Biochemical and Structural Studies of Histone and Associated Proteins from Chick and Human Nuclei

by

Qin Qin Zhuang

A thesis submitted in partial fulfilment of the requirement of

Liverpool John Moores University for the degree of Doctor of Philosophy

Supervisors: Professor Colin D. Reynolds

Professor John P. Baldwin

Dr. Elaine M. Hemers

May 2012

THE FOLLOWING PUBLICATIONS, FIGURES, PLATES AND TABLES HAVE BEEN EXCLUDED ON INSTRUCTION FROM THE UNIVERSITY.

FIG 1.1 P2, FIG 1.2 P6, FIG 1.3 P7,FIG 1.4 P10, FIG1.5 P11, FIG 1.6 P12, FIG 1.7 P 15, TABLE 1.1 P16 TABLE 1.2 P17, FIG 1.8 P21, FIG1.13 P 33, 1.15 P37, FIG 1.16 P38, FIG 1.18 P41, FIG 1.19 P44, FIG1.20 P45, FIG 1.21 P52, FIG 1.22 P53, FIG 1.23 P56, FIG 1.24 P57, TABLE 1.3 P62, FIG 1.25 P64, FIG 1.26 P65, FIG 1.27 P66, FIG 1.28, P67 AND FIG 1.29 P68

Acknowledgment

First of all, I would like to give all my thanks to my supervisors Professor John Baldwin and Professor Colin Reynolds. Both of you have given me all the possible helps in these 4 years and especially during this thesis writing. Thank you very much John and Colin! Really hope we could keep working together in the future.

Special thanks to Dr. Stan Lambert, who was the technician of our laboratory. Sadly he passed away before this thesis is completed. He did not only teach me the laboratory skills, but also his way of thinking, planning, operating and analyzing. I want to say thank you very much Stan! You will be missed forever.

Many thanks to all the colleagues in the LJMU chromatin research group: Dr Chris Wood, Dr Katie Evans, Mr. Larus Foulger and Mr. Hugh Smallman. All of you have given me irreplaceable helps during my PhD work.

Many thanks to Dr James Nicholson and Dr Mark Dickman for providing very important data for my PhD work. Also thanks to all the project students who worked with me.

Finally I want to give all my thanks to my father Mr. Xiaojun Zhuang, my mother Mrs. Hong Wang, my wife Mrs. Yanzhen Sun and all my family! Without your love and support, I would not be able to complete this thesis.

Thank you all very much! Wish you all be healthy and happy every day!

QinQin Zhuang
December 2011
Liverpool, England

Abstract

Chromatin is the complex of DNA, histones and non-histone proteins that make up chromosomes and the region of dispersed chromosomes during the interphases and S phase of the cell cycle. It is found inside cell nuclei in eukaryotic cells. There are many important proteins in the nuclei which are involved in DNA replication, gene expression, DNA repair, etc. Core histones, linker histones and HMGs are the most important proteins in this group. In this thesis, a new, quick and efficient method is developed, to extract and purify proteins from cell nuclei, which is named "forward technology".

By using the "forward technology", ultra-pure native core histone octamers, dimers and tetramers were extracted from chick erythrocytes. By using the pure histone dimers and tetramers, the complexes of NAP1-dimer and NAP1-tetramer were obtained in collaboration with others. Crystals of pure histone octamers with a higher resolution than before were produced (Chapter 2).

A low KCl/phosphate wash of chick erythrocyte nuclei removed up to 1 g of proteins without nuclei lysis (cePNE1 proteins). The high KCl/phosphate soluble fraction of the cePNE1 is rich in HMG proteins, peptidyl-prolyl isomerases (FKBP3) and heatshock protein 70 which were fractionated by cation-exchange chromatography and anion-exchange chromatography (Chapter 3).

Valuable (in terms of function) high-molecular-weight proteins were enriched, and another group of nucleoproteins called cePNE5 is fractionated by the "forward technology" from chick erythrocyte nuclei. It was confirmed that HMGB proteins prefer to bind to core histone H2A-H2B dimers (Chapter 4).

It has proved possible to fractionate HMG-rich PNE1 proteins separate from a linker-histone rich nucleoprotein extract, separate from pure core histones (not as octamers) from human tissue culture cells. The method has been applied to human leukocytes obtained from the National Blood Service. This enrichment of particular groups of proteins will be useful for proteomic studies (Chapter 5).

ABBREVIATIONS

2DE	2-dimensional gel electrophoresis
ACF	ATP-utilising chromatin assembly and remodeling factor
ATP	Adeonosine Triphosphate
APS	Ammonium Persulphate
bp	base pair
cePNE	chick erythrocyte Protein Nuclear Extract
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetra-acetic Acid
FPLC	Fast protein liquid chromatography
G_0 phase	Gap 0 phase
G_1 phase	Gap 1 phase
G_2 phase	Gap 2 phase
HAT	Histone acetyl transferase
HCl	Hydrochloric Acid
HDAC	Histone deacetylase
HiTrap™ SP FF	HiTrap™ SP Sepharose Fast Flow
HiTrap™ DEAE FF	HiTrap™ DEAE Sepharose Fast Flow
HMG	High Mobility Group protein
HMGB	High Mobility Group Box
HPLC	High performance liquid chromatography
K_2HPO_4	di-Potassium Hydrogen Phosphate
KCl	Potassium Chloride
KH_2PO_4	Potassium di-Hydrogen Phosphate
M phase	Mitotic phase
MALDI	Matrix-assisted laser desorption/ionization
mAU	mili Absorbance Unit
$MgCl_2$	Magnesium Chloride
MW	Molecular Weight
NaCl	Sodium Chloride

NAP-1	Nucleosome Assembly Protein 1
OH	Hydroxyl Group
pI	Isoelectric point
PPI	Peptidylprolyl Isomerases
PTM	Post Translational Modification
RNA	Ribonulceic Acid
S phase	Synthesis phase
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SP	Sulphopropyl
SUMO	Small ubiquitin-like modifier
TEMED	N,N,N',N'-Tetramethylethylenediamine
UV	Ultraviolet

Amino acid Abbreviations

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Units

A	Absorbance
Å	Angstrom
C	Celsius
cm	Centimetre
Da	Dalton
g	Earth gravity
kDa	Kilo Dalton
L	Litre
µL	Micro-litre
µm	Micrometre
M	Molar (unit of concentration)
m	Metre
mAU	Milli-absorbance unit
mg	Milligram
min	Minute
ml	Millilitre
mM	Milli-molar (unit of concentration)
mm	Millimetre
nm	Nanometre
M.W.	Molecular Weight
RCF	Relative centrifugal force
rpm	Revolution per minute
V	Volt

Content

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Chromatin

Chromatin is the complex of DNA, histones and non-histone proteins that make up chromosomes (Figure 1.1) and the region of dispersed chromosomes during the interphases and S phase of the cell cycle. Chromosomes were considered to determine inherited characteristics in the end of the nineteenth century. Studies carried out during that time were mainly based on cytological observations with the light microscope. At the same time RNA (first identified as yeast nucleic acid), DNA (nucleic acid) and histones were all discovered as parts of genes (Johns, 1964). Investigators first considered that proteins in chromatin were the major component of the genes, until the double helical structure of DNA with its immediate implications for self-duplication was discovered by Watson and Crick in 1953 (Watson and Crick, 1953), followed soon after by deciphering the genetic code (Crick, *et al.*, 1961). Nowadays we recognise DNA as containing genetic information, but how this information is controlled was not fully understood.

In eukaryotes, chromatin is found inside the nuclei of eukaryotic cells, whereas in prokaryotes, the chromatin is held within the nucleoid. The main function of chromatin structure is to pack very long DNA strands into small nuclei. For example, the human genome of 3×10^9 base pairs (bp) would extend over a metre if unravelled, but actually this is compacted into a nucleus of less than 10^{-5} m in diameter.

In most major types of eukaryotic cells the main proteins involved in this DNA-protein complex are the histone proteins. B-form DNA surrounds the core histone to form a nucleosome core particle. Then, by adding linker histones and 20 to 60 base pairs of linker DNA (species dependent) whole nucleosomes are formed. The nucleosomes form chain-like beads on a string and they zig zag from one nucleosome to the next. The nucleosomes pack more tightly to form the 30 nm fibre. By adding some further scaffold proteins, chromosomes are formed. (Figure 1.1) (Wolffe, 1999)

Figure 1.1 The structure of chromatin

The major structures in DNA compaction: DNA, the nucleosome, the 10 nm "beads-on-a-string" fibre, the 30nm fibre and the metaphase chromosome (http://en.wikipedia.org/wiki/Chromatin).

1.2 Nucleosome

In the 1970s, based on the development of a series of technologies including nuclease digestion (Hewish and Burgoyne, 1973), protein-protein cross-linking (Kornherg and Thomas, 1974), electron microscopy (Olins and Olins, 1974), neutron scattering (Baldwin et al., 1975; Suau *et al.*, 1979) and sedimentation analysis, it was discovered that chromatin is built up with a repetitive fundamental nucleoprotein complex which was then named the nucleosome (Hewish and Burgoyne, 1973; Olins and Olins, 1974; Kornherg and Thomas, 1974; Baldwin, *et al.*, 1975; Suau, *et al.*, 1979). Structure studies of nucleosomes were carried out continuously to the present time (Baldwin *et al.*, 1973; Pardon *et al.*, 1972; Richmond *et al.*, 1984; Luger *et al.*, 1997; Moudrianakis *et al.*, 1985). Researchers now have realised that the nucleosome is the basic structural unit of chromatin (Luger *et al.*, 1997). In most types of eukaryotic cells, 146 base pairs of DNA are wrapped around a core-histone octamer plus a single molecule of linker histone to form a nucleosome. Nucleosomes play three major roles within the cell nucleus. First, they provide the lowest level of compaction which is required to fit double stranded B-form DNA into the cell nucleus. Secondly, they are important in the regulation of transcription by preventing RNA polymerase from unnecessarily accessing the promoter regions of genes which are not needed by the cell. Thirdly, remodelling of nucleosomes is necessary for all the processes of the cell cycle. When gene expression is required, special enzymes known as remodelling factors can remodel the nucleosomes in the region of a gene to allow access to the DNA (Kornberg, 1974; Allis *et al.*, 2007; Turner *et al.*, 1991; Kouzarides, 2007).

The first break through in development of nucleosome structure was the nuclease digestion experiment. DNA in chromatin was digested into a regular size of fragments length 180-200 bp (Williamson, 1970; Hewish and Burgoyne, 1973). By analytical ultracentrifuge technology, these fragments were separated and measured to have a mass about 20,000 Da which contain about 11,000 Da of proteins. Then electron microscopic analysis was introduced and the micrographs showed a 'beads on a string' structure of these repeating units. Each repeating particle was approximately 10 nm in diameter (Thoma, Koller and Klug, 1979). From all these observations, Kornberg suggested a nucleosome model that each particle consisted

of DNA and histones. DNA is wrapped around a core histone octamer consisting of a histone $(H3\text{-}H4)_2$ tetramer and two histone H2A-H2B dimers (Kornberg, 1974; Kornberg and Thomas, 1974). Micrococcal nuclease cleavage experiments indicated that 146 bp of DNA interacted very strongly with core histones, while the DNA in between nucleosome core particles (which was lost in separating nucleosomes during nuclease cleavage) is called linker DNA (Figure 1.2). The scaling of the components of the nucleosome, the DNA being on the outside of the histone core was proved by contrast variation in neutron scattering from nucleosome core particles (Suau *et al.*, 1979).

Later, by using native-gel electrophoresis technology, Simpson isolated a protein-DNA complex from chromatin called chromatosome which contains a histone octamer, a linker histone H1 and about 160 bp of DNA (Simpson, 1978). Studies of these chromatosomes and comparing them with H1-depleted particles showed that linker histone H1 not only interacts with linker DNA but also plays very important roles in stabilising DNA with core histones (Simpson, 1978). Researchers also found that linker histone H1 may influence the formation of the chromatin fibre (Figure 1.1) and the interaction of *trans*-acting factors with DNA, which are factors from outside the structure as opposed to *cis*-acting factors from within the structure.

The structure with DNA wrapped around the outside of the core histone octamer was determined first by neutron-scattering and shortly after by DNase I digestion experiments. Investigators found that the DNA with core histones was nicked by DNase I once per 10-11 bp for the entire 146 bp in nucleosomes (Lutter, 1978). This indirectly revealed that the whole 146 bp of DNA must lie on the surface of core histones. Evidence was given by neutron and X-ray crystallography of the nucleosome core particles. The nucleosome core particles in aqueous solutions, DNA has a larger radius of gyration than the protein component which indicated DNA was wrapped around the histones. Further structure studies were carried out by Klug and colleagues who analysed the crystals to 7 Å resolution and discovered the disc-like shape of nucleosome. The disc was 11 nm in diameter and 5.6 nm in height, DNA was wrapped in 1.75 turns of a left handed double helix around the histone core (Figure 1.3) (Wolffe, 1999). It was originally thought that the bending of the DNA around the histone core was not uniform because of the difference in DNA sequence,

but later high-resolution X-ray studies of the nucleosome core particle and histone octamer in KCl/phosphate (Luger *et al.*, 1997; Chantalat *et al.*, 2003) showed it to be an exuberant structure that (except in some cases) was not dependent on DNA sequence. There were very sharp bends approximately one and four helical turns to either side away from the centre of the nucleosomal DNA.

Figure 1.2 The organization of nucleosome core particle and the position of linker DNA

The structure of nucleosome core particle, DNA wrapped around core histone and located as a 'zig zag' shape in chromatin (Thomas, Koller and Klug, 1979). Linker histone is shown as the black molecule stabilising the nucleosome core particle and interacting with linker DNA in between two nucleosomes (Wolffe, 1999).

Figure 1.3 The organization of DNA in the nucleosome

In this diagram A shows the 'disc' shape of a nucleosome core particle, where the hatched cylinder is the core histone octamer and the open tube is DNA. B shows one turn of DNA wrapped around the histone octamer, the number refers to turns of the DNA helix away from the dyad axis. C shows that the pathway for transcription factors (big circles) to access the DNA (small circles) is restricted by core histones (hatched square) (Wolffe, 1999).

1.3 Histone proteins

Histone proteins (Johns, 1964) are the major proteins involved in the DNA-protein complex of chromatin. Core histones pack DNA into nucleosomes; linker histone and linker DNA connect nucleosomes together to form a further tightly-packed structure in order to form chromatin as described. Apart from associating with DNA, histone proteins also undertake protein-protein interactions between themselves and other proteins. These interactions contribute to maintain their high degree of evolutionary conservation and also allow them to carry out the function of regulating gene expression.

Although histone proteins are highly conserved in different species, not all eukaryotic cells pack their DNA with histone. For example, dinoflagellates are reported to package the majority of their DNA with small basic proteins completely unlike histones (Vernet *et al.*, 1990); and in spermatozoa of mammalian species DNA is packed with basic molecules known as protamines.

There are five major histone classes which are known as histone H1, H2A, H2B, H3 and H4. The core histone octamer is formed by a $(H3\text{-}H4)_2$ tetramer and a pair of (H2A-H2B) dimers. DNA wraps around the core histone octamer in a left-handed superhelix to form a nucleosome (Figure 1.4). The major interactions between core histones and DNA are electrostatic forces binding positively charged lysines and arginines on the histones to negatively charged phosphates on DNA. Core histones can be separated from DNA by high salt concentrations. By increasing the salt concentration, histones H2A and H2B are removed from the nucleosome structure first and then followed by H3 and H4. In the absence of DNA, histones H2A and H2B form stable dimers; two pairs of histones H3 and H4 dimers form a stable tetramer (Kornberg, 1974; Kornberg and Thomas, 1974). The core histone proteins are small basic proteins containing large amounts of lysine and arginine (more than 20% of the total amino acids). Histones H2A (13,960 Da) and H2B (13,774 Da) contain more lysine, and histones H3 (15,273 Da) and H4 (11,236 Da) contain more arginine. Histones are very highly conserved proteins, in core histone octamers, H3 and H4 are the most conserved based on their important roles in the structure and function of nucleosomes. Histones H2A and H2B are slightly less conserved compared with H3 and H4 but still highly conserved compared with many other proteins. The four

core histones all consist of a C-terminal histone-fold domain containing three α-helices and an N-terminal charged tail rich in lysine residues. These N-terminal charged tails are involved with many histone post-translational modifications (see section 1.4). The reason for the C-terminal histone-fold domain to be highly conserved might be their central structural roles in a nucleosome; for the N-terminal tails it might be because of their roles in post-translational modifications and interactions with other proteins that organize higher-order chromatin structures.

The structure of the central histone-fold domain of four histones is very similar. Each of them contain three α-helices; the central long α-helix is connected to two short α-helices by loop segments. These loop segments have some β-strand structure. The four histones form heterodimers with each other (H2A-H2B, H3-H4) by a 'hand-shake' model where their central α-helix acts as the interface (Figure 1.4). These heterodimers then interact with each other to form the octamer. Two H3-H4 dimers bind together to form a tetramer and each of them bind to a H2A-H2B dimer to finally give the octamer. The interface between two H3-H4 dimers is less extensive compared with the interface between the H3-H4 dimer and the H2A-H2B dimer, but the latter is more accessible to solvent and is consequently less stable (Eickbusch and Moudrianakis, 1978; Karantza *et al.*, 1996) (Figure 1.4). Also this is confirmed by higher-resolution crystallography of nucleosome core particles (1.9Å structure, Davey, 2002) and histone octamers (Chantalat *et al.*, 2003; Wood *et al.*, 2005)

Each nucleosome is attatched to a linker histone and between two nucleosomes there is a region of DNA called linker DNA to which the linker histone binds (Figure 1.5) (Turner, 2001). Compared with highly conversed core histones, linker histones are much more variable. The most common linker histone is histone H1, in some species linker histone variants exist such as histone H5 in chick erythrocyte cells. Linker histones are also highly basic proteins, rich in lysine residues and slightly larger than core histones (>20,000 Da) (Wolffe, 1999). Linker histones contain a central domain which consists of three α-helices attached to a three stranded β-sheet and highly basic charged tails rich in lysine residues at both the N-terminal and C-terminal regions (Figure 1.6). The central domain plays important roles in stabilising the nucleosome structure and the tails interact with the linker DNA between nucleosomes.

Figure 1.4 The structure of core histone octamer

The structure and formation of core histone octamer: histone H3 and H4 form dimers, H2A and H2B form another dimer. Then two (H3-H4) dimer form a tetramer, a tetramer binds to two (H2A-H2B) dimers to form a histone octamer (Wheeler, 2005).

Figure 1.5 The structure and function of linker histones

One model for poly-nucleosome chain is where linker histones bind to the entry/exit sites of nucleosomes to pack nucleosomes more tightly. They also interact with linker DNA to package nucleosomes into 30 nm fibre,the structure of which, including the location of the linker histone in relation to the nucleosome-core-particle, is still not determined in detail(http://www.cbs.dtu.dk/staff/dave/roanoke/genetics980218.html).

Figure 1.6 The structure of linker histone H5

The structure of linker histone H5, the core particle domain consists of three α-helices and β-sheets. Two highly basic charged tails are found at both N- and C-terminal regions (Wolffe, 1999).

1.4 Nucleosome remodelling

Because the DNA double strands are very tightly packed in the eukaryotic cell nuclei by histone proteins, other proteins can not readily interact with DNA under these conditions. So when the DNA replication or gene expression needs to take place, the DNA needs to be released from the histones first. This can be done by nucleosome remodelling. There are three ways in which nucleosomes are remodelled to allow access of proteins to access specific regions of the DNA sequence (the genome) at specific timings within the cell cycle (Figure 1.7).

Firstly, the remodelling can be ATP dependent structure changes to the chromatin fibre made by groups of enzymes (eg. the SWI/SNFcomplex). This way involves histone octamer movements as a whole or movement or depletions of histone dimers and tetramers. Our group developed the interactions between two histone octamers since we obtained high resolution histone octamer crystals (Wood, *et al.*, 2005). There is a very strong interaction between the acid rich patch of one octamer with the basic lysines and arginines of the next.

Secondly, histones (probably as whole tetramers or whole dimers), with slightly different sequences from the majority in a species and cell type, can be inserted in place of the normal proteins. These are called histone variants and they are coded by their own separate genes. The variants are correlated with specific function such as histone H2A.X (replacing histone H2A).

Thirdly, the major types of histone post-translational modifications (PTMs) such as acetylation, deacetylation, methylation, demethylation, phosphorylation, dephosphorylation, ubiquitylation and several more modifications are correlated with specific functions (Kouzarides, 2007; Reynolds, 2008). There are more than 60 different modifications of histone residues which were identified by either specific antibodies or by mass spectrometry and this number has a huge potential to be increased (Kouzarides, 2007) (Table 1.1 and 1.2).

Experiments indicated that the level of these post-translational modifications change with different phases during the cell cycle (Figure 1.7). During S phase of the

cell cycle, all the core histones are acetylated but the key modifications of histones H3 and H4 are mono- and diacetylation. In G2 phase, histones H3 and H4 become hyperacetylated and in mitosis all four core histones are deacetylated. Histone H2A is phosphorylated throughout the cell cycle. Histone H3 is phosphorylated during mitosis. Histone H1 phosphorylation occurs in S phase, increases in G2 phase and becomes maximal at metaphase. Histone ubiquitylation and ADP-ribosylation occurs in S phase and become maximal in G2 phase (Wolffe, 1999).

Figure 1.7 The relationship between cell cycle and histone modifications

This diagram shows the modifications of histones occurring associated with different stages in the cell cycle (Wolffe, 1999).

Table 1.1 Different Classes of Modifications Identified on Histones

This table shows the classes of different modifications on histones and their related functions (Kouzarides, 2007).

Keys: *ac = acetyl group*

me = methyl group

ph = phosphroyl group

ub = ubiquitin

su = SUMO

ar = ADP ribose

Table 1.2 Histone modification enzymes

This table shows more than 60 modifications on histones which have been identified and the enzymes that catalyse these modifications in humans and yeast are shown (Kouzarides, 2007).

1.4.1 Histone variants

There are many histone variants which are found in different species or in different stages during the cell cycle in the same species. These histone variants are coded by separate genes, to replace the normal histone proteins in certain regions of chromatin to carry out specific functions.

One of the first examples of histone variant studies is in the sea urchin. The sea urchin contains seven different H2A variants expressed separately for the cleavage, blastula and gastrula stages of embryogenesis; it also contains four different H2B variants which are developmentally regulated as well. Sea urchin sperm has an H2B variant containing an N-terminal tail of 21 amino acids extension, which may interact with linker DNA to provide special stability for sperm nucleosomes (Bavykin *et al.*, 1990; Hill and Thomas, 1990). Studies also show that histones H3 and H4 in sea urchin do not have any variants during development which indicates the central role of these histones in nucleosome structure and chromatin assembly. Variations of histones H2A and H2B can change the compaction of DNA in nucleosomes and chromatin fibres by directly influencing the nucleosome structure or through changing the binding of histone H1 to the nucleosome core particle.

A highly conserved histone H2A variant known as H2A.Z in mammals (also known as H2A.F/Z in chickens, hv1 in *Tetrahymena* and H2A.vD in *Drosophila*) is found to associate with actively transcribed chromatin. It has an N-terminal tail which is post-translationally modified through acetylation to a greater extent than normal H2A. Another conserved histone H2A variant called H2A.X contains an extended C-terminal tail beyond the histone-fold domain. Wheat H2A1 is another example of an H2A variant that has a C-terminal tail with 19 amino acids. Although these histone variants have a wide range of different sequences, they form a family of proteins that share the histone-fold domain and might be incorporated into a nucleosomal structure.

The histone-fold domain consists of three α-helices, all the core histone proteins and their variants are developed from a basic DNA-binding protein

containing only a histone-fold domain and lack of any tails which can interact with DNA to form a nucleosome-like structure. The tails of different histone variants provide them with specific functions. The tails of core histones in eukaryotic nuclei give them the ability to interact with other proteins, DNA or complexes. Other regulatory proteins with histone-fold domains can replace normal histones in the nucleosome to change the nucleosome structure, in order to carry out certain functions.

Some regulatory proteins with histone-fold domains such as CENP-A (found at the centromere) and rat macro H2A use their histone-fold domains to confer specialised properties on individual nucleosomes through replacing normal histones within chromatin. CENP-A is similar to histone H3 but with a very different N-terminal tail (Figure 1.8). It is found to form heterodimers with histone H4 in nucleosomes. Some other transcriptional regulatory proteins use their histone-fold domains to bind the DNA and direct specific protein-protein interactions. This group of histone variants including the TATA-binding protein associated factors $TAF_{II}40$ and $TAF_{II}60$, and CCAAT-box-binding proteins CBF-A, -B and -C and HAP2, 3 and 5. $TAF_{II}40$ and $TAF_{II}60$ are related to histones H3 and H4. They replace normal histones H3 and H4 in the TATA box and maintain the DNA in a semi-compacted state for transcription (Kokubo *et al.*, 1993; Nakatani *et al.*, 1996; Xie *et al.*, 1996). CBF-A, B and C and HAP2, 3 and 5 are evolutionarily related proteins in different species. They are transcriptional activators where CBF-A and C, HAP3 and 5 are similar to histones H2B and H2A. Together with the third component (CBF-B or HAP2) they can bind to the DNA.

Figure 1.8 Histones and their variants containing the histone fold domain

This Figure shows the core histones and their variants which contain histone fold domains but with different tail domains.A is showing several histone variants compared with original histones. B is showing how histone fold domains bind to each other in histones H2A-H2B and H3-H4 dimers (Wolffe, 1999).

1.4.2 Histone acetylation and deacetylation

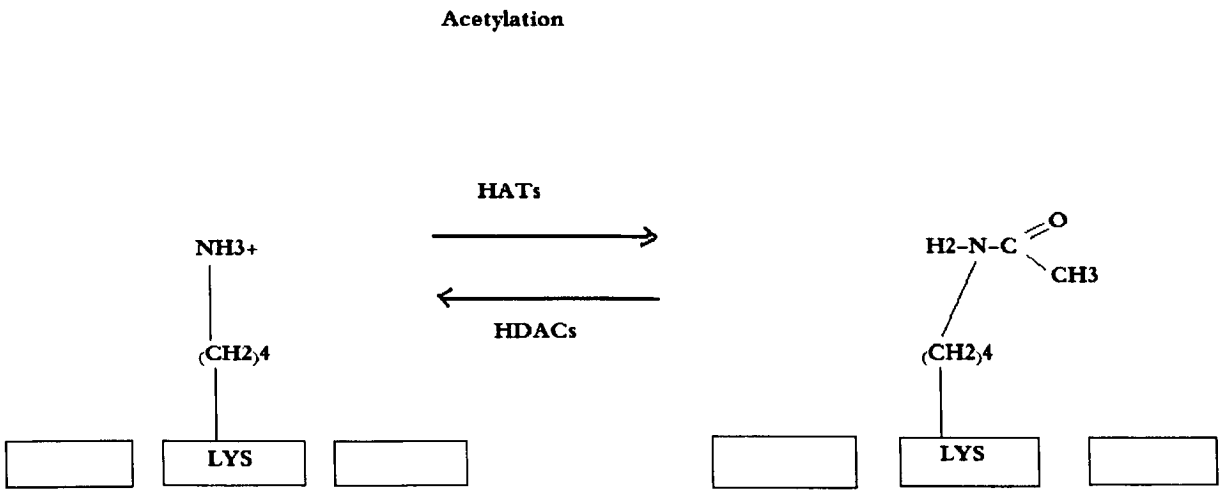
The acetylation and deacetylation of histones occurs at specific lysine amino-acid residues at the N-terminal tails of the histone molecules which are covalently bonded to an acetyl group ($COCH_3$). Histone acetylation of the four core histones occurs in all animal and plant species examined (Csordas, 1990). These modifications lead to a loosening of DNA from the histone proteins to activatetranscription. Histone-acetyltransferase and Histonedeacetylase are the enzymes which catalyse these reactions (Berg *et al.*, 2002) (Figure 1.9).

There is a balance of the activities of histone acetyltransferase (HAT) and histone deacetylase (HDAC) which together give the final histone acetylation level. For example histone H4 with four acetylated lysine residues would be considered as hyperacetylated, whereas the protein containing one or no acetylated lysine residues is described as hypoacetylated. The acetylated lysines in the histone tails lead to a more relaxed chromatin state which gives a higher activity of gene expression. On the other hand, the presence of deacetylated lysine residues in histone tails gives a tight model of chromatin and is associated with gene silencing. Histone deacetylation can increase the ionic interaction between positively charged histone and negatively charged DNA which decreases the accessibility of the transcription machinery to the DNA.

The hyperacetylated state of core histones not only increases the accessibility of proteins to DNA, but is also significant for protein-protein interactions. Experiments show that the N-terminal tails of the core histones are accessible to trypsin, which indicates that they are exposed on the outside of the nucleosomal core particle. Further studies of cross-linking showed that the N-terminal tail of H2B can have weak interactions with linker DNA. Protein NMR experiments also indicate that the N-terminal tails of histones H3 and H4 have interactions with the DNA in nucleosome core particles (Wolffe, 1999; Bradbury *et al.*, 1974). Researchers proved that acetylation or removal of the histone tails can increase the accessibility of *trans*-acting factors to their recognition sites in nucleosomes (Lee *et al.*, 1993; Vettese-Dadey *et al.*, 1996). Histone acetylation might have a major regulatory role in transcription and this was proved by Hubbes *et al.* (1988). They used antibodies to

acetylated lysines in nucleosomes with immune-precipitation to separate the highly acetylated nucleosomes. These were then analysed by removing the histones and analysing the DNA sequences. These were shown to be from regions of the genome active in transcription in chick erythrocytes (globin genes). This was one of the early chromatin immuno-precipitation (ChIP) experiments. The chromatin immune-precipitation together with gene array technology (ChIP on CHIP) was applied to analyse acetylation and lysine methylation of histones. The results show that these modifications are not uniformly distributed in chromatin, acetylation is found to be enriched at specific sites in the promoter and 5' end of the coding regions (Kouzarides, 2007).

Figure 1.9 Histone acetylation and deacetylation



This Figure shows the acetylation and deacetylation of the lysine residues. An acetyl group is added to or removed from the lysine residue. These reactions are catalysed by enzymes called HATs (histoneacetyltransferase) and HDACs (histone deacetylase) (Figure produced by the Author).

1.4.3 Histone phosphorylation

Histone phosphorylation occurs at serine and theonine residues on the N-terminal tails of histone molecules and is catalysed by a family of kinases (Figure 1.10). It has mostly the same effect as acetylation. Phosphorylation of serine at a certain position can enhance the activity of histone acetyl transferases (called 'cross talk').

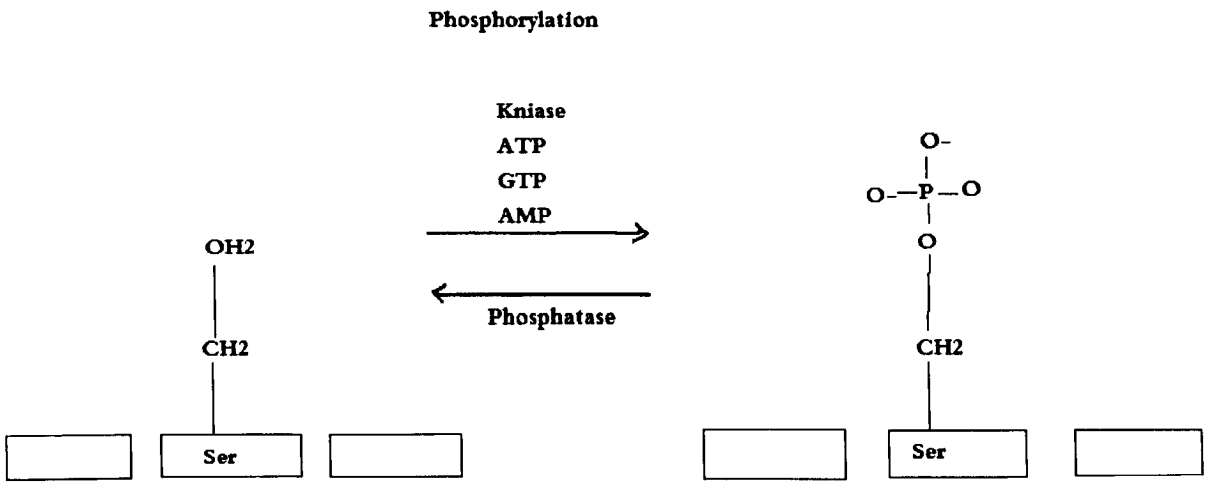
The serine residues in the histone H3 basic amino terminal domains are rapidly phosphorylated when some extracellular signals such as growth factors or phorbol esters stimulate quiescent cells to proliferate (Mahadevan *et al.*, 1991). This phosphorylation may lead to conformation changes of nucleosomes or even the higher-order structure of chromatin, and the related transcriptional activity would be increased. Experiments which introduce phorbol esters to proto-oncogenes *c-fos* and *c-jun* show that their transcriptional activity level was rapidly increased and sulphydryl groups in histone H3 cysteine residues were exposed. Because cysteine residues of histone H3 are normally buried inside of nucleosomes, exposure of sulphydryl groups of these cysteine residues may indicate that a major disruption of nucleosome structure such as dissociation of histones H2A-H2B dimers has occurred. These changes might be associated with histone H3 phosphorylation and acetylation.

Phosphorylation also occurson histones H4 and H2A in the cytoplasm shortly after histone synthesis (Sung and Dixon, 1970; Ruiz-Carillo *et al.*, 1975; Jackson *et al.*, 1976; Dimitrov *et al.*, 1994). Studies show that this phosphorylation together with deacetylation of histone H4 may target them to molecular chaperones involved in DNA replication and chromatin assembly. Another histone H2A variant, H2A.X, also can be phosphorylated. Experiments show that this phosphorylation plays a role in nucleosome spacing during chromatin assembly. However, nucleosome spacing would not be affected when H2A.X phosphorylation is removed by phosphatases (Dimitrov *et al.*, 1994)

Linker-histone phosphorylation is one of the widely studied post-translational modifications of histones. Phosphorylation of histone H1 tails occurs at a highly-

conserved motif: S/T-P-X-K-R (serine/threonine - proline - any amino acid – lysine-arginine). This motif exists several times along the basic charged tails on each side of the central globular domain of histones H1and H5 (Snijders *et al.*, 2008). Studies of mammalian cell culture discovered that rapidly dividing cells have the highest level of histone H1 phosphorylation and the level decreases in non-proliferating cells (Wolffe, 1999). During the cell cycle, histone H1 phosphorylation level is lowest in G1 phase; it increases in S phase and mitosis and reaches the maximum at metaphase. This indicates that histone H1 phosphorylation level is increasing when the chromosome is getting more condensed, and suggests that there are some relationships between histone H1 phosphorylation and chromosome compaction (Bradbury *et al.*, 1974). Phosphorylation of histone H1 would weaken the interaction between the histone H1 and linker DNA, because of the basic charged tails of histone H1 are neutralized. But surprisingly, this modification influences the binding of the protein to chromatin more than to DNA (Hill *et al.*, 1991). This may suggest that phosphorylation of histone H1 could weaken the interaction between linker histone and chromatin therefore allowing other trans-acting factors to bind to DNA or the fibre in order to change the chromosomal architecture.

Figure 1.10 Histone phosphorylation



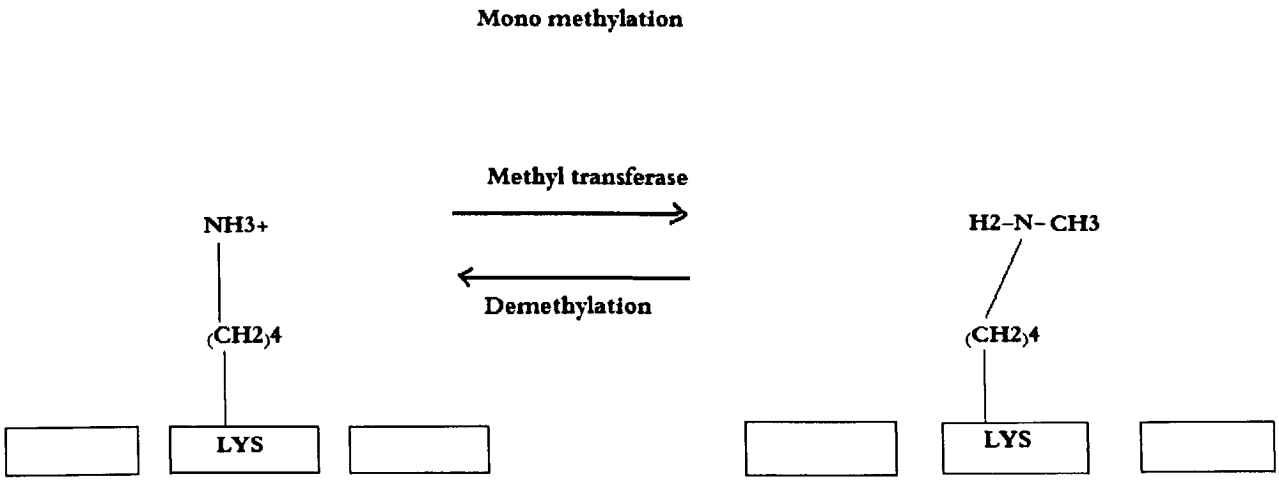
This Figure shows the phosphorylation of a serine residue. A phosphate group is added to and removed from the serine residue by a family of kinases and phosphatases respectively (Figure produced by the Author).

1.4.4 Histone methylation

Histone methylation is another possible modification that occurs to control gene expression. This may happen by adding a methyl group to a specific lysine or arginine residue (Figure 1.11). This modification particularly appears in histone H3 and H4 and it is catalysed by histone methyltransferases. It has recently been found that demethylases are also present to maintain the balance between methylation and demethylation (Vargason, 2000).

The function of histone methylation is not clear yet. Most methylation occurs on histone H3 at lysines 9 and 27 and histone H4 at lysine 20. It was found that up to three methyl groups can be added to each lysine residues of histone H3; meanwhile histone H4 contains maximum two methyl groups. Histone methylation of lysine residues occurs after nucleosome assembly. It seems to occur preferentially to the histone H3 and H4 which are already acetylated (Kouzarides, 2007).

Figure 1.11 Histone methylation

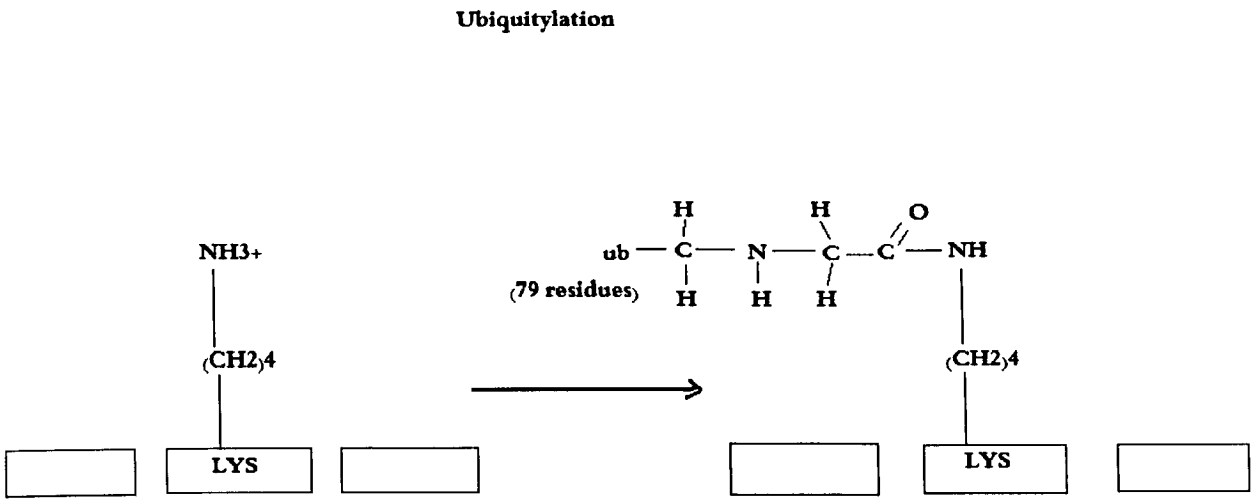


This Figure shows the mono methylation of a lysine residue. A methyl group is added to the lysine residue by methyl transferase. Methylation can be mono-, di- or tri-. The lysines in histone H3 are found can be added with up to three methyl groups, but in histone H4 they contain maximally two methyl groups (Figure produced by the Author).

1.4.5 Histone ubiquitylation

Ubiquitin is a highly conserved small regulatory protein with 76 amino acids that is ubiquitous in eukaryotes. Ubiquitylation refers to the post-translational modification of a protein by the covalent attachment of one or more ubiquitin monomers. The most prominent function of ubiquitin is labelling proteins for proteasomal degradation. So the function of ubiquitylation is to control the stability, function, and intercellular localization of a wide variety of proteins. For the histone proteins, ubiquitin may be attached to lysine residues of the C-terminal tail of histone H2A and H2B (Figure 1.12). This modification might be linked to the repression or activation of gene expression. For example, ubiquitylation of H2B stimulates gene expression since it promotes the methylation of H3 on Lys 4 and Lys 79. Ubiquitylation of histone H2A tail domain may disrupt the interaction of linker histones with nucleosomal DNA and may also disrupt higher-order chromatin structures by impeding internucleosomal interactions (Kouzarides, 2007).

Figure 1.12 Histone ubiquitylation



This Figure shows that an ubiquitin molecule with 79 residues is added to a lysine residue. This is found to occur on the C-terminal tail of histones H2A and H2B (Figure produced by the Author).

1.4.6 Histone sumoylation

Sumoylation is a newly identified modification of lysine residues on histone tails. It is found to be related with transcriptional repression (Shiio and Eisenman, 2003). Small ubiquitin-related modifier (SUMO) is a polypeptide that has similar 3D structure to ubiquitin (18% of sequence similarity). It is covalently attached to other proteins through the activities of an enzyme cascade (E1-E2-E3) similar to that of ubiquitylation (Shiio and Eisenman, 2003). SUMO is reported to bind to many different proteins; in most of the cases it decreases the transcriptional activity (e.g. Elk1, Sp3, c-Myb, and c-Jun) (Verger, *et al.*) and represses the gene. Shiio and Eisenman's study shows that the sumoylation of histone H4 is correlated to transcriptional repression as well. They genetically fused SUMO to histone H4 and found that SUMO-H4 associates with chromatin and can be immuno-precipitated with endogenous histone deacetylase1 (HDAC1) and heterochromatic protein 1 (HP1). These data indicate that sumoylation of histone H4 is associated with gene repression.

According to Shiio and Eisenman's study, a model of the role of sumoylation of histone H4 is given (Nathan *et al.*, 2003). It suggests that once a gene is expressed it must then be attenuated and finally repressed. So acetylation itself might be a signal for recruiting sumoylation enzymes. Experiments show that H4 sumoylation increases with increasing H4 acetylation, and then the acetyl group is removed by HDAC and gene repression is caused by histone methylation, in order to build up chromatin condensation. In this process, sumoylation may actually be the signal that initiates a sequence of attenuation followed by repression (Figure 1.13).

Figure 1.13 A model of the function of histone H4 sumoylation

This Figure shows a model of the function of histone H4 sumoylation. SUMO binds to histone H4 as a signal to recruit HDAC to deacetylate the protein and in order to attenuate and repress the gene (Nathan et al., 2003).

Keys: Act = activator

ac = acetyl group

HAT = histone acetyltransferase

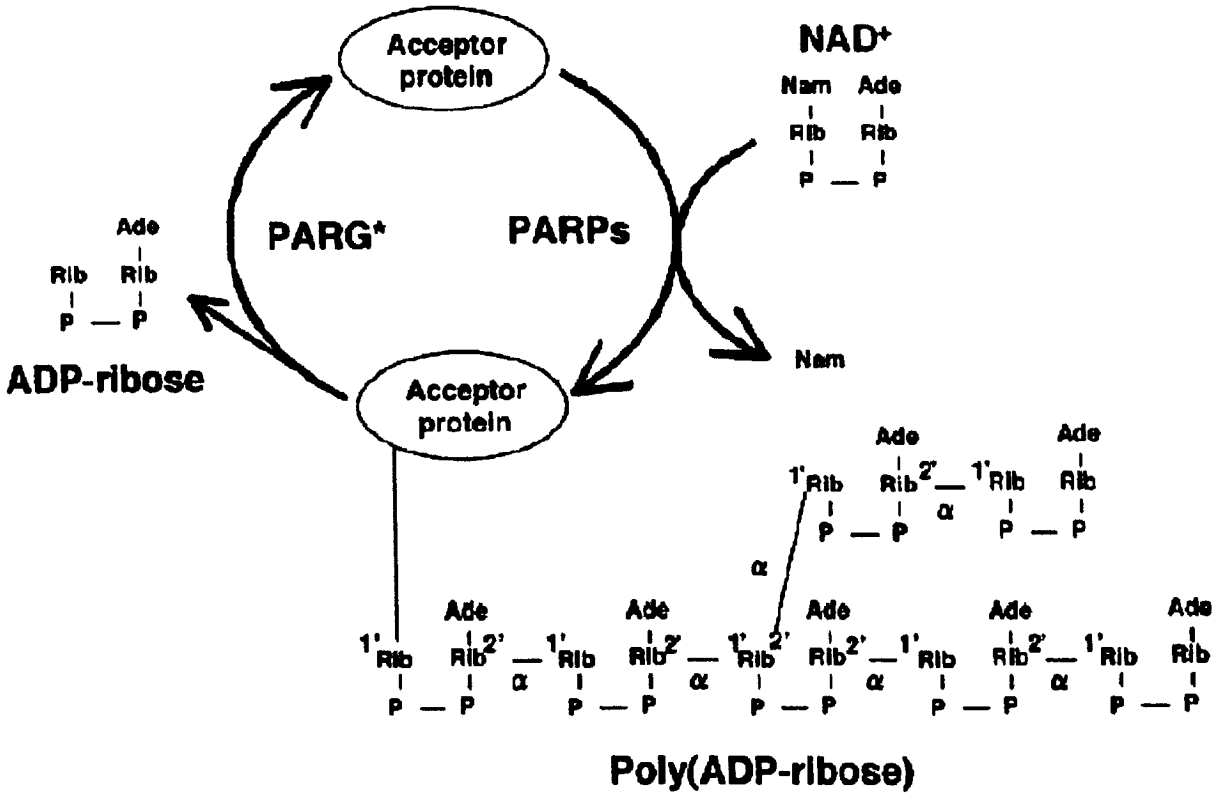
Rep = repressor

HDAC = histone deacetylase

1.4.7 Histone ADP ribosylation

ADP ribosylation is a covalent post-translational modification catalysed by ADP-ribosyltransferase (Figure 1.14). Poly(ADP-ribose) polymerase 1 (ARTD1/PARP1) is one of the key enzymes of ADP ribosylation and it is responsible for most of the cellular poly(ADP-ribose) formation. ADP ribosylation has been found to be related to histone proteins *in vivo* for many years (Ueda, 1975). Histones isolated from rat liver nuclei and HeLa cells treated with radioactive NAD^+ indicated that all core histones H2A, H2B, H3, H4 and linker histone H1 are ADP ribosylated to a variable extent (Otake *et al.*, 1969; Adamietz *et al.*, 1978; Minaga *et al.*, 1979 and Ogata *et al.*, 1980). Association of long negatively charged chains of ADP-ribose with histones may contribute to the disruption of nucleosomes by exchange of histones to this competitor polyanion. ADP-ribosylation may lead to unfolding of the chromatin fibre and may play an important role in DNA repair (Wolffe, 1999). It is also involved in cell cycle regulation, DNA replication and transcription. ADP-ribosylation can be read by the zinc finger motif, in order to regulate chromatin structure and transcription (Hottiger, 2011). Recent studies suggest that lysine residues of all the core histones are the acceptor sites of ADP-ribosylation instead of previously assumed glutamic acid residues (Messner *et al.*, 2010).

Figure 1.14 ADP-ribosylation



This diagram shows that how the poly(ADP-ribose) formed and how it is modifying the target proteins (Miwa, et al., 2006).

Keys: PARG = poly(ADP-ribose) glycohydrolase

PARP = poly(ADP-ribose) polymerase

1.4.8 Peptidylprolyl isomerisation

In drug development studies, several organic molecules (with molecular weights far less than those of nearly all proteins) were found empirically to suppress immune reactions and therefore to be effective in stopping rejection of human transplants. Three immuno-suppressive small drug molecules, FK506, rapamycin and cyclosporin are shown in Figure 1.15. These drugs were found to bind to specific bio-molecules that inhibit the mechanisms of T-cell and/or B-cell control. These bio-molecules belong to a class of enzymes called peptidylprolyl isomerases (PPIs), at least some of which (and possibly nearly all) control the immunological processes of lymphocytes (in T-cell and B-cell lymphocytes in particular).

The PPIs that bind to FK506 or rapamycin are called FK506-binding proteins abbreviated to FKBPX, where X is the approximate molecular weight of the protein. Those binding to cyclosporin are called cyclophilins, mostly A (cypA) or B (cypB). The FKBPX family mainly includes FKBP 13 (12a), 25, 38, 52 and 51. The FKBP proteins, in nearly all known cases, consist of the PPI/drug-binding region plus other regions specific for a particular FKBP protein. FKBP 12 is called the canonical FKBP protein because it contains only the PPI/drug-binding region (Figure 1.16).

These PPIs change the 3D structure of proteins by targeting very specific proline aminoacid residues. This changes the proline from the usual *trans* conformation into its isomeric *cis* conformation (Figure 1.17).

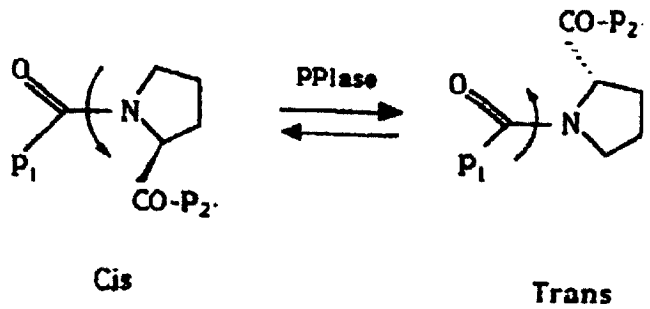
Figure 1.15 Three immune-suppressive molecules, cyclosporin, rapamycin and FK506

This Figure shows the structure of three immune-suppressive small drug molecules which A is cyclosporin A, B is rapamycin and C is FK506 (Galat, 1993).

Figure 1.16 Topography for different FKBPs

This Figure shows the topography of different FKBPs. The pink region is the canonical region for drug binding (Leclercq et al., 2000).

Figure 1.17 The*trans* and *cis* forms of proline



This Figure shows the trans and cis forms of proline. Naturally there are about 6% of prolins in cis form and others are in trans form (Galat, 1993).

Keys: P_1 = *peptide 1*

P_2 = *peptide 2*

All the FKBP proteins contain a region where a five-stranded β sheet has a right-handed twist around the only α helix in that part of an FKBP. The structure forms a hydrophobic pocket where the drug molecule binds (Figure 1.18). This region of the proteins contains the PPI activity and the drug binding site(s) (FK506, rapamycin etc).

Peptidylprolyl isomerisation is a PTM for the target protein at a particular proline and it changes the conformation of the protein because of the change in trajectory of ensuing amino-acid residues. Therefore it must cause at least local changes of protein three-dimensional conformation that results in binding other factors and causing function changes.

