An experimental approach to determine the binding mode of yeast linker histone, Hho1p

Elizabeth M. Smith

SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF SCIENCE

IN THE DEPARTMENT OF MOLECULAR AND CELL BIOLOGY
FACULTY OF SCIENCE
UNIVERSITY OF CAPE TOWN

2003

Contents

Abstract

In this study, we have developed an assay to investigate of the binding mode of yeast linker histone Hho1p to nucleosomes, using magnetic bead pull-down assays. To understand the structural and regulatory role of, Hho1p, the linker histone in *Saccharomyces cerevisiae*, and that of higher eukaryotes (Thoma *et al.,* 1979), it is essential to determine the manner in which these histones associate with chromatin. It was previously proposed that Hho1p may bind to two nucleosomes simultaneously, where globular domain one and globular domain two bind to the adjacent nucleosomes. The identified lysine-rich region, linking the globular domains, was proposed to interact with the DNA linking the neighbouring nucleosomes.

To test the binding of the separate globular domains of Hho1p to DNA, globular domain one and globular domain two (GD1 and GD2) were over-expressed in *E. coli* and purified. It was found that the MTE leader peptide shown to improve expression of the chicken linker histone, GH5 in *E. coli* (Gerchman *et al.,* 1993), was not required for the over-expression of the separate yeast globular domains in *E. coli.* Expression and purification studies also revealed that GD2 may be structurally less stable than GD1. Both proteins were expressed with a C-terminal hexa histidine-tag which could be cleaved. This would be important in assessing the influence of the charged region on the nucleosomal binding by the globular domains.

Two modified proteins will serve as controls in the magnetic bead assays; the first of these control proteins was designed as a copy of Hho1p and contains a protease site in the lysine-rich linker region of the protein. This site can be cleaved to separate two complexes to which the two globular domains may bind. This protein was over-expressed in *E. coli* and purified. The second control, which has not been successfully purified, would contain two chicken linker histone globular domains (GH5) flanking the Hho1p lysine-rich linker region. A quantification method for the biotin-labeled DNA was designed, and the initial testing of the specificity of the magnetic bead assay was also determined. In this study we have therefore designed and prepared some of the reagents necessary to study directly the proposed binding of yeast Hho1p to two adjacent nucleosome cores.

Abbrevations

Å	angstrom
°C	degrees centigrade
Acn	Acetonitrile
ADP	Adenosinediphosphate
AGGA	DNA sequence containing adenine-guanine-guanine-adenine
Arg	L-Arginine
ATP	Adenosinetriphosphate
AT-rich	DNA rich in adenosine and thymidine
bp	Base pair
BSA	Bovine Serum Albumin
$CaCl_2$	Calcium Chloride
CAP	Catabolite Activator Protein
CAPS	(3-[Cyclohexylamino]-1-propanesulfonic acid)
$CoCl_2$	Cobalt Chloride
CpG	5'-cytosyl-guanidine-dinucleotide
cm	centimeter
C-terminus	Carboxyterminus
Da	Dalton
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DNAse I	Deoxyribonuclease I
rDNA	DNA coding for ribosomal RNA
DTT	DL-Dithiothreitol
E. coli	*Escherichia coli*
EDIA	N-ethyldiisopropylamine
EDTA	Ethylene-diamine-tetra-acetic acid
GC-rich	DNA rich in guanidine and cytosine
GD1	Globular domain one of yeast linker histone, Hho1p
GD2	Globular domain two of the yeast linker histone, Hho1p

GH1	Globular domain of linker histone H1
GH5	Globular domain of linker histone H5
GuHCl	Guanidium hydrochloride
hr	hour
H1-CS	Cleavage stage H1
H1-GFP	H1 fused to Green Fluorescent Protein
H1t	Testes-specific H1
HCl	Hydrogen chloride
His	L-Histidine
His-tag	Carboxy/Amino-terminal tract of 6 consecutive histidine residues
Hho1p	Yeast linker histone
HNF-3γ	Hepatocyte Nuclear Factor 3γ
HPLC	High Performance Liquid Chromatography
HTH	Helix-Turn-Helix
IMAC	Immobilised Metal Affinity Chromatography
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobase
kDa	kilodalton
KCl	Potassium Chloride
KH_2PO_4	Potassium Phosphate
LB	Luria Broth
Lys	L-Lysine
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation/ Time Of Flight
mg	milligram
min	minute
µg	microgram
µm	micrometer
µM	micromolar
µmol	micromole
mm	millimeter
mM	millimolar
mmol	millimole

mol	mole
Mg^{2+}	Magnesium Ion
$MgSO_4$	Magnesium Sulphate
MNase	Micrococcal Nuclease
MS	Mass Spectrometry
MTE	Methionine – Threonine - Glutamic acid
Na^+	Sodium Ion
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NaH_2PO_4	Sodium Phosphate
NH_4HCO_3	Ammonium Bicarbonate
$NiSO_4$	Nickel Sulphate
NMR	Nuclear Magnetic Resonance
nm	nanometer
N-terminus	Amino Terminus
OD	Optical Density
OD_{600}	Optical Density at 600 nm
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PMSF	Phenylmethansulfonylfluoride
poly(dA-dT)	DNA sequence containing consecutive adenine-thymidine base steps
hRFX-DBD	DNA-Binding Domain of human Recognition Factor of X-box
rHho1p	Recombinant Hho1 Protein
rpm	Revolutions Per Minute
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SPKK	Serine-Proline-Lysine-Lysine
U	Unit
TFA	Trifluoroacetic Acid

TLC	Thin Layer Chromatography
TRE	Thyroid Hormone Response Element
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
V	Volt
v/v	volume per volume
w/v	weight per volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

List of tables

List of figures

Acknowledgements

Firstly, I would like to thank my supervisor, Hugh Patterton for all his support, advice and contagious scientific enthusiasm. I would also like to thank Wolf Brandt and George Lindsay for their stimulating discussions and valuable advice with the protein purifications. I am extremely grateful towards all the technicians in the department of Molecular and Cell Biology at UCT, especially Faezah Davids and Pat Thompson, for practical help. I would like to thank Rob Karreman for help with the MALDI-TOF. Furthermore, I wish to acknowledge the Wellcome Trust and the National Research Foundation for funding.

I would also like to thank my colleagues, especially Tim, Rochan, Sachin, Tony and the Illing Laboratory.

A special thanks to my parents and friends for all the love and support, especially Carol, Michelle, Gerrit, Ceri, Hilton and Steve.

Lastly, I would like to thank David, who means the world to me.

Chapter 1: Introduction.
The binding of H1 to chromatin.

1.1 Chromatin.

Chromatin is the packaged form of the eukaryotic genome. The different structural components that make up chromatin facilitate the organization of DNA into different levels of compaction, ranging from the winding of approximately 168 bp of DNA into two negative superhelical turns of 110 Å diameter to the very dense DNA-protein complex found in the 30 nm chromatin fiber and the metaphase chromosome. Although chromatin plays an essential structural role, it is not a static structure since it influences various functions of DNA that are regulated by chromatin structure. To understand the mechanistic details of these regulatory effects, the interactions between the structural components of chromatin must be described.

1.2 The nucleosome.

The basic unit of chromatin is the nucleosome, and our understanding of chromatin structure is the most comprehensive for this fundamental unit. Each nucleosome consists of two copies each of the core histones H2A, H2B, H3 and H4 (Kornberg, 1974; Thomas and Kornberg, 1975), which forms the histone octamer around which approximately 168 bp of DNA is wrapped in two turns of a negative superhelix. In chromatin, adjacent nucleosomes are joined by a variable length of DNA known as linker DNA (McGhee *et al.*, 1980).

The linker histone binds to the nucleosome cores and promotes the organization and the stabilization of the 30 nm chromatin fiber (Thoma *et al.*, 1979). This binding of the linker histone leads to a metastable intermediate structure which is observed during micrococcal nuclease digestion (Thomas, 1984). This structure, the chromatosome (Figure 1.1), consists of 168 bp of DNA coiled around the core octamer with the ends bound by the linker histone (Simpson, 1978). Subjecting the chromatosome to further digestion by micrococcal nuclease produces the 146 bp core

particle lacking the linker histone. Thus in binding to the nucleosome, the linker histone protects an additional 20 bp of nucleosomal DNA.

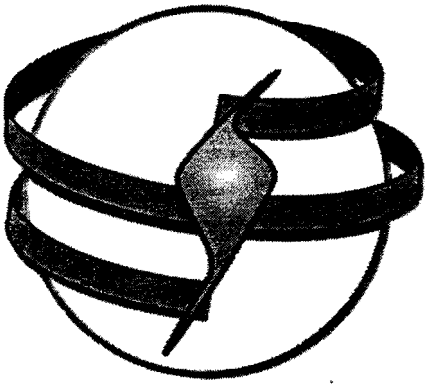


Figure 1.1 The Chromatosome. The chromatosome is generated by micrococcal nuclease digestion (MNase) and consists of 168 bp of DNA wrapped around the histone octamer, with H1 associated with this structure. (Travers, 1999)

The core histones are highly conserved (De Lange *et al.,* 1969) and this level of conservation is probably related to their structural and functional roles in both the nucleosome structure and various chromosomal processes. All four core histones are structurally similar and contain an extended histone fold domain, histone fold extensions and the N-terminal histone tails. The histone fold contains three α-helices connected by two loops (Arents *et al.,* 1991; Arents and Moudrianakis, 1995). It is through this domain that the histone-histone interactions take place, stabilizing the constituent histones in the central octamer of the nucleosome.

The initial crystal structure of the nucleosome core was solved at 7 Å resolution (Richmond *et al.,* 1984), and showed the non-uniform curvature of DNA in the nucleosome which confirmed the disc-like shape of the core particle (Finch, 1977), that resulted from the left-handed supercoiling of the DNA around the core. A subsequent study by Richmond and co-workers refined the crystal structure to 2.8 Å resolution (Luger *et al.*, 1997). The histone octamer has a two-fold symmetry (Arents *et al.,* 1991). Each of the core histones forms a heterodimer with its partner in a "handshake" motif which leads to a crescent-shaped structure. These are organized in

such a way that two H2A-H2B dimers flank an H3-H4 tetramer. The histone fold regions of the H2A-H2B dimers and H3-H4 tetramer organize 121 bp of nucleosomal DNA. The H3-H4 tetramer is in contact with the central turns of DNA and the H2A-H2B dimers are in contact with regions at the ends of the nucleosome core DNA (Luger *et al.*, 1997).

The structure of the N-terminal tails of the histones, are of particular interest. These charged tails are the targets of most of the post-translational modifications (Grunstein, 1997) to the histones and have been implicated in the modulation of the chromatin structure through interaction with other structural chromatin proteins. The tails of H3 and H2B pass between the gyres of the DNA superhelix, in a channel created by the minor grooves of the adjacent gyres. The N-terminal domain of H2A associates with the minor groove on the outside of the DNA superhelix. The N-terminal tail regions of H4 have divergent structures which may have been a function of the crystallization of these tails (Luger *et al.*, 1997).

1.3 Linker histones.

Apart from the core histones, which form the central core onto which the DNA is spooled in a nucleosome, there are also the linker histones. These histones, forming the H1 family, are thought to partially neutralize the charge of the linker DNA connecting adjacent nucleosomes. The H1 family is more variable than the core histones, probably as a result of the different structural functions preformed by these histones (Wolffe, 1995).

1.3.1 Origin and variants.

The linker histone H1 is evolutionarily less conserved than the core histones and most organisms express several different H1 isotypes in different cell types and in different developmental stages. It was suggested that this family of histones has a prokaryotic origin (Kasinsky *et al.*, 2001). In contrast to the lysine-rich carboxyl region of linker histones, the globular winged helix was only formed later in eukaryotic evolution, which might reflect the developmental variability and complexity in the evolution of eubacteria to higher eukaryotes (Kasinsky *et al.*, 2001).

There are numerous species and tissue-specific variants of the linker histone, with different biochemical and biophysical properties. The variants differ predominantly in the net charge and length of the C-terminal domains. Based on the biochemical and biophysical behavior of linker histones, Van Holde (1989) proposed the existence of four classes of linker histones, namely H1, H1°/H5, H1t and H1-CS. Khochbin (2001) arranged these groups as the cleavage-stage linker histones (H1-CS), differentiation-associated linker histones (H1° and H5), testis-specific linker histones (H1t) and replication-associated linker histones (H1), respectively.

Of particular interest is the tissue-specific linker histone H5. This linker histone has been used in various studies on both the structure and the function of the linker histones (Ramakrishnan *et al.,* 1996; Simpson, 1978). H5 accumulates during the process of terminal differentiation of some nucleated erythrocytes and has been implicated in both chromatin compaction and transcriptional inactivation (Ramakrishnan, 1997).

The existence of the various variants and the evolutionary development of the linker histones might be indicative that they have more targeted regulatory functions in both chromatin structure and gene regulation.

1.3.2 Linker histone structure.

Linker histones are basic lysine-rich proteins (Zlatanova &Van Holde, 1996), which exhibit distinct tripartite structure (Hartman *et al.*, 1977). The central structured region contains most of the hydrophobic residues and has been shown to be able to protect the additional 20 bp of nucleosomal DNA from micrococcal nuclease digestion in the same manner as the full-length linker histones (Allan *et al.,* 1980; Buckle *et al.,* 1992). The N- and C-terminal tails are lysine rich and are unstructured in solution (Aviles *et al.,* 1978). These basic regions are implicated in non-specific interactions with the linker DNA joining adjacent nucleosomes (Staynov & Crane-Robinson, 1988).

Two linker histones have been identified that lack this tripartite structure. The central globular domain is absent in the macronuclear H1 of the heterokaryon, *Tetrahymena* (Hayashi *et al.,* 1987). The other exception is the *Saccharomyces cerevisiae* H1, which contains two globular domains separated by a lysine-rich region (Landsman, 1996; Patterton *et al.,* 1997).

The structures of the globular domains of both H1 (GH1) and H5 (GH5) were solved using 2D-NMR (Cerf *et al.,* 1994) and X-ray crystallography (Ramakrishnan *et al.,* 1993), respectively. Results from these studies correlate well.

From the 2.5 Å resolution X-ray structure (Ramakrishnan *et al.,* 1993) of the globular domain of the chicken erythrocyte H5 (GH5), it was found to consist of a three-helix bundle (I-III) with a β-hairpin at the C-terminus. Between helices I and II there is a short three-residue stretch which is in the β-conformation and, together with the C-terminal β-hairpin, this strand forms the third strand of an antiparallel β-sheet. In the asymmetric crystal structure, the structure of the β-hairpin in the two molecules is quite distinct and might suggest the flexibility of this motif in solution (Ramakrishnan *et al.,* 1993).

NMR studies on the globular domain of linker histone H1 showed the structural similarities between GH1 and GH5 (Cerf *et al.,* 1994). There was, however, a difference in the conformation of the loop connecting helices I and III. This difference was in residue His 62 which is situated in this loop in GH5 and was shown to crosslink to chromatin (Mirzabekov *et al.,* 1989) and to GH5-DNA complexes (Goytisolo *et al.,* 1996), yet is lacking in GH1. The absence of His 62 in GH1 gives this loop a negative charge which might result in the observed different affinities of these domains for DNA. H5 was shown to bind more strongly than H1 to the nucleosome (Sun *et al.,* 1990), and the isolated GH5 has been shown to bind more strongly than GH1 to DNA (Thoma *et al.,* 1983).

1.4 Linker histone binding to DNA.

NMR studies on the protein fold of GH5 (Clore *et al.*, 1987) demonstrated the structural similarities between the globular domain of H5 and the C-terminal DNA binding domain of the cAMP receptor protein (CAP) of *E. coli.* The CAP DNA binding domain also exists as a three-helix bundle with a C-terminal hairpin. Structural similarities between GH5 and HNF-3γ, another member of the winged helix family of helix-loop-helix transcription factors (Clark *et al.,* 1993) also showed a similar structure, and a primary DNA-binding site. This site was suggested to consist of helix III, the C-terminal β-hairpin and the loop between helices II and III.

Comparison with the crystal structure of GH5 confirmed these similarities and a model for the binding of GH5 to DNA was proposed, based on the mode of DNA binding by CAP (Ramakrishnan *et al.,* 1993).

The proposed model of DNA binding by the globular domain is in agreement with various biochemical studies. Protection from reductive methylation (Thomas *et al.,* 1986) was used to determine which H5 lysine residues were in contact with the chromatin fiber. That study implicated Lys 69 and in particular Lys 85, situated in the globular domain of the linker histone H5, in binding to DNA. The crystal structure of GH5 showed that Lys 69 and Arg 73 are in positions to interact with one strand of DNA and that Lys 85 would be able to interact with the other strand of DNA. All these residues have corresponding residues in CAP that bind the DNA strands in a similar manner in the CAP DNA co-crystal structure.

Buckle and co-workers (1992), using site-directed mutagenesis of the globular domain of H5, determined that mutation of Lys 85 led to the loss of the protection of the chromatosomal DNA, again implicating this residue in DNA binding. The model proposed by Ramakrishnan and colleagues was further supported by chemical crosslinking studies of His 25 and His 62 (Mirzabekov *et al.,* 1989) to DNA in chromatin. In the crystal structure these residues are in a position to interact with nucleosomal DNA.

1.4.1 Binding preferences.

Direct competition experiments showed that linker histone H1 preferentially binds superhelical, as compared to linear or relaxed circular DNA (Yaneva *et al.*, 1990). The ability of the globular domain to bind more than one DNA segment was observed with the cooperative DNA binding of the isolated globular domains from H1 and H5 (Draves *et al.,* 1992; Thomas *et al.,* 1992). This binding resulted in the continuous arrays of globular domains bridging tramlines of DNA.

Synthetic DNA four-way junctions exhibit a strong resemblance to the nucleosomal entry/exit DNA and this makes these structures ideal for *in vitro* studies pertaining to linker histone nucleosomal binding (Lilley *et al.*,1992). Chicken erythrocyte linker histones H1 and H5 form defined complexes with these four-way junctions (Varga-Weisz *et al.*, 1993). The linker histones and the isolated globular domains exhibited a higher affinity for the four-way junction DNA compared to linear double-stranded

DNA that have the same sequence information as the four-way junction DNA (Zlatanova and Van Holde, 1996). The suggestion from all these studies was that GH5 contained more than one DNA-binding site.

1.4.2 Secondary DNA binding site.

Observations made from the crystal structure of GH5 (Ramakrishnan *et al.,* 1993) also suggested the existence of a secondary DNA binding site which consisted of Lys 40, Arg 42, Lys 52 and Arg 94. These residues are situated on the opposite side of the molecule, 25-30 Å away from the primary DNA-binding site. The seven residues that are implicated in the possible secondary binding site are highly conserved in the linker histone H1 family (Crane-Robinson and Ptitsyn, 1989; Wells and McBride, 1991).
When the conserved basic residues in either the primary DNA binding site or the putative secondary binding sites were mutated to alanines or glutamic acids, the resulting mutants were unable to preferentially bind to synthetic four-way junctions (Goytisolo *et al.*,1996). These GH5 mutants could not confer the protection of the additional 20 bp of nucleosomal DNA. These results strongly suggested that these residues were important for the correct binding of GH5 to the nucleosomal DNA.

1.4.3 Structure of C- and N-terminal tails and association with DNA.

The C- and N-terminal tail domains of the linker histones were shown to be unstructured in solution (Aviles *et al.,* 1978). The C-terminal tail region of the linker histone is basic, and rich in lysine, serine, proline and alanine. This region has been shown to be involved in the condensation of the nucleosomes into the 30 nm filament (Thoma *et al.,* 1983). The C-terminal tail is known to assume an α-helical conformation in perchlorate, where the anion may mimic the negatively charged phosphate groups of the DNA backbone (Clark *et al.*, 1988). This α-helical conformation was suggested to track the phosphate backbone of the linker DNA, and in this way determine the conformation of the linker DNA between nucleosomes (Zlatanova *et al.,* 1996).
The C-terminal tail region also contains the phosphorylation sites of the linker histones, a modification that was implicated in changes in chromatin structure that affect transcription (Roth and Allis, 1992).

Little is known about the structure and the association of the N-terminal domain of the linker histone, but it was suggested to be involved in the correct positioning of the linker histone globular domain on the nucleosome (Allan *et al.,* 1980).

1.5 Nucleosomal binding by the globular domain of the linker histone.

1.5.1 Models for the nucleosomal binding.

1.5.1.1 The Allan (1980) model.

The original model for the binding of the lysine-rich linker histone H1 to the nucleosome was proposed by Allan and colleagues (1980). In this model (Figure 1.2) the trypsin generated globular domain of the linker histone H1 was shown to be sufficient for the protection of the additional 20 bp of DNA present in the chromatosome, but that complete condensation of the chromatin required the intact linker histone H1.

The 20 bp protection was shown to be at the exiting and entering 10 bp DNA duplexes, and a symmetrical setting for the globular domain over the dyad was proposed. In this way the globular domain would be in a position to contact both ends of the nucleosomal DNA directly. It was also suggested that the globular domain was bound to the outward facing minor groove of the nucleosomal DNA, which is located at the center of the nucleosome. Experimental support for this hypothesis was provided by DNase I footprinting studies (Staynov and Crane-Robinson, 1988), which indicated a high degree of protection afforded by H1 and H5 at the nucleosomal dyad, and the moderate protection at the ends of the nucleosomal DNA duplexes.

This symmetrical model (Allan *et al.,* 1980) was widely accepted as the model for linker histone binding to the nucleosome in bulk chromatin, where it seemed clear that the chromatosomes in bulk chromatin represented an extension of 2 × 10 bp from the core particle. However, it was still not clear how the H1 molecule, which in itself is asymmetrical (N-terminus, globular domain, extended C-terminus), was accommodated on the symmetrical nucleosome (presuming that one histone molecule binds per nucleosome). In particular, did all the H1 molecules assemble so that the C-terminal tails pointed in the same direction on the chromatin fiber?

Resolution of the structure of the globular domain by the X-ray crystallography (Ramakrishnan *et al.*, 1993) and NMR studies (Cerf *et al.*, 1994) indicated that symmetrical binding by the globular domain at the pseudodyad axis was unlikely. The comparison of the globular domain of the linker histone with other DNA binding proteins suggested that the globular domain binds in the major groove of the nucleosomal DNA. Interpretation (Ramakrishnan *et al.*, 1993) of the winged-helix fold of the GH5 from the crystal structure implied that the globular domain bound to the nucleosomal DNA by inserting helix III into a major groove of the nucleosomal DNA. However, in the nucleosome core the minor groove faces outwards at the dyad position. This implied that the two exiting DNA duplexes might be asymmetrically and imprecisely moved to accommodate this natural asymmetry of the linker histones. This setting would not be expected, since the primary mode of chemical crosslinking between linker histone molecules were head-to-tail (Lennard *et al.*, 1985), which implies regular and not random placement in the chromatin.

The adequacy of the symmetrical model for the binding of the globular domain of the linker histone to the nucleosome was also brought into question by neutron scattering studies (Lambert *et al.*, 1991). Results suggested a more interior location of the globular domain on the chromatosome at a position closer to the mass centre of the structure between two central gyres of the nucleosomal DNA, and thus slightly away from the dyad axis of symmetry.

1.5.1.2 The Wolffe-Hayes model (Hayes *et al.*, 1996; Pruss *et al.*, 1996).

Both the original symmetrical model for the binding of the globular domain to the nucleosome (Allan *et al.*, 1980) and the bridging model (Zhou *et al.*, 1998), discussed in section 1.5.1.3 below, places GH5 on or close to the chromatosomal dyad axis of symmetry. A drastically different asymmetric model was proposed using chromatosomes positioned on the unique *Xenopus borealis* 5S rRNA gene. Hydroxyl radical and DNase I footprinting experiments (Hayes *et al.*, 1993) suggested that the minor groove of the DNA at the dyad axis remained accessible to reagents after linker histone association. This indicated that the dyad axis was not protected by the globular domain of the linker histone as proposed by the symmetrical model (Allan *et al.*, 1980).

Allan Model (1980) **Bridging Model (Zhou *et al.*, 1996)** **Wolffe-Hayes Model (Hayes *et al.*, 1996; Pruss *et al.*, 1996)**

Figure 1.2 The models proposed for the nucleosomal binding by the globular domain of the linker histone.

The typical chromatosome (monomer) pause of 20 bp above that of the 146 bp of the core particle observed upon limited micrococcal nuclease digestion, was, however observed. The extension of the DNA protection was asymmetric representing 5 bp and 15 bp (Hayes *et al.,* 1993). The asymmetric protection shown in this study could be an artefact of the micrococcal nuclease digestion, rather than protection afforded by the binding of the linker histone, due to a preference of this nuclease for cleavage following an AT base pair (McGhee *et al.,* 1983). The majority of the bulk chromosomal sequences cloned and sequenced for the Zhou study contained AT base pairs at the ends of the fragments (Muyldermans and Travers, 1994).

Results from the Hayes study (Hayes *et al.,* 1993) were confirmed by using the zero length histone-DNA crosslinking (Mirzabekov *et al.,* 1989). It was shown that there was no change in the positions of the core histone crosslinking and a GH5 crosslink was observed at ~6 bp from the end of the nucleosomal DNA, which is on the same side as the 15 bp micrococcal nuclease protection site (Hayes *et al.,* 1994). The conclusion was that this crosslinking between the globular domain of the linker histone and the nucleosomal DNA occurred at a single site at a position away from the dyad axis of the nucleosome.

Further demonstration of the off-axis location of GH5 on this specific reconstituted nucleosome was given by Pruss *et al.* (1996). In this study, involving histone-DNA crosslinking with chromatosomes reconstituted on the 5S rRNA gene, the most intense crosslinking was found for the major groove contacts 60-68 bp from the dyad axis of the core particle. This supported the position of the globular domain of H5 at a single site inside the one gyre of the nucleosomal DNA, approximately 65 bp away from the dyad axis of symmetry of the nucleosomal core (Pruss *et al.,* 1996; Hayes, 1996). With this placement, no contact was made by the globular domain with either the nucleosomal dyad axis or the terminal DNA duplexes. The secondary DNA binding site of the globular domain was in contact with the H2A-H2B dimer thus making no contact with the nucleosomal DNA.

This internal position is in agreement with the findings that H2A cross-links with the globular domain of linker histones (Boulikas *et al.*, 1980). Furthermore, it agrees with the DNA co-crystal structure of the DNA binding transcription factor HNF-3/*fork* head domain where the bound DNA bends towards the globular domain (Clark *et al.,* 1996). The neutron scattering studies (Lambert *et al.,* 1991), which places the globular domain near the histone octamer, also supports this model.

Hayes and co-workers (1996) used site-directed cleavage mutagenesis to introduce a cysteine residue into the linker histone H1° of *X. borealis* at the boundary between the globular and C-terminal domains, and mapped the approximate position of the cysteine residue by hydroxyl radical cleavage. Mapping of the single-strand breaks showed a cleavage site ~67 bp from the dyad axis, which placed the globular domain binding site on the inside of the DNA superhelix, near the border of the nucleosome complex. The globular domain was shown to interact asymmetrically with the nucleosomal DNA, in the major groove of the DNA. This placement corresponded well with the crosslinking experiments that placed the globular domain 60-68 bp away from the dyad axis (Pruss *et al.,* 1996; Hayes, 1996).

To test whether the binding of the linker histone with the 5S nucleosome was an exceptional case, which is restricted to this specific nucleosome, or if it was of a more general nature, the analysis was shifted to another positioned nucleosome, which contained the DNA sequence from the *Xenopus laevis* thyroid hormone receptor βA gene (TrβA). In this study (Guschin *et al.,* 1998) the histone DNA crosslinking of this positioned nucleosome, which contained the thyroid hormone response element (TRE), was investigated. It was concluded that the globular domain of the linker histone H1° from *X. laevis* associated asymmetrically with the DNA at the one edge of the TrβA nucleosome. This association led to rearrangement of the core histone-DNA contacts at the dyad axis of the nucleosome. These observations supported the bridging model.

The Wolffe-Hayes model (Pruss *et al.,* 1996; Hayes, 1996) made the assumption that the somatic 5S rDNA chromatosome is uniquely positioned on the two different 5S rDNA fragments, and that in each case there is a separate preferred dyad position (Travers, 1999). Two studies have shown that more than a single position can be occupied by chromatosomes reconstituted with the same *X. borealis* somatic 5S rDNA sequence. Zlatanova and colleagues (1998) questioned the Wolffe-Hayes model (Pruss *et al.,* 1996; Hayes, 1996), and showed that the DNA sequence that was used were prone to artefacts in nuclease digestion when the protection given by linker histones was examined. Mapping dyad positions directly (Panetta *et al.,* 1998), revealed a difference in the locations of the positioned nucleosomes on somatic and oocyte 5S rDNA. This resulted in altered accessibility of the transcription factor binding sites in these two nucleosomes. It was concluded that the population of core particles and chromatosomes, which were formed on the 5S rDNA, was a mixture of

several different translational settings (Travers, 1999). The model proposed by Wolffe-Hayes (Pruss *et al.*, 1996; Hayes, 1996) could correspond with the bridging model proposed by Zhou and co-workers (1998) if at least some of their chromatosomes occupied the positions observed by Rhodes and colleagues (1998) (Travers, 1999).

