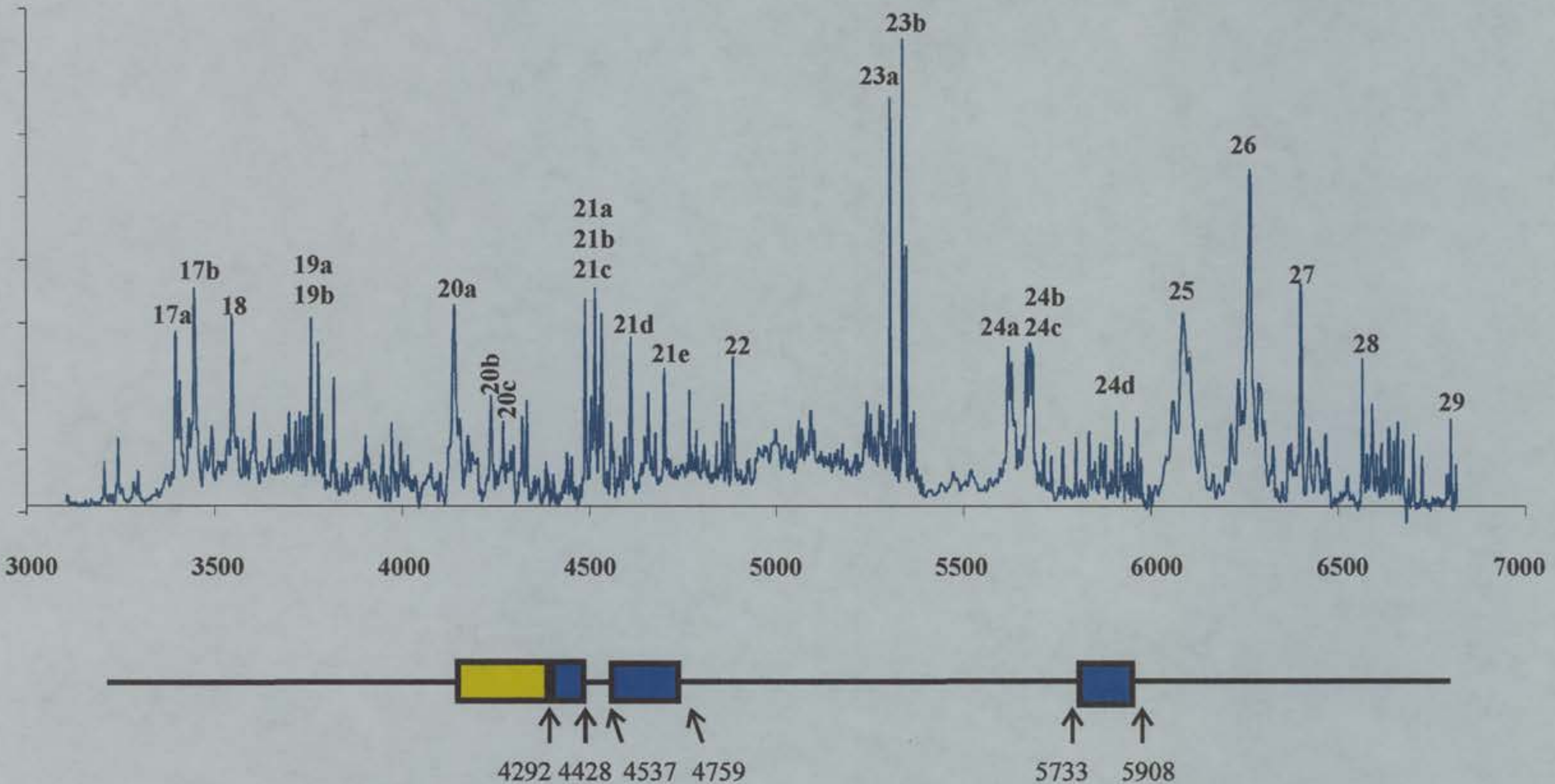


Figure 3.14a Map of the histone octamer (nucleosome) positions on the epsilon globin gene region

The centre of each positioning site is shown relative to the underlying epsilon gene. The sequence (x axis) is numbered with respect to the cap site of the beta-adult globin gene, to allow it to be joined to the beta-adult map already produced (Davey et al., 1995). The y axis indicates the relative strength of the positioning sites.

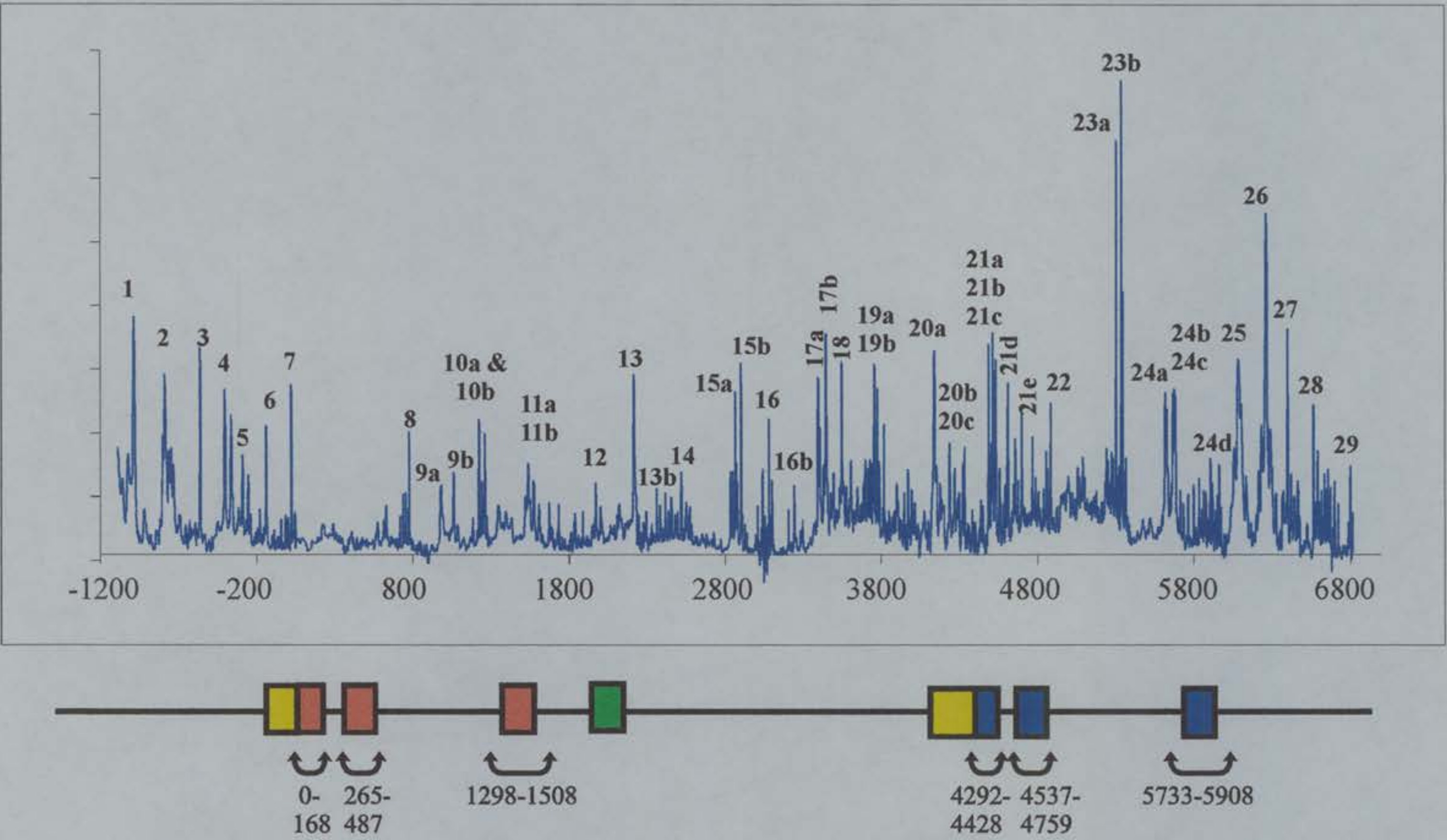


Figure 3.14b Combined map of the histone octamer (nucleosome) positions on the region containing the beta-adult and epsilon globin genes

The centre of each positioning site is shown relative to the underlying epsilon globin gene. The sequence (x axis) is numbered with respect to the cap site of the beta-adult globin gene. The y axis indicates the relative affinities of the positioning sites.

striking similarities, and differences, between the ε and β^A genes. These are discussed in the sections which follow.

3.3.2.5 Features of the nucleosome positioning map for the β^A and ε genes

This nucleosome positioning map covers a 7.2 kb region, encompassing two globin genes and the regions flanking them. This region contains both coding and non-coding sequences, regulatory regions (two promoters and an enhancer), as well as 5′ and 3′ flanking and intergenic DNAs. A comparison of features in the nucleosome positioning map for the ε and β^A globin genes would be of interest because the two genes evolved from a common ancestral gene, so there are some sequence similarities, though the two genes have diverged a great deal, especially in the non-coding regions (Reitman et al., 1993). In addition, they show some similarities in gene expression (they occur within the same chromatin domain, and are both expressed in erythroid cells) and some differences (they are expressed at different developmental stages) (Felsenfeld et al, 1993; Orkin, 1995).

Long-range patterns in nucleosome placement can be detected, because a long (7.2 kb) stretch of DNA has been mapped. The data have been analysed for the occurrence of long-range periodicities in the placement of nucleosomes, and the distribution of sites between introns, exons and flanking regions has been evaluated.

The map also reveals short-range features of the chromatin structure. It shows nucleosome placement over particular sites on the gene, including promoters or intron-exon boundaries.

3.3.2.5.1 Distribution of strong positioning sites within the epsilon gene

The strongest positioning sites are found in the introns and flanking regions of the

gene, rather than in the coding regions. Of the twenty most prominent positioning sites (Figure 3.14a), six are in the 5' flanking region (peaks 17a, 17b, 18, 19a, 19b and 20a), three are in the first intron (peaks 21 a, b and c), six are in the second intron (peaks 22, 23a, 23b, 24a, 24b and 24c) and four are in the 3' flanking region (peaks 25, 26, 27 and 28). By contrast, no prominent positioning sites occur in the first, (4292 to 4428) and third (5733 to 5908) exons. The centres of two strong sites (21b and 21c) occur in the first intron, but close to the boundary of the second exon, and a slightly weaker site (21d, at 4612) occurs within it. As the exons represent 14 % of the total number of base pairs mapped, three of the twenty strongest positioning sites on the ε-globin gene might be expected within the exons. In fact, only one fairly strong site is found in the exons, within exon two, but two strong sites are placed close to its upstream boundary. From this analysis, it cannot be argued in fact that there is a marked under-representation of nucleosome positioning sites within the exons. Integration of the area under the ε and β^A-globin exons in Figure 3.14b for the ε-globin gene represents 10.7 % of the total area under the whole map, whilst the exons represent 13.8 % of the total sequence in the map. By this criterion, it could be suggested that nucleosome positioning sites are under-represented in the exons.

3.3.2.5.2 Long-range periodicities of nucleosome positioning sites

Inspection of the nucleosome positioning map for the ε-globin gene (Figure 3.14a) reveals that strong positioning sites display a notable periodicity, a feature shared with the β^A-globin gene (Davey et al., 1995). Autocorrelation analyses of the ε-globin data show that strong positioning sites are commonly separated by an interval of ~360 and ~540 bp (Figure 3.15A). This suggests a long-range periodic function, possibly based on ~180 bp repeats. Fourier transformation of the autocorrelation function for ε-globin reveals two prominent peaks at 177 and 186 bp, which confirms an average periodicity of around ~180 bp (Figure 3.15B)

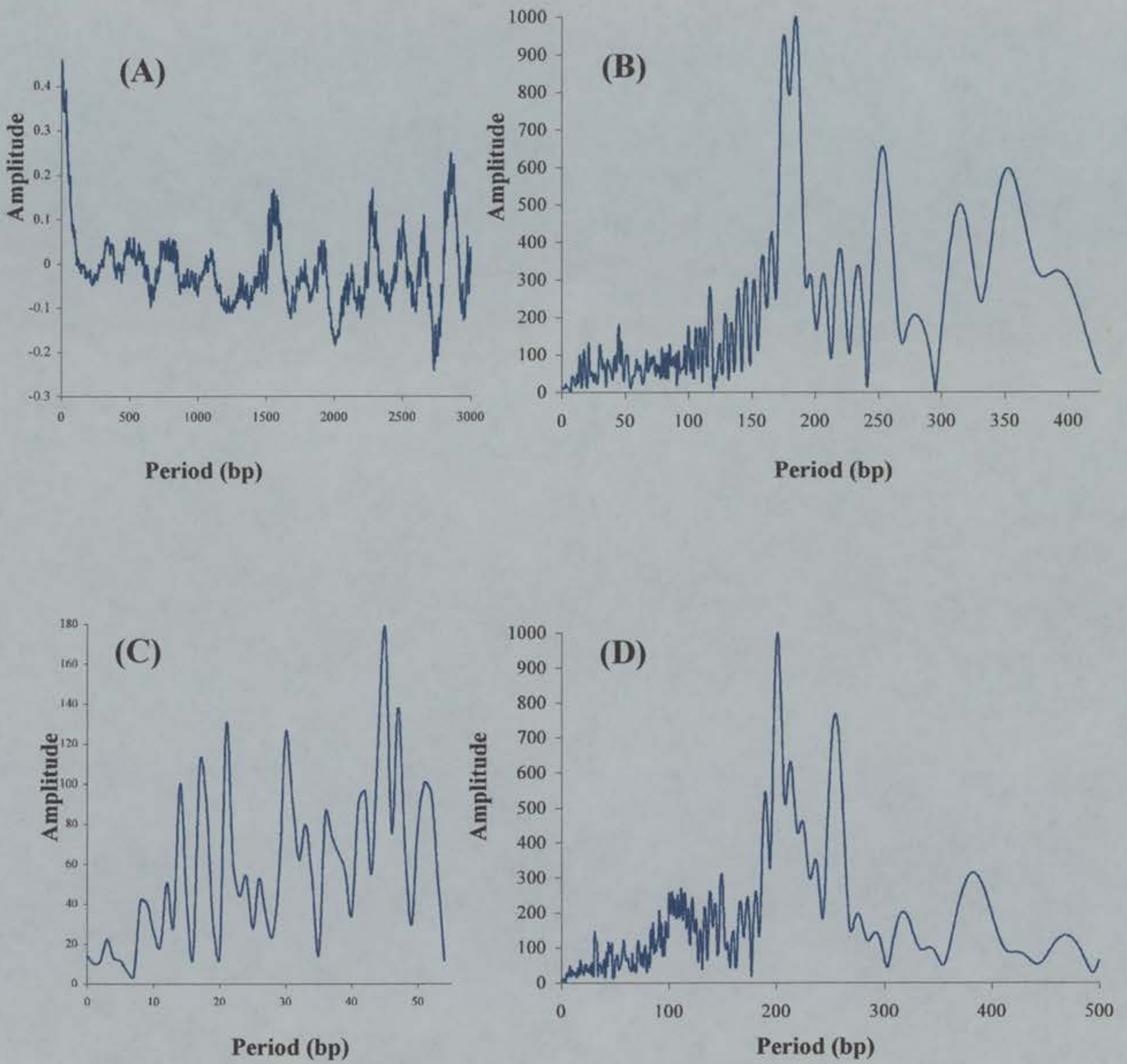


Figure 3.15 Autocorrelation and Fourier analyses of monomer DNA extension data

(A) Autocorrelation analysis of ε-globin. (B) Fourier transformation of the autocorrelation function for ε-globin. (C) An expansion of the 0 to 60 bp region of B. (D) Fourier transform of the autocorrelation function for β^A-globin.

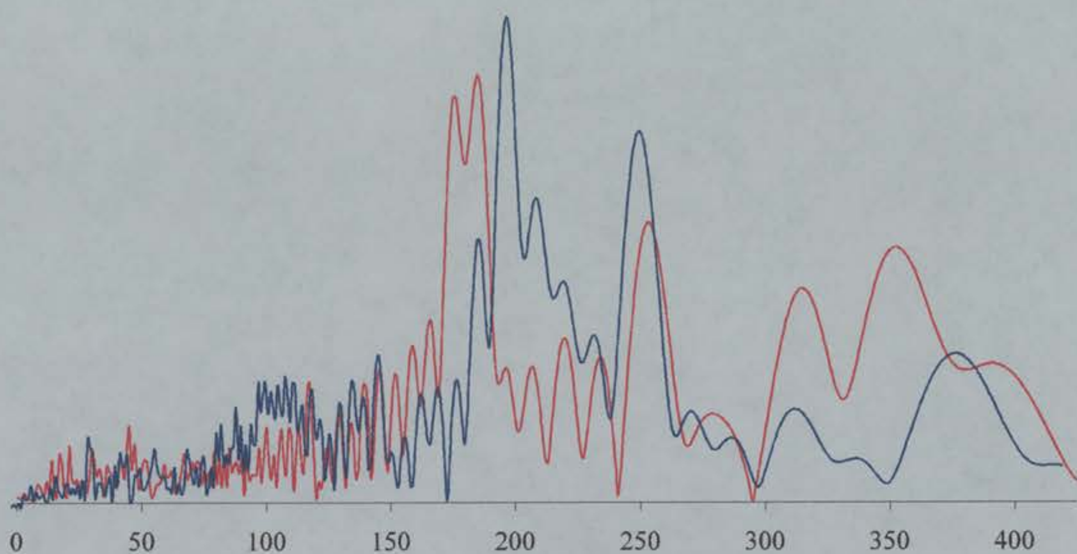


Figure 3.16 Fourier analyses of monomer DNA extension data

The Fourier transformations of the autocorrelation functions for ε-globin (red) and β^A-globin (blue) are overlaid.

By contrast, it has been shown by Fourier analysis that the strongest sites on the β^A-globin gene occur with a 203 bp periodicity (Davey et al., 1995- Figure 3.15D). Fourier transformations of the autocorrelation function for the ε-globin and β^A-globin genes is compared in Figure 3.16. Fourier transformation of the autocorrelation analysis carried out on the combined β^A/ε globin map revealed prominent peaks at 177, 186 and ~200 bp; as would be expected from a combination of the two individual maps (data not shown).

The strong positioning sites on the ε globin gene, which exhibit a periodicity based on multiples of ~ 185 bp, are in register across the whole gene. Positioning sites in the 5' flanking region of ε globin (numbered 17a, 18, 19a or b, and 20a in Figure 3.14b) are in phase with sites in the first intron (21a or b) in the second intron (22, 23a or b, 24a or b) and in the 3′ flanking region (25, 26, 27, 28 and 29). The positioning sites on the β^A globin gene which exhibit the 200 bp periodicity (numbered 1-16 in Figure 3.14b), are also in register across the introns and flanking regions (Davey et al., 1995).

3.3.2.5.3 Short-range periodicities shown by multiple overlapping sites

Inspection of the ε-globin nucleosome positioning map reveals that alternative overlapping positioning sites may be separated by distances of ~45, ~30 and ~20 bp. The Fourier analysis confirms this observation (Figure 3.15C), though it should be noted that the amplitudes of these peaks are weak when viewed in the context of the entire analysis (Figure 3.15B). The Fourier analysis also reveals that these short-range periodicities are much less marked than the long-range periodicities. On the β^A-globin gene, alternative overlapping sites were noted by short-range periodicities of 20 and 40 bp, which again were less prominent than the long-range periodicities (Davey et al., 1995).

The short-range periodicities may arise from a pattern in the DNA sequence and/or from an aspect of the core particle structure (3.4.5).

3.3.2.5.4 Positioning over the epsilon globin promoter

The ε-globin nucleosome positioning map reveals three reasonably prominent alternative nucleosome positions on the ε-globin promoter (Figure 3.14a, sites 20a, 20b and 20c). Site 20a is the strongest of these three sites. The location of these positioning sites is shown, relative to the underlying ε-globin promoter DNA sequence and important transcription factor binding sites, in Figure 3.17.

The dyad (centre) of nucleosome 20a occurs at 4141, so that the CACCC and Sp1 recognition sites occur within thirty base pairs of the dyad. The dyad of nucleosome 20b occurs at 4237; the Sp1 and CCAAT recognition sites occur within 10 and 20 base pairs of the dyad respectively; the GATA and TATA sites within 35 base pairs of the dyad. The dyad of nucleosome 20c is at 4270; the CCAAT site occurs within the terminal 20 base pairs of the nucleosome and the Sp1 site lies 40 base pairs from the dyad; the GATA and TATA sites within 20 base pairs of the centre.

It is known that the location of regulatory elements within the nucleosome can modulate the access of transcription factors to their recognition sites (reviewed in Wolffe et al., 1994). Therefore, the map indicates that information within the DNA sequence, by directing nucleosome positioning, has the potential to regulate access of protein factors to their recognition sites, and thereby may contribute to the regulation of ε-globin gene expression.

3.3.2.5.5 Occurrence of certain DNA sequence patterns within strong nucleosome positioning sites

Statistical analysis of nucleosome positioning sites has provided substantial evidence for the involvement of certain sequence patterns in positioning nucleosomes (Travers

20a 20b 20c

4100 4120 4140 4160 4180 4200 4220 4240 4260 4280 4300

Sp1/ CACCC   CCAAT Sp1   TATA/ GATA

```
4150                         4175
.           Sp1 & chicken EKLF  .
CTTGGCTGCTGCCACACCCCTGACCACCCCCTCCTGCCCTCCAGCACAAAC
            Sp1/ CACCC

4200                         4225
.                            .           Sp1
CCCCACACATCCCAGCCAATCGACCCCACAGCGTGGGTGGGGAGGAGCTG
                CCAAT                    Sp1

4250                         4275
.           GATA 1 & TFIID    .
TCAGCGGTGGATAAAAGCCCCGGGGGTCCGCAGCTCCGCTC
            GATA/ TATA
```

Figure 3.17 Nucleosome positioning over the epsilon globin promoter *in vitro*

Numbering is relative to the beta-adult cap site. The names beneath the sequence represent motifs shown to be important to epsilon expression by mutational analysis. Underlines indicate the extent of the footprint at these motifs in extracts from primitive and definitive red blood cells (Mason et al., 1996). The arrow indicates the transcription start site.

and Klug, 1987; Travers and Muyldermans, 1996; Ioshikhes et al., 1996; Bolshoy et al., 1996). Nucleosome positioning sequences within the β^A and ε-globin genes were searched for DNA sequence elements suggested as positioning signals by previous studies. The motifs considered in this analysis included RRRNNRRR, where R = A or G, and N = A, C, G or T; MMMNNMMM, where M = A or C, (Ulyanov and Stormo, 1995); repeats of $(A/T)_3NN(G/C)_3NN$ (Shrader and Crothers, 1989); CACACA (Widlund et al., 1997); bend site consensus sequences (Wada-kiyama et al., 1996) and oligoadenosine tracts (Mahloogi and Behe, 1997).

Thirty-seven of the most prominent nucleosome positioning sites on the region encompassing the β^A and ε genes were studied (these positioning sites are those labelled in bold in Figure 3.14b, except for 1-7 and 13b, 20c, which were omitted from this analysis because they were investigated in a previous study (1 to 7, Shen, 1997) or because they were less prominent, 13b, 20c). The sequences of these thirty-seven positioning sites, extending 75 bp either side of the nucleosome dyad (denoted as 0 bp), were aligned at their dyad axes. The occurrences and position of a each sequence motif was studied, and the results are summarised in Figures 3.18a, b and c.

Ulyanov and Stormo (1995) suggested that the MMMNNMMM sequence motif (where M = A or C) was preferentially placed 40 bp either side of the dyad axis. The data do not support this for the β-globin sequence (Figure 3.18a).

The motif RRRNNRRR (Figure 3.18a) shows a notable polar distribution, and a slight preference for placement about 35-40 bp 3′ from the dyad. This result does not agree with the suggestion (Ulyanov and Stormo, 1995) that it is placed symmetrically, 15 bp away from the dyad axis. My observations fit better with a previous study of the region upstream of the β^A globin gene, which showed a polar distribution for the sequence ~ 40 bp from the dyad (Shen, 1997).

Shrader et al. (1989) showed that $(A/T)_3NN(G/C)_3NN$ tandem repeats can act as a strong nucleosome positioning signal. This motif was found only in single copies within the ε and β^A globin genes, so it is unlikely to significantly influence positioning on these

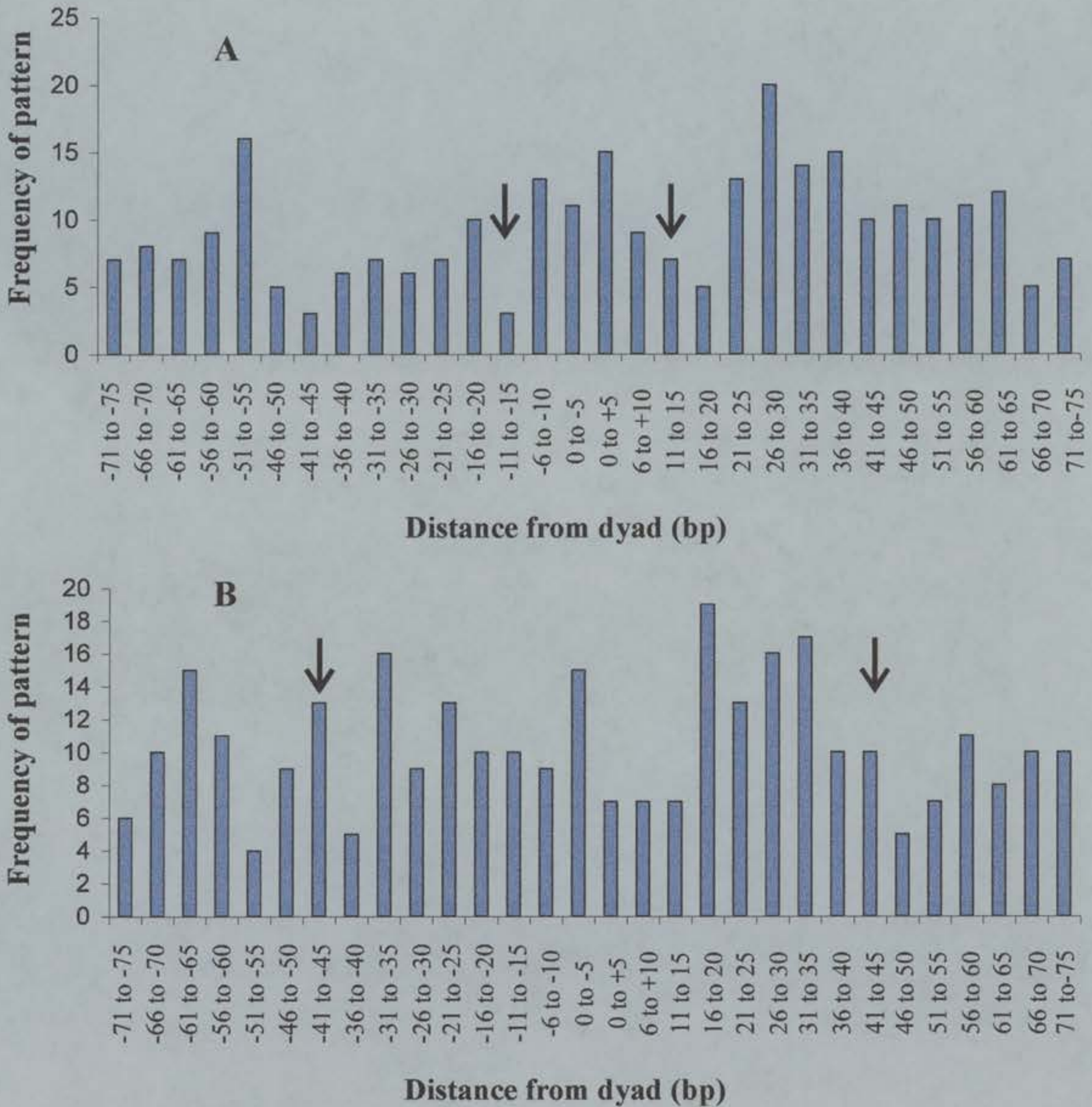


Figure 3.18a Occurrence of the sequence motifs RRRNNRRR and MMMNNMMM within nucleosome positioning sites on the beta-adult and epsilon globin genes

Thirty-seven of the most prominent positioning sites were used for this analysis. The occurrence of the pattern is shown relative to the centre of the nucleosome, which is located at '0 bp'. The data for the RRRNNRRR motif are shown in (A), and those for MMMNNMMM in (B). The arrows indicate the locations where Ulyanov and Stormo (1995) suggested these motifs were preferentially placed.

genes.

It has been proposed that DNA bend sites, which occur at ~700 bp intervals within the human β^A and ε-globin genes, act as nucleosome phasing sequences (Wada-Kiyama et al., 1994; 1995; 1996). Almost all of these sites were accompanied by AT-rich sequences with the consensus $A_2N_8A_2N_8A_2$. Each $A_2N_8A_2N_8A_2$ site, however, does not necessarily create a bend site, so the occurrence of the more stringent bend-site consensus, $A_3N_7A_3N_7A_3$ was also investigated. Six putative bend sites were found within the chicken β^A and ε gene regions (Figure 3.18b) (however, one of these putative bend sites, at 5427 relative to the β^A cap site, is unlikely to be an actual site of bending, because the bend-site sequence is an oligoadenosine tract $(A)_{14}$). The *in vitro* nucleosome mapping analysis reveals that four of the 6 putative bend sites occur within a nucleosome positioning site, but only one of these is a strong site (Figure 3.18b).

Mahloogi and Behe (1997) proposed that the oligoadenosine tracts (15 to 30 contiguous adenosines) found in eukaryotic genomes favour nucleosome formation. One tract containing 14 contiguous adenosines occurred at position 5427 relative to the β^A gene. This site did not occur within a strong nucleosome positioning site *in vitro*.

Widlund et al. (1997) characterised 87 nucleosome positioning sequences *in vivo* in the mouse. They observed that 23 of these were enriched in 5′-CA-3′ (equivalent to 5′-TG-3′) and that these occurred in runs, which were sometimes very long. The β^A and ε globin sequence was searched for $(5'\text{-CA-}3')_3$. The sequence was found six times, five of which occurred within a positioned nucleosome (Figure 3.18c). Four of these positioning sites were also of a relatively high affinity. It is also noticeable that in four out of the five cases, the CACACA motif occurs within 10 bp of the dyad.

3.3.2.5.6 Positioning at exon-intron boundaries

Denisov et al. (1997) used the characteristic ~10 bp periodicity of AA(TT) nucleotides in nucleosomal DNA to predict nucleosome positioning at eukaryotic splice

Figure 3.18b Table showing nucleosome positioning at putative DNA bend sites

The table shows the position of bend sites, predicted on the basis of sequence, relative to the β^A cap site, the position of any prominent nucleosome positioning site identified *in vitro* within 50 bp and the strength of this *in vitro* positioning site. The location of the bend sites is also given relative to the ε cap site, because Wada-Kiyama et al. (1996) observed that bend sites in the mouse and human globin sequences that they studied occupy conserved positions relative to the cap site of the adjacent gene. The schematic below the table shows nucleosome positioning sites identified *in vitro* and it shows which of these occur close to putative bend sites (indicated by arrows), and their position relative to the cap β^A site.

Sequence of putative bend site	Bend site position relative to β^A cap site	Nucleosome positioning site *in vitro* ± 50 bp	Strength of *in vitro* positioning site *	Bend site position relative to ε cap site
$A_2N_8A_2N_8A_2$	1800	1837	+	-2492
$A_2N_8A_2N_8A_2$	2688	-	-	-1594
$A_3N_7A_3N_7A_3$	2868	2866	+ + + +	-1424
$A_2N_8A_2N_8A_2$	5032	5027	+ +	+740
$A_3N_7A_3N_7A_3$	5427	-	-	+1150
$A_2N_8A_2N_8A_2$	6652	6659	+ +	+2360

*+ = 5 to 10 % of strength of strongest positioning site (23b in Figure 3.14b), ++ = 10 to 20 %, + + + = 20 to 30 %, + + + + = 30 to 40 %, + + + + + = 40 to 100 %.

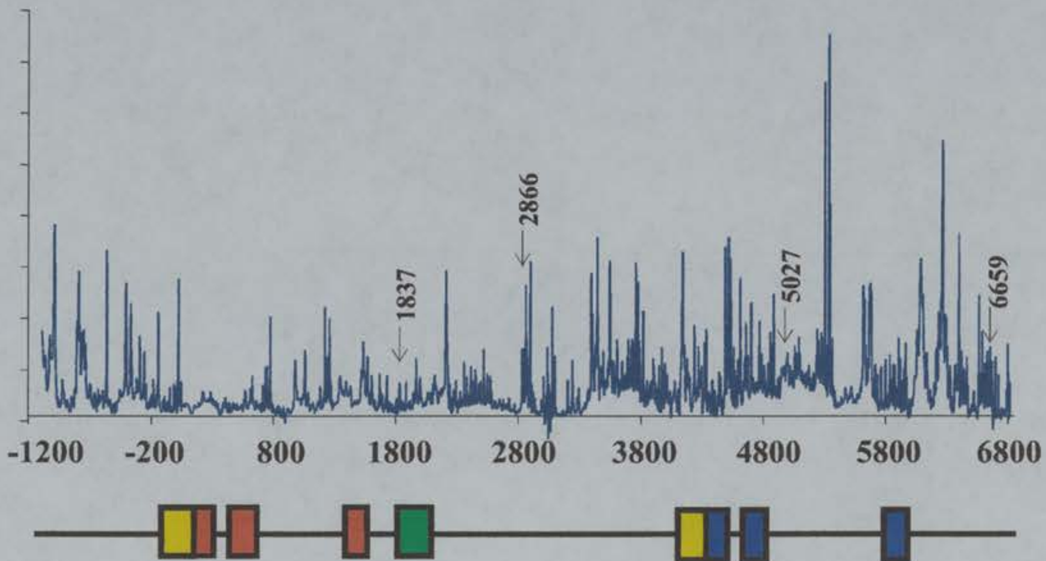
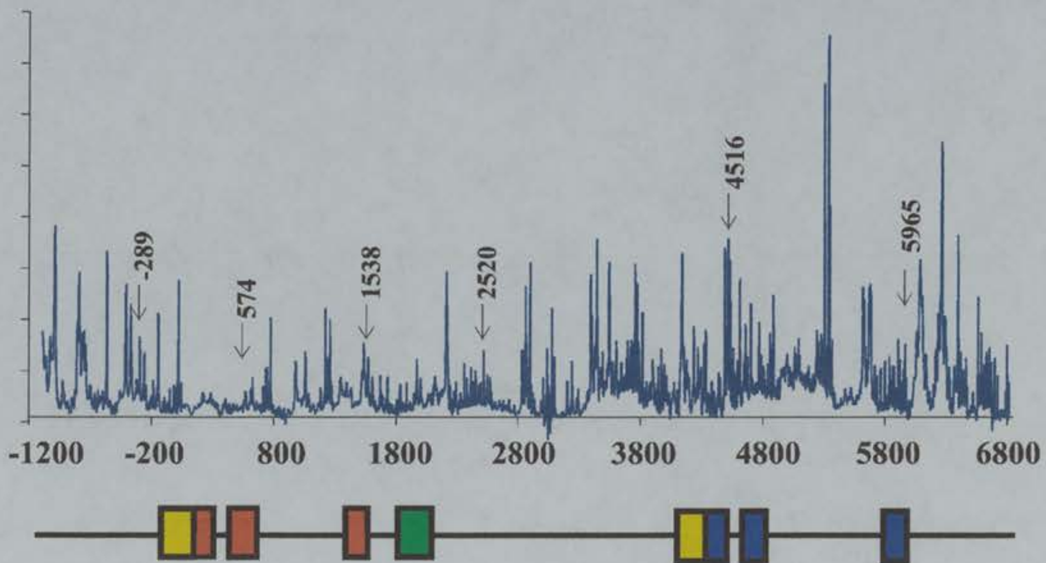


Figure 3.18c Table showing nucleosome positioning at CACACA sequence motifs

The position of the CACACA site is given relative to the to β^A cap site. The closest prominent nucleosome position *in vitro* is given; a number in brackets indicates the number given to this position for the purposes of discussion in the map of the globin genes in Figure 3.14b. The schematic diagram below the table shows nucleosome positioning on the β and ε-globin gene region *in vitro*, and the positioning sites which occur close to CACACA motifs are indicated by arrows.

Position of CACACA relative to β^A cap site	Closest nucleosome positioning site *in vitro*	Strength of *in vitro* positioning site *
-319	-289	+++
+572	+574	+
+1538	+1538 (11)	+ + +
2519	2520 (14)	+ +
4516	4516 (21b)	+ + + +
5964	5965	+ + +

*+ = 5 to 10 % of strength of strongest positioning site (23b in Figure 3.14b), ++ = 10 to 20 %, + + + = 20 to 30 %, + + + + = 30 to 40 %, + + + + + = 40 to 100 %.



sites. They demonstrated that splice junctions which were predicted to occur within nucleosomes were located preferentially within the central 10-15 bp around the centre of the nucleosome.

Nucleosome positions identified on the ε and β^A globin genes *in vitro* at splice sites were compared with the nucleosome position as predicted by Denisov et al. (1997). The data are summarised in Figure 3.19. The first column of the table in this figure gives the position, relative to the β^A cap site, of the AG (at exon-intron junctions) or GT (at intron-exon junctions) which form an invariant part of the splice site consensus. Denisov et al. (1997) suggest that the AG or GT dinucleotide occurs within the central 10-15 bp of the nucleosome. The second column gives the position of the closest prominent nucleosome positioning site identified *in vitro* which is within 40 bp of the AG/GT, and the third column gives the strength of this *in vitro* positioning site. Denisov et al. (1997) proposed that AA and TT dinucleotides in the splice site signal itself could contribute to the 10 bp periodicity of A/T which was used for nucleosome prediction. For example, the splice site consensus at exon-intron boundaries is (C/A)AG<u>GT</u>(A/G)AGT, where GT (underlined) is invariant. This consensus could potentially be GTAAGT. At intron-exon boundaries the consensus is $(T/C)_3$N(C/T)<u>AG</u>(G/A), where AG (underlined) is invariant, so there may be TT dinucleotides upstream of the AG. The final column of the table in Figure 3.19 shows the sequence at and around the splice site consensus, to show if there are any AA or TT dinucleotides which have the potential to contribute to positioning.