Our group is familiar with PTMs of histone proteins and there is one protein FPR4 that has some homology with FKBPs. This protein has been found by the Kouzaredes's Group to isomerise histone H3 proline 38 that seems to be necessary for the recognition and methylation of H3 Lysine 36 via the SET (Su(var) 3-9, Enhancer of Zeste, Trithorax domain) methyl transferase (Nelson, *et al.*, 2006). Otherwise, until present, the detailed pinpointing of prolines in the targets for the FKBPs is not established. Also the mechanisms of the immunophilin function or detailed structures of the FKBP proteins complexed with their target proteins remain unknown. Even such structures of FPR4 with portions of histone H3 or the methyl transferase seem not to have been done and recorded in the RCSB PDB (Research Collaboratory for Structural Bioinformatics, Protein-structure Data Bank).

Figure 1.18 Structure of the PPI-β-sheet region

This Figure shows the structure of the PPI-β-sheet region of all FKBP family proteins (a and b, canonical region which is really the complete FKBP12). The FK506 drug molecule is shown bound to the FKBP12 through the yellow region in c and d (Kang et al., 2008).

1.5 HMG proteins

High Mobility Group (HMG) proteins are a group of chromosomal proteins that help with transcription, replication, recombination and DNA repair. HMG proteins are subdivided into three superfamilies which are HMGB (HMG-box) family, HMGN (nucleosomal binding domain) family and HMGA (A-T hook motif for binding to DNA) family.

HMGs were first fractionated with linker histone H1 by perchloric acid extraction. The proteins soluble in perchloric acid except linker histone H1 were then separated by trichloroacetic acid (2%) into an insoluble group of large proteins named low mobility group (LMG, determined by gel electrophoresis) proteins and a soluble group of small size proteins named high mobility group (HMG) proteins. HMG proteins were then subdivided into three classes which were originally called HMG1/2, HMG14/17 and HMGI/Y by their size. These classes were then found to contain different motifs and renamed as HMGB (HMG1/2, containing HMG-box), HMGN (HMG14/17, containing nucleosomal binding domain) and HMGA (HMGI/Y, containing A-T hook motif).

HMGB

The HMGB proteins are the major family of HMGs. They play important roles in a range of chromatin-related processes such as transcription, replication, and they are also involved in determining the cell phenotype (Bustin, 1999; Thomas and Travers, 2001; Agresti and Bianchi, 2003; Bianchi and Agresti, 2005; Hock *et al.*, 2007). They also have functions as extracellular signalling molecules (Lotze and Tracey, 2005; Bianchi and Manfredi, 2007). Three major variants are involved in this family which are HMGB1, HMGB2 and HMGB3.

HMGB proteins contain a basic N-terminal tail and an acidic C-terminal tail. The N-terminal tail contains two HMG-boxes, called HMG box A and HMG box B, which are protein domains consisting of three α-helices and creating two independent DNA-binding surfaces (Read *et al.*, 1993; Weir *et al.*, 1993; Gerlitz, *et al.*, 2009) (Figure 1.19). This structure of multiple DNA-binding sites on the same protein is similar to linker histones. Experiments showed that HMGB proteins can replace

linker histones in chromatin (Nightingale *et al.*, 1996; Ura *et al.*, 1996). The HMG box also provides the binding site for chromatin.

HMGN

HMGNs are highly conserved proteins containing N-terminal basic regions. These basic regions are believed to interact with nucleosomal DNA. They were examined to have higher affinity for nucleosomal DNA than for naked DNA. They may contribute to the folding of chromatin and indirectly increase the accessibility of regulatory complexes to RNA polymerase. There are five major variants in this family which are HMGN1, HMGN2, HMGN3a, HMGN3b and HMGN4. Their sizes are all around 10 kDa (Bustin, 2001). The proteins in this family all contain a 30 amino acid nucleosomal binding domain (NDB). A highly conserved 8 amino acids motif is embedded in this domain which is **RRSARL**SA. The bolded RSRL amino acids in this sequence are essential for the specific binding of the HMGN protein to the nucleosome core particle (Gerlitz*et al.*, 2009).

HMGA

HMGA family proteins contain three conserved AT-hook DNA-binding domains and all have an acidic C-terminal tail. The AT-hook domain is a small motif containing the PRGRP sequence which is flanked by additional positively charged residues (Reeves and Nissen, 1990; Huth *et al.*, 1997). The central RGR sequence binds to the minor groove of AT base pairs where the arginine residues form electrostatic and hydrophobic contacts with the floor of the minor groove DNA (Figure 1.20). After binding to the DNA, conformational changes of HMGA proteins take place. This gives them the ability to bend the DNA (Reeves and Beckerbauer, 2001).

HMG proteins are thought to play an important role in many sorts of human disorders. Disruptions and rearrangements in the genes coding for some of the HMG proteins are related to some common benign tumors. Antibodies to HMG proteins are found in patients suffering from autoimmune disease (Iannaconeet *al.*, 2007).

Figure 1.19 A model for the binding of HMG box A and B to linker DNA and chromatin

This Figure shows how the HMGB proteins bind to the DNA through HMG boxes A and B (Gerlitz, et al., 2009).

Figure 1.20 A model showing how AT-hooks might bind to the minor groove in AT rich DNA

This Figure shows that how AT-hooks might bind to the minor groove in AT rich DNA (Gerlitz, et al., 2009).

1.6 Aim of the project

The need for nucleoproteomics and nucleoprotein purification

The Association for Clinical Biochemistry (ACB) was founded in 1953, and is one of the oldest such Associations in the world. Based in the United Kingdom, it is a professional body dedicated to the practice and promotion of clinical science. The Association has medical and non-medical members in all major UK healthcare laboratories, in many university departments and in several commercial companies. The links with its Corporate Members leads to a fruitful relationship with the clinical diagnostics industry. The Association liaises with and is consulted by many national and international organisations on issues relating to Clinical Biochemistry.

The ACB has circulated the following statement showing the importance of isolating and developing protein extraction procedures (including nucleoproteins), so that new protein biomarkers can be developed as a national priority:

"It is a government priority that NHS research should be directed towards more immediate patient benefit and, with the clinical demand for more and better markers of disease, Clinical Biochemistry should be ideally placed to play an important role in this, particularly in the translation of protein biomarkers into the clinic. However, doubts have been raised regarding the capacity and resources of hard pressed routine of the Clinic Biochemistry laboratories to do this, and there may be a lack of awareness of the ongoing work in clinical proteomics for the discovery of new protein biomarkers".

The above stresses the need for biomarkers for disease. There is also the need to understand the mechanisms of all processes of cell cycle control in terms of the functions of all the proteins in the cell nucleus. So there is a need for the following:

1. Targeted extraction of **all** nucleoproteins in normal and disease cells;
2. Special attention at targeted extraction of those nucleoproteins that are closely involved in the primary architecture of chromatin such as core histones, linker histones, HMG proteins, so called immunophilins, and molecules associated with nuclear scaffold, such as topoisomerases;

3. Targeted extraction of proteins with known changes or likely changes associated with cancer and other diseases.

Objectives of this Thesis

In this thesis, nucleoprotein purification is developed by a cheap integrated approach using KCl/phosphate solutions, in order to address the issues of the last section. KCl/phosphate is used for the following reasons:

1. The pH of the extraction solutions can be controlled by changing the ratio of K_2HPO_4/KH_2PO_4 over a wide range;
2. The ionic strength of the extracting solutions can be controlled by the concentration of KCl and phosphate to values as high as 2 M KCl, 2 M phosphate or higher in the case of KCl.

As a result of 1 and 2:

a. The solution conditions for extracting DNA binding molecules are rich in phosphate, simulating the native conditions;
b. Precipitation of nucleoproteins to separate the soluble from the insoluble is possible by varying the ionic strength, while being in control of a very stable pH;
c. Control of low ionic strength KCl/phosphate condition enables the extraction of nucleoproteins (that we term cePNE1 proteins and they are rich in high mobility group proteins and many others, see later) that diffuse through the nuclear membrane into the extraction solution maintaining the protein depleted nuclei intact. This was discovered in our laboratory when preparing chick erythrocyte nuclei. Use of detergent and nuclei washing solutions in dilute KCl-tris, pH 7.5 enabled pure white nuclei to be prepared (see later). However, replacing the tris with phosphate at the same pH and ionic strength resulted in reddish nuclei because of the passage of haemoglobin from the cytoplasm into the nuclei. This indicates that nuclear membranes are permeable to proteins in the presence of phosphate.
d. It is possible to control pH and ionic strength, after the extraction of cePNE1 proteins, to break (lyse) cell nuclei and extract sequentially DNA binding proteins, firstly linker histones preferentially and then, in higher ionic strength,

to strip core histones from the DNA into the extraction solution (as opposed to acid extraction, see later).

In the chapters 2, 3 and 4 of this thesis, the principles established above are used for an integrated series of nucleoprotein extraction procedures using also conventional chromatography methods. In each chapter, brief examples of procedures used by others are given. Many proteins have been purified in this way paying particular attention to histones, HMGs, peptidylprolyl isomerases (PPIs) and other nucleoproteins, for example, topoisomerases.

In the context of nucleoproteins as biomarkers for disease, the methods, having been developed for chick erythrocyte nuclei, were applied to human leukocyte nuclei, extracted from human buffy coats supplied by the National Blood Service (chapter 5). These experiments were carried out in national laboratories, formerly the CCLRC Daresbury laboratory and latterly the Diamond national X-ray source at the Harwell/Rutherford laboratory sites. These laboratories have stringent safety procedures in place and national ethics approval was also obtained.

1.7 Review: DNA replication and histone chaperones

1.7.1 DNA replication and chromatin assembly

DNA replication in part of the cell cycle is a basic process in all living organisms. The two strands of B-form DNA hold the same genetic information because both strands can serve as templates for the reproduction of the complementary strand. During DNA replication, the two strands are separated at specific sites in the genome called 'origins' of replication. RNA primers attached to single stranded DNA and the enzyme DNA polymerase extends the primers to form new strands of DNA by adding nucleotides complementaryto the template strand (Figure 1.21). One DNA polymerase produces the leading strand copy; another DNA polymerase binds to the lagging strand and makes discontinuous segments called Okazaki fragments. This happens because the DNA polymerase can only extend a DNA strand from the 5' end to 3' end, and the Okazaki fragments will be linked together by DNA ligase afterwards.

The first step of replication initiation involves the assembly of a pre-replication complex (pre-RC) (Sun and Kong, 2010). DNA replication is initiated by the binding of the replicator to special regions of DNA known as origins. In prokaryotes, for example *E.coli*, the replication origin named *oriC* is a DNA sequence of ~250 bp. The origin-recognition protein DnaA binds to several 9 bp highly-conserved sequences in *oriC*. Beside these 9 bp sequences there is a second conserved AT-rich sequence which is the site for opening the duplex for loading DnaB helicase and subsequent initiation of DNA synthesis. In eukaryotes, such as budding yeast *Saccharomyces cerevisiae*, the replication origin is a 100-150 bp sequence similar to the structure of *oriC*. These origins contain an essential 11 bp consensus sequence of 5'-(A/T)TTTA(T/C)(A/G)TTT(A/T)-3' called the 'A' element and a non-conserved B region that is also important for origin activity (Marahrens and Stillman, 1992). The binding site of the origin-recognition complex (ORC) is located at the A region and the A region, together with a part of B region beside it, is used for the assembly of pre-RC.

The pre-RC assembly occurs at the late M and early G_1 phase of the cell

cycle when the activity of cyclin-dependent kinase (CDK), which phosphorylates pre-RC components and inhibits their activity, is low (Nguyen *et al.*, 2001). In budding yeast *S. cerevisiae*, ORC binds to the DNA replication origin first and loads Cdc6 onto the DNA at the early G_1 phase. Following that, ORC and Cdc6 together recruit Cdt1-MCM (Mcm2p, Mcm3p, Mcm4p, Mcm5p, Mcm6p and Mcm7p) to origins to form a pre-RC (Diffley *et al.*, 1994; Speck *et al.*, 2005; Chen *et al.*, 2007) (Figure 1.22).

In the cell cycle, eukaryotic cells must replicate large amounts of genomic DNA distributed on all the different chromosomes. In doing that, replication needs to initiate throughout S phase at multiple origins along each chromosome (Diffley and Labib, 2002). Initiation from those origins must be at the same time so that no region of the genome is left unreplicated and no region is replicated more than once. DNA is guaranteed to be replicated only once in a cell cycle by a simple mechanism which dismantles the pre-RC and destroys or inhibits all its components after the DNA replication is initiated (Stillman, 2004). DNA replication must also be coordinated with chromosome segregation to ensure that each daughter cell receives a complete and unaltered complement of genetic information. Any mistakes made in either DNA replication or chromosome segregation can result in loss or duplication of the genetic information, and these events can play an important role in the production of cancer cells.

After the assembly and binding of the pre-RC to the origins, DNA replication starts at the S phase of the cell cycle. At this stage, the DNA helix has been opened and unwound to form "replication forks" (Figure 1.21). In prokaryotes, this is done by DnaA when it binds to the origin and initiates DNA replication. In eukaryotes, the DNA helix is unwound by an enzyme known as DNA helicase which is a component of the pre-RC, but the mechanisms and the exact timing of the melting of the origin DNA are not well understood. Some recent structure studies of Cdc6/ORC-double-stranded-DNA complexes revealed that the double-stranded DNA was bent but not melted. This suggested that the DNA melting may require additional components (such as MCM) (Gai *et al.*, 2010). The pre-RC is assembled on the origin during G1 phase but activated at the start of S phase. The opened DNA double strands are prevented from reannealing by a single strand binding protein known as replication

protein A (RPA) in eukaryotes. Then a group of proteins including the most important DNA polymerases are recruited to the replication forks, which then move outwards from the replication fork in both directions to synthesise new DNA strands by a semi-conserved mechanism (Morgan, 2007).

DNA polymerase can only add new deoxyribonucleotides to the 3' end of a previously existing strand that is base-paired with the template strand. So the new DNA strand is initiated with a short polynucleotide primer, synthesised by the polymerase called primase, which copies a short fragment of the DNA template into RNA. Then the primase is replaced by DNA polymerase which completes the new DNA strand. In eukaryotes, the primase and DNA polymerase activities are both carried out by a protein complex called Pol α-primase. Also the newly synthesised nucleotides can only be attached to the 3' end of the new DNA strand. That means in the replication fork, only one of the newly synthesised DNA strands can grow from the 5' end to the 3'end directly (called leading strand). The other one has to be synthesised from the 5' end to the 3' end towards the opposite direction (called lagging strand) as a series of discontinuous fragments called Okazaki fragments. These fragments can then be linked together by DNA ligase (Figure 1.21). The moving of DNA polymerase is controlled by an accessory protein called PCNA in eukaryotes (Figure 1.27). It forms a ring around the DNA template strand and can move along it freely. It gives the DNA polymerase the ability to continuously synthesise long DNA (Morgan, 2007).

When the DNA replication machine reaches the end of the chromosome in eukaryotes, the leading strand can be completed directly but the lagging strand cannot, because of the lack of space for the primer of another Okazaki fragment. This problem is solved by telomerase. Telomeres are long segments of repeating DNA sequences at the end of chromosomes. The enzyme generating these telomeres is called telomerase. Telomerase contains an RNA template allowing them to add extra DNA repeats on the end of the chromosome. These repeats give space for the primer of newly synthesised Okazaki fragments (Morgan, 2007).

Figure 1.21 DNA replication

This diagram shows the model of DNA replication. Two strands are separated and form a replication fork. Each strand acts as a template. One DNA polymerase produces the leading strand copy; another DNA polymerase binds to the lagging strand and makes discontinuous segments called Okazaki fragments (http://www.replicationfork.com/).

Figure 1.22 A model of the assembly of pre-RC in budding yeast *S. cerevisiae*

This diagram shows the assembly of the pre-replication complex in budding yeast S. cerevisiae. This only occurs at late M phase and early G1 phase when the activity of CDK is down-regulated by the anaphase promoting complex (APC). ORC binds to the DNA replication origin first and loads Cdc6 onto the DNA at early G_1. Following that, ORC and Cdc6 together recruit Cdt1-MCM to origins to form a pre-RC. After the DNA replication starts, the Cdc6 and Cdt1 are hydrolysed to avoid the DNA being replicated more than once in a cell cycle.
(http://dnareplication.csc.mrc.ac.uk/research.html)

The eukaryotic genome is packaged into a nucleoprotein structure called chromatin. The basic repeating unit of chromatin is known as the nucleosome (explained in chapter 1), this comprises approximately 1.7 turns (~146 bp) of DNA wrapped around two molecules of each of the core histone proteins (H2A, H2B, H3 and H4) and, as also explained there are varying lengths of linker DNA together with linker histones in between each pair of nucleosomes. Highly-condensed metaphase chromosomes are formed by multiple folding of the zig-zagging chains of nucleosomes. This highly condensed structure provides protection and organization of DNA, but also permits rapid access to the DNA during its transcription, replication, repair and recombination. These dynamic transitions in the chromatin structure are regulated by hundreds of proteins largely organised in groups including histone chaperones, non-histone DNA binding proteins, ATP-dependent chromatin remodelling complexes and post-translational modification enzymes (Choudhary and Varga-Weisz, 2007; Campos and Reibnberg, 2009; Das *et al.*, 2010). During DNA replication, DNA must first be loosened from core histones by chromatin disassembly. After replication the new DNA strands need to be packed again into chains of nucleosomes leading to chromatin reassembly.

Histone proteins are highly conserved among eukaryotic organisms. Recent research has shown that the structure of a yeast nucleosome is similar to a metazoan nucleosome (Wolffe, 1999). Even in some archaebacteria their nucleosomes contain 80 bp of DNA wrapped around two molecules of each of the two archaeal histones. The similarity of nucleosome structures among eukaryotic organisms indicates that the mechanism of chromatin assembly is highly conserved across all eukaryotes. In our laboratory the histone octamer from chick erythrocytes has been crystallised (Chantalat *et al.*, 2003; Wood *et al.*, 2005) and the structure solved to a resolution of 1.90 Å. This structure is essentially identical to that of the octamer in the nucleosome core particle (Luger *et al.*, 1997) made from cloned Xenopus histones, even to many bound water molecules.

Chromatin assembly is not a well understood process in the cell. The majority of chromatin assembly occurs immediately after DNA replication, where nucleosomes are disrupted by the passage of the replication machinery. The newly replicated daughter DNA strand is rapidly assembled into chromatin by a multistep

process. Firstly, histones H3 and H4 are bound to the DNA, followed by incorporation of two histone H2A-H2B dimers to complete the nucleosome (Das *et al.*, 2010) (Figure 1.23). The original nucleosomes before replication are disassembled into H3-H4 tetramers and H2A-H2B dimers, and then they are randomly distributed to the two daughter DNA duplexes by an unknown mechanism. The parental chromatin only provides half of the histones which are required, the other half of the histones are synthesized during S phase of the cell cycle.

Newly synthesised core histones enter the nucleus mediated by a network of soluble transport receptors called karyopherins that bind to the nuclear localization signal in the N-terminal domain of each core histone. Before assembly into chromatin, these newly synthesized core histones are acetylated at lysine residues 5 and 12 of H4, and lysine residue 14 of H3; and then rapidly deacetylated after deposition onto newly replicated DNA. The reason why this histone acetylation and subsequent deacetylation occurs during chromatin assembly is as yet unknown. *In vitro*, the N-terminal domain of histone H3 and H4 are unnecessary for chromatin assembly. The acetylation of histone N-terminal domains is not required in deposition of histone onto the DNA but recent studies showed that it facilitates the regular spacing of nucleosomes during chromatin assembly *in vitro*.

Histone proteins are rich in the positively-charged basic amino acids lysine and arginine. This gives them an intrinsic affinity for the negatively-charged phosphate groups of DNA in the histone-DNA interactions. *In vitro*, mixing histones and DNA at physiological salt concentration leads to the rapid formation of undefined insoluble aggregates (Figure 1.24). Chromatin assembly occurs in a regulated and ordered manner by the help of some additional anionic factors which shield the charge of the histones from DNA. These additional negatively charged molecules act as histone chaperones to allow the H3-H4 tetramers to bind to DNA first because of their higher affinity for DNA compared with H2A-H2B dimers. H2A-H2B dimers are deposited subsequently, having greater affinity for subnucleosomal particles containing DNA and H3-H4 tetramers than for either histone chaperones or DNA.

These *in vitro* experiments show that chromatin assembly is mediated by histone chaperones. (Figures1.23 and 1.24)

Figure 1.23 The mechanism of chromatin assembly and disassembly

This diagram shows the stepwise assembly and disassembly of nucleosomes mediated by histone chaperones. Histone H2A is depicted in yellow, H2B in red, H3 in blue and H4 in green (Das et al., 2010). In chromatin assembly, it is believed that one histone H3-H4 dimer is attached to the DNA strand first, followed by another H3-H4 dimer to form a 'tetrasome'. Then a histone H2A-H2B dimer is attached to the 'tetrasome', and finally another H2A-H2B dimer is recruited to form a nucleosome. The chromatin disassembly is progressed in the completely opposite way of chromatin assembly. All these processes are mediated by histone chaperones.

Figure 1.24 The chromatin assembly is mediated by histone chaperones

This diagram shows the mechanism of histone assembly (Das et al., 2010).

A. *Directly mixed histone and DNA in vitro will rapidly form insoluble aggregates because of the affinity between positively charged histones and negatively charged DNA.*
B. *Histone chaperones could shield the positive charge of histone proteins and let chromatin assembly occur in a regulated manner.*
C. *An ATP-dependent chromatin remodelling machine is required in the regulation of the space between newly synthesised nucleosomes.*

1.7.2 Histone chaperones

Histone chaperones are negatively charged molecules that bind to histone proteins and facilitate nucleosome formation (Tyler, 2002). Previously, histone chaperones were considered as simple histone binding proteins which were involved in histone transport into the nucleus and in nucleosome assembly. Later experiments indicated that they help to maintain a delicate balance between nucleosome assembly and partial or complete disassembly (Figure 1.25; Table 1.3). Recently, histone chaperones have been implicated in histone removal and exchange in addition to their traditional role of aiding nucleosome assembly. Their role in maintaining chromatin structure and dynamics in non-replicating cells is increasingly recognized (Das *et al.*, 2010).

Many molecules can act as histone chaperones *in vitro* by shielding the basic charge of histone proteins from DNA to mediate the process of chromatin assembly. Such molecules include pectin, RNA, polyglutamic acid and even salt. Most of these molecules just have a little or even no relevance for the assembly of chromatin in the cell. However, some histone chaperones still can be identified by using a cell free chromatin assembly system coupled to ongoing DNA replication. Histone chaperones have traditionally been grouped into chaperones that are 'specific' for H2A-H2B (e.g. NAP1, FACT, Chz1, nucleophosmin, nucleoplasmin, nucleolin), or H3-H4 (e.g. Asf1, Spt6, HIRA, or CAF1) (Table 1.3). Some chaperones are highly specific such as Chz1, which preferentially bind H2A.Z-H2B over H2A-H2B. But some other chaperones, like NAP1, bind all histones including H1 with high affinity.

Many structural studies have been carried out on these histone chaperones, mostly not complexed to their target histones. The structure of several variants of nucleoplasmin, NAP1 and the related SET/TAF-1 and the H3-H4 chaperone Asf1 are now known (Burgess and Zhang, 2010). The structure of HIRA and CAF-1 are also predicted through homology modelling (Burgess and Zhang, 2010). The overall structures of these chaperones are without similarities. However, a four-stranded β-structure is common in the four types of structures (Figure 1.26). Studies show that this region directly interacts with histones in Asf1, and it is also suggested that it might be the primary histone-recognition motif for the NAP1 family (Burgess and

Zhang, 2010).

Asf1 (CIA1) is the most conserved histone chaperone in eukaryotes. It has been shown to interact with the C-terminal of histone H3 by NMR (Mousson *et al.*, 2005), and the later crystallisation studies showed that it binds to histone H3-H4 dimer with interactions to both H3 and H4 (English *et al.*, 2006; Natsume *et al.*, 2007; Ishikawa *et al.*, 2011). Although the exact functions of Asf1 in chromatin disassembly and re-assembly are not yet clear, it is believed to play important roles in the deposition of H3-H4 tetramers onto DNA during nucleosome assembly. The crystal structure of the histone H3-H4-Asf1 complex showed that the complex made a mutually exclusive interaction to another histone H3-H4 dimer. This indicates that Asf1 may play a role in nucleosome disassembly as well (Natsume *et al.*, 2007). Studies by Douglas *et al.* (2011) showed that Asf1 aids both deposition of histone H3-H4 tetramers onto DNA and removal of the histone H3-H4 tetramer from DNA. But the latter process needs additional factors to go to completion (Douglas *et al.*, 2011). The isolated histone H3-H4-Asf1 complex was also found to be a CAF-1 stimulator (Natsume *et al.*, 2007).

Chromatin assembly *in vitro* from proteins extracted from *Xenopus* oocytes and human cells proceeds in a stepwise manner, similar to the chromatin assembly occurring in the cell. A heterotrimeric protein complex termed chromatin assembly factor-1 (CAF-1) was extracted and identified from human cells. It acts as a histone chaperone to deposit newly synthesised histone H3 and H4 onto newly replicated DNA *in vitro* (Stillman, 1986).

CAF-1 mediates chromatin assembly by the interactions between CAF-1 and the nucleoprotein called "proliferating cell nuclear antigen" or PCNA which is a component of the DNA replication machinery. It localises to sites of ongoing DNA replication in the cell and is required for efficient chromain-mediated transcriptional silencing. However, yeast lacks CAF-1 but can still grow well, so this indicates that CAF-1 is not the only histone chaperone in the cell. A novel histone chaperone termed "replication-coupling assembly factor" (RCAF), was identified to have an ability to facilitate CAF-1-mediated assembly of nucleosomes onto newly replicated DNA *in vitro* (Burgess and Zhang, 2010). RCAF is a complex of the *Drosophila* homologue of the yeast antisilencing function 1 (Asf1) protein and histone H3 and H4.

Asf1 appears as a histone chaperone for newly synthesised histones H3 and H4 *in vitro*. The study showed that the function of Asf1 is to bind to CAF-1 to enhance chromatin assembly. To do this, CAF-1 targets Asf1 to the DNA replication fork but does not directly mediate the process (Figure 1.27). CAF-1 also binds H3-H4 by interacting with a component of the DNA replication machinery, PCNA (see above). This facilitates nucleosome assembly following DNA replication and repair (Burgess and Zhang, 2010). Lack of CAF-1 in yeast cells leads to reduced silencing at telomeres and higher sensitivity to DNA damaging agents (Kaufman *et al.*, 1997; Linger and Tyler, 2005).

Nucleosome assembly protein (NAP) is another class of histone chaperone. NAP1 and Vps75 are two proteins in this class which have been shown to associate with all four core histones(Krogan *et al.*, 2006). NAP1 is reported to have roles in transporting histones into the nucleus (Chang *et al.*, 1997; Mosammaparast *et al.*, 2001) and, together with ATP-dependent chromatin remodelling factors, to play roles in disassembly and re-assembly of nucleosomes *in vitro* (Lorch *et al.*, 2006; Tyler *et al.*, 1999). Recently, some studies showed that two histone H3-H4 dimers bind to a single NAP1/Vps75 dimer (Andrews *et al.*, 2008, 2010; Park *et al.*, 2008) which may suggest a new model of deposition and removal of histone H3-H4 onto and from the DNA as an integrity tetramer rather than split into dimers (Bowman *et al.*, 2011) (Figure 1.29).

It is currently believed that during chromatin assembly after DNA replication, H3-H4 or $(H3\text{-}H4)_2$ is first deposited to the DNA, followed by the deposition of H2A-H2B dimers (Burgess and Zhang, 2010). There are two potential procedures to deposit H3-H4 to the DNA. One possible pathway is that two H3-H4 dimers sequentially deposit onto the DNA. Another possibility can be that two H3-H4 dimers form a tetramer on a histone chaperone before being deposited onto the DNA (Das *et al.*, 2010). Studies show that the histone chaperone Asf1 can bind to H3-H4 dimers through an H3 interface which is involved in the formation of $(H3\text{-}H4)_2$ tetramers (English *et al.*, 2006). However, in this structural study, cloned histones were involved and the H3 N terminal helical region was not present. Asf1 was also found to be required for acetylation of histone H3 on lysine 56, which is an important mark of newly-synthesized histones in DNA replication and repair (Recht *et al.*, 2006;

Chen *et al.*, 2008; Li *et al.*, 2008). Rtt106 is another histone H3-H4 chaperone found in yeast (Burgess and Zhang, 2010). Studies have shown that Rtt106 is involved in DNA replication and repair (Huang *et al.*, 2007; Li *et al.*, 2008). A possible model for the deposition of H3-H4 is that H3-H4 dimers are transferred from Asf1-H3-H4 to CAF-1 and Rtt106, which in turn deposit H3-H4 onto replicating DNA for nucleosome formation (Figure 1.28) (Burgess and Zhang, 2010).

Table 1.3 Summary of histone chaperones (Modified from Das *et al.*, 2010)

Figure 1.25 A summary of potential histone chaperone functions

This Figure shows the potential functions of histone chaperones. The functions are determined by the affinity between a chaperone and histones compared to the affinity between the histone and DNA. These affinities can be controlled by post-translational modification of histones or the chaperones (Das et al., 2010).

Figure 1.26 The structures of some histone chaperones

The structures of NAP1, Nucleoplasmin, Asf1 and a homology model of HirA or CAF1 is shown. They share a β-sheet motif, but their overall structures are not similar (Das et al., 2010).

Figure 1.27 The mechanisms of CAF-1 and ASF1 in the mediation of chromatin assembly

Assembly of newly replicated DNA into chromatin requires both Asf1 and CAF-1 in vitro. CAF-1 interacts with the ring-shaped PCNA component of the replication machinery and the recently identified interaction between Asf1 and CAF-1 may target Asf1 to the DNA replication fork to facilitate the process(Burgess and Zhang, 2010).

Figure 1.28 A model of how histone H3-H4 deposits onto DNA

This diagram shows the coordination of histone chaperones for replication-dependent nucleosome assembly. In the replication fork, newly-synthesised H3-H4 is deposited onto the DNA first and followed by deposition of H2A-H2B dimers to form the nucleosome. The deposition of H3-H4 is mediated by histone chaperones CAF-1, Asf1 and Rtt106, but how the Asf-H3-H4 complex interacts with CAF-1 and Rtt106 and whether CAF-1 and Rtt106 bind to H3-H4 dimers or tetramers still remains unknown (Burgess and Zhang, 2010).

Figure 1.29 A new model of deposition and removal of histone H3-H4 as an integrity tetramer: $(H3\text{-}H4)_2$

This diagram shows an alternative mode of histone H3-H4 to be deposited or removed onto or from DNA as an intact tetramer form instead of dimer form (Bowman *et al.*, 2011).

1.7.3 ATP-dependent chromatin remodelling

Histone chaperones alone are not enough to generate regular arrays of nucleosomes with 180-200 base pair spacing *in vitro* (Figure 1.24). Research has shown that chromatin assembly in crude extracts requires ATP hydrolysis in order to generate regular arrays of physiologically spaced nucleosomes. The study of crude *Drosophila* embryos identified a second key component of the chromatin assembly machinery which is an ATP-dependent chromatin remodelling factor (ACF) (Steger and Workman, 1996).

Histone chaperones and ACF work together as a combination to generate regular arrays of nucleosomes in an ATP-dependent manner. The energy required is provided by ATP-hydrolysis which allows the histone octamer to move along the DNA until regular spacing between nucleosomes is achieved.

A similar protein complex to *Drosophila* ACF has been identified in humans, called the remodelling and spacing factor (RSF). Compared with ACF, RSF has more functions including both deposition of histones and spacing nucleosomes. The other difference between ASF and RSF is that the ASF molecule can assemble multiple DNA molecules into chromatin but each RSF only assembles one. This may be due to the ability of RSF to bind to histones H3-H4 and act as a histone chaperone and also a remodelling factor at the same time (Figure 1.24).

1.7.4 Our studies

As described in this review, many mechanisms in DNA replication are still unclear. Some unanswered questions are:How do histone chaperones bind to histones? Do histones H3-H4 bind to their chaperones as a dimer or a tetramer? How were histones H3-H4 deposited onto the DNA with their chaperones?What exactly happens in the replication fork during histone disassembly and re-assembly? Purified histone dimers (H2A-H2B) and tetramers $(H3\text{-}H4)_2$ with a high yield are described in Chapter 2. That gives a great potential to study how histone dimers and tetramers interact with their chaperones and DNA by, for example, crosslinking experiments (Chapter 3.3). The purification of pure histone octamers, dimers and tetramers and the use of them to study the related structures and functions are explained in Chapters 2, 3, 4 and 5.

In collaboration with Dr. Emily Newman and Dr. John MeGeehan in Portsmouth University, we have produced a complex of histone chaperone NAP1 with our pure native chick histones H3-H4 tetramers and studied the structure of the NAP1-$(H3\text{-}H4)_2$ tetramer complex. The complex was first successfully obtained by using our high yield and high purity native H3-H4 tetramers made in this study and their recombinant NAP1. This allowed them to find the correct conditions for forming a meaningful structure allowing a whole range of structure studies of the low resolution NAP1-$(H3\text{-}H4)_2$ tetramer complex. In this work, the histones H3-H4 tetramers were intact and with all its native post translational modifications. This is quite different from other studies (English *et al.*, 2006) when they studied the complex of histone H3-H4 tetramers with the histone chaperone Asf1. In the studies of English *et al.*, (2006) the α-helix and the whole N-terminal tail of histone H3 were not present. Based on these results, the Portsmouth group has produced a complex of NAP1 and their recombinant histones H3-H4 tetramers and compared the complexes containing native intact histones H3-H4 tetramers and recombinant tetramers. These studies are described and discussed in chapter 2.

In the studies of native nucleoproteins obtained from chick erythrocytes with our extraction methods, we found that the linker histone H5 protein co-fractionated with histone tetramers under their native, intact states. This may suggest a new

model of nucleosome assembly in which the histone H3-H4 tetramers deposit onto the DNA first, followed by the deposition of linker histones and nucleosome spacing, and finally the deposition of histone H2A-H2B dimers. These studies are described and discussed in chapter 4.

The methods developed and described in this thesis, can be used to extract and purify nucleoproteins in their native state in a very quick and efficientmanner. This gives a very good opportunity to study histones and their associated proteins, such as histone chaperones, and in order to study the mechanisms of chromatin disassembly and re-assembly, DNA replication and gene expressionin diseases such as cancer.

1.8 Materials and methods

1.8.1 Centrifugation

Centrifugation is a technique used for the separation of suspensions or even macromolecules from solutions according to their size, shape and density by centrifugal force. Basically during centrifuging, the more-dense components migrate away from the centrifuge axis and the less-dense components migrate toward the axis, if they are less dense than the centrifugation medium (Atkinson and Mavituna, 1983).

In a laboratory, centrifugations are normally carried out with small volumes between 1 to 5,000ml of liquid. A rotor holds a number of tubes of sample balanced to pairs, which is placed into the centrifuge machine and an artificial gravitational field would be induced by spinning the rotor. Typical rotation speeds of laboratory centrifuges, used for centrifuging particles such as cell organelles, range from 1,000 – 15,000 rpm. Ultracentrifuges are used for biomolecules (Geankoplis, 2003). The magnitude of the induced gravitational field is measured in terms of the G value, which indicates how many times of gravity is induced into the system. The G value is also referred to as the RCF (relative centrifugal force) value which depends on the rotation speed and the manner in which the centrifuge tubes are held by the rotor:

$$G = r\omega^2/g = 2\pi r n^2/g$$

Where:

r = distance from the axis of rotation (m)

ω = angular velocity (radians/s)

g = acceleration due to gravity (m/s^2)

n = rotation speed (revolutions per minute, rpm)

The distances, r, from the axis of rotation to the top and the bottom of the centrifuge tubes are different, so the bottom of the tube has the highest G value and the top of the tube has the lowest G value. So as the particles migrate to the bottom of the tube they will experience an increasing G value. Normally the average G value

(a mean of the maximum and minimum values) is used for nomograms giving the relationship between rotation speed and G value which is provided by centrifuge manufacturers for each type of rotor. Typical G values range in laboratory centrifuges are from 1,000 – 20,000. Ultracentrifuges can provide G values up to 200,000 and even higher.

Because of the time spent on speed up and slow down of the rotor; to accurately calculate the complete sedimentation time is almost impossible. Commonly there is an empirical correlation which is used to calculate the complete centrifugation time:

$$t = k/S$$

where:

t = complete sedimentation time

k = k-factor of the centrifuge

S = Svedberg coefficient of the material being precipitated

One Svedberg unit (S) is equal to 10^{-13}s (sedimentation coefficient). The k-factor can be calculated using the empirical correlation:

$$k = 2.53\times10^{11}(\ln(r_{max}-r_{min})/rpm_{max}^{2})$$

Where:

r_{max} = radial distance from the axis to the bottom of the tube (cm)

r_{min} = radial distance from the axis to the top of the tube (cm)

rpm_{max} = maximum rotation speed (/min)

An ultracentrifuge which ranges from 30,000 – 100,000 rpm can be used if a higher speed separation is required. It is normally used for separating molecules such as DNA/RNA or proteins. Because of the very high speed, these special-designed devices include provision of vacuum and cooling system to avoid the heat generated during the high speed spinning. The centrifuge tubes with sample need to be balanced very accurately and carefully, to within 0.01 g (Ghosh, 2006).

1.8.2 Dialysis

Dialysis is a common method to separate molecules in solution by the difference in their rates of diffusion through a semipermeable membrane. Typically a solution of several types of molecules is placed into a semipermeable dialysis bag, such as a cellulose membrane with pores that only molecules smaller than a certain molecular weight can pass through. This could typically be 10,000 daltons or, at greater cost, as low as 3,000 daltons. Then the sealed dialysis bag is placed in a container containing different buffers, or pure water. The pores on the semipermeable membrane of the dialysis bag only allow small molecules, often water and salts, to pass through. Large molecules like proteins or DNA are retained inside the dialysis bag.

The solution into which the dialysis bag is placed will be changed after several hours when the salt concentration in the bag has been changed. The new solution changed will lead to a further change of the salt in the protein solution. After several changes, the new salt conditions will be fully established.

This technique is usually used to remove the salt from a protein solution; it will not distinguish between proteins effectively. The main applications of dialysis technique include:

1. Remove of acid or alkali from products
2. Remove of salts and low molecular weight components from solutions of macromolecules
3. Concentration of macromolecules
4. Haemodialysis

During dialysis, the solute transport through the dialysis membrane is driven by the concentration difference between each side of the membrane. The solute flux J can be calculated by:

$$J = SD_{eff}(\Delta C/\delta_m)$$

Where:

S = dimensionless solute partition coefficient (-)

D_{eff} = effective diffusivity of solute within the membrane (m^2/s)

δ_m = membrane thickness (m)

S, D_{eff} and δ_m can be combined to calculate the membrane mass transfer coefficient which gives:

$$J = K_M \Delta C = \Delta C / R_M$$

Where:

K_M = membrane mass transfer coefficient (m/s)

R_M = membrane resistance (s/m)

The dialysis membrane creates a resistance known as R_M which is a constant for a particular solute membrane system. Because normally during dialysis both side of dialysis membrane are liquids, the liquid boundary layers on either side of the membrane also create resistance to solute transport. The total three barriers can be added:

$$R_O = R_M + R_1 + R_2$$

Where:

R_O = overall resistance (s/m)

R_1 = resistance of upstream side boundary layer (s/m)

R_2 = resistance of downstream side boundary layer (s/m)

This equals to:

$$1/K_O = 1/K_M + 1/K_1 + 1/K_2$$

Where:

K_O = overall mass transfer coefficient (m/s)

K_1 = upstream side mass transfer coefficient (m/s)

K_2 = downstream side mass transfer coefficient (m/s)

So the overall solute transport from the upstream to the downstream side can be calculated by:

$N = K_O (C_1 - C_2)$

Where:

C_1 = upstream concentration (kg/m^3)

C_2 = downstream concentration (kg/m^3)

The solution outside the dialysis bag is stirred to keep always the maximum concentration difference between the inside and outside of the bag.

1.8.3 Precipitation

Precipitation is a traditional purification method which is widely used in bioseparations. The basic mechanism of precipitation is to separate the proteins with different solubility in the solution. By adding precipitating agent into the solution, the components which tend to be insoluble under the new condition will be precipitate out from the solution. Then the solid-liquid separation techniques such as centrifugation will be used to separate the insoluble components and the soluble parts will remain in the supernatant.

The precipitation factors are mainly including cooling, pH adjustment, addition of solvents, addition of anti-chaotropic salts, addition of chaotropic salts and addition of biospecific reagents.

The solubility of proteins in aqueous solutions depends on the temperature, and different proteins have different sensitivities to the temperature of their solubility. For example, at a certain temperature one protein may largely be precipitated out but another protein can be still largely in the solution.

pH is another key element in protein solubility. The minimum solubility of a protein is at its isoelectric point (pI). Different proteins have different pI, so if the solution is maintained at the pH equal to the pI of one protein which would largely precipitate out, meanwhile other proteins would remain in the solution. However, proteins may not precipitate out at solvent pH equal to their pI: it depends on the abundance of charged, hydrophilic, amino-acid residues such as lysines, arginines, histidines (basic), asparagine-acid and gultamic-acid present. The presence of high solvent salts ionic strength therefore promotes precipitation (salting out, see later).

Some organic solvents such as ethanol or acetone can be added into the solution to precipitate out proteins or DNA/RNA. These organic solvents can reduce the dielectric constant of the solution and decrease the solubility of the proteins in the solution. This can be calculated by:

$$\ln(S/S_w) = (A/RT)(1/e_w - 1/e)$$

Where:

S = solubility of the protein in solvent (kg-moles/m^3)

S_w = solubility of protein in water (kg-moles/m^3)

A = a constant

e = dielectric constant of the medium (-)

e_w = dielectric constant of water (-)

For example, the dielectric constant of water at 25 °C is 78.3 and ethanol at the same temperature is 24.3. So when ethanol is added into the solution the solubility of the proteins in the solution would be decreased and finally precipitate out. The main problem inthe using of this method is that the organic solvents can denature proteins. To minimize the denaturation this method is normally carried out at low temperatures and the final concentration of the organic solvents in the solution cannot be very high.

Adding salts to the solution can also reduce the solubility of the proteins in the solution. Salts can expose hydrophobic patches on proteins by removing the highly structured water layer which covers the patches in solution. This allows the proteins to interact with each other causing aggregation and precipitation. Meanwhile salts can shield charged groups on proteins which also leads to precipitation. The behaviour of proteins is different in salt addition which makes this a successful purification method. Salt induced protein precipitation is governed by the Cohn equation:

$$\ln(S) = B - K_s C_s$$

Where:

B = a constant which is the natural log of the theoretical solubility of the protein in salt free water.

K_s = salting out constant (m^3/kg-moles)

C_s = salt concentration (kg-moles/m^3)

Step-by-step precipitation is a simple but very useful method in protein purification. Careful design of the precipitation methods can allow large amounts of

different proteins to be divided into small groups.Even pure single protein can be isolated. Crystallization is a special type of precipitation which requires highly optimized and carefully controlled conditions. Crystallization is very important because the protein crystals can then be analysed by X-ray diffraction to study the protein structure (Ladisch, 2001).

1.8.4 Size-exclusion chromatography

Size-exclusion chromatography is a widely used technique to separate proteins by their size. It also can be applied to many biological studies such as removal of non-protein components (DNA), removal of protein aggregates, the study of biological interactions and protein folding (Bailey and Ollis, 1986).

The mechanism of size-exclusion chromatography is basically using a column containing a matrix of porous beads (100-250 μm). When proteins in solution pass through the column they will be separated according to their size. Smaller proteins would enter the pores of the beads which can delay them in passing through; meanwhile larger molecules would directly pass through the column because they are not able to enter the pores of the beads. The result ends up with the larger proteins being washed out first and smaller proteins later.

The ability of separation depends on the type of beads. The pore size of the beads needs to be closely controlled, and the beads need to have high physical-chemical stability. Also the beads need to be hydrophilic and inert, to minimize chemical interactions between sample and matrix (Wheelwright, 1991).

The equipment of size-exclusion chromatography normally includes a pump, a sample injector, the column, a detector and a fraction collector. Before use, the column should be equilibrated with running buffer for at least 1 column volume. Then the protein sample will be applied into the system through the injector. The flow rate of size-exclusion chromatography should be carefully controlled, normally between 5 – 15 cm/h. Too high a flow rate leads to incomplete separation and band spreading. Too slow a flow rateleads to diffusion and band spreading. The sample volume has to be controlled as well. The loading volume is restricted to <5% (typically 2%) of the column volume in order to maximize resolution. Because of that, the protein concentration of the sample should be quite high such as 2 – 20 mg/ml. The solubility depends on the characteristic proteins present so that the sample needs to be centrifuged or filtered to remove any precipitations before loading onto the column, otherwise the column could be blocked. All the solvents used in size-exclusion chromatography need to be degassed while filtering, to reduce the effect of gas bubbles. This applies to all chromatographic procedures (Culter, 2004).

1.8.5 Ion-exchange chromatography

Ion-exchange chromatography is a technique used to distinguish molecules by their net charge. If the pH of the sample is below its pI, the molecules carry a positive net charge and will bind to a cation exchange column; if the pH of the sample is above its pI, the molecules carry a negative net charge and will bind to an anion exchange column. Column contains either cation-exchange resin or anion-exchange resin depending on the net charge of the sample. The sample solution will pass through the column and the target molecules will bind to the column; other molecules will pass through the column with the solution. The target molecule could be washed out afterwards (Helfferich, 1962).

Ion-Exchange Chromatography is based on the interaction between the charged molecules in the sample and the matrix contained in the ion-exchange column. Proteins contain different charged amino-acid residues. Lysine, arginine and histidine have a positive charge at physiological pH; aspartic acid and glutamic acid have a negative charge at physiological pH. The net charge of a protein depends on the overall chargeof its positively charged amino acid groups and negatively charged amino acid groups at physiological pH. The net charge of each charged amino acid varies according to the pH of the solution and it will affect the overall charge of a protein. Each protein has an isoelectric point (pI) which is a certain pH where the protein has a net charge of zero. So if the solution pH is below a protein's pI, the protein will has a positive net charge; on the other hand if the solution pH is above a protein's pI, the protein will be negatively charged (Belter *et al.*, 1988).

There are two major types of ion-exchange matrices which are cation-exchange resins and anion-exchange resins (Table 1.4). Cation-exchange resins have a negatively charged surface and are able to bind to positively charged proteins; anion-exchange resins have a positively charged surface and will bind to negatively charged proteins. These resins are available as dry granular material or as preswollen loose beads, and now prepacked columns are very commonly used.

An increasing salt (e.g. KCl) solution is passed through or a changing pH in the direction of the pI of the protein to be fractionated. In a cation-exchange

chromatography, the protein sample is injected to the column. Buffer (called binding buffer) with a low pH and low concentration of KCl is passed through the column. The proteins with a pI above this pH would be positively charged and bind to the column, other proteins would be directly washed out as the non-bind fraction. Then an elution buffer with an increased KCl concentration and increased pH is passed through the column. The increased positively charged potassium ions, K^+, will compete with the positively charged proteins for binding to the column; in the meanwhile the increased pH towards the pI of the proteins will lower their net positive charge and be released from binding to the column and eluted into fraction-collector, progressively as the salt in the elution buffer is increased. The elution buffer progressive salt increase is produced by control of a mixer device. This mixes the high-salt buffer, B, with the binding buffer, A. Similar effects happen in an anion-exchange chromatography: the increased negatively charged chloride ion in KCl will compete with the negatively charged proteins and the decreased pH will lowerproteins' negative charge and in order to be washed out.

In our group, size-exclusion chromatography and ion-exchange chromatography are all carried out by Fast Protein Liquid Chromatography (FPLC) with the AKTA system (made by G.E. Healthcare). FPLCis a system developed from High-Performance Liquid Chromatography (HPLC), which is designed for a full range of chromatography modes such as ion-exchange chromatography, size-exclusion chromatography, Affinity Chromatography, etc.

Table 1.4 Ion-exchange columns

Ion-exchange Type	Strong exchanger	Weak exchanger
Cation exchangers	SP (sulfopropyl)	CM (carboxymethyl)
	S (Methyl sulfonate)	
Anion exchangers	Q (quaternary ammonium)	DEAE (diethylaminoethyl)
	QAE (quaternary aminoethyl)	

The ion-exchange columns used in the Author's experiments are all supplied by G.E. Healthcare.

1.8.6 Gel electrophoresis

Gel electrophoresis is an analytical technique which separates biomolecules by their electrical charge and size. The biomolecules move under the influence of an electrical field and migrate through a gel at a rate depending on their size and charge.

The gel used to separate molecules is usually a crosslinked polymer whose composition and porosity is chosen based on the weight and composition of the target of the analysis. For proteins or small nucleic acids, the gel is usually made with different concentrations of acrylamide and a cross linker, producing different size mesh networks of polyacrylamide. For large nucleic acids which have greater than a few hundred bases, usually an agarose gel is used. During electrophoresis, a voltage across the gel produces an electric-field profile that exerts forces on the charged molecules allowing them to migrate to various positions in the gel. Molecules move from one end to the other end of the electric field by the net charge they contain. Because the gel provides resistance for their movement, the molecules with smaller size will move faster than the molecules with larger size. The molecules with different sizes and shapes will so be separated (Berg, 2002).

For protein separation, sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is commonly used. The protein solution to be analysed is first mixed with SDS, an anionic detergent which denatures secondary and non-disulfide-linked tertiary structures and gives a negative charge to each protein in proportion to its mass. Without SDS, different proteins with similar molecular weights would migrate differently due to differences in folding, as differences in folding patterns would cause some proteins to better fit through the gel matrix than others.

Another treatment of the protein sample, needed to be done before running the gel, is to heat the sample to near boiling in the presence (or sometimes absence) of a reducing agent such as dithiothreitol (DTT) or 2-mercaptoethanol which further denatures the proteins by reducing disulfide linkages. This is known as reducing SDS-PAGE. Sometimes non-reducing SDS-PAGE is needed when it is important to see if a protein or group of proteins interact via disulphide bridges. PAGE gels can be used, without SDS and without boiling a sample, to allow proteins to migrate

through the gel, linked acrylamide, according to its native conformation and charge.

In SDS PAGE, denatured proteins will be loaded onto the gel carefully, mixed with "sample buffer" containing glycerol and the reducing agent. This allows the solutions to sink to the bottom of the loading pocket formed in the gel. The polyacrylamide gel is inside a tank of two slides of glass and submerged in a suitable buffer. An electric voltage is applied across the gel; the negatively charged proteins will start to move from the top of the gel to the bottom if the bottom electrode is positive. Short proteins will more easily flow through the pores in the gel, while larger ones will have more difficulty, so the small proteins will move faster than larger ones. Following the electrophoresis, the gel may be stained to allow the protein bands to become visible. Stain not bound to the SDS-protein bands is removed by washing in a destaining solution. In our experiments coomassie stain was used and destaining was in distilled water. The recipes and method of SDS-PAGE which were used in our experiments are described in Appendix 3.

1.8.7 Mass spectrometry

Mass spectrometry is an analytical technique used to measure the mass-to-charge ratio of ions. It is usually used to find the composition of a physical sample by generating a mass spectrum representing the masses of sample components. The mass spectrum is measured by a mass spectrometer.

A spectrometer usually involves three main parts which are an ion source, a mass analyzer and a detector system. Firstly the mass spectrometer will produce ions from the sample; then it separate ions of differing masses; after that it detects the number of ions of each mass produced and finally it collects the data and generates the mass spectrum.

This technique could be used to identify unknown compounds by the mass of the compound molecules (Bruins, 1991).