1.5.1.3 The Bridging model (Zhou *et al.*, 1998).

An asymmetric model for the location of the globular domain of H5, using bulk chromatin, was suggested by Muyldermans and colleagues (1998). In this model, which was derived using a site-specific protein-DNA photo-crosslinking technique (Pendergast *et al.*, 1992), the globular domain was shown to bridge one terminus of DNA and the DNA at a position close to the dyad axis of symmetry of the nucleosomal core. Helix III of GH5 was suggested to bind to the major groove of the first helical turn of the chromatosomal DNA duplex, and the secondary DNA binding site to the DNA close to the nucleosomal dyad. Helix I and III of the globular domain were proposed to be turned to the solvent and to the nucleosome, respectively. This orientation would place the C-terminal region of the globular domain in a position to contact the internucleosomal linker DNA as the strands entered and exited the nucleosome. It was argued that the results could support either asymmetric or symmetric extension of the DNA protection afforded by the binding of the linker histones in the nucleosome. This was due to the fact that binding of the globular domain helix III at one terminus of the chromatosome and binding of the histone tails at the other terminus could support the symmetrical 10 bp extension of nucleosomal DNA protection. However, H5 binding may also result in an asymmetric extension of protection of 20 bp at one side and 0 bp at the other (An *et al.*, 1998).

The bridging model was also supported by other studies: binding of two DNA duplexes by globular domain of histone H5 (Thomas et al., 1992); the existence of two DNA binding sites which are both necessary for the chromatosome formation (Goytisolo *et al.*, 1996); neutron scattering studies, which proposed a binding by the globular domain of two superhelical turns between one terminus and the dyad (Lambert *et al.*, 1991); and crosslinking of the histidine 25 residue to the DNA in the chromatin (Mirzabekov *et al.*, 1989).

1.6 Linker histones and chromatin higher order structure.

1.6.1 The 30 nm chromatin fiber.

Linker histones are involved in stabilising both the nucleosome and the chromatin higher-order structure (Zlatanova, 1996). Understanding both the nucleosomal and chromosomal association, is essential to understand the processes that are regulated by the chromatin structure.

Electron microscopy studies of the salt-dependent conformational transitions of soluble chromatin fragments (Klug *et al.*, 1976; Thoma *et al.*, 1979) indicated that the formation of well-defined 30 nm fibres required the presence of lysine-rich linker histones. Chromatin depleted of H1 also condensed with increasing ionic strength, but no definite structures were observed. The structure of the extended fragment at low ionic strength was also dependent on the presence of linker histones. The H1-containing chromatin fiber had a zig-zag appearance, where the DNA entered and exited the nucleosome on the same side. In H1-depleted chromatin, the entering and exiting points on the nucleosomal DNA were more random, which led to the extended "beads-on-a string" conformation. This led to the conclusion that linker histones mediated the transition between the extended 10 nm nucleosomal filament and the well-defined folded 30 nm filament. These observations were confirmed by more recent scanning-force microscopy studies (Leuda *et al.*, 1994; Yang *et al.*, 1994).

Recent studies have investigated the relationship between linker histones and the core histone tail domains during chromatin condensation (Carruthers and Hansen, 2000). It was shown that the core histone N-termini and the linker histones functioned independently during condensation of the chromatin fiber, where the core histone amino termini specify the formation of the folded chromatin fiber and the linker histones stabilized this folded structure.

The linker histone H5 bound to and restrained the entering and exiting nucleosomal DNA in chromatin model systems that lack the core histone N-termini (Carruthers *et al.*, 2000). However these tailless arrays were unable to fold or undergo intermolecular fiber-fiber interactions. Thus, it seemed apparent that the linker histones are required for the stabilization of the chromatin fiber, but not essential in the condensation process.

1.6.2 Models for the structure of the 30 nm fiber.

On the basis of existing data, various models for the involvement of the linker histones in organization of the nucleosomes into higher order structure have been proposed. Based on images obtained by electron microscopy (Klug *et al.*, 1979), the early model of the 30 nm fiber was proposed, which was a solenoidal structure in which a helical array of nucleosomes, 6-8 per turn, were linked by histone H1 with its globular domain binding to the dyad region located in the interior of the solenoid. This interior location was in agreement with the neutron scattering results (Graziano *et al.*, 1994). The linker DNA connecting the adjacent nucleosomes in the helix was believed to be bent (Yao *et al.*, 1990). In this model, the filament diameter is constant, independent of linker DNA (Thomas *et al.*, 1986). A model where the linker DNA assumed a bent configuration was suggested by Felsenfeld and colleagues (1983). The DNA follows the path of the nucleosomes and is not confined to the centre of the 30 nm fibre, as shown in the solenoidal model.

"Superbeads" were observed by Franke and colleagues (1984), although the results were subsequently attributed to an artefact of the preparation methodology (van Holde, 1989). A zig-zag array was observed, where the nucleosomal arrays do not resemble a filament, but clusters, referred to as "superbeads", which contained varying amounts of nucleosomes. These structures were found at low ionic strength and were dependent on the presence of linker histones. Rattner and colleagues (1984) found that this zig-zag formed a condensed ribbon containing two parallel rows of nucleosomes, coiling of this ribbon generated 30 nm fiber.

1.7 Yeast linker histone H1 (Hho1).

1.7.1 Is there a linker histone H1 in yeast?

Although gene deletion, replacement, and mutagenesis studies in *Saccharomyces cerevisiae* have contributed significantly to our understanding of the influence of chromatin structure and the role of the core histones in DNA function (Grunstein 1990), very little work has been preformed on the linker histone H1. This is mostly due to the fact that endogenous yeast H1 has not been isolated from yeast, and another

contributing factor is the differences in the chromatin structure of yeast and that of higher eukaryotes, as described below.

When compared to that of higher eukaryotes, the yeast genome is 40% transcriptionally active as opposed to approximately 5% (Hereford & Rosbash, 1977). With DNase I digestion studies (Lohr & Hereford, 1979), it was shown that most of the genome is sensitive to digestion by DNase I, which indicates that the chromatin is predominantly in an unfolded conformation. This was confirmed in the recent finding that yeast chromatin is acetylated to a much greater degree than that of higher eukaryotes (Davie *et al.*, 1981; Waterborg, 2000). Another important difference in chromatin structure is the nucleosomal repeat length. In yeast this repeat length is unusually short at ~165 bp (Thomas & Furber, 1976; Lohr *et al.*, 1977).

No biochemical evidence for the existence of a linker histone in yeast has been provided, and with the apparent differences in chromatin structure it was assumed that there is no need for the partial neutralization of the DNA phosphate backbone by a linker histone. Immunological evidence of H1-like proteins in yeast was, however, reported, with the use of anti-mouse H1 antibodies (Srebreva *et al.*, 1987; Smith *et al.*, 1984).

The completion of the *S .cerevisiae* genome sequence (Goffeau *et al.*, 1996; Bussey *et al.*, 1997) allowed searches for homology through homology matching with parts of the genomes of other organisms. The presence of a linker histone in yeast was suggested by such a homology search with the sequence of human H1.1 which identified a homologues sequence between positions 308827 and 309603 of chromosome XVI, subsequently named *HHO1*.

HHO1 encodes a protein (Hho1p) of 258 amino acids in length (Ushinsky *et al.*, 1997) with a predicted molecular weight of ~28 kDa. Homology searches for genes of similar sequence to *HHO1* also indicated only a single copy of this gene in yeast compared to the core histone genes, which are present in two nearly identical copies. Alignment of the protein sequence with itself revealed the presence of a putative second globular domain (Landsman, 1996), which makes it different from the tri-partite structure of the canonical linker histones. The overall structure of the yeast H1 is that of two globular domains which are connected by a 42 amino acid lysine-rich linker region. This region is similar to the C-terminal tail of other linker histones, and contains 12 lysines, 10 alanines and 4 prolines out of the 42 residues (Landsman, 1996).

1.7.2 Hho1p biochemical evidence.

Additional evidence suggesting that Hho1 was a true linker histone included the nuclear localization, using transformants expressing a fusion protein of Hho1p and green fluorescent protein (GFP) (Ushinsky *et al.*, 1997).

With the use of a recombinant Hho1p, biochemical evidence proposing that Hho1p was a linker histone, was provided by Patterton and co-workers (1998). A characteristic property of linker histone binding to the nucleosome is the protection of an additional 20 bp of linker DNA from limit micrococcal nuclease digestion (Simpson, 1978). The ability of the rHho1p to confer a chromatosome stop was tested by the MNase digestion of H1-stripped chromatin, reconstituted with rHho1p. In the presence of Hho1p, the characteristic transient pause was observed as in higher eukaryotes (Noll *et al.*, 1977), indicative of the protection of two full superhelical turns of the nucleosomal DNA from endonucleolytic cleavage.

As shown previously, linker histones bind to the nucleosome, with and without the C- and N-terminal tail regions, although the exact binding mode and position is still unresolved (Allan *et al.*, 1981; Crane-Robinson, 1997). To test whether the purified recombinant (Hho1p) could bind to reconstituted nucleosome cores *in vitro*, the recombinant protein was added at a range of concentrations to a 390 bp radio-labeled fragment containing the tandem repeat of the *Lytechinus variegatus* 5S rDNA reconstituted into di-nucleosomes. The resulting reconstitution products showed the formation of two slow-migrating complexes upon electrophoretic analysis. This implied that the protein formed a stable tertiary complex with a reconstituted di-nucleosome core *in vitro*, similar to the complexes formed by linker histones in higher eukaryotes.

1.7.3 Secondary structure.

Linker histones have been shown to have two DNA binding sites. The globular domains of both H1 and H5 contain a cluster of positively charged residues that constitute the second DNA binding site that is located on the opposite side to the primary DNA binding site (Ramamkrishnan *et al.,* 1993). Both proposed binding sites have been shown to be necessary for the correct binding of the globular domain to the nucleosome. GH5 mutants lacking the residues in the secondary DNA binding site could not confer the characteristic kinetic pause at ~168 bp upon nuclease digestion (Goytisolo *et al.,* 1996, Duggan and Thomas, 2000).

Sequence alignment of the amino acid sequence of the two globular domains of Hho1p with globular domains of various other linker histone isotypes indicate that the globular domain one of yeast linker histone Hho1p contained six of the seven basic residues that constitute the two DNA binding sites of GH5. Globular domain two of this linker histone only contains four out of the seven residues of the two proposed DNA binding sites. Both these domains lack the equivalent of Lys 69 from site one, the primary DNA binding site. Additional alignment analysis showed that this residue is also absent in other linker histones, and might indicate that the residue is not essential for the correct binding of the globular domain to the nucleosomal DNA.

Using a structure prediction algorithm, nnPredict, it was shown (Patterton *et al.,* 1998) that the proposed primary globular domain of Hho1p could assume a secondary structure similar to that of the chicken linker histone H5, shown in the crystal structure (Ramakrishnan *et al.,* 1993). It was suggested that the predicted secondary structure of the assigned globular domain of Hho1p contains a single-winged helix protein fold similar to that observed in GH5 (Ramakrishnan *et al.,* 1993). This also corresponds to the winged helix fold in other DNA binding proteins like CAP and HNF-3γ (Ramakrishnan *et al.,* 1993 and Cerf *et al.,* 1994).

1.8 Project aims.

The ultimate aim of this project is to determine whether the yeast linker histone can bind to two nucleosomes simultaneously, where each of the globular domains bind to an adjacent nucleosomal core, and the lysine-rich inter globular region contacting the linker DNA between the nucleosomes (Figure 1.3).

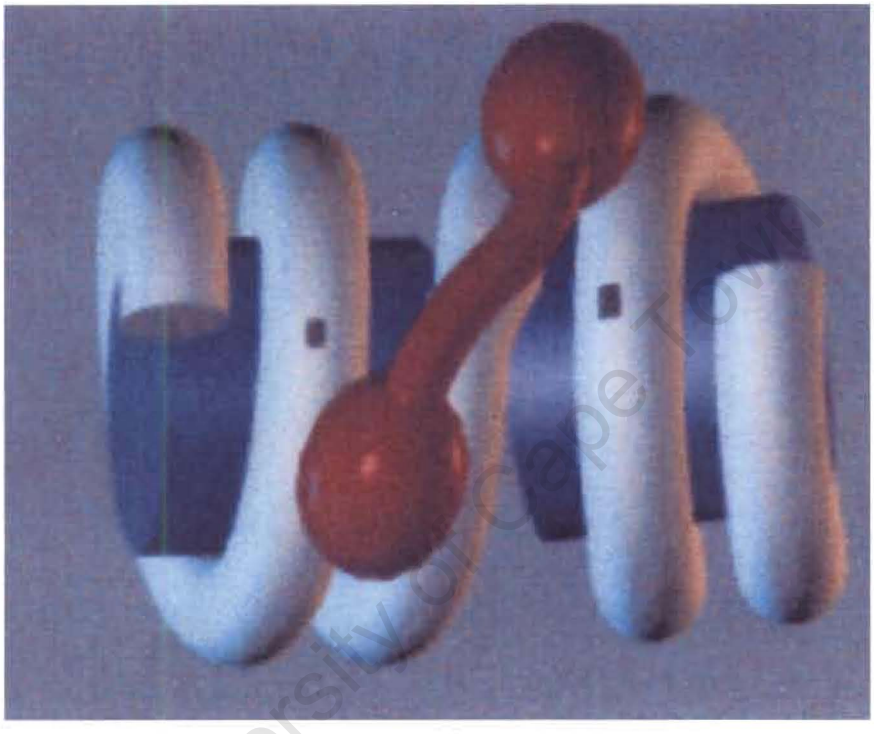


Figure 1.3 Proposed model for the nucleosomal binding by yeast linker histone, Hho1p.

In the part of the study reported in this thesis, a structural assay is developed that could be used to investigate the binding by the yeast linker histone in the manner proposed in the model. Specific reagents were prepared and certain techniques have been optimized to facilitate a detailed study.

The separate globular domains and one control construct to be used in the magnetic bead pull-downs was constructed, over-expressed in *E. coli* and purified (Chapter 3).

A second control construct which will contain two GH5 coding regions connected by the coding region of the linker region from the yeast linker histone Hho1p, was designed and the preliminary cloning was completed, but the cloning of this fragment

into an expression vector still needs to be completed before recombinant protein can be expressed in *E. coli* and purified.

Preliminary DNA binding studies and the magnetic bead technology is illustrated (Chapter 4). DNA binding studies with GH5, which serves as positive control in these studies, and also with the two globular domains, are shown. The conditions for these assays need to be optimized. The optimization of the magnetic bead pull-downs with radio- and biotin-labeled nucleosomes is described and an adapted dot-blot technique for the quantification and detection of the biotin-labeled nucleosomes was set-up (Chapter 4). These reagents will allow definitive experiments which would contribute to structural assays that would attempt to determine if the yeast linker histone (Hho1p) binds to two nucleosome cores simultaneously, and contribute to an understanding of the molecular mechanism by which histone H1 contributes to chromatin condensation and regulates the expression of specific genes.

Chapter 2: Materials and Methods.

2.1 Cloning of the Hho1 domains, including two modified Hho1 constructs.

2.1.1 Plasmid purification.

All molecular biological manipulations were performed according to established protocols (Ausubel *et al.,* 1995). Chemicals and reagents were Molecular Biology Grade. Small-scale plasmid purifications were performed using a QIAquick Plasmid Purification preps (Qiagen, GmbH, Hilden, Germany) and a Wizard Maxiprep kit (Promega Corp., Madison, USA) was used for large-scale purifications, using protocols supplied with each kit.

2.1.2 Primer design.

Synthetic oligonucleotide primers were supplied by the Oligonucleotide Synthesis Facility, Department of Molecular and Cell Biology, University of Cape Town.

2.1.2.1 The native globular domains.

Table 2.1 Sequences of oligonucleotides used for the amplification of GD1 and GD2 fragments.

Name	Sequence
GD1 forward primer (GD1F)	5′-GGGCGCC**ATG**TCCAAGAGTTACAGGGAGTT-3′ *
GD1 reverse primer (GD1R)	5′-GGGCTCGAG**TCA**TTTTCTTCTTGGCCAGTT-3′ #
GD2 forward primer (GD2F)	5′-GGGGCGCC**ATG**TCCTCTTTCGGGTTCTTCA-3′ *
GD2 reverse primer (GD2R)	5′-GGGGCTCGAG**TCA**TTTGACCTTCTTCTGT-3′ #

* The sequence corresponding to the *Sfo I* restriction site is underlined and the start codon is in bold.
\# The sequence corresponding to the *Xho I* restriction site is underlined, and the sequence corresponding to a stop codon is in bold.

2.1.2.2 The modified *HHO1* lysine-rich region.

Table 2.2 Sequences of oligonucleotides used for the amplification of the modified *HHO1* construct; *HHO1*-FXa with the protease recognition site in the lysine-rich region.

Name	Sequence
Flanking primer **Hho1 A** (forward primer)	5′-GGGTGATCATTGAAGGGCTC-3′ *
Flanking primer **Hho1 D** (reverse primer)	5′-GGGAGGCCTTCTTGGCGGTA-3′ #
Template mismatch oligonucleotide **Hho1 C** (forward)	5′-ACGAAGCTA**ATCGAAGGTCGT**GCGCCAAAG-3′ a
Template mismatch oligonucleotide **Hho1 B** (reverse)	5′-CTTTGGCGC**ACGACCTTCGAT**TAGCTTCGT-3′ a

* The sequence corresponding to the *Bcl I* restriction site is underlined.
\# The sequence corresponding to the *Stu I* restriction site is underlined.
a The sequence corresponding to the Factor Xa Protease recognition sequence is in bold.

2.1.2.3 The GH5-linker-GH5 fragment.

Table 2.3 Sequences of oligonucleotides used for the amplification of the modified GH5 and *HHO1* lysine-rich region (*HHO1*-linker fragment), for the construction of GH5-*HHO1*-linker fragment.

Name	Sequence
link-mag F (forward primer)	5′-GGGGTCGACAAGAAAGAAAAAGAGGT-3′ *
link-mag R (reverse primer)	5′-GGGCTCGAGTCTTGGCGGTA-3′ #
GH5-mag F	5′-GGGCTCGAGATGACGGAGTCGGCATCGCACCCCAC-3′ #
GH5-mag R	5′-GGGCTCGAGGGACCTCTTGGCCTTGT-3′ #

* The sequence corresponding to the *Sal I* restriction site is underlined.
\# The sequence corresponding to the *Xho I* restriction site is underlined.

Table 2.4 Sequences of oligonucleotides used for the amplification of the modified GH5 and GH5-*HHO1*-linker region, for the construction of GH5-linker-GH5.

Name	Sequence
GH5-link *NdeI* (forward primer)	5′-GGGCATATGAAGAAAGAAAAAGAGGT-3′ *
GH5-link *Xho I* (reverse primer)	5′-GGGCTCGAGTCTTGGCGGTA-3′ #
GH5-*Xho I*	5′-GGGCTCGAGATGACGGAGTCGGCATCGCACCCCAC-3′ #
GH5-*Sal I*	5′-GGGGTCGACGGACCTCTTGGCCTTGT-3′ a

* The sequence corresponding to the *Nde I* restriction site is underlined.
\# The sequence corresponding to the *Xho I* restriction site is underlined.
a The sequence corresponding to the *Sal I* restriction site is underlined.

2.1.3 Polymerase Chain Reaction (PCR).

2.1.3.1 Amplification of native domains.

The amplification of the coding regions for the globular domains (GD1 & GD2) were performed by the Polymerase Chain Reaction using 10 ng of pRS413-*HHO1* (Patterton *et al.*, 1998) as template in a 50 µℓ volume containing 1 × *Pfu* DNA polymerase reaction buffer (Promega), 3 U *Pfu* DNA polymerase, 150 µM of the four deoxyribonucleotide triphosphates, and 0.5 µM of each primer. The following cycling parameters were used: an initial denaturation step at 94 °C for 1 min, followed by 25 cycles consisting of 94 °C incubation for 1 min, an annealing step at 60 °C for 1 min, and an extension step at 72 °C for 1.5 min. A final step of 72 °C for 5 min was included to complete partial extension reactions. All PCR products were verified by agarose gel electrophoresis (section 2.2.1.5).

2.1.3.2 Template-mismatched PCR mutagenesis.

This procedure was based on the methods by Higuchi and co-workers (1988) and Ho and co-workers (1989). The first step was the separate amplification of overlapping regions of the coding region for *HHO1*. In these PCR reactions, 12 ng of pRS413-*HHO1* (Patterton *et al.*, 1998) was used as template in a 50 µℓ volume containing 1 × *Pfu* DNA polymerase reaction buffer (Promega), 3 U *Pfu* DNA polymerase (Promega), 150 µM of the four deoxyribonucleotide triphosphates and 0.5 µM of mismatch oligonucleotide needed to introduce the Factor Xa protease site, and 0.5 µM of the a flanking oligonucleotide Hho1 C (forward) or Hho1 B (reverse).

In the first reaction, Hho1 A was used as flanking primer and Hho1 B as mismatch primer. In the same way Hho1 D was used with the mismatch primer Hho1 C in a separate reaction. PCR products were verified by agarose gel electrophoresis (section 2.2.1.5). The amplified fragments were recovered form a 1 % (w/v) low melting agarose gel. DNA was purified using QIAquick Gel Extraction kit (Qiagen) using the manufacturer's supplied protocol (section 2.2.1.6).

In the subsequent overlap extension PCR reaction, the purified PCR products were used as template for the amplification of the full-length coding sequence of the

modified Hho1. In the overlap extension PCR, 2 ng of the template DNA (equal concentrations of the appropriate overlapping PCR products mentioned above) was used in a 48 µℓ volume containing 1 × *Pfu* DNA polymerase reaction buffer (Promega), 3 U *Pfu* DNA polymerase and 150 µM of the four deoxyribonucleotide triphosphates. This was followed by a extension step of 2 cycles of denaturation for 1 min at 94 °C, annealing of the template at 55 °C for 1 min and extension by *Pfu* DNA polymerase at 72 °C for 1 min. This two-step PCR cycling was found to minimize the amplification of contaminants, and allowed for the formation of the full-length template prior to the addition of the flanking primers.

Following this initial 2 cycle PCR, Hho1 A (forward flanking primer) and Hho1 D (reverse flanking primer) were added to a final concentration of 0.5 µM and the reaction allowed to continue at the following cycling parameters: an initial denaturation at 94 °C for 1 min followed by 25 cycles consisting of a denaturation step at 94 °C for 1 min, an annealing step at 55 °C for 1 min, and an extension step at 72 °C for 1.5 min. The amplification was concluded by incubating the sample at 72 °C for 5 min to complete partial extension reactions, and the sample then transferred to 4 °C. PCR products were verified using agarose gel electrophoresis (section 2.2.1.5).

2.1.3.3 The GH5-linker-GH5 fragment.

The amplification of coding regions for GH5 and the *HHO1* lysine-rich region (linker) were performed by PCR using 10 ng of each of the coding region GH5 and 10 ng of the coding region of *HHO1* as templates for the GH5 and *HHO1*-linker fragments respectively. These coding regions were obtained by the *Nde I / Xho I* digestion of pET-20b(+) and pRS413-*HHO1*, respectively. The digested fragments were subsequently isolated from a 1 % (w/v) low melting agarose (section 2.2.1.6). An aliquot of 0.5 µM of each primer were used (indicated in Table 2.3) in a 50 µℓ reaction volume containing 1 × *Pfu* DNA polymerase reaction buffer (Promega), 3 U *Pfu* DNA polymerase and 150 µM of the four deoxyribonucleotide triphosphates. The following cycling parameters were used: an initial denaturation step at 94 °C for 1 min, followed by 25 cycles consisting of 94 °C incubation for 1 min, an annealing

step at 60 °C for 1 min, and an extension step at 72 °C for 1.5 min. A final step of 72 °C for 5 min was included to complete partial extension reactions.

All PCR products were verified by agarose gel electrophoresis (section 2.2.1.5). The same reaction conditions were used for the amplification of coding regions for GH5 and the GH5-*HHO1*-linker fragment using 10 ng of GH5 and GH5-*HHO1*-linker fragments as templates respectively. PCR products were verified using agarose gel electrophoresis (section 2.2.1.5).

2.1.4 Restriction and ligation of the coding sequences into *E. coli* expression vectors.

2.1.4.1 Globular domain one and globular domain two.

Expression vectors for globular domain 1 and 2 (GD1 and GD2) of Hho1p containing a N-terminus hexa histidine-tag were constructed in pPRO Ex A (Novagen Life Technologies, Inc.).

The amplified fragments corresponding to GD1 and GD2 were recovered from a 1 % (w/v) low melting agarose gel (section 2.2.1.6.) and purified with the QIAquick Gel Extraction Kit (Qiagen). The procedure was performed according to the manufacturer's recommendations. The isolated fragments were then digested with 10 U of the restriction enzymes *Sfo I* and *Xho I* (Promega) for 4 hr at 37 °C. The digested fragments were recovered from a 1 % (w/v) low melting agarose gel, and purified with the QIAquick Gel Extraction Kit (Qiagen).

A plasmid linearized by *Sfo I* / *Xho I* cleavage was dephosphorylated in a 30 µℓ reaction volume consisting of 1 × Shrimp Alkaline Phosphatase reaction buffer (Promega), 300 ng pPRO EX A and 0.3 U of Shrimp Alkaline Phosphatase. The mixture was incubated at 37 °C for 15 min, and the enzyme inactivated by incubation of the reaction at 65 °C for 15 min. Approximately 16 ng of the purified fragment (in five-fold molar excess to vector) was mixed with 50 ng dephosphorylated of pPRO Ex A (Novagen), 1 µℓ 10 × Ligation Buffer (Promega) and 1 µℓ of T4 DNA Ligase (Promega, 3 Weiss U/µℓ) in a total volume of 10 µℓ. An experimental control (background control) was included where the digested GD1 or GD2 insert fragment was omitted. The ligation reaction was allowed to proceed overnight at 4 °C.

Following ligation, 5 μℓ of the ligation mix was added to 100 μℓ of calcium chloride competent DH5α *E. coli* cells (Invitrogen, Gibco BRL, U.S.A.). The suspension was incubated on ice for 30 min after which the cells were heat-shocked, according to the methodology of Inoue and co-workers (1990) at 42 °C for 45 sec. The transformed cells were immediately placed on ice (4 °C) for a further 2 min and 900 μℓ of room temperature SOC medium [0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM $NaCl_2$, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$ and 20 mM glucose] was added. Cells were incubated with shaking (150 rpm) at 37 °C for 1 hr. Following incubation, 100 μℓ of the transformed cells were plated onto LB agar plates containing 50 μg/mℓ ampicillin. Plates were incubated for 14 hr at 37 °C.

Plasmid DNA was isolated (see section 2.1.1) from overnight cultures containing 10 mℓ LB medium supplemented with 50 μg/mℓ ampicillin and colonies chosen randomly from the incubated plates. Positive constructs were identified by restriction enzyme screening with the restriction enzymes *Sfo I* and *Xho I* (Promega), and verified by a secondary PCR screening with the oligonucleotide primers initially used to amplify the inserts. The correct fragments from both procedures were verified by electrophoresis on a 1 % (w/v) agarose gel stained with ethidium bromide (0.5 μg/mℓ) as described below (section 2.2.1.5). Plasmids found to contain correct inserts were sequenced by automated nucleotide sequencing (section 2.2.1.6), and denoted pPRO Ex A-GD1 (globular domain 1) and pPRO Ex A-GD2 (globular domain 2). Glycerol stocks were prepared of the *E. coli* DH5α cells containing the pPRO Ex A plasmids with the coding regions for GD1 and GD2 and stored at -70 °C.

pPRO Ex A-GD1 and pPRO Ex A-GD2 were transformed into *E. coli* strain BL21 (DE3) containing a plasmid encoding T7 lysozyme (pLysS) as described above. The transformed cells were plated onto LB agar containing ampicillin (50 μg/mℓ) and chloramphenicol (34 μg/mℓ).

2.1.4.2 Modified *HHO1*-FXa.

The method used for the cloning of the modified *HHO1* into an expression vector was as described for the cloning of the globular domains into pPRO Ex A (2.1.4.1), with a few modifications. The expression vector used for the modified Hho1p, namely *HHO1*-FXa, was pET-20b(+), containing a C-terminus hexa histidine-tag (Novagen).