Figure 3.16 reveals that, of the eight sites contained within the β^A and ε globin genes, three have a nucleosome dyad positioned within 10 base pairs of the splice junction; one site has a dyad 30 bp distant and the remainder are far from strong positioning sites. This lends some support to the suggestion (Denisov et al., 1997) that splice junctions, when they happen to reside within the nucleosome, have preference for the central region of the nucleosome. Of the three sites where the dyad lies close to the splice junction, two have TT dinucleotides, but the other has a TCT tri-nucleotide. These examples, whilst indicating that the DNA sequence at splice junctions may contribute to nucleosome

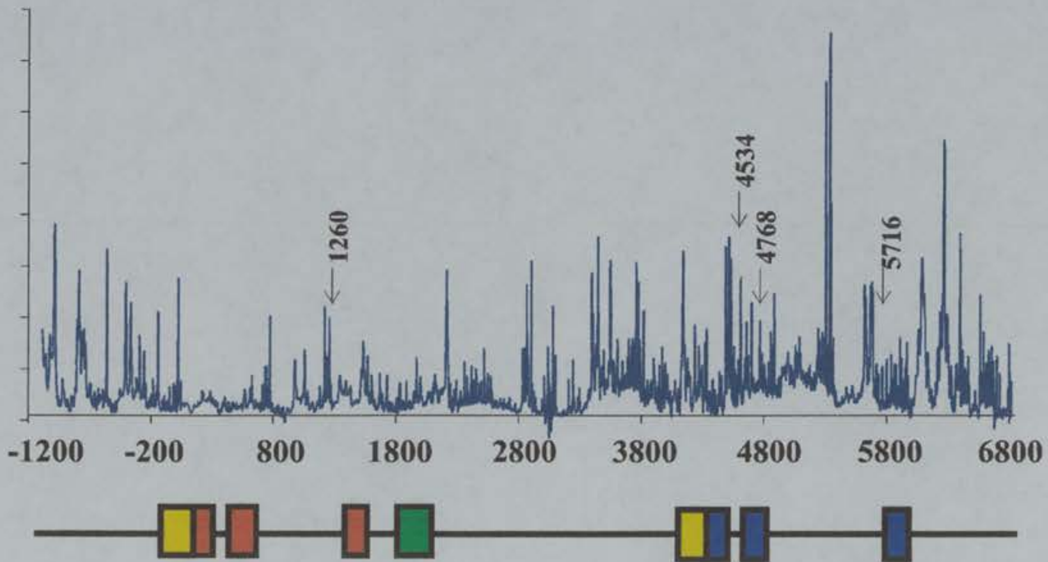
Figure 3.19 Table showing nucleosome positioning at splice sites

The table below gives the position (relative to the β^A cap site) of the AG or GT at the splice site junction, the position of any prominent positioning site *in vitro* within 40 bp of these, the strength of these positioning sites, and the sequence at and around the splice sites. In the schematic the position of the *in vitro* positioning sites which occur within 40 bp of a splice site are indicated on the map of nucleosome placement *in vitro* on the entire β^A and ε-globin gene region.

Junction (position of AG or GT)	Nucleosome positioning site *in vitro* (± 40bp)	Strength of nucleosome positioning site*	AA or TT at splice site?
173	None	-	tggccaggtnggt
263	None	-	atctctctacaggctg
488	None	-	ttcagggtgagat
1296	1260 (10b)	+ + +	ttcccacagctcct
4429	None	-	ccctggccaggtagg
4535	4534	+ + + +	ccctctgcaggc
4760	4768	+ + +	ttcagggtgaga
5731	5716	+ +	tttctcctgcagct

*+ + = 10 to 20 % of strength of strongest positioning site (23b, in Figure 3.14b),

+ + + = 20 to 30 %, + + + + = 30 to 40 % and + + + + + = 40 to 100 %.



positioning, are too few in number to draw firm conclusions.

3.4 DISCUSSION

3.4.1 Production of the nucleosome positioning map for the chicken ε-globin gene by monomer extension

In this study, core histone octamer positions were mapped over a 3.8 kb region encompassing the ε-globin gene. The resulting map, (Figure 3.14a), shows the precise placement and relative strength of nucleosome positioning sequences with respect to the underlying globin DNA sequence. The study was carried out using monomer DNA extension, as this can map nucleosome positions over a long distance, at base pair resolution. Both original and oligonucleotide-based monomer DNA extension techniques were used. The oligonucleotide-based method was developed as part of this project to reduce the amount of subcloning required for mapping. An experiment which demonstrated the reliability and validity of this technique is presented in this chapter.

3.4.2 Features of the ε-globin gene nucleosome positioning map: Comparison with the β^A map

The map for the ε-globin gene adjoins that for the β^A-globin gene (Davey et al., 1995). The data for ε-globin were normalised with respect to those for β^A-globin using common core histone octamer positioning sites located within pBluescript vector sequence, to produce a map covering 7.2 kb of continuous sequence containing the β^A and ε genes and their shared enhancer (Figure 3.14b). Both sections of the map were produced from monomer extension data by the same method of analysis (described in section 3.5.2).

Comparison of the β^A and ε-globin regions of the map reveals some strong similarities. In both cases, many of the possible positioning sites are occupied at low

(essentially background) frequencies, whilst a subset of sites are occupied at much higher frequency, indicating that the latter are much stronger nucleosome positioning sequences. A similar range in the strength of nucleosome positioning sites was seen over the β^A and ε-globin regions; the strength of one of the strongest nucleosome positioning sites in the β^A-globin region (numbered 3 in Figure 3.14b) is ~300 times greater than background, whilst the strongest site in the ε-globin gene (23b in Figure 3.14b) is ~1000 greater than background.

There are some striking similarities, and differences, between the ε and β^A genes in terms of the long and short-range features of the nucleosome positioning map. Firstly, the strongest positioning sites on both genes tend to be absent from exons, and are found more frequently in introns and flanking regions. This supports a role for non-coding DNA sequence in influencing chromatin structure and thereby contributing to regulation of gene expression. Secondly, both gene regions show a notable periodicity in strong positioning sites. On the β^A region, there is a 203 bp periodicity (Davey et al., 1995), whilst on the ε region there is a periodicity based on a 177 and 186 bp repeat. The difference in nucleosome repeat length could create differences in the packaging of these genes into higher-order chromatin structure, which could lead to differences in folding and unfolding of chromatin which in turn could affect gene repression and activation. Thirdly, alternative overlapping positioning sites on the β^A and ε gene regions are related by short-range periodicities, which could reflect short-range repeats in the DNA sequence or the core histone octamer structure. Finally, there is the potential for nucleosome positioning on the ε-globin promoter to contribute to globin gene regulation, as proposed for the nucleosome positioned on the β^A promoter (Buckle et al., 1991).

3.4.3 Distribution of strong nucleosome positioning sites between introns and exons

Strong nucleosome positioning sites may be under-represented in the coding regions

of the ε-globin gene, a feature shared with the β^A-globin gene (Davey et al, 1995). Previous studies have shown that this is not due to an under-representation of the coding regions in the core particles used in monomer extension (Davey et al, 1995). Therefore the absence of strong positioning sites in the exons may denote an incompatibility between sequence requirements for coding and nucleosome positioning for the β^A- and ε-globin genes. This observation may extend to other genes; studies predicting nucleosome positioning by looking for the characteristic ~10 bp periodicity of AA(TT) nucleotides in the DNA sequence, have suggested that strong nucleosome positioning sites may occur less frequently over the central region of exons (Denisov et al., 1997; Long et al., 1995).

The unequal distribution of strong positioning sites between coding and non-coding regions suggests a possible role for introns and flanking regions of genes as organisers of chromatin structure across the whole gene. Whether non-coding regions perform a function, or whether they are just 'junk' sequence, which accumulates accidentally during evolution has long been a contentious issue (reviewed in Holmes, 1993). In the human genome, about 3% of the sequence comprises coding regions and introns, about 2% are regulatory elements which control gene expression, about 15% consists of telomeres or centromeres and about 80% has no known function. Even in the domains which contain active genes only a small percentage of the DNA actually codes for protein. In the chicken β-globin gene cluster, for example, the coding regions represent under 4 % of the ~25 kb domain; the introns ~8 %; the regulatory regions ~ 5 %, and CR1 repeat elements 16% of the total.

Many researchers consider that the non-coding regions of the genome, which constitute such a large proportion of the total sequence, do have important functions to perform. Researchers have for a long time suggested that non-coding regions may play a role as organisers of chromatin structure (Zuckerandl, 1981; 1997). Positioned nucleosomes can influence the positions of surrounding nucleosomes (Thoma, 1992; Lu et al., 1994). Liu et al. (1997) demonstrated that two positioned nucleosomes, 1000 bp

apart, were capable of organising a positioned array of nucleosomes between them. Therefore it seems probable that strong positioning signals in introns and flanking regions could regulate the formation of chromatin structure across genes and possibly whole chromatin domains. In agreement with this proposal, and our observation that strong nucleosome positioning sites are located in the chicken β^A/ε globin intergenic region, Liu et al., (1993) demonstrated that this intergenic region facilitates nucleosome array formation *in vitro*. Further evidence for introns as chromatin organisers is provided by experiments with ovalbumin and rat growth hormone transgenes in mice which showed that removal of the introns led to a disruption in chromatin structure, and, as a possible consequence of this, a decrease in expression levels (Lauderdale and Stein, 1992; Liu et al., 1995). Other researchers argue that the non-coding DNA cannot have an important function as species such as the puffer fish, which has a genome one eighth the size of vertebrates, does not show any ill effects from its lack of 'junk' DNA. However, this species still possesses introns, though they are somewhat shorter (reviewed in Holmes, 1993). An understanding of the function of non-coding DNA will require deeper investigation of the relationship between DNA sequence and nucleosome positioning, and nucleosome positioning and higher order structure.

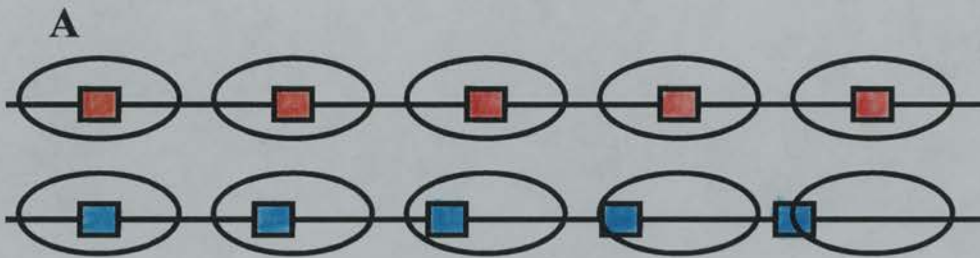
3.4.4 Long-range periodicity of strong nucleosome positioning sites

Strong positioning sites across the ε-globin gene occur with a periodicity based on a 177 and 186 bp repeat (Figure 3.15B). This fits with the data from *in vitro* reconstitution studies which showed that the ε/β^A intergenic region and the 5' end of the epsilon globin gene set up an array of nucleosomes with a ~180 bp spacing over the ε and β^A gene region (Liu et al., 1993). By contrast, there is a 200 bp periodicity of strong positioning sites on the β^A gene (Davey et al., 1995).

Is it possible for these regularly spaced positioning sites to be occupied *in vivo*, so that a regularly spaced array of positioned nucleosomes is set up? This will be dependent

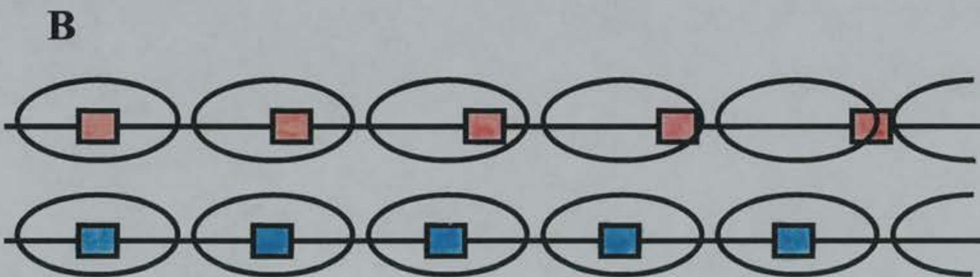
upon the nucleosome density *in vivo*, which is reflected in the nucleosome spacing. For example, if the nucleosome density is such that there are 5 nucleosomes every 1 kb, then nucleosomes will be spaced on average every 200 bp. Therefore, a set of nucleosome positioning sequences spaced at 200 bp intervals could potentially be occupied, but sites at ~ 182 intervals could not be optimally occupied. This is shown in the schematic (A), below, where two stretches of DNA are shown, both 1 kb in length, occupied by nucleosomes (circles) at a density of 5 nucleosomes per 1 kb. The red boxes on one indicate positioning signals at 200 bp intervals and the blue boxes the positioning signals at 182 bp intervals.

A



If the nucleosome density were to increase by only 10 %, so that there were now 5.5 instead of 5 nucleosomes every 1 kb, the nucleosomes would occur at ~ 182 bp intervals (see (B) below).

B



Comparing (A) with (B), it can be seen that this creates a radical alteration in the 'picture' of DNA packaging; regions of the DNA sequence which were previously folded up into nucleosomes are now exposed in linker DNA, and vice versa. It is also clear now, that in B, with the nucleosomes spaced at 182 bp intervals the positioning

sites spaced at 182 bp intervals could in theory be optimally occupied, but the sites spaced at 200 bp intervals could not. Therefore a change of only 10 % in nucleosome density can radically alter the packaging of DNA into chromatin.

It is conceivable that cells modulate nucleosome density in order to alter chromatin structure, possibly in some cases by determining whether sequence-encoded positional information is used effectively or not. It has been suggested previously that the alteration in nucleosome spacing which occurs during erythroid development could facilitate the formation of distinctive higher-order chromatin structures, which in turn, would affect globin gene expression (Liu et al., 1993; Evans et al., 1990).

In this context, the difference between the ε and β^A genes in the spacing of strong positioning sites is intriguing, because the spacing on ε globin *in vitro* reflects the spacing of nucleosomes seen on the β-globin cluster *in vivo* when ε globin is inactivated post 5 days of development, and the spacing on β^A globin *in vitro* reflects the spacing in adult red blood cells and non-erythroid tissues where it is inactive/close to becoming inactive.

Villeponteau et al. (1992) determined the nucleosome spacing in 5-day red blood cells, where ε globin is active, and β^A globin inactive, to be in the range 170-182 bp over the globin genes; in 12-day red blood cells where ε is inactive and β^A active, the spacing is ~ 160 ± 10 bp over the globin genes, and finally in adult red blood cells, where ε- and β^A–globin are both inactive, the spacing is ~ 200 bp. Therefore, the 177 and 186 spacing *in vitro* on the ε gene is close to the spacing *in vivo* early in development, when ε is initially active, and then repressed. The *in vitro* ~200 bp spacing on the β^A gene reflects the spacing in cells where this gene is inactive, both in non-erythroid cells, and in erythroid cells late in development.

One possible interpretation of these results is that the periodic distribution of strong positioning sites on the ε- and β^A-globin genes may facilitate their packaging into an inactive higher order structure at the correct developmental stage. For example, the spacing *in vivo* in 12 day red blood cells where ε globin is inactive, is such that the

strong nucleosome positioning sites on ε-globin could potentially be occupied, so that a regularly spaced array of nucleosomes could facilitate the formation of a stable, regular higher order structure. It has been suggested that only chromatin with a very regular linker length could form a regular higher order structure, with its extensive internucleosomal contacts, which confer great stability (Woodcock, 1994; Ramakrishnan, 1997).

By contrast, the nucleosome spacing in 12 day red blood cells, where β^A is active, is not compatible with the strong nucleosome positioning sites on the β^A-globin gene being occupied. Therefore, it is possible that the higher order structure is less stable, because its formation is not directed by an array of nucleosomes, precisely positioned by information encoded within the DNA sequence. It is known that transcription of β^A and the other globin takes place in the context of the 30 nm fibre, with only transient unfolding of the chromatin taking place (Ericsson et al., 1990; Andersson et al., 1992). Therefore, lower stability of the 30 nm fibre could be an important factor in the transcription of the β^A globin gene in 12 day red blood cells. In adult red blood cells and non-erythroid cells where β^A is not transcribed, the nucleosome spacing is compatible with the strong positioning sites *in vitro* being occupied. Here, there is the potential for a regular positioned nucleosomal array to form, which could facilitate the formation of a stable higher order structure.

The proposal above indicates how the size of the interval between nucleosomes in a regular array may influence gene expression, and may explain why spacing varies during development. The linker length may have the potential to affect gene expression in other ways. Several of the models proposed for the higher order structure of chromatin suggest that linker length may influence the stability of the 30 nm structure. One version of the solenoid model proposes that the structure could not accommodate an array of nucleosomes spaced at 163-176 bp intervals (Butler, 1984). The nature of the higher order structure has not been fully defined, so the impact of different linker lengths on the stability of higher order structure cannot yet be effectively predicted.

The proposal discussed here suggests that positioning signals encoded in the DNA could help determine chromatin structure in some tissues, and be subordinate to other determinants of nucleosome positioning in other tissues.

The role of sequence-directed nucleosome positioning in determining nucleosome spacing *in vivo* is still under debate. In one study, nucleosome repeat length was seen to be the same on randomly selected chicken genomic sequences *in vitro* and *in vivo*, indicating that the DNA sequence is leading to the formation of positioned arrays (Liu and Stein, 1997). Blank and Becker (1996) reconstituted nucleosomes using an embryonic Drosophila extract onto an artificial sequence containing a strong nucleosome positioning signal repeated at 172 and 208 bp intervals. In contrast, they found that the separation of nucleosome positioning signals did not determine nucleosome repeat length; it is subordinate to other, predominantly electrostatic factors. This experiment was criticised on the basis that the extract used was embryonic, (Liu and Stein, 1997). In *Drosophila* embryos, the chromatin structure must allow for fast DNA replication rates, therefore the situation may not be representative of other cells. A comparison of nucleosome placement on genes *in vivo* and *in vitro* may help to resolve these questions, and a study of the kind is presented in the next chapter.

3.4.5 Short-range periodicities relating multiple overlapping positioning sites

This analysis reveals alternative overlapping positioning sites on the ε-globin gene, which tend to be separated by 45, 30 and 20 bp intervals (Figure 3.15C). On the β^A-globin gene, alternate overlapping positions are separated by 20 and 40 bp intervals. These short-range periodicities may arise from preferential placement of DNA sequence patterns at certain positions either side of the core particle dyad, in combination with the symmetrical and repetitive nature of the structure of the core histone octamer itself, or alternatively, from repetition of certain sequence motifs in the DNA.

It is thought that certain DNA sequence patterns may occur preferentially at particular positions within the 146 bp of core particle DNA (Ulyanov and Stormo, 1995). As the core particle exhibits a high degree of symmetry about its dyad axis, sequence patterns may occur preferentially at the same positions on *both* sides of the dyad. Thereby, a single occurrence of a sequence motif could give rise to two overlapping nucleosome positions; one with the sequence motif 5′ to its dyad, and the other with the motif 3′ to its dyad. For example, it was found using a multi-alphabet consensus algorithm that RRRNNRRR (where R=A or G, and N= A,C,G or T) occurs preferentially 15 bp either side of the dyad, where the DNA must be able to accommodate being bent sharply on the nucleosome surface (Ulyanov and Stormo, 1995). The distortion of the DNA at these sites is created by the $\alpha 1 \alpha 1$ DNA binding sites of the two H3:H4 dimers on either side of the dyad, as discussed in chapter one (Luger at al., 1997). A single occurrence of the RRRNNRRR motif could give rise to 2 positioned nucleosomes, separated by 30 bp, one with the motif 5′ to its dyad, and the other with the motif 3′ to its dyad. Positioned nucleosomes at 30 bp intervals within the map were checked for an instance of this, but none were found.

It may be that other sequence elements, as yet unknown, which occur preferentially at positions either side of the dyad, have given rise to these short-range periodicities. The database of nucleosome positioning sequence produced in this study could be searched for sequences such as these. As an alternative approach, it may be possible to identify other repeated motifs within the DNA binding sites in the core histone octamer which could generate short-range periodicities. For example, inspection of the core particle structure at high resolution (Luger et al., 1997) reveals that the histone amino-terminal tails protrude into the minor groove every 20 bp. To create a 20 bp short-range periodicity, such as that seen on the β^A gene, the amino terminal tails of the core histones would have to recognise the same feature in the DNA sequence, but they may be insufficiently homologous to do this. It has been suggested that when nucleosomes are reconstituted by salt dialysis, nucleosome assembly is essentially complete before the

histone tails can associate with the DNA (Blank and Becker, 1996) so this suggests that if they do make a contribution to nucleosome positioning, it may only represent 'fine-tuning'.

It has also been shown that repeated DNA sequence patterns can be responsible for alternative overlapping nucleosome positions (reviewed in Thoma, 1992). Numerous studies have indicated that alternating A/T and G/C di or tri-nucleotides, occurring at ~10 bp intervals, play a role in nucleosome positioning, and can give rise to a set of rotationally positioned overlapping nucleosomes, separated by 10 bp intervals (Satchwell et al., 1986; Bolshoy et al., 1996). The 10 bp periodicity of A/T and G/C facilitates the bending of the DNA around the core histone octamer. This repeated sequence pattern seems unlikely to contribute to the creation of overlapping alternative nucleosome positions on the globin genes, as Fourier analysis on the combined ε/β^A map shows that a 10 bp periodicity of positioning sites is fairly infrequent. This may indicate that the globin DNA has a different, say 20 bp periodicity of A/T and G/C nucleotides, which is contributing to nucleosome positioning. However, analysis of strong positioning sites on 1.5 kb of the beta-adult gene (Chang-Hui Shen, pers. comm.) did not reveal any obvious pattern in the distribution of AA and TT dinucleotides. Therefore, a 10 bp periodicity in A/T and G/C nucleotides may not be a significant influence on nucleosome positioning over the ε- and β^A- globin genes.

3.4.6 Nucleosome positioning over the epsilon globin promoter

Nucleosomes positioned over regulatory elements such as promoters can play a decisive role in regulating gene expression by controlling the access of transcription factors to their recognition sites (reviewed in Wolffe et al., 1994). Positioned nucleosomes can regulate transcription factor access in four ways (reviewed in Thoma, 1992); firstly, they can occlude recognition sites, and restrict transcription factor access; secondly, they can render other sites accessible by leaving them exposed in linker DNA;

thirdly, transcription factor binding can be enhanced when the recognition site is 'presented' on the surface of the nucleosome, and finally, a nucleosome can bring together two positioning sites which are far apart on the DNA so that the transcription factors can bind co-operatively (reviewed in Wolffe, 1994; Thoma, 1992). In the first case, where a nucleosome is positioned over a transcription factor binding site, the site is not necessarily completely inaccessible; accessibility of the site to protein factors depends on its translational and rotational position. Sites occurring within the central 60 bp, or facing the nucleosome core, are generally less accessible (reviewed in Wolffe et al., 1994).

This study reveals that sequence-directed nucleosome positioning has the potential to regulate access of transcription factors to their recognition sites on the ε-globin promoter. Figure 3.17 shows the three prominent alternative nucleosome positions (20a, 20b and 20c) identified on the ε-globin gene promoter, and the placement of these nucleosomes relative to the GATA, TATA, Sp1, CCAAT and CACCC transcription factor binding sites. These sites have been shown to be important in regulating ε-globin gene expression. GATA-1 and TFIID bind the non-canonical TATA box. GATA-1 binding may recruit TFIID to the TATA box, and in this way confer erythroid specificity on formation of the transcription complex at globin promoters (Mason et al., 1996; Fong and Emerson, 1992). The Sp1 site, where the protein footprint extends from 4226 to 4240, is bound by the ubiquitous transcription factor Sp1. Deletion of the CCAAT site (centred at ~ 4218) reduces promoter activity, but no protein footprints have been obtained over this region (Mason et al., 1996). The Sp1/CACCC site, where the footprint extends from ~ 4150 to ~ 4177, binds Sp1 and a chicken homologue of EKLF, (Mason et al., 1996). EKLF is an erythroid-specific transcription factor expressed at comparable levels in primitive and definitive erythroid cells in mice (reviewed in Orkin, 1995). Although expressed at all developmental stages, it may contribute in the switch from foetal to adult globin expression in humans (Orkin, 1995). Transient transfection assays have shown that the GATA/TATA, Sp1 and CCAAT motifs contribute to promoter

activity, and the CACCC/Sp1 site enhances transcription in the presence of an enhancer (Mason et al., 1996).

If nucleosome 20a occupies the epsilon promoter, the CACCC/Sp1 recognition sites lie within the central, less accessible region of the nucleosome, whilst the remaining transcription factor binding sites are exposed in the linker DNA. Alternatively, if nucleosome 20b were positioned over the epsilon promoter, then the Sp1 (4226-4240), CCAAT (centred at 4218), and GATA /TATA sites would occur within its central region, whilst the CACCC/Sp1 sites (4150-4177) would occur outside, but close to, the boundary of this nucleosome. If the promoter were occupied by nucleosome 20c, the Sp1, GATA and TATA sites would lie within the central region of the nucleosome, and the CCAAT site would occur in its more accessible terminal 20 base pairs.

The sequence-directed nucleosome positioning could contribute to the repression of ε-globin in non-erythroid cells. Nucleosomes 20b and 20c occupy the main transcription factor binding sites, including the GATA site which may help to confer erythroid specificity on the TATA box. The fact that the GATA site is unoccupied in non-erythroid cells could favour the formation of a positioned nucleosome on the ε-globin promoter, which would then occlude the binding sites of transcription factors like Sp1 which are present in non-erythroid cells. Many other factors, such as hypoacetylation of the chromatin, hypermethylation of the DNA, and a tightly packaged chromatin domain, are associated with repression in non-erythroid cells, but a positioned nucleosome on the promoter may still make an important contribution.

Sequence-directed positioning also has the potential to contribute to the switch from ε- to β^A-globin expression which happens at around day 8 in development. There is strong evidence to suggest that the temporal control of the β^A- and ε-globin genes is effected by a competition between these genes for interaction with their shared enhancer (Choi and Engel, 1988; Foley and Engel, 1992). The relative stability of the ε-globin promoter:enhancer complex and the β^A-globin promoter:enhancer complex decides whether β^A- or ε-globin competes most effectively for the enhancer. This is determined

by the strength of the interaction between transcription factors bound at the enhancer and the promoter, so a positioned nucleosome which regulates the access of these transcription factors to the DNA could potentially play a role in globin gene switching. It has been suggested that many of the promoter transcription factor binding sites are capable of mediating interaction with the enhancer, and do so in a redundant manner. In the intact promoter it is proposed that the CACCC/Sp1 site is preferentially involved in the interaction with the enhancer. It has been suggested that the chicken EKLF-like protein at the ε-globin promoter interacts with GATA-1 at the enhancer (Mason et al., 1996). Therefore nucleosome 20a, which is positioned over the CACCC/Sp1 site could play a role in reducing the stability of the promoter: enhancer interaction, and thereby contribute to switching. Nucleosomes 20b and 20c cover the proximal sites on the promoter, which may also contribute to the stability of the ε-globin promoter:enhancer interaction. The precise positioning of nucleosome 20b with the CACCC/Sp1 site at its boundary, or of nucleosome 20c, with the CCAAT site at its boundary may have relevance for the removal of the nucleosome prior to activation of ε-globin. It has been proposed that in some cases nucleosomes may be removed from regulatory elements by the co-operative binding of multiple transcription factors (reviewed in Felsenfeld, 1992). The binding of a transcription factor at an element located in the more accessible boundary could therefore provide a first step in the removal or rearrangement of the nucleosome. This is speculative, as the interaction of these particular transcription factors with nucleosomes have not been studied.

There is evidence which suggests that there is a decrease in the affect of the enhancer on ε-globin gene in definitive red blood cells, where ε-globin is no longer expressed. This may result from a decrease in the levels of Sp1 and GATA1 in definitive cells (Minie et al., 1992). The reduced occupation of these sites on the ε-globin globin promoter and/or on the enhancer, may lower the stability of the ε-globin promoter:enhancer complex in definitive cells, and thereby contributes to ε/β^A switching (Mason et al., 1996). The reduced occupancy of the GATA1 and Sp1 sites could

facilitate nucleosome formation over these sites, thereby possibly excluding the remaining transcription factors from their sites. Conversely, the presence of GATA1 and Sp1 in primitive cells, where ε-globin is expressed, could favour the removal of a repressive nucleosomal architecture from the ε-globin promoter.

Therefore, a nucleosome positioned on the epsilon globin promoter has the potential to contribute to the regulation of ε/β^A switching. However, the contribution of chromatin structure has not yet been fully evaluated, as many of the studies have involved transient transfection, where the mechanism of switching was studied in the absence of a proper chromatin structure. These studies suggested that increased stability of the β^A promoter:enhancer, due to the binding of the stage-specific proteins to the β^A-globin promoter is the decisive factor in ε/β^A switching (Foley and Engel, 1992). These factors, NFE-4 and BCTF which are specific to the definitive red blood cells where β^A-globin is expressed, bind a stage selector element in the β^A promoter and increase the stability of its interaction with the enhancer. This does not rule out a role for chromatin structure in the temporal regulation of ε- and β^A-globin expression; Foley and Engel (1992) suggested that chromatin structure, which was not considered in their study, could make a contribution. In agreement with this, it was proposed that the precisely positioned nucleosome on the β^A- or ε-globin promoter which occludes most of the important transcription factor binding sites could be displaced by the binding of NFE-4 (Buckle et al., 1991). Therefore, there is a precedent for a positioned nucleosome being involved in the switch from ε to β^A expression. As discussed above, there is some evidence that changes in transcription factor (Sp1 and GATA1) binding at the enhancer and ε-globin promoter may affect the stability of the ε-globin promoter:enhancer complex, which in turn could contribute to β^A/ε switching (Minie et al., 1992; Mason et al., 1996). This means that a positioned nucleosome on the ε-globin promoter has the scope to play a significant role.

3.4.7 DNA sequence patterns within strong nucleosome positioning sites

Statistical analysis of nucleosome positioning sites has provided substantial evidence for the involvement of certain sequence patterns in positioning nucleosomes (Travers and Klug, 1987; Travers and Muyldermans, 1996; Ioshikhes et al., 1996; Bolshoy et al., 1996). Therefore nucleosome positioning sequences within the beta-adult and epsilon globin genes were searched for a number of proposed positioning signals.

This study does not support the suggestion by Ulyanov and Stormo (1995) that the MMMNNMMM sequence motif (where M = A or C) was preferentially placed 40 bp either side of the dyad axis (Figure 3.18a).

The motif RRRNNRRR, where R = A or G, shows a polar distribution, and a slight preference for placement 35-40 bp from one side of the dyad (Figure 3.18a); this fits with one previous study (Shen, 1997) but does not agree with Ulyanov and Stormo (1995), who placed it 15 bp either side of the dyad axis.

The $(A/T)_3NN(G/C)_3NN$ sequence occurs in single copies, at too low a frequency to contribute to nucleosome positioning.

It has been proposed that DNA bend sites, which occur at ~ 700 bp intervals within the human β^A and ε-globin genes, have a high affinity for nucleosomes (Wada-Kiyama et al., 1994; 1995; 1996). In the present study, six putative bend sites were found in the chicken β^A and ε-globin genes, and four of these occurred within positioned nucleosomes *in vitro*. Of these four , three are relatively weak, and only one is a strong prominent site (Figure 3.18b). Therefore, from this small sample, it can be tentatively concluded that DNA bend sites make only a small contribution to sequence-directed nucleosome positioning on the globin genes.

On the mouse and human globin genes the DNA bend sites were found at ~ 700 bp, +/- 250 bp intervals (Wada-Kiyama et al., 1996). The putative bend sites found on the

These sites do not display a ~700 bp periodicity, but all of the bend-site spacings (apart from 888) were multiples of ~ 200 bp.

Wada-kiyama et al., (1996) also observed that the positions of bend sites are conserved upstream of a variety of mouse and human β-globin genes, despite the lack of sequence conservation. If these bend-sites were also conserved in chicken, it would suggest that they played a fundamental, conserved role in determining the chromatin structure of the globin genes. Of the bend sites observed at ~ -1600, -1200 and ~ -300 in human, only the first corresponded to a site in chicken.

To draw firm conclusions about the contribution of DNA bend sites to nucleosome positioning, an experimental mapping of the chicken globin genes by the circular permutation assay would be necessary to define actual bend sites, as the search for a bend-site consensus may overlook some sites, and include other sites which are not actually bent in solution (Wada-Kiyama et al., 1995).

Widlund et al. (1997) observed that mouse nucleosome positioning sequences were enriched in runs of 5′-CA-3′. The data produced in this study provides some support for this, though the number of examples is small; 5′-CACACA-3′ occurred 6 times, and 5 of these times it occurred within a nucleosome positioning site; on four of these occasions, it occurred within 10 bp of the dyad (Figure 3.18c).

Mahloogi and Behe (1997) proposed that the oligoadenosine tracts (15 to 30 contiguous adenosines) found in eukaryotic genomes favour nucleosome formation. They demonstrated that *in vitro*, the free energy of nucleosome formation on oligoadenosine tracts is ~ 1 kcal/mol higher at 60 °C than on heterologous sequence DNA. They suggest that the high temperature is thought to favour nucleosome formation by disrupting the spine of hydration in the minor groove of the oligoadenosine tract, and enhancing its flexibility. They argue that factors other than non-physiological high temperatures, such as the concentrated macromolecular solution found *in vivo*, could also increase flexibility in the same way, and thereby favour nucleosome formation on oligoadenosine tracts. They also suggest that shorter oligoadenosine tracts which are

found ~ every 20 kb *in vivo* could have a smaller, but still significant role *in vivo*. The oligoadenosine tract found within the β-globin cluster was not located in a strong nucleosome positioning site *in vitro*. This does not rule out the possibility that the tract could be folded up in a nucleosome *in vivo*, as the reconstitution conditions used *in vitro* may not alter the flexibility of the tract if the mechanism proposed by Mahloogi et al. (1997) is correct.

Overall, therefore, there is little evidence to suggest that most of the motifs investigated play a substantial role in determining nucleosome positioning on the β-globin genes epsilon and beta-adult. A few, such as CACACA runs, seem to make a contribution, but from the small number of examples found.

There are however other sequences not analysed in this study which could contribute to nucleosome positioning. These include several sequences which have the strongest evidence to suggest that they are determinants of nucleosome positioning; namely, the ~10.4 bp periodicity of dinucleotides AA and TT seen by Ioshikhes et al. (1996), or the 10.2 periodicity of alternating A/T and G/C di or tri-nucleotides described by Satchwell et al. (1986). However, these motifs occur at a high frequency, and the manual methods used in this analysis are not therefore suitable, because analysing data from a sufficiently large number of nucleosomes would be unfeasible. These motifs await future analysis by techniques, such as multiple alignment algorithms, hidden Markov models, or other computational methods.

The data produced in this study could be searched for novel sequence patterns which contribute to nucleosome positioning. In future analyses, the nucleosome positioning sequences compiled will provide a superior database for determining the sequence patterns which contribute to nucleosome positioning. This database is larger than those currently in the literature (Ioshikhes et al., 1993), the nucleosome positions are mapped at high resolution, and the relative binding affinities are known, so that a 'weighting' can be applied for the patterns found in stronger positioning sites.

3.4.8 Positioning of nucleosomes at splice-sites

The three nucleosomes positioned at splice sites within the ε- and β^A-globin genes (Figure 3.16) give some credence to the suggestion that when splice junctions are incorporated into nucleosomes, they preferentially occur close to its dyad axis (Denisov et al., 1997) It was proposed that inclusion in a nucleosome would reduce DNA damage to the critical splice junctions. In addition, they propose that the AA and TT dinucleotides at the splice site may contribute to positioning; two of the three splice sites which occur close to the dyad of a positioned nucleosome contain TT dinucleotides, and the other has a TCT nucleotide (Figure 3.16). It is interesting to note that all three examples occur on the ε globin gene.

These examples, whilst indicating that the DNA sequence at splice junctions may contribute to nucleosome positioning, are too few to draw firm conclusions. Furthermore, it could be argued that the presence of a splice site would be insufficient to decide the position of the nucleosome dyad, because in general, the additive contribution of multiple DNA sequence patterns, rather than a single sequence pattern like the splice site consensus, is thought to determine nucleosome positioning (reviewed in Thoma, 1992). In some cases, however, a small number of nucleotides within the 146 bp core particle DNA may make a decisive contribution to nucleosome positioning. For example, it was possible to abolish nucleosome formation at a strong nucleosome positioning site identified *in vitro*, upstream of the chicken β^A-globin gene, by methylation of a $(CpG)_3$ motif within this nucleosome. Nucleosome exclusion probably involves methylation-induced changes in DNA structure at and around the (CpG) motif which reduces the flexibility of the sequence, so that the energy of nucleosome formation becomes prohibitive (Davey et al., 1997).

CHAPTER 4: MAPPING NUCLEOSOME POSITIONS ON THE EPSILON AND BETA-ADULT GLOBIN GENES OF CHICKEN *IN VIVO*

4.1 Introduction

Chromatin structure makes a decisive contribution to the regulation of gene expression (reviewed in Grunstein, 1990; Wolffe et al., 1994; Wolffe and Hayes, 1999). The β-globin genes in chicken, and other vertebrates, are part of a small gene family which is clustered in a ~30 kb domain (Stalder et al., 1980; Hebbes et al., 1994); their expression is restricted to erythroid cells, and the individual genes exhibit a stage-specific pattern of expression during development (reviewed in Orkin et al., 1995). There is strong evidence that chromatin structure makes an important contribution to this temporal and tissue-specific regulation (Larsen and Weintraub, 1982; Felsenfeld, 1993; Orkin et al., 1995). Nucleosome positioning, specifically, may play a role; positioned nucleosomes on the β^A promoter and its enhancer may be involved in inactivating the gene in non-erythroid tissues, and in the switch in expression from ε to β^A which occurs around day eight in development (Emerson et al., 1985b; Buckle et al., 1991).

It has been demonstrated *in vitro* (this study, chapter 3) that the DNA sequence of the ε-globin promoter has the capacity to precisely position nucleosomes with respect to transcription factor binding sites shown to be important to ε-globin expression; if nucleosomes were positioned at these sites *in vivo*, they would have the potential to contribute to both the switch from ε to β^A-globin expression, and tissue-specific regulation. In addition, *in vitro* studies on nucleosome positioning have shown that a long-range periodicity in the placement of nucleosomes could potentially influence the packaging of the globin genes into an inactive higher order structure, and thereby affect globin gene regulation.

The purpose of this study was to determine nucleosome placement on the ε and β^A -globin genes in adult red blood cells and in brain to investigate the extent to which the DNA sequence directs nucleosome positioning *in vivo* where numerous other factors can modulate the recognition or interpretation of this sequence-encoded positioning information. The intention was to determine the degree to which nucleosome positioning, and specifically sequence-directed positioning, contributes to globin gene regulation *in vivo*.

4.2 Materials and Methods

Methods were adapted from Sambrook et al. (1989), unless otherwise stated. Chemical reagents were obtained from BDH, unless otherwise stated.

4.2.1 Materials

For preparation of nuclei:-

Buffer A: This contains 0.25 M sucrose, 6 mM $MgCl_2$, 50 mM Tris-Cl pH 7.5, 0.5 mM EGTA and 0.2 mM PMSF.

PBS buffer: Phosphate buffered saline (PBS), containing 0.01 M phosphate buffer, pH 7.4, 0.0027 M potassium chloride and 0.137 M sodium chloride, was obtained from Sigma.

For digestion of nuclei:-

a) Cleavage with cuprous-phenanthroline (Quivy and Becker, 1996).

1,10 Phenanthroline: 1,10 Phenanthroline (Sigma) was dissolved at a final concentration of 40 mM in ethanol and stored at -20°C.

NBII Buffer for digestion with cuprous-phenanthroline: NBII buffer, in which cuprous-phenanthroline cleavage of chromatin is carried out, contains 60 mM KCl, 15 mM NaCl pH 8.0, 15 mM Tris-Cl pH 7.4, 0.3 M sucrose, 0.5 mM spermidine and 0.15 mM spermine in distilled water. The buffer is filter-sterilised, and 0.1 mM PMSF added.

Neocuproine: Neocuproine (Sigma) was dissolved at a final concentration of 28 mM in ethanol and stored at -20°C.