The mass spectrometry might be the most useful technology in proteomics. During this process, proteins will be digested by enzyme (typically using trypsin) into small peptides, which contain three to eight amino-acids residues. The technology produce tiny nanometer-sized droplets, and introduce them so that the peptides are in a charged state in the vacuum of an electo-optical system that determines the mass of a peptide to an accuracy of better than 0.01% (typically +/- 0.01 to 0.001 Daltons). Such amazing accuracy easily distinguishes which amino-acid residues are present in a peptide and also which amino-acid residues have particular modifications, because the options for modifications are limited in number and they introduce uniquely-defined changes in molecular weight. Proteins can be identified through this technology by matching the observed peptide compositions to proteins encoded on the genome map.

Our group is collaborating with Dr. Mark Dickman of The University of Sheffield for the mass spectrometry analysis of proteins separated in SDS-PAGE gels. We pre-treat the gel sample in our laboratory and then send them to Sheffield University for analysis in their Mass Spectrometer.

There are basically five steps in treating the gel sample for mass spectrometry. First of all, the SDS-PAGE gel was carefully rinsed with deionised water and the bands which were going to be analysed were cut from the gel. Then the bands were washed and incubated with ammonium bicarbonate and acetonitrile solution at 37°C to remove the commassie blue stain. After that the gel pieces were reduced to break disulphide bonds and they were then alkylated to prevent disulphide bonds reforming. Then the gel pieces were treated by trypsin digestion which digests the peptides in the gel pieces into approximately five amino acidresidues. Finally the peptides were extracted from the gel pieces and dried. They were then ready to be sent to Sheffield University for mass spectrometry analysis.

1.8.8 Lyophilization (Freeze-drying)

The proteins in solution could suffer from many reactions due to physical, chemical and biological factors. Physical reactions such as aggregation and precipitation, chemical reactions such as glycation of lysine residues, and biological reaction such as degradation by enzymes will affect the further study or analysis of the protein. On the other hand, proteins under a dry state could preserve all or most of its initial biological activity which offers many advantages for long-term storage. Meanwhile to dry the protein sample can be also considered as a concentrating process since the protein can be then dissolved into a certain amount of buffer to provide a certain concentration.

Lyophilization, also known as freeze-drying method is a widely used technique to dry the protein samples. A freeze-drying machine normally consists of a set of shelves, a condenser, and a high performance vacuum pump. The protein sample in solution will first be cooled down to about -40°C, meanwhile the condenseris also cooled down to usually -60 °C. Then the sample will be placed into a chamber where the vacuum pump will reduces the pressure in it. This allows sublimation of the water molecules in the sample solution, and the condenser acts as a lower temperature trap for the sublimed water to be collected as ice.

Freeze-drying is quite a time consuming process. Depending on the total volume of the sample, process for 72 hours or more is not unusual.

Chapter 2

Extraction and crystallisation of histone octamers, dimers and tetramers from chick erythrocyte nuclei

2.1 Extraction of histone octamers, dimers and tetramers from chick erythrocyte nuclei

Introduction

The Liverpool John Moores University Chromatin Research Group has developed the methods for making pure histone octamers, dimers and tetramers and published high resolution histone octamer structure in high KCl/phosphate (1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4) .

A previous student Dr. Sirirath Sodngam (Sodngam, 2007) has established a method of histone octamer purification as a first stage in histone dimers and tetramers purification. With her method histone octamers were purified from most of the other proteins. She also fractionated histone dimers and tetramers at low ionic strength by dialysing histone octamers into 100 mM KCl, 100 mM equimolar phosphate to purify them in a S100 size-exclusion column. The tetramers eluted first from the column and the dimers afterwards, but the two peaks were not well separated.

My tasks are firstly to produce extra-pure histone octamers and then crystallise them and attempt to gain higher resolution structure than 1.9 Å already achieved by our group (Wood *et al.*, 2005). Also we want to improve on the previous student's work by using the extra-pure histone octamers to dialyse against 100mM KCl, 100mM equimolar phosphate to extract pure histone dimers and tetramers by using cation-exchange chromatography, and try to crystallise these histone dimers

and tetramers in order to analyse their structure by X-ray diffraction studies (Chapter 2.1 & 2.2). Finally we want to use the pure native histone dimers and tetramers in collaboration with others in order to produce specific complexes with histone chaperones (Chapter 2.3).

During the course of the work a publication by Santos-Rosa (Santos-Rosa*et al.*, 2008) showed that there is a specific endogenous protease that cleaves the core histone H3. In my hands the use of benzamidine hydrochloride as an inhibitor of proteases in general, and this one in particular, worked well and pure uncleaved histone octamers could be produced. Latterly, the inclusion of this protease inhibitor in the last gel exclusion chromatography, even though this affected the OD_{278} ultraviolet-absorption monitor, gave the final step to fractionate ultra-pure octamers. Benzamidine hydrochloride was also effective for pure core histones in tissue-culture extracts without the need for extra protease inhibitors, for example the "complete" cocktail of inhibitors from the company Merck.

Materials and Methods:

Buffers:

(add benzamidine hydrochloride freshly, keep in 4 °C before use; all buffers were filtered by 0.4 micron filters before use)

1. Nuclei washing buffer: 1 L of 10 mM KCl, 5 mM K_2HPO_4, 5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride.

2. Nuclei Lysing buffer: 400 ml of 2.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride.

3. Histone octamer precipitation buffer: 100 ml of 2 M KCl, 1 M K_2HPO_4, 1 M KH_2PO_4, 2.5 mM benzamidine hydrochloride.

4. Octamer maintenance buffer: 2 L of 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4.

5. Ion-exchange running buffer:

   Buffer A: 1 L of 100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4.
   Buffer B: 1 L of 2 M KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4.

Procedure:

Four tubes (about 160 ml) of the chick nuclei (prepared as in Appendix 1) were thawed in the 4 °C cold room. After thawing, the chick nuclei were partitioned into two 50 ml centrifuge tubes, balanced and then centrifuged at 2000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) (900 xg) for 20 mintues. This speed and time was determined by experiments, if the speed is too high centrifugation would cause nuclei lysis. After centrifugation, the supernatant was discarded and the nuclei pellets were mixed with equal volume of nuclei washing buffer, dispersed by sucking in and out with a submerged Pasteur pipette and made up to 160 ml. Then the nuclei sample was centrifuged again at 2000 rpm (Beckman

Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 mintues and the supernatant was discarded, the nuclei pellets were mixed with equal volume of nuclei washing buffer, dispersed by sucking in and out with a submerged pipette and made up to 75 ml. Then 50 ml of nuclei lysis buffer was put into each of six Ti 45 centrifuge tubes, 12.5 ml of nuclei sample was carefully layered on top of the nuclei lysis buffer in each of six Ti 45 centrifuge tubes and the tubes were balanced in pairs within 0.01 g with nuclei washing buffer. The six tubes were quickly shaken to mix the nuclei and lysis buffer. The nuclei lysed to form a gel in the centrifuge tubes. The balanced tubes were placed into a Beckman Ti 45 rotor to ultra-centrifuge at 35,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 45 rotor) (100,000 xg) for 16 hours. The layering and sudden shaking of the centrifuge tubes were new in my methods, since nuclei take a short time before lysing, by which time the mixed conditions (2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4) had been established for each nucleus.

After ultra-centrifugation, the DNA pellets were kept in the freezer; the supernatant was concentrated down to a certain volume V which the concentration was 20 mg/ml. There was about 480 mg of total proteins in the supernatant so V=24 ml this time. A volume of 0.9×V ml (21.6 ml) histone octamer precipitation buffer was added dropwise while stirring on ice to the concentrated 24 ml of sample. After 10 minutes stirring on ice, a light precipitate was formed and the sample was centrifuged at 12,000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 30 minutes to separate the supernatant and precipitate. The precipitate was kept as a group of proteins we called cePNE2 and further analysis could be done for them. The supernatant was collected and a further 1.5×V ml (36 ml) histone octamer precipitation buffer was added dropwise while stirring on ice to the centrifuged supernatant. After 10 minutes stirring on ice, a heavy precipitate was formed and the sample was centrifuged again at 12,000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 30 minutes to separate the supernatant and precipitate. This time a dense pellet was formed on top of the supernatant after centrifuging; the supernatant containing mainly linker histones was sucked out by a syringe and kept in the freezer. The octamer crusts were dissolved in 10 ml of octamer maintenance buffer and loaded onto a S100 gel exclusion column to run at 0.6 ml/minute flow rate for 500 ml (1.5 column volume) (Figure 2.1).

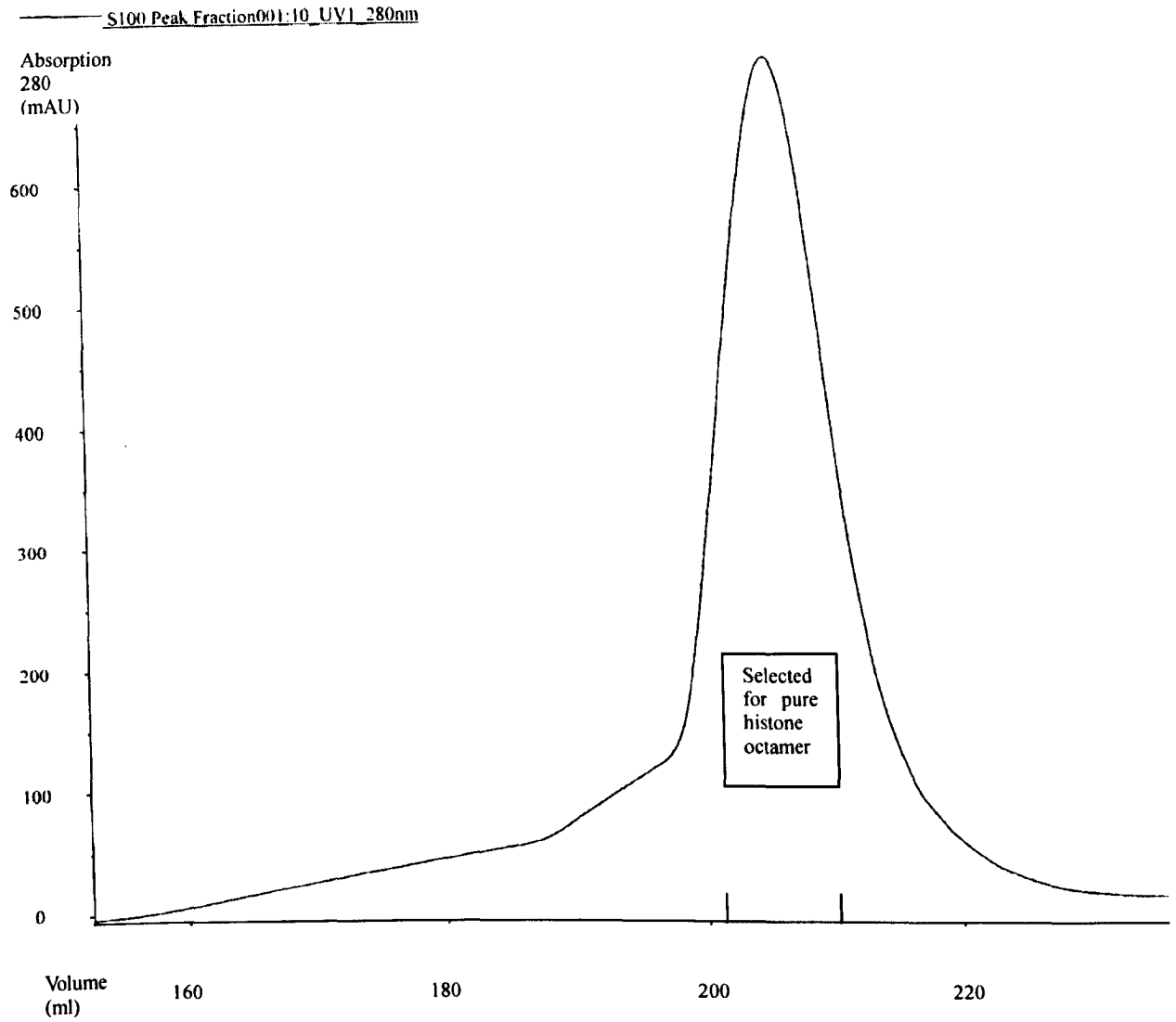
The main peak fractionated from the gel exclusion which contained pure histone octamers (Figure 2.1b) was dialysed against cation exchange buffer A (100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4), then the sample was loaded onto four series of 5 ml Sephadex SP columns (high trap, fast flow, GE Healthcare) to run cation exchange chromatography with buffer A: 100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4; and buffer B: 2 M KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4. (Figure 2.2) There were two peaks separated from this cation exchange which the first one was pure histone dimers and the second one was pure histone tetramers, 20% SDS-PAGE was run for pure histone octamers, dimers and tetramers (Figure 2.3 & 2.4). Note that this is the reverse order from size-exclusion chromatography where the dimers and tetramers were partially separated. This proves that the isolated H3-H4 complex is definitely in the tetramer form.

The principle of this method was that the dimers and tetramers were ultra-pure because they were prepared from pure histone octamers. Therefore they were not contaminated by other proteins which feature in other methods for preparing histone dimers and tetramers.

Results:

The chromatograms of the size-exclusion chromatography for purifying histone octamers and of the cation-exchange chromatography for histone dimers and tetramers are given in Figure 2.1 and 2.2, which are self-explanatory and show high-resolution fractionation of the octamers, dimers and tetramers.

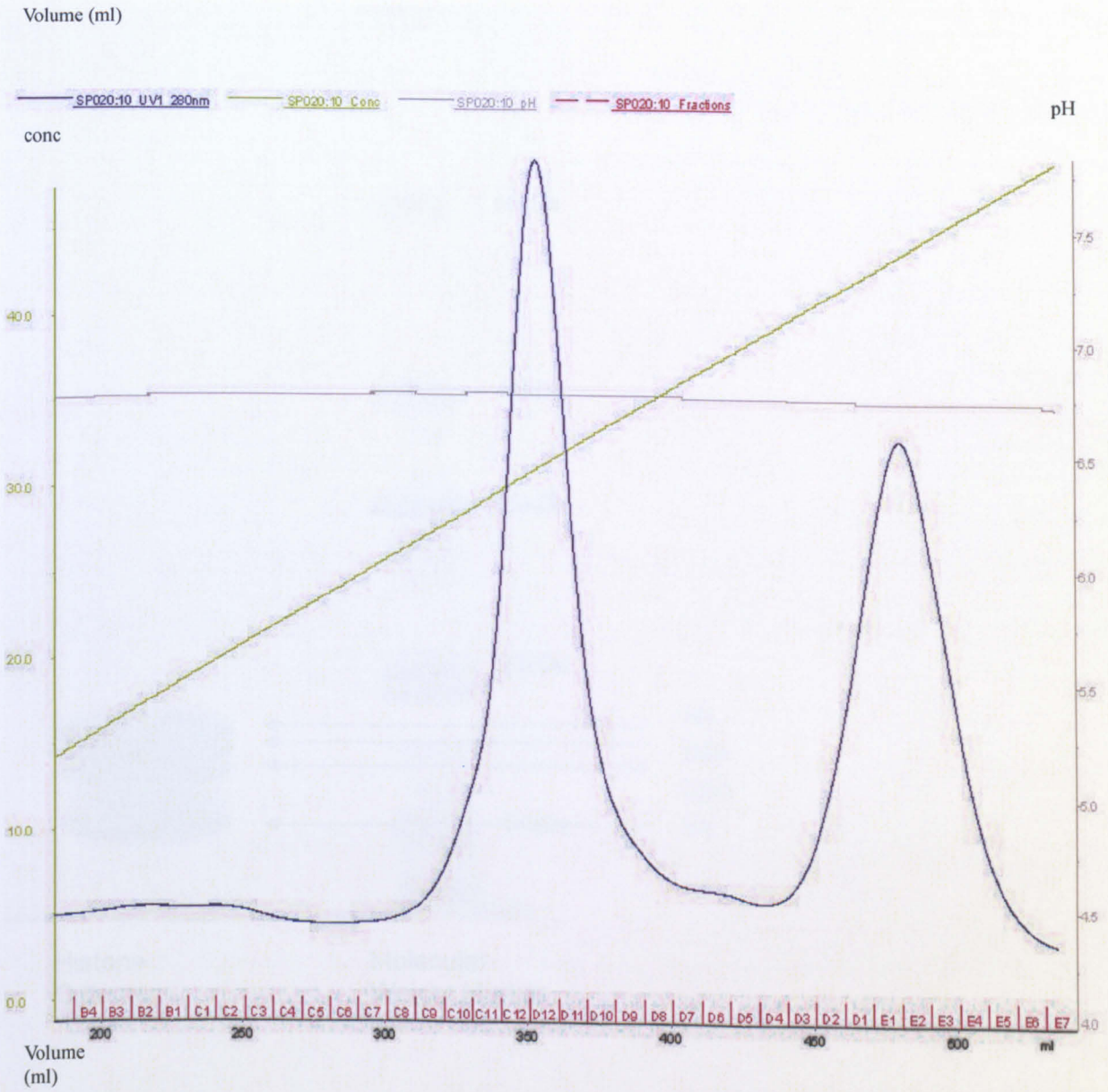
Figure 2.1 Pure histone octamers fractionated from S100 size exclusion chromatography



Four tubes of chick nuclei were lysed by nuclei lysis buffer (2.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride) and ultracentrifuged at 100,000 xg in the Beckman Ti 45 rotor for 16 hours. After that the supernatant was taken and concentrated down to 20 mg/ml. Then after two steps of precipitation with histone octamer precipitation buffer (2 M KCl, 1 M K_2HPO_4, 1 M KH_2PO_4, 2.5 mM benzamidine hydrochloride), most of histone octamers were precipitated out and then re-dissolved into 10 ml of 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4. Then the histone octamer sample was loaded onto a S100 size exclusion column to run at 0.6 ml/minute flow rate for 500 ml (1.5 column volume) with 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4. This is the chromatogram of the S100 size exclusion.

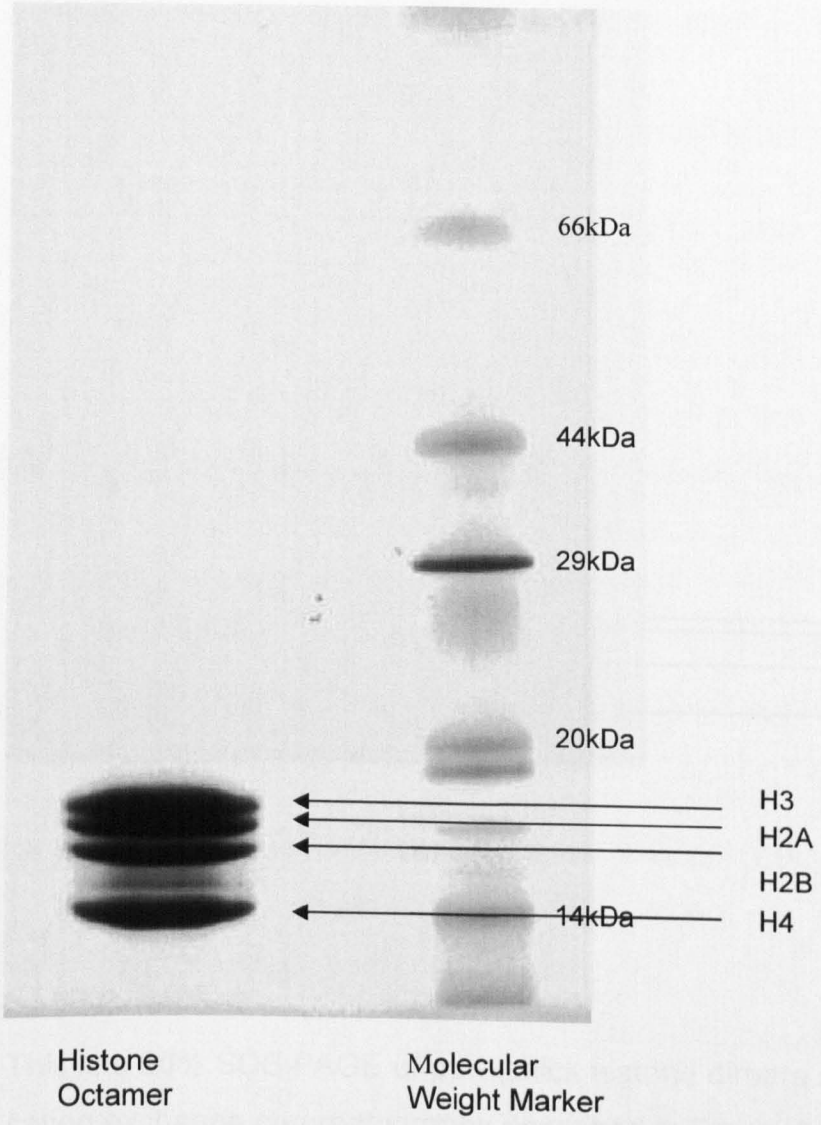
From previous experiments it was known that the middle part of the main peak of this size exclusion contains the most pure histone octamers. So only the central part of the main peak shown in Figure 2.1 was selected for histone octamers production and further histone dimers and tetramers purification. A sample was taken from this selected histone octamer sample and loaded onto a 20% SDS-PAGE gel, the result is shown in Figure 2.3.

Figure 2.2 Histone dimers and tetramers fractionated by cation exchange chromatography



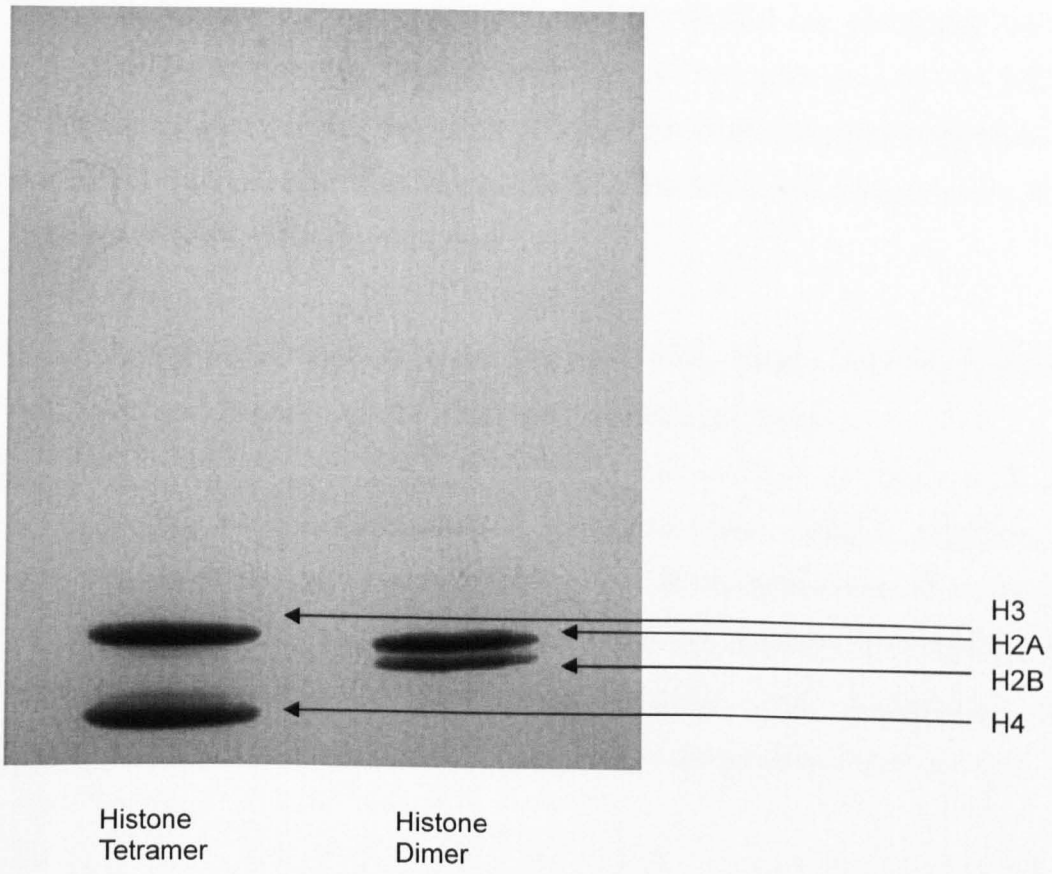
The chick histone octamer sample was loaded on to four series of 5-ml Sephadex SP columns (Hightrap™, fast flow, G.E. Healthcare) running with a salt gradient from 0.1 M KCl, 0.05 M K_2HPO_4, 0.05 M KH_2PO_4 (buffer A) to 2 M KCl, 0.05 M K_2HPO_4, 0.05 M KH_2PO_4 (buffer B). The first peak is eluted with 30% buffer B which is pure chick histone dimer and the second peak is eluted with 40% buffer B which is pure chick histone tetramer. SDS-PAGE of each fraction is shown in Figure 2.4.

Figure 2.3 20% SDS-PAGE of pure chick histone octamer sample selected from Figure 2.1



This is a 20% SDS-PAGE of pure chick histone octamer sample selected from Figure 2.1b. Four bands from top to bottom are histone H3, H2A, H2B and H4.

Figure 2.4 20% SDS-PAGE of pure chick histone dimers and tetramers selected from Figure 2.2



This is a 20% SDS-PAGE of pure chick histone dimers and tetramers fractionated by cation exchange chromatography described in Figure 2.2.

Conclusions and discussion

Very pure histone octamers were produced by removing cePNE1 proteins (Chapter 3) before the lysis of cell nuclei, the removal of cePNE2 proteins by precipitation and careful selection of fractions from the size-exclusion S100 column. A second fractionation after concentrating the fractions chosen from the first column gave even purer histone octamers.

Histone dimers and tetramers were very clearly separated by using four SP 5ml, fast flow, Hitrap cation-exchange columns in series.

In the next section, 2.2, it is shown how histone octamers, dimers and tetramers were set up for crystallisation and X-ray diffraction. Our native pure histone tetramers were the first product to produce the specific complexes with NAP1 – a histone chaperone – in collaboration with Dr. John McGeehan and Dr. Emily Newman from Portsmouth University. This is discussed in section 2.3.

2.2 Crystallisation of Chick Histone Octamers, Dimers and Tetramers

Introduction

When the chick histone octamers, dimers and tetramers had been purified, the next step of study was to try to crystallise them. Our group has successfully crystallised chick histone octamers in 2 M KCl/1.35 M phosphate pH 6.9 at 2.15 Å resolution (Lambert *et al.*, 1999; Chantalat *et al.*, 2003) and at 1.9 Å resolution (Wood *et al.*, 2005). The crystals of chick histone dimers and tetramers have not been crystallised in any research so far. The aim of this experiment was firstly to try to crystallise histone octamers to achieve even higher resolution by using the extra pure histone octamer samplesproduced by the new purification method developed; secondly to try to crystallise chick histone dimers and tetramers and study their structures by X-ray crystallography. This enabled us to appreciate the very different solubilities of histone dimers and tetramers in KCl/phosphate.

Methods:

The chick histone octamers, dimers and tetramers samples used for crystallisation were purified in our laboratory and the methods were described in chapter 2.1. Dialysis technology (chapter 1.8.2) was used for the crystallisation trials of chick histone octamers, dimers and tetramers. The crystallisation conditions were:

Crystallisation solutions:

1. Octamer crystallisation buffer a: 2 M KCl, 0.675 M K_2HPO_4, 0.675 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
2. Octamer crystallisation buffer b: 1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
3. Dimer crystallisation buffer a: 1 M KCl, 0.5 M K_2HPO_4, 0.5 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
4. Dimer crystallisation buffer b: 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
5. Dimer crystallisation buffer c: 0.8 M KCl, 0.4 M K_2HPO_4, 0.4 M KH_2PO_4, 2.5

mM benzamidine hydrochloride, pH 6.8.

6. Dimer crystallisation buffer d: 0.7 M KCl, 0.35 M K_2HPO_4, 0.35 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
7. Dimer crystallisation buffer e: 0.6 M KCl, 0.3 M K_2HPO_4, 0.3 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
8. Dimer crystallisation buffer f: 0.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
9. Tetramer crystallisation buffer a: 1 M KCl, 0.5 M K_2HPO_4, 0.5 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
10. Tetramer crystallisation buffer b: 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
11. Tetramer crystallisation buffer c: 0.8 M KCl, 0.4 M K_2HPO_4, 0.4 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
12. Tetramer crystallisation buffer d: 0.7 M KCl, 0.35 M K_2HPO_4, 0.35 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
13. Tetramer crystallisation buffer e: 0.6 M KCl, 0.3 M K_2HPO_4, 0.3 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.
14. Tetramer crystallisation buffer f: 0.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.

Procedures:

20 ml of 1 mg/ml pure chick histone octamers sample was concentrated down to 1ml (20 mg/ml). Six 10 ml bottles were prepared and labeled, three of them contained 10 ml of octamer crystallisation buffer a (2 M KCl, 0.675 M K_2HPO_4, 0.675 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8), three of them contained 10 ml of crystallisation buffer b (1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8). The crystallisation buffer a was used by Chantalat *et al.* (2003) for chick histone octamer crystallisation and they successfully obtained chick histone octamer crystals at 2.15 Å.

0.05 ml of chick histone octamers sample was loaded onto a crystallisation button, then covered by a piece of dialysis membrane and sealed with rubber

rings.The sealed buttons were put into each bottle containing octamer crystallisation buffer a & b and the bottles were sealed and stored at 4 °C fridge.

Each chick histone dimers and tetramers in 100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4, 2.5 mM benzamidine hydrochloride were concentrated down to 20 mg/ml. Three 10ml bottles were prepared for each different conditions (dimer crystallisation buffer a to e and tetramer crystallisation buffer a to e) and 0.05 ml of samples were loaded to the crystallisation buttons, sealed and stored at 4 °C fridge. All the buttons were checked under the microscope in every 7 days.

Results and discussion

The chick histone octamers formed many crystals in octamer crystallisation buffer b (1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8) (Figure 2.5 a and b), X-ray diffraction analysis showed crystals giving the highest resolution,1.8 Å (Figure 2.6 a and b, Table 2.1). No crystals but a lot of protein precipitations formed in octamer crystallisation buffer a. The Author prepared the histone octamers and produced the crystals. The X-ray diffraction on beam line at the National Light Source Diamond was carried out by Dr. James Nicholson, a member of Liverpool John Moores University group and of the Diamond Source.

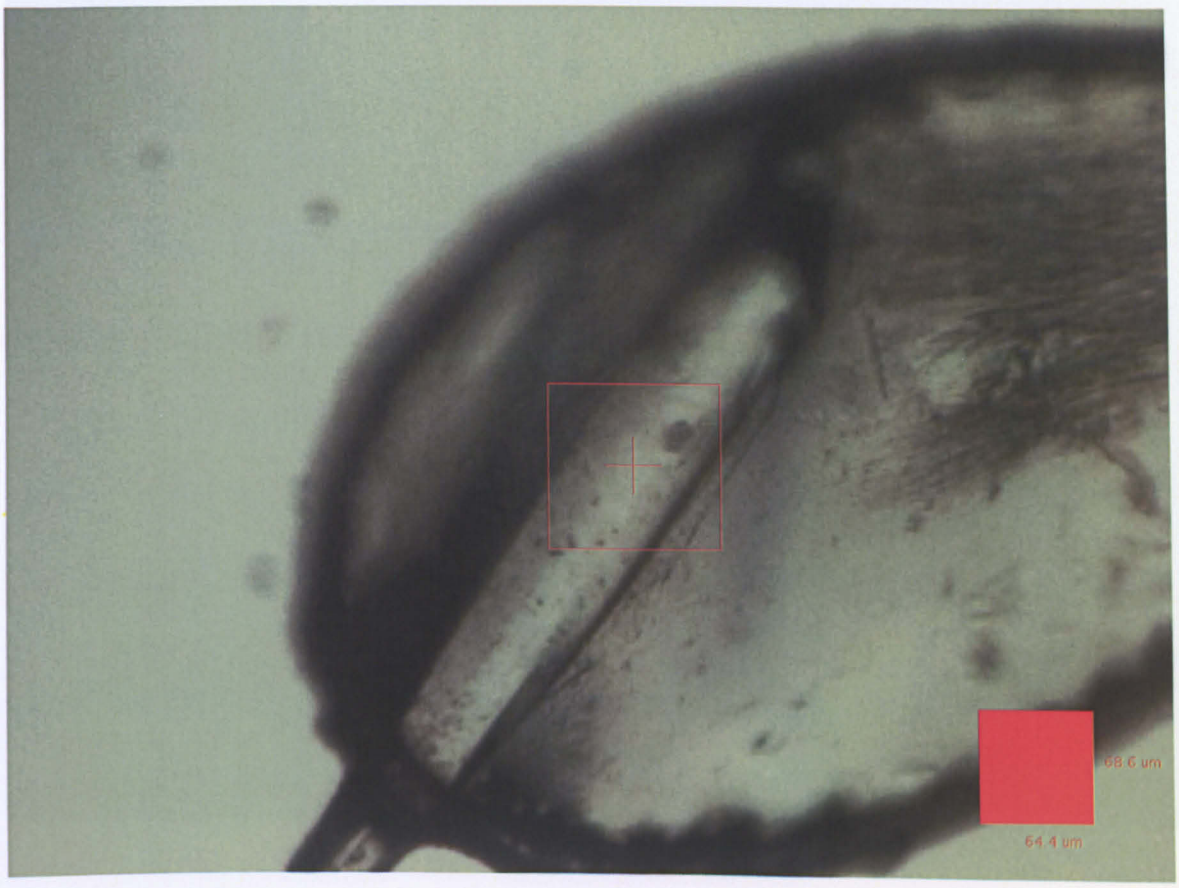
The dimers did not form any crystals or precipitations in any dimer crystallisation buffers. They seem to have very high solubility even in high ionic strength buffer as 1 M KCl, 0.5 M K_2HPO_4, 0.5 M KH_2PO_4, pH 6.8, and much higher. The tetramers did not form any crystals in all the tetramer crystallisation buffers, but a lot of protein precipitations were formed in tetramer crystallisation buffer a (1 M KCl, 0.5 M K_2HPO_4, 0.5 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.) and b (0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8) (Figure 2.7a). A suspected tetramer crystal was formed in tetramer crystallisation buffer b but the analysis showed that it was just a salt crystal (Figure 2.7b).

In a later chapter (chapter 3), where soluble proteins were prepared by direct lysis of chick-erythrocyte nuclei, HMG proteins, linker histones and histone dimers were produced in 1.6 M KCl, 1.22 M equimolar phosphate, but no histone tetramers. Further experiments at high KCl/phosphate, will be made to try to crystallise histone dimers, but internationally it is appreciated that histone tetramers are unlikely to crystallise on their own.

The pIs calculated for histone dimers and tetramers are 10.58 and 11.42 respectively and the total numbers of lysines or arginines in each case are 54 and 112 respectively. It therefore is remarkable that the dimers and tetramers differ in their precipitation conditions. However, this maybe correlated with synchrotron-

radiation circular dichroism studies of the dimers and tetramers of an earlier colleague, Dr. Sirirath Sodngam (Sodngam, 2007). The percentage α helix of the tetramers was appreciately less than calculated from the octamer structure, suggesting changes in the α_N helix of the tetramers when isolated from the histone dimers.

Figure 2.5a The chick histone octamer crystal formed in 1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8. The crystal is at liquid nitrogen temperature and mounted in the loop for insertion into the X-ray diffraction camera



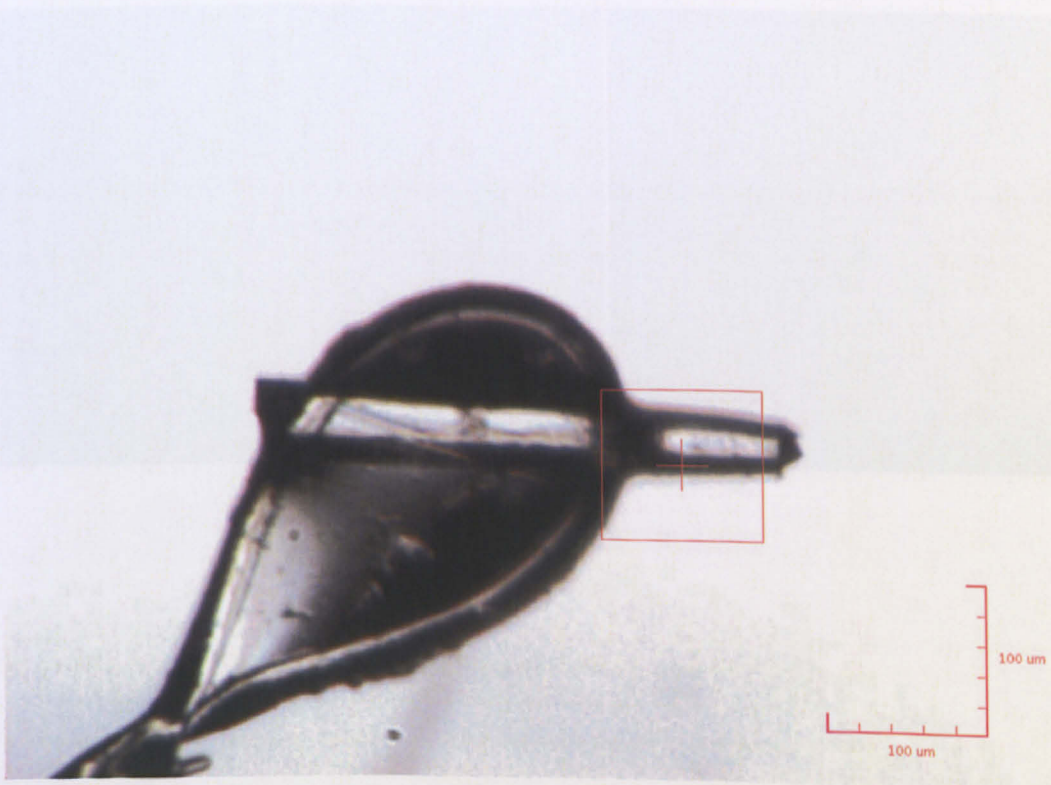
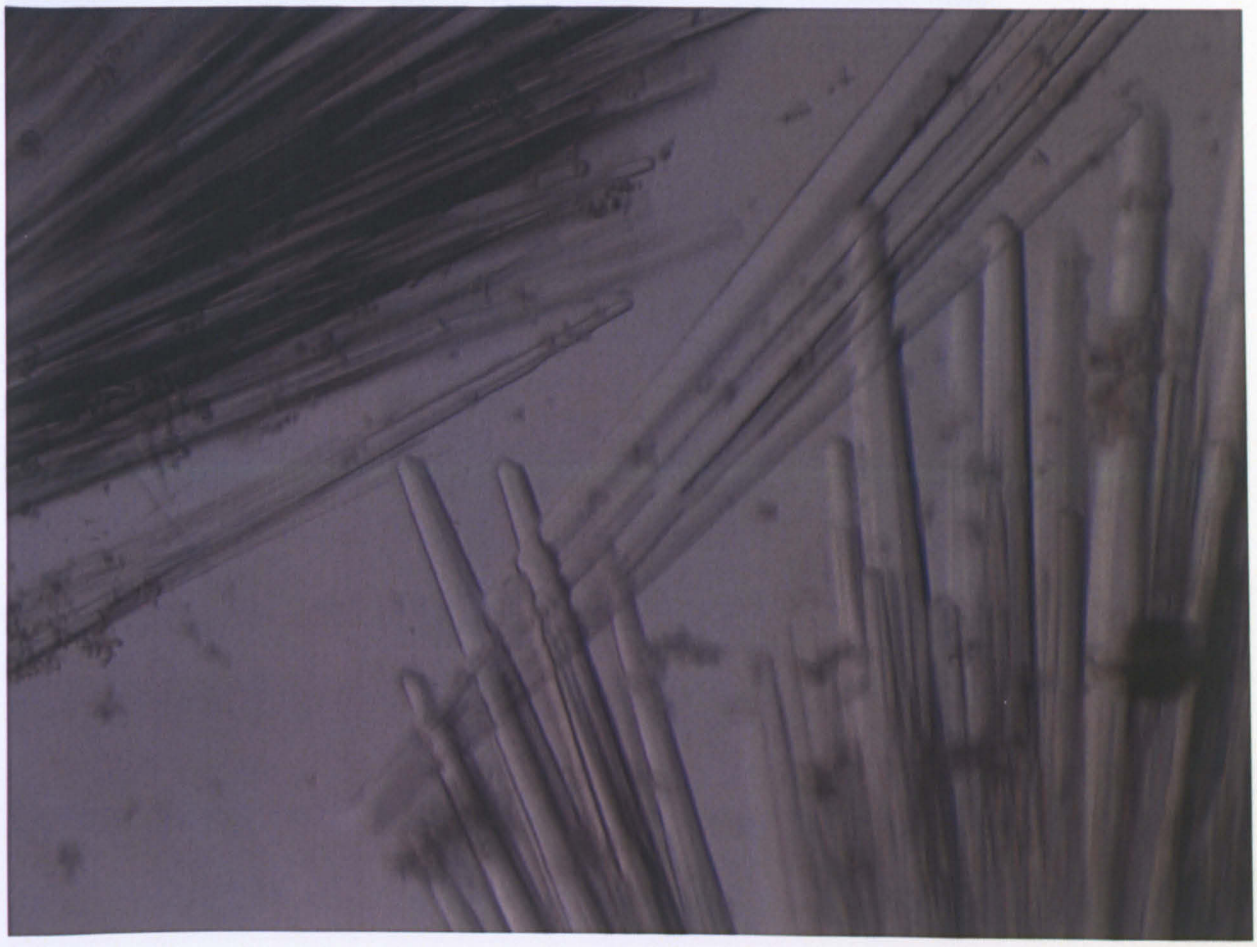
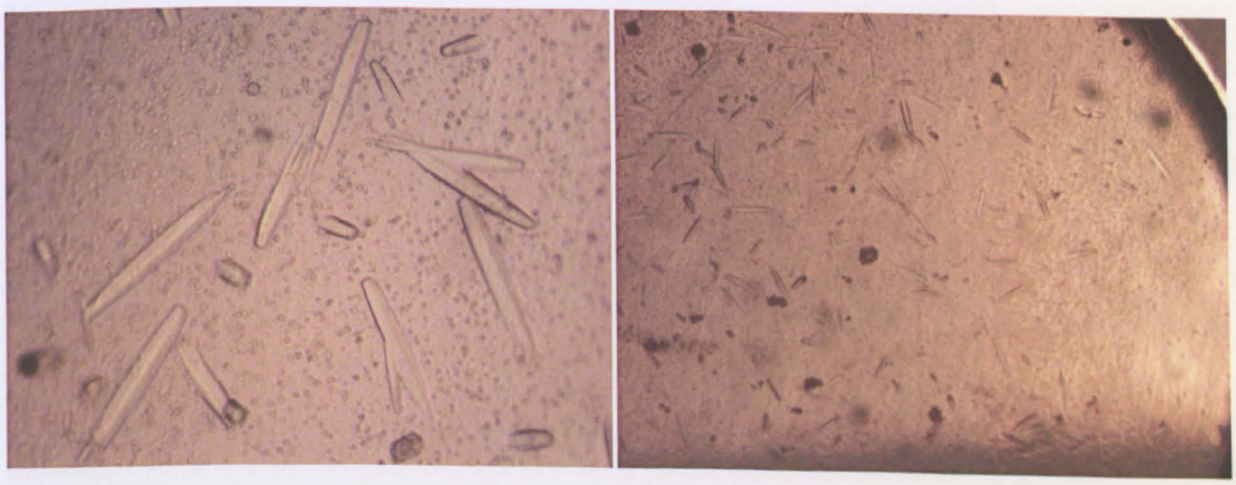
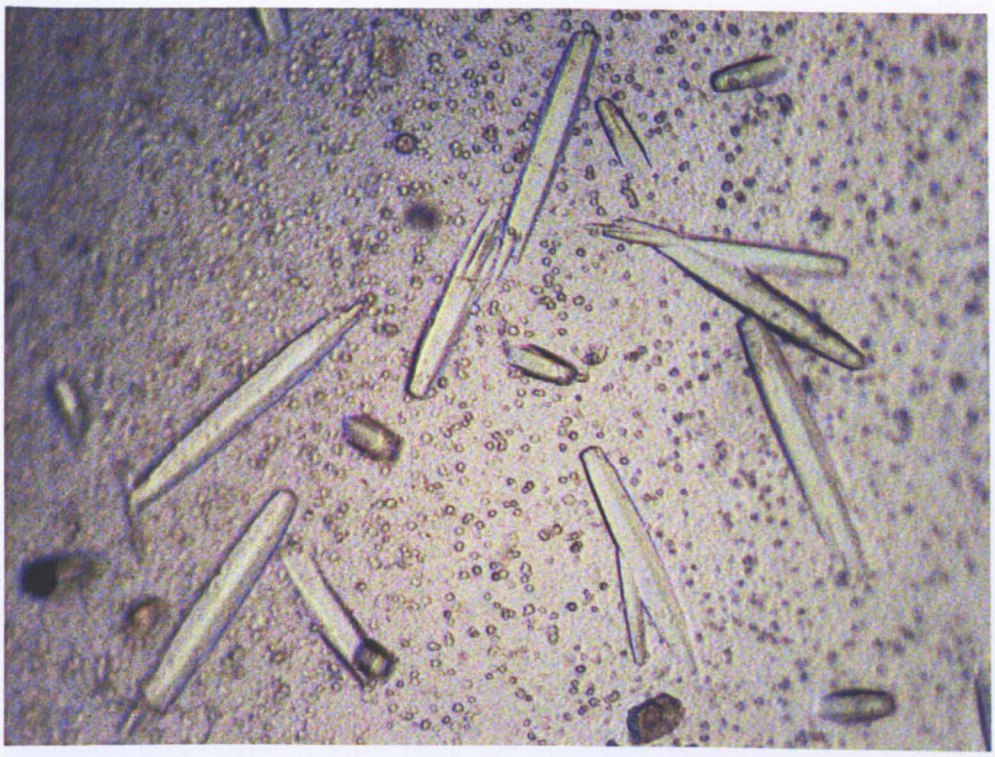


Figure 2.5b Examples of chick histone octamer crystals in crystallisation cells



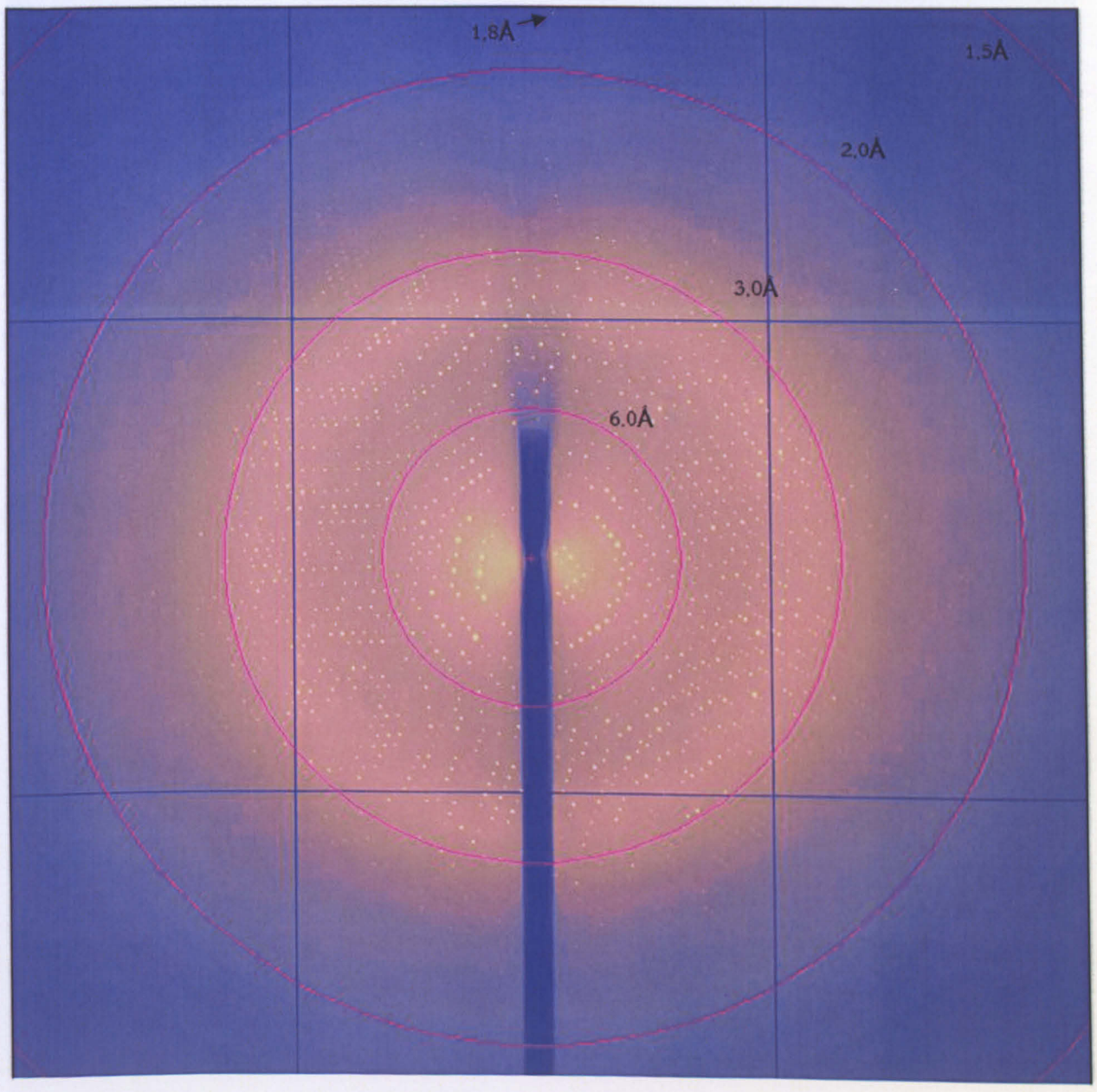




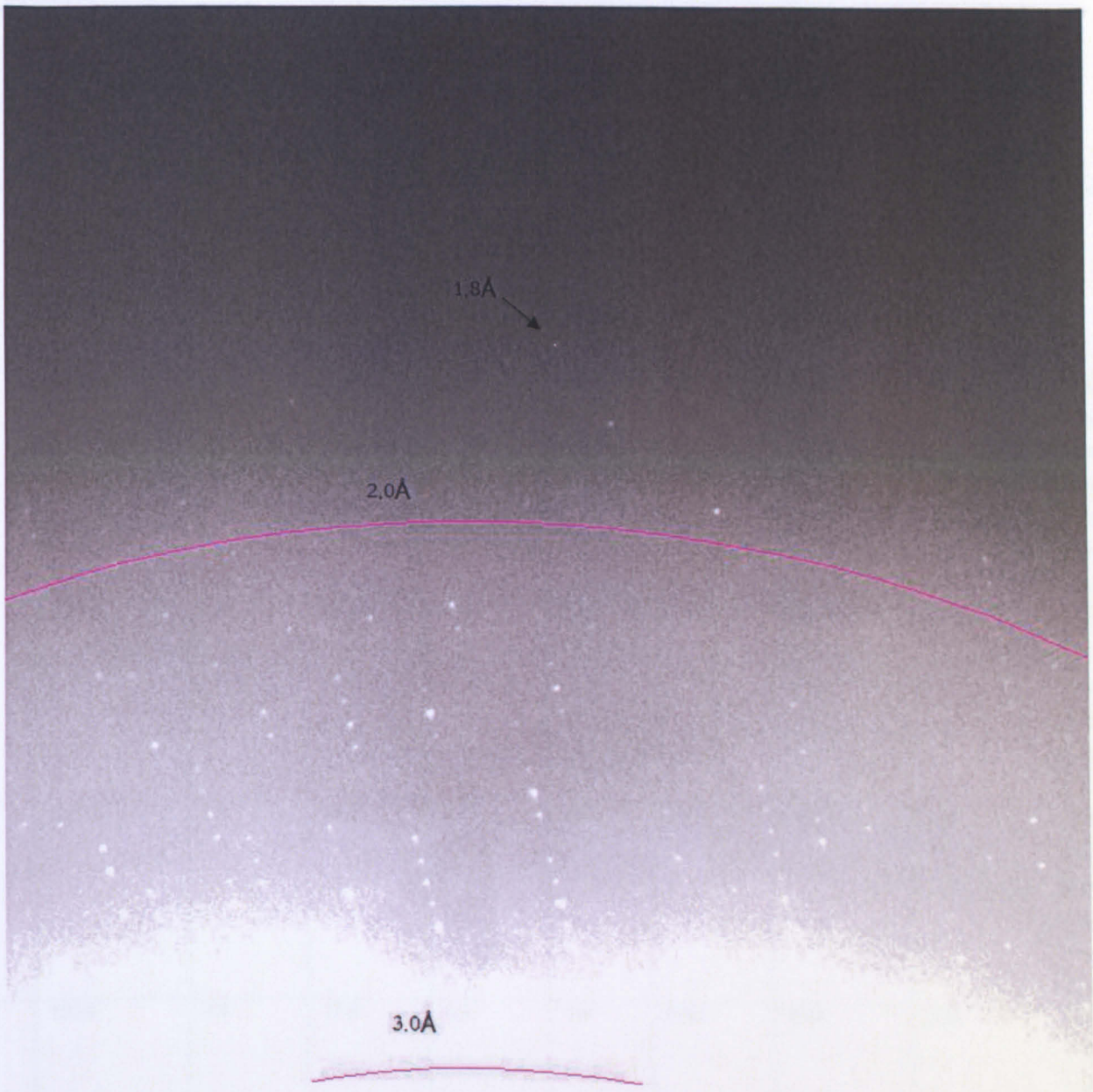
Several images are showing the chick histone octamer crystals growing in 1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8.

