

PURIFICATION AND INTERACTION STUDIES OF
HISTONE, HMGB, AND PPI PROTEINS, FACILITATED BY
THE DIPOLAR NATURE OF HMGBS

HUGH SMALLMAN

A thesis submitted in partial fulfilment of the requirement of
Liverpool John Moores University for the degree of Master of Philosophy

September 2017

Acknowledgments

First of all, I would like to give all my thanks to my supervisors Dr Chris Wood, Dr Katie Evans and Professor John Baldwin. You have given me all help I needed during the last few years and have shown great patience.

Special thanks to Dr Stan Lambert, who was our laboratory manager for a while. Sadly, he passed away before this thesis was completed. He was always helpful and explained to me how to apply SDS-PAGE.

Many thanks to my other colleagues in the LJMU chromatin research group, especially Dr Qin Qin Zhuang. He and I worked together on the experimental work covered in Chapter 5.

Many thanks to Professor Mark Dickman and his group (Sheffield University) and Dr James Nicholson (Chester University) for providing important data for my MPhil work. Thanks also to Dr Kehinde Ross for help with cell culture, even though it does not figure in this thesis.

Finally I want to give all my thanks to my wife Lian-ju Smallman, who has not seen much of me recently due to my writing this thesis.

Thank you. I wish you all the best of times.

Hugh Smallman
Liverpool, England

The publishers (Elsevier Inc.) have kindly agreed that information from Zhuang Q, Smallman H, Lambert SJ, Sodngam SS, Reynolds CD, Evans K, Dickman MJ, Baldwin JP, Wood CM (2014) "Cofractionation of HMGB proteins with histone dimers" *Anal Biochem.* **447**: 98-106 can be included in the following thesis figures and tables: Figure 5.3(b), 5.4, 5.5, 5.6; Table 5.4,5.5, 5.7, 5.9. The paper can be obtained via Science Direct at <https://www.sciencedirect.com/>

Abstract

High mobility group box (HMGB) proteins are the most abundant non-histone proteins in the nuclei of eukaryotic cells and are highly conserved in animals. They bind dynamically to chromosomal DNA, interact with other proteins including transcription regulators, and have key roles in gene transcription and DNA repair. These fundamental activities require HMGB proteins to interact with the histone proteins found in chromatin, but the precise mechanisms remain unclear - the work here provides steps towards their clarification.

Acid extraction of histones from chicken erythrocytes was explored, as this provides high yields without the use of ultracentrifugation. Histone protein markers (MW range 11.4 – 22.5kDa) suitable for use with SDS-PAGE were prepared using sulphuric acid extraction. For future research, almost pure histone octamers were also prepared, using phosphoric acid extraction combined with potassium chloride to ensure the octamers remained intact and in a nuclear-like environment.

The remaining work was based on proteins extracted by mild methods from chicken erythrocytes to retain post-translational modifications. A nuclear protein set was isolated containing almost entirely histones and HMGB proteins, and was subject to cation-exchange chromatography, with phosphate buffers chosen to partially simulate nuclear conditions. HMGB molecules have a C-terminal acidic tail which in free solution is folded back onto the remaining basic part of the protein. Analysis of the chromatogram peaks suggested the HMGB proteins had unfolded into a dipolar configuration, with their basic parts binding to the cation-exchange column and their acidic tails binding to the histones. Gel filtration chromatography applied to fractions eluted from the cation-exchange column suggested the presence of one or more unidentified complexes.

Native HMGB proteins were isolated using a novel method based on their dipolar nature. At high concentration, HMGB1 proteins were used to pull out a potential HMGB1/FKBP3 complex from a pool of nuclear proteins. Technology to significantly enhance the concentration of HMGB1 in chicken plasma was also developed, as the basis for prompt, low cost measurement of HMGB1 in biofluids (this has medical applications, such as in cancer diagnosis and prognosis). Further exploiting the dipolar nature of HMGB proteins, a method was developed for isolating the peptidyl-prolyl isomerases FKBP3 and Cyp B.

Abbreviations

ACF	ATP-utilising chromatin assembly and remodeling factor
ATP	Adenosine Triphosphate
APS	Ammonium Persulphate
AU	Acid Urea
bp	base pair
cePNE	chick erythrocyte Protein Nuclear Extract
DEAE	Diethylaminoethyl
dH ₂ O	de-ionised water
DNA	Deoxyribonucleic Acid
e	The charge on a proton
EDC	1-ethyl-3[3-dimethylaminopropyl]carbodiimide
EDTA	Ethylenediaminetetraacetic Acid
EM	Electron Microscopy
FLIP	Fluorescence Loss In Photobleaching
FPLC	Fast Protein Liquid Chromatography
HDAT	Histone acetyl transferase
HCl	Hydrochloric Acid
HDAC	Histone deacetylase
HMG	High Mobility Group (protein)
HMGB	High Mobility Group Box (protein)
HPLC	High performance liquid chromatography
IEX	Ion-exchange chromatography
IM	Intestinal metaplasia
ISG	International Scientific Group
K ₂ HPO ₄	Dibasic potassium phosphate
KCl	Potassium Chloride
kDa	kilo Daltons
KH ₂ PO ₄	Monobasic potassium phosphate
LJMU	Liverpool John Moores University
mAU	milli Absorbance Unit

MgCl ₂	Magnesium chloride
mgf	mascot generic files
MS	Mass spectroscopy
MW	Molecular Weight
MWCO	Molecular Weight Cut Off
NaCl	Sodium chloride
NCP	Nucleosome core particle
PBE	Polybuffer exchanger
pI	Isoelectric Point
PPI	Peptidyl-Prolyl Isomerase
PTM	Post-Translational Modification
SDS-PAGE	Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SUMO	Small ubiquitin-like modifier
TAU	Triton Acid Urea
TCA	Trichloroethane
TLR	Toll-like receptor
TPS	Thiopropyl-Sepharose-6B
TEMED	N,N,N',N'-Tetramethylethylenediamine
UV	Ultraviolet

Contents list

	Page	
Title page	1	
Acknowledgements	2	
Abstract	3	
Abbreviations	4	
Contents list	6	
1. Introduction		7
2. Background on cell type selection, and on histone, HMGB, and FKBP3 proteins		11
3. Methods used in our work, including a review of similar methods		58
4. Development of acid extraction methods to prepare histone markers and histone octamers		97
5. Investigation of potential HMGB/histone protein complexes		115
6. A method for the efficient separation of native HMGB proteins from other nuclear proteins in chicken erythrocytes		162
7. Research applications for purified native HMGB1		184
8. A method for the efficient separation of native FKBP3 and Cyp B from other nuclear proteins		222
9. Conclusions and recommendations		242
10. References		244

Chapter 1

Introduction

<u>Contents</u>	Page
1.1 Background	8
1.2 Scope of the thesis	8
1.3 The dynamic nucleosome	10

1.1 Background

Our research group at Liverpool John Moores University (LJMU) has been investigating nuclear proteins interactions for more than a decade, based on proteins from chicken erythrocytes. The group is seeking to clarify the workings of the dynamic nucleosome which lies at the heart of all critical cell nuclear functions. Interactions between the nucleosome and other proteins are transitory, and thus difficult to investigate. The group is working towards establishing conditions which place nucleosome related complexes into metastable intermediary states suitable for thorough investigation. Using native histones the group successfully crystallized a histone octamer structure at 1.9 Angstrom (Wood *et al.*, 2005) which closely matches the octamer structure found in nucleosome core particles assembled using histones produced by recombinant DNA technology. The group has also developed a range of methods for efficiently isolating nuclear proteins, such as described by Foulger *et al.* (2012).

The HMGB protein family are ubiquitous in animals, well conserved, abundant in the nucleus, and involved in many processes related to chromatin, yet their precise mechanisms of action are not known. A major objective of the group is to improve our understanding of the interaction between HMGB proteins and chromatin; for example, our future work may include the crystallisation of an HMGB/octamer structure, if feasible.

The author joined the above research group as a part time student several years ago. This thesis provides a snapshot of some of the work he has carried out in conjunction with others in the group, with the above objective in mind. In particular, the transient dipolar nature of HMGB proteins has been explored, such as its relevance in potential HMGB/histone complexes. This dipole nature has also been used as the basis for devising efficient methods for isolating native nuclear proteins, in particular the HMGB proteins themselves, and peptidyl prolyl isomerase (PPI) proteins. Apart from some limited work based on the acid extraction of histones, the nuclear proteins used in these studies have been obtained from chicken erythrocytes using mild extraction methods to produce "native" proteins which retain their post translational modifications (PTMs).

1.2 Scope of the thesis

Chapter 2 briefly examines the extent to which a study based on chicken cells is relevant to the human situation, and provides background information about the molecules in this

thesis. Chapter 3 outlines the experimental methods employed, complementing the additional experimental details found in later chapters. Chapters 4 – 8 describe experimental work. Conclusions and recommendations are provided in Chapter 9, and a master list of references is provided in Chapter 10. The most used protocols for extracting nuclear proteins from chicken erythrocytes are provided in Appendices A, B, C, with additional protocols in some of the experimental sections. Appendices D and E provide protocols for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and for preparing SDS-PAGE samples for mass spectroscopy (MS), respectively.

Chapter 4 describes two approaches to acid extraction of histones. Firstly, a histone marker set, molecular weight (MW) range 11.4 - 22.5kDa and suitable for SDS-PAGE, was efficiently prepared from chicken erythrocytes using a simple sulphuric acid extraction method. Secondly, a high concentration of nearly pure histone octamers was prepared using phosphoric acid extraction combined with potassium chloride; this was designed to retain the octamers in an intact state throughout the extraction process.

Chapter 5 investigates a set of potential HMGB/histone complexes. A model is proposed for these interactions based on the HMGB proteins acting as dipoles. HMGB proteins have strongly acidic, unstructured C-terminal tails which fold back onto the remaining basic part of the protein which includes two DNA binding "HMG boxes". The HMGB proteins become strong dipoles when conditions are such that the acidic tail flicks away from the rest of the molecule. In this circumstance, the acidic tail and the remaining basic part of the protein both become available for ionic interaction with other proteins, and also the HMG boxes become available to bind to DNA.

Chapter 6 provides a method for the large scale, efficient separation of native HMGB proteins based on their dipolar nature.

Chapter 7 demonstrates how an excess of purified HMGB1 can be used to pull potential HMGB/protein complexes from a pool of proteins. In addition, purified HMGB1 was employed in the preliminary development of an efficient method for measuring HMGB1 in blood plasma; the value of such measurements in cancer diagnosis and prognosis is explained. The chapter also suggests using HMGB1 in complex with an octamer, for X-ray crystallography or electron microscopy.

Chapter 8 provides a method for the separation of native FKBP3 and Cyp B from the other nuclear proteins, again exploiting the dipolar nature of the HMGB proteins.

1.3 The dynamic nucleosome

Cells operate under instructions stored as permanent, coded four-letter information contained in deoxyribonucleic acid (DNA). DNA is located in the cell nucleus as a limited number of very long molecules, each comprising $\sim 10^8$ pieces of code in continuous string. Short sections of the DNA are wound around cylindrical structures formed from 8 core histone proteins (histone octamers), this unitary arrangement protects the DNA and forms part of a system controlling access to the code. The units are referred to as nucleosome core particles, and since the DNA is continuous, each particle is linked to the next unit by a short length of DNA. A linker histone protein also bridges the gap between particles. A particle plus associated linker DNA and linker histone is called a nucleosome, and the full string of nucleosome associated with one DNA molecule is called a chromosome, and is referred to as "beads on a string". Further coded instructions, with a varying degree of permanence, are provided by small chemical additions to the DNA and histones, called Post Translational Modification (PTMs).

A chromosome is a highly dynamic molecule - the nucleosome structure determined by X-ray crystallography may be regarded as a ground state, from which transitional states arise. Li *et al.* (2006) have shown that in free solution DNA remains fully wrapped on a nucleosome for only ~ 250 ms before spontaneously unwrapping off the histone surfaces, and spontaneously rewrapping within 10 - 50ms. Histones come and go from the nucleosome - for an example, using fluorescent tags, Misteli *et al.* (2000) have shown the residence time of individual linker histone H1 molecules in chromatin is only minutes. Nucleosomes are actively moved along the DNA strand by ATP-dependent remodelling complexes (Turner, 2001; Teif and Rippe, 2009). Proteins transiently assemble into complex molecular structures on nucleosomes, to promote essential nuclear functions related to DNA, such as transcription, repair, and replication. These interactions between nucleosomes and other proteins are also short lived. For example, Scaffidi *et al.* (2002) have shown the residence time of HMGB1 molecules on chromatin is shorter than the residence time of H1. Moreover, although the existence of nucleosomes has been known for decades, the relevance of chromatin structure is still being determined. Only recently has it been shown (Maeshima *et al.*, 2014; Ou *et al.*, 2017) that some of the controls on gene accessibility more commonly depend on the density of disordered "beads on a string", rather than on access being restricted by a regular higher order chromatin structures, as previously thought. Chapter 2 further discusses DNA and histones (and their PTMs), nucleosomes, chromatin, and other molecules relevant to this study.

Chapter 2

Background on cell type selection, and on histone, HMGB, and FKBP3 proteins.

<u>Contents</u>	Page
2.1 Introduction	13
2.2 Chicken and human similarities	15
2.3 Chicken erythrocytes	15
2.4 DNA, proteins, chromatin, and the dynamic nucleosome	17
2.5 Histones	23
2.6 HMGB proteins	36
2.7 FKBP3	49
<u>Tables</u>	
Table 2.1 Some databases providing protein information	14
Table 2.2 Histone molecular weights and charges	23
Table 2.3 Histone structures	24
Table 2.4 Histone stoichiometry in chicken erythrocyte nuclei	25
Table 2.5 Histone PTMs	34
Table 2.6 Some HMG family members (old and new names)	37
Table 2.7 Examples of mammalian HMG-box proteins	40
Table 2.8 HMGB molecular weights, charges, and acidic tail lengths	42
Table 2.9 Molecular functions determined for human HMGB1	44
Table 2.10 Biological processes determined for human HMGB1	45
Table 2.11 HMGB PTMs	48
Table 2.12 FKBP3 molecular weights and charges	50
Table 2.13 FKBP3 PTMs	50
Table 2.14 Molecular functions determined for human FKBP3	55
Table 2.15 Biological processes determined for human FKBP3	56
Table 2.16 Molecular functions and biological processes for chicken FKBP3	56
Table 2.17 Predicted functional partners for human and chicken FKBP3	57

Figures

	Page
Figure 2.1 Simplified production route for erythrocytes in adult chicken	16
Figure 2.2 Amino acid condensation	18
Figure 2.3 Structure of histone H2A	26
Figure 2.4 Structure of histone dimer H2A-H2B	27
Figure 2.5 Structure of histone octamer	28
Figure 2.6 Structure of nucleosome core particle	29
Figure 2.7 Structure of the globular region of histone H5	30
Figure 2.8 Examples of histone PTMs	35
Figure 2.9 HMGB1 human/chicken pairwise sequence alignment	38
Figure 2.10 Structure of HMGB1	41
Figure 2.11 Structure of FKBP12	51
Figure 2.12 Location of FKBP3 in human cell line U-251 MG	52
Figure 2.13 Comparison of sequences between human and chicken FKBP3	53

2.1 Introduction

This chapter provides background to the purification and interaction studies of the histone, HMGB, and peptidyl-prolyl isomerase (PPI) eukaryotic nuclear proteins covered by this thesis. The aim is to understand further the mechanisms in the cell nuclei of vertebrate animals, and ultimately this work contributes to an understanding of the human organism, with potential benefits in medicine.

For this study, chicken erythrocytes were selected as the source of nuclear proteins since the chicken and human genome and proteins have similarities, as discussed in Sections 2.2 and 2.3. Chicken erythrocytes are easily extracted from fresh blood, which can be obtained in large quantities from a chicken farm.

Nuclear activities are dominated by the involvement of chromatin which incorporates DNA and histone proteins, discussed in Sections 2.4 and 2.5. The other main proteins in this study (HMGB and FKBP3 proteins) are discussed in Sections 2.6 and 2.7. Numerous freely available databases (Table 2.1) provide further extensive data on these proteins.

Table 2.1 Some databases providing protein information.

Parent organisation and webpage	Scope of databases
NCBI www.ncbi.nlm.nih.gov NCBI Resource Coordinators (2017)	Genome and gene expression. Protein sequence, structure, function, and roles in disease. Includes RefSeq.
EMBL-EBI www.ebi.ac.uk Kanz <i>et al.</i> (2005)	Genome, RNA and gene expression. Macromolecular complexes. Protein classification and structure (ePDB). Includes the ENSEMBLE genome browser, API, and database, providing access to genome annotation for animals, plants, bacteria, etc. Part of the STRING consortium database of known and predicted protein interactions.
UniprotKB (Uniprot) www.uniprot.org Chen <i>et al.</i> , 2017	Protein sequence and function: <ul style="list-style-type: none"> - manually annotated (Swiss-Prot) - automatically annotated (TrEMBL).
RCSB PDB www.rcsb.org Berman <i>et al.</i> (2000)	3D shapes of proteins, nucleic acids, and complex assemblies.
GeneCards www.genecards.org Rebhan <i>et al.</i> (1998)	Integrated gene-centric human data from ~125 web sources, including genomic, transcriptomic, proteomic, genetic, clinical, and functional information. Maintained by Weizmann Institute of Science.
Human Protein Atlas www.proteinatlas.org Uhlen <i>et al.</i> (2015)	Prevalence of proteins: <ul style="list-style-type: none"> - within cell location (Cell Atlas). - within tissue (Tissue Atlas) - within cancer cells (Cancer Atlas)
PhosphoSite Plus www.phosphosite.org Hornbeck <i>et al.</i> (2012)	Curated post-translational modifications on proteins from human, mouse, and related species.
Chemspider www.chemspider.com Pence and Williams (2010)	Chemical structure database.
ExPASy https://www.expasy.org/ Gasteiger E <i>et al.</i> (2005)	SIB Bioinformatics Resource Portal providing access to scientific databases and software tools in different areas of life sciences including proteomics, genomics, phylogeny, systems biology, etc. Includes the "Lalign" sequence alignment software.

Notes

Multiple databases are provided by the above parent organisations, and each usually provides links to further databases.

2.2 Chicken and human similarities

Chickens and humans are vertebrate animals, diverging from a common ancestor about 310 million years ago. Chicken was the first livestock species to be sequenced. DNA was taken from a female red jungle fowl (*Gallus gallus*) (Hillier *et al.* 2004) (Wallis JW *et al.* 2004). Burt (2005) has reviewed differences between the chicken and human genomes. The distribution of DNA amongst chromosomes differs between the two species (79 chicken chromosomes, 46 human), and the chicken genome is only about one third the size of the human genome. 20 – 25% of size difference is due to low frequency of repeats, pseudogenes, and duplications, but otherwise the reason is not clear. Despite the size difference, there are similarities between the two genomes. The chicken genome is estimated to contain 20,000 – 23,000 genes, similar to the human genome. Examination of known functional sequences indicates that perhaps 75% of protein-coding regions are conserved across the two species. When the chicken genome was used to identify novel genes missed from the human gene set, only 37 potentially new human genes were predicted (Eyras *et al.* 2005).

Like human blood, chicken blood contains erythrocytes (red blood cells), leukocytes (white blood cells), thrombocytes (called platelets in human blood), and plasma (containing proteins, lipids, dissolved nutrients, and waste products). Human blood comprises mainly plasma (~54% by volume) and erythrocytes; chicken blood is similar. However, chicken erythrocytes contain a nucleus whereas human erythrocytes do not.

Section 2.5 compares some characteristics of human and chicken histones. Sections 2.6 and 2.7 include sequence diagrams which show the extent of correspondence between human and chicken HMGB1 residue sequences, and between human and chicken FKBP3 sequences.

2.3 Chicken erythrocytes

The chicken erythrocyte is ovoid with longest axis of about 10 microns, and in common with other non-mammalian vertebrates contains a centrally located ovoid nucleus. The nucleus contains DNA, RNA, and many hundreds of different proteins (Zhuang, 2011). The function of both chicken and human erythrocytes is to transport oxygen, held in combination with haemoglobin in the cell, around the organism. Erythrocytes are red in colour due to the haemoglobin.

Blood cells derive from self-renewing multi-potent hematopoietic stem cells (HSCs) in a complex process known as haematopoiesis, in which each blood cell type arises from the replication and differentiation of a series of intermediate progenitor cell types. Hematopoietic progenitor cell locations in an organism vary considerably during development. However, in both adult human and adult chicken, HSCs are located in bone marrow which provides a factory for the manufacture of most blood cells (Figure 2.1). However, unlike human erythrocytes, for unknown reasons, chicken erythrocytes have a nucleus. Most chicken erythrocytes are in the same, relatively inactive, state, leading to low levels of protein and RNA synthesis (Cameron and Prescott, 1963; Kabat and Attardi, 1966). This has the experimental advantage that the cellular functions which are determined can be reliably assigned to a single cell type and cell cycle point, rather than being a composite behaviour which does not apply to any specific cell or cell status.

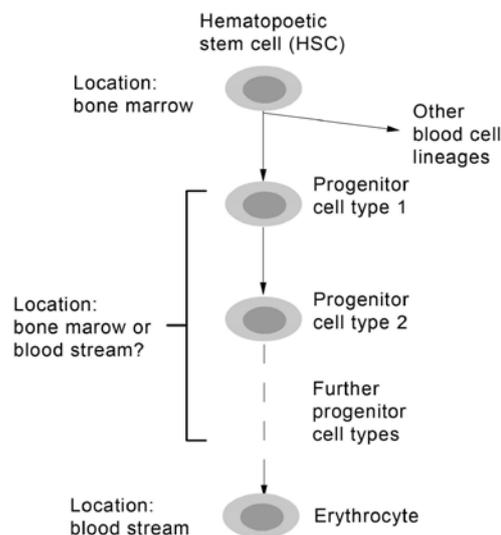


Figure 2.2 Production route for erythrocytes in adult chicken

Figure 2.1 Simplified production route for erythrocytes in adult chicken

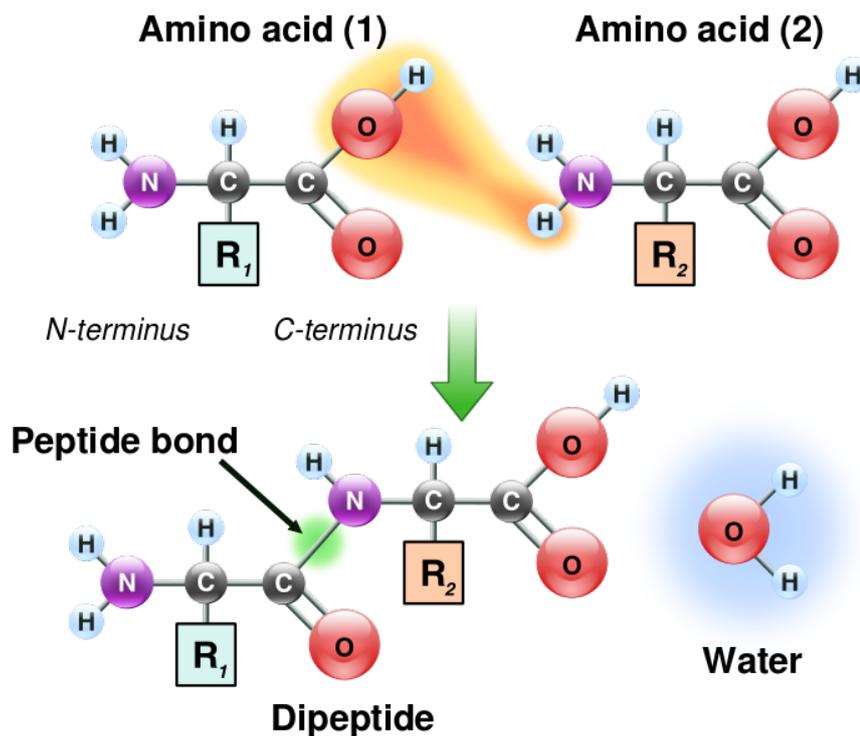
2.4 DNA, proteins, chromatin, and the dynamic nucleosome

2.4.1 DNA

Cells operate under instructions stored as permanent, coded four-letter information contained in deoxyribonucleic acid (DNA). Each letter, A, C, G, and T, corresponds to a cyclic organic base adenine (A), cytosine (C), guanine (G), and thymine (T). DNA is a long un-branched polymer in which these bases are each linked to a five-carbon sugar (deoxyribose), and the sugars form long chains with alternating phosphate groups. DNA can be either single stranded, or double stranded, in which the bases link together in specific Watson-Crick pairs AT or CG (although occasionally there may be non-standard pairings – Leontis *et al.*, 2002). Double stranded DNA forms a three different helical structures (A, B and Z), the most common of which in cells is B-DNA (Wing *et al.*, 1980). DNA is used as a pattern for making temporary intermediate instructions in the form of further coded four-letter information contained in messenger ribonucleic acid (mRNA), in a process called transcription. mRNA is also a long, un-branched polymer, either single stranded or double stranded. However, RNA differs from DNA in that its sugar (ribose) has a hydroxyl group at the 2' carbon position in place of a hydrogen atom found in the DNA sugar, and thymine (T) is replaced by uracil (U) (Gros *et al.*, 1961). RNA in turn is used as a pattern by the ribosomes for the cellular manufacture of proteins based on those parts of the chromatin DNA which are being transcribed. RNA is used in several other roles, including incorporation in ribosomes, transfer of amino acids to the ribosomes for incorporation into a protein peptide, and post-transcriptional regulation.

2.4.2 Proteins

A protein comprises a string of amino acid residues, formed when amino acids (of which there are 21 common types) condense together losing water molecules as they do so (Figure 2.2). The selection and order of residues, together with post-translational modifications (PTMs), determine the protein properties, including how it folds into a three dimensional structure. Proteins have essential roles in cell structure and many cellular processes (Whitford, 2005).



Wikipedia image in public domain: image is own work by Yassine Mrabet.

Figure 2.2 Amino acid condensation

R represents the amino acid side chain.

Repeated condensation, leading to exclusion of a water molecule between each amino acid generates an amino acid residue string called a peptide. The string will typically fold into a three-dimensional structure called a protein (Whitford, 2005).

2.4.3 Chromatin

Double stranded DNA forms helical structures. In an animal cell, the structures contain ~100 million DNA bases and would extend ~2m if stretched out (Maeshima *et al.*, 2014). However, most of the DNA is wound around protein core structures which occupy little space (Luger *et al.*, 1997; Turner, 2001). This protects the DNA, and helps control access to individual stretches of DNA. Each core structure is assembled from 8 histone proteins, hence the name "histone octamer". Typically, 147 DNA base pairs are wrapped 1.65 times around each octamer, this assembly is called a nucleosome core particle. The continuous string of DNA base pairs links the nucleosome core particles together, with a linker histone and linker DNA (20 – 80 DNA base pairs, depending on species) between each core particle forming a "beads on string" structure, also called the 10nm Fibre (Maeshima *et al.*, 2014). The combination of a nucleosome core particle, linker histone and linker DNA is called a nucleosome, and a collection of linked nucleosomes is called chromatin.

When cells replicate, new nucleosomes have to be assembled. Firstly, histones H3 and H4 are bound to the DNA, followed by incorporation of two histone H2A-H2B dimers to complete the nucleosome, this involves histone chaperones as discussed by Das *et al.* (2010). Chaperones are further discussed in Section 2.4.5.

Chromatin is involved in regulating the major nuclear processes, including transcription, replication, DNA repair and recombination (Bonaldi *et al.*, 2002). Access to the DNA for these purposes is influenced to some extent by the degree of nucleosome packing, which varies depending on the cellular activity. For example, Widom and Klug (1985) found that *in vitro* the beads on a string pack into a regular structure called the 30nm Fibre. However, this type of packing was found in chromatin artificially formed *in vitro* by reconstituting purified DNA and histones, or in permeabilized cells from which other components had been extracted (Ou *et al.*, 2017). Recent studies *in vivo* have shown that in many cells the additional packing comprises an irregularly folded 10nm Fibre (Maeshima *et al.*, 2014). Exceptionally, chicken erythrocytes appear to contain some chromatin in an ordered 30nm Fibre, possibly associated with gene silencing (Scheffer *et al.*, 2011; Maeshima *et al.*, 2014).

Chromatin has been classified into two main types – euchromatin which is less compact and from which genes are more likely to be transcribed, and the more compact heterochromatin where transcription is repressed. Filion *et al.* (2015) suggest a finer classification since they have found five types of chromatin (which they name using colours) in fruit flies (*Drosophila*), based primarily on the proteins found bound to the chromatin DNA. Some of their findings are summarised below.

Blue is one of two types of heterochromatin, both of which can cover long stretches of DNA. It is marked with PcG attached proteins and the methylated histone H3K27me2 (see Section 2.5.5 for more details of histones modifications). Green is a second type of heterochromatin, and is marked with HP1 attached protein and methylated histone H3K9me2. Black is the most highly repressive and abundant type of chromatin. Domains tend to be longer than for the other types. It includes attached proteins H1, D1, IAL, and SUUR. Red is one of two types of euchromatin, the other is Yellow; both have hallmarks of transcriptionally active euchromatin. Both are marked by HDACs, RPD3, SIR2, and SIN3A attached proteins, and high levels of H3K4me2 and H3K79me3. However, Red is also marked by the nucleosome-remodeling ATPase Brahma (BRM), regulator of chromosome structure SU(VAR)2-10, the Mediator subunit MED31, and the 55 kDa subunit of CAF1 – these marks are almost entirely absent from the other chromosomal categories. Red and Yellow are associated with different types of genes which may be regulated by different pathways. Genes more broadly expressed throughout embryonic stages and across tissues are more highly enriched in Yellow chromatin, and so are genes for universal cellular functions, such as DNA repair, and nucleic acid metabolic processes.

Ou *et al.* (2017) have developed a system which combines multi-tilt electron microscopy tomography with mild OsO₄ labelling of DNA. This has allowed the 3D organisation of individual chromatin chains, in chicken erythrocytes and in human interphase and mitotic cells, to be directly visualized. They found ordered 30nm chromatin fibers were present in nuclei purified from hypotonically lysed chicken erythrocytes treated with MgCl₂, but in fixed human cells (osteosarcoma U2OS cells and small airway epithelial cells, SAECS) they found disordered 5 - 24nm diameter granular chains - these were packed together at different concentration densities in interphase and mitotic cells. Their data suggests that more usually the assembly of domains in the nucleus with different chromatin concentrations determines the accessibility and activity of DNA, rather than the presence of higher-order regular chromatin arrays.

Gene transcription requires the involvement of many proteins, for example, transcription factors (proteins which attach to short sections of DNA; Turner, 2001), and seems to take place in a limited number of "transcription factories" (Chambeyron and Bickmore, 2004). Transcription factories ensure local concentrations (co-localisation) of related molecules; these can subsequently be processed together, without having to provide the necessary molecular concentrations across the whole nuclear region (Wood, 2010). For example, during interphase it has been found that chromosomes are partitioned into very large groups called topologically associated domains (TADs) (Dixon *et al.*, 2015).

2.4.4 *The dynamic nucleosome*

Transient assemblies of specialised molecules (such as proteins, RNA, and nucleotides), operating in multi-stage processes, provide the machinery to convert the DNA information in chromatin into specific products required for essential cellular processes. There are mechanisms which access, select and transcribe, recombine, or repair, at correct times, the appropriate part of the DNA necessary for the specific cell type. These mechanisms include concentration/dispersal of the nucleosomes (Ou *et al.*, 2017), repeated remodelling of the nucleosome by dynamic transitory processes (see Section 2.4.5, also Turner, 2001; Li *et al.*, 2005; Teif and Rippe, 2009), and localised concentrations of interacting molecules within the nucleus (Dixon *et al.*, 2015). Further processes modify the output from transcription, en route to providing the final molecules required by the cell. Thus, although X-ray crystallography shows an apparently stable nucleosome core particle, it is only a snapshot of a stable intermediary in an essentially dynamic situation (Kowalski and Palyga, 2011).

2.4.5 *Chromatin remodelling*

Chromatin remodelling is a dynamic process which provides access to sections of the DNA string which may otherwise be buried within the nucleosome core particle or be inaccessible due to the presence of linker histones. As described in Chapter 1, DNA spontaneously and transiently unwraps from and rewraps onto histone octamer structure, thus exposing sections of DNA (Li *et al.*, 2006).

Segal *et al.* (2006) have claimed there is a genomic code for nucleosome position within the DNA sequence, which if true suggests that the same sections of DNA may repeatedly be buried within the nucleosome core particles. This claim is disputed (Zhang *et al.*, 2010). In any event, remodelling complexes powered by the energy released from the breakdown of ATP can actively move nucleosomes along the DNA strand. Turner (2001) lists the following remodellers - *S. cerevisiae* SW1/SNF family, *X. laevis* Mi-2, *H. sapiens* NuRD, and the *D. melanogaster* ISWI family. More remodellers have since been found, and in more organisms; these include SWR1 and INO80, both of which have human versions (Conaway and Conaway, 2008). Remodelling complexes have different roles, for example, INO80 and SWI/SNF-family remodellers have a role in DNA repair. Post translation modifications (PTMs) of the histones and related proteins (discussed in Section 2.5.5), the replacement of canonical histones by their variants, and histone chaperones all affect nucleosome remodelling (Petesch and Lis, 2012), (Horn *et al.* 2002).

Histone chaperones assist chromatin assembly and disassembly, and help replace histones with new ones or variants as necessary (Das *et al.*, 2010). Chaperones work in concert with remodellers. An example is the NAP-1 protein which together with the ACF remodeller (ISWI family) assembles chromatin into periodic arrays (Turner, 2001). HMGB1 has been described as a DNA chaperone by Bonaldi *et al.* (2002), who propose that the DNA bending ability of HMGB1 assists the linker DNA to loop out, and in doing so assists ACF/CHRAC-dependent nucleosome sliding. Ju *et al.* (2006) provide some support for this model, showing that HMGB1 replaces linker histone H1 during a transcription event.

Residence times of the molecules within a remodelling complex are a key parameter. Where affinities for the complex are low, residence times of individual molecules will be short so that only the correct complexes will exist for long enough to carry out significant functions (Wood, 2010).

2.5 Histones

2.5.1 Histone families

There are five main families of histone proteins found in nuclei: H1/H5, H2A, H2B, H3, and H4. Histones are well conserved across most eukaryotes (Zhuang *et al.*, 2014). Table 2.2 compares representative human and chicken histone proteins. Variants arise in specific circumstances, such as H2AX which is associated with DNA repair (Millar, 2013). Takami *et al.* (1996) identify chicken histone variants.

Table 2.2 Human and chicken histone molecular weights and charges

Histone version	Identifier UniprotKB	N	MW* kDa	pI	Net charge at pH 7.1**
H1.0 human	P07305	194	20.86	10.84	+53.2
H1 chicken	P09987	218	21.88	11.12	+60.9
H5 chicken	P02259	190	20.73	12.19	+61.4
H2A_IV chicken	P02263	129	13.94	10.90	+17.5
H2A 1B/E human	P04908	130	14.13	11.05	+17.7
H2B 7 chicken	P0C1H5	126	13.96	10.32	+18.5
H2B 1-J human	P06899	126	13.90	10.32	+18.5
H3.2 chicken	P84229	136	15.39	11.27	+20.3
H3.2 human	Q71DI3	136	15.39	11.27	+20.3
H4_VI chicken	P62801	103	11.37	11.36	+18.3
H4 human	P62805	103	11.37	11.36	+18.3

Notes

N number of amino acid residues in protein. *Isotopically averaged molecular weight.

**Assumed pH in nucleus of inactive chicken erythrocyte.

Data calculated using Scripps Protein Calculator (2016): pI (Isoelectric Point) estimate assumes all residues have pKa values equivalent to the isolated residues. For a folded protein this is not valid, but the approximated value can be useful for planning protein purifications. Similar considerations apply to the calculated net charge.

2.5.2 Core histones

H2A, H2B, H3, and H4 each include three similar α -helices linked by two loops/bends (with some beta strand character), together with unstructured tails, plus one or more additional helices (Table 2.3). The basic structure is called a histone fold – see Figures 2.3 to 2.5 for more details. The histone octamer comprises two dimers (H2A-H2B), and a single tetramer (H3-H4)₂ - prevalence per nucleosome is consistent with two molecules each of H2A, H2B, H3, and H4 (Table 2.4). Specific docking sequences form strong links between dimer and tetramer, stabilising the octamer structure (Chantalat *et al.*, 2003). The histone tails are mainly located outside the DNA/core octamer. Figures 2.3 to 2.6 show the following structures: H2A histone; H2A histone in a dimer with H2B; histone octamer; and a complete nucleosome core particle. Various databanks provide histone structures, peptide sequences, annotation, and other data (Table 2.1).

Table 2.3 Histone structures

Histone	Sequence identifier (UniprotKB)	N	Example of RCSC PDB source structure	Helices in source structure
H1.0 human	P07305	194	-	-
H5 chicken	P02259	190	1HST	3
H1 yeast	P53551	258	1UST	3
H2A_IV chicken	P02263	129	1TZY	6
H2A 1B/E human	P04908	130	5B0Z	6
H2B 7 chicken	P0C1H5	126	1TZY	4
H2B 1-J human	P06899	126	5B0Z	4
H3.2 chicken	P84229	136	1TZY	4
H3.2 human	Q71DI3	136	5B0Z	4
H4_VI chicken	P62801	103	1TZY	4
H4 human	P62805	103	5B0Z	4

Notes

N Number of amino acid residues in the protein.

No PDB structure was found for H1 human, and only a partial structure (globular region) was found for H5 chicken.

Table 2.4 Histone stoichiometry in chicken erythrocyte nuclei

Histone	Molar ratio	Stoichiometry
H1A	0.08 (0.00)	1.28 (0.02)
H1B	0.12 (0.00)	
H5	0.44 (0.01)	
H2A	1.04 (0.01)	8.00 (0.08)
H2B	1.03 (0.02)	
H3	0.93 (0.03)	
H4	1.00 (0.02)	

Notes

Results selected from Table 2 of Bates and Thomas (1981).

Standard error is shown in brackets.

See also Kowalski and Palyga (2011) for discussion of H1/H5 ratio.

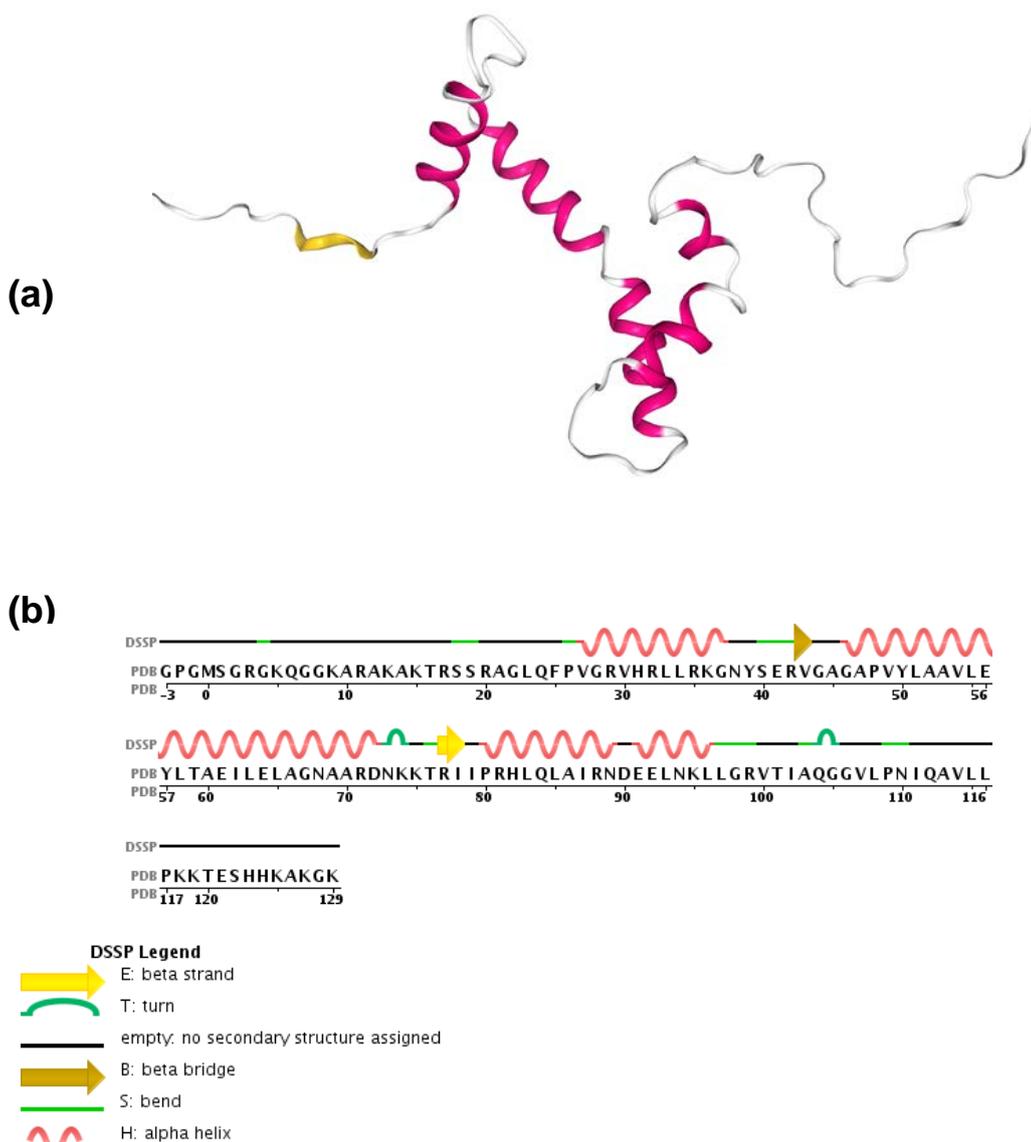


Figure 2.3 Structure of histone H2A

(a) Structure of a human histone H2A, selected from RCSB PDB Entry 2RVQ, using Deepview (SwissProt – Guex and Peitsch, 1997) to select just the atoms in the H2A chain, followed by NGL rendering (Rose and Hildebrand, 2015). Image shows each alpha helix in red. PDB Entry 2RVQ was derived by Moriwaki *et al.* (2016). The structure contains a histone fold, with four α -helices and two loops (with β -strand character), together with long disordered N- and C-terminal tails. See Figure 2.4 for more details.

(b) Sequence chain view of H2A copied directly from the PDB sequence display for 2RVQ.

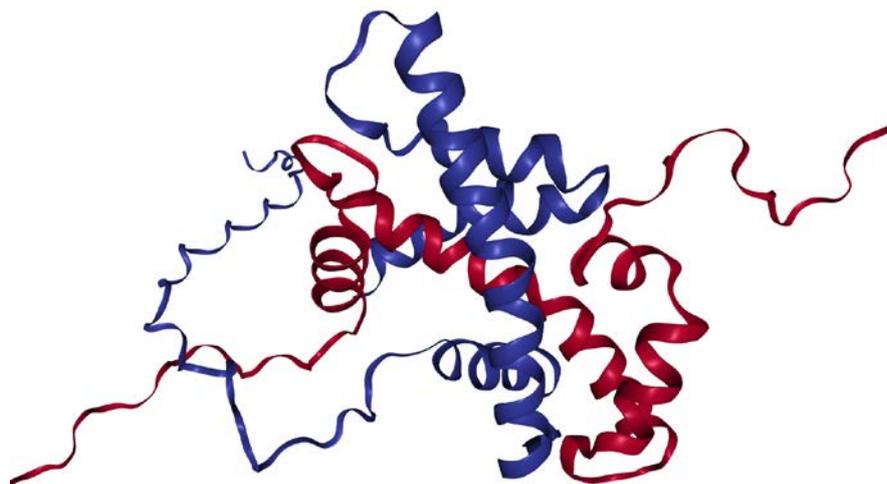


Figure 2.4 Structure of histone dimer H2A-H2B

Structure of human histone dimer H2A-H2B from RCSB PDB Entry 2RVQ, showing H2A and H2B in separate colours. The image was obtained using NGL rendering (Rose and Hildebrand, 2015).

Moriwaki *et al.* (2016) derived 2RVQ using Solution NMR applied to the histone dimer H2A-H2B, isolated from human cells. The above structure is just one of several possible structures which fit the NMR parameters. Moriwaki states "H2A and H2B each contain a histone fold, comprising four α -helices and two β -strands (α 1- β 1- α 2- β 2- α 3- α C), together with the long disordered N- and C-terminal H2A tails and the long N-terminal H2B tail".

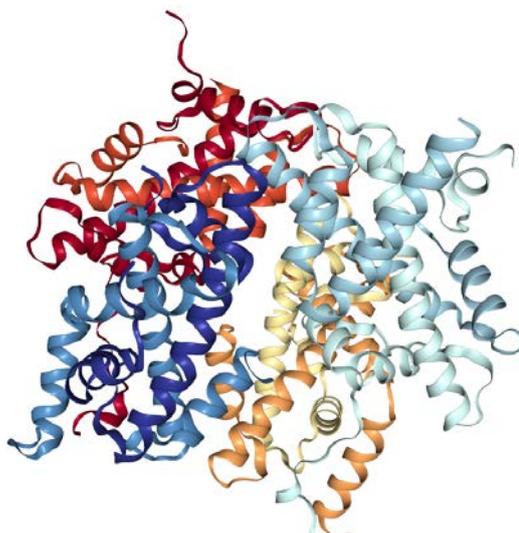


Figure 2.5 Structure of histone octamer

X-ray crystal structure of domestic chicken histone octamer from RCSB PDB Entry 1TZY (Wood *et al.*, 2005). The image was obtained using NGL rendering (Rose and Hildebrand, 2015). Each histone is shown in a separate colour.

1TZY was derived using X-ray diffraction at 1.9 Angstrom resolution, from chicken (*Gallus gallus domesticus*) erythrocytes. Each 8 histone octamer is composed of two histone dimers, (H2A-H2B) and (H2A'-H2B'), and one tetramer, (H4-H3).(H3'-H4'). The dimer and tetramer sub-structures are formed by histone-fold pairing (Arents *et al.*, 1991). The four histone-fold pairs are connected together by three four-helix bundles in the order (H2A-H2B)-(H4-H3)-(H3'-H4')-(H2A'-H2B') (Chantalat *et al.*, 2003).

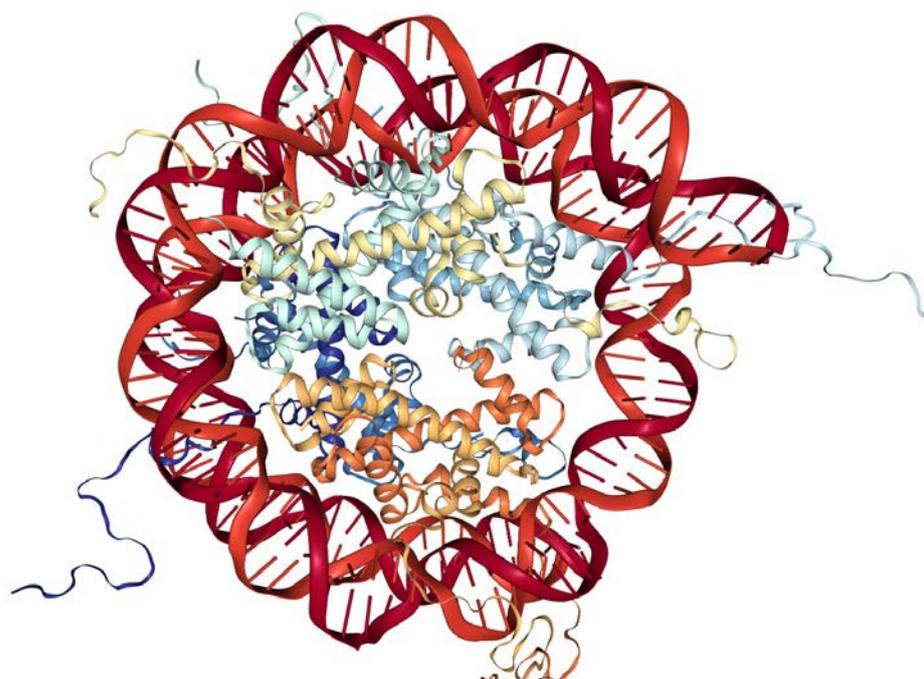


Figure 2.6 Structure of nucleosome core particle

X-ray crystal structure at 1.9 Angstrom resolution of the nucleosome core particle from RCSB PDB Entry 1KX5 (Davey *et al.*, 2002). The image was obtained using NGL rendering (Rose and Hildebrand, 2015).

147 base pairs of human palindromic DNA are wound around a histone octamer assembled from recombinant histones from an African Clawed Frog. Histone tails are shown extending beyond the DNA. These disordered tails are the main repository of PTMs (see Section 2.5.5). Histone interactions with DNA include: (i) electrostatic attraction between the histone helices positive charges and DNA negatively charged phosphate groups, (ii) hydrogen bonds between histone main chain amides and the DNA backbone, (iii) non-polar interactions (hydrophobic effect) which assembles a structure which maximises contact between the polar residues and water molecules (iv) salt bridges and hydrogen bonds between basic amino acids and DNA phosphate oxygens, and (v) H3 and H2B N-terminal tail insertion into DNA minor grooves (Davey *et al.*, 2002).

2.5.3 Linker histones

Each nucleosome core particle is connected to the next by linker DNA in association with a linker-histone protein H1 or a variant such as H5 (Bates and Thomas, 1981; Lambert *et al.*, 1991; Thomas *et al.*, 1992). H1 is the main linker-histone in humans; H5 is more common in birds. There are numerous linker histone variants (Snijders *et al.*, 2008). Figure 2.7 shows the globular region of histone H5. H5/H1 prevalence exceeds one molecule per nucleosome (Table 2.4). Koutzamani *et al.* (2002) indicate that the accumulation of H5/H1 coincides with cessation of cell proliferation and with increasingly compact chromatin. Parseghian (2015) includes a review of linker histone residence times on nucleosomes.

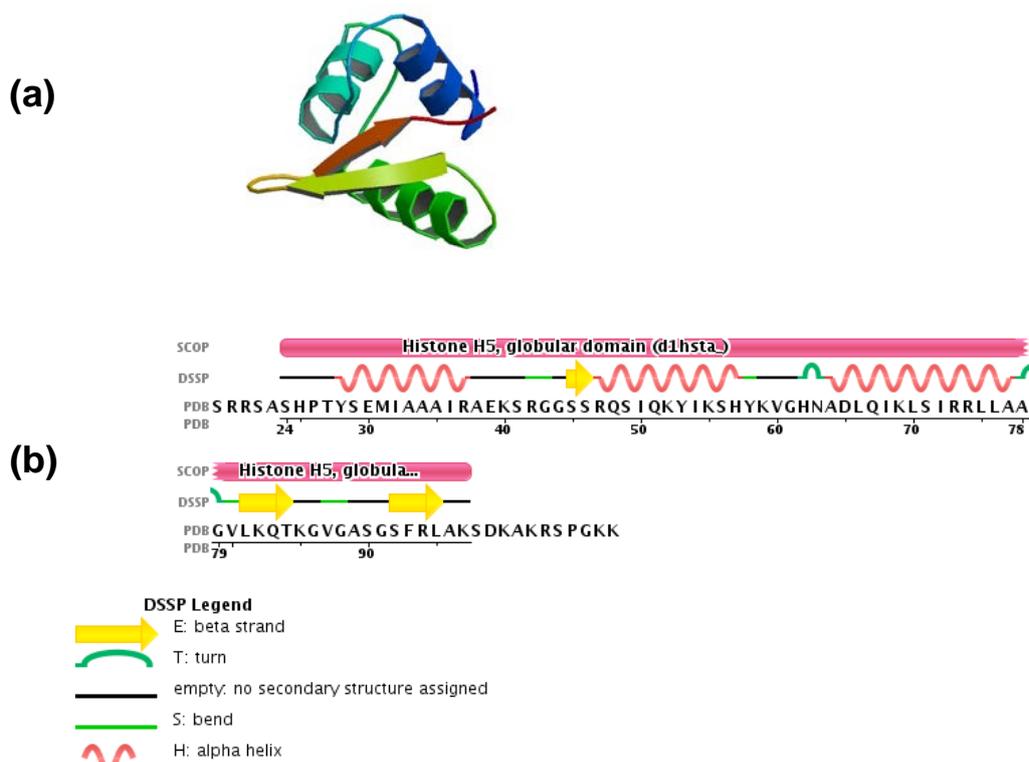


Figure 2.7 Structure of the globular region of histone H5

(a) X-ray structure at 2.6 Angstrom resolution of the H5 globular region; image is as shown in RCSB PDB Entry 1HST (Ramakrishnan *et al.*, 1993). 1HST covers protein residues 24 to 97 of a domestic chicken histone H5 protein cloned in *E-coli* (complete protein has 190 residues).

(b) Sequence chain view copied directly from the PDB sequence display.

2.5.4 Histone interactions

Histone octamer interactions with DNA are described in Figure 2.6.

The cell nuclear pH influences the electrical charges on the histones and other amphoteric molecules in the nucleus, and these charges influence which molecules associate together. The pH of cell organelles varies according to function, and it has proved difficult to find pH data specific to chicken erythrocytes. However, work using pH sensitive fluorescent probes by Bright *et al.* (1987) and Llopis *et al.* (1998) on Swiss 3T3 and HeLa cells respectively indicated that the pH in the nucleus is the same as the surrounding cytoplasm. 3T3 nuclear pH was 7.09 (0.01 SEM – Standard Error of the Mean) in quiescent cells and 7.35 (0.01 SEM) in active cells. HeLa nuclear pH (living cells) was ~7.34 (~0.01 SEM) – estimated from Figure 1b of Llopis *et al.* (1998). As indicated earlier, chicken erythrocytes are generally quiescent. Consequently, it is assumed that chicken erythrocyte nuclear pH is ~7.1.

The 147 base pairs coiled onto the histone octamer have a total charge of $-294e$, with about half this charge on the inside of the coil. Perhaps not co-incidentally, the histone octamer [2 x (H2A, H2B, H3, H4)] total charge is about $+149e$, roughly equal to the coil internal charge. Similarly, the charge on the linker DNA is at least partially balanced by a positive charge on the linker histone. The remaining DNA charge on the nucleosome is balanced by the polar nature of water and various cations (Maeshima *et al.*, 2014).

Histone tails have many important DNA-related functions (Kan and Hayes, 2007; Kan *et al.* 2007; Kan *et al.* 2009; Kowalski and Palyga 2011). The histone tails are the primary sites for post-translational modifications (PTMs), further details in next section.

Figure 2.6 shows the unstructured histone tails extending beyond the DNA wound on the histone octamer. The N-terminal tails of histones H3 and H4 are strongly positive at physiological pH and have a role in folding of the nucleosome array (beads on a string) into more compact structures. Unsurprisingly, the tails have an affinity for DNA, which is strongly negative. In addition, residues 16 – 24 of the H4 N-terminal tail domain will bind to a precisely sculptured acidic patch on histone H2A in an adjacent nucleosome core particle (Luger *et al.*, 1997), leading to compaction when this binding is possible. Several other proteins can bind to the H2A acidic patch; in competing with the H4 tail they modify nucleosome compaction, and hence transcription (Kalashnikova *et al.*, 2012).

2.5.5 Histone Post-Translation Modifications (PTMs)

Dynamic cellular processes involving enzymes (writers) can add small chemical groups onto specific residues in a protein, or otherwise modify a specific residue, after the protein has been translated from RNA (Felsenfeld and Groudine, 2003); this is called a post-translational modification (PTM). The histone acetyltransferases (HATs) are examples of writers. Other enzymes (erasers) can remove PTMs (Hebbes *et al.*, 1994) - the histone deacetylases (HDACs) are examples of erasers. DNA is also subject to post-translational modification, especially methylation. Many histone PTMs have now been identified, usually by MS which detects the additional PTM mass (Zhang *et al.* 2002). Some native PTMs may be removed by the protocols for isolating histones (Rodriguez-Collazo *et al.*, 2009). For example, acid-labile PTMs are compromised by acid-based protein extraction methods (Sut and Biterge, 2017). Apart from the work described in Chapter 4, our work (Chapters 5 – 8) has avoided acid extraction.

However, cellular enzymes for adding and removing PTMs may modify the PTMs during extraction of the histones.

PTMs are variable and transitory, and depend on numerous factors including species, cell type, cell status (for example, whether cell is quiescent), and residue type. PTMs impact on the direct interactions between histones and DNA, or can provide docking sites for PTM readers (Bannister and Kouzarides, 2011; Musselman *et al.*, 2012). Binding of a reader may recruit or stabilize the various components of the protein complexes which mediate transcription, recombination, replication and repair. Histone PTMs are also involved with cell cycle regulation and with cell differentiation. The protein complexes can contain multiple readers, so that the complex can target more specific nucleosomes and associated genomic sites. Many readers can recognise the residue sequence surrounding the PTM.

Histone PTMs provide a control on gene expression, and collectively they are referred to as the histone code (Felsenfeld and Groudine, 2003; Kiefer *et al.*, 2008). Where they are retained from one cell generation to the next, they effectively carry genetic information, and, together with DNA methylation, are referred to as the epigenetic code.

Some of the more common histone PTMs and their readers are briefly discussed below based on a review by Musselman *et al.* (2012). The PTMs comprise:

- (i) Attachment of 1, 2 or 3 methyl groups to lysine and arginine residues, leaving the charge unchanged, but altering the residue hydrophobicity.
- (ii) Attachment of a single acetyl group to lysine residues, neutralising the lysine charge.
- (iii) Attachment of a single phosphoryl group to serine and threonine residues, adding a negative charge to the residues.

The PTMs occlude different amounts of space; for example, a phosphoryl group is larger than a methyl group.

Modified residues are identified below. For a specific modification, the terminology is as follows: histone identification, residue type and sequence position, type of group (me, ac, or ph), and the number of groups attached, if more than one.

Lysine methylation (mono-, di-, or tri) H3(K4, K9, K26, K27, K36, K79); H4K20; H1K26.

These are the best studied PTMs. About a dozen readers have been found. Different readers differentiate between the degree of methylation (mono-, di-, or tri). Examples of a PTM role: the transcriptional complex TFIID can bind to H3K4me3 leading to activation of transcription; alternatively, if the mSin3a histone deacetylase binds to H3K4me3 the gene is repressed.

Arginine methylation (mono-, di-) H3R2, H3R8, H3R17, H3R26; H4R3; H2AR11, H2AR29.

There is little information about these PTMs. The TRDRD3 Tudor domain is thought to recognise the asymmetric forms of H3R17me2 and H4R3me2; this may contribute to activation of transcription. The MLL1/WDR5 complex may recognise H3R2me2 (symmetric form).

Lysine acetylation Numerous lysines on H3, H4 and H2B are acetylated, mostly in their N-terminal tails; also H2AK5 and H2AK9. Readers include the bromodomain protein family and Dpf3b, also the histone chaperone Rtt106 may bind to H3K56ac. It was once thought that the role of lysine acetylation was solely to disrupt the interface between histones and DNA, but the presence of readers has indicated otherwise. The BRD transcriptional regulator family bind via their bromodomains. Recognition of H3K56ac by Rtt06 is necessary for gene silencing and DNA damage response. The SWI/SNF and ACF1/ISWI remodelling complexes include bromodomains.

Serine and threonine phosphorylation H3S10, H3S28; H4S1; H2AS1; H2AXS139; H2BS14; H3T3, H3T6, H3T11, H3T45; H2AT120. These PTMs have a role in DNA damage response, mitosis and transcriptional regulation. The BRCT domain in several proteins in a DNA repair pathway recognises yeast H2AXS139ph which is phosphorylated in response to DNA damage. Phosphorylation of H3S10 by the Aurora B kinase during mitosis is necessary for chromosome condensation and segregation. The 14-3-3 family has some isoforms which recognizes H3S10ph and H3S28ph110; this family has a role in the transcription activation of *GAL1* and *HDAC1*. The BIR domain in Survivin (a member of the inhibitor of apoptosis (IAP) family) recognises H3T3ph.

Table 2.5 summarises histone PTMs, and Figure 2.8 provides some examples. Databases such as PhosphoSite Plus (Table 2.1) record many other PTMs.

Section 2.7.2 includes an outline of the interaction between proline isomerisation and lysine methylation and acetylation.

Table 2.5 Histone PTMs

PTM	Typical target residues
Acetylation	Lysine
Phosphorylation	Serine, threonine, tyrosine, lysine, arginine, histidine
Methylation (mono, di, tri)	Lysine, arginine
Deimination [citrullination]	Arginine
□-N-acetylglucosamination	Serine, threonine
ADP ribosylation (mono, poly)	Glutamate, arginine
Ubiquitylation (mono, poly)	Lysine, cysteine, serine
SUMOylation	Lysine
Histone tail clipping	Tail residues
Proline isomerisation	Proline
Propionylation	Lysine
Butyrylation	Lysine

The table is based on Bannister and Kouzarides (2011), Chen *et al.* (2007), and Sut and Biterge (2017). PTMs are mainly located on histone tails. See also Felsenfeld and Groudine (2003), Horn *et al.* (2002), Kiefer *et al.* (2008), Petesch and Lis (2012), Zhang *et al.*, (2002).

(a)

001 MAGG**X**AG**X**DS G**X**AKAKAVSR SQRAGLQFPV GRIHRHLKTR TTSHGRVGAT
 051 AAVYSAAILE YLTAEVLELA GNASKDLKVK RITPRHLQLA IRGDEELDSL
 101 IKATIAGGGV IPHIH**X**SLIG KKGQQKTA

(b)

001 M**X**SGRG**X**QGG**X** ARA**X**AK**X**SRSS RAGLQFPVGR VHRLLRKGNY AERVGAGAPV
 051 YLAAVLEYLT AEILELAGNA ARDNKKTRII PRHLQLAIRN DEELNKLKLGK
 101 VTIA**X**QGGVLP NIQAVLLPK**X** TDSHKAKAK

Figure 2.8 Examples of histone PTMs

Key to PTMs: **X** Acetylation or phosphorylation, **X** Acetylation, **X** Di-methylation,
X Ubiquitination, **X** Methylation.

(a) *Human histone H2A.V* - PTMs as found for underlined residues by Olszowy *et al.*, 2015; Uniprot sequence Q71UI9-1.

(b) *Chicken histone H2A.IV* – PTMs as recorded in UniProt database for sequence P02263.

2.6 HMGB proteins

2.6.1 HMG super-families

Goodwin and Johns (1973) isolated two groups of proteins from calf thymus chromatin. One group of proteins was easily soluble in 10% trichloroacetic acid and migrated unusually rapidly on a polyacrylamide gel, hence they were called “high-mobility group” (HMG) proteins. After histones, HMG proteins are the most abundant and ubiquitous group of chromosome-related proteins (Kang *et al.*, 2014), and many additional HMG type proteins have been found since 1973. As shown in the Human Protein Atlas (Table 2.1), HMG family members are present in many animal tissues. Most HMG proteins have a molecular weight below 30KDa. They include three HMG super-families: HMGA, HMGB, and HMGN. Table 2.6 identifies some family members, using both modern and old nomenclature.

HMGA proteins contain AT-hook domains (usually 3 domains) which bind to AT-rich DNA at the minor groove. They can also bind non-B-form DNA with specific features such as bent and supercoiled DNA (Reeves, 2010). Animal HMGB protein contain two HMG-box domains which can bind and bend DNA. HMGN proteins contain a domain which binds the nucleosome core particle (Zhang and Wang, 2008). Collectively, the HMGA, HMGB and HMGN super-families are referred to as ‘architectural factors’, and family members are involved in changing nucleosome and chromatin structure, and in the assembly of molecular complexes for processes such as embryonic development, transcription, replication, and DNA repair (Reeves, 2010), and control the expression of some genes (Bianchi and Agresti, 2005).

There are equivalent HMG proteins across a wide range of species, although plant HMGB-type proteins typically have only a single DNA binding box (Pedersen and Grasser, 2010), and HMGB proteins in yeast differ from multicellular species (Zhuang *et al.*, 2014). Within animals, HMG proteins are highly conserved. For example, for selected chicken and human HMGB1 sequences, the chicken residues replicate 90.7% of the human residues, and 98.6% are identical or conservative replacements which are likely to conserve structure and function (Figure 2.9).