Stop Mix I : Stop mix I, for treatment of nuclei following cuprous-phenanthroline cleavage, contained 20 mM Tris-Cl pH 8.0, 20 mM NaCl and 1% SDS dissolved in distilled water. The solution was filter-sterilised. 600 μg/ml of proteinase K was added just prior to use.

Stop Mix II: Stop mix II, for treatment of nuclei following cuprous-phenanthroline cleavage, was prepared by dissolving NaCl at a concentration of 150 mM in distilled water and adding EDTA pH 8.0 to a final concentration of 5 mM.

b) Cleavage with Methidium Propyl EDTA (MPE) (Cartwright and Elgin, 1984).

MPE: MPE was dissolved in distilled water at a final concentration of 1mM and stored at -20°C.

MPE Buffer: MPE buffer, in which cleavage of chromatin by MPE is carried out, contains 15 mM Tris-Cl pH 7.4, 15 mM NaCl, 60 mM KCl, 1 mM EDTA, 0.25 M sucrose, 0.15 mM spermine and 0.5 mM spermidine in distilled water. The buffer was filter-sterilised before use.

c) Digestion with Micrococcal Nuclease (MNase) (Yenidunya, 1992; adapted from Nedospasov and Georgiev, 1980).

MNase: from Worthington was used for preparation of core particles.

Micrococcal Nuclease Digestion Buffer: This contains 0.2 M sucrose, 5 mM Tris-Cl pH 7.4, 80 mM NaCl and 0.2 mM PMSF in distilled water and filter-sterilised.

1mM $CaCl_2$ was added just prior to digestion.

Micrococcal Nuclease Termination Buffer: This contains 5 mM Tris-Cl pH 7.6, 10 mM EDTA, 80 mM NaCl and 0.2 mM PMSF in distilled water and filter-sterilised.

For Southern blotting and filter hybridisation:-

Protocols were adapted from Ausubel et al., 1997.

Nylon filter papers for southern blotting: Hybond N (Amersham) and zeta probe (Biorad) were used for southern blotting.

OLB buffer: OLB labelling was adapted from Sambrook et al. (1989). OLB buffer is prepared by mixing 100 μl of solution A (1.25 M Tris-Cl pH 7.5, 125 mM $MgCl_2$ and 1.8 μl of β-mercaptoethanol) with 250 μl of solution B (2 M Hepes, pH 6.6) and 150 μl of solution C (Hexadeoxyribonucleotides (Pharmacia) at 4.5 mg/ml). The OLB mix was stored at -20°C.

TPE/Ficoll loading buffer: Samples which were to be run on TPE gels were resuspended in 1 x TPE/Ficoll loading buffer. 5 x TPE loading buffer contains 12.5% Ficoll (400), 0.25% bromophenol blue and 25 mM EDTA in 5 x TPE buffer. It was left on a roller overnight to dissolve and sonicated before use.

4.2.2 Probes used for indirect end-labelling

The probes used for indirect end labelling were produced by PCR, as described in section 4.2.9.1. The positions of the 6 probes and 9 restriction enzymes used for mapping, relative to the beta-adult and epsilon genes, are shown in Figure 4.1. The Dra probe, which was shorter than the rest was ligated to itself (section 4.2.9.2) to increase signal. The names, sizes and exact positions of the probes, relative to the beta-adult cap site, are given below.

1. Eco probe: This extends from -1054 to -739 (315 bp)

2. Hind probe: This extends from 869 to 1257 (388bp).

3. Dra probe: This extends from 2484 to 2623 (139 bp), and is self-ligated.

4. Bam probe: This extends from 3351 to 3717 (366 bp).

5. Eag probe: This extends from 4828 to 5114 (286 bp).

6. Rsa probe: This extends from 5909 to 6193 (284 bp).

4.2.3 Preparation of nuclei from chicken erythrocytes

Blood (10-15 ml), was collected through two layers of surgical gauze into an equal volume of ice-cold PBS (phosphate buffered saline) containing 0.5 mM EGTA, 0.2 mM PMSF and 25 units/ml of heparin (Sigma). The blood was made up to 30 ml with PBS, and centrifuged at 1200 x g for 3 minutes. The supernatant and 'buffy coat' (white blood cells) were removed using a vacuum line. The pellet was

resuspended in 30 ml of PBS and the spin repeated. After removal of the supernatant, the pellet was made up to 5 ml with buffer A. The blood was added dropwise to 35 ml (7 volumes) of ice-cold buffer A containing 1% v/v Triton X-100. The blood was mixed on a roller at 4°C for 15 minutes. Following centrifugation at 1000 x g for 4 minutes and removal of the supernatant, the nuclei were resuspended in 30 ml of buffer A and spun again (twice). The pellet was resuspended in 4 ml of buffer A and kept at 4°C. Cleavage reactions were carried out as soon as possible on the fresh nuclei.

Alternatively, if the nuclei were being prepared for storage, the pellets were mixed with 1 volume of sterile glycerol and stored at -70°C.

4.2.4 Preparation of nuclei from brain

Brain tissue from adult chickens was stored from fresh in liquid nitrogen until preparation of nuclei could be carried out. The brain tissue, kept frozen with additional liquid nitrogen, was ground up in a pestle and mortar. The ground tissue was gradually made up to 30 ml with PBS. The brain tissue was homogenised with two strokes in a Dounce homogeniser and filtered through autoclaved muslin. The homogenate was then centrifuged at 1000 x g for three minutes at 4 °C (all subsequent centrifugations were carried out at this speed and temperature). Following this (and all subsequent centrifugations) the supernatant was removed using a vacuum line. The pellet was resuspended in 30 ml of PBS and the spin repeated. The pellet was resuspended in PBS and spun again. The pellet was then resuspended in 30 ml of buffer A and centrifuged again. The pellet was then resuspended in 30 ml of buffer A / 0.2% Triton X-100 to lyse the brain cells. This wash in buffer A / Triton was repeated. The pellet was then washed three times in 30 ml of buffer A. The pellet was finally resuspended in 4 ml of buffer A, stored at 4°C, and cleavage reactions carried out as quickly as possible on the fresh nuclei.

4.2.5 Determining concentration of nuclear DNA in preparation of nuclei

5 μl of nuclei were made up to 100 μl with NB/C buffer (10 mM Tris pH 7.5, 85 mM KCl, 5.5 % sucrose, 5 mM spermidine, 1.5 mM $CaCl_2$ and 250 μM PMSF) and 20 U of DNase I, and incubated for 10 minutes at 37°C. After adding 400 μl of sonication buffer (10 mM phosphate buffer, pH 6.8 (Sambrook et al., 1989), 5 M urea, 2 M NaCl), nuclei concentration was determined by spectrophotometry, assuming that 50 μg/ml of dsDNA has an OD of 1 at 260 nm.

4.2.6 Cleavage of chromatin in nuclei with DNA cleavage reagents

4.2.6.1 Cleavage of chromatin in nuclei by cuprous phenanthroline

This protocol was adapted from Quivy and Becker (1996). Nuclei in buffer A were centrifuged at 1000 x g for 30 seconds, and the pellet was resuspended in NBII buffer. The nuclei were spun again, and resuspended in NBII buffer. 250 μg of nuclei were made up to 0.5 ml with NBII buffer, and incubated in a water bath at 22°C for 2-3 minutes.

The cuprous-phenanthroline (OP-Cu) complex was made by mixing 0.5 ml of 40 mM 1,10-phenanthroline with 0.5 ml of 9 mM copper (II) sulphate. After incubation at room temperature for 1 minute the complex was diluted to 20 ml with water. 50 μl of the complex was added to the nuclei, and the reaction initiated by addition of 50 μl of 58 mM 3-mercaptopropionic acid (Sigma) and incubated at 22°C for the required time (generally 5-8 minutes).

The reaction was quenched by adding 1/12 volume of 28 mM neocuproine and leaving the nuclei on ice for 5 minutes. The nuclei were spun at 1000 x g for 1 minute and the supernatant removed. The pellet was resuspended in 100 μl of stop mix I, containing proteinase K, and 50 μl of stop mix II, and incubated at 37 °C for 3 hours. RNase A was added to a final concentration of 100 μg/ml, and incubated for

30 minutes. The DNA was deproteinated by phenol/chloroform extraction, followed by ethanol precipitation. The DNA was resuspended in TE.

Control digests on naked genomic DNA were carried out in the same way, with the following modifications; the reaction was carried out at a DNA concentration of 600 μg/ml, and only 30 μl of the complex was added per ml of reaction; after quenching with neocuproine, the DNA was purified by phenol/chloroform extraction, followed by ethanol precipitation and resuspension in TE.

4.2.6.2 Cleavage of nuclear chromatin with methidium propyl EDTA (MPE)

This protocol was adapted from Cartwright and Elgin (1984). Nuclei in buffer A were centrifuged at 1000 x g for 30 seconds, and the pellet was resuspended in 1 ml of MPE cleavage buffer. The nuclei were spun again, and resuspended in MPE cleavage buffer. After determining the concentration of nuclei, nuclei eqquivalent to 250 μg of DNA were made up to 0.5 ml with MPE cleavage buffer, and incubated at 25°C for 2-3 minutes.

The MPE complex was prepared by mixing equal volumes of 1 mM MPE and freshly prepared 1 mM ferrous ammonium sulphate, diluting ten-fold with MPE cleavage buffer and adding DTT to a final concentration of 2 mM. 500 μl of MPE complex was added to the 250 μg of nuclei and incubated at 25°C for the required time (10-14 minutes)

The reaction was quenched by adding bathophenanthroline (Sigma) to a final concentration of 5 mM. The reactions were supplemented with 3 volumes of a buffer containing 25 mM EDTA, 20% SDS, and 266 μg/ml of freshly added proteinase K and incubated for 3 hours at 37°C. The samples were phenol/chloroform extracted, ethanol precipitated, resuspended in TE and digested with 100 μg/ml of RNase A for 30 minutes at 37°C.

4.2.6.3 Cleavage of chromatin in nuclei with micrococcal nuclease (MNase)

This protocol was adapted from Yenidunya (1992). Nuclei in buffer A were centrifuged at 1000 x g for 30 seconds, and the pellet was resuspended in 1 ml of MNase buffer. The nuclei were spun again, and resuspended in MNase buffer. After determining the concentration of nuclei, 250 μg of nuclei were made up with MNase buffer to 0.5 ml, and incubated at 22°C for 2-3 minutes.

After adding $CaCl_2$ to a final concentration of 1 mM and MNase to a final concentration of 2U/ml, the samples were incubated at 22°C for the required time.

The reactions were terminated by the addition of 1 volume of MNase termination buffer. After addition of RNAse A to 100 μg/ml, the samples were incubated at 37°C for 30 minutes. Then SDS was added to a final concentration of 0.5 % and proteinase K to 200 μg/ml, and the samples incubated at 37°C for at least 3 hours.

4.2.7 Southern Blotting

The DNA fragments produced by treating nuclei with a DNA cleavage reagent were digested for 8 to 15 hours with the appropriate restriction enzyme. After adding a further 5 units of enzyme, the reaction was mixed well and incubated for an extra hour. The DNA was purified by phenol/chloroform extraction, ethanol precipitated, resuspended in TE and left to redissolve overnight at 4°C.

After adding 1/5 volume of 5 X TPE-Ficoll loading buffer, samples which contained approximately 20 μg of DNA were dry-loaded and run on a 1.3% TPE-agarose gel, with ^{32}P-labelled 1 kb and 100 bp ladders (Promega), for two hours at 20 mA, then 19 hours at 60 mA.

Blotting was carried out as described by Southern (1975). Briefly, after electrophoresis, agarose gels were denatured in 0.5 M NaOH/ 1.5 M NaCl twice for 20 minutes on a rotary shaker at room temperature. The gels were rinsed with distilled water and neutralised twice in 3 M sodium acetate pH 5.5, for 20 minutes.

While the gel was neutralising, the following apparatus was set up. A solid

support was placed in a container of 20 x SSC (sufficient for blotting overnight), and a sheet of 17 mm Whatman paper (with wicks which dipped down into the liquid) was placed on top of the support. This paper was wetted thoroughly with 20 x SSC, and a smaller sheet of Whatman paper, pre-wetted in 20 x SSC, was placed on top of this and the neutralised gel was placed, well-side down, on top of this. Nylon hybridisation filter paper (one sheet), soaked in water, was laid onto the gel, taking care to avoid air bubbles. The edges of the paper were covered with parafilm and another sheet of Whatman paper, soaked in water, was placed on top of the filter and a dry sheet of Whatman paper was placed on top of this. Absorbent paper towels were placed onto the Whatman paper and a glass plate placed on top of the towels. An even pressure was applied to the gel by placing a 500 g weight onto the glass plate. The transfer was carried out overnight.

4.2.8 Filter Hybridisation

4.2.8.1 Cross-linking the DNA to the filter

Following completion of the transfer, the DNA was cross-linked to the filter by exposing it to 70,000 Joules/cm^2 of UV light in a Stratalinker (Stratagene) (alternatively, where stated, the DNA was cross-linked to the filter by baking at 80°C for 2 hours or by exposure on a UV light box for three minutes). The filter was washed in 2 x SSC for 2 minutes and further washed in 6 x SSC for 2 minutes. If necessary, the filter was stored at 4°C, sealed in a plastic envelope.

4.2.8.2 Prehybridisation of the filter

The filter was removed from the plastic envelope and placed on a piece of plastic mesh cut to the same size. The mesh and filter were rolled up and placed inside a hybridisation bottle (Amersham). 50 ml of prehybridisation buffer, preheated to 68°C, were added to the prehybridisation bottle (prehybridisation buffer contains 3x

SSC, 10 mM EDTA, 0.1% SDS, 0.2% polyvinylpyrolidine (PVP), 0.2% ficoll, 0.2% BSA, 100 μg/ml denatured salmon sperm DNA and 500 μg/ml of heparin), and the bottle incubated at 68°C for at least 2 hours. Meanwhile, the DNA probe was ^{32}P radiolabelled by the OLB method.

4.2.8.3 Preparation of ^{32}P-labelled probes by the OLB method

DNA fragments chosen as probes were ^{32}P radiolabelled by the OLB method (Sambrook et al., 1989). This involves annealing random DNA hexanucleotides to single-stranded DNA templates, followed by Klenow extension incorporating ^{32}P labelled dNTP's (Amersham).

Prior to labelling, approximately 20 ng DNA was dissolved in 24 μl distilled water and boiled for 3 minutes. The denatured DNA was added to a solution containing 1 μl each of 10 mM dGTP and dTTP, 5 μl of $\alpha[^{32}P]$dCTP (6000 Ci/mmole), 5 μl of $\alpha^{32}P$dATP (6000 Ci/mmole), 4U of Klenow fragment of DNA polymerase (NBL) and 10 μl of OLB buffer.

The reaction was incubated at 19°C for 30 minutes and for a further 30 minutes at 30°C. The reaction was stopped by the addition of 150 μl of OLB stop buffer (25 mM NaCl, 25 mM Tris-Cl pH 7.5, 5 mM EDTA and 1% SDS).

Unincorporated radioactive label was removed from the probe by spinning through a G-50 Sephadex column as described in chapter two.

4.2.8.4 Hybridisation of the filter

The ^{32}P-labelled probe was denatured at 98°C for 4 minutes and added immediately into 25 ml of hybridisation buffer preheated to 68°C (hybridisation buffer contains prehybridisation buffer supplemented with 9% dextran sulphate). After pouring out the prehybridisation buffer, the hybridisation buffer containing the denatured probe was added to the bottle containing the filter. The bottle was incubated overnight in a rotary incubator at 68°C.

To remove non-hybridising probe, the filter was then washed four times for 20 minutes with 2x SSC/0.5% SDS and twice with 0.5 x SSC/0.5% SDS at 68°C. If the washes were still radioactive the washes in 0.5x SSC/0.5% SDS were repeated as necessary. Higher stringency washes were used if required. After removing excess liquid the filter was sealed in a thin plastic envelope and exposed to Kodak XomatAR high intensity X-ray film with screens at -70°C.

4.2.9 Preparation of probes for southern blotting

4.2.9.1 Making probes by PCR

Probes were prepared by reactions, carried out as described in section 2.16. Briefly, template DNA (the appropriate clone from those described in section 3.2) was mixed with two appropriate oligonucleotide primers (designed so that they form a PCR product with the desired upstream and downstream boundaries), dNTP's, buffer and Taq polymerase. This mixture was taken through repeated cycles of denaturation, annealing and synthesis to produce a probe of the desired size and sequence.

4.2.9.2 Ligation of probe

The short Dra probe was self-ligated prior to OLB labelling to enhance the signal it generated.

'Polishing' the ends of the probe with Vent

Preparation of probes with Taq polymerase leaves 3′ A and T overhangs which could interfere with ligation of the probe. These overhangs were removed using Vent polymerase.

500 ng of probe was made up to 30 µl with 3 µl of 10 x Vent buffer, 0.5 units of

Vent DNA polymerase, 3 µl of 2 mM dNTP's and double-distilled water, and incubated at 72 °C for 20 minutes. After extraction with buffered phenol:chloroform and chloroform:IAA, the probe was precipitated and resuspended, ready for phosphorylation.

Phosphorylating the ends of the probe

Probe DNA (< 3 µg) was made up to 30 µl with 3 µl of 10 X blunt premix (500 mM imidazole-Cl (pH 6.4), 180 mM $MgCl_2$ and 50 mM DTT) and double-distilled water, heated to 70 °C for 3 minutes and then cooled immediately on ice. After the addition of 5 µl of 10 x blunt kinase buffer, 5 µl of 10 mM ATP, 30 U of T4 polynucleotide kinase (Pharmacia) and 6 µl of double-distilled water, the reaction was incubated at 37 °C for 1 hour. After supplementing the reaction with 40 µl of TE, it was incubated at 68 °C to inactivate the polynucleotide kinase. After the free ATP was removed by purification down a G-25 sephadex column, the DNA was purified and ethanol precipitated.

Self-ligating the probes

The probe DNA was resuspended in 1 X ligation buffer (50 mM Tris-HCl, (pH 7.4), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP and 0.025 mg/ml BSA) at 50 ng/µl, and 10 U of DNA ligase was added for every µg of DNA. The reaction was incubated for 12 to 36 hours at 12 °C. To observe the efficiency of the self-ligation, the probe was run on a 1.3 % agarose gel.

4.3 RESULTS

4.3.1 Introduction

This study has investigated nucleosome placement over a 7.5 kb region containing the chicken ε and β^A globin genes *in vivo*, in brain, where both genes are silent, and in adult red blood cells where epsilon is silent, but beta-adult has been expressed. The results are summarised in Figures 4.15 and 4.16. The mapping was carried out by indirect end-labelling (Wu, 1980; Nedospasov and Georgiev, 1980). The accuracy of this technique, which is in general ± 20 bp, is sufficient to say with some degree of certainty whether nucleosome positioning sites identified on the ε and β^A-globin genes *in vitro* (chapter 3) by the monomer extension technique are also occupied *in vivo*.

Comparison of the new *in vivo* data, with the *in vitro* data, which revealed the distribution of nucleosomes as determined by the underlying DNA sequence (Figure 3.14b), has shown the extent to which the DNA sequence of the globin genes determines their chromatin structure in active and inactive tissues. This study has revealed that the many of the strong nucleosome positioning sites identified on the β^A-globin *in vitro*, which occur with a 200 bp periodicity, may be occupied in brain and red blood cells. This regular array of positioned nucleosomes may facilitate the folding of long stretches of the chromatin fibre into an inactive higher order structure. By contrast, fewer of the strongest nucleosome positioning sites identified on the ε-globin gene *in vitro*, which are separated by a smaller (~180 bp) periodicity, are occupied in the brain and red blood cells.

These data also suggest that the placement of individual nucleosomes on the ε-globin promoter in the inactive state has the potential to regulate access of protein factors and the stability of promoter-enhancer interactions, a decisive factor in the competition of ε and β^A for their shared enhancer.

4.3.2 Strategy for mapping nucleosome placement on the ε and β^A genes

Nucleosome placement on the ε and β^A-globin genes in chicken was mapped in brain and in red blood cells from adult chickens, by digestion of nuclei with a DNA cleavage reagent, followed by indirect end-labelling. This technique will be discussed in detail below, but briefly, nuclei were prepared from fresh tissue and treated with a DNA cleavage reagent such as cuprous-phenanthroline (OPCu) which cuts the exposed DNA in the linker DNA between nucleosomes. Then, the DNA recovered from the nuclei was cut with a restriction enzyme and run with size markers and the appropriate controls on an agarose gel. The DNA was transferred to a nylon filter and hybridised with a radiolabelled probe. After washing off non-specific signal, the filter was exposed to X-ray film. In addition, densitometer scans were made for some of the gels following Phosphor-imager analysis (Raytech). The size of each fragment represents the distance from an OPCu cleavage site in the linker DNA to the restriction enzyme site adjacent to the probe. The fragments were sized to define the position of nucleosome boundaries relative to the restriction enzyme site.

In the first stage of the analysis, linker DNA between nucleosomes in the nucleus is cut with a cleavage reagent. Three cleavage reagents, micrococcal nuclease, cuprous-phenanthroline and methidium propyl EDTA (MPE), were considered for use in this study. Time-courses for digestion of adult red blood cell nuclei with micrococcal nuclease, MPE and OPCu are shown in Figures 4.2A, 4.2B and 4.2C.

Micrococcal nuclease has a much higher sequence specificity in its cleavage of DNA than either OPCu or MPE (Cartwright et al., 1983). In addition, micrococcal nuclease, unlike MPE, does not always detect nucleosomes on active genes, and may only yield a smear instead of a nucleosome ladder (Benezra et al., 1986). Therefore, the chemical cleavage reagents, MPE and OPCu were considered to be better candidates for use in this study. These reagents generate single-stranded cuts in the phosphate backbone of the DNA, and in both cases bound protein protects the DNA

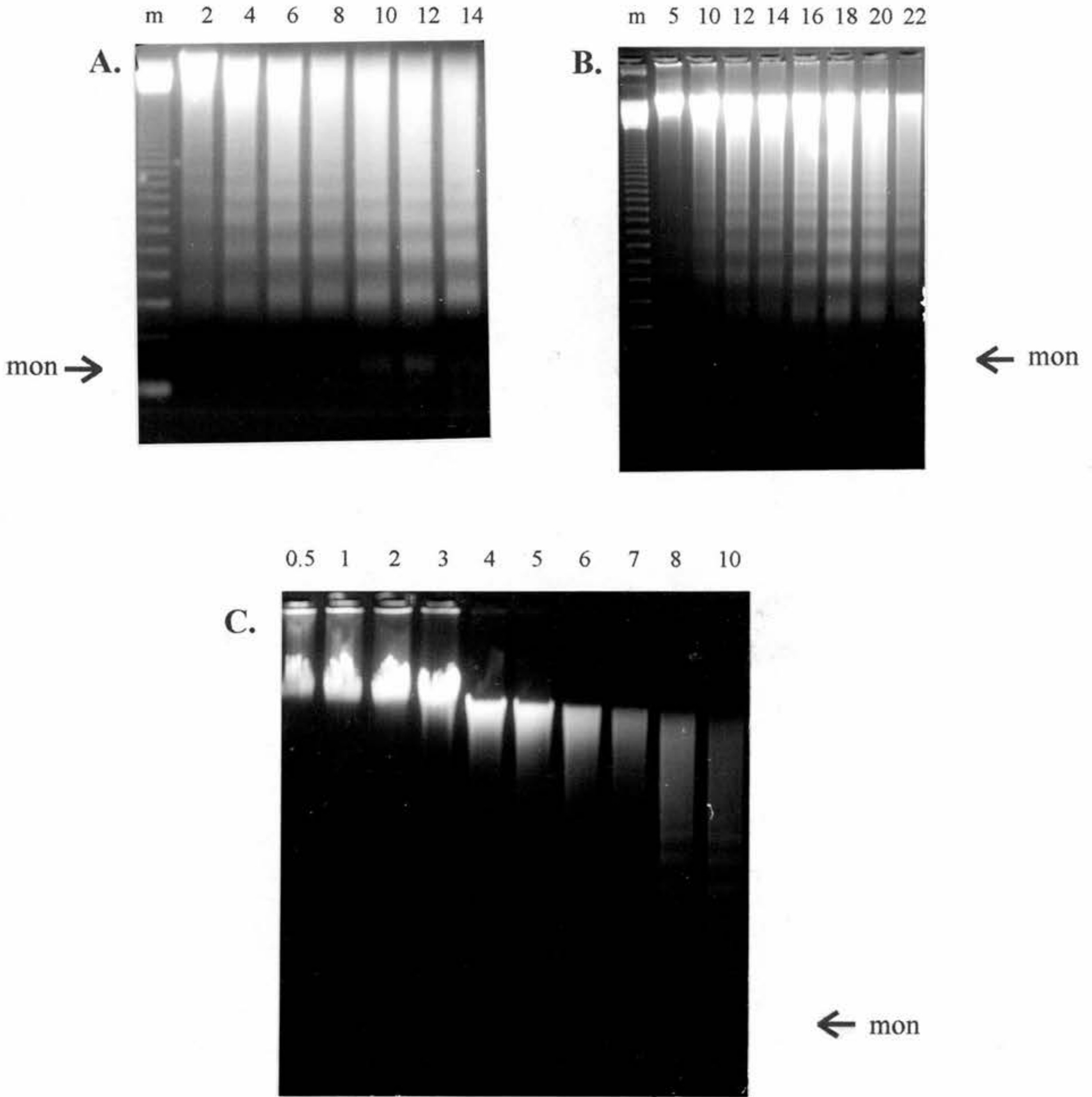


Fig. 4.2 Time-course of digestion of red blood cell nuclei by DNA cleavage reagents and micrococcal nuclease.

Gel A shows nuclei digested with micrococcal nuclease at 2 U/ml for 2, 4, 6, 8, 10, 12 and 14 minutes. Gel B shows nuclei digested with MPE for 5, 10, 12, 14, 16, 18 and 20 minutes . Gel C shows nuclei digested with cuprous phenanthroline for 0.5, 1, 2, 3, 4, 5, 6 ,7, 8 and 10 minutes. The marker (m) is a 123 bp ladder. The position of mononucleosome DNA is indicated (mon).

from cleavage. MPE has been used to map nucleosome positions in a number of studies (for example, Cartwright et al., 1984), and OPCu has been used to determine the protein architecture of the *Drosophila* hsp27 gene promoter (Quivy and Becker, 1996). Some studies have indicated that OPCu has a higher sequence specificity than MPE (see Cartwright et al., 1983). Others, however, suggest that OPCu has essentially no primary sequence specificity (Marshall et al., 1981) and limited secondary structure specificity (Perrin et al., 1996), and this was confirmed by this project and another similar study on the sheep β-lactoglobulin gene (S.Boa, pers. comm.). Therefore, OPCu and MPE appeared equally suitable for this study. The analyses presented in this chapter were carried out with OPCu.

In this study, nucleosome placement was analysed in both brain and red blood cells from adult chickens. The globin genes are expressed only in erythroid cells, and the individual genes are expressed only at specific times in development (reviewed in Felsenfeld, 1993). In brain, a non-erythroid tissue, the globin genes are inactive, and the chromatin domain which contains them is thought to be packaged into an inaccessible higher order structure, because the DNA is resistant to nucleases and hypermethylated (Wood and Felsenfeld, 1982; Mandel and Chambon, 1979). In erythroid tissues, such as red blood cells, the chromatin domain encompassing the globin genes has a more open structure, as indicated by increased nuclease sensitivity (Stalder et al., 1980; reviewed in Felsenfeld, 1993). In adult circulating red blood cells the globin domain retains the characteristics of active chromatin, such as nuclease sensitivity, although transcription has almost completely ceased in the bulk of cells. In these cells ε-globin is not active, but the β^A-globin gene has been recently transcribed and its promoter retains a nuclease hypersensitive site (Mathis et al., 1980; Bellard et al., 1980). If time had allowed, red blood cells at other stages in development would have been studied, in particular, 5-day primitive red blood cells, where the embryonic epsilon gene is active, and the beta-adult gene is silent.

Following OPCu digestion of the nuclei from brain and red blood cells, nucleosome placement was analysed by indirect end-labelling. The indirect end-labelling technique was chosen because the large region encompassing the β^A and ε globin genes could be mapped by this method. The resolution achieved with this

globin genes could be mapped by this method. The resolution achieved with this technique (generally ± 20 bp) was sufficient to say with some confidence whether a nucleosome positioning site identified *in vitro* by monomer extension was occupied *in vivo* or not. The resolution of the technique depends upon several factors, and the ability of the agarose gel to resolve fragments is the most important of these. The resolution is also affected by the accuracy with which fragments can be sized. To obtain as accurate an estimate as possible, the fragments in each lane of a gel were sized using sets of markers to the right and left of the samples, both by hand, and using the AIDA (Raytek) software package. This programme, which can analyse either the densitometry data or a scan of the X-ray film, identifies the position of fragments automatically, and the positions can be adjusted manually if required. When the size of marker fragments on the gel has been entered into the AIDA programme, an equation is generated which relates the position of a fragment to its size in bp. AIDA uses this equation automatically to convert the position of fragments into DNA size. By defining the minimum and maximum values for the size of each fragment, it is possible to estimate the limits of accuracy for each nucleosome position. Certain factors may increase the error in sizing fragments by hand or using AIDA. For example, 'smiling' of bands makes size estimation more difficult. The viscosity of the DNA can also affect the apparent size of fragments. One set of markers was mixed with 20 μg of sheared chicken DNA so that it had the same viscosity as the other samples. These markers showed that the viscosity of the samples caused 'smiling' of the larger fragments, and the top edge of such bands, rather than the centre, provided a more accurate size estimate.

The nucleosome mapping technique used in this study requires that certain controls are included in each analysis:-

- Naked genomic DNA was cut with restriction enzyme, to ensure that the probe anneals specifically to only one fragment.

- Naked genomic DNA was digested with OPCu and then restricted with the appropriate restriction enzyme to reveal any sequence-specific cuts created by OPCu in naked DNA.

•,Finally, nuclei were digested with OPCu alone to indicate the location of fragments created by an OPCu cut at each end. Any fragment which was indistinguishable from a fragment in this control was eliminated from the analysis because it might represent the distance between two OPCu cut sites rather than the distance from an OPCu cleavage site to the restriction enzyme site.

A dozen indirect end-labelling analyses, using 6 probes, were required to map nucleosome placement on the ~ 7 kb β^A and ε-globin gene region in brain and red blood cells. The probes were chosen so that they were separated by ~ 1800 bp intervals, so that the fragments analysed were < 2000 and therefore able to be sized more accurately. The probes were placed so that suitable restriction enzymes occurred at the 5′ and 3′ ends of the probe, allowing nucleosome placement to be mapped both upstream and downstream of the probe. The probes are shown in Figure 4.1. In some cases, difficulties with restriction enzymes giving only partial cutting meant that only one direction was mapped. The probes were designed to be 250 to 390 bp long. Probes below 250 bp were found to give a very weak signal; therefore the short (139 bp) Dra probe required multimerisation by self-ligation (section 4.2.9.2) before it gave a sufficiently strong signal. It is preferable to use probes which are as short as possible, because the longer the probe the greater the chance that it will detect fragments with OPCu cuts at both ends.

4.3.3. Generation of a nucleosome placement map for the ε and β^A genes in brain and red blood cells

Twelve mapping analyses were carried out to map nucleosome placement over the chicken β^A- and ε-globin genes in brain and red blood cells. A typical example, which maps nucleosome placement using the Hind probe over the β^A-globin gene in brain, is discussed in detail first. The samples for the mapping experiment were run on an agarose gel (Figure 4.3) and the DNA transferred to a nylon filter which was hybridised with the Hind probe (Figure 4.4). An enlargement of part of Figure 4.4 is shown in Figure 4.5.

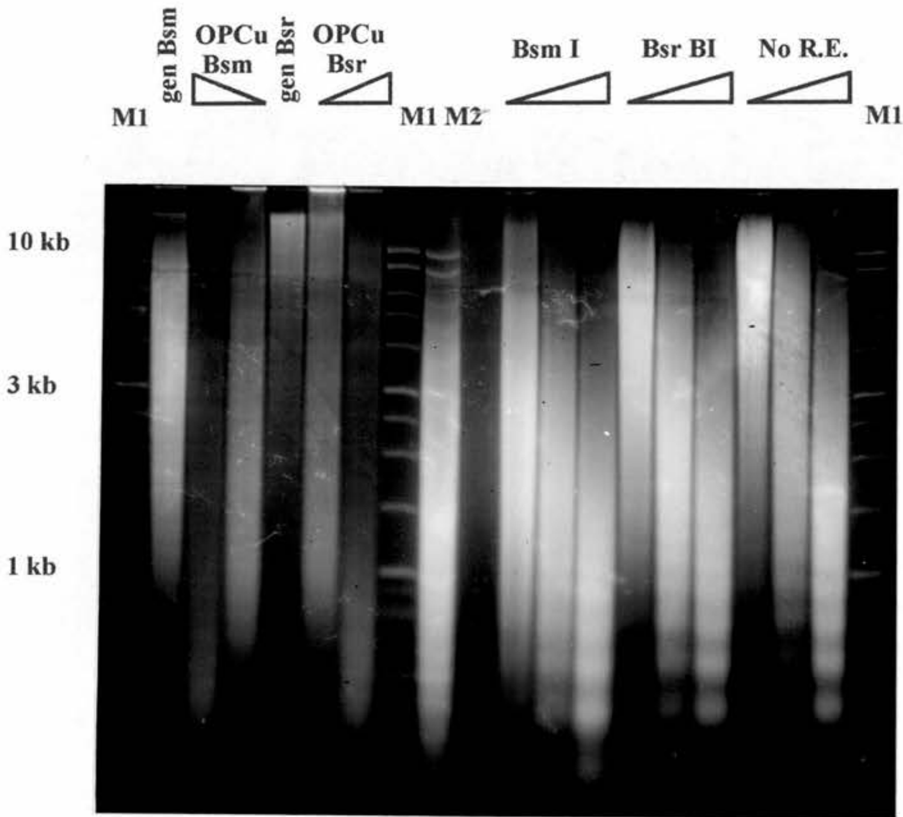


Fig. 4.3 A typical agarose gel prepared for mapping nucleosome placement by indirect end-labelling.

DNA from nuclei treated with OPCu was restricted with either Bsm I or with BsrBI (right hand side of gel, Bsm I and BsrBI; the gradient indicates increasing times for OPCu digestion). DNA from nuclei digested with OPCu, but not restricted is shown (No RE). Naked genomic DNA was restricted with Bsm I (gen Bsm) or BsrBI (gen Bsr). Naked genomic DNA, digested with OPCu, was restricted with Bsm I (OPCu Bsm) or BsrBI (OPCu Bsr). Some DNA marker sizes are indicated next to the M1 marker lane. M2 contains the same marker DNA, mixed with sheared chicken genomic DNA.

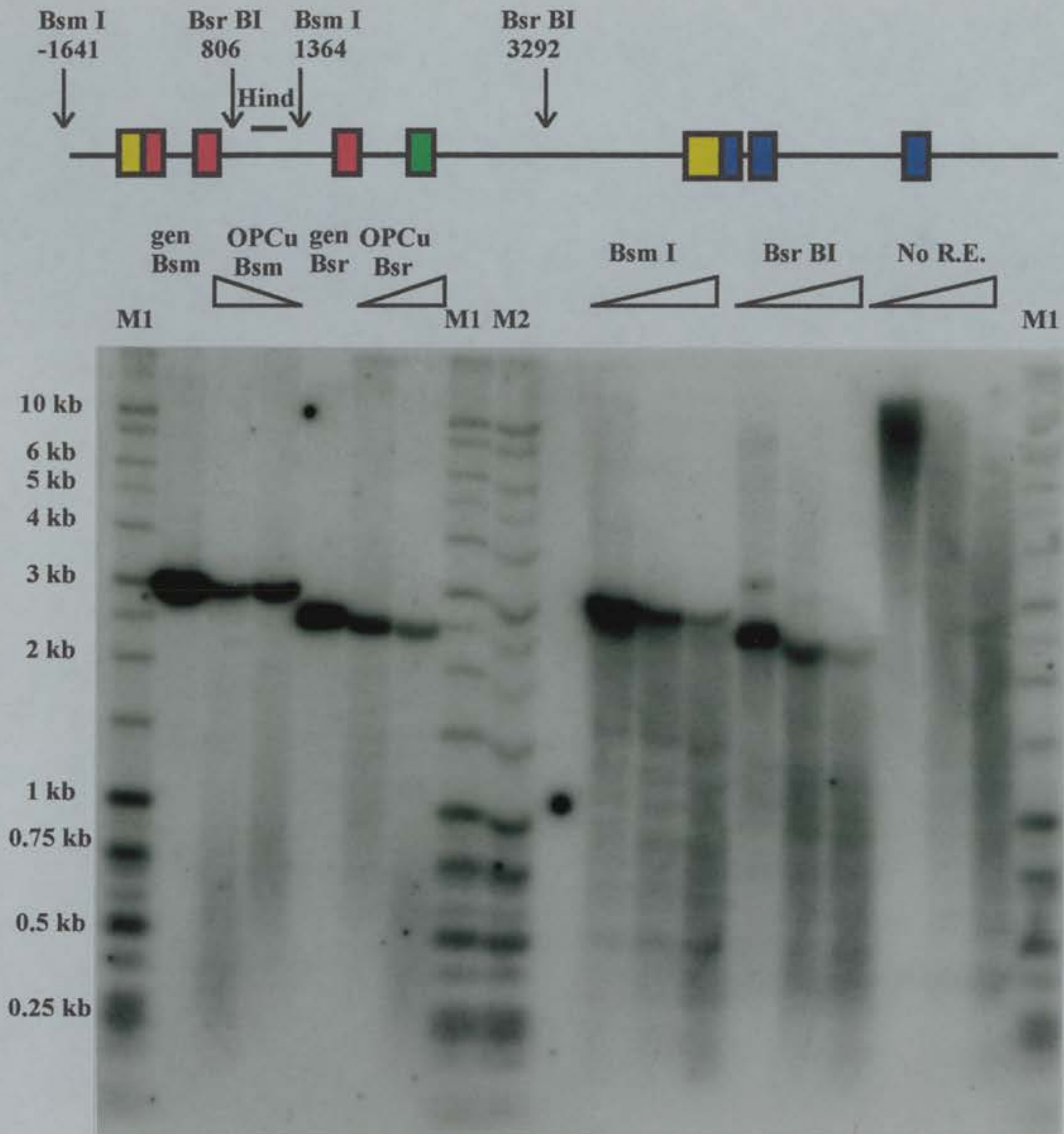


Fig. 4.4 Nucleosome mapping on brain nuclei using the Hind probe.
DNA from nuclei cut with OPCu was restricted with either Bsm I or with BsrBI (Bsm and Bsr). In this and all the subsequent figures wedge shapes above the lanes indicate the use of increasing amounts of OPCu in the treatment of samples.. DNA from nuclei digested with OPCu, but not restricted is shown (No RE). Naked genomic DNA was restricted with Bsm I (gen Bsm) or BsrBI (gen Bsr). Naked genomic DNA, digested with OPCu, was restricted with Bsm I (OPCu Bsm) or BsrBI (OPCu Bsr). The DNA marker size is indicated next to the marker lane (M1). M2 contains the same marker DNA, mixed with sheared chicken genomic DNA. The schematic above the blot shows the position of the Hind probe, relative to the β^{A}-globin gene (red), the ε-globin gene (blue)and the enhancer (green).