Figure 2.6 An oscillation image of the chicken histone octamer crystal recorded on the CCD detector at Diamond

a.



b.



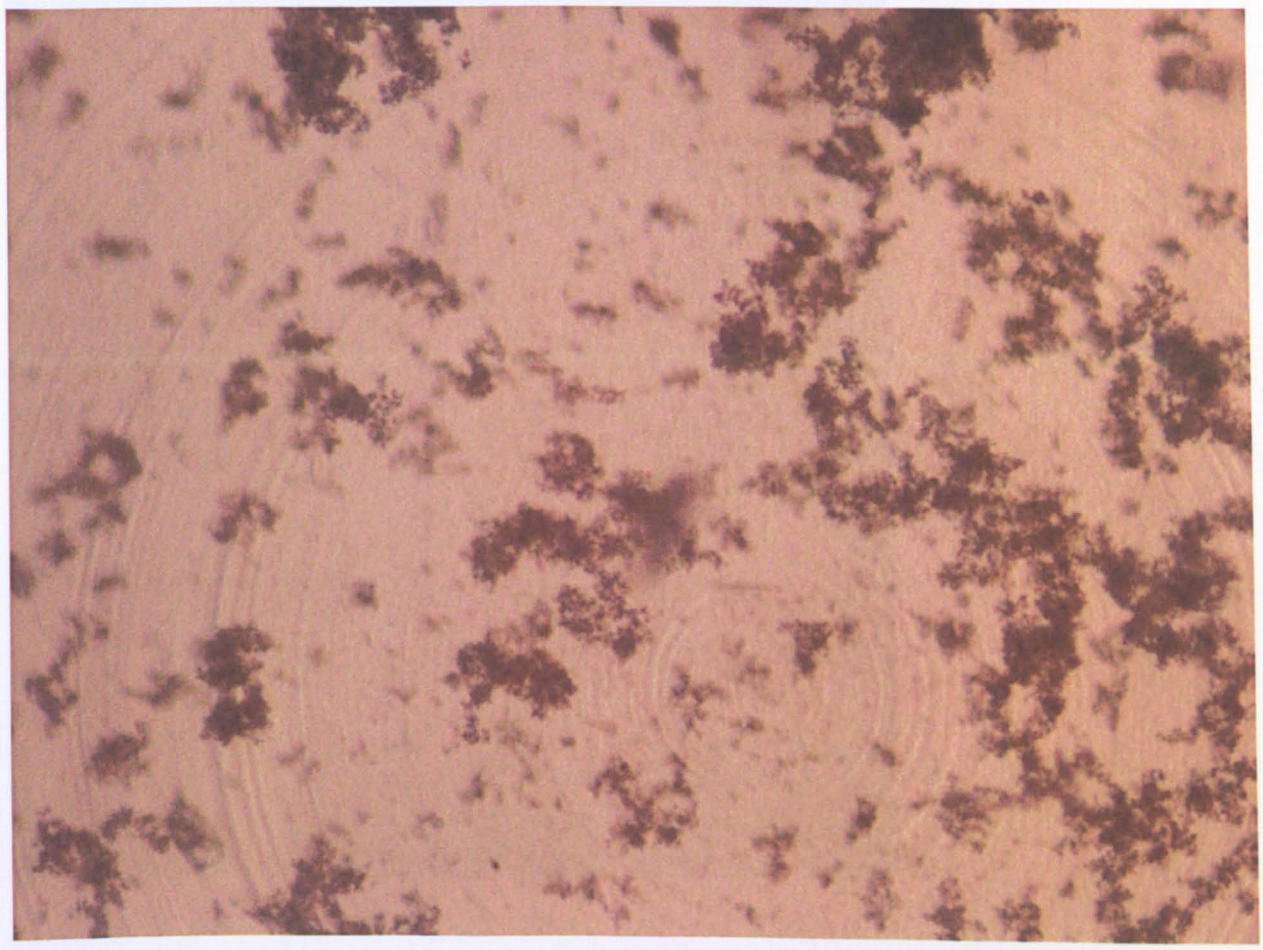
Details of an X-ray diffraction image of the chick histone octamer crystals grown from 1.6 M KCl, 0.8 M K_2HPO_4, 0.8 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8. There are a few diffractions that show the crystal has a resolution higher than 2.0 Å, with some reaching 1.8 Å spacing, which has never been achieved before. Although images of diffraction spots are "noisy" for the highest resolution spots, computer methods enable remarkably accurate determinations of spot intensities.

Table 2.1 Details of the X-ray diffraction studies of four chick histone octamer crystals at Diamond

Strategies to resolution 2.00 Å

Run #	Predicted completeness (%)	Predicted I/sigma	Predicted I/sigma (overall)	Predicted multiplicity	# of images	Total exposure time (min:sec)	Total collection time (min:sec)	Maximum resolution (Å)	Description
1	100.0	1.9	24.3	3.0	532	3:39	25:49	2.07	Standard Anomolous Dataset Multiplicity=3 I/sig=2 Maxlifespan=2 83 s
2	100.0	2.0	32.3	5.1	520	4:07	25:47	2.00	Standard Native Dataset Multiplicity=3 I/sig=2 Maxlifespan=2 83 s
3	100.0	2.1	34.6	11.8	800	3:37	36:57	2.03	strategy with target multiplicy=16, target I/sig=2 Maxlifespan=2 83 s
4	100.0	2.0	17.4	2.9	134	0:34	6:09	2.29	Gentle: Target Multiplicity=2 and target I/Sig 2 and Maxlifespan=2 8

Figure 2.7a The protein precipitation of pure histone H3-H4 tetramers in 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8



This picture shows the protein precipitation of pure histone H3-H4 tetramers in 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8. Although there are no crystals growing, a lot of protein precipitation formed under this condition, which means this condition is close to where crystals would form if that was possible.

Figure 2.7b A suspected chick histone tetramer crystal obtained from 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM benzamidine hydrochloride, pH 6.8

A crystal formed in 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4, 2.5 mM Benzamidine hydrochloride, pH 6.8 histone H3-H4-tetramer crystallisation tray. However, further analysis showed that this was a salt crystal.

2.3 The studies of NAP1-histone tetramer complex in collaboration with Portsmouth group

Introduction

As described in chapter 1.7, histone chaperones are negatively charged molecules that bind to histone proteins and facilitate nucleosome formation (Tyler, 2002). Previously, histone chaperones were considered as simple histone binding proteins which were involved in histone transport into the nucleus and in nucleosome assembly. Later experiments indicated that they help to maintain a delicate balance between nucleosome assembly and partial or complete disassembly (Figure 1.25; Table 1.3). Recently, histone chaperones have been implicated in histone removal and exchange in addition to their traditional role of aiding nucleosome assembly. Their role in maintaining chromatin structure and dynamics in non-replicating cells is increasingly recognized.

Many molecules can act as histone chaperones *in vitro* by shielding the basic charge of histone proteins from DNA to mediate the process of chromatin assembly. Such molecules include pectin, RNA, polyglutamic acid and even salt. Most of these molecules just have little or even no relevance for the assembly of chromatin in the cell. However, some histone chaperones still can be identified by using cell free chromatin assembly system coupled to ongoing DNA replication. Histone chaperones have traditionally been grouped into chaperones that are 'specific' for H2A-H2B (e.g. NAP1, FACT, Chz1, nucleophosmin, nucleoplasmin, nucleolin), or H3-H4 (e.g. Asf1, Spt6, HIRA, or CAF1) (Table 1.3). Some chaperones are really specific such as Chz1, which preferentially bind H2A.Z/H2B over H2A-H2B. However, some other chaperones, like NAP1, bind all histones including H1 with high affinity.

Histone chaperone NAP1

NAP1 (nucleosome assembly protein 1) family proteins interact with a wide range of other molecules to carry out multi functions including histone shuttling (Mosammaparast *et al.*, 2002), cell-cycle control (Kellogg *et al.*, 1995), transcriptional regulation (Walter *et al.*, 1995), histone variant exchange (Okuwaki *et al.*, 2005) and facilitating nucleosome sliding (McQuibban *et al.*, 1998). The crystal structure of

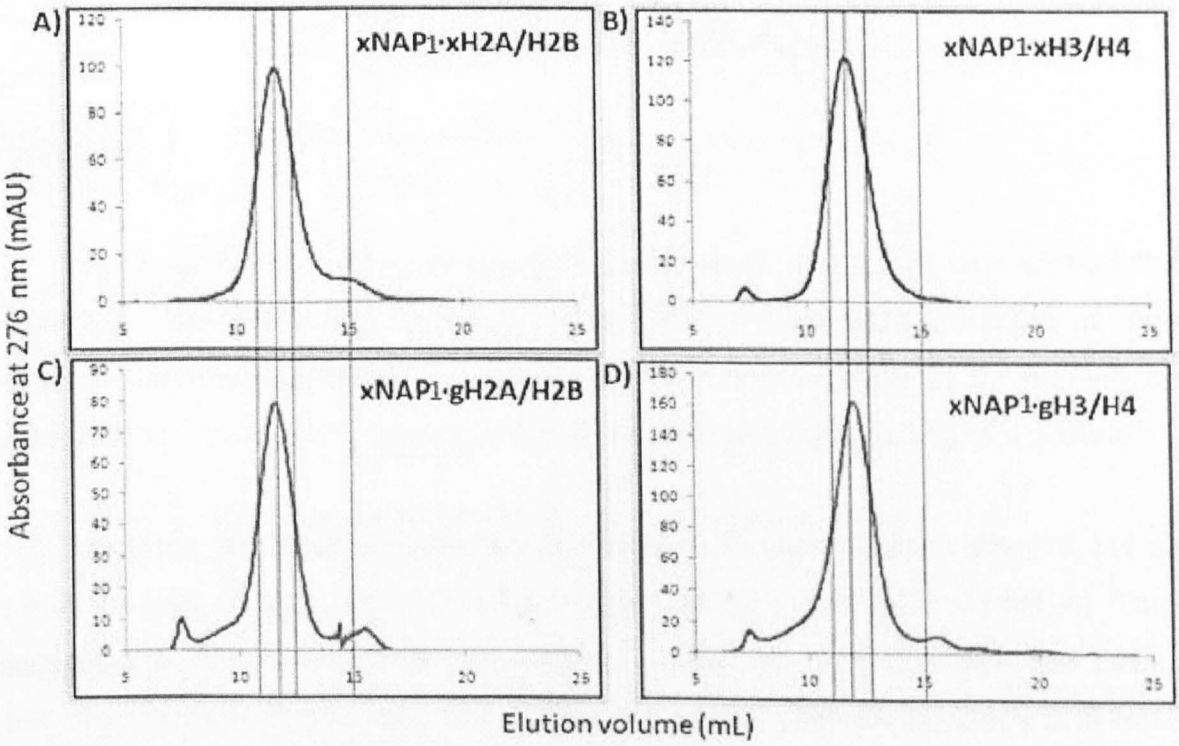
Saccharomyces cerevisiae NAP1 (yNAP1) was obtained by Park and Luger in 2006 (Park and Luger, 2006). The yNAP1 contains unstructured N- and C-terminal regions and a central region with two functional domains, I and II, where domain I consists of three α-helices and domain II consists of an anti-parallel β-sheet, a protruding β-hairpin and three α-helices. The central domain has been shown to be necessary for both histone binding and nucleosome assembly (Fujii-Nakata *et al.*, 1992).

Studies of NAP1-H3-H4 tetramer complex in collaboration with Portsmouth group

In collaboration with LJMU and Dr. Emily Newman and Dr. John MeGeehan in Portsmouth University, they have produced complex of histone chaperone NAP1 with our pure native chick histones H3-H4 tetramers and studied the structure of the NAP1-H3-H4 tetramer complex by protein NMR. The complex was first successfully obtained by using our high yield and high purity native H3-H4 tetramers made in this thesis work and their recombinant NAP1. This allowed them to find the correct conditions for forming a meaningful structure allowing a whole range of structure studies of the low resolution NAP1-H3-H4 tetramer complex. In this work, the histones H3-H4 tetramers were intact and with all its native post translational modifications. This is quite different from other studies (English *et al.*, 2006) when they studied the complex of histone H3-H4 tetramers with the histone chaperone Asf1. In the studies of English *et al.* (2006), the α-helix and the whole N-terminal tail of histone H3 were not present. Based on these results, the Portsmouth group has produced a complex of NAP1 and their recombinant histones H3-H4 tetramers and compared the complexes containing native intact histones H3-H4 tetramers and recombinant tetramers.

In comparing the complex obtained with native histones H3-H4 with those from the recombinant histones, they showed in the size-exclusion on a Superose 6 column, that the elution profiles are similar but there is a slight shift of the maximum between the native and recombinant NAP1-tetramer structure (Figure 2.8). This suggests that the native and recombinant histones form complexes with NAP1 with similar hydrodynamic radii. It also indicates that the presence of PTMs on the histones do not inhibit the formation of large NAP/histone assemblies.

Figure 2.8 The size exclusion chromatography (superpose 6) comparing the difference between NAP1-native histones and NAP1-recombinant histones



This diagram shows the size exclusion chromatography results of NAP1-recombinant histones (xNAP1-xH2A-H2B, xNAP1-xH3-H4) and NAP1-native histones (xNAP1-gH2A-H2B, xNAP1-gH3-H4) (Newman, 2011).

Chapter 3

Chick erythrocyte nucleoproteins studies by using the "Forward technology"

Introduction

The development of technologies to break open cell nuclei and extract their proteins has been part of the research at the JMU Laboratory for a number of years. This is an important technology because it allows the proteins to be purified and studied in detail, particularly to determine structures and remodelling of a protein.

The most important proteins are the histone proteins (H2A, H2B, H3, H4 and H1) and the high-mobility-group (HMG) proteins (HMGB, HMGN and HMGA). These molecules are called "architectural proteins", because they package the DNA in higher organisms (eukaryotes). This packaging of the DNA by histones and HMGs also protects the DNA from any action by other proteins. When cells go through the cell cycle particular proteins and groups of proteins have to get through this protection to access the DNA sequence (chapter 1). The cell cycle is well explained in reviews and on the internet. Some of the intricate mechanisms requiring access to the DNA sequence include transcription, cell-cycle check points, DNA replication and repair, mitosis, planned cell death (apoptosis), changes in cell development (epigenetics), etc.

The histone proteins and the HMGs are remodelled in well defined ways and at precisely defined positions in the total DNA sequence of the species being studied. So for example individual sequences of DNA within specific genes can be transcribed to produce mRNA at particular instants in the timing of the cell cycle.

This intricate control of DNA access to particular proteins involves three types of remodelling:

a. Changes in structure of individual nucleosomes or groups of nucleosomes (this is done with the use of remodelling enzymes with energy provided by ATP) (Thomas,2001);
b. Use of slightly changed aminoacid sequences of HMG or histone molecules (called variants, e.g. use of H2A.Z in place of the usual H2A), inserted in specific nucleosomes in specific genes. The variant proteins are put in place of the usual proteins at specific places on the DNA, at specific times and in very specific mechanisms. Many enzymes are involved in the insertion of variants which themselves are derived from mRNA from their own genes.
c. Detailed changes (modifications) to proteins by adding specific groups to particular aminoacid residues in an HMG or histone protein, after the molecules have been made at the ribosomes in the cytoplasm. So a methyl group (methylation) or an acetyl group (acetylation) may be added to a lysine residue, a phosphoryl group may be added covalently to a serine or threonine amino-acid residue and there are several other modifications. The modifications occur in different groups positioned at particular sites in patterns that are correlated with different functions of nucleosomes within a gene (e.g. within histone H3 during transcription). The modifications are effected using enzymes (acetyl transferases, histone deacetylases (HDACs), methyl transferases, kinases that phosphorylate, phosphatases that dephosphorylate. Modification patterns, at specific proteins and places on the DNA, occur at specific times and to enable very specific mechanisms (Chapter1, Kouzarides, 2008; Turner, 2001).

The remodelling described above is fundamental to the ways in which cells work in disease (cancer cells, cells in sepsis etc.) and in health, so studies of the remodelling in all sorts of situations in the cell cycle are a major study in molecular biology. It is amazingly intricate and detailed in the control and even the existence of cells, organelles in the cells, whole organs and complete species.

Studies of the remodelling of histone proteins and HMG proteins form part of the fields of proteomics and biomolecularstructure determination. This project will take these studies forward to develop efficient ways of extracting and purifying groups of

connected proteins from cell nuclei: The proteins that enable the remodelling of nucleosomes come in groups that are connected by lots of weak connections (i.e. non-covalently connected) and the proteins in a connected group have closely-related functions. For example one group is called the "enhanceosome", which contains among its proteins histone acetyl transferases.

In the LJMU Chromatin Research Group, two major technologies are used for these nucleoprotein studies, one called the "Forward technology" and the other called the "Reverse technology". The "Forward technology" means to separate nucleoproteins step by step into small groups by increasing the KCl/phosphate salt concentration in the extraction buffer, and then do further purification or studies on each of them. "Reverse technology" means to target a group of high-salt soluble proteins, separate them from most of the other proteins in the very first several steps by a relatively high KCl/phosphate extraction buffer to make sure that we can purify the target proteins in the minimum of time. "Forward technology" lyses nuclei at relatively lower concentrations of phosphate/salt buffers compared with the "Reverse technology". This allows most of the nucleoproteins, including histones, HMGs, etc., to be separated from DNA/RNA and dissolved into the lysing buffer. They can then be separated into each group of proteins, step by step, in later purification procedures (Figure 3.1).

A lysing buffer with high phosphate salt concentration is used in a single step in the "Reverse technology". Under this condition, only proteins with the highest solubility, such as HMGs, would remain in the lysing buffer after nuclei lysis. All the other components would be precipitated out with the DNA/RNA pellet.

In the studies of chick erythrocyte nucleoproteins by using the "Forward technology" (the preparation of pure chick erythrocyte nuclei is attached as Appendix 1), the first group of proteins separated from the nuclei is called "cePNE1" proteins.

When the chick nuclei were produced (Appendix 1) with a low salt buffer (100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride), some protein on the surface of the nuclei and inside the nuclei which are soluble under the low salt conditions, could be dissolved into the buffer and

washed out from the nuclei. During this process, a certain amount of $MgCl_2$ was present in the washing buffer to keep the nuclei intact by stabilising the nuclear membrane. These proteins washed out from the chick nuclei are called cePNE1 (chick erythrocyte Proteins Nuclear Extraction) proteins.

A long SDS-PAGE plate-gel of cePNE1 proteins is given in Figure 3.2 and the proteins were analysed by mass spectrometry (Appendix 2), before the work of the Author commenced. There are some quite important proteins involved in this group (Table 3.1), so some further purification was decided to be applied to this group of proteins. First the cePNE1 sample was directly loaded onto a cation-exchange chromatography column to try to separate the proteins. Then another experiment was done which was to precipitate out some proteins from the cePNE1 proteins first and then to run both cation-exchange and anion-exchange to try to separate the proteins remaining soluble after precipitation. These two experiments are explained in chapter 3.1. Then, after the lysis of nuclei, another group of proteins are precipitated out with the DNA/RNA as "cePNE3". Then during the purification of histone octamers, dimers and tetramers (Chapter 2), "cePNE2", "cePNE4" (Chapter 4.1) and "cePNE5" (Chapter 4.2) are separated step by step and further purification and studies were introduced to these groups of proteins (Figure 3.1). The results of these studies are given in each sections of this chapter.

Figure 3.1 Purification of nucleoproteins by using the "Forward technology"

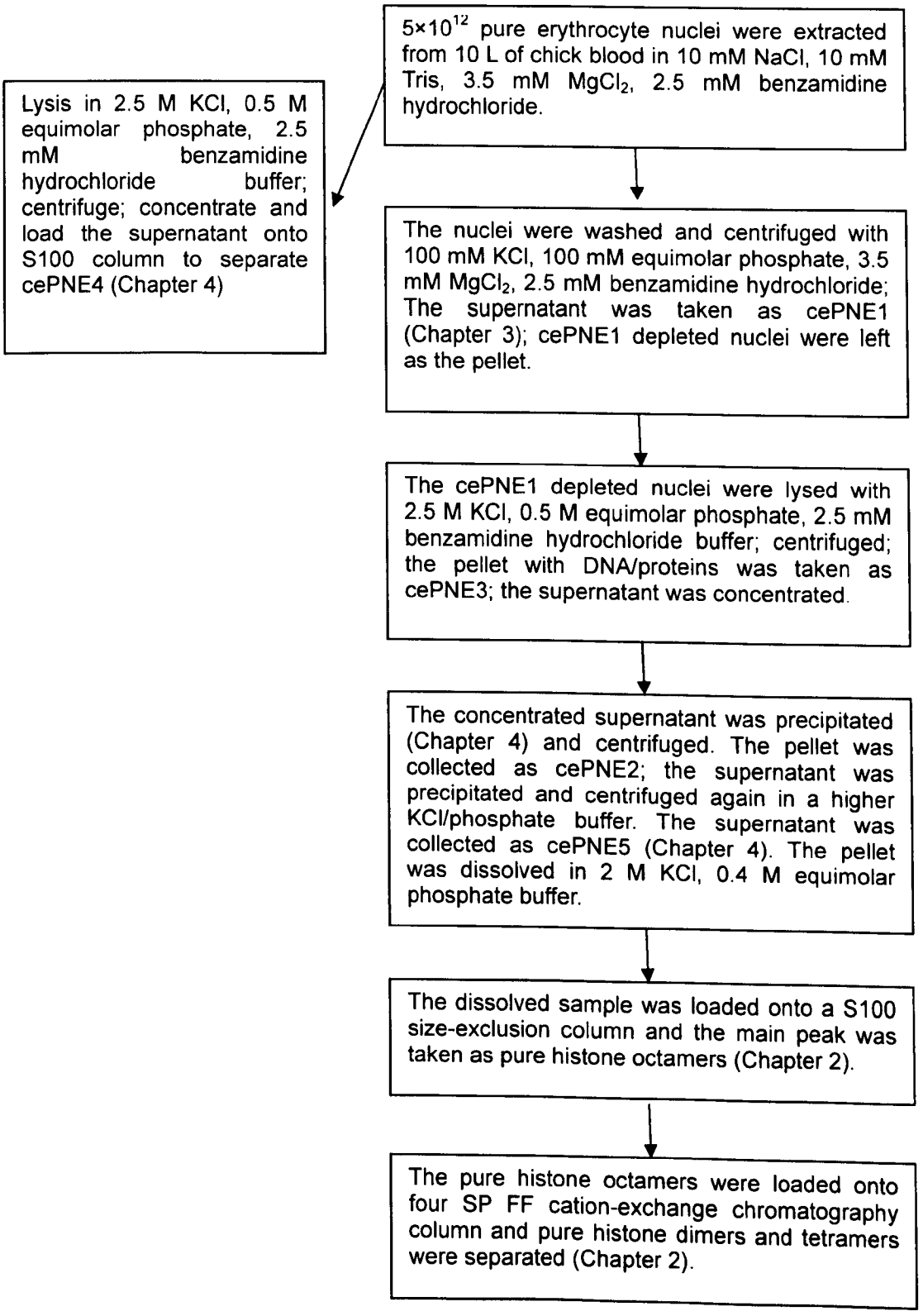
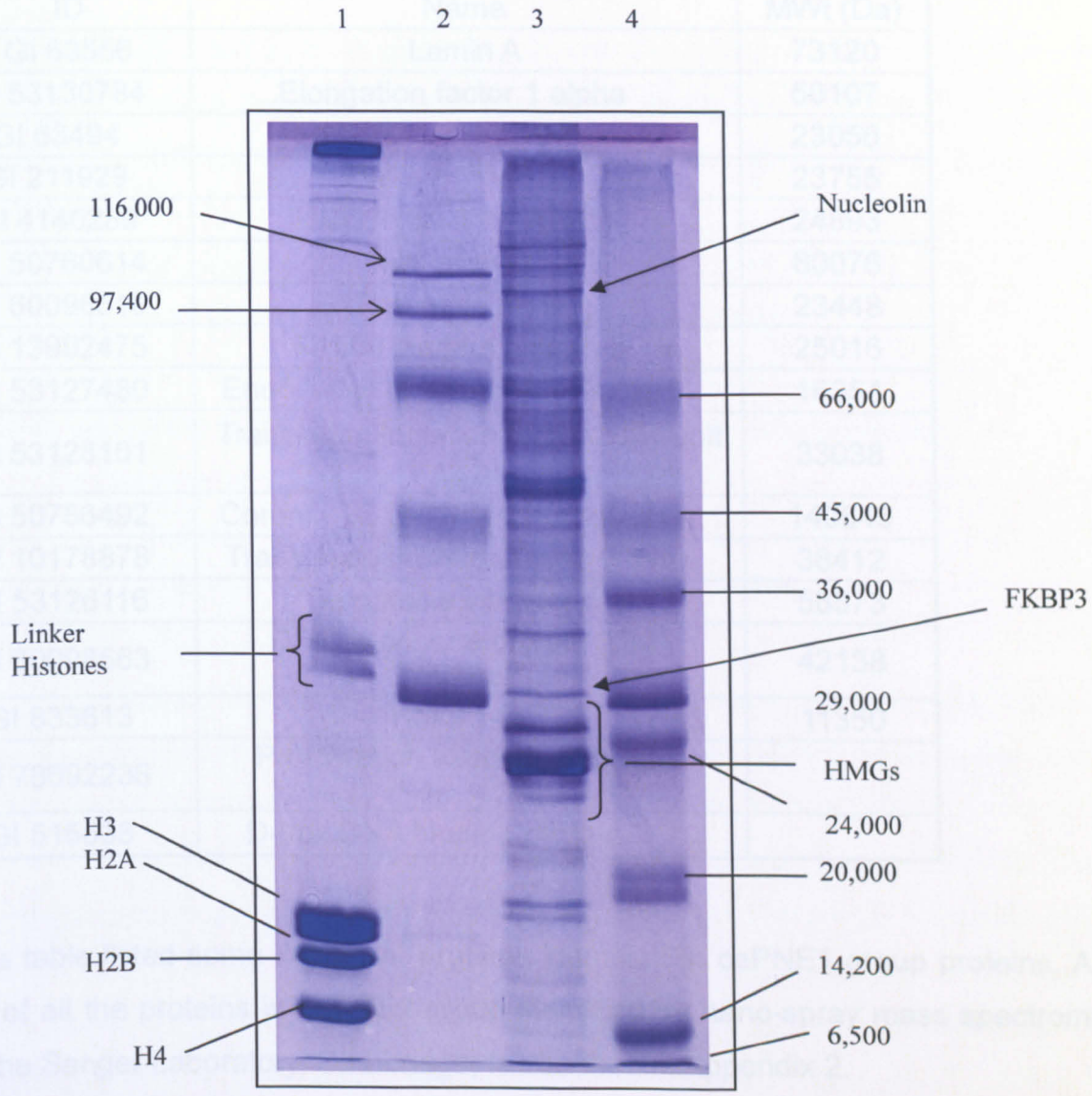


Figure 3.2 An SDS-PAGE of cePNE1 proteins



A 175 mm long plate SDS-PAGE gel of the proteins present in the fractions obtained after washing in HMG extraction buffer (100 mM potassium chloride, 100 mM equimolar mono and dibasic potassium phosphate, 3.5 mM magnesium chloride). The sample was precipitated with TCA, washed in acetone and dissolved in sample buffer containing 2-mercaptoethanol. Lane 1: Chick erythrocyte histone standard. Lane 2: High-range molecular-weight marker (from top to bottom myosin 205000 Da, β-galactosidase 116000 Da, phosphorylase b 97400 Da, albumin bovine 66000 Da, albumin egg 45000 Da and carbonic anhydrase 29000 Da). Lane 3: Sample. Lane 4 Low molecular weight marker (from top to bottom albumin bovine 66000 Da, albumin egg 45000 Da, glyceraldehydes-3-phosphate dehydrogenase 36000 Da, carbonic anhydrase 29000 Da, trypsinogen 24000 Da, trypsin inhibitor 20000 Da, α-lactalbumin 14200 Da, aprotinin 6500 Da). The HMG proteins were clearly within the molecular weight range ca. 17,000 to 27,000Da. All the proteins in cePNE1 group were identified by mass spectrometry (Table 3.1, Appendix 2) (Gel did by previous student Joanne Bebbington, bands identified by mass spectrometry).

Table 3.1 Some important proteins identified in cePNE1

ID	Name	MWt (Da)
GI 63556	Lamin A	73120
GI 53130784	Elongation factor 1 alpha	50107
GI 63494	HMG 2a (HMGB3)	23056
GI 211929	HMG 2 (HMGB2)	23755
GI 4140289	HMG Type 1 (HMGB1)	24893
GI 50760614	Tau protein	80076
GI 60098519	SAP domain protein	23448
GI 13992475	FK506 binding protein 3	25016
GI 53127480	Endothelial differentiation factor1	16354
GI 53128101	Translation Initiation factor 2 subunit 1 alpha	33038
GI 50756492	Coronin 1C (hCRNn4 homologue)	140645
GI 10178878	Translation initiation factor 2 beta	38412
GI 53126116	Apoptosis inhibitor 5	58573
GI 60098563	DEK oncogene protein (DNA binding)	42138
GI 833613	HMG 14a	11350
GI 78692236	Proliferation associated protein 1	
GI 515838	D-box DNA binding protein	

This table listed some important proteins identified in cePNE1 group proteins. A full list of all the proteins in cePNE1 group identified by nano-spray mass spectrometry at the Sanger Laboratory, Cambridge, is attached in Appendix 2.

The mass spectrometry of these proteins, in table 3.1, was done by Professor Hornby's group at Sheffield University using MALDI-TOF.

3.1 Cation-exchange of cePNE1 proteins

Cation-exchange chromatography of cePNE1 proteins was carried out using the following buffers (benzamidine hydrochloride was added and kept freshly at 4 °C before use; all buffers were filtered by 0.4 micron filters before use):

1. Nuclei washing buffer: 1 L of 10 mM KCl, 5 mM K_2HPO_4, 5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride.
2. cePNE1 extraction buffer: 200 ml of 80 mM KCl, 40 mM K_2HPO_4, 40mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride.
3. Cation-exchange running buffer:
   Buffer A: 1 L of 80 mM KCl, 80 mM KH_2PO_4, pH 4.5.
   Buffer B: 1 L of 1 M KCl, 80 mM K_2HPO_4, pH 8.

Four tubes (about 160 ml) of chick-erythrocyte nuclei were thawed in the cold room, then the chick nuclei were partitioned into four centrifuge tubes, balanced and centrifuged at 2000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) (650 xg) for 20 minutes. The supernatants were discarded and the nuclei pellets were mixed with an equal volume of nuclei washing buffer (10 mM KCl, 5 mM K_2HPO_4, 5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride), dispersed and made up to 160 ml in four tubes, each containing 40 ml. The samples were centrifuged at 1,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) (359 xg) for 15minutes, the supernatants were discarded.

An equal volume of cePNE1-extraction buffer (80 mM KCl, 40 mM K_2HPO_4, 40 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride) was added to each nuclei pellet and the mixture was homogenised with a submerged plastic pipette. The suspensions were then transferred to two separate tubes through a 22-gauge syringe needle to disperse the nuclei and prevent clumping. The tubes were then centrifuged at 1,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 20 minutes and the supernatants were retained.

This cePNE1 wash was repeated twice, the supernatants of all three runs were combined as cePNE1 sample. The combined sample was then ultracentrifuged

at 100,000 xg (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) overnight with six tubes in the Ti 70 rotor to avoid any DNA/RNA remaining in the sample.

Extracting the proteins three times with a small amount of PNE1 extracting buffer allowed the ionic strength to be increased slowly and to keep the concentration of protein in the supernatant as high as possible. One wash with a large volume of extraction buffer to extract as many proteins as possible would have led to a much larger dilution of proteins in the supernatant.

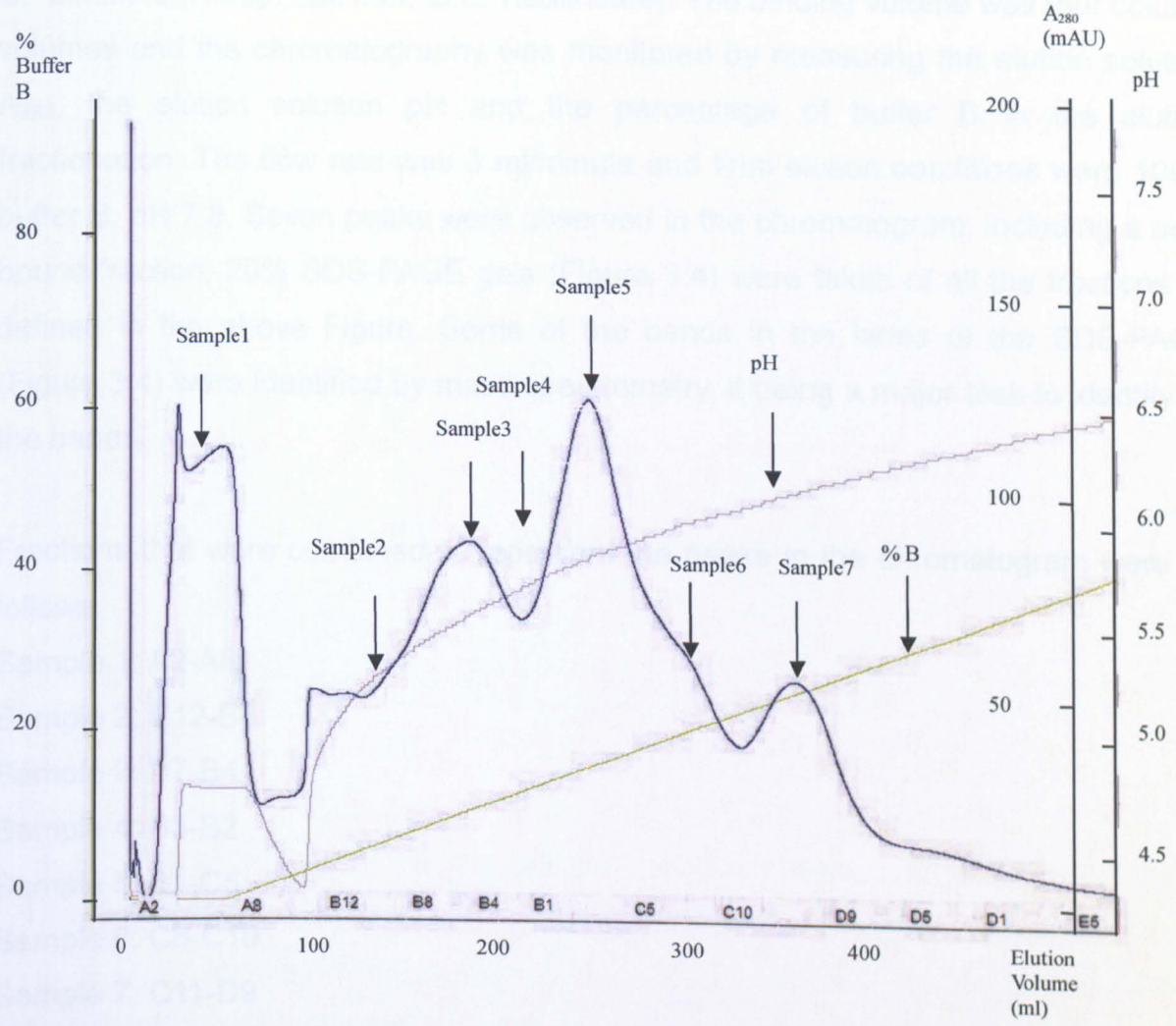
The cePNE1 proteins were dialysed against cation-exchange buffer A (80 mM KCl, 80 mM KH_2PO_4) and then concentrated down to 40 ml (4 mg/ml, 160 mg of total proteins). The sample after dialysis and concentration was then loaded onto four SP cation-exchange columns (5-mls each column, Hitrap, Fast Flow, G.E. Healthcare) and run at 3 ml per minute with a linear gradient between buffer A and elution buffer B. The chromatogram is shown in Figure 3.3; seven peaks were observed in the chromatogram, including a non-bound fraction. 20% SDS-PAGE gels (Figure 3.4a) were taken of all the fractions as defined in Figure 3.3. Some of the bands in the lanes of the SDS-PAGE (Figure 3.4b) were identified by mass spectrometry.

The areas under the seven peaks in Figure 3.3 show that the proteins in peaks 3 and 4 are quite abundant in view of the large quantity of proteins extracted by one-step removal of proteins from nuclei from 10 litre of blood (~1 g). Heat-shock protein 70 (Hsp70) is very abundant. The peptides from the Hsp70 mass spectrometry show that they come from the N-terminal half of the protein as listed in the NCBI data for chick erythrocytes (residues 1 to 390). This was clear from mapping the mass spectrometry peptides against the NCBI database for Hsp70

Peak 5 in Figure 3.3 contains mainly HMGB proteins, which are therefore abundant in the cePNE1 proteins, at a preparative level.

There are no histones in the cePNE1 proteins except a very small quantity of H2Az histone shown up in the nanospray mass spectrometry.

Figure 3.3 The cation-exchange chromatogram of cePNE1 proteins



The colour code in this Figure refers to the A280 (blue), % buffer B (green) and pH (grey). The chick erythrocyte nuclei were prepared as described in Appendix 1 in 10 mM Tris pH 7.5, 10 mM NaCl, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride. The nuclei were washed three times in 80 mM KCl, 40 mM K_2HPO_4, 40 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride, centrifuging at 1,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) in a Beckman JS 13.1 swing-out rotor each time. The supernatants were retained, containing chick erythrocyte protein nuclei extract 1 (cePNE1 proteins). The nuclei remained intact and were later used for histone isolation and purification, so the procedure was an important purification step for this in removing the cePNE1 proteins. The cePNE1 proteins themselves were dialysed into filtered 80 mM KCl, 80 mM KH_2PO_4 as buffer A in cation-exchange chromatography, with buffer B being

- Legend to Figure 3.3 continued-

filtered 1 M KCl, 80 mM K_2HPO_4.The cation-exchange resin was four 5-ml Sephadex SP columns (Hitrap, fast flow, G.E. Healthcare). The binding volume was four column volumes and the chromatography was monitored by measuring the elution solution A_{280}, the elution solution pH and the percentage of buffer B in the elution fractionation. The flow rate was 3 ml/minute and final elution conditions were 100% buffer B, pH 7.8. Seven peaks were observed in the chromatogram, including a non-bound fraction. 20% SDS-PAGE gels (Figure 3.4) were taken of all the fractions as defined in the above Figure. Some of the bands in the lanes of the SDS-PAGE (Figure 3.4) were identified by mass spectrometry, it being a major task to identify all the bands.

Fractions that were combined to represent the peaks in the chromatogram were as follows:

Sample 1: A2-A8

Sample 2: B12-B8

Sample 3: B7-B4

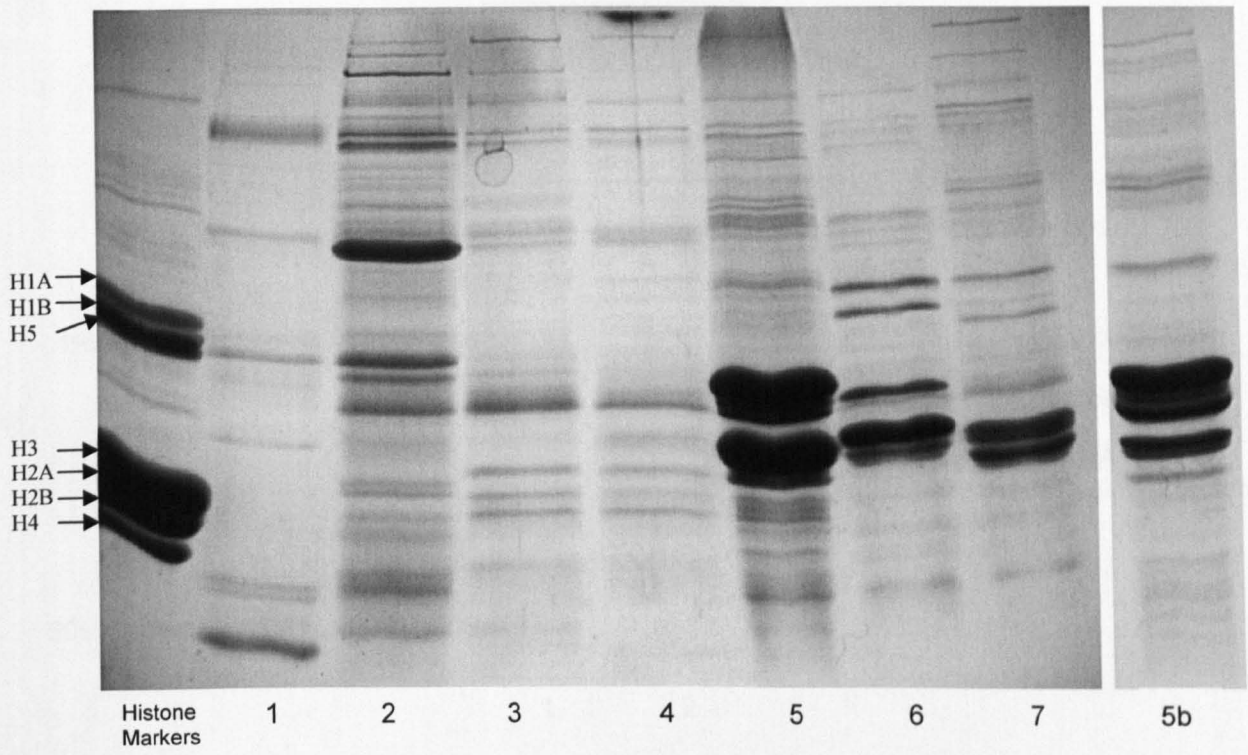
Sample 4: B3-B2

Sample 5: B1-C5

Sample 6: C6-C10

Sample 7: C11-D9

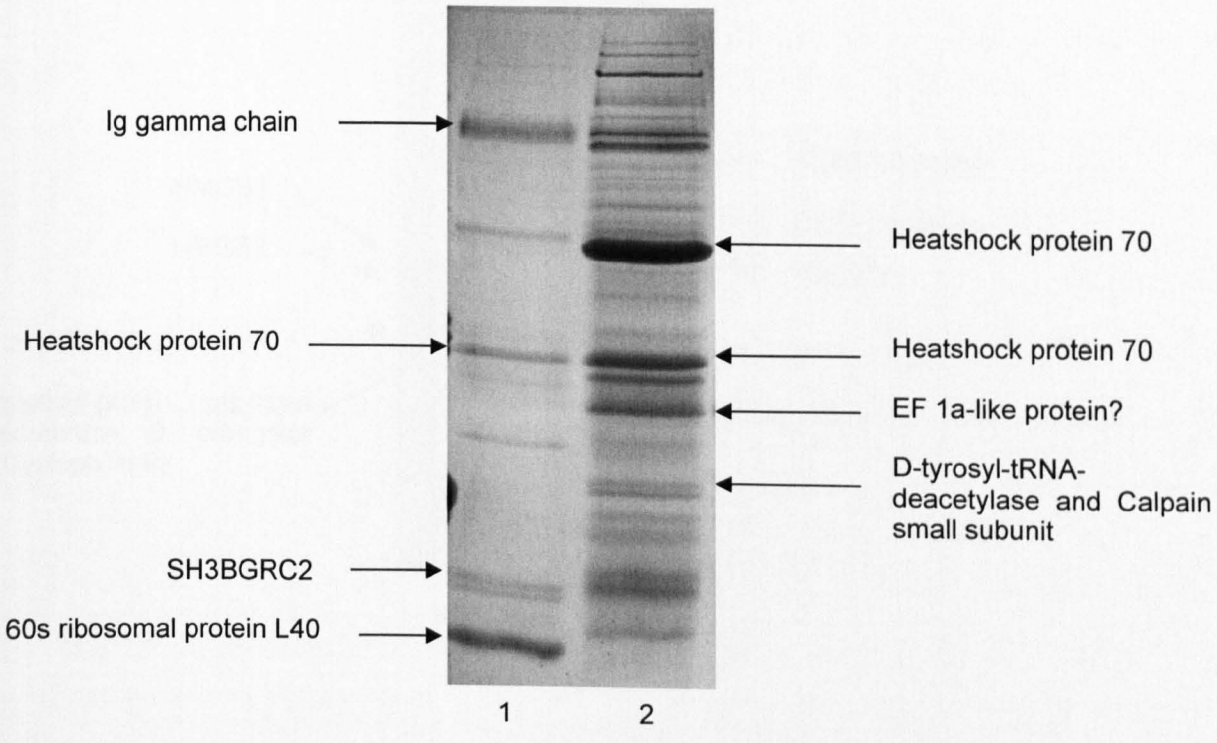
Figure 3.4a 20% SDS-PAGE of the proteins selected from cePNE1 cation-exchange chromatography (Figure 3.3)



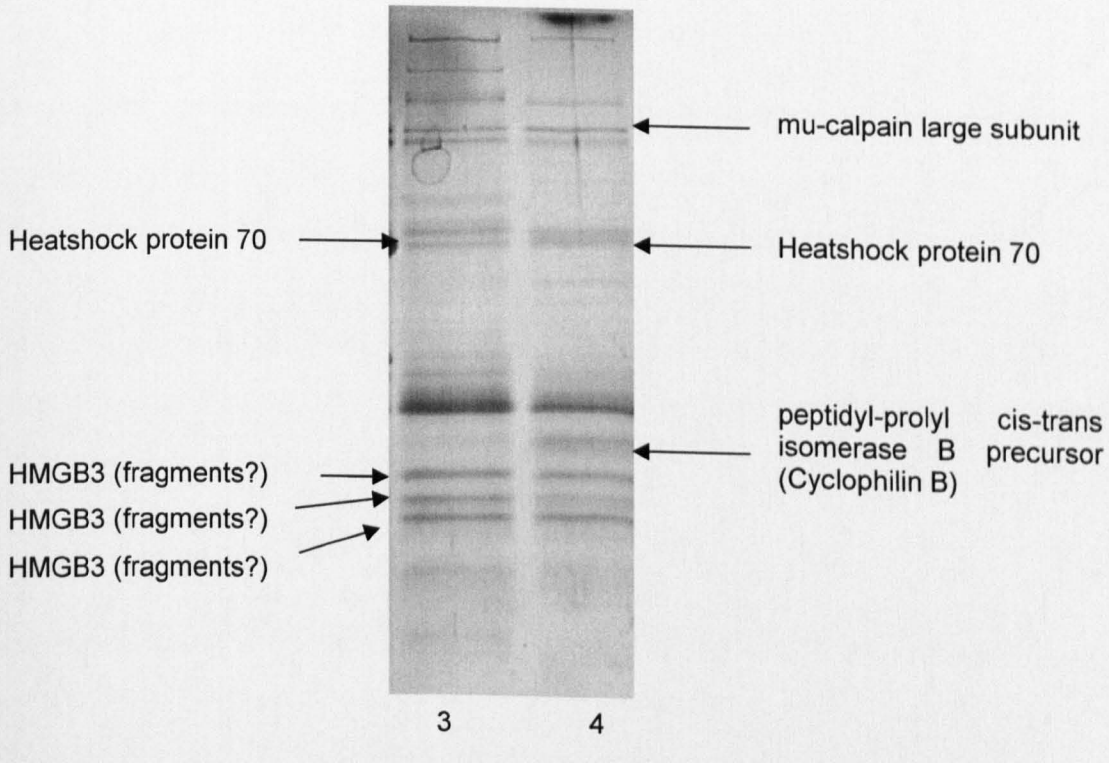
This is a 20% SDS-PAGE of the samples selected from the cation-exchange of cePNE1 proteins (Figure 3.3). Lanes 1 to 7 refer to peaks 1 to 7 in Figure 3.3 (lane 5b is a lower loading of the original lane 5 which was slightly overloaded). Some bands were cut out and sent for mass spectrometry, the results are shown in Figure 3.4b.

Figure 3.4b The bands which were cut from the gel shown in Figure 3.4a and analysed by mass spectrometry

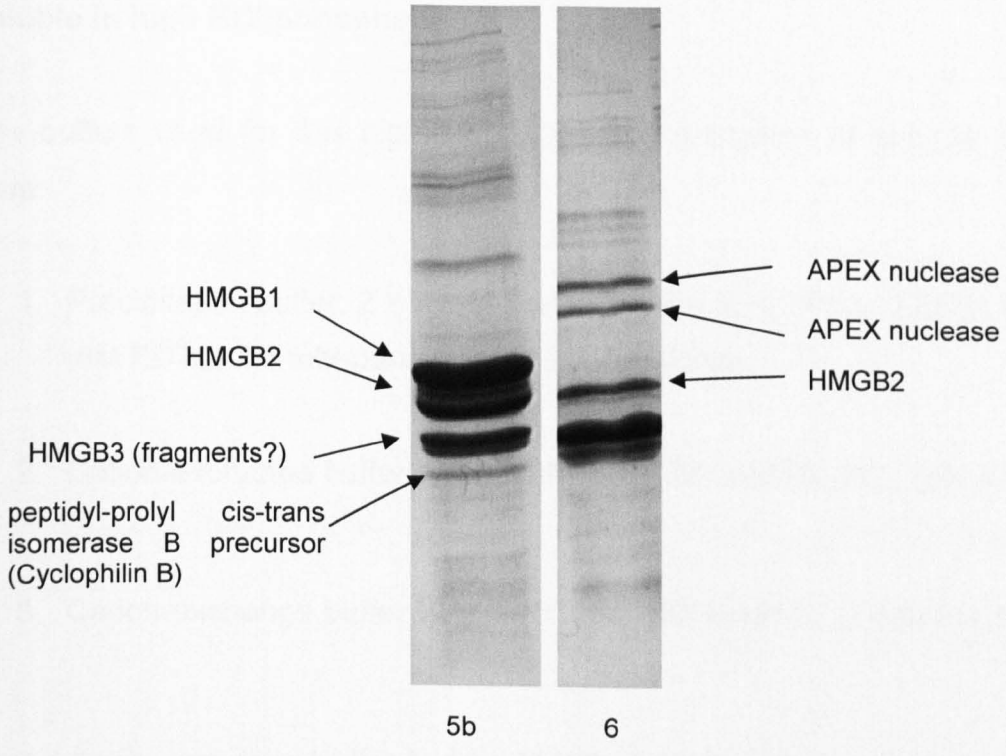
a.



b.



c.