Table 2.6 Some HMG family members (old and new names)

New name	Old (alternate) name
Generic	
HMGA1, 2,n	HMG-I/HMG-Y/HMG-C
HMGB1, 2,n	HMG-1/HMG-2 etc
HMGN1, 2,n	HMG-14/HMG-17
Specific examples	
HMGA1a	HMG-I; HMGI; HMG I; HMG-I/Y; a-protein
HMGA1b	HMG-Y; HMGY; HMG Y
HMGA1c	HMG-I/R
HMGA2	HMGI-C
HMGB1	HMG1; HMG-1;HMG 1; amphoterin; p30
HMGB2	HMG2; HMG-2; HMG 2
HMGB3	HMG2a; HMG-2a; HMG 2a (HMG-4)
HMGB4	-
HMGN1	HMG-14; HMG14; HMG-14a; HMG-14b
HMGN2	HMG-17; HMG17; HMG 17
HMGN3	Trip7
HMGN3a	-
HMGN4	-
HMGN5	NSBP1, NBD-45

Notes

Nomenclature is based on Bustin (2001) and HMG nomenclature website (2016). In modern terminology, small letters denote splice variants, eg HMGA1a. Some of the proteins originally containing HMG as part of their name have now been found not to be part of the HMG super-families. For example, HMG20 was later found to be ubiquitin.

2.6.2 HMGB proteins

The HMGB proteins are the most abundant HMG proteins. They are strongly represented in the nucleus, with about one molecule of HMGB1 and HMGB2 for every 10 – 15 nucleosomes (Thomas, 2001). They are part of a wider group of proteins equipped with one or more HMG-boxes all of which are DNA binding (Table 2.7). HMGB proteins are expressed to varying degrees in all human tissues, except for HMGB4 which is almost entirely limited to the testis (Human Protein Atlas). HMGB proteins form part of various cellular pathways, and in addition appear to globally regulate chromatin structure. For example, Agresti and Bianchi (2003) argue that HMGB proteins promote dynamic nucleosome sliding, with repeated binding to chromatin for fractions of a second.

HMGB1, HMGB2, HMGB3, and HMGB4 proteins each have two HMG box domains which are largely positively charged. Each box domain has a similar L-shaped structure formed from three α -helices (Thomas, 2001; Catena *et al.*, 2009; Stros, 2010) (Figure 2.10). The domain nearest the N-terminal is denoted Box A, with the second domain Box B. HMGB1, 2 and 3 proteins each have a long C-terminal acidic tail, although of varying length (Table 2.8). Despite sequence and structure similarity, HMGB proteins have distinct non-redundant functions. Separate knockouts of *HMGB1*, *2*, and *3* genes produce mice with separate distinct health problems, even though these proteins share about 80% sequence identity. Mice lacking HMGB1 die shortly after birth. Mice survive the absence of HMGB2, but have reduced fertility. Mice also survive the absence HMGB3, but the balance between hematopoietic stem cell self-renewal and differentiation is disturbed (Ronfani *et al.*, 2001; Pedersen and Grasser, 2010).

HMGB proteins associate with chromatin and bind and bend linear DNA, but prefer binding to bent or distorted DNA; binding is DNA sequence-independent (Thomas, 2001), (Kang *et al.*, 2014). Despite similar structure, the Box A and Box B domains perform differing functions, and function appears to vary significantly within the HMGB family. Based on rat and bovine HMGB1 and circularisation assays by Teo *et al.* (1995) and Paull *et al.* (1993), Zhang and Wang (2008) indicate that HMGB1 Box B is necessary for significant DNA bending. On the other hand, based on human HMGB2 and atomic force microscopy, Zhang *et al.* (2009) show that HMGB2 Box A is sufficient to provide a 78° DNA bend.

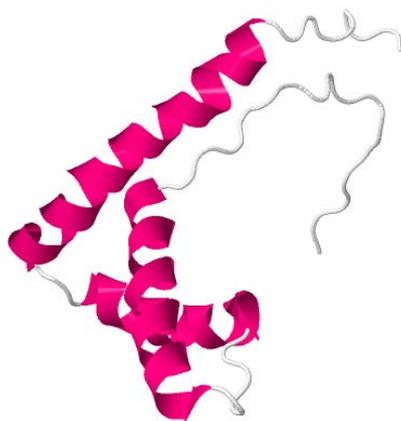
Although the full significance of the HMGB acidic tails is still to be determined, Ueda *et al.* (2004) have shown that the last five amino acid residues DDDDE at the C-terminal end of HMGB1 are essential for transcription stimulation and for binding of HMGB1 to nucleosome linker DNA, and that the DDDDE sequence interacts with the core histone H3 N-terminal tail. Belgrano *et al.* (2013) show that the HMGB1 acidic tail improves protein stability and increases DNA bending. Lee and Thomas (2000) indicate that removal of the tail increases affinity of the HMG boxes for DNA, and abolishes functional differences between HMGB1 and HMGB2. Shirakawa *et al.* (1997) show the acidic tail of HMGB2 is required if HMGB2 is to be retained in the nucleus of COS-7 cells.

Table 2.7 Examples of mammalian HMG-box proteins

Protein	Number of HMG-boxes	Sequence-specific DNA binding
UBF	6	No
HMGB1	2	No
HMGB2	2	No
HMGB3	2	No
HMGB4	2	No
TFAM	2	No
SP100-HMG	2	No
LEF1	1	Yes
TCF	1	Yes
TOX	1	Yes
SRY	1	Yes
SOX	1	Yes
BAF57	1	Yes
PB1	1	Yes
WHSC1	1	Yes
HPB1	1	Yes
BBX	1	Yes
CIC	1	Yes
SSRP1	1	Yes

Based on Kang *et al.* (2014).

(a)



(b)

GSH	MGKGDPKKPR	GKMS	SYAFFV	QTCREEHKKK	HPDASVNF	SE	FSKKCSERWK	50
TMS	AKEKGGKF	EDMAKADKAR	YEREMK	TYIP	PKGETKKKFK	DPNAPKRPPS		100
	AFFLFCSEYR	PKIKGE	HPGL	SIGDVAKCLG	EMWNN	TAADD	KQPYEKKA	150
	LKEKYEKDIA	AYR	AKGKPDA	AKKGVVKAEK	SKKKKEEEED	EDEEDEEEEE		200
	EDEEDEEEEE	DDDDE						215

Figure 2.10 Structure of HMGB1

(a) *HGB1 Box A structure*. NMR solution structure of protein RCSB PDB 2RTU, comprising residues GSH + residues 1- 84 of human HMGB1, Uniprot P09429 (GSH residues arise from the *E. coli* expression vector); NGL rendered (Rose and Hildebrand, 2015). There are three red α -helices in a twisted L-shaped fold a, plus bends, turns, and some residues with no secondary structure assigned. Oxidation conditions have generated a disulphide bond between cysteines C23 (α -Helix I) and C45 (α -Helix II). A full structure is not available because of molecule size and the unstructured tail which confounds crystallisation. Structure citation: Wang *et al.* (2013b).

(b) *HMGB1 full sequence for P09429*. Helices are highlighted in yellow, with the three Box A α -helices residues in red text, and the three Box B α -helices residues in green text. The acidic C-terminal tail is highlighted in grey. However, in the databases there is some variation as to which residues are identified as part of helices and as part of the HMG-box domains. For a slightly different sequence, Stros (2010) indicates the following: Box A - G2 to I79; Box B - F89 to R163.

Table 2.8 HMGB molecular weights, charges, and acidic tail lengths

Protein	Uniprot identifier	N	MW* kDa	pI	Net charge at pH 7.1** e	Residues in acidic tail***
HMGB1 human	P09429	215	24.89	5.73	- 4.6	30
HMGB1 chicken	Q9YH06	215	24.91	5.73	-4.6	30
HMGB2 human	P26583	209	24.03	7.62	+1.6	22
HMGB2 chicken	P26584	207	23.83	8.38	+3.6	20
HMGB3 human	O15347	200	22.98	8.36	+3	20
HMGB3 chicken	P40618	202	23.06	8.12	+2	21
HMGB4 human#	Q2WW32	186	22.49	10.2	+28.5	0

Notes

Data was calculated using Scripps Protein Calculator (2016)(Version 3.3: CD Putnam). The pI estimate assumes all residues have pKa values that are equivalent to the isolated residues. For a folded protein this is not valid. However, the approximate value can be useful for planning protein purifications. Similar considerations apply to the calculated charge.

No sequence found for HMGB4 chicken.

N number of amino acid residues.

* Isotopically averaged molecular weight.

** Assumed pH in nucleus of inactive chicken erythrocyte.

*** Number of contiguous acidic residues.

2.6.3 HMGB1 function and biological processes

As summarised in Table 2.9, HMGB1 is involved in chromatin/DNA related molecular functions, including DNA binding and bending, RNA binding, and interactions with transcription factors. Examples include its function as a chaperone (see Section 2.4.5), and the replacement of H1 in nucleosomes - see example (b) below. It also functions within the immune system, particularly when cells are stressed (including cytokine activity, and binding to specific molecules such as the RAGE receptor). Immune system aspects are only briefly discussed here, since the experimental focus is on HMGB1 in the nucleus. Similarly, HMGB1 is involved in the biological processes listed in Table 2.10. Most processes relate to chromatin or the immune system. However, three of the processes listed relate to cellular development, and one involves the catabolic pathway.

Examples of processes involving HMGB1 include:

- (a) DNA repair (Lange *et al.*, 2008).
- (b) Transcription – example: Ueda *et al.* (2004) indicate HMGB1 may stimulate transcription by binding to linker DNA; Cato *et al.* (2008) describe an interaction between recombinant HMGB1 and linker histones; Ju *et al.* (2006) show that HMGB1 replaces H1 in the nucleosome during TopoII β /PARP-1–dependent regulated gene transcription.
- (c) Secretion of HMGB1 from stressed cells (necrotic and apoptic cells, and cells stimulated with endotoxin) as a signalling molecule with pro-inflammatory cytokine activity (Scaffidi *et al.*, 2005). In vitro, antibodies raised against the HMGB1 B-box domain inhibit the cytokine activity (Li *et al.*, 2003). This cytokine activity has been extensively studied, for example see Diener *et al.* (2013), Stros (2010), Abrahams (2009), Cohen *et al.* (2009), Klune *et al.* (2008), Zhang and Wang (2008), Urbonaviciute *et al.* (2007), Kokkola *et al.* (2003), and Wang *et al.* (1999).
- (d) V(D)J recombination (part of the adaptive immune system) (Swanson, 2002).
- (e) Cell migration, for example via binding to RAGE (Huttunen *et al.*, 2002).
- (f) Regulation of telomerase. HMGB1 knockout in mouse embryonic fibroblasts (MEFs) reduces telomerase activity and telomere length, whereas overexpression of HMGB1 enhances telomerase activity (Polanska *et al.*, 2012). Conversely, HMGB2 knockout increased telomerase activity in MEFs. This may arise from subtle sequence variations between HMGB1 and HMGB2 sequences - the amino-acid residues immediately either side of the A-box in HMGB1 are slightly different to the residues in the same locations in HMGB2 (Foulger *et al.*, 2012)].

Table 2.9 Molecular functions determined for human HMGB1

Bubble DNA binding	Source: AgBase
Calcium-dependent protein kinase regulator activity	Source: Ensembl
Chemoattractant activity	Source: UniProtKB
C-X-C chemokine binding	Source: UniProtKB
Cytokine activity	Source: UniProtKB
Damaged DNA binding	Source: UniProtKB
DNA binding, bending	Source: UniProtKB
DNA polymerase binding	Source: UniProtKB
Double-stranded DNA binding	Source: UniProtKB
Double-stranded RNA binding	Source: Ensembl
Four-way junction DNA binding	Source: AgBase
Lipopolysaccharide binding	Source: UniProtKB
Lyase activity	Source: UniProtKB
Phosphatidylserine binding	Source: UniProtKB
Poly(A) RNA binding	Source: UniProtKB
Protein kinase activator activity	Source: Ensembl
RAGE receptor binding	Source: UniProtKB
Repressing transcription factor binding	Source: UniProtKB
Single-stranded DNA binding	Source: UniProtKB
Single-stranded RNA binding	Source: Ensembl
Supercoiled DNA binding	Source: AgBase
Transcription factor activity, sequence-specific DNA binding	Source: UniProtKB
Transcription factor binding	Source: UniProtKB

Notes

List was derived from the “Function” section for P09429 (human HMGB1) in the UniProtKB database (Table 2.1).

Text indicates function related to DNA, RNA, and transcription processes.

Text indicates function related to immune system and stressed cells.

Table 2.10 Biological processes determined for human HMGB1

Page 1 of 2

Activation of innate immune response	Source: UniProtKB
Apoptotic cell clearance	Source: UniProtKB
Apoptotic DNA fragmentation	Source: Reactome
Autophagy	Source: UniProtKB-KW
Base-excision repair	Source: Ensembl
Chromatin assembly	Source: Ensembl
Dendritic cell chemotaxis	Source: UniProtKB
DNA geometric change	Source: AgBase
DNA ligation involved in DNA repair	Source: UniProtKB
DNA recombination	Source: UniProtKB
DNA topological change	Source: UniProtKB
Endothelial cell chemotaxis	Source: Ensembl
Endothelial cell proliferation	Source: Ensembl
Eye development	Source: Ensembl (2)
Inflammatory response	Source: CACAO
Inflammatory response to antigenic stimulus	Source: UniProtKB
Innate immune response	Source: Reactome
Lung development	Source: Ensembl (2)
Macrophage activation involved in immune response	Source: Ensembl
Myeloid dendritic cell activation	Source: UniProtKB
Negative regulation of apoptotic cell clearance	Source: Ensembl
Negative regulation of blood vessel endothelial cell migration	Source: CACAO
Negative regulation of CD4-positive, alpha-beta T cell differentiation	Source: UniProtKB
Negative regulation of interferon-gamma production	Source: UniProtKB
Negative regulation of RNA polymerase II transcriptional preinitiation complex assembly	Source: UniProtKB
Negative regulation of transcription from RNA polymerase II promoter	Source: UniProtKB
Neuron projection development	Source: UniProtKB (2)
Neutrophil clearance	Source: UniProtKB
Plasmacytoid dendritic cell activation	Source: Ensembl
Positive regulation of activated T cell proliferation	Source: UniProtKB
Positive regulation of apoptotic process	Source: UniProtKB
Positive regulation of cysteine-type endopeptidase activity involved in apoptotic process	Source: UniProtKB
Positive regulation of cytosolic calcium ion concentration	Source: UniProtKB
Positive regulation of dendritic cell differentiation	Source: UniProtKB
Positive regulation of DNA binding	Source: UniProtKB
Positive regulation of DNA ligation	Source: UniProtKB

Table 2.10 Biological processes determined for human HMGB1

Page 2 of 2

Positive regulation of ERK1 and ERK2 cascade	Source: UniProtKB
Positive regulation of glycogen catabolic process	Source: Ensembl (3)
Positive regulation of interferon-alpha production	Source: Ensembl
Positive regulation of interferon-beta production	Source: Ensembl
Positive regulation of interleukin-10 production	Source: UniProtKB
Positive regulation of interleukin-12 production	Source: UniProtKB
Positive regulation of interleukin-1 beta secretion	Source: Ensembl
Positive regulation of interleukin-1 secretion	Source: UniProtKB
Positive regulation of interleukin-6 secretion	Source: UniProtKB
Positive regulation of JNK cascade	Source: UniProtKB
Positive regulation of MAPK cascade	Source: UniProtKB
Positive regulation of mismatch repair	Source: UniProtKB
Positive regulation of monocyte chemotaxis	Source: UniProtKB
Positive regulation of myeloid cell differentiation	Source: Ensembl (2)
Positive regulation of NIK/NF-kappaB signaling	Source: Ensembl
Positive regulation of sprouting angiogenesis	Source: Ensembl
Positive regulation of toll-like receptor 2 signaling pathway	Source: Ensembl
Positive regulation of toll-like receptor 4 signaling pathway	Source: Ensembl
Positive regulation of toll-like receptor 9 signaling pathway	Source: UniProtKB
Positive regulation of transcription from RNA polymerase II promoter	Source: UniProtKB
Positive regulation of tumor necrosis factor production	Source: Ensembl
Positive regulation of wound healing	Source: Ensembl
Regulation of autophagy	Source: UniProtKB
Regulation of nucleotide-excision repair	Source: Ensembl
Regulation of restriction endonuclease activity	Source: UniProtKB
Regulation of T cell mediated immune response to tumor cell	Source: UniProtKB
Regulation of tolerance induction	Source: UniProtKB
Regulation of transcription from RNA polymerase II promoter	Source: UniProtKB
Response to glucocorticoid	Source: Ensembl (4)
T-helper 1 cell activation	Source: UniProtKB
T-helper 1 cell differentiation	Source: UniProtKB
Tumor necrosis factor secretion	Source: UniProtKB
V(D)J recombination	Source: UniProtKB

Notes

List was derived from the “Function” section for P09429 (human HMGB1) in the UniProtKB database (Table 2.1).

Text indicates function related to DNA, RNA, and transcription processes.

Text indicates function related to immune system and stressed cells.

2.6.4 HMGB1 in development and disease

Hock *et al.* (2007) summarise some of the relationships between HMGs and development and disease. HMG levels vary throughout development and have a critical impact. Although HMG knockout mice embryos all survive the loss of a single HMG and the mice are born, all such mice suffer some abnormalities. For example, loss of HMGB1 leads to mice dying with 24 hrs of birth and increased chromosomal instabilities. HMGB1 is upregulated in colon, breast, gastric and gastrointestinal cancer cells. Kang *et al.* (2014) review HMGB1 in health and disease based on ~2000 references; they identify changes in HMGB1 expression in ~100 diseases, including various cancers. Hence HMGB1 attracts interest as a drug intervention target, and as a diagnostic and prognostic tool (Gibot *et al.*, 2007). One focus is seeking to ameliorate the pro-inflammatory impact of HMGB1, as a means of avoiding lethal sepsis (Zhu *et al.*, 2009; Rosas-Ballina *et al.*, 2009; Wang *et al.*, 1999).

2.6.5 Functions of other HMGB proteins

HMGB2 influences many processes, including:

- Inhibition of chondrocyte cell differentiation (Taniguchi *et al.*, 2011)
- Suppression of pathologic cell growth (unlike HMGB1) (Franklin *et al.*, 2012)
- Cancer progression (Kwon *et al.*, 2010; Lee *et al.*, 2010)
- Sex cycle of the *Plasmodium* parasites (Gissot *et al.*, 2008)
- Like HMGB1, HMGB2 also associates with RAGE (Wang *et al.*, 2013b)

There is less information for HMGB3 and HMGB4 than for HMGB1 and HMGB2. HMGB3 is found on the X chromosome. HMGB3 has been shown to have a role in regulating hematopoietic stem cell self-renewal and differentiation, and in the Wnt signaling pathway (Nemeth *et al.*, 2003, 2005, 2006). Fusion of the NUP98-HMGB3 genes has been implicated in acute myeloblastic leukemia (AML) (Petit *et al.*, 2010). HMGB4 is similar to the other HMGB proteins except that it lacks the acidic tail, and is expressed only in testis and weakly in the brain. HMGB4 is a transcriptional repressor (Catena *et al.*, 2009).

2.6.6 HMGB PTMs

HMG proteins are subject to PTMs (Table 2.11) with various functions. For example, methylation of HMGB1 Lys-42 disrupts the Box A structure, reduces DNA binding, and encourages HMGB1 to diffuse into the cytoplasm (Zhang and Wang, 2008).

Phosphorylation or acetylation within the two nuclear localization signal (NLS) domains in HMGB1 inhibits transfer into the nucleus and mediates secretion (Kang *et al.*, 2014). De-phosphorylation of serine residues close to the acidic tail of fruit fly embryo HMGB1 greatly reduces binding to four-way junction DNA (Wiśniewski *et al.*, 1999). Some HMGB PTMs are added/removed by the same enzymes as histone PTMs (Zhang and Wang, 2008), suggesting that HMGB proteins may overlap in function with histones (other evidence of overlap includes the model described by Ueda *et al.* (2004) in which HMGB1 binds to linker DNA).

Table 2.11 HMG PTMs

PTM	Typical target residues
Acetylation	Lysine, Glutamic acid
Phosphorylation	Arginine, Lysine, Serine, Threonine, Tyrosine
Methylation (mono, di)	Lysine, Arginine
Succinylation	Lysine
Sumoylation	Lysine
Ubiquitylation	Lysine
Disulphide bond	Cysteine
Cysteine sulphonic acid	Cysteine

Notes

The above PTMs are a composite of those for (i) HMGB1, 2, and 3 (mammal) in the PhosphoSite Plus database, and (ii) for P09429 (human HMGB1) in the UniProtKB database. Databases are as per Table 2.1.

2.7 FKBP3

Our group has been investigating whether the FKBP3 protein (also known as FKBP25) forms a complex with HMGB proteins as indicated by Leclercq *et al.* (2000). Based on proteins from chicken erythrocytes, our experiments so far have not confirmed such a complex (Zhuang, 2011). However, the efficient method for isolating FKBP3 described in Chapter 8 will assist our further work in this area.

2.7.1 FKBP3 molecular weight, charge, structure, and sequence

Human and chicken FKBP3 molecular weights and charges are given in Table 2.12, and PTMs are given in Table 2.13. Figure 2.11 shows the NMR solution structure 2MPH for human FKBP3. The N-terminal comprises a compact tilted bundle of 5 alpha helices. Helander *et al.* (2014) propose that the bundle is involved in DNA binding. The C-terminal end contains the so called FKBP binding domain; this is responsible for the most studied FKBP3 interactions. This is a five-strand anti-parallel beta sheet (one strand is split) and an alpha helix. The C-terminal includes two helices which occur before the last beta strand in the residue sequence, but they do not interfere with the beta sheet. Human FKBP3 is expressed across all tissues. Its main location is the cytosol - see Human Protein Atlas (Table 2.1) data for human cell lines U-2 OS, A-431, and U-251 MG. However, some limited FKBP3 is also present in the nucleus (Figure 2.12).

As of January 2017, a search in the US, European, or Japanese structural databases (RCSB PDB, PDBe, PDBj) found no experimental structure for chicken FKBP3. However, the UniProtKB sequence Q90ZK7 is available for chicken FKBP3. The human and chicken FKBP3 sequences are very similar - 92.4% of chicken FKBP3 residues are the same or equivalent residues to those in human FKBP3 (Figure 2.13).

In fact, FKBP3 is widely conserved, especially within mammals. For human, rabbit, mouse, and bovine FKBP3 sequences, only residues 77, 91, 104, and 215 are non-equivalent (Leclercq *et al.*, 2000).

Table 2.12 FKBP3 molecular weights and charges

Protein	UniprotKB identifier	N	MW* kDa	pI	Net charge at pH 7.1** e
FKBP3 human	Q00688	224	25.18	9.27	+8.9
FKBP3 chicken	Q90ZK7	227	25.03	9.19	+7.7

Notes

Data was calculated using Scripps Protein Calculator (2016). The pI estimate assumes all residues have pKa values that are equivalent to the isolated residues. For a folded protein, this is not valid. However, the approximate value can be useful for planning protein purifications. Similar considerations apply to the calculated charge.

N number of amino acid residues.

*Isotopically averaged molecular weight.

**Assumed pH in nucleus of inactive chicken erythrocyte.

Table 2.13 FKBP3 PTMs

Composite PTMs (1)	Typical target residues
Acetylation	Lysine, Alanine
Phosphorylation	Serine, Tyrosine, Threonine
Methylation (mono)	Lysine, Arginine
Succinylation	Lysine
Ubiquitylation	Lysine
Q00688 PTMs (2)	
Acetylation	Alanine (Residue 2)
Phosphorylation	Serine (Residue 36)
Acetylation	Lysine (Residue 99)
Phosphorylation	Serine (Residue 152)
Acetylation	Lysine (Residue 170)

Notes

Composite PTMs are based on various cell lines and tissues, from human, mouse, and rat in the PhosphoSite Plus database, accessed Jan 2017. References supporting these PTMs are included in the database. The other PTMs apply to UniProtKB Q00688 (FKBP3_human), taken from the UniprotKB database Jan 2017. See Table 2.1 for databases.



Figure 2.11 Structure of FKBP3

(a) NMR solution structure 2MPH for human FKBP3 for residues 1- 224, shown in cartoon style using the JSMOL viewer (<http://www.jmol.org/>). Beta strands are shown in yellow, alpha helices in red, pi helices in purple. At the N-terminal end there is a bundle of 5 helices. At the C-terminal end, there is an FKBP binding domain comprising an anti-parallel beta sheet (4 full strands and one split strand) and an alpha helix, and two pi helices.

(b) Schematic for 2MPH residues 2-224 showing the helices and beta strands corresponding to the above structure. Beta strands are shown in yellow, helices in red (all are alpha helices except last two which are pi helices). Annotation marks relate to PTMs.

Images citation: RCSB PDB 2MPH, deposited by Shin J, Prakash A, Yoon H in 2014. Report still to be published, but see Prakash *et al.* (2016b).

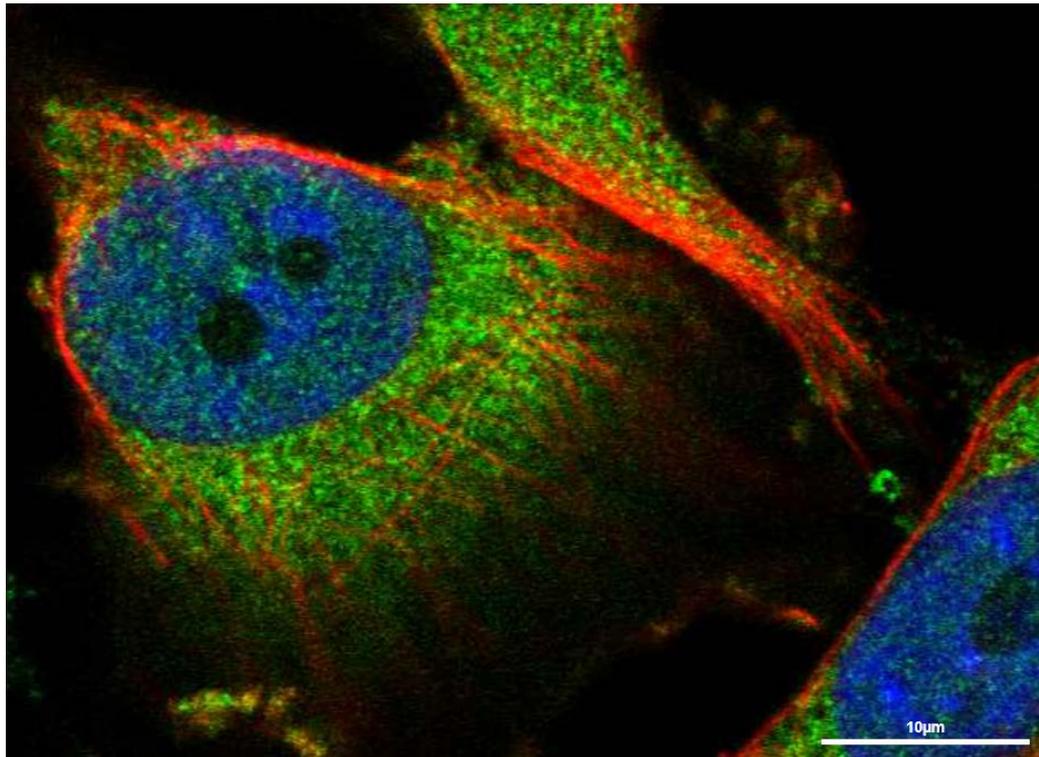


Image credit: Human Protein Atlas V16 accessed 14 Jan 2017; Uhlen *et al.* (2015).

Figure 2.12 Location of FKBP3 in human cell line U-251 MG

Antibody green stain (Abcam CAB012232) shows location of FKBP3 in a human cell (cell line U-251 MG human glioblastoma astrocytoma). Other stains show nucleus in blue, microtubules in red. Green speckles indicate some FKBP3 is present in the nucleus, although at low level.

```

          10          20          30          40          50          60
EMBOSS  AAVPQRAWTVEQLRSEQLPKKDI IKFLQEHGSDSFLAEHKLLGNIKNVAKTANKDHLVT
          :.:.:. . :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.:
EMBOSS  AATAPAQPWSAEELRSEALPKKDI IKFLQEHAAQAFLAEHRLLGQVKNVAKTANKEQLIA
          10          20          30          40          50          60

          70          80          90          100         110
EMBOSS  AYNHLFETKRFKGTESISKVVSEQVKNVCLN--EDKPKETKSEETLDEGPPKYTKSVLKKG
          :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.:
EMBOSS  AYTQLFHTQRFKGTGAERAAEKAKPGKAEGEKEKDKAAKAEPAEEGPPKYTKSILKKG
          70          80          90          100         110         120

120      130      140      150      160      170
EMBOSS  DKTNFPKKGDVVHCWYTGTLQDGTVFDTNIQTSAKKKKNAKPLSFKVGVGKVIKRWDEAL
          :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.:
EMBOSS  DKTNFPKKGDTVHCWYTGKLDGTVFDTNVQTSKKKKAAKPLSFKVGVGKVIKRWDEAL
          130      140      150      160      170      180

180      190      200      210      220
EMBOSS  LTMSKGEKARLEIEPEWAYGKKKQPDAKIPPNAKLTFEVELVDID
          :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.: :.:.:.:
EMBOSS  LTMSKGEKAQLEIEPEWAYGKKKQPDAKIPPNAKLFFEVELVDIE
          190      200      210      220

```

Image citation: Result from Lalign web-based service, alignment algorithm accessed Dec 2016 - Huang and Miller (1991).

Figure 2.13 Comparison of sequences between human and chicken FKBP3

Residues 2 - 224 for FKBP3 human (UniprotKB Q00688) are shown in **yellow**. Other residues are 3 - 227 for FKBP3 chicken (Uniprot Q90ZK7). There are 168 identical peptides from (74.7%), and 208 identical or conservative replacement peptides (92.4%) (based on 225 total residues).

Key

- “.” identical peptide.
- “.” conservative replacement of peptide (similar function).
- “ “ non-conservative replacement of peptide (dissimilar function).

2.7.2 FKBP3 function and biological processes

FKBP3 binds to various molecules, especially to macrolides. It is also a peptidyl-prolyl isomerase (PPI). Macrolides are a class of natural molecule which have a large ring-shaped structure to which one or more sugars are attached. From 1971 onwards several macrolides were discovered which inhibited pathways that lead to lymphocyte activation; hence they suppress the immune response. Macrolides with this property include cyclosporin A, FK506 (tacrolimus), and rapamycin (sirolimus). These molecules have important medical uses, such as preventing rejection of organ transplants. Subsequently, proteins called immunophilins were found which bind to these macrolides, countering their immunosuppressive effects (Schreiber, 1991). One family of immunophilins which bound to FK506 was designated FKBP3 (FK506 Binding Proteins). The family also bound rapamycin. FKBP3 was the first immunophilin to be found in cell nuclei (Helander *et al.*, 2014).

In addition to cytosolic activities such as rapamycin binding, human FKBP3 is involved in nuclear activities, including binding to various nuclear proteins and also to RNA and DNA – see Tables 2.14 and 2.15. The UniProtKB database (Table 2.1) lists only two molecular functions and one biological process for the chicken FKBP3 Q90ZK7 (Table 2.16). However, the EMBL tool STRING v10 (Table 2.1) predicts three proteins as a common partner for both chicken and human FKBP3 - transcriptional repressor YY1, the serine/threonine kinase MTOR involved in regulation of cellular metabolism, and the LARP6 protein involved in regulation of translation (Table 2.17).

Aside from their binding ability, PPIs can catalyse the rotation (reversibly between cis/trans positions) of the peptidyl-prolyl amide bond in protein substrates. Schreiber (1991) refers to this as rotamase activity, but more recently the term used is peptidyl-prolyl isomerase (PPIase) activity. Protein peptide bonds can align to provide either a cis or trans isomer, but most peptide bonds adopt the trans isomer, since this offers less steric repulsion to the preceding C α atom than the cis isomer. However, for peptide bonds to proline, there are almost equal steric clashes so the cis isomer may be nearly as common as the trans isomer. Proline residues are synthesized in the ribosome in trans isomer form. PPIs are found in both eukaryotes and prokaryotes.

There is evidence that PPIs interact directly with chromatin. Fpr4 is a member of an FK506 binding protein family (*Saccharomyces cerevisiae*) and is regarded as a histone chaperone (Monneau *et al.*, 2013). *In vitro* studies by Nelson *et al.* (2006) indicated that, via interaction with the histone tail and isomerisation of H3P38, isomerization by Fpr4 hinders the ability of Set2 to methylate H3K36. Consistent with this, removal (by a single nucleotide mutation) of Fpr4 catalytic activity *in vivo* increased the levels of H3K36 methylation. This interfered with the transcription of specific genes in yeast. A second interaction involving Fpr4 and histone H3 has been examined by Howe *et al.* (2014). They show H3K14 is important for isomerization at H3A15-H3P16 - this controls, via a subunit of the Set1 K4 methyltransferase complex, methylation of H3K4. H3K14 acetylation increases the A16-P16 *trans* configuration, and this reduces H3K4me3.

Table 2.14 Molecular functions determined for human FKBP3

Function	Reference
Rapamycin and mTOR binding	Lee <i>et al.</i> (2016)
FK506 binding.	Hung and Schreiber (1992) Prakash <i>et al.</i> (2016a)
Peptidyl-prolyl cis-trans isomerase activity*	GO_Central - inferred
Poly(A) RNA binding	Castello <i>et al.</i> (2012)
DNA binding	Helander <i>et al.</i> (2014)
Receptor activity	Hung and Schreiber (1992)
Histone binding	Galat and Thai (2014)
YY1 binding	Yang <i>et al.</i> (2001)
HDAC binding	Yang <i>et al.</i> (2001)
Nucleic acid recognition	Prakash <i>et al.</i> (2016b)
Binding to nucleolin engaged with RNA Casein kinase II binding	Gudavicius <i>et al.</i> (2014)

Table 2.15 Biological processes determined for human FKBP3

Function	Reference
Stimulates auto-ubiquitylation of MDM2 leading to its degradation, so that p53 is induced.	Ochocka <i>et al.</i> (2009)
FKBP proteins form part of a mechanism for translocating proteins	Putyrski and Schultz (2012) Lee <i>et al.</i> (2016)
Chaperone-mediated protein folding	GO_Central – inferred from biological aspect of ancestor.
Regulates transcription factors, histone chaperone activity, and chromatin structure	Yau <i>et al.</i> (2011)
Interaction with ribosomal proteins, ribosomal processing factors, and chromatin modifiers. Involvement with ribosome biogenesis.	Gudavicius <i>et al.</i> (2014)

Table 2.16 Molecular functions and biological processes for chicken FKBP3

Molecular Function	Reference
FK506 binding.	GO_Central - inferred from biological aspect of ancestor.
Peptidyl-prolyl cis-trans isomerase activity*	GO_Central - inferred from biological aspect of ancestor.
Biological processes	
Chaperone-mediated protein folding	GO_Central – inferred from biological aspect of ancestor.

Note

Above information was taken from the UniProtKB database for Q990ZK7.

Table 2.17. Predicted functional partners for human and chicken FKBP3

Protein	Predicted functional partner	
	Human FKBP3 (UniProtKB Q00688)	Chicken FKBP3 (UniProtKB Q90ZK7)
HDAC2	Yes	
YY1	Yes	Yes
MTOR	Yes	Yes
UBC	Yes	
SMG1	Yes	
BOLA2	Yes	
LARP6	Yes	Yes
ENSG00000261740	Yes	
BOLA2B	Yes	
RPTOR		Yes
SMG1		Yes
FKBP10		Yes
SNRPE (ribonucleoprotein)		Yes
NFX1		Yes
WTAP		Yes
ENSGAL0000002399		Yes
GTF2H1		Yes
NFXL1		Yes
LOC420093		Yes
PRKDC		Yes

Note

Partners were identified by the EMBL tool STRING v10 (Table 2.1).

Chapter 3

Methods used in our work, including a review of similar methods

<u>Contents</u>	Page
3.1 Introduction	59
3.2 Basic methods for isolating and assaying nuclear proteins	59
3.3 Our approach to isolation and assay of nuclear proteins	63
3.4 Buffers, protease inhibitors, centrifugation, filtration, dialysis, freeze drying, and general laboratory practice	65
3.5 Preparation of chicken erythrocyte nuclei, and isolation of nuclear protein sets	73
3.6 Protein separation and assay using chromatography	76
3.7 Proteins separation and assay using SDS-PAGE	81
3.8 Protein identification using MS	82
<u>Tables</u>	
Table 3.1 Papers consulted to establish the basic methods used by others for isolating and assaying nuclear proteins	60
Table 3.2 Our AKTA equipment and user documentation	78
Table 3.3 Example of change in protein charge with pH	80
<u>Figures</u>	
Figure 3.1 Our approach to isolating nuclear proteins	64
Figure 3.2 Equipment for supplying de-ionised water	66
Figure 3.3 Buffer filtering	67
Figure 3.4 Ultrafiltration arrangement_	69
Figure 3.5 Effect of dialysis on buffer concentrations	71
Figure 3.6 Devices used to mildly homogenise nuclei to prevent clumping	73
Figure 3.7 pH of different mixtures of monobasic and dibasic potassium phosphate	74
Figure 3.8 Schematic of our AKTA Purifier 10 chromatography system	77
<u>Appendices</u>	
Appendix 3.1 SDS-PAGE protocol	84
Appendix 3.2 Sheffield University Protocol for treatment of excised SDS-PAGE gel bands containing proteins, in preparation for MS	95

3.1 Introduction

This chapter outlines the methods we used to isolate and investigate proteins and complexes, including a review of similar methods used by others (Section 3.2).

Section 3.3 introduces the general approach used in the present work, as developed over a period of time by a research group at Liverpool John Moores University - for example, see Pongdam (2008), Foulger (2011), and Zhuang (2011). Laboratory methods are outlined in Section 3.4, with more complex methods (chromatography, SDS-PAGE, and MS) in Sections 3.5 – 3.8. Detailed protocols for preparation of chicken erythrocyte nuclei, extraction of proteins from intact nuclei, nuclei lysis, and precipitation of sets of proteins are provided in the Chapters which follow.

Our source material was fresh chicken blood, available as a waste product from a chicken farm. Erythrocytes (red blood cells) were prepared from the blood as a source of nuclear proteins closer to *in-vivo* proteins than those produced in cell lines or by recombinant DNA methods.

3.2 Basic methods for isolating and assaying nuclear proteins

To put the present work into context, some basic alternative methods for isolating and assaying nuclear proteins were reviewed; the methods are summarised below. Table 3.1 lists the papers consulted. Advanced methods such as *in-vivo* studies using fluorescence (Coats *et al.*, 2013), and ChIP-seq such as used by Song *et al.* (2016) were not reviewed.

Table 3.1 Papers consulted to establish the basic methods used by others for isolating and assaying nuclear proteins

Isolation and assay of histone proteins

Chauveau *et al.* (1956) - nuclei preparation

Murray (1966) - acid extraction from cells

Johns (1967) – separation/assay/quantification using acidic gels

Panyim and Chalkley (1968) – inhibition of proteolytic enzymes; AU gels

Chen *et al.* (1974) – high-salt extraction

Eickbush and Moudrianakis (1978) – high salt extraction

Gurley *et al.* (1983) – cell/nuclei lysis; acid extraction

Hake *et al.* (2006) – acid extraction

Shechter *et al.* (2007) – acid extraction, high-salt extraction

Rodriguez-Collazo *et al.* (2009) – urea/acid extraction/purification

Isolation and assay of non-histone proteins

Goodwin and Johns (1973) – combined or separate lysis of nuclei/cells; salt extraction.

Goodwin *et al.* (1975) – combined lysis of nuclei and cells; salt extraction.

Sanders (1977) – combined lysis of nuclei and cells, acid extraction

Isackson *et al.* (1980) - separate lysis of cells and nuclei; salt extraction

Garg and Reeck (1998) - separate lysis of cells and nuclei; high salt extraction.

Leclercq *et al.* (2000) - combined lysis of nuclei and cells; salt extraction

Luo *et al.* (2014) - lysis of cells only; salt extraction

3.2.1 Protease suppression

During protein work, researchers have universally kept temperatures low, typically 4°C. Proteases and other enzymes have been suppressed by adding one or more reagents such as sodium bisulphite, ClHgPhSO₃, chelating agents, PMSF, urea, and phosphatase inhibitors.

3.2.2 Lysis of cells and nuclei

Lysis methods applied either individually, or in combination, have been: mechanical means (shear or centrifugal force), the application of non-ionic detergent and/or chelating agent and/or urea to weaken the cellular membranes, and osmosis. Cells and nuclei have been lysed in one step, but more commonly cells have been lysed and cytosolic components removed as a supernatant by centrifugation, prior to lysing the nuclei. This two-step process avoids the need to separate cytosolic and nuclear proteins, and removes enzymes which may otherwise modify the nuclear proteins.

3.2.3 Dissociation of proteins from nuclear structures, especially from chromatin

Strongly-ionised solutions (either acid or salt) have been typically used to compete with the electrostatic charges holding the nuclear structures together. Acid has the disadvantage that some histone PTMs are acid labile (Chen *et al.*, 1974). However, acid competes effectively with the acidic DNA for histones. Once pH is below the pI of DNA (~5), the DNA is no longer negatively charged, removing ionic bonding to histones (although other types of bond remain). In sufficiently acidic conditions the DNA phosphodiester bonding fails, producing fragmented DNA which is easily separated from the proteins by low speed centrifugation. Acid is considered more effective than salt in suppressing protein-modifying enzymes (Shechter *et al.*, 2007). While less damaging to PTMs, salt leaves the DNA intact and difficult to remove (see below).

3.2.4 Separation of the dissociated proteins from other nuclear components

Following nuclei lysis, centrifugation has been used to sediment DNA, RNA, and other nuclear components, leaving nuclear proteins in the supernatant. However, with salt extraction, the intact DNA forms a gel which has required ultracentrifugation for removal (Goodwin *et al.*, 1975). This problem was avoided by Rodriguez-Collazo *et al.* (2009) who used sonication combined with 8M urea to fragment the DNA; however, this approach denatures the proteins. As an alternative to nuclei lysis, soluble mobile proteins have been washed out through the nuclear membrane pores, leaving the intact nuclei to be removed by centrifugation (Gurley *et al.*, 1983).

3.2.5 Purification of the separated proteins

Nuclear proteins, in particular histones and HMG proteins, have often been purified by precipitation in a solvent (TCA, acetone, or ethanol).

3.2.6 Separation of the nuclear proteins, one from another

The first-choice method for separating proteins has been chromatography, in most cases using a cation-exchange medium. Hake *et al.* (2006) and others have used reversed-phase chromatography so that only hydrophobic molecules bind to the column. Rodriguez-Collazo *et al.* (2009) used a TPS column which provides covalent S-S bonding with the sulphur in histone H3 cysteine residues (no other histones contain cysteine).

3.2.7 Protein separation and assay using gel electrophoresis

Gel electrophoresis has been widely used to separate proteins and to estimate their molecular weights. Murray (1966) used electrophoresis applied to a starch gel to examine the proteins in fractions obtained from cation-exchange chromatography of extracts from calf thymus chromatin. He compared the band pattern with earlier patterns for histones. Later researchers have used electrophoresis applied to acid urea (AU), triton acid urea (TAU), and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels. Panyim and Chalkley (1968) used AU gels to separate histones according to size and charge. AU gels are still used since some PTMs, such as acetylation and phosphorylation, introduce additional charges which are reflected in migration speed on the gel. TAU gels have been used to provide an additional separation mechanism since migration speed is influenced by the binding of the detergent Triton X-100 to hydrophobic regions of the proteins (Shechter *et al.*, 2007). However, SDS-PAGE as described by Laemmli (1970) has been most popular, usually on its own, but in some cases in combination with separation according to protein isoelectric point, to provide a two-dimensional gel (2D gel). SDS-PAGE migration speed is an indication of protein molecular weight. The mechanism of SDS-PAGE is summarised in Section 3.7.

3.2.8 Determination of protein sequence

Early work (Murray, 1966; Goodwin *et al.*, 1975; Sanders, 1977) included amino acid analysis of histones, in which the histone was hydrolysed, and the amount of each amino acid was determined, but not the sequence of the individual amino acid residues. Edman degradation sequencing became available, via automatic sequencers. This involved identifying individual amino acid residues as they were removed one at a time from the protein N-terminal. More recent studies have used mass spectrometry to determine protein residue sequence (Leclercq *et al.*, 2000; Hake *et al.*, 2006) – see Section 3.8.3.

3.3 Our approach to isolating and assaying nuclear proteins

With some variations as specified in later chapters, Figure 3.1 summarises our approach to isolating nuclear proteins; this produces specific protein groups. Proteins were further separated and assayed by chromatography and SDS-PAGE. MS provided by Professor Dickman (Sheffield University) was used to identify/confirm proteins.

Protein isolation methods were selected:

- (i) To minimise enzymatic degradation of cells, nuclei, and proteins - the chicken erythrocytes were processed without delay, buffers were chilled (typically 4 – 6°C) and included a protease inhibitor where practicable, final products were stored at -20°C or below, and a freezing buffer was applied as necessary.
- (ii) To be sufficiently mild to minimise damage to the wild-type protein structure (with the exception of the acid extraction methods in Chapter 4).
- (iii) To provide stable conditions similar to those in an intact nucleus - buffers were mainly near to biological pH and ionic strength and included phosphate charges to mimic those due to DNA.

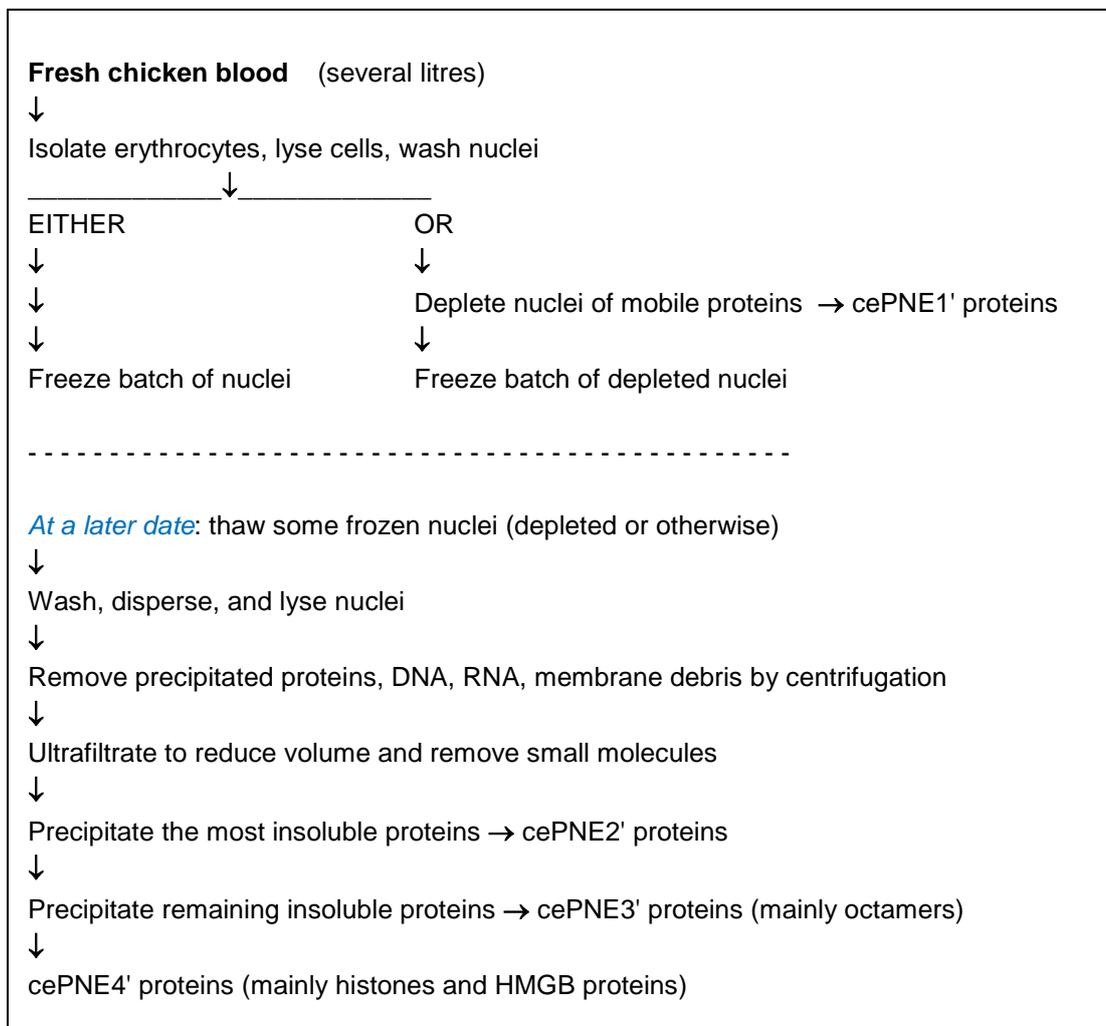


Figure 3.1 Our approach to isolating nuclear proteins

Note: the above approach was simplified for the production of a Histone Standard and histone octamers (Chapter 4).

3.4 Buffers, protease inhibitors, centrifugation, filtration, dialysis, freeze drying, and general laboratory practice

3.4.1 Buffers

Buffers were used for cell and nuclei lysis, pH control, and as a carrier and eluent during chromatography. Buffers were mainly made up from stock solutions, previously prepared from the primary reagents. Solutions were made up using high-quality de-ionised water (dH₂O) with conductivity ~1 μ S/cm, using an ISG Deioniser (Figure 3.2) or equivalent. Buffers were chilled before use. To avoid blocking the chromatography stationary phase, buffers and samples for chromatography were filtered (0.45micron) using a water vacuum pump (Figure 3.3). This method also de-gassed the buffers.

3.4.2 Protease inhibitors

Where possible, 2.5mM benzamidine hydrochloride was included in buffers because it inhibits trypsin-like enzymes and serine proteases (Sigma-Aldrich data sheet). It was made up in solution immediately prior to use, since it deteriorates in solution. It has the following limitations, since it strongly absorbs UV radiation at the protein monitoring wavelength of 280nm:

- (i) Benzamidine hydrochloride can be used for anion-exchange chromatography because it does not bind to the positive stationary phase, but all mobile phases must have an identical benzamidine hydrochloride concentration to maintain a fixed UV absorbance offset.
- (ii) It is unsuitable for cation-exchange chromatography because it binds to the negative stationary phase and is subsequently released in the eluent together with the sample proteins, creating a large UV absorbance peak which tends to mask protein absorbance. The binding is presumably caused by the dissociation of the chloride ion, leaving a positively charged benzamidine ring which initially binds to the negative stationary phase until eluted by an increasing ion concentration.

The above problem was solved by using benzamidine in place of benzamidine hydrochloride for the cation-exchange buffers, since benzamidine does not bind to the column. Benzamidine has been previously used to suppress enzymic degradation of proteins (Claeys and Collen, 1978).

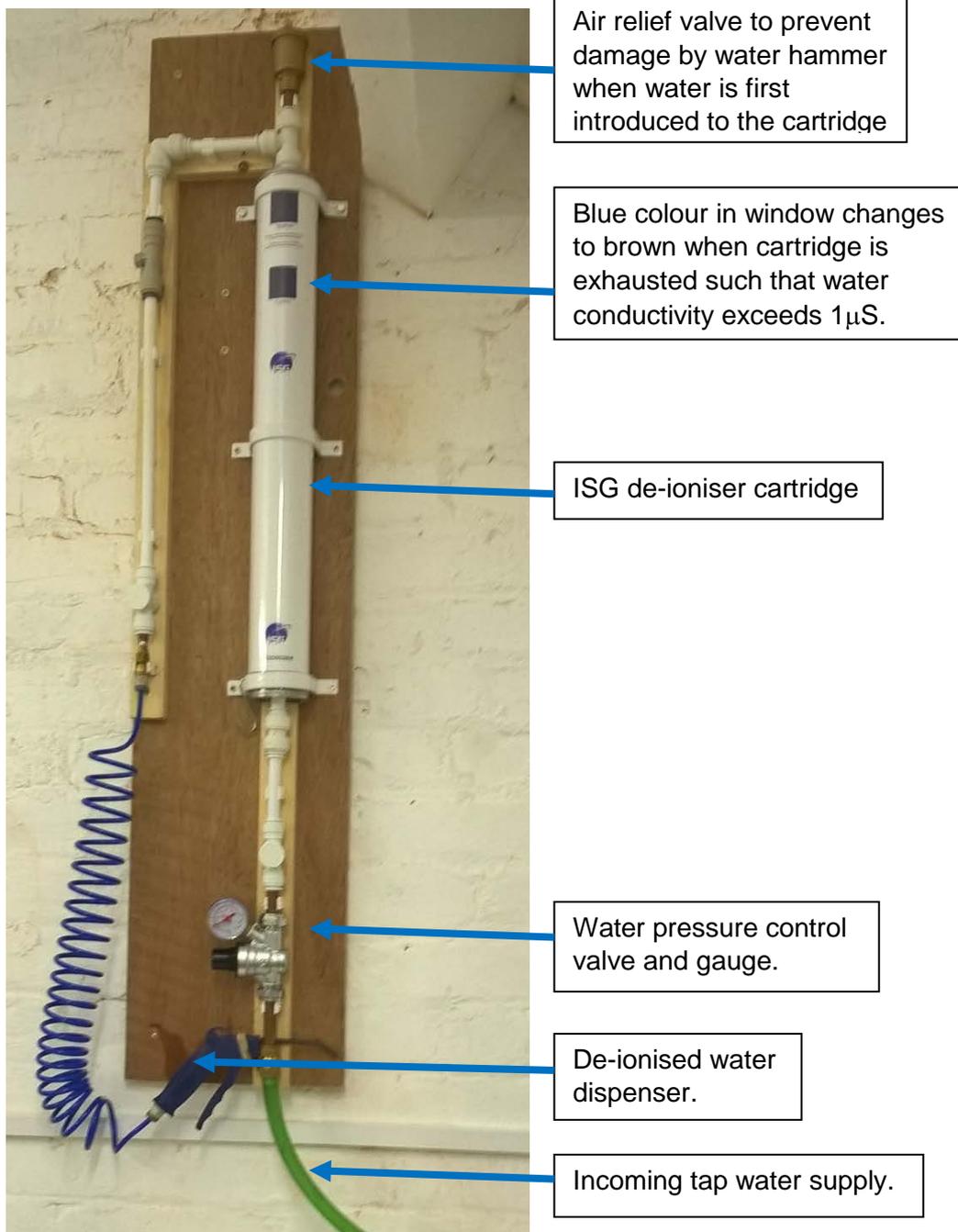


Figure 3.2 Equipment for supplying de-ionised water

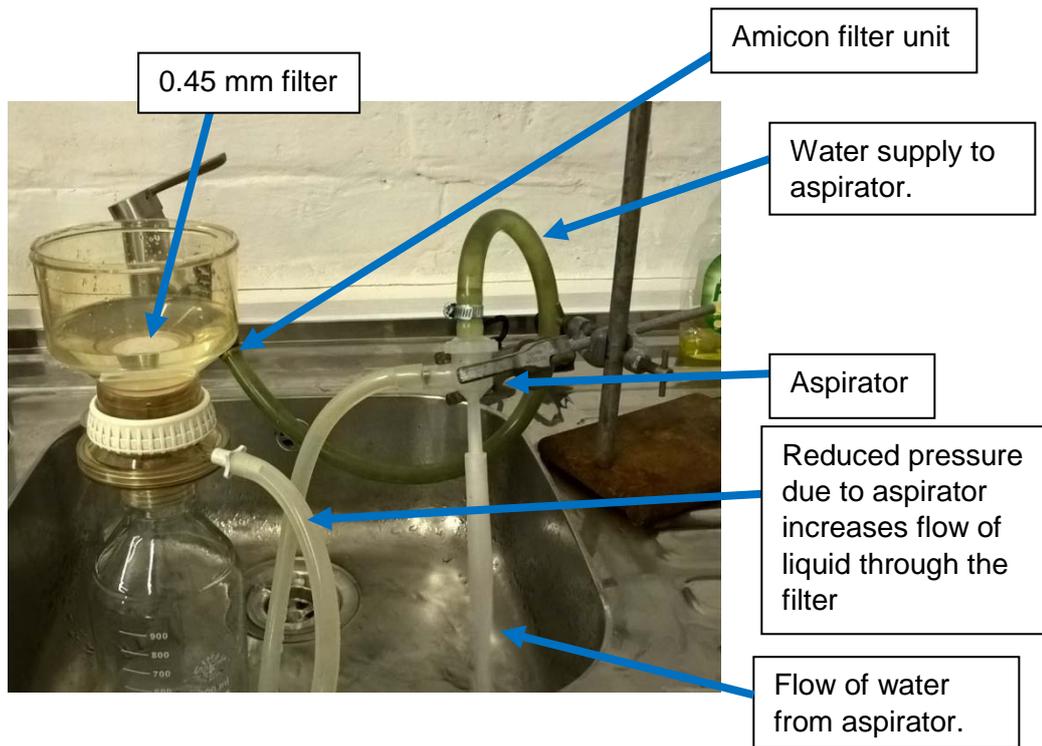


Figure 3.3 Buffer filtering

3.4.3 Centrifugation

In a mixture of materials, more dense material will sediment due to the force imposed by gravitational acceleration g acting on the mass M of material. The rate of sedimentation depends on the difference in densities of the materials, and on the size of particles – large particles sediment more quickly since they are subject to less drag than small particles.

Centrifugation can impose on the sample a much greater acceleration a_c than gravity. This allows sedimentation to take place in much shorter times. The acceleration depends on the effective radius of the centrifuge rotor r (that is, the distance between the centre of rotation and the contents of the centrifuge tube or bucket), and on how quickly the rotor rotates R . With acceleration a_c in units of g , r in metres, and R in rpm (revolutions per minute): $a_c = 0.00112r.R^2 g$

Centrifugation was used to pellet cells, nuclei, and precipitates, and to remove debris. Centrifuge tubes/bottles were carefully balanced to avoid destroying the centrifuge. A balance was struck between minimising sedimentation times, and centrifuge speeds (in association with rotor radii), which risked accelerations (g-forces) which could cause premature lysis of cells or nuclei. Centrifugation was also used to pool small samples into the bottom of Eppendorf tubes.

3.4.4 Filtration

Filtration was used to remove debris and some precipitates. Filter pore size and filter material were chosen to separate out the required particle size, while minimising filter blockage and hence filtration time. 0.45micron nylon filters were found satisfactory for most applications. Most filtration was conducted using a water vacuum pump (Figure 3.3). Syringe filters were initially used when injecting a sample into a chromatography column, but later it was found more efficient to pre-filter the sample using the water vacuum pump.

Ultrafiltration driven by nitrogen gas pressure (Figure 3.4) was used to concentrate proteins using low protein-binding modified polyethersulfone filters with 10kDa molecular weight cut off (MWCO)(the proteins of interest in the study all have MW >10 kDa). Small proteins and protein fragments below 10kDa passed through filter and were discarded. Both water and buffer reagent molecules passed through the filter, so constant buffer molarities were maintained while protein concentration increased.

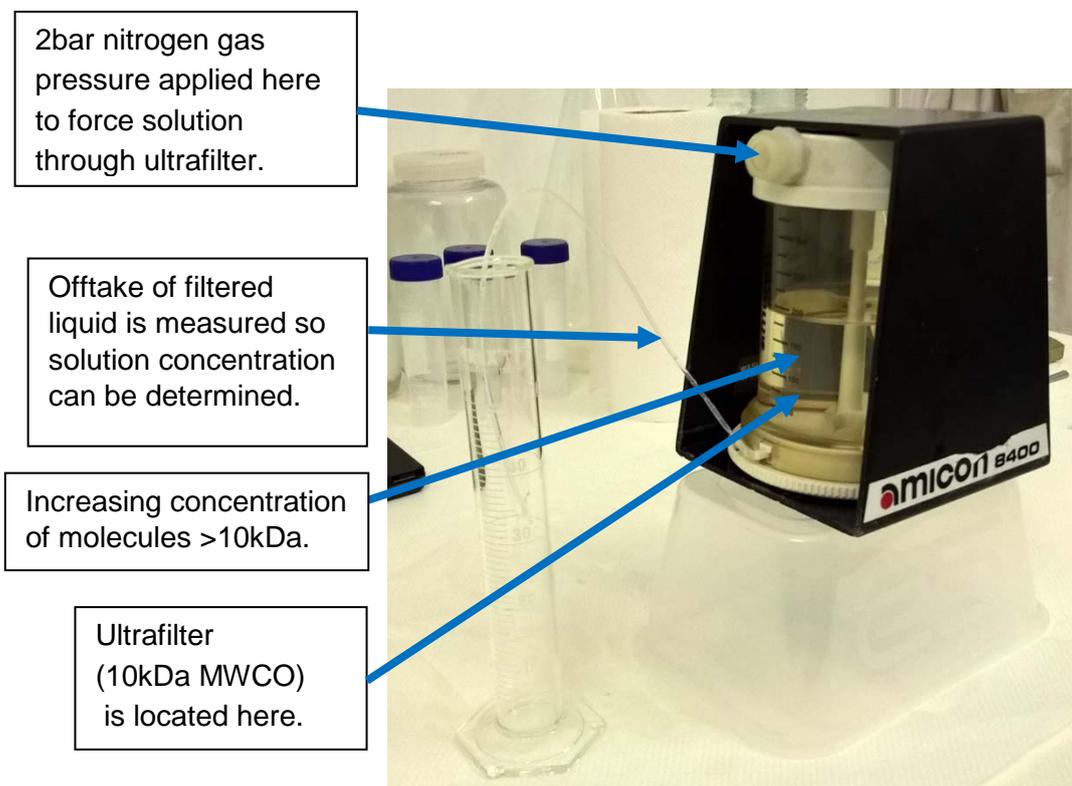


Figure 3.4 Ultrafiltration arrangement

3.4.5 Dialysis

Dialysis was used to change buffers in a sample, say from Buffer A to Buffer B. Buffer A and the sample protein are placed in the dialysis bag, and the bag is placed in Buffer B in a beaker (Figure 3.5). The bag has pores which are small enough to retain the protein entirely inside the bag, but big enough to allow Reagent A and Reagent B (from Buffer A and Buffer B respectively) to diffuse through. The pores are bigger than water molecules, but surface tension retains water in the bag when it is out of the beaker.

At the start of dialysis, Reagent A is confined to volume V in the dialysis bag, and Reagent B is confined to volume V in the beaker outside the bag. At the end of dialysis, the concentration of Reagent A is much reduced, whereas Reagent B is now present in the bag at close to its original concentration. From Figure 3.5, the equations which apply are:

$$\text{Change in concentration of Reagent A is} = v/(v+V) \quad (1)$$

$$\text{Change in concentration of Reagent B is} = V/(v+V) \quad (2)$$

For example, if v is 100mL and V is 1000mL, the concentration of Reagent A in the dialysis bag is reduced by $100/(100 + 1000)$, ie to 9% of its original concentration. Meanwhile the concentration of Reagent B in the dialysis bag has increased from zero to $1000/(100 + 1000)$ of the original Reagent B concentration (90%). This process can be repeated. For example, a second, identical dialysis cycle will reduce the concentration of Reagent A in the bag to 9% of 9% = 0.81% concentration and will further increase Reagent B concentration.

The above equations are approximations, for the following reasons:

- (i) The equations are based on complete dialysis, which would take infinite time. In practice, dialysis for a few hours each time using a magnetic stirrer, together with 2 or 3 cycles of dialysis, was considered sufficient for these studies. More stringent dialysis was required when preparing samples for freeze drying.
- (ii) The equations assume that particle concentrations in bag and beaker are the same, but they will differ due to the retained protein molecules and due to differences in buffer concentration. When they differ, osmotic pressure causes the volume of liquid in the dialysis bag to change. This alters the final protein concentrations, and efficiency of dialysis.