Three controls were included in this analysis (Figure 4.4), and were routinely used in all other mapping analyses.

• Naked genomic DNA was cut with restriction enzyme, either Bsm I or BsrBI. These controls indicate that the Hind probe anneals to the correct fragment and not elsewhere, as fragments of the expected size only are observed (3005 bp, lane gen Bsm and 2486 bp, gen Bsr).

• Naked genomic DNA was cut with OPCu and then restricted with Bsm I or BsrBI. A smear is observed in these lanes (OPCu Bsm and OPCu Bsr), rather than any discrete fragments. This demonstrates that OPCu does not make any sequence-specific cuts in the region being mapped in this analysis.

•Nuclei were digested with OPCu, and the DNA recovered but not restricted (lane no R.E.). This control shows the position of fragments created by an OPCu cut at each end. Any fragment which was indistinguishable from a fragment in this control was eliminated from the analysis because it might represent the distance between two OPCu cut sites rather than the distance from an OPCu cleavage site to the restriction enzyme site; two such fragments which were eliminated from this analysis are indicated in Figure 4.5. However, comparison of this control (middle lane of No R.E.) with the samples digested with both OPCu and restriction enzyme, at the same time-point in OPCu digestion, (middle lane of Bsm or Bsr) shows that fragments in the restriction enzyme-digested samples are much smaller, and they occur in a part of the gel where there is little material in this control cut with OPCu only. Therefore, the fragments in lanes Bsm and Bsr are the result of restriction enzyme digestion, and they reflect the position of OPCu-cut sites in linker DNA relative to the restriction enzyme cut site (apart from the exceptions indicated). This demonstrates unambiguously that there are positioned nucleosomes in this region.

For mapping nucleosome placement in this example, DNA recovered from brain nuclei digested with OPCu was restricted with either Bsm I or BsrBI (Figure 4.5, lanes Bsm and Bsr respectively) and hybridised to the Hind probe. (see the schematic at the top of Figure 4.5). Fragments in lanes Bsm and Bsr (~ 0.2 to 2 kb in size) can be used to determine the placement of positioned nucleosomes 5′ and 3′ of the probe,

Fig. 4.5 Nucleosome mapping on brain nuclei using the Hind probe.

A Southern analysis of the DNA recovered from brain nuclei cut with OPCu and then restricted with either Bsm I or with BsrBI is presented. The schematic above the blot shows the position of the Hind probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles), and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. In this, and all subsequent figures, the nucleosome positions were inferred using the lane representing the earliest time point in the OPCu treatment of the nuclei. The arrows indicate the position of OPCu cleavage sites and the fletched arrows indicate sites which were not included in the analysis because they occur in the 'No R.E.' control. The markers are as in Figure 4.4.

respectively. The largest band in these lanes represents the fragment created by two restriction enzyme cuts. The other, smaller fragments in these lanes, which define the position of an OPCu-cleavage site in linker DNA, occupy discrete positions, indicating the presence of positioned nucleosomes in this region. The size of each of these fragments represents the distance from an OPCu cut site in linker DNA to the restriction enzyme site adjacent to the probe. Therefore the position of the linker DNA either side of a nucleosome can be defined, and in this way the placement of the centre of each positioned nucleosome can be estimated, relative to the underlying globin DNA sequence. The schematic at the side of the blot in Figure 4.5 shows the inferred nucleosome positions determined in this analysis.

The mapping with Bsm I shows positioning over the first intron, second exon and second intron. Two nucleosomes occupy positions, at 780 and 958 bp relative to the β^A cap site, which are consistent with strong nucleosome positioning sites identified *in vitro* at 775 and 979 (nucleosomes 8 and 9a, and Figure 4.16). The analysis also reveals a positioned nucleosome on the β^A promoter (at -116). The deduced nucleosome position centred at 569 is not included in the final summary of the positioning data (Figure 4.15) because there is some uncertainty about it; the gap between the two OPCU-cleavage sites is rather small (~ 155 bp).

The mapping with BsrBI reveals nucleosome placement over part of the second intron, the third exon, and the region around the enhancer. The nucleosome positioned at 1270 suggests that the prominent position identified *in vitro* at ~ 1260 (nucleosome 10b in Figure 4.16) may be occupied in brain. The nucleosome located at ~ 1985 over the enhancer may also occupy the position identified *in vitro* at 1971 (nucleosome 12 in Figure 4.16). This placement is consistent with nucleosome positioning data produced by a previous study on the enhancer in brain (Buckle et al., 1991). Two of the positioned nucleosomes (at ~ 1358 and ~ 1740) may occupy minor nucleosome positioning sites detected *in vitro* at ~ 1348 and ~ 1732 (Figure 4.16). Cleavage sites marked by a feathered arrow are ignored as they are indistinguishable from fragments in the OPCu-cut control.

For future reference, the OPCu-cut control shows that the region probed has an average nucleosome spacing of ~ 200 bp (Figure 4.4, no R.E.). The monomer, dimer

and trimer are ~ 200, ~ 400 and ~ 600 bp.

Similar analyses were carried out in brain and red blood cells to map nucleosome placement over the entire β^A and ε globin gene region in these tissues. The indirect end-labelling data from all these analyses, showing the centre of nucleosome positions, are presented in Figures 4.6 to 4.14 and the results for all the analyses are summarised in Figures 4.15 and 4.16.

Nucleosome mapping on red blood cell nuclei with the Hind probe

Figure 4.6 reveals nucleosome placement over the same region as Figure 4.5, but in adult red blood cells. The most obvious difference between the two tissues is the strong OPCu cut sites over the β^A promoter region, which reflect the loss of the nucleosome positioned at ~ -116, and the consequent nuclease hypersensitivity of the β^A promoter in this tissue where the gene has been active. It is also noticeable that other positioned nucleosomes around the promoter appear to be disrupted, and those remaining appear less prominent (compare nucleosome 768 in Figure 4.6 with 780 in Figure 4.5). From the mapping with BsrBI, the enhancer also appears to exhibit hypersensitivity to OPCu.

Nucleosome mapping on brain nuclei with the Eco probe

Figure 4.7 shows nucleosome placement, again mapped with the Eco probe, over the 5′ flanking region, the promoter and the first exon of the β^A gene in brain. Though these data have a high background, there are clearly a number of positioned nucleosomes, including one on the promoter, centred at ~ -110. These data are consistent with the nucleosome occupying the positioning site identified *in vitro* at ~ -139 (nucleosome 5, Figure 4.16). It is also consistent with the results of another study on nucleosome placement on the β^A-globin gene in brain (Buckle et al., 1991).

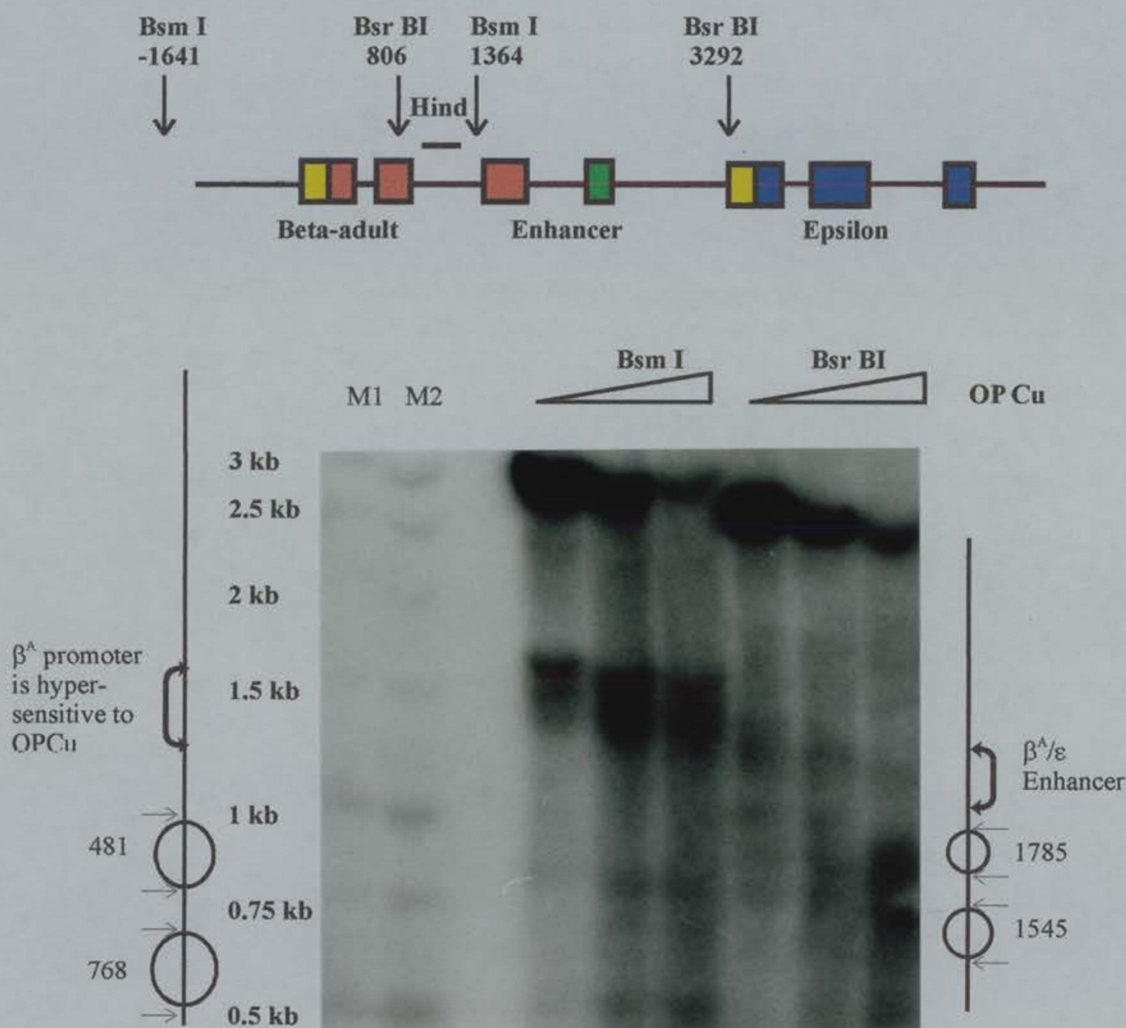


Fig. 4.6 Nucleosome mapping on red blood cell nuclei with the Hind probe.

A Southern analysis of the DNA recovered from red blood cell nuclei cut with OPCu and restricted with either Bsm I or with BsrBI. The schematic above the blot shows the position of the Hind probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers asociated with these indicate the centre of each nucleosome relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. Markers are as in Figure 4.4.

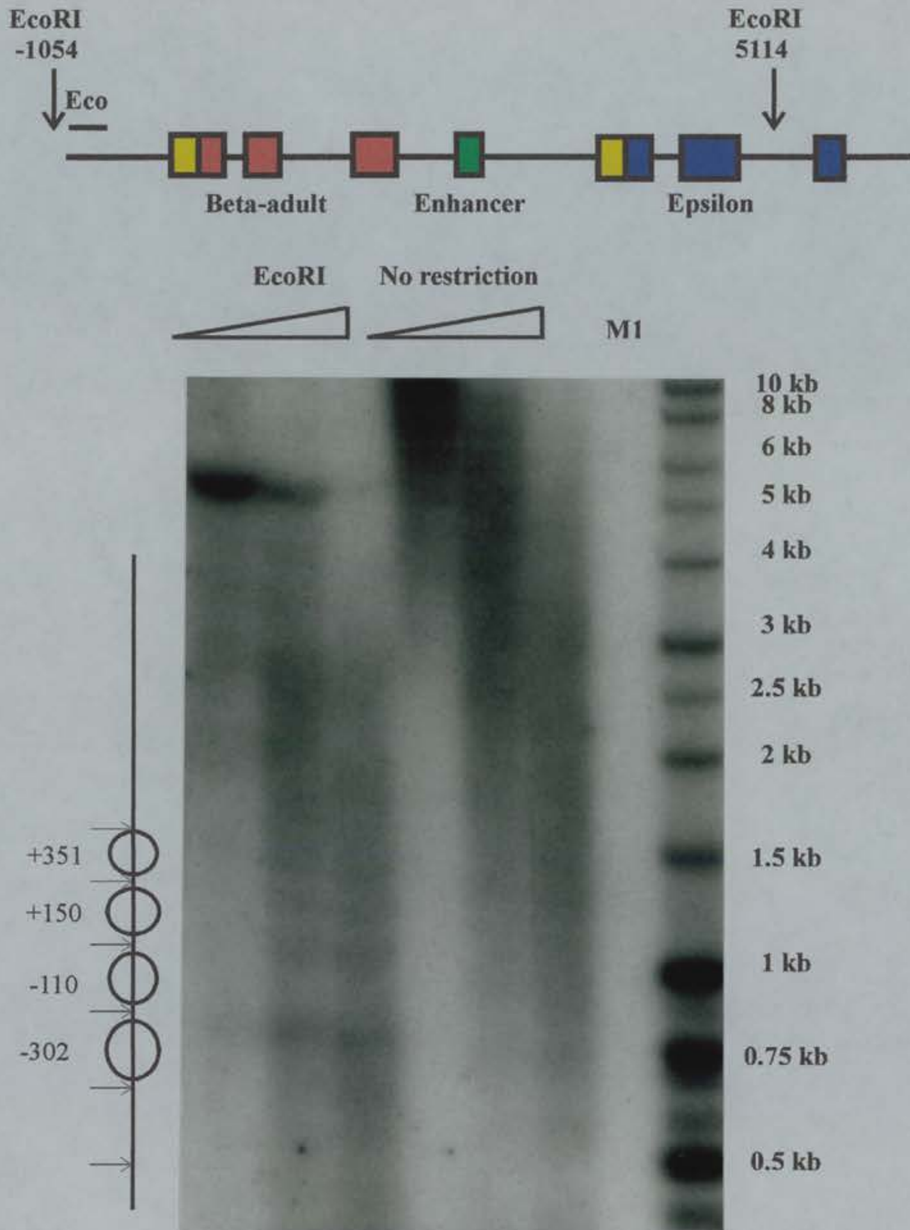


Fig. 4.7 Nucleosome mapping on brain nuclei using the Eco probe.

A Southern analysis of the DNA recovered from brain nuclei cut with OPCu and then restricted with EcoRI is presented. DNA from nuclei cut with OPCu but not restricted is also shown. The schematic above the blot shows the position of the Eco probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the positions of OPCu cleavage sites. The markers are as in Figure 4.4.

Nucleosome mapping on red blood cell nuclei using the Eco probe

Figure 4.8 shows the same region as Figure 4.7, mapped with the Eco probe, in red blood cells. As in Figure 4.6, positioned nucleosomes can no longer be detected on or around the β^A-globin promoter. There may be a nucleosome centred at –503, but the cleavages at the boundary are very weak. There are strong cleavage sites within the promoter at ~ -195, ~ -90, and ~ 15, which reflect its nuclease hypersensitivity.

Nucleosome mapping on brain nuclei using the Dra probe

Figure 4.9 shows the region around the enhancer in brain, mapped using the Dra probe. The enhancer is occupied by a nucleosome positioned at ~ 1987, which is consistent, within the limits of accuracy of the technique, with the data from Figure 4.5 where a positioned nucleosome was found at 1985. This is consistent with occupation of the positioning site identified *in vitro* at ~ 1967 (12 in figure 4.16). The nucleosomes positioned at ~2195 and ~2372 suggest that the sites identified *in vitro* at ~2215 and ~2361 (13 and 13b in Figure 4.16) are occupied in brain.

Some of the fragments in Figures 4.9 are close in size to fragments in the control where the nuclei were cut with OPCu, but not restricted. These fragments were not eliminated from the analysis because when the same time-point in OPCu digestion is compared for the control (middle lane, no restriction) and the restriction enzyme sample (middle lane, Sac I), it is evident that the fragments in the restriction enzyme-digested samples are much smaller, and they occur in a part of the gel where there is little material in the control. Therefore, the fragments in the Sac I lanes are in fact restriction enzyme-dependent. A similar consideration prevented the elimination of certain fragments from the analysis of Figure 4.10 (below).

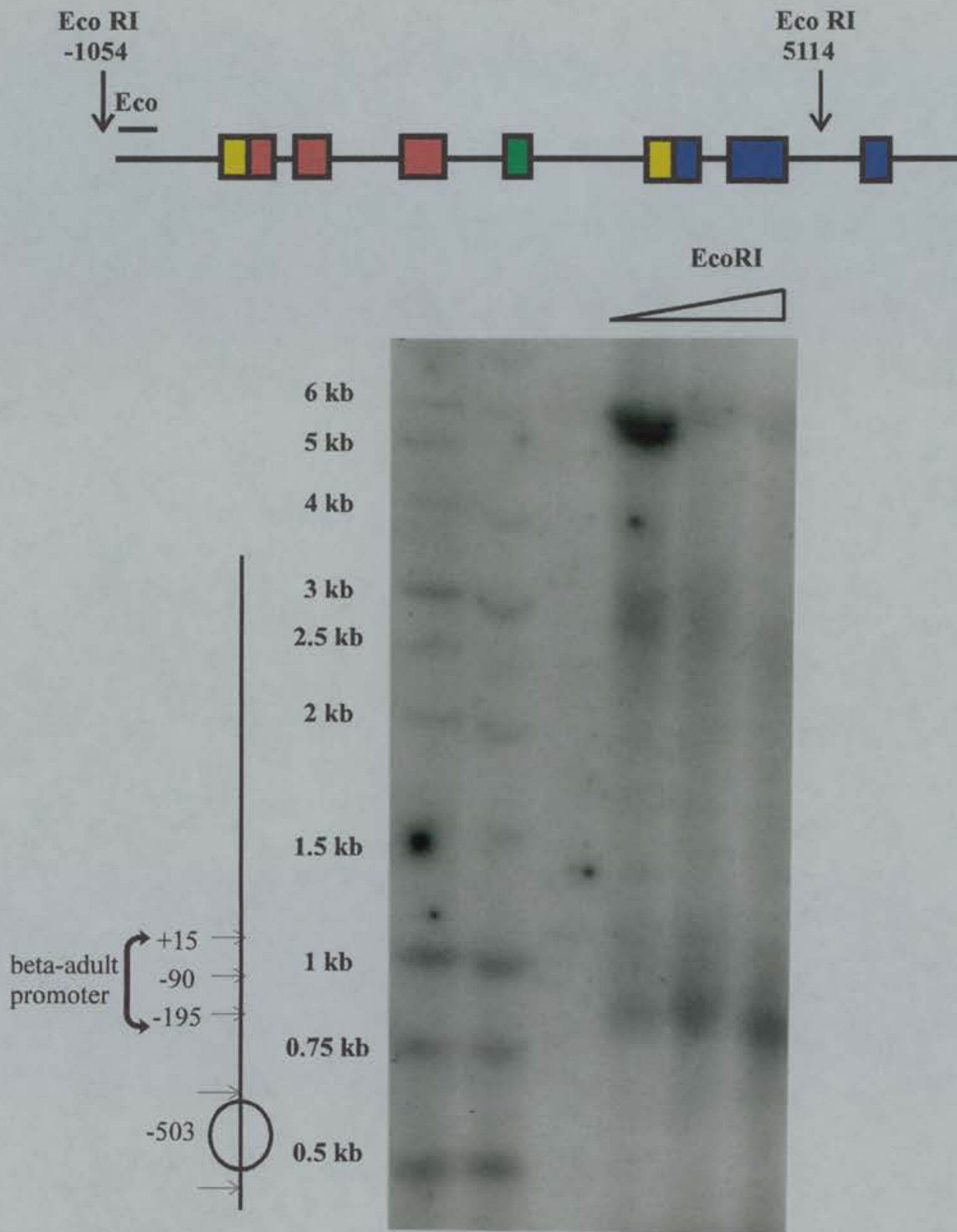


Fig. 4.8 Nucleosome mapping on red blood cell nuclei using the Eco probe.

A Southern analysis of the DNA recovered from red blood cell nuclei digested with OPCu was restricted with EcoRI. The schematic above the blot shows the position of the Eco probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. The markers are as in Figure 4.4.

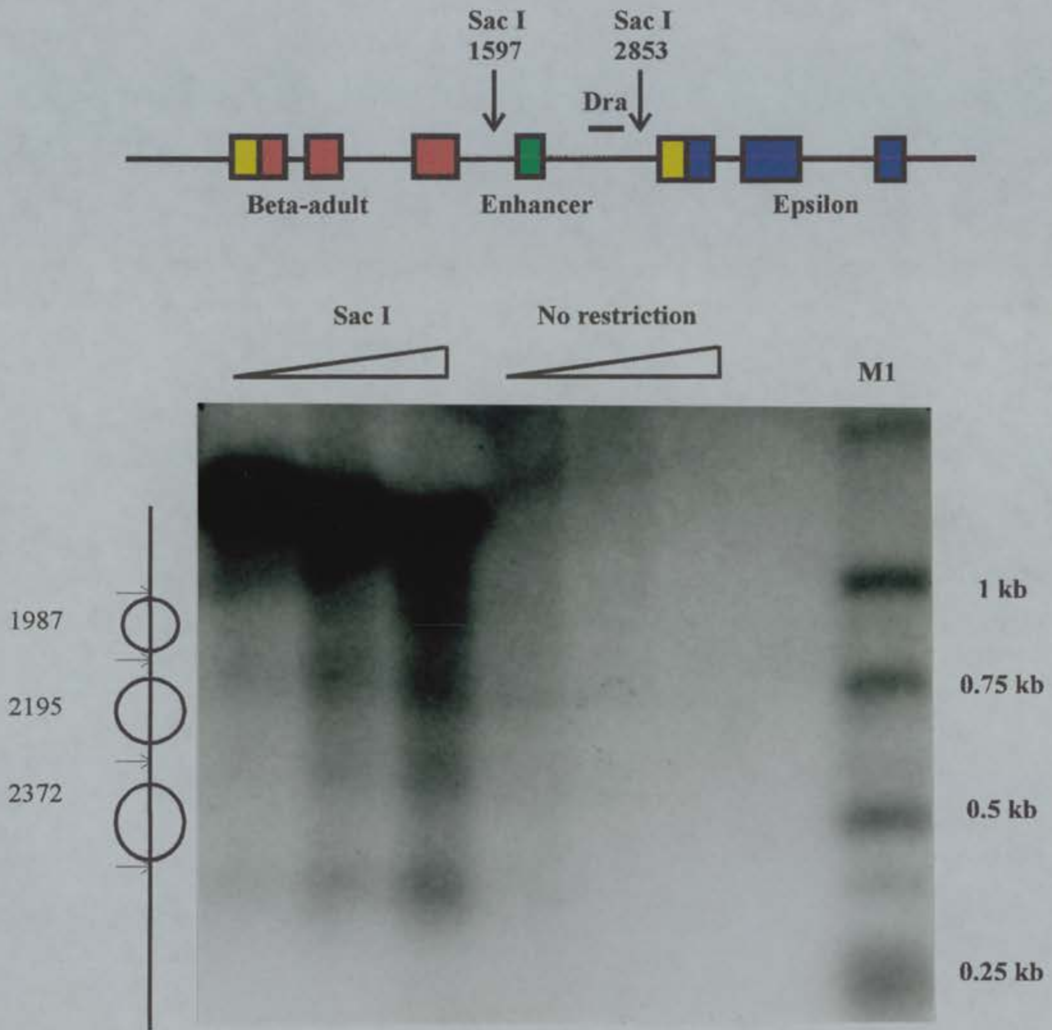


Fig. 4.9 Nucleosome mapping on brain nuclei using the Dra probe.

A Southern analysis of the DNA recovered from brain nuclei cut with OPCu and then restricted with Sac I is presented. DNA from nuclei cut with OPCu but not restricted is also shown. The schematic above the blot shows the position of the Dra probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the positions of OPCu cleavage sites. The markers are as in Figure 4.4.

Nucleosome mapping on red blood cell nuclei using the Dra probe

Figure 4.10 shows the enhancer region, mapped in red blood cells with the Dra probe. The nucleosome positioned on the enhancer (nucleosome 1987 in Figure 4.9) has been disrupted. Extra OPCu cut sites occur over the enhancer, at 2103, 2016 and 1936, reflecting its nuclease sensitivity in red blood cells. However, the nucleosome which was positioned 3′ of the enhancer in brain also occurs in red blood cells (at ~2385).

Nucleosome mapping on brain nuclei using the Bam probe

Figure 4.11 shows nuclcosome placement 5′ of the ε-globin gene. The mapping with Msc I enzyme and the Bam probe reveals a nucleosome positioned at ~ 3096, which is consistent with occupation of the *in vitro* positioning site identified at 3080 (16, Figure 4.16). There may also be a nucleosome positioned at ~ 2981, but the gel is rather smeary so it is difficult to say if there is a clear gap between the potential nucleosome boundaries. The nucleosome positioned at ~ 2370 has a greater than typical uncertainty in its position (± 35 bp) than other sites, as the boundaries either side of it are rather unclear.

The mapping with BamHI reveals nucleosomes positioned at 3780, 4210 and 4786 which may occupy positioning sites identified *in vitro* (nucleosomes 19b, 20b and 4786 respectively in Figure 4.15 and 4.16).

Nucleosome mapping on red blood cell nuclei using the Bam probe

Figure 4.12 shows nucleosome placement on the same region in red blood cells. The nucleosome positioned at 3121 in red blood cells may be the nucleosome seen in brain at 3096. There at some differences with the map in brain, but the poor quality of the data in brain makes it difficult to ascribe significance to these differences.

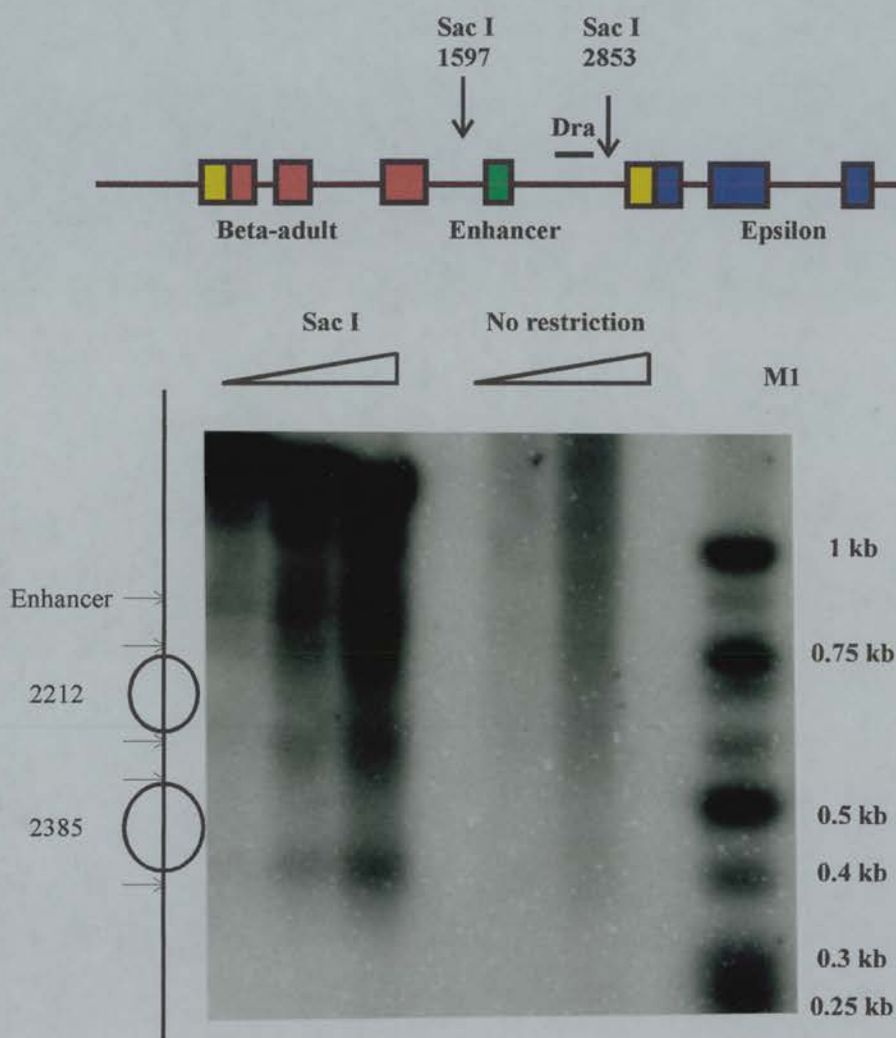


Fig. 4.10 Nucleosome mapping on red blood cell nuclei using the Dra probe.

A Southern analysis of the DNA recovered from red blood cell nuclei cut with OPCu and then restricted with Sac I is presented. DNA from nuclei, cut with OPCu, but not restricted is also shown. The schematic above the blot shows the position of the Dra probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. The markers are as in Figure 4.4.

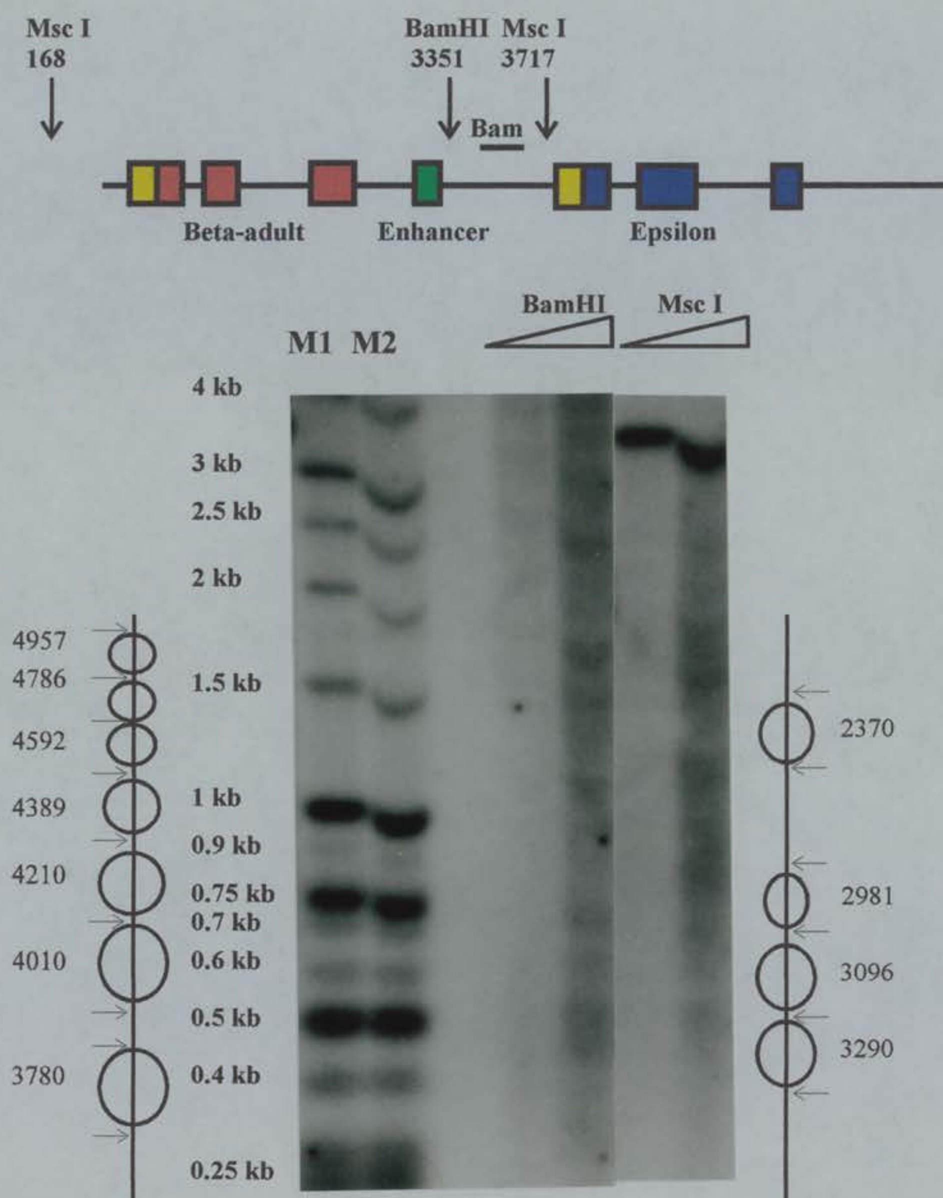


Fig. 4.11 Nucleosome mapping on brain nuclei using the Bam probe

A Southern analysis of DNA recovered from brain nuclei cut with OPCu and then restricted with either BamHI or with Msc I is presented. The schematic above the blot shows the position of the Bam probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the positions of OPCU cleavage sites. Markers are as in Figure 4.4.

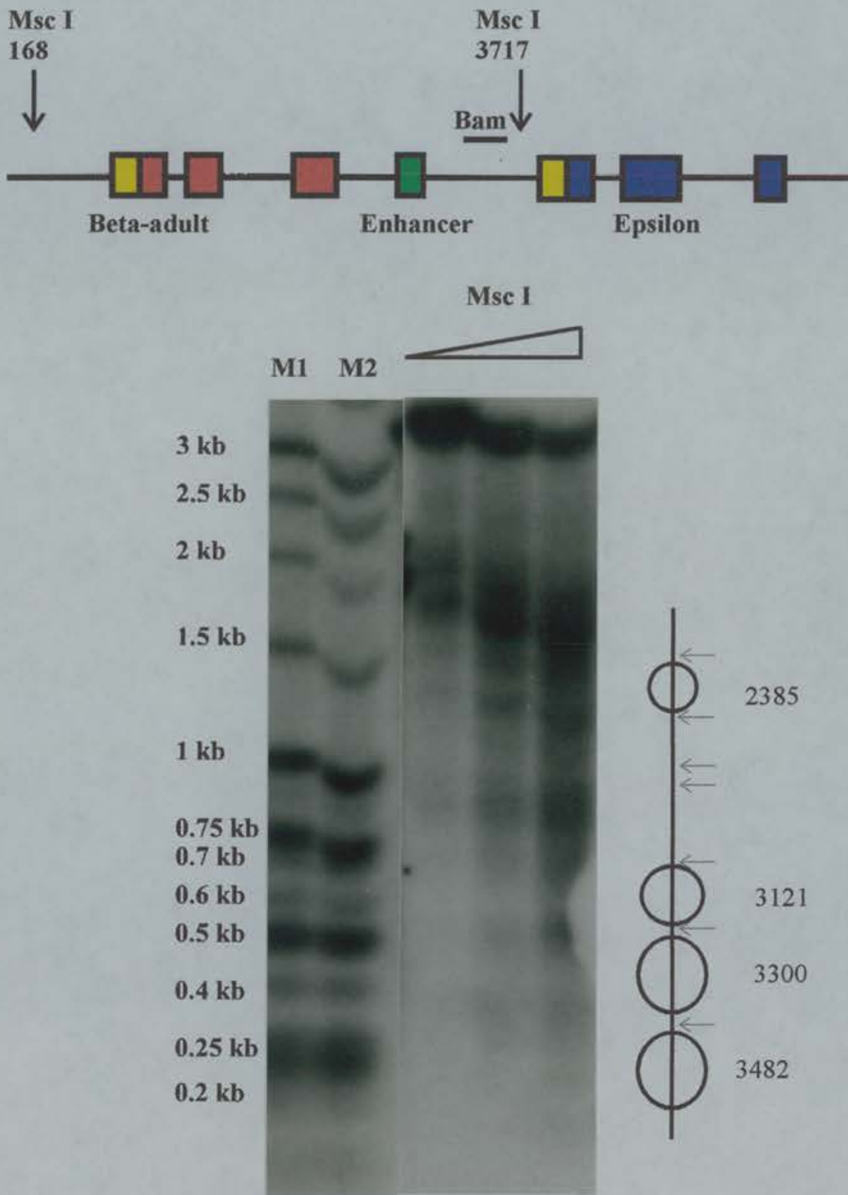


Fig. 4.12 Nucleosome mapping on red blood cell nuclei using the Bam probe

A Southern analysis of DNA recovered from red blood cell nuclei cut with OPcu was restricted with Msc I. The schematic above the blot shows the position of the Bam probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of the nucleosome, relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. The markers are as in Figure 4.4.

Nucleosome mapping on red blood cell nuclei with the Eag probe

Figure 4.13 shows the location of nucleosomes around the ε-globin promoter and first exon in red blood cell nuclei, mapped using the Eag probe and EcoRI enzyme. The nucleosome positioned on the promoter at ~ 4212 suggests that the minor nucleosome position identified *in vitro* at 4237 (20b in Figure 4.16) may be occupied in brain. The other nucleosomes positioned around the promoter do not correspond to positions detected *in vitro*. Strong cutting over the enhancer, reflecting its hypersensitivity, can also be seen. In brain, the placement of nucleosomes around the promoter is essentially the same (data not shown). As discussed earlier for Figures 4.9 and 4.10, where the enhancer is shown at higher resolution, there are differences around the enhancer between brain and red blood cells.

Cleavage site (1) in Figure 4.13 was eliminated from the analysis because its size is indistinguishable from that of a fragment occurring in one of the controls. The fragment in the control reaction, which represents a very small percentage of the total cutting, is thought to be generated by star activity. Although the control was more heavily digested with EcoRI than the other samples, the fragment was not included in the analysis to err on the side of caution. A small amount of star activity would not affect the evaluation of the other fragments in the mapping analysis.

Nucleosome mapping on brain nuclei using the Rsa probe

Figure 4.14 shows nucleosome placement in the second intron, third exon and 3′ flanking region of ε, using the Rsa probe. The mapping with the BstEII enzyme reveals four positioned nucleosomes. One of these at ~ 5281 may be occupying the very strong nucleosome positioning site identified *in vitro* at 5304 (23a in Figure 4.16). The nucleosome at ~ 4924 is bracketed, as one of its boundaries may coincide with a fragment in the OPCu-cut control. It is difficult to be certain, as the OPCu-cut control bands are quite broad and are close to merging above 1 kb.