HMGB1
HMGB2
HMGB3 (fragments?)
peptidyl-prolyl cis-trans isomerase B precursor (Cyclophilin B)
APEX nuclease
APEX nuclease
HMGB2
5b
6

3.2 Cation-exchange and anion-exchange chromatography of cePNE1 proteins soluble in high KCl/phosphate

The buffers used for this high KCl/phosphate extraction of soluble nucleoproteins were:

1. Precipitation buffer: 2 × 2 L of 2 M KCl, 0.65 M K_2HPO_4, 0.65 M KH_2PO_4, 5 mM EDTA, 2.5 mM benzamidine hydrochloride.

2. Cation-exchange buffer A: 100 mM KCl, 100 mM KH_2PO_4 (pH 4.5).

3. Cation-exchange buffer B: 2 M KCl, 92 mM K_2HPO_4, 8 mM KH_2PO_4 (pH 8).

4. Anion-exchange buffer A: 10 mM KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4 (pH8).

5. Anion-exchange buffer B: 1 M KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4 (pH8).

The experiment started with 40 ml of concentrated cePNE1 sample which was washed out from chick-erythrocyte cell nuclei with 100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4, 3.5 mM $MgCl_2$ and 2.5 mM benzamidine hydrochloride without lysing (section 3.1). The concentration of the sample was 11.2 mg/ml and the total volume was 40 ml. Total protein in this sample was 450 mg. The cePNE1 sample was then dialysed against precipitation buffer (2 M KCl, 0.65 M K_2HPO_4, 0.65 M KH_2PO_4, 5 mM EDTA and 2.5 mM benzamidine hydrochloride) overnight.

The dialysed cePNE1 sample was centrifuged at 12,000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 20 minutes to remove precipitation. The supernatant was taken as sample P2. The concentration was 6.13 mg/ml, the volume was 10 ml, and total protein was 61 mg. The majority of the total 450 mg had therefore precipitated, leaving the 61 mg of soluble proteins in the supernatant,

mostly HMG proteins.

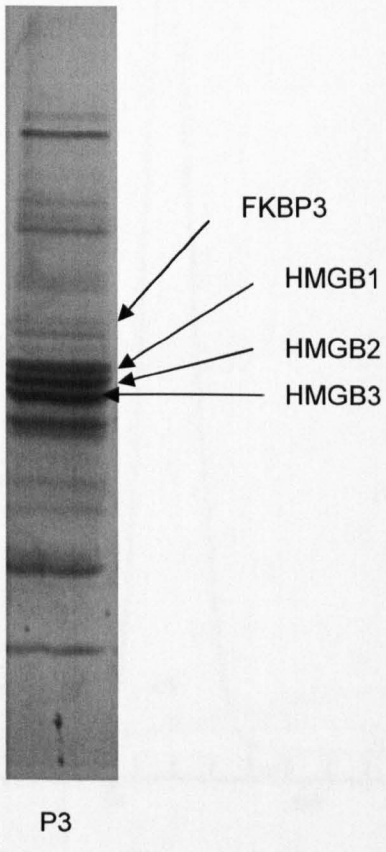
Sample P2 was then split into two parts each contained 5 ml labelled as sample P3 and P4. P3 was dialysed against cation-exchange buffer A (100 mM KCl, 100 mM KH_2PO_4, pH 4.5) and P4 was dialysed against anion-exchange buffer A (10 mM KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4, pH 8.0) overnight.

Cation-exchange chromatography of the high KCl/phosphate supernatant

After dialysis, UV spectrophotometry was taken for P3 as well as SDS-PAGE (Figure 3.5). The concentration was 2.6 mg/ml, the volume was 6.5 ml and total protein was 17 mg. The HMG proteins have gone from being a significant minority in the P1 sample to the majority in the soluble fraction of the P3 sample.

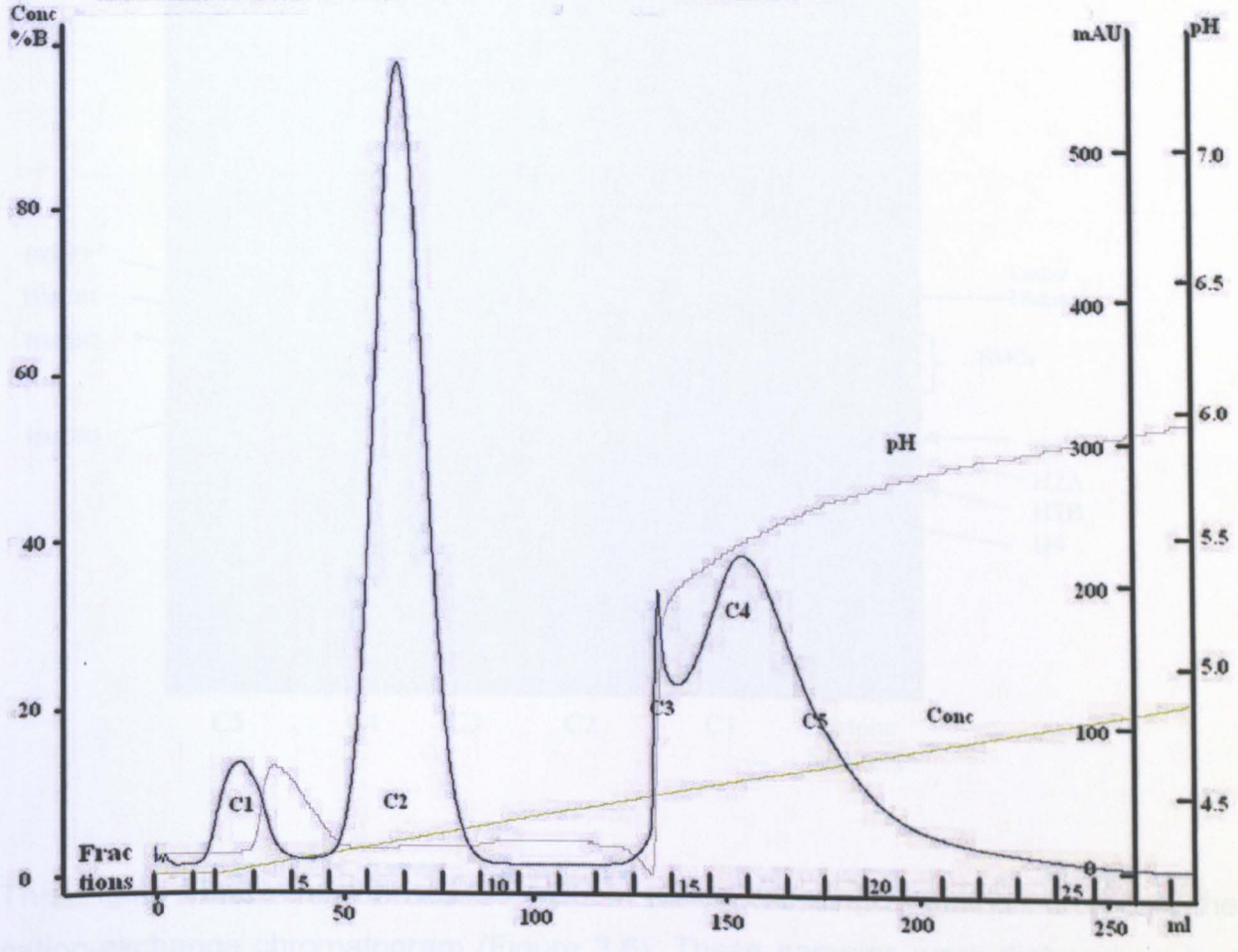
P3 was bound to four 5 ml SP-Sephacryl cation-exchange columns (G.E.Healthcare) with binding buffer (buffer A): 100 mM KCl, 100 mM KH_2PO_4, pH 4.5 and elution buffer (buffer B): 2 M KCl, 92 mM K_2HPO_4, 8 mM KH_2PO_4, pH 8.0. Five peaks were separated as samples C1-C5 (Figure 3.6). They were then dialysed against distilled water and freeze-dried for SDS-PAGE (Figure3.7). The UV spectrophotometry was taken for each sample, C1-C5, after dialysis against distilled water (Table 3.2).

Figure 3.5 20% SDS-PAGE of the sample P3, selected for cation-exchange chromatography from the soluble fraction of cePNE1



Some bands were cut from this gel and analysed by mass spectrometry. The results are shown in this Figure (bands identified by compared with the histone marker).

Figure 3.6 The chromatogram of sample P3 using cation-exchange chromatography



The colour code in this Figure refers to the A_{280} (blue), % buffer B (green) and pH (grey). The precipitated cePNE1 sample was dialysed against cation-exchange buffer A (100 mM KCl, 100 mM KH_2PO_4, pH 4.5) and loaded onto four series of 5 ml SP Fast Flow columns to run cation-exchange chromatography. The binding buffer was also buffer A and the elution buffer (buffer B) was 2 M KCl, 92 mM K_2HPO_4, 8 mM KH_2PO_4 (pH 8.0). The chromatography was run from 0% buffer B to 100% buffer B. Note that the sharp rise of at elution volume 135 mls may caused by blockage of pump B. The buffering was therefore not stabilised with phosphate at this pH.

Five peaks/shoulders were taken separately and labeled as sample C1 to C5:

C1: Fractions 3-4, 20ml

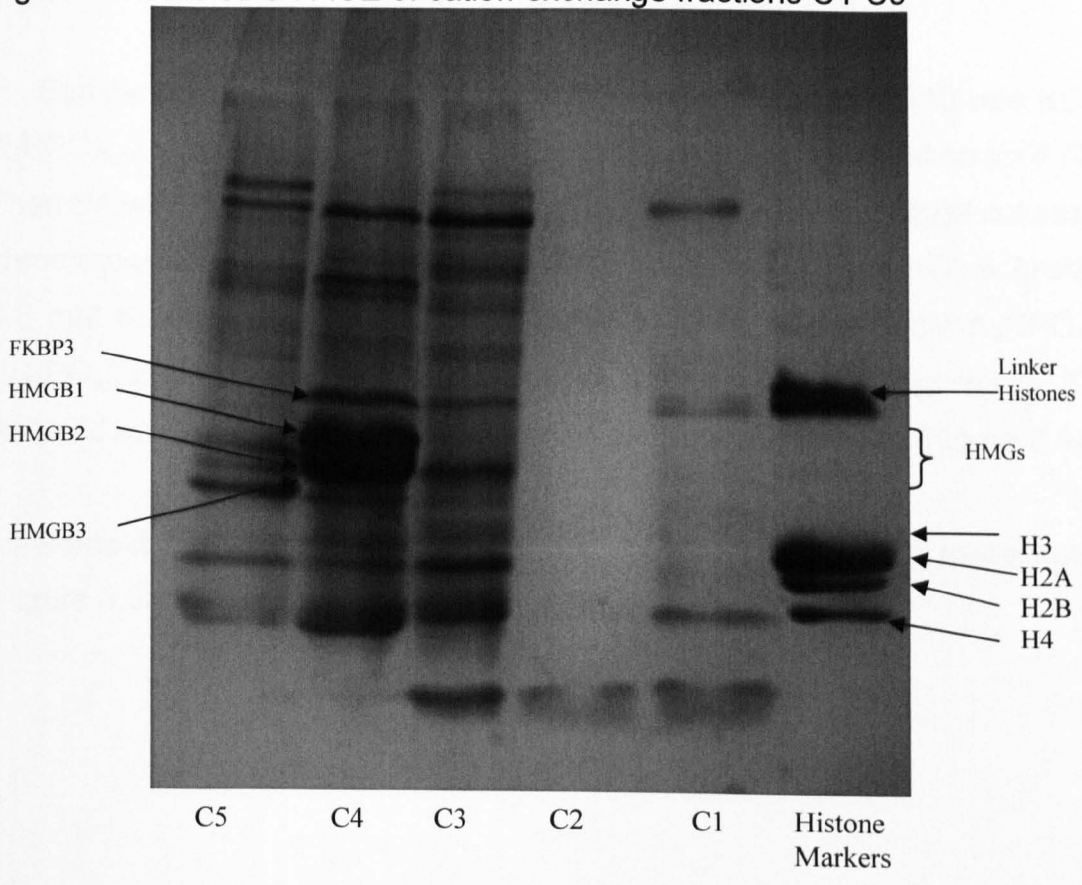
C2: Fractions 6-9, 40ml

C3: Fraction 14, 10ml

C4: Fractions 15-17, 30ml

C5: Fractions 18-20, 30ml

Figure 3.7 20% SDS-PAGE of cation-exchange fractions C1-C5



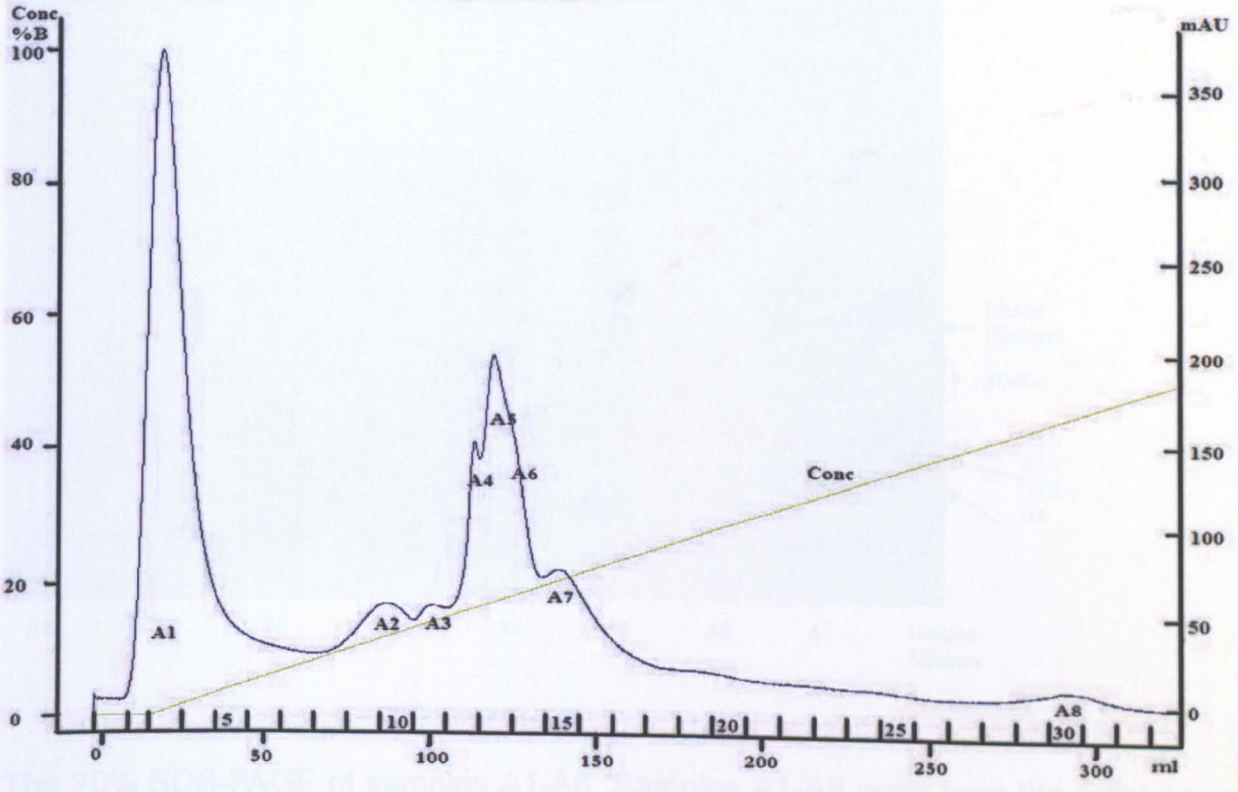
This Figure shows the 20% SDS-PAGE of cation-exchange fractions C1-C5 of the cation-exchange chromatogram (Figure 3.6). These samples were dialysed against distilled water overnight and then freeze-dried. Before the gel was run, a little bit of powder of each sample was taken and put into five 1ml tubes labelled C1-C5. 20 μl of 1.0 M Tris-HCl buffer (pH 6.8) was added to each tube. If the powder in the tube was not fully dissolved, another 20 μl of 1.0 M Tris-HCl buffer (pH 6.8) was added until the powder was just dissolved. Then the 3X Laemmli buffer with a half volume of the sample was added to each sample and they were mixed and denatured in 90°Cwater bath for 2 minutes (Appendix 3). After that 20 μl of each sample was loaded onto the gel. Note the strong HMG bands in gel lanes C4 and C5. The proteins in the C4 and C5 peaks are the majority of proteins in the supernatant after precipitation procedure (Figure 3.6 and Table 3.2) (bands identified by compared with the histone marker).

Anion-exchange chromatography of the high KCl/phosphate supernatant

Sample P4 was dialysed against anion-exchange buffer A (10 mM KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4) and an UV spectrophotometry was taken for it (Table 3.2). Then P4 was loaded onto four series of 5ml DEAE anion-exchange columns and the chromatography was run with binding buffer (buffer A): 10 mM KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4 and elution buffer (buffer B): 1 M KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4. Eight peaks were separated as samples A1-A8 (Figure3.8). They were dialysed against distilled water and freeze-dried for SDS-PAGE (Figure 3.9a).

Some bands from the gels were cut out and analysed by mass spectrometry (Figure 3.9a and b), as in the legend to Figure 3.9a.

Figure 3.8 The chromatogram of sample P4 using anion-exchange



The colour code in this Figure refers to the A_{280} (blue), % buffer B (green) and pH (grey).The precipitated cePNE1 sample was dialysed against anion-exchange buffer A (10 mM KCl, 9.2 mM KH_2PO_4, 0.8 mM KH_2PO_4, pH 8) and loaded onto four series of 5ml DEAE anion-exchange columns and run. The binding buffer (buffer A) of this anion-exchange was 10 mM KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4 (pH 8); the elution buffer (buffer B) was 1 M KCl, 9.2 mM K_2HPO_4, 0.8 mM KH_2PO_4 (pH 8). The chromatography was run from 0% buffer B to 100% buffer B.

Eight peaks were taken separately and labeled as sample A1 to A8:

A1: Fractions 2-5, 40ml

A2: Fractions 9-10, 20ml

A3: Fraction 11, 10ml

A4: Fraction 12, 10ml

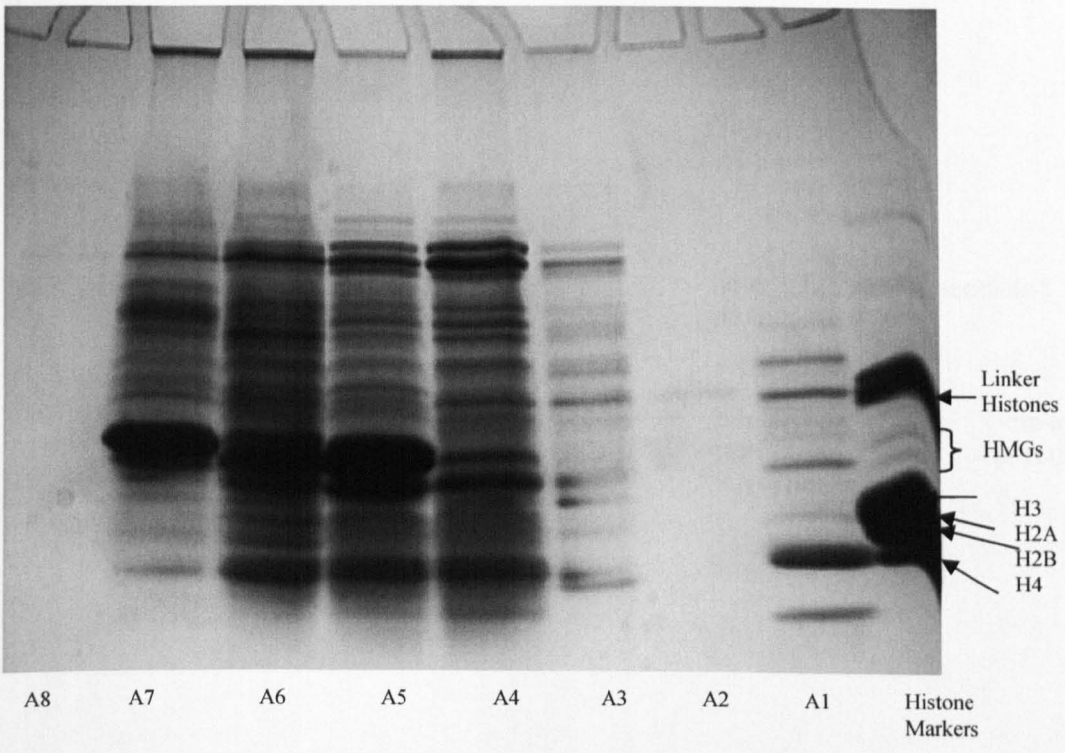
A5: Fraction 13, 10ml

A6: Fraction 14, 10ml

A7: Fractions 15-16, 20ml

A8: Fractions 30-31, 20ml

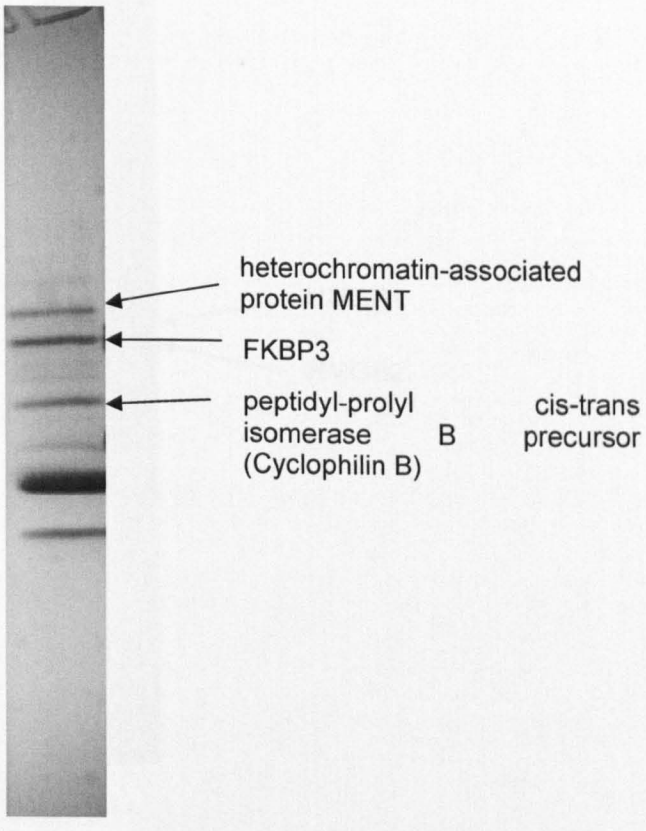
Figure 3.9a 20% SDS-PAGE of sample A1-A8



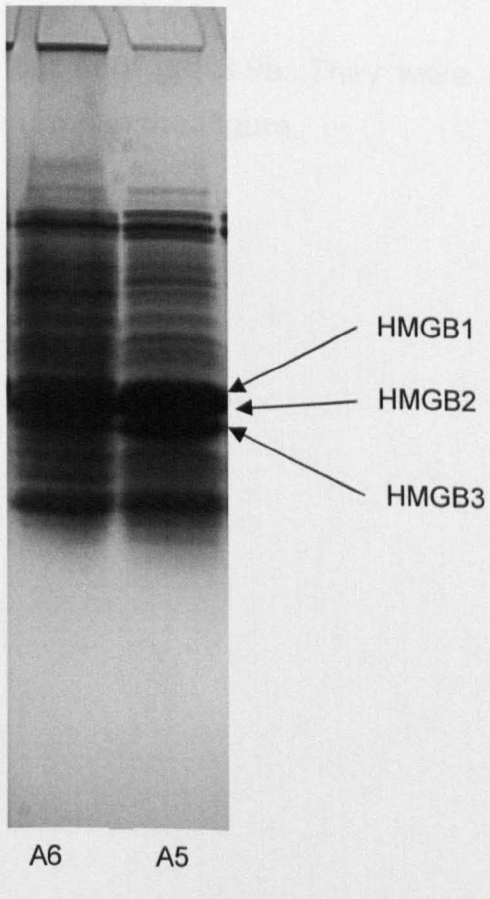
The 20% SDS-PAGE of samples A1-A8. Samples A1-A8 were from the eight peaks of anion-exchange chromatography (Figure 3.8). The samples were dialysed against distilled water overnight and then freeze-dried. They were treated the same as described in Figure 3.7 and 20 μl of each sample was loaded onto the gel.

Figure 3.9b The bands analysed by mass spectrometry from gel 3.9a

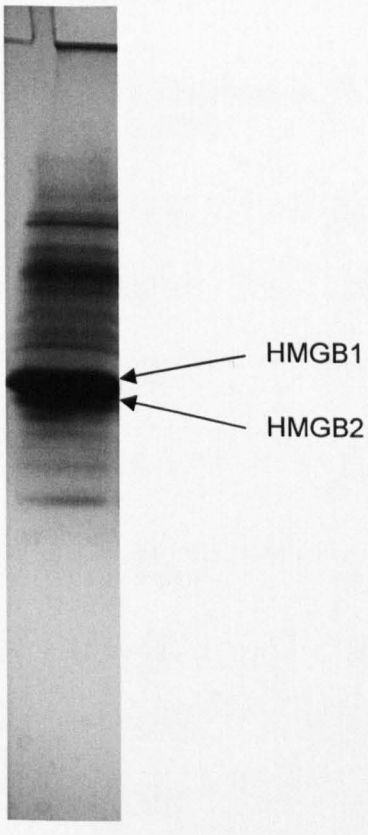
a.



b.



c.



A7

These Figures show the bands cut from gel 3.9a. They were analysed by mass spectrometry and the results are shown in the Figure.

Table 3.2 The yield of samples in each step of the experiment

Sample	Protein Description	Volume (ml)	Concentration (mg/ml)	Yield (mg)
P1	cePNE1	40	11.2	447
P2	Supernatant after precipitation	10	6.13	61.3
P3	Sample P2 in buffer A cation-exchange	6.5	2.6	17
P4	Sample P2 in buffer A anion-exchange	9	2	18
C1	Fraction 1, cation-exchange	23	0.043	1
C2	Fraction 2, cation-exchange	N/A	N/A	N/A
C3	Fraction 3, cation-exchange	N/A	N/A	N/A
C4	Fraction 4, cation-exchange	35	0.23	8.25
C5	Fraction 5, cation-exchange	35	0.17	5.93
A1	Fraction 1, anion-exchange	40	0.1	4
A2	Fraction 2, anion-exchange	20	0.02	0.4
A3	Fraction 3, anion-exchange	10	0.06	0.6
A4	Fraction 4, anion-exchange	10	0.15	1.5
A5	Fraction 5, anion-exchange	10	0.27	2.7
A6	Fraction 6, anion-exchange	10	0.16	1.6
A7	Fraction 7, anion-exchange	20	0.1	2
A8	Fraction 8, anion-exchange	20	0.044	0.89

3.3 Cross-Linking of cePNE1 proteins

After some initial studies of the cePNE1 proteins, we found that there are hundreds of different proteins in this group. The most important proteins in this group for our studies are the HMG proteins, and a surprising quantity of the FK-binding protein 3, a peptidyl-prolyl isomerase. As described in the earlier chapters, HMGs play very important roles in nucleosome remodelling, in order to contribute to the regulation of DNA replication, transcription and repair. All these processes require interactions between HMGs and other proteins. So it was considered important to try to cross link these protein assemblies in the cePNE1 group and identify the cross-linked proteins by gel electrophoresis and mass spectrometry.

Solutions:

a. 2 ml aliquots of cePNE1 proteins in 100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4.
b. Cross-linking buffer without EDC: 500 mls of 100 M KCl, 100 mM phosphate pH 7.5, by dilution from stock solution 3.
c. Cross-linking buffer with EDC: dissolve 0.5 gm of EDC in 20 ml of cross-linking buffer, and make up to 25 ml. Freeze at -80°C in 1 ml aliquots.

One 2 ml microfuge tube of cePNE1 proteins was thawed by leaving the tube in the cold room for one hour. The tube of cePNE1 proteins was then dialysed against cross-linking buffer (Working solution b) overnight. Meanwhile, the water bath was set at the correct temperature, 23 °C, and two electrophoresis gels were made up at the same time, one 20% acrylamide and one 12% acrylamide ready for loading, each with a 5% acrylamide stacking gel.

Eight labelled microfuge tubes were prepared, each with 10 µl of 3X Laemmli buffer (Stock solution 2).

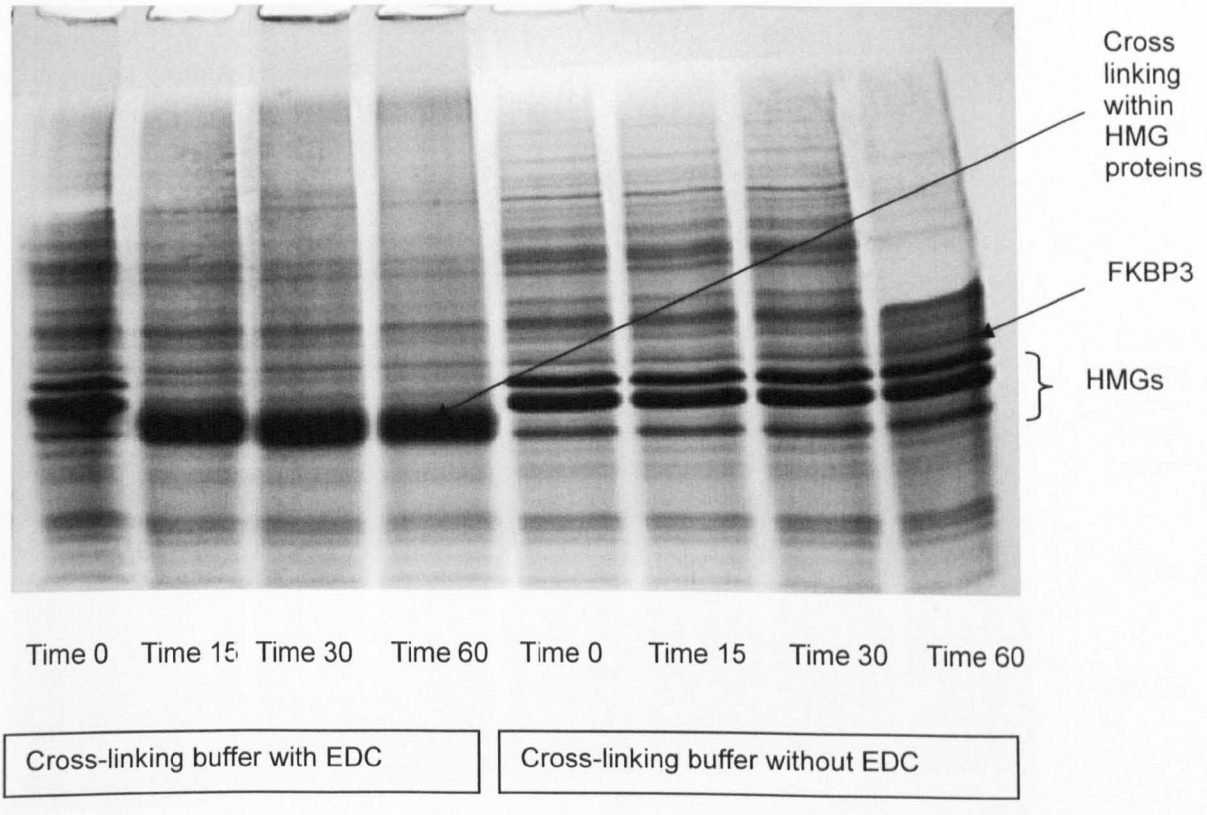
0.5 ml of the dialysed sample of cePNE1 proteins was mixed with 0.5 ml of cross-linking buffer without EDC (Working solution b) and was put into water-bath

temperature, 23 °C. 20 µl of mixed sample was taken at time 0 minute, 15 minutes, 30 minutes and 1 hour, immediately put into the 10 µl of hot 3X Laemmli buffers (the 3X Laemmli buffer was heated to 100 °C just before mixing). These set of samples are the control of the experiment.

0.5 ml of the dialysed sample of cePNE1 proteins was mixed with 0.5 ml of EDC cross-linking buffer (Working solution c) and was put into a 23 °C water-bath at the same time as the control. 20 µl from the sample with EDC buffer and 20 µl from the sample without EDC buffer were taken. This was done at time t = 0 minutes, 15 minutes, 30 minutes and 1 hour, immediately put into the 10 µl of 3X Laemmli buffer at 95 °C.

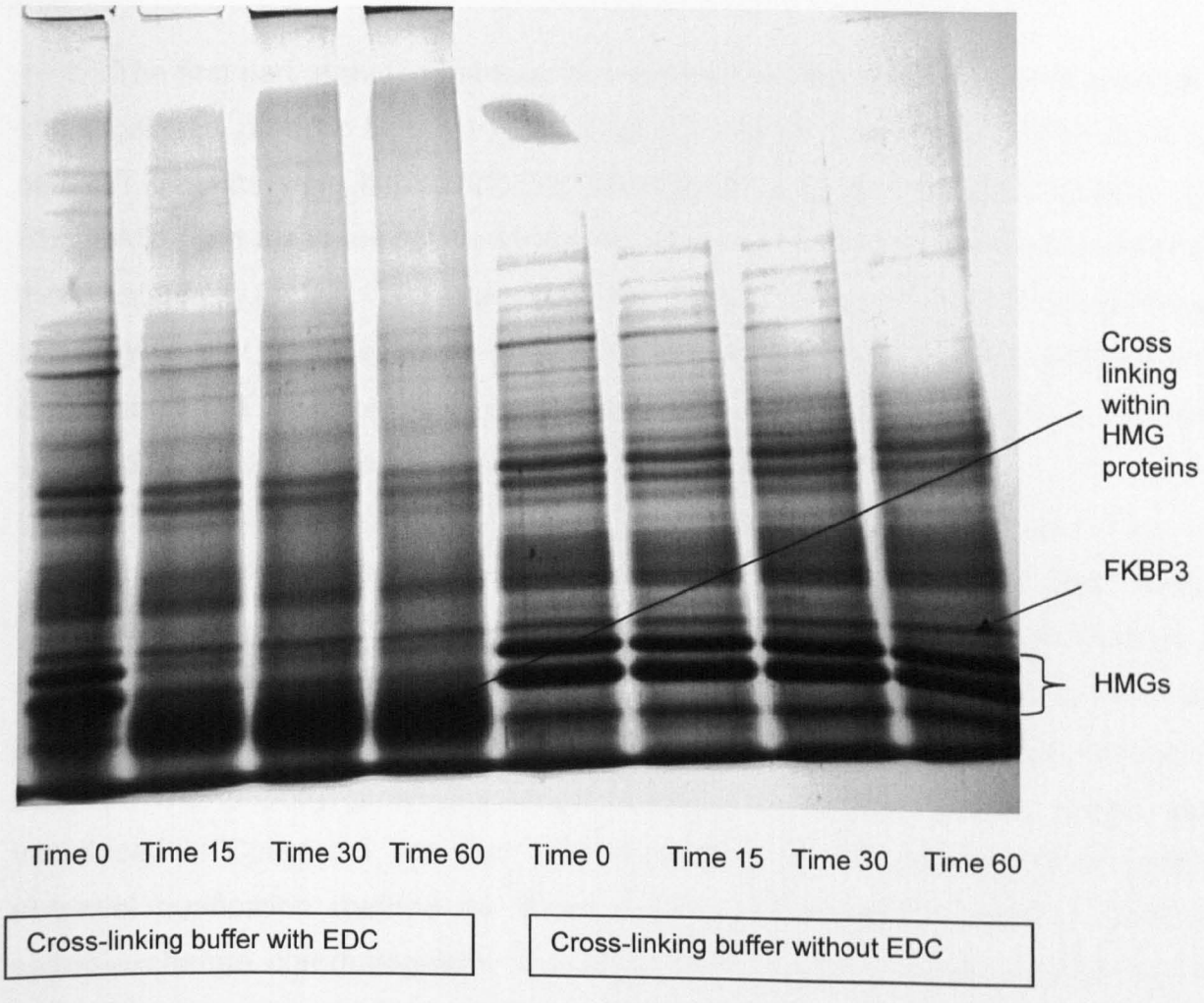
All the two sets of samples were then loaded onto the 20% and 12.5% electrophoresis gels and run (Figure 3.10a and Figure 3.10b).

Figure 3.10a 20% SDS-PAGE of cePNE1 cross linking



The 20% gel of cePNE1 cross-linking results. The lanes from left to right are: time 0.15 min, 30 min, and 1 hr with EDC; time 0, 15 min, 30 min, and 1 hr without EDC (bands identified by compared with the histone marker).

Figure 3.10b 12.5% SDS-PAGE of cePNE1 cross linking



The 12.5% gel of cePNE1 cross-linking results. The lanes from left to right are time 0, 15 min, 30 min, and 1 hr with EDC, time 0, 15 min, 30 min, and 1 hr without EDC (bands identified by compared with the histone marker).

Conclusion and discussion

The first part of the experiment in this Chapter describes the direct separation of the cePNE1 proteins by cation-exchange chromatography. In this experiment, the cePNE1 proteins were separated into several not very well separated peaks. The SDS-PAGE and mass spectrometry results show that a group of proteins containing the heatshock protein 70, eluted at the early stage of the cation-exchange chromatography. Then a group of proteins containing the important peptidyl-prolyl isomerase FKBP3 were eluted in peak/shoulder 3 and 4. Finally a group of proteins enriched in HMGBs were washed out in peak/shoulder 5 and 6.

Heatshock protein 70 is a highly conserved protein family with multiple fundamental functions. Their main functions include enhancing cell survival by its ability to catalyse the reassembly of damaged ribosomal protein (Pelham, 1984) and preventing proteins from degradation. These functions are important for cancer studies. The peptidyl-prolyl isomerase FKBP3 and HMG proteins which were introduced in Chapter 1 are also very important. So this experiment provided a potential purification method for these functionally important proteins. Also this cation-exchange chromatography for native chick nucleoproteins showed that the FKBP3 is not co-fractionated with any HMG proteins, which related to some issues about the work, did by Leclercq *et al.* (2000), and this is explained later in this conclusion.

The second part of this Chapter is a further experiment extended from the first part. The cePNE1 proteins were first precipitated in a high KCl/phosphate buffer, only the most soluble proteins remained in the solution. After that both cation-exchange and anion-exchange chromatography were applied. Unfortunately, some mass spectrometry results for this part of work have not been completed. From the SDS-PAGE and mass spectrometry results available, it seems that the major proteins remaining after precipitation are the HMGBs. This experiment may lead to some future work for extraction and studies of pure HMGB proteins.

The final part of this Chapter is an experiment of chemical cross-linking of the cePNE1 proteins. The method of cross linking was a valuable training in the control

of protein cross linking experiments and the use of controls of the PhD. of the Author. As mentioned earlier, the method used by Cato, *et al.,* (2008) was in the presence of Dithiothreitol (DTT). Therefore in their experiments, the disulphide bridges were not present. During the Author's cross linking experiment, it is seen that as time progressed the HMG protein bands disappeared and a new diffuse band was built up. This is described by Cato, *et al.,* (2008). The gels confirm the work of a colleague working on the "reverse technology" in our laboratory that the FKBP3 pepitidyl-prolyl isomerase did not cross link to any proteins. This is clear in both 20% and 12.5% SDS-PAGEs. There are publications (Leclercq *et al.*, 2000) showing that HMG proteins and FKBP3 run together in the pH dimension of a 2 dimensional electrophoresis gel (2DE). So these cross linking experiments did not confirm the presence of complexes of HMGs with FKBP3.

Some other studies that have been done in our laboratory, suggested that in nuclear extracts, such as the linker histone rich extract from human cell culture, show more cross linking information. And it should be noted that linker histones are not present in cePNE1 extract described in this Chapter. It has been pointed out by Stros & Kolibalova (1987) and Stros (2010) that cross-linking of nucleoproteins to extract molecular assemblies must be done with native proteins rather than recombinant proteins which are without their PTMs or with acid-extracted proteins.

Experiments have not been done in the presence of DTT as was done by Cato *et al.,* (2008), and it is not clear that cross linking experiments between proteins without disulphide bridges are meaningful for the contacts.

Chapter 4

Extraction and Studies of the cePNE4 and cePNE5 proteins from Chick-erythrocyte Nuclei by using the "Forward Technology"

The "Forward Technology", as described in Chapters 2 and 3, is a very quick and efficient method to separate nucleoproteins into small groups. Each of these groups of proteins is enriched in some important nucleoproteins such as linker histones, core histones, HMG proteins, etc. After separation and further purification, studies can be applied to each group (named as cePNE1, 2, 3, 4, 5, etc.).

In this Chapter, some studies of these groups of proteins, which have been done by the Author in collaboration with project students, are described. The Author was pleased for the opportunity to share some research projects with students undergoing projects as a part of their honours degree programs (Mr Damien Long & Mr Arabarb Mirza), a taught MSc program (Mr Zhao) and new PhD program (Mr Hugh Smallman). I considered it is very useful to gain experience in training others in research methods. The projects of this Chapter were part of this work and include firstly, the fractionation and characterisation of high-molecular-weight proteins soluble with histone octamers in 2 M KCl, 0.4 M equimolar phosphate (section 4.1) and secondly the analysis of proteins that were still soluble after the procedure of Chapter 2, where the histone octamers were precipitated (section 4.2).

.

4.1 Extraction of High Molecular Weight proteins (cePNE4) from chick nuclei

During the preparation of histone octamers, dimers and tetramers, histone octamers were prepared and dissolved in 2 M KCl, 0.4 M equimolar phosphate and loaded onto the S100 size-exclusion column. The histone octamers should be eluted first and then any linker histones, because linker histones are smaller than histone octamers. However, when we ran the SDS-PAGE for the fractions of the gel exclusion we found that there were some proteins eluted out before the main peak for histone octamers. That might be because some linker histones form complexes with other proteins to give them an increased molecular weight or that some high-molecular-weight proteins themselves were eluted. So it was decided to do the following experiment to try to work out which proteins may form complexes with linker histones and also which proteins fractionate before the histone octamers in this way. Because of the presence of high-molecular-weight proteins in the cePNE1 extraction (Figure 3.1), these proteins were left in the nuclei before lysis so that they could be extracted with the histone-octamer lysis procedure. That is to say we leave out the nuclei in 100 mM KCl, 100 mM equimolar phosphate described earlier in the thesis. So we extract pure chick-erythrocyte nuclei in 2 M KCl, 0.4 M equimolar phosphate. The procedures and results are given below:

Buffers:

(add benzamidine hydrochloride freshly, keep at 4°C before use)

1. Nuclei washing buffer: 1 L of 10 mM KCl, 5 mM K_2HPO_4, 5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride.

2. Nuclei Lysing buffer: 400 ml of 2.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride.

3. Size-exclusion running buffer: 2 L of 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4.

Procedure:

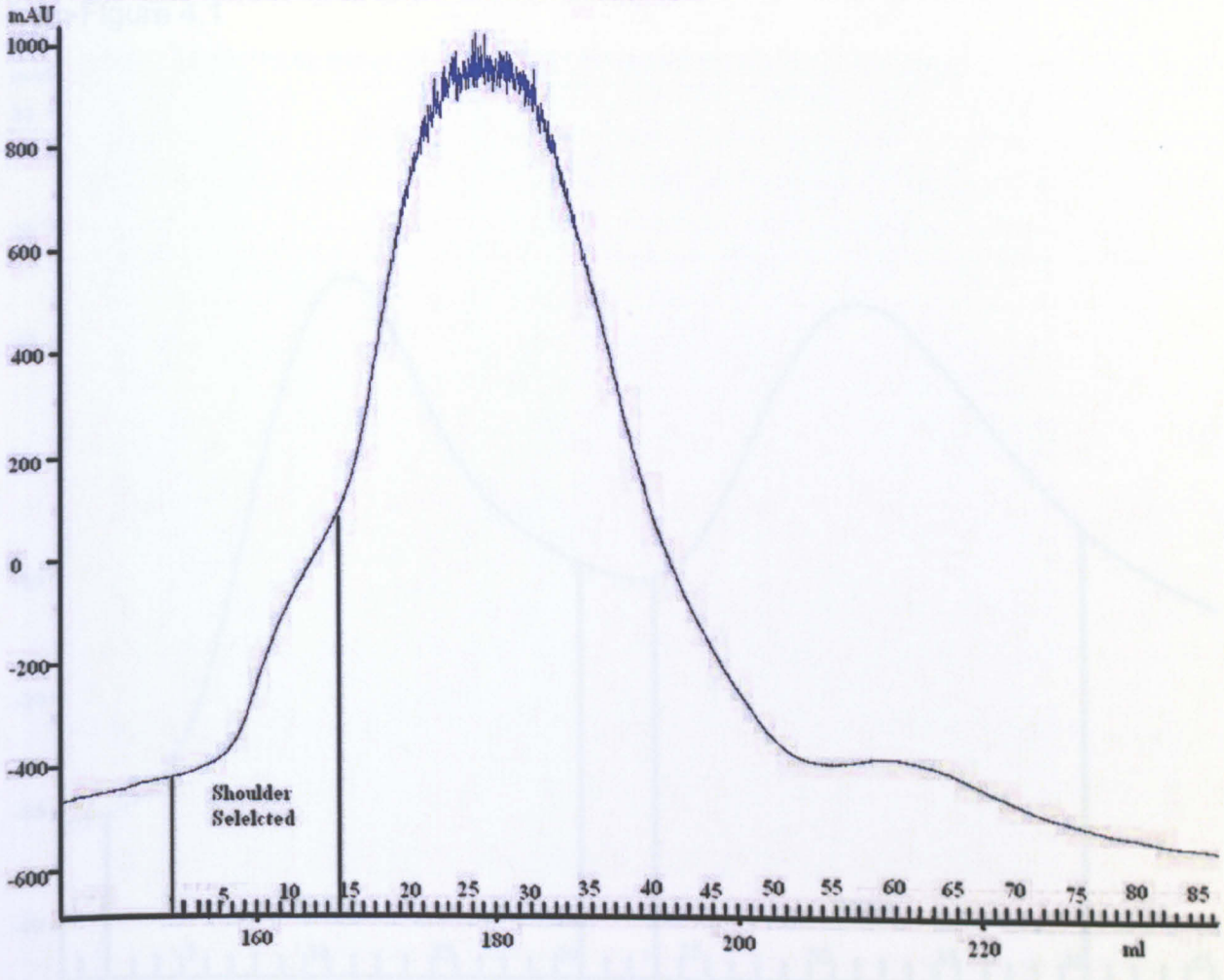
120 ml of pure chick erythrocyte nuclei (Appendix 1) were thawed in the cold room. After thawing, the chick nuclei were partitioned into four 50 ml centrifuge tubes, balanced, and centrifuged at 2,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 20 minutes. The supernatant was discarded and the nuclei pellets were mixed with an equal volume of nuclei washing buffer (10 mM KCl, 5 mM K_2HPO_4, 5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride), dispersed by sucking in and out the suspension of nuclei with a submerged plastic Pasteur pipette. Then the suspension was made up to 75 ml with nuclei washing buffer, stirring in a beaker on ice. 60 ml of cold nuclei lysing buffer (2.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride) was put into each of six Beckman Ti 45 centrifuge tubes on ice. 12.5 ml of nuclei sample were carefully layered on the top of the nuclei lysing buffer in each centrifuge tube using a syringe with 22-gauge needle. The centrifuge tubes were then capped and quickly shaken to mix up the nuclei and lysing buffer, making a uniform gel. The centrifuge tubes were then balanced and ultra centrifuged at 36,000 (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 45 rotor) (100,000 xg) for 16 hours.

After centrifugation, the DNA pellets were kept, the supernatant was concentrated down to 10 ml. The sample was then dialysed against gel exclusion running buffer (2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4) overnight and loaded onto an S100 column to run (Figure 4.1).

The chromatogram shows a main peak and a clear shoulder before the main peak (Figure 4.1). The shoulder was collected and concentrated down to 10 ml and re-loaded to the S100 column for another size-exclusion under the same conditions as above. This time two clear peaks were given (Figure 4.2). Samples collected from the two peaks were dialysed against distilled water and freeze-dried. Then they were loaded onto a 20% SDS-PAGE gel. The first peak contains some high molecular weight proteins and the second peak contains mainly histone proteins (Figure 4.3a & 4.3b).

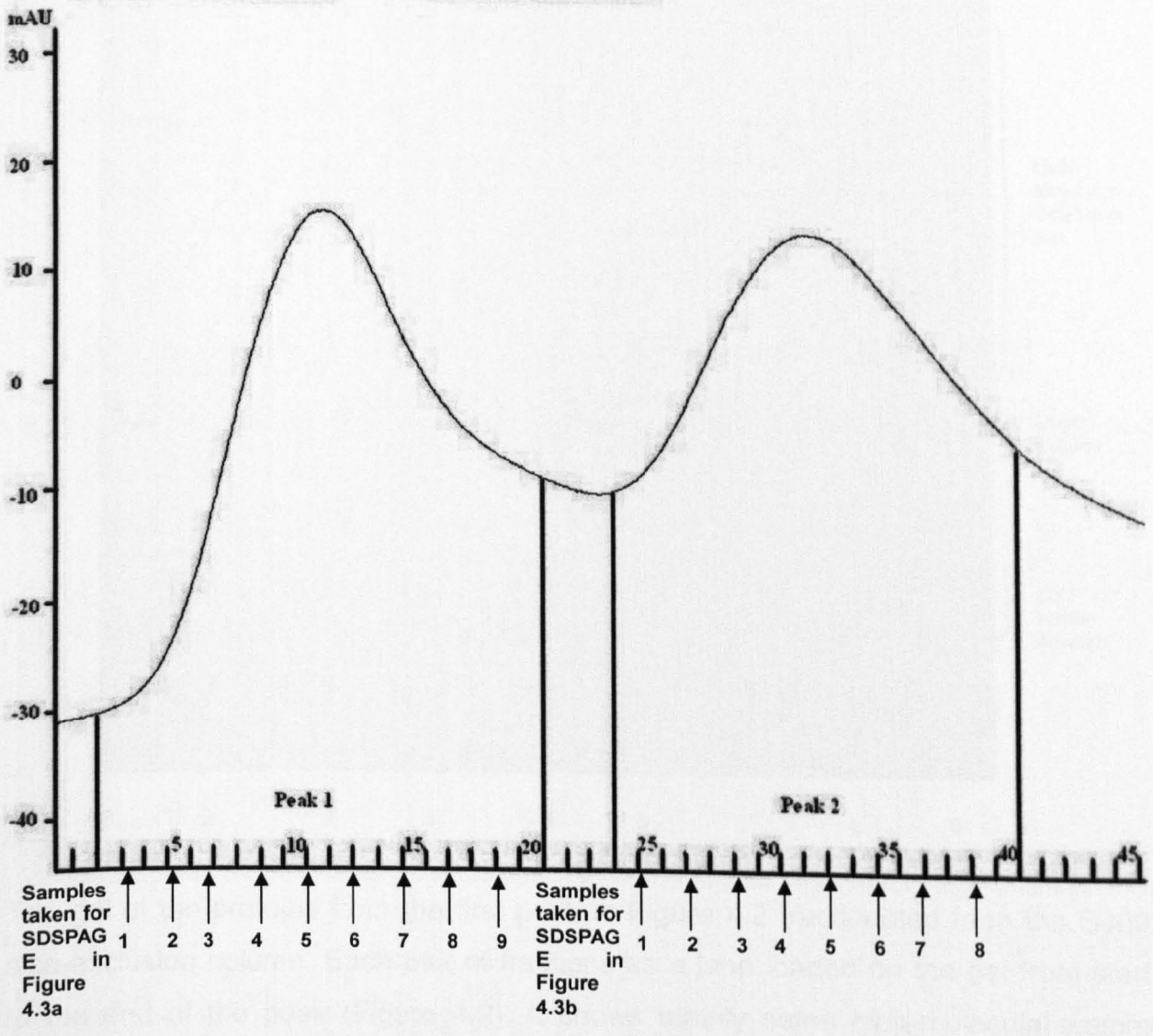
The sample from the first peak was sent to Sheffield University to be analysed by mass spectrometry (Table 4.1).