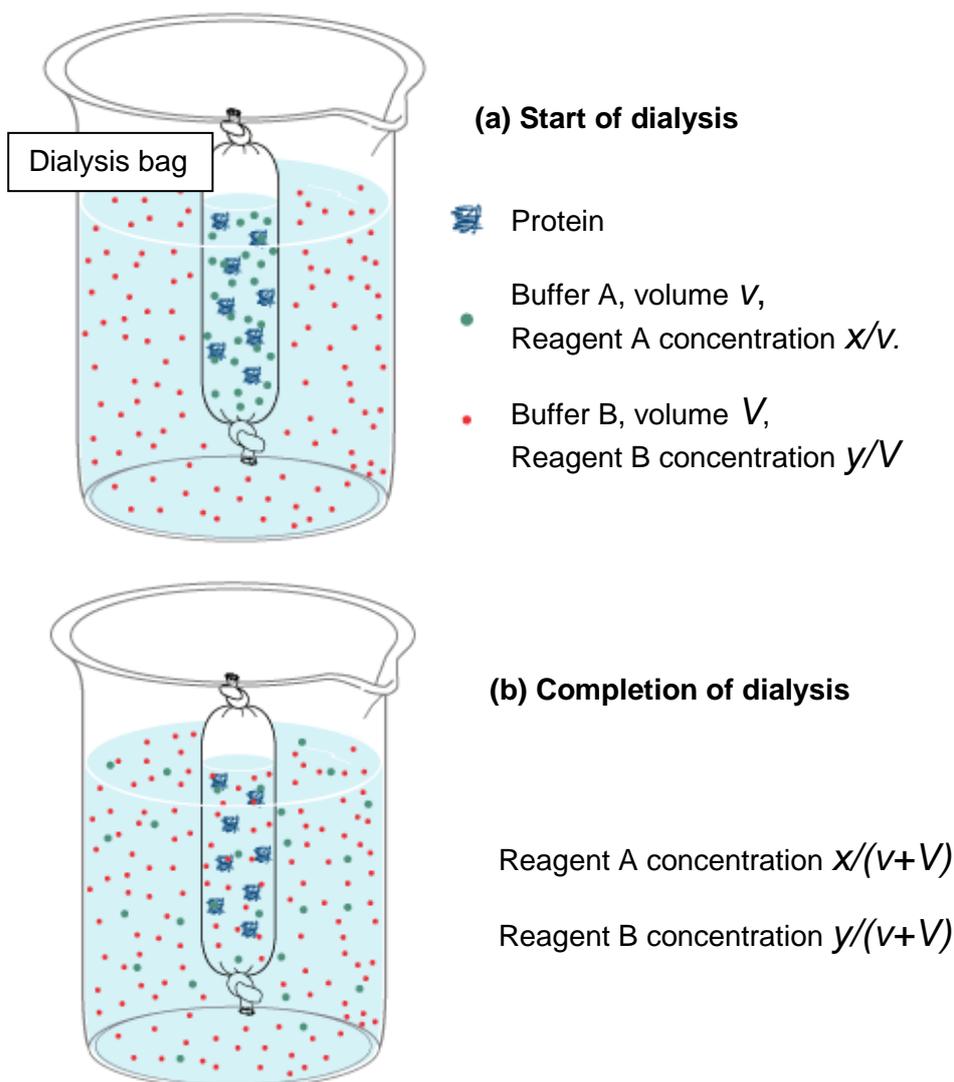


Figure 3.5 Effect of dialysis on buffer concentrations

Note: this is a simplification, which does not take account of osmosis.

Graphics credit: Potcherboy at English Wikipedia 11/05/2012, licensed under CCA 3.0.

(a) At start of dialysis, Reagent A mass X is confined to volume V in the dialysis bag, and the new Reagent B mass Y is confined to volume V in the beaker outside the bag.

(b) At the end of a complete dialysis, the protein remains in dialysis bag (dialysis tubing with a low molecular weight cut-off is used to avoid loss of histones and other small proteins), but Reagent A is evenly diffused throughout the total volume of the bag and liquid in the beaker, $v+V$, so:

$$\text{Change in concentration of Reagent A is } (x/(v+V))/(x/v) = v/(v+V) \quad (1)$$

$$\text{Similarly, change in concentration of Reagent B is } (y/(v+V))/(y/V) = V/(v+V) \quad (2)$$

3.4.6 Freeze drying (lyophilisation)

Freeze drying was used earlier in the study to concentrate proteins for SDS-PAGE (later, it was found more efficient to concentrate the samples by ultrafiltration instead). Freeze drying involved exposing the frozen sample to air at sub-zero temperature and low pressure (millibar). The low air pressure caused water to sublime away, while low air temperature maintained the sample liquid as ice so that water did not boil and remove the sample. After freeze drying for hours or days, depending on sample size and air pressure, all of the water was removed leaving the sample as a dry powder. Prior to freeze drying, it was necessary to reduce sample buffer concentrations to very low levels by several cycles of dialysis, since freeze drying increases buffer ionic strength which can lead to smeared and distorted SDS-PAGE bands.

3.4.7 General laboratory practice

Prudent laboratory practice (NRC, 2011) was generally applied, included attention to the following:

- Pre-planning of each experiment.
- Thorough cleaning of glassware etc. in hot water, and with detergent if protein present. Final rinse in deionised water immediately prior to use.
- Review of (Material) Safety Data Sheets for all reagents.
- Workplace safety, including safety of electrical systems, and nitrogen supplies.
- Lab coat and gloves: always worn.
- Glasses and breathing protection: worn as necessary.
- Clear work area before and after the experiment.
- Disposal routes for used materials.
- Security of materials.

3.5 Preparation of erythrocyte nuclei, and isolation of protein sets

3.5.1 Preparation of erythrocyte nuclei

Johnson and Swarbrick (Lancashire) process 8 week old chickens for the consumer market. Fresh blood from this process was collected, and erythrocyte nuclei were obtained by lysing the cytosolic membrane using a combination of osmotic pressure and a lysis buffer containing a non-denaturing, non-ionic detergent (Triton X-100) together with $MgCl_2$, KCl, Tris/HCl, and benzamidine hydrochloride (to inhibit proteases). Throughout the procedure, cell/nuclei clumping was prevented by a combination of stirring and gentle homogenisation (passing the cells through a 1 mm diameter syringe needle or the 2 mm diameter nozzle of a plastic Pasteur pipette – Figure 3.6). Triton X-100 dissolves and disrupts the cytoplasmic membrane (Luo *et al.*, 2014). In mammalian cells, magnesium ions have a concentration of 0.8mM (Lodish *et al.*, 2007). 3.5 mM $MgCl_2$ was included in the extraction buffer to provide an Mg^{+2} environment which would otherwise have been provided by the cytosol; the Mg^{+2} ions stabilize the nuclear membrane by complexing with phospholipids, reducing membrane fluidity and lowering membrane permeability (Durlach, 1985). 10mM KCl was used to assist nuclei stability (Rodriquez-Collazo *et al.* (2009) - this compares with the ~12mM NaCl concentration found in the cytosol of mammalian cells (Lodish *et al.*, 2008); also it avoided compromising the hypotonic status of the lysis buffer. Tris/HCl was used as a buffer to provide the required pH.



Figure 3.6 Devices used to mildly homogenise nuclei to prevent clumping

Top: plastic Pasteur pipette with 2 mm diameter nozzle. Middle: syringe with 1 mm diameter needle to transfer supernatant, Lower: syringe with 0.8 mm diameter needle to disperse nuclei.

3.5.2 Washing proteins out of the intact nuclei (optional, not included in all experiments)

In this optional step (Figure 3.1), the nuclei were mixed with an extraction buffer so that the soluble, loosely attached proteins were washed out through the existing pores in the nuclear membrane, leaving the nuclei intact, but partially depleted. A process of mixing the nuclei with extraction buffer and centrifuging was repeated three times, usually the supernatant was collected from the second and third extractions. The proteins in the supernatants were designated the cePNE1' group (Chicken Erythrocyte Nuclear Extract 1). Throughout the procedure, nuclei clumping was prevented by a combination of stirring and gentle homogenisation as per Section 3.5.1. The extraction buffer contained a salt (typically 90mM KCl) which provided ions to compete with the electrostatic charges holding the nuclear structures together. The buffer also contained a mixture of monobasic and dibasic potassium phosphate (KH_2PO_4 and K_2HPO_4) to provide the desired pH (Figure 3.7), and to provide an environment similar to DNA with its phosphate groups. The buffer also included MgCl_2 to maintain nuclei membrane integrity.

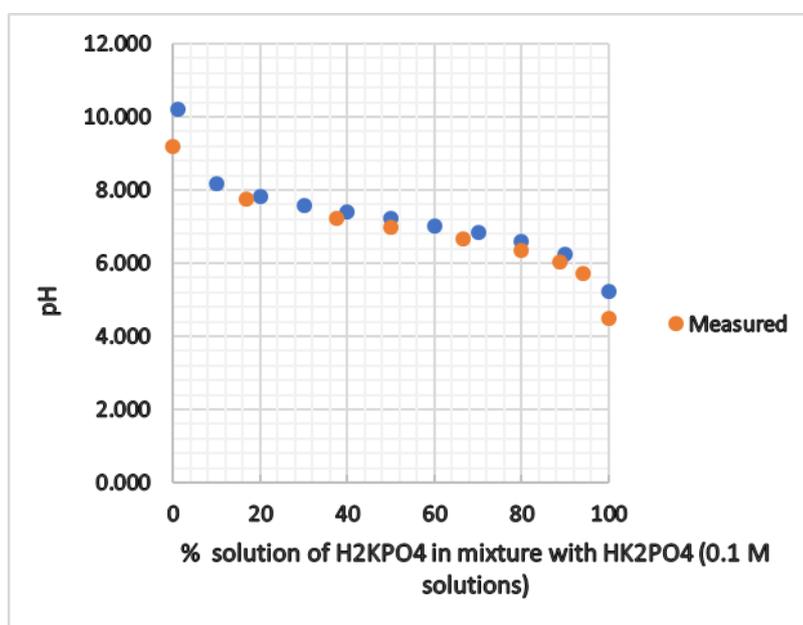


Figure 3.7 pH of different mixtures of monobasic and dibasic potassium phosphate

Orange - single measurements made by H Smallman March 2011.

Blue – theoretical values.

3.5.3 Lysis of nuclei and primary separation of nuclear proteins

High salt conditions, were used to lyse the nuclei and dissociate nuclear proteins from the nuclear structures. pH was controlled by a mixture of monobasic and dibasic potassium phosphate. The high concentration of ionic particles in the nuclei lysis buffer collapsed the nuclei by osmosis, and the absence of magnesium ions (normally provided by the now absent cytosol) compromised the nuclear membrane. All nuclei were simultaneously exposed to the same high salt lysis condition, to ensure the histone octamer maintained its structure despite the removal of DNA from the nucleosome (Zhuang, 2011). This was achieved by (i) preventing nuclei clumping by a combination of stirring and gentle homogenisation as per Section 3.5.1, and (ii) layering nuclei on top of an appropriate volume of lysis buffer in a container which was then shaken vigorously

The proteins obtained from nuclei lysis were separated into three groups – designated cePNE2', cePNE3', and cePNE4', on the basis of solubility, by applying varying amounts of a Precipitation Buffer. Phosphates were included in the buffer to mimic the environment provided by DNA. Together with the phosphates, KCl was also included to ensure the core histones were retained as complete octamers (Wood *et. al.*, 2005).

3.6 Protein separation and assay using chromatography

3.6.1 Principles of chromatography

In the present work, fast protein liquid chromatography (FPLC) was used to separate and purify nuclear molecules according to differences in molecular size and charge. GE Healthcare (2006), (2007), (2014), and (2016) explain the principles and methods used in chromatography which are based on protein properties, these can include size, charge, hydrophobicity, and various affinities. In FPLC, the sample in solution (mobile phase) is passed through a column containing a bed of particles. The stationary phase comprises either the particles themselves or specific molecules covalently attached to the particles. Some or all of the sample molecules initially adhere to the stationary phase and are either released after a delay (as in gel filtration), or are released when a solution of different composition is passed down the column.

3.6.2 Chromatography equipment

The chromatography equipment was an AKTA Purifier 10 system supplied by Amersham Biosciences (subsequently taken over by GE Healthcare). The system included Unicorn Version 5 software running over the Windows XP operating system on a PC.

Figure 3.8 shows a schematic of the AKTA equipment. Table 3.2 lists our AKTA equipment parts and user documentation. Pumps were under software control, allowing the composition of the mobile phase to be varied over time by flowing (and mixing) differing volumes of buffer solutions A and B. Material exiting the column was automatically collected in fixed amounts (called fractions) into numbered test tubes in a fraction collector under software control. A further key component was the computer controlled injection valve which transferred sample from a pre-charged sample loop into the mobile phase at the start or a run. The Unicorn software recorded the buffer flow rates (determined from the pump flow rates), mobile phase UV absorption, conductivity, pH, and fraction collection position relative to flow. UV absorption (set at 280nm) was a particularly useful measure of when and how much of a protein was passing through the column. For early work, the sample loop was loaded by syringe at a sample injection point on the sample control valve. This was a difficult operation. For subsequent work the sample loop was loaded with sample offline, and then installed in the AKTA system.

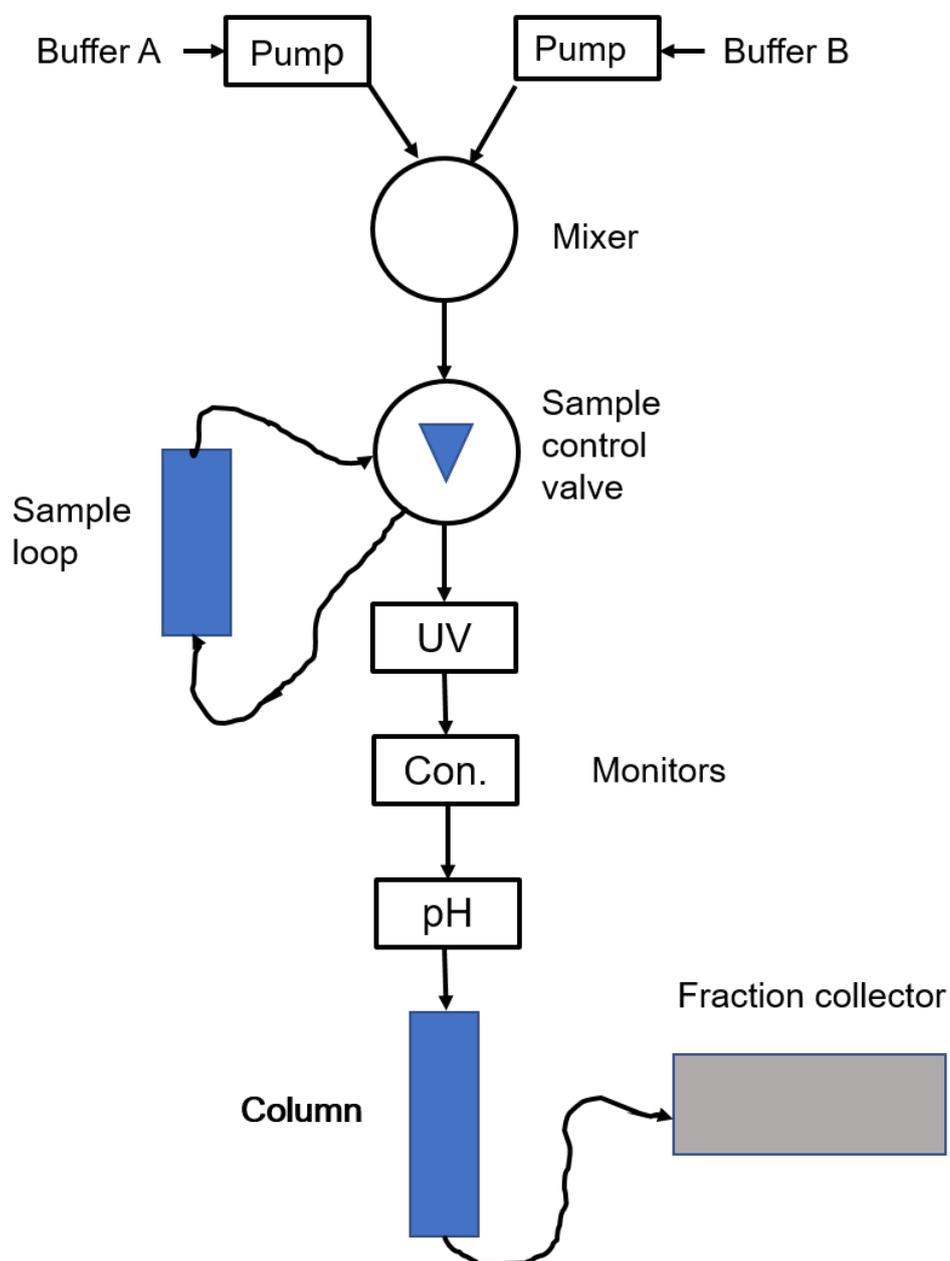


Figure 3.8 Schematic of our AKTA Purifier 10 chromatography system

Items not shown include computer control arrangement, waste solution routes.

Con. Conductivity monitor

Table 3.2 Our AKTA equipment and user documentation

Identifier	Function	Supplier
AKTA Purifier 10 including: P-900 UV-900 pH/C-900 M-925 PV-907 (INV-907) PV-908 Box 900	FPLC system, including - System pump unit - UV monitor - pH, and conductivity monitor - Mixer - Motorised 7-port injection valve - Motorised 8-port valve for selection of columns, samples, or buffers. - Storage box	AB
CU-950	Control unit for fraction collector	AB
Frac-950	Fraction collector	AB
Super-loop 50ml AKTA design	50 mL sample loop	AB
Super-loop 150ml 18-1023-85	150 mL sample loop	GE
Accessory kit	Filters, tubing, cables, connectors	AB
Unicorn 5.01 XP upgrade	FPLC control software	AB
PC with Windows XP	PC running Unicorn software	Various
Hiprep 16/60 Sephacryl S-100 HR	Gel filtration column	GE
Hiprep 16/60 Sephacryl S-300 HR	Gel filtration column	GE
HiTrap DEAE FF 5mL	Anion-exchange column	GE
HiTrap SP Sepharose FF 5mL	Cation-exchange column	GE
UM Box 18-1141-00	Set of User Manuals (2004), inc. Unicorn getting started; instructions for Purifier 10 components.	AB
Doc. 18-1140-44	UM-AKTApurifier installation guide	AB
Doc. 18-1140-79 2003	AKTApurifier Making Your First Run	AB
Doc. 03-0014-90/91 AB 2004	Unicorn 5.0 User Reference Manual	AB
Doc 28-4079-54 AA 2006	AKTApurifier Getting Started	GE
Doc 28407953 AC 2008	AKTApurifier User Guide	GE
Doc 28-9579-57 AB 2013	AKTApurifier Operating Instructions	GE

Supplier key: AB - Amersham Biosciences; GE - GE Healthcare Bio-Sciences.

3.6.3 *Gel filtration chromatography*

GE Healthcare (2014) describes gel-filtration chromatography, using the more modern term, size exclusion chromatography (SEC). The stationary phase consists of a bed of porous particles which remain in the column. The mobile phases comprise (i) buffer solution to equilibrate the column, and (ii) the sample in the same buffer solution. When the bed is equilibrated the buffer fills the pores. When the sample is injected the smallest molecules most easily penetrate the pores, and so they are most delayed in flowing through the bed. Larger molecules are less delayed since they penetrate the pores less easily. Thus, different groups of proteins pass through the column at different times in accordance with molecule size, leading to a series of peaks in UV absorption. Proteins which are bigger than the pore size are not delayed at all and pass through the column first.

3.6.4 *Ion-exchange chromatography*

GE Healthcare (2016) describes ion-exchange chromatography, referring to it as IEX. The stationary phase consists of a bed of permanently charged particles which remain in the column; these particles in a cation-exchange column are negatively charged; those in an anion-exchange column are positively charged. The mobile phases comprise (i) a buffer solution to equilibrate the column, (ii) the sample in the same buffer solution, and (iii) a different buffer solution to change conditions so that sample molecules are selectively released from the column. When the bed is being equilibrated the stationary phase charged particles attract, and pull from solution, the opposite charged molecules in the buffer. Similarly, when the sample is injected, the stationary phase charged particles attract, and pull from solution, the opposite charged sample molecules. Provided the buffer ionic strength is sufficiently low, most of the charged sample molecules (of the appropriate polarity) are attracted and pulled from solution - their higher ionic strength, and in some case greater charge, enables them to dynamically compete with buffer molecules for attachment to the charges on the stationary phase. A different buffer solution (the eluent) is then passed through the column to release (elute) the sample molecules attached to the stationary phase particles. Increased eluent ionic strength can be used to release the sample from the column, or a change in pH, or both. As shown by the example in Table 3.3, change in pH markedly changes the net charge on a protein, so changes its binding strength to the column.

Table 3.3 Example of change in protein charge with pH

Protein: HMGB3_Chick (Uniprot P40618)

pH	Protein charge
4.00	+35.3
5.00	+12.0
6.00	+3.8
7.00	+2.1
8.00	+0.4
8.12 (pI)	0.0
9.00	-5.4
10.00	-25

Protein charge was calculated using the Scripps Protein Calculator (2016). The calculated values assume all residues have pKa values that are equivalent to the isolated residues. For a folded protein this is not valid. However, the calculated values illustrate the principle that protein charge depends on pH.

3.6.5 Chromatography as an assay tool

In addition to protein separation, chromatography assists with protein identification and quantification. Ion-exchange chromatograms reflect the charge on proteins, and gel filtration chromatograms reflect the sizes of proteins and complexes.

The UV absorption of the eluted proteins is proportional to their concentrations, and the areas of the peaks on the chromatograms as plotted by the AKTA UV monitor provide an indication of the amount of each protein on the column. Proteins absorb UV radiation at 280nm, mainly depending on the prevalence of amino acid residues with aromatic rings (such as tryptophan), so UV absorption at 280nm provides a measure of protein concentration, provided there are no other molecules which significantly absorb UV at 280nm. Nucleic acid contamination of the protein must be avoided (although nucleic acid absorption peaks at 260nm, it also strongly absorbs at 280nm). Gill and Hippel (1989) attribute most protein absorption of UV radiation at 280nm to tyrosine, tryptophan, and cystine residues (cystine is the oxidised form of cysteine found when cysteines form a disulphide bond to fold a protein). Phenylalanine residues also absorb 280nm UV radiation.

3.7 Protein separation and assay using SDS-PAGE

An electric field can be used to pull charged molecules through a porous medium (Reed *et al.*, 2003). This is electrophoresis applied to a molecular sieve. Molecules of dissimilar size and/or charge can be separated since they move at different speeds through the media. Although starch and agarose have been used, polymerised acrylamide (with bis-acrylamide cross-links) is the most popular porous medium - it is chemically inert, the pore size is uniform, and pore size can be selected by varying the concentration of acrylamide (Berg *et al.*, 2015), (Reed *et al.*, 2003). Particular properties can be produced by combining the polymerised acrylamide with different reagents (Section 3.2.7). The porous medium is usually cast as a slab and is referred to as a gel. Pore size can be varied by changing the acrylamide concentration and is selected to match the sizes of molecules being separated, smaller proteins require smaller pore sizes for effective separation.

For the present work, SDS polyacrylamide gel electrophoresis (SDS-PAGE) similar to that developed by Laemmli (1970) was used extensively to separate and estimate the molecular weights (MWs) of proteins. Our SDS-PAGE equipment configuration included a vertically located gel slab, as illustrated in "A guide to polyacrylamide gel electrophoresis and detection" at <http://www.bio-rad.com/en-uk/applications-technologies/protein-electrophoresis-methods>. Our gels contained 20% acrylamide. SDS was added to the protein, to denature it and to add a constant negative charge per unit length of each resulting peptide. Samples were placed into wells at the top of the gel. A power supply was connected to wires in buffer solutions at the top and bottom of the gel, to provide the electrical field. The negatively charged peptides migrated towards the positive electrode at the bottom of the gel. A key feature was the need to ensure uniformity of the electrical field across the gel, to provide clearly defined horizontal bands of peptides beneath each well. This meant running the gels sufficiently slowly to avoid heat distortion of the gel. Peptide positions on the gel were highlighted by staining with Coomassie Blue. Our detailed SDS-PAGE protocol is given in Appendix 3.1.

Berg *et al.* (2015) present results from Weber and Osborn (1975) which show that the speed at which a typical protein migrates through the gel is approximately inversely proportional to the logarithm of the protein MW (however, HMG proteins run at a different speeds than expected from their MW). To assist with protein identification and MW measurement, we have used a set of histone proteins as MW markers. These were run on the SDS-PAGE gels in adjoining lanes to the proteins of interest (Chapter 4).

3.8 Protein identification using MS

3.8.1 Principles of using MS to identify proteins

A mass spectrometer comprises an ion source, a mass analyser, and an ion detector (Whitford, 2005). The principle is that electric and magnetic fields change the trajectory and time of flight of a charged particle depending on its charge and mass. This enables the m/Q ratio of each particle of mass m charge Q to be determined with great accuracy. Furthermore, Q is the sum of a few equal sized packets of charge, due to one or more excess electrons or protons. This enables the mass of a particle to be determined. Before MS is applied to proteins, the protein is fragmented into peptides of a few residues each, typically using a protease such as trypsin. Once the mass of a fragment has been accurately determined, the residues in that fragment can be identified because each residue (and each PTM type) has a fixed, unique mass, except for leucine and isoleucine which have equal mass. Furthermore, since there are many molecules in a sample, fragments of different lengths are produced, and fragments can be related together because the mass of a larger peptide is the sum of the masses of its smaller fragments. This enables the sequences of peptides in a fragment to be determined. Finally, the fragment sequences can be compared with known sequences of proteins in a database, enabling identification of the sample proteins (Gygi and Aebersold, 2000).

3.8.2 MS experimental work

We obtained protein samples for MS either by excising protein bands from our SDS-PAGE gels (gel pieces typically 1.5mm x 10mm), or by chromatography alone. We have on occasion ourselves prepared the proteins within a gel piece for MS using a Sheffield University protocol, but in most recent cases just the gel pieces were supplied to Professor Dickman's group at Sheffield University, who then applied the protocol. This "in-gel" protocol denatured the protein and fragmented the peptides with trypsin. The peptide fragments were then recovered in solution from the gel piece, dried in a vacuum concentrator, and sent for MS. Alternatively, our chromatography fractions were supplied directly to Sheffield University, where the fractions were prepared for MS using an "in-solution" protocol, similar to the "in-gel" protocol. The "in-gel" protocol is given in Appendix 3.2.

Protein identities were obtained by MS analysis by Professor Mark Dickman's group at Sheffield University with the aid of an Amazon ETD (Bruker Daltonics) through an online nano liquid chromatography system (Ultimate 3000 RSLC, Dionex). Peptides were separated using a 5 mm x 300 μ m trapping column and 75 μ m x 15cm analytical PepMap C18 reverse-phase column. Tryptic peptide elution was achieved through a 55 min linear gradient from 94% solvent A (0.1% (v/v) formic acid) to 40% solvent B [0.1% (v/v) formic acid, 80% (v/v) acetonitrile] at a flow rate of 300 nL/min. Profile scans with mass spectra ionization (MS1) (m/z 300-1800) were done in enhanced positive mode followed up by 4 collision induced dissociation (CID) fragmentation ultra-scan mode (m/z 100-1800). Fragmentation involved loading the trap to a target value of 200,000 m/z with a maximum accumulation time of 200 minutes. The width of precursor isolation was set to 4.0 and singly charged precursors were excluded. Online acquisition of mass spectra was done using "Profile MS" with automatic dependent MS/MS scans. MS signals for each band or sample were converted to mascot generic files (mgf) using the software Bruker DataAnalysis 4.2. MS converted mgfs were used to query standard databases (NCBI nr or Swiss-Prot) with bony vertebrates selected as taxonomy. Queries were submitted using Mascot Server v.2.2, 01 (Matrix Science). The following parameters were typically used: peptide mass tolerance = \pm 0.6 Da, fragment MS/MS tolerance = \pm 0.6 Da, peptide charge = 2+. Electrospray ionisation trap (ESI-TRAP) was selected as instrument. Tryptic enzyme specificity with up to 2 missed cleavages was applied to all searches. Oxidized methionine were selected as variable modifications respectively for tryptic digest. Mascot calculated peptide ion score cut-off of \geq 20 was used to filter and protein identification was usually based on a minimum of two unique peptides. Peptide matches above homology or identity threshold were selected (for significance threshold, $P < 0.05$).

Specific MS parameters are included with each set of MS results used in this thesis.

In Mascot, the ion score for an MS/MS match is based on the calculated probability, P , that the observed match between the experimental data and the database sequence is a random event. The reported score is $-10\log(P)$. Thus, the larger the score, the more likely it is that the protein has been correctly identified. See

http://www.matrixscience.com/help/interpretation_help.html for more details.

Appendix 3.1

SDS-PAGE Protocol

Contents

1. Introduction.
2. Selection of protein concentration to run on SDS-PAGE.
3. Preparation of the gel.
4. Setting up the electrophoresis tank.
5. Conditioning of samples prior to running.
6. Running samples on the gel.
7. Treatment of the gel after electrophoresis.

Table

1. Solutions.
2. Reagent sources.

Figures

1. Assembled glass plates and PTFE comb.
2. Electrophoresis tank, tank cover, and PSU.

1. Introduction

SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) separates proteins according to their size (molecular weight) and permits an estimate of protein sizes (molecular weights). This Appendix contains a protocol for SDS-PAGE which has been developed by the LJMU chromatin group from a method described by Laemmli (1970). It is based on polymerising 20% acrylamide, leading to a porous matrix (gel) which effectively separates protein sizes in the range 8 – 30 KDa (acrylamide is a carcinogen, so standard laboratory precautions including wearing of gloves are important). Reed *et al.* (2003) and Encor (2017) explain the SDS-PAGE concept, and various features of a suitable protocol. These features are similar to our own protocol and are discussed below.

SDS has a hydrophobic lipid-like dodecyl part and highly charged sulphate part. The dodecyl part interacts with the protein hydrophobic residues, disrupting the protein folding, and attaching the negatively charged sulphate part to the protein. Combined with β -mercaptoethanol cleavage of protein disulphide bonds, this ensures the protein is a peptide string with an approximately even charge along its length. Consequently, electrophoretic migration speed through a suitable porous matrix depends on the peptide string length, with longer strings (hence larger proteins) migrating more slowly. Migration speed is roughly inversely proportional to protein molecular weight, although proteins with below average hydrophobic residues per unit length run more slowly than expected on SDS-PAGE. PTMs can also cause proteins to run more slowly.

Acrylamide polymerisation is caused by the addition of ammonium persulphate (APS) which decays to produce free radical SO_4^- ions which open the double bond in acrylamide, enabling acrylamide molecules to link together in linear chains. These chains can then be cross-linked by bis-acrylamide molecules. The process is speeded up by adding tetramethylethylenediamine (TEMED) which catalyses APS decay. During polymerisation, exposure to oxygen is minimised since it can inhibit polymerisation by mopping up the free radical SO_4^- ions.

2. Selection of protein concentration to run on SDS-PAGE

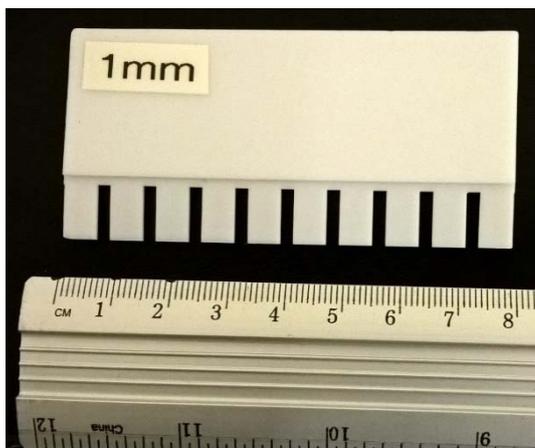
A sample containing one or more proteins is inserted into a well in the gel, and each protein is drawn vertically down the gel by the electrostatic force. Proteins of differing molecular weight become located in separate bands in the gel, made visible by staining. The visual detection limit for Coomassie Blue stain is $\sim 0.2\mu\text{g}$ of protein per band (Reed *et al.*, 2003). However, ideally each band will contain up to $10\mu\text{g}$ of protein, providing a strongly visible band, without overloading and smearing the band. The appropriate sample concentration often needs to be determined by trial and error, with the same sample being run at different concentrations. However, a starting estimate of the appropriate protein concentration is based on the number of separate bands expected from a given sample. For example, if the sample contains 4 proteins of different molecular weights, the sample in the well should ideally contain $4 \times 10\mu\text{g} = 40\mu\text{g}$. The sample volume in a well is typically $40\mu\text{L}$, so a protein concentration of 1g/L will provide the required $40\mu\text{g}$ of protein in the well. However, the initial protein sample is diluted by a buffer in the volume ratio of 2:1 prior to insertion in the well. The target, then, is an initial protein concentration of 1.5g/L (1.5mg/mL), since this dilutes to 1g/L .

3. Making a gel

The gel has two components: a fine grained “resolving gel”, and above it a coarser “stacking gel” which helps provide a uniform electrical field within the resolving gel. The stacking gel is indented with wells into which samples can be placed.

A pair of glass plates, held 1 mm apart by spacers, and with three edges sealed with a flexible orange rubber strip, are assembled using 3 clips (Figure 1). This provides a mould for the gel. Before assembly, the glass plates and other parts are washed with hot water, wiped with 95% ethanol, washed with dH_2O , and allowed to dry. One plate has a cut-out at the top so that a PTFE comb can be inserted to make sample wells in the gel. For our work, the glass plate dimensions were usually $\sim 100 \times 100\text{ mm}$, giving a gel slab $85 \times 85 \times 1\text{ mm}$.

Appendix 3.1



(a)



(b)

Figure 1 Assembled glass plates with PTFE comb

(a) A PTFE comb marked 1 mm (it is designed for gels 1 mm thick) which provides 10 wells, each 6 mm wide by 10 mm deep.

(b) A pair of plates with PTFE comb in position and plates set up for gel polymerisation (however, gel solutions are not included, these would be poured in between the glass plates before inserting the comb).

The resolving gel is based on a mono-acrylamide solution made up as per Table 1, with final acrylamide concentration of 20%. APS and TEMED are mixed in last, since they catalyse polymerisation of the acrylamide. The solution is poured into the gap between the glass plates until the level of solution reaches 20mm below the cut-out in the plate. A small amount of water-saturated iso-butanol is then poured onto the top of the solution to a depth of about 5mm. This reduces surface tension so that the top edge of the gel will be sufficiently flat not to disturb the electric field, and to exclude oxygen which otherwise would hinder the polymerisation. The plates are left to stand while polymerisation occurs. A portion of the solution is put into an Eppendorf which can be tipped to confirm that the solution has turned into a gel (polymerised); this occurs in about 1 hour. The water-saturated iso-butanol is then poured away from above the resolving gel, and traces of it are removed with blotting paper.

The stacking gel is based on a 5% mono-acrylamide solution (Table 1). APS and TEMED are again mixed in last. The stacking gel solution is poured into the remaining gap between the plates, up to the level of the cut-out in the glass plate. A PTFE comb (Figure 1) is immediately inserted into the cut-out, so that the comb teeth displace some of the solution; the teeth form the wells, each 10 mm deep by 6 mm wide, into which samples can be introduced. The PTFE comb is carefully inserted, taking care that no bubbles are trapped under the comb which would otherwise disrupt the electric field. The gel plate assembly is left to stand; after about 40 mins the stacking solution becomes polymerised.

Ideally, proteins are run on the gel immediately, but a gel can be used the following day if the gel plate assembly is put in a cool place and wrapped in a damp tissue.

Table 1. Solutions

Reagent (1)(2)	Resolving gel (3)	Stacking gel (4)	Running buffer (5)	3X buffer (6)
30% acrylamide	10mL	1.25mL	-	-
1% bis-acrylamide	975 μ L	1.95mL	-	-
1.5M tris-HCl pH8.8	3.75mL	-	-	-
1.0M tris-HCl pH6.8	-	1.85mL	-	1.85mL
20% SDS solution (7)	150 μ L	75 μ L	-	-
SDS powder	-	-	0.5g	0.6g
10% APS solution (8)	100 μ L	50 μ L	-	-
TEMED solution	10 μ L	10 μ L	-	-
Bromophenol blue 0.1%	-	300 μ L	-	300 μ L
2-mercapto-ethanol.	-	-	-	1.5mL
Glycerol	-	-	-	3mL
Glycine powder	-	-	14.4g	-
Tris powder	-	-	3.03g	-
dH ₂ O	-	2mL	Make up to 500mL	Make up to 10mL

Notes

1. Reagent sources are identified in Table 2.
2. Percentages are by weight, but more usually dH₂O is simply added to make up the appropriate volume, since 1mL of the solution weighs about 1g.
3. This makes up 15 mL of 20% acrylamide Resolving Gel solution, sufficient for two gels. Slowly mix gel reagents for 3 mins without catalysts (APS + TEMED), then slowly mix for 1 min with catalysts before pouring. Slow mix reduces entrapment of air which hinders polymerisation – this also applies to making up the Stacking Gel solution.
4. This makes 7.5mL of 5% acrylamide Stacking Gel solution, sufficient for two gels. Use mask when handling tris powder.
5. This makes 500mL of Running Buffer, sufficient for one run in electrophoresis tank. When making up, solution is warmed to aid dissolution of the glycine.
6. 10mL of 3X buffer is made up and stored frozen in 1mL aliquots.
7. 50mL SDS solution is made up from powder and kept at room temperature, to prevent SDS coming out of solution. Use mask when handling SDS powder.
8. 10mL of APS solution is made up from powder and stored frozen in 1mL aliquots.

Table 2 Reagent sources

Reagent	Product Reference.
30% Acrylamide solution	Severn Biotech 20-2030-10
1% Bis-acrylamide solution	Severn Biotech 20-2050-10
1.5M Tris-HCl pH 8.8 solution	Severn Biotech 20-7900-10
1.5M Tris-HCl pH 6.8 solution	Severn Biotech 20-7901-10
SDS powder	Sigma L3771, for electrophoresis
APS powder	Sigma A3678
TEMED solution	Sigma T9281
Colloidal Coomassie Blue stain	Severn Biotech 30-38-50
Bromophenol blue 0.1%	Not known
2-mercaptoethanol (1)	Sigma electrophoresis reagent M-7154
Glycerol	VWR Technical 24397.365
Glycine powder	VWR AnalaR 10119CU
Tris powder	Severn Biotech 30-20-60

Abbreviations

APS Ammonium persulphate

SDS Sodium Dodecyl Sulphate

TEMED Tetramethylethylenediamine

Note

1. Synonym: β -mercaptoethanol

4. Setting up the electrophoresis tank.

The clips and orange rubber strip are removed from the glass plates which are now stuck together by the polymerised gel, and the plates are placed in the electrophoresis tank as shown in Figure 2. Running buffer is poured into the top and bottom compartments, to provide electrical connections to top and bottom of the gel. A syringe with a bent needle is used to squirt running buffer under the plates - necessary to displace any air bubbles, since bubbles interfere with the electrical field. The running buffer (Table 1) is made up at least the day before if possible, which reduces trapped air that can accumulate in bubbles under the bottom edge of the plates.

Usually a pair of gels (in their glass plates) are fitted to the tank. Alternatively, a single gel with a blank plate opposite can be fitted.

5. Conditioning of samples prior to running.

The sample is normally made available in a 1M Tris-HCl pH 8.8 buffer, and it is treated with 3X buffer immediately before running on a gel.

3X buffer (Table 1) includes the following reagents:

- (i) 2-mercaptoethanol to reduce/cleave disulphide bridges between protein sub-units.
- (ii) SDS to denature the protein and impart a negative charge along the peptide.
- (iii) Glycerol to increase sample density, so sample will sink to bottom of sample well.
- (iv) bromophenol blue to impart a blue stain to the proteins.
- (v) tris-HCl buffer to maintain the required pH.

The Eppendorf tube containing the sample, and the tube containing the 3X buffer are briefly heated in water near boiling point, prior to vortexing. Heating has two functions (i) to assist with full denaturing of the protein, and (ii) to make the glycerol more fluid - glycerol is viscous at room temperature.

One volume of 3X buffer is added to two volumes of sample. After vortex mixing, the samples are centrifuged at 13,000 rpm for ~ 2 minutes to sediment particles which could distort the electrical field.

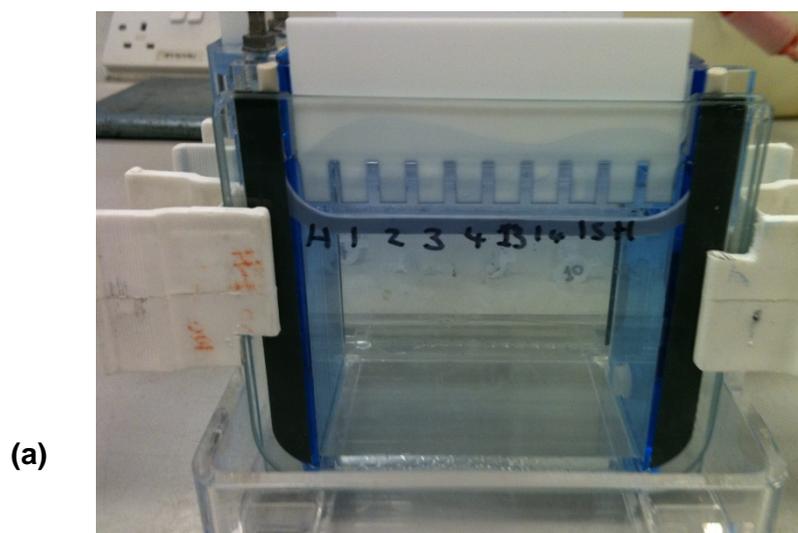


Figure 2 Electrophoresis tank, tank cover, and PSU

- (a) A gel within a pair of plates with PTFE comb in position in the electrophoresis tank. Wells are marked with sample identification codes.
- (b) The pair of plates, the tank cover ready to be put in place, and the electrophoresis power supply unit (PSU). However, gel solution and running buffer are omitted.

6. Running samples on the gel.

The PTFE comb is carefully removed from the gel in the electrophoresis tank, leaving wells in the stacking gel. Using a syringe with a small needle (about 0.8mm diameter), running buffer is squirted into each well to remove any loose or un-polymerised acrylamide.

A pipette with a long slim tip (VWR 732-0508 capillary tips for gel loading) is used to load about 40 μ L of a conditioned sample into each sample well. Typically, there are 10 wells in a gel, two of which are usually reserved for standard reference proteins. Histones were used in this study as reference proteins (see Chapter 4).

The tank cover is put in place. It includes a pair of insulated cables which make electrical connection through the cover with the top and bottom buffer-filled compartments. The positive cable connects to the upper compartment, the negative cable to the lower; this draws down the proteins since they have a negative charge due to the addition of SDS. The cables are connected to a supply which provides the options of an adjustable constant voltage, and an adjustable constant current.

When the power supply is turned on, proteins preceded by a blue band are drawn down the gel. Once the blue band reaches the bottom of the gel the electrical supply is switched off.

Typically, for the work here, the following regime is adopted:

- initially: 26mA constant current per gel.
- later (when blue band ~1 cm above gel bottom): constant current is switched to constant voltage to reduce overheating of the gel.

Under the above regime, the sample runs relatively quickly, in at about 2 - 3 hours. However, there is a potential penalty in terms of some limited gel over-heating and distortion. A slower run or better gel cooling system is necessary to ensure no distortion of the gel.

7. Treatment of the gel after electrophoresis.

The glass plate assemblies are removed from tank, carefully prised apart using a plastic spatula, and the gels are peeled away. One corner of each gel is slightly clipped (triangular or rectangular piece of gel removed) to ensure gels can later be distinguished apart, and the position of protein bands identified correctly.

The gels are rinsed in dH₂O, and then soaked in Coomassie Blue stain for at least two hours, rocking the stain tray every 15 mins or so, such that the stain is evenly absorbed by the protein bands. The gel is then rinsed in dH₂O and can be stored in dH₂O in a fridge (~6°C).

The stained protein bands are visible in normal light, but the gel is usually photographed under UV light, which more clearly shows the Coomassie Blue stain.

If MS identification of a protein is required, the section of gel containing the protein band can be cut out with a scalpel and MS applied. Processing of such gel pieces is described in Chapter 3 Section 3.8.2.

Appendix 3.2

Sheffield University Protocol for treatment of excised SDS-PAGE gel bands containing proteins, in preparation for MS

This protocol was kindly supplied by Professor Mark Dickman.

The excised gel band pieces were de-stained with 200µl of 200mM ammonium bicarbonate in 40% acetonitrile at 37°C for 30min. The supernatant was carefully pipetted off and the de-staining procedure was repeated until the gel pieces became colourless. The excised gel pieces were dehydrated with 200µl acetonitrile by incubating at 37°C for 15min, and the acetonitrile was pipetted off carefully. The gel pieces were rehydrated by incubating with 50mM ammonium bicarbonate in 50% acetonitrile in water at 37°C for 15min. The supernatant was taken off and the samples were vacuum-dried for 15-30min. After drying, 200 µl of 10mM DTT was added to the gel pieces and incubated for 1hr at 56°C. Thereafter the gel pieces were centrifuged at 14200 x g for 10sec and all liquids were discarded. Aqueous 55mM iodoacetamide (200µl) was added to the gel pieces and incubated at ambient temperature for 30min in the dark. After incubation all liquid was discarded. The gel pieces were washed three times with 200µl of 50mM ammonium bicarbonate at ambient temperature for 15min at 37°C. Then the gel pieces were centrifuged at 14200 x g for 10sec and dried in a vacuum concentrator for 30min at 37°C.

Trypsin working solution was prepared by dissolving 20µg of trypsin in 100µl of 1mM HCl and 900µl of 40mM ammonium bicarbonate in 9% acetonitrile. 40µl of this was added for in-gel tryptic digestion, then 50µl of 40mM ammonium bicarbonate in 9% (v/v) acetonitrile in water and the mixture was incubated at 4°C for 15 min to allow the trypsin to diffuse into the gel piece. Enough trypsin buffer (50mM ammonium bicarbonate in 9% acetonitrile) was added to the gel pieces to properly immerse them during incubation at 37°C overnight. The sample was centrifuged at 14200 x g for 10sec and with the aid of a pipette the supernatant was removed and placed in a labelled siliconized peptide low bind collection tube (Eppendorf). Peptide extraction was continued by adding 50µl of 50% acetonitrile to the gel slices, vortexing and incubating at 37°C for 15min. 20µl of 5% (v/v) formic acid was added to the gel pieces and vortexed briefly and the incubation at 37°C was continued for a further 15min. The tube was vortexed briefly and centrifuged at 14200 x g for 10sec.

The supernatant was removed and pooled in the labelled collection tube. Peptide extraction was continued with the addition of a further 50µl of 50% acetonitrile in 5% formic acid and incubation at 37°C for 30min. The tube was vortexed briefly and centrifuged at 14200 x g for 10 sec. The supernatant was recovered and added to the labelled collection tube. The pooled supernatants were vacuum-dried to concentrate the peptides and stored at -20°C. On the day of analysis, the dried peptides were re-suspended in 10µl 5% formic acid and 3-6µl was injected into the reverse-phase column of coupled LC/MS for protein analysis.

Chapter 4

Development of acid extraction methods to prepare histone markers and histone octamers

<u>Contents</u>	Page
4.1 Introduction	98
4.2 Previous studies	98
4.3 Preparation of the Histone Standard	99
4.4 Preparation of nearly pure histone octamers	108
4.5 Conclusions	114
<u>Tables</u>	
Table 4.1 Histone number of residues and molecular weight	106
Table 4.2 HMGB number of residues and molecular weight	107
<u>Figures</u>	
Figure 4.1 Summary of the acid extraction method for preparing the Histone Standard to provide markers for SDS-PAGE	100
Figure 4.2 SDS-PAGE gel showing the Histone Standard and Octamers prepared by simple acid extraction methods from chicken erythrocytes	105
Figure 4.3 Summary of the acid extraction method for preparation of nearly pure histone octamers	109
Figure 4.4 Floating precipitate containing octamers	111

4.1 Introduction

This chapter describes the development of two acid extraction methods. The first required only an acid extraction and a dialysis to obtain a set of individual histones. The second provided pure histone octamers. Acid extraction has the disadvantage that it compromises acid-labile histone PTMs (Chen *et al.*, 1974; Sut and Biterge, 2017), but it has the advantage that large quantities of histones can be obtained with simple steps and without requiring an ultracentrifuge. When using acid extraction, the histones should be used in circumstances in which the potentially degraded acid-labile PTMs do not matter.

In the first method, acid extraction of histones from chicken erythrocytes was used to provide a set of highly reproducible protein markers covering the molecular weight range 11.4 – 22.5kDa. These markers were included in SDS-PAGE gels used in the research covered by this thesis and are referred to as the "Histone Standard". A high yield of markers was obtained since histones are the most prolific nuclear proteins. Degradation of the histone PTMs was not a problem when used as SDS-PAGE markers - PTMs are only a small fraction of total histone weight, and changes in protein charge due to PTM removal are swamped by the charge on SDS attached to the protein for SDS-PAGE.

In the second method, a very high yield of nearly pure histone octamers was prepared from chicken erythrocytes, using phosphoric acid extraction in high salt conditions. It was originally hoped that extraction could be performed in more mild acid conditions to preserve PTMs, but, in practice, extraction at pH2 was necessary. However, the use of KCl and phosphates in suitable concentrations provided conditions conducive to maintaining intact octamers throughout the extraction/isolation process, paving the way for future structural studies.

4.2 Previous studies

As discussed in Chapter 3, acid extraction of histones is not new, but using this to prepare markers as described here seems to have been largely confined to our research group at LJMU. A former member of the group, Connie Sin, examined this method and published part of the protocol given here (Sin, 2009).

The second method is unusual in using phosphoric acid in high salt conditions to extract histones. It is a recent development of other work by our research group (Zhuang, 2011).

4.3 Preparation of the Histone Standard

4.3.1 Summary

Figure 4.1 summarises the main steps applied to prepare the Histone Standard. To support a range of experiments, chicken erythrocyte nuclei were prepared from about 4.5L of fresh chicken blood and were frozen in a glycerol freezing buffer to prevent ice crystals forming. The procedure provided about 20 Falcon tubes each containing 40mL of loosely compacted frozen nuclei. Subsequently, the nuclei were recovered from one tube of frozen nuclei, lysed in Histone Extraction Buffer, and dialysed into a solution suitable for running the Histone Standard on an SDS-PAGE gel. A gel was run to confirm the expected histone content.

4.3.2 Preparation of erythrocyte nuclei

Chilled buffers were prepared as below. Benzamidine hydrochloride degrades in solution so it was added to the buffer immediately before use. The lysis buffers were hypotonic (leading to osmotic pressure within the cell), and contained a detergent to degrade cells walls, and MgCl₂ to maintain the nuclear membrane.

Anti-Clotting Buffer	10% w/v tri-sodium citrate, 2.5mM benzamidine hydrochloride.
Cell Lysis Buffer 1	1.25% w/v Triton X100, 10mM KCl, 10mM Tris-HCl pH 6.8, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.
Cell Lysis Buffer 2	0.5% w/v Triton X100, 10mM KCl, 10mM Tris/HCl pH 6.8, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.

4.5L blood was collected into 300mL Anti-Clotting Buffer in a tray as the chickens passed by, with the tray continually rocked to ensure thorough mixing. The blood was filtered on site through two layers of butter muslin into a container. The blood container was kept cold by embedding in ice during transport to the laboratory.

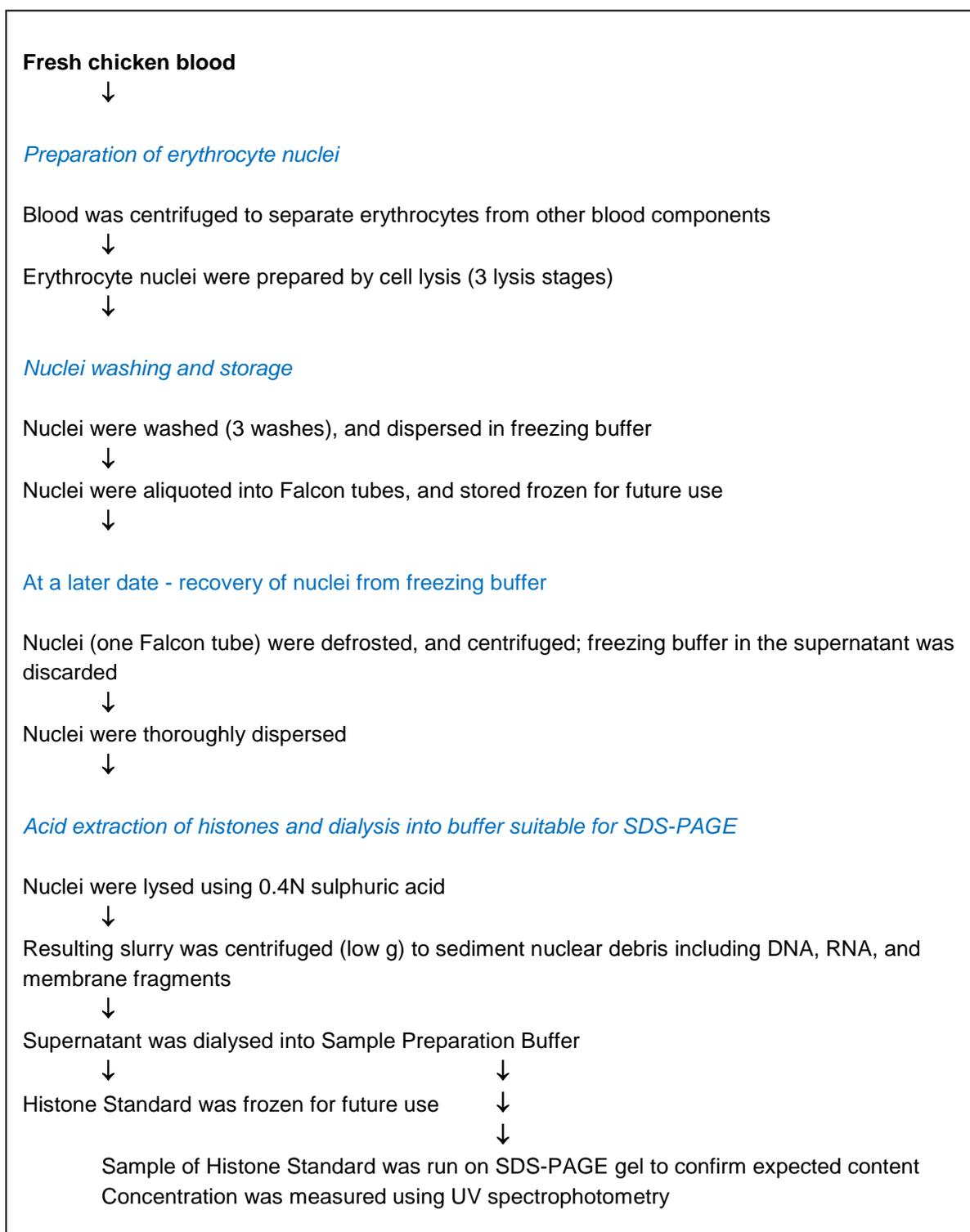


Figure 4.1 Summary of the acid extraction method for preparing the Histone Standard to provide markers for SDS-PAGE

See Section 4.3.2 onwards for further details. All operations were conducted at ~4°C unless otherwise stated. Centrifuging was carried out using a Sorvall RTH-750 rotor.

Back at the laboratory, the blood was centrifuged in six 0.75L bottles at 2750rpm (Sorvall RTH-750 rotor) for 20mins to separate the erythrocytes from the other blood components. As with all centrifuging, this was conducted at 4°C. The blood separated into three layers; the top layer (plasma) was removed with a water pump and discarded. The next layer, a "skin" of buffy coat (containing white blood cells and platelets), was moved to the side of the bottle with a spatula to assist removal and was similarly removed. The remaining layer contained erythrocytes. They were lysed in three stages,

For the first lysis stage, the loose pellets of intact red blood cells in each of the six bottles were agitated with a rod to release cells stuck to the bucket bottoms and were combined to form ~0.75L loose slurry. The slurry was poured into 7.25L Cell Lysis Buffer 1 in a bucket cooled by ice bags around the side, and the solution was stirred briskly with a large motor-driven stirring bar. After 1hour, the solution was centrifuged in twelve 0.75L bottles at 2500rpm for 20mins. Most supernatant was discarded using a water pump, but a layer of supernatant was left above the loose pellets, equal in volume to the loose pellet.

For the second lysis stage, Cell Lysis Buffer 2 was added to each of the twelve bottles, about equal in volume to the liquid remaining in each bottle. Each bottle was swirled to provide gentle mixing. To prevent aggregates, which would compromise equal exposure of cells, the suspensions were thoroughly homogenised, by gently shaking, and by sucking in and out the slurry with a submerged plastic pipette. The contents of the 12 bottles were then transferred to four 0.75L bottles. The content of each of the four bottles was made up to 0.75L with Cell Lysis Buffer 2, gently mixed (as above), and centrifuged at 2,000rpm for 20mins. The supernatant was discarded.

For the third lysis stage, pellets from the above procedure, including a little supernatant, were mixed with an equal volume (approx.) of Cell Lysis Buffer 2, homogenised again by squeezing in and out with a submerged plastic pipette, and transferred to four new 0.75-litre bottles via a syringe through a 21-gauge needle, to further disperse clumps of cells and nuclei. The content of each of the four bottles was made up to 0.75L with Cell Lysis Buffer 2, swirled to provide gentle mixing, and centrifuged at 2000rpm for 20mins. The supernatant was discarded, leaving loose pellets of nuclei.

4.3.3 Nuclei washing and storage

Chilled buffers were prepared as below.

Nuclei Washing Buffer 10mM KCl, 10mM Tris/HCl pH 6.8,
3.5mM MgCl₂, 2.5mM benzamidine hydrochloride.

Nuclei Freezing Buffer 10mM KCl, 10mM Tris/HCl pH 6.8,
3.5mM MgCl₂, 2.5mM benzamidine hydrochloride.
All components were dissolved in 40% w/v glycerol.

Pellets from the cell lysis procedure were mixed with an equal volume (approx.) of Nuclei Washing Buffer in the four 0.75L bottles. They were gently shaken and homogenised by squeezing in and out with a submerged plastic pipette and were transferred to two new 0.75L bottles via a syringe through a 21-gauge needle. The content of each bottle was made up to 0.75L with Nuclei Washing Buffer, shaken thoroughly, and centrifuged at 2000 rpm for 20 minutes. Most of the supernatant was discarded leaving ~100mL of nuclei in each bottle. The above procedure was repeated twice more so that the nuclei were almost white (indicating most of the red haemoglobin had been removed).

To prepare the nuclei for storage, the above nuclei pellets were dispersed by stirring with an equal volume of Nuclei Washing Buffer. The suspension was then stirred with an equal volume of Nuclei Freezing Buffer. The nuclei, now in 20% glycerol, were gently shaken and homogenised by squeezing in and out with a submerged plastic pipette and were transferred via a syringe through a 21-gauge needle into ~20 Falcon tubes (each containing ~40mL). The tubes were frozen at -20°C.

4.3.4 Recovery of nuclei from freezing buffer

A chilled buffer was prepared as below.

Nuclei Washing Buffer 10mM KCl, 10mM Tris/HCl pH 6.8,
3.5mM MgCl₂, 2.5mM benzamidine hydrochloride.

One Falcon tube containing ~40mL of nuclei in Nuclei Freezing Buffer was thawed slowly at 6°C and centrifuged at 1,500rpm for 20mins. The tube contained about 15mL of sedimented nuclei and 25mL of clear supernatant above. 20mL of the supernatant was removed with a water vacuum pump, and the nuclei were shaken gently and dispersed with a plastic Pasteur pipette in the remaining solution. The supernatant was discarded, and the sedimented nuclei pellet (~13mL) was mixed with an equal volume of Nuclei Washing Buffer and dispersed by gentle swirling. The mixture was made up to 80mL with Nuclei Washing Buffer and stirred. The nuclei in suspension were then thoroughly dispersed by transferring via a syringe through a 21-gauge needle to a separate stirred beaker.

4.3.5 Acid extraction of histones and dialysis into buffer suitable for SDS-PAGE

Chilled buffers were prepared as below.

Histone Extraction Buffer: 0.4N sulphuric acid.

Sample Preparation Buffer: 10mM Tris/HCl pH6.8, 2.5mM benzamidine hydrochloride.

Normality (N) reflects the number of hydrogen ions available. 0.4N = 0.2M since sulphuric acid provides two hydrogen ions.

20mL of Histone Extraction Buffer was put into a new Falcon Tube, and 20mL of dispersed nuclei was gently layered on top of the buffer, and the tube was shaken suddenly – this ensured an even exposure of nuclei to lysis conditions, so that all nuclei were lysed. The acid lysed the nuclei and dissociated the DNA attached to the histones. The resulting slurry was centrifuged at 3,000rpm for 30mins and the supernatant was retained as the histone extract. The histone extract (~15mL) was dialysed into the Sample Preparation Buffer (three changes of buffer), aliquoted into Eppendorf tubes, and frozen.

4.3.6 Sample of Histone Standard run on SDS-PAGE gel to confirm expected content

10 μ L samples from the frozen histone extract were diluted in various amounts of dH₂O. From these solutions, 40 μ L samples which included 33% 3X buffer (see protocol in Chapter 3) were run on SDS-PAGE gels. Figure 4.2 shows the gel obtained for two different nominal dilutions, marked H 30:1 and H 45:1; the Histone Standard comprised linker and core histones, with molecular weights as shown.

4.3.7 The Histone Standard: discussion

Histone bands in the gel

Chromatin comprises a string of nucleosomes, and each nucleosome includes one histone octamer made from two histone dimers (H2A.H2B) and one histone tetramer (H3.H4)₂, together with one linker histone. Consistent with this, Bates and Thomas (1981) showed that, within their accuracy of measurement, chicken erythrocyte nuclei contained equal quantities of H2A, H2B, H3, and H4. They also showed that the erythrocyte chromatin contained linker histone variants, designated H1A, H1B, and H5, with about 1.3 linker histones (of one variant or another) for every eight core histones (see Chapter 2, Table 2.4). Most histone molecules in the nucleus are incorporated into chromatin. For example, most H1 is bound to chromatin, as shown by Misteli *et al.* (2006), although as reviewed by Parseghian (2015) the linker histones are all in dynamic equilibrium between the bound and unbound states. In view of the histone stoichiometry, each histone in the Histone Standard appeared on the gel with roughly the same intensity (Figure 4.2).

In chicken erythrocytes, the linker histone may be H5 (most common) or one of several linker histone variants. These variants fall into two categories, which we have designated H1a or H1b. H1a variants contain 224 or 225 residues, and H1b variants contain 218 – 220 residues. Consistent with their limited ranges of residues, the H1a and H1b variants form just two bands on our gel (Figure 4.2). A third band was formed by H5, this has lower molecular weight than the H1a/H1b variants. Table 4.1 shows the relevant residue numbers and MWs, with chicken H1.03 taken as representative of the H1a variants, and chicken H1 as representative of the H1b variants; the table also includes information for the core histones H2A, H2B, H3, and H4.