The amplified fragment corresponding to *HHO1*-FXa was recovered from a 1 % (w/v) low melting agarose gel (section 2.2.1.6), and purified with the QIAquick Gel Extraction Kit (Qiagen). The fragments were digested with 10 U of the restriction enzymes *Bcl I* and *Stu I* (Promega) for 4 hr at 37 °C. The digested fragments were recovered from a 1 % (w/v) low melting agarose gel and purified with the QIAquick Gel Extraction Kit (Qiagen). The isolated plasmid cleaved with *Bcl I* and *Stu I*, was dephophorylated by alkaline phosphatase treatment, as described above.

The *Bcl I / Stu I* digested *HHO1*-FXa fragment was mixed with the treated vector at 5-fold molar excess, and ligated as described above. Following ligation 5 μℓ of the ligation mix was added to 100 μℓ of calcium chloride competent DH5α *E. coli* cells and transformed as described above.

Plasmids containing the correct mutated region in the *HHO1* lysine-rich region were identified using automated nucleotide sequencing (section 2.2.1.7). Plasmids were denoted pET-20b(+)-*HHO1*-FXa. Glycerol stocks were prepared of the *E. coli* DH5α cells containing the pET-20b(+) plasmids, and stored at -70 °C.

pET-20b(+)-*HHO1*-FXa was transformed into *E. coli* strain BL21 (DE3) containing a plasmid encoding T7 lysozyme (pLysS) using the methodology of Inoue and colleagues (described above). The transformed cells were plated onto LB agar containing ampicillin (50 μg/mℓ) and chloramphenicol (34 μg/mℓ).

2.1.4.3 GH5-linker-GH5.

The amplified fragments corresponding to GH5 and *HHO1*-linker (lysine-rich region) were recovered from a 1 % (w/v) low melting agarose gel (section 2.2.1.6.) and purified with the QIAquick Gel Extraction Kit (Qiagen). The procedure was performed according to the manufacturer's recommendations. The isolated fragments were then digested with 10 U of the restriction enzymes *Sal I* and *Xho I* (Promega) for 4 hr at 37 °C. The digested fragments were recovered from a 1 % (w/v) low melting agarose gel, and purified with the QIAquick Gel Extraction Kit (Qiagen). These fragments were ligated at a 1:5 GH5 to *HHO1*-linker molar ratio with 1 μℓ 10 × Ligation Buffer (Promega) and 1 μℓ of T4 DNA Ligase (Promega, 3 Weiss U/μℓ) in a total volume of 10 μℓ. The ligation product was isolated from a 1 % (w/v) low melting agarose gel (section 2.2.1.6.).

The isolated GH5-*HHO1*-linker fragment was digested with 10 U of *Taq I* restriction enzyme for 4 hr at 37 °C. The digested fragments were verified by agarose gel electrophoresis (section 2.2.1.5).

Following the amplification and 1 % (w/v) low melting agarose isolation (section 2.2.1.6.) of the GH5 and GH5-*HHO1*-linker fragments, the fragments were ligated at a 1:1 molar ratio with 1 μℓ 10 × Ligation Buffer (Promega) and 1 μℓ of DNA T4 Ligase (Promega, 3 Weiss U/μℓ) in a total volume of 10 μℓ. The GH5-linker-GH5 ligation product was isolated from a 1 % (w/v) low melting agarose gel (section 2.2.1.6) and digested with 10 U of *Xho I* for 4 hr at 37 °C. The digested fragments were verified by agarose gel electrophoresis (section 2.2.1.5). The restriction enzyme digestion identified the GH5-linker-GH5 fragment ligated in the correct orientation. This fragment was sequenced verified with automated nucleotide sequencing.

2.1.5 Agarose gel electrophoresis.

All DNA samples including PCR products, restriction enzyme digestion products and screening products, were verified by agarose gel electrophoresis through a 1 % (w/v) agarose gels in 1 × TAE electrophoresis buffer [40 mM Tris acetate (pH 8.3), 2 mM EDTA]. Samples were electrophoresed at constant voltage of 130 V for 45 min at room temperature. Gels were stained with ethidium bromide (0.5 μg/mℓ). DNA was visualised by short wavelength UV transillumination (254 nm).

2.1.6 Gel isolation and purification of DNA.

Samples were electrophoresised on a 1 % low melting point agarose gel in 1 × TAE electrophoresis buffer (containing 0.5 μg ethidium bromide/mℓ) at 75 V for 60 min. DNA was visualised by brief exposure to low wavelength UV (320 nm). The appropriate bands were excised from the gels with a sterile blade. Gel slices were melted by incubation at 70 °C in a waterbath and DNA purified using the QIAquick Gel Extraction kit (Qiagen), according to the manufactures instructions. DNA concentrations were determined from the absorbance of samples at 260 nm using the conversion factor for double stranded DNA (1 OD unit = 50 μg/mℓ).

2.2 Over-expression of recombinant proteins in *E. coli*.

2.2.1 Growth and induction time course.

To establish whether the recombinant proteins were stably expressed by BL21(DE3) *E. coli* cells transformed with pLysS, a single colony of BL21(DE3)pLysS containing either of pET-20b(+)-*HHO1*, pET-20b(+)-*HHO1*-FXa, pPRO Ex A-GD1 or pPRO Ex A-GD2 was inoculated into 10 mℓ liquid LB media containing ampicillin (50 μg/mℓ) and chloramphenicol (34 μg/mℓ). Cultures were grown with shaking (500 rpm) at 37 °C to an OD_{600} of 0.7. A 1 mℓ volume of the culture was removed to serve as a pre-induction control, and expression induced from the T7 promoter by addition with isopropyl-β, D-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. Cultures were returned to 37 °C. 1 mℓ aliquot was removed at time intervals 60, 90, 120 and 180 min. Aliquots were centrifuged for 20 sec in a benchtop microcentrifuge, the supernatant decanted, and the cell pellets frozen at -20 °C.

Following the induction time-course, 20 μℓ of 5 × SDS-PAGE sample application buffer [0.5 M Tris·Cl (pH 6.8), 10 % (w/v) sodium dodecyl sulphate (SDS), 0.1 % (w/v) bromophenol blue and 10 % (v/v) glycerol], to which β-mercaptoethanol (Merck) had been added to a final concentration of 2 % (v/v), was added to each pellet. Samples were denatured at 90 °C on a heating block, and loaded onto a 15 % (w/v) polyacrylamide gel. Gels were electrophoresed at 200 V for 45 min. Protein was visualized by Coomassie stain as described (section 2.4.1).

2.3 Protein purification of recombinant proteins.

Large scale protein expression was achieved by inoculating a single colony of BL21(DE3)pLysS containing either pET-20b(+)-*HHO1*-FXa, pPRO Ex A-GD1, pPRO Ex A-GD2 or pET-20b(+)-*HHO1* in 10 mℓ LB medium, containing ampicillin (50 μg/mℓ) and chloramphenicol (34 μg/mℓ). The cultures were grown at 37 °C to an OD_{600} of 0.7. This whole culture was then added to 1-2 ℓ of fresh LB media supplemented with the appropriate antibiotics, and the cells cultured to an OD_{600} of 0.7. These bacterial cells were harvested in a JA 14 rotor (Beckman) at 8 000 × g for 10 min at 4 °C. The cell pellets were resuspended in fresh LB media supplemented

with the appropriate antibiotics. Expression was induced with a final concentration of 0.4 mM IPTG and cultures grown for further 3 hr at 37 °C. In a separate experiment 1 ℓ cultures in 5 ℓ flasks (well aerated) were grown to an OD_{600} of 0.4, induced with 0.4 mM IPTG and then harvested. Recombinant protein expression using these conditions was vastly improved.

The bacterial cells were harvested by centrifugation in a JA 14 rotor (Beckman), using a J2-21 centrifuge (Beckman) at 10 000 × g for 20 min at 4 °C. The cell pellets were resuspended in 40 mℓ per litre of binding buffer [5 mM imidazole, 500 mM NaCl, 20 mM Tris Cl, pH 7.9]. The cells were lysed by sonication using a 0.4 cm tip (red) fitted on a Bronwill Biosonik III sonicator (Bronwill Scientific, Inc., New York) at an intensity setting of 30. Sonication of the cells was performed on ice for sixty 45 sec pulses with 15 sec rest on ice, or until samples were no longer viscous. The cellular debris was pelleted by centrifugation at 39 000 × g in a JA 20 rotor (Beckman) for 20 min at 4 °C.

2.3.1 Immobilised metal affinity chromatography.

The recombinant proteins were purified by immobilized metal affinity chromatography (IMAC) using a His Bind kit (Novagen). A 50 % slurry of nickel agarose resin (bed volume of 2.5 mℓ) was packed into a plastic column provided by the manufacturer. The storage buffer was allowed to drain to the bed level after which 7.5 mℓ milli-Q H_2O was added. The column was then charged with 12.5 mℓ charge buffer (50 mM $NiSO_4$) and equilibrated in 7.5 mℓ of binding buffer. The crude extract was applied to the column after which the column was washed with 25 mℓ of binding buffer, followed by 15 mℓ of wash buffer [40 mM imidazole, 500 mM NaCl, 20 mM Tris Cl, pH 7.9]. Recombinant proteins were then eluted with 15 mℓ elution buffer [1 M imidazole, 0.5 M NaCl, 20 mM Tris Cl, pH 7.9].

Fractions of 1 mℓ were collected. All fractions were stored at -20 °C and a 20 µℓ aliquot from each step (crude extract, flow through, binding wash, wash and elution fractions) was loaded onto a 15 % (w/v) SDS-PAGE gel, and electrophoresed at 200 V for 1 hr. Protein was visualized by Coomassie stain as described (section 2.4.1).

2.3.2 Ion-exchange chromatography of Hho1p, Hho1-FXap and GD2 on Biorex-70 cation exchange resin.

The Ni-agarose fractions collected above were also purified using Biorex-70 (Bio-Rad, Bio-Rad Laboratories Ltd., United Kingdom). The Ni-agarose fractions were pooled and the buffer changed in a 10 kDa molecular Centricon Centrifugal Ultrafilter column (Millipore, Millipore Corporation).

The different buffering conditions employed were:

1. 0.1 M KH_2PO_4 (pH 7.2), 1 mM EDTA
2. 0.1 M KH_2PO_4 (pH 7.2), 1 mM EDTA, 8 % GuHCl
3. 20 mM CAPS (pH 10.5), 1 mM EDTA, 8 % GuHCl
4. 20 mM CAPS (pH 10.5), 1 mM EDTA
5. 10 mM Tris Cl (pH 7.5), 1 mM EDTA

The samples were loaded onto a 0.2-0.4 mℓ bed-volume Biorex-70 cation exchange column, equilibrated with 10 bed volumes of the same buffer used for the preparation of the protein sample. Columns had a 0.7 cm diameter with a height of 8.5 cm. Dependant on the bed volume and total protein concentration, 50-100 µg of the recombinant protein of interest was loaded on the equilibrated column. Unbound protein was washed from the column with 3 bed volumes of the loading buffer. Protein applied to the columns that were pre-equilibrated with the potassium phosphate and CAPS buffers, were eluted with a linear gradient of 2 bed volumes of a 10-25 % (w/v) guanidinium chloride (GuHCl) followed by a final elution step of 2 bed volumes 50 % (w/v). Columns and protein samples equilibrated with the 10 mM Tris Cl (pH 7.5) buffer, were eluted with a linear 50 mM-2 M NaCl gradient, where increasing concentrations had a 2 × bed volume.

All the guanidinium solutions were buffered at pH 7.2 with 100 mM potassium phosphate, 1 mm EDTA and all the NaCl solutions were buffered to pH 7.5 with 10 mM Tris Cl. Fractions containing GuHCl were dialyzed against 1 mM Tris Cl, 1 mM EDTA, 0.1 mM PMSF using Dispo-Biodialyzer (AmiKa Corporation, Columbia, USA) 5 kDa molecular weight cut-off micro dialysis system. This allowed for the dialysis of small volumes. Aliquots of 100 µℓ of each sample were applied to the

membrane and dialysis proceeded overnight at 4 °C with constant circulation of 1 ℓ dialysis buffer.

Fractions were collected (fraction sizes collected were 2 bed volumes) and 20 $\mu\ell$ of each fraction were analysed by 15 % (w/v) SDS-PAGE gel and electrophoresed at 200 V for 45 min. Protein was visualised by Coomassie staining as described (section 2.4.1).

2.3.3 Reverse-phase high performance liquid chromatography of Hho1p, Hho1-Fxap, GD1 and GD2.

Reverse-phase HPLC (RP-HPLC) was employed using a Jupiter C18 column (Phenomenex, U.S.A), with a 300 Å pore-size and a 5-10 μm range of particle sizes. The column was washed and equilibrated in 1 % trifluoroacetic acid (TFA).

Concentrated samples form the Ni-agarose elution were prepared by washing the protein sample in 0.1 % TFA and centrifugation at 8 000 × g in a benchtop microcentrifuge to remove any debris. Aliquots of 100 $\mu\ell$ of the supernatant from this centrifugation were injected in the column. The linear gradient of 0.1 % aqueous TFA to 70 % (v/v) acetonitrile in 0.1 % TFA (v/v), with a flow-rate 0.7 mℓ/min was used. The low pH (2.0) of this system aids in the suppression of undesirable ionic interactions between the protein and the packing material, which is caused by the presence of non-derivatized silanols (negatively charged above pH 4-4.5) (Guo *et al.*, 1986 and Mant *et al.*, 1991).

The absorbance of the eluent was recorded at 229 nm. Fractions of 1.5 mℓ were collected, and 20 $\mu\ell$ of each fraction loaded onto a 15 % (w/v) SDS-PAGE gel, electrophoresed at 200 V for 45 min. Protein was visualised by Coomassie staining as described (section 2.4.1). The remaining volume of the collected fraction was frozen at -70 °C for 1 hr and lyophilized overnight. The lyophilized protein was resuspended in the appropriate buffer described in the text and stored at -20 °C.

2.3.4 Gel filtration of Hho1p and GD2.

Bio-Gel P-10 (fractionation range of 1.5-20 kDa) was used for the separation of GD2 and Bio-Gel P-60 (fractionation range of 3-60 kDa) was used for Hho1p and Hho1-FXap separations. Column sizes were chosen according to the volume of the sample relative to the bed volume and a range of different bed volumes were employed. Resin preparation was according to the manufacturer's instructions, and after degassing and packing, the resin was equilibrated in 500 mM Tris Cl (pH 7.5), 1 mM EDTA.
Fractions were transferred to 500 mM Tris Cl (pH 7.5), 1 mM EDTA using a Centricon ultrafiltration device as described and loaded and eluted with 7-10 bed volumes of the same buffer. The volume of the fractions to be collected was determined by the concentration of the desired protein that was loaded, and the bed volume of the column.
Aliquots of 20 $\mu\ell$ of each fraction was loaded onto a 15 % (w/v) SDS-PAGE gel and electrophoresed at 200 V for 45 min. Protein was visualised by Coomassie staining as described (section 2.4.1).

2.3.5 Concentration and buffer changes of all purified proteins.

Protein fraction from the various purification steps were concentrated and buffers changed by using a Centricon Centrifugal Filter Devices YM10 and YM30 MW Membranes (Millipore), according to the manufacturer's instructions. Briefly, 2 mℓ of the sample was loaded into the sample reservoir and centrifuged at 5000 × g using a JA 20.1 rotor (Beckman). All centrifugation steps were performed at room temperature for 1 hr.

2.3.6 Protein quantification.

Protein concentrations were determined using a microassay procedure of the Bradford method (Bradford, 1976) (Bio-Rad). Five dilutions of a BSA standard (Bio-Rad), was used in the linear range from 0.1 mg/mℓ-0.5 mg/mℓ. Incubation of the protein and dye (Coomassie Brillant Blue G-250) was performed at room temperature for 5 min, and the absorbance measured at 595 nm in a microplate reader (Titertek Multiskan PLUS,

Flow Laboratories, Finland). Comparison of the test samples to a standard curve provided the protein concentration.

2.3.7 Removal of N-terminal hexa histidine-tag.

The N-terminal hexa histidine-tag was removed from the purified GD1 and GD2 by digestion with recombinant TEV protease (Invitrogen, Life Technologies, U.S.A.). A 20 μg amount of the purified GD1 and GD2 as digested in a 150 μℓ reaction volume containing 7.5 μℓ 20 × rTEV buffer, 20 U rTEV Protease and 0.1 M DTT at 30 °C for 4 hr.

Aliquots of 25 μℓ from each sample were loaded onto a 15 % (w/v) SDS-PAGE gel and electrophoresed at 200 V for 45 min. Protein was visualised by Coomassie staining as described (see section 2.4.1).

2.3.8 Cleavage of Hho1-FXap with Factor Xa protease.

The modified Hho1-FXap which contains the cleavage site for Factor Xa protease (Promega) in the lysine-rich region between the two globular domains of Hho1p, was digested in a 100 μℓ reaction volume consisting of 35 μg Hho1-FXap, 2 U Factor Xa Protease, 40 mM Tris Cl (pH 7.3) and 80 mM NaCl. The digestion was performed at 25 °C for 1 hr.

An aliquot of 25 μℓ of sample was loaded onto a 15 % (w/v) SDS-PAGE gel, and electrophoresed at 200 V for 45 min. The digested fragments were visualised by Coomassie staining as described (see section 2.4.1).

2.4 Characterization of Hho1p domains.

2.4.1 SDS-PAGE gel electrophoresis of proteins.

All proteins were analysed by SDS-PAGE electrophoresis as described by Laemmli (1970). Gels consisted of a 15 % (w/v) polyacrylamide (acrylamide:bisacryamide 29:1) containing 375 mM Tris Cl (pH 8.8) and 0.1 % (w/v) SDS. The stacking gel consisted of 6 % (acrylamide:bisacryamide 29:1) (w/v) polyacrylamide containing

125 mM Tris Cl (pH 6.8) and 0.1 % (w/v) SDS. Samples were denatured at 95 °C in 1 × sample application buffer (83 mM Tris Cl (pH 6.8), 1.6 % (w/v) SDS, 0.016 % (w/v) Bromophenol blue and 1.6 % glycerol (v/v) containing 25 % (v/v) β-mercaptoethanol. Gels were electrophoresed at 200 V for 1 hr in a MiniProtean III (Bio-Rad) electrophoresis system.

Following electrophoresis, gels were stained with Coomassie Blue stain solution [0.25 % (w/v) Coomassie Brilliant Blue in 100 % methanol] for 1 hr and destained overnight in destain solution [7 % (v/v) acetic acid, 25 % (v/v) ethanol].

Ethanol was substituted with methanol in the destain solution where the protein samples were to be analysed by matrix assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF).

2.4.2 In-gel trypsin digest and MALDI-TOF analysis.

After SDS-PAGE electrophoresis and Coomassie Brillant Blue staining of the recombinant proteins, the gels were destained overnight in 7 % (v/v) acetic acid, 25 % (v/v) methanol. Bands corresponding to the proteins of interest were excised form the gel for analysis by in-gel trypsin digestion.

The positive experimental control used was the band corresponding to carboxy anhydrase in the marker lane containing the Broad Range Molecular Marker (Bio-Rad Laboratories, Hercules, CA, USA) and a blank gel piece was used as negative control. All gel slices were washed twice with 200 μℓ 50 % (v/v) acetonitrile, 25 mM NH_4HCO_3. The solution was removed after a 15 sec incubation step. Gel slices were then washed with 200 μℓ (v/v) 100 % acetonitrile and dried on a rotary evaporator for 30 min. Samples were digested overnight at 37 °C with trypsin (20 μℓ of 20 μg/mℓ, Sigma Aldrich) in 25 mM NH_4HCO_3. Acetonitrile and TFA [50 % (v/v) and 5 % (v/v), respectively] were added to a final volume of 70 μℓ and samples incubated for 30 min. The remaining solution was transferred to a clean eppendorf tube and dried.

Samples were subsequently dissolved in 60 μℓ 0.05 % (v/v) TFA, 5 % (v/v) acetonitrile and mixed in a 1:1 dilution with α-cyano-4-hydroxycinnamic acid [10 mg/mℓ in 60 % (v/v) acetonitrile, 0.3 % (v/v) TFA]. An aliquot (2 μℓ) was spotted on a gold coated MALDI protein grid. The grid was air-dried and inserted into a MALDI-TOF (Perspective Biosystems Voyger DE.Pro) for analyses.

Analyses were performed at an accelerating voltage of 20 kV and a grid voltage of 74 %, with the negative-ion selector switched off. Proteins were identified by comparing the patterns of obtained fragment $^{m}/_{z}$ –values to databases of predicted patterns (Clauser *et al.*, 1999) using the MS-Fit Program from the UCSF Mass Spectrometry facility available at ***http://prospector.ucsf.edu***.

2.4.3 Circular Dichroism (CD).

Proteins were concentrated as described (section 2.3.5) and washed in three separate buffers: 10 mM $NaPO_4$ (pH 7), 0.5 M NaCl and in 100 mM $NaPO_4$. Protein concentration was determined by Bradford assay (section 2.3.6) and adjusted to 0.2 mg/mℓ. CD spectra were recorded in the range 190-250 nm with measurements every 0.5 nm. Spectra were recorded on a JASCO J-180 CD spectropolarimeter in cuvettes with a 1 mm path length. Spectra represent the average of three scans.

2.5 Preparation of H1-stripped long chromatin from chicken erythrocyte nuclei.

Where possible, all procedures were performed on ice, or in a cold room at 4 °C.

2.5.1 Isolation of chicken erythrocyte nuclei.

Fresh chicken blood (200 mℓ) was collected in a plastic beaker filled with 70 mℓ of ACD [15.6 mM citrate, 89 mM Na_3citrate, 16 mM NaH_2PO_4 and 129 mM glucose]. The suspension was filtered through 3 layers of cheesecloth into a 500 mℓ bottle, and immediately placed on ice. Eight JA 20 (Beckman) centrifuge tubes were filled to half maximal volumes with chicken blood, and topped up with 1 × SSC [10 mM trisodium citrate, 150 mM NaCl]. Cells were washed three times by centrifugation in a JA 20 rotor at 4 °C for 5 min at 1000 × g followed by gentle resuspension with a glass rod. The supernatants were removed by aspiration. Hereafter, cells were washed repeatedly in SSC buffer supplemented with 0.1 % (v/v) Triton X-100 until only cream-coloured nuclei was visible in the pellet.

Nuclei were washed three times in buffer A [15 mM Tris Cl, 65 mM NaCl, 60 mM KCL, 0.15 mM spermine, 0.5 mM spermidine, 0.2 mM EDTA, 0.2 mM EGTA, 5 mM

2-mercaptoethanol, 0.2 mM PMSF] and stored in buffer A containing 50 % (v/v) glycerol at -20 °C.

2.5.2 Isolation of H1-stripped long chicken chromatin.

Nuclei were washed three times in buffer A to remove glycerol, after which the absorbance of the nuclei suspension was determined in 0.1 M NaOH at 260 nm. The nuclei were adjusted to 50 OD/mℓ (5 mg/mℓ chromatin) in buffer A.

2.5.3 Trial micrococcal nuclease digestion.

A trial MNase (micrococcal nuclease) digestion was performed in order to determine the appropriate time for digestion to produce long chromatin. A 1 mℓ aliquot of the nuclei was adjusted to 1 mM $CaCl_2$ and incubated at 37 °C for 3 min. 40 U MNase (Sigma) was added, and 200 µℓ aliquots were removed at 0, 5, 10, 15 and 20 min. The digestion was quenched at each of these time points by the addition of EDTA to a final concentration of 2.5 mM and placing the samples on ice. Samples were centrifuged at 5000 × g for 10 min in a benchtop microcentrifuge.

The supernatant was discarded and the pellet resuspended in 1 mℓ 10 mM Tris Cl, 0.2 mM PMSF, 0.2 mM EDTA. Samples were shaken gently (setting 1 on a Vortex Genie) at 4 °C for 30 min in order to release the soluble chromatin. Samples were centrifuged at 5000 × g for 10 min in a benchtop microcentrifuge, after which 500 µℓ aliquots were transferred to clean eppendorf tubes.

SDS and Proteinase K were added to a final concentration of 0.1 % (v/v) and 0.5 % (w/v) respectively. Protein was removed by extracting twice with equal volumes neutralized phenol, and once with an equal volume chloroform:isoamylalcohol (24:1) (v/v). Samples were ethanol precipitated according to standard procedure (Ausubel *et al.*, 1995) and resuspended in 15 µℓ TE buffer [10 mM Tris Cl (pH 7.5), 1 mM EDTA], and analysed by electrophoresis on a 1 % (w/v) agarose gel (as described in section 2.1.5).

2.5.4 Bulk digestion and gel filtration chromatography.

After determining the ideal digestion time (section 2.5.3) to produce long chromatin the bulk digestion was performed. The nuclei solution (at 50 OD/mℓ) was adjusted to 1 mM $CaCl_2$ and 40 U/mℓ MNase(Worthington) added. The digestion proceeded for the determined time (typically 1-2 min) at 37 °C, after which EDTA was added to a final concentration of 2 mM. The sample was returned to ice, divided into two equal aliquots and centrifuged at 4 °C at 5000 × g for 10 min. The supernatant was discarded, and samples resuspended in half the original volume 10 mM Tris Cl (pH 8.0), 0.2 mM EDTA, 0.2 mM PMSF. This solution was gently shaken at 4 °C in a cold room using a Vortex Genie (Scientific Industries, Inc, U.S.A) on setting 1 for 30 min, after which it was centrifuged at 5000 × g for 10 min. The supernatant was collected and the pooled volume was measured accurately using a sterile 25 mℓ plastic pipette. This volume was adjusted to a NaCl of 650 mM by slowly adding 5 M NaCl drop-wise in order to dissociate the linker histones.

Gel-filtration chromatography was used for size-fractionation of chromatin and removal of linker histones. The entire sample was loaded onto a Sepharose 4B (Biorad) column (with a bed volume of approximately 300 mℓ, the resin had been degassed prior to the packing of the column) equilibrated in 10 column volumes of 630 mM NaCl, 10 mM Tris Cl (pH 8), 0.2 mM PMSF and 0.2 mM EDTA. A total of 100 fractions of approximately 8 mℓ were collected under gravity flow.

The protein content of the eluted fractions was determined at 230 nm. Selected fractions were analysed on a 15 % SDS-PAGE gel in order to monitor the elution of both nucleosomal arrays as well as linker histone.

The DNA content of the same fractions was determined by analysis on a 1 % (w/v) agarose gel in 1 × TAE buffer following phenol extraction and precipitating the DNA with ethanol according to a standard technique (Ausubel *et al.*, 1995). Appropriate long nucleosomal fragments free of linker histone were dialyzed overnight in 10 mM Tris Cl, 0.2 mM PMSF, 0.2 mM EDTA with one buffer change, in order to remove the glycerol and NaCl before concentration. Chromatin was concentrated to 7 OD_{260}/mℓ using YM-30 Centricon ultrafiltration spin columns (Millipore). Aliquots of 3 mℓ were concentrated at a time at 15,000 × g for 30 min intervals. This chromatin was then made up to 15 % (v/v) glycerol, and 100 μℓ aliquots were flash-frozen in

eppendorf tubes in a bath made of solid CO_2 and ethanol. Aliquots were then immediately placed at -70 C for storage.

2.5.5 Preparation of the *Lytechinus variegatus* 5S rRNA 207-bp fragment.

The p5S207-12 plasmid (Simpson, 1985) was transformed into the *E. coli* SURE competent cells using the heat-shock method (Inoue *et al.*, 1990) as described (section 2.1.4.1). Following transformation, a single colony was inoculated into 400 mℓ LB medium supplemented with ampicillin (50 µg/mℓ) and grown with shaking at 37 °C for 12 hr. Plasmid DNA was isolated using the Midiprep Purification System following the manufacturer's instructions (Promega). Plasmid DNA was further concentrated via an ethanol precipitation using a standard procedure (Ausubel *et al.*, 1995). The DNA pellet was resuspended in 1 × TE buffer, and the concentration determined spectrophotometrically at 260 nm. Plasmid DNA was digested in a 30 µℓ reaction containing 1 × buffer B (Promega) 1 µg/mℓ of BSA (bovine serum albumin) and 20 U of *Ava I* restriction enzyme (Promega). This reaction was incubated 37 °C for 4 hr.

The digestion products were electrophoresed on a 3.5 % nondenaturing polyacrylamide gel at 80 V for 1 hr in 1 × TBE electrophoresis buffer. Following electrophoresis, the gels were stained in 1 × TBE buffer containing ethidium bromide (0.5 µg/mℓ). DNA was visualised by brief exposure to long wavelength UV light (320 nm). The appropriate band corresponding to the 207-bp fragment was excised from the gel with a sterile blade. Gel slices were transferred to a 1.5 mℓ eppendorf, covered with elution buffer [0.5 M ammonium acetate, 1 mM EDTA, pH 8.0] and incubated at 37 °C for 3 hr. Gel fragments were pelleted at 500 × g for 1 min in a benchtop microcentrifuge. The supernatant was recovered and the remaining polyacrylamide gel fragments were rinsed with additional elution buffer to recover residual DNA. These fragments were re-centrifuged and the supernatants combined.