The mapping with Nsi I reveals four positioned nucleosomes in the region 3′ of

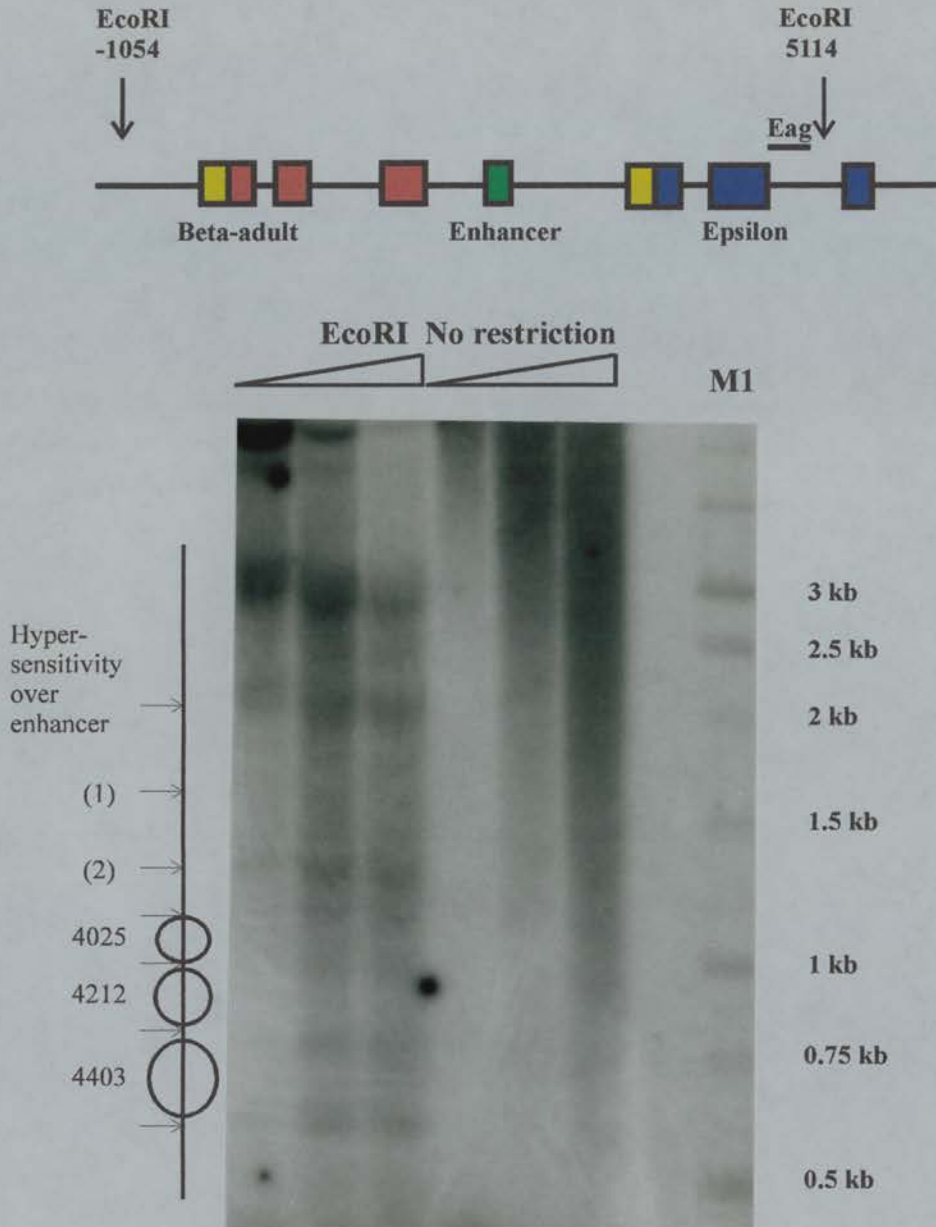


Fig. 4.13 Nucleosome mapping on red blood cell nuclei using the Eag probe.

A Southern analysis of the DNA recovered from red blood cell nuclei cut with OPCu and then restricted with Eco RI is presented. DNA from nuclei digested with OPCu, but not restricted is also shown. The schematic above the blot shows the position of the Eag probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. The markers are as in Figure 4.4.

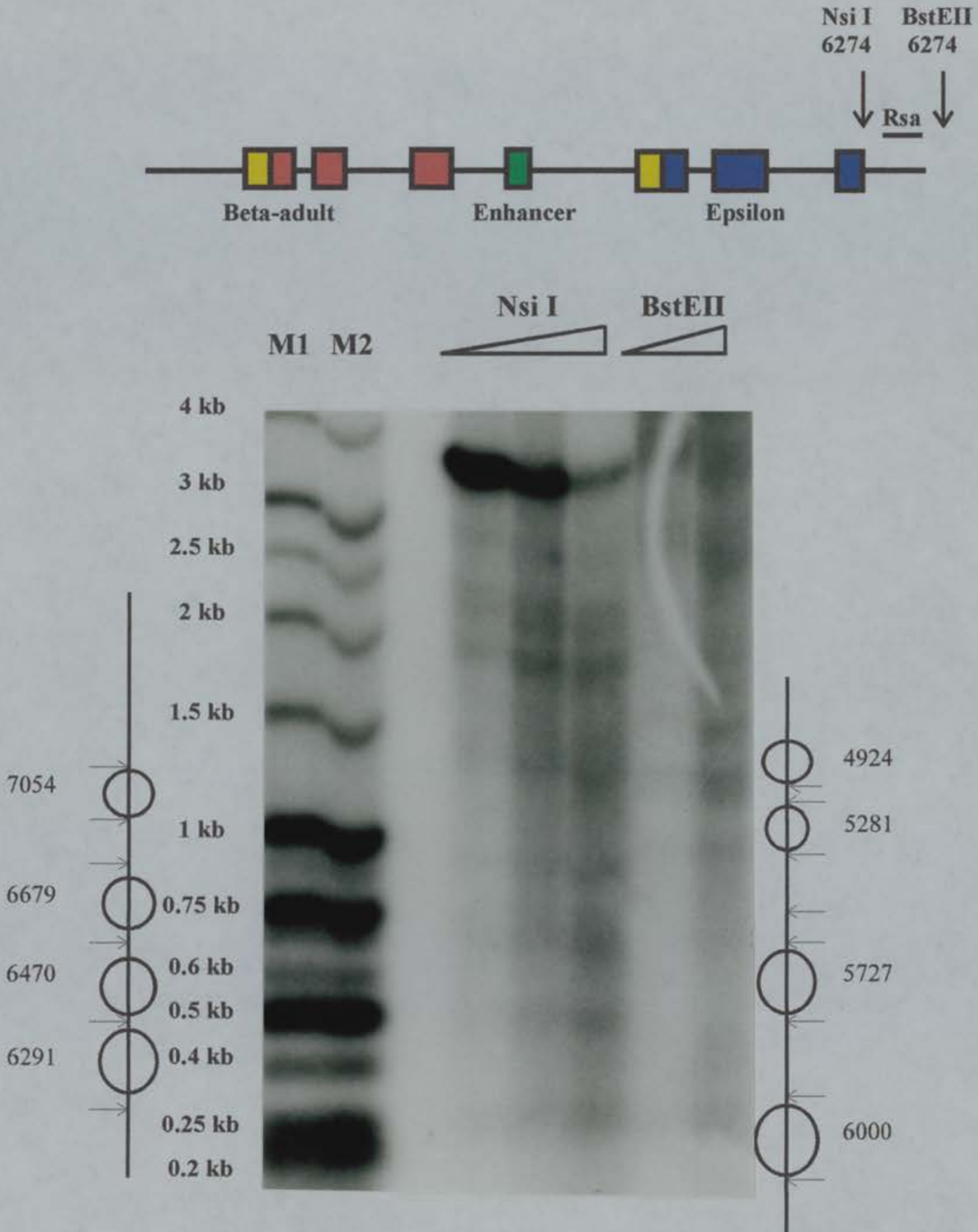


Fig. 4.14 Nucleosome mapping on brain nuclei using the Rsa probe.

A Southern analysis of the DNA recovered from brain nuclei cut with OPCu and then restricted with either BstEII or Nsi I is presented. The schematic above the blot shows the position of the Rsa probe and the restriction sites relative to the beta-adult and epsilon globin genes. The schematic at the side of the blot shows inferred nucleosome positions (circles) and the numbers associated with these indicate the centre of each nucleosome, relative to the beta-adult cap site. The arrows indicate the position of OPCu cleavage sites. The markers are as in Figure 4.4.

Figure 4.15 Nucleosome placement on the β and ε globin genes in brain and red blood cells

+ + + + + = 40 to 50 % of strongest positioning site (23b, Figure 4.16), + + + + = 30 to 40%, + + + = 20 to 30%, + + = 10 to 20% and + = 5 to 10 %.

Placement of nucleosome in brain	Placement of nucleosome in RBC	Limits of accuracy (bp)	Corresponding positioning site *in vitro*	Strength of positioning site *in vitro*
-302		± 20	-289 (5)	+ + +
-116		± 25	-139 (6)	+ + + +
118		± 25	-	-
349		± 20	-	-
	481	± 25	-	-
780	768	± 20	775 (8)	+ + + +
958		± 20	979 (9a)	+ +
1270		± 25	1260 (10b)	+ + + +
1358		± 20	1348	+ +
	(1545)	± 25	1538	+ +
1740		± 25	1732	+ +
	1785	± 25	-	-
1987		± 25	1971 (12)	+ +
2195	2212	± 25	2215 (13)	+ + + +
2370	2385	± 30	2361	+ +
2981		± 25	-	-
3096	3121	± 20	3080 (16)	+ + + +
3290	3300	± 20	3296	+
	3482	± 20	-	-
3780-		± 20	3776 (19b)	+ + +
4010	4025	± 25	-	-
4217	4212	± 25	4237 (20b)	+ +
4389	4403	± 25 / ± 20	4404	+
4592		± 25	4612 (21d)	+ + + +
4786		± 25	4768	+ + + +
4924	4952	± 30	-	-
5281	5270	± 30/ ± 35	5304 (23a)	+ + + + +
5727	5734	± 20	5716/5735	+ +
6000	6010	± 25	-	-
6291	6289	± 25	6266 (26)	+ + + + +
6470	6485	± 20	6466	+ +
6679	6653	± 20	6674/6659	+ +
7054	7048	± 25	off end of map	

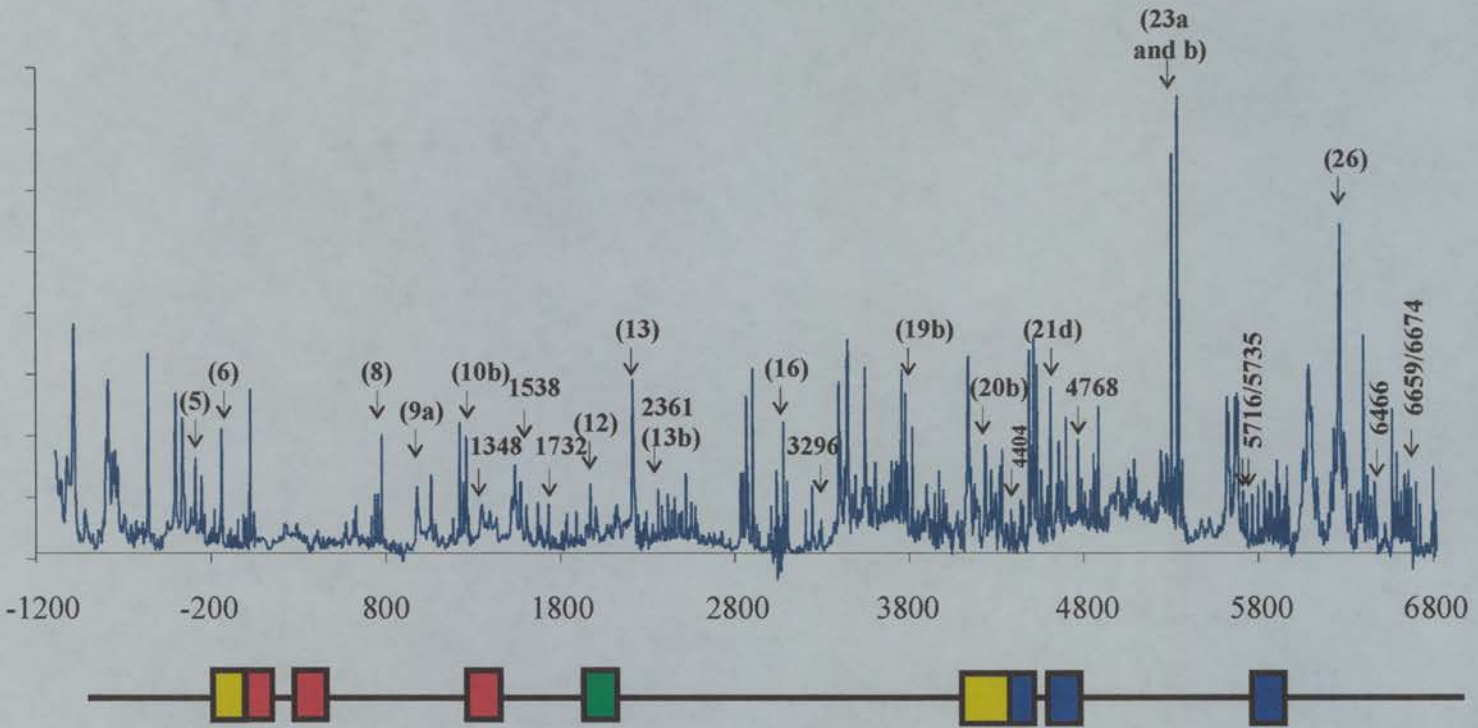


Figure 4.16 Map showing which *in vitro* core histone octamer positioning sites may be occupied *in vivo* on the beta-adult and epsilon globin genes

The map shows the centre of each core histone octamer positioning site, relative to the underlying epsilon and beta-adult genes. Sites which may be occupied in brain and in red blood cells (Figure 4.15) are labelled either with their position relative to the β^A cap site, or their number (in brackets) if they were given one for the purposes of discussion, as in Figure 3.14b.

site identified *in vitro* at 6266 (26 in Figure 4.16), as the difference between the two sites lies just within the range of experimental error. There is some uncertainty concerning the nucleosome positioned at 7054 because one nucleosome boundary is rather weak, and may coincide with a fragment in the OPCu-cut control.

4.3.4 General features of nucleosome placement in brain and red blood cells on the ε and β^A gene region, and a comparison with features of the *in vitro* nucleosome positioning map

Nucleosome placement has been mapped over most of the ~ 8 kb region containing the β^A and ε globin genes (Figures 4.15 and 4.16). Nucleosomes are spaced by ~ 200 on these genes in adult red blood cells (Villeponteau et al., 1992) and in brain (this study, Figure 4.4, lane 'No RE'). Therefore ~ 40 nucleosomes would be expected to occur across this 8 kb region. In brain, out of 40 expected nucleosomes, mapping by indirect end-labelling has identified 29 nucleosomes as being positioned. In red blood cells, 19 positioned nucleosomes have been identified (Figure 4.15).

There are some interesting similarities, and differences, between nucleosome placement in brain and red blood cells. There are fewer positioned nucleosomes across the genes in red blood cells than in brain (Figure 4.15) because nucleosomes around the β^A-globin promoter and the enhancer have been disrupted, due to the recent expression of the β^A-globin gene in this tissue. However, except for this disruption around the promoter and enhancer, the placement of nucleosomes across the β^A and ε genes in brain and red blood cells is very similar.

Nucleosome placement *in vivo* can now be compared with placement *in vitro* where only the DNA sequence and the core histones can influence positioning. This has revealed the degree to which DNA sequence directs nucleosome positioning *in vivo*. Nucleosome placement has been determined *in vivo* with an accuracy of ~ ± 20 bp (Figure 4.15), so if a position identified *in vitro* occurs within this range, is suggests that the *in vitro* position is occupied *in vivo*. Due to the lower resolution of

the end-labelling technique, it cannot be proved that this is the case, but the data are taken as being consistent with this interpretation. Of the 29 positioned nucleosomes in brain, 11 may occupy very strong (+ + + + +/+ + + +) and strong (+ + +) positioning sites identified *in vitro*, while 8 may occupy fairly strong (+ +) positioning sites (the definition of very strong, strong and fairly strong is given in Figure 4.15). In red blood cells, of the 19 positioned nucleosomes, 5 may occupy very strong and strong positioning sites, and 5 may occupy fairly strong positioning sites (Figures 4.15 and 4.16). Therefore it could be argued that DNA sequence makes a large contribution to nucleosome positioning *in vivo*. On the 4 kb region mapped on the β^A-globin gene, there are 25 fairly strong to very strong positioning sites *in vitro*. For a position *in vivo* to be considered to coincide with one of these positions it would have to occur within ± 25 bp; ~ 30 % of the β^A sequence is within 25 bp of a strong positioning site *in vitro*. Therefore, on the basis of chance, as positioned nucleosomes *in vivo* are distributed fairly evenly across the β^A region, it would be expected that ~30% of the sites identified *in vivo* on this region (~5 out of 14) would coincide with a position *in vitro* by probability. In fact, 11 of the nucleosomes positioned on the gene *in vivo* coincide with *in vitro* positioning sites. By applying the same elementary analysis, of the 14 nucleosomes positioned *in vivo* on the region of ε-globin gene mapped *in vitro*, 8 would be expected to coincide with fairly strong to strong positioning sites identified *in vitro*; it is found that 9 coincide. Therefore the nucleosomes positioned on the ε-globin gene in brain coincide with approximately the expected number of *in vitro* positioning sites, as expected by chance.

Comparison of specific features of the *in vivo* and *in vitro* data, which will be discussed in the next section, has revealed several important ways in which DNA sequence-directed nucleosome positioning has the potential to contribute to the regulation of β^A and ε-globin gene expression.

4.3.5 Specific features of nucleosome placement in brain and red blood cells

4.3.5.1 Long-range patterns in the placement of nucleosomes in brain and red blood cells

In brain

Many of the nucleosomes positioned on the β^A and ε-globin genes in brain, particularly those over the β^A gene region, appear, within the limits of resolution of the indirect end-labelling technique, to occupy sites identified as strong positioning sites *in vitro*, (Figure 4.15 and 4.16). Therefore, many of the strong positioning sites identified *in vitro*, particularly those over the β^A gene region which exhibit a 200 bp periodicity, are employed *in vivo* in the brain, where the nucleosome spacing is ~ 200 bp. This agrees with the proposal made in chapter three, that the regular distribution of strong positioning sites over the β^A and ε gene *in vitro* could facilitate the packaging of the gene *in vivo* into an inactive higher order chromatin structure.

On the ε-globin gene fewer of the most prominent positioning sites are used in brain. These prominent sites are spaced by 177 and 186 bp, whilst the average nucleosome spacing in brain is ~200 bp, so optimal occupation of the majority of these strong positioning sites may be incompatible with the nucleosome density in this tissue.

In red blood cells

The placement of positioned nucleosomes over the β^A and ε gene region in red blood cells is very similar to that seen brain, except for some disruption around the promoter of the beta-adult gene and the enhancer. Many of the nucleosome positioning sites occupied in brain also seem to be occupied in red blood cells, allowing for the resolution of the mapping technique (Figure 4.15). The periodicity

```
4150                          4175
.          Sp1 & chicken EKLF  .
CTTGGCTGCTGCCACACCCCTGACCACCCCCTCCTGCCCTCCAGCACAAAC
           Sp1/ CACCC

4200                            4225
.                               .          Sp1
CCCCACACATCCCAGCCAATCGACCCCACAGCGTGGGGTGGGGAGGAGCTG
                CCAAT                    Sp1

4250                          4275
.          GATA 1 & TFIID     .
TCAGCGGTGGATAAAAGCCCCGGGGGTCCGCAGCTCCGCTC
           GATA/ TATA
```

Fig. 4.17 Nucleosome positioning over the ε globin promoter in brain and red blood cells

The positioned nucleosome on the promoter *in vivo* may occupy the positioning site identified *in vitro* at 4237 (20b in Figure 3.14a) relative to the beta-adult cap site. The schematic at the top of this Figure indicates the postion occupied by nucleosome 20b relative to the transcription factor binding sites within the promoter. The sequence of the epsilon globin promoter is shown below this. Sequence motifs shown to be important to ε globin expression by mutational analysis are indicated. Underlines indicate the extent of the footprint at these motifs in extracts from primitive and definitive red blood cells. Names above the line indicate protein factors thought to bind the recognition sites (Mason et al., 1996). The arrow indicates the transcription start.

is less marked than in brain, as disruption of the pattern occurs around the regulatory regions of the beta-adult gene; the promoter region (from ~ -250 to ~150 base pairs) and enhancer (~ 1820 to 2050 base pairs) do not contain positioned nucleosomes, and show increased sensitivity to cuprous phenanthroline.

4.3.5.2 Nucleosome positioning over the ε-globin promoter

A positioned nucleosome occurs on the ε-globin promoter *in vivo* in adult red blood cell nuclei and in brain nuclei at ~ 4212 (Figure 4.13). This is consistent with it occupying the positioning site determined accurately *in vitro* at 4237 (nucleosome 20b in Figure 4.16). Due to the low resolution of the indirect end-labelling technique, it cannot be stated categorically that the nucleosome *in vivo* occupies this position, but the data are taken as consistent with this interpretation. The possibility exists that the nucleosome occupies a site equivalent to the much weaker positioning site identified *in vitro* at 4178. The placement of these nucleosomes relative to the transcription factor binding sites in the promoter is shown in Figure 4.17.

4.3.5.3 Distribution of nucleosomes between introns, exons and flanking regions

Nucleosome placement on the β^A-globin gene in brain and red blood cells (Figure 4.16), reveals that positioned nucleosomes *in vivo* occur in introns and flanking regions, but are generally absent from the coding regions.

On the ε-globin gene, nucleosomes are positioned within introns, exons and flanking regions *in vivo*. Figure 4.18 summarises the nucleosome positions occurring within the exons of the beta-adult and epsilon genes.

4.3.5.4 Nucleosome positioning at splice sites

Theoretical prediction of nucleosomal positioning at eukaryotic splice sites, relying upon the characteristic ~10 bp periodicity of AA(TT) nucleotides in nucleosomal DNA, has suggested that these splice junctions are preferentially

Gene	Position of Exon	Nearest nucleosome dyad in red blood cells	Nucleosome dyad in brain	Corresponding nucleosome *in vitro*
Beta-adult	4 to 162	-	-	-
	265 to 487	481	349	-
	1290 to 1508	-	1270 (weak)	1260
Epsilon	4292 to 4428	4403	4389	4404 (very minor position)
	4537 to 4759	Undetermined	4592	4612
	5733 to 5908	5734	5727	5716 or 5735

Figure 4.18 Nucleosome positioning over the exons of the beta-adult and epsilon globin genes.

The second column (from the left) shows the boundaries of the exon. Columns three and four show the actual nucleosome positions determined *in vivo* in red blood cells and brain. Column 5 gives the *in vitro* nucleosome positioning site to which the position occupied *in vivo* may correspond; where two *in vitro* positions are given, the resolution of the mapping technique used *in vivo* means the data are consistent with either of the sites being occupied. All positions are given relative to the beta-adult cap site.

located within a few (+/- 15) bp of the nucleosome dyad axis (Denisov et al., 1997). The results in brain and in red blood cells (summarised in Figure 4.19) suggest that positioning signals at splice junctions do not make an important contribution to nucleosome positioning on epsilon and beta-adult globin *in vivo*. Of the eight splice sites in brain and red blood cells, only three may have a nucleosome positioned at the predicted place.

In brain, there are no nucleosomes positioned with their dyads close (± 35 bp) to the splice sites at 173, 263, 488, 4429 and 4535 . In red blood cells, there are no nucleosomes positioned at the splice sites at 173, 263, 1296, or 4760. The nucleosome at 481 in red blood cells may have its dyad at the splice site at 488 (within the limits of accuracy of the *in vivo* mapping technique). The nucleosome mapped at 1270 in brain could, within the limits of accuracy of the technique (+/- 25 bps), have its dyad close to the splice site, but if it corresponds to the nucleosome positioning site at 1260 determined *in vitro*, then its dyad is more distant from the splice site than predicted. The nucleosome positioned at 4403 in red blood cells could be located close to the splice site at 4429 The dyad of the nucleosome at 4786 in brain could potentially be close to the splice site at 4760. The dyad of the nucleosome mapped at 5727 in brain (Figure 4.14) and 5734 in red blood cells (Figure 4.15, data not shown) may have its dyad at the splice site at 5731. If this nucleosome corresponds to the nucleosome mapped *in vitro* at 5716, then its dyad does indeed lie close to the splice site.

Junction (position of AG or GT)	Closest nucleosome position in brain	Closest nucleosome position in RBC	Corresponding nucleosome position *in vitro*	AA or TT at splice site?
173	None	None	None	tggccaggtnggt
263	349	None	None	atctctctacaggct
488	None	481 (weak)	None	ttcagggtgagat
1296	1270 (weak)	None	1260	ttcccacagctcct
4429	4389	4403	4404 (Very minor)	ccctggccaggtagg
4535	4592	N/A	4534	ccctctgcaggc
4760	4786	None	4768	ttcagggtgaga
5731	5727	5734	5716 or 5735	tttctcctgcagct

Figure 4.19 Nucleosome positioning at splice sites in brain and red blood cells

The first column (from the left) indicates the predicted position of the nucleosome at the splice site according to Denisov et al., 1997. Columns two and three show the actual nucleosome positions determined *in vivo* in brain and red blood cells. Column four gives the *in vitro* nucleosome positioning site to which the position occupied *in vivo* may correspond; where two *in vitro* positions are given, the resolution of the mapping technique used *in vivo* means the data is consistent with either of the sites being occupied. The fifth column indicates the presence at the splice site of an AA or TT dinucleotide, which is thought to contribute to the positioning signal.
All positions are given relative to the beta-adult cap site.

4.4 DISCUSSION

4.4.1 General features of nucleosome placement in brain and red blood cells

Nucleosome placement has been mapped over most of the ~ 8 kb region containing the β^A and ε-globin genes (Figures 4.15 and 4.16). Nucleosome placement appears to be very similar in brain and red blood cell nuclei. A few differences exist because of the disruption of nucleosomes around the β^A promoter and the enhancer in red blood cells, which was required for the recent expression of the β^A gene in this tissue. Disruption of nucleosomes on the β^A promoter and enhancer precedes gene expression and ensures that transcription factor binding sites are accessible (reviewed in Wolffe, 1992); this is a feature common to many poised and active genes (Felsenfeld, 1992).

Differences in nucleosome placement between the two tissues could be due to factors such as the presence of different transcription factors, and modifications to the histones (for example, acetylation). In brain, a non-erythroid tissue, the globin genes are inactive, and the chromatin domain which contains them is thought to be packaged into an inaccessible higher order structure, because the DNA is resistant to nucleases and hypermethylated (Wood and Felsenfeld, 1982; Mandel and Chambon, 1979). By contrast, in adult circulating red blood cells, transcription has almost completely ceased, but the β^A-globin gene has been recently transcribed and the globin domain retains the characteristics of active chromatin, such as nuclease sensitivity (Mathis et al., 1980; Bellard et al., 1980). The hypersensitive site on the β^A-globin promoter and the enhancer is indicative of nucleosome disruption, a common feature of the regulatory elements of poised or transcribed genes (reviewed in Felsenfeld, 1992). Various mechanisms for the disruption of nucleosomes on regulatory elements have been proposed. These include displacement of nucleosomes by the binding of transcription factors (reviewed in Felsenfeld, 1992). It has been suggested that the stage-specific factors NFE-4 and BGP1 may facilitate

the removal of the nucleosome on the beta-adult promoter (Buckle et al., 1991). Hyperacetylation, which is a feature of the chromatin in erythroid cells (Hebbes et al., 1994 and references therein), may destabilise nucleosomes and enhance the accessibility of DNA bound by nucleosomes to transcription factors (Wolffe, 1997; Wade et al., 1997; Wolffe and Hayes, 1999). HMG14/17, which are preferentially associated with the active globin domain in erythroid cells (Postnikov et al., 1991), are thought to interfere with linker histone binding (Alfonso et al., 1994), which in turn could increase nucleosome mobility and facilitate transcription factor access (Meersseman et al., 1992; reviewed in Wolffe et al., 1997).

The similarity in nucleosome positioning would be created by features shared by the two tissues. The average nucleosome spacing is the same, ~ 200 bp, in brain (this study, Figure 4.4) and adult red blood cells (Villeponteau et al., 1992). This has the potential to influence nucleosome positioning, as discussed in section 3.6.4. If nucleosomes were positioned with respect to a 'boundary' in some regions of the globin gene, then a defined nucleosome spacing would influence where the other nucleosomes were positioned. Various elements have been suggested to act as a 'boundary'; DNase I hypersensitive sites, proteins and other positioned nucleosomes (reviewed in Thoma, 1992). The DNA sequence, of course, is common to both tissues, and its influence has been investigated in this study.

One major aim of this study was to map nucleosome placement on the ε and β^A-globin genes *in vivo*, so that a comparison could be made with the nucleosome positioning data obtained *in vitro* from the same region (this study and Davey et al., 1995). This comparison of the *in vitro* and *in vivo* nucleosome positioning data has revealed the extent to which DNA sequence directs nucleosome positioning in brain and adult red blood cells *in vivo*. Previous studies have shown that DNA sequence directs the positioning of a subset of nucleosomes within the genome, and that these can play an important role in regulating gene expression (reviewed in Simpson, 1990; Thoma, 1992). A more recent example is the study on the *Xenopus* 5S rRNA genes. Differences in sequence-directed nucleosome positioning on the oocyte and somatic 5S rRNA genes are thought to be responsible for the different developmental expression of these two genes (Panetta et al., 1998). Another recent

example is the rat protamine 1 gene, where a nucleosome positioned on the promoter *in vivo* coincides with a positioning site identified *in vitro* that may contribute to repression of the gene (Adroer et al., 1998).

The role of DNA sequence in phasing nucleosomes over a long stretch of DNA, however, is more contentious. Liu and Stein (1997) observed that similar arrays formed on a random selection of 2 or 10 kb chicken DNA fragments *in vitro* and *in vivo*, which would indicate that DNA sequence was responsible for setting up the array. On one fragment they suggested that two positioned nucleosomes, 1000 bp apart, created a phased array of nucleosomes between them in the presence of linker histone. By contrast, Blank and Becker (1996) argue that DNA sequence is subordinate to other factors in determining nucleosome spacing, and its role is limited to 'fine-tuning' nucleosome positions within a repeat pattern established by other principles. They reconstituted nucleosomes onto artificial DNA containing a strong nucleosome positioning signal, repeated at regular intervals of 172 or 208 bp, from *Lytechinus variegatus*. They demonstrated that nucleosome spacing was independent of the spacing of the positioning signal, and that it was determined by electrostatic principles. However, the mechanism of array formation in the *Drosophila* extract used may be rather different, to accommodate the rapid DNA replication and chromatin assembly which occurs in the developing fly. Therefore, this extract may not be representative of cells in general (Liu and Stein, 1997)

My study is important in assessing the contribution of DNA sequence to nucleosome positioning *in vivo*, because it is unique in comparing nucleosome placement on a long (~ 8 kb) stretch of a naturally occurring DNA sequence. Previous attempts have been made to assess the extent to which the DNA sequence contributes to its own packaging and these suggest that 95% of the genome has an affinity for DNA which is no higher than randomly synthesised DNA (Lowary and Widom, 1997). The data from the present study provides an opportunity to analyse how DNA sequence contributes to nucleosome positioning on an 8 kb region containing two genes which show tissue-specific and developmentally regulated expression.

Comparison of specific aspects of the *in vivo* and *in vitro* data, which will be

discussed in greater detail in the next section, reveals that DNA sequence contributes to nucleosome positioning on the β^A and ε genes, and thereby has the potential to regulate their expression. For example, the *in vivo* and *in vitro* data have raised the possibility that the DNA sequence is not limited to a role in 'fine tuning' nucleosome positions. Instead, the conditions in cells are altered so that DNA sequence is predominant in some tissues and subordinate in others, so that positional information in the DNA sequence is used in a flexible and dynamic way. In addition, short-range features of the map, like positioning over promoters, reveal the potential role played by sequence-encoded positional information in determining local chromatin architecture and thereby influencing transcription factor access to the DNA. Furthermore, the distribution of strong positioning sites between the introns, exons and flanking regions provides some clues about the way DNA sequence accommodates its different roles, which include being transcribed, replicated, providing information for splicing, coding for the globin proteins and generating chromatin structure.

4.4.2 Specific features of nucleosome placement in brain and red blood cells

4.4.2.1 Long-range patterns in the placement of positioned nucleosomes

The data presented here shows that many of the strong positioning sites identified *in vitro* on the β^A gene region, which exhibit a 200 bp periodicity (Davey et al., 1995), appear to be occupied *in vivo* (Figures 4.15 and 4.16). This agrees with the proposal that sequence-directed nucleosome positioning leads to a regular distribution of nucleosomes over the β^A gene which facilitates the packaging of the gene in brain and adult red blood cells into an inactive higher order chromatin structure.

By contrast, only a few of the strongest *in vitro* positioning sites on the ε-globin

gene region appear to be occupied in brain and adult red blood cells (Figures 4.15 and 4.16). These strong *in vitro* positioning sites on epsilon globin have a 177 and 186 bp periodicity (this study, chapter 3, Figure 3.15), so occupation of these sites in brain and red blood cells is not so compatible with the average 200 bp spacing (Villeponteau et al., 1992; Sun et al., 1986) which is observed in these tissues. This suggests that the positional information encoded within the ε-globin gene sequence is subordinate in brain and red blood cells to other determinants of nucleosome positioning. The nucleosome spacing observed *in vivo* in definitive (12-day) erythroid cells is ~ 180 bp (Villeponteau et al, 1992); this spacing would be more compatible with the optimal occupation of the strong positioning sites identified *in vitro* on ε-globin. It would be interesting to investigate whether more of the strong *in vitro* positioning sites are in fact occupied in 8-day red blood cells, where epsilon globin is inactive. If this were the case, then the sequence-encoded information would have the potential to create a regular nucleosomal array across the ε-globin gene, which could have the capacity to facilitate its packaging into an inactive higher order structure at that developmental stage.

In summary, the nucleosome spacing observed in definitive erythroid cells where the ε-globin gene is inactive, and β^A-globin active, is more compatible with optimal occupation of the strong positioning sites identified *in vitro* on the ε-globin gene, and less compatible with optimal occupation of the strong *in vitro* positioning sites on β^A-globin. The nucleosome spacing observed in adult red blood cells and brain, where β^A-globin is inactive, is more compatible with optimal occupation of the strong positioning sites identified *in vitro* on the β^A-globin gene, and in fact many of these *in vitro* positioning sites do appear to be occupied in these tissues (Figures 4.15 and 4.16). By contrast, the nucleosome spacing observed in brain and red blood cells is less compatible with optimal occupation of the strong *in vitro* positioning sites identified on the ε-globin gene, and it does indeed appear that only a few of the strongest *in vitro* positioning sites are occupied in these tissues. Optimal occupation of the strong positioning sites which occur with a regular spacing may favour the formation of a regular nucleosome array. Therefore, the periodicity of strong

nucleosome positioning sequences on the globin genes is such that it appears to facilitate formation of regular nucleosome arrays on ε-globin at day 12 in development when ε-globin is inactive, and β^A active, whilst regular array formation is facilitated on β^A at a later stage in development, when β^A is inactive.

The increase in nucleosome spacing on the globin genes, in combination with the difference in spacing of the nucleosome positioning sequences on the embryonic and adult genes, may create an array of strongly positioned nucleosomes on the two genes at different developmental stages. What functional consequence might this have? Others have hypothesised that differences in nucleosome spacing on the globin genes may lead to distinctive higher order structures (Liu et al., 1993; Evans et al., 1990) The sequence-directed formation of nucleosome arrays at different stages could in turn facilitate the folding of the β^A and ε-globin genes into a more stable higher order structure at different times in development. It has been proposed that a regular array of positioned nucleosomes would form a more stable higher order higher order structure because a regular linker length allows a regular higher order structure to be formed with more extensive internucleosome contacts than in an irregular structure (reviewed in Ramakrishnan et al., 1997). Gene expression is thought to take place in the context of a higher order structure with only transient unfolding during transcription itself (Andersson et al., 1982; Ericsson et al., 1990). Therefore a less stable higher order structure would facilitate the unfolding of active genes, and a more stable higher order structure created by an array of positioned nucleosomes would reinforce the repression of an inactive gene. It is also possible that the difference in spacing of nucleosome positioning sequences on the β^A and ε-globin genes has further significance. With the current level of understanding, it is uncertain what structural and functional differences in chromatin structure there might be between nucleosomal arrays with repeat lengths of 180 or 200 bp. It is possible that nucleosome spacing could affect the stability of the higher order structure formed; for example, it was suggested that one of the models proposed for higher order chromatin structure could not accommodate nucleosomes spaced at 163 to 176 bp intervals (Butler, 1984). Until the nature of the higher order structure is

better defined it is uncertain whether the there is a significant difference between arrays with nucleosomes spaced at 186 and 200 bp.

In summary, these data have led to the hypothesis that certain conditions, such as nucleosome density, differ between different cell types, so that DNA sequence is predominant in some cell types and subordinate in others. The *in vitro* positioning sites can be considered as the 'default' positions, and other factors *in vivo* modulate the recognition and interpretation of the positioning information in the globin DNA sequence. In this way, the positional information in the DNA sequence is used in a flexible and dynamic way to influence the higher order chromatin structure of the globin genes, and thereby control their expression.

4.4.2.2 Nucleosome positioning on the epsilon promoter

Globin gene expression is tissue-specific, and the expression of individual genes is restricted to specific times during development. The data presented here show that there is a positioned nucleosome on the epsilon globin promoter which could potentially contribute to the tissue-specific and temporal regulation of the epsilon globin gene. This positioned nucleosome occurs on the epsilon promoter *in vivo* in adult RBC and in brain at ~ 4212 (Figure 4.13), which is consistent with it occupying the positioning site determined accurately *in vitro* at 4237 (20b in Figure 4.16). This nucleosome would occlude the Sp1 recognition site (located within a few base pairs of the dyad), the CCAAT site (~ 20 bp from dyad), and the GATA and TATA motifs (~25 bp from dyad). These sites are involved in activating epsilon in the presence and absence of the enhancer. The Sp1/CACCC site would lie right at the boundary of this nucleosome; these sites are required for interaction with the enhancer (Figure 4.17). Many DNA elements and transcription factors, in addition to those above, are involved in regulating epsilon, and the influence of a positioned nucleosome must be understood within this context.

In erythroid tissues where globin genes are expressed, the globin gene cluster is contained within a 30 kb domain which exhibits many of the modifications

associated with active chromatin, such as hyperacetylation (Hebbes et al., 1994), sensitivity to nucleases (Stalder et al., 1980; Hebbes et al., 1994) and hypersensitivity to nucleases (reviewed in Wolffe, 1992). In non-erythroid tissues, the domain is highly methylated (Mandel and Chambon, 1979), and more resistant to nucleases than in its active state (Wood and Felsenfeld, 1982). The positioned nucleosome has the potential to play a repressive role in non-erythroid cells, because it is located over a number of the important transcription factor binding sites, so it has the capacity to restrict the binding of transcription factors necessary for gene expression. Many other factors, such as methylation (Mandel and Chambon, 1979), loss of nuclease sensitivity (Wood and Felsenfeld, 1982), hypoacetylation and loss of nuclease hypersensitivity at promoters, the enhancer and upstream sites correlate with inactivation of the globin gene cluster (reviewed in Wolffe, 1992), but the positioned nucleosome may still make a contribution. GATA 1 is a tissue-specific transcription factor, more or less restricted to the erythroid lineage (reviewed in Orkin, 1995), therefore reduced occupancy of the GATA site could aid the formation of a positioned nucleosome on the epsilon promoter which would in turn occlude the binding sites of ubiquitous transcription factors like Sp1.

There is strong evidence to suggest that the temporal control of the epsilon and beta-adult genes is effected by a mutually exclusive competition between these genes for interaction with their shared enhancer (Choi and Engel, 1988; Foley and Engel, 1992). Epsilon competes more effectively for the shared enhancer in 5 day (primitive) red blood cells, so that it is expressed whilst beta-adult is silent; at about 8 days in development, a shift in the relative stabilities of the promoter-enhancer complexes occurs, and epsilon expression decreases whilst beta-adult is upregulated, until in 12-day (definitive) red blood cells, epsilon is fully silenced.