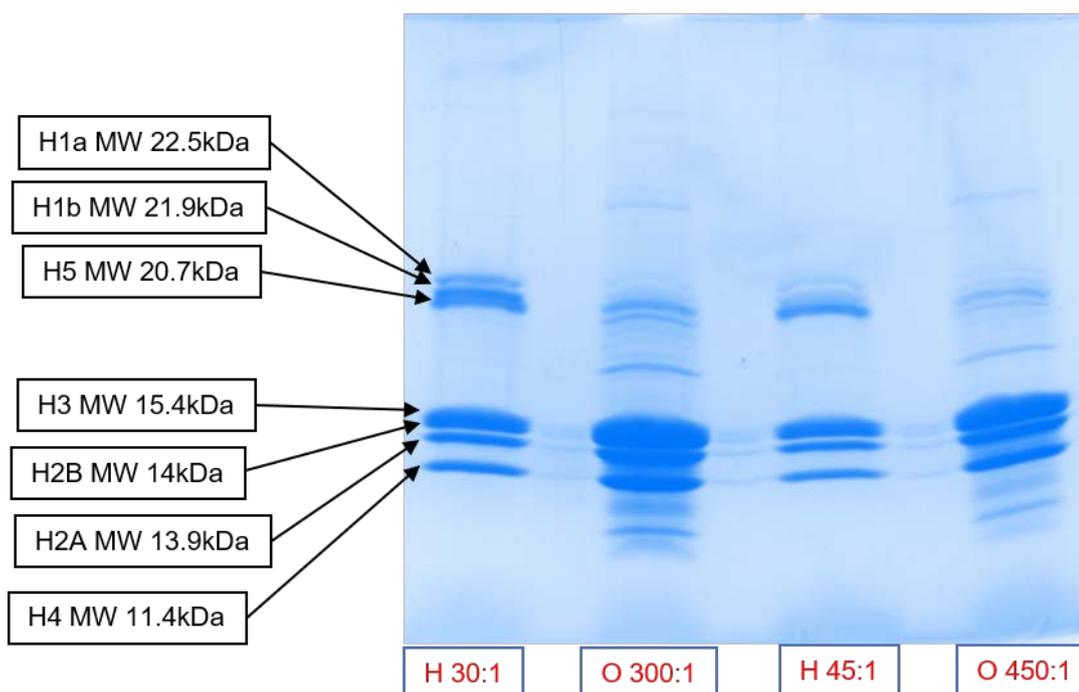


Figure 4.2 SDS-PAGE gel showing the Histone Standard and Octamers prepared by simple acid extraction methods from chicken erythrocytes

H: Histone Standard;

O: Octamers - the strong bands show mainly core histones are present, indicative of a high yield of almost pure octamers.

Nominal dilutions of samples in dH₂O were as shown above in red. However, a review of the strength of the gel bands indicated that the effective dilutions may be less than shown – see Section 4.3.7 "Histone standard yield".

Stain:	Coomassie Blue (although photographed in black and white)
Stacking gel:	5% acrylamide
Resolving gel:	20% acrylamide
Running buffer:	glycine/tris/SDS
Sample treatment:	1 part 3X Buffer to 2 parts sample
Sample loading:	40µL per well
SDS-Page protocol:	as described in Chapter 3

Table 4.1 Histone number of residues and molecular weight

Histone	Uniprot identifier	Residues	MW* (kDa)	Position on the gel (Figure 4.2)
H1.03 (representative of H1a variants)	P08285	224	22.5	1 (top)
H1 (representative of H1b variants)	P09987	218	21.9	2
H5	P02259	190	20.7	3
H3	P84229	136	15.4	4
H2B	P0C1H5	126	14.0	5
H2A	P02263	129	13.9	6
H4	P62801	104	11.4	7 (bottom)

* From UniprotKB database.

4.3.7 (continued) *The Histone Standard: discussion*

Location of histone bands in the gel and usefulness for studying HMGB proteins

Berg *et al.* (2015) present results from Weber and Osborn (1975) which show that, as a general rule, the speed at which a protein migrates through an SDS-PAGE gel is approximately inversely proportional to the logarithm of the protein MW, although some proteins are major exceptions to this rule. Bates and Thomas (1981) have shown each core histone in the Histone Standard appears on an SDS-PAGE gel in the order expected from the order of their MWs (Table 4.1). The linker histones bands on the gel are likewise ordered. However, the spacing between bands in Figure 4.2 indicates some limited variation from the speed rule.

The Histone Standard was particularly useful for studying HMGB proteins. The HMGB proteins are heavier than the histone proteins (Table 4.2), so might be expected to run slower than any of the histones. However, the HMGB proteins run anomalously fast on an SDS-PAGE gel, hence the designation "High Mobility Group" (HMG) proteins. For this reason, when present, the HMGB proteins provide gel bands at positions roughly midway between the H5 and H3 bands. We have not found any documented explanation for the higher mobility of HMGB proteins, but it may be related to the enhanced negative charge provided by their acidic tails.

Table 4.2 HMGB number of residues and molecular weight

Protein	Uniprot identifier	Residues	MW* (kDa)
HMGB1	Q9YH06	215	24.9
HMGB2	P26584	207	23.8
HMGB3	P40618	202	23.1

* From UniprotKB database.

Histone Standard yield

As indicated in Section 4.3.5, after dialysis the Histone Standard was present in Sample Preparation Buffer (~15mL). A UV measurement at 280nm indicated a concentration of proteins of about 4mg/mL, indicating an estimated total yield of ~60mg.

In preparing samples for Figure 4.2, the Histone Standard was diluted by the addition of various amounts of dH₂O followed by vortexing, leading to the nominal dilutions shown in Figure 4.2. According to Reed *et al.* (2003), the detection limit using Coomassie Blue stain (R-250) on SDS-PAGE gels is ~0.2µg protein, and our group has previously assumed that a strong band on a gel contains ~10µg of protein. However, given the above histone protein yield of 4mg/mL, each histone band (7 bands) clearly visible on the gel at dilution 30:1 (Figure 4.2) contained only ~0.5 µg of the histone - less than expected on past experience. It is possible that, due to insufficient mixing, the effective dilutions were less than the nominal dilutions in Figure 4.2. On the other hand, overnight staining in fresh Coomassie Blue stain was applied, and a modern Biorad Chemidoc illuminator with automatic intensity compensation was used to photograph the stained gel; this may have improved the visibility of the bands compared with some of our previous work. It may be appropriate for us to re-evaluate the sensitivity afforded by our current gel staining and gel photography procedures.

4.4 Preparation of nearly pure histone octamers

4.4.1 Summary

Figure 4.3 summarises the main steps in preparing the octamers. As before, a batch of chicken erythrocyte nuclei was prepared and frozen; later two tubes of nuclei were recovered from the freezing buffer for the work here. The nuclei were lysed, dialysed into a buffer designed to maintain the histone octamers in an intact state, and the solution was concentrated. Addition of Octamer Precipitation Buffer enabled almost pure octamers to be isolated as a floating precipitate. A sample of the octamers was dissolved in a buffer suitable for SDS-PAGE, and a gel was run to confirm the octamer content.

4.4.2 Preparation and freezing of a batch of nuclei, and later recovery of two Falcon tubes of nuclei from the freezing buffer

These steps were as described in Sections 4.3.2 - 4.3.4.

4.4.3 Acid extraction, dialysis into Octamer Maintenance Buffer, and volume reduction

Chilled buffers were prepared as below.

Acidic Nuclei Lysis Buffer: (pH2)	3.3M KCl, 166mM H ₂ KPO ₄ , 166mM phosphoric acid, 4mM benzamidine.
Octamer Maintenance Buffer: (pH6.8)	2M KCl, 200mM HK ₂ PO ₄ , 200mM H ₂ KPO ₄ , 2.5mM benzamidine hydrochloride.

4.4.3 (continued) Acid extraction, dialysis into Octamer Maintenance Buffer, and volume reduction

Acidic Nuclei Lysis Buffer (30mL) was put into a new Falcon Tube, and 20mL of dispersed nuclei from the first Falcon tube was gently layered on top of the Buffer, and the tube was shaken suddenly – this ensured an even exposure of the nuclei to lysis conditions. This was repeated for the nuclei in the second Falcon Tube. A slurry containing proteins, DNA, RNA, and membrane debris was produced. After standing for 20min, the slurry was centrifuged at 3000rpm for 20min. The supernatants in the two tubes were combined and dialysed into Octamer Maintenance Buffer. The resulting solution was reduced to 17.5mL using an Amicon concentrator (10kDa MWCO ultrafiltration membrane).

4.4.4 Selective precipitation of octamers

Chilled buffers were prepared as below.

Octamer Precipitation Buffer: 2M KCl, 1M HK_2PO_4 , 1M H_2KPO_4 ,
(pH6.8) 2.5mM benzamidine hydrochloride.

Sample Preparation Buffer: 10mM Tris/HCl, 2.5mM benzamidine hydrochloride.
(pH6.8)

Octamer Precipitation Buffer (58mL) was added to the above concentrated solution (17.5mL). Two precipitates were formed, one less dense than the solution, and the other more dense. After centrifugation at 3000rpm for 20min, a precipitate containing octamers was seen floating on the supernatant (Figure 4.4). The floating precipitate was skimmed off and dissolved in pH6.8 Sample Preparation Buffer (10mL). The resulting solution was reduced to 1.8mL using an Amicon concentrator (10kDa MWCO ultrafiltration membrane) and retained for assay and for future experiments.

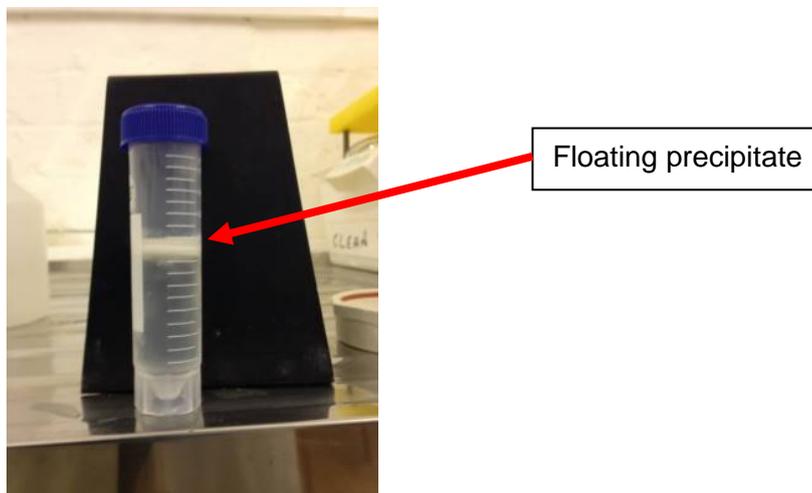


Figure 4.4. Floating precipitate containing octamers

4.4.5 Sample of Octamer run on SDS-PAGE gel to confirm expected content

5 μ L samples from the concentrated octamer solution were diluted in various amounts of dH₂O. From these solutions, 40 μ L samples which included 33% 3X buffer (see protocol in Appendix D) were run on SDS-PAGE gels. Figure 4.2 shows the gel obtained for two different nominal dilutions (300:1 and 450:1). The octamer was dissociated into the component core histones H2A, H2B, H3, and H4. As seen in Figure 4.2, the gel bands for these core histones are overloaded compared with other, much feinter bands in the same lanes. This indicated the precipitate contained nearly pure octamers.

4.4.6 Method for preparing octamers: discussion

Maintaining intact octamers

The octamer preparation method was designed to keep the histones in a suitable high salt and phosphate environment which, from experience (Zhuang, 2012), would keep the octamers intact. This required that the Acidic Nuclei Lysis Buffer contained 3.3M KCl, in addition to the phosphates and protease inhibitor. This very high salt concentration was maintained in solution by slightly warming the solution during make-up and using the solution without delay. The group's first choice protease inhibitor, benzamidine hydrochloride, could not be used because it was insoluble in the high salt buffer. Benzamidine was used instead; this appeared effective – the core histones were seen to be intact (Figure 4.2). The octamer preparation was carried more recently than the work covered by other chapters in this thesis; this was the first time we employed benzamidine, although it has been used by others (Chapter 8.4.4).

On this occasion, to facilitate SDS-PAGE assay, the floating precipitated octamers were dissolved in low ionic strength Sample Preparation Buffer, dissociating the octamers into dimers (H2A.H2B) and tetramers (H3.H4)₂. However, for future use the floating precipitate would be dissolved in the Octamer Maintenance Buffer. This high salt, moderately high phosphate buffer (2M KCl, 0.2M H₂KPO₄, 0.2M HK₂PO₄) provides sufficient ionic strength and phosphate ions to retain intact octamers, while avoiding the formation of octamer crystals - Chantalat *et al.* (2003) found that crystals were formed in buffer containing 2M KCl, 0.675M H₂KPO₄, 0.675M HK₂PO₄. In a separate experiment, we have applied gel filtration chromatography to the histone octamers in Octamer Maintenance Buffer to confirm that the octamers remained intact – as the larger unit, the octamers eluted from the gel filtration column ahead of histone dimers and tetramers.

Maintaining intact histone octamers has practical value, since it avoids the time consuming step of assembling the separate histone dimers (H2A.H2B) and tetramers (H3.H4)₂ into octamers. Luger *et al.* (1999) assembled octamers by dialysing the core histones into high salt conditions, followed by gel filtration. The dialysis had to be carried out precisely since high salt is required to counteract the positive repulsion between the histone dimers and tetramers, yet the tetramers are insoluble in high salt and, unless induced into the octamers, will come out of solution leaving an imbalance in dimers. The Luger method generally leaves some un-combined histones, which must then be removed by gel filtration.

Octamer quality and yield

The strong bands in Figure 4.2 (lanes marked O 300:1 and O 450:1) show mainly core histones are present, indicative of a high yield of octamers. The presence of other, feinter bands in the octamer lanes indicated the octamer precipitate was not quite pure. However, we have previously shown that gel filtration chromatography can be used to improve the purity of octamers obtained in a similar precipitate (Zhuang, 2011).

Using an alternative octamer extraction method (Zhuang *et al.*, 2014), we have previously prepared histone octamers with a total yield of ~36mg from two Falcon tubes of nuclei. Here a UV absorption measurement at 278nm on a 50:1 diluted sample (part of the 1.8mL sample, Section 4.4.4) indicated an absorbance of 0.6au for a 1cm pathlength. Taking the extinction coefficient for histones to be 0.464 on average, this yields a concentration of histones of 1.3mg/mL. Thus, the total yield of proteins here, from two tubes of nuclei, was significantly higher than our previous method, at $1.8\text{mL} \times 50 \times 1.3\text{mg/mL} = 117\text{mg}$.

Impact of buffer acidity on extraction of proteins from lysed nuclei

Lysis of the nuclei left a mixture of proteins, DNA, RNA, and nuclei membrane debris. We had hoped that if the acidity of the nuclei lysis buffer was just below the pI of DNA (about 5 - Saeki and Kunito, 2010), the reduced electrostatic attraction between the positively charged histones and the negatively charged DNA would permit histones to be separated from the DNA at low g acceleration (low relative centrifugal force). Unfortunately, our trials showed that in lysis conditions less acidic than pH2, the nuclear extract formed a gel; ultracentrifugation is then necessary if proteins are to be separated from the gel.

The Acidic Nuclei Lysis Buffer at pH2 produced a slurry from the lysed nuclei instead of a gel, and we found that the proteins could be separated from the gel at low g acceleration. We applied 3000rpm for 20 minutes to sediment the less soluble proteins, DNA, RNA, and nuclei membrane debris, leaving the more soluble proteins such as histones in the supernatant. For the Sorvall RTH-750 rotor we used, samples are situated at a maximum radius of 18.5cm, so 3000rpm corresponded to 1864g. This compares with 100,000g we applied overnight to obtain the soluble proteins from the gel which arose when a mildly acidic nuclei lysis buffer was applied (see Chapter 5).

In sufficiently strong acid, the DNA phosphodiester bonding fails, producing highly fragmented DNA - this may be the mechanism that replaced the intransigent gel with a slurry more amenable to low g centrifugation. However, we have no information about the fragmented status of the DNA in our experiment.

4.5 Conclusions

1. Using a simple sulphuric acid extraction of histone proteins, large quantities of SDS-PAGE markers (MW range 11.4 - 22.5kDa) have been efficiently prepared from chicken erythrocytes.
2. A high concentration of nearly pure octamers was also efficiently extracted from chicken erythrocytes. In this case, phosphoric acid was used for the extraction, as part of a strategy to maintain the octamers in an intact state. This has the practical advantage of avoiding additional time-consuming steps associated with reassembling separate histone dimers and tetramers into octamers.
3. Benzamidine was used by us for the first time as a protease inhibitor instead of our more usual choice of benzamidine hydrochloride. It appeared effective in minimising degradation of the octamers. As indicated in Chapter 8.4.4, this choice of protease inhibitor will have value elsewhere in our future work.

Chapter 5

Investigation of potential HMGB/histone protein complexes

<u>Contents</u>	Page
5.1 Introduction	117
5.2 Previous studies	117
5.3 Preparation of nuclear protein subset cePNE4' and investigation of potential complexes within	122
5.4 A model to explain the cation-exchange chromatography result	149
5.5 Further support for the proposed model	158
5.6 Conclusions and recommendations	159
<u>Tables</u>	
Table 5.1 Some previous studies of HMGB/linker histones interactions	119
Table 5.2 Some previous studies of HMGB/core histones interactions	120
Table 5.3 Status of HMGBs and histones in the nuclear protein sets	130
Table 5.4 (a) Summary of MS results for peaks in cation-exchange chromatogram of cePNE4' proteins (Figure 5.3b)	135
Table 5.4 (b) Parameters used to obtain MS results for fractions from chromatography of cePNE4' proteins	138
Table 5.5 Proteins in each peak fraction of the cation-exchange chromatogram of cePNE4' proteins (Figure 5.3(a) and (b))	139
Table 5.6 Some features of chicken HMGB and histone proteins	140
Table 5.7 Summary of MS results for fractions 5a1 – 5a5 and 5c1 – 5c4 from gel filtration chromatography applied to fractions 5a and 5c from the cation-exchange chromatography applied to cePNE4' proteins	144
Table 5.8 Summary of proteins in fractions 5a1 – 5a5 of the gel filtration chromatogram (Figure 5.5a) for Fraction 5a obtained by cation-exchange chromatography of cePNE4' proteins	147
Table 5.9 Summary of proteins in fractions 5c1 – 5c5 of the gel filtration chromatogram (Figure 5.5b) for Fraction 5c obtained by cation-exchange chromatography of cePNE4' proteins	148
Table 5.10 Summary of MS results for histone H5	152
Table 5.11 Charges on proteins and complexes relevant to the cation-exchange chromatogram	154

<u>Figures</u>	Page
Figure 5.1 Summary of the method for isolating cePNE4' proteins from chicken erythrocytes	123
Figure 5.2 Summary of the chromatography applied to the cePNE4' proteins	124
Figure 5.3 Cation-exchange chromatogram of cePNE4' proteins	131
Figure 5.4 SDS–PAGE gels of peak fractions from cation-exchange chromatography of cePNE4' proteins	133
Figure 5.5 Gel filtration chromatograms for proteins in Fractions 5a and 5c from cation-exchange chromatography of cePNE4' proteins (Figure 5.3b)	142
Figure 5.6 SDS–PAGE gels of peak fractions from gel filtration chromatography of Fractions 5a and 5c from cation-exchange chromatography of cePNE4' proteins	143
Figure 5.7 Proposed arrangement for attachment of histone/HMGB to cation-exchange stationary phase	155
Figure 5.8 A model for the release of proteins from the cation-exchange column	156

5.1 Introduction

This chapter describes an investigation of potential HMGB/histone protein complexes in the nuclei of chicken erythrocytes, extending work by Sin (2009) and Zhuang (2011). Experiments were conducted in an environment containing phosphate ions, to simulate the presence of charged phosphates due to DNA. The work is based on a subset of the nuclear proteins extracted from erythrocytes in fresh chicken blood. This provides a simplified environment in which to examine potential complexes, while recognising that it may preclude some potential complexes which could be mediated by the excluded nuclear proteins. The simplified environment greatly assisted the assay of proteins using SDS-PAGE and MS.

The experimental work was carried out jointly by the author and Dr Q Zhuang.

5.2 Previous studies

5.2.1 The bigger picture

HMGB proteins, especially HMGB1, have very important interactions with the chromatin machinery. Applying Fluorescence Loss in Photobleaching (FLIP) to HeLa cells, Scaffadi *et al.* (2002) showed that HMGB1 binds transiently to chromatin. Thomas and Travers (2001) have identified potential models for HMGB1 as a chaperone for transcription factors. *In vitro*, Bonaldi *et al.* (2002) showed HMGB1 increases the efficiency of ACF-dependent nucleosome sliding. Ju *et al.* (2006) obtained evidence of replacement of H1 by HMGB1 when investigating protein complexes at the pS2 promoter site in MCF-7 cells. Celona *et al.* (2011) showed HMGB1 increases the rate of assembly of nucleosomes. Thus, binding of HMGB proteins to histones is a part of a bigger picture, in which HMGB proteins interact with chromatin, either by binding of the HMG boxes to DNA, or by binding to linker or core histones, or by binding to both. Several PDB structures are available in which one or two HMG boxes are bound to DNA, but there is no structure for a complete HMGB molecule bound either to DNA or a nucleosome.

HMGB/ histone protein interactions have been investigated for more than 30 years, although some studies may have been compromised. HMGB proteins prepared by a standard method (acid extraction followed by acetone precipitation) became oxidised

within two hours (Kohlstaedt *et al.*, 1986), forming intramolecular disulphide bonds which hindered HMGB1/H1 binding (Kohlstaedt *et al.*, 1994). Selected studies are discussed below, using the modern terminology HMGB1, HMGB2, and HMGB3 in place of the older HMG group terminology used in some papers.

5.2.2 Interactions between HMGB proteins and linker histones (H1 and H5)

Table 5.1 indicates some HMGB/linker histone interactions. In buffers at biological pH and ionic strength, HMGB1 bound to histone H1 in ratio of about 1:1. HMGB1 also bound to H5. However, Cary *et al.* (1979) found no evidence of HMGB2 binding to H1, perhaps because HMGB2 is more basic than HMGB1, or because the HMGB2 C-terminal is less accessible (shown by reduced susceptibility to proteolytic cleavage).

Cary *et al.* (1979) showed that increased buffer strength (in excess of that found in a cell) compromised HMGB1/H1 binding, and Cato *et al.* (2008) found that HMGB1/H1 binding in free solution is primarily through electrostatic attraction between the HMGB1 acidic C-terminal tail and the H1 basic C-terminal tail.

Catez *et al.* (2004) provided evidence of HMGB1 competing with H1 for nucleosomal binding sites. HMGB1 and H1 appear to bind to opposite sides of bent DNA, and may do so simultaneously (An *et al.*, 1998; Zlatanova and van Holde, 1998).

5.2.3 Interactions between HMGB proteins and core histones (H2A, H2B, H3, H4)

Table 5.2 indicates some HMGB/core histone interactions. Studies have consistently shown that buffer concentrations above ~250mM compromised the formation of an HMGB/core histone complex. Apart from this, the picture has been confusing. Earlier studies (Bernues *et al.*, 1983, 1986) found HMGB/histone dimer complexes, and binding between the HMGB1 N-terminal and the histone dimer (H3)₂. Using HMGB1 in solution with nucleosomes, Stros (1987) also found an interaction between the HMGB1 N-terminal and (H3)₂. In contrast, Ueda *et al.* (2004) and Watson *et al.* (2014) found an interaction only between the HMGB1 C-terminal and H3 N-terminal.

Table 5.1. Some previous studies of HMGB/linker histone interactions

HMGB	Histone	Complex	Comment
HMGB1	H1	HMGB1/H1	<p>Proteins in free solution.</p> <p>Complex is found in buffers up to 200mM, complex is compromised by greater buffer concentrations.</p> <p>Ratio of HMGB1:H1 is ~1:1.</p> <p>No complex is found if HMGB1 cysteine residues are oxidised.</p> <p>Interaction is between tails of HMGB1 C-terminal (-ve) and H1 C-terminal (+ve). Deletion of more than 10 residues from HMGB1 C-terminal compromises the complex.</p> <p>Cary <i>et al.</i> (1979), Kohlstaedt <i>et al.</i> (1987, 1994) Cato <i>et al.</i> (2008).</p>
HMGB1	H1	HMGB1/H1/ linker DNA	<p>Proteins in solution with chromatin.</p> <p>HMGB1 and H1 both bind to linker DNA; at the same time or in competition.</p> <p>An <i>et al.</i> (1998), Zlatanova and van Holde (1998), Thomas and Travers (2001), Ueda <i>et al.</i> (2004), Catez <i>et al.</i> (2004).</p>
HMGB1	H5	HMGB1/H5	<p>Proteins in free solution.</p> <p>Cary <i>et al.</i> (1979), Cato <i>et al.</i> (2008).</p>
HMGB2	H1	No complex	<p>Proteins in free solution.</p> <p>Cary <i>et al.</i> (1979)</p>
HMGB proteins	Linker histones	HMGB1/ H1.11R. HMGB1/ H5.	<p>Proteins in free solution.</p> <p>Sin (2009)</p>

Table 5.2 Some previous studies of HMGB/core histone interactions

(Page 1 of 2)

HMGB	Histone	Complex	Comment
HMGB1	Core histones	HMGB1/ H2A.H2B	Proteins in free solution. Complex found in buffers up to 225mM, complex is compromised by greater buffer concentrations. Interaction is between central/C-terminal region of HMGB1, and H2A.H2B Bernues <i>et al.</i> (1983, 1986)
HMGB2	Core histones	HMGB2/ H2A.H2B	Proteins in free solution. Complex found in buffers up to 225mM, complex is compromised by greater buffer concentrations. Interaction is between central/C-terminal region of HMGB2, and H2A.H2B Bernues <i>et al.</i> (1986)
HMGB1	Core histones	HMGB1/ H3.H4	Proteins in free solution. Bernues <i>et al.</i> (1983)
HMGB1	Core histones	HMGB1/(H3) ₂	Proteins in free solution. Interaction is between HMGB1 N-terminal, and (H3) ₂ Bernues <i>et al.</i> (1986)
HMGB2	Core histones	HMGB2/(H3) ₂	Proteins in free solution. Interaction is between HMGB2 N-terminal, and (H3) ₂ Bernues <i>et al.</i> (1986)
HMGB1	Core histones	Tetramers and dimers	Proteins in free solution. Complex with HMGB1 not found – potential reasons (i) pH9.5 buffer, and (ii) core histones were extracted by mild means (not acid extraction). Bonne-Andrea <i>et al.</i> (1984)

Table 5.2 Some previous studies of HMGB/core histone interactions

(Page 2 of 2)

HMGB	Histone	Complex	Comment
HMGB1	Octamers Individual core histones	HMGB1/H3	Proteins in free solution. The complex was found in buffers up to 200mM, complex is compromised by greater buffer concentrations. Interaction is between C-terminal of HMGB1 C-terminal, and H3 N-terminal. Only weak interaction of HMGB1 with individual H2A, H2B, H4. Ueda <i>et al.</i> (2004)
HMGB1	Nucleosomes	HMGB1/(H3) ₂	HMGB1 in solution with nucleosomes. Interaction is between N-terminal of HMGB1, and (H3) ₂ . No interaction with H2A, H2B, or H4. Stros (1987)
HMGB1	Nucleosomes	HMGB1/ nucleosome	HMGB1 in solution with nucleosomes Ueda <i>et al.</i> (2004).
HMGB1	Chromatin (linker-histone-depleted)	HMGB1/H3	HMGB1 in solution with chromatin. Interaction is between HMGB1 C-terminal and H3 N-terminal region; most of HMGB C-terminal tail can be deleted without compromising the complex. No binding to H2A, H2B, or H4 in nucleosomes. Watson <i>et al.</i> (2014)
HMGB proteins	Linker and core histones	Possible HMGB/H2A.H2B	Proteins in free solution. HMGB proteins, H2A, and H2B appear to cofractionate on a cation-exchange column (proteins as identified by gel band positions). Zhuang (2011)

5.3 Preparation of nuclear protein subset cePNE4' and investigation of potential complexes within

5.3.1 Summary

Figures 5.1 and 5.2 summarise the main steps applied to prepare the nuclear protein subset cePNE4' and to investigate potential complexes within. To provide complete lysis, the chicken erythrocytes were lysed in three stages, using a progressively lower concentration of detergent at each stage to ensure the nuclei remained intact. The nuclei were washed, depleted of the more mobile proteins (cePNE1' group), washed again, and stored frozen in a number of Falcon tubes for future use. Subsequently, one Falcon tube of nuclei was defrosted, and the nuclei were lysed. Various unwanted molecules were removed by ultracentrifugation and ultrafiltration. Two further nuclear protein subsets (cePNE2' and cePNE3' groups) were removed by precipitation. The remaining proteins, designated the cePNE4' group and containing predominantly HMGB proteins and histones, were subjected to cation-exchange chromatography to investigate for potential complexes. Some potential complexes thus found were investigated by gel filtration chromatography. A key feature was the use of ultracentrifugation to isolate proteins from the DNA and RNA. This relatively mild approach was more likely to preserve histone post translational modifications (PTMs), In contrast, as also discussed in Chapter 4, acidic extraction can degrade PTMs, for example by removing phosphorylation from some histone arginines, lysines, and histidines (Chen *et al.*, 1974; Sut and Biterge, 2017).

5.3.2 Preparation of erythrocyte nuclei

Chilled buffers were prepared as below. Benzamidine hydrochloride degrades in solution so it was added to the buffer immediately before use. The lysis buffers were hypotonic (leading to osmotic pressure within the cell), and contained a detergent to degrade cells walls, and MgCl₂ to maintain the nuclear membrane.

Anti-Clotting Buffer	10% w/v tri-sodium citrate.
Cell Lysis Buffer 1 (pH7.5)	1.25% w/v Triton X100, 10mM NaCl, 10mM Tris-HCl pH7.5, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.
Cell Lysis Buffer 2 (pH7.5)	0.5% w/v Triton X100, 10mM NaCl, 10mM Tris/HCl pH7.5, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.

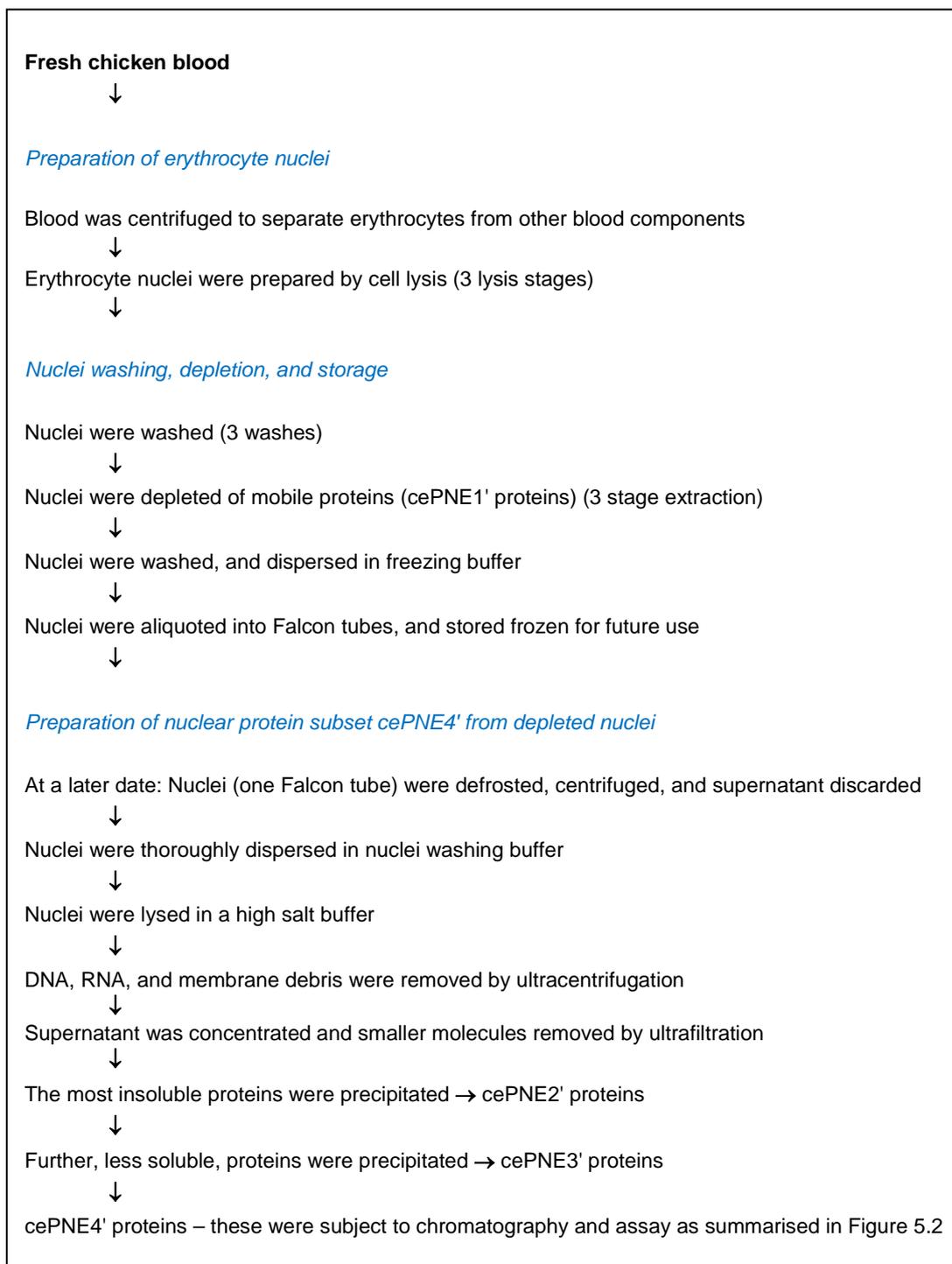


Figure 5.1 Summary of the method for isolating cePNE4' proteins from chicken erythrocytes

Notes:

1. All operations were conducted at ~4°C unless otherwise stated.
2. Where rpm are given, centrifuging was carried out using a 6 place swing-out bucket rotor with max radius ~200mm.
3. See Section 5.3.2 onwards for further details.

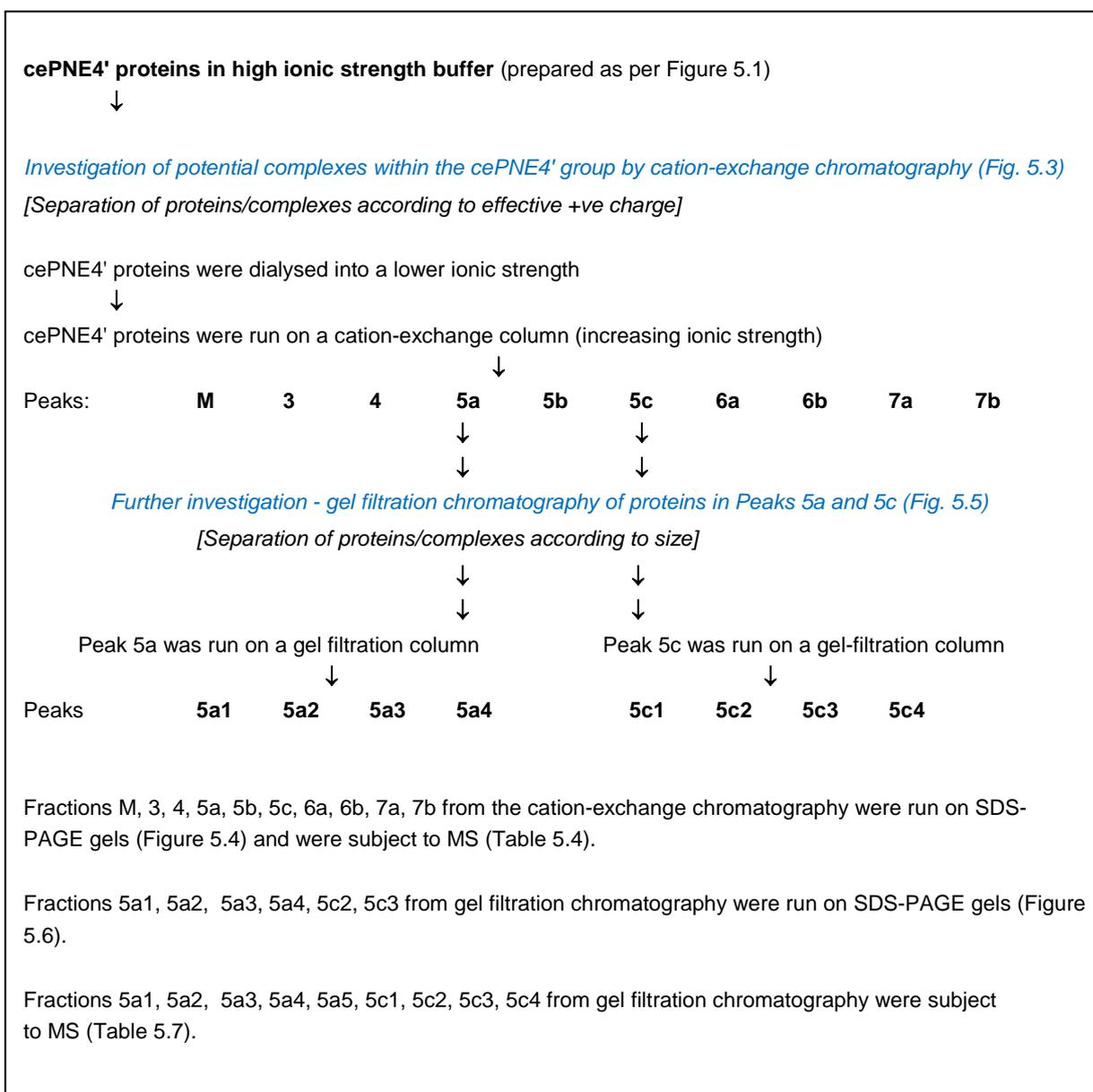


Figure 5.2. Summary of the chromatography applied to the cePNE4' proteins

Notes

1. Operations were conducted at ~4°C unless otherwise stated.
2. Buffers used for chromatography were filtered (0.45µm) to avoid clogging the column.
3. See Section 5.3.5 onwards for further details.

5.3.2 Continued: Preparation of erythrocyte nuclei

4L blood was collected into 400mL Anti-Clotting Buffer in a tray as the chickens passed by, with the tray continually rocked to ensure thorough mixing. The blood was filtered on site through two layers of butter muslin into a container. The blood container was kept cold by embedding in ice during transport to the laboratory.

Back at the laboratory, the blood was centrifuged in four 1L bottles at 2000rpm for 20min to separate the erythrocytes from the other blood components. As with all centrifuging, this was conducted at 4°C. The blood separated into three layers; the top layer (plasma) was removed with a water pump and discarded. The next layer, a "skin" of buffy coat (containing white blood cells and platelets), was moved to the side of the bottle with a spatula to assist removal, and was similarly removed. The loose pellets of intact red blood cells in each bottle were agitated with a rod to release cells stuck to the bottom and were combined to form ~1L loose slurry. The centrifuging was repeated. Residual buffy coat was removed as before. The remaining layer contained erythrocytes (red blood cells).

The red blood cells were lysed in three stages. For the first cell lysis stage, the red cells were partitioned equally among three 1L containers, making sure to stir with a rod to remove cells stuck to the bottom of the containers. A small quantity of Cell Lysis Buffer 1 was added to the red blood cells in each container and the mixture was shaken thoroughly, but not violently, to ensure dispersion. The containers were then filled with Cell Lysis Buffer 1 and shaken to produce a uniform suspension. The contents of each container was emptied into a separate 2L Duran bottle and more Cell Lysis buffer 1 was added, to make up three full 2L Duran bottles. The slurries of red blood cells were briskly stirred for one hour with magnetic fleas, and then were centrifuged at 2500rpm for 20min in six 1L centrifuge bottles. Supernatants were discarded, leaving loose pellets.

For the second stage, the above pellets were mixed with an equal volume of Cell Lysis Buffer 2 and the nuclei were homogenised by shaking and squeezing in and out with a submerged plastic pipette. The homogenates were partitioned equally into two 1L centrifuge bottles using a 50ml syringe with 22-gauge needle. The bottles were filled with Cell Lysis Buffer 2, shaken to produce a uniform suspension, and centrifuged at 2000rpm for 15min. The supernatant was discarded.

The third stage was a repeat of the second stage.

5.3.3 Nuclei washing, cePNE1' depletion, and storage of frozen nuclei

Chilled buffers were prepared as below. The extraction buffer was mildly hypertonic (leading an outflow of liquid from the nuclei) and contained MgCl₂ to maintain the nuclear membrane.

Nuclei Washing Buffer (pH7.5)	10mM NaCl, 10mM Tris/HCl pH7.5, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.
cePNE1'-extraction Buffer (pH6.8)	80mM KCl, 40mM H ₂ KPO ₄ , 40mM HK ₂ PO ₄ 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.
Nuclei Freezing Buffer (pH6.8)	10mM KCl, 5mM H ₂ KPO ₄ , 5mM HK ₂ PO ₄ 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride. All components are dissolved in 27% w/v glycerol.

Nuclei prepared as described in Section 5.3.2 were washed three times. The loosely pelleted nuclei were combined with an equal volume of Nuclei Washing Buffer added to each centrifuge bottle. The bottles were gently shaken, and the nuclei were homogenised by squeezing in and out with a submerged plastic pipette. They were then transferred to two new 1L bottles via a syringe through a 22-gauge needle. The content of each bottle was made up to 1L with Nuclei Washing Buffer, shaken thoroughly, and centrifuged at 2000rpm for 15min. Most of the supernatant was discarded leaving ~100mL of loosely pelleted nuclei in each centrifuge bottle. The washing procedure was repeated twice more so that the nuclei were almost white, indicating most of the red haemoglobin had been removed.

cePNE1' proteins (more mobile, loosely bound nuclear proteins) were extracted in three stages. First, the loosely pelleted nuclei were combined with an equal volume of cePNE1'-extraction Buffer added to each centrifuge bottle. They were gently shaken and homogenised by squeezing in and out with a submerged plastic pipette and transferred to two 0.25L centrifuge bottles via a syringe with a 22-gauge needle. The suspensions were centrifuged at 2000rpm for 15min. The supernatant was discarded, leaving ~100mL of loosely pelleted nuclei in each bottle.

For the second extraction, an equal volume of cePNE1'-extraction Buffer was added to the nuclei. The mixture was homogenised, transferred to two clean 0.25L bottles, and centrifuged at 2000rpm for 15min. This time, the supernatant containing cePNE1' proteins and the nuclei were both retained. The third extraction was a repeat of the second. By this time the nuclei were almost white, since most of the red haemoglobin had been removed.

The supernatants from the second and third extractions were combined and retained for future study. The depleted nuclei were dispersed by stirring with three volumes of Nuclei Freezing Buffer, increasing the total volume of nuclei in suspension to about 800mL. The nuclei, now in 20% glycerol, were gently shaken and homogenised by squeezing in and out with a submerged plastic pipette, and were aliquoted, via a syringe with a 22-gauge needle, into 20 Falcon tubes each containing ~40mL. The tubes were frozen at -80°C.

5.3.4 Preparation of cePNE4' proteins from frozen depleted nuclei

Chilled buffers were prepared as below. The absence of Mg⁺⁺ ions from the Nuclei Lysis Buffer was designed to weaken the nuclear membrane, and the strongly hypertonic buffer was intended to collapse the nuclei, and to dissociate the chromatin.

Nuclei Washing Buffer (pH6.8)	10mM KCl, 10mM Tris/HCl pH6.8, 3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.
Nuclei Lysis Buffer (pH6.8)	2500mM KCl, 250mM H ₂ KPO ₄ , 250mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride.
Precipitation Buffer (pH6.8)	2000mM KCl, 1000mM H ₂ KPO ₄ , 1000mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride.
PNE2'/octamer Dissolving Buffer (pH6.8)	2000mM KCl, 200mM H ₂ KPO ₄ , 200mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride

One Falcon tube containing ~40mL of nuclei in Nuclei Freezing Buffer was thawed in the cold room, balanced with a Falcon tube containing buffer, and centrifuged at 1500rpm for 15min. The supernatant was discarded, and the sedimented nuclei pellet (~13mL) was mixed with an equal volume of Nuclei Washing Buffer and dispersed by gentle swirling. The mixture was made up to 80mL with Nuclei Washing Buffer and stirred. The nuclei in suspension were then thoroughly dispersed by transferring via a syringe with a 22-gauge needle into a separate stirred beaker.

The dispersed nuclei were lysed in a high salt solution. By running the nuclei down the side of the tube, 13mL samples of nuclei (6 samples in all), taken from the still-stirring nuclei suspension, were carefully layered onto 52 mL of Nuclei Lysis Buffer in each of six ultra-centrifuge tubes (70mL polycarbonate tubes for use with a Beckman 45Ti rotor). Each tube was sealed and quickly shaken (tube cap downward) to evenly mix the nuclei and Nuclei Lysis Buffer.

The six centrifuge tubes were centrifuged at 100,000g overnight (~12hrs), and the pellets containing DNA, RNA, and nuclear membrane debris were discarded. The supernatant from the six tubes was pooled and concentrated for 4hrs down to 25mL using a cooled (~6°C) 400mL Amicon concentrator pressurised by nitrogen at 2bar. Water and smaller molecules (such as protein fragments) flowed through the concentrator ultrafiltration membrane (10kDa MWCO). The reduced volume facilitated subsequent steps, which involved precipitating the less soluble nuclear proteins.

The 25mL concentrated sample of proteins was stirred on ice, and 22.5mL of Precipitation Buffer was added very slowly, in a dropwise fashion, increasing the phosphate concentrations (see Chapter 8 Table 8.1). A cloudy precipitate of proteins formed after a while; the precipitate was pelleted by centrifugation at 10,000rpm for 5min (Beckman Ti70 rotor). The supernatant was retained for a second precipitation stage. The pelleted proteins, designated the cePNE2' group, were dissolved in 10mL PNE2'/octamer Dissolving Buffer, mixed with 10mL of the same buffer in 40% glycerol, and frozen at -80°C for future use.

About 47ml of supernatant collected as above was set to stir on ice. 37.5mL of Precipitation Buffer was added, dropwise as before, further increasing phosphate concentrations (see Table 8.1 for details). The solution was centrifuged at 3000rpm for 40min. A second precipitate less dense than the solution was formed; it floated as a soft crust on top of the solution. The solution, containing proteins designated the cePNE4' group, was sucked from under the crust using a large-capacity syringe. This group was used in the investigation of potential complexes as described later in this chapter. The floating crust, designated the cePNE3' group of proteins, was dissolved in 10mL PNE2 'octamer Dissolving Buffer, mixed with 10mL of the same buffer in 40% glycerol, and frozen at -80°C for future use.

Table 5.3 summarises the status of the HMGB and histone proteins throughout the processes leading up to the isolation of the cePNE4' group. Although part of the HMGB and histone populations were lost during these processes, HMGB and core histone proteins were represented in the cePNE4' group due to their initial abundance. Linker histones were also well represented due to their high solubility, including H5 since this variant is prevalent in non-replicating chicken erythrocytes.

5.3.5 Investigation of potential complexes within the cePNE4' group by cation-exchange chromatography.

The cePNE4' proteins, in about 80mL of solution containing 2M KCl, ~1.5M phosphate, were dialysed into the Cation-exchange Buffer A defined in Figure 5.3 - this had a low ionic strength, closer to native nuclei conditions. In these conditions, the octamers would have dissociated into dimers and tetramers. Tetramers are less soluble than dimers, and some will have precipitated (Zhuang, 2011) and would have been removed by the 0.45 μ m filtration applied to all samples and buffers prior to chromatography. Dialysis expanded the solution to 175mL, which was reduced to about 100mL using an Amicon concentrator (10kDa MWCO ultrafiltration membrane); this also eliminated smaller molecules below 10kDa MW. The cePNE4' solution was run on a cation-exchange column, using 5 cation-exchange columns, with a linear ionic strength gradient as defined in Figure 5.3a. Later, fresh cePNE4' solution was prepared as before and run again on a cation-exchange column, under the same conditions as before except that, to improve the resolution between peaks, the flow rate was reduced from 3mL/min to 2mL/min and 8 columns were used instead of 5. The chromatograms are shown in Figures 5.3a and 5.3b.

Table 5.3. Status of HMGBs and histones in the nuclear protein sets

Stage	Principal components of the solution.		Status of HMGBs and histones
	KCl mM	Total phosphate mM	
Depletion of cePNE1 proteins from intact nuclei.	90	90	Part removal of the HMGB protein population due to their mobility in the nucleus (Scaffidi <i>et al.</i> , 2002). However, some HMGBs remained in nuclei. Most linker and core histones remained in nuclei bound to chromatin.
Lysis of nuclei, ultracentrifugation, and ultrafiltration.	2000	400	HMGB proteins remained in solution. DNA, RNA and membrane debris were removed by ultracentrifugation, leaving soluble proteins including histone octamers and linker histones in solution.
First precipitation - removal of cePNE2 proteins.	2000	1160	HMGB proteins remained in solution. Histone octamers and linker histones remained in solution.
Second precipitation – removal of cePNE3 proteins.	2000	1520	Most histone octamers were precipitated, but some may have remained in solution. HMGB proteins remained in solution. Linker histones remained in solution.
Dialysis of cePNE4 proteins into Cation-Exchange Buffer A	100	100	Any histone octamers would have dissociated into histone dimers and tetramers, leading to dimers and some tetramers in solution. Tetramers are less soluble than dimers and some would have been precipitated and filtered out during preparation for chromatography. HMGB proteins remained in solution. Linker histones remained in solution.

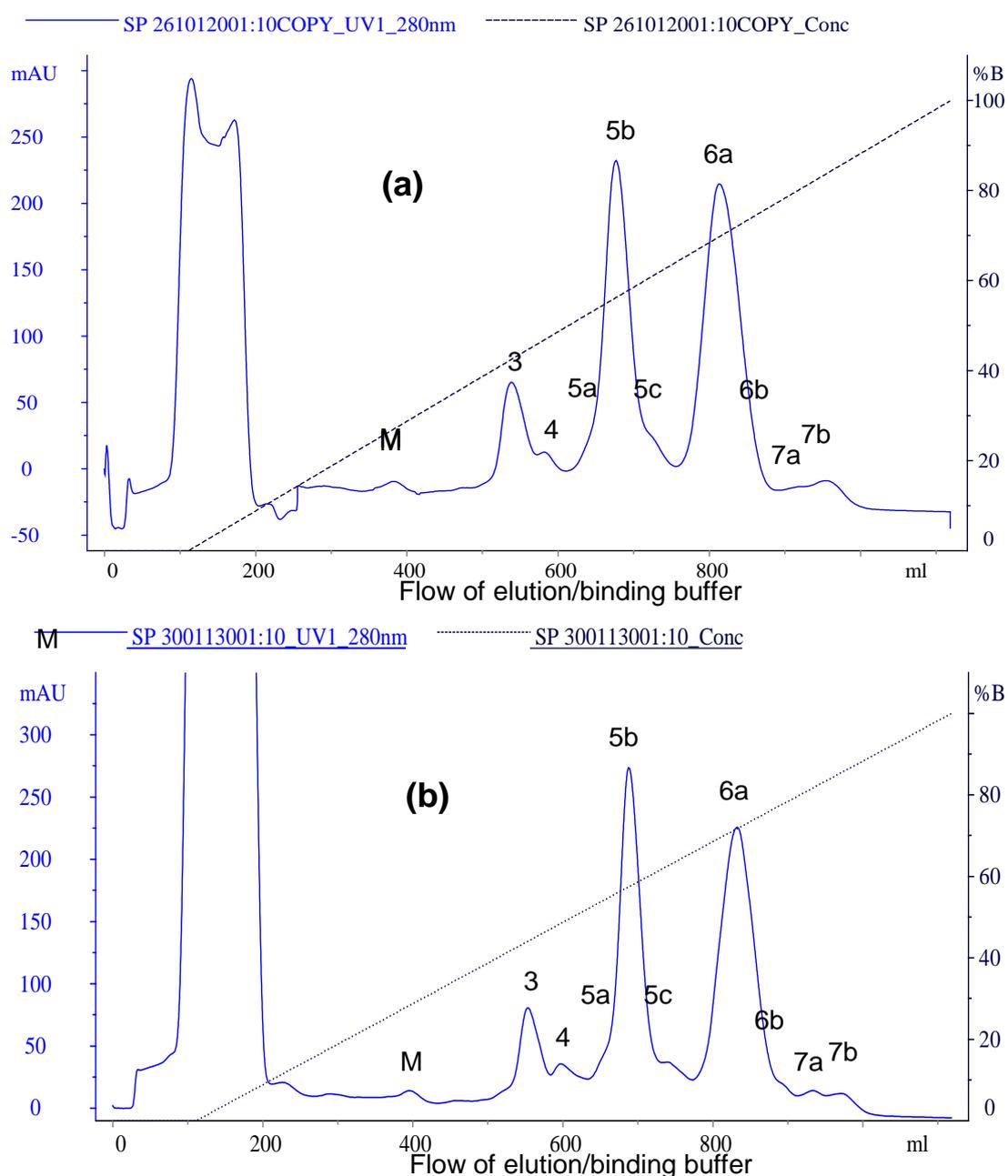


Figure 5.3 Cation-exchange chromatograms of cePNE4' proteins

(a) Initial run using 5 columns (HiTrap SP Sepharose FF 5mL) and flow rate 3mL/min

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B:

- at start of elution Buffer A 100%, Buffer B 0%; - at end of elution Buffer A 0%, Buffer B 100%.

Cation-exchange Buffer A 100mM KCl, 50mM H₂KPO₄, 50mM HK₂PO₄, pH6.8

Cation-exchange Buffer B 1100mM KCl, 50mM H₂KPO₄, 50mM HK₂PO₄, pH6.8

Fraction size 10mL

Chromatogram identifier SP261012001

Fractions subject to assay M, 3, 4, 5a, 5b, 5c, 6a, 6b, 7a, 7b (SDS-PAGE)

(b) Revised run to improve resolution using 8 columns and flow rate 2mL/min

Other parameters As Run (a)

Chromatogram identifier SP300113001

Fractions subject to assay M, 3, 4, 5a, 5b, 5c, 6a, 6b, 7a, 7b (MS)

5.3.5 Continued.

Investigation of possible complexes within the cePNE4' group by cation-exchange chromatography.

Measurement of UV absorption at 278nm of a sample from the biggest peak (Peak 5b) indicated a concentration of 0.4mg/mL.

Peaks in the chromatograms were marked M, 3, 4, 5a, 5b, 5c, 6a, 6b, 7a, and 7b (letters after these numbers indicated sub-divisions of a combined peak). The associated fractions from the 5 column chromatography (Figure 5.3a) were run on SDS-PAGE gels, shown in Figures 5.4(a) - 5.4(d).

The corresponding fractions from the 8 column chromatography (Figure 5.3b) were dialysed into ammonium bicarbonate (ABC), freeze dried, and sent to Professor Dixon's Group at the University of Sheffield for MS. His group kindly provided the MS results which are summarised in Table 5.4. Based on the gel positions and MS results the prominent proteins in each Figure 5.3 fraction are summarised in Table 5.5. The HMGB proteins were strongly represented in Peak 5 together with the dimer histones H2A and H2B, despite significantly different net charges (Table 5.6). Consequently, the Peak 5 proteins were further investigated by gel filtration chromatography.

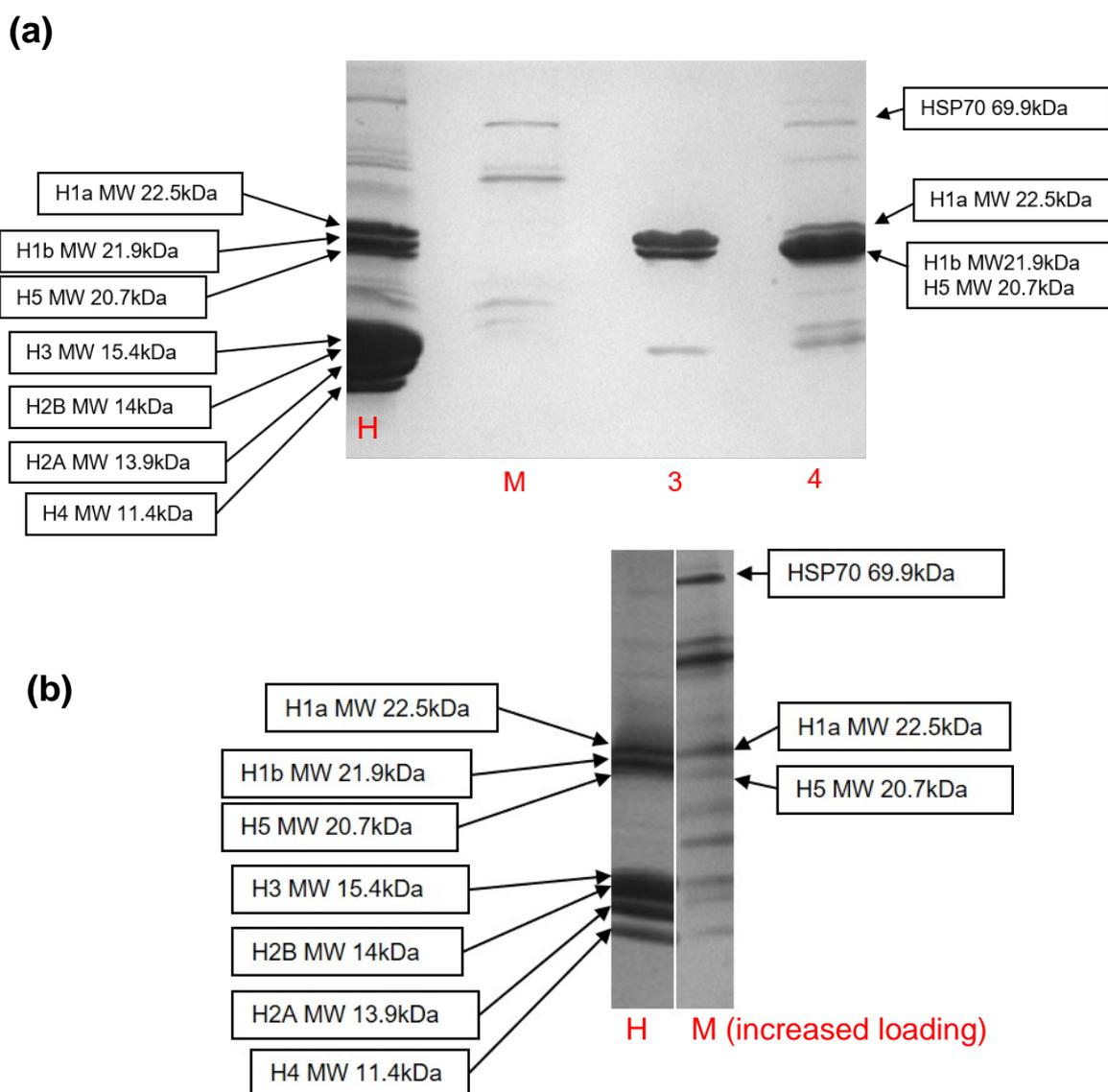


Figure 5.4 (Page 1 of 2) SDS-PAGE gels of peak fractions from cation-exchange chromatography of cePNE4' proteins

(a) Gel of fractions M, 3, 4

(H is the Histone Standard – see Chapter 4)

The chromatogram is shown in Figure 5.3a

Stain: Coomassie Blue (although photographed in black and white)

Stacking gel 5% acrylamide; Resolving gel 20% acrylamide; Running buffer glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol: as described in Chapter 3

(b) Gel of fraction M with increased loading to show more protein bands

Other parameters as Gel (a)

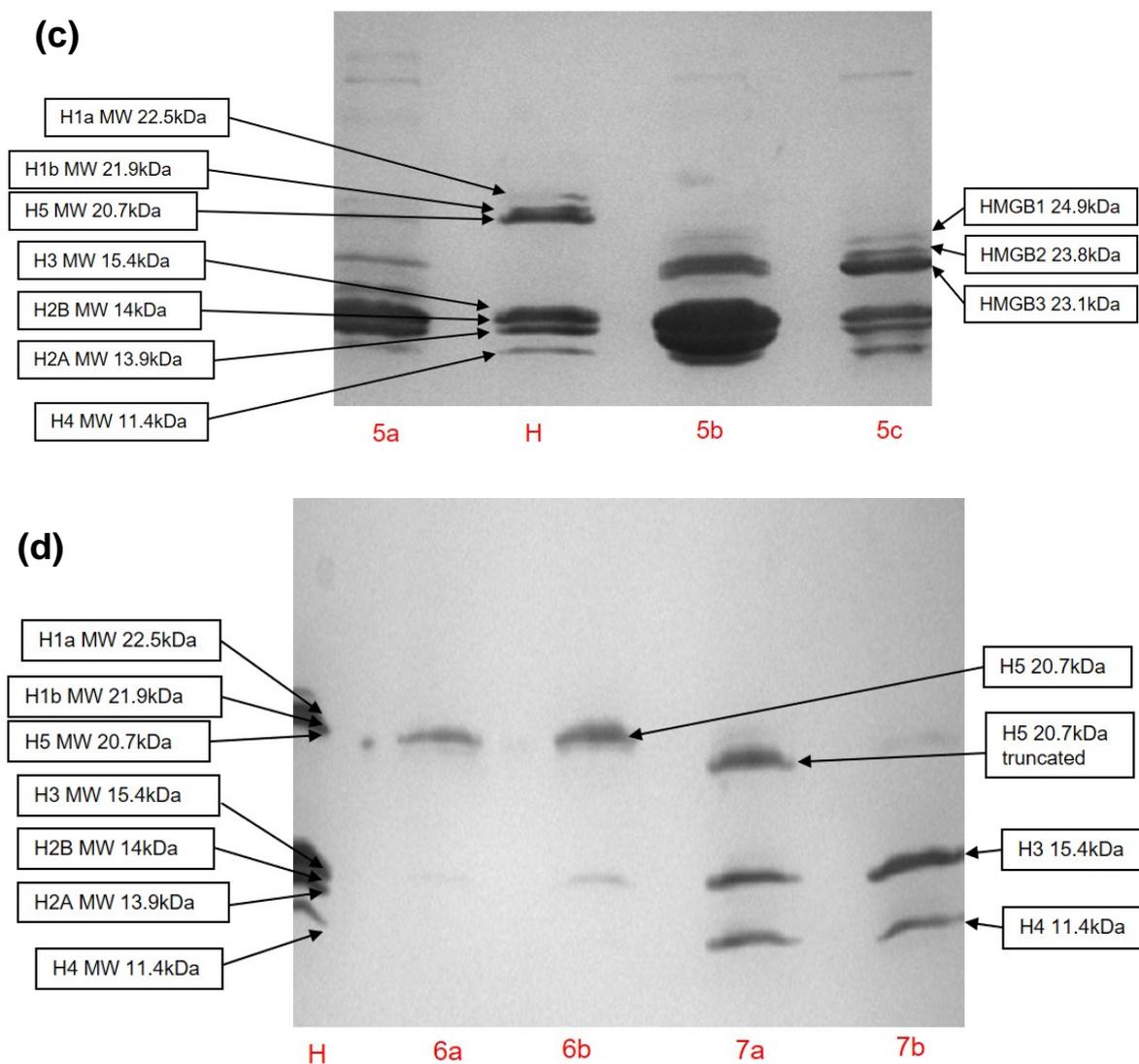


Figure 5.4 (Page 2 of 2) SDS-PAGE gels of peak fractions from cation-exchange chromatography of cePNE4' proteins

(c) Gel of fractions 5a, 5b, 5c

(H is the Histone Standard – see Chapter 4)

The chromatogram is shown in Figure 5.3a

Stain: Coomassie Blue (although photographed in black and white)

Stacking gel 5% acrylamide; Resolving gel 20% acrylamide; Running buffer glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol: as described in Chapter 3

(d) Gel of fractions 6a, 6b, 7a, 7b

In this gel, the lanes H (histone standard) and 7b were distorted by gel overheating

Other parameters as Gel (c)

Table 5.4(a) Summary of MS results for fractions from the cation-exchange chromatogram of cePNE4' proteins (Figure 5.3b)

(Page 1 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
M	H1.11R	P08288	1	42	1	44
	H1.11L	P08287	2	36	2	38
	H1.01	P08284	3	35	3	37
	H1	XP425456	4	31	4	33
	H1.03	P08285	5	34	5	35
	HSP70	P08106	6, 7	25	6	26
	H5	P02259	13	21	15	23
3	H1.03	P08285	1	53	1	74
	H1.11R	P08288	2	56	2	74
	H1.11L	P08287	3	63	3	71
	H1.01	P08284	4	46	4	48
	H1	P09987	-	-	5	54
	H5	P02259	5	35	6	40
	HSP70	P08106	-	-	10	8
4	H1	P09987	1	26	13	22
	H1	XP425456	-	-	1	38
	H1.01	P08284	2	45	2	50
	H1.11L	P08287	3	45	3	46
	H1.10	P08286	4	38	4	41
	H1.11R	P08288	5	40	5	54
	H1.03	P08285	6	40	6	51
	H5	P02259	7	27	7	37
	H2A	Q92069	8	30	8	33
	H2A.Z	P02272	13	33	17	34
	H2B	P0C1H3	9	32	10	35
H4	P62801	10	31	11	37	
HSP70	P08106	12	7	14	8	

Table 5.4(a) Summary of MS results for fractions from the cation-exchange chromatogram of cePNE4' proteins (Figure 5.3b)

(Page 2 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
5a	MS results were available only after gel filtration chromatography – Table 5.7					
5b	H5	P02259	-	-	1	37
	H2A	Q92069	-	-	2	39
	H4	P62801	-	-	3	70
	H2B	P0C1H3	-	-	4	46
	H2B VII	P0C1H5	-	-	6	46
	H2A.Z	P02272	-	-	9	34
	HMGB3	NP990626	-	-	10	30
	HMGB2	P26584	-	-	13	34
	HMGB1	NP990233	-	-	15	12
	H3	Q92068	-	-	19	32
5c	MS results were available only after gel filtration chromatography – Table 5.7					
6a	H2B	P0C1H3	1	44	1, 2, 12	31, 39, 25
	H2B VII	P0C1H5	-	-	7	46
	H2A	Q92069	2	38	3	39
	H2A.X	XP004947974	4	24	6	25
	H2A.Z	P02272	-	-	11	33
	H5	P02259	3	33	4	36
	H4	P62801	5	57	8	62
	HMGB3	NP990626	-	-	13	12
6b	H5	P02259	1	35	1	37
	H4	P62801	2	49	2	53
	H2B	P0C1H3	3	21	4	23

Table 5.4(a) Summary of MS results for fractions from the cation-exchange chromatogram of cePNE4' proteins (Figure 5.3b)

(Page 3 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
7a	H4	P62801	1	64	1	68
	H5	P02259	2	37	2	53
	H2A IV	P02263	3	29	4	15
	H3	P84229	4	21	6	23
7b	H4	P62801	1	63	1	66
	H5	P02259	2	33	2	36
	H3	P84229	3	34	3	37
	H2A	Q92069	4	23	-	-
	H2A.Z	P02272	5	16	7	18

Notes

1. Some proteins have not been listed due to low MS score and/or low sequence coverage, or obsolete sequence, or not *Gallus gallus*.