Figure 4.1 S100 Size-exclusion Chromatogram of chick erythrocyte nucleoproteins



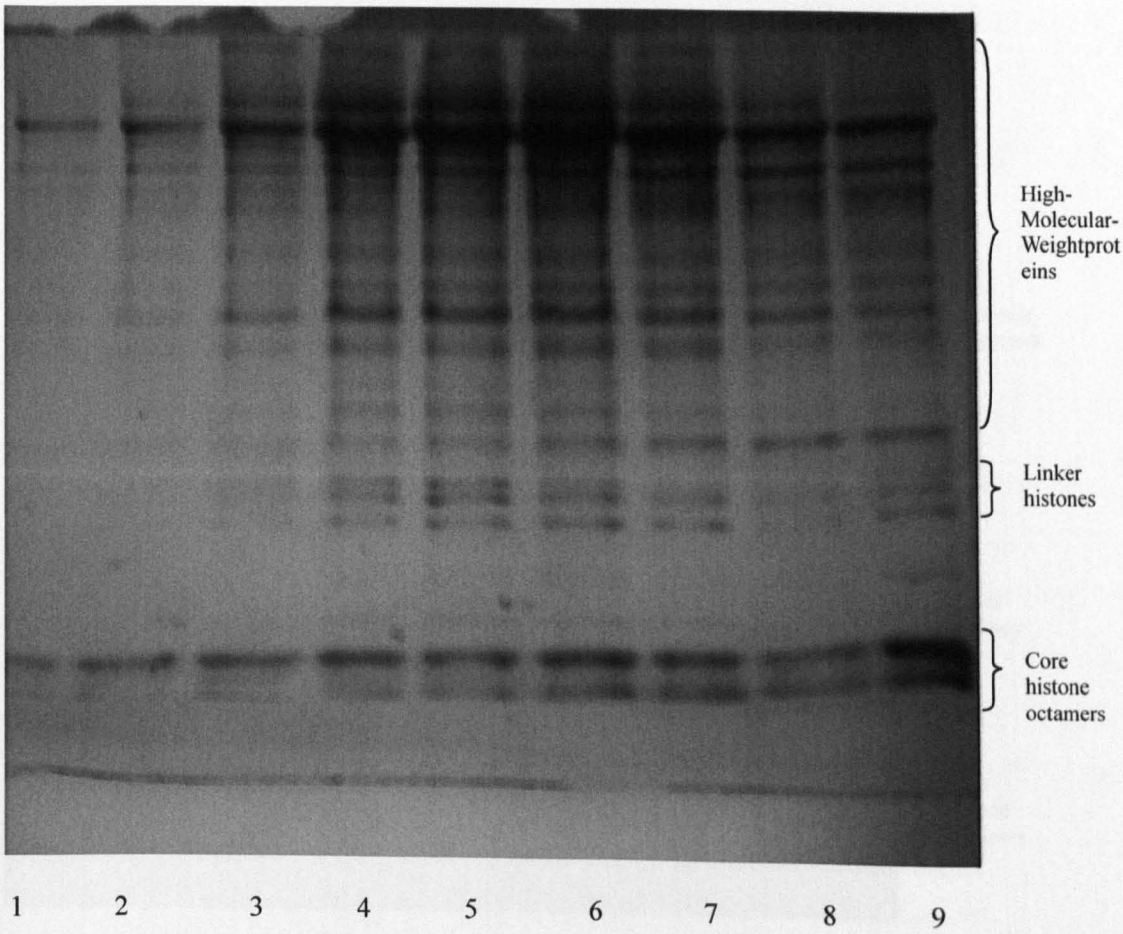
This Figure shows the fractionation by size-exclusion chromatography of the lysate of chicken-erythrocyte nuclei. The chick erythrocyte nuclei were lysed by 2.5 M KCl, 0.25 M K_2HPO_4, 0.25 M KH_2PO_4, 2.5 mM benzamidine hydrochloride buffer and ultracentrifuged at 100,000 xg for 16 hours. The supernatant were collected and concentrated down to 10 ml,then dialysed in 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4 and loaded onto a S100 size-exclusion column to run with 2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4 at 0.6 ml/minute flow rate for 500 ml (1.5 column volume). The shoulder before the main peak fractions (1-14) was selected and loaded on the gel exclusion column again under the same condition.

Figure 4.2 The chromatogram of the second gel exclusion using the sample selected from Figure 4.1



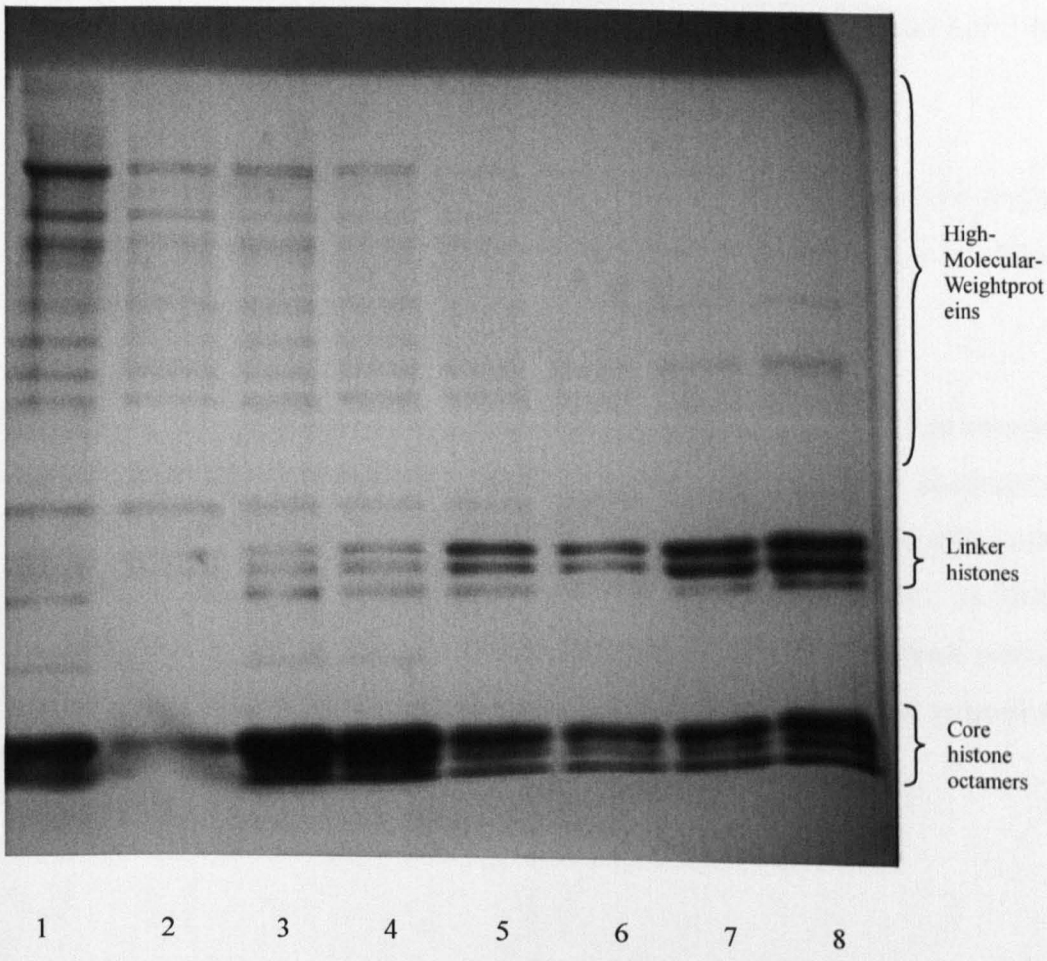
The shoulder selected from the first gel exclusion (Figure 4.1) was re-loaded onto the S100 column again under the same condition as described in Figure 4.1. The shoulder of first gel exclusion was further separated into two peaks. These samples was selected and loaded on SDS-PAGE with each pair of fractions as a lane on the gel. The results are shown in Figure 4.3a and Figure 4.3b.

Figure 4.3a 20% SDS-PAGE of proteins from peak 1 (cePNE4) of Figure 4.2



The gel of the proteins from the first peak in Figure 4.2 fractionated from the S100 size-exclusion column. Each pair of fractions as a lane loaded on the gel from start to the end of the peak (Figure 4.2). It shows mainly some high-molecular-weight proteins in this peak. Some linker histone and histone octamer proteins remain in the sample but these are minimal compared with the high-molecular-weight proteins.

Figure 4.3b 20% SDS-PAGE of proteins from peak 2 of Figure 4.2



The gel of the second peak fractionated from S100 gel exclusion column. Each pair of fractions as a lane loaded on the gel from start to the end of the peak (Figure 4.2). It shows that the high molecular weight proteins are decreasing and mainly histone proteins are in this peak.

The results of mass spectrometry are in Table 4.1 where the proteins of fractions 2-20 in Figure 4.2 were combined and freeze-dried and sent for electrospray mass spectrometry at Sheffield University (Ms Helen Phillips and Dr. Mark Dickman). The analysis was a total result for the combined proteins, not for selected band from the electrophoresis.

The molecular mass of each protein was given in the second column in the Table 4.1. Therefore some indication of where the proteins are positioned for example in lane 6 of Figure 4.3a.

The initial analysis of the mass spectrometry identified many "unnamed gallus gallus protein product". The Author therefore entered the peptide sequences from these products and entered them into the NCBI BLAST-search routine comparing with the NCBI gallus gallus data bank. In doing this, three further proteins were identified. The first 12 proteins in the list therefore are clearly identified and the final column of the table gives a short description of the function of the proteins (Table 4.1).

Table 4.1 The mass spectrometry results of the proteins in the first peak of the second gel exclusion run (cePNE4) (Figure 4.3a)

Protein	Mass	Score	Queries matched	emPAI	Description	Function
gi|29837126	142904	2988	154	4.85	SMC1 protein cohesin subunit [Gallus gallus]	Regulation of the separation of sister chromatids during cell division
gi|45383139	141482	2493	168	4.95	chondroitin sulfate proteoglycan 6 (bamacan) [Gallus gallus]	SMC3, regulation of the separation of sister chromatids during cell division
gi|63711	75594	1240	75	4.27	nucleolin [Gallus gallus]	A eukaryotic nucleolar phosphoprotein, is involved in the synthesis and maturation of ribosomes
gi|45384180	183131	800	28	0.27	topoisomerase (DNA) II beta 180kDa [Gallus gallus]	Cuts both strands of the DNA helix simultaneously in order to manage DNA tangles and supercoils
gi|53128417	79976	581	35	1.24	FACT complex subunit SSRP1 [Gallus gallus]	Affects eukaryotic RNA polymerase II (Pol II) transcription elongation
gi|134800	285188	418	28	0.22	Spectrin alpha chain, brain (Spectrin, non-erythroid alpha chain) (Fodrin alpha chain)	An actin crosslinking and molecular scaffold protein that links the plasma membrane to the actin cytoskeleton
gi|60302806	152988	387	28	0.35	similar to KIAA0648 protein [Gallus gallus]	
gi|50757376	287699	374	26	0.17	PREDICTED: similar to spectrin alpha chain, brain - chicken [Gallus gallus]	An actin crosslinking and molecular scaffold protein that links the plasma membrane to the actin cytoskeleton
gi|60098709	143760	344	34	0.29	sister chromatid cohesion protein PDS5 homolog A [Gallus gallus]	Regulator of cohesion maintenance
gi|50810565	56972	232	12	0.31	PREDICTED: similar to type II alpha-keratin IIC [Gallus gallus]	Abundant in keratinocytes in the cornified layer of the epidermis
gi|5074599	186901	211	13	0.10	PREDICTED: similar to THO complex subunit 2 (Tho2) [Gallus gallus]	part of the TREX (transcription/export) complex

gi|45382553	134857	184	12	0.23	SMC2 structural maintenance of chromosomes 2-like 1 [Gallus gallus]	A subunit of condensin I and II, large protein complexes involved in chromosome condensation
gi|46399073	62423	164	8	0.10	type II alpha-keratin IIA [Gallus gallus]	
gi|60302824	36633	112	3	0.09	heterogeneous nuclear ribonucleoprotein H3 (2H9) [Gallus gallus]	
gi|5080677	84352	98	6	0.08	PREDICTED: similar to keratin 6 irs3 [Gallus gallus]	
gi|50748121	153598	96	4	0.06	PREDICTED: similar to KIAA0097 gene product; colonic and hepatic tumor over-expressed [Gallus gallus]	
gi|50760853	96952	81	3	0.07	PREDICTED: similar to type I hair keratin KA31 [Gallus gallus]	
gi|50728598	20946	80	5	0.15	PREDICTED: similar to H2B histone family, member F [Gallus gallus]	
gi|50745996	188082	72	5	0.05	PREDICTED: similar to hypothetical protein [Gallus gallus]	
gi|15636898	119483	72	3	0.05	formin binding protein 11-related protein [Gallus gallus]	
gi|63448	13946	70	4	0.52	unnamed protein product [Gallus gallus]	
gi|50728590	49063	68	4	0.13	PREDICTED: similar to germinal histone H4 gene [Gallus gallus]	
gi|45383996	32612	60	5	0.32	nucleophosmin [Gallus gallus]	
gi|50748796	27914	58	2	0.11	PREDICTED: similar to Mthfd1-prov protein [Gallus gallus]	
gi|571460	90179	54	2	0.07	protein kinase [Gallus gallus]	
gi|7189561	100744	48	4	0.06	serine/arginine repetitive matrix 1 [Gallus gallus]	

gi|50760887	82798	48	3	0.04	PREDICTED: similar to keratin K12 [Gallus gallus]	
gi|50749292	314115	47	10	0.04	PREDICTED: similar to hypothetical protein FLJ10839; cell-cycle and apoptosis regulatory protein 1 [Gallus gallus]	
gi|50734049	199324	47	5	0.02	PREDICTED: similar to Desmoplakin (DP) (250/210 kDa paraneoplastic pemphigus antigen) [Gallus gallus]	
gi|5073108	342733	45	3	0.01	PREDICTED: similar to ATM [Gallus gallus]	
gi|50748318	43531	45	3	0.07	PREDICTED: similar to Arginase II, mitochondrial precursor (Non-hepatic arginase) (Kidney-type arginase) [Gallus gallus]	
gi|50738743	84848	43	2	0.04	PREDICTED: similar to KIAA1469 protein [Gallus gallus]	
gi|71896647	124068	42	2	0.03	ubiquitin protein ligase E3C [Gallus gallus]	
gi|5074908	249733	42	3	0.01	PREDICTED: similar to ninein isoform 2; GSK3B-interacting protein; ninein centrosomal protein; glycogen synthase kinase 3 beta-interacting protein [Gallus gallus]	
gi|50757442		42	2		PREDICTED: similar to cis-Golgi matrix protein GM130 [Gallus gallus]	
gi|50758915	192581	42	1	0.02	PREDICTED: similar to cadherin 22 precursor; ortholog of rat PB-cadherin [Gallus gallus]	

gi|50748243	32054	41	3	0.21	PREDICTED: similar to RIKEN cDNA 2610015J01 [Gallus gallus]	
gi|50740806	89979	40	1	0.03	PREDICTED: similar to kinesin-related protein 3A [Gallus gallus]	
gi|50759167	203960	39	1	0.02	PREDICTED: similar to MEGF6 [Gallus gallus]	
gi|50756669	87623	38	3	0.04	PREDICTED: similar to ring finger protein 10 [Gallus gallus] describtion	

Conclusions and discussion

By using the method described in this section, a group of native proteins with a relatively high molecular weight are extracted. The functions of the proteins with highest scores in this group are given in Table 4.1. These include: SMC1 and SMC3, subunits of the cohesin complex, which regulate the separation of sister chromatids during cell division; nucleolin, which is involved in the synthesis and maturation of ribosomes; topoisomerase (DNA) II, which cuts both strands of the DNA helix simultaneously in order to manage DNA tangles and supercoils; FACT complex subunit SSRP1, which affects eukaryotic RNA polymerase II (Pol II) transcription elongation; etc.

These are all fundamental proteins with important functions. The method described in this section,is able to produce a significant amount of these proteins in their native state within less than 48 hours. This gives a potential to introduce further purification and studies of these proteins.

4.2 Extraction and studies of cePNE5 proteins

During the extraction and purification of chick erythrocyte histone octamers, dimers and tetramers, another group of proteins produced is the cePNE5. This group of proteins is extracted during the last stage of precipitation of pure histone octamer (chapter 2). After centrifugation, the histone octamers precipitate out as a pellet, meanwhile the supernatant are collected as the cePNE5 proteins.

Previous work in the LJMU Group concerned proteins that were soluble in high KCl/phosphate solutions, above precipitation conditions for histone octamers (Snijders *et al.*, 2008). Some cation-exchange-chromatography fractionation of a similar protein extract produced in one step by lysing total chick nuclei (without removing cePNE1 proteins) was carried out by Mr Larus Foulger (Foulger L.F. Ph.D. thesis, submitted). The Author took part in some of this work that was carried out by sweeping the elution-buffer pH from a low value to pH 8.0.

Work was undertaken on the group of proteins of Snijders *et al.* (2008) with Mr Hugh Smallman who was at an early stage of his Ph.D. The objectives were to carry out the cation-exchange chromatography at a stable pH of 6.8 and to identify proteins that co-eluted with the majority proteins in each peak of the chromatogram.

The first part of the procedure is described in Chapter 2, where core-histone-enriched protein was precipitated in 2 M KCl, 1.5 M equimolar phosphate, pH 6.8. About 80 mg of precipitated protein, pelleted after ultracentrifugation, was dissolved in "octamer maintenance buffer" for size-exclusion chromatography to produce about 30 mg of pure histone octamers (chapter 2).

The supernatant from the ultracentrifugation (cePNE5) contained about 100 mg of protein and was dialysed into 0.1 M KCl, 0.1 M equimolar phosphate pH 6.8 for analysis in the present work with Mr Smallman.

Four 5-ml SP sephadex, HITRAP, Fast-flow cation-exchange columns were equilibrated with buffer A (100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4). The 40-ml sample of cePNE5 was introduced into the loading loop of the AKTA liquid

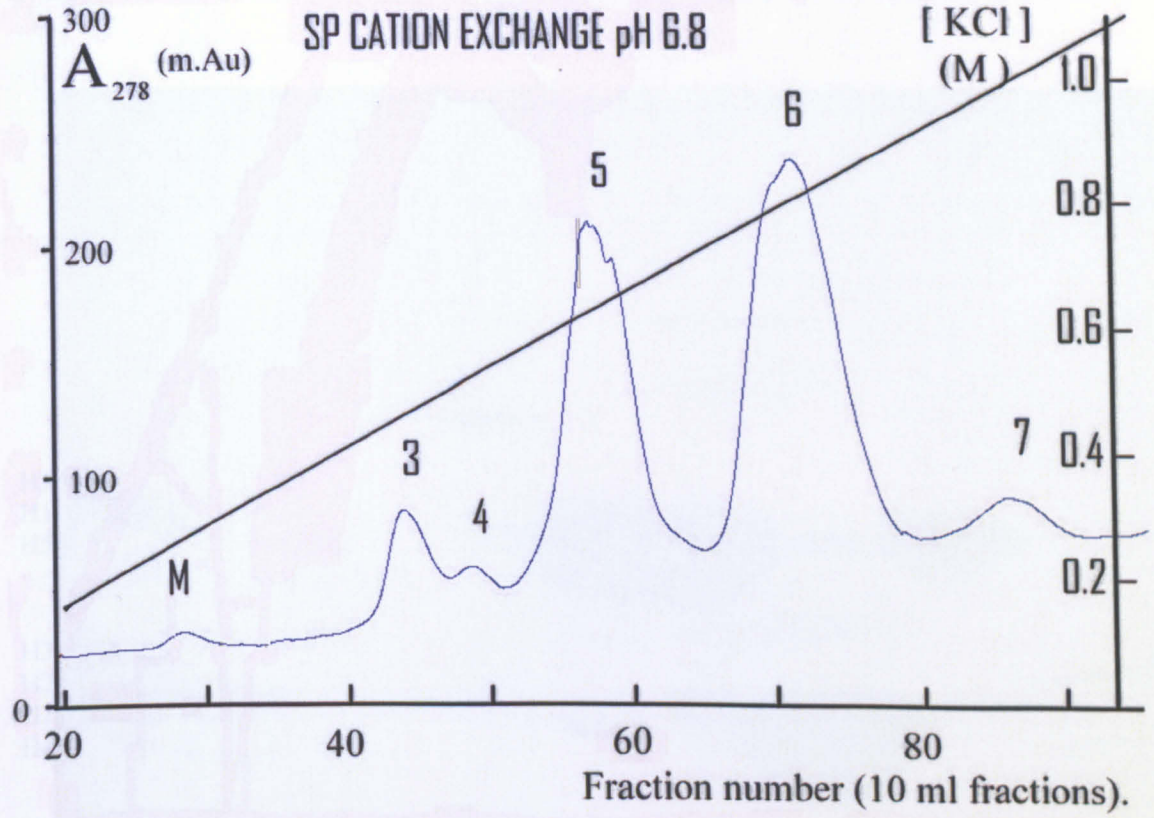
chromatography equipment.

Cation-exchange chromatography was carried out with a linear gradient of the elution-buffer B (1 M KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4). The chromatogram is in Figure 4.4. Six well-resolved peaks were assigned and the major proteins in each peak were as follows: Peak M (mystery peak) - to be determined; Peak 3 - linker histones 1a and 1b; Peak 4 - linker histone H1b; Peak 5 - core histone dimers; Peak 6 - linker histone H5; Peak 7 - residual histone tetramers that are insoluble at the high salt before dialysis into buffer A.

Fractions from Peak 6 were combined and concentrated to seven mls in an Amicon pressurised-ultrafiltration cell with a 10 kDa cut-off molecular weight. Ultra-violet spectrophotometry of this solution gave a yield of Peak 6, histone H5, of approximately 23 mg.

Each peak of the chromatogram of Figure 4.4, were then dialysed into distilled water and freeze dried, so that SDS-PAGE could be carried out as in Figure 4.5.

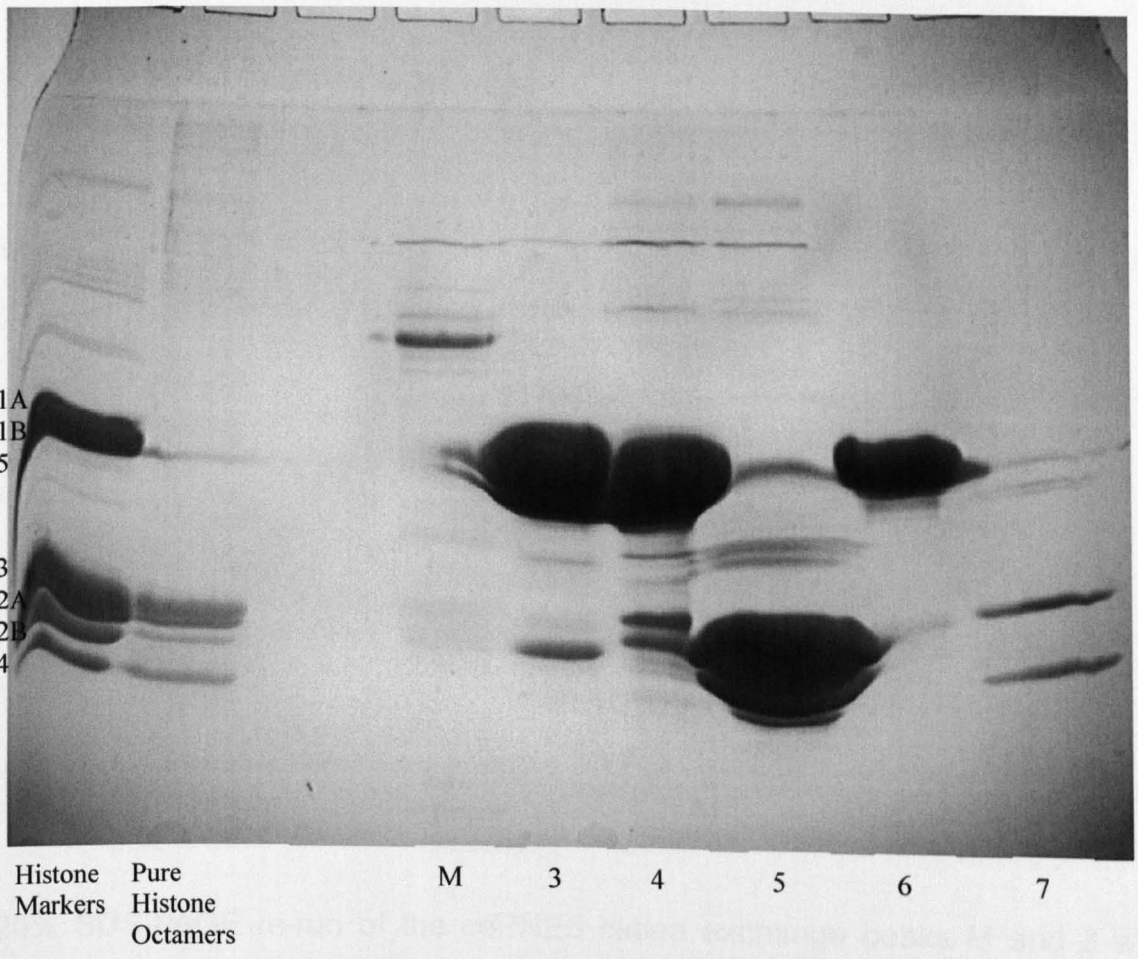
Figure 4.4 The chromatogram of the cePNE5 proteins using cation exchange



This diagram is showing the chromatogram of the cation exchange of cePNE5. 40 ml of cePNE5 sample was dialysed against cation exchange buffer A (100 mM KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4) and then loaded to four series of 5ml SP sephadex, HITRAP, Fast-flow cation-exchange columns. The chromatography was carried out with elution buffer B (1 M KCl, 50 mM K_2HPO_4, 50 mM KH_2PO_4). Seven well-resolved peaks were assigned as follows: Peak N is the non-binding protein (not shown in this Figure), Peak M (mystery peak) and peaks 3, 4, 5, 6 and 7.

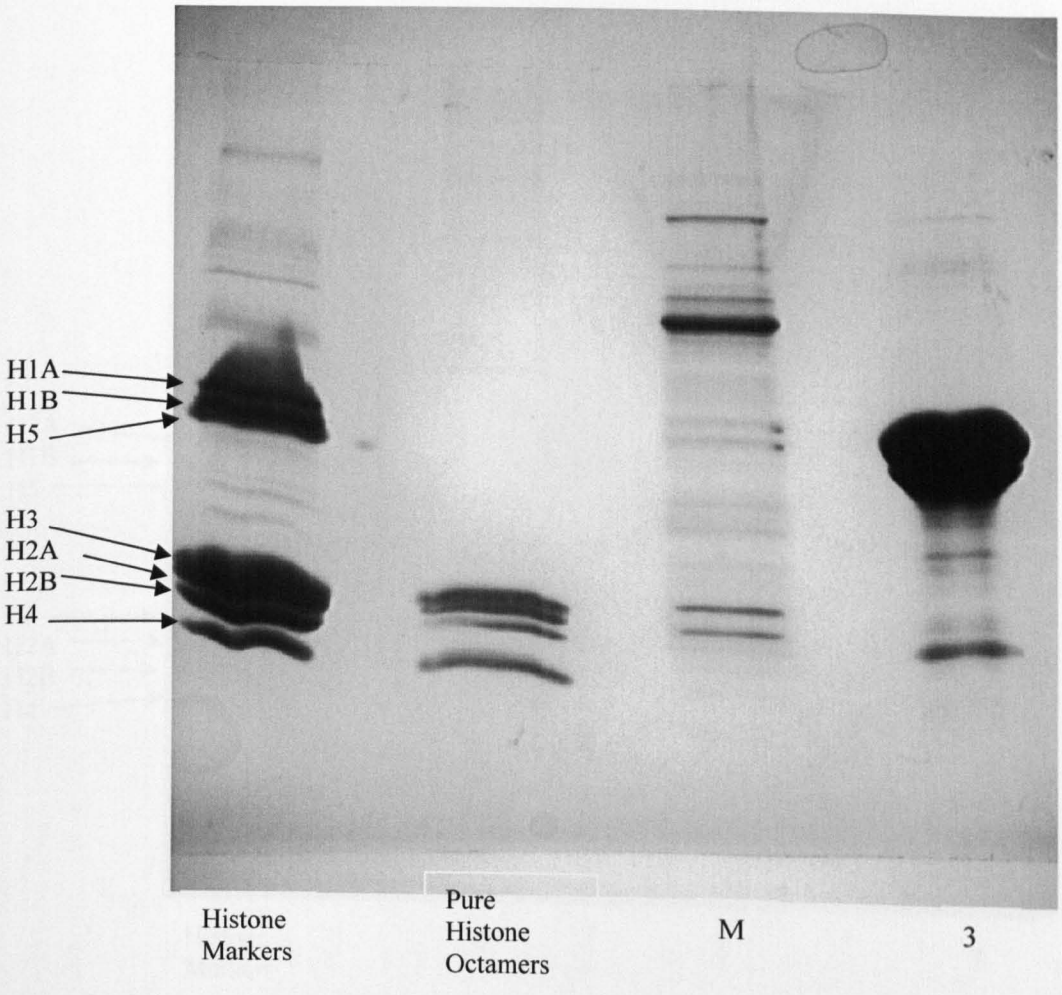
Figure 4.5 20% SDS-PAGE of peaks M, 3, 4, 5, 6 and 7

A. 20% SDS-PAGE of the cePNE5 cation exchange peaks M, 3, 4, 5, 6 and 7.



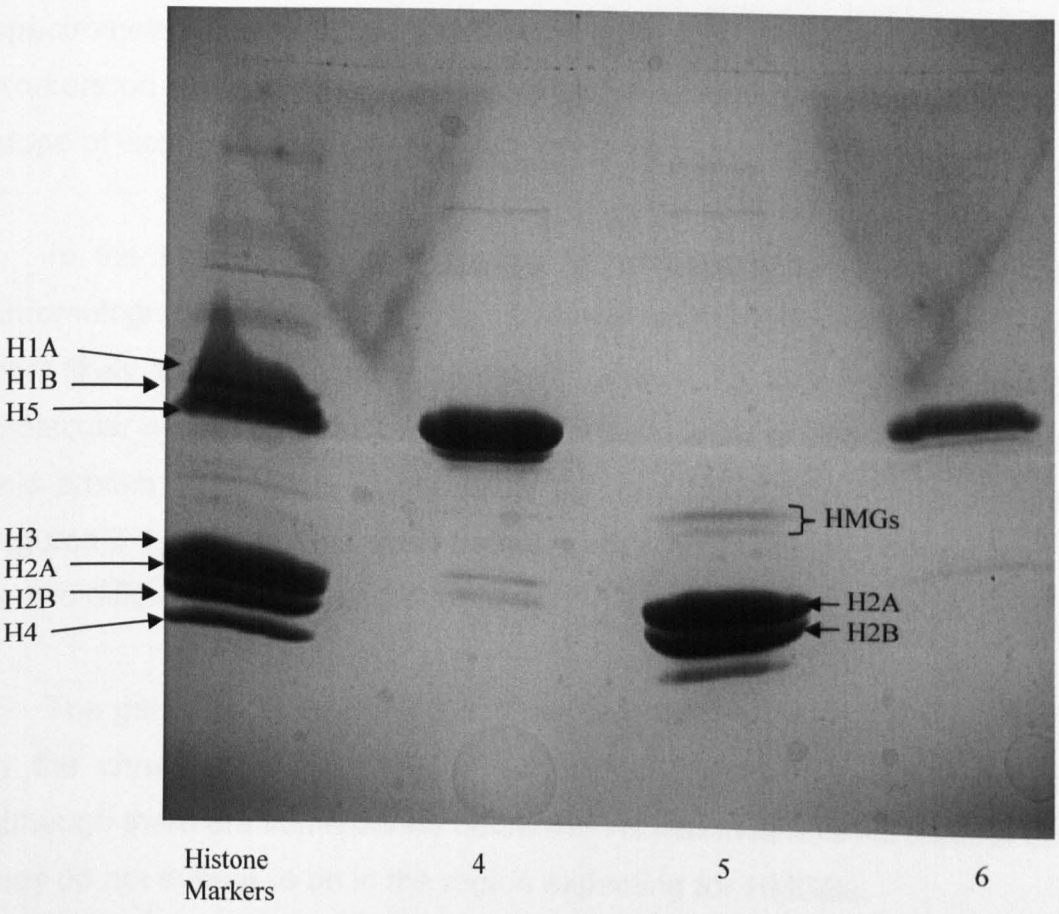
This Figure shows the 20% SDS-PAGE of the cePNE5 proteins cation exchange peaks M, 3, 4, 5, 6 and 7, compare with histone marker and pure histone octamer sample. Sample 3, 4, 5 and 6 are very heavily loaded to show all the proteins in the samples, but this caused some bands to spill over to other lanes. So two other 20% SDS-PAGE (Figure 4.5B and C) were run to decrease the loading of these overloaded samples and gaps were left in between each lane to avoid any possible spill over.

B.



A 20% SDS-PAGE re-run of the cePNE5 cation exchange peaks M and 3 with a higher loading of M and a lower loading of 3. Gaps were left in between each lane.

C.



A 20% SDS-PAGE re-run of the cePNE5 cation exchange peaks 4, 5 and 6 with a lower loading. Gaps were left in between each lane. This indicates that a linker histone like band of sample 5 in Figure 4.5A was actually a spill over from sample 4 (bands identified by compared with the histone marker).

Conclusion and discussion

The work described in this section 4.2 has not yet been completed with mass spectrometry data from the bands on Figure 4.5. However, in discussion with co-workers on this project we may summarise the interpretation of the results at this stage of work as follows:

In the Figures the lane labelled M corresponding to the small peak of the chromatogram contains a prominent protein and in comparison with Figure 3.4. It is very likely that this protein is heatshock protein 70, by comparing two gels and the molecular weight of these two proteins. Future work will be needed to establish why this protein fractionates under wildly different conditions in the chromatograms of Figures 3.4 and 4.6. This could be due to the modification of the protein, for example, or two different variants of the protein.

The gel lanes in Figure 4.6 suggest that the HMGs co-fractionate with the peak in the chromatogram in Figure 4.4 corresponding to histone dimers. However, although there are some bands below the H1 histones and H5 histone in the Figures, they do not appear to be in the region expecting for HMGBs.

Although recent literature makes almost no note of the possibility for H2A-H2B interacting with HMGBs, older papers of Bernues *et al.* (1983 and 1986) detect chemical crosslinking of the two components in free solution. It is clear from Figure 4.5 that the H2A-H2B dimers are in large excess compared with the HMGs. It will be of some interest in future experiments to see if specific modifications or variants of the HMGs bind to the histone dimers. It is noticeable in the chromatogram of Figure 4.4 that peak 5 corresponding to histone dimers (and also peak 6 for histone H5) is not symmetrical and it will be interesting in future experiments to run gels from the individual fractions across those peaks. If there are significant differences it will be useful to re-fractionate by cation-exchange chromatography different regions of peak 5 (and also peak 6).

Chapter 5

Extraction and Studies of Nucleoproteins from Human-Leukocyte and Human Tissue-culture Nuclei by using the "Forward Technology"

As described in Chapter 3, the two main technologies that are used in the LJMU chromatin research group to purify and study nucleoproteins are called "forward technology" and "reverse technology". The "forward technology" enables the separation of a small group of proteins (PNE1) from the cell nuclei without lysing them. Then nucleoproteins are separated from the nuclei lysate progressively by increasing its KCl/phosphate concentration (Chapter 2 and 3). The "reverse technology" involves the direct lysis of the nuclei by high KCl/phosphate "histone octamer precipitation conditions", which precipitate out most nucleoproteins including histone octamers in the very beginning first step. Only the most soluble proteins such as HMGs remained in the solution. Further fractionation, purification and analysis were then applied to these soluble proteins. Both technologies were applied to chick-erythrocyte nuclei and discussed in chapter 3. How the "forward technology" was applied to human-leukocytesand human cancer cell line (tissue-culture) nuclei is described in this chapter.The Author collaborated with supervisors in the development of this work and with Dr. James Nicholson of STFC Daresbury laboratory, where finances and laboratory facilities were available. The majority of the laboratory work was done by the Author.

The first step in studying human-leukocyte nucleoproteins was to obtain pure human-leukocyte cells which were separated from human buffy coats by centrifugation. Then the pure human-leukocyte cells were lysed by cell-lysis buffer containing detergent and the nuclei were separated from the other cell components by centrifugation and washing. This step of the experiment is described in section 5.1.

After the pure human-leukocyte nuclei were obtained, they were washed by PNE1 washing buffer to remove a group of proteins through the nuclear membrane,

called the huPNE1 (hu = human) proteins (see chapter 3.1 for the studies of the cePNE1 proteins). This step was done while the nuclei were kept intact by inclusion of $MgCl_2$ in the PNE1 washing buffer. The human-leukocyte nuclei were then lysed and the "forward technology" and the "reverse technology" were applied to study the nucleoproteins. The technologies were developed from the studies of chick-erythrocyte nucleoproteins, but some difficulties occurred and the technologies were slightly altered while applying them to the human-leukocyte nucleoproteins studies. The studies of human-leukocyte nuclei were mainly using the "forward technology" as described in section 5.2.

The "forward technology" was also applied to HT29 (human colon adenocarcinoma) human tissue-culture cell nuclei. The tissue-culture cells were harvested and lysed to produce nuclei by colleagues Dr. Katie Evans, Dr. Elaine Hemers and their project students. The nucleoproteins were then extracted from the nuclei by using the "forward technology". This is described in section 5.3.

The objective of this experiment was to apply the methods developed in the studies of chick-erythrocyte nucleoproteins to human-leukocyte nucleoproteins, to extract and enrich certain groups of proteins separately such as histone octamers (Carcia *et al.*, 2007; Kornberg and Thomas, 1974), linker histones and HMGs (Davey *et al.*, 2002) from human-leukocyte nuclei. Thus, methods for studying human nucleoproteins (Hake *et al.*, 2006; Khan and Krishnamurthy, 2005; Nightingale *et al.*, 2006) related to human disease such as cancer could be undertaken by using leukocytes from patients with disease in comparison with healthy people.

5.1 Extraction and purification of nuclei from human-leukocyte cells and human HT29 tissue-culture cells

Introduction

The human-leukocyte cells and nuclei were prepared from human buffy coat. Because it is a human product, an ethics approvement was needed. The ethics approved for this project was obtained from the Cheshire Research Ethics Committee. The work was carried out with the Science and Technology Facilities Councillaboratory at Daresbury, Cheshire, where a safety risk assessment was carried out. The buffy coats had been tested for diseases and were purchased for non-clinical use. At least 24 hours delay was incurred for disease-testing of the blood before it was available from the National Blood Service.

The human HT29 tissue-culture cells were prepared and harvested and the nuclei from them were extracted and purified by colleagues Dr. Katie Evans, Dr. Elaine Hemers and their project students. For the cell culture, HT-29 cells (purchased from ATCC) were used and cultured in a 'full' growth medium consisting of 500 ml of advanced RMPI media 1640 (Sigma-Aldrich), 50 ml foetal bovine serum (FBS or FCS, where C is short for calf), 5 ml penicillin/streptomycin (pen/step) and 5 ml L-glutamine- 200 mM (Gibco cat # 25030). The cells were maintained in a 95% O_2, 5% CO_2 humidified atmosphere at 37^0C. The cells were grown until there were 16 confluent flasks in total from which the nuclear HMGs could be extracted. Trypsin-EDTA (Sigma-Aldrich) was used to divide the cells in to culture flasks. CCD18 cells (myofibroblasts) were also cultured using this media and grown for stimulation with varying HMGB1 doses. The procedure used to produce nuclei from HT29 tissue culture was similar to that from human-leukocyte cells described below.

Methods for Producing Human Leukocyte-cells and Cell Nuclei:

Buffers:

(Benzamidine hydrochloride was added freshly and keptat 4 °C before use; all buffers were filtered by 0.4 μm filters before use)

Stock solutions:

1. 1 M Tris, pH 7.5, 1 M NaCl, 100 ml.
2. 10% Triton-X-100, 100 ml.
3. 1M $MgCl_2$.
4. Complete protease-inhibitor pills (Roche), 1 tablet per 40 mls of protein solution.

Working solutions:

1. Cell-lysis buffer: 100 ml of 10 mM Tris pH 7.5, 10 mM NaCl, 1.1 % Triton X100, 3.5 mM $MgCl_2$: add 1 ml of stock solution 1, 11 ml of stock solution 2, and 0.35 ml of stock solution 3 to distilled water and made up to 100 ml, followed by mixing the total solution thoroughly. The solution was made up immediately before use and filtered. 20 ml of this solution was used for the cell lysis, containing one dissolved protease-inhibitor tablet.

Procedure:

Immediately after arrival, two buffy coats were transferred to a beaker containing two Roche protease-inhibitor tablets and stirred untill the tablets dissolved.The blood was then transferred into two 50 ml Falcon tubes (about 45 ml per tube). The tubes were centrifuged at 3,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 8 minutes.This speed and time was determined by colleagues Prof. J. Baldwin and Dr. K. Evans, which is short enough so that the red cells were intact and the white cells not contaminated or tinged with red cells. The profile of cells (Figure 5.1) allows separation of the white cells and platelets with plasma at the top. Under this was a fairly thick layer of white cells (leukocytes) distributed in the plasma, then a thin layer of platelets with clearly-defined top and bottom boundaries. The thick bottom layer of red cells, at 8 minutes, was separated from the other layers. Note that the plasma was cloudy (containing white cells still). The platelet layer had sharp boundaries and was narrow (Figure 5.1).

Figure 5.1 Buffy coats centrifuged at 3,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 8 minutes

The buffy coats (about 45 ml) with a pill of complete protease inhibitor were centrifuged at 3,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 8 minutes immediately after arrival. The centrifugation conditions were critical so that the blood was separated into three main layers after centrifugation. The cloudy top layer was plasma still mixed with white blood cells; the middle thin layer of platelets had clearly-defined top and bottom boundaries; the bottom thick layer was non-nucleated red blood cells.

The cloudy plasma layers containing the white cells in each Falcon tube were carefully removed down to the top of the platelet layers and put into two 50 ml Falcon tubes that were balanced. The two Falcon tubes were centrifuged at 3,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 25 minutes and the white-cell region at the bottom of the tubes was retained while removing the clear plasma layer above (Figure 5.2).

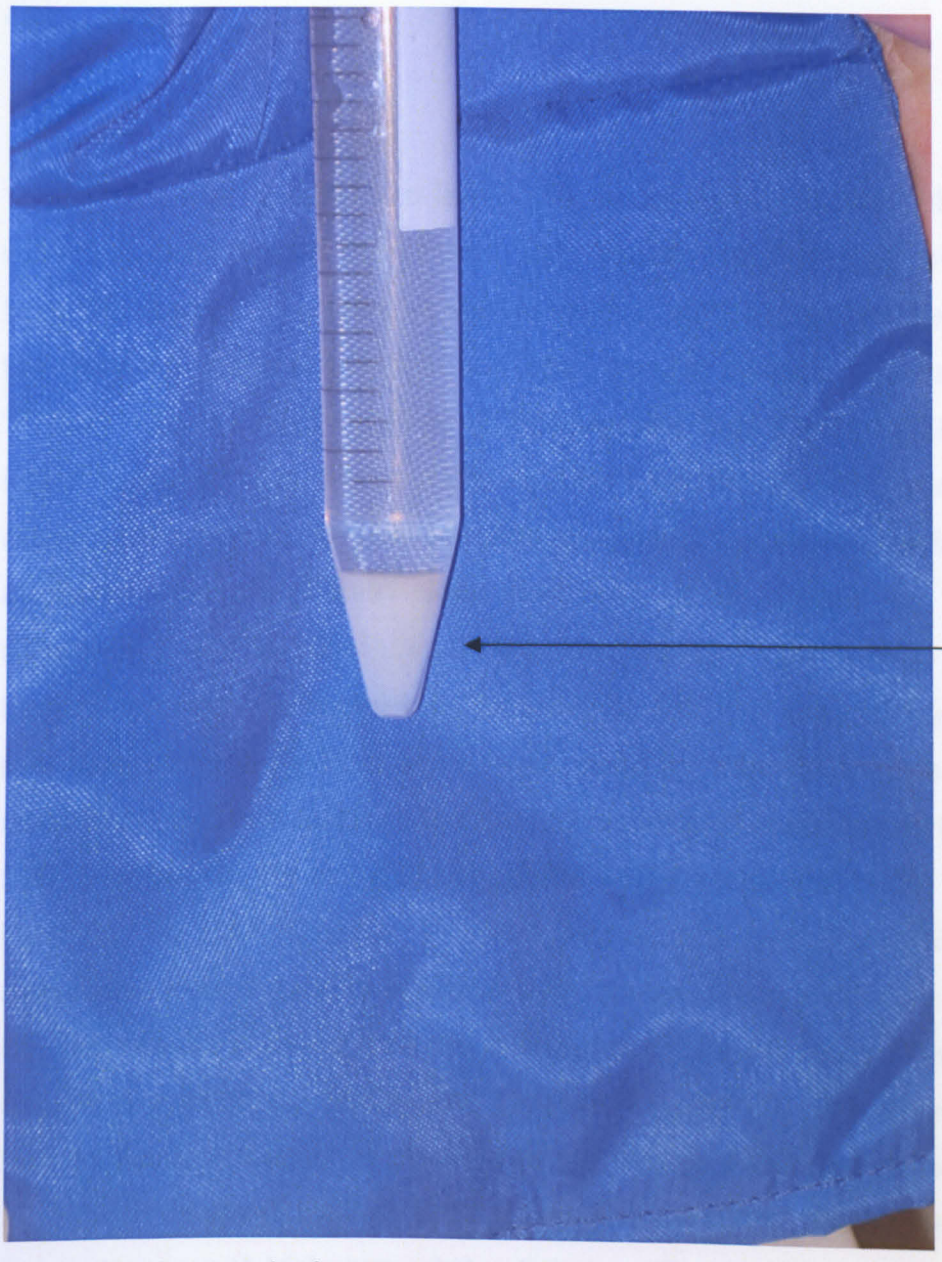
In principle, the initial centrifugation could have been even shorter, so that the erythrocytes were at the bottom of the centrifugation tube and the leukocytes and platelets still distributed in the plasma for later removal. Then the plasma containing leukocytes and platelets could be centrifuged exactly as described in the previous paragraphs to separate platelets from plasma containing leukocytes.

Figure 5.2 The mixture of plasma and white blood cells were then centrifuged at 3,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 25 minutes

The mixture of plasma and white blood cells was centrifuged at 3,500 rpm for 25 minutes. The white blood cells were pelleted to the bottom of the tube. The clear top layer supernatant was the plasma. The small red region down the bottom of the tube was a small amount of red blood cells remaining in the sample, or more probably residual platelets.

The remaining leukocytes at the bottom of the tube were transferred to be dispersed in 20 ml of cell lysis buffer, stirring in a small beakerfor 30 minutes to lyse the cells. The lysed cells were partitioned into two fresh 15 ml Falcon tubes, balanced and centrifuged at 2,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 25 minutes. The two supernatants of cytoplasmic material were retained for the extraction of cytoplasmic proteins. The two white pellets of pure human-leukocyte nuclei at the bottom of the tubes were then ready for the next step of the experiment (Figure 5.3).

Figure 5.3 Preparation of the pure human-leukocyte nuclei



The lysed human-leukocyte cells were centrifuged at 2,500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 25 minutes, and then the pure nuclei were obtained at the bottom of the tube.

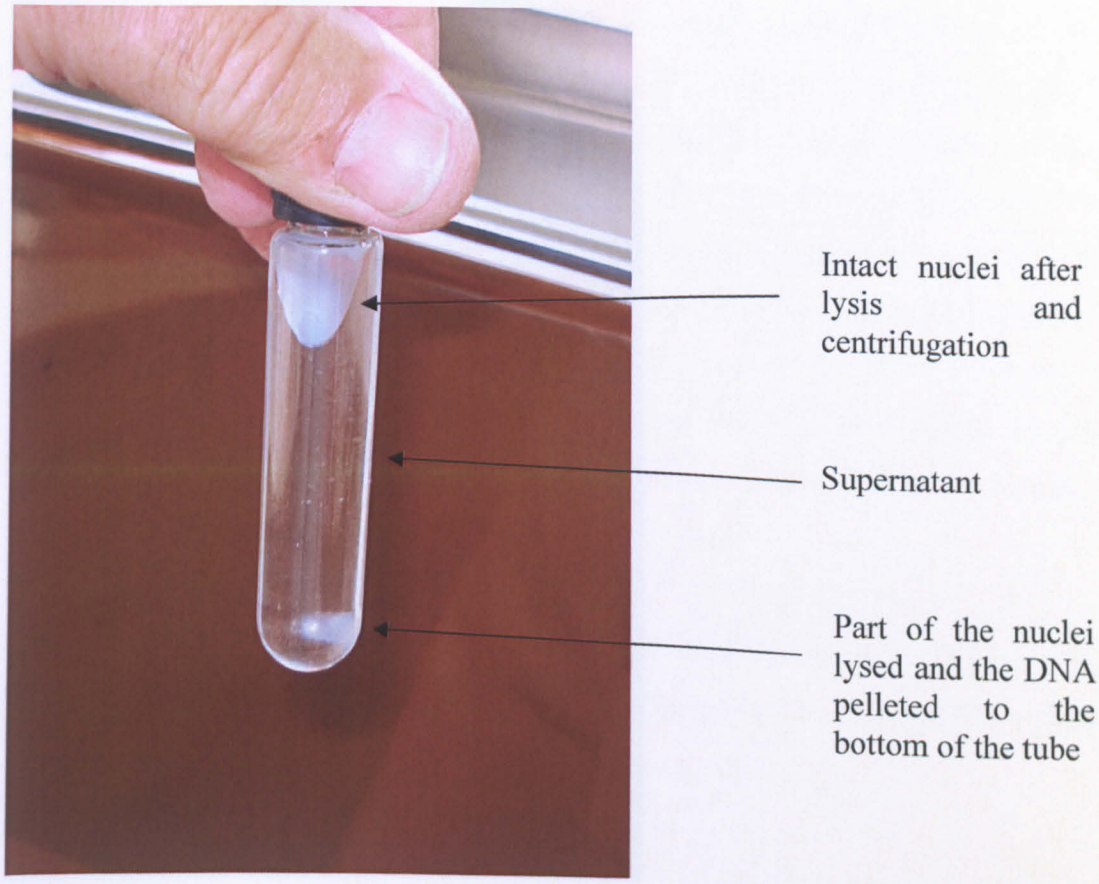
5.2 Extraction of nucleoproteins from human-leukocyte nuclei by using the "forward technology"

After the pure human leukocyte nuclei were obtained, the next step was to extract nucleoproteins from them. The "forward technology", which was developed from the studies of chick-erythrocyte nuclei, was applied to these human nuclei. In the "forward technology", the first step was to extract a group of proteins called "PNE1" from the nuclei by low KCl/phosphate buffer with $MgCl_2$ to keep the nuclei intact. Following that, the nuclei were then lysed in high ionic strength KCl/phosphate nuclei lysis buffer to remove histones and other nucleoproteins from the DNA. These nucleoproteins were then separated into small groups by further purifications (Chapters 2 and 3). When this "forward technology" was applied to human nuclei, the high KCl/phosphate nuclei lysis buffer (2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4, 2.5 mM benzamidine hydrochloride) was used in attempt to lyse the nulcei, but the results showed that much of the human nuclei kept intact under this condition (Figure 5.4). Meanwhile, the low KCl/phosphate PNE1 extraction buffer (70 mM KCl, 35 mM K_2HPO_4, 35 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride), developed from the study of chick-erythrocytes, was introduced to the human nuclei; but the results showed that the human nuclei lysed under this low KCl/phosphate condition even when $MgCl_2$ was present (Figure 5.5). Therefore, a new technology had to be established for these human nuclei. From a series of experiments, we found that although the low KCl/ phosphate condition could lyse the nuclei, the ionic strength was not high enough to dislodge the histones in which case the histones remained with the DNA. Finally, a new "forward technology" was established by lysing the nuclei under the low KCl/phosphate conditions while the histones still remained with DNA as a chromatin gel. Then the lysate was centrifuged through a second high ionic strength KCl/phosphate buffer to remove the histones from the DNA.