The competition model for globin gene switching described above suggests transcription factors which bind the promoter and interact with proteins bound at the enhancer will decide the relative stability of the epsilon promoter-enhancer interaction and the beta-adult promoter-enhancer interaction, and thereby regulate temporal expression of these genes. Positioned nucleosomes which regulate the access of these transcription factors to the DNA could potentially play a role in

globin gene switching. Many of the studies carried out on the developmental regulation of β^{A} and ε-globin (for example, Foley and Engel, 1992) have involved transient transfections, where the mechanism of switching will take place in the absence of a proper chromatin structure. Researchers have suggested that the binding of the stage-specific factors NFE-4 and βCTF to the beta-adult promoter, and their interaction with proteins at the enhancer in definitive red-blood cells is the decisive factor in beta-adult/epsilon switching (Gallarda et al., 1989). However, it was suggested that chromatin structure, which was not considered in their study, could still play a role (Foley and Engel, 1992). In agreement with this, it was proposed that the precisely positioned nucleosome on the beta-adult promoter which occludes most of the important transcription factor binding sites could be displaced by the binding of NFE-4 (Buckle et al., 1991).

A decrease in the stability of the epsilon promoter-enhancer interaction may also contribute in the switch from epsilon to beta-adult expression; enhancement of epsilon expression in a construct containing the enhancer and epsilon is only 10-fold less in definitive cells than primitive cells (Mason et al., 1996). No stage-specific factors have been identified on the epsilon promoter, but changes in the levels of non stage-specific factors may influence switching; GATA 1 and Sp1 levels are higher in primitive than definitive cells (Minie et al., 1992) whilst GATA 3 is higher in definitive cells (Leonard et al., 1993). It has been proposed that the presence of GATA 1 at the enhancer increases the stability of its interaction with the epsilon promoter, whilst GATA 3 at the enhancer stabilises its interaction with the beta-adult promoter. The decrease in GATA 1 levels could also affect the occupancy of the GATA site on the epsilon promoter. The nucleosome positioned on the epsilon promoter in red blood cells and brain which was revealed by this study occludes this GATA 1 / TATA site and most of the major transcription factor binding sites (the Sp1 site at 4234 and the CCAAT site), (Figure 4.17). As GATA 1 and Sp1 levels are lower in definitive red blood cells, it is possible that this allows nucleosomes to compete more effectively with transcription factors for binding sites on the promoter, resulting in a positioned nucleosome on the epsilon promoter which restricts the access of the remaining GATA 1, Sp1 and also chicken EKLF, and other

transcription factors. In this way, the nucleosome on the epsilon promoter could make a contribution to repressing epsilon transcription in definitive cells.

The precise positioning of the nucleosome may also be important for the activation of epsilon globin in primitive cells. Studies have suggested that an active promoter conformation may be achieved in several ways; competition between transcription factors and nucleosomes for binding sites on the promoter following replication may lead to the loss, rearrangement or repositioning of nucleosomes. Alternatively, some transcription factors bind to recognition sites within nucleosomes, leading to rearrangement or displacement of the nucleosome, allowing other transcription factors to bind subsequently. The mechanisms which activate the epsilon promoter are uncertain, but the data here allow speculations to be made. The occupancy of the GATA/TATA site by GATA 1 in erythroid cells may help displace/rearrange the nucleosome. GATA 1 is capable of accessing a site in nucleosomal DNA, and 3 GATA sites within a nucleosome have been shown to be capable of causing nucleosome rearrangement (Boyes et al., 1998). There is only one GATA site at the epsilon promoter, but it is possible that it could interact with other proteins bound at the epsilon promoter to alter nucleosome structure. It is interesting that the CACCC site, which appears to be important in the promoter-enhancer interaction (Mason et al., 1996), occurs right at the nucleosome boundary. It would be interesting to investigate whether the interaction of the EKLF-related protein and Sp1 at this site, with a protein at the CCAAT or GATA site could affect the stability of the positioned nucleosome. By contrast, a similar mechanism involving GATA1 could not remove the positioned nucleosome from the beta-adult promoter as the GATA site on this promoter occurs in the linker DNA outside the positioned nucleosome.

4.4.3 Distribution of nucleosomes between introns, exons and flanking regions

The nucleosome positioning maps of the beta-adult gene in brain and red blood

cells reveal that positioned nucleosomes *in vivo* occur in introns and flanking regions, but are less frequent than expected in the coding regions. This may result partly from an incompatibility between the sequence requirements for coding and nucleosome positioning, as discussed in section 3.4.4. However, the other determinants of nucleosome positioning which operate *in vivo* have also failed to create positioned nucleosomes on the exons. The presence of nucleosomes on the introns, however, shows its recent transcription has not substantially disrupted nucleosome positioning on the beta-adult gene. This agrees with previous work which suggested that any unfolding of the globin genes for transcription is only transient (Andersson et al., 1982; Ericsson et al., 1990;). The restriction of strong positioning sites to the introns and flanking regions may enable the beta-adult gene to accommodate its roles in creating chromatin structure and encoding the globin protein. It has been proposed previously that introns and flanking regions may play a role as chromatin organisers (Zuckerandl, 1991; 1997). Expression of certain transgenes without introns in mice has resulted in reduced levels of expression, possibly due to a disruption in the chromatin structure of the gene (Liu et al., 1995).

On the epsilon globin gene, nucleosomes are positioned within introns, exons and flanking regions. The epsilon globin gene also showed a lower-than-expected occurrence of nucleosomes on exons *in vitro*, but other determinants of nucleosome positioning which operate *in vivo* have created positioned nucleosomes.

4.4.4 Nucleosome placement over splice sites

Theoretical prediction of nucleosome positioning has suggested that splice junctions are preferentially located within a few (+/- 10) base pairs of the nucleosome dyad axis, and that part of the splice site consensus forms this positioning signal (Denisov et al., 1997).

The data presented here indicates that positioning signals at splice junctions do not make an important contribution to nucleosome positioning on epsilon and beta-adult globin brain and red blood cells. It has been proposed that a single motif like the splice-site signal is rarely the major factor in sequence-directed nucleosome

positioning; in general, multiple sequence elements in the DNA wrapped around the nucleosome make an additive contribution (Thoma, 1992), and furthermore *in vivo* many other factors can modulate nucleosome positioning .

Thesis: Final Perspective

Nucleosome placement has now been mapped over ~ 8kb of the chicken β^A and ε-globin gene region *in vitro* and *in vivo*, in brain and red blood cells (this study, and Davey et al., 1995). These data were gathered to answer three main questions. Firstly, to define what DNA sequence patterns played a role in nucleosome positioning. Secondly, to determine what contribution DNA sequence made to nucleosome positioning on the globin genes. Thirdly, to elucidate the role played by nucleosome positioning and, in particular, by sequence-directed nucleosome positioning, in regulating globin gene expression.

The *in vitro* nucleosome positioning data generated in this study can be used to determine which sequence patterns contribute to nucleosome positioning. A limited analysis has been carried out thus far, searching the strongest nucleosome positioning sequences for a number of sequence patterns which previous studies have suggested as nucleosome positioning signals. In the future, the nucleosome positioning data from the ε globin gene (this study), will be combined with data from the β^A-globin gene region (Davey et al., 1995) and data from the β-lactoglobulin gene, which is currently being mapped in this laboratory. This will produce a database of approximately 18,000 nucleosome positioning sequences. This will be superior to those in the literature (for example, Satchwell et al., 1986; Ulyanov et al., 1995; Ioshikhes et al., 1993), because it is much larger, the nucleosome positions are all mapped at high resolution, and the relative binding affinities are known, so that a 'weighting' can be applied for the patterns found in stronger positioning sites. This database can be searched for sequences which have been suggested to contribute to nucleosome positioning by previous studies (Satchwell et al, 1986; Ulyanov & Stormo; 1995; Godde et al., 1996; Ioshikhes et al., 1996; Wang et al., 1996; Widlund et al., 1997; Lowary and Widom, 1998). It can be analysed for novel positioning signals by techniques such as multiple alignment algorithms (Ioshikhes et al., 1996), multi-alphabet consensus algorithms (Ulyanov and

Stormo, 1995), hidden Markov models (Baldi et al., 1996), or other computational methods.

Numerous studies have shown that DNA sequence may make an important contribution to nucleosome positioning *in vivo* (reviewed in Thoma et al., 1992; Simpson et al., 1990; Butinelli et al., 1993). Nucleosome placement on the β^A and ε-globin gene region has been compared *in vivo* and *in vitro* to elucidate the role played by DNA sequence in directing nucleosome positioning *in vivo*. This has revealed that many of the strongest nucleosome positioning sites identified *in vitro* on the β^A globin gene region, and a smaller number of strong positioning sites on the ε-globin gene region, appear to be occupied in brain and red blood cells (Figure 4.15). To quantify the contribution made, future experiments could use competitive reconstitution (Shrader and Crothers, 1989) to calculate the free energy of nucleosome formation for some of the some the nucleosome positions on the β^A and ε globin genes, so that the relative affinities over the entire region (I am assuming my scale is linear) can be expressed as ΔG for nucleosome formation. It has been demonstrated that, in some cases at least, manipulating the DNA sequence to cause relatively small differences in ΔG for nucleosome formation (of ~0.4 kcal/mol) was sufficient to cause changes in rotational or translational positioning (Butinelli et al., 1995).

Comparison of nucleosome positioning *in vitro* and *in vivo* reveals that a number of nucleosomes on β^A and ε globin genes are unlikely to be determined by the DNA sequence. Comparison of positioning in brain and red blood cells also suggests that certain differences between the two cell types are capable of modulating whether nucleosome positioning signals in the DNA sequence are used to position nucleosomes or not. In red blood cells, the DNA is hypomethylated, the chromatin is hyperacetylated and associated with HMG14/17, and the globin domain has a more open chromatin structure, whilst in brain the chromatin domain has a more closed structure, and the DNA is hypermethylated and the core histones are hypoacetylated. The differences between brain and red blood cell which have the potential to influence chromatin

structure include greater methylation of the DNA (Davey et al., 1997), the presence of linker histone (reviewed in Wolffe et al., 1997), the presence of HMG proteins (Alfonso et al., 1994), hyperacetylation of the core histones (Wolffe et al., 1997), the presence of cell-specific transcription factors and the presence of DNase hypersensitive sites, other positioned nucleosomes, and proteins that act as boundary elements (reviewed in Thoma et al., 1992). The influence of some of these factors, such as the presence of linker histone, acetylation of the core histones and methylation of the DNA on nucleosome positioning can be studied *in vitro*, using monomer DNA extension. For example, it has been demonstrated that methylation of the β^A–globin promoter greatly reduces the relative strength of one particular nucleosome positioning sequence (Davey et al., 1997). It is possible to study the influence of non-histone proteins, such as HMG's or transcription factors, on nucleosome positioning *in vitro* (Panetta et al., 1998).

In brain and red blood cells, the nucleosome density on the DNA, a determinant of the nucleosome repeat length, is very similar. However, nucleosome repeat length on the globin genes varies during development; for example it is shorter in primitive (5-day) reticulocytes than in mature (19-day) and adult erythrocytes. Therefore, nucleosome density may cause differences in nucleosome positioning between erythroid cells at different stages in development. For example, strong positioning sites have been identified *in vitro* at 186 bp intervals on the ε-globin gene. In 8 day red blood cells, where a nucleosome spacing of ~ 180 bp is observed (Villeponteau et al., 1992), it would be possible for these strong *in vitro* positioning sites to be occupied. By contrast, in adult red blood cells, where the observed nucleosome spacing is 200 bp (Villeponteau et al., 1992), the nucleosome density is not compatible with an array of positioned nucleosomes occupying these sites. Therefore, it can be hypothesised that a small increase in nucleosome density could radically alter the pattern of nucleosome positioning. Using the *in vitro* positioning map for the β^A and ε-globin genes which shows the relative strength of all potential nucleosome positioning sequences, a prediction of nucleosome positioning over the two genes for a given nucleosome repeat

length could be made, based on an optimisation of the free energy for the formation of the array of positioned nucleosomes.

This study on nucleosome placement has revealed additional ways in which chromatin structure could potentially contribute to regulating expression of the globin genes. Previous studies have demonstrated that the globin genes are expressed only in erythroid cells, and the individual genes are expressed at specific times in development (Orkin, 1990; Orkin 1996; Felsenfeld, 1993). The activation of the chicken globin genes in erythroid cells is preceded by the formation of a more 'open' chromatin structure on the domain which encompasses the β-globin cluster, as indicated by the increased nuclease sensitivity of the globin genes in erythroid tissues. This 'opening up' of the domain plays the crucial role in controlling the erythroid specificity of globin expression (Felsenfeld, 1993). The positioned nucleosome found on the ε-globin promoter in brain, however, which occludes the major ubiquitous and erythroid-specific transcription factors could still reinforce repression in non-erythroid cells. My data also reveal a second way in which DNA-sequence directed nucleosome positioning could contribute to erythroid specificity. The DNA sequence of the β^A-globin gene directs the formation of a regular nucleosome array on the gene which could facilitate its packaging into a higher order chromatin structure in inactive tissues like brain (Davey et al., 1995). Since these nucleosome positioning signals reside primarily in the introns and flanking regions, it would be intriguing to see if a stably transfected copy of the gene without introns would show a different pattern of tissue-specific expression.

Previous work has demonstrated that the stage-specific regulation of the genes is regulated by the interaction of enhancer elements with individual globin promoters, which involves changes in chromatin structure and binding of transcription factors at these regulatory elements (Orkin, 1995; Felsenfeld, 1993). It has been proposed that temporal expression of the β^A and ε-globin genes is regulated by competition for their shared enhancer (Choi and Engel, 1988; Foley and Engel, 1992). This competition is influenced by the relative stabilities of the promoter-enhancer complexes, and it is

thought to play a decisive role in controlling the switch in expression from ε to β^A which occurs at ~ day eight in development. The promoter-enhancer stability, and therefore the switch, is influenced by the binding of stage-specific transcription factors such as NFE-4 to the β^A–globin promoter which enhances the stability of its interaction with the enhancer (Foley and Engel, 1992). There is evidence to show that a decrease in the stability of the ε promoter:enhancer interaction, due to a decrease in the levels of non-stage specific factors such as GATA1 and Sp1, may play a role in the switch (Mason et al., 1996).

The positioned nucleosome on the ε-globin promoter revealed by this study, which occludes most of the major transcription factor binding sites, has the potential to contribute to the β^A/ε switch in expression by decreasing the stability of the promoter-enhancer interaction. This nucleosome has certain transcription factor binding sites very close to its boundary, and it is interesting to speculate if this facilitates its disruption prior to gene activation. There is a precedent for this; it has been suggested that binding of the transcription factors BGP1 and NFE-4 to the nucleosome on the β^A promoter could bring about its displacement (Buckle et al., 1991). It is possible to study the interaction of transcription factors and nucleosomes *in vitro* (Panetta et al., 1998), so the mechanisms by which nucleosomes on the β^A and ε-globin gene promoters were displaced could be investigated. The nucleosome positioned on the ε globin promoter may occupy a strong positioning site identified *in vitro*. It would be interesting to confirm placement of this nucleosome at a strong nucleosome positioning sequence by mapping the position of the nucleosome *in vivo* with a technique such as primer extension which gives higher resolution data (Mueller and Wold, 1989).

This study also raises the possibility that the difference in the spacing of strong *in vitro* positioning sites observed on the β^A and ε genes serves to facilitate their packaging into a more stable higher order structure at the correct developmental stage. To further investigate this hypothesis, it will be necessary to map nucleosome placement on the β^A and ε-globin genes in 5-day red blood cells. In addition, it would be informative to map

nucleosome placement *in vitro* and *in vivo* on the other embryonic gene ρ-globin, and the foetal gene β^{H}–globin to see if the sequence-directed formation of nucleosome arrays could facilitate packaging of these genes at the correct developmental stage.

It is difficult to evaluate the contribution made by the aspects of chromatin structure discovered in this study to temporal globin gene expression because most of the previous work on the mechanism of switching was carried out in the absence of chromatin using transient transfection methods. It is interesting that Mason et al. (1996) found that competition between the β^{A} and ε-globin genes for their enhancer was not required for the correct temporal expression of these genes when the chicken β-globin cluster was introduced into mice. This could reflect differences between mice and chicken in the regulation of the globin genes, or it could indicate that the chromatin structure of the globin genes influences their temporal expression. Future work with stably transfected constructs, transgenic systems, or synthetic nuclei (Barton and Emerson, 1994) will deepen our understanding of the role chromatin structure plays in regulating globin gene expression.

REFERENCES

Abruzzo, L.V. & Reitman, M. (1994) Enhancer activity of upstream hypersensitive site 2 of the chicken β-globin gene cluster is mediated by GATA-1 sites. J. Biol. Chem., 269, 32565-32571.

Adachi, Y., Luke, M. & Laemmli, U.K. (1991) Chromosome assembly in vitro – topoisomerase II is required for condensation. Cell, 64, 137-148.

Adams, C.C. & Workman, J.L. (1993) Nucleosome displacement in transcription. Cell, 72, 305-308.

Adams, C.C. & Workman, J.L. (1995) Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. Mol. Cell. Biol., 15, 1405-1421.

Adroer, R. and Oliva, R. (1998) Nucleosome positioning in the rat protamine 1 gene *in vivo* and *in vitro*. Biochim. Biophys. Acta, 1442, 252-260.

Affolter, M., Cote, J., Renaud, J., and Ruizcarrillo, A. (1987). Regulation of histone and beta-globin gene-expression during differentiation of chicken erythroid-cells. Mol. Cell. Biol., 7, 3663-3672.

Aizawa, S., Nishino, H., Saito, K., Kimura, K., Shirakawa, H. & Yoshida, M. (1994) Stimulation of transcription in cultured-cells by high-mobility group protein-1 - essential role of the acidic carboxyl-terminal region. Biochemistry, 33, 14690-14695.

Alfonso, P.J., Crippa, M.P., Hayes, J.J. & Bustin, M. (1994) The footprint of chromosomal-proteins HMG-14 and HMG-17 on chromatin subunits. J. Mol. Biol., 236, 189-198.

Allan, J., Hartman, P.G., Crane-Robinson, C. & Aviles, F.X. (1980) The structure of histone H1 and its location in chromatin. Nature, 288, 675-679.

Allan, J., Mitchell, T., Harborne, N., Bohm, L. and Crane-Robinson, C. (1986) Roles of H1 domains in determining higher-order chromatin structure and H1 location. J. Mol. Biol., 187, 591-601.

Almer, A., Rudolph, H., Hinnen, A. & Horz, W. (1986) Removal of positioned nucleosomes from the yeast pho5 promoter upon pho5 induction releases additional upstream activating DNA elements. EMBO J., 5, 2689-2696.

Andersson, K., Mahr, R., Björkroth, B. and Daneholt, B. (1982) Rapid deformation of the thick chromosome fibre upon completion of RNA synthesis at the Balbiani ring genes in *Chironomous tentans*. Chromosoma, 87, 33-48.

Archer, T.K., Cordingley, M.G., Wolford, R.G. & Hager, G.L. (1991) Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary-tumor virus promoter. Mol. Cell. Biol., 11, 688-698.

Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E. & Moudrianakis, E.N. (1991) The nucleosomal core histone octamer at 3.1 Å resolution - a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. USA, 88, 10148-10152.

Arents, G. & Moudrianakis, E.N. (1993) Topography of the histone octamer surface - repeating structural motifs utilized in the docking of nucleosomal DNA. Proc. Natl. Acad. Sci. USA, 90, 10489-10493.

Arents, G. & Moudrianakis, E.N. (1995) The histone fold - a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc. Natl. Acad. Sci. USA, 92, 11170-11174.

Aubin, R.J., Frechette, A., De-Murcia, G., Mandel, P., Lord, A., Grondin, G. & Poirier, G.G. (1983) Correlation between endogenous nucleosomal hyper(ADP-ribosyl)ation of histone H1 and the induction of chromatin relaxation. EMBO J., 2, 1685-1693.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (eds) (1997) Current Protocols in Molecular Biology, volumes 1-3 (Wiley).

Aviles, J., Chapman, G.E., Kneale, G.G., Crane-Robinson, C. & Bradbury, E.M. (1978) The conformation of histone H5. Eur. J. Biochem., 88, 363-371.

Baldi, P., Brunak, S., Chauvin, Y. & Krogh, A. (1996) Naturally occurring nucleosome positioning signals in human exons and introns. J. Mol. Biol., 263, 503-510.

Baron, M.H. (1996). Developmental regulation of the vertebrate globin multigene family. Gene Expression 6, 129-137.

Barton, C.B. and Emerson, B.M. (1994). Regulated expression of the beta-globin gene locus in synthetic nuclei. Genes Dev., 8, 2453-2465.

Bavykin, S.G., Usachenko, S.I., Zalensky, A.O. & Mirzabekov, A.D. (1990) Structure of nucleosomes and organization of internucleosomal DNA in chromatin. J. Mol. Biol., 212, 495-511.

Becker, P.B. & Wu, C. (1992) Cell-free system for assembly of transcriptionally repressed chromatin from Drosophila embryos. Mol. Cell. Biol., 12, 2241-2249.

Bellard, M., Kuo, M.T., Dretzen, G. and Chambon, P. (1980) Nucl. Acids Res., 12, 2737-2750.

Benezra R., Cantor, C.R. and Axel, R. (1986) Nucleosomes are phased along the mouse beta-major globin gene in erythroid and non-erythroid cells. Cell, 44, 697-704.

Bergman, L.W. (1986) A 278 bp DNA fragment containing the upstream activator sequence determines precise nucleosome positioning of the yeast PHO5 gene. J. Cell Biochem., No. S10D, 111.

Bina, M. (1994) Periodicity of dinucleotides in nucleosomes derived from simian virus 40 chromatin. J. Mol. Biol., 235, 198-208.

Bird, A. (1992) The essentials of DNA methylation. Cell, 70, 5-8.

Blank T.A. and Becker, P.B. (1995) Electrostatic mechanism of nucleosome spacing. J. Mol. Biol., 252, 305-313.

Bolshoy, A. (1995) CC dinucleotides contribute to the bending of DNA in chromatin. Nature Struc. Biol., 2, 446-448.

Bolshoy, A., Ioshikhes, I. & Trifonov, E.N. (1996) Applicability of the multiple alignment algorithm for detection of weak patterns: periodically distributed DNA pattern as a study case. CABIOS, 12, 383-389.

Bolshoy, A., McNamara, P., Harrington, R.E. & Trifonov, E.N. (1991) Curved DNA without AA - experimental estimation of all 16 DNA wedge angles. Proc. Natl. Acad. Sci. USA, 88, 2312-2316.

Boulikas, T. (1994). Transcription factor binding sites in the matrix attachment region (MAR) of the chicken alpha globin gene. J. Cell. Biochem. 55, 513-529.

Bouvet, P., Dimitrov, S. & Wolffe, A.P. (1994) Specific regulation of Xenopus chromosomal 5S rRNA gene transcription in vivo by histone H1. Genes Dev, 8, 1147-1159.

Boyes, J. and Felsenfeld, G. (1996). Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. EMBO J., 15, 2496-2507.

Boyes, J., Omichinski, J., Clark, D., Pikaart, M. and Felsenfeld, G. Perturbation of nucleosome structure by the erythroid transcription factor GATA-1. J. Mol. Biol., 279, 529-544.

Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D. & Broach, J.R. (1993) Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. Genes Dev, 7, 592-604.

Bresnick, E.H., Bustin, M., Marsaud, V., Richardfoy, H. & Hager, G.L. (1992) The transcriptionally active MMTV promoter is depleted of histone H1. Nucl. Acids Res., 20, 273-278.

Broders, F., Razin, S., Farache, G., Moreau, J., and Scherrer, K. (1986). Correlations of repetitive and AT-rich DNA segments within the chicken globin gene domains. Molecular Biology Reports 11, 177-187.

Browne, I.R. and Sutcliffe, J.G. (1987) A typical nucleosome spacing of rat neuronal identifier elements in non-neuronal chromatin. Nucl. Acids Res., 15, 3563-3567.

Brownell, J.E. & Allis, C.D. (1996) Special hats for special occasions - linking histone acetylation to chromatin assembly and gene activation. Curr. Opin. Genet. Dev., 6, 176-184.

Brukner, I., Sanchez, R., Suck, D. & Pongor, S. (1995a) Trinucleotide models for DNA bending propensity - comparison of models based on DNaseI digestion and nucleosome packaging data. J. Biomol. Struc. Dynam., 13, 309-317.

Brukner, I., Sanchez, R., Suck, D. & Pongor, S. (1995b) Sequence-dependent bending propensity of DNA as revealed by DNase I parameters for trinucleotides. EMBO J., 14, 1812-1818.

Buckle, R., Balmer, M., Yenidunya, A., and Allan, J. (1991). The promoter of the inactive chicken Beta-Globin gene contains precisely positioned nucleosomes. Nucl. Acids Res. 19, 1219-1226.

Buschhausen G., Wittig, B., Graessman, M. & Graessman, A. (1987) Chromatin structure is required to block transcription of the methylated HSV thymidine kinase gene. Proc. Natl. Acad. Sci. USA, 84, 177-1181.

Bustin, M., Trieschmann, I., and Postnikov, Y.V. (1995) The HMG-14/17 chromosomal protein family- architectural elements that enhance transcription from chromatin templates. Seminars in Cell Biology, 6, 247-255.

Butler,P.J.G. (1983) Higher order structure of nucleosome oligomers from short repeat chromatin. CRC Critic. Rev. in Biochem. 15 57-91.

Butler, P.J.G. (1984) A defined structure of the 30-nm chromatin fiber which accommodates different nucleosomal repeat lengths. EMBO J., 3, 2599-2604.

Butler P.J.G. and Thomas, J.O. (1990) Changes in chromatin folding in solution J. Mol. Biol., 140, 505-529.

Butler, P.J.G. and Thomas, J.O. (1998) Dinucleosomes show compaction by ionic strength, consistent with bending of linker DNA. J. Mol. Biol., 281, 401-407.

Buttinelli, M., Negri, R., Di Marcotullio, L. & Di Mauro E. (1995) Changing nucleosome positions through modification of the rotational information. Proc. Natl. Acad. Sci. USA, 92, 10747-10751.

Buttinelli, M., Di Mauro, E. and Negri, R. (1993) Multiple nucleosome positioning with unique rotational setting for the *Saccharomyces cerevisiae* 5S rRNA gene *in vitro* and *in vivo*. Proc. Natl. Acad. Sci. USA, 90, 9315-9319.

Cacchione, S., Cerone, M.A., De Santis, P. & Savino, M. (1995) Superstructural features of the upstream regulatory regions of 2 pea rbcs genes and nucleosomes positioning - theoretical prediction and experimental evaluation. Biophys. Chem., 53, 267-281.

Cacchione, S., De Santis, P., Foti, D., Palleschi, A. & Savino, M. (1989) Periodical polydeoxynucleotides and DNA curvature. Biochemistry, 28, 8706-8713.

Cairns, B.R., Kim, Y.J., Sayre, M.H., Laurent, B.C. and Kornberg, R.D. (1994) A mutisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene-products isolated from yeast. Proc. Natl. Acad. Sci. USA, 91, 1950-1954.

Calladine, C.R. & Drew, H.R. (1986) Principles of sequence-dependent flexure of DNA. J. Mol. Biol., 192, 907-918.

Calladine, C.R. & Drew, H.R. (1996) A useful role for static models in elucidating the behavior of DNA in solution. J. Mol. Biol., 257, 479-485.

Calladine, C.R., Drew, H.R. & McCall, M.J. (1988) The intrinsic curvature of DNA in solution. J. Mol. Biol., 201, 127-137.

Callan, H.G. (1986) Lampbrush chromosomes. Springer-Verlag, Berlin.

Camerini-Otero, R.D., Sollner-Webb, B. and Felsenfeld, G. (1976) Organisation of histones and DNA in chromatin: evidence for an arginine-rich kernel. Cell, 8, 333-347.

Campoy, F.J., Meehan, R.R., McKay, S., Nixon, J. and Bird, A. (1995) Binding of Histone H1 to DNA is indifferent to the methylation of CpG sequences. J. Biol. Chem., 270, 26473-26481.

Caplan, A., Kimura, T., Gould, H., and Allan, J. (1987). Perturbation of chromatin structure in the region of the adult beta- globin gene in chicken erythrocyte chromatin. J. Mol. Biol., 93, 57-69.

Cartwright, I.L., Hertzeberg, R.P., Dervan, P.B. and Elgin, S.C.R. (1983) Cleavage of DNA with Methidiumproplyl-EDTA.Iron(II). Proc. Natl. Acad. Aci. USA 80, 3213-3217.

Cartwright, I.L., and Elgin, S.C.R. (1984) Chemical footprinting of 5S RNA chromatin in embryoes. EMBO J., 3, 3101-3108.

Cartwright, I.L. and Elgin, S.C.R. (1986) Nucleosomal instability and induction of new upstream protein-DNA associations accompany activation of four small heat- shock protein genes. Mol. Cell Biol., 6, 779-791.

Cattini, P.A., Harborne, N. & Allan, J. (1988) The linker histone stoichiometry of chicken erythrocyte chromatin in situ. J. Electron Microscopy, 37, 31-37.

Cerf, C., Lippens, G., Ramakrishnan, V., Muyldermans, S., Segers, A., Wyns, L., Wodak, S.J. & Hallenga, K. (1994) Homonuclear and heteronuclear 2-dimensional NMR studies of the globular domain of histone H1 - full assignment, tertiary structure, and comparison with the globular domain of histone H5. Biochemistry, 33, 11079-11086.

Chasman, D.I., Lue, N.F., Buchman, A.R., Lapointe, J.W., Lorch, Y. & Kornberg, R.D. (1990) A yeast protein that influences the chromatin structure of UASG and functions as a powerful auxiliary gene activator. Genes Dev, 4, 503-514.

Choi, O.R. & Engel, J.D. (1986) A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. Nature, 323, 731-734.

Choi, Ok-Ryun,B. and Engel, J.D. (1988). Developmental regulation of beta globin gene switching. Cell, 55, 17-26.

Chung, J.H., Whiteley, M., and Felsenfeld, G. (1993). A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid-cells and protects against position effect in drosophila. Cell, 74, 505-514.

Clarke, D.J. and Felsenfeld, G. (1991) Formation of nucleosomes on positively supercoiled DNA. EMBO J., 10, 387-395.

Clore, G.M., Gronenborn, A.M., Nilges, M., Sukumaran, D.K. & Zarbock, J. (1987) The polypeptide fold of the globular domain of histone H5 in solution - a study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. EMBO J., 6, 1833-1842.

Cole, R.D. (1984) A minireview of microheterogeneity in H1 histone and its possible significance. Annals Biochem., 136, 24-30.

Cook, P.R. and Brazell, I.A. (1975) Supercoils in human DNA. J. Cell Sci. 19, 261-279.

Cook, P.R. (1989) The nucleoskeleton and the topology of transcription. Eur. J. Biochem, 185, 487-501.

Costanzo, G., Dimauro, E., Salina, G. & Negri, R. (1990) Attraction, phasing and neighbor effects of histone octamers on curved DNA. J. Mol. Biol., 216, 363-374.

Costello, J.F., Futscher, B.W., Tano, K., Graunke, D.M. & Piper, R.O. (1994) Graded methylation in the body and promoter of the of the O^6-Methylguanine DNA methyltransferase (MGMT) gene in human glioma cell lines. Mol. Cell. Biol. 14, 6515-6521.

Cote, J., Quinn, J., Workman, J.L. & Peterson, C.L. (1994) Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science, 265, 53-60.

Craddock, C.F., Vyas,P., Sharpe, J.A.,Ayyub, H.,Wood, W.G., Higgs, D.R., (1995). Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. EMBO J., 14, 1718-1726.

Crane-Robinson, C. & Ptitsyn, O.B. (1989) Binding of the globular domain of linker histones H5/H1 to the nucleosome - a hypothesis. Protein Engineering, 2, 577-582.

Crane-Robinson, C. (1997) Where is the globular domain of linker histone located on the nucleosome? Trends Biochem. Sci., 22, 75-77.

Crothers, D.M., Haran, T.E. & Nadeau, J.G. (1990) Intrinsically bent DNA. J. Biol. Chem., 265, 7093-7096.

Csordas, A. (1990) On the biological role of histone acetylation. Biochem. J., 265, 23-38.

Czelusinak, J., Goodman, M., Hewett-Emmett, D., Weiss, M.L., Venta P.J.& Tashian, R.E. (1982) Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature, 298, 297-300.

Dasso, M., Dimitrov, S. & Wolffe, A.P. (1994) Nuclear assembly is independent of linker histones. Proc. Natl. Acad. Sci. USA, 91, 12477-12481.

Davey, C., Pennings, S. & Allan, J. (1997) CpG methylation remodels chromatin structure in vitro. J. Mol. Biol., 267, 276-288.

Davey, C., Pennings, S., Meersseman, G., Wess, T.J. & Allan, J. (1995) Periodicity of strong nucleosome positioning sites around the chicken adult β^A-globin gene may encode regularly spaced chromatin. Proc. Natl. Acad. Sci. USA, 92, 11210-11214.

Davie, J.R. (1996) Histone modifications, chromatin structure, and the nuclear matrix. J. Cell. Biochem., 62, 149-157.

De Santis, P., Palleschi, A., Savino, M. & Scipioni, A. (1990) Validity of the nearest-neighbor approximation in the evaluation of the electrophoretic manifestations of DNA curvature. Biochemistry, 29, 9269-9273.

De Santis, P., Palleschi, A., Savino, M. & Scipioni, A. (1992) Theoretical prediction of the gel electrophoretic retardation changes due to point mutations in a tract of SV40 DNA. Biophys. Chem., 42, 147-152.

De Santis, P., Fua, M., Palleschi, A. & Savino, M. (1993) Relationships between intrinsic and induced curvature in DNAs -theoretical prediction of nucleosome positioning. Biophys. Chem., 46, 193-204.

De Santis, P., Fua, M., Palleschi, A. & Savino, M. (1995) Influence of dynamic fluctuations on DNA curvature. Biophys. Chem., 55, 261-271.

De-Murcia, G., Huletsky, A., Lamarre, D., Gaudreau, A., Pouyet, J., Daune, M. & Poirier, G.G. (1986) Modulation of chromatin superstructure induced by poly(ADP-ribose) synthesis and degradation. J. Biol. Chem., 261, 7011-7017.

DeLange, R.J., Fambrough, D.M., Smith, E.L. & Bonner, J. (1969) Calf and pea histone IV. III. complete amino acid sequence of pea seedling histone IV: comparison with the homologous calf thymus histone. J. Biol. Chem., 244, 5669-5679.

Denisov, D.A., Shpigelman, E.S and Trifonov, E.N. (1997) Protective nucleosome centering at splice sites as suggested by sequence-directed mapping of the nucleosomes. Gene, 205, 145-149.

Diekmann, S. & Mclaughlin, L.W. (1988) DNA curvature in native and modified EcoRI recognition sites and possible influence upon the endonuclease cleavage reaction. J. Mol. Biol., 202, 823-834.

Diekmann, S. (1987) Temperature and salt dependence of the gel migration anomaly of curved DNA fragments. Nucl. Acids Res., 15, 247-265.

Dimitrov, S. & Wolffe, A.P. (1996) Remodeling somatic nuclei in Xenopus laevis egg extracts - molecular mechanisms for the selective release of histones H1 and H1° from chromatin and the acquisition of transcriptional competence. EMBO J., 15, 5897-5906.

Dimitrov, S.I., Russanova, V.R. & Pashev, I.G. (1987) The globular domain of histone H5 is internally located in the 30-nm chromatin fiber - an immunochemical study. EMBO J., 6, 2387-2392.

Dlakic, M. & Harrington, R.E. (1996) The effects of sequence context on DNA curvature. Proc. Natl. Acad. Sci. USA, 93, 3847-3852.

Dodgson, J.B., Stadt, S.J., Choi, O.R., Dolan, M., Fischer, H.D. and Engel, J.D. (1983). The nucleotide-sequence of the embryonic chicken beta-type globin genes. J. Biol. Chem., 258, 2685-2692.

Dolan, M., Dodgson, J.B., and Engel, J.D. (1983). Analysis of the adult chicken beta-globin gene - nucleotide-sequence of the locus, microheterogeneity at the 5'-end of beta-globin. J. Biol. Chem., 258, 3983-3990.

Dolan, M., Sugarman, B.J., Dodgson, J.B., and Engel, J.D. (1981). Chromosomal arrangement of the chicken beta-type globin Genes. Cell, 24, 669-677.

Drew, H.R. & Calladine, C.R. (1987) Sequence-specific positioning of core histones on an 860 base-pair DNA - experiment and theory. J. Mol. Biol., 195, 143-173.

Drew, H.R. & Travers, A.A. (1985) DNA bending and its relation to nucleosome positioning. J. Mol. Biol., 186, 773-790.

Du, W., Thanos, D. & Maniatis, T. (1993) Mechanisms of transcriptional synergism between distinct virus-inducible enhancer elements. Cell, 74, 887-898.

Dunphy, W.G. & Newport, J.W. (1988) Mitosis inducing factors are present in a latent form during interphase in the Xenopus embryo. J. Cell. Biol., 106, 2047-2056.

Durrin, L.K., Mann, R.K., Kayne, P.S. & Grunstein, M. (1991) Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. Cell, 65, 1023-1031.

Edmondson, D.G., Smith, M.M. & Roth, S.Y. (1996) Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. Genes Dev, 10, 1247-1259.

Eickbusch, T.H. & Moudrianakis, E.N. (1978) Compaction of DNA helices into either continuous supercoils or folded-fiber rods and toroids. Cell, 13, 295-306.

Eisfeld, K., Candau, R., Truss, M. & Beato, M. (1997) Binding of NF1 to the MMTVpromoter in nucleosomes: influence of rotational phasing, translational positioning and histone H1. Nucl. Acids Res., 25, 3733-3742.

Elefanty, A.G., Antoniou, M., Custodio, N., Carmofonseca, M., and Grosveld, F.G. (1996). GATA transcription factors associate with a novel class of nuclear- bodies in erythroblasts and megakaryocytes. EMBO J., 15, 319-333.

Elfring, L.K., Deuring, R., McCallum, C.M., Peterson, C.L. & Tamkun, J.W. (1994) Identification and characterization of Drosophila relatives of the yeast transcriptional activator SNF2/SWI2. Mol. Cell. Biol., 14, 2225-2234.

Elgin, S.C.R. (1988) The formation and function of DNaseI hypersensitive sites in the process of gene activation. J. Biol. Chem., 263, 19259-19262.

Ellis, J., Tanun, K.C., Harper, A., Yannoutsos, N., Fraser, P., Michalovich, D., Philipsen, S., and Grosveld, F. (1995). Evaluation of human beta-globin gene-therapy constructs by expression in single-copy transgenic mice. J.Cell. Biochem., No. S21A, p376.

Ellis, J., Tanun, K.C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S., and Grosveld, F. (1996). A dominant chromatin-opening activity in 5'-hypersensitive site-3 of the human beta-globin locus-control region. EMBO J., 15, 562-568.

Emerson, B.M. and Felsenfeld, G. (1984). Specific factor conferring nuclease hypersensitivity at the 5' end of the chicken adult beta-globin gene. Proc. Natl. Acad. Sci.USA, 81, 95-99.

Emerson, B.M., Jackson, P.D., and Felsenfeld, G. (1985). Specific protein-DNA interactions associated with active and inactive forms of the chicken adult beta-globin gene. Federation Proceedings 44, 1769.