DNA was precipitated after the addition of 2 volumes of 100 % ethanol at -20 °C for 20 min. DNA was pelleted in a benchtop microcentrifuge for 5 min. After removal of the supernatant the pellet was re-dissolved in 100 µℓ TE buffer (pH 7.4), 10 µℓ volume of 3 M sodium acetate (pH 5.2) was added, and the DNA re-precipitated after addition of 2 volumes 100 % ethanol and recovered by centrifugation. Pellets were

rinsed in 70 % ethanol, dried, and resuspended in 1 × TE (pH 7.4) buffer. DNA concentrations were determined spectrophotometrically at 260 nm.

2.5.6 [α-^{32}P]dCTP and biotin-16ddUTP labeling of the isolated 207-bp fragment.

The isolated 207-bp fragment was labeled at the 5' and 3' ends with [α-^{32}P]dCTP and biotin-16ddUTP, respectively. The *Ava I* restriction enzyme produces dissimilar overhangs following cleavage from the p5S207-12 plasmid (Simpson *et al.,* 1985) allowing labeling of only one the 5' end with [α-^{32}P]dCTP. The fragment was labeled in this manner for use in the subsequent gel electromobility shift assays. The 5' end of 5 µg of the *Ava I* digested fragment was labeled in a 50 µℓ reaction volume which contained 5 µℓ 10 × Klenow buffer (Promega), 5 U of the DNA polymerase I large (Klenow) fragment, 120 mM of dATP, dTTP, dGTP and 50 µCi of [α-^{32}P]dCTP. The labeling was allowed to proceed at 37 °C for 1 hr. The unincorporated nucleotides were removed with a SigmaSpin Post-Reaction Purification Column according to the manufacture's protocol (Sigma). The percentage incorporation of the radioactive nucleotide ranged from 50-70 %, and was determined by liquid scintillation counting, using a Beckman LS 5000 TD Liquid Scintillation Counter, of 2 µℓ aliquots of pre- and post-spin fractions. [α-^{32}P]dCTP-labeled 207-bp fragment was stored at 4 C for no longer than 24-48 hr.

In a separate reaction the 3' end of the *Ava I* digested 207-bp fragment was labeled with biotin-16-2',3'-dideoxy-uridine-5'-triphosphate (Roche Diagnostics GmbH, Germany). A 30 µℓ reaction volume containing 1 × reaction buffer, 10 µg of the digested 207-bp fragment (~ 200 pmol 3' ends), 70 U terminal transferase (Roche), 2.5 mM $CoCl_2$ and 200 pmol biotin-16ddUTP was incubated at 37 °C for 1 hr. The unincorporated biotin was removed using a SigmaSpin Post-Reaction Purification Column according to the manufacture's protocol (Sigma). The percentage incorporation was determined, as described in section 2.6, with a streptavidin-HRP conjugate and chemiluminescent detection.

2.6 Determination of the biotin-16ddUTP incorporation.

A synthetic oligonucleotide was produced, which had a biotin moiety covalently attached to the 5' end of the fragment. The sequence was completely random as follows: 5' **biotin**-CCGGAATGGATAGTCTGCAACGTCTGGAGA- 3'. Biotin used for the production of the oligonucleotide was BiotinTEG Phosphoramidite (Glen Research, Cambio Ltd, United Kingdom).

Biotin incorporation was determined using a modified dot blot protocol. Briefly, 2 μℓ of exponentially increasing amounts of the biotin oligonucleotide standard and biotin 16-ddUTP labeled 207-bp fragment were manually blotted onto a Hybond-N nylon membrane (Amersham Pharmacia Biotech UK Ltd, England, United Kingdom). The membrane was left to air dry. The membrane was blocked with 150 mM NaCl, 100 mM Tris Cl (pH 7.5) containing 5 % (v/v) Liquid Block (Amersham) for 30 min with gentle agitation. Membranes were briefly rinsed in 150 mM NaCl, 100 mM Tris Cl (pH 7.5). The streptavidin-HRP conjugate (Dako, Denmark) was diluted 1:3000 in 400 mM NaCl, 100 mM Tris Cl (pH 7.5) containing 0.5 % (w/v) BSA (Sigma), and the membranes incubated in this reagent for 30 min at room temperature, with gentle agitation. The excess streptavidin-HRP conjugate was rinsed off by washing four times with 400 mM NaCl, 100 mM Tris Cl (pH 7.5) using an excess of the wash solution. Detection solution contained equal volumes of 100 mM Tris Cl (pH 8.5), 90 mM p-coumaric acid, 50 mM hydrogen peroxide and 100 mM Tris Cl (pH 8.0), 250 mM 3-aminophalhydrazin. A volume of 4 mℓ of the mixed detection reagent is required for 4 cm^2 of membrane. The detection solution was added directly to the membrane and incubated at room temperature for 1 min. The signal was visualized using chemiluminescence with a Gene Gnome (Syngene Bio Imaging). Quantification was performed with Quantity One software (Bio-Rad).

2.7 Nucleosome reconstitutions by stepwise salt dilution.

Nucleosome cores were reconstituted onto the [α-^{32}P]dCTP and biotin-16ddUTP-labeled 207-bp fragment by stepwise reduction of the ionic strength (Hayes and Lee, 1997). The isolated chicken erythrocyte H1-stripped long chromatin (section 2.5.2) served as donor histone. An initial 20 μℓ reaction volume contained 150 ng of [α-

^{32}P]dCTP or biotin-16ddUTP-labeled 5S fragment and 12 µg of donor chromatin (17 OD/mℓ) in 1 M NaCl. The salt concentration was reduced in a stepwise fashion by dilution with reconstitution buffer [10 mM Tris Cl (pH 7.5), 0.2 mM EDTA, 0.2 mM PMSF, 0.4 % (v/v) IGEPAL (Sigma), 0.5 mM DTT, 15 µg/mℓ BSA] followed by incubation at 37 C for 30 minute intervals. The NaCl was reduced to a final concentration of 25 mM. Glycerol was added to a concentration of 10 % (v/v) at the final dilution step to allow loading onto analytical gels (section 2.11).

2.8 Quantification of reconstitution efficiency of nucleosome cores on biotin-labeled DNA fragment.

Analysis of the biotin-labeled, reconstituted 5S cores was performed by means of Southern transfer and detection using the streptavidin-HRP conjugate. An amount of 5 pmol of the biotin-labeled 5S fragment, unlabeled 5S fragment and 40 µℓ of the nucleosome cores reconstituted on the biotin-labeled 5S fragment (~ 0.25 pmol biotin-labeled 5S fragments) was electrophoresised on 0.7 % (w/v) agarose gel in 0.5 × TBE electrophoresis buffer at 90 V for 1 hr. Duplicate gels were electrophoresis in this manner.

Pre-transfer DNA content was analyzed through ethidium bromide staining (0.5 µg/mℓ). DNA was visualised by short wavelength UV transillumination (254 nm). The DNA from the duplicate gel was transferred onto a nylon membrane (Hybond N) by electrotransfer at 30 V (100 mA) in 0.5 × TBE buffer using a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad). This transfer was performed overnight at 4 °C with continuous buffer recirculation.

2.9 Reconstitution of GH5, GD1 and GD2 containing chromatosomes.

The reconstitution procedure was identical to that of assembling nucleosomes, with the following modifications: The globular domain of histone H5 (GH5) was added at 50 mM NaCl, followed by incubation of the reconstitution sample on ice for 30 min. Glycerol, which improves the electrophoresis process, was added to a final concentration of 10 % in the last dilution step from 50 to 25 mM NaCl. This step was followed by incubation on ice for 30 min.

2.10 Sucrose density gradient centrifugation.

The basic method for the separation of the reconstituted 5S cores from free labeled 5S fragment and long chromatin was sedimentation through 5-20 % (w/v) linear sucrose gradients. The gradient was formed at a flow-rate of 1 mℓ. h^{-1} using a peristaltic pump (PUMP P-1, Pharmacia Biotech, Sweden) and gradient former in 12 mℓ ultracentrifugation tubes by self-diffusion of equal volumes of 5 % and 26 % sucrose (w/v) solutions in the same buffer used for the reconstitution reactions [10 mM Tris Cl (pH 7.5), 0.2 mM EDTA, 0.2 mM PMSF, 0.4 % (w/v) IGEPAL (Sigma), 0.5 mM DTT, 15 µg/mℓ BSA]. A volume of 2 mℓ of the reconstitution reactions were loaded onto the top of 10 mℓ 5-20 % sucrose gradients. The gradients were centrifuged at 67 200 × g in SW 40.1 rotor (Beckman) for 16 hr at 4 °C. Fractions (1 mℓ) were collected from the ultracentrifugation tubes using displacement with 50 % sucrose.

Fractions from the [α-^{32}P]dCTP-labeled 5S cores were analyzed by scintillation counting and fractions from the biotin-16-ddUTP-labeled 5S cores were analyzed with dot blotting followed by detection with streptavidin-HRP (section 2.6).

Fractions were pooled and dialyzed overnight in 25 mM NaCl, 1mM Tris Cl (pH 7.5), 0.1 mM PMSF, 0.1 mM EDTA with one buffer change to remove the sucrose. Labeled, reconstituted mononucleosomes were concentrated to 10-20 OD_{260}/mℓ using YM-30 Centricon ultrafiltration spin columns (Millipore).

2.11 Gel electrophoresis of reconstitution samples.

The reconstituted 5S cores were analyzed by electrophoresis on a 0.7 % (w/v) agarose gel in 0.5 × TBE at 90 V for 1 hr, and by electrophoresis on 8 % native polyacrylamide [40 % (w/v), 60:1 bis:bisacrylamide] in 0.5 × TBE electrophoresis buffer at 200 V for 3 hr, at 4 °C with continuous buffer recirculation.. Agarose and polyacrylamide gels were dried under vacuum at 75 C for 1 hr and then exposed overnight to a storage phosphorimager screen. Screens were scanned using a phosphorimager (Biorad Molecular Imager Personal FX).

2.12 Magnetic bead technology.

A volume 10-100 μℓ (0.1-1 mg) streptavidin magnetic beads (Boehringer Mannheim, GmBh), washed three times in 10 mM Tris Cl (pH 7.5), 1 mM EDTA, 100 mM NaCl binding buffer (TEN_{100}), was incubated with 1-20 pmol biotin-labeled mononucleosomes in the same buffer, for 30 min at 4 °C. Beads were prevented from settling at the bottom of the 1.5 mℓ eppendorf tube by occasional gentle resuspension. After the incubation, particles were washed two times with 10 mM Tris Cl (pH 7.5), 1 mM EDTA, 1 M NaCl (TEN_{1000}). The magnetic bead separations were performed by applying the magnetic particle separator to the bottom of the tube for 5 min. The immobilized biotin-labeled nucleosomes were eluted with 6 M GuHCl. Aliquots were taken from the supernatant at the various steps and assayed using liquid scintillation counting.

Chapter 3: Cloning and expression of globular domain 1, globular domain 2 and two modified constructs of the yeast linker histone Hho1p.

3.1 Introduction.

The presence of a linker histone in yeast was first suggested with the completion of the sequencing of the yeast genome (Goffeau *et al.*, 1996; Bussey *et al.*, 1997). The endogenous protein has yet to be purified, but the predicted structure of this candidate linker histone, when compared to that of other linker histones, indicates the presence of two globular domains (Landsman, 1996; Patterton *et al.*, 1998). The functionality of these two globular domains, and indeed the yeast linker histone Hho1p, must still be elucidated.

In order to perform functional studies on the yeast linker histone, and investigate the association of the two globular domains with nucleosomal DNA, both the separate globular domains as well as two modified proteins had to be expressed and purified. To produce sufficient quantities for biochemical analyses, recombinant proteins were over-expressed in *E. coli*. The DNA fragments encoding the separate globular domains of the yeast linker histone Hho1p were amplified using a construct that contained the coding region of Hho1, and were cloned into a T7 expression vector.

Two constructs were designed to express proteins that were used as control peptides in the binding study of yeast Hho1p to nucleosomes. The first of these control proteins contained the yeast Hho1p lysine-rich region with a centrally located protease site. The second protein contained two chicken histone H5 globular domains (GH5) connected by the lysine-rich region of the yeast linker histone. The DNA encoding these modified proteins were cloned into T7 expression vectors.

Constructs were sequenced to verify the correct sequence and to confirm that the cloned regions were in-frame with a C-terminal hexa histidine-tag, which is required for the purification of the recombinant proteins. After expression in *E. coli*, the proteins were purified using immobilized metal affinity chromatography (IMAC). Further purifications were performed using gel filtration chromatography, ion

exchange chromatography and high performance liquid chromatography (HPLC). The identity of all proteins was verified by MALDI-TOF analysis.

3.2 Results

3.2.1 Construction of expression vectors for yeast Hho1p globular domain one (GD1) and two (GD2).

The Hho1p globular domains (GD1 and GD2) were used to test the ability of the individual domains to associate with nucleosomal DNA. In order to sub-clone and express the two globular domains of Hho1p in *E. coli*, the corresponding coding regions for the two separate proteins (Figure 3.1) were amplified by PCR using primer pairs that introduce *Sfo I* and *Xho I* restriction enzyme sites, which would allow the subsequent ligation into the expression vector digested with the same set of enzymes.
It was previously shown (Gerchman *et al.*, 1994) that the globular domains of chicken histones H5 and H1 could not be efficiently over-expressed in *E. coli*. However, addition of the amino acid leader motif methionine-threonine-glutamate (MTE) at the N-terminus of the H5 globular domain was sufficient for effective over-expression. Due to the high degree of sequence homology between the yeast globular domains and GH5, we decided to investigate the effect of this MTE leader peptide on the expression of the globular domains. Primer pairs (Table 2.1) were designed to amplify GD1 and GD2 of yeast *HHO1* by PCR, introducing the nucleotide sequences encoding the MTE leader peptide motif. To eliminate the expression of a region of the pPRO Ex A vector, it was also necessary to incorporate a stop codon (TGA) in the reverse primers of both domains.
The amplification produced the desired 255 bp globular domain 1 and 274 bp globular domain 2 (Figure 3.2). The fragments were isolated from a low melting agarose gel and digested with the appropriate enzymes. The digested fragments were isolated in the same manner as the PCR products (Figure 3.3 A) and subsequently ligated into the pPRO EX A vector digested with the same restriction enzyme set as those used on the amplified GD1 and GD2 (see Figure 3.3 B for vector construction scheme).
pPRO EX A was used due to the capability of removing the C-terminal hexa histidine-tag from the protein expressed from this vector. This gave us the option of

having the globular domains with and without the hexa histidine-tag, and thus the opportunity to investigate if this hexa amino acid motif had any effect on the binding by the globular domains to nucleosomal DNA.

The constructs were transformed into DH5α *E. coli* cells, and colonies screened by restriction enzyme digestion. The sequences of inserts in positive constructs were subsequently verified by automated nucleotide sequencing. The constructs are referred to as pPRO EX A-GD1 and pPRO EX A-GD2.

3.2.2 Construction of expression vectors for Hho1-FXa and GH5-linker-GH5.

3.2.2.1 Construction of expression vector for Hho1-FXa.

To test the hypothesis that a single yeast linker histone can bind to two nucleosomes via each of the two proposed globular domains, Hho1p and two modified versions of this protein were expressed. One of these modified proteins, Hho1-FXa, was a copy of the original protein (Hho1p) with a protease site in the lysine-rich region between the two globular domains (see Figure 3.1). This protease site will allow the cleavage of the full-length protein into the separate globular domains after binding to the nucleosomes. In order to introduce the Factor Xa protease site in the lysine-rich region, a sequence encoding the four amino acids (isoleucine, glutamate, glycine and arginine) was incorporated into the original template by mutagenesis via overlap PCR extension (template–mismatched PCR mutagenesis), a technique previously used for the introduction of new sequences into native templates (Higuchi *et al.,* 1988; Ho *et al.,* 1989). To allow optimal expression of the protein containing the incorporated residues in *E. coli*, highly utilized codon sequences were chosen from the *E. coli* codon preference tables, obtained from DNAssist (Patterton *et al.,* 2000).

20 40 60 80
MAPKKSTTKTTSKGKKPATSKGKEKSTSKAAIKKTTAKKEEASSKSYRELIIEGLTALKERKSSRPALKKFIKENYPIVGSASNF
100 120 140 154 160
DLYFNNAIKKGVEAGDFEQPKGPAGAVKLAKKKSPEVKKEKEVSPKPKQAATSVSATASKAKAASTKLAPKKVVKKKSPTVT
180 200 220 240
AKKASSPSSLTYKEMILKSMPQLNDGKGSSRIVLKKYVKDTFSSKLKTSSNFDYLFNSAIKKCVENGELVQPKGPAGIIKLNKKK
260
VKLSTLE

Figure 3.1 Hho1p sequence showing the coding regions for globular domain 1 and for globular domain 2. Globular domain 1 spanning residues 38-119, and globular domain 2 spanning residues 173-256. The lysine-rich region spans amino acids 120-172. The position (154) of the inserted FXa cleavage site is indicated by the arrow.

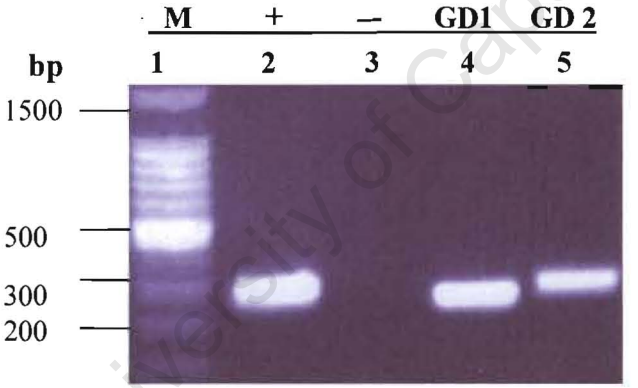


Figure 3.2 Products of PCR amplification of globular domain 1 and globular domain 2. 1 % (w/v) agarose gel in 1 × TAE containing ethidium bromide and visualized using short wavelength UV transillumination. Lane **1**, 100 bp DNA molecular weight marker.; lane **2**, positive control. 288 bp GH5 PCR product the using primer set globF and globR (Koorsen, 2001 MSc thesis, UCT); lane **3**, negative control, lacking template, with oligonucleotides GD1F and GD1R (see Table 2.1); lane **4**, 255 bp GD1 PCR product amplified with oligonucleotides GD1F and GD1R; lane **5**, 274 bp GD2 PCR product amplified with oligonucleotides GD2F and GD2R.

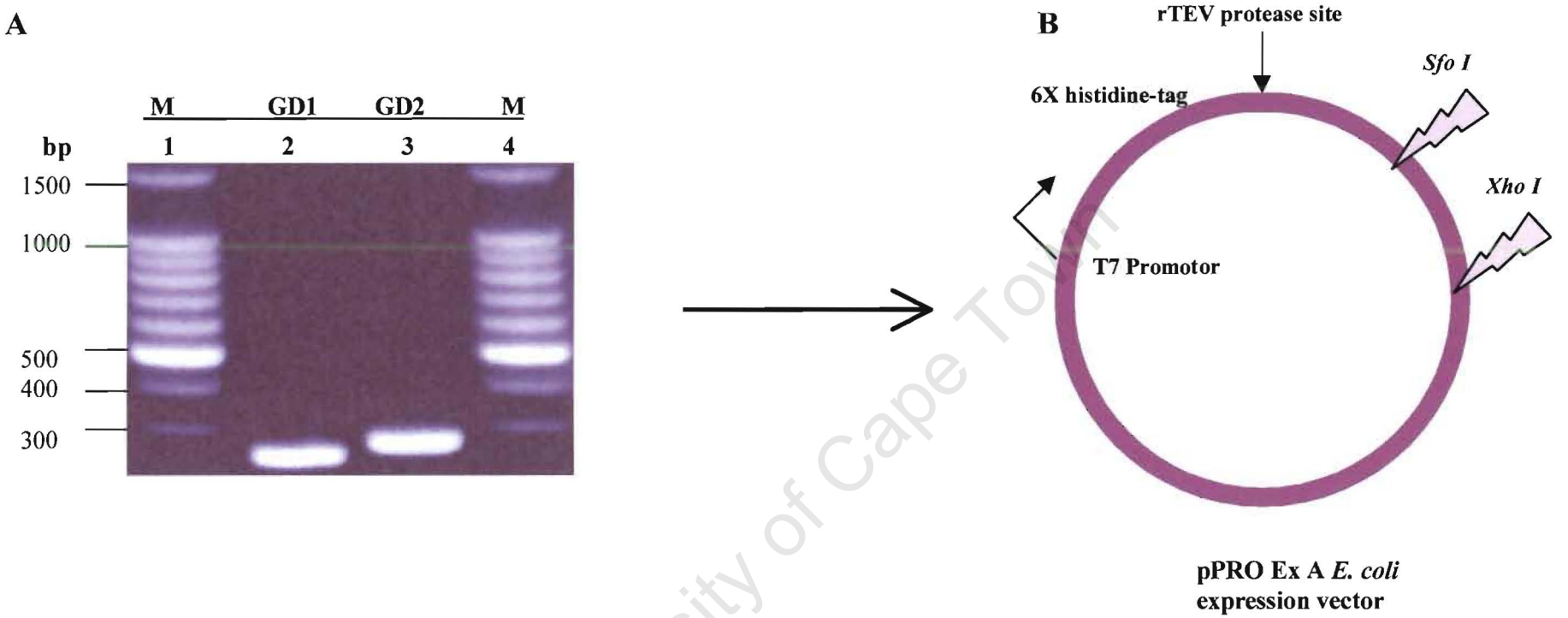
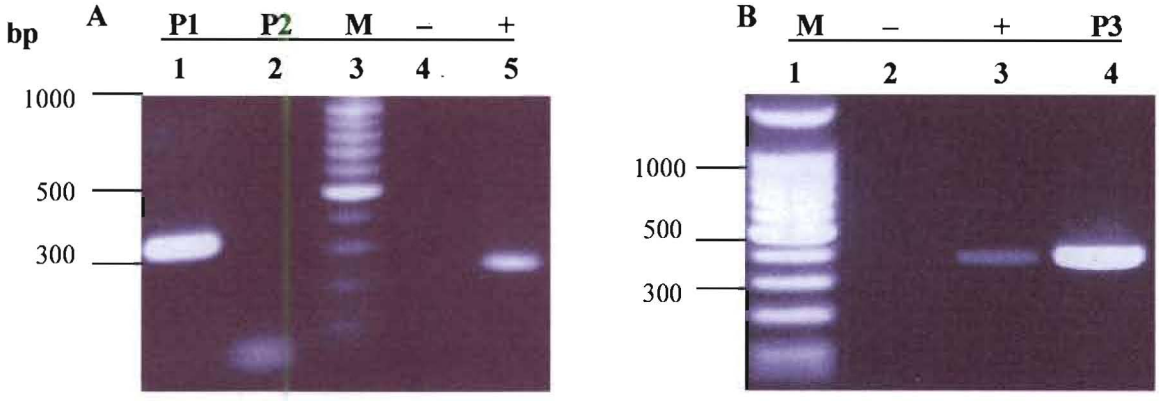


Figure 3.3 Restriction fragments of globular domain 1 and globular domain 2 to be ligated into the expression vector pPRO Ex A.
Panel A: isolated PCR fragments and **Panel B:** expression vector map. In Panel A, 1 % (w/v) agarose gel in 1 × TAE containing ethidium bromide was visualized using short wavelength transillumination. Lanes **1** and **4**, 100 bp DNA molecular weight marker; lane **2**, GD1 PCR product digested with *Sfo I* and *Xho I* and isolated from a 1 % (w/v) low melting agarose gel; lane **3**, GD2 PCR product digested with *Sfo I* and *Xho I* and isolated from a 1 % (w/v) low melting agarose gel.

The first step was the separate amplification of two fragments of an overlapping region of the coding region for Hho1p (residues 453 to 471 of full-length Hho1p). The template used in this first amplification was a *HHO1* plasmid template that contained the full-length *HHO1* gene in pET-20b(+). An oligonucleotide containing the coding sequences for isoleucine, glutamate, glycine and arginine together with an oligonucleotide complementary to either of the Hho1p flanking regions was employed in each of the pair of PCR reactions to produce overlapping segments. Flanking oligonucleotides were designed in such a way that restriction sites present in the Hho1p coding sequence were utilised instead of incorporating new restriction sites for the subsequent ligation. The amplification (Figure 3.4 A) produced the desired 310 bp P1 (product from the amplification using oligonucleotide Hho1 A and Hho1 B, refer to Table 2.2) and 60 bp P2 (product using oligonucleotides Hho1 C and Hho1 D, Table 2.2). The PCR fragments were isolated from a low melting agarose gel. The products from the first round of PCR shared a segment of identical sequence which served as templates in the second round of PCR for the amplification of the entire region (residues 148-581 of Hho1p) by overlap extension PCR, using the flanking oligonucleotides (Hho1 A and Hho1 D). To optimize this reaction and minimize non-specific products, the two fragments were subjected to two PCR cycles without the flanking oligonucleotides, which were then added subsequently. The amplification produced a 369 bp P3 fragment (Figure 3.4 B). This fragment was isolated from a low melting agarose gel and the incorporation of the protease site coding region was verified by automated nucleotide sequencing.

The fragment (P3) was digested with the *Bcl I* and *Stu I* and the digested fragments were isolated in the same manner as the PCR products. The pET-20b(+)-*HHO1* vector was digested with the same restriction enzymes as those used on the amplified P3 fragment, and isolated from a low melting agarose gel to separate the digested vector from the 369 bp wild-type fragment. The digested vector and 369 bp P3 mutated fragment were ligated and the resulting expression plasmids containing the coding regions for the protease site were transformed into the DH5α *E. coli* strain for plasmid amplification and isolation. The presence of the correct insert in the final vector was verified by automated nucleotide sequencing of the coding strand (Figure 3.5). The plasmid from the positive clone, pET-20b(+)-*HHO1*-FXa, was transformed into *E. coli* BL21(DE3) cells containing the pLysS plasmid, and used in subsequent protein expression.



3.4 Products of template-mismatched PCR mutagenesis.
Amplification of P1, P2 and P3 fragments used in the construction of the expression vector pET-20b(+)-*HHO1*-FXa. **Panel A** and **B**, 1 % (w/v) agarose gel in 1 × TAE containing ethidium bromide and visualized using short wavelength UV transillumination. Lane **M**, in **A** and **B:** 100 bp DNA molecular weight marker (Promega). Panel **A:** PCR products from the first round of amplification to construct the overlapping fragment with the peptidase site incorporated in the wild-type sequence. Lane **1**, 310 bp PCR product (P1) using primer set Hho1 A (flanking primer) and Hho1 B (template mismatch reverse primer); lane **2**, 60 bp PCR product (P2) using primer set Hho1 C (template mismatch forward primer) and Hho1 D (flanking primer); lane **3**, 100 bp DNA molecular weight marker; lane **4**, negative control with primers globF and globR; lane **5**, positive control. 288 bp GH5 PCR product using the primer set globF and globR. Panel **B:** PCR product form the second round of PCR using fragments from the first round of PCR as templates. Lane **1**, 100 bp DNA molecular weight marker; lane **2**, negative control containing no template with primers globF and globR; lane **3**, positive control. 288 bp GH5 PCR product using the primer set globF and globR; lane **4**, 369 bp P3 PCR product amplified with primers Hho1 A and Hho1 D. Refer to Table 2.2 for primer sequences.

Wild-type Hho1p: 5'-CTT-TGG-CGC-TAG-CTT-3'

Lys Pro Ala Arg Gly Glu Ile Leu Lys
Hho1-FXap: 3'-GAA-ACC-GCG-TCG-TGG-AAG-CTA-ATC-GAA-5'
5'-CTT-TGG-CGC-ACG-ACC-TTC-GAT-TAG-CTT-3'

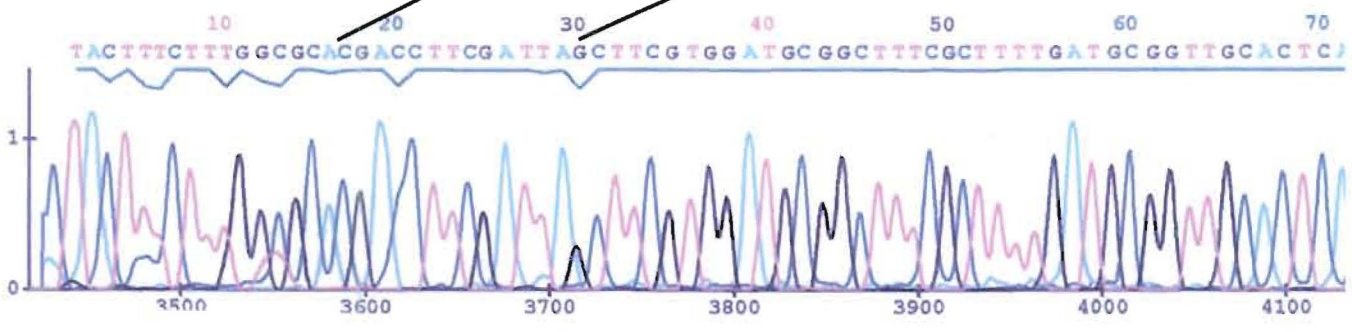


Figure 3.5 Electropherogram of pET-20b(+)-*HHO1*-FXa indicating Factor Xa peptidase site. The coding strand of pET-20b(+)-*HHO1*-FXa was sequenced with the forward primer used in the PCR mutagenesis of the original *HHO1* template (wild-type *HHO1*). The mutated sequence (mutated *HHO1*-FXa) consisting of the incorporated amino acid (underlined) coding regions, resulting from this mutagenesis, are indicated.