Another problem that occurred in these experiments was that of protein degradation. As with chick-erythrocytes, histones and other chromatin-fibre architecture proteins are subject to degradation due to endogenous proteases. One of these proteases has been identified by the group of Kouzarides. It cleaves the tail

of histone H3 (Santo-Rosa*et al.*, 2008). Therefore it was important to use 2.5 mM benzamidine hydrochloride or the complete protease-inhibitor cocktail from the company "Roche". Since the latter is expensive and we found that the cost could be minimized by using our method of leukocyte preparation.

Figure 5.4 The human-leukocyte nuclei were lysed in high ionic strength KCl/phosphate buffer



This Figure shows that the human-leukocyte nuclei only partially lysed under the high ionic strength KCl/phosphate condition (2 M KCl, 0.2 M K_2HPO_4, 0.2 M KH_2PO_4, 2.5 mM benzamidine hydrochloride). The top pellet was the nuclei which remained intact after lysis and centrifugation. The bottom pellet was the DNA from the nuclei which were lysed.

Figure 5.5 The nuclei lysed under the low ionic strength KCl/phosphate conditions

This Figure shows the pure human-leukocyte nuclei mixed with PNE1 extraction buffer (80 mM KCl, 40 mM K_2HPO_4, 40 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride) and centrifuged at 2,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 15 minutes. The nuclei were lysed under this low ionic strength condition and chromatin formed in the tube.

Figure 5.4 shows that the human-leukocyte nuclei were just partially lysed under the high ionic strength KCl/phosphate lysing condition developed for chick-erythrocyte nuclei (Chapters 2 and 3). On the other hand, Figure 5.5 shows that they were lysed under the low ionic strength KCl/phosphate condition usually used to extract the PNE1 proteins without nuclei lysis. The Author helped to supervise two project students, Mr. Arabarb Mirza and Mr. Damien Long,in their experiments with chick-erythrocyte cell nuclei. These showed that the cePNE1 proteins (Chapter 3) could be extracted without nuclei lysis at lower KCl/phosphate ionic strength than 0.9 M KCl, 0.45 M K_2HPO_4, 0.45 M KH_2PO_4. Their experiments also showed that nuclei lysis at a relatively lower ionic strength (0.2 M KCl, 0.1 M K_2HPO_4, 0.1 M KH_2PO_4) were possible. In that case, linker histones were released into solution, while core histones remained attached to DNA.

Conditions close to this proved effective in isolating huPNE1 proteins, linker histones and core histones separately from cell-culture cells. The results showing this are in section 5.3.

Work on the leukocytes needs to be continued to reproduce the conditions of section 5.3. However, working in a laboratory outside the John Moores University, necessary for safety containment and ethics approved, does present problems of access and this has been particularly difficult because of changes to the CCLRC Daresebury Laboratory. However, the work on human leukocytes will continue at the Diamond Light Source Laboratory near Oxford, where new ethics approval and safety conditions have just been completed.

So while nucleoproteins can be released from DNA-phosphate interactions by competition with an increased KCl/phosphate concentration, histone octamers in a highly enriched form remain attached to the DNA. They can then be removed by acid or salts in a highly enriched form.

According to the results and conclusions made above, a new "forward technology" was developed and applied to human-leukocyte nuclei. This time 8 ml of new nuclei-lysis buffer (125 mM KCl, 62.5 mM K_2HPO_4, 62.5 mM KH_2PO_4, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride) were placed into a 10 ml Ti 70 Beckman

centrifuge tube. The 2 ml of nuclei, now well dispersed in cell-lysis buffer (10 mM tris pH 7.5, 10 mM NaCl, 1.1 % Triton X100, 3.5 mM $MgCl_2$), were gently layered on top of the new-nuclei lysis buffer, using a glass Pasteur pipette (Figure 5.6). The tube was then carefully sealed without disturbing the layering and then shaken suddenly to lyse the nuclei. Sudden shaking instantly distributes the nuclei uniformly before the lysis. The centrifuge tube was balanced against a second tube containing buffer and centrifuged at 36,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) (100,000 xg) for 18 hours.

After centrifugation, the supernatant containing the nucleoproteins extracted from the human-leukocyte nuclei was collected. Approximately 30 mg of total proteins were extracted from 2 ml of the nuclei. The supernatant was then dialysed against distilled water and freeze-dried. The sample was then re-dissolved into 1 M Tris-HCl, pH 6.8 buffer and 20% SDS-PAGE was run for it (Figure 5.7).

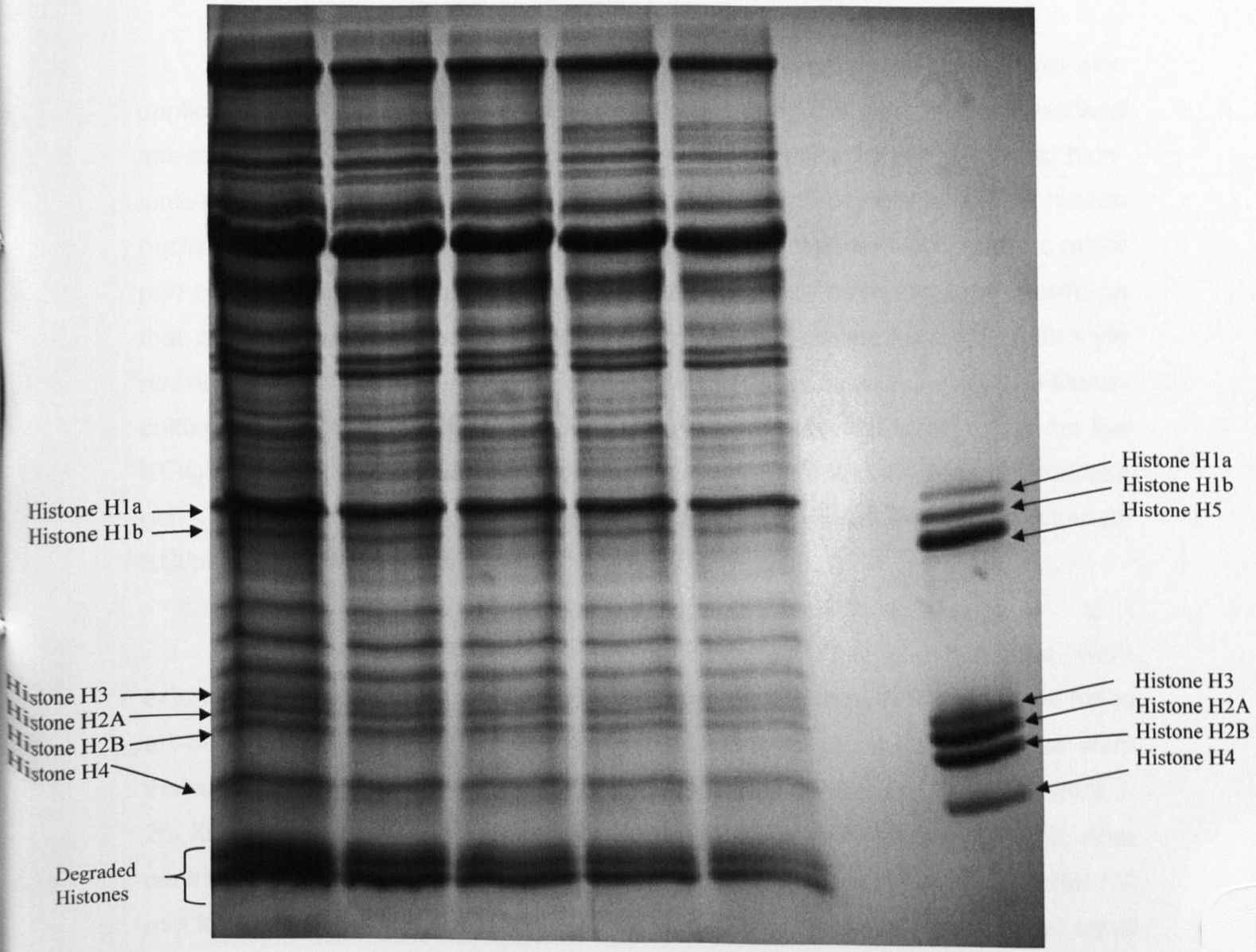
Figure 5.6 The lysis of human-leukocyte nuclei

a. b.

The 2 ml of human-leukocyte nuclei suspension was dispersed (a) and layered on the top of the nuclei lysis buffer (125 mM KCl, 62.5mM K_2HPO_4, 62.5 mM KH_2PO_4, 5 mM EDTA) with a glass Pasteur pipette (b). A sudden shaking was then applied to the tube to lyse the nuclei. After that the tube was balanced and centrifuged at 36,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 18 hours. The supernatant was then taken and a SDS-PAGE was run for it (Figure 5.7).

TEXT BOUND INTO THE SPINE

Figure 5.7 20% SDS-PAGE of the nucleoproteins extracted from the human-leukocyte nuclei



This Figure shows a 20% SDS-PAGE of the nucleoproteins extracted from the human-leukocyte nuclei. Core-histone octamers and linker histones were extracted out from the nuclei (bands identified by compared with the histone marker), but the yield was not as high as expected. That was because the major part of core histones still remained attached to DNA as explained above and in section 5.3. This problem will be solved by applying a further altered "forward technology" to the human leukocyte nuclei as for HT29 tissue-culture cell nuclei (section 5.3).

5.3 Extraction of nucleoproteins from human tissue-culture-cell nuclei by using the "forward technology"

The "forward technology", developed from chick-erythrocyte nuclei, was also applied to HT29 human tissue-culture cell nuclei. During the experiments described above (Figure 5.4 and 5.5), we found that the human nuclei did not lyse under high-ionic-strength nuclei-lysis buffer in contrast to chick-erythrocyte nuclei. The human nuclei could be lysed under a low *ionic strength* KCl/phosphate buffer, but the major part of the histones were not extracted out and remained bound to DNA. Based on that, a new "forward technology", altered from the previous one for chick-erythrocyte nuclei described in Chapters 2 and 3, was established for these human HT29 tissue-culture cell nuclei. In this new technology, the nuclei were first lysed under the low KCl/phosphate condition while the histones still remained with DNA as a chromatin gel. Then the lysate was centrifuged through a second high ionic strength KCl/phosphate buffer to dislodge out the histones.

The human HT29 tissue-culture cells were harvested and the nuclei were extracted as described in section 5.1. The nuclei were stored in 50 ml Falcon tubes at -80 °C in glycerol maintenance buffer (~40 ml in each tube). The nuclei were thawed at 4 °C before use, and then centrifuged at 3,000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes to remove the glycerol. After centrifugation, the supernatant was removed, a volume of nuclei washing buffer (10 mM tris, 10 mM NaCl, 3.5 mM $MgCl_2$, and 2.5 mM benzamidine hydrochloride) equal to that of the nuclei was added to the tube and thoroughly mixed with the nuclei by a submerged Pasteur pipette. The total volume was then made up to 40 ml and the nuclei were centrifuged at 2,000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes. The supernatant was removed and the nuclei were washed again with nuclei washing buffer and pelleted by centrifugation. Approximately 2 ml of nuclei in each tube were then kept in 4 °C cold room before use.

A tube of nuclei (2 ml) was used for an acid extraction to provide a control to examine the total protein in the nuclei. The nuclei were dispersed in nuclei washing buffer through a 22-gauge needle and an equal volume of 1 N sulphuric acid was

added to the tube. The tube was shaken well to mix and left on ice for an hour to let the acid extraction take place. After that the mixture was centrifuged at 13,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 20 minutes, the supernatant was collected and dialysed against distilled water. Then the protein concentration of the supernatant was measured by UV spectrophotometry (0.93 mg/ml, total protein = 3.71 mg). The proteins were then freeze-dried and ready for SDS-PAGE gel analysis (Figure 5.8 and 5.10).

Then, the huPNE1 proteins were extracted from another tube of nuclei, dispersed in PNE1 extraction buffer (75 mM KCl, 75 mM equimolar potassium phosphate, 3.5 mM $MgCl_2$, 2.5 mM benzamidine hydrochloride). The mixture was then centrifuged at 2,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 10 minutes; the supernatant was collected as the huPNE1 proteins. The PNE1 extraction procedure was repeated twice and three supernatants were combined together and ultra-centrifuged at 35,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 18 hours to remove any small amount of RNA which may have leaked through the nuclear membrane. The supernatant was then dialysed against distilled water and the protein concentration was measured by UV spectrophotometry (0.33 mg/ml, total protein = 3.77 mg). The huPNE1 sample was then freeze-dried and ready for SDS-PAGE gel analysis (Figure 5.8).

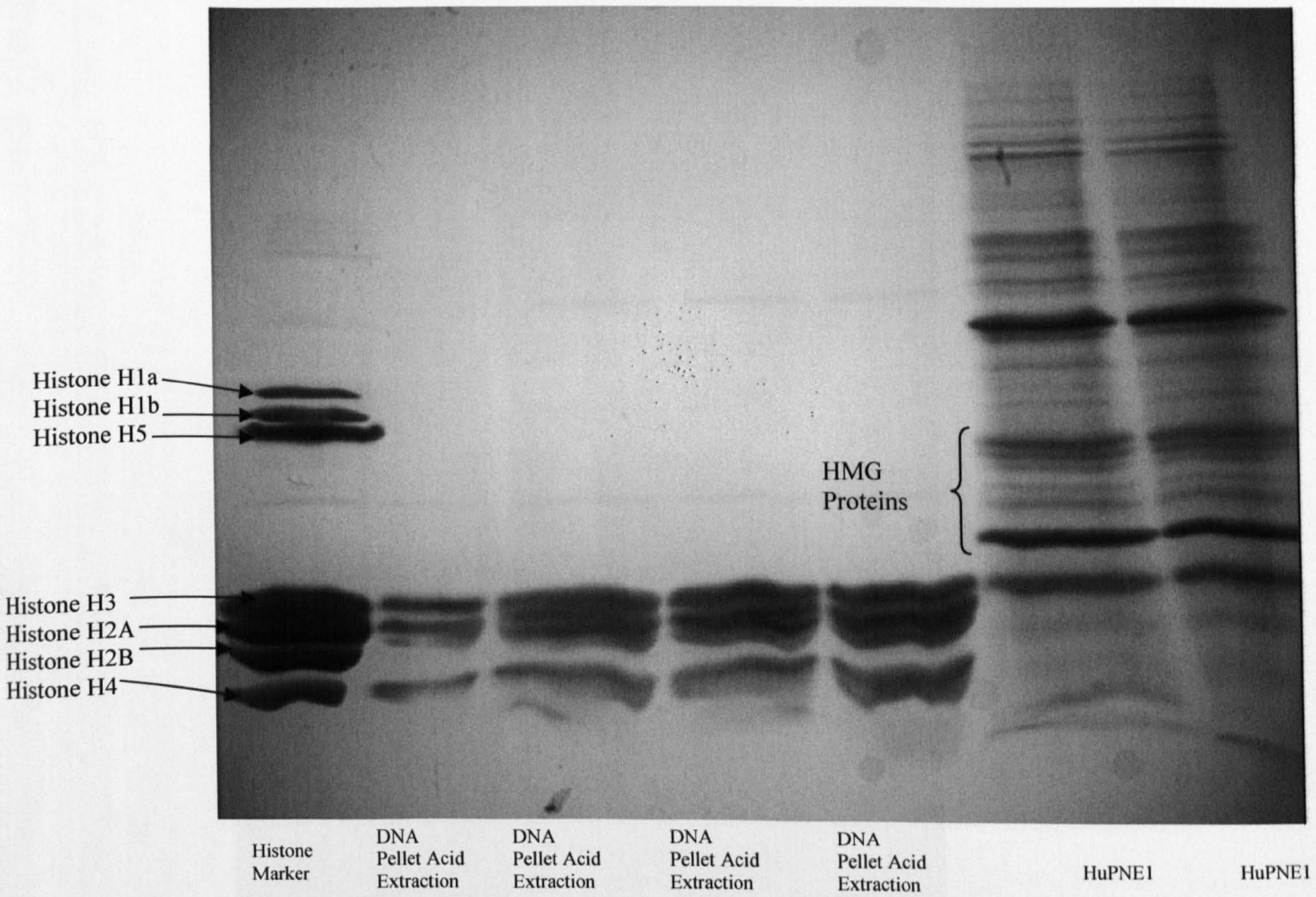
The PNE1-depleted nuclei were re-suspended in 1 ml of PNE1 extraction buffer and gently layered on top of 4 ml of nuclei lysis buffer 1 (0.2 M KCl, 0.2M equimolar potassium phosphate, 2.5 mM benzamidine hydrochloride). The tube was then shaken vigorously and the nuclei lysed in about 30 seconds. After that, the nuclei lysate was dispersed and carefully layered on top of 5 ml of nuclei lysis buffer 2 (2 M KCl, 0.4 M equimolar potassium phosphate, 2.5 mM benzamidine hydrochloride) (Figure 5.9). Then the tube was directly ultra-centrifuged at 35,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 18 hours without disturbing the two layers. After centrifugation, the supernatant was taken and separated as the top layer, the middle layer and the bottom layer (the layers were not clearly defined, 3 ml of each layer was taken). The three samples were then dialysed against distilled water and the protein concentrations were measured by UV spectrophotometry (top: 0.36 mg/ml, total protein = 2.0 mg; middle: 0.41 mg/ml, total

protein = 1.32 mg; bottom: 0.47 mg/ml, total protein = 3.49 mg). The samples were then freeze-dried and ready for SDS-PAGE gel analysis (Figure 5.8 and 5.10).

The above sample procedure allows separation (enrichment) of the important chromatin architectural proteins: HMGs, linker histones and core histones in a simple procedure, and allows the analysis of nucleoproteins for proteomics of cell-culture nucleoproteins. The methods should be directly transferable to leukocyte nucleoproteins to determine their yields and proteomics, with further enrichment of nucleoproteins by the methods described in chapter 2 and 3.

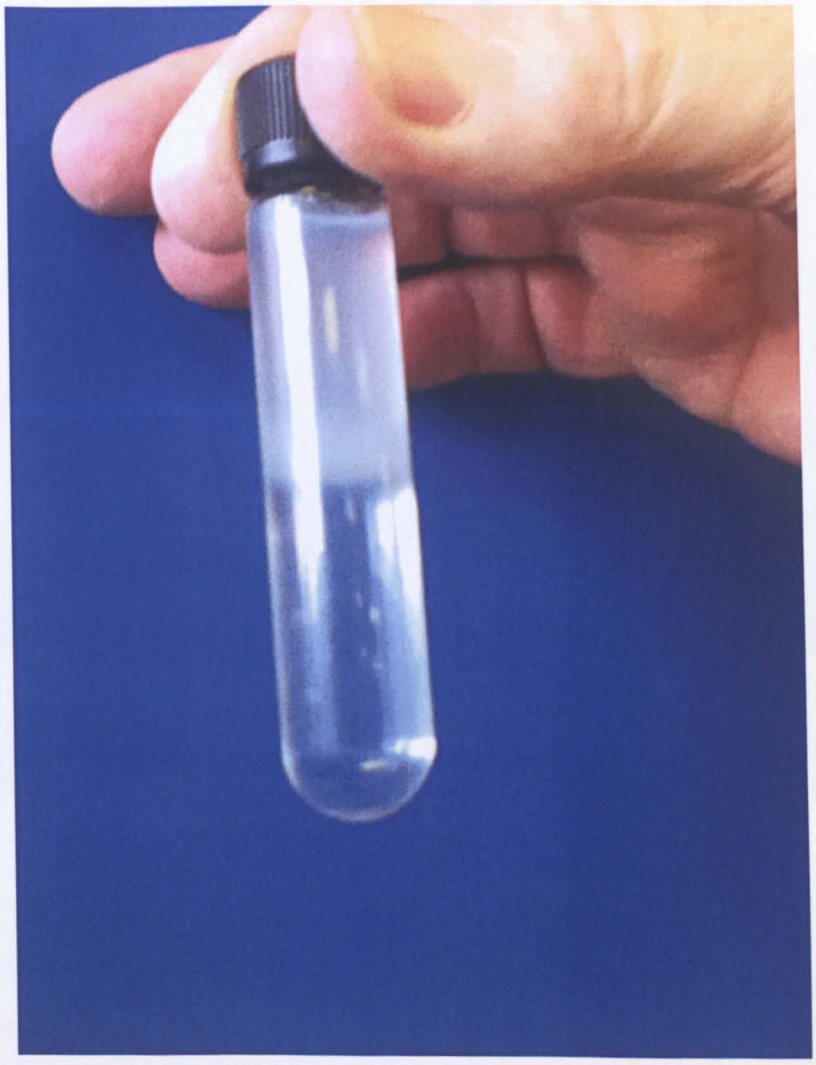
Results:

Figure 5.8 The 20% SDS-PAGE of the huPNE1 proteins which were extracted from the HT29 human tissue-culture nuclei



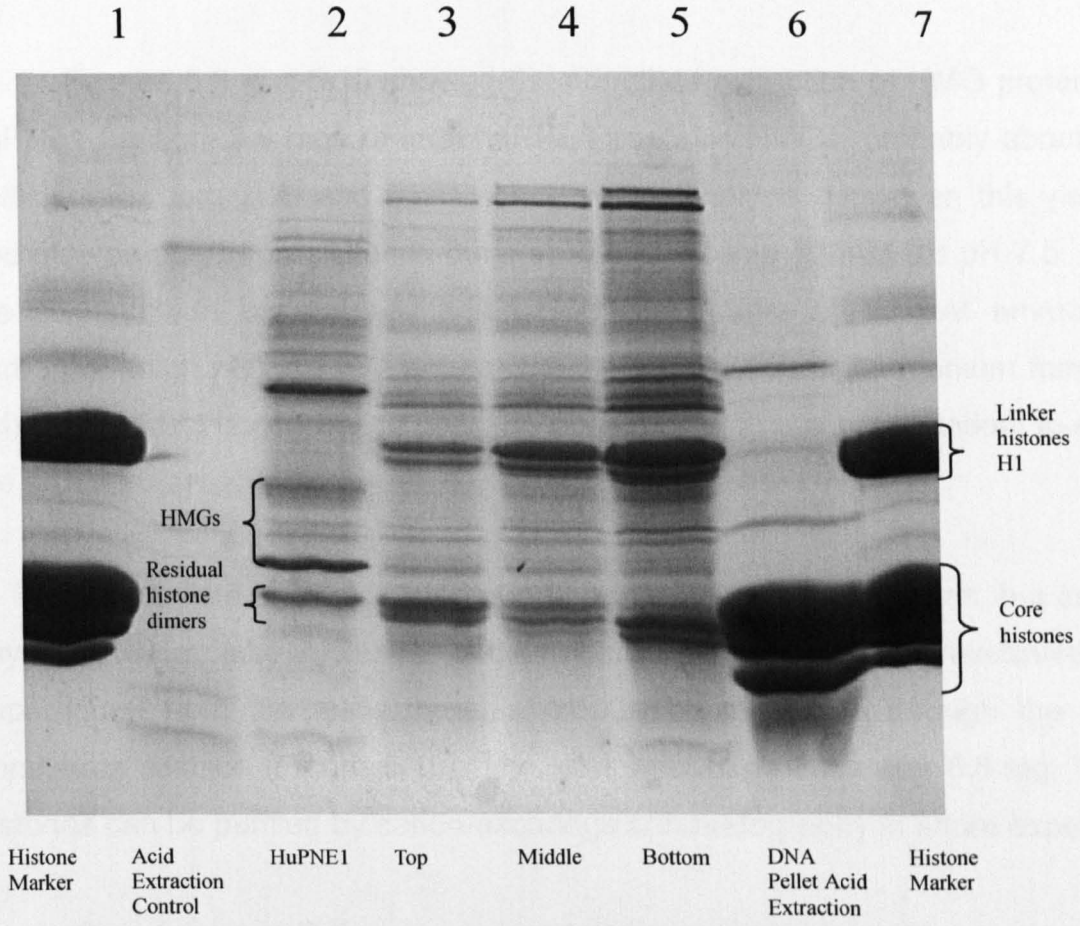
This Figure shows the 20% SDS-PAGE of the huPNE1 proteins which were extracted from the HT29 human tissue-culture nuclei (in 75 mM KCl, 75 mM equimolar phosphate extraction buffer), and compare with the histones acid extracted from the pellet after centrifugation that is composed of pure core histones (bands identified by compared with the histone marker).

Figure 5.9 The lysate of nuclei lysed in nuclei lysis buffer 1 (0.2 M KCl, 0.2 M equimolar phosphate) was then carefully layered on top of the nuclei lysis buffer 2 (2 M KCl, 0.4 M equimolar phosphate)



This Figure shows that how the lysate of nuclei lysed in nuclei-lysis-buffer 1 (0.2 M KCl, 0.2 M equimolar phosphate) was then carefully layered on top of the nuclei-lysis-buffer 2 (2 M KCl, 0.4 M equimolar phosphate). This tube was then ultra-centrifuged at 35,000 rpm (Beckman OPTIMA™ L-80 preparative ultracentrifuge, Ti 70 rotor) for 18 hours without disturbing these two layers.

Figure 5.10 The 20% SDS-PAGE of the three layers of nucleoproteins which were extracted from the HT29 human tissue-culture nuclei



This Figure shows 20% SDS-PAGE of the top layer, the middle layer and the bottom layerof the nucleoproteins extracted from the HT29 human tissue-culture nuclei. Note that the HMG proteins were enriched in the HuPNE1 fraction (lane 2), the linker histones were enriched in the supernatant after nuclei lysis and ultracentrifugation (lane 3, 4 and 5), and while the core histones remained attached to the DNA pellet to be acid extracted (lane 6) bands identified by compared with the histone marker.

Conclusions and discussion

HT29 Cell-Culture Cells:

Figures 5.8 and 5.10 showed the complete extraction of HMG proteins in the huPNE1 fraction. 3.8 mgs of total huPNE1 contains HMGs, probably about 300 ng, sufficient for both 2DE and mass spectrometry analysis. However, this yield is low and may be due to precipitation during the dialysis into 10 mM tris pH 7.5. In future, the solution can be dialysed against volatile buffers of 10 mM ammonium bi-carbonate (high pH), 10 mM ammonium acetate (pH 7) or ammonium formate (low pH). Then direct freeze drying could be done including any precipitation, to extract all the proteins.

A high yield of linker histones and some of the other proteins, but including a tiny quantity of core histones, were extracted completely and exclusively in the supernatant from the nuclei lysis, after ultra-centrifugation through the high-KCl-phosphate cushion (Figure 5.10). The total yield of proteins was 6.8 mg. The linker histones can be purified by cation-exchange chromatography in future experiments.

Core histones (3.5 mg) were exclusively targeted by acid extraction from the chromatin pellet after ultracentrifugation. Histone dimers and tetramers can be extracted by cation-exchange chromatography for further targeting. The chromatin pellet can be cut and the proteins can be extracted in KCl/phosphate, rather than acid extraction. The yield would be lower, but there would be plenty of extract for 2DE and PTM analysis, without exposure to acid conditions.

Remarkably the core histones were not degraded at all. This indicated that the expensive complete protease inhibitors are not necessary to extract tissue culture proteins intact. Benzamidine hydrochloride was enough to protect histones from the various proteases. The procedure therefore opens the door for a very efficient, fast and inexpensive identification of PTMs and variants of selective core histones, selective linker histones and selected HMG proteins and many other nucleoproteins.

Human Blood Leukocytes:

In the work explained in this chapter, the isolation of cells from human blood from the National Blood Services was completed without using any solvents, except the introduction of heparin by the National Blood Services to prevent blood clotting. The procedure was simply to use low speed centrifugation and careful separation to extract leukocytes dispersed in the blood plasma, leaving the platelets layer and erythrocytes as clear bands. This is a fast, efficient procedure which in our opinion compares well with other existing methods (Graziani-Bowering *et al.*, 1997; Okuzaki *et al.*, 2011).

There are many strategies for separating blood components using centrifugation and others have also used the principle of distributing platelets and leukocytes in blood plasma allowing separation from the bottom red cells component (Internet: OnFocus Recipe Box, 2009). In our procedure we have started with human buffy-coats from the National Blood Services. If we started from whole blood it would be possible to allow the red cells to fractionate to the bottom of the tube allowing the platelets and leukocytes to be extracted, distributed in the plasma. It would be dangerous to remove any plasma top layer in case of losing the platelets and leukocytes distributed in that layer. Then the procedure outlined in this chapter would be very efficient for obtaining pure platelets and pure leukocytes without exposing the components to any solvents except for the heparin introduced initially into the blood, to prevent clotting. The leukocytes could then be fractionated into the component cells (Neutrophils, most abundant; Eosinophils, 2-4%; Basophils, 1%; Lymphocytes, 25%; and Monocytes, 3-8%) using density gradient centrifugation, using, for example, percol or "Optiprep" as density gradient media (Axis-Shielddensitygradientmedia, "nicomed" Application Sheet C43; 3rd, July 2011). We are aware that pure plasma after high speed centrifugation could be used as a density gradient media and this would mean that there would be no intervention by any outside solutions since the collection of blood apart from the addition of heparin.

Human leukocytes were extracted in one experiment and tissue-culture cells in the second one. Nuclei were produced as described in this chapter, based on

early papers by Garrard and Hancock (Garrard and Hancock, 1978) and by Hancock, Faber and Fakan (Hancock *et al.*, 1977). The methods have been developed in many laboratories (Simon and Felsenfeld, 1979; Moudrianakis*et al.*, 1985; Goodwin *et al.*, 1979) including the LJMU Chromatin Research Group (Prof. J. Baldwin) with experience from work with colleagues at Portsmouth University (Suau *et al.*, 1979). Comparing each of these technologies developed by others we believe that our method is unique in the use of phosphate combined with KCl. Indeed this combination established conditions for crystallisation of intact histone octamers. The crystals with space group $P6_5$ are unique to the LJMU Chromatin Research Group and are discussed in chapter 2 (Chantant *et al.*, 2003; Wood *et al.*, 2005).

In our method of extracting cell nuclei from human buffy coat, only one solution is required reduce which is the time of the experiment by using one cell lysis stage. After centrifugation, the nuclei were simply taken up in buffer and proteins are extracted as described earlier.

In the procedures described in this chapter, we were able to produce the huPNE1 proteins (enriched in HMG proteins), separate from a supernatant after centrifugation enriched in linker histones. Again separately, core histones were extracted from the sedimented chromatin pellet. The procedure for the tissue-culture cells can be a standard procedure for enriching this group of proteins greatly reducing the clutter of spots in 2D gel electrophoresis and enabling directly mass spectrometry determination of the enriched proteins separately in each of the three extractions.

The methods for cePNE1-proteins extraction and purification of HMGs from chick erythrocytes have been described in chapter 3. Shechter *et al.* (2007) have reviewed isolation, purification and analysis of histone proteins particularly applied to mass spectrometry and other studies in the group of Allis (Shechter *et al.*, 2007). The method described in this chapter allows core histones, linker histones, HMG proteins and other proteins such as peptidylprolylisomerases to be studied by mass spectrometry, following one extraction from human tissue-culture cells or human leukocytes.

The procedures developed in this chapter could be used to study biomarkers of diseases, by studying proteins from diseased cells in comparison with normal cells used as control.

Conclusion and Discussion

The major target of this thesis work was to develop a fast and efficient technology to extract, purify and study nucleoproteins from chick and human cell nuclei. Proteins were extracted by successively increasing the ionic strength of KCl/phosphate solution to high values. The methods together are called the "forward technology" to distinguish them from procedures where proteins are extracted abruptly at very high ionic strengths to produce only the most soluble nucleoproteins.

In Chapter 2, ultra-pure histone octamers, dimers and tetramers were extracted from chick erythrocyte nuclei within 3 days by using this "forward technology". The pure histone octamers were crystallised to reach the highest resolution, 1.8 Å. The pure histone H2A-H2B dimers and $(H3\text{-}H4)_2$ tetramers were complexed with histone chaperone NAP1 in collaboration with Dr. Emily Newman and Dr. John MeGeehan of Pothmouth University.

In Chapter 3, the "forward technology" was used to enrich important nucleoporteins such like HMGs, into small groups, to allow further purification and studies of these proteins. A group of proteins rich in HMGs was extracted from chick erythrocyte nuclei by the "forward technology" called cePNE1 proteins. Ion-exchange chromatography was applied to the cePNE1 proteins to fractionate them. During these experiments, an important peptidyl prolyl isomerase, FKBP3, was found to be co-fractionated with HMGB1 which may indicate a potential complex between them.

In Chapter 4, two other groups of proteins were extracted from chick erythrocyte nuclei called cePNE4 and cePNE5 proteins. CePNE4 proteins mainly contain some high molecular weight proteins including some functionally important ones such as DNA topoisomerase II. CePNE5 proteins contain mainly linker histones. Cation exchange chromatography was used to fractionate this group of proteins. Results showed that some histone dimers co-fractionated with HMGB proteins. This indicates complexes between HMGB proteins and histone dimmers (H2A-H2B) which could be purified further for crystallography and other structural studies.

In Chapter 5, the "forward technology", developed from chick nuclei, was applied to human-leukocyte and human-tissue-culture nuclei. Pure human tissue-culture nuclei were produced in a fast and efficient manner. HMG proteins were enriched in the huPNE1 group proteins and human core histones were acid extracted (Figure 5.8).

These works will lead to a wide range of further studies of nucleoproteins:

The complex between HMGBs and FKBP3 will need to be purified. Following that, further extraction, purification, crystallization and other functional or disease related studies can be carried out for this complex.

The same can be done for the complex between HMGBs and histone H2A-H2B dimers. Crystallization of the ultra-pure histone dimers and tetramers, which have never been achieved, should be carried out.

The methods of extracting and studying of human nucleoproteins in this thesis need to be extended, particularly for structural studies of nucleoprotein assemblies.

References

Adamietz P., Bredehorst R. & Hilz H. (1978), ADP-ribosylated histone H1 from HeLa cultures. Fundamental differences to $(ADP\text{-}ribose)_n$-histone H1 conjugates formed into vitro. *Eur. J. Biochem.*, **91**, 317-326.

Adams, C.R. & Kamakaka, R.T. (1999) Chromatin assembly: biochemical identities and genetic redundancy. *Curr. Opin.Genet. Dev.*, **9**, 185-190.

Agresti A. & Bianchi M.E. (2003), HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.*, **13**, 170-178.

Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., Shilatifard, A., Workman, J. & Zhang, Y. (2007), New nomenclature for chromatin-modifying enzymes. *Cell.*, **131**, 633-636.

Atkinson, B., & Mavituna, F. (1983), Biochemical Engineering and Biotechnology Handbook, *Macmillan Publishers Ltd.*, Surrey.

Andrews, A.J., Downing, G., Brown, K., Park, Y.J. & Luger, K. (2008), A thermodynamic model for Nap1-histone interactions. *J. Biol. Chem.*, **283,** 32412–32418.

Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A. & Luger, K. (2010), The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions. *Mol. Cell.*, **37,** 834–842.

Bailey, J.E. & Ollis, D.F. (1986), Biochemical engineering Fundamentals, 2nd edition, *cGraw Hill*, New York.

Baldwin, J.P., Bradbury, E.M., Butler-Browne, G.S. & Stephens R.M. (1973), Histone interactions with prokaryote DNA. *FEBS Lett.*, **34**, 133-136.

Baldwin, J.P., Boseley P.G., Bradbury, E.M. & Ibel, K. (1975), The subunit structure of the eukaryotic chromosome. *Nature*, **253**, 245-249.

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.

Belter, P.A., Cussler, E.L. &Hu, W. S. (1988), Bioseparations: Downstream Processing for Biotechnology. *John Wiley and Sons*, New York.

Berg, J.M., Tymoczko, J.L. & Stryer, L. (2002), Biochemistry. 5th edition.*W. H. Freeman*, New York.

Bernués, J., Espel, E. & Querol, E. (1986), Identification of the core-histone-binding domains of HMG1 and HMG2. *Biochim. Biophys. Acta.*, **866,** 242-251.

Bernués, J., Querol, E., Martinez, P., Barris, A., Espel, E. & Lloberas, J. (1983), Detection by chemical cross-linking of interaction between high mobility group protein 1 and histone oligomers in free solution. *J. Biol. Chem.*, **258,** 11020-11024.

Bianchi M.E. & Agresti A. (2005), HMG proteins: dynamic players in gene regulation and differentiation. *Curr. Opin. Genet. Dev.*, **15**, 496-506.

Bianchi M.E. & Manfredi A.A. (2007), High-mobility group box 1 (HMGB1) protein at the crossroads between innate and adaptive immunity. *Immunol. Rev.*, **220**, 35-46.

Blank, T.A. & Becker P.B. (1995), Electrostatie mechanism of nucleosome spacing. *J. Mol. Biol.*, **252**, 305-313.

Bohm, V., Hieb, A.R., Andrews, A.J., Gansen, A., Rocker, A., Toth, K., Luger, K. & Langowski, J. (2010), Nucleosome Accessibility Governed by the Dimer/Tetramer Interface. *Nucl. Acid. Res.*, **39**, 3093-3102.

Bowman, A., Ward, R., Wiechens, N., Singh, V., El-Mkami, H., Norman, D.G. & Owen-Hughes, T. (2011), The Histone Chaperones Nap1 and Vps75 Bind Histones H3 and H4 in a Tetrameric Conformation. *Mol. Cell.*, **41,** 398-408.

Bradbury, E.M. Inglis, R.J. & Matthews, H.R. (1974), Contral of cell division by very lysine rich histone (F1) phosphorylation. *Nature (London)*, **247**, 257-261.

Bruins, A.P. (1991), Liquid chromatography-mass spectrometry with ionspray and electrospray interfaces in pharmaceutical and biomedical research. *J. Chromatogr.*, **554,** 39-46.

Burgess, R.J. & Zhang, Z. (2010), Histones, histone chaperones and nucleosome assembly. *Protein Cell*, **1**, 607-612.

Bustin M. (1999), Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Mol. Cell. Biol.*, **19**, 5237-5246.

Bustin M. (2001), Chromatin unfolding and activation by HMGN(*) chromosomal proteins. *Trends. Biochem. Sci.*, **26**, 431-437.

Campos, E.I. & Reinberg, D. (2009), Histones: annotating chromatin. *Annu. Rev. Genet.*, **43**, 559-599.

Cato, L., Stott, K., Watson, M. &Thomas, J.O. (2008), The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J. Mol. Biol.*, **384,** 1262-1272.

Chang, L., Loranger, S.S., Mizzen, C., Ernst, S.G., Allis, C.D.& Annunziato, A.T. (1997), Histones in transit: cytosolic histone complexes and diacetylationof H4 during nucleosome assembly in human cells. *Biochem.*, **36,** 469–480.

Chantalat, L., Nicholson, J.M., Lambert, S.J., Reid, A.J., Donovan, M.J., Reynolds, C.D., Wood, C.M. & Baldwin, J.P. (2003), Structure of the histone-core octamer in KCl/phosphate crystals at 2.15Å resolution. *Acta Crystallogr. D. Biol. Crystallogr.*, **59**, 1395-1407.

Chen, C.C., Carson, J.J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J. & Tyler, J.K. (2008), Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell*, **134**, 231-243.

Chen, S., de Vries, M.A. & Bell, S.P. (2007), Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes Dev.*, **21**, 2897-2907.

Choudhary, P. & Varga-Weisz, P. (2007), ATP-dependent chromatin remodelling: action and reaction. *Subcell. Biochem.*, **41**, 29-43.

Crick, F.H.C., Leslie Barnett, F.R.S., Brenner, S. & Watts-Tobin, R.J. (1961), General nature of the genetic code for proteins. *Nature*, **192**, 1227-1232.

Csordas, A. (1990), On the biological role of histone acetylation. *Biochem. J.* **265**, 23-38.

Culter, P. (2004), Protein Purification Protocols, 2nd edition, *Human Press Inc.*, New Jersey.

Das, C., Tyler, J.K. & Churchill, M.E.A. (2010), The histone shuffle: histone chaperones in an energetic dance. *Trends. Biochem. Sci.*, **35**, 476-489.

Daugaard, M., Rohde, M. & Jaattela, M. (2007), The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Letters.*, **581,** 3702-3710.

Davey, C.A., Sargent, D.F., Luger, K., Maeder A.W. & Richmond, T.J. (2002), Solvent mediated interactions in the structure of the nucleosome core particle at 1.9Å resolution. *J. Mol. Biol.*, **319**, 1097-1113.

Dimitrov, S., Dasso, M.C. & Wolffe, A.P. (1994), Remodelling sperm chromatin in Xenopus laevis egg extracts: the role of core histone phosphorylation and linker histone B4 in chromatin assembly. *J. Cell. Biol.*, **126**, 591-601.

Diffley, J.F., Cocker, J.H., Dowell, S.J. & Rowley, A. (1994), Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell*, **78**, 303-316.

Diffley, J.F. & Labib, K. (2002), The chromosome replication cycle. *J. Cell. Sci.*, **115,** 869-872.

Donham, D.C. 2nd, Scorgie, J.K. &Churchill, M.E. (2011), The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histoneH3-H4-DNA complexes. *Nucl. Acid. Res.*, **39,** 5449-58.

Douglas, C., Donham, I.I., Jean, K., Scorgie, & Mair, E.A. (2011), The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3-H4-DNA complexes. *Nucl. Acid. Res.***10,** 1-10.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E. & Tyler, J.K. (2006), Structural basis for the histone chaperone activity of Asf1. *Cell*, **127**, 495-508.

Fearon, P., Lonsdale-Eccles, A.A., Ross, K., Todd, C., Sinha, A., Allain, F., & Reynolds, N.J. (2010), Keratinocyte Secretion of Cyclophilin B via the

Constitutive Pathways is Regulated through Its Cyclosporin-Binding Site. *J. Inves.Dermat.*, **0**, 000-000.

Fong, S., Mounkes, L., Liu, Y., Maibaum, M., Alonzo, E., Desprez, P., Thor, A.D., Kashani-Sabet, M. & Debs, R.J. (2003), Functional Identification of Distinct Set of Antitumor Activities Mediated by the FKBP Gene Family. *PNAS*, **100**, 14253-14258.

Ford, T.C. & Rickwood, D. (1990), A new one-step method for the isolation of human mononuclear cells. *J. Immunol. Meth.*, **134,** 237-241.

Fujii-Nakata, T., Ishimi, Y., Okuda, A. & Kikuchi, A. (1992), Functional analysis of nucleosome assembly protein, NAP-1. The negatively charged COOH-terminal region is not necessary for the intrinsic assembly activity. *J. Biochem.*, **267,** 20980-20986.

Galat A. (1993), Peptidylproline *cis-trans*-isomerases: immunophilins. *Eur. J. Biochem.*, **216,** 689-707.

Gai, D., Chang, Y.P. & Chen, X.S. (2010), Origin DNA melting and unwinding in DNA replication. *Curr. Opin. Struct. Biol.*, **20,** 756-762.

Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D. & Hunt, D.F. (2007), Organismal differences in post-translational modifications in histones H3 and H4. *J. Biol. Chem.*, **282,** 7641-55.

Garrard, W.T. & Hancock, R. (1978), Preparation of chromatin from animal tissues and cultured cells. *Meth. Cell. Biol.* **17,** 27-50.

Gasser, R., Koller, T. & Sogo, J.M. (1996), The stability of nucleosomes at the

replication fork. *J. Mol. Biol.*, **258**, 224-239.

Geankoplis, C.J. (2003), Transport Processes and Separation Process Principles, 4th edition, *Prentice Hall*, Upper Saddle River.

Gerlitz G., Hock R., Ueda T. & Bustin M. (2009), The dynamics of HMG protein-chromatin interactions in living cells. *Biochem. Cell. Biol.*, **87**, 127-137.

Ghosh, R. (2006), Principles of bioseparations engineering, *World Science Publishing Co. Pte. Ltd.*, McMaster University, Canada.

Goldberg, A.D., Allis, C.D. & Bernstein, E. (2007), Epigenetics: a landscape takes shape. *Cell,* **128,** 635-638.

Goodwin, G.H., Mathew, C.G., Wright, C.A., Venkov, C.D. & Johns, E.W. (1979), Analysis of the high mobility group proteins associated with salt-soluble nucleosomes. *Nucl. Acid. Res.* **7,** 1815-1835.

Graziani-Bowering, G.M., Graham, J. & Filion, L.G. (1997), *A quick, easy and inexpensive method for the isolation of human peripheral blood monocytes. J. Immunol. Meth.,* **207**, 157-168.

Hake, S.B., Garcia, B.A., Duncan, E.M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D.P., Allis, C.D. & Hunt, D.F. (2006), Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.,* **281,** 559-568.

Hancock, R., Faber, A.J. & Fakan, S. (1977), Isolation of interphase chromatin structures from cultured cells. *Meth. Cell. Biol.* **15,** 127-147.

Helfferich, F.G. (1962), Ion Exchange. *Science,* **138,** 133.

Hewish, D.R. & Burgoyne, L.A. (1973), Chromatin sub-structure. The Digestion of Chromatin DNA at Regularly Spaced Sites by a Nuclear Deoxyribonuclease. *Biochem. Biophys. Res. Commun.*, **52**, 504-510.

Hill, C.S. & Thomas, J.O. (1990), Core histone-DNA interactions in sea urchin sperm chromatin. The N-terminal tail of H2B interacts with linker DNA. *Eur. J. Biochem.*, **187**, 145-153.

Hock R., Furusawa T., Ueda T. & Bustin M. (2007), HMG chromosomal proteins in development and disease. *Trends.Cell. Biol.*, **17**, 72-79.

Hottiger, M.O. (2011), ADP-ribosylation of histones by ARTD1: An additional module of the histone code? *FEBS Lett.*, **585**, 1595-1599.

Huang, S., Zhou, H., Tarara, J. & Zhang, Z. (2007), A novel role for histone chaperones CAF-1 and Rtt106 in heterochromatin silencing. *EMBO. J.*, **26**, 2274-2283.

Hubbes 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.

Huth J.R., Bewley C.A., Nissen M.S., Evans J.N., Reeves R., Gronenborn A.M. & Clore G.M. (1997), The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat. Struct. Biol.*, **4**, 657-665.

Iannacone M., Sitia G., Narvaiza I., Ruggeri Z.M. & Guidotti L.G. (2007), Antiplatelet drug therapy moderates immune-mediated liver disease and inhibits viral clearance in mice infected with a replication-deficient adenovirus. *Clin. Vaccine. Immunol.*, **14**, 1532-1535.

Isackson, P.J., Debold, W.A. & Reeck G.R. (1980), Isolation and Separation of Chicken Erythrocyte High Mobility Group Non-Histone Chromatin Proteins by

Chromatography on Phosphocellulose. *Elsevier/North-Holland Biomedical Press*, **119**, 337-342.

Ishikawa, K., Ohsumi, T., Tada, S., Natsume, R., Kundu, L.R., Nozaki, N., Senda, T., Enomoto, T., Horikoshi, M. &Seki, M. (2011), Roles of histone chaperone CIA/Asf1 in nascent DNA elongation during nucleosome replication. *Genes Cells.*, **16,** 1050-1062.

Jackson, V., Shires, A., Tanphaichitr, N. & Chalkley, R. (1976), Modification to histones immediately after synthesis. *J. Mol. Biol.*, **104**, 471-483.

Johns, E. (1964), Studies on histones. *J. Biochem.*, **93**, 161-163.

Kang, C.B., Ye, H., Dhe-Paganon, S. & Yoon, H.S. (2008), FKBP Family Proteins: Immunophilins with Versatile Biological Functions. *Neurosignals.* **16**, 318-325.

Kaufman, P.D. & Btchan, M.R. (1994), Assembly of nucleosomes: do multiple assembly factors mean multiple mechanisms? *Curr.Opin.Genet. Dev.*, **4**, 229-235.

Kaufman, P.D. (1996), Nucleosome assembly: the CAF and the HAT. *Curr.Opin. Cell Biol.*, **8**, 369-373.

Kaufman, P.D., Kobayashi, R. & Stillman, B. (1997), Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-1. *Genes Dev.*, **11**, 345-357.

Kellogg, D.R., Kikuchi, A., Fujii-Nakata, T., Turck, C.W. & Murray, A.W. (1995), Members of the NAP/SET family of proteins interact specifically with B-type cyclins. *J. Cell. Biol.,* **130,** 661-673.

Khan, A.U. & Krishnamurthy, S. (2005), Histone modifications as key regulators of transcription. *Front Biosci.,* **10,** 866-872.

Kokubo, T., Gong, D.W., Wootton, J.C., Horikoshi, M. & Roeder, R.G. (1993), Molecular cloning, structural relationships and interactions of Drosophila TFIID subunits. *Nature*, **367**, 484-487.

Kornberg, R. (1974), Chromatin structure: a repeating unit of histones and DNA. *Science*, **184**, 868-871.

Kornberg, R. & Thomas, J.O. (1974), Chromatin structure: oligomers of histones. *Science*, **184**, 856-858.

Kouzarides, T. (2007), SnapShot: Histone-modifying enzymes. *Cell.*, **131**, 822.

Kouzarides, T. (2007), Chromatin Modifications and Their Functions. *J. Cell*, **128**, 693-705.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., *et al.* (2006), Global landscape of proteincomplexes in the yeast Saccharomyces cerevisiae. *Nature,* **440,** 637–643.

Krude, T. & Knippers, R. (1991), Transfer of nucleosomes from parental to replicated chromatin. *Mol. Cell Biol.*, **11**, 6257-6267.

Krude, T. (1999), Chromatin assembly during DNA replication in somatic cells. *Eur. J. Biochem.*, **263**, 1-5.

Ladisch, M.R. (2001), Bioseparations Engineering: Principles, Practice and Economics, *John Wiley and Sons*, New York.

Lambert, S.J., Nicholson, J.M., Chantalat, L., Reid, A.J., Donovan, M.J. & Baldwin, J.P. (1999), Purification of histone core octamers and 2.15 Å X-ray analysis of crystals in KCl/phosphate. *Acta Crystallogr. D. Biol. Crystallogr.*, **55,** 1048-1051.

Leclercq, M., Vinci, F. & Galat, A. (2000), Mammalian FKBP-25 and its associated proteins. *Arch. Biochem. Biophys.*, **380,** 20-28.

Lee, D.Y., Hayes, J.J., Pruss, D. & Wolffe, A.P. (1993), A positive role for histone acetylation in transcription factor binding to nucleosomal DNA. *Cell*, **72**, 73-84.

Lee, J., & Kim, S.S. (2010), Current Implications of Cyclophilins in Human Cancers. *J. Exper. Clin. Canc. Res.*, **29**, 97.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A. & Zhang, Z. (2008), Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell*, **134**, 244-255.

Linger, J. & Tyler, J.K. (2005), The yeast histone chaperone chromatin assembly factor 1 protects against double-strand DNA-damaging agents. *Genet.*, **171**, 1513-1522.