Emerson, B.M., Lewis, C.D. and Felsenfeld, G. (1985b) Interaction of specific nuclear factors with the nuclease hypersensitive region of the chicken beta-adult gene- nature of the binding domain. Cell, 41, 21-30.

Emerson, B.M., Nickol, J.M., Jackson, P.D., and Felsenfeld, G. (1987). Analysis of the tissue-specific enhancer at the 3' end of the chicken adult beta-globin gene. Proc. Natl. Acad. Sci. USA 84, 4786-4790.

Engel, J.D. (1993). Developmental regulation of human Beta-globin gene transcription: a switch of loyalties. Trends Genet., 9, 304-309.

Ericsson, C., Grossbach, U., Björkroth, B. and Daneholt,B. (1990) Presence of H1 on an active Balbiani ring gene. Cell, 60, 73-83.

Evans, T., Reitman, M., and Felsenfeld, G. (1988). An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. Proc.Natl. Acad. Sci.USA, 85, 5976-5980.

Evans, T., Felsenfeld, G., and Reitman, M. (1990) Control of globin gene transcription. Annu. Rev. Cell. Biol., 6, 95-124.

Evans, T. and Felsenfeld, G. (1991). Transactivation of a globin promoter in non-erythroid cells. Mol. Cell. Biol. 11, 843-853.

Fascher,K.D., Schmitz, J. and Horz, W. (1990) Role of trans-activating proteins in generation of an active site at the PHO5 promoter in *S. cerevisiae*. EMBO J., 9, 2523-2528.

Fedor, M.J., Lue, N.F. & Kornberg, R.D. (1988) Statistical positioning of nucleosomes by specific protein binding to an upstream activating sequence in yeast. J. Mol. Biol., 204, 109-127.

Felsenfeld, G. (1992) Chromatin as an essential part of the transcriptional mechanism. Nature, 355, 219-224.

Felsenfeld, G. (1993). Chromatin structure and the expression of the globin genes. Gene, 135, 119-124.

Felsenfeld, G. (1996) Chromatin unfolds. Cell, 86, 13-19.

Finch, J.T. & Klug, A. (1976) Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. USA, 73, 1897-1901.

Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M. & Klug, A. (1977) Structure of nucleosome core particles of chromatin. Nature, 269, 29-36.

Finch, J.T., Noll, M. & Kornberg, R.D. (1975) Electron microscopy of defined length of chromatin. Proc. Natl. Acad. Sci. USA, 72, 3320-3322.

Fitzgerald, D.J. and Anderson, J.N.(1998) Unique translational positioning of nucleosomes on synthetic DNA. Nucl. Acids Res. 26, 2526-2535.

Flaus, A., Luger, K., Tan, S. & Richmond, T.J. (1996) Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. Proc. Natl. Acad. Sci. USA, 93, 1370-1375.

Foley, K.P. and Engel, J.D. (1992). Individual stage selector element mutations lead to reciprocal changes in β^A versus ε globin gene transcrption: genetic confirmation of promoter competition during globin gene switching. Genes Dev 6, 730-744.

Foley, K.P., Pruzina, S., Winick, J.D., Engel, J.D., Grosveld, F., and Fraser, P. (1994). The chicken beta/epsilon-globin enhancer directs autonomously regulated, high-level expression of the chicken epsilon-globin gene in transgenic mice. Proc. Natl. Acad. Sci. USA, 91, 7252-7256.

Fong, T.C. and Emerson,B.M. (1992) An erythroid specific protein cGATA-1 mediates distal enhancer activity through a specialised beta-globin TATA box. Genes Dev., 6, 521-532.

Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. & Turner, D.H. (1986) Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA, 83, 9373-9377.

Gallarda, J.L., Foley, K.P., Yang, Z. and Engel, J.D. (1989) The beta-globin stage selector element factor is erythroid-specific promoter/enhancer binding protein NFE-4. Genes Dev., 3, 1845-1859.

Gallo, C.V.D., Vassetzky, Y.S., Targa, F.R., Georgiev, G.P., Scherrer, K., and Razin, S.V. (1991). The presence of sequence-specific protein-binding sites correlate with replication activity and matrix binding in a 1.7 kb-long DNA fragment of the chicken alpha-globin gene domain. Biochem. Biophys. Res. Comm. 179, 512-519.

Garciaramirez, M., Rocchini, C. & Ausio, J. (1995) Modulation of chromatin folding by histone acetylation. J. Biol. Chem., 270, 17923-17928.

Gasser, S.M. & Laemmli, U.K. (1986) The organization of chromatin loops: characterization of a scaffold attachment site. EMBO J., 5, 511-518.

Gasser, S.M. and Laemmli, U.K. (1987) A glimpse at chromosomal order, Trends Genet., 3, 16-22.

Garrick, D., Sutherland, H., Robertson, G., and Whitelaw, E. (1996). Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. Nucl. Acids Res., 24, 4902-4909.

Georgakopoulos, T., Gounalaki, N. & Thireos, G. (1995) Genetic-evidence for the interaction of the yeast transcriptional coactivator proteins Gcn5 and Ada2. Mol. Gen. Genet., 246, 723-728.

George, K.M., Leonard, M.W., Roth, M.E., Lieuw, K.H., Kioussis, D., Grosveld, F., and Engel, J.D. (1994). Embryonic expression and cloning of the murine gata-3 gene. Development 120, 2673-2686.

Gerchman, S.E., Ramakrishnan, V. (1987) Chromatin higher-order structure studied by neutron scattering and scanning electron microscopy. Proc. Natl. Acad. Sci. USA 84, 7802-7806.

Gilman, J.G. and Josifovska, O. (1997). AT repeats in beta globin locus control region HS2 bind nuclear factor hmbi(y) and diminish enhancement due to adjacent GATA-1 site. Amer. J. Human Gen. 61, No. 4 SS, 988.

Ginder, G.D., Whitters, M.J., and Pohlman, J.K. (1984). Activation of a chicken embryonic globin gene in adult erythroid-cells by 5-azacytidine and sodium-butyrate. Proc. Natl. Acad. Sci. USA, 81, 3954-3958.

Glauber, J.G., Wandersee, N.J., Little, J.A., and Ginder, G.D. (1991). 5'-flanking sequences mediate butyrate stimulation of embryonic globin gene-expression in adult erythroid-cells. Mol. Cell. Biol. 11, 4690-4697.

Godde, J.S., Nakatani, Y. & Wolffe, A.P. (1995) The amino terminal tails of the core histones and the translational position of the TATA box determine TBP/TFIIA association with nucleosomal DNA. Nucl. Acids Res., 23, 4557-4564.

Godde, J.S., Kass, S.U., Hirst, M.C. & Wolffe, A.P. (1996) Nucleosome assembly on CTG triplet repeats. J. Biol. Chem., 271, 24325-24328.

Gong, Q.H. and Dean, A. (1993). Enhancer-dependent transcription of the epsilon-globin promoter requires promoter-bound GATA-1 and enhancer-bound AP-1/NF-E2. Mol. Cell.Biol. 13, 911-917.

Gong, Q.H., Mcdowell, J.C., and Dean, A. (1996). Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene in-vivo by 5'-hypersensitive site-2 of the beta-globin locus-control region. Mol. Cell. Biol. 16, 6055- 6064.

Gong, Q.H., Stern, J., and Dean, A. (1991). Transcriptional role of a conserved GATA-1 site in the human epsilon-globin gene promoter. Mol. Cell. Biol. 11, 2558-2566.

Goodsell, D.S. & Dickerson, R.E. (1994) Bending and curvature calculations in B-DNA. Nucl. Acids Res., 22, 5497-5503.

Gottesfeld. M.J., Murphy, R.F., and Bonner, J. (1975) Structure of transcriptionally active chromatin. Proc. Natl. Acad. Sci. USA, 72, 4404-4408.

Goytisolo, F.A., Gerchman, S.E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V. & Thomas, J.O. (1996) Identification of 2 DNA-binding sites on the globular domain of histone H5. EMBO J., 15, 3421-3429.

Graessman A., Sanderg, G., Guhl, E. & Graessman, M. (1994). Methylation of single sites within the herpes simplex virus *tk* coding region and simian virus 40 T-antigen intron causes gene inactivation. Mol. Cell. Biol., 14, 2004-2010.

Grant , P.A., Duggan, L., Cote, J., Roberts, S.M., Brownell, J.E., Candau, R., Ohba, R., OwenHughes, T., Allis, C.D., Winston, F., Berger, S.L. and Workman, J.L. (1997) Yeast Gcn5 functions in two mutli-subunit complexes to acetylate nucleosomal histones. Genes Dev., 11, 1640-1650.

Graziano, V., Gerchman, S.E., Schneider, D.K. & Ramakrishnan, V. (1994) Histone H1 is located in the interior of the chromatin 30-nm filament. Nature, 368, 351-354.

Green, G.R., Lee, H.J. & Poccia, D.L. (1993) Phosphorylation weakens DNA-binding by peptides containing multiple SPKK sequences. J. Biol. Chem., 268, 11247-11255.

Griffith, J., Bleyman, M., Rauch, C.A., Kitchin, P.A. & Englund, P.T. (1986) Visualization of the bent helix in kinetoplast DNA by electron microscopy. Cell, 46, 717-724.

Grigoryev, S.A., Spirin, K.S., and Krasheninnikov, I.A. (1990). Loosened nucleosome linker folding in transcriptionally active chromatin of chicken-embryo erythrocyte nuclei. Nucl. Acids Res., 18, 7397-7406.

Gross, D.S. & Garrard, W.T. (1988) Nuclease hypersensitive sites in chromatin. Annu. Rev. Biochem, 57, 159-197.

Grosschedl, R., Giese, K. & Pagel, J. (1994) HMG domain proteins - architectural elements in the assembly of nucleoprotein structures. Trends Genet., 10, 94-100.

Grosveld, F., Blom von Assendelft, G., Greaves, D.R. and Kolias, G. (1987) Cell, 51, 975-985.

Grosveld, F., Antoniou, M., Berry, M., Deboer, E., Dillon, N., Ellis, J., Fraser, P., Hurst, J., Imam, A., Meijer, D., Philipsen, S., Pruzina, S., Strouboulis, J., and Whyatt, D. (1993). Regulation of human globin gene switching. Cold Spring Harbor Symposia On Quantitative Biology 58, 7-13.

Groudine, M. and Weintraub, H. (1981). Activation of globin genes during chicken development. Cell 24, 393-401.

Grunstein, M. (1990) Nucleosomes: regulators of transcription. Trends Genet., 6, 395-400.

Grunstein, M., Durrin, L.K., Mann, R.K., Fisher-Adams, G. & Johnson, L.M. (1992) Histone: regulation of transcription in yeast. In: Transcriptional regulation, edited by McKnight, S. & Yamamoto, K. Cold Spring Harbor Laboratory Press, New York, p. 1295-1315.

Gualerzi, C.O. & Pon, C.L. (1990) Initiation of mRNA translation in prokaryotes. Biochemistry, 29, 5881-5889.

Gualerzi, C.O., Pon, C.L., Pawlik, R.T., Canonaco, M.A., Paci, M. & Wintermeyer, W. (1986) Role of initiation factors in Escherichia coli translational initiation. In: Structure, function, and genetics of ribosomes, edited by Hardesty, B. & Kramer, G. Springer-Verlag, New York, p. 621-641.

Guo, B., Odgren, P.R., Vanwijnen, A.J., Last, T.J., Nickerson, J., Penman, S., Lian, J.B., Stein, J.L. & Stein, G.S. (1995) The nuclear matrix protein Nmp-1 is the transcription factor YY1. Proc. Natl. Acad. Sci. USA, 92, 10526-10530.

Hager, G., Smith, C., Svaren, J. & Horz, W. (1995) Initiation of expression: remodelling genes. In: Chromatin structure and gene expression, edited by Elgin, S.C.R. IRL press, Oxford, p. 89-103.

Hagerman, P.J. & Ramadevi, V.A. (1990) Application of the method of phage T4 DNA ligase catalyzed ring closure to the study of DNA structure: a computational analysis. J. Mol. Biol., 212, 351-362.

Hagerman, P.J. (1988) Flexibility of DNA. Annu. Rev. Biophys. Biophys. Chem., 17, 265-286.

Hall, J.M. & Cole, R.D. (1986) In nondividing cells, histone H1(0) is synthesized and deposited onto chromatin without accompanying phosphorylation. Biochemistry, 25, 491-495.

Hall, M.N., Gabay, J., Debarbouille, M. & Schwartz, M. (1982) A role for mRNA secondary structure in the control of translation initiation. Nature, 295, 616-618.

Han, M. and Grunstein, M. (1988) Depletion of Histone H4 and nucleosomes activates the PHO5 genes in *Saccharomyces cerevisiae*. Cell, 55,1137-1145.

Hanahan, D. Studies on transformation of E. Coli with plasmids. J. Mol. Biol., 166, 557-580.

Hannon, R., Evans, T., Felsenfeld, G., and Gould, H. (1991). Structure and promoter activity of the gene for the erythroid transcription factor GATA-1. Proc. Natl. Acad. Sci. USA, 88, 3004-3008.

Haran, T.E., Kahn, J.D. & Crothers, D.M. (1994) Sequence elements responsible for DNA curvature. J. Mol. Biol., 244, 135-143.

Harrington, M.A., Jones, P.A., Imagawa, M. & Karin, M. (1988) Cytosine methylation does not affect binding of transcription factor SP1. Proc. Natl. Acad. Sci. USA, 85, 2066-2070.

Hartman, P.G., Chapman, G.E., Moss, T. & Bradbury, E.M. (1977) Studies of the role and mode of operation of the very lysine rich histone H1 in eukaryote chromatin. Eur. J. Biochem., 77, 45-51.

Hattori, M. & Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. Annals Biochem., 152, 232-238.

Hayes, J.J., Tullius, T.D. & Wolffe, A.P. (1990) The structure of DNA in a nucleosome. Proc. Natl. Acad. Sci. USA, 87, 7405-7409.

Hayes, J.J., Clark, D.J. & Wolffe, A.P. (1991) Histone contributions to the structure of DNA in the nucleosome. Proc. Natl. Acad. Sci. USA, 88, 6829-6833.

Hayes, J.J. & Wolffe, A.P. (1992a) The interaction of transcription factors with nucleosomal DNA. BioEssays, 14, 597-603.

Hayes, J.J. & Wolffe, A.P. (1992b) Histones H2A/H2B inhibit the interaction of transcription factor IIIA with the Xenopus borealis somatic 5S RNA gene in a nucleosome. Proc. Natl. Acad. Sci. USA, 89, 1229-1233.

Hayes, J.J. & Wolffe, A.P. (1993) Preferential and asymmetric interaction of linker histones with 5S DNA in the nucleosome. Proc. Natl. Acad. Sci. USA, 90, 6415-6419.

Hayes, J.J., Pruss, D. & Wolffe, A.P. (1994) Contacts of the globular domain of histone H5 and core histones with DNA in a chromatosome. Proc. Natl. Acad. Sci. USA, 91, 7817-7821.

Hayes, J.J. (1996) Site-directed cleavage of DNA by a linker histone-Fe(II) EDTA conjugate - localization of a globular domain binding site within a nucleosome. Biochemistry, 35, 11931-11937.

Hayes, J.J., Kaplan, R., Ura, K., Pruss, D. & Wolffe, A. (1996b) A putative DNA-binding surface in the globular domain of a linker histone is not essential for specific binding to the nucleosome. J. Biol. Chem., 271, 25817-25822.

Hebbes, T.R., Thorne, A.W. & Crane-Robinson, C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J., 7, 1395-1402.

Hebbes, T.R., Thorne, A.W., Clayton, A.L., and Crane-Robinson, C. (1992). Histone acetylation and globin gene switching. Nucl. Acids Res., 20, 1017-1022.

Hebbes, T.R., Clayton, A.L., Thorne, A.W., and Crane-Robinson, C. (1994) Core-histone acetylation co-maps with generalised DNase I sensitivity in the chicken β-globin chromosomal domain. EMBO J., 13, 1823-1830.

Hecht, A., Laroche, T., Strahlbolsinger, S., Gasser, S.M. & Grunstein, M. (1995) Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins - a molecular model for the formation of heterochromatin in yeast. Cell, 80, 583-592.

Hesse, J.E., Nickol, J.M., Lieber, M.R. & Felsenfeld, G. (1986) Regulated gene expression in transfected primary chicken erythrocytes. Proc. Natl. Acad. Sci. USA, 83, 4312-4316.

Hill, C.S., Martin, S.R. & Thomas, J.O. (1989) A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-bound histone H1 from sea urchin sperm. EMBO J., 8, 2591-2599.

Hirano, T. & Mitchison, T.J. (1993) Topoisomerase II does not play a scaffolding role in the organization of mitotic chromosomes assembled in Xenopus egg extracts. J. Cell. Biol., 120, 601-612.

Hirano, T. & Mitchison, T.J. (1994) A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. Cell, 79, 449-458.

Hiraoka, Y., Minden, J.S., Swedlow, J.R., Sedat, J.W. & Agard, D.A. (1989) Focal points for chromosome condensation and decondensation revealed by 3-dimensional in vivo time lapse microscopy. Nature, 342, 293-296.

Hirschhorn, J.N., Brown, S.A., Clark, C.D. & Winston, F. (1992) Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. Genes Dev., 6, 2288-2298.

Holmes, B. (1993) Message in a genome. New Scientist, 147, No. 1990, 30-33.

Holt, C.L. & May, G.S. (1996) An extragenic suppressor of the mitosis defective bimd6 mutation of Aspergillus nidulans codes for a chromosome scaffold protein. Genetics, 142, 777-787.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P. & Bradbury, E.M. (1993) Studies of the DNA-binding properties of histone H4 amino terminus – thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 tail to DNA. J. Biol. Chem., 268, 305-314.

Horiuchi, J., Silverman, N., Marcus, G.A. & Guarente, L. (1995) Ada3, a putative transcriptional adapter, consists of 2 separable domains and interacts with Ada2 and Gcn5 in a trimeric complex. Mol. Cell. Biol., 15, 1203-1209.

Horowitz, R.A., Agard, D.A., Sedat, J.W. and Woodcock, C.L. (1994) The 3D architecture of chromatin in situ. J. Cell. Biol., 125, 1-10.

Ikonomi, P., Noguchi, C.T., and Schechter, A.N. (1996). GATA-1 participates in the silencing of the epsilon-globin gene. Blood, 88, 585.

Ikuta, T., Papayannopoulou, T., Stamatoyannopoulos, G., and Kan, Y.W. (1996). Globin gene switching - in-vivo protein-DNA interactions of the human beta-globin locus in erythroid- cells expressing the fetal or the adult globin gene program. J. Biol. Chem., 271, 14082-14091.

Iler, N., Goodwin, A.J., and Lowrey, C.H. (1996). Artificial manipulation of beta-globin promoter chromatin structure yields increased expression and inducibility but not position- independent expression. Blood 88, 581.

Imaizumischerrer, T., Rangini, Z., Vesque, C., Duquesnoy, P., and Scherrer, K. (1989). Presence of globin gene transcripts in chicken oocytes and of a partially processed globin RNA in early embryos. Differentiation, 41, 22-33.

Imbalzano, A.N., Kwon, H., Green, M.R. & Kingston, R.E. (1994) Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature, 370, 481-485.

Ioshikhes, I., Bolshoy, A. & Trifonov, E.N. (1992) Preferred positions of AA-dinucleotides and TT-dinucleotides in aligned nucleosomal DNA-sequences. J Biomol. Struc. Dynam., 9, 1111-1117.

Ioshikhes, I. and Trifonov, E.N. (1993) Nucleosomal DNA sequence database. Nucl. Acids Res. 21, 4857-4859.

Ioshikhes, I., Bolshoy, A., Derenshteyn, K., Borodovsky, M. & Trifonov, E.N. (1996) Nucleosome DNA sequence pattern revealed by multiple alignment of experimentally mapped sequences. J. Mol. Biol., 262, 129-139.

Izaurralde, E., Kas, E. & Laemmli, U.K. (1989) Highly preferential nucleation of histone H1 assembly on scaffold associated regions. J. Mol. Biol., 210, 573-585.

Jackson, D.A., McCready, S.J. and Cook, P.R. (1984) Replication and transcription depend on attachment to the nuclear cage. J. Cell Sci. Suppl. 1, 59-79.

Jackson, J.R. & Benyajati, C. (1993) DNA-histone interactions are sufficient to position a single nucleosome juxtaposing Drosophila Adh adult enhancer and distal promoter. Nucl. Acids Res., 21, 957-967.

Jane, S.M., Amrolia, P., and Cunningham, J.M. (1995a). Developmental regulation of the human beta-globin cluster. Austral. and New Zeal. J. Medicine 25, 865-869.

Jane, S.M., Nienhuis, A.W., and Cunningham, J.M. (1995b). Hemoglobin switching in man and chicken is mediated by a heteromeric complex between the ubiquitous transcription factor CP2 and a developmentally specific protein. EMBO J., 14, 97-105.

Johnston, M. & Dover, J. (1987) Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA-binding domain. Proc. Natl. Acad. Sci. USA, 84, 2401-2405.

Juan, L.J., Utley, R.T., Vignali, M., Bohm, L. & Workman, J.L. (1997) H1-mediated repression of transcription factor binding to a stably positioned nucleosome. J. Biol. Chem., 272, 3635-3640.

Kamakaka, R.T. & Thomas, J.O. (1990) Chromatin structure of transcriptionally competent and repressed genes. EMBO J., 9, 3997-4006.

Kandolf, H. (1994) The H1a histone variant is an in vivo repressor of oocyte type 5S gene transcription in Xenopus laevis embryos. Proc. Natl. Acad. Sci. USA, 91, 7257-7261.

Kas, E., Izaurralde, E. & Laemmli, U.K. (1989) Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin - the role of oligo(dA)(oligo(dT) tracts. J. Mol. Biol., 210, 587-599.

Kass, S.U., Pruss, D. and Wolffe, A.P. (1997) How does DNA methylation repress transcription? Trends Genet., 13, 444-449.

Kaufman, P.D. & Botchan, M.R. (1994) Assembly of nucleosomes: do multiple assembly factors mean multiple mechanism? Curr. Opin. Genet. Dev., 4, 229-235.

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F. & Grunstein, M. (1988) Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell, 55, 27-39.

Kefalas, P., Gray, F.C. & Allan, J. (1988) Precise nucleosome positioning in the promoter of the chicken β^A-globin gene. Nucl. Acids Res., 16, 501-517.

Keshet, I., Leiman-Hurwitz, J. & Cedar, H. (1986) DNA methylation affects the formation of active chromatin. Proc. Natl. Acad. Sci. USA, 82, 2560-2564.

Kielman, M.F., Barradeau, S., Smits, R., Harteveld, C.L., and Bernini, L.F. (1996). Characterization and localization of the mprox1 gene directly upstream of the mouse alpha-globin gene cluster: identification of a polymorphic direct repeat in the 5'UTR. Mammalian Genome, 7, 877-880.

Klug, A., Rhodes, D., Smith, J., Finch, J.T. & Thomas, J.O. (1980) A low resolution structure for the histone core of the nucleosome. Nature, 287, 509-516.

Knezetic, J.A. & Luse, D.S. (1986) The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell, 45, 95-104.

Knezetic, J.A. and Felsenfeld, G. (1993). Mechanism of developmental regulation of alpha-pi, the chicken embryonic alpha-globin gene. Mol. Cell. Biol. 13, 4632-4639.

Ko, L.J. and Engel, J.D. (1993). DNA-binding specificities of the gata transcription factor family. Mol. Cell. Biol. 13, 4011-4022.

Koo, H.S. & Crothers, D.M. (1988) Calibration of DNA curvature and a unified description of sequence- directed bending. Proc. Natl. Acad. Sci. USA, 85, 1763-1767.

Koo, H.S., Wu, H.M. & Crothers, D.M. (1986) DNA bending at A/T tracts. Nature, 320, 501-506.

Kornberg, R.D. (1974) Chromatin structure: a repeating unit of histones and DNA. Science, 184, 868-871.

Kornberg, R.D. & Thomas, J.O. (1974) Chromatin structure: oligomers of histones. Science, 184, 865-868.

Kornberg, R.D. & Stryer, L. (1988) Statistical distributions of nucleosomes - nonrandom locations by a stochastic mechanism. Nucl. Acids Res., 16, 6677-6690.

Kraevskii, V.A., Mikhailov, V.S., and Razin, S.V. (1992). Stable bent and low-melting DNA-sequences in the region of the replication origin of chicken alpha-globin gene domain. Molecular Biology 26, 672-678.

Kralovics, R., Fajkus, J., Kovarik, A. & Bezdek, M. (1995) DNA curvature of the tobacco grs repetitive sequence family and its relation to nucleosome positioning. J Biomol. Struc. Dynam., 12, 1103-1119.

Kropp, B., Leoni, L., Sampaolese, B. & Savino, M. (1995) Influence of DNA superstructural features and histone amino terminal domains on nucleosome positioning. FEBS Lett, 364, 17-22.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. & Allis, C.D. (1996) Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature, 383, 269-272.

Kwon, H., Imbalzano, A.N., Khavari, P.A., Kingston, R.E. & Green, M.R. (1994) Nucleosome disruption and enhancement of activator binding by a human SWI/SNFcomplex. Nature, 370, 477-481.

Larsen, A. and Weintraub, H. (1982) An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell, 29, 609-622.

Lauderdale, J.D. and Stein, A. (1992) Introns of the ovalbumin gene promoter promote nucleosome alignment *in vitro*. Nucl.Acids Res. 20, 6589-6596.

Laurent, B.C., Treich, I. & Carlson, M. (1993) The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. Genes Dev., 7, 583-591.

Lee, D.Y., Hayes, J.J., Pruss, D. & Wolffe, A.P. (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell, 72, 73-84.

Leonard, A.C. & Thomas, J.O. (1985) The arrangement of H5 molecules in extended and condensed chicken erythrocyte chromatin. EMBO J., 4, 3455-3462.

Leonard, M.W., Lim, K.C., and Engel, J.D. (1993). Expression of the chicken GATA factor family during early erythroid development and differentiation. Development, 119, 519-531.

Leuba, S.H., Yang, G.L., Robert, C., Samori, B., van Holde, K., Zlatanova, J. & Bustamante, C. (1994) 3-dimensional structure of extended chromatin fibers as revealed by tapping mode scanning force microscopy. Proc. Natl. Acad. Sci. USA, 91, 11621-11625.

Leuba, S.H., Zlatanova, J. & van Holde, K. (1993) On the location of histone H1 and histone H5 in the chromatin fiber - studies with immobilized trypsin and chymotrypsin. J. Mol. Biol., 229, 917-929.

Lewin, B. (1997) Genes V, (5th edition), published by Wiley.

Lewis, C.D. and Laemmli, U.K. (1982) Higher-order metaphase chromosome structure. Cell, 29, 171-181.

Lewis, C.D., Clark, S.P., Felsenfeld, G., and Gould, H. (1988). An erythrocyte-specific protein that binds to the poly(dG) region of the chicken beta-globin gene promoter. Genes Dev., 2, 863-873.

Lewis, E.A. & Reams, R.R. (1983) Histone H1 – ultra centrifugation studies of the effects of ionic strength and denaturants on the solution conformation. Arch. Biochem. Biophys., 223, 185-192.

Li, B.Y., Adams, C.C. & Workman, J.L. (1994) Nucleosome binding by the constitutive transcription factor Sp1. J. Biol. Chem., 269, 7756-7763.

Li, J., Noguchi, C.T., Miller, W., Hardison, R., and Schechter, A.N. (1998). Multiple regulatory elements in the 5'-flanking sequence of the human epsilon-globin gene. J. Biol. Chem., 273, 10202-10209.

Li, Q. & Wrange, O. (1993) Translational positioning of a nucleosomal glucocorticoid response element modulates glucocorticoid receptor affinity. Genes Dev., 7, 2471-2482.

Li, Q., Blau, C.A., Clegg, C.H., and Stamatoyannopoulos, G. (1997a). Elements of developmental regulation of the human epsilon-globin gene reside within the proximal promoter. Blood, 90, 567.

Li, W., Nagaraja, S., Delcuve, G.P., Hendzel, M.J. & Davie, J.R. (1993) Effects of histone acetylation, ubiquitination and variants on nucleosome stability. Biochem. J. 296, 737-744.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L. & Ward, D.C. (1988) Delineation of individual human chromosomes in metaphase and interphase cells by in situ supression hybridization using recombinant DNA libraries. Hum Genet, 80, 224-234.

Liebhaber, S.A., Wang, Z., Cash, F.E., Monks, B., and Russell, J.E. (1996). Developmental silencing of the embryonic zeta globin gene: concerted action of the promoter and the 3' flanking region combined with stage specific silencing by the transcribed segment. Mol. Cell. Biol. 16, 2637-2646.

Lin Sun, Xu,Y.Z.,Bellard,M.,Chambon P. (1986). Digestion of the Chicken Beta-Globin gene chromatin with Micrococcal Nuclease reveals the presence of an altered nucleosomal array characterised by an atypical ladder of DNA fragments. EMBO J., 5, 293-300.

Liu, K.Y., Lauderdale, J.D., and Stein, A. (1993). Signals in chicken beta-globin DNA influence chromatin assembly in-vitro. Mol. and Cell. Biol. 13, 7596-7603.

Liu, K.Y., Sandgren, E.P., Palmiter, R.D. and Stein, A. (1995) Rat growth hormone gene introns stimulate nucleosome alignment *in vitro* and in transgenic mice. Proc. Natl. Acad. Sci. USA. 92, 7724-7728.

Liu, K and Stein, A. (1997) DNA sequence encodes information for nucleosome array formation. J. Mol. Biol., 270, 559-573.

Locklear, L., Ridsdale, J.A., Bazettjones, D.P., and Davie, J.R. (1990). Ultrastructure of transcriptionally competent chromatin. Nucl. Acids Res., 18, 7015-7024.

Long, M., Rosenberg, C., and Gilbert, W. (1997) Intron phase correlations and the evolution of intron/exon structure of genes. Proc. Natl. Acad. Sci. USA, 92, 12495-12499.

Lois, R. and Martinson, H.G. (1989). Chicken globin gene-transcription is cell lineage specific during the time of the switch. Biochemistry, 28, 2281-2287.

Lorch, Y., Lapointe, J.W. & Kornberg, R.D. (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell, 49, 203-210.

Losa, R. & Brown, D.D. (1987) A bacteriophage RNA polymerase transcribes in vitro through a nucleosome core without displacing it. Cell, 50, 801-808.

Lowary, P.T. and Widom, J. (1997) Nucleosome packaging and nucleosome positioning Proc. Natl. Acad. Sci. USA, 94, 1183-1188.

Lowary, P.T. & Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol., 276, 19-42.

Lowman, H. & Bina, M. (1990) Correlation between dinucleotide periodicities and nucleosome positioning on mouse satellite DNA. Biopolymers, 30, 861-876.

Lu, Q., Wallrath, L.L. and Elgin, S.C.R. (1994) Nucleosome positioning and gene regulation. J. Cellular Biochem., 55, 83-92.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature, 389, 251-260.

Lutter, L.C. (1979) Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel electrophoresis. Nucl. Acids Res., 6, 41-56.

Mahloogi and Behe (1997) Oligoadenosine tracts favour nucleosome formation. Biochem. Biophys. Res. Comm., 235, 663-668.

Makarov, V., Dimitrov, S., Smirnov, V. & Pashev, I. (1985) A triple helix model for the structure of chromatin fiber. FEBS Lett., 181, 357-361.

Maman, J.D., Yager, T.D. & Allan, J. (1994) Self-association of the globular domain of histone H5. Biochemistry, 33, 1300-1310.

Mandel, J.L. and Chambon, P. (1979) DNA methylation: organ-specific variations in the methylation pattern within and around ovalbumin and other chicken genes. Nucl. Acids Res., 7, 2081-2103.

Marcus, G.A., Silverman, N., Berger, S.L., Horiuchi, J. & Guarente, L. (1994) Functional similarity and physical association between Gcn5 and Ada2 - putative transcriptional adapters. EMBO J., 13, 4807-4815.

Marshall, L.E., Graham, D.R., Reich, K.A. and Sigman, D.S. (1981) Cleavage of DNA by the 1,10-Phenanthroline-Cuprous complex. Hydrogen peroxide requirement and primary and secondary structure specificity. Biochemistry, 20, 244-250.

Martin, D.I.K., Fiering, S., and Groudine, M. (1996). Regulation of beta-globin gene-expression - straightening out the locus. Curr. Op. Genet.Dev., 6, 488-495.

Mason, M.M., Lee,E., Westphal, H. and Reitman, M. (1995) Expression of the Chicken β-globin gene cluster in mice: Correct developmental expression and distributed control. Mol. Cell. Biol., 15, 407-414.

Mason, M.M., Grasso, J.A., Gavrilova, O., and Reitman, m. (1996). Identification of functional elements of the chicken epsilon-globin promoter involved in stage-specific interaction with the beta/epsilon enhancer. J. Biol. Chem., 271, 25459-25467.

Matallana, E., Franco, L. & Perezortin, J.E. (1992) Chromatin structure of the yeast suc2 promoter in regulatory mutants. Mol. Gen. Genet., 231, 395-400.

Mathis, D., Oudet, P. and Chambon, P. (1980). Prog. Nucl. Acids Res. Mol. Biol. 24, 1-55

Mathog, D., Hochstrasser, M., Gruenbaum, Y., Saumweber, H. & Sedat, J. (1984) Characteristic folding pattern of polytene chromosomes in Drosophila salivary gland nuclei. Nature, 308, 414-421.

McArthur, M. & Thomas, J.O. (1996) A preference of histone H1 for methylated DNA. EMBO J., 15, 1705-1714.

McGhee, J.D. & Felsenfeld, G. (1980) Nucleosome structure. Annu. Rev. Biochem., 49, 1115-1156.

McGhee, J.D., Nickol, J.M., Felsenfeld, G. & Rau, D.C. (1983) Higher order structure of chromatin - orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length. Cell, 33, 831-841.

McGhee, J.D., Rau, D.C., Charney, E. & Felsenfeld.G. (1980) Orientation of the nucleosomes within the higher order structure of chromatin. Cell, 22, 87-96.

McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D., and Felsenfeld, G. (1981). A 200-base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. Cell, 27, 45-55.

McNamara, P.T., Bolshoy, A., Trifonov, E.N. & Harrington, R.E. (1990) Sequence-dependent kinks induced in curved DNA. J Biomol. Struc. Dynam., 8, 529-538.

McPherson, C.E., Shim, E-Y., Friedman, D.S. and Zaret, K.S. (1993) An active tissue-specific enhancer bound transcription factors existing in a precisely positioned nucleosome array. Cell, 75, 387-398.

Meehan, R., Lewis, J., Cross, S., Nan, X., Jeppeson P. & Bird, A. (1992) Transcriptional repression by methylation of CpG. J. Cell Sci., S16, 9-14.

Meersseman, G., Pennings, S. & Bradbury, E.M. (1991) Chromatosome positioning on assembled long chromatin linker histones affect nucleosome placement on 5S rDNA. J. Mol. Biol., 220, 89-100.

Meersseman, G., Pennings, S. & Bradbury, E.M. (1992) Mobile nucleosomes - a general behavior. EMBO J., 11, 2951-2959.

Minie, M., Clark, D., Trainor, C., Evans, T., Reitman, m., Hannon, R., Gould, H., and Felsenfeld, G. (1992). Developmental regulation of globin gene-expression. J. Cell Sci. 15-20.

Mirkovitch, J., Gasser, S.M. & Laemmli, U.K. (1988) Scaffold attachment of DNA loops in metaphase chromosomes. J. Mol. Biol., 200, 101-109.

Mirkovitch, J., Mirault, M.E. & Laemmli, U.K. (1984) Organization of the higher order chromatin loop - specific DNA attachment sites on nuclear scaffold. Cell, 39, 223-232.

Mirzabekov, A.D., Shick, V.V., Belyavsky, A.V. & Bavykin, S.G. (1978) Primary organization of nucleosome core particle in chromatin: sequence of histone arrangement along DNA. Proc. Natl. Acad. Sci. USA, 75, 4184-4189.

Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J. (1996) Cell, 87, 1261-1270.

Moudrianakis, E.N. & Arents, G. (1993) Structure of the histone octamer core of the nucleosome and its potential interactions with DNA. Cold Spring Harbor Symp. Quant. Biol., 58, 273-279.

Mueller, R.D., Yasuda, H. & Bradbury, E.M. (1985a) Phosphorylation of histone H1 through the cell cycle of Physarum polycephalum: 24 sites of phosphorylation at metaphase. J. Biol. Chem., 260, 5081-5086.

Mueller, R.D., Yasuda, H., Hatch, C.L., Bonner, W.M. & Bradbury, E.M. (1985b) Identification of ubiquitinated histones 2A and 2B in Physarum polycephalum. Disappearance of these proteins at metaphase and reappearance at anaphase. J. Biol. Chem., 260, 5147-5153.

Mueller, P.R. and Wold B. (1989) In vivo footprinting of a muscle-specific enhancer by ligation-mediated PCR. J. Mol. Biol., 256, 503-516)

Muller, M.M., Gerster, T. and Schaffner, W. (1988) Enhancer Sequences and the regulation of gene transcription. Eur. J. Biochem., 176, 485-495.

Muyldermans, S. & Travers, A.A. (1994) DNA sequence organization in chromatosomes. J. Mol. Biol., 235, 855-870.

Nacheva, G.A., Preobrazhenskaya, O.V., Melnikova, A.F. & Karpov, V.L. (1989) Structural changes in chromatin of the Drosophila heat shock protein 70 gene during transcription. Mol Biol. 23, 695-703.

Nagai, H., Yuzawa, H. & Yura, T. (1991) Interplay of 2 cis-acting mRNA regions in translational control of sigma 32 synthesis during the heat shock response of *Escherichia coli.* Proc Natl Acad Sci USA. 88, 10515-10519.

Nan, X.S., Ng, HH., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by methyl CpG binding protein MeCP2 involves a histone deacetylase complex. Nature, 393, 386-389.

Nedospasov, S.A. and Georgiev, G.P., (1980) Non-random cleavage of SV40 DNA in the compact minichromosome and free in solution by micrococcal nuclease. Biochem. Biophys. Res. Comm., 92, 532-539.

Nelson, D.A., Ferris, R.C., Zhang, D.E., and Ferenz, C.R. (1986). The beta-globin domain in immature chicken erythrocytes -enhanced solubility is coincident with histone hyperacetylation. Nucl. Acids Res., 14, 1667-1682.

Neubauer, B., Linxweiler, W. & Horz, W. (1986) DNA engineering shows that nucleosome phasing on the African green monkey alpha-satellite is the result of multiple additive histone-DNA interactions. J. Mol. Biol., 190, 639-645.

Nickel, B.E., Allis, C.D. & Davie, J.R. (1989) Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. Biochemistry, 28, 958-963.

Nickol, J., Behe, M. & Felsenfeld, G. (1982) Effect of the B-Z transition in poly(dG-m5dC).poly(dG-m5dC) on nucleosome formation. Proc. Natl. Acad. Sci. USA, 79, 1771-1775.

Nickol, J.M. and Felsenfeld, G. (1983). DNA conformation at the 5' end of the chicken adult beta-globin gene. Cell, 35, 467-477.

Nickol, J.M. and Felsenfeld, G. (1988). Bidirectional control of the chicken beta-globin and epsilon-globin genes by a shared enhancer. Proc. Natl. Acad. Sci. USA 85, 2548-2552.