3.2.2.2 Construction of expression vector for GH5-linker-GH5.

The modified construct that would serve as a positive experimental control in the magnetic bead binding assays, was designed to contain two chicken linker histone globular domains (GH5) flanking the lysine-rich region of the yeast linker histone (Hho1p). This protein was designed to test the efficacy of the assay, testing the ability of peptides known to bind to nucleosome cores to bind to two cores simultaneously, and be detected to bind in this manner by the assay.

The strategy for the construction of expression vector was to amplify firstly the lysine-rich region using the isolated *HHO1* coding region as template, and to amplify separately GH5, using the GH5 coding region expression vector, pET-20b(+)-GH5. The isolated GH5 and *HHO1* coding regions were utilized as templates in the amplification to diminish non-specific amplification products. GH5 was amplified using primer pairs (GH5-mag F and GH5-magR, see Table 2.3) which introduced *Xho I* restriction enzyme sites on both the 5' and 3' ends of the amplified fragment. The *HHO1* lysine-rich region was amplified with a primer set (link-mag F and link-mag R, see Table 2.3) that introduced *Sal I* and *XhoI* sites on the 5' and 3' ends, respectively. The amplification produced the desired 152 bp *HHO1*-linker fragment and the 288 bp GH5 fragment (Figure 3.6). The amplified fragments were isolated from a low melting agarose gel, digested with the appropriate enzymes and ligated at a 1:5 GH5 to *HHO1*-linker molar ratio. Ligation of the *Xho I* and *Sal I* sites produced a *Taq I* site which could be used to identify positive clones. The ligation product was isolated from a low melting agarose gel and digested with *Taq I*. From this restriction enzyme screening, fragments of 158 and 265 base pairs (shown in Figure 3.7) were expected from the ligation product which contained the GH5 and the *HHO1*-linker region ligated in the correct orientation. Several attempts to ligate this fragment into the *Xho I* digested pET-20b(+)-GH5 proved unsuccessful. Consequently, the next step was to construct the GH5-linker-GH5 fragment, which involved using GH5-*HHO1*-linker and GH5 as templates in separate PCR amplifications. The GH5-*HHO1*-linker fragment was re-amplified using oligonucleotides (GH5-link *Nde I* and GH5-link *Xho I*, see Table 2.4) which introduced the *Nde I* and *Xho I* restriction sites. GH5 was re-amplified with oligonucleotides which incorporated *Sal I* and *Xho I* restriction sites (GH5-*Xho I* and GH5-*Sal I*, Table 2.4). The *Sal I* site was introduced (using the oligonucleotide, GH5-*Sal I*) to be able to isolate GH5-linker-GH5 fragments after

ligation, by *Xho I* cleavage. This step cuts unwanted ligation products. The desired 288 bp and 420 bp amplification products are shown in Figure 3.8 A, representing the amplified GH5 and GH5-*HHO1*-linker fragments (Figure 3.8 B). Sequences were verified using automated nucleotide sequencing after which the fragments were ligated at a 1:1 molar ratio. Following ligation of the GH5 and GH5-*HHO1*-linker fragments, the sample was digested with *Xho I*, and the fragment isolated from low melting agarose gel. The sequence of this GH5-linker-GH5 fragment was verified by automated nucleotide sequencing. This fragment will be ligated into a vector digested with the appropriate enzymes and transformed into a recombination deficient *E. coli* strain.

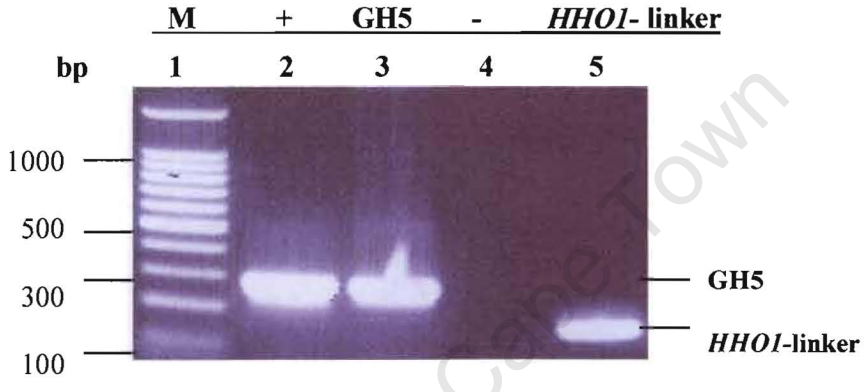


Figure 3.6 Products of PCR amplification of GH5 and *HHO1*-linker. A 1 % (w/v) agarose gel in 1 × TAE containing ethidium bromide and visualized using short wavelength UV transillumination is shown. Lane **1**, 100 bp DNA molecular weight marker; lane **2**, positive control. 288 bp GH5 PCR product from the primer set globF and globR; lane **3**, 288 bp GH5 PCR product amplified with oligonucleotides GH5-mag F and GH5-mag R; lane **4**, negative control with oligonucleotides globF and globR; lane **5**, 152 bp *HHO1*-linker PCR product amplified with oligonucleotides link-mag F and link-mag R (Table 2.3). The positions of GH5 and *HHO1*- linker fragments are indicated.

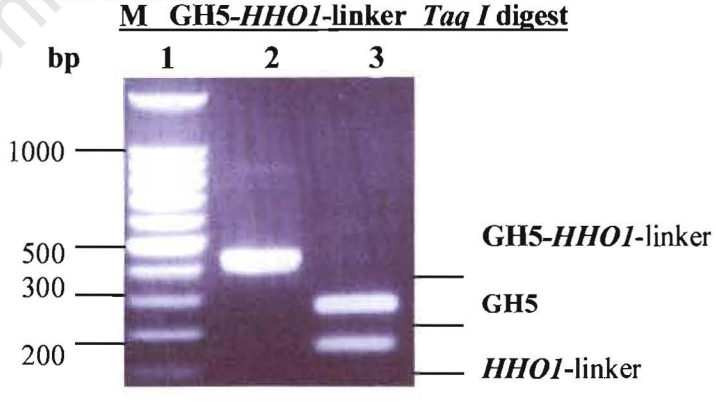


Figure 3.7 Products of restriction enzyme screening of GH5-*HHO1*-linker ligation product. 1 % (w/v) agarose gel in 1 × TAE containing ethidium bromide and visualized using short wavelength UV transillumination. Lane **1**, 100 bp DNA molecular weight marker; lane **2**, 423 bp isolated GH5-linker ligation product; lane **3**, *Taq I*-digested 265 bp and 158 bp GH5-linker fragments. The positions of GH5 GH5-*HHO1*-linker and *HHO1*-linker are indicated.

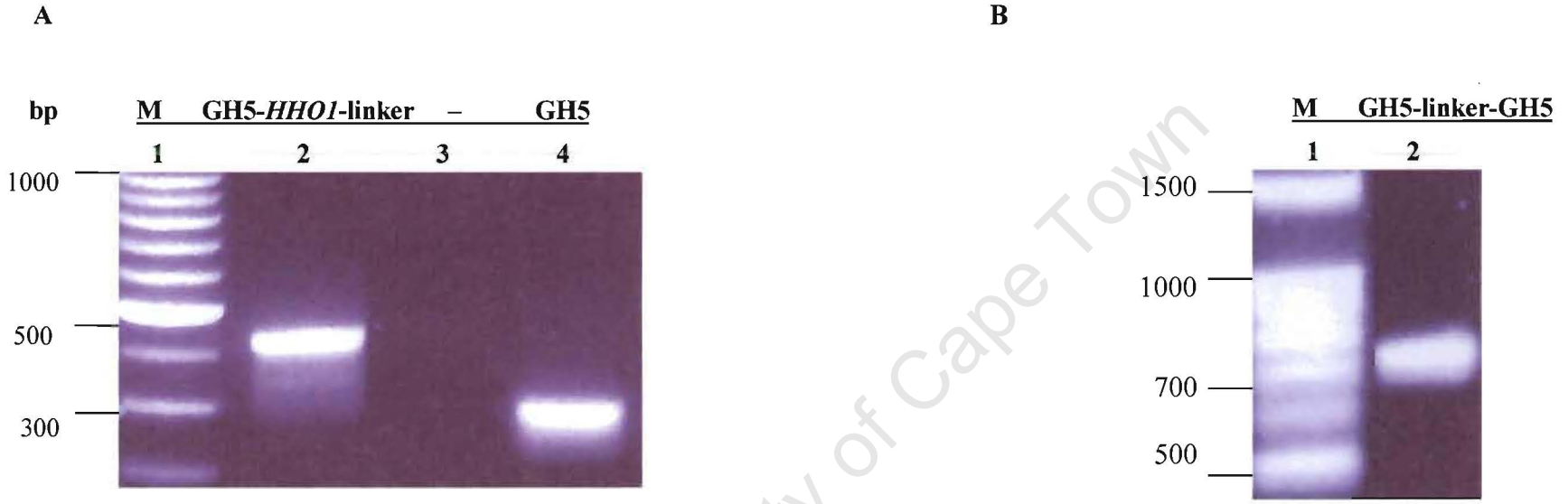


Figure 3.8 Construction of GH5-linker-GH5. Panel A, Products of PCR amplification of GH5-*HHO1*-linker and GH5**. Panel B,** Digested and isolated GH5-linker-GH5**.** 1 % (w/v) agarose gels in 1 × TAE containing ethidium bromide and visualized using short wavelength UV transillumination. **Panel A:** lane **1**, 100 bp DNA molecular weight marker; lane **2**, 422 bp GH5-*HHO1*-linker PCR product amplified with GH5-link *Nde I* and GH5-link *Xho I*, (Table 2.4); lane **3**, negative control containing no template; lane **4**, 288 bp GH5 PCR product amplified with oligonucleotides GH5-*Sal I* and GH5-*Xho I*. **Panel B:** Lane **1**, DNA molecular weight marker; lane **2**, Isolated 683 bp GH5-linker-GH5 fragment.

3.2.3 Expression of GD1, GD2, Hho1p and Hho1-FXap in *E. coli.*

The pPRO EX A bacterial expression system was utilized for the expression of the yeast Hho1p globular domains in *E. coli.* The full-length native linker histone and the modified linker histone, containing the peptidase recognition site, was expressed using the pET-20b(+) expression vector in *E. coli*. Expression using both the pPRO Ex A and pET-20b(+) vectors was driven by an IPTG-inducible T7 promotor. In this way, expression can be stringently regulated in the presence of T7 lysozyme, which is expressed from the pLysS plasmid. T7 lysozyme inhibits basal levels of T7 polymerase present prior to induction, and which can lead to counter-selection of plasmids from which toxic peptides are expressed (Baneyx, 1999). IPTG is a synthetic analog of lactose, which is a natural inducer of expression from the lac operon. IPTG is used for induction of protein expression since it is not hydrolysed by β-galactosidase, and therefore the levels of the inducer remain constant throughout the expression period.

The optimal induction time to obtain the maximal yield for all recombinant proteins was established by an induction time-course. *E. coli* cultures containing pPRO EX A-GD1, pPRO EX A-GD2, pET-20b(+)-*HHO1* and pET20b(+)-*HHO1*-FXa expression plasmids, respectively, were induced. Aliquots were withdrawn at half-hour or hour intervals, and the total cellular protein analysed by SDS-PAGE electrophoresis.

As shown in Figure 3.9 A and B expression of both GD1 and GD2 was induced 30 mins (Lane 3) after the addition of IPTG. A two-hour induction period produced a maximal amount of both recombinant GD1 and GD2 in *E. coli*, which was visible in lane 6 in both panels A and B (labeled as 120 mins). Figure 3.10 A and B indicates that both Hho1p and Hho1-Fxap were induced upon IPTG induction. As with GD1 and GD2, the maximal amount of both recombinant proteins was obtained after a 1-2 hr induction period.

All expressed peptides exhibited anomalous migration on 15 % (w/v) SDS-PAGE gels. GD1, with a predicted size of 11.8 kDa (obtained from the peptide calculator at ***http://prowl.rockefellar.edu***), migrated at a region corresponding to approximately 14 kDa, and GD2 with an expected size of 12.5 kDa, migrated at approximately 16 kDa. Although the size of Hho1p is 28 kDa, the isolated histidine-tagged protein migrated at a size corresponding to approximately 33 kDa. This was also the case for Hho1-

Fxap which has a predicted weight of 29.3 kDa and migrated at approximately 33 kDa. Linker histone migration on SDS-PAGE is usually atypical, and this behaviour was shown in linker histone migration of other species (Cole *et al.,* 1989). This retardation on SDS-PAGE is attributed to the highly charged N- and-C terminal tails and basic nature of the proteins. GD1 and GD2 are both highly hydrophobic but lack the charged tail regions. Both recombinant proteins contained a charged N-terminal hexa histidine-tag, which may have some influence on the anomalous migration on SDS-PAGE gels.

3.2.3.1 Effect of the N-terminal amino acids on the efficiency of expression.

As mentioned previously, the GD1 and GD2 coding regions were constructed in such a way as to introduce an MTE leader peptide motif at the N-terminal of the expressed proteins to facilitate the investigation of the effect that this motif would have on the expression of GD1 and GD2. However, this leader peptide had no effect on the expression of either of the proteins as was indicated by the similar degrees of expression of both GD1 and GD2 in an independent induction study using pET-20b(+) vectors containing GD1 and GD2 with and without the MTE leader peptide motif.

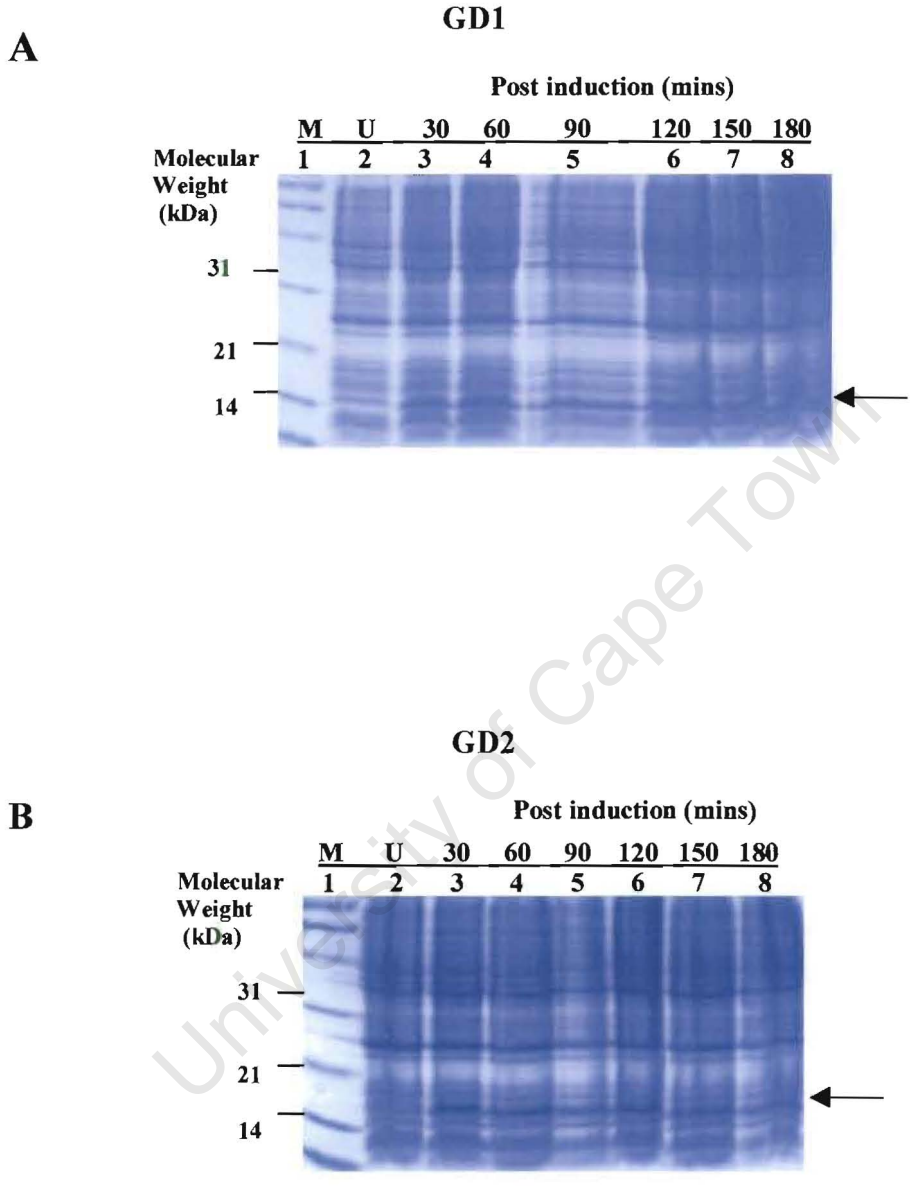
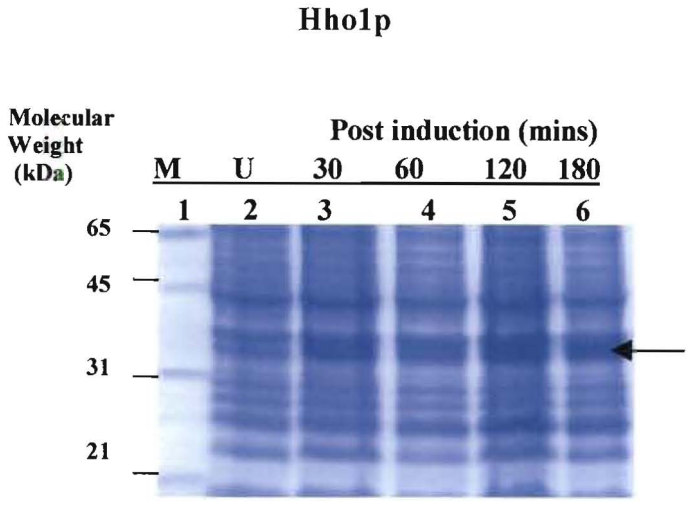


Figure 3.9 Induction of recombinant protein in *E. coli* strain BL21(DE3)pLysS
SDS-PAGE (15 %) gels were visualized by Coomassie staining. In **A** and **B**; lane **1**, protein molecular weight marker, with sizes indicated at the left. All other lanes contain total *E .coli* protein where cells were lysed prior to loading by addition of sample application buffer. Accumulation of GD1 and GD2 are shown.

A



B

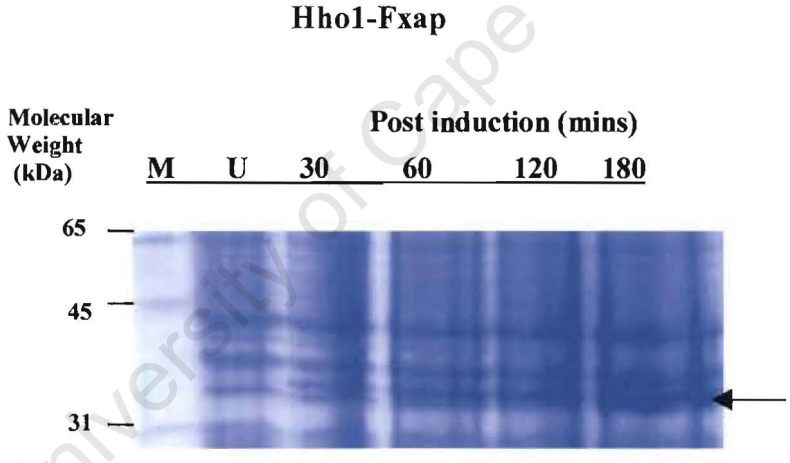


Figure 3.10 Protein induction in *E. coli* strain BL21(DE3)pLysS. Panel A; induction of Hho1p and Panel B induction of Hho1-Fxap.
SDS-PAGE (15 %) gels, protein was visualized by Coomassie staining. In **A** and **B**; lane **1**, Protein molecular weight marker, with sizes indicated on the left. All other lanes contain total *E. coli* protein where cells were lysed prior to loading by addition of sample application buffer. Accumulation of Hho1p and Hho1-Fxap are indicated.

3.2.4 Protein purification by immobilized metal affinity chromatography.

For the large-scale purification, 2 ℓ cultures for the expression of GD2, and 1 ℓ cultures for the expression of GD1, Hho1p and Hho1-Fxap, respectively, were grown to an cell density OD_{600} of 0.6. After a 2 hr induction period, the cells were harvested and lysed. The first purification step for histidine-tagged recombinant GD1, GD2, Hho1p and Hho1-Fxap was by immobilized metal affinity chromatography (IMAC). The resin used was nickel-agarose charged with nickel sulphate (Ni^{2+}). The 6 × histidine-tag on the N-terminus of the peptides binds to Ni^{2+} cations. Unbound proteins are washed away, and the target peptides are recovered by elution with imidazole, which displaces the 6 × histidine-tagged peptides from the Ni^{2+} ions (Sulkowski, 1989). The recommended elution conditions were altered by decreasing the imidazole concentration in the wash buffer from 60 mM to 40 mM to prevent premature elution of the histidine-tagged peptides. All recombinant proteins were eluted from the Ni^{2+} charged column with 1 M imidazole.

Separations of GD1 and GD2 are shown in Figure 3.11. The Coomassie-stained SDS-PAGE gels indicate that GD1 was predominately recovered in elution fractions 3-6 (lanes 8-11) and GD2 was predominantly recovered in elution fractions 4-6 (lanes 9-11). The final yield of GD1 and GD2 differed significantly, even though the culture volume used for GD2 was larger. This effect could be attributed to some instability of the expressed GD2 (Ali, 2000, unpublished data). Under the conditions of the purification, GD1 is most likely to be in a structured state whereas GD2 might be in an unfolded state and thus subject to proteolysis. Purified Hho1p and Hho1-Fxap displayed similar elution profiles (Figure 3.12), and both proteins were eluted predominantly at 1M imidazole (fractions 2-5, lanes 7-10).

In all elution fractions additional contaminating proteins were also eluted, evident by the other bands present. MALDI-TOF mass spectrometry was used to confirm that the proteins, initially identified after induction by SDS-PAGE analysis, were indeed GD1, GD2, Hho1p and Hho1-Fxap, as described below.

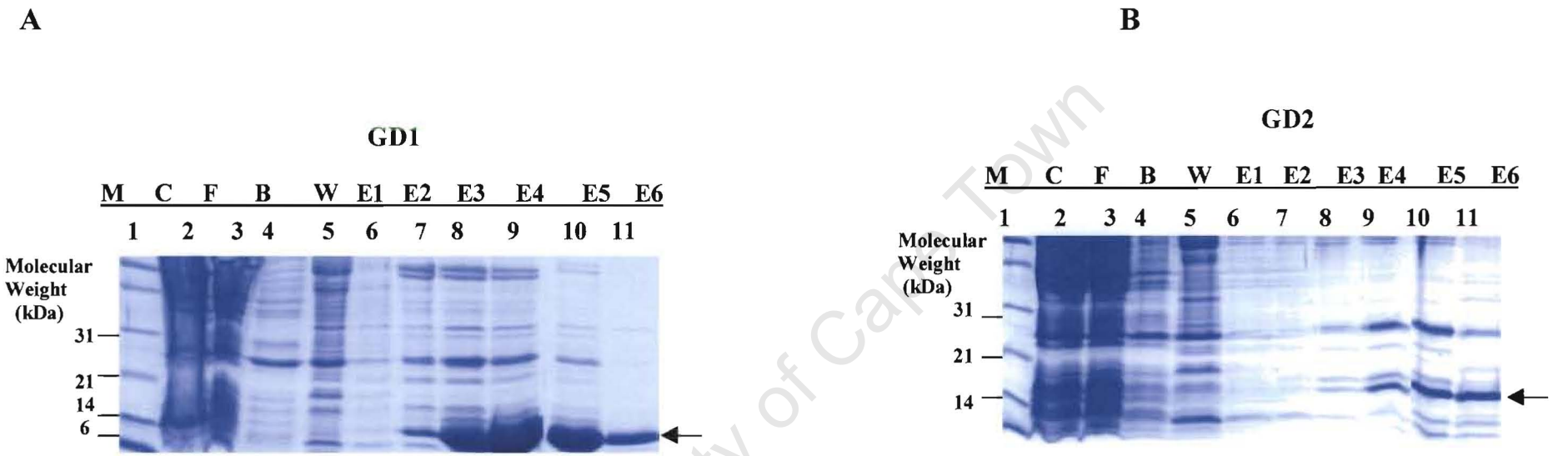


Figure 3.11 Purification of recombinant GD1 and GD2 using Ni-agarose chromatography.

Panel A, Elution of globular domain 1 and **Panel B,** elution of globular domain 2. Aliquots (20 μℓ) of fractions eluted during Ni-chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein visualised by Coomassie stain. In **A** and **B:** lane **1**, protein molecular weight marker, with sizes indicated at the left; lane **2**, crude protein extract loaded onto the column; lane **3**, binding flow-through; lane **4**, binding buffer eluate; lane **5**, wash buffer eluate; lanes **6-11**, eluted fractions. The positions of GD1 and GD2 are indicated.

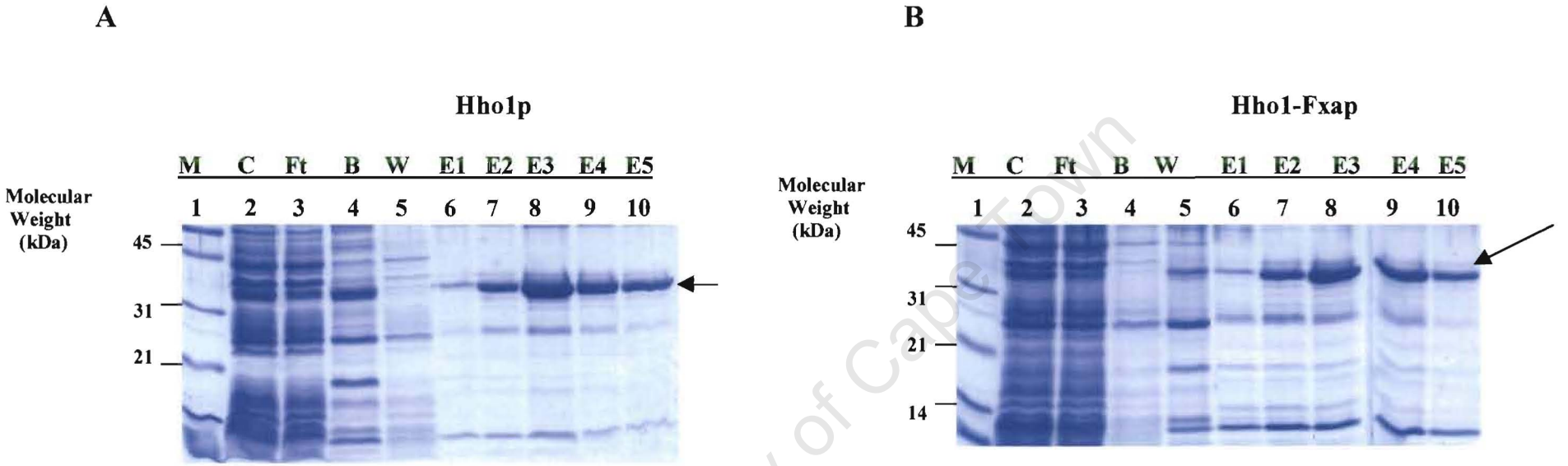


Figure 3.12 Purification of recombinant Hho1p and Hho1-Fxap using Ni-agarose chromatography.
Panel A, Elution of Hho1p and **Panel B**, elution of Hho1-Fxap. Aliquots (20 $\mu\ell$) of fractions eluted during Ni-chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. In **A** and **B: lane 1**, protein molecular weight marker, with sizes indicated at the left; lane **2**, crude protein extract; lane **3**, binding flow-through; lane **4**, binding buffer eluate; lane **5**, wash buffer eluate; lanes **6-10**, eluted fractions. The positions of Hho1p and Hho1-Fxap are indicated.