Lorch, Y., Maier-Davis, B. & Kornberg, R.D. (2006), Chromatin remodeling by nucleosome disassembly in vitro. *Proc. Natl. Acad. Sci. USA*, **103,** 3090–3093.

Lotze M.T. & Tracey K.J. (2005), High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat. Rev. Immunol.*, **5**, 331-342.

Luger, K., *Maeder, 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. (1978), Kinetic analysis of deoxyribonuclease I cleavage sites in the nucleosome core: evidence for a DNA superhelix. *J. Mol. Biol.*, **124**, 391-420.

Mahadevan, L.C., Willis, A.C. & Barrah, M.J. (1991), Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid and protein synthesis inhibitors. *Cell*, **65**, 775-783.

Marahrens, Y. & Stillman, B. (1992), A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science*, **255**, 817-823.

Mark, E.E. (2000), Where the cell cycle and histones meet. *Genes & Dev.*, **14**, 2265-2270.

McQuibban, G.A., Commisso-Cappelli, C.N. & Lewis, P.N. (1998), Assembly, remodelling, and histone binding capabilities of yeast nucleosome assembly protein 1. *J. Biol.Chem.*, **273,** 6582-6590.

Minaga T., Romaschin A.D., Kirsten E. & Kun E. (1979), The in vivo distribution of immunoreactive larger than tetrameric polyadenosine diphosphoribose in histone and non-histone protein fractions of rat liver. *J. Biol. Chem.*, **254**, 9663-9668.

Messner S., Altmeyer M., Zhao H., Pozivil A., Roschitzki B., Gehrig P., Rutishauser D., Huang D., Caflisch A. & Hottiger M.O. (2010), PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucl. Acid. Res.*, **38**, 6350-6362.

Miwa, M., Kanai, M., Uchida, M., Uchida, K. & Hanai, S. (2006), Roles of Poly(ADP-Ribose) Metabolism in the Regulation of Centrosome Duplication and in the Maintenance of Neuronal Integrity. *Land. Biosc. Mad. Cur. Biosc. Dat.*

Morgan, D.O. (2007), The Cell Cycle. *New Science Press. U.S.A.*

Moudrianakis, E.N., Love, W.E., Wang, B.C., Xuong, N.G. and Burlingame, R.W. (1985), Crystallographic structure of the octamer histone core of the nucleosome. *Science*, **229,** 1110-1112.

Mosammaparast, N., Ewart, C.S. & Pemberton, L.F. (2002), A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. *EMBO J.*, **21,** 6527-6538.

Mosammaparast, N., Jackson, K.R., Guo, Y., Brame, C.J., Shabanowitz, J., Hunt, D.F. & Pemberton, L.F. (2001), Nuclear import of histone H2A and H2B is mediated by a network of karyopherins. *J. Cell Biol.,* **153,** 251–262.

Mousson, F., Lautrette, A., Thuret, J.Y., Agez, M., Courbeyrette, R., Amigues, B., Becker, E., Neumann, J.M., Guerois, R., Mann, C. & Ochsenbein, F. (2005), Structural basis for the interaction of Asf1 with histone H3 and its functionalimplications. *Proc. Natl. Acad. Sci. USA,* **102,** 5975–5980.

Murray, K. (1966), The acid extraction of histones from calf thymus deoxyribonucleoprotein. *J. Mol. Biol.,* **15,** 409-419.

Nakatani, Y., Bagby, S. & Ikura, M. (1996), The histone folds in transcription factor TFIID. *J. Mol. Biol.*, **271**, 6575-6578.

Nathan D., Sterner D.E. & Berger S.L. (2003), Histone modifications: Now summoning sumoylation. *Proc. Natl. Acad. Sci. USA*, **100**, 13118-13120.

Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M. &Senda, T. (2007), Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature,* **446,** 338-41.

Nelson, C.J., Sanos-Rosa, H., & Kouzarides, T. (2006), Proline isomerization of Histone H3 regulates lysine methylation and gene expression. *Cell.*, **126**, 905-916.

Newman, E. (2011), PhD thesis, University of Portsmouth.

Nguyen, V.Q., Co, C. & Li, J.J. (2001), Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature,* **411**, 1068-1073.

Nightingale K., Dimitrov S., Reeves R. & Wolffe A.P. (1996), Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. *EMBO J.*, **15**, 548-561.

Nightingale, K.P., O'Neill, L.P. & Turner, B.M. (2006), Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. *Curr. Opin. Genet. Dev.*, **16,** 125-136.

Ogata N., Ueda K. & Hayaishi O. (1980), ADP-ribosylation of histone H2B. Identification of glutamic acid residue 2 as the modification site. *J. Biol. Chem.*, **255**, 7610-7615.

Okuwaki, M., Kato, K., Shimahara, H., Tate, S. & Nagata, K. (2005), Assembly and disassembly of nucleosome core particles containing histone variants by human nucleosome assembly protein 1. *Mol. Cell. Biol.*, **25,** 10639-10651.

Okuzaki, D., Kimura, S., Yabuta, N., Ohmine, T. & Nojima H. (2011), LeukoCatch, a quick and efficient tool for the preparation of leukocyte extracts from blood. *BMC Clin. Pathol.*, **11,** 9.

Olins, A.L., and Olins, D. E. (1974), Spheroid chromatin units (v bodies). *Science,* **183**, 330–332.

Otake H., Miwa M., Fujimura S. & Sugimura T. (1969), Binding of ADP-ribose polymer with histone. *J. Biochem.*, **65**, 145-146.

Park, Y.J. & Luger, K. (2006), The structure of nucleosome assembly protein 1. *Proc. Natl. Acad. Sci. U.S.A.*,**103,** 1248-1253.

Park, Y.J., Sudhoff, K.B., Andrews, A.J., Stargell, L.A.& Luger, K. (2008), Histone chaperone specificity in Rtt109 activation. *Nat. Struct. Mol. Biol.*, **15,** 957–964.

Quesniaux, V., Erard, F. & Ryffel, B. (2010), *Adjuvant activity on murine and human macrophages* In Vaccine Adjuvants (ed. Davies, G.). *Methods Mol. Biol.*, **626**, 117-130.

Read C.M., Cary P.D., Crane-Robinson C., Driscoll P.C. & Norman D.G. (1993), Solution structure of a DNA-binding domain from HMG1. *Nucleic. Acids. Res.*, **21**, 3427-3436.

Recht, J., Tsubota, T., Tanny, J.C., Diaz, R.L., Berger, J.M., Zhang, X., Garcia, B.A., Shabanowitz, J., Burlingame, A.L., Hunt, D.F., *et al.* (2006), Histone chaperone Asf1 is required for histone H3 Lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc. Natl. Acad. Sci. U S A*, **103**, 6988-6993.

Reeves R. & Nissen M.S. (1990), The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins.A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.*, **265**, 8573-8582.

Reeves R. & Beckerbauer L. (2001), HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochem. Biophys. Acta.*, **1519**, 13-29.

Reynolds, C., Nicholson, J., & Baldwin, J.P. (2008), Chromatin: Cracking the Chemical Code. *Eur. Pharm. Rev.*, **4**, 75-81.

Ruiz-Carillo, A., Wangh, L.J. & Allfrey, V.G. (1975), Processing of newly synthesized histone molecules. *Science*, **190**, 117-128.

Santos-Rosa, H., Krimizis, A., Nelson, C., Bartke, T., Saksouk, N., Cote, J. & Kouzarides, T. (2008), Histone H3 tail clipping regulates gene expression. *Nat. Struct. Mol. Biol.*, **16**, 17-22.

Satoh, K., Shimokawa, H., & Bradford C. (2010), Cyclophilin A – Promising New Target in Cardiovascular Therapy. *Circ. J.*, **74**, 2249-2256.

Shechter, D., Dormann, H.L., Allis, C.D. & Hake, S.B. (2007), Extraction, purification and analysis of histones. *Nat. Protoc.*, **2,** 1445-57.

Shiio Y., & Eisenman R.N. (2003), Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. USA*, **100**, 13225-13230.

Simon, R.H. &Felsenfeld, G. (1979), A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxylapatite. *Nucl. Acid. Res.*, **6,** 689-696.

Simpson, R.T. (1978), Structure of the chromatosome, a chromatin core particle containing 160 base pairs of DNA and all the histones. *J. Bio. Chem.*, **17**, 5524-5531.

Snijders, A.P.L., Pongdam, S., Lambert, S.J., Wood, C.M., Baldwin, J.P. & Dickman, M.J. (2008), Chracterisation of Post-Translational Modifications of the Linker Histones H1 and H5 from Chicken Erythrocytes Using Mass Spectrometry. *J. Biol. Chem.*, **10**, 4326-4335.

Sodngam, S. (2007), A Coherent procedure for fractionating proteins from chick erythrocyte cell nuclei. *PhD Thesis,* Liverpool John Moores University.

Speck, C., Chen, Z., Li, H. & Stillman, B. (2005), ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat. Struct. Mol. Biol.*, **12**, 965-971.

Steger, D.J. & Workman, J.L. (1996), Remodeling chromatin structures for transcription: what happens to the histones? *Bioess.*, **18**, 875-884.

Stillman, B. (1986), Chromatin assembly during SV40 DNA replication *in vitro*. *Cell*, **45**, 555-565.

Stillman, B. (2005), Origin Recognition and Chromosome Cycle. *FEBS Lett.*, **579,** 877-884.

Stros, M. (2010), HMGB proteins: interactions with DNA and chromatin. *Biochem. Biophys. Acta.*, **1799,** 101-113.

Stros, M.&Kolíbalová, A.(1987), Interaction of non-histone proteins HMG1 and HMG2 with core histones in nucleosomes and core particles revealed by chemical cross-linking. *Eur. J. Biochem.*, **162,** 111-118.

Suau, P, Bradbury E.M. & Baldwin J.P. (1979), Higher-order structures of chromatin in solution. *Eur. J. Biochem.*, **97**, 593-602.

Sun, J.Y., & Kong, D.C. (2010), DNA replication origins, ORC/DNA interaction, and assembly of pre-replication complex in eukaryotes. *Akta. Biochem. Biophys. Sci.*, **42**, 433-439.

Sung, M.T. & Dixon, G.H. (1970), Modification of histones during spermiogenesis in trout: a molecular mechanism for altering histone binding to DNA.*Proc. Natl Acad. Sci. USA*, **67**, 1616-1623.

Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabanowitz, J., Chait, B.T., Hunt, D.F. & Allis, C.D. (2007), Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc. Natl. Acad. Sci. U S A.*, **104,** 2086-2091.

Thoma, F., Koller, T. & Klug, A. (1979), Involvement of histone H1 in the organization of the nucleosome and the salt-dependent superstructures of chromatin. *J. Cell. Biol.*, **83**, 402-427.

Thomas J.O. (2001), HMG1 and 2: architectural DNA-binding proteins. *Biochem. Soc. Trans.*, **29**, 395-401.

Thomas, J.O.& Travers, A.A. (2001), HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends. Biochem. Sci.*, **26,** 167-174.

Turner, B.M. (2001), Chromatin and gene regulation.*Blackwell Science Ltd.*Tyler, J.K. (2002), Chromatin assembly. Cooperation between histone chaperones and ATP-dependent nucleosome remodelling machines. *Eur. J. Biochem.*, **269**, 2268-2274.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T.&Kadonaga, J.T. (1999), The RCAF complex mediates chromatin assemblyduring DNA replication and repair. *Nature,* **402,** 555–560.

Ueda K., Omachi A., Kawaichi M. & Hayaishi O. (1975), Natural occurrence of poly(ADP-ribosyl) histones in rat liver. *Proc. Natl. Acad. Sci. USA*, **72**, 205-209.

Ura K., Nightingale K. & Wolffe A.P. (1996), Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. *EMBO J.*, **15**, 4959-4969.

Vargason, J.M., Eichman B.F. & Ho P.S. (2000), The extended and eccentric E-DNA structure induced by cytosine methylation or bromination. *Nat. Struct. Biol.*, **7**, 758-761.

Verger, A., Perdomo J. & Crossley M. (2003), Modification with SUMO.A role in transcriptional regulation. *EMBO Rep.*, **4**, 137-142.

Vettese-Dadey, M., Walter, P., Chen, H., Juan, L-J.& Workman, J.L. (1994), Role of the histone amino termini in facilitated binding of a transcription factor, GAL4-AH to nucleosome cores. *Mol. Cell. Biol.*, **14**, 970-981.

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.

Watson, J.D. & Crick, F.H.C. (1953), A structure for deoxyribosenucleic acids. *Nature (London),* **171**, 737-738.

Weir H.M., Kraulis P.J., Hill C.S., Raine A.R., Laue E.D. & Thomas J.O. (1993), Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.*, **12**, 1311-1319.

Wheelwright, S.M. (1991), Protein Purification, *John Wiley and Sons*, New York.

Willianmson, R. (1970), Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *J. Mol. Biol.*, **51**, 157-168.

Wolffe, A. (1999), Chromatin Structure and Function 3rd Edition.*National Institutes of Health, Bethesda, Maryland, USA. Academic press.*

Wood C.M., Nicholson, J.M., Lambert, S.J., Chantalat, L., Reynolds, C.D. & Baldwin J.P. (2005), High-resolution structure of the native histone octamer. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **61**, 541-545.

Xie, X., Kokubo, T., Cohen, S.L., Mirza, U.A., Hoffmann, A., Chait, B.T., Roeder, R.G., Nakatani, Y. & Burley, S.K. (1996), Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature*, **380**, 316-322.

Zeng, B., Macdonald, J.R., Bann, J.G., Beck, K., Gambee, J.E., Boswell, B.A. & Bachinger, H.P. (1998), Chicken FK506-binding Protein, FKBP65, A Member if the FKBP Family of Peptidylprolyl *cis-trans* Isomerases, Is Only Partially Inhibited by FK506. *J. Biochem.*, **330**, 109-114.

Appendix 1

Isolation of chick erythrocyte nuclei

Introduction

Chick erythrocyte cell nuclei are one of the main resources of our experiment. Although chick's cells and nuclei may behave differently compared with mammalian's cells and nuclei, there are still many advantages to use chick erythrocyte nuclei. Firstly chick erythrocyte cells contain nuclei which mammalian erythrocyte cells don't, and the yield of erythrocyte cells is much greater than leukocyte cells. Secondly the nucleosome structure and histone proteins which we are focusing on are stably inherited in different species. Finally chick blood as the resource of erythrocyte cells is very cheap and easy to obtain which allow us to do large amount of works on it.

Chick erythrocyte nuclei preparation is the first step of our purification of chick histone octamers, dimers, tetramers and other associated nuclei proteins. Our group developed this method to purify the nuclei from fresh chick blood collected from chicken farm.

Method:

Buffers:

(All buffers stored at 4°C before use)

1. Chick blood anti clotting buffer: 2 L of 10% Tris-sodium Citrate.

2. Cell lysis buffer: 20 L of 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$, 1.25% Triton X-100.

3. Detergent nuclei washing buffer: 30 L of 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$, 0.5% Triton X-100.

4. Nuclei washing buffer: 30 L of 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$.

5. Nuclei freezing buffer: 10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$, 50% glycerol.

Procedure:

Twelve litres of chicken blood was collected from the chicken farm. 10% Tris-sodium citrate was added into the blood at a ratio of 100 ml buffer per litre of blood to stop the blood clotting and the blood was allowed to cool slowly on the way to the laboratory. Then the blood was filtered through two layers of butter muslin which pre-wetted with 10% Tris-sodium citrate to filter out most of the leukocyte cells. Then the blood was partitioned into twelve 1 litre centrifuge tubes and centrifuged at 3000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes to separate the cells. After centrifugation the buffy coat was sucked off , the erythrocytes were pool together and lysed in cell lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$, 1.25% Triton X-100) to a final volume of 20 litres. The erythrocytes with lysis buffer were stirred at 4 °C for one hour to allow the cells to be lysed.

After lysis the lysed cells were centrifuged at 2500 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes to separate the nuclei and other materials. Then the supernatant was sucked off down to the pellet, the nuclei pellets were pool together and washed in detergent nuclei washing buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM $MgCl_2$, 0.5% Triton X-100) with a ratio of 1/3 nuclei to 2/3 buffer and then the nuclei were centrifuged at 2000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes and the supernatant was removed. This stage is to lyse any unlysed cells remain in the nuclei pellet; it might be repeat one or two more times to lyse all the cells, each time the nuclei pellets were dispersed by squirting through the nozzle of a 20 ml syringe.

Then the nuclei were washed with nuclei washing buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3.5 mM MCl_2, 2.5 mM benzamidine hydrochloride) by suspend the nuclei into the buffer with a ratio of 1/3 nuclei to 2/3 buffer and then the nuclei were centrifuged at 2000 rpm (Beckman Acanti® J-26 XP centrifuge, JS 13.1 rotor) for 15 minutes and the supernatant was removed. This step is to remove the haemoglobin remain in the nuclei; it might be repeated several times until the nuclei pellets become pure white. Then the nuclei were dispersed by squirting through the nozzle of a 20 ml syringe and mixed with an equal volume of nuclei freezing buffer. Finally the nuclei were partitioned into 50 ml tubes and flash freezed by liquid N_2. They can then be stored in -80 °C freezer ready for further purification of nuclei proteins.

Appendix 2

cePNE1-protein identification by nanospray mass spectrometry

Note that our methods extract gram quantities of the mixture of proteins so that purification of significant quantities of the most abundant proteins will be possible for mass spectrometry and structural studies.

	Proteins (accession number/molecular weight/score/queries matched/name)	Category of protein
1.	gi|63556 Mass: 73120 Score: 3820 Queries matched: 409	19
	unnamed protein product [Gallus gallus] lamin A	
2.	gi|50748706 Mass: 545623 Score: 2901 Queries matched: 99	2
	PREDICTED: similar to Dynein heavy chain, cytosolic (DYHC) (Cytoplasmic dynein heavy chain 1) (DHC1) [Gallus gallus]	
3.	gi|50751217 Mass: 267232 Score: 2829 Queries matched: 87	9
	PREDICTED: similar to Tpr [Gallus gallus]	
4.	gi|53136476 Mass: 100168 Score: 2576 Queries matched: 95	2
	hypothetical protein [Gallus gallus] methylenetetrahydrofolate dehydrogenase	
5.	gi|13384736 Mass: 531690 Score: 2403 Queries matched: 83	17
	dynein, cytoplasmic, heavy chain 1 [Mus musculus]	
6.	gi|2516264 Mass: 81360 Score: 2262 Queries matched: 100	2
	muCL [Gallus gallus]	
7.	gi|37590083 Mass: 69870 Score: 1942 Queries matched: 86	3
	heat shock protein Hsp70 [Gallus gallus]	
8.	gi|63560 Mass: 67900 Score: 1919 Queries matched: 96	19
	unnamed protein product [Gallus gallus] lamin B2	
9.	gi|29171734 Mass: 97146 Score: 1907 Queries matched: 84	10
	eukaryotic translation initiation factor 2C, 2 [Homo sapiens]	
10.	gi|50731932 Mass: 96379 Score: 1873 Queries matched: 84	10
	PREDICTED: similar to eukaryotic translation initiation factor 2C, 2; argonaute 2 [Gallus gallus]	
11.	gi|45544523 Mass: 70827 Score: 1707 Queries matched: 68	3
	heat shock protein [Numida meleagris]	
12.	gi|57525483 Mass: 51044 Score: 1654 Queries matched: 80	9
	similar to eukaryotic translation initiation factor 2, subunit 3, structural gene X-linked [Gallus gallus]	
13.	gi|57529969 Mass: 33038 Score: 1622 Queries matched: 97	10
	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa [Gallus gallus]	
14.	gi|3122072 Mass: 50125 Score: 1597 Queries matched: 452	10
	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor Tu) (EF-Tu)	
15.	gi|57529974 Mass: 58573 Score: 1596 Queries matched: 69	13
	apoptosis inhibitor 5 [Gallus gallus]	
16.	gi|6753738 Mass: 51033 Score: 1538 Queries matched: 71	10
	eukaryotic translation initiation factor 2, subunit 3, structural gene X-linked [Mus musculus]	
17.	gi|50759866 Mass: 167812 Score: 1429 Queries matched: 48	1
	PREDICTED: similar to Microtubule-actin crosslinking factor 1, isoforms 1/2/3 (Actin cross-linking family protein 7) (Macrophin 1) (Trabeculin-alpha) (620 kDa actin-binding protein) (ABP620) [Gallus gallus]	
18.	gi|50748798 Mass: 74889 Score: 1371 Queries matched: 49	2
	PREDICTED: similar to Mthfd1-prov protein [Gallus gallus]	

19.	gi|45382971 Mass: 38412 Score: 1366 Queries matched: 64	10
	eukaryotic translation initiation factor 2, subunit 2 beta, 38kDa [Gallus gallus]	
20.	gi|56377788 Mass: 50123 Score: 1361 Queries matched: 347	9
	elongation factor-1 alpha (EF-1alpha) [Pelodiscus sinensis]	
21.	gi|4758256 Mass: 36089 Score: 1359 Queries matched: 84	10
	eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa [Homo sapiens]	
22.	gi|7242148 Mass: 51098 Score: 1280 Queries matched: 54	10
	eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked [Mus musculus]	
23.	gi|53136508 Mass: 53174 Score: 1275 Queries matched: 87	1
	hypothetical protein [Gallus gallus] coronin, actin binding protein, 1C	
24.	gi|71896353 Mass: 76972 Score: 1253 Queries matched: 44	7
	heterogeneous nuclear ribonucleoprotein M [Gallus gallus]	
25.	gi|45383133 Mass: 145766 Score: 1253 Queries matched: 32	2
	SMC4 structural maintenance of chromosomes 4-like 1 [Gallus gallus]	
26.	gi|27881747 Mass: 51265 Score: 1252 Queries matched: 53	10
	Eif2s3-prov protein [Xenopus laevis]	
27.	gi|50768419 Mass: 40824 Score: 1239 Queries matched: 67	24
	PREDICTED: similar to proliferation-associated protein 1, partial [Gallus gallus]	
28.	gi|50759864 Mass: 282157 Score: 1220 Queries matched: 34	1
	PREDICTED: similar to Microtubule-actin crosslinking factor 1, isoform 4 [Gallus gallus]	
29.	gi|50745996 Mass: 188082 Score: 1212 Queries matched: 47	12
	PREDICTED: similar to hypothetical protein [Gallus gallus] stromal antigen 2	
30.	gi|1362713 Mass: 67266 Score: 1206 Queries matched: 57	2
	calpain (EC 3.4.22.17) large chain 1 - chicken (fragment)	
31.	gi|6179570 Mass: 68512 Score: 1175 Queries matched: 75	1
	radixin [Gallus gallus]	
32.	gi|50759868 Mass: 170652 Score: 1169 Queries matched: 48	1
	PREDICTED: similar to Microtubule-actin crosslinking factor 1, isoform 4 [Gallus gallus]	
33.	gi|60302698 Mass: 105393 Score: 1151 Queries matched: 42	2
	TAO kinase 3 [Gallus gallus]	
34.	gi|6273778 Mass: 613652 Score: 1128 Queries matched: 39	1
	trabeculin-alpha [Homo sapiens]	
35.	gi|50758833 Mass: 388014 Score: 1114 Queries matched: 33	9
	PREDICTED: similar to PPAR-alpha interacting complex protein 285 [Gallus gallus]	
36.	gi|60098709 Mass: 143760 Score: 1088 Queries matched: 43	
	hypothetical protein [Gallus gallus] PDS5, regulator of cohesion maintenance, homolog B	
37.	gi|8118590 Mass: 105224 Score: 1087 Queries matched: 41	
	KFC [Gallus gallus]	
38.	gi|50750286 Mass: 59229 Score: 1085 Queries matched: 26	8
	PREDICTED: similar to Nucleolar protein NOP5 (Nucleolar protein 5) (NOP58) (HSPC120) [Gallus gallus]	
39.	gi|2832596 Mass: 72326 Score: 1077 Queries matched: 80	
	OTTHUMP00000028920 [Homo sapiens]	
40.	gi|29436810 Mass: 36152 Score: 1034 Queries matched: 51	10
	Eukaryotic translation initiation factor 2, subunit 1 alpha [Danio rerio]	
41.	gi|15824727 Mass: 669735 Score: 1024 Queries matched: 36	1
	macrophin 1 isoform 4 [Homo sapiens]	
42.	gi|33526989 Mass: 41768 Score: 1023 Queries matched: 54	1
	beta-actin [Monopterus albus]	
43.	gi|63711 Mass: 75594 Score: 1019 Queries matched: 41	8
	unnamed protein product [Gallus gallus] nucleolin	
44.	gi|50759792 Mass: 45614 Score: 988 Queries matched: 49	
	PREDICTED: similar to PTB-associated splicing factor [Gallus gallus]	
45.	gi|1083569 Mass: 51010 Score: 969 Queries matched: 33	
	kappa-B motif-binding phosphoprotein - mouse	
46.	gi|50754301 Mass: 186642 Score: 968 Queries matched: 29	18
	PREDICTED: similar to Nup210 protein [Gallus gallus]	
47.	gi|53127738 Mass: 118615 Score: 966 Queries matched: 39	2
	hypothetical protein [Gallus gallus]ubiquitin specific peptidase 48	
48.	gi|17318569 Mass: 66027 Score: 959 Queries matched: 69	19

	keratin 1 [Homo sapiens]	
49.	gi|71896565 Mass: 87658 Score: 947 Queries matched: 30	2
	similar to Putative pre-mRNA splicing factor RNA helicase (DEAH box protein 15) [Gallus gallus]	
50.	gi|33415848 Mass: 41752 Score: 928 Queries matched: 56	1
	cytoplasmic actin type 5 [Rana lessonae]	
51.	gi|46048916 Mass: 25016 Score: 921 Queries matched: 50	12
	FK506 binding protein 3, 25kDa [Gallus gallus]	
52.	gi|50749186 Mass: 29826 Score: 905 Queries matched: 28	
	PREDICTED: similar to DnaJ homolog, subfamily C, member 9; DnaJ protein SB73 [Gallus gallus]	
53.	gi|71896967 Mass: 23448 Score: 901 Queries matched: 112	
	hypothetical protein LOC425058 [Gallus gallus]	
54.	gi|60098563 Mass: 42138 Score: 897 Queries matched: 48	
	hypothetical protein [Gallus gallus] DEK oncogene (DNA binding)	
55.	gi|74004805 Mass: 49853 Score: 882 Queries matched: 277	10
	PREDICTED: similar to eukaryotic translation elongation factor 1 alpha 1 [Canis familiaris]	
56.	gi|55956899 Mass: 62027 Score: 876 Queries matched: 57	19
	keratin 9 [Homo sapiens]	
57.	gi|50755234 Mass: 126319 Score: 871 Queries matched: 46	9
	PREDICTED: similar to transcription elongation regulator 1; transcription factor CA150; TATA box binding protein (TBP)-associated factor, RNA polymerase II, S, 150kD; TATA box-binding protein-associated factor 2S [Gallus gallus]	
58.	gi|5815432 Mass: 24778 Score: 868 Queries matched: 184	5
	high mobility group protein HMG1 [Gallus gallus]	
59.	gi|63494 Mass: 23056 Score: 860 Queries matched: 251	5
	HMG2a [Gallus gallus]	
60.	gi|50750688 Mass: 60638 Score: 858 Queries matched: 26	2
	PREDICTED: similar to Protein disulfide isomerase A5 precursor (Protein disulfide isomerase-related protein) [Gallus gallus]	
61.	gi|63558 Mass: 66490 Score: 858 Queries matched: 23	19
	unnamed protein product [Gallus gallus] lamin B1	
62.	gi|391636 Mass: 22914 Score: 858 Queries matched: 281	5
	HMG-1 [Gallus gallus]	
63.	gi|53126280 Mass: 89269 Score: 855 Queries matched: 24	
	hypothetical protein [Gallus gallus]	
64.	gi|496902 Mass: 46803 Score: 854 Queries matched: 23	10
	translation initiation factor [Homo sapiens]	
65.	gi|57530393 Mass: 33962 Score: 849 Queries matched: 24	9
	similar to transcription elongation factor IIS - mouse [Gallus gallus]	
66.	gi|50746196 Mass: 41423 Score: 846 Queries matched: 34	
	PREDICTED: similar to hypothetical protein FLJ20534 [Gallus gallus] hypothetical protein	
67.	gi|61097989 Mass: 104127 Score: 837 Queries matched: 37	
	adaptor-related protein complex 2, alpha 2 subunit [Gallus gallus]	
68.	gi|71897277 Mass: 47249 Score: 836 Queries matched: 32	7
	hypothetical protein LOC426516 [Gallus gallus]heterogeneous nuclear ribonucleoprotein K	
69.	gi|50753573 Mass: 141198 Score: 833 Queries matched: 22	
	PREDICTED: similar to KIAA0017 protein [Gallus gallus]	
70.	gi|51704021 Mass: 68976 Score: 832 Queries matched: 53	
	LOC398649 protein [Xenopus laevis]	
71.	gi|45382027 Mass: 22399 Score: 810 Queries matched: 53	2
	peptidylprolyl isomerase B (cyclophilin B) [Gallus gallus]	
72.	gi|2506056 Mass: 23900 Score: 801 Queries matched: 52	2
	calpain small subunit [Gallus gallus]	
73.	gi|51859448 Mass: 104108 Score: 800 Queries matched: 34	
	Adaptor protein complex AP-2, alpha 2 subunit [Rattus norvegicus]	
74.	gi|57524964 Mass: 50164 Score: 797 Queries matched: 20	
	RuvB-like 1 [Gallus gallus]	
75.	gi|53136648 Mass: 48972 Score: 787 Queries matched: 31	2
	hypothetical protein [Gallus gallus]Spliceosome RNA helicase BAT1 (DEAD box protein UAP56) (56 kDa U2AF65-associated protein).	
76.	gi|62738728 Mass: 88891 Score: 787 Queries matched: 22	

	Chain C, VcpP97 COMPLEXED WITH ADP	
77.	gi|57529634 Mass: 65852 Score: 786 Queries matched: 25	2
	phenylalanine-tRNA synthetase-like, beta subunit [Gallus gallus]	
78.	gi|53129083 Mass: 42956 Score: 772 Queries matched: 28	2
	hypothetical protein [Gallus gallus] flap structure-specific endonuclease 1	
79.	gi|5453842 Mass: 43785 Score: 771 Queries matched: 40	24
	proliferation-associated 2G4, 38kDa [Homo sapiens]	
80.	gi|50756645 Mass: 40435 Score: 756 Queries matched: 21	
	PREDICTED: similar to GTP-binding protein [Gallus gallus]	
81.	gi|50748121 Mass: 153598 Score: 746 Queries matched: 29	15
	PREDICTED: similar to KIAA0097 gene product; colonic and hepatic tumor over-expressed [Gallus gallus]	
82.	gi|50806270 Mass: 284058 Score: 739 Queries matched: 44	20
	PREDICTED: similar to ankyrin [Gallus gallus]	
83.	gi|600761 Mass: 24848 Score: 729 Queries matched: 141	5
	high mobility group 1 protein	
84.	gi|45382755 Mass: 23813 Score: 726 Queries matched: 88	5
	non-histone chromosomal protein [Gallus gallus]chicken HMGB2	
85.	gi|12004634 Mass: 51155 Score: 718 Queries matched: 22	
	reptin [Xenopus laevis]	
86.	gi|50758733 Mass: 236832 Score: 716 Queries matched: 21	2
	PREDICTED: similar to transglutaminase y [Gallus gallus]	
87.	gi|28317 Mass: 59492 Score: 711 Queries matched: 31	
	unnamed protein product [Homo sapiens]	
88.	gi|55741612 Mass: 50151 Score: 710 Queries matched: 37	10
	thymopoietin [Gallus gallus]	
89.	gi|50747169 Mass: 117573 Score: 694 Queries matched: 20	
	PREDICTED: similar to chromosome condensation protein G, partial [Gallus gallus]	
90.	gi|4098678 Mass: 210415 Score: 690 Queries matched: 61	9
	C1 transcription factor [Mus musculus]	
91.	gi|45383053 Mass: 82432 Score: 686 Queries matched: 20	
	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1 [Gallus gallus]	
92.	gi|4139170 Mass: 49079 Score: 685 Queries matched: 264	10
	elongation translation factor 1 alpha [Cyanophora paradoxa]	
93.	gi|23712 Mass: 45220 Score: 682 Queries matched: 34	12
	myoblast antigen 24.1D5 [Homo sapiens]	
94.	gi|4699643 Mass: 16325 Score: 679 Queries matched: 71	4
	Chain D, R-State Form Of Chicken Hemoglobin D	
95.	gi|46048961 Mass: 35681 Score: 676 Queries matched: 20	2
	glyceraldehyde-3-phosphate dehydrogenase [Gallus gallus]	
96.	gi|55420670 Mass: 45150 Score: 673 Queries matched: 181	10
	elongation factor-1 alpha [Chloreuptychia arnaca]Eukaryotic translation elongation factor 1 alpha 1	
97.	gi|71895845 Mass: 54033 Score: 669 Queries matched: 29	5
	similar to non-POU domain containing, octamer-binding; Nuclear RNA-binding protein, 54-kD; non-Pou domain-containing octamer (ATGCAAAT) binding protein; non-POU-domain-containing, octamer-binding [Gallus gallus]	
98.	gi|211929 Mass: 23755 Score: 646 Queries matched: 70	5
	high-mobility group-2 protein	
99.	gi|71894943 Mass: 51136 Score: 641 Queries matched: 19	
	similar to BAG-family molecular chaperone regulator-5 (BAG-5) [Gallus gallus]	
100.	gi|57529813 Mass: 37256 Score: 640 Queries matched: 34	
	BUB3 budding uninhibited by benzimidazoles 3 homolog [Gallus gallus]	
101.	gi|40018610 Mass: 244392 Score: 629 Queries matched: 17	
	U5 snRNP-specific protein, 200 kDa [Mus musculus]	
102.	gi|32693297 Mass: 41203 Score: 623 Queries matched: 128	10
	translation elongation factor 1-alpha [Saccharomyces cerevisiae]	
103.	gi|45383127 Mass: 271672 Score: 614 Queries matched: 20	20
	talin 1 [Gallus gallus]	
104.	gi|50779470 Mass: 35374 Score: 613 Queries matched: 27	1
	PREDICTED: similar to plectin, partial [Gallus gallus]	

105.	gi|50755565 Mass: 17015 Score: 610 Queries matched: 39	
	PREDICTED: similar to G10 protein homolog (EDG-2) [Gallus gallus]	
106.	gi|60302704 Mass: 56901 Score: 604 Queries matched: 15	18
	nucleoporin 93kDa [Gallus gallus]	
107.	gi|71896957 Mass: 64463 Score: 602 Queries matched: 19	
	similar to CDA02 protein [Gallus gallus]	
108.	gi|50760614 Mass: 80076 Score: 591 Queries matched: 155	17
	PREDICTED: similar to microtubule-associated protein tau; Tau microtubule-associated protein [Gallus gallus]	
109.	gi|58865420 Mass: 93243 Score: 588 Queries matched: 16	18
	nucleoporin 93 (predicted) [Rattus norvegicus]	
110.	gi|63524 Mass: 53607 Score: 576 Queries matched: 34	12
	unnamed protein product [Gallus gallus] Ig gamma chain (clone 36) - chicken (fragment)	
111.	gi|50748193 Mass: 177312 Score: 567 Queries matched: 14	
	PREDICTED: similar to pincher; pinocytic chaperone [Gallus gallus]	
112.	gi|27436951 Mass: 67647 Score: 560 Queries matched: 33	19
	lamin B2 [Homo sapiens]	
113.	gi|71897255 Mass: 59796 Score: 560 Queries matched: 16	8
	hypothetical protein LOC426574 [Gallus gallus] nucleolar protein 5A (56kDa with KKE/D repeat)	
114.	gi|60302806 Mass: 152988 Score: 560 Queries matched: 19	
	similar to KIAA0648 protein [Gallus gallus]	
115.	gi|50758110 Mass: 36944 Score: 559 Queries matched: 13	2
	PREDICTED: similar to malate dehydrogenase, mitochondrial; malate dehydrogenase 2; Malate dehydrogenase 2 NAD (mitochondrial) [Gallus gallus]	
116.	gi|71895795 Mass: 20837 Score: 552 Queries matched: 11	
	hypothetical protein LOC428588 [Gallus gallus] hypothetical protein LOC428588	
117.	gi|181402 Mass: 65825 Score: 552 Queries matched: 18	19
	epidermal cytokeratin 2 [Homo sapiens]	
118.	gi|71897251 Mass: 45199 Score: 551 Queries matched: 46	
	MHC B-G antigen isoform 2 precursor [Gallus gallus]	
119.	gi|50749292 Mass: 314115 Score: 546 Queries matched: 22	13
	PREDICTED: similar to hypothetical protein FLJ10839; cell-cycle and apoptosis regulatory protein 1 [Gallus gallus]	
120.	gi|125090 Mass: 46645 Score: 544 Queries matched: 14	19
	Keratin, type I microfibrillar 48 kDa, component 8C-1 (Low-sulfur keratin)	
121.	gi|40788963 Mass: 94570 Score: 542 Queries matched: 13	
	KIAA0095 [Homo sapiens]	
122.	gi|53127674 Mass: 70507 Score: 531 Queries matched: 14	
	hypothetical protein [Gallus gallus]chicken hypothetical protein	
123.	gi|53136520 Mass: 56365 Score: 529 Queries matched: 13	
	hypothetical protein [Gallus gallus]chicken hypothetical protein	
124.	gi|9789893 Mass: 47399 Score: 527 Queries matched: 13	
	BRG1/brm-associated factor 53A [Mus musculus]	
125.	gi|45382089 Mass: 50121 Score: 525 Queries matched: 17	
	tubulin, beta 3 [Gallus gallus]	
126.	gi|53133424 Mass: 49519 Score: 520 Queries matched: 13	2
	hypothetical protein [Gallus gallus] Phenylalanyl-tRNA synthetase alpha chain (Phenylalanine - tRNA ligase alpha chain) (PheRS).	
127.	gi|27124646 Mass: 65045 Score: 517 Queries matched: 26	3
	heat shock protein 70 [Crassostrea gigas]	
128.	gi|1098573 Mass: 41885 Score: 515 Queries matched: 25	1
	actin	
129.	gi|52138645 Mass: 15685 Score: 512 Queries matched: 21	
	adult alpha D globin [Gallus gallus]	
130.	gi|61857876 Mass: 54596 Score: 509 Queries matched: 11	19
	PREDICTED: similar to hair type II keratin intermediate filament protein [Bos taurus]	
131.	gi|67549 Mass: 23978 Score: 504 Queries matched: 1857	2
	trypsin (EC 3.4.21.4) precursor - bovine	
132.	gi|65165 Mass: 49847 Score: 501 Queries matched: 13	
	alpha-tubulin [Xenopus laevis]	

133.	gi|45382329 Mass: 19518 Score: 501 Queries matched: 29	15
	tumor protein, translationally-controlled 1 [Gallus gallus]	
134.	gi|71896527 Mass: 72006 Score: 500 Queries matched: 26	8
	similar to Double-strand-break repair protein rad21 homolog (hHR21) (Nuclear matrix protein 1) (NXP-1) (SCC1 homolog) [Gallus gallus]	
135.	gi|46048671 Mass: 22072 Score: 498 Queries matched: 14	21
	calcium-binding protein [Gallus gallus]	
136.	gi|57525242 Mass: 16354 Score: 497 Queries matched: 84	
	endothelial differentiation-related factor 1 [Gallus gallus]	
137.	gi|127095 Mass: 35614 Score: 490 Queries matched: 15	
	Myeloid protein 1 precursor (p33)	
138.	gi|5881792 Mass: 41825 Score: 490 Queries matched: 29	1
	actin [Artemia franciscana]	
139.	gi|50755811 Mass: 350958 Score: 488 Queries matched: 18	
	PREDICTED: similar to mitotic checkpoint protein [Gallus gallus]	
140.	gi|65932 Mass: 33061 Score: 486 Queries matched: 11	2
	malate dehydrogenase (EC 1.1.1.37), mitochondrial - pig	
141.	gi|71895711 Mass: 61158 Score: 486 Queries matched: 16	
	similar to MGC68847 protein [Gallus gallus]	
142.	gi|1703106 Mass: 41705 Score: 485 Queries matched: 37	1
	Actin-1	
143.	gi|38453896 Mass: 51422 Score: 485 Queries matched: 197	10
	translation elongation factor 1 alpha [Nematostella vectensis]	
144.	gi|4530086 Mass: 41190 Score: 481 Queries matched: 156	10
	elongation factor 1-alpha [Semibalanus balanoides]Eukaryotic translation elongation factor 1 alpha 1	
145.	gi|63359 Mass: 11703 Score: 478 Queries matched: 13	14
	cytochrome C [Gallus gallus]	
146.	gi|63014 Mass: 15419 Score: 477 Queries matched: 44	
	alpha-A globin [Gallus gallus]	
147.	gi|50750519 Mass: 124958 Score: 474 Queries matched: 21	
	PREDICTED: similar to melanoma differentiation associated protein-5; DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide [Gallus gallus]	
148.	gi|50751020 Mass: 165461 Score: 473 Queries matched: 14	
	PREDICTED: similar to complement component factor h [Gallus gallus]	
149.	gi|9725 Mass: 17837 Score: 470 Queries matched: 35	
	unnamed protein product [Manduca sexta]	
150.	gi|45384078 Mass: 177965 Score: 465 Queries matched: 12	
	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2 [Gallus gallus]	
151.	gi|71895651 Mass: 109408 Score: 461 Queries matched: 14	7
	similar to 116 kDa U5 small nuclear ribonucleoprotein component (U5 snRNP-specific protein, 116 kDa) (U5-116 kDa) [Gallus gallus]	
152.	gi|54258 Mass: 269665 Score: 461 Queries matched: 14	
	talin [Mus musculus]	
153.	gi|23274114 Mass: 37063 Score: 460 Queries matched: 13	
	Hnrpa3 protein [Mus musculus]	
154.	gi|72007072 Mass: 50782 Score: 460 Queries matched: 27	10
	PREDICTED: similar to eukaryotic translation initiation factor 2, subunit 3 gamma, 52kDa [Strongylocentrotus purpuratus]	
155.	gi|53130612 Mass: 76846 Score: 460 Queries matched: 14	
	hypothetical protein [Gallus gallus] Kinesin-like protein KIF2A	
156.	gi|50750722 Mass: 199631 Score: 459 Queries matched: 10	15
	PREDICTED: similar to hypothetical protein FLJ40597 [Gallus gallus] similar to B-aggressive lymphoma 2B	
157.	gi|71895037 Mass: 166171 Score: 458 Queries matched: 14	
	hypothetical protein LOC419128 [Gallus gallus] hypothetical protein LOC419128	
158.	gi|65987 Mass: 35686 Score: 456 Queries matched: 13	2
	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - pig	
159.	gi|1334733 Mass: 48143 Score: 455 Queries matched: 15	12
	unnamed protein product [Gallus gallus] Ig mu chain C region.	
160.	gi|50732081 Mass: 211721 Score: 453 Queries matched: 12	

	PREDICTED: similar to zinc finger protein 27 [Gallus gallus]	
161.	gi|10719701 Mass: 19471 Score: 452 Queries matched: 35	22
	ubiquitin [Scyliorhinus torazame]	
162.	gi|38648948 Mass: 88588 Score: 452 Queries matched: 16	2
	Phosphofructokinase, platelet [Xenopus tropicalis]	
163.	gi|41400354 Mass: 66010 Score: 448 Queries matched: 11	23
	RNase L inhibitor; HP68; RLI [Macaca fascicularis]	
164.	gi|73980410 Mass: 33004 Score: 447 Queries matched: 181	10
	PREDICTED: similar to eukaryotic translation elongation factor 1 alpha 1 isoform 5 [Canis familiaris]	
165.	gi|50755457 Mass: 477902 Score: 446 Queries matched: 24	9
	PREDICTED: similar to transformation/transcription domain-associated protein [Gallus gallus]	
166.	gi|22597134 Mass: 36010 Score: 446 Queries matched: 198	10
	elongation factor-1 alpha [Eubosmina coregoni]Eukaryotic translation elongation factor 1 alpha 1	
167.	gi|62658950 Mass: 475622 Score: 446 Queries matched: 17	
	PREDICTED: similar to KIAA0614 protein [Rattus norvegicus]	
168.	gi|50759301 Mass: 34397 Score: 443 Queries matched: 13	
	PREDICTED: similar to DNA fragmentation factor alpha subunit (DNA fragmentation factor 45 kDa subunit) (DFF-45) (Inhibitor of CAD) (ICAD) [Gallus gallus]	
169.	gi|50762325 Mass: 269190 Score: 441 Queries matched: 14	9
	PREDICTED: similar to transcriptional regulator [Gallus gallus]	
170.	gi|45331057 Mass: 53850 Score: 436 Queries matched: 17	3
	70 kDa heat shock protein [Megachile rotundata]	
171.	gi|55420740 Mass: 45177 Score: 435 Queries matched: 161	10
	elongation factor-1 alpha [Neomaenas monachus]Eukaryotic translation elongation factor 1 alpha 1	
172.	gi|6467317 Mass: 29963 Score: 432 Queries matched: 62	
	elongation factor 1-a [Diphyllobothrium stemmacephalum]	
173.	gi|50752650 Mass: 125697 Score: 429 Queries matched: 17	
	PREDICTED: similar to mSin3A [Gallus gallus]	
174.	gi|45383564 Mass: 38449 Score: 425 Queries matched: 12	
	DNA fragmentation factor, 40kDa, beta polypeptide (caspase-activated DNase) [Gallus gallus]	
175.	gi|50751518 Mass: 22300 Score: 424 Queries matched: 21	
	PREDICTED: similar to peroxiredoxin 1 [Gallus gallus]	
176.	gi|56123383 Mass: 37691 Score: 423 Queries matched: 130	10
	elongation factor-1 alpha [Maculinea arion]Eukaryotic translation elongation factor 1 alpha 1	
177.	gi|20302635 Mass: 45145 Score: 423 Queries matched: 157	10
	elongation factor-1 alpha [Antherina suraka]Eukaryotic translation elongation factor 1 alpha 1	
178.	gi|71895487 Mass: 59151 Score: 421 Queries matched: 10	
	similar to Retinoblastoma binding protein 5 [Gallus gallus]	
179.	gi|61969374 Mass: 71440 Score: 417 Queries matched: 20	3
	heat shock protein 70 [Cryptosporidium andersoni]	
180.	gi|27819634 Mass: 51219 Score: 416 Queries matched: 12	
	RuvB-like 2 [Danio rerio]	
181.	gi|89882 Mass: 53647 Score: 415 Queries matched: 9	19
	keratin, type II, component 7c, cytoskeletal - sheep (tentative sequence)	
182.	gi|211114 Mass: 15378 Score: 412 Queries matched: 39	
	alpha-globin-A	
183.	gi|50728740 Mass: 39784 Score: 409 Queries matched: 29	
	PREDICTED: similar to DEAD box polypeptide 17 isoform 3 [Gallus gallus]	
184.	gi|3970714 Mass: 27240 Score: 409 Queries matched: 14	
	eIF-2gA [Homo sapiens]	
185.	gi|50730570 Mass: 147292 Score: 407 Queries matched: 12	10
	PREDICTED: similar to Eukaryotic translation initiation factor 5B (eIF-5B) (Translation initiation factor IF-2) [Gallus gallus]	
186.	gi|4557469 Mass: 104486 Score: 403 Queries matched: 21	
	adaptor-related protein complex 2, beta 1 subunit isoform b [Homo sapiens]	
187.	gi|1362732 Mass: 37641 Score: 401 Queries matched: 13	
	single stranded D box binding factor 2 - chicken	
188.	gi|57530465 Mass: 59247 Score: 399 Queries matched: 15	2
	tyrosyl-tRNA synthetase [Gallus gallus]	

189.	gi|50728556	Mass: 238084 Score: 395 Queries matched: 11	18
	PREDICTED: similar to Nuclear pore complex protein Nup205 (Nucleoporin Nup205) (205 kDa nucleoporin) [Gallus gallus]		
190.	gi|71896753	Mass: 36962 Score: 392 Queries matched: 11	7
	similar to heterogeneous nuclear ribonucleoprotein A2/B1 isoform 2 [Gallus gallus]		
191.	gi|4103604	Mass: 107046 Score: 392 Queries matched: 9	
	putative mitochondrial outer membrane protein import receptor [Homo sapiens]		
192.	gi|22597128	Mass: 35171 Score: 391 Queries matched: 155	10
	elongation factor-1 alpha [Simocephalus vetulus]Eukaryotic translation elongation factor 1 alpha 1		
193.	gi|5670229	Mass: 30083 Score: 389 Queries matched: 138	10
	elongation factor 1 alpha [Cinara etsuhoe]Eukaryotic translation elongation factor 1 alpha 1		
194.	gi|63517257	Mass: 115553 Score: 387 Queries matched: 20	2
	PREDICTED: ubiquitin specific protease 48 [Mus musculus]		
195.	gi|56118984	Mass: 42619 Score: 384 Queries matched: 8	1
	ARP1 actin-related protein 1 homolog A, centractin alpha [Gallus gallus]		
196.	gi|50759639	Mass: 29596 Score: 384 Queries matched: 14	
	PREDICTED: similar to Dnajc8 protein [Gallus gallus]		
197.	gi|1575536	Mass: 123039 Score: 380 Queries matched: 10	
	regulator of nonsense transcript stability [Homo sapiens]		
198.	gi|47420845	Mass: 52329 Score: 379 Queries matched: 13	
	FUS/TLS [Gallus gallus]		
199.	gi|50753682	Mass: 151122 Score: 373 Queries matched: 15	
	PREDICTED: similar to retinoblastoma-like 2 (p130) [Gallus gallus]		
200.	gi|7657381	Mass: 55146 Score: 366 Queries matched: 13	
	PRP19/PSO4 pre-mRNA processing factor 19 homolog [Homo sapiens]		

Appendix 3

Recipes of SDS-PAGE

The SDS-PAGE gel was made up as the following table in LJMU chromatin research group:

Resolving gel:

Percentage gel	5%	7.5%	10%	12.5%	15%	17.5%	20%
30% Acrylamide	2.5 ml	3.75 ml	5 ml	6.25 ml	7.5 ml	8.75 ml	10 ml
1% Bis-acrylamide	3.9 ml	2.6 ml	1.95 ml	1.56 ml	1.3 ml	1.115 ml	0.975 ml
1.5 M Tris-HCl pH 8.8	3.75 ml	3.75 ml	3.75 ml	3.75 ml	3.75 ml	3.75 ml	3.75 ml
20% SDS	0.15 ml	0.15 ml	0.15 ml	0.15 ml	0.15 ml	0.15 ml	0.15 ml
Water	4.6 ml	4.35 ml	4.05 ml	3.19 ml	2.2 ml	1.135 ml	0.025 ml
10% APS	0.1 ml	0.1 ml	0.1 ml	0.1 ml	0.1 ml	0.1 ml	0.1 ml
TEMED	0.01 ml	0.01 ml	0.01 ml	0.01 ml	0.01 ml	0.01 ml	0.01 ml

Stacking gel:

Percentage gel	5%
30% Acrylamide	2.5 ml
1% Bis-acrylamide	3.9 ml
1.0 M Tris-HCl pH 6.8	3.75 ml
20% SDS	0.15 ml
1% bromophenol blue	0.6 ml
Water	4.0 ml
10% APS	0.1 ml
TEMED	0.02 ml

Running buffer:

14.4 g Glysine, 3.03 g Tris base and 0.5 g SDS, dissolve in 500 ml distilled water.

3X Laemmli sample buffer:

1.875 ml of 1M Tris pH 6.8 (187.5 mM)
0.6 g of SDS
3 ml of glycerol
1.5 ml of 2-mercaptoethanol
0.3 ml of 0.1% bromphenoblue
Make up to 10 ml with distilled water