Nissen, M.S. & Reeves, R. (1995) Changes in superhelicity are introduced into closed circular DNA by binding of high-mobility group protein I/Y. J. Biol. Chem., 270, 4355-4360.

Nobile, C., Nickol, J. & Martin, R.G. (1986) Nucleosome phasing on a DNA fragment from the replication origin of simian virus 40 and rephasing upon cruciform formation of the DNA. Mol. Cell. Biol., 6, 2916-2922.

Noll, M. & Kornberg, R.D. (1977) Action of micrococcal nuclease on chromatin and the location of histone H1. J. Mol. Biol., 109, 393-404.

Noll, M. (1974) Internal structure of the chromatin subunit. Nucl. Acids Res., 1, 1573-1578.

Norton, V.G., Imai, B.S., Yau, P. & Bradbury, E.M. (1989) Histone acetylation reduces nucleosome core particle linking number change. Cell, 57, 449-457.

Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995) Defective hematopoiesis in fetal liver resulting from inactivation of the eklf gene. Nature, 375, 316-318.

Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. & Nakatani, Y. (1996) The transcriptional coactivators p300 and cbp are histone acetyltransferases. Cell, 87, 953-959.

Ohalloran, T.V., Lippard, S.J., Richmond, T.J. & Klug, A. (1987) Multiple heavy-atom reagents for macromolecular X-ray structure determination - application to the nucleosome core particle. J. Mol. Biol., 194, 705-712.

Ohsumi, K., Katagiri, C. & Kishimoto, T. (1993) Chromosome condensation in Xenopus mitotic extracts without histone H1. Science, 262, 2033-2035.

Oliviero, S. & Struhl, K. (1991) Synergistic transcriptional enhancement does not depend on the number of acidic activation domains bound to the promoter. Proc. Natl. Acad. Sci. USA, 88, 224-228.

Olson, W.K., Marky, N.L., Jernigan, R.L. & Zhurkin, V.B. (1993) Influence of fluctuations on DNA curvature - a comparison of flexible and static wedge models of intrinsically bent DNA. J. Mol. Biol., 232, 530-551.

Orkin, S.H. (1995). Regulation of globin gene-expression in erythroid-cells. Eur. J.Biochem. 231, 271-281.

Ortiz, B.D., Cado, D., Chen, V., Diaz, P.W., and Winoto, A. (1997). Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. EMBO J., 16, 5037-5045.

Oudet, P., Gross-Bellard, M. & Chambon, P. (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell, 4, 281-300.

Owen-Hughes, T. & Workman, J.L. (1996) Remodeling the chromatin structure of a nucleosome array by transcription factor targeted trans displacement of histones. EMBO J., 15, 4702-4712.

Owen-Hughes, T., Utley, R.T., Cote, J., Peterson, C.L. & Workman, J.L. (1996) Persistent site specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. Science, 273, 513-516.

Panetta, G., Butinelli, M., Flaus, A., Richmond, T.J. and Rhodes, D. (1998) Differential nucleosome positioning on the Xenopus oocyte and somatic 5S rRNA genes determines both TFIIIA and H1 binding: A mechanism for selective H1 repression. J. Mol. Biol., 282, 683-697.

Paranjape, S.M., Kamakaka, R.T. & Kadonaga, J.T. (1994) Role of chromatin structure in the regulation of transcription by RNA polymerase II. Annu. Rev. Biochem., 63, 265-297.

Paulson, J.R. & Laemmli, U.K. (1977) The structure of histone-depleted metaphase chromosomes. Cell, 12, 817-828.

Pazin, M.J., Kamakaka, R.T. & Kadonaga, J.T. (1994) ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. Science, 266, 2007-2011.

Pederson, D.S., Thoma, F. & Simpson, R.T. (1986) Core particle, fiber, and transcriptionally active chromatin structure. Annu. Rev. Cell. Biol., 2, 117-147.

Pehrson, J.R. (1989) Thymine dimer formation as a probe of the path of DNA in and between nucleosomes in intact chromatin. Proc. Natl. Acad. Sci. USA, 86, 9149-9153.

Penner, C.G. and Davie, J.R. (1992). Multisubunit erythroid complexes binding to the enhancer element of the chicken histone-H5 gene. Biochem. J., 283, 905-911.

Pennings, S., Meersseman, G. & Bradbury, E.M. (1991) Mobility of positioned nucleosomes on 5S rDNA. J. Mol. Biol., 220, 101-110.

Pennings, S., Meersseman, G. & Bradbury, E.M. (1994) Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. Proc. Natl. Acad. Sci. USA, 91, 10275-10279.

Pennings, S., Muyldermans, S., Meersseman, G. & Wyns, L. (1989) Formation, stability and core histone positioning of nucleosomes reassembled on bent and other nucleosome-derived DNA. J. Mol. Biol., 207, 183-192.

Pérez-Ortin, J.E., Estruch, F., Matallana, E. and Franco, L. (1986) Sliding-end-labelling. FEBS, 208, 31-33.

Perezortin, J.E., Matallana, E. & Franco, L. (1989) Chromatin structure of yeast genes. Yeast, 5, 219-238.

Perlmann, T. & Wrange, O. (1988) Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. EMBO J., 7, 3073-3079.

Perrin, D.M. et al., (1996) Oxidative chemical nucleases. Prog. in Nucl. Acids Res. and Mol. Biol., 52, 123-150.

Perry, C.A., Dadd, C.A., Allis, C.D. & Annunziato, A.T. (1993) Analysis of nucleosome assembly and histone exchange using antibodies specific for acetylated H4. Biochemistry, 32, 13605-13614.

Pfeffer, U., Ferran, N. Tosetti, F. and Vidali, G. (1989) Histone Acetylation in conjugating *Tetrahymena thermophila*. J.Cell Biol., 109, 1007-1014.

Phi-Van, L. and Stratling, W.H. (1988) The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. EMBO J., 7, 655-664.

Pina, B., Bruggemeier, U. & Beato, M. (1990) Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. Cell, 60, 719-731.

Polach, K.J. & Widom, J. (1995) Mechanism of protein access to specific DNA-sequences in chromatin - a dynamic equilibrium model for gene regulation. J. Mol. Biol., 254, 130-149.

Postnikov, Y.V., Shick, V.V., Belyavsky, A.V., Khrapko, K.R., Brodolin, K.L., Nikolskaya, T.A., and Mirzabekov, A.D. (1991). Distribution of high mobility group proteins 14/17 and linker histone-H1 and histone-H5 on transcribed and non-transcribed regions of chicken erythrocyte chromatin. Nucl. Acids Res., 19, 717-725.

Prunell, A. and Kornberg, R.D. (1978) Relation of nucleosomes to DNA sequence. Cold Spring Habour Symposia on Quan. Biol., 42, 103-108.

Pruss, D. & Wolffe, A.P. (1993) Histone DNA contacts in a nucleosome core containing a Xenopus 5S rRNA gene. Biochemistry, 32, 6810-6814.

Pruss, D., Hayes, J.J. & Wolffe, A.P. (1995) Nucleosomal anatomy - where are the histones? BioEssays, 17, 161-170.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N. & Wolffe, A.P. (1996) An asymmetric model for the nucleosome - a binding site for linker histones inside the DNA gyres. Science, 274, 614-617.

Pruzina, S., Antoniou, M., Hurst, J., Grosveld, F., and Philipsen, S. (1994). Transcriptional activation by hypersensitive site 3 of the human beta-globin locus-control region in murine erythroleukemia-cells. Biochim. Biophys. Acta, 1219, 351-360.

Quinn, J., Fyrberg, A.M., Ganster, R.W., Schmidt, M.C. & Peterson, C.L. (1996) DNA-binding properties of the yeast SWI/SNF complex. Nature, 379, 844-847.

Quivy, J.P. and Becker, P.B., (1996) The architecture of the heat-inducible Drosophila hsp27 promoter in nuclei. J. Mol. Biol., 256, 249-263.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. & Sweet, R.M. (1993) Crystal-structure of globular domain of histone H5 and its implications for nucleosome binding. Nature, 362, 219-223.

Ramakrishnan, V. (1995) The histone fold - evolutionary questions. Proc. Natl. Acad. Sci. USA, 92, 11328-11330.

Ramakrishnan, V., (1997) Histone H1 and chromatin higher order structure. Critical Rev. Euk. Gene Expression, 7, 215-230.

Ramsay, N. (1986) Deletion analysis of a DNA sequence that positions itself precisely on the nucleosome core. J. Mol. Biol., 189, 179-188.

Razin, S.V., Rzeszowskawolny, J., Moreau, J., and Scherrer, K. (1985). Localization of sites of DNA attachment to the nuclear matrix in the domain of the chicken alpha-globin genes in functionally active and functionally inactive nuclei. Molecular Biology 19, 376-384.

Razin, A. & Cedar, H. (1991) DNA methylation and gene repression, Microbiol. Rev. 55, 451-458.

Reese, J.C., Apone, L., Walker, S.S., Griffin, L.A. & Green, M.R. (1994) Yeast TAF(II)s in a multisubunit complex required for activated transcription. Nature, 371, 523-527.

Reeves, R. & Nissen, M.S. (1993) Interaction of high mobility group I(Y) nonhistone proteins with nucleosome core particles. J. Biol. Chem., 268, 21137-21146.

Reik, A., Schutz, G. & Stewart, A.F. (1991) Glucocorticoids are required for establishment and maintenance of an alteration in chromatin structure - induction leads to a reversible disruption of nucleosomes over an enhancer. EMBO J., 10, 2569-2576.

Reitman, M. and Felsenfeld, G. (1988). Mutational analysis of the chicken beta-globin enhancer reveals 2 positive-acting domains. Proc. Natl. Acad. Sci. USA 85, 6267-6271.

Reitman M. and Felsenfeld, G. (1990) Developmental regulation of Topoisomerase II sites. Mol. Cell. Biol., 10, 2772-2786.

Reitman, M., Grasso, J.A., Blumenthal, R. and Lewit, P. (1993) Primary sequence evolution, and repetitive elements of the Gallus gallus (chicken) β-globin cluster. Genomics, 18, 616-626.

Reitman, M., Lee, E., Westphal, H., and Felsenfeld, G. (1990). Site-independent expression of the chicken β^A-globin gene in transgenic mice. Nature, 348, 749-752.

Renz, M. & Day, L.A. (1976) Transition from noncooperative of cooperative and selective binding of histone H1 to DNA. Biochemistry, 15, 3220-3228.

Renz, M., Nehls, P. & Hozier, J. (1977) Involvement of histone H1 in the organization of the chromosome fiber. Proc. Natl. Acad. Sci. USA, 74, 1879-1883.

Rhodes, D. (1985) Structural analysis of a triple complex between the histone octamer, a Xenopus gene for 5S RNA and transcription factor III A. EMBO J., 4, 3473-3482.

Rich, A., Nordheim, A. and Wang, A.H.-J. (1984) The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem., 53, 791-846.

Richard-Foy, H. & Hager, G.L. (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. EMBO J., 6, 2321-2328.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. & Klug, A. (1984) Structure of the nucleosome core particle at 7Å resolution. Nature, 311, 532-537.

Ridsdale, J.A., Hendzel, M.J., Delcuve, G.P. & Davie, J.R. (1990) Histone acetylation alters the capacity of the H1 histones to condense transcriptionally active competent chromatin. J. Biol. Chem., 265, 5150-5156.

Ring, D. & Cole, R.D. (1983) Close contacts between H1-histone molecules in nuclei. J. Biol. Chem., 258, 5361-5364.

Ronchi, A., Berry, M., Raguz, S., Imam, A., Yannoutsos, N., Ottolenghi, S., Grosveld, F., and Dillon, N. (1996). Role of the duplicated CCAAT box region in gamma-globin gene-regulation and hereditary persistence of fetal hemoglobin. EMBO J., 15, 143-149.

Roth, S.Y. & Allis, C.D. (1992) Chromatin condensation - does histone H1 dephosphorylation play a role? Trends Biochem. Sci., 17, 93-98.

Roth, S.Y., Dean, A. & Simpson, R.T. (1990) Yeast α-2 repressor positions nucleosomes in trp1 ars1 chromatin. Mol. Cell. Biol., 10, 2247-2260.

Roth, S.Y., Schulman, I.G., Richman, R., Cook, R.G. & Allis, C.D. (1988) Characterization of phosphorylation sites in histone H1 in the amitotic macronucleus of *Tetrahymena* during different physiological states. J. Cell. Biol., 107, 2473-2482.

Runkel, L. and Nordheim, A. (1986). Conformational DNA transition in the *in vitro* torsionally strained chicken beta-globin 5' region. Nucl. Acids Res., 14, 7143-7158.

Rupp, R.A.W., Nicolas, R.H., Borgmeyer, U., Lobanenkov, V.V., Plumb, M.A., Sippel, A.E., and Goodwin, G.H. (1988). TGGCA protein is present in erythroid nuclei and binds within the nuclease-hypersensitive sites 5' of the chicken β^A-globin and β^H-globin genes. Eur. J. Biochem. 177, 505-511.

Ruppert, S. & Tjian, R. (1995) Human TAF(pi)250 interacts with rap74 - implications for RNA polymerase II initiation. Genes Dev., 9, 2747-2755.

Russanova, V.R., Dimitrov, S.I., Makarov, V.L. & Pashev, I.G. (1987) Accessibility of the globular domain of histone H1 and histone H5 to antibodies upon folding of chromatin. Eur. J. Biochem., 167, 321-326.

Sambrook, J., Fritsh, E.F. & Maniatis, T. (1989) Molecular cloning. A laboratory manual. 3rd Ed., Cold Spring Harbor Laboratory Press, New York,

Sandeen, G., Wood, W.I. & Felsenfeld, G. (1980) The interaction of high mobility group proteins HMG14 and 17 with nucleosomes. Nucl.Acids Res., 8, 3757-3778.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, 74, 5463-5467.

Satchwell, S.C. & Travers, A.A. (1989) Asymmetry and polarity of nucleosomes in chicken erythrocyte chromatin. EMBO J., 8, 229-238.

Satchwell, S.C., Drew, H.R. & Travers, A.A. (1986) Sequence periodicities in chicken nucleosome core DNA. J. Mol. Biol., 191, 659-675.

Sauer, F., Wassarman, D.A., Rubin, G.M. & Tjian, R. (1996) TAF(II)s mediate activation of transcription in the drosophila embryo. Cell, 87, 1271-1284.

Sazanov, A.A., Dukelskaya, A.D., and Alexeyevich, L.A. (1992). Chromosome localization of the chicken beta-like globin gene-cluster by in situ hybridization. Tsitologiya 34, 35-40.

Schild, C., Claret, F.X., Wahli, W. & Wolffe, A.P. (1993) A nucleosome dependent static loop potentiates estrogen-regulated transcription from the Xenopus vitellogenin B1 promoter invitro. EMBO J., 12, 423-433.

Schulze, E. et al., (1993) Structural and functional differences between Histone variants with differential intranuclear distribution. Proc. Natl. Acad. Sci. USA, 90, 2481-2485.

Schneider, T., Stormo, G.D., Gold, L. & Ehrenfeuchs, A. (1986) Information content of binding sites on nucleotide sequences. J. Mol. Biol., 188, 415-431.

Shapiro, S.G. (1991). Uniformity in the nonsynonymous substitution rates of embryonic beta-globin genes of several vertebrate species. J. Mol. Evol., 32, 122-127.

Shelton, D.A., Stegman, L., Hardison, R., Miller, W., Bock, J.H., Slightom, J.L., Goodman, M., and Gumucio, D.L. (1997). Phylogenetic footprinting of hypersensitive site 3 of the beta-globin locus control region. Blood 89, 3457-3469.

Shen, C-H, (1997) The role of linker histone globular domains in chromatosome function. Thesis. University of Edinburgh.

Shen, X.T. & Gorovsky, M.A. (1996) Linker histone H1 regulates specific gene-expression but not global transcription *in vivo*. Cell, 86, 475-483.

Shen, X.T., Yu, L.L., Weir, J.W. & Gorovsky, M.A. (1995) Linker histones are not essential and affect chromatin condensation *in vivo*. Cell, 82, 47-56.

Shick, V.V., Belyavsky, A.G., Bavykin, S.G. & Mirzabekov, A.D. (1980) Primary organization of the nucleosome core particles. Sequential arrangement of histones along DNA. J. Mol. Biol., 139, 491-517.

Shimizu, M., Roth, S.Y., Szentgyorgyi, C. & Simpson, R.T. (1991) Nucleosomes are positioned with base pair precision adjacent to the α-2 operator in Saccharomyces cerevisiae. EMBO J., 10, 3033-3041.

Shpigelman, E.S., Trifonov, E.N. & Bolshoy, A. (1993) Curvature - software for the analysis of curved DNA. CABIOS, 9, 435-440.

Shrader, T.E. & Crothers, D.M. (1989) Artificial nucleosome positioning sequences. Proc. Natl. Acad. Sci. USA, 86, 7418-7422.

Sigman, D.S., Kuwarba, M.D., Chen, C-H.B. and Bruce, T.W. (1990) Nuclease activity of the 1,10-Phenanthroline-copper in the study of protein DNA interactions. Meths. Enzymol., 208, 414-425.

Silverman, J.A. & Hill, B.A. (1995) Characterization of the basal and carcinogen regulatory elements of the rat mdr1b promoter. Mol. Carcinigenesis, 13, 50-59.

Simpson, R.T. (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry, 17, 5524-5531.

Simpson, R.T. (1990) Nucleosome positioning can affect the functioning of a cis acting DNA element *in vivo*. Nature, 343, 387-389.

Simpson, R.T. (1991) Nucleosome positioning - occurrence, mechanisms, and functional consequences. Prog. Nucl. Acids Res. Mol. Biol., 40, 143-184.

Sinden, R.R. (1994) Left-handed Z DNA. In: DNA structure and function. New York, Academic Press.

Singal R., Ferris, R., Little, J.A., Wang, S.J. & Ginder, G.D. (1997) Methylation of the minimal promoter of an embryonic globin gene silences transcription in primary erythroid cells. Proc. Natl. Acad. Sci. USA, 94, 13724-13729.

Sivolob, A.V. & Khrapunov, S.N. (1995) Translational positioning of nucleosomes on DNA - the role of sequence-dependent isotropic DNA bending stiffness. J. Mol. Biol., 247, 918-931.

Southern, E.M. (1975) Detection of specific sequences among fragments separated by gel electrophoresis. J. Mol. Biol., 98, 503-517.

Smith, Seale, R.L. and Yu, J. (1983) Transcribed DNA exhibits an altered nucleosome spacing. Proc. Natl. Acad. Sci. USA, 80, 5505-5509.

Spiker, S. (1982) Histone variants in plants - evidence for primary structure variants differing in molecular weight. J. Biol. Chem., 257, 4250-4255.

Sprous, D., Zacharias, W., Wood, Z.A. & Harvey, S.C. (1995) Dehydrating agents sharply reduce curvature in DNAs containing A-tracts. Nucl. Acids Res., 23, 1816-1821.

Srinivasan, A.R., Torres, R., Clark, W. & Olson, W.K. (1987) Base sequence effects in double helical DNA .1. potential energy estimates of local base morphology. J Biomol. Struc. Dynam., 5, 459

Stalder, J., Larsen, A., Engel, J.D., Dolan, M., Groudine, M. & Weintraub, H. (1980) Tissue-specific cleavages in the globin chromatin domain introduced by DNase I. Cell, 20, 451-460.

Stamatoyannopoulos, J.A., Goodwin, A., Joyce, T., Lowrey, C.H. (1995). NF-E2 and GATA binding motifs are required for the fromation of DNase I hypersensitive site 4 of the human beta globin locus control region. EMBO J., 14, 106-116.

Staynov, D.Z. & Crane-Robinson, C. (1988) Footprinting of linker histones H5 and histones H1 on the nucleosome. EMBO J 7, 3685-3691.

Staynov, D.Z. (1983) Possible nucleosome arrangements in the higher order structure of chromatin. Intl. J. Biol.Macromolec., 5, 3-9.

Stefanovsky, V.Y., Dimitrov, S.I., Russanova, V.R., Angelov, D. & Pashev, I.G. (1989) Laser-induced crosslinking of histones to DNA in chromatin and core particles - implications in studying histone-DNA interactions. Nucl. Acids Res., 17, 10069-10081.

Steger, D.J. & Workman, J.L. (1996) Remodeling chromatin structures for transcription: what happens to the histones? BioEssays, 18, 875-884.

Stein, A. & Bina, M. (1984) A model chromatin assembly system - factors affecting nucleosome spacing. J. Mol. Biol., 178, 341-363.

Stein, A. & Kunzler, P. (1983) Histone H5 can correctly align randomly arranged nucleosomes in a defined in vitro system. Nature, 302, 548-550.

Studitsky, V.M., Clark, D.J. & Felsenfeld, G. (1995) Overcoming a nucleosomal barrier to transcription. Cell, 83, 19-27.

Sugarmen, B.J., Dodgson, J.B. & Engel, J.D. (1983) Genomic organization, DNA sequence, and expression of chicken embryonic histone genes. J. Biol. Chem., 258, 9005-9016.

Sun, Y.L., Xu, Y.Z., Bellard, M., and Chambon, P. (1986). Digestion of the chicken beta-globin gene chromatin with micrococcal nuclease reveals the presence of an altered nucleosomal array characterized by an atypical ladder of DNA fragments. EMBO J., 5, 293-300.

Suau, P., Bradbury, E.M. and Baldwin, J.P. (1979) Higher order chromatin structure in solution. Eur. J. Biochem., 97, 593-602.

Svaren, J. & Horz, W. (1996) Regulation of gene expression by nucleosomes. Curr. Opin. Genet. Dev., 6, 164-170.

Svaren, J. & Horz, W. (1997) Transcription factors vs nucleosomes: regulation of the Pho5 promoter in yeast. Trends Biochem. Sci., 22, 93-97.

Swedlow, J.R., Sedat, J.W. & Agard, D.A. (1993) Multiple chromosomal populations of topoisomerase-II detected in vivo by time-lapse, 3-dimensional wide-field microscopy. Cell, 73, 97-108.

Tanaka, S., Zatchej, M. and Thoma, F. (1992) Artificial nucleosome positioning sequences tested in yeast minichromosomes: a rotational setting is not sufficient to position nucleosomes *in vivo*. EMBO J., 11, 1187-1193.

Taunton, J., Hassig, C.A. & Schreiber, S.L. (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator rpd3p. Science, 272, 408-411.

Taylor, I.C.A., Workman, J.L., Schuetz, T.J. & Kingston, R.E. (1991) Facilitated binding of Gal4 and heat shock factor to nucleosomal templates - differential function of DNA-binding domains. Genes Dev., 5, 1285-1298.

Tewari, R., Gillemans, N., Harper, A., Wijgerde, M., Zafarana, G., Drabek, D., Grosveld, F., and Philipsen, S. (1996). The human beta-globin locus control region confers an early embryonic erythroid-specific expression pattern to a basic promoter driving the bacterial lac Z gene. Development, 122, 3991-3999.

Thatcher, T.H. & Gorovsky, M.A. (1994) Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. Nucl. Acids Res., 22, 174-179.

Thoma, F. & Koller, T.H. (1977) Influence of histone H1 on chromatin structure. Cell, 12, 101-107.

Thoma, F., Koller, T.H. & Klug, A. (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J. Cell. Biol., 83, 403-427.

Thoma, F. & Koller, T. (1981) Unravelled nucleosomes, nucleosome beads and higher order structures of chromatin - influence of non-histone components and histone H1. J. Mol. Biol., 149, 709-733.

Thoma, F., Losa, R. & Koller, T. (1983) Involvement of the domains of histones H1 and H5 in the structural organization of soluble chromatin. J. Mol. Biol., 167, 619-640.

Thoma, F., Bergman, L.W. & Simpson, R.T. (1984) Nuclease digestion of circular TRPARS1 chromatin reveals positioned nucleosomes separated nuclease sensitive regions. J. Mol. Biol., 177, 715-733.

Thoma, F. & Simpson, R.T. (1985) Local protein-DNA interactions may determine nucleosome positions on yeast plasmids. Nature, 315, 250-252.

Thoma, F. (1986) Protein-DNA interactions and nuclease sensitive regions determine nucleosome positions on yeast plasmid chromatin. J. Mol. Biol., 190, 177-190.

Thoma, F. & Zatchej, M. (1988) Chromatin folding modulates nucleosome positioning in yeast minichromosomes. Cell, 55, 945-953.

Thoma, F. (1991) Structural-changes in nucleosomes during transcription - strip, split or flip. Trends Genet., 7, 175-177.

Thoma, F. (1992) Nucleosome positioning. Biochim. Biophys. Acta, 1130, 1-19.

Thomas, J.O. & Kornberg, R.D. (1975) An octamer of histones in chromatin and free in solution. Proc. Natl. Acad. Sci. USA, 72, 2626-2630.

Thomas, J.O. (1984) The higher-order structure of chromatin and histone H1. J. Cell Sci. Suppl. 1, 1-20.

Thomas, G.H. and Elgin, S.C.R. (1988) Protein architecture of the DNase I-hypersensitive region of the *Drosophila* hsp-26 promoter. EMBO J., 7, 7291-7301.

Thompson, J.S., Johnson, L.M. & Grunstein, M. (1994) Specific repression of the yeast silent mating locus HMR by an adjacent telomere. Mol. Cell. Biol., 14, 446-455.

Thorne, A.W., Kmiciek, D., Mitchelson, K., Sautiere, P. & Crane-Robinson, C. (1990) Patterns of histone acetylation. Eur. J. Biochem., 193, 701-713.

Travers, A. (1989) DNA structure - curves with a function. Nature, 341, 184-185.

Travers, A.A. & Klug, A. (1987) The bending of DNA in nucleosomes and its wider implications. Phil. Trans. Royal Soc. London Series B- Biological Sciences, 317, 537-561.

Travers, A.A. & Muyldermans, S.V. (1996) A DNA-sequence for positioning chromatosomes. J. Mol. Biol., 257, 486-491.

Tremethick, D.J. & Molloy, P.L. (1988) Effects of high mobility group protein-1 and protein-2 on initiation and elongation of specific transcription by RNA polymerase II invitro. Nucl. Acids Res., 16, 11107-11123.

Trifonov, E.N. (1980) Sequence-dependent deformational anisotropy of chromatin DNA. Nucl. Acids Res., 8, 4041-4053.

Trifonov, E.N. (1990) DNA as a chromatin organiser in Beveridge, D.L. and Lavery, R. (eds.) Theoretical Biochem. Mol. Biol., Albany New York, Adenine press.

Truss, M., Bartsch, J., Schelbert, A., Hache, R.J.G. & Beato, M. (1995) Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo. EMBO J., 14, 1737-1751.

Tsukiyama, T. & Wu, C. (1995) Purification and properties of an ATP-dependent nucleosome remodeling factor. Cell, 83, 1011-1020.

Tsukiyama, T., Becker, P.B. & Wu, C. (1994) ATP-dependent nucleosome disruption at a heat shock promoter mediated by binding of GAGA transcription factor. Nature, 367, 525-532.

Tuan, D., Solomon, W., Li, Q. and London, I.M. (1985) The beta-like-globin gene domain in humans. Proc. Natl. Acad. Sci. USA, 77, 3816-3820.

Turnell, W.G. & Travers, A.A. (1992) Algorithms for prediction of histone octamer binding sites. Meths. Enzymol., 212, 387-399.

Turner, B.M. & O'Neill, L.P. (1995) Histone acetylation in chromatin and chromosomes. Seminars In Cell Biology, 6, 229-236.

Uemura, T., Ohkura, H., Adachi, Y., Morino, K., Shiozaki, K. & Yanagida, M. (1987) DNA topoisomerase II is required for condensation and separation of mitotic chromosomes in *S.* pombe. Cell, 50, 917-925.

Ulyanov, A.V. & Stormo, G.D. (1995) Multi alphabet consensus algorithm for identification of low specificity protein-DNA interactions. Nucl. Acids Res., 23, 1434-1440.

Ulyanov, N.B. & Zhurkin, V.B. (1984) Sequence-dependent anisotropic flexibility of B-DNA - a conformational study. J Biomol. Struc. Dynam., 2, 361-385.

Ura, K. & Wolffe, A.P. (1996) Reconstruction of transcriptionally active and silent chromatin. Meths. Enzymol., 274, 257-271.

Ura, K., Hayes, J.J. & Wolffe, A.P. (1995) A positive role for nucleosome mobility in the transcriptional activity of chromatin templates - restriction by linker histones. EMBO J., 14, 3752-3765.

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

van Holde, K.E. and Zlatanova, J., (1995) Chromatin higher order structure: chasing a mirage. J. Biol. Chem., 270, 8373-8376.

Varga-Weisz, P., van Holde, K. & Zlatanova, J. (1994) Competition between linker histones and HMG1 for binding to four-way junction DNA: implications for transcription. Biochem. Biophys. Res. Comm., 203, 1904-1911.

Varga-Weisz, P.D., Blank, T.A. & Becker, P.B. (1995) Energy-dependent chromatin accessibility and nucleosome mobility in a cell-free system. EMBO J., 14, 2209-2216.

Verreault, A. and Thomas, J.O. (1993). Chromatin structure of the Beta globin chromosomal domain in adult chicken erythrocytes. Cold Spring Harbor Symposia On Quan. Biol., 58, 15-24.

Vettese-Dadey, M., Grant, P.A., Hebbes, T.R., Crane-Robinson, C., Allis, C.D. & Workman, J.L. (1996) Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J., 15, 2508-2518.

Vettesse-Dadey, M., Walter, P., Chen, H., Juan, L.J. & Workman, J.L. (1994) Role of the histone amino termini in faciliated binding of a transcription factor, GAL4-AH, to nucleosome cores. Mol. Cell. Biol., 14, 970-981.

Villeponteau, B. and Martinson, H. (1981). Isolation and characterization of the complete chicken beta-globin gene region - frequent deletion of the adult beta-globin genes in lambda. Nucl. Acids Res., 9, 3731-3746.

Villeponteau, B., Brawley, J., and Martinson, H.G. (1992). Nucleosome spacing is compressed in active chromatin domains of chick erythroid-cells. Biochemistry, 31, 1554-1563.

Villeponteau, B., Landes, G.M., Pankratz, M.J., and Martinson, H.G. (1982). The chicken beta globin gene region. J. Biol. Chem., 257, 11015-11023.

Wada-Kiyama Y., Kiyama R. (1996). An intrachromosomal Repeating Unit Based on DNA Bending. Mol. Cell. Biol. 16, 5664-5673.

Wada-Kiyama, Y. and Kiyama, R. (1994). Periodicity of DNA bend sites in human epsilon-globin gene region. J. Biol. Chem., 269, 22238-22244.

Wadakiyama, Y. and Kiyama, R. (1995). Conservation and periodicity of DNA bend sites in the human beta-globin gene locus. J. Biol. Chem., 270, 12439-12445.

Wade, P.A., Pruss, D. & Wolffe, A.P. (1997) Histone acetylation: chromatin in action. Trends Biochem. Sci., 22, 128-132.

Wall, G., Varga-Weisz, P.D., Sandaltzopoulos, R. & Becker, P.B. (1995) Chromatin remodeling by GAGA factor and heat shock factor at the hypersensitive Drosophila hsp26 promoter in vitro. EMBO J., 14, 1727-1736.

Walmsley, M.E., Buckle, R.S., Allan, J. & Patient, R.K. (1991) A chicken red cell inhibitor of transcription associated with the terminally differentiated state. J Cell Biol, 114, 9-19.

Walter, P.P., Owen-Hughes, T.A., Cote, J. & Workman, J.L. (1995) Stimulation of transcription factor binding and histone displacement by nucleosome assembly protein-1 and nucleoplasmin requires disruption of the histone octamer. Mol. Cell. Biol., 15, 6178-6187.

Wandersee, N.J., Ferris, R.C., and Ginder, G.D. (1996). Intronic and flanking sequences are required to silence enhancement of an embryonic beta-type globin gene. Mol. Cell. Biol. 16, 236-246.

Wang, J., Hogan, M. & Austin, R.H. (1982) DNA motions in the nucleosome core particle. Proc. Natl. Acad. Sci. USA, 79, 5896-5900.

Wang, J.N.C. and Hogan, M. (1985). An equilibrium between distorted and undistorted DNA in the adult chicken beta-a-globin gene. J. Biol. Chem., 260, 8194-8202.

Wang, Y. & Griffith, J. (1996) The $[(G/C)_3NN]_n$ motif. A common DNA repeat that excludes nucleosomes. Proc. Natl. Acad. Sci. USA, 93, 8863-8867.

Watson, G.D. and Crick, F.H.C. (April 25, 1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature, No. 4356, 737-738.

Wellman, S.E., Sittman, D.B. & Chaires, J.B. (1994) Preferential binding of H1e histone to GC-rich DNA. Biochemistry, 33, 384-388.

Widlund, H.R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P.E., Kahn, J.D., Crothers, D.M. and Kubista, M. (1997) Identification and Characterisation of nucleosome positioning sequences. J. Mol. Biol., 267, 807-817.

Widom, J., Finch, J.T. & Thomas, J.O. (1985) Higher order structure of long repeat chromatin. EMBO J., 4, 3189-3194.

Widom, J. & Klug, A. (1985) Structure of the 300 A chromatin filament: X-ray diffraction from oriented samples. Cell, 43, 207-213.

Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F., and Fraser, P. (1996). The role of eklf in human beta-globin gene competition. Genes Dev., 10, 2894-2902.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics *in vivo*. Nature, 377, 209-213.

Williams, S.P., Athey, B.D., Muglia, L.J., Schappe, R.S., Gough, A.H. & Langmore, J.P. (1986) Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys. J., 49, 233-248.

Williamson, R. (1970) J. Mol. Biol., 51, 157-168.

Wilson, C.J., Chao, D.M., Imbalzano, A.N., Schnitzler, G.R., Kingston, R.E. & Young, R.A. (1996) RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodelling. Cell, 84, 235-244.

Wisniewski, J.R. & Schulze, E. (1994) High-affinity interaction of dipteran high mobility group (HMG) proteins-1 with DNA is modulated by C-terminal regions flanking the HMG box domain. J. Biol. Chem., 269, 10713-10719.

Wolffe, A.P. (1991) Implications of DNA replication for eukaryotic gene expression. J. Cell Sci., 99, 201-206.

Wolffe, A.P. (1992) Chromatin: Structure and Function. Academic press.

Wolffe, A.P. (1994) Nucleosome positioning and modification: chromatin structures that potentiate transcription. Trends Biochem. Sci., 19, 240-245

Wolffe, A.P., Khochbin, S. & Dimitrov, S. (1997) What do linker histones do in chromatin? BioEssays, 19, 249-255.

Wolffe, A.P. and Hayes, J.J. (1999) Chromatin disruption and modification. Nucleic Acids Research, 27, 711-720.

Wood, W.I. and Felsenfeld, G. (1982). Chromatin structure of the chicken beta-globin gene region - sensitivity to DNase-I, micrococcal nuclease, and DNase I. J. Biol. Chem., 257, 7730-7736.

Wood, W.L.,Felsenfeld,G. (1981). Chromatin Structure of the Chicken Beta Globin Gene Region. J. Biol. Chem., 257, 7730-7736.

Woodcock, C.L., Grigoryev, S.A., Horowitz, R.A. & Whitaker, N. (1993) A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. Proc. Natl. Acad. Sci. USA, 90, 9021-9025.

Woodcock, C.L.F., Frado, L.L.Y. & Rattner, J.B. (1984) The higher order structure of chromatin - evidence for a helical ribbon arrangement. J. Cell Biol., 99, 42-52.

Woodcock, C.L. (1994) Chromatin fibers observed in situ in frozen hydrated sections -native fiber diameter is not correlated with nucleosome repeat length. J. Cell Biol., 125, 11-19.

Worcel, A., Strogatz, S. & Riley, D. (1981) Structure of chromatin and the linking number of DNA. Proc. Natl. Acad. Sci. USA, 78, 1461-1465.

Workman, J.L. & Kingston, R.E. (1992) Nucleosome core displacement in vitro via a metastable transcription factor nucleosome complex. Science, 258, 1780-1784.

Wu, C. (1980) The 5′ end of Drosophila heat shock gene in chromatin is hypersensitive to DNase I. Nature, 286, 854-860.

Wu, H.M. & Crothers, D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. Nature, 308, 509-513.

Wu, T.C. & Lichten, M. (1994) Meiosis induced double-strand break sites determined by yeast chromatin structure. Science, 263, 515-518.

Wu, C. (1997) Chromatin remodelling and control of gene expression. J. Biol. Chem., 272, 28171-28174.

Xie, X., Kokubo, T., Cohen, S.L., Mirza, U.A., Hoffman, A., Chait, B.T., Roeder, R.G., Nakatani, T. & Burley, S.K. (1996) Structural similarity between TAF's and the heterotetrameric core of the histone octamer. Nature, 380, 316-322.

Xu, C.F., Brown, M.A., Chambers, J.A., Griffiths, B., Nicolai, H. & Solomon, E. (1995) Distinct transcription start sites generate 2 forms of brca1 mRNA. Human Mol. Genet., 4, 2259-2264.

Yakoto, H., van den Engh, G., Hearst, J.E., Sachs, R.K. & Trask, B.J. (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. J. Cell Biol., 130, 1239-1249.

Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. & Nakatani, Y. (1996) A p300/cbp-associated factor that competes with the Adenoviral oncoprotein e1a. Nature, 382, 319-324.

Yao, J., Lowary, P.T. & Widom, J. (1990) Direct detection of linker DNA bending in defined length oligomers of chromatin. Proc. Natl. Acad. Sci. USA, 87, 7603-7607.

Yao, J., Lowary, P.T. & Widom, J. (1991) Linker DNA bending by the core histones of chromatin. Biochemistry, 30, 8408-8414.

Yenidunya, A. (1992) Thesis for King's College, University of London: Nucleosome positioning studies on the human ζ2-globin gene promoter.

Yenidunya, A., Davey, C., Clark, D., Felsenfeld, G. & Allan, J. (1994) Nucleosome positioning on chicken and human globin gene promoters *in vitro*: novel mapping techniques. J. Mol. Biol., 237, 401-414.

Yoon, Y.S., Kim, J.W., Kang, K.W., Kim, Y.S., Choi, K.H. & Joe, C.O. (1996) Poly (ADP-ribosyl)ation of histone H1 correlates with internucleosomal DNA fragmentation during apoptosis. J. Biol. Chem., 271, 9129-9134.

Zentgraf, H. & Franke, W.W. (1984) Differences of supranucleosomal organization in different kinds of chromatin: Cell type specific globular subunits containing different numbers of nucleosomes. J. Cell Biol., 99, 272-285.

Zhurkin, V.B. (1985) Sequence-dependent bending of DNA and phasing of nucleosomes. J Biomol. Struc. Dynam., 2, 785-804.

Zlatanova, J. & Doenecke, D. (1994) Histone H1°- a major player in cell-differentiation. FASEB J., 8, 1260-1268.

Zweidler, A. (1984) Core histone variants of the mouse: primary structure and differential expression. In: Histone genes, edited by Stein, G.S., Stein, J.L. & Marzluff, W.F. John Wiley & sons, New York, p. 339-369.

Zuckerandl, E. (1991) A general function of non-coding polynucleosome sequences. Molec. Biol. Rep., 31, 161-162.

Zuckerandl, E., (1997) Junk DNA and sectorial repression. Gene, 205, 323-343.