3.2.4.1 In-gel trypsin digest and MALDI-TOF analyses of GD1, GD2, Hho1p and Hho1-Fxap.

In-gel trypsin digests and subsequent MALDI-TOF analyses were used to confirm the identity of all recombinant proteins. The purified proteins were electrophoresed on SDS-PAGE gels and protein bands corresponding to GD1, GD2, Hho1p, Hho1-Fxap and carbonic anhydrase, used as a positive control for the experiment, were excised from the gel, and treated with the proteolytic enzyme, trypsin, which catalyses the cleavage of polypeptide chains next to lysine and arginine residues. Each protein has a specific pattern of fragment sizes generated by limited tryptic digestion. The resulting trypsin-digested mixtures were analysed using MALDI-TOF mass spectrometry. The digested fragments were applied to a matrix, α-cyano-4-hydroxycinnamic acid (α-CHCA). This matrix produces intense signals for peptides and proteins (Beavis *et al.*, 1992). The co-precipitate of this UV-light absorbing matrix and the digested sample is irradiated with a nanosecond pulse laser. The ionised molecules are accelerated in an electric field and enter a flight tube. Molecules are separated depending on the acceleration which is a function of the mass to charge (m/z) ratio.

Carbonic anhydrase was used as a positive control. Table 3.1 and Figure 3.13 show the expected theoretical and the experimental fragments obtained for this protein. These fragments were used as control in subsequent database searches (Clauser *et al.*, 1999), using the MS-Fit Program from the UCSF Mass Spectrometry facility (available at http://prospector.ucsf.edu). Most of the mass peaks detected for GD1, GD2, Hho1p and Hho1-Fxap, could be assigned to peptides which were generated by theoretical tryptic digests of these proteins. PeptIdent, a peptide mass fingerprinting tool from Expasy, was used to identify all the recombinant proteins. This program uses the fragments sizes obtained by MALDI-TOF and compares this to existing proteins sequences in the Swiss-Prot database. In this way, the mass spectrometry data can be assigned to the protein of highest sequence similarity. The database search with the digested mass spectrum of recombinant GD1 (Figure 3.14) identified the trypsin-digested peptide as part of Hho1p from *Saccharomyces cerevisiae.* Of the 9 experimental peptide sizes entered, 6 were matched to Hho1p. 11 peptide sizes from the 12 entered for GD2 (Figure 3.15) were matched to that of Hho1p. A database search with the digested mass spectrum of Hho1p and Hho1-Fxap (Figures 3.16 and

3.17) also identified the trypsin-digested peptides as part of Hho1p from *Saccharomyces cerevisiae*.

Table 3.1 Table showing the expected and experimental masses of peptides obtained by tryptic digest of carbonic anhydrase.

Segment	Theoretical Mass (Da)	Experimental value (Da)
1	980.07	981.231
2	1013.14	1015.06
3	1582.80	1584.17
4	2100.19	2100.58
5	2254.55	2255.91
6	2513.48	2514.64
7	2854.30	2855.99

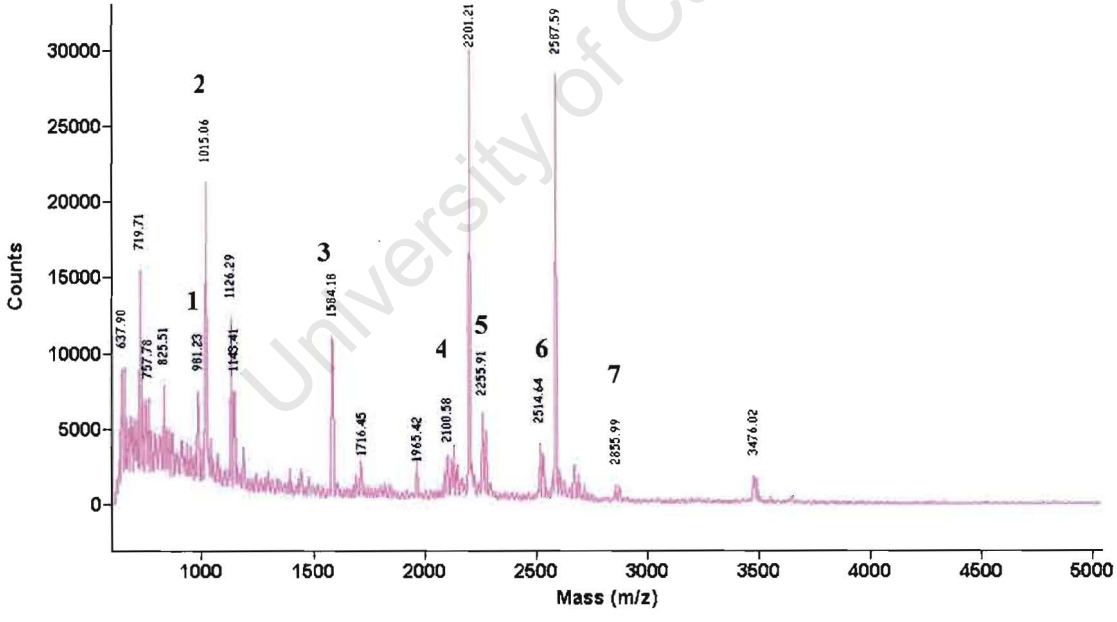


Figure 3.13 MALDI-TOF analysis of carbonic anhydrase.
Mass spectrum of carbonic anhydrase tryptic digest. Mass peaks corresponding to the peptides are indicated.

Table 3.2 Table showing the expected and experimental mass of peptides obtained by tryptic digest of GD1.

Segment	Theoretical Mass (Da)	Experimental value (Da)
1	815.95	816.983
2	1177.26	1178.21
3	1200.46	1201.45
4	1305.43	1306.43
5	1485.76	1487.36
6	1757.94	1759.81
7	2377.62	2379.91
8	2505.79	2508.22
9	2766.13	2768.66

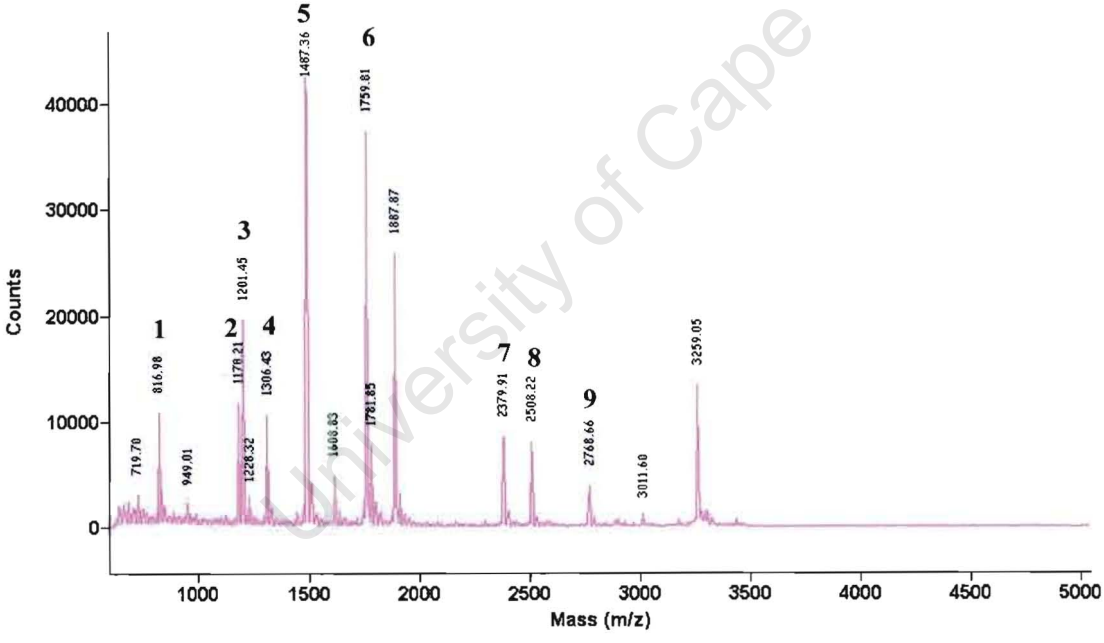


Figure 3.14 MALDI-TOF analysis of GD1.
Mass spectrum of GD1 tryptic digest. Mass peaks corresponding to the peptides are indicated.

Table 3.3 Table showing the expected and experimental mass of peptides obtained by tryptic digest of GD2 (asterisks indicate values obtained from partial digestion).

Segment	Theoretical Mass (Da)	Experimental value (Da)
1	671.81	672.6054
2	*1027.25	*1027.72
3	*1075.21	*1075.698
4	*1155.43	*1155.861
5	*1203.38	*1203.744
6	*1316.54	*1317.334
7	*1378.70	*1378.099
8	1607.76	1608.022
9	*1735.93	*1736.475
10	*1849.09	*1849.757
11	*1977.27	*1978.109
12	*2351.80	*2352.005

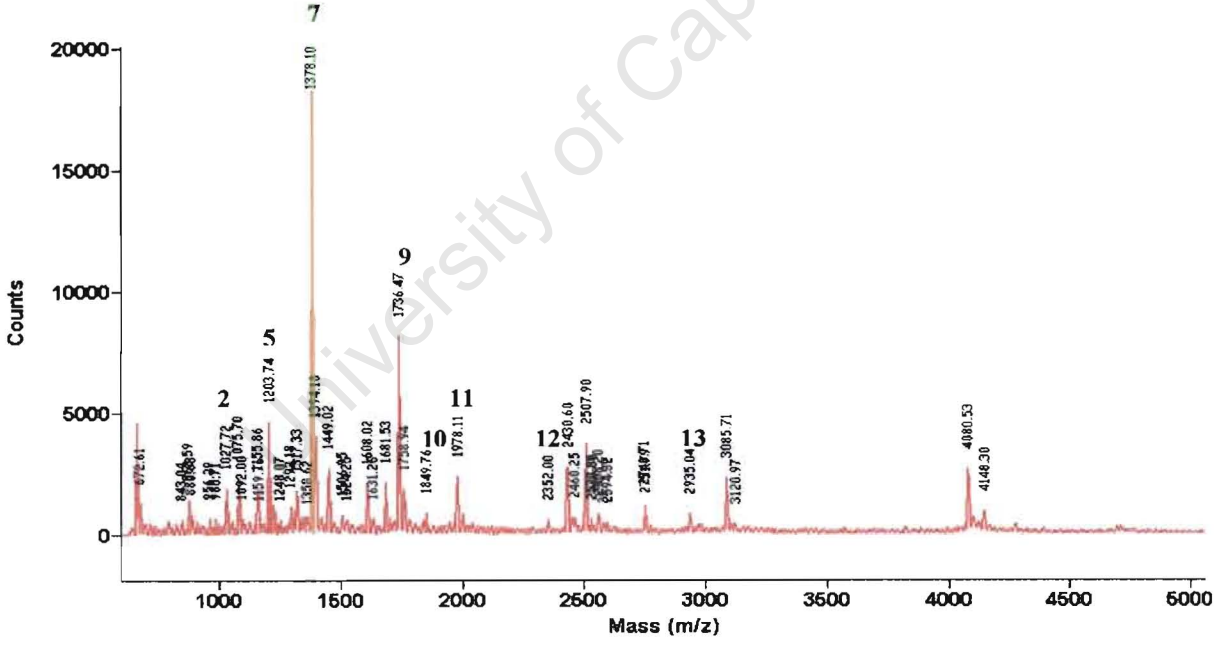


Figure 3.15 MALDI-TOF analysis of GD2.
Mass spectrum of GD2 tryptic digest. Mass peaks corresponding to the peptides are indicated.

Table 3.4 Table showing the expected and experimental masses of peptides obtained by tryptic digest of Hho1p.

Segment	Theoretical Mass (Da)	Experimental value (Da)
1	944.30	944.2261
2	1041.15	1042.11
3	1075.21	1075.251
4	1169.32	1169.207
5	1200.46	1200.345
6	1305.43	1305.414
7	1377.52	1377.481
8	1485.45	1485.644
9	1655.95	1655.821
10	1735.93	1735.499
11	1757.94	1757.588
12	2505.79	2504.95

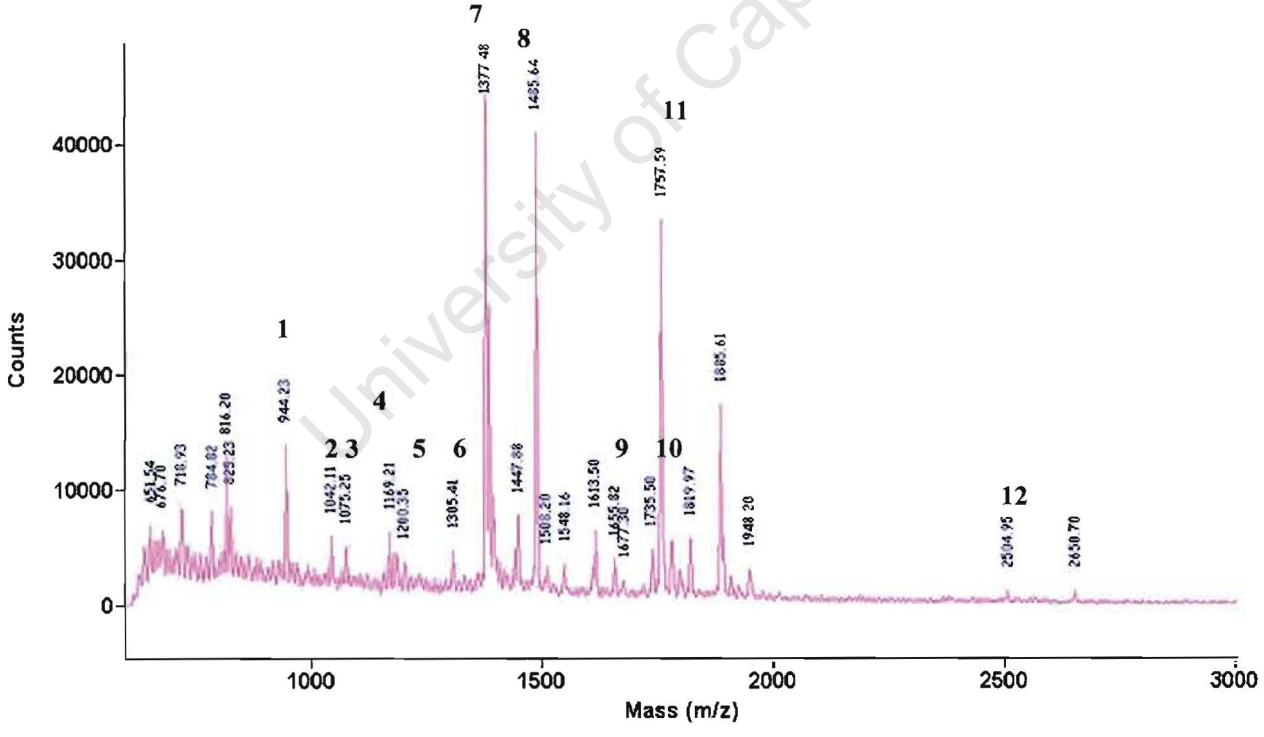


Figure 3.16 MALDI-TOF analysis of Hho1p.
Mass spectrum of Hho1p tryptic digest. Mass peaks corresponding to the peptides are indicated.

Table 3.5 Table showing the expected and experimental masses of peptides obtained by tryptic digest of Hho1-Fxap.

Segment	Theoretical Mass (Da)	Experimental value (Da)
1	944.12	943.735
2	1200.46	1200.08
3	1305.43	1305.03
4	1377.52	1376.95
5	1485.76	1485.16
6	1735.93	1735.43
7	2505.79	2503.36

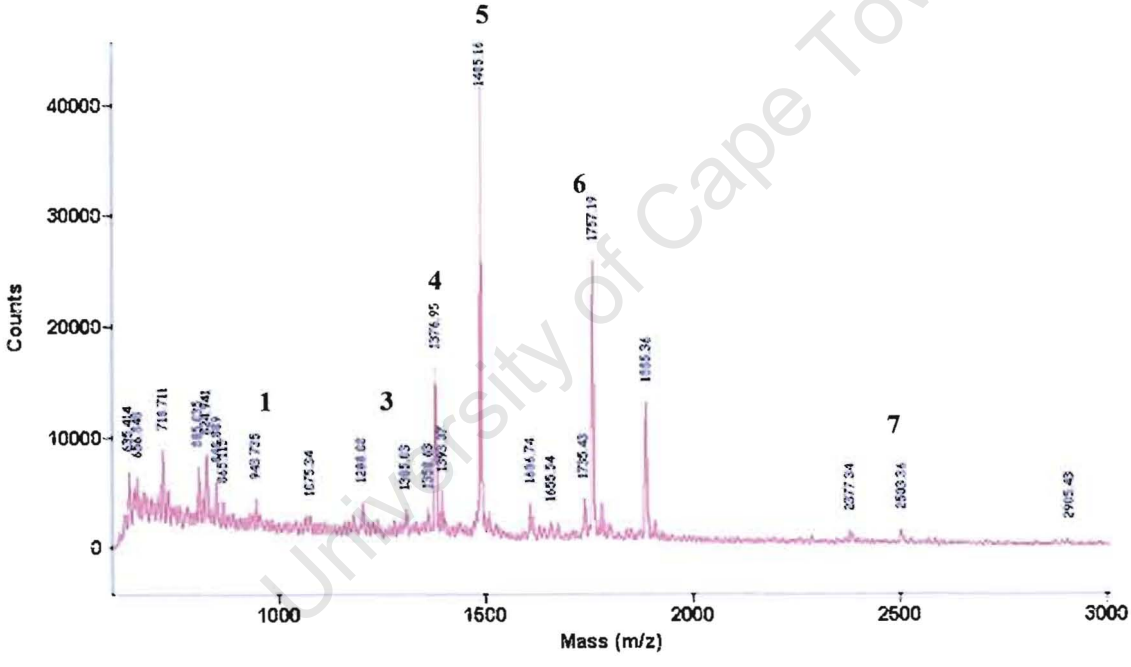


Figure 3.17 MALDI-TOF analysis of Hho1-Fxap.
Mass spectrum of Hho1-Fxap tryptic digest. Mass peaks corresponding to the peptides are indicated.

3.2.5 High performance liquid chromatography of GD1, GD2, Hho1p and Hho1-Fxap.

Following elution of GD1, GD2, Hho1p and Hho1-Fxap from the Ni^{2+} column, several contaminating proteins that co-purified with the protein of interest were present. In order to purify the recombinant proteins away from these contaminating proteins, an additional chromatographic step was used. Reverse-phase HPLC (RP-HPLC) was employed using a C-18 column. Before injecting the protein sample on the column, the protein fractions recovered from the metal-chelation chromatography were pooled and transferred to the buffer in which the column was equilibrated. Aliquots of the crude sample were injected onto the column and the recombinant proteins were eluted form the column using a linear acetonitrile gradient.

GD1 and GD2 had different retention times on the column (Figure 3.18), with GD1 eluting much earlier in the gradient. The retention time is based on the hydrophobic association between the proteins and the hydrophobic surface of the stationary phase. GD2 and GD1 have similar hydrophobicity indexes of 34.5 % and 36.38 %, respectively (DNAssist). The retention time of GD2 could be due to the conformation of the protein, where more hydrophobic residues are exposed to the hydrophobic surface, leading to a stronger association than that of GD1. The elution profiles of Hho1p and Hho1-Fxap were very similar (Figure 3.19). The very late elution of both proteins, are most likely due to the high degree of hydrophobicity for both these proteins, a characteristic of linker histones.

This method purified both the globular domains (GD1 and GD2 removing all contaminating *E. coli* proteins (Figure 3.18 C). As with the IMAC purification, GD1 seemed more stable and was obtained in much higher yield. Structural analysis by means of circular dichroism (CD) (data not shown), indicated that GD1 was in a highly folded conformation, and RP-HPLC was therefore used to purify GD1 following Ni-agarose eluted chromatography.

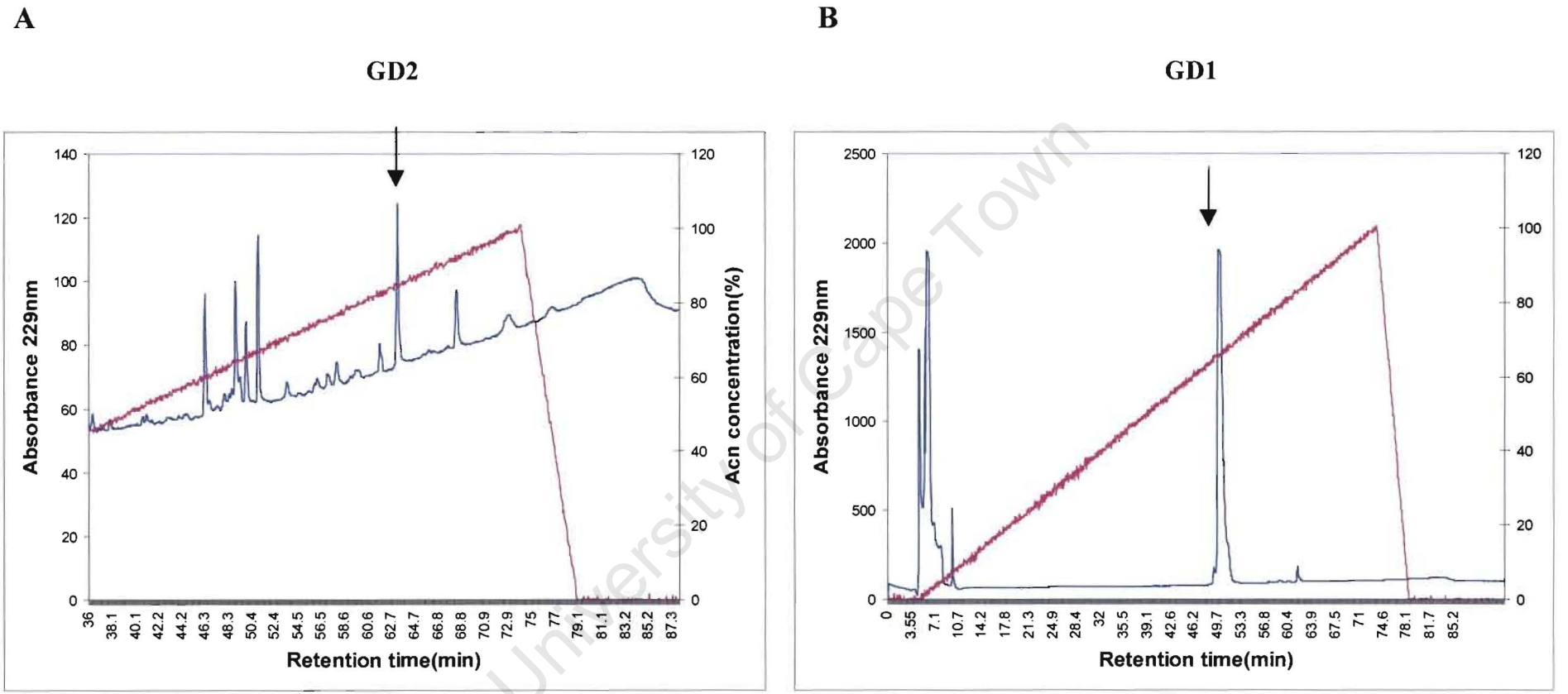


Figure 3.18 Chromatograms of the purification of recombinant GD1 and GD2 by reverse-phase HPLC. Panel A, indicates the peak corresponding to GD2 (retention time = ~ 62 min). **Panel B,** indicates the peak corresponding to GD1 (retention time = ~ 50 min). Both panels have the gradient of Buffer B (70 % Acetonitrile) superimposed on the chromatograms.

C

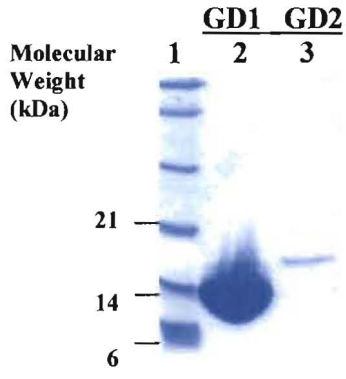


Figure 3.18 Panel C, GD1 and GD2 purified by reverse-phase HPLC. Aliquots (20μℓ) of fractions eluted during RP-HPLC were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1,** protein molecular weight marker, with sizes indicated at the left; lanes **2** and **3,** proteins eluted from the column corresponding to the peaks in Panels **B** and **A**.

There was a substantial loss of GD2 using RP-HPLC. The loss of the protein during the post RP-HPLC processes, *e.g.,* lyophilizing and buffer changes, suggested that this domain was in an unstructured conformation. The method also purified both rHho1p and rHho1-Fxa (Figure 3.19, Panels C and D), but these two proteins were also unstable in the post RP-HPLC processes. The use of the organic solvent acetonitrile is expected to unfold the proteins and it is not known if GD2, Hho1p and Hho1-Fxap will refold into the native conformation. The functional assays require that the proteins be in a native, biologically relevant conformation. Thus, an alternative method to purify GD2, Hho1p and Hho1-Fxap following initial enrichment by IMAC was investigated.

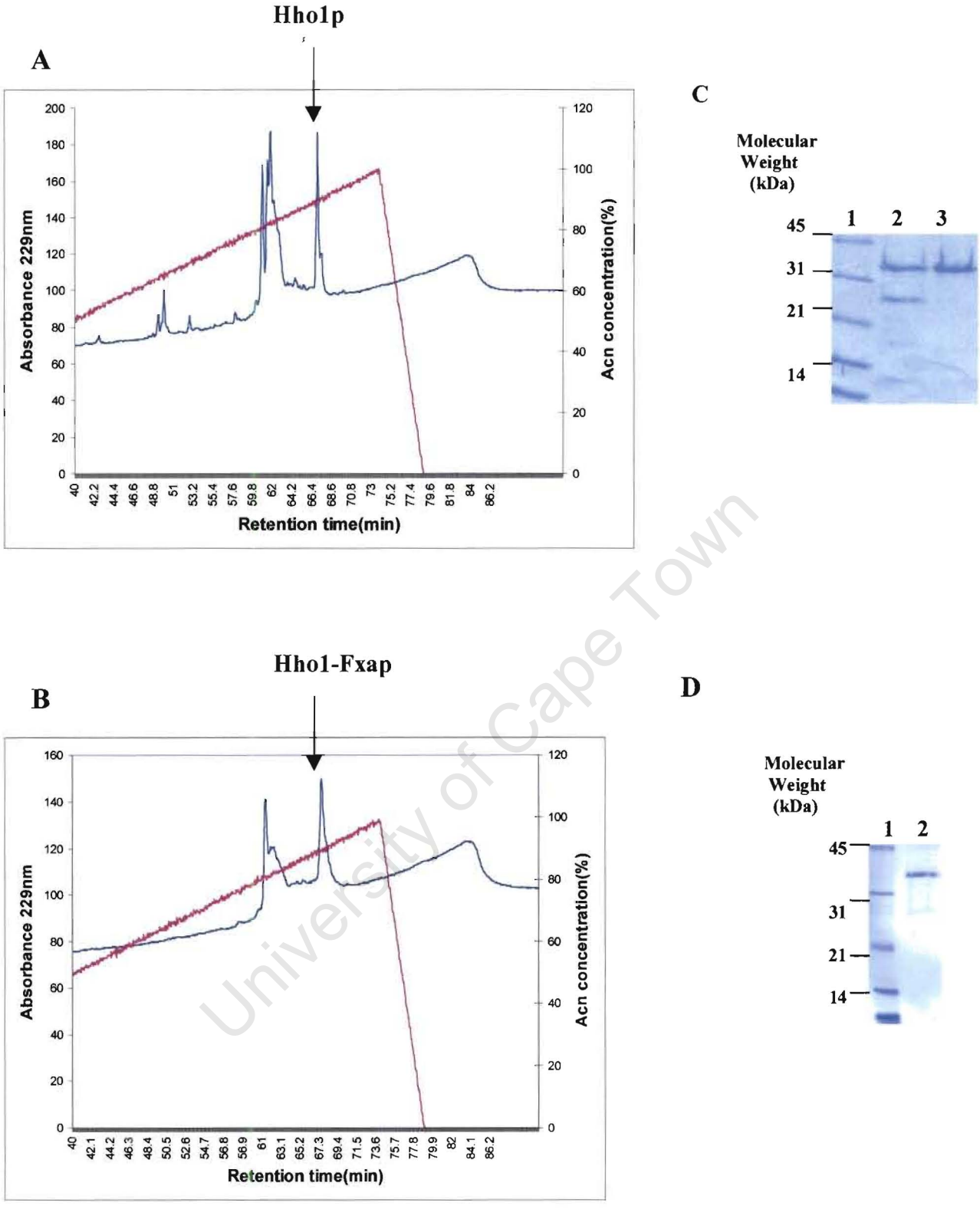


Figure 3.19 Chromatograms of the purification of recombinant Hho1p and Hho1-Fxap by reverse-phase HPLC. Panel A, the peak of Hho1p (retention time = ~ 65 min). **Panel B,** the peak of Hho1-Fxap (retention time = ~ 68 min). Both panels have the gradient superimposed on the chromatograms. **Panels C** and **D,** 15 % SDS-PAGE gels of the purified recombinant Hho1p and Hho1-Fxap, respectively. Proteins visualized by Coomassie staining. In **C** and **D;** lane **1,** protein molecular weight marker, with sizes indicated at the left. Panel **C;** lane **2,** crude protein sample; lane **3**, purified Hho1p. Panel **D;** lane **2**, purified Hho1-Fxap.