2. MS results were provided by Professor Dickman's group, Sheffield University, with the MS parameters as per Table 5.4(b).

Table 5.4(b) Parameters used to obtain MS results for fractions from chromatography of cePNE4' proteins

Parameters for MS using the NCBI sequence database

Search title : Submitted from JB CEX all chick ncbi nr 250313 by Mascot Daemon on DELL_T3400
 Example of MS data file: C:\Documents and Settings\Administrator\My Documents\MS Data\2013\March 13\amazon\JB CEX 3 140313_GC2_01_158.d\JB CEX 3 140313_GC2_01_158.mgf
 Database : NCBI nr 10.5 (11049075 sequences; 3763785421 residues)
 Taxonomy : bony vertebrates (1131674 sequences)
 Timestamp : 25 Mar 2013 at 10:53:59 GMT

Parameters for MS using the SwissProt sequence database

Search title : Submitted from JB CEX all chick swissprot 250313 by Mascot Daemon on DELL_T3400
 Example of MS data file: C:\Documents and Settings\Administrator\My Documents\MS Data\2013\March 13\amazon\JB CEX 3 140313_GC2_01_158.d\JB CEX 3 140313_GC2_01_158.mgf
 Database : SwissProt 10.5m5 (516615 sequences; 181923321 residues)
 Taxonomy : bony vertebrates (80741 sequences)
 Timestamp : 25 Mar 2013 at 12:22:54 GMT

Common parameters for MS

Enzyme : Trypsin/P
 Variable modifications : Acetyl (Protein N-term), Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 0.6 Da
 Fragment Mass Tolerance : ± 0.4 Da
 Max Missed Cleavages : 2
 Instrument type : ESI-TRAP

Fractions representing each peak in the chromatogram Figure 5.3(b) were dialysed into ammonium bicarbonate (ABC) and freeze-dried before dispatch to Sheffield for MS. See Chapter 3 Section 3.8 for details of the further preparation for MS, and the MS equipment.

The MS generated sets of sequence fragments which were matched against either NCBI nr or SwissProt databases by MASCOT software to identify likely proteins, and the MS Score provided by MASCOT was an indicator of how likely the identification was correct. The data provided by MASCOT included the sequence fragments found by MS. This data was too extensive to include in this thesis. Instead, two parameters have been provided in the Summary table. First, the results are shown in the order of the Score (gaps in the order are due to proteins other than *Gallus gallus*). Second, the sequence fragments were manually marked up against the full sequence, to provide a percentage of the protein sequence covered. For the unpublished MS results (SwissProt database) summarised above, the sequence coverage includes N-terminal K or R residues for each sequence fragment which were found by MASCOT matches to the full protein sequence. These N-terminal residues were not part of the sequence fragments found by MS, since trypsin cuts residue sequences downstream of a K or R residue.

Table 5.5 Proteins in each peak fraction of the cation-exchange chromatogram of cePNE4' proteins (Figure 5.3(a) and (b)).

(Based on a review of MS-PAGE gels and MS results - Figures 5.4 and 5.6; Tables 5.4 and 5.7)

	Increasing ionic strength applied to displace protein or protein complex →									
	Peak									
Protein	M	3	4	5a	5b	5c	6a	6b	7a	7b
HMGB1				x	x	x				
HMGB2				x	X	X				
HMGB3				x	X	X	x			
H2A			x	X	X	X	x	x	x	x
H2B			x	X	X	X	x	x	x	
H3				x	x	x		x	X	X
H4			x	x	X	X	x	x	X	X
H5	x	x	x	x	x	x	X	X	X	x
H1a/b	X	X	X	x						
HSP70	x	x	x	x						

Notes

1. X indicates the stronger SDS-PAGE bands/high scoring MS results. Some proteins as above were identified by MS, but do not show on the gel, and vice versa. A few low scoring proteins other than HMGB and histone proteins were also identified by MS but are not shown above.

2. Histone H1 variants falls into two groups which run at different speeds on a gel: *Group H1a* have 224-225 residues, *Group H1b* have 218 to 220 residues.

Table 5.6 Some features of chicken HMGB and histone proteins

Protein	Features	N	MW (kDa)	Net charge (e)	
				pH7.1	pH6.8
HMGB1	+ve N-terminal tail, Box A, Box B, and connecting region. -ve C-terminal tail.	215	24.9	-4.6	-3.9
HMGB2	As above	207	23.8	+3.6	+4.4
HMGB3	As above	202	23.1	+2.0	+2.9
H2A	+ve N-terminal tail, globular region, and C-terminal tail.	129	13.9	+17.5	+18.0
H2B	+ve N-terminal tail and globular region (very short separate C-terminal).	126	14.0	+18.5	+19.0
H3	As above	136	15.4	+20.3	+20.6
H4	As above	104	11.4	+18.3	+18.6
H5	+ve throughout, especially C-terminal half. N-terminal (27bp), globular region (28 – 96bp), and region with unknown structure (97-190bp). 30% of residues are hydrophobic.	190	20.7	+61.4	+61.9
H1	+ve throughout, especially C-terminal half. SwissModel and NESG theoretical models covering Residues 36 – 112 provide a similar globular structure to H5. 37% of residues are hydrophobic.	218	21.9	+60.9	+60.9

Notes

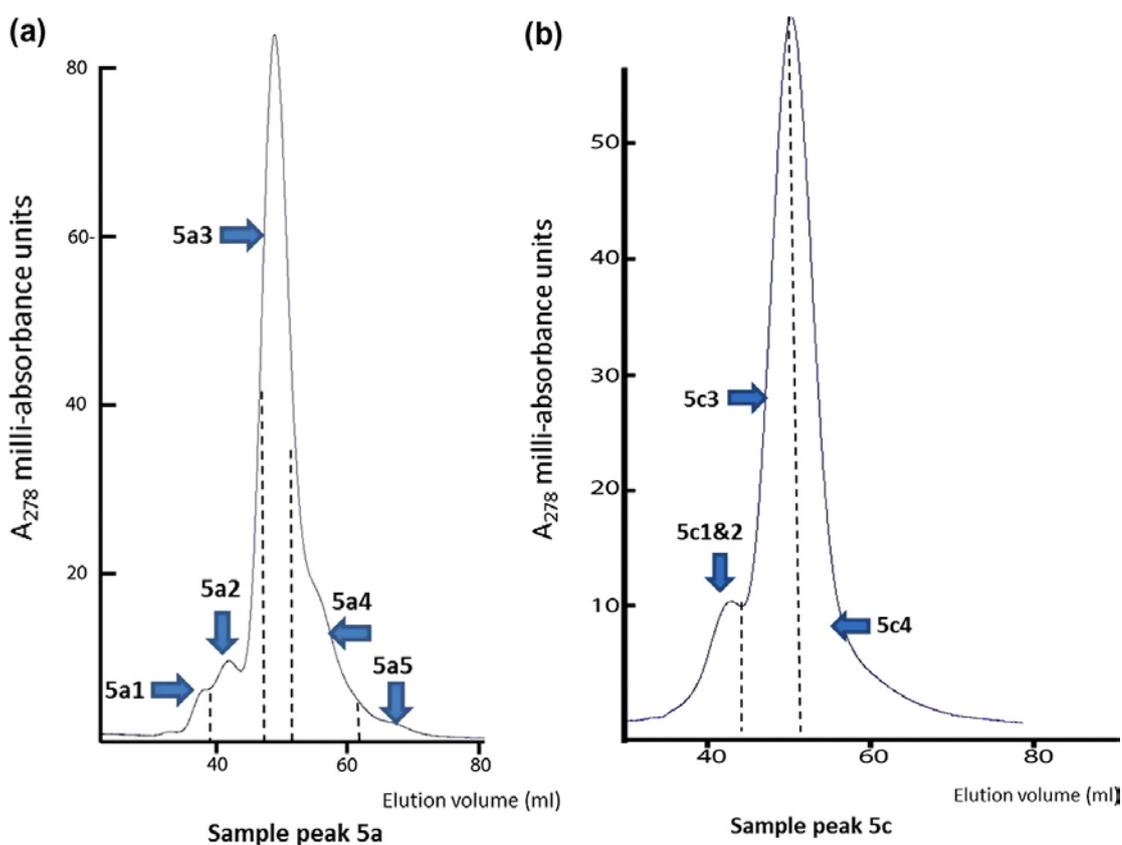
N: number of residues

HMGB1, 2 and 3 details are based on UniprotKB Q9YH06, P26584, and P40618 respectively. Core histone details are based on P02263, P0C1H5, P84229, and P62801 respectively. H5 and H1 details are based on P02259 and P09987 respectively.

Net charges calculated using Scripps Protein Calculator (2017) are only approximate since the calculator assumes all residues have pKa values that are equivalent to the isolated residues; for a folded protein this is not valid. In event, only part of the charge will be exposed to another molecule.

5.3.6 Further investigation - gel filtration chromatography of proteins in Peaks 5a and 5c

Cation-exchange Fractions 5a and 5c from the 8 column chromatography (Figure 5.3b) were each concentrated to 4 ml and individually run on a gel-filtration column. The Gel Filtration Buffer (as defined in Figure 5.5) was selected to match the mean elution conditions which existed for these two fractions during the cation-exchange chromatography. Each gel filtration chromatogram had a number of peaks (Figure 5.5). The fractions comprising these peaks were run on SDS-PAGE gels (Figure 5.6), and subjected to MS at Sheffield University, the results are summarised in Table 5.7. Based on the gel positions and MS results, the prominent proteins in each gel filtration fraction are summarised in Tables 5.8 and 5.9.



The above plots have previously been published in Zhuang *et al.* (2014)

Figure 5.5 Gel filtration chromatograms for proteins in Fractions 5a and 5c from cation-exchange chromatography of cePNE4' proteins (Figure 5.3b)

(a) Fraction 5a gel filtration chromatogram (identifier: S100 Peak Fraction 050213001)

Gel filtration Buffer 700mM KCl, 50mM H_2KPO_4 , 50mM HK_2PO_4
2.5mM benzamidine hydrochloride.

pH 6.8

Gel Filtration column Hiprep 16/60 Sephacryl S-100 HR

Total column volume 120mL

Fraction size 10mL

Flow rate 0.5mL/min

Fractions subject to assay 5a1, 5a2, 5a3, 5a4, 5a5

(b) Fraction 5c gel filtration chromatogram (identifier: S100 Peak Fraction 050213002)

Parameters as Chromatogram (a)

Fractions subject to assay 5c1, 5c2, 5c3, 5c4

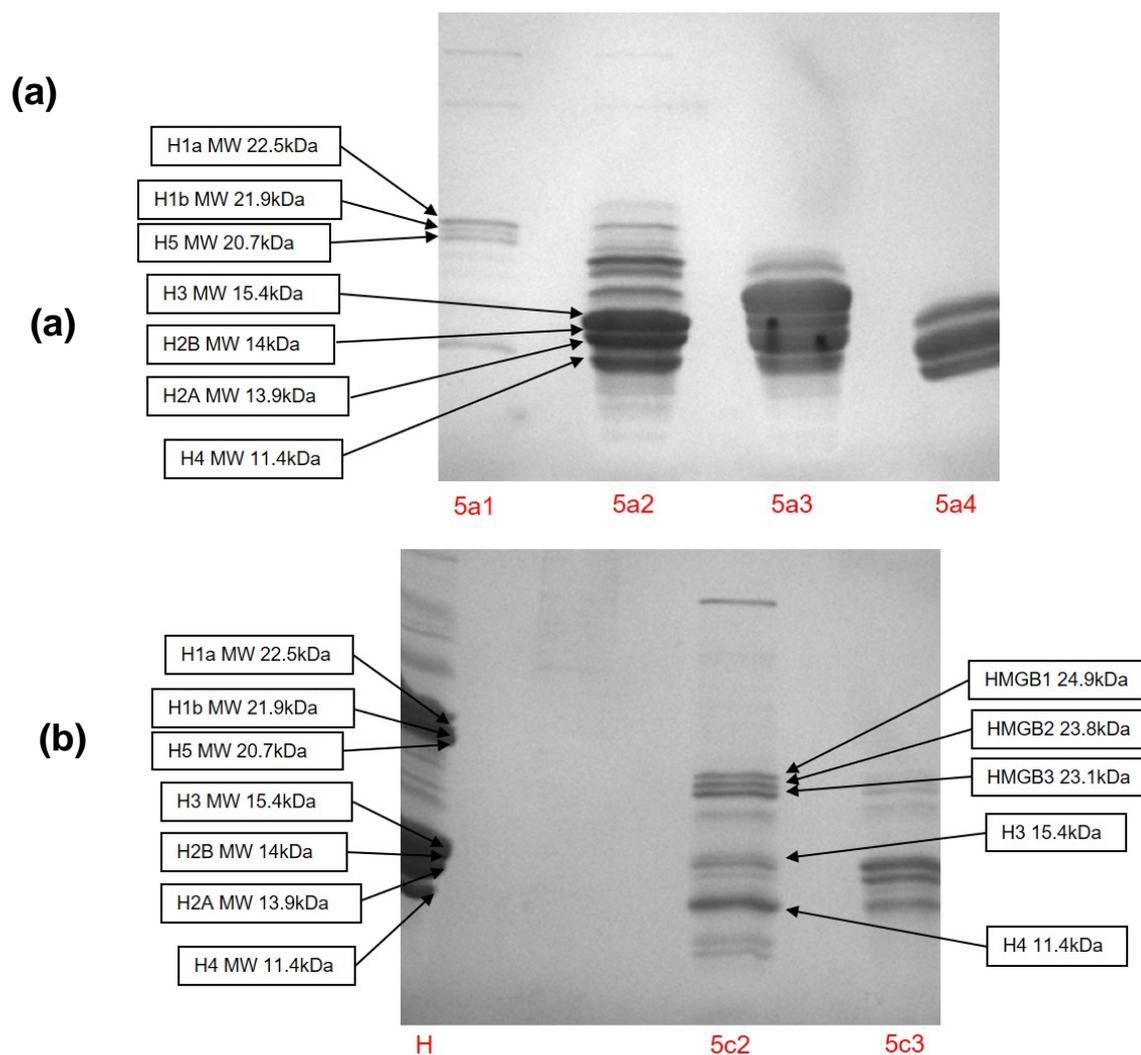


Figure 5.6 SDS-PAGE gels of peak fractions from gel filtration chromatography of Fractions 5a and 5c from cation-exchange chromatography of cePNE4' proteins

(a) Gel of fractions 5a1, 5a2, 5a3, 5a4

(H is the Histone Standard – see Chapter 4)

The chromatogram is shown in Figure 5.5(a)

Stain: Coomassie Blue (although photographed in black and white)

Stacking gel 5% acrylamide; Resolving gel 20% acrylamide; Running buffer glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol: as described in Chapter 3

(b) Gel of fractions 5c2, 5c3

The chromatogram is shown in Figure 5.5(b)

Other parameters as Gel (a)

Table 5.7 Summary of MS results for fractions 5a1 – 5a5 and 5c1 – 5c4 from gel filtration chromatography applied to fractions 5a and 5c from the cation-exchange chromatography applied to cePNE4' proteins (Page 1 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
5a1	H1.01	P08248	-	-	2	23
	H1.11R	P08288	-	-	3	20
	HP1B3	Q5ZM33	-	-	5	17
	H4	P62801	-	-	7	33
	H3	P84229	-	-	12	21
5a2	H1.11R	P08288	-	-	1	27
	H1.01	P08248	-	-	2	25
	H1	P09987	-	-	3	24
	H4 VIII	P70081	-	-	6	67
	HSP70	P08106	-	-	7	22
	H2B	P0C1H3	-	-	8	44
	H1.X	XP425158	-	-	13	18
	H2A	Q92069	-	-	15	24
	HMGB3	NP990626	-	-	16	19
	HMGB1	NP990233	-	-	18	28
	HMGB2	P26584	-	-	21, 23	12, 21
	H3	P84229	-	-	24	34
	H5	P02259	-	-	25	16
	HP1B3	Q5ZM33	-	-	26	8
5a3	H2A	Q92069	-	-	1, 5	48, 25
	H2B	P0C1H3	-	-	2, 6	67, 49
	H5	P02259	-	-	3	38
	H2A.V	P02272	-	-	4	34
	H2B VII	P0C1H5	-	-	7	67
	H4 VIII	P70081	-	-	8	60
	H1.11R	P08288	-	-	9	26
	HMGB3	NP990626	-	-	11	12

Table 5.7 Summary of MS results for fractions 5a1 – 5a5 and 5c1 – 5c4 from gel filtration chromatography applied to fractions 5a and 5c from the cation-exchange chromatography applied to cePNE4' proteins (Page 2 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
5a4	H5	P02259	-	-	1	53
	H2B	P0C1H3	-	-	2	34
	H2A	Q92069	-	-	4	19
5a5	H2A.V	P02272	-	-	8	23
	H5	P02259	-	-	9	36
	H2B	P0C1H3	-	-	11, 14	44, 32
	H2A	Q92069	-	-	12	39
5c1	H2A.V	P02272	1	20	1	23
	H4	P62801	2	46	2	53
	H2B	P0C1H3	3	19	3	20
5c2	H5	P02259	1	29	1	33
	H2B	P0C1H3	2	55	2, 4	59, 40
	H2A	Q92069	3	36	3	39
	H4	P62801	4	52	6	59
	H2B VII	P0C1H5	5	55	7	59
	HMGB2	P26584	6	26	8, 10	12, 27
	HMGB3	P36194	7	12	9	14
	H3	Q92068	8	24	11	29

Table 5.7 Summary of MS results for fractions 5a1 – 5a5 and 5c1 – 5c4 from gel filtration chromatography applied to fractions 5a and 5c from the cation-exchange chromatography applied to cePNE4' proteins (Page 3 of 3)

Peak	Protein <i>Gallus gallus</i>	Identifier	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)		Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)	
			In order of MS score	% sequence coverage	In order of MS score	% sequence coverage
5c3	H5	P02259	1	35	1	39
	H2A	P02263	2	42	2	46
	H2B	P0C1H3	3	74	3, 4	74, 55
	H2A.X	XP004947974	4	32	7	30
	H2A.V	P02272	5	20	8	23
	H4	P62801	6	59	9	23
	H3	P84229	7	20	10, 11	16, 24
	H2B VII	P0C1H5	-	-	5	75
5c4	H5	P02259	1	31	1, 15	39., 24
	H2B	P0C1H3	2	41	4, 8	32, 25
	H2A	Q92069	3	36	7	39
	H2A.V	P02272	4	32	9	33
	H4	P62801	5	39	10	45
	RLP26	NP001264027	6	35	11	40
	H3	P84246	7	15	-	-
	RLP23A	XP415820	8	15	-	-

Notes

1. Some proteins have not been listed due to low MS score and/or low sequence coverage, or an obsolete sequence, or not *Gallus gallus*.

2. MS results were provided by Professor Dickman's group, Sheffield University, with the MS parameters as per Table 5.4(b).

Table 5.8 Summary of proteins in fractions 5a1 – 5a5 of the gel filtration chromatogram (Figure 5.5a) for Fraction 5a obtained by cation-exchange chromatography of cePNE4' proteins

(Based on a review of MS-PAGE gels and MS results - Figure 5.6(a) and Table 5.7)

	Smaller protein or protein complex →				
	Gel filtration fraction				
Protein	5a1	5a2	5a3 (main peak)	5a4	5a5
HMGB1		x			
HMGB2		X			
HMGB3		X	x		
H2A		X	X	X	x
H2B		X	X	X	x
H3	x	x			
H4	x	x	x		
H5		x	X	X	x
H1	x	X	X		
HSP70		x			

Notes

1. X indicates the stronger SDS-PAGE bands/high scoring MS results. Some proteins as above were identified by MS, but do not show on the gel, and vice versa. A few low scoring proteins other than HMGB and histone proteins were also identified by MS, but are not shown above.

proteins other than HMGB and histone proteins were also identified by MS, but are not shown above.

2. Protein H1, H2A, and H2B in the above table represent MS results for one or more of the following variants:

H1: UniprotKB P09987

H1.11R: UniprotKB P08288

H1.01: UniprotKB P08284

H2A: UniprotKB Q92069

H2A.V or H2A.Z: UniprotKB P02272 or Q5ZMD6

H2B: UniprotKB P0C1H3

H2B VII: UniprotKB P0C1H5

Table 5.9 Summary of proteins in fractions 5c1 – 5c5 of the gel filtration chromatogram (Figure 5.5b) for Fraction 5c obtained by cation-exchange chromatography of cePNE4' proteins

(Based on a review of MS-PAGE gels and MS results - Figure 5.6(b) and Table 5.7)

	Smaller protein or protein complex →			
	Gel filtration peak			
Protein	5c1	5c2	5c3	5c4
			(main peak)	
HMGB1		x		
HMGB2		X		
HMGB3		X		
H2A	x	X	X	x
H2B	x	x	X	x
H3		x	x	
H4	x	X	X	x
H5		x	x	x

Notes

1. X indicates the stronger SDS-PAGE bands/high scoring MS results. Some proteins as above were identified by MS, but do not show on the gel, and vice versa. A few low scoring proteins other than HMGB and histone proteins were also identified by MS, but are not shown above.

2. Proteins HMGB3, H2A, and H2B in the above table represent MS results for one or more of the following protein variants;

HMGB3: UniprotKB P36194 or NCBI NP_990626.2

H2A: UniprotKB P02263

H2A.V or H2A.Z: UniprotKB P02272 or Q5ZMD6

H2A: NCBI XP_004947974

H2B: UniprotKB Q92069

H2B VII: UniprotKB P0C1H5

5.4 Discussion: a model to explain the cation-exchange chromatography results

5.4.1 The general picture

For the cation-exchange chromatography, the protein sample was submitted to the column in an environment similar to biological conditions (see Figure 5.3 for buffers), in terms of pH, ionic concentration, presence of phosphates, and presence of anchored negative charges. Biological conditions in chicken erythrocyte nuclei are assumed to include pH~7.1 and ionic concentration ~300milliosmoles (Chapter 2 Section 2.5.4; Lodish *et al.*, 2007 - Table 11.2), and the nuclear DNA includes phosphates and associated negative charges.

Table 5.5 shows the following:

- (i) H1 histone variants were only moderately attached to the negatively charged stationary phase of the cation-exchange column. They eluted mainly in Peaks M, 3, and 4 [490 – 740mM].
- (ii) H2A, H2B histones and HMGB proteins were more attached than the H1 proteins. They eluted mainly in Peak 5 (a, b, and c) [740 - 825mM].
- (iii) Some H3 and H4 histones eluted over the same range as H2A, H2B histones and HMGB proteins, but a significant amount were more strongly attached, eluting in Peak 7 (a and b) [950 - 1030mM].
- (iv) H5 histones eluted across a wide range from 350 – 1000mL, but mainly eluted in Peaks 6 and 7 [buffer 870 - 1030mM].

5.4.2 Issues to be resolved

The above results have a number of unusual features:

1. The net charge on the H1 variants is large, $\sim 60e$, and similar to that on a chicken H5 molecule (Table 5.6), yet the H1 molecules were significantly less attached to the cation-exchange column than most of the H5 molecules.

2. The net charge on an HMGB1 protein (Table 5.6) is negative, and the net positive charges on HMGB2 and HMGB3 are small. Consequently, it could be expected that HMGB1 would not have bound to the column, and that the other HMGB proteins would have been only weakly bound. In contrast, all three HMGB proteins bound as strongly to the column as some of the core histones, which have large net positive charges. Buffer conditions were such that most core histones would have formed dimers and tetramers. From Table 5.6, the relevant net charges at pH6.8 are:

HMGB1	$-3.9e$
HMGB2	$+4.4e$
HMGB3	$+2.9e$
H2A.H2B	$+37e$
(H3.H4) ₂	$+78.4e$
Linker histones	$\sim 60e$

3. Consistent with their high net charge of $\sim 60e$, some H5 histones were strongly attached to the cation-exchange column, hence eluting in buffer 870 - 1030mM. However, other H5 histones were eluted in ionic strength as low as 490mM, indicating an exceptionally wide spectrum of attachment of H5 molecules to the cation-exchange column.

4. The MS results (Table 5.4) suggest more H4 than H3 molecules, but the difference is small and may not be significant (H4 molecules are equal in number to H3 molecules in chromatin, and there is nothing in our procedures which would obviously preferentially remove H3 proteins). Consequently, this effect is not considered further.

5.4.3 Factors considered in seeking to resolve the issues

The possibility of H1 fragmentation was considered (a protease inhibitor was omitted when running the cation-exchange column, for the reason described in Section 8.4.4). However, the experimental evidence refuted this - MS found H1 residue fragments spanning most of the full H1 sequence, and SDS-PAGE gels (Figure 5.4(a)) showed bands as expected for complete H1 molecules.

H5 fragmentation was considered. Except for Fractions 7a and 7b, the SDS-PAGE gel evidence (Figure 5.4(d)) suggested that mainly full length H5 molecules were present in the chromatography, so fragmentation is not very likely to have caused H5 to bind to the column for the whole spectrum of eluent ionic strengths, although it cannot be completely ruled out. The gel in Figure 5.4(d) showed that H5 in Fractions 7a and 7b was slightly shorter than the full length molecule, since H5 progressed slightly more quickly down the gel than full length H5 (although the H5 gel band for Fraction 7b is very faint). Some degradation of H5 was not unexpected, since our preferred protease inhibitor was omitted from the cation-exchange chromatography.

Only the N-terminal half of H5 was identified by the MS results (Table 5.10), but MS may miss many residues if reliance is placed on a single trypsin digestion, as was applied for the MS results we have used here. As further discussed in Section 8.4.3, Snijders *et al.* (2008) have shown that significantly better sequence coverage of linker histones can be achieved by applying a series of partial trypsin digestions. Thus, the MS results we describe here do not preclude the presence of both full length and fragmented H5 molecules. Also, H5 was present in high concentration, so only a small proportion of H5 molecules would need to have been fragmented to produce the wide spectrum of attachment indicated by the MS results (see Tables 5.4 and 5.5).

Table 5.10 Summary of MS results for histone H5

Peaks are those in the cation-exchange chromatogram Figure 5.3

Peak	Summary based on published MS results (NCBI database) (Zhuang <i>et al.</i> , 2014)			Summary based on unpublished MS results (SwissProt database, and including N-terminal K/R residues)		
	In order of MS score	N-terminal half or C-terminal half of protein	% sequence coverage	In order of MS score	N-terminal half or C-terminal half of protein	% sequence coverage
M	13	N-terminal	21%	15	N-terminal	23%
3	5	N-terminal	35%	6	N-terminal	40%
4	4	N-terminal	27%	7	N-terminal	37%
5a1	No data	-	-	-	Not found	None
5a2	No data	-	-	25	N-terminal	16%
5a3	No data	-	-	3	N-terminal	38%
5a4	No data	-	-	1	N-terminal	53%
5a5	No data	-	-	9	N-terminal	36%
5b	No data	-	-	1	N-terminal	37%
5c1	-	No H5	None	-	No H5	None
5c2	1	N-terminal	29%	1	N-terminal	33%
5c3	1	N-terminal	24%	1	N-terminal	39%
5c4	1	N-terminal	31%	1	N-terminal	39%
6a	3	N-terminal	33%	4	N-terminal	36%
6b	1	N-terminal	33%	1	N-terminal	37%
7a	2	N-terminal	37%	2	N-terminal	53%
7b	2	N-terminal	33%	2	N-terminal	37%

MS results were provided by Professor Dickman's group, Sheffield University, with the MS parameters as per Table 5.4(b).

A further factor considered was a potential mechanism for binding between the HMGB proteins, histones, and the (acidic) negatively charged cation-exchange column. Cato *et al.* (2008) proposed a model in which HMGB1 boxes attach to DNA, leaving the HMGB1 acidic tail free to bind to the basic tail of a linker histone. In fact, all the histones have basic unstructured tails with potential to link to an HMGB C-terminal acidic tail. Furthermore, Watson *et al.* (2007) have provided evidence that, although the HMGB1 C-terminal acidic tail in free solution is folded back onto the basic part of the molecule (including both HMG boxes) most of the time, this configuration is in dynamic equilibrium with an open dipolar configuration, in which the acidic tail is not bound to the HMG boxes. Thus, anomalous binding of an HMGB molecule to the cation-exchange column can be explained if the molecule has flicked open (perhaps facilitated by the strong negative charge on the column), with the basic part of HMGB (typically, residues 1 – 185) binding to the stationary phase on the column. This would leave the HMGB acidic tail free to bind to the basic tail of a histone. These considerations have led to the derivation of a simple model, described below, to account for the proteins found in the cation-exchange chromatography peaks (Table 5.5).

5.4.4 A model which accounts for the cation-exchange chromatography results

A model for the release of proteins from the cation-exchange column is proposed, based on:

- (a) Initial retention of a portion of the linker and core histones indirectly on the column stationary phase, via complexes with the HMGB proteins. Table 5.11 shows the net charges on the complexes;
- and (b) Initial retention of the remaining portion of histones directly on the column.

The model assumes a basic (positive) part of the HMGB molecule has bound to the negative charges on the column stationary phase, releasing the HMGB C-terminal acid tails to bind to a histone (Figure 5.7). In this model, the starting situation (beginning of elution) on the column is as shown in the bottom line in Figure 5.8. Available histones are shown bound to the column, but also shown are the complexes in which an HMGB protein is bound both to the column and to a linker histone, or a histone dimer, or a histone tetramer.

Table 5.11 Charges on proteins and complexes relevant to the cation-exchange chromatogram

Protein or protein complex	Net charge at pH6.8 (e)
H1	60.9
H5	61.9
H2A.H2B	37
(H3.H4) ₂	78.4
H1/HMGB	57 – 65.3
H5/HMGB	58 – 66.3
H4/HMGB	14.7 - 23
H2A.H2B/HMGB	33.1 – 41.4
(H3.H4) ₂ /HMGB	74.5 – 82.8

Notes

Net charges are based on those shown in Table 5.6.

HMGB1 reduces the net charge on the complex by 3.9e, whereas HMGB2 or HMGB3 increase the net charge by 4.4e or 2.9e respectively.

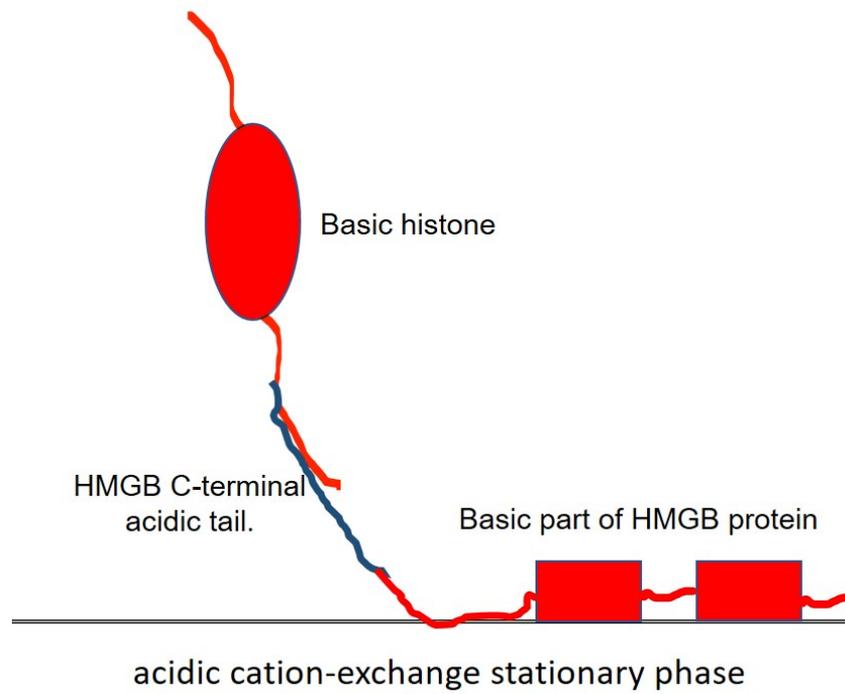


Figure 5.7 Proposed arrangement for attachment of histone/HMGB complex to cation-exchange stationary phase

Peaks 6b, 7a, 7b

H2A.H2B	(H3.H4) ₂	H5				
↑	↑	↑				
-	-	-	-/-	-/-	-/-	-/-

Column stationary phase

Peak 6a

		H5		H2A.H2B/HMGB		
H2A.H2B	(H3.H4) ₂	↑ H5		↑ -/-		
			-/-		-/-	-/-

Column stationary phase

↑

Peak 5a, 5b, 5c

			H1/HMGB	H2A.H2B/HMGB	(H3.H4) ₂ /HMGB	H5/HMGB
			↑	↑	↑	↑
H2A.H2B	(H3.H4) ₂	H5	-/-	H2A.H2B/HMGB	-/-	-/-

Column stationary phase

↑

Peak 4

			H1	H2A.H2B		H5
			↑	↑		↑
H2A.H2B	(H3.H4) ₂	H5	H1/HMGB	H2A.H2B/HMGB	(H3.H4) ₂ /HMGB	H5/HMGB

Column stationary phase

↑

Peaks M, 3

			H1		H5
			↑		↑
H2A.H2B	(H3.H4) ₂	H5	H1 /HMGB	H2A.H2B/HMGB	(H3.H4) ₂ /HMGB

Column stationary phase

↑

After sample injection

H2A.H2B	(H3.H4) ₂	H5	H1/HMGB	H2A.H2B/HMGB	(H3.H4) ₂ /HMGB	H5/HMGB
---------	----------------------	----	---------	--------------	----------------------------	---------

Column stationary phase

↑

Figure 5.8 A model for the release of proteins from the cation-exchange column

-/- Indicates complex no longer present on the column.

↑ Part or all of a protein released from the column. ↑ Increasing eluent ionic strength.

In the proposed model, depending on the relative binding strengths involved, waves of proteins are released from the column as eluent ionic strength increases. Taking account of the net charges involved (Table 5.11) and given that the SP Sepharose column provides a strong cation-exchange stationary phase, it is assumed that some of the weaker binding is within the histone/HMGB complexes. Some of the complexes will also be less bound to the column than the directly bound histones.

The experimental data matches the detailed model represented by Figure 5.8. It shows the first wave of proteins, released at relatively low ionic strength in Peak M and Peak 3, contained H1 and H5 histones - it is proposed that these molecules are released from complexes on the column. In the second wave, H1 and H5 are joined by H2A.H2B released from part of the H2A.H2B/HMGB complex on the column. In the next wave, all of the complexes, except some H2A.H2B/HMGB, are completely released from the column. Finally, histones directly bound to the column become released. These waves of proteins map onto the contents of the various peaks in the cation-exchange chromatogram, from Peak M up to Peak 7b (Table 5.5, Figure 5.8).

The model relies on the initial creation of four histone/HMGB complexes (Figure 5.7), via binding of the HMGB C-terminal acidic tails to the basic histones. The complex is formed either prior to the sample being injected, or when assisted by the negative charges on the column stationary phase (this may help the HMGB C-terminal acidic tail detach from the rest of the molecule). This latter option may partially reflect the situation which arises in the nucleus in the presence of DNA which is negative, as is the cation-exchange stationary phase.

5.5 Further support for the proposed model

The existence of potential HMGB/histone complexes on the cation-exchange column was investigated by taking HMGB-rich fractions (Peaks 5a and 5c) from the cation-exchange chromatogram and running them on gel filtration columns, as described in Section 5.3.6. This approach also provided information about the central Peak 5b due to its overlap with the adjoining fractions. The gel filtration chromatography separated the proteins (or proteins complexes) according to size, for which molecular weight is an approximate proxy. Larger molecules/complexes eluted earlier; smaller molecules/complexes eluted later.

Because buffer conditions supported the presence of core histone dimers and tetramers, the relevant molecular weights (kDa), based on MWs in Table 5.6, were:

(H3.H4) ₂	53.6
H2A.H2B	27.9
H3.H4	26.8
HMGB proteins	23.1 – 24.9
H1	21.9
H5	20.7

The presence of proteins in the gel filtration peaks showed a consistent pattern (Tables 5.8 and 5.9). The HMGB proteins are almost completely absent from the two main gel filtration peaks (5a3 and 5c3/5c4), even though their lower individual molecular weights should place them to the right of the histone dimers. Instead, the HMGB proteins appear with histone dimers and linker histones in small peaks (5a2 and 5c2) to the left of the peak histone distributions. This result suggests the existence of complexes which involve an HMGB protein and have a molecular weight which exceeds that of the individual HMGB proteins. This support for HMGB complexes is very interesting, not least since the ionic strength of the gel filtration buffer (700mM molarity) is above the ionic strength at which HMGB/histone interactions have previously been reported (see Section 5.2). Just as H1 is known to reconfigure on interaction with DNA to increase binding to the DNA (Harshman *et al.*, 2013), the model suggests that the HMGB molecules may reconfigure into their open dipolar configurations on interaction with the cation-exchange stationary phase, so that their strongly negatively-charged acidic tails then become available to bind the positively charged histones.

The gel filtration results provide evidence of HMGB proteins being in complex either with histones, or as oligomers, at a relatively high ionic strength. However, sedimentation studies by Marekov *et al.* (1984) mitigate against HMGB oligomers – at low ionic strength (50mM tris-HCl) native calf thymus HMGB1 in isolation formed an oligomer, but no oligomer was found in 200mM NaCl.

5.6 Conclusions and recommendations

1. The cePNE4' group of proteins, prepared as described in this chapter from chicken erythrocytes, contains primarily histones and HMGB proteins. This has facilitated an analysis of their interactions, using chromatography, SDS-PAGE, and MS.
2. A simple coherent model has been derived which provides an explanation for the main proteins in each peak of a cation-exchange chromatogram obtained for the cePNE4' proteins. The model is based on the proposed existence of numerous histone/HMGB complexes arising from ionic interaction involving the HMGB C-terminal acid tails, with the remainder of the HMGB protein initially bound to the cation-exchange column at ionic strength and pH near to biological conditions (see Cation-exchange Buffer A, Figure 5.3). However, we have yet to investigate whether any of these proposed complexes existed prior to running the cePNE4' proteins on the cation-exchange column.
3. Some limited evidence in support of the above complexes was found using gel filtration chromatography, despite the buffer strength (see Figure 5.5) exceeding a level at which histone/HMGB complexes have previously been reported (Section 5.2).
4. Previous studies (Section 5.2) indicated the existence of a limited number of histone/HMGB interactions, mainly involving HMGB1. The above model suggests that such interactions may be more widespread, involving all of the histones and each HMGB (1,2 or 3), at least in the presence of negative charges, such as are found on a strong cation-exchange stationary phase, or on DNA.

5. At this stage the model should be regarded as speculative. However, the experimental system described provides a basis for several ways in which the model can be verified. These include:

(i) Re-running the experimental work, but with more attention to details such as establishing the quantities of each protein, checking SDS-PAGE gels for evidence or otherwise of full length proteins, better calibration of the gel-filtration chromatography (using pure proteins) to permit the size of complexes to be estimated, measures to ensure oxidation of the HMGBs is prevented, and arrangements to ensure protease inhibition during the cation-exchange chromatography (such as described in Section 8.4.4). The experiments could also be simplified as necessary by using samples containing fewer proteins, such as a sample containing only HMGB2 and H5. In addition, the effect of peak broadening (due to limited resolution) in the cation-exchange chromatogram could be taken into account by preparing SDS-PAGE gels to cover all fractions, not just those fractions associated with peaks in the chromatogram.

(ii) Use of immuno-precipitation to confirm whether complexes existed before or after the chromatography, and if so which specific histones and HMGBs have formed complexes. However, there is a dearth of commercial antibodies for linker histones. For example, there are relatively few specific anti-H1 antibodies, because there are so many H1 variants (>10 in humans) and H1 PTMs (Harshman *et al.*, 2013). A search for commercial antibodies on the biocompare.com website did not reveal any anti-chicken antibodies which specifically targeted H5 (however, there is a supplier of anti-chicken H1/H5 antibody).

(iii) Preparing a similar sample, but with one or more HMGBs removed, prior to obtaining a cation-exchange chromatogram for comparison with the one shown here. The next Chapter explores an efficient way of removing the HMGBs.

(iv) Running the experiment with a less positively charged cation-exchange stationary phase, to elucidate whether a strongly positive stationary phase is necessary for the proposed HMGB/histone complexes to form.

6. Should the model be broadly confirmed, it provides a limited step towards further insight into a possible more general role of HMGB proteins in the nucleus, in addition to the known role of HMGB1 in assisting access to the DNA in chromatin. It would also shift the spotlight onto the role of initial unstructured ionic interactions between biological molecules, which may be essential in advance of the assembly of the biological machines which in contrast rely for final configuration on specific structural details. A related question is the extent to which ionic strength within the nucleus is localised, since this parameter affects the strength of ionic attraction. Chromatin function is already known to be influenced by its density which varies from one location to another within a nucleus (Ou *et al.*, 2017). Ionic strength could also be localised and under cellular control.

Chapter 6

A method for the efficient separation of native HMGB proteins from other nuclear proteins in chicken erythrocytes

<u>Contents</u>	Page
6.1 Introduction	163
6.2 Previous studies	163
6.3 Experimental method for the efficient isolation and separation of native HMGB proteins	166
6.4 Impact of pH on the method for isolation and separation of native HMGB proteins	173
6.5 Discussion of results	175
6.6 Conclusions	182
<u>Tables</u>	
Table 6.1 Summary of methods for isolating HMGB proteins based on anion-exchange chromatography	165
Table 6.2(a) Summary of MS results for SDS-PAGE bands in Figure 6.3	171
Table 6.2(b) Parameters used to obtain MS results for SDS-PAGE gel bands in Figure 6.3	172
Table 6.3 Absorbance and MWs of chicken HMGB and histone proteins	178
Table 6.4 Chicken erythrocyte HMGB pls and charges	180
<u>Figures</u>	
Figure 6.1 Summary of a simple method for preparing native HMGB proteins from chicken erythrocytes	167
Figure 6.2. Anion-exchange chromatogram of cePNE1' proteins (pH6.8)	168
Figure 6.3 SDS-PAGE gel of fractions from anion exchange of cePNE1' proteins	170
Figure 6.4. Anion-exchange chromatogram of cePNE1' proteins (pH5.6)	174
Figure 6.5 Scheme for investigating the partitioning of HMGB1 proteins between the cytosol, free in the nucleus, and attached to chromatin	177
Figure 6.6 Proposed model for attachment of HMGB protein to anion-exchange column	181

6.1 Introduction

A method based on anion-exchange chromatography and exploiting the dipolar nature of HMGBs, has been devised to permit a simple and efficient separation of almost pure native HMGBs from the other nuclear proteins in chicken erythrocytes. This includes the separation of HMGB1 from HMGB2/HMGB3. The method is well suited to large scale, cost effective production of native HMGB proteins, for use in further studies as explored in Chapter 7.

The method exploits insights from histone and HMGB purification studies described by Foulger *et al.* (2012). Foulger used more complex steps than those described here, including ultracentrifugation and cation-exchange chromatography to obtain moderately pure HMGB proteins from chicken erythrocytes. HMGB1 and HMGB3 were present in a single peak of the cation-exchange chromatogram, and HMGB2 was present in a separate peak. A further step involving anion exchange chromatography permitted the separation of HMGB3 with high purity.

6.2 Previous studies

As a part of wider studies, numerous previous methods have been applied to obtain HMGB proteins, either by extraction from animal tissues or cell lines, or by recombinant methods. As indicated by the varied methods discussed in Chapter 3, there does not appear to be a standard protocol for obtaining HMGB proteins.

Several methods have used acid extraction, but this has the disadvantage of degrading the acid-labile PTMs (Chen *et al.*, 2007), and compromises interactions with DNA (Marekov *et al.*, 1984). Recombinant HMGB proteins are also likely to suffer from a lack of appropriate PTMs. As with other proteins, PTMs have a key influence on HMGB1 function; HMGB PTMs fine-tune interactions of the proteins with chromatin and determine their relocation from the nucleus to the cytoplasm and secretion (Stros, 2010).

Many of the methods have used cation-exchange chromatography to purify the HMGB proteins (see Chapter 3). While this method separates the HMGB proteins from other nuclear proteins, it is likely to involve more steps than the method described here, and it provides only limited success in separating HMGB proteins one from another.

A few studies have sought efficient methods for production of native HMGB proteins which avoid acid extraction and do use anion-exchange chromatography. For example, Marekov *et al.* (1984) obtained HMGB proteins using the following steps: (i) cells and nuclei were lysed, and the chromatin was sedimented, with any soluble proteins being discarded, (ii) ionic strength was increased to extract the less bound proteins from the chromatin, (iii) ionic strength was further increased to precipitate unwanted proteins from the extract, and (iv) the remaining extract was run on an anion-exchange column (see Table 6.1 for details). In addition to the non-binding fractions, the anion-exchange chromatogram showed two main peaks - a double peak referred to as pure HMGB2, and a later single peak containing HMGB1.

Adachi *et al.* (1990) used a similar approach to Marekov, also summarised in Table 6.1. The anion-exchange chromatogram they obtained showed three main peaks. The early peak was identified as containing HMGN proteins, the later second and third peaks were identified by antibodies as containing HMGB2 and HMGB1 respectively.

Neither Marekov nor Adachi found any HMGB3. This may be because they used calf and porcine thymus tissue respectively. The Human Protein Atlas (2017) indicates that in humans, high levels of HMGB3 are mainly found in organs associated with reproduction or the lung (although RNA for HMGB3 is found in all organs).

The method described in this Chapter also relies on anion-exchange chromatography, but is considered a particularly simple and effective method, since it produces a high yield of relatively pure HMGBs quickly from cell nuclei. One key aspect is that many unwanted proteins remain within the nuclei, since the nuclei are not lysed.

6.1 Summary of methods for isolating HMGB proteins based on anion-exchange chromatography

Step	Marekov <i>et al.</i> (1984) method applied to calf thymus tissue.	Adachi <i>et al.</i> (1990) method applied to porcine thymus tissue.
1. Lysis of cells and nuclei to provide access to chromatin	Mechanical shearing. Hypotonic buffer. EDTA.	Mechanical shearing. Hypotonic buffer.
2. Extraction of proteins loosely bound to chromatin	In solution with 165mM $(\text{NH}_4)_2\text{SO}_4$.	In solution with 350mM NaCl.
3. Precipitation of unwanted proteins	In solution with 2.6M $(\text{NH}_4)_2\text{SO}_4$.	In solution with 2% TCA.
4. Remaining proteins run on anion-exchange column	Dialysis into running buffer, pH7.2. NaCl gradient from 0 to 0.4M. DEAE-cellulose column.	Precipitation in 10% TCA. followed by dissolving in running buffer, pH7.8. NaCl gradient from 0 to 1.0M PBE 94 column.

Additional reagents, which were included in the buffers to fix pH and to suppress proteases, are not shown in the above table.

EDTA sequesters divalent metal ions and in certain conditions solubilizes cell membrane lipopolysaccharides, compromising the integrity of the membrane (Gray and Wilkinson, 1965).

6.3 Experimental method for the efficient isolation and separation of native HMGB proteins

6.3.1 Summary

Figure 6.1 summarises the main steps applied to isolate and separate the HMGB proteins from chicken erythrocytes. Erythrocyte nuclei were prepared, and the cePNE1' group of proteins was extracted from the intact nuclei. The cePNE1' proteins were then run on an anion-exchange column, providing peaks with high concentrations of HMGB proteins.

6.3.2 Preparation of erythrocyte nuclei

Erythrocytes were separated from the other components of fresh chick blood by centrifugation, and intact nuclei were obtained by lysing the cells in three stages in a hypotonic buffer containing a detergent, as described in Section 5.3.2 (Chapter 5) .

6.3.3 Nuclei washing and depletion

Nuclei were washed and depleted of cePNE1' proteins using a mildly hypertonic extraction buffer with pH6.8 as described in Section 5.3.3 except that a slightly higher ionic strength extraction buffer was now employed, as follows:

cePNE1'-extraction Buffer	90mM KCl, 45mM H ₂ KPO ₄ , 45mM HK ₂ PO ₄
(pH6.8)	3.5mM MgCl ₂ , 2.5mM benzamidine hydrochloride.

The cePNE1' proteins contained substantial quantities of HMGB proteins for the reasons set out in Section 6.5.1.

6.3.4 Anion-exchange chromatography to isolate and separate HMGB proteins

The cePNE1' proteins were present in cePNE1'-extraction Buffer at pH6.8. This buffer was also used as Buffer A for anion-exchange chromatography, except that MgCl₂ was omitted from Buffer A – see Figure 6.2 for the buffer composition. After filtration (0.45µm) to remove any particulate which might clog the column, a sample of the proteins was injected directly into the anion-exchange column (8 x HiTrap DEAE FF 5mL units). Proteins were progressively eluted from the column using a linear KCl gradient from 90 to 600mM. Figure 6.2 shows the chromatogram.

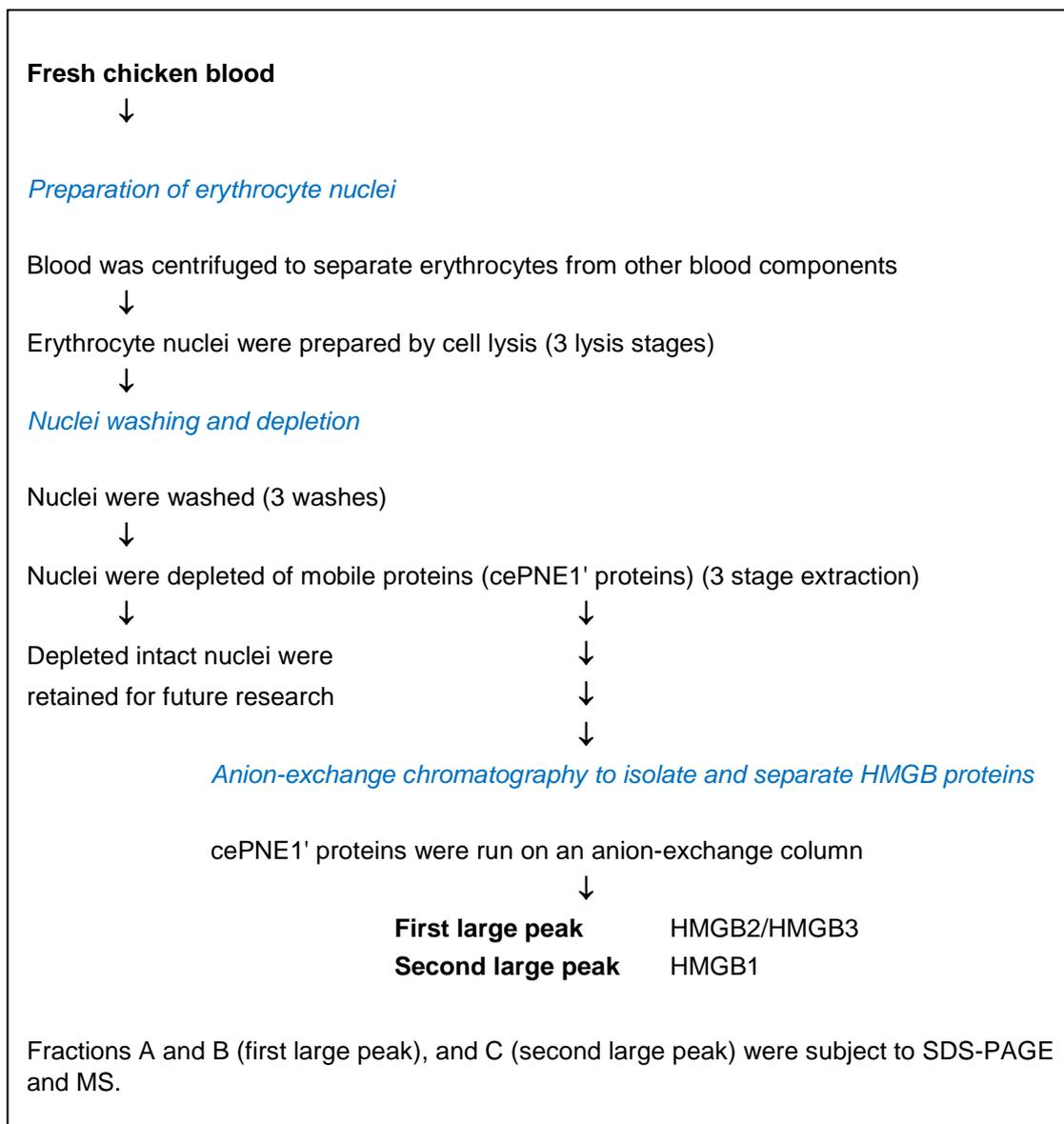


Figure 6.1 Summary of a simple method for preparing native HMGB proteins from chicken erythrocytes

See Section 6.3.2 onwards for further details of the method. Operations were conducted at $\sim 4^{\circ}\text{C}$ unless otherwise stated. Buffers used for chromatography were filtered ($0.45\mu\text{m}$) to avoid clogging the column.

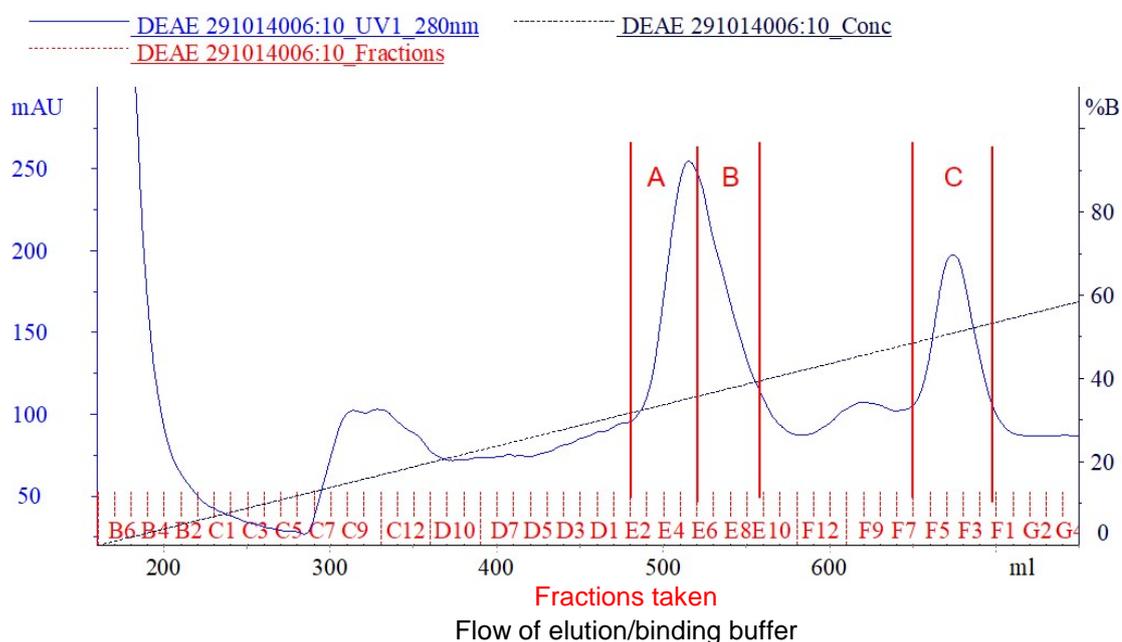


Figure 6.2. Anion-exchange chromatogram of cePNE1' proteins (pH6.8)

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B.

- at start of elution: Buffer A 100%, Buffer B 0%;

- at end of elution: Buffer A 0%, Buffer B 100%.

Buffer A (pH6.8) 90mM KCl, 45mM H₂KPO₄, 45mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Buffer B (pH6.8) 600mM KCl, 45mM H₂KPO₄, 45mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Columns 8 x HiTrap DEAE FF 5mL (GE Healthcare)

Total column volume 40mL

Fraction size 10mL (X-axis includes marked fractions)

Chromatogram identifier DEAE 291014006

Fractions subject to assay A, B, C (see Figure 6.3, Table 6.2)

The chromatogram showed several peaks. It was suspected that the largest peak might contain both HMGB2 and HMGB3, so Fractions A and B were taken from the left and right hand sides respectively of that peak. Fraction C was taken from the second largest peak, and all three fractions were subject to SDS-PAGE, with the result shown in Figure 6.3. Bands were cut out from the SDS-PAGE gels, prepared for MS as described in Chapter 3, and were subject to MS at Sheffield University. The MS results summarised in Table 6.2 indicated HMGB proteins were present, but did not conclusively differentiate between the different HMGB proteins. However, taking account of the gel band positions, it was possible to assign HMGB1, 2 or 3 to each of the bands in Figure 6.3. On this basis, the largest peak in Figure 6.2 contained both HMGB2 and HMGB3, with HMGB2 being slightly more bound to the column than HMGB3. The second largest peak contained HMGB1. The peaks also provided an indication of the yields of HMGB proteins, this is further discussed in Section 6.5.1.

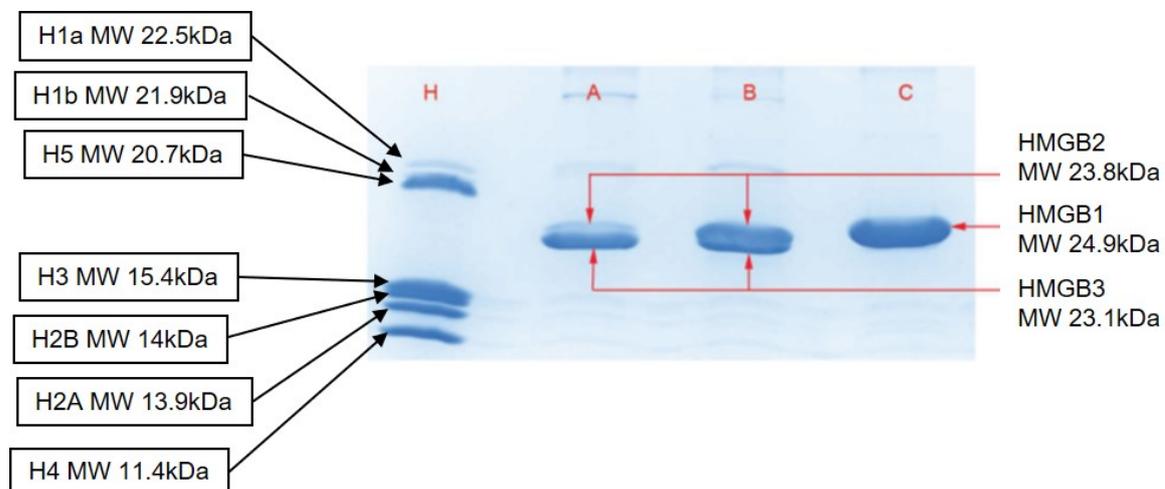


Figure 6.3 SDS-PAGE gel of fractions from anion exchange of cePNE1' proteins

H Histone standard, see Chapter 4.

A, B, C Fractions as defined in Figure 6.2.

Bands are marked up with protein as determined by protein position and MS results.

Stain: Coomassie Blue (although photographed in black and white)

Stacking gel: 5% acrylamide

Resolving gel: 20% acrylamide

Running buffer: glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol: as described in Chapter 3

Table 6.2(a) Summary of MS results for SDS-PAGE bands in Figure 6.3

Band as per Figure 6.3 (MS identifier)	Proteins identified by MS	Score (NCBI database)	Score (SwissProt database)	As interpreted in Figure 6.3 based on band position.
Fraction A Upper band (P3.11)	HMGB1 or HMGB3	197		HMGB2
	HMGB3		197	
Fraction A Lower band (P3.10)	HMGB1 or HMGB3	233		HMGB3
	HMGB3		222	
Fraction B Upper band (P3.8)	HMGB3 or HMGB1	77		HMGB2
	HMGB3		77	
	HMGB2	58	44	
Fraction B Lower Band (P3.9)	HMGB3-like [<i>Bos taurus</i>]	140		HMGB3
	HMGB3		186	
	HMGB2	63		
Fraction C (P3.7)	HMGB1 (<i>Tupaia chinensis</i>)	217		HMGB1
	HMGB1		210	
	H4		77	

Notes

1. Results are for *Gallus gallus*, unless stated otherwise. A few proteins have not been listed due to low MS score.
2. MS results were provided by Professor Dickman's group, Sheffield University, with the MS parameters as per Table 6.2(b).
3. Unfortunately, for an unknown reason, the gel cut-out samples submitted to Sheffield University provided poor MS sequence coverage. For example, the MS coverage of Fraction C (P3.7) using the SwissProt database was only 18.6% (excludes K or R residues not found by MS, but assigned by MASCOT software on the basis of a protein match), and MS of Fraction C using the NCBI database was unable to identify *Gallus gallus* HMGB1 (although it did find *Tupaia chinensis* HMGB1). Coupled with poor sequence coverage, the similarity of HMGB sequences has made it difficult for the MASCOT sequence matching program to distinguish between HMGB proteins, as can be seen by the above results.

Table 6.2(b) Parameters used to obtain MS results for SDS-PAGE gel bands in Figure 6.3

Parameters for MS using the NCBI sequence database

```

Search title           : JB bands May 2016 (C:\ProgramData\Matrix
Science\Mascot Daemon\parameters\Alison gallus Amazon.par), submitted
from Daemon on MJDHOTDESK-PC
MS data file          : D:\MJD\March 16\P 3.7_RA7_01_5587.mgf
Database 1            : contaminants 20160129 (247 sequences; 128130
residues)
Database 2            : NCBIInr 20140730 (47310802 sequences; 16880015641
residues)
Taxonomy 2            : bony vertebrates (4148099 sequences)
Timestamp             : 17 May 2016 at 15:13:01 GMT

```

Parameters for MS using the SwissProt sequence database

```

Search title           : JB bands May 2016 swissprot
(C:\ProgramData\Matrix Science\Mascot Daemon\parameters\Alison gallus
Amazon.par), submitted from Daemon on MJDHOTDESK-PC
MS data file          : D:\MJD\March 16\P 3.7_RA7_01_5587.mgf
Database 1            : contaminants 20160129 (247 sequences; 128130
residues)
Database 2            : SwissProt 2015_12 (550116 sequences; 196219159
residues)
Taxonomy 2            : bony vertebrates (83931 sequences)
Timestamp             : 27 Sep 2016 at 13:14:29 GMT

```

Common parameters for MS

```

Enzyme                 : Trypsin
Variable modifications : Carbamidomethyl (C),Oxidation (M)
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.6 Da
Fragment Mass Tolerance: ± 0.6 Da
Max Missed Cleavages   : 1
Instrument type         : ESI-TRAP
Number of queries      : 1098

```

Bands were cut from the SDS-PAGE gel in Figure 6.3 and were subject to MS at Sheffield University after preparation for MS as described in Chapter 3.