3.2.6 Further purification of the recombinant proteins.

The elution profiles obtained for all the recombinant proteins using IMAC purification illustrated the presence of approximately 10 kDa and 25 kDa *E. coli* protein that co-elute with the recombinant proteins from the Ni^{2+} charged columns at 1 M imidazole. These two *E. coli* proteins were the predominant contaminating proteins. The contaminating protein that migrated at a size roughly corresponding to 10 kDa on a SDS-PAGE gel was removed by using Ni^{2+} charged columns, and eluting in a step-wise imidazole gradient, in 100 mM steps. The aim was to determine at what imidazole concentration some of the weakly bound contaminating proteins eluted from the Ni^{2+} charged columns.

The weakly bound 10 kDa *E. coli* protein was eluted at 400-500 mM imidazole (Figure 3.20, lane 8), showing that an imidazole gradient can be used to remove the 10 kDa contaminating protein. However, this method was not sufficient to remove all contaminating proteins.

3.2.6.1 Gel filtration.

The first approach for the further purification of the proteins of interest from the remaining contaminating *E. coli* protein, was protein separation by means of gel filtration chromatography. Bio-Gel P-10, with a typical fractionation range of 1.5 to 20 kDa, was employed for GD2, and Bio-Gel P-60, with a fractionation range 3-60 kDa, was employed for the 28 kDa Hho1p. The Bio-Gel matrix contains porous polyacrylamide beads that are hydrophilic and essentially free of charge (BioRad). The fractions from the IMAC columns were concentrated in 500 mM Tris·Cl (pH 7.5), 1 mM EDTA and applied to the P-10 and P-60 columns, which were equilibrated in the same buffer. However, separation of both the proteins using gel filtration, proved unsuccessful (data not shown). GD2 seemed to have a weak interaction, presumably hydrophobic, with the matrix and was co-eluted from the column with the contaminating protein. Hho1p seemed to act in the opposite manner, and could not be eluted from the column at all, which indicated that the interaction, presumably hydrophobic, between the protein and the matrix was exceedingly strong.

Various parameters, *e.g.*, the bed volume of the column, fraction size and washing conditions were altered in an attempt to obtain separation, but proved unsuccessful.

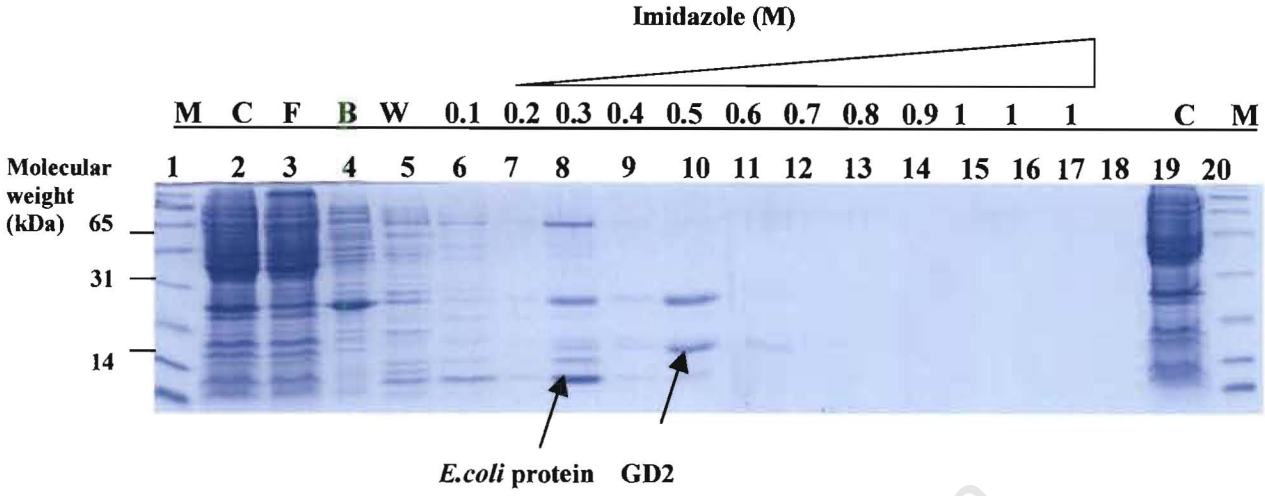


Figure 3.20 Purification of recombinant GD2 using Ni-agarose chromatography with modified elution conditions. Removal of an ~10 kDa contaminating protein.
Aliquots (20 μℓ) of fractions eluted during Ni-chromatography steps were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1**, protein molecular weight marker, with sizes indicated at the left; lanes **2,** crude protein extract; lane **3,** binding flow-through; lane **4**, binding buffer eluate; lane **5**, wash buffer eluate; lanes **6-17** protein fractions eluted with a linear gradient of 0.1-1 M imidazole. The position of the contaminating *E. coli* protein and GD2 are indicated.

3.2.6.2 Elution with NaCl using Biorex-70.

The next approach was ion-exchange chromatography. Biorex-70 is a weakly acidic cation exchange resin which contains carboxylic acid exchange groups on the acrylic polymer lattice (BioRad). This matrix is hydrophilic and free of nonspecific binding or denaturing effects on proteins. An ion-exchange procedure was followed where the counterions on the matrix were replaced by the sample ions that had the same charge. This method had been employed previously to purify linker histones (D'Anna *et al.*, 1979; Thompsom *et al.*, 1976). It was also been used for the removal of trace ions from linker histone preparations (Hayes, 1996).

The first protocol used with the Biorex-70 cation exchange resin was simply to apply the protein sample to the resin followed by elution with an increasing salt concentration. The isolated GD2 and Hho1p recovered from the metal-chelation chromatography was pooled and transferred to 10 mM Tris·Cl (pH 7.5), using a 10 kDa MWCO ultracentrifugation column, passed over the cation exchange resin, and eluted using a linear 0.05-2 M NaCl gradient. Using these conditions (Figure 3.21 B) GD2 exhibited weak binding to the resin, evident from the loss of the protein in the

flow-through fraction (lane 3). GD2 was mainly recovered in the wash fraction (lane 4-5). However, the NaCl gradient did not separate GD2 from the contaminating proteins. The Hho1p (Figure 3.21 A) had a strong interaction with the matrix and minimal protein was lost in the flow-through of wash fraction. However, when the Hho1p eluted from the column, it did so with the contaminating proteins. Increasing the bed volume of the packed column to improve this separation did not alter the elution profile (data not shown).

A

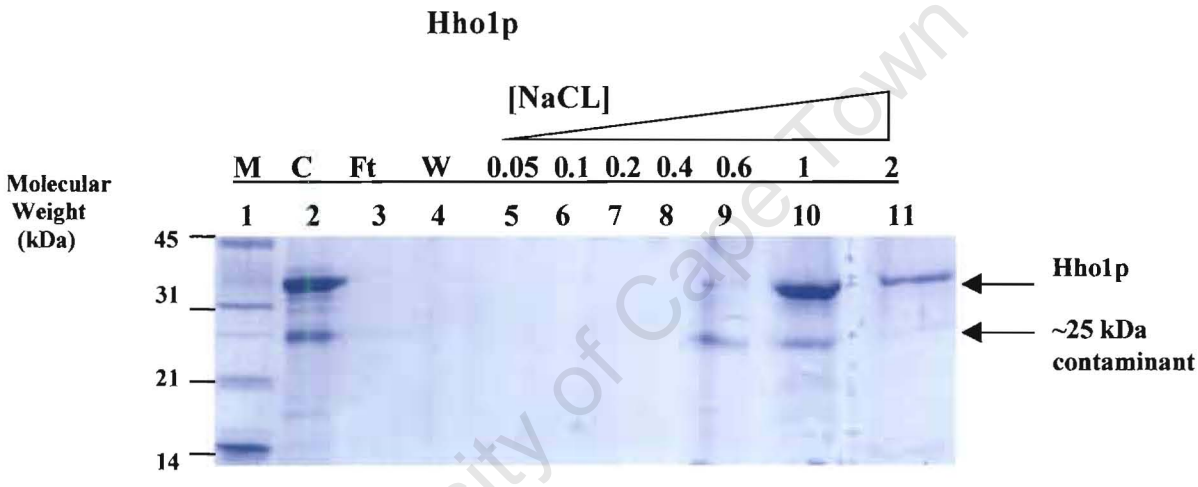


Figure 3.21 Panel A, Purification of recombinant Hho1p using Biorex-70, eluting with increasing NaCl gradient.

Aliquots (20 μℓ) of fractions eluted during ion-exchange chromatography steps were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1,** protein molecular weight marker, with sizes indicated at the left; lane **2,** crude protein extract; lane **3,** binding flow-through; lane; lane **4,** wash eluate; lanes **5-11** protein fractions eluted at increasing NaCl concentration. Positions of Hho1p and the ~ 25 kDa contaminant are indicated.

B

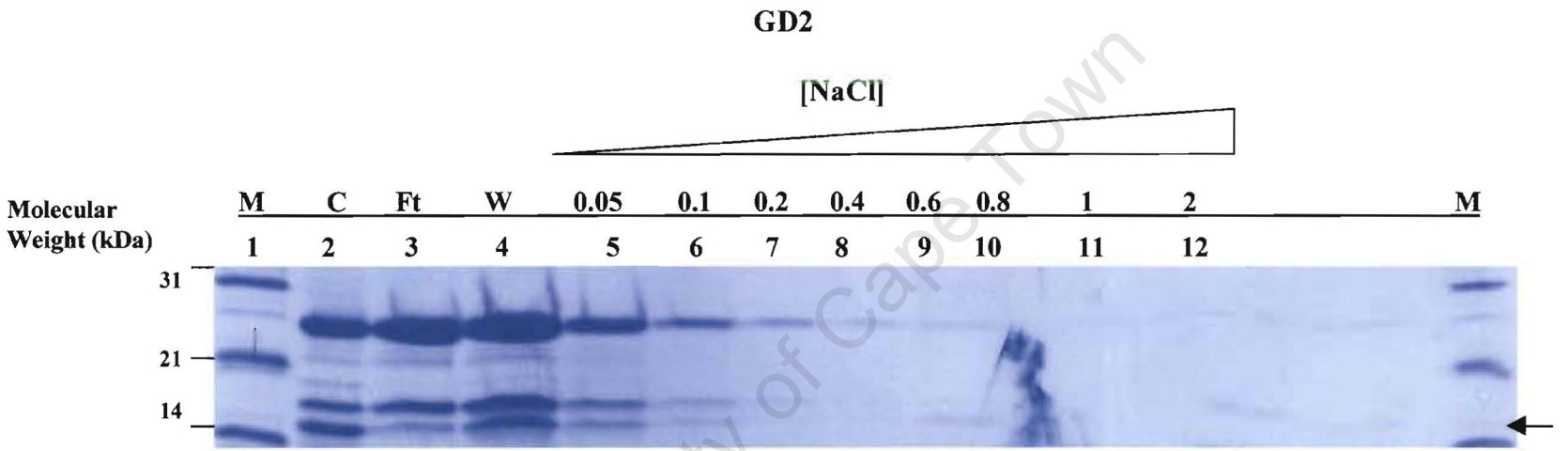


Figure 3.21 Panel B, Purification of recombinant GD2 using Biorex-70, eluting with increased NaCl concentration.
Aliquots (20 μℓ) of fractions eluted during ion-exchange chromatography, were electrophoresed on a 15 % SDS-PAGE gel, and the protein visualised by Coomassie stain. Lane **1**, protein molecular weight marker, with sizes indicated at the left; lane **2**, crude protein extract; lane **3**, binding flow-through; lane **4**, wash eluate; lanes **5-12**, elution fractions with increasing NaCl concentration. The position of GD2 is indicated.

3.2.6.3 Elution using a GuHCl linear gradient.

We next attempted a protocol, previously used to separate H1 from HMG1 and HMG2 proteins, namely chromatography on Biorex-70 (Kincade & Cole, 1966).

The Biorex-70 resin and pooled Ni^{2+} elution fractions were equilibrated in 8 % guanidinium hydrochloride (GuHCl) and buffered at pH 6.8 with 0.1 M K_2PO_4. The protein sample was loaded onto the cation exchange resin, and eluted with a linear GuHCl gradient from 10-50 % (w/v), containing 0.1 M K_2PO_4 (pH 6.8). Using these conditions, Hho1p was seperated from the contaminating proteins (Figure 3.22 B). However, there was a significant loss of Hho1p in the flow-through and wash fractions. There was also a loss of GD2 in the flow-through and wash fractions, but no separation between GD2 and the contaminants were obtained under these conditions (Figure 3.22 A).

The loss of protein in the flow-through fraction indicated that the binding to the matrix was not sufficient. In an attempt to optimize interaction between the hydrophobic proteins and the matrix, a buffer with a higher pH was used. After equilibrating both the protein samples and the matrix in 20 mM CAP (pH 10.5), 8% GuHCl, the same elution conditions were followed. The elution profiles for both proteins using the pH 10.5 buffer are shown in Figure 3.23. It is evident that the higher pH value did not increase the binding of either of these proteins to the matrix. This confirmed that the weak hydrophobic interaction observed between the recombinant proteins and the Biorex-70 matrix was not due to the pH of the buffer, but was a result of the initial concentration of GuHCl, with which both the crude protein sample and the matrix had been equilibrated.

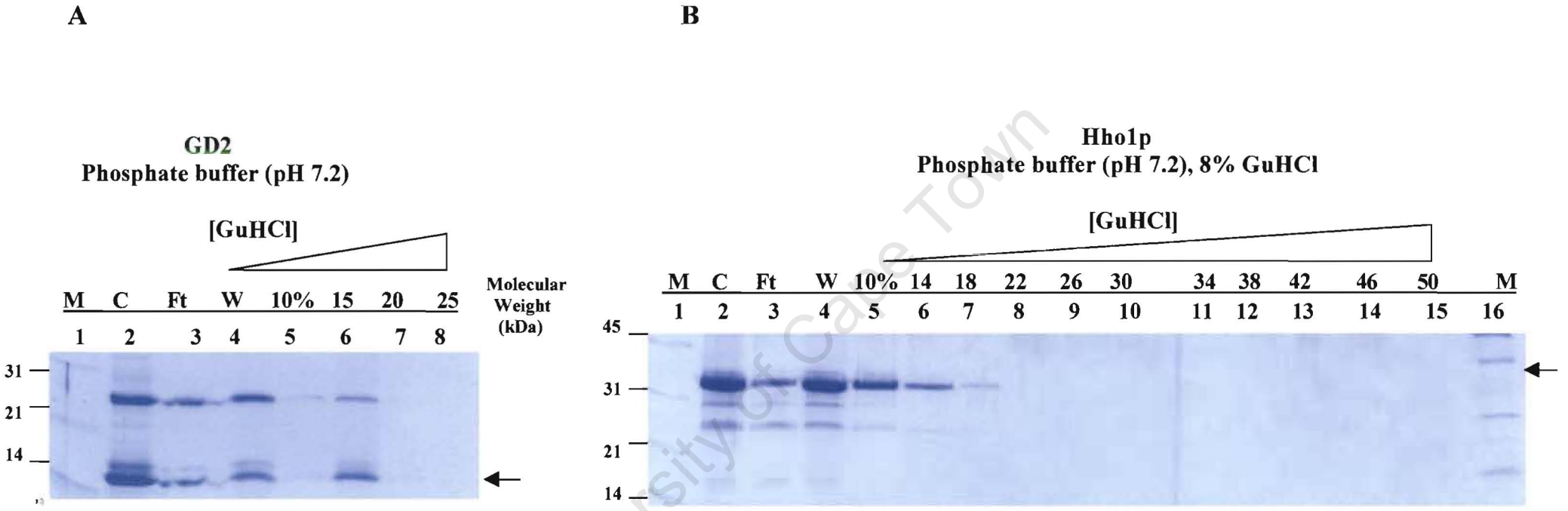


Figure 3.22 Purification of GD2 and Hho1p using Biorex-70, eluting with an increasing GuHCl concentration.
Resin and protein sample equilibrated in 0.1 M phosphate buffer (pH 7.2), 8 % GuHCl. **Panel A**, elution of GD2 and **Panel B**, elution of Hho1p.
Aliquots (20 μℓ) of fractions eluted during ion-exchange chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lanes **1** and **16,** protein molecular weight marker, with sizes indicated at the left; lane **2,** crude protein extract; lane **3,** binding flow-through; lane **4,** wash eluate; Panel **A** lanes **5-8** and Panel **B** lanes **5-16;** protein fractions eluted at increasing GuHCl concentration. The position of GD2 and Hho1p are indicated.

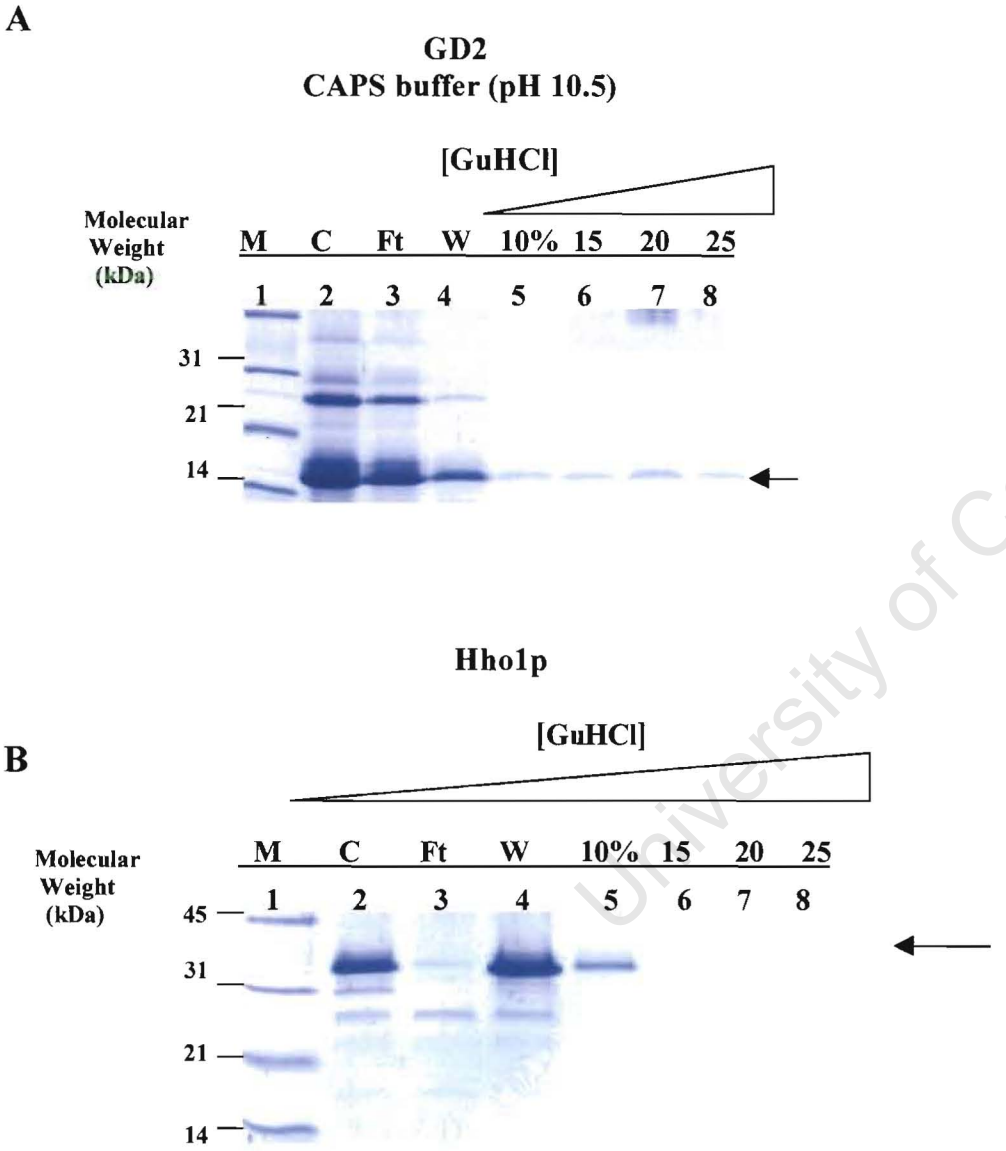


Figure 3.23 Purification of GD2 and Hho1p using Biorex-70 resin.
Resin and crude protein sample equilibrated in 20 mM CAPS buffer (pH 10.5), 8 % GuHCl. The protein was eluted with an increasing GuHCl concentration. **Panel A**, elution of GD2. **Panel B**, elution of Hho1p. Aliquots (20 $\mu\ell$) of fractions eluted during the chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1**, protein molecular weight marker, with sizes indicated at the left; lane **2**, crude protein extract; lane **3**, binding flow-through; lane **4**, wash eluate; lanes **5-8** elution fractions with increasing GuHCl concentration. Positions of GD2 and Hho1p are indicated.

Omitting the 8 % guanidium hydrocholoride from the loading buffer and protein samples, led to a strong interaction between the proteins and the resin bed at both 7.2 and 10.5 pH values (Figure 3.24). This method separated the proteins from all contaminants and in an attempt to maintain the native charge and structure of the proteins, the 0.1 M phosphate buffer (pH 7.2) was used in future purifications of all recombinant proteins. Hho1-Fxa was separated using this optimized protocol (Figure 3.25).

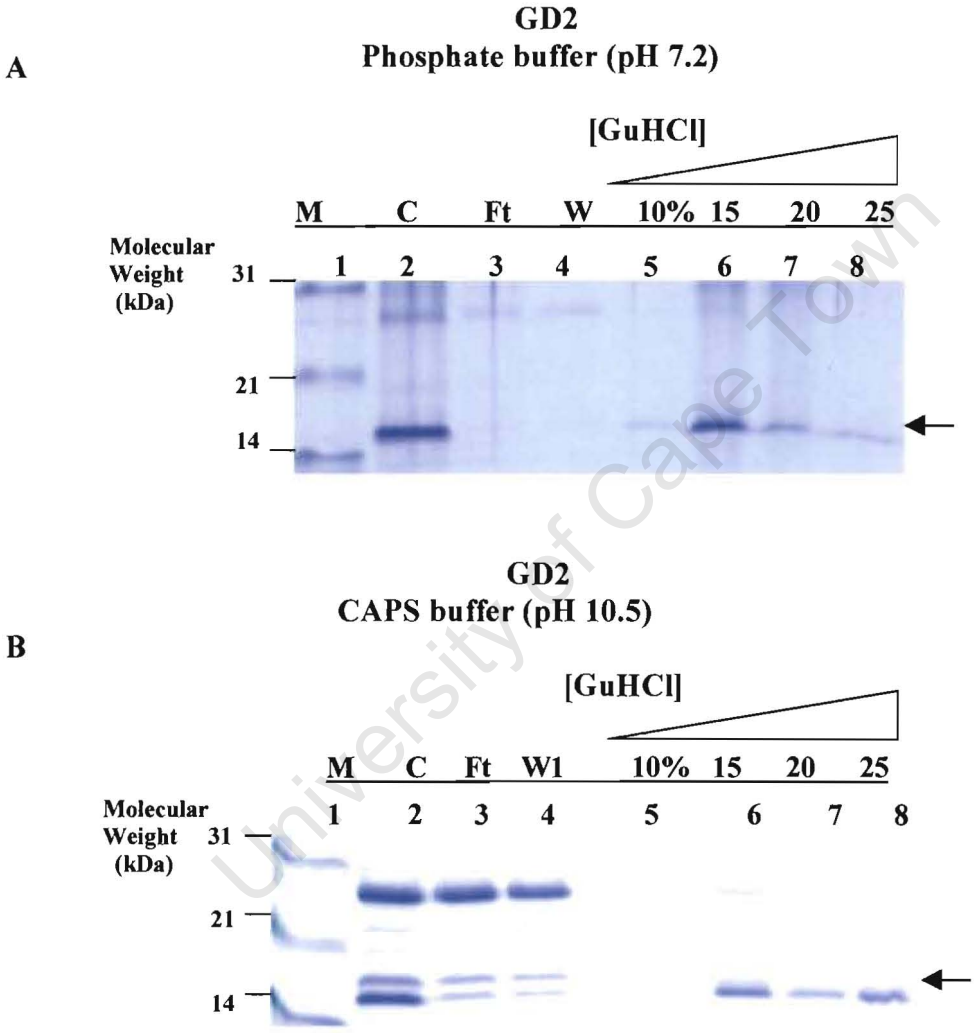


Figure 3.24 Panel A and B, Purification of GD2 using Biorex-70 resin.
Resin and crude protein sample equilibrated in 0.1 M phosphate buffer in **Panel A** and 20 mM CAPS buffer in **Panel B**. Both samples (A and B) were eluted with an increasing GuHCl concentration. Aliquots (20 μℓ) of fractions eluted during chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lanes **1**, protein molecular weight marker, with sizes at left; lanes **2**, crude protein extract; lane **3**, binding flow-through; lane **4**, wash eluate; lanes **5-8**, protein fraction eluted at increasing GuHCl concentration. Position of GD2 is indicated.

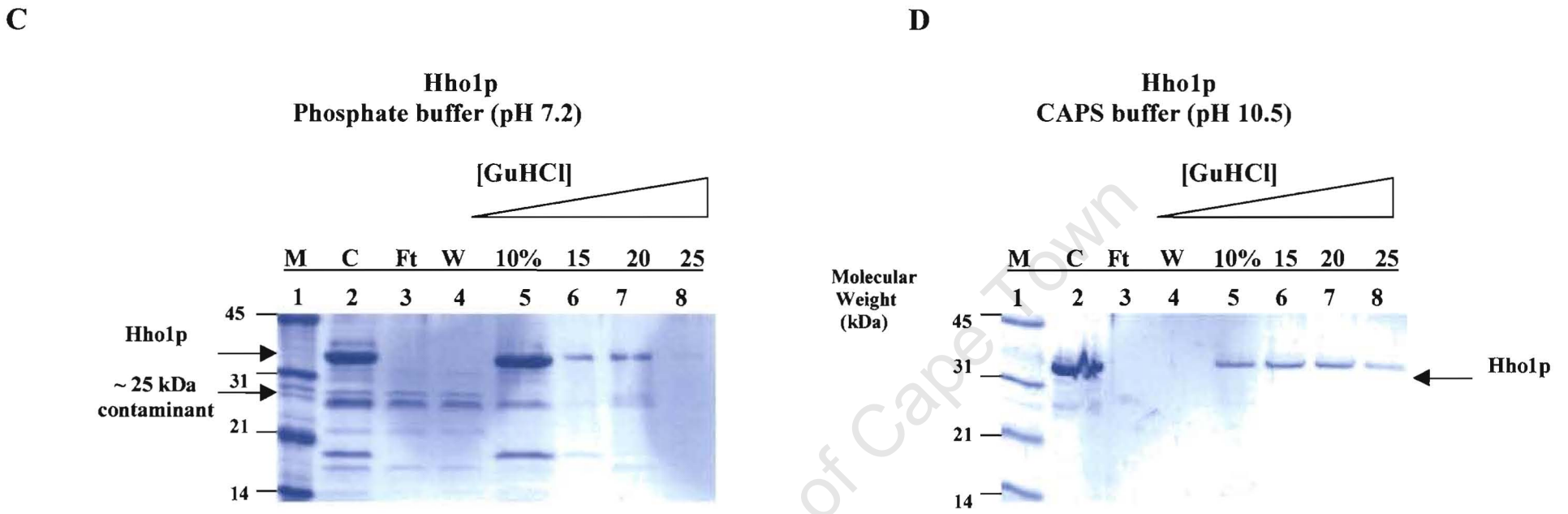


Figure 3.24 Panel C and D, Purification of recombinant Hho1p using Biorex-70 resin.
Resin and crude protein sample equilibrated in 0.1 M phosphate buffer in **Panel A** and 20 mM CAPS buffer in **Panel B**. Both samples (A and B) were eluted with an increasing GuHCl concentration. Aliquots (20 $\mu\ell$) of fractions eluted during chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lanes **1**, protein molecular weight marker, with sizes indicated at left; lanes **2**, crude protein extract; lane **3**, binding flow-through; lane **4**, wash eluate; lanes **5-8**, 10 % – 25 % indicate the GuHCl concentration in the buffer used to elute the proteins. The position of Hho1p and an ~ 25 kDa contaminant are indicated.