6.4 Impact of pH on the method for isolation and separation of native HMGB proteins

To further explore the extent to which the above results were influenced by the pI of the HMGBs, and hence whether the HMGBs were acting as dipoles, the previous work was repeated using pH5.6 buffers, below the pI of HMGB1 (5.73).

The cePNE1' proteins were extracted with a revised cePNE1'-extraction Buffer; this was the same as the original buffer except buffered at pH5.6, achieved by reducing the proportion of HK_2PO_4 in the 90mM phosphate mixture. This lower pH was selected to just avoid lysing the nuclei.

This buffer was also used as Buffer A for anion-exchange chromatography – see Figure 6.4 for the buffer composition. As before, after filtration (0.45 μm) a sample of the proteins was injected directly into the anion-exchange column. Proteins were progressively eluted from the column, again using a linear KCl gradient from 90 to 600mM. The chromatogram and running conditions are shown in Figure 6.4. However, no SDS-PAGE gels or MS results were obtained; these were considered unnecessary because the purpose was to show that the isolation of HMGB proteins was independent of pH, and because an intermediate result at pH5.8 (see below) enabled the two biggest peaks in this chromatograph to be mapped onto the HMGB proteins.

The chromatogram at pH5.6 again shows several peaks, but the two biggest peaks have shifted to the right compared with the chromatogram at pH6.8, indicating that the proteins in these peaks have become more strongly bound. For example, at pH6.8, HMGB1 eluted at 330mM KCl, whereas at pH5.6 HMGB1 eluted as 364mM KCl.

The above work was repeated at pH5.8, again by varying the proportion of HK_2PO_4 . The chromatogram again showed 3 large peaks, similar to those in Figure 6.4. An SDS-PAGE gel was run with fractions from the second and third large peaks. The second peak produced a strong band at the same position as that for Fractions A and B in Figure 6.3, indicating that it contained HMGB2/3, and the third peak produced a strong band slightly higher up the gel, indicating it contained HMGB1. However, no MS results were obtained.

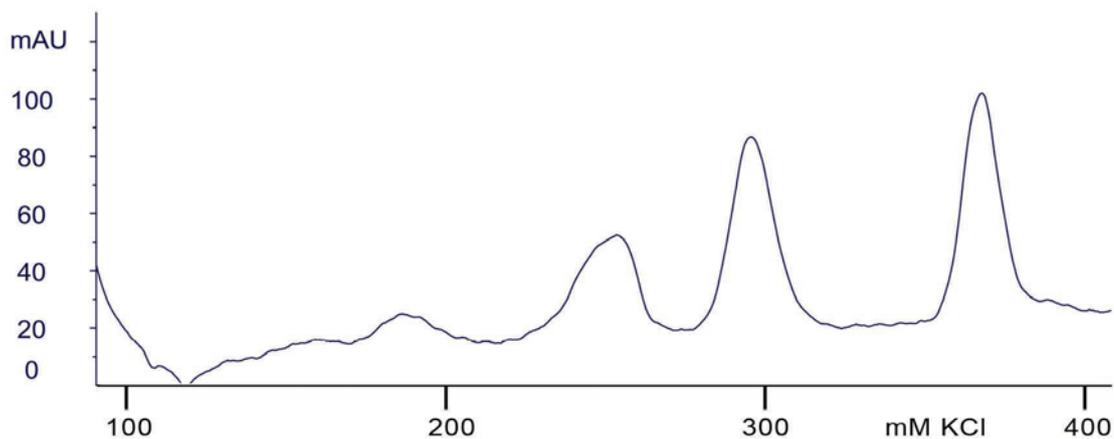


Figure 6.4. Anion-exchange chromatogram of cePNE1' proteins (pH5.6).

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%;

At end of elution: Buffer A 0%, Buffer B 100%.

Buffer A (pH5.6) 90mM KCl, 85.5mM H₂KPO₄, 4.5mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Buffer B (pH5.6) 600mM KCl, 85.5mM H₂KPO₄, 4.5mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Anion-exchange columns 8 x HiTrap DEAE FF 5mL (GE Healthcare)

Total column volume 40mL

Fraction size 10mL

Chromatogram identifier DEAE 291014003

No SDS-PAGE gel or MS results were obtained for this chromatogram for the reason outlined in Section 6.4.

6.5 Discussion of the results

6.5.1 Distribution HMGB proteins, and yields afforded by this method

HMGB1 proteins retained in cells are distributed between the cytosol, free in the nucleus, and attached to chromatin in the nucleus (Einck and Bustin, 1985). Scaffidi *et al.* (2002) show that HMGB1 is highly mobile in the nucleus, turning over rapidly between chromatin-bound and soluble states (see also Chapter 5 discussion). Although some HMGB1 will have been retained on chromatin, and not all the protein solution will have been extracted from the nucleus, it was expected that the extracted cePNE1' proteins would contain a significant proportion of the HMGB1 nuclear population; this is supported by calculations below. Large peaks in Figures 6.2 and 6.4 also indicate a significant amount of HMGB2 and HMGB3 in the cePNE1' protein group, providing evidence that HMGB2 and HMGB3 are likewise prevalent and have high mobility in the chicken erythrocyte nucleus.

Some of the previous HMGB production methods, as described in Section 6.2, discard any HMGB proteins which are not bound to chromatin, and so may be less efficient than the method described here. On the other hand, cells which have gone into apoptosis retain HMGB1 on the chromatin (Scaffidi *et al.*, 2002), so the relative efficiency of the method described here will depend on the proportion of cells in apoptosis at the time of extracting the cePNE1' proteins. The extent of this was not determined, but the proportion of cells in apoptosis may be low due to maintaining cold conditions, a short time (hours) between collecting the fresh chicken blood, and prevention of blood clotting (by the addition of tri-sodium citrate).

As shown by Figures 6.2 - 6.4, DEAE anion-exchange chromatography achieves unique selectivity in isolating almost pure HMGB proteins from 200 proteins in the cePNE1' group identified by nanospray MS (Appendix 2 of Zhuang, 2011). The area under the HMGB1 peak in Figure 6.2 (Fraction C) is difficult to judge, but perhaps equals ~3 Au.mL. Table 6.3 shows that the extinction coefficient for HMGB1 is approximately $0.8 \text{ (mg/mL)}^{-1}/\text{cm}$. On this basis, the yield of HMGB1 in the experiment was ~2.4 mg. However, the experiment as conducted allowed some loss of HMGB1. For example, only two from three extraction washes of chick-cell nuclei were retained to provide the cePNE1' extract.

HMGB protein yield could have been increased by utilising all three washes, and by adding further washes, taking care to avoid nuclei lysis. It is judged that these steps could have increased yield by a factor of about 2, leading to yield of HMGB1 of ~4.8mg. This compares with about 300Thmg of pure histone octamers which can be purified with high yield from a similar quantity of nuclei, $\sim 5 \times 10^{11}$ nuclei (Zhuang *et al.*, 2014). As HMGB1 is about a quarter the molecular weight of a histone octamer (Table 6.3), this yield represents about one molecule of HMGB1 for every 15 extracted octamer molecules. This compares to one HMGB1 molecule per ~ 10 nucleosomes thought to be present in cell nuclei (Stros, 2010), although this ratio varies depending on species and cell type (Einck and Bustin, 1985). The area under the HMGB2/HMGB3 peak (Fractions A, B in Figure 6.2) is about double the HMGB1 area. Together with the SDS-PAGE result in Figure 6.3, this suggests that the method also isolates ~ 2.4 mg each of HMGB2 and HMGB3. In a future experiment, this could be confirmed by running the combined A and B Fractions in Figure 6.2 on a cation-change column. Figures 3 and 4 in Foulger *et al.* (2012) show that in the right conditions, cation-exchange chromatography efficiently separates HMGB2 from HMGB3.

Provided potential losses of material can be correctly taken into account (these could be estimated by using single proteins to calibrate the efficiency of the method), the isolation method described here could be extended to investigate the partitioning of HMGB1 proteins between the cytosol, free in the nucleus, and attached to chromatin in the nucleus. This partitioning is important, not least since HMGB1 has a role in the immune system (see Tables 2.9 and 2.10, Chapter 2), requiring HMGB1 to be released into the extracellular environment. Moreover, previous experimentally measured ratios of cytoplasmic to nuclear HMGB1 have varied considerably, from about 1 to 2.9 (Einck and Bustin, 1985). As indicated in Figure 6.5, HMGB1 in the whole cell could be measured by applying anion-exchange chromatography to the whole cell lysate, after removal of DNA, RNA, and debris, followed by high salt conditions to dismantle the chromatin and precipitate less soluble proteins. Likewise, HMGB1 attached to chromatin alone could be measured by applying the same method to the nuclei depleted of cePNE1' proteins. The method could also be extended to establish the quantity and location of HMGB2 and HMGB3.

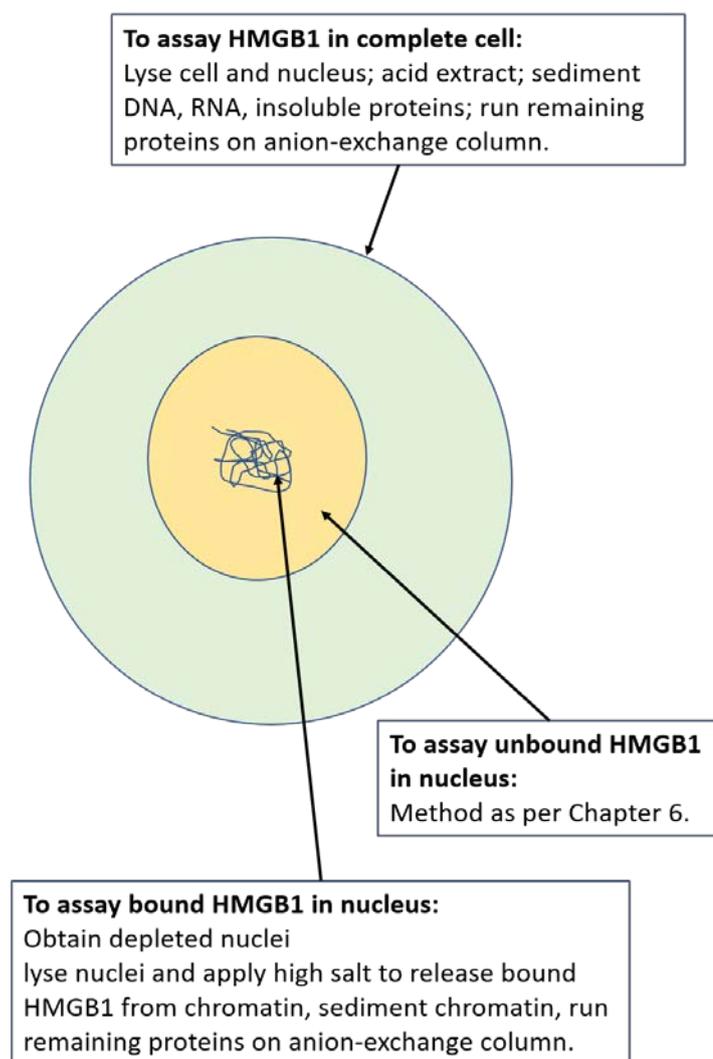


Figure 6.5 Scheme for investigating the partitioning of HMGB1 proteins between the cytosol, free in the nucleus, and attached to chromatin

This figure illustrates how the quantities of HMGB1 in different parts of the cell may be assayed. If the total HMGB1 in the complete cell is T, in the cytosol is C, free in the nucleus is F, and bound to chromatin is B, then $C = T - F - B$.

Table 6.3 Absorbance and MWs of chicken HMGB and histone proteins

Protein	Uniprot identifier	Total residues	MW (kDa)	Extinction Coefficient at 278nm (mg/mL) ⁻¹ /cm
HMGB1	Q9YH06	215	24.9	0.792
HMGB2	P26584	207	23.8	0.887
HMGB3	P40618	202	23.1	0.856
H5	P02259	190	20.7	0.203
H1	P09987	218	21.9	0.064
H2A	P02263	129	13.9	0.301
H2B	P0C1H5	126	14.0	0.501
H3	P84229	136	15.4	0.273
H4	P62801	104	11.4	0.493

MW and Extinction Coefficient were calculated using the Scripps Protein Calculator. The Calculator equates Absorbance with the Extinction Coefficient, and estimates the Extinction Coefficient by the method of Gill and von Hippel (1989) in which lyophilized proteins were used to establish an absorbance curve based on the number of tryptophans, tyrosines, and disulphide bonds.

6.5.2 Evidence of dipolar nature of HMGBs

The estimated pI of HMGB1 is 5.73 (Table 6.4). In a pH6.8 solution, the net charge on HMGB1 is negative, and so HMGB1 will bind to the positive charges on the stationary phase of the anion-exchange column. However, the net charge on HMGB1 is small at pH6.8 (Table 6.4), so strong binding would not be expected. In any event, HMGB2 and HMGB3 have estimated pIs of 8.38 and 8.12 respectively, and so have positive net charges at pH6.8 (Table 6.4). Consequently, they would not be expected to bind to the anion-exchange column. In contrast, all three HMGB proteins bind to the column.

At pH5.6, even HMGB1 would not be expected to bind to the anion-exchange column. The estimated pI of HMGB1 is just above 5.6, and so HMGB1 has a small net positive charge (Table 6.4). In contrast, not only does HMGB1 bind to the column, but the two main peaks in the anion-exchange chromatogram at pH5.6 are shifted to the right compared with the chromatogram at pH6.8, showing that at pH5.6 a higher proportion of HMGBs are strongly bound to the anion-exchange column.

The above results clearly show that the net charges and associated estimated pIs for the complete HMGB proteins are not responsible for the binding behaviour of HMGBs to an anion-exchange column. A credible explanation is that the HMGB molecules are acting as dipoles, so that their strongly acidic tails bind to the basic (+ve) anion-exchange stationary phase (Figure 6.6). Consistent with this, HMGB1 with the longest C-terminal acidic tail (30 acidic residues) binds more strongly than the other HMGB proteins (C-terminal acid tail 20-21 acidic residues).

As positive ions increased (due to the more acidic conditions at pH5.6) the HMGB proteins counter-intuitively bound more strongly to the positive charge on an anion-exchange column, compare with pH6.8. However, this could be explained as follows. As indicated by Table 6.4, in these circumstances the HMGB C-terminal acidic tails remained very strongly negative, although slightly less negative. Given the increase in positive ions in the eluent, it is likely that some other negative ions previously bound to the positive column stationary phase were removed, effectively preventing them from competing with the HMGB proteins for binding sites on the column. Consequently, the HMGB C-terminal acidic tails bound more strongly to the column.

Table 6.4 Chicken erythrocyte HMGB pIs and charges

Protein	Uniprot identifier	pH	Estimated pI	Net charge
HMGB1	Q9YH06	6.8	5.73	-3.9
As above	As above	5.6	As above	+1.0
30 residue C-terminal acidic tail	As above	6.8	-	-29.9
30 residue C-terminal acidic tail	As above	5.6	-	-28.2
HMGB2	P26584	6.8	8.38	+4.4
As above	As above	5.6	As above	+9.3
22 residue C-terminal acidic tail	As above	6.8	-	-21.0
22 residue C-terminal acidic tail	As above	5.6	-	-19.8
HMGB3	P40618	6.8	8.12	+2.9
As above	As above	5.6	As above	+5.6
21 residue C-terminal acidic tail	As above	6.8	-	-21.0
21 residue C-terminal acidic tail	As above	5.6	-	-19.8

Data calculated using Scripps Protein Calculator (2017) for the pH values as specified in the table. pI (Isoelectric Point) is an approximate value since assumes all residues have pKa values that are equivalent to the isolated residues, for a folded protein this is not valid. Similar considerations apply to the calculated charge, only part of which will be exposed to another molecule.

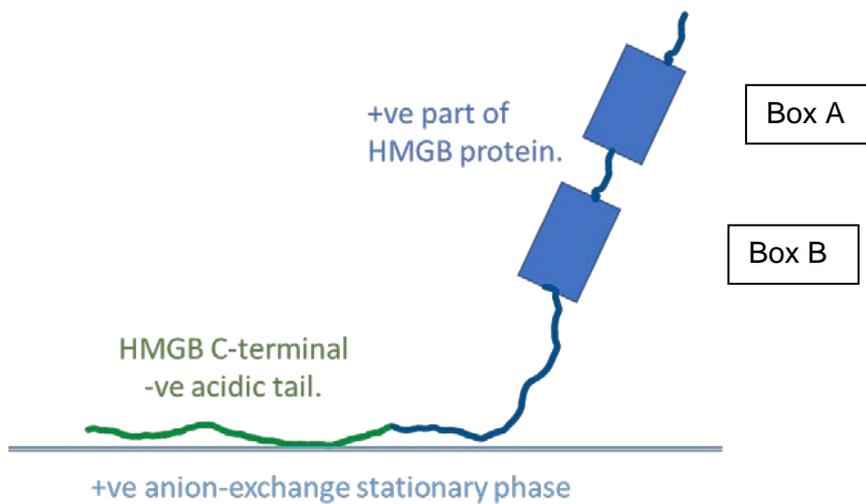


Figure 6.6 Proposed arrangement for attachment of HMGB protein to anion-exchange column

6.5.3 Effective isolation of HMGB proteins

The strengths of the gel bands in Figure 6.3 indicate that the two main peaks in the anion-exchange chromatogram at pH6.8 contain almost pure HMGB proteins. Of the HMGB proteins, HMGB1 apparently has the most important functions: these include its interaction with chromatin, its use as an extra-cellular signal, its role in disease (Kang *et al.*, 2014), and its necessity for life (Stros, 2010). It is particularly useful, therefore, that this method fully separates HMGB1 protein from the other two HMGB proteins. The method does not separate HMGB2 from HMGB3 – this is no surprise if they bind to the column only via their acidic tails (as shown in Table 6.4, the net charges on their respective tails are identical at pH6.8 and pH5.6). Fortunately, work by Foulger *et al.* (2012) indicates that a mixture of HMGB2 and HMGB3 can easily be separated by cation-exchange chromatography.

6.5.4 Impact on pH on the relative proportions of the three HMGB proteins

Examination of the size of the peaks in the anion-exchange chromatograms in Figures 6.2 and 6.4 shows that the relative proportions of the three HMGB proteins change according to pH. It is not known whether this effect is due to differing proportions of HMGB proteins in the cePNE1' protein group, or due to differing column binding efficiency.

6.6 Conclusions

1. The method described here provides a quick and efficient means of isolating HMGB proteins from the other nuclear proteins in chicken erythrocytes; all steps can be completed within two days from collecting chicken blood or (in the future) collecting mammalian leukocytes. Unlike some other methods, it does not require use of an ultracentrifuge, and it avoids acid degradation of HMGB PTMs. Thus, it provides a valuable resource for HMGB studies (for example, as described in Chapter 7). In particular, it complements a long term strategy of our research group, which is aimed at facilitating structural studies of HMGB proteins in complex with histone octamers (as a simplified model) and nucleosomes.

2. The results obtained with this method strongly suggest that the HMGB "flick knife" structure has opened, releasing the acidic tail to bind to the anion-exchange positive (basic) stationary phase.

3. The HMGB preparation method described here could be extended to investigate the partitioning of HMGB proteins between the cytosol, free in the nucleus, and attached to chromatin in the nucleus.

4. Foulger *et al.* (2012) showed the presence of HMGB3 in chicken erythrocyte nuclei, and the work here indicated its presence in significant quantities (although there is a small possibility that the molecule was a truncated form of HMGB2, given the ambiguity of the MS results shown in Table 6.2). The Human Protein Atlas indicates that high levels of HMGB3 are mainly found in human tissue associated with the lung or reproduction, and HMGB3 is also found in blood plasma. Thus, the presence of HMGB3 may provide some clue to its function in chicken erythrocytes; the human lung, human blood, and chicken erythrocytes are all involved in oxygen transfer.

Chapter 7

Research applications for purified native HMGB1

<u>Contents</u>	Page
7.1 Introduction	186
7.2 Previous studies	187
7.3 Example 1: use of an excess of purified HMGB1 to probe for potential complexes with other proteins	189
7.4 Example 2: use of purified HMGB1 to provide an experimental platform for developing cheap and quick methods to measure HMGB1 in biofluids	205
7.5 Example 3: proposed use of purified HMGB1 to form a complex with the histone octamer suitable for structural studies	216
7.6 Example 4: proposed use of purified HMGB1 to explore staining suitable for electron microscopy of a complex involving HMGB1 and chromatin	220
7.7 Conclusions and recommendations	221
<u>Tables</u>	
Table 7.1 Potential application of HMGB1 measurement in bio-fluids to cancer diagnosis or prognosis	188
Table 7.2(a) Summary of MS results showing the principle components of three bands in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)	192
Table 7.2(b) Detailed MS result for the protein in the upper faint band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)	193
Table 7.2(c) Detailed MS result for the protein in the strong middle band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)	194
Table 7.2(d) Detailed MS result for the protein in the third band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)	195
Table 7.3 Detailed MS result for the protein in the strongest band in the SDS-PAGE gel (Figure 7.4) for Fraction 3 from the cation-exchange chromatogram at pH4.5 (Figure 7.3)	199

<u>Tables continued</u>	Page
Table 7.4 Chicken erythrocyte HMGB1, FKBP3, Cyp B: MWs, residue number, pls, and net charges	200
Table 7.5 MS running conditions for gel band W8 (Figure 7.7) from Peak 2 of the anion-exchange chromatogram of plasma spiked with HMGB1 (Figure 7.6)	209
Table 7.6 Concentrations of proteins in plasma and serum	211
Table 7.7 Progressive increase in HMGB1 concentration in human blood serum with stages of gastric cancer	211
<u>Figures</u>	
Figure 7.1. Cation-exchange chromatogram with ionic strength/pH gradient, using HMGB1 to select for complexes from a subset of cePNE1' proteins	190
Figure 7.2 SDS-PAGE gel of the proteins in Fractions 1 and 2 from the cation exchange chromatogram (Figure 7.1)	191
Figure 7.3 Cation-exchange chromatogram with fixed pH4.5, using HMGB1 to select for complexes from a subset of cePNE1' proteins	197
Figure 7.4 SDS-PAGE gel of the proteins in Fraction 3 from the cation exchange chromatogram at pH4.5 (Figure 7.3)	198
Figure 7.5 Proposed arrangement for attachment of FKBP3/HMGB1 to cation-exchange stationary phase	202
Figure 7.6. Anion-exchange chromatogram of plasma spiked with HMGB1	207
Figure 7.7 SDS-PAGE gel of Fractions 1 and 2 from the anion-exchange chromatogram of plasma spiked with HMGB1 (Figure 7.6)	208
Figure 7.8 Octamer structure from crystallized histone octamer superimposed on DNA structure from crystallized nucleosome core particle	218
Figure 7.9 Comparison of DNA/octamer interactions in a crystallized nucleosome core particle with phosphate/octamer interactions found in the crystallized octamer	219

7.1 Introduction

Chapter 6 demonstrated an efficient method for isolating HMGB1 from the other nuclear proteins in the nuclei of chicken erythrocytes without degrading the acid-labile PTMs. This chapter examines some of the uses to which this molecule may be put.

In the first example, an excess of HMGB1 was used to probe for potential complexes with other proteins. Assembly of a complex is a dynamic process, with the amount of complex produced being the result of the balance of reaction rates for complex association and complex dissociation. Adding a plentiful supply of a HMGB1 to a target mixture increases the likelihood that HMGB molecules will encounter the other molecules required for a complex and will drive the equilibrium mixture towards a larger amount of that complex.

In the second example, HMGB1 was added to chicken blood plasma, to provide an experimental platform for developing methods to sensitively detect HMGB proteins in plasma and other bio-fluids quickly and at low cost. As discussed below, the determination of HMGB proteins in bio-fluids has great potential in disease diagnosis and prognosis, including for cancer. This work extends that described by Foulger (2012).

The third example is theoretical. It is known that HMGB1 interacts with DNA and chromatin, and that the HMGB1 C-terminal acid tail binds to the N-terminal tail of the H3 histone in the octamer (see Chapter 5 Section 5.2), but there remains uncertainty about the precise functioning of HMGB1 in relation to chromatin. It is proposed that HMGB1 be combined with histone octamers, to seek a complex which can be crystallized.

The fourth example is also theoretical. It is proposed that the purified native HMGB1 prepared as described in Chapter 6 could be used as a substrate for developing a heavy metal stain suitable for electron microscopy studies of HMGB1 complexes, particularly with chromatin.

Although this chapter is focussed on HMGB1, in general the approaches described here could also be applied to native purified HMGB2 and HMGB3, also obtained as described in Chapter 6.

7.2 Previous studies

HMGB1 is already the most studied of the HMGB proteins, but there is great scope for further investigation of this protein. HMGB1 is essential for mammalian life (although cells can survive the absence of Hmgb1, mice lacking HMGB1 died shortly after birth – Ronfani *et al.*, 2001), yet the fundamental mechanisms involving HMGB1 are still to be fully explained. HMGB1 is the most abundant of the widely expressed HMGB nuclear proteins, and it influences multiple nuclear processes including transcription, replication, recombination, and DNA repair – see Chapter 2 Section 2.6. These processes generally involve chromatin, and association of HMGB1 with chromatin is highly dynamic (Scaffidi *et al.*, 2002). Transient interactions with nucleosomes include displacement of histone H1 (Ju *et al.*, 2006), and facilitation of nucleosome remodelling (Bianchi and Agresti, 2005). HMGB1 also organises the cellular stress response. It provides an extracellular signal as an alarmin (also called a damage associated molecular pattern molecule - DAMP); in conjunction with other proteins it is involved in cytokine activity, and hence the inflammatory and immune response (Gibot *et al.*, 2007; Cohen *et al.*, 2009; Abraham, 2009; Diener *et al.*, 2013).

Kang *et al.* (2014) reviewed about 2000 papers related to HMGB1 in health and disease, identifying more than 40 intracellular binding partners, and 14 extracellular receptors including RAGE and several TLR receptors. They identified about 70 papers which describe HMGB-targeting therapeutic strategies for treatment of sepsis. They indicate the value of further investigating HMGB1 localization, structure, post-translational modification (PTM), and identification of additional partners.

Kang *et al.* (2014) identified many papers indicating that HMGB1 measurement has a role in diagnosis and/or prognosis for numerous diseases, including cancer. Table 7.1 lists some cancers for which HMGB1 measurement in biofluid may be relevant. In most studies, serum was selected as the biofluid. Serum is obtained as a supernatant after centrifuging clotted blood. However, for convenience, plasma was selected as the biofluid for the initial study described here (Example 2 below). Plasma contains a few more proteins than serum because it is obtained by centrifuging blood which has not clotted and therefore retains more clotting factors. Various studies have sought to improve methods for such measurements in biofluids. Shi *et al.* (2012) improved the quantification of very small quantities of biomarkers in human plasma/serum, but required relatively expensive

equipment including capillary liquid chromatography facilities and two triple quad mass spectrometers. To improve the detection and measurement of low level biomarkers, Chertov *et al.* (2004) successfully applied solvent precipitation to selectively remove unwanted proteins from serum. More recently, Geyer *et al.* (2016) have examined the use of MS to identify and quantify many hundreds of proteins in human plasma and have developed an MS based system which can process, identify, and quantify the concentrations of ~300 proteins from a single 5 μ L blood sample within 3 hrs, or 1040 proteins if the MS measurement time is extended to 16hrs. The relevance of this to our work is discussed in Section 7.4.3.

Table 7.1 Potential application of HMGB1 measurement in bio-fluids to cancer diagnosis or prognosis

The examples below are based on studies reviewed by Kang *et al.* (2014).

Cancer which was studied	Biofluid	Potential benefit from measuring HMGB1 in the biofluid	References
NSCLC	Serum	Prognosis	References to the applicable studies are listed in Table 5 of Kang <i>et al.</i> (2014).
Squamous cell	Serum	Diagnosis, Prognosis	
Pancreatic	Serum	Diagnosis, Prognosis	
Colorectal	Serum	Diagnosis, Prognosis	
Gastric	Serum	Diagnosis, Prognosis	
Oesophagus	Serum	Prognosis	
Hepatocellular	Serum	Prognosis	
MPM	Serum	Prognosis	
ALC	Serum	Prognosis	
Peritoneal	Serum	Diagnosis	
Head and neck	Serum	Prognosis	
Breast	Plasma, serum	Prognosis	

7.3 Example 1: use of an excess of purified HMGB1 to probe for potential complexes with other proteins

7.3.1 Investigation of potential complexes with HMGB1 using cation-exchange chromatography with an ionic strength-pH gradient.

HMGB1 was mixed with a group of other nuclear proteins, and was subjected to cation-exchange chromatography, to probe for proteins in the group which might form a complex with HMGB1. This was a proof of concept experiment.

HMGB1 in this experiment was prepared using an anion-exchange column as described in Chapter 6 (the HMGB1 was that contained in Peak C in Figure 6.2). As indicated in Chapter 6, HMGB1 is strongly represented in the cePNE1' proteins, and it was not considered necessary to enhance the concentration of HMGB1 for the work described here, although this could easily have been done. The group of proteins being probed for potential HMGB1 partners was a subset of the cePNE1' proteins, as contained in the non-bind fractions (the large peak on the left hand side of Figure 6.2) on the same anion-exchange column used to isolate HMGB1. This group contained only those cePNE1' proteins which were positively charged. In addition, the group excluded any unfolded HMGB proteins since their dipolar configuration caused them to bind to the anion-exchange column, later eluting in Fractions A, B and C in Figure 6.2.

HMGB1 and the package of non-bind proteins (the cePNE1' subset) were combined, dialysed into cation-exchange Buffer A (pH4.5, 100mM KCl, 100mM H₂KPO₄), and concentrated to 100mL using an Amicon pressurised concentrator at 2bar with a 10kDa MWCO membrane. The sample was then injected onto a cation-exchange column and eluted with an increasing proportion of Buffer B (pH6.8, 600mM KCl, 50mM H₂KPO₄, 50mM HK₂PO₄), providing a linear ionic strength gradient combined with a pH gradient.

The chromatogram (Figure 7.1) exhibited a sharp peak at about 250mM KCl, pH6.05. Fractions 1 and 2 were taken from the left and right hand sides of the peak, and both fractions were subject to SDS-PAGE (Figure 7.2). Bands were cut out of the SDS-PAGE gel for Fraction 2, prepared for MS as per Chapter 3, and sent to Sheffield University for MS. The MS results are summarised in Table 7.2. Fraction 1 contained both HMGB1 and FKBP3; Cyp B was also present, but its advanced position on the gel indicated it was substantially truncated. Fraction 2 also contained HMGB1 and the truncated Cyp B, but only a weak band for FKBP3.

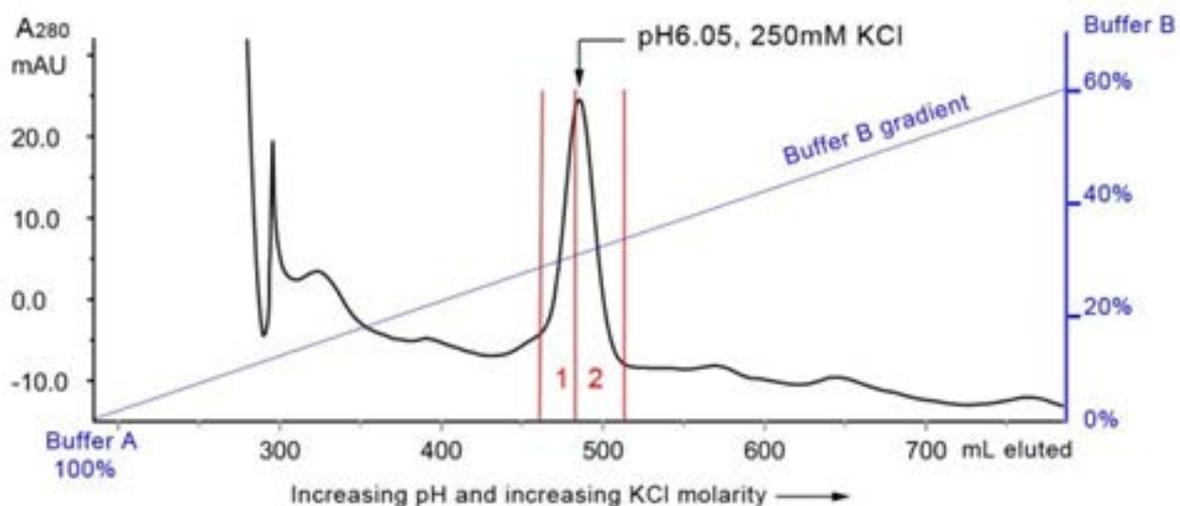


Figure 7.1. Cation-exchange chromatogram with ionic strength/pH gradient, using HMGB1 to select for complexes from a subset of cePNE1' proteins

A linear ionic strength/pH gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%;

At end of elution (1200mL): Buffer A 0%, Buffer B 100%.

Buffer A (pH4.5)	100mM KCl, 100mM H ₂ KPO ₄
Buffer B (pH6.8)	600mM KCl, 50mM H ₂ KPO ₄ , 50mM HK ₂ PO ₄
Cation-exchange columns	8 x HiTrap SP Sepharose FF 5mL (GE Healthcare)
Total column volume	40mL
Total elution volume	1200mL (120 x 10mL fractions)
Flow rate	1.5mL/min
Chromatogram identifier	SP301014003
Fractions subject to assay	1, 2 (SDS-PAGE gel is shown in Figure 7.2)

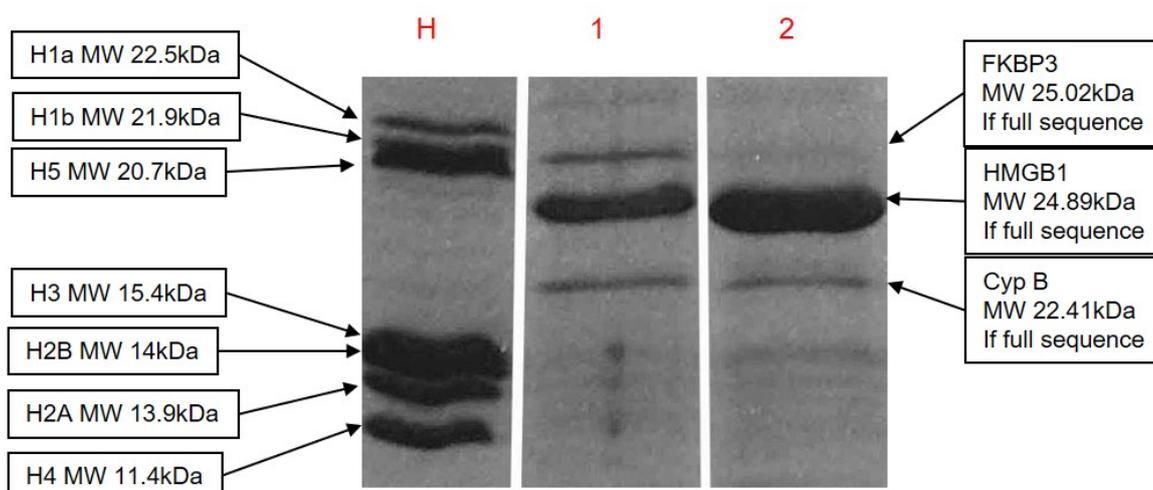


Figure 7.2 SDS-PAGE gel of the proteins in Fractions 1 and 2 from the cation exchange chromatogram (Figure 7.1)

Stain:	Coomassie Blue (although photographed in black and white)
Stacking gel:	5% acrylamide
Resolving gel:	20% acrylamide
Running buffer:	glycine/tris/SDS
Sample treatment:	1 part 3X Buffer to 2 parts sample
Sample loading:	adjusted to provide visible and separate bands
SDS-Page protocol	as described in Chapter 3

H: Histone Standard (see Chapter 4 for more details)

1: Fraction 1 from the cation-exchange chromatogram (Figure 7.1)

2: Fraction 2 from the cation-exchange chromatogram (Figure 7.1)

Proteins in Fraction 2 were identified by MS (Table 7.2).

Table 7.2(a) Summary of MS results showing the principle components of three bands in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)

Band in Lane 2	Protein (<i>Gallus gallus</i>)	Coverage of full sequence by fragments found by MS	MASCOT Score	Full sequence
Weak band	FKBP3 GI:46048916 (NCBINR)	55%	695	Table 7.2(b)
Strong band	HMGB1 GI45382473 (NCBINR)	59%	1016	Table 7.2(c)
Lower distinct band	Cyclophilin B P24367 (SwissProt)	55%	691	Table 7.2(d)

Table 7.2(b) Detailed MS result for the protein in the upper feint band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)

The following sequences were found by MS and mapped to the full FKBP3 sequence as below. Ignoring gaps, the sequences found run from K28 to E227 (this sequence of 200 residues has a molecular weight of 22.15kDa - Scripps Protein Calculator).

R.LLGQVK.N
 K.AAKPLSFK.V
 K.KAAKPLSFK.V
 K.AEEPAAEAGPPK.Y
 R.GWDEALLTMSK.G
 K.LFFEVELVDIE.
 K.AQLEIEPEWAYGK.K
 K.LQDGTVFDTNVQTSSK.K
 K.FLQEHAAQAF LAEHR.L
 K.EQLIAAYTQLFHTQR.F
 K.TANKEQLIAAYTQLFHTQR.F

Sample	Protein	Identifier	Mass (Da)	Score	Coverage
Fraction 2 Upper feint band	FKBP3 (<i>Gallus gallus</i>)	GI:46048916 (NCBINR)	25016	695	55%
227 residues full sequence, with 126 residues found by MS shown in red below.					
1 MAAATAPAQP WSAEELRSEA LPKKDIIKFL QEHAAQAF LA EHRLLGQVKN VAKTANKEQL 61 IAAYTQLFHT QRFKGTGAE RAAEKAKPGK AEKEKDKA AKAEPAEEG PPKYTKSILK 121 KGDKTNFPFK GDTVHCWYTG KLQDGTVFDT NVQTSSK KKK AAKPLSFKVG VGKVI RGWDE 181 ALLTMSKGEK AQLEIEPEWA YGKKGQDAK IPPNAKLFFE VELVDIE					

Other gallus gallus proteins found in the same band (scores in brackets) were histone H5 (145), HMGB1 (110), and 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase eta-2 (74).

Source of MS data	Mark Dickman Group, Sheffield University
Search title	ncbinr KT 1-26 131214 (C:\ProgramData\Matrix Science\Mascot Daemon\parameters\Alison gallus Amazon.par), submitted from Daemon on MJDHOTDESK-PC
MS data file	D:\OAN Raw data\ALISON DEC 2014 KT\KT1-26 DEC 2014\KT6_RC6_01_1228.d\KT6_RC6_01_1228.mgf
Database 1	A-mel 20131104 (10583 sequences; 5145280 residues)
Database 2	NCBINr 20140730 (47310802 sequences; 16880015641 residues)
Taxonomy 2	bony vertebrates (4148099 sequences)
Timestamp	6 Jan 2015 at 20:36:51 GMT
Enzyme	Trypsin
Variable modifications	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	: ± 0.6 Da
Fragment Mass Tolerance	: ± 0.6 Da
Max Missed Cleavages	2
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

Table 7.2(c) Detailed MS result for the protein in the strong middle band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)

The following sequences were found by MS and mapped to the full HMGB1 sequence as below. Ignoring gaps, the sequences found run from K12 to E164 (this sequence of 153 residues has a molecular weight 17.78kDa - Scripps Protein Calculator).

K.DIAAYR.A
 K.FKDPNAPK.R
 K.GKFEDMAK.A
 K.KFKDPNAPK.R
 K.YEKDIAAYR.A
 K.NYVPPKGETK.K
 K.NYVPPKGETKK.K
 K.GEHPGLSIGDVAK.K
 K.MSSYAFFVQTCR.E
 K.HPDASVNFSEFSK.K
 K.IKGEHPGLSIGDVAK.K
 K.KHPDASVNFSEFSK.K
 K.KLGEMWNNTAADDK.Q
 K.IKGEHPGLSIGDVAKK.L
 K.KKHPDASVNFSEFSK.K
 K.KHPDASVNFSEFSK.C
 K.KKHPDASVNFSEFSK.K
 K.RPPSAFFLFCSEFRPK.I
 K.LGEMWNNTAADDKQPYEK.K
 K.LGEMWNNTAADDKQPYEKK.A
 K.KLGEMWNNTAADDKQPYEK.K

Sample	Protein	Identifier	Mass (Da)	Score	Coverage
Fraction 2 Strong central band	HMGB1 (<i>Gallus gallus</i>)	GI:45382473 (NCBINR)	24893	1016	59%
215 residues full sequence, with 126 residues found by MS shown in red below.					
1	MGKGDPKKPR	GKMSSYAFFV QTCREHKKK	HPDASVNFSE	FSKKC	SERWK TMSSEKGGKF
61	EDMAKADKLR	YEKEMKNYVP	PKGETKKKFK	DPNAPKRPPS	AFFLFCSEFR PKIKGEHPGL
121	SIGDVAKKLG	EMWNNTAADD	KQPYEKKA	LKEKYEKDIA	AYRAKGVDA GKKVVAEK
181	SKKKKEEEEED	EDEDEDEEED	EEEEEEEEED	DDDDE	

Other gallus gallus proteins found in the same band (scores in brackets) were alpha-globin-A (196), hemoglobin alpha-A chain (151), elongation factor 1-alpha 1 (93), and HMGB3 (70).

Source of MS data	Mark Dickman Group, Sheffield University
Search title	ncbinr KT 1-26 131214 (C:\ProgramData\Matrix Science\Mascot Daemon\parameters\Alison gallus Amazon.par), submitted from Daemon on MJDHOTDESK-PC
MS data file	D:\OAN Raw data\ALISON DEC 2014 KT\KT1-26 DEC 2014\KT7_RC7_01_1230.d\KT7_RC7_01_1230.mgf
Database 1	A-mel 20131104 (10583 sequences; 5145280 residues)
Database 2	NCBINr 20140730 (47310802 sequences; 16880015641 residues)
Taxonomy 2	bony vertebrates (4148099 sequences)
Timestamp	6 Jan 2015 at 20:36:51 GMT
Enzyme	Trypsin
Variable modifications	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	: ± 0.6 Da
Fragment Mass Tolerance	: ± 0.6 Da
Max Missed Cleavages	2
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

Table 7.2(d) Detailed MS result for the protein in the third band in the SDS-PAGE gel (Figure 7.2) for Fraction 2 from the cation-exchange chromatogram (Figure 7.1)

The following sequences were found by MS and mapped to the full Cyp B sequence as below. Ignoring gaps, the sequences found run from K36 to K172 (this sequence of 137 residues has a molecular weight 15.04kDa - Scripps Protein Calculator).

K.GFGFK.G
 K.HVVFGK.V
 R.VVIGLFGK.T
 R.FPDENFK.L
 K.VLEGMDVVR.K
 K.DFMIQGGDFTR.G
 K.TVENFVALATGEK.G
 K.DTNGSQFFITTVK.T
 K.HYGPWVSMANAGK.D
 K.SIYDRFPDENFK.L
 R.VIKDFMIQGGDFTR.G

Sample	Protein	Identifier	Mass (Da)	Score	Coverage
Fraction 2 Lower band	Cyp B (PPIB) (<i>Gallus gallus</i>).	P24367 (SwissProt)	22410	691	55%
207 residues full sequence, with 115 residues found by MS shown in red below.					
<pre> 1 MKALVAATAL GPALLLLLP A SRADERKKG PKVTAKVFFD LRVGEEDAGR VVIGLFGKTV 61 PKTVENFVAL ATGEKGF GSKFHRVIKD FMIQGGDFTR GDGTGGKSIY GDRFPDENFK 121 LKHYGPWVS MANAGDTNG SQFFITTVKT AWLDGKHVVV GKVLEGMDVV RKVENTKTDS 181 RDKPLKDVTI ADCGTIEVEK PFAIAKE </pre>					

Other *Gallus gallus* proteins found in the same band (scores in brackets) were HMGB1 (323) H4 (230) Peroxiredoxin-1 (230), Hemoglobin subunit beta (114) Elongation factor 1-alpha 1 (111), H3 (67), H2B (56), and H5 (41).

<u>Source of MS data</u>	<u>Mark Dickman Group, Sheffield University</u>
Search title	Copy of Gallus KT 1-26 131214 (C:\ProgramData\Matrix Science\Mascot Daemon\parameters\Alison gallus Amazon.par), submitted from Daemon on MJDHOTDESK-PC
MS data file	D:\OAN Raw data\ALISON DEC 2014 KT\KT1-26 DEC 2014\KT8_RC8_01_1231.d\KT8_RC8_01_1231.mgf
Database	SwissProt 2014_07x (546000 sequences; 194259968 residues)
Taxonomy	bony vertebrates (83609 sequences)
Timestamp	6 Jan 2015 at 13:56:09 GMT
Enzyme	Trypsin
Variable modifications	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	: ± 0.6 Da
Fragment Mass Tolerance	: ± 0.6 Da
Max Missed Cleavages	2
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

7.3.2 Further investigation using cation-exchange chromatography at a fixed pH4.5

Figures 7.1 and 7.2 show HMGB1 co-fractionated with FKBP3 in a solution containing 250mM KCl at pH6.05. This potential complex was tested by preparing another cation-exchange chromatogram at a different pH. This was on the basis that it was unlikely that the two proteins would be affected identically by a change of pH, so they were unlikely to significantly co-fractionate again unless they were in a complex (although, as discussed in Section 7.3.3, this was not an effective strategy, after all). Thus, the experiment in Section 7.3.1 was repeated at pH4.5 throughout the elution, leading to the results in Figures 7.3 and 7.4, and Table 7.3. This time, a much smaller quantity of FKBP3 fractionated with HMGB1. As expected, shifting to a more acidic pH4.5 condition increased the positive charges on HMGB1 and FKBP3 (Table 7.4) so that they became more attached to the cation-exchange column, and consequently a higher ionic strength (605mM KCl) was required for their elution.

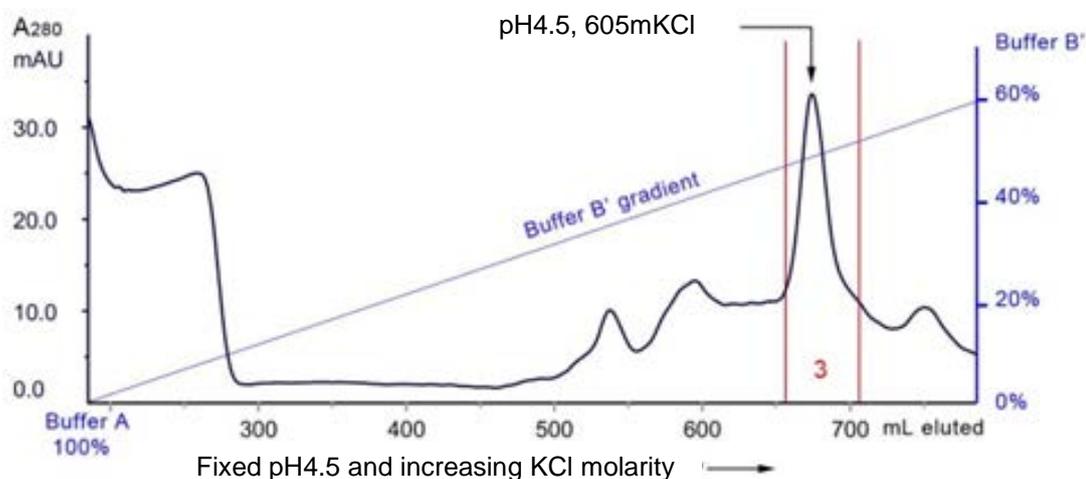


Figure 7.3 Cation-exchange chromatogram with fixed pH4.5, using HMGB1 to select for complexes from a subset of cePNE1' proteins

A linear ionic strength/pH gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%.

At end of elution (1200mL): Buffer A 0%, Buffer B 100%.

Buffer A	100mM KCl, 100mM H ₂ KPO ₄
Buffer B	850mM KCl, 100mM H ₂ KPO ₄
pH	4.5
Cation-exchange columns	8 x HiTrap SP Sepharose FF 5mL (GE Healthcare)
Total column volume	40mL
Total elution volume	1200mL (120 x 10mL fractions)
Flow rate	1.5mL/min
Chromatogram identifier	SP301014007
Fraction subject to assay	Peak 3

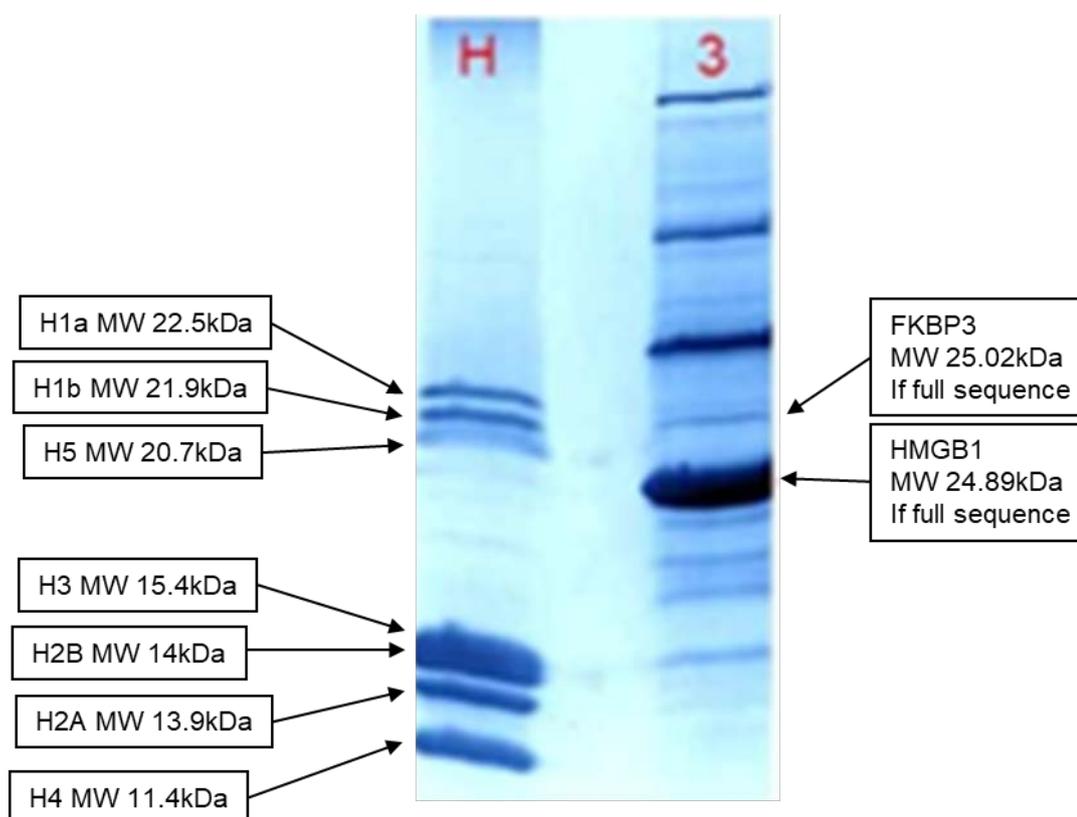


Figure 7.4 SDS-PAGE gel of the proteins in Fraction 3 from the cation exchange chromatogram at pH4.5 (Figure 7.3)

H: Histone Standard

3: Fraction 3 from the cation exchange chromatogram at pH4.5 (Figure 7.3)

HMGB1 was identified by MS (Table 7.3)

FKBP3 was identified by its position on the gel (no MS result available).

Stain: Coomassie Blue

Stacking gel: 5% acrylamide

Resolving gel: 20% acrylamide

Running buffer: glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol as described in Chapter 3

Table 7.3 Detailed MS result for the protein in the strongest band in the SDS-PAGE gel (Figure 7.4) for Fraction 3 from the cation-exchange chromatogram at pH4.5 (Figure 7.3)

The following sequences were found by MS and mapped to the full HMGB1 sequence as below. Ignoring gaps, the sequences found run from K29 to A164 (this sequence of 136 residues has a molecular weight of 15.7kDa - Scripps Protein Calculator).

K.YEKDIAAYR.A
 K.NYVPPKGETK.K
 K.HPDASVNFSEFSK.K
 K.KHPDASVNFSEFSK.K
 K.IKGEHPGLSIGDVAK.K

Sample	Protein	Identifier	Mass (Da)	Score	Coverage
Fraction 3 strongest band	HMGB1 (<i>Gallus gallus</i>)	GI:5815432 (NCBINR)	24778	373	26%
214 residues full sequence, with 56 residues found by MS shown in red below.					
<pre> 1 MGKGDPKKPR GKMSYAFFV QTCREEHKK HPDASVNFSE FSKKCSERWK TMSKEKGKF 61 EDMAKADKLR YEKEMKNYVP PKGETKKKFK DPNAPKRPPS AFFLFCSEFR PKIKGEHPGL 121 SIGDVAKKLG EMWNNTAADD KQPYEKKA LKEKYEKDIA AYRAKGVDA GKKVVAKA EK 181 SKKKKEEEEE EDEDEDEED EEEEEEEEE DDDE </pre>					

Another *Gallus gallus* protein found in the same band was FRG1 (score 108).

Source of MS data	Mark Dickman Group, Sheffield University
Search title	JB bands May 2016 (C:\ProgramData\Matrix Science\Mascot Daemon\parameters\Alison gallus Amazon.par), submitted from Daemon on MJDHOTDESK-PC
MS data file	D:\MJD\March 16\P 3.20_RC4_01_5613.mgf
Database 1	contaminants 20160129 (247 sequences; 128130 residues)
Database 2	NCBI nr 20140730 (47310802 sequences; 16880015641 residues)
Taxonomy 2	bony vertebrates (4148099 sequences)
Timestamp	17 May 2016 at 19:39:43 GMT
Enzyme	Trypsin
Variable modifications	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	± 0.6 Da
Fragment Mass Tolerance	± 0.6 Da
Max Missed Cleavages	1
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

Table 7.4 Chicken erythrocyte HMGB1, FKBP3, Cyp B: MWs, residue number, pIs, and net charges

Protein	Uniprot identifier	MW KDa	No of residues	Estimated pI	Net charge <i>e</i> (applicable pH)
HMGB1	Q9YH06	24.89	215	5.73	-1.6 (6.05)
Charge on HMGB1 excluding the 30 residue C-terminal acidic tail.	Q9YH06			9.86	+27.8 (6.05)
FKBP3	Q90ZK7	25.02	227	9.19	+10.7 (6.05)
Cyp B	P24367	22.41	207	9.39	+9.8 (6.05)
Cyp B with C-terminal tail removed from V173 onwards.	P24367	18.57	172	9.79	+11.6 (6.05)
HMGB1	Q9YH06	24.89	215	5.73	+22.8 (4.5)
Charge on HMGB1 excluding the 30 residue C-terminal acidic tail.	Q9YH06			9.86	+39.5 (4.5)
FKBP3	Q90ZK7	25.02	227	9.19	+25.6 (4.5)
Cyp B	P24367	22.41	207	9.39	+21.5 (4.5)

Notes

1. Data calculated using Scripps Protein Calculator (2017). pI (Isoelectric Point) is an approximate value since assumes all residues have pKa values that are equivalent to the isolated residues, for a folded protein this is not valid. Similar considerations apply to the calculated charge, only part of which will be exposed to another molecule.

2. The above proteins have been selected as typical examples, there are variants.

7.3.3 Discussion of the results of using HMGB1 to probe for potential complexes

The objective of Section 7.3 was to use HMGB1 (prepared as described in Section 6) to probe for potential complexes with other proteins. A cation-exchange chromatogram of HMGB1 mixed with a parcel of other proteins (prepared as described in Section 7.3.1) showed that HMGB1 eluted with FKBP3 and Cyp B at pH6.05 (Figures 7.1 and 7.2).

At pH6.05, FKBP3 has a net positive charge of +10.7e (Table 7.4), and so would be expected to bind moderately well to the negatively charged cation-exchange column. This was consistent with the chromatogram, which showed that FKBP3 was eluted when the buffer ionic strength was formed by 250mM KCl and 100mM phosphate. Similar binding would be expected, and was found, for Cyp B (Figure 7.2), given that this molecule has a net positive charge of +9.8e at pH6.05 (Table 7.4). Surprisingly, however, despite its slightly lower charge, the gel in Figure 7.2 shows that Cyp B bound slightly more strongly than FKBP3 - compare the gel bands for Fractions 1 and 2 which represent either side of the peak in the chromatogram.

As discussed in Section 5.4.3 (Chapter 5), Watson *et al.* (2007) have shown that at low ionic strength the HMGB1 C-terminal acidic tail is folded back onto the positive part of the molecule (including both HMGB boxes) most of the time, but is in dynamic equilibrium with an open dipolar configuration. HMGB1 has an estimated pI of 5.73, and a net negative charge of -1.6e at pH6.05 (Table 7.4). The net charge is likely to be appropriate to the folded HMGB1 molecules, which would therefore not have bound to the column. On the other hand, binding of HMGB1 to the cation-exchange column can be explained if the molecule has flicked open (perhaps facilitated by the strong negative charge on the column), with the positive part of HMGB1 (typically, residues 1 – 185) binding to the stationary phase on the column, due to its charge of about +27.8e at pH6.05 (Table 7.4). This charge is more than double the net charges on FKBP3 and Cyp B, so independent fully unfolded HMGB1 molecules would be expected to elute later than FKBP3 and Cyp B, which is not indicated by the results. Thus, the elution of HMGB1 together with FKBP3 and Cyp B suggests a complex, or complexes. It is proposed that the negative tails on numerous open HMGB1 molecules captured many of the FKBP3 molecules, for example by binding to the positive central part of FKBP3. This possibility for FKBP3 is illustrated by Figure 7.5; a similar explanation would apply to the co-elution of the Cyp B molecules.

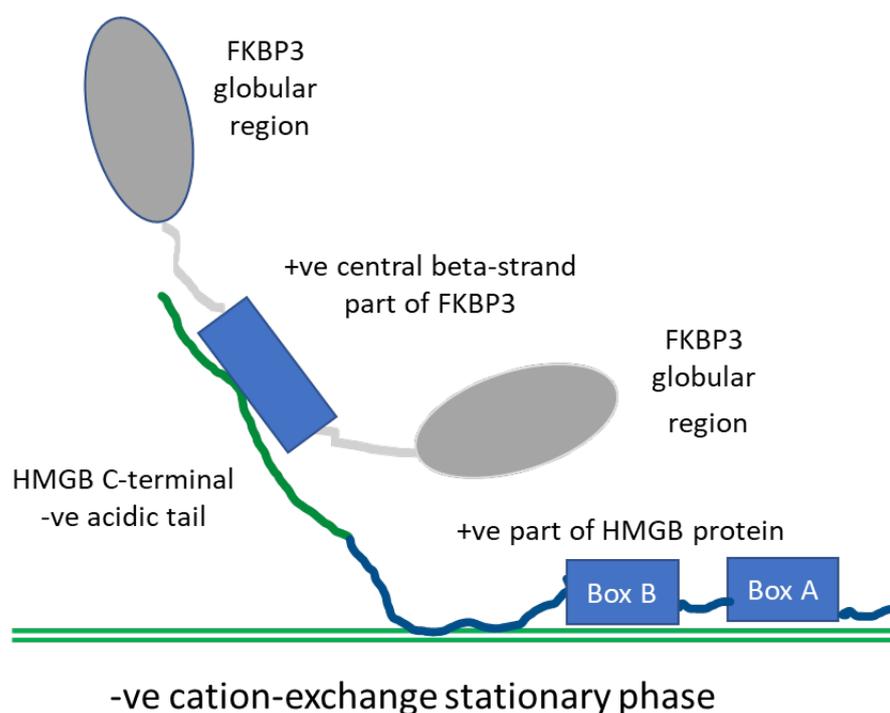


Figure 7.5 Proposed arrangement for attachment of FKBP3/HMGB1 to the cation-exchange stationary phase

In free solution, HMGB1 exists in a dynamic equilibrium between a closed (tail folded back onto the HMG boxes) and an open (tail free) configuration (Watson *et al.*, 2007). It is proposed that, in open configuration, the positive part of HMGB1 (residues 1 – 185) binds to the cation-exchange column, allowing the C-terminal acidic tail to capture FKBP3. Binding is most likely to the FKBP3 central region containing beta-strands, which is strongly basic.

Berg *et al.* (2015) present results from Weber and Osborn (1975) which show that the speed at which a typical protein migrates through an SDS-PAGE gel is approximately inversely proportional to the logarithm of the protein MW. However, this does not apply to all proteins. In particular, HMG proteins are called High Mobility Group proteins because they run faster on a gel than expected from their MW. In light of this, the gel result in Figure 7.2 is consistent with the presence of full length FKBP3 and HMGB1 proteins. However, the gel indicates that the Cyp B molecules were either truncated or, like HMGB1, ran anomalously fast on a gel. Table 7.2(d) shows that no residues were identified by MS for the Cyp B N-terminal section up to A35, nor were residues identified for the C-terminal tail from V173 onwards. These omissions could be an MS artefact (see discussion in Chapter 8 Section 8.4.3). However, if the C-terminal tail was removed by a protease, it would have increased the net positive charge on Cyp B to +11.6e at pH6.05 (Table 7.4) - this could account for Cyp B binding slightly more strongly to the column than FKBP3. In addition, the truncated protein MW would be 18.57kDa (Table 7.4) - this is consistent with the Cyp B gel mobility shown in Figure 7.2. The possibility of Cyp B truncation, and the means to prevent it, are further discussed in Section 8.4.3 and 8.4.4). [Note that the methods in both Chapter 7 and Chapter 8 led to the Cyp B proteins being subject to cation-exchange chromatography which omitted any protease, potentially leading to the truncation].

Figures 7.3 and 7.4 indicate that at pH4.5, FKBP3 eluted from the column in the same peak as HMGB1. At this pH, HMGB1 (folded) and FKBP3 would have bound similarly to the column, due to their similar net charges, +22.8e and +25.6e respectively (Table 7.4). Consequently, this does not shed light on the possible existence of an HMGB1/FKBP3 complex. It was not investigated whether any other band in Figure 7.4 was Cyp B.

Leclercq *et al.* (2000) refer to porcine HMG-II proteins A34719 and A28897 as forming a complex with porcine FKBP3. In modern nomenclature, A34719 and A28897 correspond to HMGB2 (Uniprot KB P17741) and HMGB1 (UniprotKB P12682) respectively. Leclercq showed that A34719 bound to FKBP3 on a 2-D gel. In the first (isoelectric-focussed) dimension, FKBP3 and HMGB were both present at a pI from 9.2 to 10.0, despite the HMGB protein having a substantially lower pI of 5.73 (Table 7.4). In the second (SDS-PAGE) dimension, sequencing confirmed both FKBP3 and the HMGB were present in the same band on the gel; unsurprising since their molecules have roughly the same molecular weight - FKBP3 and HMGB1 (UniprotKB P12682) both have a molecular

weight of ~25kDa. Unfortunately, Leclercq's result is difficult to interpret, and may not be relevant to biological conditions. Apart from the high pH in the isoelectric dimension, the samples were prepared in denaturing conditions provided by a 9.8M urea buffer.

Using a buffer comprising 100mM KCl and 100mM phosphates, Foulger *et al.* (2012) applied gel filtration chromatography at pH6.0 to a mixture containing HMGB1, HMGB3 and FKBP3 and found no complex involving these proteins. This compares with the elution conditions here of 250mM KCl and 100mM phosphates at pH6.05. Thus, the Foulger work did not support the possibility of a complex between HMGB1 and FKBP3 which we found here. However, it is possible that the higher ionic strength, or the presence of a strongly negative cation-exchange stationary phase is necessary before many of the HMGB proteins unfold to present a dipole which binds to FKBP3.

In summary, the result obtained here suggested that at pH6.05 HMGB1 forms a complex with FKBP3 and/or Cyp B, at least in the presence of a strong cation (negative) stationary phase, and that this may arise because HMGB1 C-terminal acidic tail unfolds from the rest of the molecule. Some previous work, for and against the existence of an HMGB1/FKBP3 complex as discussed above, is not conclusive since it was performed under different conditions.

The result also indicates that the Cyp B molecules prepared here suffered loss of their C-terminal from V173 onwards.

The model suggested here (ie an unfolded HMGB1 molecule binds to FKBP3) could be tested by several ways, including:

1. Run the cation-exchange chromatography again using the individual pure proteins, to rule out any co-fractionation arising solely from overlapping peaks of the individual proteins.
2. Running the cation-exchange chromatography again using a less positive cation-exchange stationary phase, to see if co-fractionation continues.
3. Applying gel filtration chromatography to mixture containing pure HMGB1 and FKBP3 in biological conditions (ionic strength ~300milliosmoles; pH7.1). If a complex exists it will elute earlier than the individual molecules.

7.4 Example 2: use of purified HMGB1 to provide an experimental platform for developing cheap and quick methods to measure HMGB1 in biofluids

7.4.1 Strategy

Some limited preliminary work was carried out using purified HMGB1 to investigate a proposed simple, low cost method for increasing the sensitivity with which HMGB1 can be measured in the bloodstream. Blood contains many proteins which could swamp small quantities of HMGB1. After separating blood plasma from fresh erythrocyte blood, the proposed method was based on removing a high proportion of these unwanted proteins, with the following steps:

Step 1. The less soluble proteins were precipitated in a high ionic-strength buffer. This selected HMGB proteins since they are highly soluble. The high ionic-strength had the added benefit of dissociating the HMGB proteins from any complexes.

Step 2. Anion-exchange chromatography was used to isolate HMGB1. This provided a second level of selectivity.

Step 3. This is the detection and quantification of HMGB1, for example by using a gel, or an antibody-based system, or MS. This step has not yet been carried out, but it will provide a third level of selectivity.

The preliminary work included spiking chicken plasma with a known quantity of HMGB1, to evaluate the precipitation step and to establish the conditions necessary to isolate HMGB1 using anion-exchange chromatography, and to establish the efficiency of the method.

7.4.2 Experimental work

Chilled buffers were prepared as follows:

Anti-Clotting Buffer	10% w/v tri-sodium citrate, 2.5mM benzamidine hydrochloride.
High ionic strength buffer pH6.8)	2000mM KCl, 700mM H ₂ KPO ₄ , 700mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride

Anion-exchange Buffers A and B pH6.8 buffers as specified in Figure 7.6

To prepare blood plasma, 6L of blood was collected into Anti-Clotting Buffer (~100mL buffer/L blood) from chickens being prepared for sale on a conveyor belt. The collection tray was continually rocked to ensure thorough mixing. The blood was filtered on site through two layers of butter muslin into a container. The blood container was kept cold by embedding in ice during transport to the laboratory. Back at the laboratory, the blood was centrifuged in eight 0.75L bottles at 3000rpm (Sorvall RTH-750 rotor) for 30mins to separate the erythrocytes from the other blood components. As with all centrifuging, this was conducted at 4°C. The blood separated into three layers; the top layer (plasma) was retained for use as below.

Step 1. 125mL of chicken plasma was filtered (0.45µm) and concentrated to 25mL using an Amicon pressurised concentrator with a 10kDa MWCO membrane. The solution was dialysed into the High Ionic Strength Buffer. This produced a slurry of precipitated protein, which was removed by filtration (0.45µm). UV absorption measurements before and after precipitation indicated about 550mg (11%) of proteins remained in the clear solution.

Step 2. The clear solution was dialysed into Anion-exchange Buffer A (specified in Figure 7.6). 0.7mg of HMGB1 was mixed with the dialysed solution, and the solution was made up to 50mL with Anion-exchange Buffer A. The sample was injected onto an anion-exchange column and eluted with a linear ionic strength gradient (details in Figure 7.6). The chromatogram (Figure 7.6) included two distinct peaks designated 1 and 2, which were subjected to SDS-PAGE (Figure 7.7). Bands W1 to W2 were cut the SDS-PAGE gels and sent to Sheffield University for MS. Table 7.5 provides an example of the MS running conditions, and the MS results are discussed in Section 7.4.3.

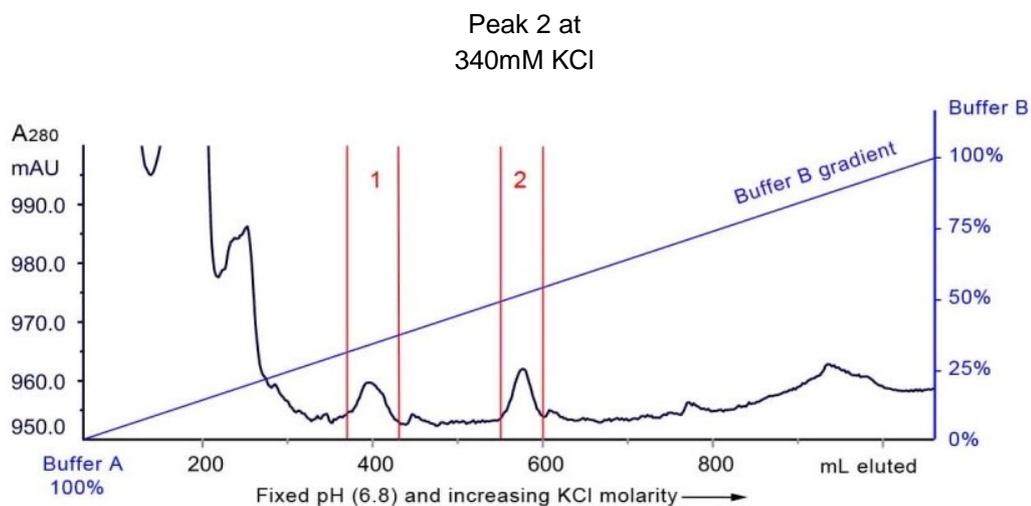


Figure 7.6. Anion-exchange chromatogram of plasma spiked with HMGB1

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%;

At end of elution: (1200mL) Buffer A 0%, Buffer B 100%.

Buffer A	100mM KCl, 50mM H ₂ KPO ₄ , 50mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride.
Buffer B	600mM KCl, 50mM H ₂ KPO ₄ , 50mM HK ₂ PO ₄ , 2.5mM benzamidine hydrochloride.
pH	6.8
Anion-exchange columns	8 x HiTrap DEAE FF 5mL (GE Healthcare)
Total column volume	40mL
Total elution volume	1200mL (120 x 10mL fractions)
Flow rate	0.9mL/min
Chromatogram identifier	DEAE low flow rate 271115001
Fractions subject to assay	1, 2 (See Figure 7.7)

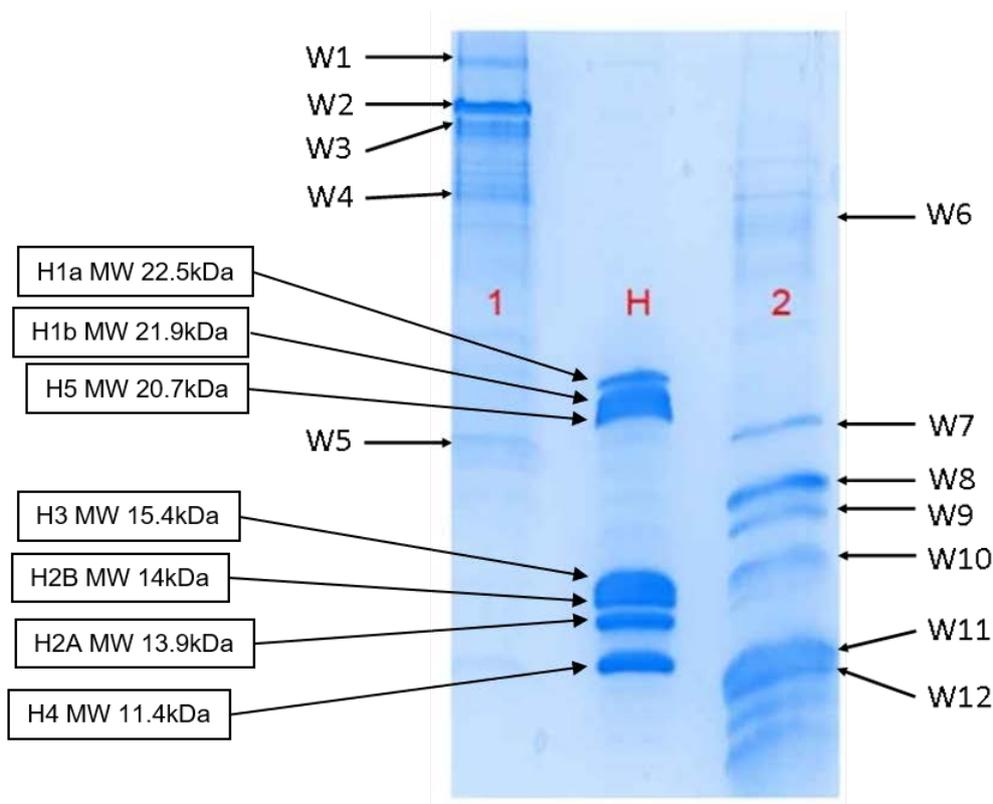


Figure 7.7 SDS-PAGE gel of Fractions 1 and 2 from the anion-exchange chromatogram of plasma spiked with HMGB1 (Figure 7.6)

H: Histone Standard

1: Peak 1 from the anion-exchange chromatogram (Figure 7.6)

2: Peak 2 from the anion-exchange chromatogram (Figure 7.6)

Stain: Coomassie Blue

Stacking gel: 5% acrylamide

Resolving gel: 20% acrylamide

Running buffer: glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-Page protocol as described in Chapter 3

MS applied to Bands 1 - 12 confirmed that Peak 2 contained HMGB1, and Peak 1 did not, but otherwise the MS results were unusual, and may have been compromised as discussed in Section 7.4.3.

Table 7.5 MS running conditions for gel band W8 (Figure 7.7) from Peak 2 of the anion-exchange chromatogram of plasma spiked with HMGB1 (Figure 7.6)

This table provides an example of the running conditions applied to all the gel bands W1 – W12 in Figure 7.7.

Search title : JPL run for MJD unmerged 20160210 (C:\ProgramData\Matrix Science\Mascot Daemon\parameters\JPL_Amazon_ncbi_vert_20160210.par), submitted from Daemon on MJDHOTDESK-PC

MS data file : D:\MJD\Feb 2016\collected_mgfs\W8 020216_RB8_01_5130.mgf

Database : NCBI nr 20140730 (47310802 sequences; 16880015641 residues)

Taxonomy : bony vertebrates (4148099 sequences)

Timestamp : 11 Feb 2016 at 02:29:27 GMT

Enzyme : Trypsin

Fixed modifications : Carbamidomethyl (C)

Variable modifications : Oxidation (M)

Mass values : Monoisotopic

Protein Mass : Unrestricted

Peptide Mass Tolerance : ± 1.2 Da

Fragment Mass Tolerance: ± 0.6 Da

Max Missed Cleavages : 1

Instrument type : ESI-TRAP

Number of queries : 1788

7.4.3 Discussion of progress in developing a method for cheap and effective measurement of HMGB1 in biofluids

Suitability of chicken erythrocytes for the study.

The work here was intended as a contribution to the development of cheap and quick methods for measuring HMGB1 in biofluids, such as plasma, serum, and urine. As shown in Chapter 2, chicken HMGB1 has 98.6% identical or conservative replacement peptides compared with human HMGB1, making it a useful tool for this purpose. Chicken plasma contains only about half the concentration of protein found in human plasma or serum (Table 7.6), but this does not invalidate its use as a model system.

HMGB1 levels to be detected

The required sensitivity to HMGB1 in plasma is the first parameter to be determined. As one of many studies focussed on measuring HMGB1 levels in cancer patients (see Table 7.1), Chung *et al.* (2009) carried out clinical studies on the correlation of HMGB1 levels in serum with stages of gastric cancer. Chung found the level of HMGB1 in the serum of "normal" patients (ie with gastric problems, but no pre-cancer indication) to be 3.9ng/mL (mean), progressively increasing through various cancer diagnostic stages to 16.5ng/mL for advanced gastric cancer (Table 7.7). Despite considerable variability from one patient to another, the level of HMGB1 was shown to provide a statistically significant difference between one diagnostic stage and the other. Also, the group of metastatic cancer patients with HMGB1 below 14ng/mL exhibited markedly longer survival times compared with those above 14ng/mL. Thus, to diagnose and determine the prognosis of gastric cancer, the goal is to accurately measure HMGB1 protein levels of 4ng/mL and above in human blood serum which has a protein concentration of ~80mg/mL in human serum (Table 7.6); this is about 1 part in 20 million.

Table 7.6 Concentrations of proteins in plasma and serum

Biofluid	Concentration mg/mL	Notes	Reference
Proteins in plasma from Attwater's Prairie Chickens	32	Quoted in online supplementary material to "Sturkie's avian physiology". 6 th Ed.	Scanes (editor) 2014
Proteins in plasma from domestic chicken	33 - 47	Quoted in "Avian physiology" 2 nd Ed.	Sturkie (1965)
Proteins in plasma from Jungle Fowl	44	-	Adnan and Amin Babjee (1985)
Proteins in avian plasma	40	Average across 100 avian species quoted in "Sturkie's avian physiology". 6 th Ed.	Scanes (editor) 2014
Proteins in human blood serum#	60-80	Reference level used by national health services	W Marshall (2012)

For comparison purposes between protein loads in chicken and human blood, human serum was taken as a representative loading, although serum has fewer proteins than plasma (serum is plasma from which the clotting proteins have been removed by allowing the blood to coagulate).

Table 7.7 Progressive increase in HMGB1 concentration in human blood serum with stages of gastric cancer

Diagnostic stage	Features	Number of patients in group	Concentration of HMGB1 ng/mL*
Normal patient	Patient has gastric problem, which may include gastritis, erosion, and ulcer, but no pre-cancer indication.	50	3.9 +-3.4
High-risk patient	IM; benign tumour.	50	6.3 +-6.3
Patient with EGC	Early gastric cancer (EGC)	40	9.9 +-11.5
Patient with AGC	Advanced gastric cancer (AGC)	45	16.5 +-27.4
Metastatic patient	Metastatic gastric cancer	42	14.1 +-13.2

* Taken from Chung *et al.* (2009), mean values, with 25 – 75% standard deviation range.

IM: Intestinal metaplasia - transformation of stomach epithelium into intestine-like epithelium.

Limitations of an ELISA method used to measure HMGB1 in plasma

Chung *et al.* (2009) showed that measurements to the required sensitivity can be done directly with a sophisticated enzyme-linked immunosorbent assay (ELISA) kit (the Shino-Test Corporation HMGB1 ELISA Kit II). This kit relies on specially prepared wells to which anti-HMGB1 antibodies have been attached to capture the HMGB1 molecules. Further anti-HMGB1 antibodies with an enzyme attached are added, to sandwich the HMGB1 molecules. The enzyme activates a chromogen substrate, the absorbance of which is measured with a spectrophotometer. Based on the Shino-Test kit protocol, the assay takes about two days. Apart from the time this assay requires, Urbonaviciute *et al.* (2006) have shown that the accuracy of ELISA assays can be compromised when applied to HMGB1 in human serum and plasma, due to other molecules binding to HMGB1, such as IgG1. Our approach of using a high ionic strength precipitation buffer may avoid this problem, since it will encourage the dissociation of complexes.

Results from the plasma proteomic workflows developed by Geyer *et al.* (2016)

Geyer *et al.* (2016) have demonstrated an MS-based rapid plasma proteomic workflow which identifies and quantifies the concentrations of ~300 proteins in a single 5 μ L blood sample within 3hr. Their simple sample preparation, which does not deplete the abundant plasma proteins, takes less than 2hr, and the MS measurement is performed in 30min. However, their rapid workflow did not detect HMGB1 protein in the blood plasma - see their supplementary Table S2. Using a more complex workflow, they were able to detect and quantify HMGB1 amongst 1040 other plasma proteins - their deep plasma dataset, supplementary Table S5. This workflow, however, involves numerous processing steps including reversed-phase chromatography and a 16hr MS measurement.

The concentration of HMGB1 given in the deep plasma dataset is ~45ng/mL, which is a factor of about 10 greater than the concentration in blood serum measured by Chung *et al.* (2009) for a healthy person – see Table 7.7. Serum has fewer proteins than plasma, so the difference in HMGB1 concentrations may represent a significant removal of HMGB1 when blood clots to form serum. On the other hand, the difference could be an artefact of the experimental methods. For example, it is not clear to what extent a concentration factor (if any) was applied by Geyer *et al.* (2016) when producing the deep plasma dataset.

Efficiency of our method

In the first HMGB1 selection step, about 90% of unwanted proteins were precipitated in the high strength buffer (Section 7.4.2). Anion-exchange chromatography was applied as the second selection. As discussed further below, the MS results appeared to have been compromised, but nevertheless were sufficient to indicate that HMGB1 was the primary protein in Peak 2 of the chromatogram (Figures 7.6). The gel (Figure 7.7) indicated that Peak 2 contained several other proteins. However, based on the intensity of the SDS-PAGE bands, none of these proteins were substantially in excess of the HMGB1. Prior to chromatography, the sample contained ~550mg of proteins including ~0.7mg of HMGB1, so other proteins exceeded HMGB1 by a factor of $550/0.7 = 786$. Since the concentrations of other proteins in Peak 2 was now no greater than that of HMGB1, this indicated that the second step alone had enhanced HMGB1 by a factor of nearly 800. Thus, the method shows promise as a means of highly enriching HMGB1 in plasma. The two steps together reduced the level of unwanted proteins by a factor approaching 10^4 . In these circumstances, it may be possible to employ a cheaper and quicker method than ELISA for completing the third HMGB1 selection and assay step.

The unusual MS results

The MS results indicated that none of the Peak 1 gel bands (W1-W5) contained HMGB1, but six from seven of the Peak 2 gel bands (W7-W12) contained primarily HMGB1. This provided confirmation that Peak 2 contained the spiked HMGB1. However, although the gel itself appeared normal (Figure 7.7), the "W" data obtained by MS seems to have been compromised. An example of the problem is that, according to the MS results, 6 separate gel bands (W7 to W12) all contained both HMGB1 (MS Scores 401 – 597) and platelet glycoprotein Ib alpha chain-like isoform X1 (MS Scores 113 – 349). This is a very unusual pattern, given the SDS-PAGE conditions usually break the bonds between proteins in a complex, and then the separated proteins typically migrate through the gel with a speed inversely proportional to the logarithm of the protein MW. For the above pattern to be a true representation of proteins correctly separated by SDS-PAGE, both HMGB1 and platelet glycoprotein Ib alpha chain-like isoform X1 would have to be present as sets of protein fragments of differing lengths with matching molecular weights, which is very unlikely. Nor was this explanation borne out by the MS sequences; for example, the HMGB1 part sequences identified by MS were almost identical for each gel band from W7

to W12, suggesting there was no systematic fragmentation. Thus, the detailed MS results seem to be an artefact, and so are not included here. Possible causes include (i) if the SDS was insufficiently mixed with the samples prior to running the gel, some of the sample molecules could have received a wide range of negative charges from the attached SDS molecules, thus becoming invisibly distributed throughout an extended part of the gel; or (ii) cross contamination of the gel pieces would also account for the unusual results.

In view of the above, it is intended to run the experiment again.

Refining our method

In a routine clinical implementation, the precipitation and chromatography steps would be refined, to permit them to be performed in minutes. The time-consuming dialysis precipitation step was to ensure a suitable volume for the chromatography for this study only. It would be replaced by adding the precipitation buffer directly to the sample. Column-less anion-exchange chromatography would then be applied, taking only a few minutes. After removing the precipitate by centrifugation, the supernatant would be diluted to a KCl ionic strength such that the HMGB1 molecules would just bind to an added insoluble DEAE anion-exchange medium. Based on Figure 7.6, a KCl buffer concentration just below 340mM would be suitable. Under this condition, none of the proteins to the left of Peak 2 in Figure 7.6 would bind to the anion-exchange medium. The medium would be pelleted by centrifugation and the supernatant discarded. A buffer with KCl concentration just above 340mM would be added to the pellet to dissociate the HMGB1. Other more strongly bound proteins would remain attached to the pellet, and centrifugation would yield a supernatant comprising almost entirely of HMGB1.

SDS-PAGE is one candidate for the third selection stage. It can be tailored to select proteins within a molecular weight range of just a few tens of Da, and the detection limit for silver-stained gels is ~1ng (Reed *et al.*, 2003). The SDS-PAGE method, as originally developed by Laemmli (1970) and used as the basis for the gels prepared here, is a relatively complex procedure which will need development if intended as part of a routine diagnostic tool kit. Effective development is likely to be possible, given that the original method has remained largely unchanged for nearly 50 years. We are looking at simplifying the sample preparation and sample injection by adopting single lane standardised pre-cast gels. We are also examining alternative gel cooling arrangements

which do not rely on liquid cooling. These would be new developments, which to our knowledge are not yet commercially available.

Further work

Further recommended work includes:

- (i) Re-running the experiment, with gels being obtained for the full spectrum of eluted proteins, and protein identities being confirmed by MS as necessary.
- (ii) Running the same experiment with just the plasma sample, to provide a baseline for the effect of spiking with HMGB1.
- (iii) Trialling the quick precipitation and column-less chromatography steps described above.
- (iv) Developing a simple and easy-to-use SDS-PAGE system along the lines outlined above.
- (v) Running further experiments to determine the accuracy with which HMGB1 levels can be measured in unspiked samples from mammals.

7.5 Example 3: proposed use of purified HMGB1 in a complex with the histone octamer suitable for structural studies

7.5.1 Structural studies of complexes between HMGB1 and chromatin or octamers

There is a lack of experimental structural studies of complexes between HMGB1 and chromatin or octamers. The interaction between HMGB1 and chromatin is usually dynamic (Scaffidi *et al.*, 2002), which makes such studies difficult. It would be worthwhile to use native HMGB1 proteins, prepared as described in Chapter 6, in conjunction with histone octamers produced by our group, and with other proteins if appropriate, to see if intermediate structures modelling the dynamic nucleosome can be stabilised for X-ray crystallography, NMR, or other structural studies. This would be one further step along the road to fully understanding the interaction of HMGB1 with chromatin.

7.5.2 Interaction between an intact histone octamer and HMGB1

We have not found any robust study which demonstrates an interaction between an intact histone octamer and HMGB1. Although Ueda *et al.* (2004) refer to cross-linking of HMGB1 with a free core histone octamer, examination of their results indicates that the cross-linking was between HMGB1 and the individual core histone H3 dissembled from the octamer. However, in addition to the binding of HMGB1 boxes to DNA, Thomas and Travers (2001), Ueda *et al.* (2004) and Watson *et al.* (2014) have used cross-linking, marker transfer, and antibodies to core histones and HMGB1, to show that HMGB1 also binds to chromatin via the HMGB1 C-terminal acidic tail binding to the N-terminal tail of core histone H3. They have shown the same HMGB1/H3 binding also occur between HMGB1 and the individual histone H3. They did not find any other interactions of HMGB1 with the core histones. [Some earlier work (Chapter 5 Table 5.2) and our own studies (Zhuang, 2011) also suggest an interaction between HMGB1 and the H2A/H2B dimer, but this is yet to be confirmed]. Since the H3 N-terminal tail extends beyond the octamer (Davey *et al.*, 2002; Chapter 2 Figure 2.6), it is possible that an HMGB1/octamer complex could be assembled via this interaction. Potentially, the high salt condition required to maintain an intact octamer (Octamer Maintenance Buffer in Section 4.4.1 contains 2M KCl and 400mM phosphates) may disrupt the HMGB1/H3 interaction. One approach would be to add HMGB1 to the core histones in low ionic strength, and hope that the HMGB1/H3 complex survives a progressive increase in buffer ionic strength up to the concentration which leads to assembly of the octamer. The bond achieved at low ionic strength could be

reinforced by using a zero length cross-linker such as 1-ethyl-3[3-dimethylaminopropyl]carbodiimide (EDC). EDC has previously been used by Bernues *et al.* (1983, 1986) to link HMGB1 to core histones.

In any event, an important aspect of this proposal is that the octamers and their environment should be a good representation of the octamers found in nucleosomes. We have previously determined the structure of the histone octamer at high resolution (2.15 and 1.9 angstrom) by crystallizing the octamer in free solution, in a buffer environment intended to simulate that found in a cell nucleus containing DNA (Chantalat *et al.*, 2003; Wood *et al.*, 2005). As with the buffers used in this study, the buffer for preparing the octamer crystals contained KCl and phosphates as a substitute for the DNA phosphates. The simulation of nuclei conditions has been generally successful, with the crystallized octamer structure being a good representation of the octamer structure found in crystallized nucleosome core particles (NCPs). Chantalat *et al.* (2003) pointed out that the histone tails in the crystallized octamer (at 2.15 angstrom) are largely disordered, whereas they are partially ordered in the crystallized NCP at 2.8 angstrom prepared by Luger *et al.* (1997). However, given the difficulty in trying to represent the fundamentally disordered histone tails in an ordered crystal, this difference may not be significant. A colleague in the research group, Dr James Nicholson, has recently used the CCP4 suite of programs to carry out a least-squares fit between the crystallized octamer (PDB 1TZY) and the octamer in the crystallized NCP (PDB 1KX5). Both structures have 1.9 angstrom resolution. He was able to superimpose the DNA from the NCP in correct registration with the crystallized octamer - there was excellent correspondence (Figure 7.8). In the crystallized NCP, there are a limited number of defined binding sites between DNA phosphates and the octamer histones. There are almost identical sites in the crystallized octamer, between the histones and either phosphate or chloride atoms, except there is one missing binding phosphate where, in the nucleosome, the DNA enters by binding to the α N helix on one of a histone H3. This is because in the octamer crystal, the acidic face of one octamer binds to the basic residues on the next octamer in the vicinity of the α N helix. Another colleague, Dr Chris Wood, has compared a sample of the DNA/octamer bonds for the crystallized NCP (PDB 1KX5) with the bonds at the same sites for the crystallized octamer (PDB 1TZY). As shown in Figure 7.9, the bonds are similar.

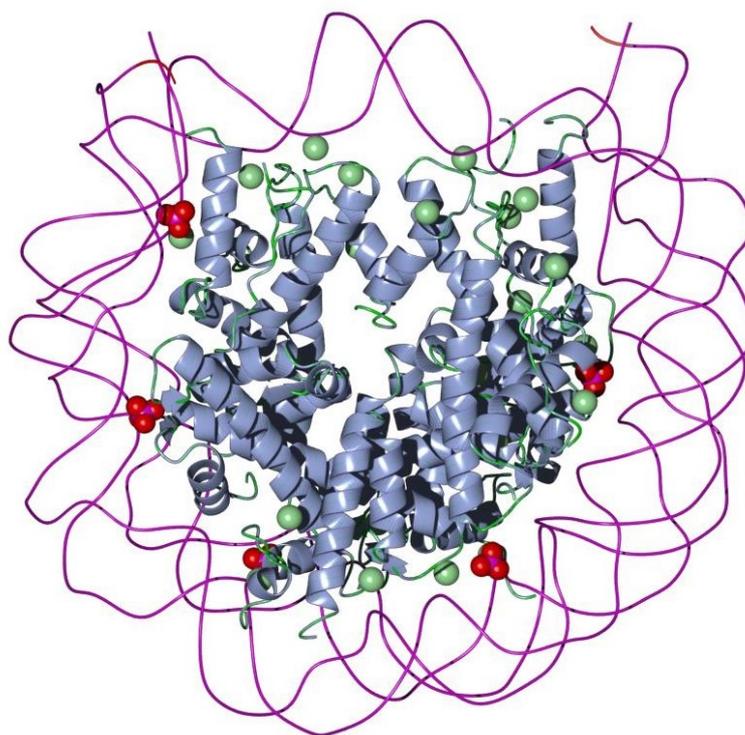


Figure 7.8 Octamer structure from crystallized histone octamer superimposed on DNA structure from crystallized nucleosome core particle

Histone octamer: RCSB PDB Entry 1TZY (Wood *et al.*, 2003); Nucleosome core particle: RCSB PDB Entry 1TZY (Davey *et al.*, 2002). Image credit: J Nicholson (University of Chester) and J Baldwin (LJMU) using RasMol (Sayle and Milner-White, 1995).

DNA backbone purple

Octamer grey

Chloride atoms green

Phosphate groups red

DNA entry and exit points (to the nucleosome core particle) at top of page.

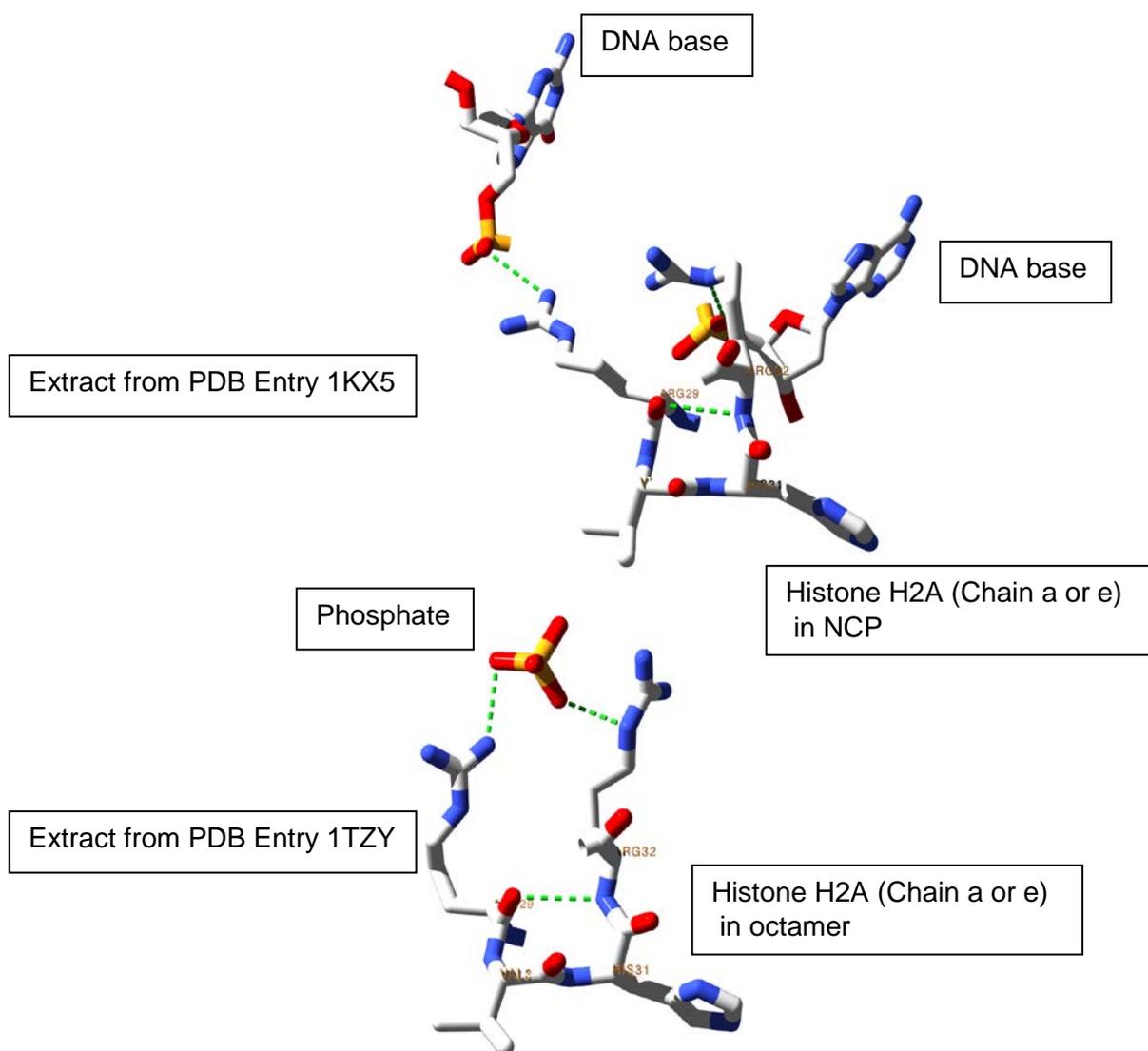


Figure 7.9 Comparison of DNA/octamer interactions in a crystallized nucleosome core particle with phosphate/octamer interactions found in the crystallized octamer

Octamer: RCSB PDB Entry 1TZY (Wood *et al.*, 2003), NCP: RCSB PDB Entry 1KX5 (Davey *et al.*, 2002). Image credit: C Wood (LJMU); images prepared using Deepview (SwissProt – Guex and Peitsch, 1997)

In the top image, the two histone arginines ARG29 and ARG32 are shown to bind to the phosphate groups of two DNA bases as found in a crystallized nucleosome core particle. In the bottom image, the same parts of the same residues ARG29 and ARG32 bind to a phosphate in the buffer of the crystallized octamer. Binding is primarily by hydrogen bonds. The similarity of location and bond type contributes to evidence that the histone octamer model 1TZY may be used to represent the octamer part of the nucleosome core particle model. In the lower image, one phosphate replaces two DNA base phosphates, so the two residue side chains adopt different rotamers from those in the upper image.

7.6 Example 4: proposed use of purified HMGB1 to explore staining suitable for electron microscopy of a complex involving HMGB1 and chromatin

There have been several recent advances in electron microscopy (EM), including better sample preparation, improved resolution, and more sophisticated image processing (conversation with Dr Paula da Fonseca at the Royal Society "Allostery and molecular machines" discussion meeting in June 2017). EM can potentially map surface electron density to individual residues, and this would be very useful in understanding complexes involving HMGB1. For example, it recently became possible using a new staining technique to provide EM images of in-situ chromatin (Ou *et al.*, 2017). We have not identified any HMGB1 structures within the PDB or NCBI databases based on electron microscopy. However, modern high performance EM could potentially be applied to complexes involving chromatin and HMGB1, once HMGB1 staining conditions have been developed which are suitable for EM and which are complementary to the chromatin staining. A ready source of purified native HMGB1 will assist with devising such staining conditions.

7.7 Conclusions and recommendations

1. A high concentration of purified native HMGB1 was added to a group of proteins, demonstrating that this provided a means of pulling out and identifying potential complexes (an HMGB1/FKBP3 complex, in this case). This work also provided evidence that HMGB1 can act as a strong dipole, given that most of the molecule is basic, including two HMGB boxes located towards the N-terminal end, balanced by a strongly acidic C-terminal tail (see Chapter 2). Watson *et al.* (2007) have shown that, at low ionic strength, the HMGB1 C-terminal acidic tail is folded back onto the positive part of the molecule (including both HMGB boxes) most of the time, but is in dynamic equilibrium with an open dipolar configuration. Results from cation-exchange chromatography support a tentative model in which the HMGB1 molecule has remained in an open configuration, with the positive (basic) part of the protein binding to the column stationary phase, and (being well in excess) the negative (acidic) tail capturing at least some of the FKBP3 molecules. A previous study using gel filtration chromatography did not find an HMGB1/FKBP3 complex (Foulger *et al.*, 2012), but it is proposed that in the lower-ionic-strength gel filtration conditions, and in the absence of a strong negative cation-exchange stationary phase, most of the HMGB1 molecules were in a closed configuration and unable to strongly bind the FKBP3.

2. Native purified HMGB1 has been employed in the preliminary development of an efficient method for measuring HMGB1 in blood plasma. Precipitation combined with anion exchange chromatography reduced the concentrations of unwanted proteins by a factor approaching 10^4 . This paves the way for a method of measuring low levels of HMGB1 which is quicker and lower cost than a typical ELISA based method, although more work is required. The development of such a method has significant medical use, for example, cancer diagnosis and prognosis (Section 7.4.3).

3. Our group has previously crystallized a histone octamer with a structure nearly identical to that found in the nucleosome core structure. It is proposed that this work be extended with a view to finding a crystallization-competent HMGB1/octamer complex, as a model structure to assist with understanding the interactions of HMGB1 with chromatin.

4. A ready source of native purified HMGB1 will assist with devising stains for electron microscopy of HMGB1 complexes.

Chapter 8

A method for the efficient separation of native FKBP3 and Cyp B from other nuclear proteins

<u>Contents</u>	Page
8.1 Introduction	223
8.2 Previous studies	223
8.3 Experimental method for the efficient isolation of FKBP3 and Cyp B	224
8.4 Discussion of results	237
8.5 Conclusions	240
<u>Table</u>	
Table 8.1 Impact of additions of Nuclear Lysis Buffer and Nuclear Precipitation Buffer during nuclei lysis and two precipitation stages	229
Table 8.2 Binding of chicken HMGB proteins, histones, FKBP3, and Cyp B to cation-exchange column at pH6.8	231
Table 8.3 MS result showing FKBP3 as the principal component of the upper band in the SDS-PAGE gel (Figure 8.4) for Fraction 3 in the anion-exchange chromatogram	235
Table 8.4 MS result showing Cyp B as the principal component of the lower band in the SDS-PAGE gel (Figure 8.4) for Fraction 3 in the anion-exchange chromatogram	236
<u>Figures</u>	
Figure 8.1 Summary of the method for preparing native FKBP3 and Cyp B (a) Preparation of the cePNE5' protein group	225
Figure 8.1 Summary of the method for preparing native FKBP3 and Cyp B (b) Selection of FKBP3 and Cyp B from the cePNE5' proteins by chromatography	226
Figure 8.2 Cation-exchange chromatogram of cePNE5' proteins extracted from chicken erythrocytes	230
Figure 8.3 Anion-exchange chromatogram of the proteins in Peak 2 from the cation-exchange chromatogram (Figure 8.2)	232
Figure 8.4 SDS-PAGE gel of the proteins in Fraction 3 from the anion-exchange chromatogram (Figure 8.3)	234

8.1 Introduction

The purpose of the thesis was to pursue purification and interaction studies of histone, HMGB, and PPI proteins. Here, a method based on a combination of cation-exchange and anion-exchange chromatography, and exploiting the dipolar nature of HMGBs, has been devised to permit efficient separation of almost pure native PPIs (FKBP3 and Cyp B) from the other nuclear proteins in chicken erythrocytes. The method is well suited to large scale, cost effective production of FKBP3, for use in further studies. The Cyp B obtained by this method was truncated, but full length Cyp B protein can probably be obtained by inclusion of a suitable protease inhibitor in the cation-exchange chromatography, as discussed later.

8.2 Previous studies

The author has not identified any specific studies aimed at large scale, cost effective production of FKBP3. As with other proteins, this protein could be produced by recombinant DNA methods, but is likely to suffer from a lack of appropriate PTMs.

Our group previously explored the principle of applying a combination of cation-exchange and anion-exchange chromatography to extract nuclear proteins (Foulger *et al.*, 2012). The method here exploits lessons from that work to provide a more effective approach to isolating FKBP3 and Cyp B. This is further discussed in Section 8.4.

8.3 Experimental method for the efficient isolation of FKBP3 and Cyp B

8.3.1 Summary

Figure 8.1(a) and (b) summarise the main steps applied to isolate FKBP3 and Cyp B (truncated) from chicken erythrocytes. Erythrocytes from fresh chicken blood were lysed, and the nuclei were washed and stored in freezing buffer. Later, some of the nuclei were lysed, and various unwanted molecules were removed by ultracentrifugation and ultrafiltration. The more insoluble proteins were precipitated and discarded. The remaining proteins were subject to cation-exchange chromatography, to isolate weakly bound positively charged proteins (including the immunophilins FKBP3 and Cyp B). Anion-exchange chromatography was then applied; this removed the HMGB proteins due to their dipolar nature leaving almost pure FKBP3 and Cyp B.

A key feature was the use of ultracentrifugation to isolate proteins from the DNA and RNA. This relatively mild approach was more likely to preserve any post translational modifications (PTMs) on FKBP3 or Cyp B. In contrast, as discussed in Chapter 4, acidic extraction can degrade PTMs, for example by removing phosphorylation from some arginines, lysines, and histidines found on histones (Chen *et al.*, 1974; Sut and Biterge, 2017). We did not find literature identifying these particular PTMs as applying to FKBP3 (Table 2.13) or Cyp B, but this does not mean that they do not exist, nor that all other FKBP3 or Cyp B PTMs are acid-stable.

8.3.2 Preparation of erythrocyte nuclei

Erythrocytes were separated from the other components of fresh chick blood by centrifugation, and intact nuclei were obtained by lysing the cells in three stages in a hypotonic buffer containing a detergent, as described in Section 4.3.2 (Chapter 4) .

8.3.3 Nuclei washing and storage

The nuclei were washed (3 washes) and stored in freezing buffer as described in Section 4.3.3.

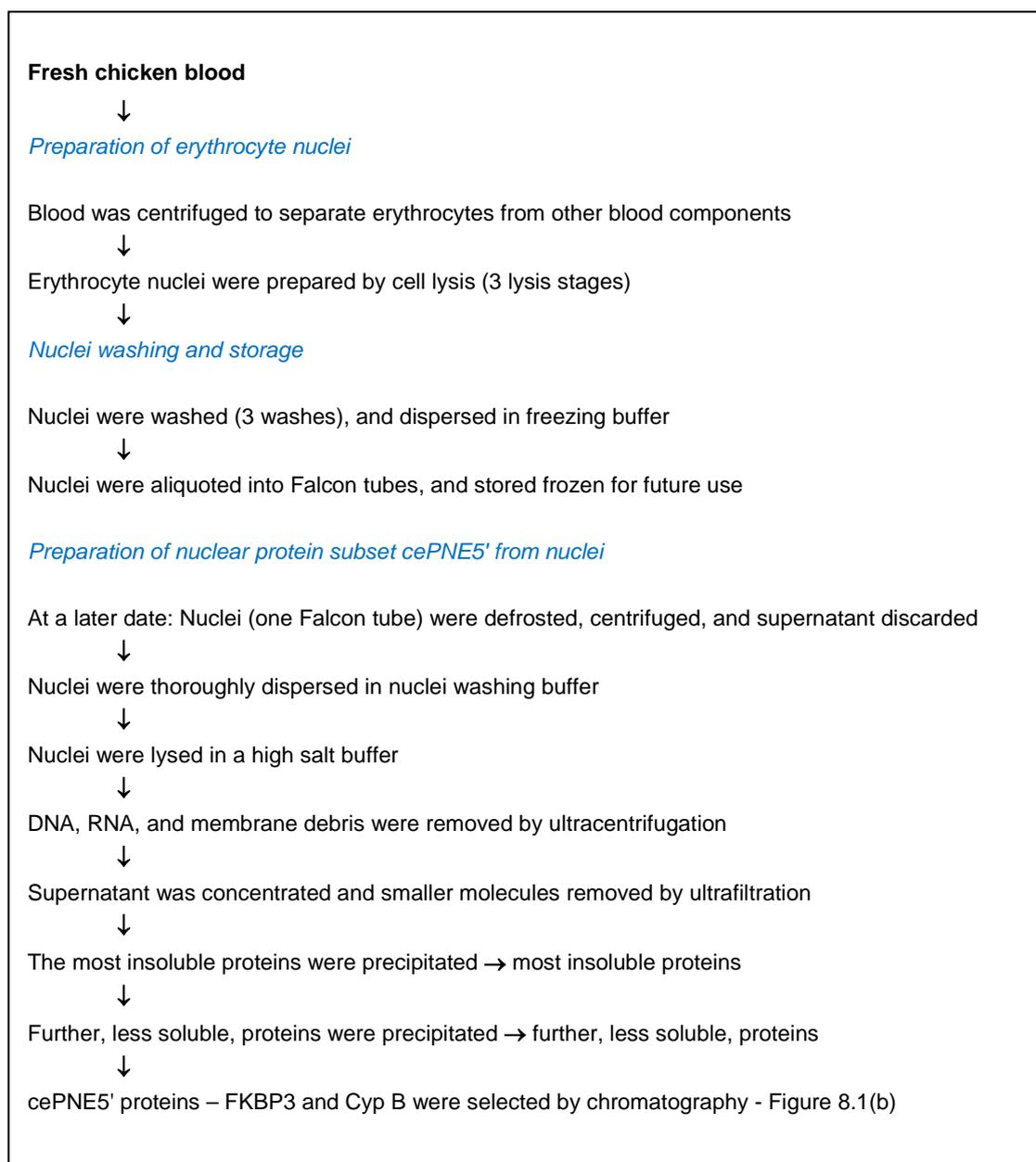


Figure 8.1 Summary of the method for preparing native FKBP3 and Cyp B

(a) Preparation of the cePNE5' protein group

See Section 8.3.2 onwards for further details of the method. All operations were conducted at $\sim 4^{\circ}\text{C}$ unless otherwise stated.

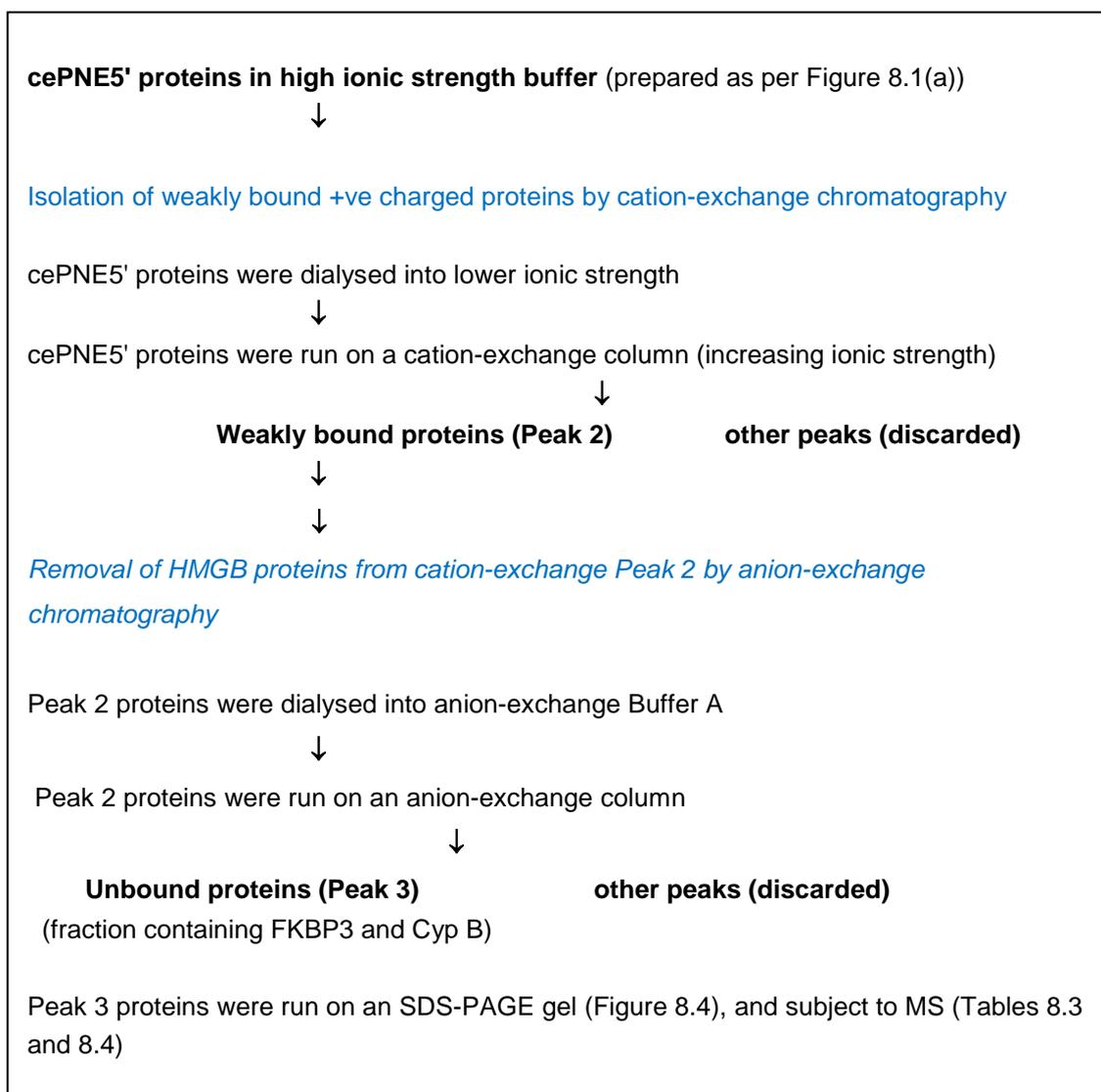


Figure 8.1 Summary of the method for preparing native FKBP3 and Cyp B

(b) Selection of FKBP3 and Cyp B from the cePNE5' proteins by chromatography

See Section 8.3.5 onwards for further details of the chromatography. Chromatography was conducted at ~4°C. Buffers used for chromatography were filtered (0.45µm) to avoid clogging the column.

8.3.4 Preparation of nuclear protein subset *cePNE5'* from nuclei

Chilled buffers were prepared as follows:

Nuclei Washing Buffer 10mM KCl, 10mM Tris/HCl pH 6.8,
3.5mM MgCl₂, 2.5mM benzamidine hydrochloride.

Nuclei Lysis Buffer 2500mM KCl, 250mM H₂KPO₄, 250mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Precipitation Buffer 2000mM KCl, 1000mM H₂KPO₄, 1000mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

To recover the nuclei from freezing buffer, one Falcon tube containing ~40mL of nuclei in Nuclei Freezing Buffer was defrosted in a fridge. The supernatant was discarded, and the sedimented nuclei pellet (~13mL) was mixed with an equal volume of Nuclei Washing Buffer and dispersed by gentle swirling. The mixture was made up to 80mL with Nuclei Washing Buffer and stirred. The nuclei in suspension were then thoroughly dispersed by transferring via a syringe through a 21-gauge needle to a separate stirred beaker.

Nuclei were lysed in a high salt solution, as follows. By running the nuclei down the side of the tube, 13mL samples of nuclei (6 samples in all), taken from the still-stirring nuclei suspension, were carefully layered onto 52 mL of Nuclei Lysis Buffer in each of six ultra-centrifuge tubes (70mL polycarbonate tubes for use with a Beckman 45Ti rotor). Each tube was sealed and quickly shaken (tube cap downward) to evenly mix the nuclei and Nuclei Lysis Buffer.

The six centrifuge tubes were centrifuged at 100,000g for 18hrs, and the pellets containing DNA, RNA, and nuclear membrane debris were discarded. The supernatant from the six tubes was pooled and concentrated for 4hrs down to 25mL using a cooled (~6°C) 400mL Amicon concentrator pressurised by nitrogen at 2bar. Water and smaller molecules (such as protein fragments) flowed through the concentrator ultrafiltration membrane (10kDa MWCO), concentrating the proteins, which helped the next stage (precipitation).

The 25mL concentrated sample of proteins was stirred on ice, and 22.5mL of Precipitation Buffer was added very slowly, in a dropwise fashion, increasing the phosphate concentrations (Table 8.1). A cloudy precipitate of proteins formed after a while; it was pelleted by centrifugation at 10,000g for 5 mins and discarded. The supernatant was retained for a second precipitation stage - 37.5mL of Precipitation Buffer was added dropwise as before, further increasing phosphate concentrations (Table 8.1). This second less dense precipitate was separated from the solution by centrifugation at 2000g for 40 min; it floated as a soft crust on top of the solution. The required protein solution (cePNE5' protein group) was sucked from under the crust, using a large-capacity syringe.

8.3.5 Isolation of weakly bound +ve charged proteins by cation-exchange chromatography of cePNE5' proteins

The cePNE5' protein solution was dialysed into cation-exchange Buffer A, and the solution was run on a cation-exchange column at fixed pH6.8 with a linear ionic strength gradient generated by progressively mixing cation-exchange Buffer B with cation-exchange Buffer A. Figure 8.2 shows the cation-exchange chromatogram and defines the buffers. Due to their strong net positive charge at pH6.8 (Table 8.2), histones and some other proteins would have bound strongly to the column, also any unfolded HMGBs would bind strongly. However, FKBP3 and Cyp B have a small net positive charges (Table 8.2), so it was expected they would bind weakly to the column, eluting in Peak 2 of Figure 8.2.

8.3.6 Removal of HMGB proteins from cation-exchange Peak 2 proteins by anion-exchange chromatography

The cation-exchange weakly bound fraction (Figure 8.2, Peak 2) was dialysed into anion-exchange Buffer A and was run on an anion-exchange column at fixed pH6.8 with a linear ionic strength gradient generated by progressively mixing anion-exchange Buffer B with anion-exchange Buffer A. Figure 8.3 shows the anion-exchange chromatogram and defines the buffers. The proteins in Fraction 3 (Figure 8.3) were subject to SDS-PAGE, with the result shown in Figure 8.4

Table 8.1 Impact of additions of Nuclear Lysis Buffer and Nuclear Precipitation Buffer during nuclei lysis and two precipitation stages

Step	Concentrations and volumes of KCl and phosphate components in the protein solution						
	KCl		H ₂ KPO ₄		HK ₂ PO ₄		
	mM	mL	mM	mL	mM	mL	
Nuclei in Nuclei Washing Buffer.	10	13	0	13	0	13	
Nuclei Lysis Buffer added (52mL).	2500	52	250	52	250	52	
Solution after adding Nuclei Lysis Buffer.	2000	65	200	65	200	65	
Proteins in solution after concentration.	2000	25	200	25	200	25	
Precipitation Buffer added (22.5mL).	2000	22.5	1000	22.5	1000	22.5	
Solution after adding Precipitation Buffer.	2000	47.5	580	47.5	580	47.5	
Proteins in solution after concentration.	2000	25	200	25	200	25	
Total Precipitation Buffer added, after adding more Precipitation Buffer (37.5mL): 22.5mL+37.5mL = 60mL total.	2000	60	1000	60	1000	60	
Solution after adding more Precipitation Buffer.	2000	85	760	85	760	85	

Notes

The less soluble nuclear proteins were progressively precipitated by increasing the total phosphate concentration from 400mM to 1160mM to 1560mM (see entries in yellow).

Concentrating the protein through an ultrafiltration membrane increased the concentration of molecules (proteins) above the filter Molecular Weight Cut-Off (MWCO), but did not alter the concentration of the molecules below the MWCO, hence buffer concentrations remained unaltered.

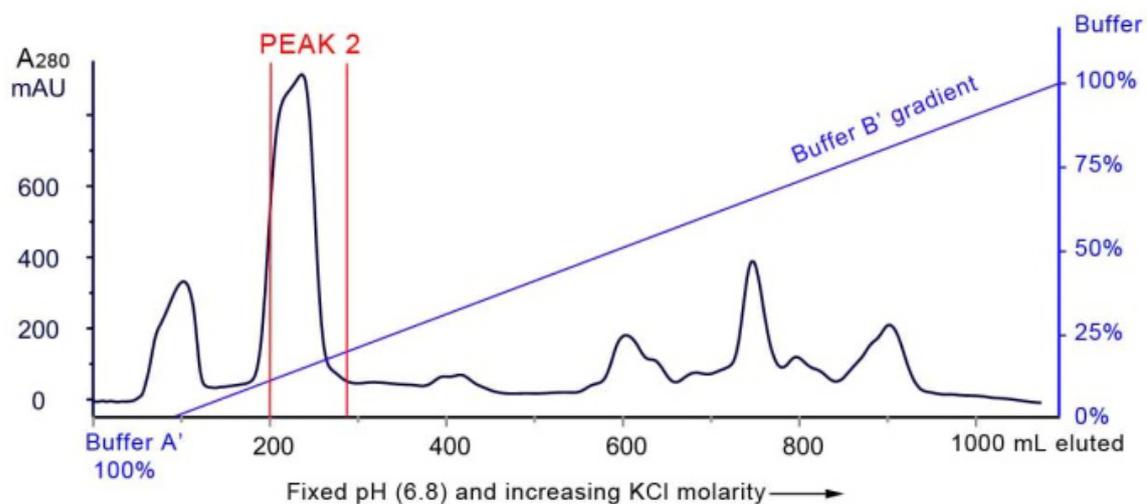


Figure 8.2 Cation-exchange chromatogram of cePNE5' proteins extracted from chicken erythrocytes

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%;

At end of elution: Buffer A 0%, Buffer B 100%.

Buffer A (pH6.8)	100mM KCl, 50mM H ₂ KPO ₄ , 50mM HK ₂ PO ₄ , 1mM MgCl ₂
Buffer B (pH6.8)	1100mM KCl, 50mM H ₂ KPO ₄ , 50mM HK ₂ PO ₄ , 1mM MgCl ₂
Cation-exchange columns	8 x HiTrap SP Sepharose FF 5mL
Total column volume	40mL
Fraction size	10mL
Flow rate	2mL/min
Chromatogram identifier	SP101013001
Fraction used in next step	Peak 2 (see Figure 8.3)

Table 8.2 Binding of chicken HMGB proteins, histones, FKBP3, and Cyp B to cation-exchange column at pH6.8

Protein	UniprotKB identifier	Effective net charge at pH6.8 (e)	Binding to cation-exchange column	Impact on the purification of FKBP3 and Cyp B
HMGB1	Q9YH06,	-3.9 if folded	Non-bind	This protein would be present in the cation-exchange chromatogram Peak 1, which was not selected.
HMGB2	P26584	+4.4 if folded	Weak	These proteins were expected in Peak 2, which was selected for anion-exchange chromatography.
HMGB3	P40618	+2.9 if folded	Weak	
FKBP3	Q90ZK7	+8.4	Weak	
Cyp B (PPIB)	P24367	+8.0	Weak	
HMGB1	Q9YH06,	+26 if unfolded, and excluding acidic tail	Strong	These proteins would have been present in later peaks, which were not selected.
HMGB2	P26584	+25.4 if unfolded, and excluding acidic tail	Strong	
HMGB3	P40618	+23.9 if unfolded, and excluding acidic tail	Strong	
H2A		+18.0	Strong	
H2B	P0C1H5	+19.0	Strong	
H3	P84229	+20.6	Strong	
H4	P62801	+18.6	Strong	
H5	P02259	+61.9	Strong	
H1	P09987	+60.9	Strong	

Net charges were calculated using Scripps Protein Calculator (2017). They are only approximate since the calculator assumes all residues have pKa values that are equivalent to the isolated residues; for a folded protein this is not valid. In any event, only part of the charge will be exposed to another molecule.

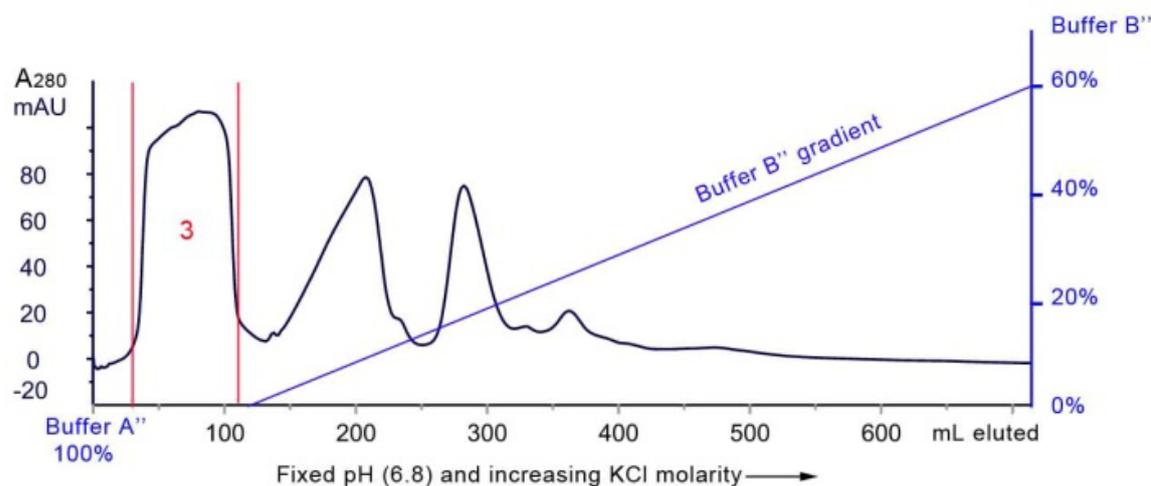


Figure 8.3 Anion-exchange chromatogram of the proteins in Peak 2 from the cation-exchange chromatogram (Figure 8.2)

A linear ionic strength gradient was generated by mixing Buffer A with Buffer B.

At start of elution: Buffer A 100%, Buffer B 0%;

At end of elution: Buffer A 0%, Buffer B 100%.

Buffer A (pH6.8) 100mM KCl, 50mM H₂KPO₄, 50mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Buffer B (pH6.8) 600mM KCl, 50mM H₂KPO₄, 50mM HK₂PO₄,
2.5mM benzamidine hydrochloride.

Anion-exchange columns 8 x HiTrap DEAE FF 5mL (GE Healthcare)

Total column volume 40mL

Fraction size 10mL

Flow rate 3mL/min

Chromatogram identifier DEAE 141013002

Fractions subject to assay 3 (SDS-PAGE gel - Figure 8.4)

8.3.7 SDS-PAGE and MS results for Fraction 3 from anion-exchange chromatography

The SDS-PAGE gel of Fraction 3 shows two distinct bands, representing two predominant proteins in the anion-exchange non-bind fraction (Figure 8.4). The bands were cut from the SDS-PAGE gel, prepared for MS as described in Chapter 3, and sent to Sheffield University for MS. The MS results in Tables 8.3 and 8.4 confirmed the bands were FKBP3 and Cyp B respectively.

The full Cyp B protein has MW 22.4kDa. The position of the Cyp B gel band partway between the bands for linker histones and core histones suggested a truncated molecule with MW between ~21.9 and 15.4kDa (lower and upper bounds of MWs of linker and core histones respectively). This is further discussed in Section 8.4.3.

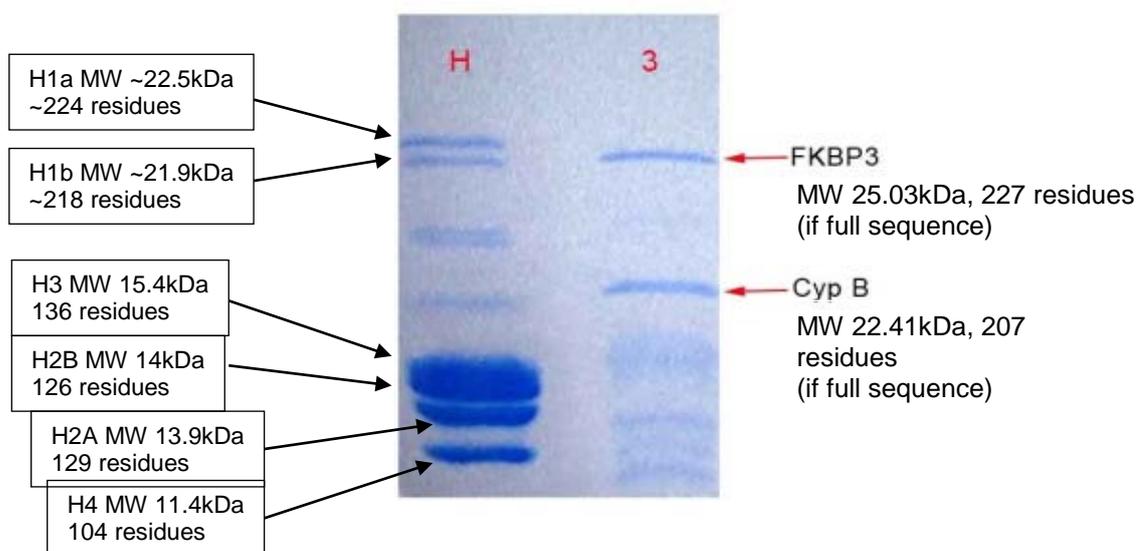


Figure 8.4. SDS-PAGE gel of the proteins in Fraction 3 from the anion-exchange chromatogram (Figure 8.3)

H: Histone Standard (see Chapter 4 for more details)

3: Fraction 3 from the anion-exchange chromatogram (Figure 8.3)

Stain: Coomassie Blue

Stacking gel: 5% acrylamide

Resolving gel: 20% acrylamide

Running buffer: glycine/tris/SDS

Sample treatment: 1 part 3X Buffer to 2 parts sample

Sample loading: adjusted to provide visible and separate bands

SDS-PAGE protocol: as described in Chapter 3

Table 8.3 MS result showing FKBP3 as the principal component of the upper band in the SDS-PAGE gel (Figure 8.4) for Fraction 3 in the anion-exchange chromatogram

The following sequences were found by MS and mapped to the full FKBP3 sequence as below. Ignoring gaps, the sequences found run from K28 to E227 (this sequence of 200 residues has a molecular weight of 22.15kDa - Scripps Protein Calculator).

K.AAKPLSFK.V
 K.KAAKPLSFK.V
 K.AEEPAAEGPPK.Y
 R.GWDEALLTMSK.G
 K.LFFEVELVDIE.
 K.AQLEIEPEWAYGK.K
 K.AQLEIEPEWAYGKK.G
 K.LQDGTVFDTNVQTSSK.K
 K.FLQEHAQAFLAEHR.L
 K.EQLIAAYTQLFHTQR.F
 K.TANKEQLIAAYTQLFHTQR.F

Sample	Protein	Identifier	Mass (Da)	Score	Coverage
Fraction 3 upper band	FKBP3 (Gallus gallus)	GI:46048916 (NCBINR)	25016	838	53%
227 residues full sequence, with 121 residues found by MS shown in red below.					
<pre> 1 MAAATAPAQP WSAEELRSEA LPKKDIIKFL QEHAAQAFLA EHRLLGQVKN VAKTANKEQL 61 IAAYTQLFHT QRFKGTDGAE RAAEKAKPGK AEGEKEKDKA AKAEPAEEG PPKYTKSILK 121 KGDKTNFPKK GDTVHCWYTG KLQDGTVFDT NVQTSSKSKK AAKPLSFKVG VGKVIRGWDE 181 ALLTMSKGEK AQLEIEPEWA YGKKGQPDAK IPPNAKLFFE VELVDIE </pre>					

Other gallus gallus proteins found in the same band (Scores in brackets - the significance of Score is explained in Section 8.3.2) were predicted talin-1 isoform X1 (461), elongation factor 1-alpha 1 (451), ras suppressor protein 1 (305), single stranded D box binding factor 2 (104), thymocyte nuclear protein 1 (73).

<u>Source of MS data</u>	<u>Mark Dickman Group, Sheffield University</u>
Search title	Alison_gallus
MS data file	D:\OAN Raw data\JANUARY 2014\29 OAN 040114_RD6_01_3211.d 29 OAN 040114_RD6_01_3211.mgf
Database	NCBI nr 20140107
Taxonomy	Bony vertebrates
Timestamp	12 Jan 2014 at 20:25:44 GMT
Enzyme	Trypsin
Variable modifications	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	± 1.7 Da
Fragment Mass Tolerance	± 0.6 Da
Max Missed Cleavages	2
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

Table 8.4 MS result showing Cyp B as the principal component of the lower band in the SDS-PAGE gel (Figure 8.4) for Fraction 3 in the anion-exchange chromatogram

The following sequences were found by MS and mapped to the full Cyp B sequence as below. Ignoring gaps, the sequences found run from run from K36 to K172. This sequence of 137 residues has a molecular weight of 15.04 kDa - Scripps Protein Calculator).

K.GFGFK.G
 K.VFFDLR.V
 R.VVIGLFGK.T
 R.FPDENFK.L
 K.VLEGMDVVR.K
 K.DFMIQGGDFTR.G
 K.TVENFVALATGEK.G
 K.DTNGSQFFITTVK.T
 K.HYGPWVSMANAGK.D
 K.SIYGDRFPDENFK.L
 R.VIKDFMIQGGDFTR.G
 K.LKHYGPWVSMANAGK.D

Sample	Protein	Uniprot Identifier Gene Name	Mass (Da)	Score	Coverage
Fraction 3 lower band	Cyp B (Gallus gallus).	P24367 (SwissProt)	22413	1849	52%
207 residues full sequence, with 109 residues found by MS shown in red below.					
<pre> 1 MKALVAATAL GPALLLLLP A SRADERKKG PKVTAKVFFD LRVGEEDAGR VVIGLFGKTV 61 PKTVENFVAL ATGEKGFVK GSKFHRVIKD FMIQGGDFTR GDGTGGKSIY GDRFPDENFK 121 LKHYGPWVS MANAGKDTNG SQFFITTVKT AWLDGKHVVF GKVLEGMDVV RKVENTKTDS 181 RDKPLKDVTI ADCGTIEVEK PFAIAKE </pre>					

Other gallus gallus proteins found in the same band (Scores in brackets - the significance of Score is explained in Section 8.3.2) were HMGB3 (142), Elongation factor 1-alpha, somatic form (53), Synaptotagmin-1 (34), DNA repair and recombination protein RAD54B (29).

Source of MS data	Mark Dickman Group, Sheffield University
Search title	Alison_gallus
MS data file	D:\OAN Raw data\JANUARY 2014\30 OAN 040114_RD7_01_3213.d\30 OAN 040114_RD7_01_3213.mgf
Database	SwissProt 2013_12
Taxonomy	bony vertebrates
Timestamp	6 Jan 2014 at 15:35:32 GMT
Enzyme	Trypsin
Variable modifications :	Carbamidomethyl (C), Oxidation (M)
Mass values	Monoisotopic
Protein Mass	Unrestricted
Peptide Mass Tolerance	± 1.7 Da
Fragment Mass Tolerance	± 0.6 Da
Max Missed Cleavages	2
Instrument type	ESI-TRAP
Preparation for MS	In-gel protein reduction, alkylation, and trypsin digestion

8.4 Discussion of results

8.4.1 Earlier work

Previously, when isolating other nuclear proteins, our group established the principle that FKBP3 and Cyp B could be isolated from other nuclear proteins if successive cation-exchange and anion-exchange chromatography stages were applied to chicken erythrocyte nuclei lysate (Foulger *et al.*, 2012). In this earlier work, an ionic strength gradient (100mM – 1000mM KCl) was simultaneously applied with a pH gradient (pH4.5 – 9), leading to a cation-exchange chromatogram with two weakly bound fractions eluting separately at ~pH5.4. At this pH, HMGB1 became positively charged, and HMGB2, and HMGB3 acquired greater positive net charges than those listed in Table 8.8 for pH6.8. The earlier work showed that FKBP3 eluted with HMGB1 and HMGB3; HMGB2 has a greater net positive charge and eluted as a separate peak with Cyp B (fragment). Subsequently, anion-exchange chromatography was used to remove the HMGB proteins from the cation-exchange weakly bound fractions. Proteins remaining in the anion-exchange non-bind fraction included FKBP3, and more weakly, Cyp B (fragment).

8.4.2 Effective isolation of FKBP3 and Cyp B (fragment)

Here, an ionic strength gradient was applied at fixed pH6.8 to the nuclei lysate, leading to a cation-exchange chromatogram with a large, weakly bound peak. Net charges at pH6.8 as shown in Table 8.2 indicated that the peak would include HMGB2, HMGB3, FKBP3 and Cyp B. Applying the same principle as in Section 8.4.1 above, anion-exchange chromatography was used to remove HMGB proteins from the non-bind fraction. As shown by the gel in Figure 8.4, this method was successful in isolating FKBP3 and Cyp B (fragment), this time in roughly equal quantities. The presence of FKBP3 and Cyp B was confirmed with high scores (an indication of the reliability of the result) by MS carried out by Sheffield University (Tables 8.3 and 8.4). Two different sequence databases (NCBI and SwissProt) were used since neither databases contained all of the relevant sequences. FKBP3 (*Gallus gallus*) was identified against the NCBI database, and Cyp B (*Gallus gallus*) was identified against the SwissProt database.

Besides FKBP3 and Cyp B (fragment), only low levels of other proteins were present, as shown by very faint additional bands in Figure 8.4, and by lower scores for the other proteins in the gel bands containing FKBP3 and Cyp B (fragment)(Tables 8.3 and 8.4). The procedure applied by Foulger *et al.* (2012) led to the inclusion of a gel band at the position expected for histone H4, but the procedure described here prevented this.

A key aspect was the inclusion of $MgCl_2$ in the cation-exchange buffers, which led to a clear separation between the non-bind and weakly-bound fractions (keeping Peak 2 separate from the first peak in Figure 8.2). When the procedure was initially performed at pH6.8 with $MgCl_2$ omitted from the buffers, the non-bind and weakly-bound fractions overlapped to some extent, leading to less effective isolation of FKBP3 and Cyp B from other nuclear proteins. The reason for the effectiveness of $MgCl_2$ inclusion is not known.

8.4.3 Protein fragmentation

As discussed in Section 7.3.3, the speed at which a typical protein migrates through the gel is approximately inversely proportional to the logarithm of the protein MW, although this does not apply to all proteins. The gel in Figure 8.4 shows that the FKBP3 protein had a gel mobility similar to that of the H1a and H1b markers. This suggests that the protein was either the full FKBP3 molecule or had lost just a few residues. MS failed to identify any sequences in the first 27 residues at the FKBP3 N-terminal end (Table 8.3), but this was not conclusive evidence of loss of all 27 N-terminal residues, nor of loss of the other unidentified residues (Table 8.3). Trypsin was applied to the protein samples to produce peptides suitable for MS. Trypsin cleaves the residue sequence on the C-terminal side of lysine and arginine residues (K and R), and sequence identification can be difficult when the trypsinised sequences are short (Harshman *et al.*, 2013). This occurs when K and R residues in the sequence are separated by only a few peptides and may have applied to the FKBP3 MS result - examination of the FKBP3 sequence in Table 8.3 indicated a relatively high density of K residues. Snijders *et al.* (2008) obtained extensive MS sequence coverage of linker histones by applying a series of limited trypsin digests since this generated longer sequence fragments. However, this technique was not applied to the MS work used here.

Figure 8.4 shows that the Cyp B protein had gel mobility intermediate between the H1b marker (MW ~21.9kDa, ~218 residues) and the H3 marker (MW 15.4kDa, 136 residues). This was evidence of a truncated Cyp B protein, since the full Cyp B sequence comprises 207 residues/MW 22.4kDa. The MS found only Cyp B sequences from K36 to K172 (ignoring gaps, 137 residues with MW 15.04kDa)(Table 8.4). If only this sequence was present, it's expected mobility would have been similar to the H3 marker. However, given that mobility was less than this (Figure 8.4), it is likely that the truncated Cyp B contained more than 137 residues. As discussed above in the context of the FKBP3 result, the MS probably did not fully identify all the available Cyp B (truncated) sequence since only a single trypsin digest was applied. However, as discussed in more detail in Section 7.3.3, an alternative experiment has found some additional evidence that the truncated Cyp B may simply have been missing the C-terminal from V173 onwards.

8.4.4 Measures to minimise protein fragmentation

Cells contain proteases (enzymes) designed to fragment proteins. In our experiments, several measures contributed to minimising this effect. Most operations were conducted in chilled conditions to reduce protease activity, and some proteases were removed by separately lysing cells and nuclei, and some may have been removed by the precipitation of the less soluble proteins. In addition, where practicable, benzamidine hydrochloride was included in the buffers. Benzamidine hydrochloride is an effective inhibitor of certain proteases in millimolar concentrations at around pH7 (Kam *et al.*, 1994) and is our usual first choice because of effectiveness and cost. However, Cyp B fragmentation may have occurred because benzamidine hydrochloride was deliberately omitted from the cation-exchange chromatography since (a) this inhibitor has very high UV absorption, and (b) it binds to the cation-exchange column and it's elution imposes a sharp UV absorption ramp on the UV measurement of proteins.

As per Chapter 4 Section 4.4.6, we have recently used benzamidine (not the hydrochloride) as an alternative inhibitor, although not yet for cation-exchange chromatography. Benzamidine has been previously used to suppress enzymic degradation of proteins (Claeys and Collen, 1978), (Luger *et al.*, 1999). Like benzamidine hydrochloride, benzamidine has a very high UV absorbance. We measured the absorbance of a 1mM solution of benzamidine in 10mM Tris/HCl pH6.8 buffer using a spectrometer scanning over the UV range 250 – 320nm. The solution had peak

absorbance below 250nm, and a secondary absorbance peak at ~270nm (~ 1AU absorbance), with absorbance reducing to 0.53AU at 280nm (our UV monitoring wavelength). For a buffer containing 2.5mM benzamidine (a typical protease inhibitor concentration), this corresponds to an absorbance of 1.33AU at 280nm, compared with chromatography absorbance peaks less than 0.1AU for some proteins in our experiments - for example, see Figure 8.3. However, provided the benzamidine does not adhere to the column, and is equally concentrated in chromatography A and B buffers, it will displace all UV measurements by a fixed amount – this does not compromise UV protein monitoring. Because the benzamidine does not incorporate an ionisable HCl component, it is likely that the benzamidine will not bind. We will be trialling it's use on a cation-exchange column.

8.4.5 Further work to explore dipolar nature of HMGB proteins.

The work described was designed to provide efficient production of the FKBP3 and Cyp B proteins. However, it could be extended to further explore the dipolar nature of HMGB proteins, in particular to determine the proportion of HMGB molecules which have unfolded into the open dipolar configuration, and under what conditions they unfold. For example, assays could be carried out to determine the extent to which the other peaks in the chromatograms in Figures 8.2 and 8.3 contain HMGB proteins. Columns with differing stationary phase charge densities could also be used, for example with a lower charge density than our HiTrap SP Sepharose FF 5mL columns which contain a strong cation-exchanger. Some comparison between the charge densities provided by the ion-exchange media and the charge density provided by DNA would also be appropriate.

8.5 Conclusions and recommendations

1. The method described here, based on a combination of cation-exchange and anion-exchange chromatography at pH6.8, has provided an efficient separation of almost pure native FKBP3 and Cyp B (in the same anion-exchange fraction) from other nuclear proteins, such as the HMGB proteins.
2. As shown by earlier work (see Section 8.4.1), a means of separating the FKBP3 and Cyp B from each other would be to run the anion-exchange fraction on a cation-exchange column at ~pH5.4. This would also have the advantage of removing some of the other unwanted proteins found at low levels in the anion-exchange fraction.

3. The method produced truncated Cyp B, probably because for practical reasons the cation-exchange chromatography was run without our first choice protease inhibitor benzamidine hydrochloride. However, benzamidine may be suitable as an alternative protease inhibitor for cation-exchange chromatography (see Section 8.4.4) and may prevent Cyp B truncation.
4. The experiment provided some limited evidence that, in certain circumstances, the HMGB molecules unfold, with their C-terminal acidic tails being released to bind to a anion-exchange column, and also releasing the basic remainder of the molecule to bind to a cation-exchange column. Section 8.4.5 discusses how the work here might be extended to further explore this dipolar nature.
5. FKBP3 and Cyp B have similar characteristics (similar net charge, molecular weight, and solubility, and eluting together). This may be a necessary condition for their function as peptidyl-prolyl isomerases, or reflect common evolutionary development, or be a coincidence.

Chapter 9

Conclusions and recommendations

1. Using a simple sulphuric acid extraction method, a set of SDS-PAGE markers based on the histone proteins (MW range 11.4 - 22.5kDa) have been efficiently prepared from chicken erythrocytes (Chapter 4). Nearly pure histone octamers have also been prepared. In this case, phosphoric acid together with potassium chloride was used for the extraction, as part of a strategy to maintain the octamers in an intact state, to facilitate future crystallisation studies.

2. A subset of chicken erythrocyte nuclear proteins containing mainly histones and HMGB proteins has been analysed for potential complexes in an environment containing negative charges provided by the strong cation-exchange stationary phase on a chromatography column, and in initial buffer conditions close to those found in a nucleus containing DNA. The proteins eluted from the column (in an ionic strength gradient) formed a pattern which could be explained by the existence of histone/HMGB complexes arising from ionic interaction of histones with the HMGB C-terminal acid tail, with the remainder of the HMGB protein initially bound to the cation-exchange column. This model is consistent with the transient dipolar nature of HMGB proteins, in which their acidic tails can become temporarily disengaged from the remainder of the protein, switching the protein into a potentially far more active state, via ionic interactions and DNA binding. Some limited evidence in support of the model was found using gel filtration chromatography. However, the model is speculative, and further confirmatory work is suggested (Chapter 5).

3. Should the above model be broadly confirmed, it provides a limited step towards further insight into a possible more general role of HMGB proteins in the nucleus, in addition to the known role of HMGB1 in assisting access to the DNA in chromatin. It would also shift the spotlight onto the role of initial unstructured ionic interactions between biological molecules, which may be essential in advance of the assembly of the biological machines which in contrast rely for final configuration on specific structural details. A related question is the extent to which ionic strength within the nucleus is localised, since this parameter affects the strength of ionic attraction. Chromatin function is influenced by its

density which varies from one location to another within a nucleus (Ou *et al.*, 2017); ionic strength may also be localised and under cellular control.

4. A method has been devised which exploits the dipolar nature of HMGB proteins to provide a quick and efficient means of isolating them from the other nuclear proteins in chicken erythrocytes (Chapter 6). All steps can be completed within two days, it does not require an ultracentrifuge, and it avoids acid degradation of PTMs. The method provides a tool for investigating the intra-cellular distribution and function of HMGB proteins. More importantly, a ready source of native HMGB proteins supports a long term aim of our research group to investigate HMGB interactions with the histone octamer and nucleosome; for example, it will support the search for HMGB/octamer structures which can be crystallized and/or subject to NMR, and the search for suitable HMGB staining such that such structures can be subject to electron microscopy (Chapter 7).

5. As described in Chapter 7, a high concentration of purified native HMGB1 was added to a group of proteins, demonstrating that this provided a means of pulling out and identifying potential complexes (a potential HMGB1/FKBP3 complex, in this case). This work also demonstrated the dipolar nature of HMGB1.

6. Native purified HMGB1 has been employed in the preliminary development of an efficient method for measuring HMGB1 in blood plasma. Precipitation combined with anion exchange chromatography reduced the concentrations of unwanted proteins by a factor approaching 10^4 . This paves the way for a method of measuring low levels of HMGB1 which is quicker and lower cost than a typical ELISA based method, although more work is required. The development of such a method has significant medical use, for example, cancer diagnosis and prognosis (Chapter 7).

7. A method has been devised, based on a combination of cation-exchange and anion-exchange chromatography, to efficiently separate pure native FKBP3 and Cyp B proteins from the other nuclear proteins in chicken erythrocytes (Chapter 8). The Cyp B protein was truncated, but this can probably be corrected by the use of benzamidine as a protease inhibitor during the cation-exchange chromatography. The method provides further evidence of the bipolar nature of HMGB proteins.

Chapter 10

References

Abraham E (2009) "Unraveling the role of high mobility group box protein 1 in severe Trauma" *Crit. Care* **13(6)**: 1004.

Adachi Y, Mizuno S, Yoshida M (1990) "Efficient large-scale purification of non-histone chromosomal proteins HMG1 and HMG2 by using Polybuffer-exchanger PBE94" *J Chromatogr.* **530(1)**: 39-46.

Adnan S, Amin Babjee SM (1985) "Haematology of the Malaysian Jungle Fowl (*Gallus gallus spadiceus*)" *Pertanika* **8(1)**: 123 -126

Agresti A, Bianchi ME (2003) "HMGB proteins and gene expression" *Curr. Opin. Genet. Dev.* **13(2)**: 170-8.

Alberts B, Johnson A, Lewis J, *et al.* (2002) "Molecular Biology of the Cell". 4th edition. *Garland Science*, New York.

An W, van Holde K, Zlatanova J (1998) "The non-histone chromatin protein HMG1 protects linker DNA on the side opposite to that protected by linker histones". *J Biol Chem.* **273(41)**: 26289-91

Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN (1991) "The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix" *Proc Natl Acad Sci U S A* **88(22)**: 10148-52

Bannister AJ, Kouzarides T (2011) "Regulation of chromatin by histone modifications" *Cell Res.* **21(3)**: 381-95.

Bates DL, Thomas JO (1981) "Histones H1 and H5: One or two molecules per nucleosome?" *Nucleic Acids Res* **9**: 5883–5894.

Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, Woodcock CL (1998). "Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin". *Proc. Natl. Acad. Sci.* **95** (24): 14173–14178.

Belgrano FS, de Abreu da Silva IC, Bastos de Oliveira FM, Fantappiè MR, Mohana-Borges R (2013) "Role of the acidic tail of high mobility group protein B1 (HMGB1) in protein stability and DNA bending" *PLoS One* **8(11)**: e79572.

Benenson Y, Adar R, Paz-Elizur T, Livneh Z, Shapiro E (2003) "DNA molecule provides a computing machine with both data and fuel" *Proc Natl Acad Sci USA* **100(5)**: 2191-6.

Berg JM, Tymoczko JL, Stryer L (2002) "Biochemistry" 5th edition. *W H Freeman*, New York.

Berman HM, Westbrook J, Feng Z, G. Gilliland, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) "The Protein Data Bank" *Nucleic Acids Research*, 28: 235-242.

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(18)**: 11020-4.

Bernués J, Espel E, Querol E (1986) "Identification of the core-histone-binding domains of HMG1 and HMG2" *Biochim Biophys Acta.* **866(4)**: 242-51.

Bianchi ME, Agresti A (2005) "HMG proteins dynamic players in gene regulation and Differentiation" *Curr Opin Genet Dev.* **15(5)**: 496-506.

Bonaldi T, Längst G, Strohner R, Becker PB, Bianchi ME (2002) "The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding" *EMBO J.* **21(24)**: 6865-73.

Bonne-Andrea C, Harper F, Sobczak J, De Recondo AM (1984) "Rat liver HMG1: a physiological nucleosome assembly factor" *The EMBO Journal* **(5)**: 1193-1199.

Boulikas T, Wiseman JM, Garrard WT (1980) "Points of contact between histone H1 and the histone octamer" *Proc. Natl. Acad. Sci.* **77(1)**: 127 -131.

Bright GR, Fisher GW, Rogowska J, Taylor DL (1987). "Fluorescence ratio imaging microscopy: temporal and spatial measurements of cytoplasmic pH". *The Journal of Cell Biology* **104** (4): 1019–1033.

Burden (2012) "Guide to the Disruption of Biological Samples – 2012" *Random Primers* **12**:1-25.

Burt DW (2005) "Chicken genome: Current status and future opportunities" *Genome Research* **15**:1692–1698.

Bustin M (2001) "Revised nomenclature for high mobility group (HMG) chromosomal Proteins" *Trends Biochem Sci.* **26(3)**: 152-153.

Cameron IL, Prescott DM (1963) "RNA and protein metabolism in the maturation of the nucleated chicken erythrocyte" *Exp Cell Res* **30**: 609–612.

Cary PD, Shooter KV, Goodwin GH, Johns EW, Olayemi JY, Hartman PG, Bradbury EM (1979) "Does high-mobility-group non-histone protein HMG 1 interact specifically with histone H1 subfractions?" *Biochem J.* **183(3)**: 657-62.

Catena R, Escoffier E, Caron C, Khochbin S, Martianov I, Davidson I (2009) "HMGB4, a novel member of the HMGB family, is preferentially expressed in the mouse testis and localizes to the basal pole of elongating spermatids" *Biol Reprod.* **80(2)**: 358-66.

Catez F, Yang H, Tracey KJ, Reeves R, Misteli T, Bustin M (2004) "Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin" *Mol Cell Biol.* **(10)**: 4321-8.

Cato L, Stott K, Watson M, Thomas JO (2008) "The interaction of HMGB1 and linker histones occurs through their acidic and basic tails" *J Mol Biol.* **384(5)**:1262-72.

Celona B *et al.* (2011) "Substantial histone reduction modulates genomewide nucleosomal occupancy and global transcriptional output" *PLoS Biol.* **9(6)**:e1001086.

Chambeyron S, Bickmore WA (2004) "Chromatin decondensation and nuclear reorganization of the *HoxB* locus upon induction of transcription" *Genes & Development* **18**: 1119 – 1130.

Chantalat L, Nicholson JM, Lambert SJ, Reid AJ, Donovan MJ, Reynolds CD, Wood CM, Baldwin JP (2003) "Structure of the histone-core octamer in KCl/phosphate crystals at 2.15 Å resolution" *Acta Crystallogr D Biol Crystallogr.* **59(8)**: 1395-407.

Chauveau J, Mouley Y, Rouiller (1956) "Isolation of pure and unaltered liver nuclei morphology and biochemical composition" *Exp Cell Res.* **11(2)**: 317-21.

Chen CC, Smith DL, Bruegger BB, Halpern RM, Smith RA (1974) "Occurrence and distribution of acid-labile histone phosphates in regenerating rat liver" *Biochemistry* **13(18)**: 3785-9.

Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, Falck JR, Peng J, Gu W, Zhao Y (2007) "Lysine propionylation and butyrylation are novel post-translational modifications in histones" *Mol Cell Proteomics* **6(5)**: 812-9.

Chen C, Huang H, Wu CH (2017) " Protein Bioinformatics Databases and Resources" *Methods Mol. Biol.* 1558: 3-39.

Chertov O *et al.* (2004) "Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry" *Proteomics* **4(4)**: 1195-1203.

Chung HW, Lee SG, Kim H, Hong DJ, Chung JB, Stroncek D, Lim JB (2009) "Serum high mobility group box-1 (HMGB1) is closely associated with the clinical and pathologic features of gastric cancer" *J Transl Med.* **7:38** 1-11.

Claeys H, Collen D (1978) "Purification and characterization of bovine coagulation factor XII (Hageman factor)" *Eur J Biochem.* **87(1)**: 69-74.

Coats JE, Lin Y, Rueter E, Maher LJ 3rd, Rasnik I (2013) "Single-molecule FRET analysis of DNA binding and bending by yeast HMGB protein Nhp6A" *Nucleic Acids Res.* **41(2)**:1372-81.

Cohen MJ, Brohi K, Calfee CS, Rahn P, Chesebro BB, Christiaans SC, Carles M, Howard M, Pittet JF (2009) "Early release of high mobility group box nuclear protein 1 after severe trauma in humans: role of injury severity and tissue hypoperfusion" *Crit Care.* **13(6)**: R174.

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW (2012) "Insights into RNA biology from an atlas of mammalian mRNA-binding proteins" *Cell* **149(6)**: 1393-406.

Crane-Robinson C, Hayashi H, Cary PD, Briand G, Sautière P, Krieger D, Vidali G, Lewis PN, Tom-Kun J (1997) "The location of secondary structure in histone H4" *Eur J Biochem* **79(2)**: 535-48.

Das C, Tyler JK, Churchill MEA (2010) "The histone shuffle: histone chaperones in an energetic dance" *Trends in Biochemical Sciences* **35**, 476-489.

Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002). "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution" *J Mol Biol.* **319(5)**: 1097-113.

Diener KR, Al-Dasooqi N, Lousberg EL, Hayball JD (2013) "The multifunctional alarmin HMGB1 with roles in the pathophysiology of sepsis and cancer" *Immunol Cell Biol.* **91(7)**: 443-50.

Deiterian-Lievre F (2005) "Commitment of hematopoietic stem cells in avian and mammalian embryos: an ongoing story" *Int. J. Dev. Biol.* **49**: 125-130.

Dill KA, MacCallum JL (2012) "The protein-folding problem, 50 years on" *Science* **338(6110)**: 1042-6.

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, Diao Y, Liang J, Zhao H, Lobanenkov VV, Ecker JR, Thomson JA, Ren B (2015) "Chromatin architecture reorganization during stem cell differentiation." *Nature* **518**: 331-6.

Durlach J (1985) "Magnesium in clinical practice" Pages 1 – 39 available in Magnesium Online Library.

Eickbush TH, Moudrianakis EN (1978) "The histone core complex: an octamer assembled by two sets of protein-protein interactions" *Biochemistry* **17(23)**: 4955-64.

Einck L, Bustin M (1985) "The intracellular distribution and function of the high mobility group chromosomal proteins" *Exp Cell Res.* **156(2)**: 295-310.

EMBL Heidelberg (2017) "Choice of lysis buffer and additives"
embl.de/pepcore/pepcore_services/protein_purification/extraction_clarification
Accessed Jan 2017.

Encor (2017) "SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)" *Encor Biotechnology Inc.* <http://www.encorbio.com/protocols/SDS-PAGE.htm>

Eyras E *et al.* (2005) "Gene finding in the chicken genome" *BMC Bioinformatics* **6(131)**: 1-12.

Felsenfeld G, Groudine M (2003) "Controlling the double helix" *Nature* **421(6921)**: 448-53.

Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B (2010) "Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells" *Cell.* **143(2)**:212-24.

Foulger LE (2011) "Development of an enrichment procedure for the isolation and analysis of nuclear proteins". Liverpool John Moores University PhD thesis.

Foulger LE, Sin CGT, Zhuang QQ, Smallman H, Nicholson JM, Lambert SJ, Reynolds CD, Dickman MJ, Wood CM, Baldwin JP, Evans K (2012) "Efficient purification of chromatin architectural proteins: histones, HMGB proteins and FKBP3 (FKBP25) immunophilin" *RSC Adv.* **2**: 10598–10604.

Franklin S, Chen H, Mitchell-Jordan S, Ren S, Wang Y, Vondriska TM (2012) "Quantitative analysis of the chromatin proteome in disease reveals remodelling principles and identifies high mobility group protein B2 as a regulator of hypertrophic growth" *Mol Cell Proteomics* **11(6)**: 1-12.

Galat A, Thai R (2014) "Rapamycin-binding FKBP25 associates with diverse proteins that form large intracellular entities" *Biochem Biophys Res Commun.* **450(4)**: 1255-60.

Garg LC, Reeck GR (1998) "Isolation and separation of HMG proteins and histones H1 and H5 and core histones by column chromatography on phosphocellulose" *Protein Expr Purif.* **14(2)**:155-9.

Gasteiger E *et al.* (2005) "Protein identification and analysis tools on the ExPASy server" from the Proteomics Protocols Handbook 2005 Edition. Walker JM (Ed.) Humana Press/Springer.

GE Healthcare (2006) "Hydrophobic Interaction and Reversed Phase Chromatography – Principles and Methods " *GE Healthcare Bio-Sciences AB, Sweden.*

GE Healthcare (2007) "Affinity Chromatography – Principles and Methods " *GE Healthcare Bio-Sciences AB, Sweden.*

GE Healthcare (2014) "Size Exclusion Chromatography – Principles and Methods " *GE Healthcare Bio-Sciences AB, Sweden.*

GE Healthcare (2016) "Ion Exchange Chromatography – Principles and Methods " *GE Healthcare Bio-Sciences AB, Sweden.*

Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M (2016) "Plasma Proteome Profiling to Assess Human Health and Disease" *Cell Syst.* **2(3)**: 185-95.

Gibot S, Massin F, Cravoisy A, Barraud D, Nace L, Levy B, Bollaert PE (2007) "High-mobility group box 1 protein plasma concentrations during septic shock" *Intensive Care Med.* **33(8)**: 1347-53.

Gill SC, von Hippel PH (1989) "Calculation of protein extinction coefficients from amino acid sequence data" *Anal Biochem.* **182(2)**: 319-26. Erratum in: *Anal Biochem.* 1990 Sep; **189(2)**: 283.

Gissot M, Ting LM, Daly TM, Bergman LW, Sinnis P, Kim K (2008) "High mobility group protein HMGB2 is a critical regulator of plasmodium oocyst development" *J Biol Chem.* **283(25)**:17030-8.

Goodwin GH, Johns EW (1973) "Isolation and Characterisation of Two Calf-Thymus Chromatin Non-Histone Proteins with High Contents of Acidic and Basic Amino Acids" *European Journal of Biochemistry* **40**: 215–219.

Goodwin GH, Nicolas RH, Johns EW (1975) "An improved large scale fractionation of high mobility group non-histone chromatin proteins" *Biochim Biophys Acta.* **405(2)**: 280-91.

Gray GW, Wilkinson SG (1965) "The Effect of Ethylenediaminetetra-acetic Acid on the Cell Walls of Some Gram-Negative Bacteria" *J gen. Microbiol.* **(39)**: 385-399.

Grdisa M, White MK (2003) "Molecular and biochemical events during differentiation of the HD3 chicken erythroblastic cell line" *Int J Biochem Cell Biol.* **35(4)**: 422-31.

Gudavicius G, Dilworth D, Serpa JJ, *et al.* (2014) "The prolyl isomerase, FKBP25, interacts with RNA-engaged nucleolin and the pre-60S ribosomal subunit" *RNA.* **20(7)**: 1014-1022.

Guedes PT, Dias de Oliveira B CEP, Paulo de Abreu Manso P, Caputo LFG, Cotta-Pereira G, Pelajo-Machado M (2014) "Histological Analyses Demonstrate the Temporary Contribution of Yolk Sac, Liver, and Bone Marrow to Hematopoiesis during Chicken Development" *PLOS ONE* **9(3)** e90975: 1 – 8.

Guex, N. and Peitsch, M.C. (1997) "[SWISS-MODEL](#) and the Swiss-PdbViewer: An environment for comparative protein modeling" *Electrophoresis* **18**, 2714-2723.

Gurley LR, Valdez JG, Prentice DA, Spall WD (1983) "Histone fractionation by high-performance liquid chromatography" *Anal Biochem.* **129(1)**: 132-44.

Gygi SP, Aebersold R (2000) "Mass spectrometry and proteomics" *Curr Opin Chem Biol.* **4(5)**: 489-94.

Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, Bazett-Jones DP, Allis CD, Hunt DF (2006) "Expression patterns and post-translational modifications associated with mammalian histone H3 variants" *J Biol Chem.* **281(1)**: 559-68.

Harshman SW, Young NL, Parthun MR, Freitas MA (2013) "H1 histones: current perspectives and challenges" *Nucleic Acids Res.* **41(21)**: 9593-609.

Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C (1994) "Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain" *EMBO J.* **13(8)**: 1823-30.

Helander S, Montecchio M, Lemak A, *et al.* (2014) "A new protein fold in human FKBP25/FKBP3 and HectD1" *Biochemical and biophysical research communications* **447(1)**: 26-31.

HMG nomenclature website (2016)

http://www.informatics.jax.org/mgihome/nomen/hmg_family.shtml.

Hillier LW, Miller W, Birne E, Warre W, Hardison R.C, Ponting C.P, Bork P, Burt DW, Groenen MA, Delany ME *et al.* (2004) "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution" *Nature* **432**: 695–716.

Hock R, Furusawa T, Ueda T, Bustin M (2007) "HMG chromosomal proteins in development and disease". *Trends Cell Biol.***17(2)**: 72-9.

Horn PJ, Carruthers LM, Logie C, Hill DA, Solomon MJ, Wade PA, Imbalzano AN, Hansen JC, Peterson CL (2002) "Phosphorylation of linker histones regulates ATP-dependent chromatin remodelling enzymes" *Nat. Struct. Biol.* **9**: 263–267.

Hornbeck PV, Kornhauser JM, Tkachev S, *et al.* (2012) "PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse" *Nucleic Acids Research* **40** (Database issue): D261-D270.

Howe FS, Boubriak I, Sale MJ, Nair A, Clynes D, Grijzenhout A, Murray SC, Woloszczuk R, Mellor J (2014) "Lysine acetylation controls local protein conformation by influencing proline isomerization" *Mol Cell.* **55(5)**: 733-44.

Human Protein Atlas (2017) - www.proteinatlas.org

Huang X, Miller W (1991) "A time efficient, linear space local similarity algorithm" *Advances in Applied Mathematics* **12**: 337–357.

Hung DT, Schreiber SL (1992) "cDNA cloning of a human 25 kDa FK506 and rapamycin binding protein" *Biochim Biophys Res Commun.* **184(2)**: 733-8.

Huttunen HJ, Fages C, Kuja-Panula J, Ridley AJ, Rauvala H (2002) "Receptor for advanced glycation end products-binding COOH-terminal motif of amphotericin inhibits invasive migration and metastasis" *Cancer Res.* **62(16)**: 4805-11.

Ingwall JS, Balschi JA (2006) "Energetics of the Na(+) pump in the heart" *J Cardiovasc Electrophysiol* **17 Suppl 1**:S127-S133.

Isackson PJ, Debold WA, Reeck GR (1980) "Isolation and separation of chicken erythrocyte high mobility group non-histone chromatin proteins by chromatography on phosphocellulose". *FEBS Lett.* **119(2)**: 337-42.

Jaffredo T, Yvernogeu L (2014) "How the avian model has pioneered the field of hematopoietic development" *Experimental Hematology* **42**: 661–668.

Johns EW (1967) "The electrophoresis of histones in polyacrylamide gel and their quantitative determination" *Biochem J.* **104(1)**: 78-82.

Ju BG, Lunyak VV, Perissi V, Garcia-Bassets I, Rose DW, Glass CK, Rosenfeld MG (2006) "A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription" *Science* **312(5781)**: 1798-802. Erratum in: *Science* (2011) **332(6030)**: 664. Erratum in: *Science* (2014) **343(6173)**: 839.

Kabat D, Attardi G (1966) "Synthesis of chicken hemoglobins during erythrocyte differentiation" *Biochim. Biophys. Acta* **138**: 382-399.

Kakhniashvili DG, Bulla LA Jr, Goodman SR (2004) "The human erythrocyte proteome: analysis by ion trap mass spectrometry" *Mol Cell Proteomics* **3(5)**: 501-9.

Kam *et al.* (1995) "Mammalian Tissue Trypsin-like Enzymes: Substrate Specificity and Inhibitory Potency of Substituted Isocoumarin Mechanism-Based Inhibitors, Benzamidine Derivatives, and Arginine Fluoroalkyl Ketone Transition-State Inhibitors" *Archives of Biochemistry and Biophysics* **316(2)**: 808-814.

Kan PY, Hayes JJ (2007) "Detection of interactions between nucleosome arrays mediated by specific core histone tail domains" *Methods.* **41(3)**: 278-85.

Kan PY, Lu X, Hansen JC, Hayes JJ (2007) "The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays" *Mol Cell Biol* **27**: 2084–2091.

Kan PY, Caterino TL, Hayes JJ (2009) "The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays" *Mol Cell Biol* **29**: 538–546.

Kang R, Chen R, Zhang Q, Hou W, Wu S, Cao L, Huang J, Yu Y, Fan XG, Yan Z, Sun X, Wang H, Wang Q, Tsung A, Billiar TR, Zeh HJ 3rd, Lotze MT, Tang D (2014) "HMGB1 in health and disease" *Mol Aspects Med.* **40**: 1-116.

Kanz C, Aldebert P, Althorpe N, *et al.* (2005) "The EMBL Nucleotide Sequence Database. *Nucleic Acids Research* **33** (Database Issue): D29-D33.

Kiefer CM, Hou C, Little JA, Dean A (2008) "Epigenetics of beta-globin gene Regulation" *Mutat Res.* **647(1-2)**: 68-76.

Klune JR, Dhupar R, Cardinal J, Billiar TR, Tsung A (2008) "HMGB1: Endogenous Danger Signaling" *Molecular Medicine* **14(7-8)**: 476-484.

Kohlstaed LA, King DS, Cole RD (1986) "Native State of High Mobility Group Chromosomal Proteins 1 and 2 Is Rapidly Lost by Oxidation of Sulfhydryl Groups during Storage" *Biochemistry* **25**: 4562-4565.

Kohlstaed LA, Cole RD (1994) "Specific Interaction between H 1 Histone and High Mobility Protein HMG1" *Biochemistry* **33**: 570-575.

Koutzamani E, Laborg H, Sarg B, Lindner HH, Rundquist I (2002) "Linker histone subtype composition and affinity for chromatin in situ in nucleated mature erythrocytes" *J Biol Chem* **277**: 44688–44694.

Kowalski A, Palyga J (2011) "Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non histone protein MENT". *Chromosome Res* **19**: 579–590.

Kwon JH, Kim J, Park JY, Hong SM, Park CW, Hong SJ, Park SY, Choi YJ, Do IG, Joh JW, Kim DS, Choi KY (2010) "Overexpression of high-mobility group box 2 is associated with tumor aggressiveness and prognosis of hepatocellular carcinoma" *Clin Cancer Res.* **16(22)**: 5511-21.

Laemmli UK (1970) "Cleavage of structural proteins during the assembly of the head of bacteriophage T4" *Nature* **227(5259)**: 680–685.

Lambert S, Muyldermans S, Baldwin J, Kilner J, Ibel K, Wijns L (1991) "Neutron scattering studies of chromatosomes" *Biochem Biophys Res Commun* **179**: 810–816.

Lange SS, Mitchell DL, Vasquez KM (2008) "High mobility group protein B1 enhances DNA repair and chromatin modification after DNA damage" *Proc Natl Acad Sci U S A.* **105(30)**: 10320-5.

Leclercq M, Vinci F, Galat A (2000) "Mammalian FKBP-25 and its associated proteins" *Arch Biochem Biophys.* **380(1)**: 20-8.

Lee KB, Thomas JO (2000) "The effect of the acidic tail on the DNA-binding properties of the HMG1,2 class of proteins: insights from tail switching and tail removal" *J Mol Biol.* **304(2)**: 135-49.

Lee D, Kwon JH, Kim EH, Kim ES, Choi KY (2010) "HMGB2 stabilizes p53 by interfering with E6/E6AP-mediated p53 degradation in human papillomavirus-positive HeLa cells" *Cancer Lett.* **292(1)**: 125-32.

Lee S-Y, Lee H, Lee H-K, *et al.* (2016) "Proximity-Directed Labeling Reveals a New Rapamycin-Induced Heterodimer of FKBP25 and FRB in Live Cells" *ACS Central Science* **2(8)**: 506-516.

Leontis NB, Stombaugh J, Westhof E (2002) "The non-Watson-Crick base pairs and their associated isostericity matrices" *Nucleic Acids Res.* **30(16)**: 3497-531.

Leung A, Jardim FP, Savic N, Monneau YR, González-Romero R, Gudavicius G, Eirin-Lopez JM, Bartke T, Mackereth CD, Ausió J, Nelson CJ (2017) "Basic surface features of nuclear FKBP facilitate chromatin binding" *Sci Rep.* **7(1)**: 3795.

Li J, Kokkola R, Tabibzadeh S, Yang R, Ochani M, Qiang X, Harris HE, Czura CJ, Wang H, Ulloa L, Wang H, Warren HS, Moldawer LL, Fink MP, Andersson U, Tracey KJ, Yang H (2003) "Structural basis for the proinflammatory cytokine activity of high mobility group box 1" *Mol Med.* **9(1-2)**: 37-45.

Li G, Levitus M, Bustamante C, Widom J (2005) "Rapid spontaneous accessibility of nucleosomal DNA" *Nat Struct Mol Biol.* **12(1)**: 46-53.

Llopis J, McCaffery JM, Miyawaki A, Farquhar MG, Tsien RY (1998). "Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins" *Proc Natl Acad Sci* **95(12)**: 6803-8.

Lodish H, Berk A, Kaiser CA, Krieger M, Scott MP, Bretscher A, Ploegh H, Matsudaira P (2007) "Molecular Cell Biology, Sixth Edition" *W. H. Freeman and Company, New York.*

Long FA, McDevit WF (1952) "Activity Coefficients of Nonelectrolyte Solutes in Aqueous Salt Solutions" *Chemical Reviews* **51(1)**:119-169

Luger K., Mader, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J. (1997) "Crystal structure of the nucleosome core particle at 2.8 Å resolution" *Nature* **389**: 251-260.

Luger K, Rechsteiner TJ, Richmond TJ (1999) "Preparation of nucleosome core particle from recombinant histones" *Methods Enzymol.* **304**: 3-19.

Luo Y, Hara T, Ishido Y, Yoshihara A, Oda K, Makino M, Ishii N, Hiroi N, Suzuki K (2014) "Rapid preparation of high-purity nuclear proteins from a small number of cultured cells for use in electrophoretic mobility shift assays" *BMC Immunol.* **15**: 586.

Maeshima K, Imai R, Tamura S, Nozaki T (2014) "Chromatin as dynamic 10-nm fibre" *Chromosoma* **123**: 225 – 237.

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) "The protein kinase complement of the human genome" *Science* **298(5600)**: 1912-34.

Marekov LN, Demirov DG, Beltchev BG (1984) "Isolation of high-mobility-group proteins HMG1 and HMG2 in non-denaturing conditions and comparison of their properties with those of acid-extracted proteins" *Biochim Biophys Acta*. **789(1)**: 63-8.

Marshall W (2012) "Analyte Monographs alongside the National Laboratory Medicine Catalogue (AMALC): Total protein" *The Association for Clinical Biochemistry and Laboratory Medicine*. <http://www.acb.org.uk/whatwedo/science/AMALC.aspx>.

Matthews HR, Huebner VD (1984) "Nuclear protein kinases" *Mol Cell Biochem*. **59(1-2)**: 81- 99.

Medvinsky A, Rybtsov S, Taoudi S (2011) "Embryonic origin of the adult hematopoietic system: advances and questions" *The Company of Biologists Ltd Development* **138**: 1017-1031.

Millar CB (2013) "Organizing the genome with H2A histone variants" *Biochem J*. **449(3)**: 567-79.

Misteli T, Gunjan A, Hock R, Bustin M, Brown DT (2000) "Dynamic binding of histone H1 to chromatin in living cells" *Nature* **408(6814)**: 877- 81.

Monneau YR, Soufari H, Nelson CJ, Mackereth CD (2013) "Structure and activity of the peptidyl-prolyl isomerase domain from the histone chaperone Fpr4 toward histone H3 proline isomerization" *J Biol Chem*. **288(36)**: 25826-37.

Moriwaki Y, Yamane T, Ohtomo H, Ikeguchi M, Kurita J, Sato M, Nagadoi A, Shimojo H, Nishimura Y (2016) "Solution structure of the isolated histone H2A-H2B heterodimer" (2016) *Sci Rep* **6**: 24999-24999.

Murray K (1966) "The acid extraction of histones from calf thymus deoxyribonucleoprotein" *J Mol Biol*. **15(2)**: 409-19.

Musselman CA, Lalonde ME, Côté J, Kutateladze TG (2012) "Perceiving the epigenetic landscape through histone readers" *Nat Struct Mol Biol.* **19(12)**: 1218-27.

NCBI Resource Coordinators (2017) "Database Resources of the National Center for Biotechnology Information" *Nucleic Acids Research.* **45** (Database issue): D12-D17.

Nelson CJ, Santos-Rosa H, Kouzarides T (2006) "Proline isomerization of histone H3 regulates lysine methylation and gene expression" *Cell* **126(5)**: 905-16.

Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, Seidel NE, Cline AP, Bodine DM (2003) "Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation" *Blood* **102(4)**:1298-306.

Nemeth MJ, Cline AP, Anderson SM, Garrett-Beal LJ, Bodine DM (2005) "Hmgb3 deficiency deregulates proliferation and differentiation of common lymphoid and myeloid progenitors" *Blood* **105(2)**: 627-34.

Nemeth MJ, Kirby MR, Bodine DM (2006) "Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation" *Proc Natl Acad Sci USA* **103(37)**: 13783-8.

NRC (2011) "Prudent practices in the laboratory". National Research Council.

Ochocka AM, Kampanis P, Nicol S, Allende-Vega N, Cox M, Marcar L, Milne D, Fuller-Pace F, Meek D (2009) "FKBP25, a novel regulator of the p53 pathway, induces the degradation of MDM2 and activation of p53" *FEBS Lett.* **583(4)**: 621-6.

Olszowy P, Donnelly MR, Lee C, Ciborowski P (2015) "Profiling post-translational modifications of histones in human monocyte-derived macrophages." *Proteome Sci.* **13**:24.

Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O'Shea CC (2017) "ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells" *Science* **357(6349)** eaag0025: 1-13.

Panyim S, Chalkley R (1969) "High resolution acrylamide gel electrophoresis of histones" *Arch Biochem Biophys.* **130(1)**: 337-46.

Parseghian (2015) "What is the role of histone H1 heterogeneity? A functional model emerges from a 50 year mystery" *AIMS Biophysics* **2(4)**: 724-772.

Paull TT, Haykinson MJ, Johnson RC (1993) "The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures" *Genes Dev.* **7(8)**: 1521-34.

Pence HE, Williams A (2010) "ChemSpider: An Online Chemical Information Resource" *Journal of Chemical Education* **87(11)**: 1123-1124.

Petesch SJ, Lis JT (2012) "Overcoming the nucleosome barrier during transcript elongation". *Trends Genet* **28**: 285–294.

Petit A, Ragu C, Della-Valle V, Lafage-Pochitaloff M, Soler G, Schluth C, Radford I, Ottolenghi C, Bernard OA, Penard-Lacronique V and Romana SP (2010) "NUP98-HMGB3, a nouvel oncogenic fusion" *Leukemia* **24**: 654–658.

Polanská E, Dobšáková Z, Dvořáčková M, Fajkus J, Štros M (2012) "HMGB1 gene knockout in mouse embryonic fibroblasts results in reduced telomerase activity and telomere dysfunction" *Chromosoma* **121(4)**: 419-31.

Pongdam S (2008) "A coherent procedure for fractionating proteins from chicken erythrocyte cell nuclei" *PhD Thesis, Liverpool John Moores University.*

Prakash A, Rajan S, Yoon HS (2016a) "Crystal structure of the FK506 binding domain of human FKBP25 in complex with FK506" *Protein Sci.* **25(4)**: 905-10.

Prakash A, Shin J, Rajan S, Yoon HS (2016b) "Structural basis of nucleic acid recognition by FK506-binding protein 25 (FKBP25), a nuclear immunophilin" *Nucleic Acids Research* **44(6)**: 2909-2925.

Pusterla T, de Marchis F, Palumbo R, Bianchi ME (2009) "High mobility group B2 is secreted by myeloid cells and has mitogenic and chemoattractant activities similar to high mobility group B1". *Autoimmunity* **42(4)**: 308-10.

Putyrski M, Schultz C (2012) "Protein translocation as a tool: The current rapamycin Story" *FEBS Lett.* **586(15)**: 2097-105.

Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet, RM (1993) "Crystal structure of globular domain of histone H5 and its implications for nucleosome binding". *Nature* **362**: 219-223.

Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D (1998) "GeneCards : A novel functional genomics compendium with automated data mining and query reformulation support" *Bioinformatics* **14(8)**: 656-664.

Reed R, Holmes D, Weyers J, Jones A (2003) "Practical skills in biomolecular sciences" 2nd ed. *Pearson Education Limited*.

Reeves R (2010) "Nuclear functions of the HMG proteins" *Biochim. Biophys. Acta* **1799**: 3–14. [published online in 2009 as "HMG Nuclear Proteins: Linking Chromatin Structure to Cellular Phenotype"].

Resnick R, Halliday D, Krane KS (2002) "Physics" 5th edition. *John Wiley & Sons, Inc.*

Rodriguez-Collazo P, Leuba SH, Zlatanova J (2009) "Robust methods for purification of histones from cultured mammalian cells with the preservation of their native modifications" *Nucleic Acids Research.* **37(11)** :e81.

Ronfani L, Ferraguti M, Croci L, Ovitt CE, Schöler HR, Consalez GG, Bianchi ME (2001) "Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2" *Development* **128(8)**: 1265-73.

Rosas-Ballina M, Goldstein RS, Gallowitsch-Puerta M, Yang L, Valdés-Ferrer SI, Patel NB, Chavan S, Al-Abed Y, Yang H, Tracey KJ (2009) "The selective alpha7 agonist GTS-21 attenuates cytokine production in human whole blood and human monocytes activated by ligands for TLR2, TLR3, TLR4, TLR9, and RAGE" *Mol Med*. **15(7-8)**: 195-202.

Rose AS, Hildebrand PW (2015) "NGL Viewer: a web application for molecular visualization" *Nucleic Acids Res*. **43**: W576-579.

Saeki T, Kunito T (2010) "Adsorptions of DNA molecules by soils and variable-charged soil constituents" *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*. Ed. A Mendez-Vilas.

Sánchez-Giraldo R, Acosta-Reyes FJ, Malarkey CS, Saperas N, Churchill ME, Campos JL (2015) "Two high-mobility group box domains act together to underwind and kink DNA" *Acta Crystallogr*. **D71**: 1423-32.

Sanders C (1977) "A method for the fractionation of the high-mobility-group non-histone chromosomal proteins" *Biochem Biophys Res Commun*. **78(3)**: 1034-42.

Saris NE, Mervaala E, Karppanen H, Khawaja JA, Lewenstam A (2000) "Magnesium. An update on physiological, clinical and analytical aspects" *Clin Chim Acta*. **294(1-2)**: 1-26.

Sayle R, Milner-White EJ (1995) "RasMol: Biomolecular graphics for all", *Trends in Biochemical Sciences (TIBS)* **20(9)**: 374.

Scaffidi P, Misteli T, Bianchi ME (2002) "Release of chromatin protein HMGB1 by necrotic cells triggers inflammation" *Nature* **418(6894)**: 191-5.

Erratum published (one reference corrected): *Nature* 2010 Sep 30; **467(7315)**: 622.

Scanes CG (Editor) (2014) "Sturkie's Avian Physiology 6th Edition". Academic Press.

Scheffer MP, Eltsov M, Frangakis AS (2011) "Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei" *Proc Natl Acad Sci* **108**:16992–16997

Schreiber SL (1991) "Chemistry and biology of the immunophilins and their immunosuppressive ligands" *Science* **251(4991)**: 283-7.

Shechter D, Dormann HL, Allis CD, Hake SB (2007) "Extraction, purification and analysis of histones" *Nat Protoc.* **2(6)**: 1445-57.

Shi T *et al.* (2012) "Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum" *Proc Natl Acad Sci U S A.* **109(38)**: 15395-400.chu

Shin JS, Pierce NA (2004) "A synthetic DNA walker for molecular transport" *J Am Chem Soc.* **126(35)**:10834-5.

Shirakawa H, Tanigawa T, Sugiyama S, Kobayashi M, Terashima T, Yoshida K, Arai T, Yoshida M (1997) "Nuclear accumulation of HMG2 protein is mediated by basic regions interspaced with a long DNA-binding sequence, and retention within the nucleus requires the acidic carboxyl terminus" *Biochemistry* **36(20)**: 5992-9.

Sin CGT (2009) "Interaction nature of HMGB1 with native linker histone variants" LJMU project report.

Smith N, Englebert VE (1969) "Erythropoiesis in chicken peripheral blood" *Canadian Journal of Zoology* **47(6)**: 1269 – 1273. [abstract only]

Snijders AP, Pongdam S, Lambert SJ, Wood CM, Baldwin JP, Dickman MJ (2008) "Characterization of post-translational modifications of the linker histones H1 and H5 from chicken erythrocytes using mass spectrometry" *J Proteome Res.* **7(10)**:4326-35.

Song L, Huang SC, Wise A, Castanon R, Nery JR, Chen H, Watanabe M, Thomas J, Bar-Joseph Z, Ecker JR. (2016) "A transcription factor hierarchy defines an environmental stress response network" *Science* **354(6312)**: 1-10.

Stoscheck CM (1990) "Quantitation of protein" *Methods Enzymol.* **182**: 50-68.

Stott K, Tang GS, Lee KB, Thomas JO (2006) "Structure of a complex of tandem HMG boxes and DNA" *J Mol Biol.* **360(1)**: 90-104.

Vol. 147, No. 1, 1987

Stros M (1987) "Binding of non-histone chromosomal protein HMG1 to histone H3 in nucleosomes detected by photochemical cross-linking" *Biochem and Biophys Research Communications* **147(1)**: 301-308.

Stros M, Polanská E, Struncová S, Pospíšilová S (2009) "HMGB1 and HMGB2 proteins up-regulate cellular expression of human topoisomerase II α " *Nucleic Acids Res.* **37(7)**: 2070-86.

Stros M (2010) "HMGB proteins: interactions with DNA and chromatin" *Biochim Biophys Acta.* **1799(1-2)**: 101-13.

Sturkie PD (1965) "Avian Physiology, 2nd Edition" Cornell University Press.

Sut I, Biterge B (2017) "Acid-labile histone phosphorylations" *MOJ Cell Sci Report* **4(3)**: 00088

Swanson PC (2002) "Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals" *Mol Cell Biol.* **22(5)**: 1340-51.

Syrjanen R (2015) "TIM family molecules in Hematopoiesis". *ACTA UNIVERSITATIS OULUENSIS Dissertation.*

Szep S, Park S, Boder ET, Van Duyne GD, Saven JG (2009) " Structural coupling between FKBP12 and buried water" *Proteins* **74(3)**: 603-11.

Szklarczyk D *et al.* (2015) "STRING v10: protein-protein interaction networks, integrated over the tree of life" *Nucleic Acids Res.* **43(Database issue)**: D447-52.

Takami Y, Higashio M, Fukuoka T, Takechi S, Nakayama T (1996) "Organization of the chicken histone genes in a major gene cluster and generation of an almost complete set of the core histone protein sequences" *DNA Res.* **3(2)**: 95-9.

Taniguchi N, Caramés B, Hsu E, Cherqui S, Kawakami Y, Lotz M (2011) "Expression patterns and function of chromatin protein HMGB2 during mesenchymal stem cell differentiation" *J Biol Chem.* **286(48)**: 41489-98.

Teif VB, Rippe K (2009) "Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities" *Nucleic Acids Research* **37(17)**: 5641-5655.

Teo SH, Grasser KD, Thomas JO (1995) "Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY" *Eur J Biochem.* **230(3)**: 943-50.

Thomas JO, Rees C, Finch JT (1992) "Cooperative binding of the globular domains of histones H1 and H5 to DNA" *Nucleic Acids Research* **20(2)**:187-194.

Thomas JO (2001) "HMG1 and 2: architectural DNA-binding proteins" *Biochem. Soc. Trans.* **29 (4)**: 395–401.

Thomas JO, Travers AA (2001) "HMG1 and 2, and related 'architectural' DNA-binding proteins" *Trends Biochem Sci.* **26(3)**: 167-74.
Erratum in: *Trends Biochem Sci.* **26(4)**: 219.

Turner BM (2001) "Chromatin and Gene Regulation: Molecular Mechanisms in Epigenetics" *Blackwell Science Ltd*, Oxford UK.

Ueda T, Chou H, Kawase T, Shirakawa H, Yoshida M (2004) "Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription Stimulation" *Biochemistry* **43(30)**: 9901-8.

Ueda T, Chou H, Kawase T, Shirakawa H, Yoshida M (2004) "Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription stimulation" *Biochemistry* **43(30)**: 9901-9908.

Ueda T, Yoshida M (2010) "HMGB proteins and transcriptional regulation" *Biochim Biophys Acta.* **1799(1-2)**: 114-8.

Uhlén M *et al.* (2015) "Tissue-based map of the human proteome" *Science* **347(6220)**:1260419.

Urb M, Sheppard DC (2012) "The Role of Mast Cells in the Defence against Pathogens". *PLoS Pathogens* **8(4)** e1002619 1-3.

Urbonaviciute V, Fürnrohr BG, Weber C, Haslbeck M, Wilhelm S, Herrmann M, Voll RE (2007) "Factors masking HMGB1 in human serum and plasma" *J Leukoc Biol.* **81(1)**: 67-74.

Van Duyne GD, Standaert RF, Schreiber SL, Clardy J (1991) "Atomic structure of the rapamycin human immunophilin FKBP-12 complex" *J.Am.Chem.Soc.* **113**: 7433-7435.

Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) "Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin" *J Mol Biol.* **229(1)**: 105-24.

Venereau E, *et al.* (2012) "Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release" *J Exp Med.* **209(9)**: 1519-28.

Von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P (2005) "STRING: known and predicted protein-protein associations, integrated and transferred across organisms" *Nucleic Acids Res.* **33** (Database issue): D433-7.

Wallis JW *et al.* (2004). "A physical map of the chicken genome". *Nature* **432**: 761–764.

Wang H, Bloom O, Zhang M, Vishnubhakat JM, Ombrellino M, Che J, Frazier A, Yang H, Ivanova S, Borovikova L, Manogue KR, Faist E, Abraham E, Andersson J, Andersson U, Molina PE, Abumrad NN, Sama A, Tracey KJ (1999) "HMG-1 as a late mediator of endotoxin lethality in mice" *Science* **285**: 248-51.

Wang J, Tochio N, Takeuchi A, Uewaki JI, Kobayashi N, Tate SI (2013a) "Redox-sensitive structural change in the A-domain of HMGB1 and its implication for the binding to cisplatin modified DNA" *Biochem Biophys Res Commun.* **441**: 701–706.

Wang W, Jiang H, Zhu H, Zhang H, Gong J, Zhang L, Ding Q (2013b) "Overexpression of high mobility group box 1 and 2 is associated with the progression and angiogenesis of human bladder carcinoma" *Oncol Lett.* **5(3)**: 884-888.

Watson M, Stott K, Thomas JO (2007) "Mapping intramolecular interactions between domains in HMGB1 using a tail-truncation approach" *J Mol Biol.* **374(5)**: 1286-97.

Watson M, Stott K, Fischl H, Cato L, Thomas JO (2013) "Characterization of the interaction between HMGB1 and H3- α a possible means of positioning HMGB1 in chromatin" *Nucleic Acids Res.* **42(2)**: 848-59.

Whitford D (2005) "Proteins: structure and function" John Wiley & Sons Ltd, UK.

Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE (1980) "Crystal structure analysis of a complete turn of B-DNA" *Nature* 287(5784): 755-8.

Wiśniewski JR, Szewczuk Z, Petry I, Schwanbeck R, Renner U (1999) "Constitutive phosphorylation of the acidic tails of the high mobility group 1 proteins by casein kinase II alters their conformation, stability, and DNA binding specificity" *J Biol Chem.* **274(29)**: 20116-22.

Wood CM (2010) "Molecular Conformations". Bookboon.com

Wood CM, 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* **61**: 541-545.

Yang W-M, Yao Y-L, Seto E (2001) "The FK506-binding protein 25 functionally associates with histone deacetylases and with transcription factor YY1" *The EMBO Journal.* **20(17)**: 4814-4825.

Yao YL, Liang YC, Huang HH, Yang WM (2011) "FKBPs in chromatin modification and Cancer" *Curr Opin Pharmacol*. **11(4)**: 301-7

Zhang J, McCauley MJ, Maher LJ, Williams MC, Israeloff NE (2009) "Mechanism of DNA flexibility enhancement by HMGB proteins" *Nucleic Acids Research*. **37(4)**: 1107-1114.

Zhang K, Tang H, Huang L, Blankenship JW, Jones PR, Xiang F, Yau PM, Burlingame AL (2002) "Identification of acetylation and methylation sites of histone H3 from chicken erythrocytes by high-accuracy matrix-assisted laser desorption ionization-time-of-flight, matrix-assisted laser desorption ionization-postsource decay, and nanoelectrospray ionization tandem mass spectrometry" *Anal Biochem*. **306(2)**: 259-69.

Zhang QQ, Wang Y (2008) "High Mobility Group Proteins and Their Post-Translational Modifications" *Biochimica et biophysica acta*. **1784(9)**: 1159-1166.

Zhang Y, Moqtaderi Z, Rattner BP, Euskirchen G, Snyder M, Kadonaga JT, Liu XS, Struhl K (2010) "Evidence against a genomic code for nucleosome positioning" *Nat Struct Mol Biol*. **17(8)**: 920-3.

Zhu S, Ashok M, Li JH, Li W, Yang H, Wang P *et al.* (2009) "Spermine protects mice against lethal sepsis partly by attenuating surrogate inflammatory markers" *Mol Med* **15**: 275–282.

Zhuang Q (2011) "Biochemical and structural studies of histone and associated proteins from chick and human nuclei" Liverpool John Moores University PhD thesis.

Zhuang Q, Smallman H, Lambert SJ, Sodngam SS, Reynolds CD, Evans K, Dickman MJ, Baldwin JP, Wood CM (2014) "Cofractionation of HMGB proteins with histone dimers" *Anal Biochem*. **447**: 98-106.

Zlatanova J, van Holde K (1998) "Linker histones versus HMG1/2: a struggle for dominance?" *Bioessays* **7**: 584-8.