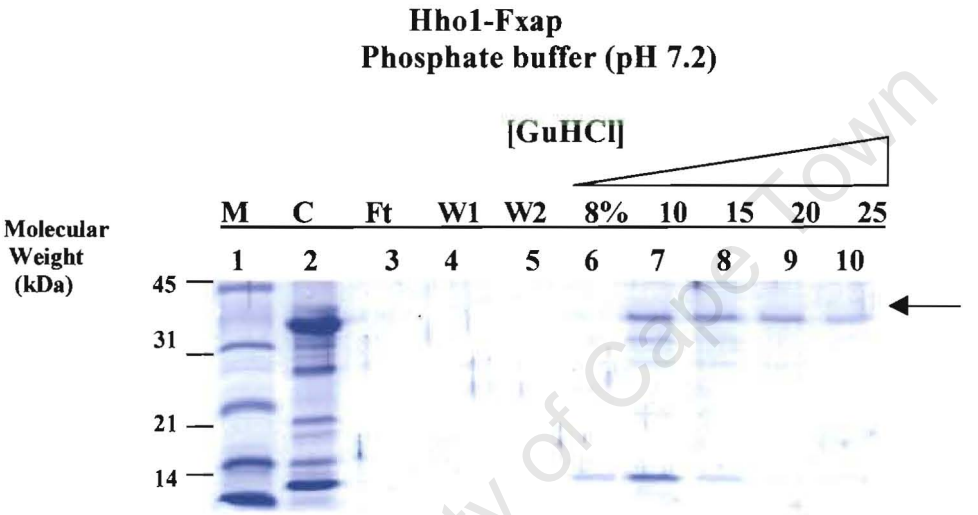
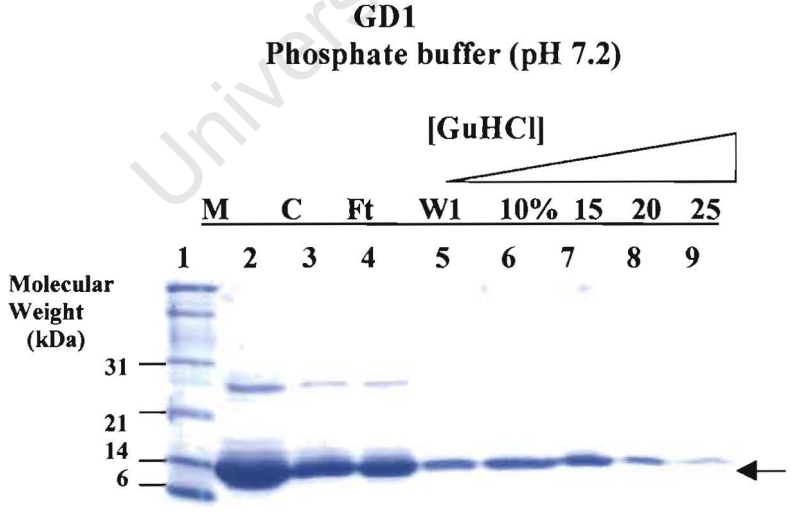


Figure 3.25 Purification of recombinant Hho1-Fxap using Biorex-70 resin.
Resin and crude protein sample equilibrated in 0.1 M phosphate buffer (pH 7.2), 8 % GuHCl. The protein sample was eluted with an increasing GuHCL concentration. Aliquots (20 $\mu\ell$) of fractions eluted during chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1**, protein molecular weight marker, with sizes indicated at the left; lane **2**, crude protein extract; lane **3**, binding flow-through; lane **4** and **5**, wash eluate; lanes **6-10**, protein fractions eluted with increasing GuHCl. The position of Hho1-Fxap is indicated.

3.2.6.4 GD1 separation using Biorex-70.

The expression and IMAC purification of the separate globular domains differed to a large degree. GD1 was also the only protein that was purified in large quantities using HPLC. To investigate if these differences between GD1 and GD2 would also extend to the ion exchange chromatography, GD1 was purified using the same conditions as those employed for the GD2 purification. Using the phosphate buffer the interaction between GD1 and the matrix was weak, evident in the loss of the protein in the flow-through and wash fractions (Figure 3.26 A, lanes 3 and 4), which is different from the strong interaction shown by GD2 for the matrix with the use of the same buffer (Figure 3.24 A). Binding of GD1 to the matrix using CAPS buffer was much stronger than that shown for GD2 in the 10.5 pH buffer (Figure 3.26 B, lanes 3 and 4 and Figure 2.24 B, lanes 3 and 4, respectively). These results underline possible differences in GD1 and GD2 structures, and further investigations into this would be useful in order to understand the nature of the separate globular domains.

A



B

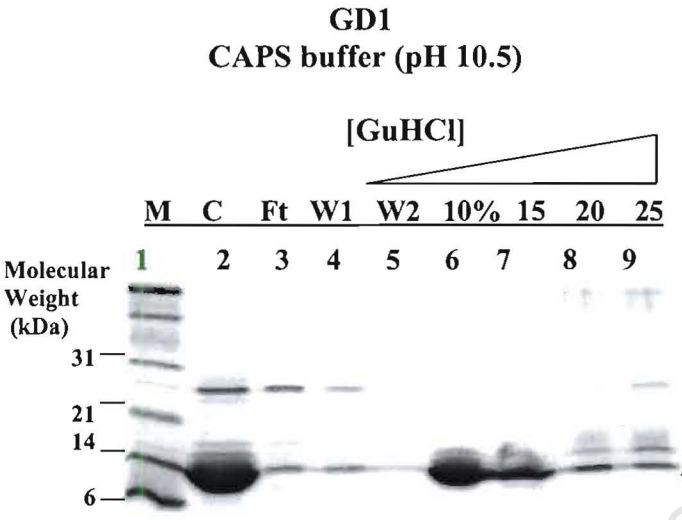


Figure 3.26 Purification of recombinant GD1 using Biorex-70 resin.
Resin and crude protein sample equilibrated in 0.1 M phosphate buffer in **Panel A** and 20 mM CAPS buffer in **Panel B**. Both protein samples (A and B) were eluted with an increasing GuHCl concentration. Aliquots (20 $\mu\ell$) of fractions eluted during chromatography were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lanes **1**, protein molecular weight marker, with sizes indicated at the left; lanes **2**, crude protein extract; lane **3**, binding flow-through; lane **4** and **5**, wash eluate; lanes **6-9**, protein fractions eluted at increasing GuHCl concentration. The position of GD1 is indicated.

3.2.7 Confirming the presence of the protease site in Hho1-Fxap.

The coding sequence for the Factor Xa protease site had been confirmed with DNA sequencing (section 3.2.2.1) and the identity of the modified Hho1p protein was verified with MALDI-TOF analysis (section 3.2.4.1). The MALDI-TOF analysis could not confirm the presence of the protease site, because on tryptic digestion the expected mass of the peptide containing this site was less than the low mass gate range of the MALDI-TOF mass spectrometer. In an attempt to further confirm that the protein did contain the protease site, the purified Hho1-FXap protein (section 3.2.6.3) was digested with Factor Xa. The expected sizes for the digestion of Hho1-FXap with

this protease are approximately 13 and 17 kDa (Figure 3.27). This digestion revealed that the site is present and accessible to digestion by the protease.

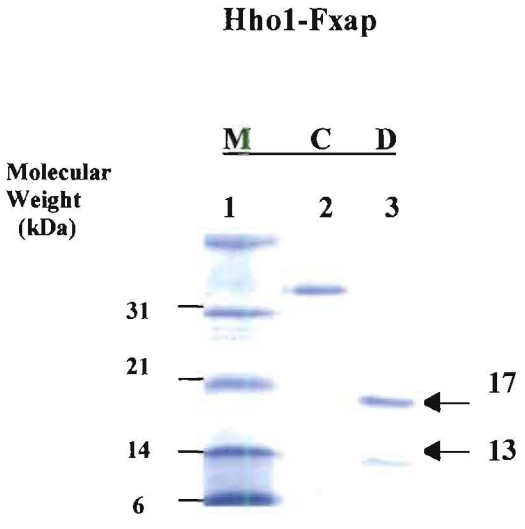


Figure 3.27 Factor Xa protease digestion of recombinant Hho1-Fxap.
Aliquots (20 μℓ) from the samples were electrophoresed on a 15 % SDS-PAGE gel and the protein was visualised by Coomassie stain. Lane **1**, protein molecular weight marker, with sizes indicated at the left; lanes **2**, crude protein; lane **3**, digested protein fragments. The expected fragment sizes of the digested protein are indicated.

3.3 Summary.

The first step in the development of the magnetic bead assay was the over-expression and purification of the required proteins, GD1, GD2, Hho1p and Hho1-FXap. All the cloning steps were based on sequence alignment data. It was verified by nucleotide sequencing that the coding regions of all the relevant constructs were free of mutations. The sequencing results also confirmed that the coding regions of GD1, GD2 and Hho1-FXap were in frame with the C- or N-terminal hexa histidine-tags.
The induction time-course experiments (section 3.2.3) indicated that protein expression was optimal at 1-2 hr for all proteins. The identity of the proteins, were verified using MALDI-TOF mass spectroscopy and the purification of the proteins involved a number of chromatographic steps (section 3.2.4).

The globular domains, GD1 and GD2, were purified in order to facilitate the study on the DNA binding characteristics of the separate domains. In the first chromatographic step, the IMAC purification, the final yield for GD1 and GD2 differed significantly. This difference may be due to the conformational state of GD2 under the purification conditions. The protein may also be subject to proteolysis if it is in an unstructured conformation. GD2 has been shown to be in a unstructured state in 10 mM Na_2PO_4 using CD spectra (Ali, 2001). In that study, the expressed protein GII (corresponding to GD2 in this study) had the same coding region as the globular domain coding region used for GD2 in our study. The difference in yield between GD1 and GD2 was observed in all subsequent purification steps and investigating the reason for this difference may contribute to the study of the binding behaviour of both the isolated globular domains and of the full-length Hho1p to nucleosomes in chromatin.

Both Hho1p and Hho1-Fxap were also purified in sufficient quantities to continue with the magnetic bead assays. The presence of the mutated sequence encoding the protease recognition and cleavage site in the coding sequence was verified by nucleotide sequencing. The identity of the expressed protein was confirmed with the MALDI-TOF analysis of the trypsin-digested fragment, and the presence of the protease site was confirmed by digestion of the purified protein with Factor Xa protease.

GH5-linker-GH5 will be an important control in the magnetic bead pull-down assay. Therefore, a requirement for the conclusion of this study is the cloning of the GH5-linker-GH5 fragment. Since this construct will contain two identical sequences repeated in proximity on the same DNA molecule, it is possible that recombination may result in the entire or partial removal of the sequence encoding GH5. For this reason, recombination deficient hosts may have to be used to successfully construct and express the GH5-linker-GH5 peptide. Several hosts are currently being tested in this regard.

Chapter 4: Development of a magnetic bead assay to study histone H1 binding.

4.1 Introduction.

Linker histones associate with nucleosomes by binding to the nucleosomal DNA, and play an important role in the stabilisation and organization of the higher order structures in chromatin (Thoma *et al.*, 1979; van Holde, 1989). It was shown that the isolated globular domain of GH5 alone protected an additional 20 bp of nucleosomal DNA from micrococcal nuclease digestion (Allan *et al.*, 1980).

As previously described (Chapter 1), the linker histone in yeast, Hho1p, exhibits an unique structure in that it contains two putative globular domains, connected by a lysine-rich peptide region. Sequence alignment of these putative globular domains indicates that globular domain one (GD1) contains six out of the seven basic amino acids that make up the two proposed DNA binding sites of GH5 (Ramakrishnan *et al.*, 1993), whereas globular domain two (GD2) contains four of these residues (Landsman, 1996; Patterton *et al.*, 1998). Despite this unique structure, Hho1p associates with the nucleosome and produces a ~168 bp kinetic pause during micrococcal nuclease digestion (Patterton *et al.*, 1998). Both these are characteristic features of the linker histones. However, the ability of the separate globular domains to independently interact with the nucleosome has not been established. This could lead to an understanding of the function of this unique linker histone. Therefore, an understanding of the binding of yeast H1 in chromatin will contribute to an insight into the role of this protein in chromatin structure in yeast and of linker histones in general.

The successful purification of the separate globular domains (GD1, GD2), Hho1 and the modified linker histone, Hho1-FXa (Chapter 3), was the first step in establishing a system for structural studies. Here, the design and optimisation of certain techniques are described that will facilitate a detailed study of the association of the separate globular domains with nucleosomal DNA.

4.2 Results.

4.2.1 Preparation of the DNA template and donor chromatin.

4.2.1.1 DNA template.

The DNA template used for the subsequent core particle and chromatosome reconstitutions was a cloned DNA, 5S 207-12, composed of 12 tandem repeats of a 207-bp fragment containing the *Lytechinus variegates* 5S rRNA gene (kindly provided by R.T. Simpson). Each of these 207-bp repeats contains a region important for the formation of a positioned nucleosome core on this gene (Simpson and Stafford, 1983).

4.2.1.2 Labeling of the 207-bp 5S fragment.

In order to investigate the association of a single Hho1p with two nucleosome cores simultaneously, two kinds of labeled nucleosome cores were required. An outline of the approach taken is shown in Figure 4.1.

Briefly, nucleosome cores reconstituted on a biotin-labeled 5S fragment and nucleosome cores reconstituted on radiolabeled fragment will be incubated with streptavidin magnetic beads, where the high binding affinity between streptavidin and biotin will be utilized. At this point, the biotin-labeled nucleosomes will be bound to the magnetic beads and the radiolabeled nucleosomes will remain unbound and free in the solution. With the addition of Hho1p and the application of a magnetic separator the binding of the yeast Hho1p to two nucleosomes will be assayed by the ability of the radio label to be removed from the supernatant in an Hho1p-dependent manner.

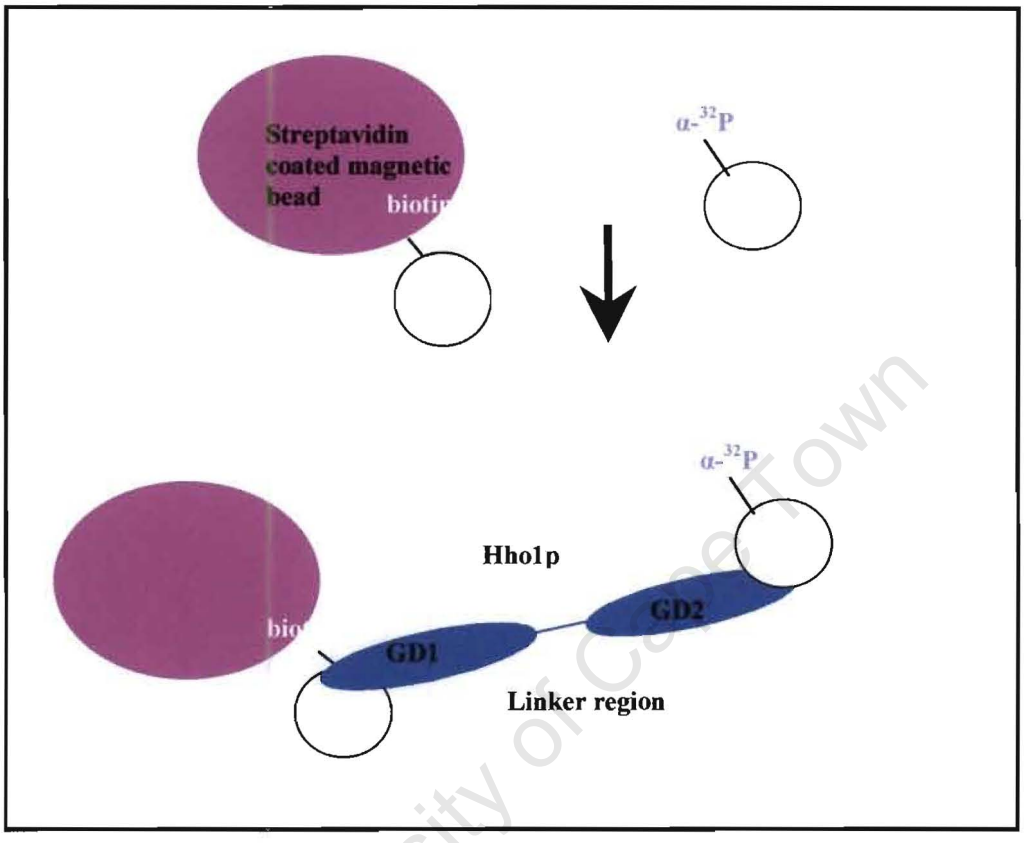


Figure 4.1 Schematic representation of the magnetic bead pull-down assay.
The diagram shows the biotin- and radiolabeled nucleosomes, with the biotin-labeled nucleosome bound to the streptavidin magnetic beads and the radiolabeled nucleosome free in solution. On addition of Hho1p, and binding of both proposed globular domains to nucleosome cores, the free radiolabeled nucleosomes may be depleted in solution in a Hho1p-association dependent manner. The use of the Hho1-FXa control protein will allow the cleavage of the lysine-rich region joining the globular domains, prior to the pull-down, and the subsequent release of the radiolabeled nucleosomes back into the solution.

To achieve this, the 3' end of the *Ava I*-digested 207-bp fragment was labeled with biotin using biotin-16ddUTP in an established procedure, and the 5' end of another DNA sample was radiolabeled with ^{32}P, using an established technique. The use of biotin-16ddUTP allowed the incorporation of a single dideoxyuridine-nucleoside at the 3' end of the 207-bp fragment. A single biotin moiety was desired because multiple biotin moieties attached to the 5S fragment could sterically impede nucleosome core formation, as well as subsequent binding of the Hho1p to the nucleosome.

4.2.1.3 Biotin end-labeling efficiency.

To determine the efficiency with which the biotin-nucleotide was incorporated into the 207-bp *Ava I* fragment, a dot blotting technique using a streptavidin-horse radish peroxidase (HRP) conjugate, with subsequent chemiluminescent detection was employed. This exploited the high binding affinity ($K_a = 10^{15}\ M^{-1}$) of biotin to streptavidin (Green, 1975) and proved to be a very sensitive quantitative assay.

A standard to quantitate biotin labeling efficiency was prepared by synthesizing a random sequence oligonucleotide primer that contained a biotin moiety at the 5' end. This primer was purified and its concentration spectrophotometrically determined.

Figure 4.2 shows an example of one quantification. The amount of biotin-labeled 207-bp fragment was quantified by comparison to the biotin-labeled standard that was blotted on the same membrane. From this, approximately 74% of the 207-bp fragments in the labeling reaction was biotin-labeled (average of five independent determinations).

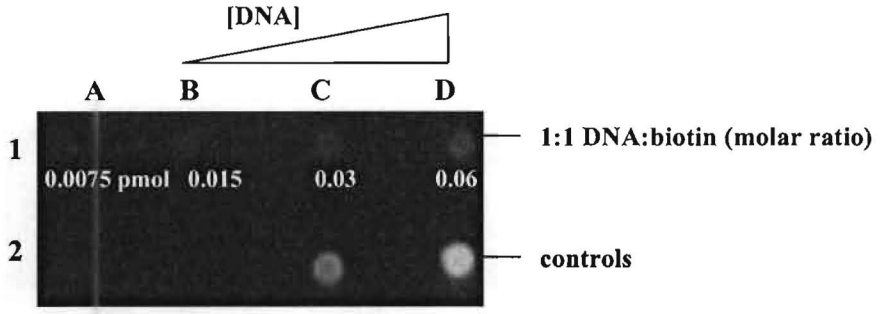


Figure 4.2 Biotin labelling effciency.
Indication of both the labeling efficiency and the percentage incorporation of the biotin-16ddUTP moiety. Chemiluminescent image of the membrane is shown. Row **1**; columns **A-D**, dilution series of increasing amounts of biotin-labeled 207-bp fragment; column **A**, 0.0075 pmol; **B**, 0.015 pmol; **C**, 0.03 pmol; **D**, 0.06 pmol. An 1:1 207-bp:biotin-16ddUTP picomolar ratio was used in the labeling reaction. Row **2**; columns **A-D**, controls used for chemiluminescent detection; column **A** and **B**; negative controls with 2 pmol of 207-bp fragment and 2 pmol of biotin-16ddUTP respectively. Positive controls are shown in columns **C** and **D**, with 0.05 and 0.1 pmol of biotin-containing oligonucleotide standard, respectively.

4.2.2 Nucleosome core reconstitutions.

The ability of the histone octamer to preferentially bind to a particular sequence is termed nucleosome positioning (van Holde, 1989). This property is illustrated with the binding of the octamer to a defined sequence of DNA from the sea urchin *Lytechinus variegates* that contains the 5S rRNA gene (Simpson, 1983). Numerous sequences have been shown to position nucleosomes, a phenomenon that is related to the sequence and sequence-dependent anisotropic flexibility of the DNA molecule (van Holde, 1989).

4.2.2.1 Gel shifts and nucleosome positioning.

Nucleosome cores were reconstituted onto the 207-bp fragment by high-salt exchange of the octamer from the H1-stripped long chromatin fragments (Rhodes *et al.*, 1985), (Figure 4.3). Approximately 75% of the 207-bp 5S DNA fragment was incorporated into the nucleosome core, determined by analysis of the phosphorimage. The same reconstituted sample, electrophoresed on a non-denaturing polyacrylamide gel, reveals multiple bands, thought to represent different translational settings of the single nucleosome core on the fragment (Figure 4.4). It was previously shown that

different translational settings resulted in structures that could be resolved by high resolution polyacrylamide electrophoresis (Patterton and Hapgood, 1996). Multiple settings on the 207-bp fragment were also previously shown (Dong *et al.,* 1990).

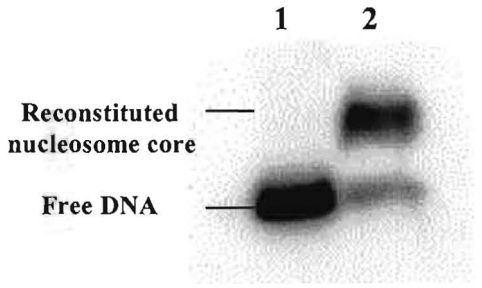


Figure 4.3 Nucleosomal core reconstitution onto [α-^{32}P]dCTP-labelled 5S fragment. Phosphorimage of samples electrophoresed in a 0.7% (w/v) agarose gel in 0.5 × TBE showing the mobility change caused by the reconstitution of the histones onto the 207-bp fragment. Lane **1**, 40 μℓ of the reconstitution reaction containing only buffer and radiolabeled 5S fragment; lane **2**, 40 μℓ of the reconstitution reaction containing H1-stripped long chromatin. The positions of the free DNA and reconstituted nucleosome cores are indicated.

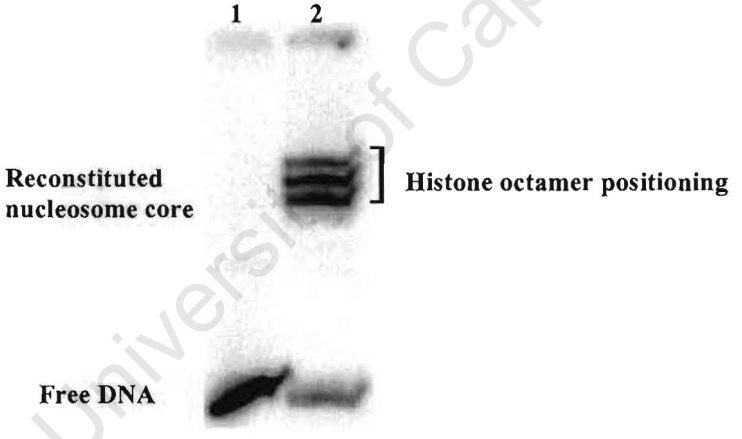


Figure 4.4 Multiple positions of the 207-bp fragment nucleosome core.
Phosphorimage of samples electrophoresed in 15% polyacrylamide gel in 0.5 × TBE. Lane **1**, 80 μℓ of reconstitution reaction containing only buffer and radiolabeled 5S fragment; lane **2**, 80 μℓ of the reconstitution reaction containing H1-stripped long chromatin. The positions of the free DNA and reconstituted nucleosome cores are indicated.

4.2.2.2 Quantification of reconstitution efficiency of nucleosome cores on the biotin-labeled 5S fragment.

To monitor the degree of the incorporation of the 16ddUTP-biotin-labeled 5S fragment into nucleosome cores, the interaction between streptavidin and biotin was utilized in the same manner as with the quantification of the biotin-labeled 5S fragment in the initial labeling reaction (section 4.2.1.3). This control was performed to ensure that biotin-labeled fragments could be reconstituted into the nucleosome as efficiently as radiolabeled fragments.

The reconstituted 5S fragment was electrophoresed in duplicate. Staining of one of these gels with ethidium bromide allowed for the pre-transfer DNA content to be analyzed through visualisation by UV transillumination. The ethidium bromide stained gel (Figure 4.5 A) represented the reconstitution efficiency of the DNA fragment but did not give an indication of the biotin-labeled fragment reconstitution efficiency.

The 5S DNA fragments in the duplicate gel were transferred to a nylon membrane by electrophoretic transfer. Subsequently, the reconstitution was assayed with chemiluminescent detection of the biotin moieties. The Southern blot phosphorimage of the reconstitution is shown in Figure 4.5 B. The fraction of biotin-labeled 5S fragment incorporated into the nucleosomes, as determined by phosphorimage analysis of the chemiluminescent image, indicated that ~73% of the biotin-labeled DNA had been reconstituted. This ratio of shifted:unshifted was similar to that observed for the DNA in the ethidium-stained gel and the ratio obtained with the reconstitution using radio-labeled 5S fragment.

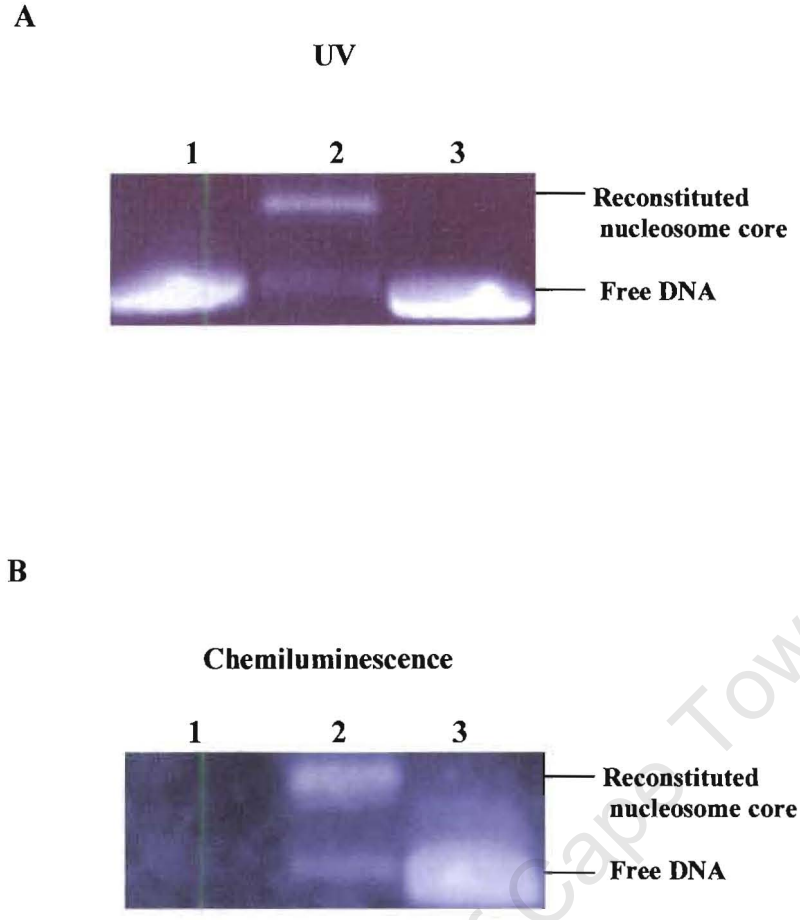


Figure 4.5 Biotin-labeled 5S histone octamer reconstitution.
Panel A; pre-transfer ethidium bromide staining. Equal quantities of biotin-labeled 5S fragment were electrophoresed on a 0.7% (w/v) agarose gel in 0.5 × TBE and stained with ethidium bromide (0.5 µg/ml). Visualisation with short-wavelength transillumination. Lane **1**, 0.4 pmol 5S fragment; lane **2**, 40 µℓ of reconstitution reaction (~0.4 pmol biotin-labeled 5S fragment); lane **3**, 0.4 pmol biotin-labeled 5S fragment. **Panel B;** Southern blot. Nylon membrane with detection of biotin using streptavidin-HRP conjugate, followed by chemiluminescent detection.

4.2.3 Reconstitution of globular domains, GD1 and GD2.

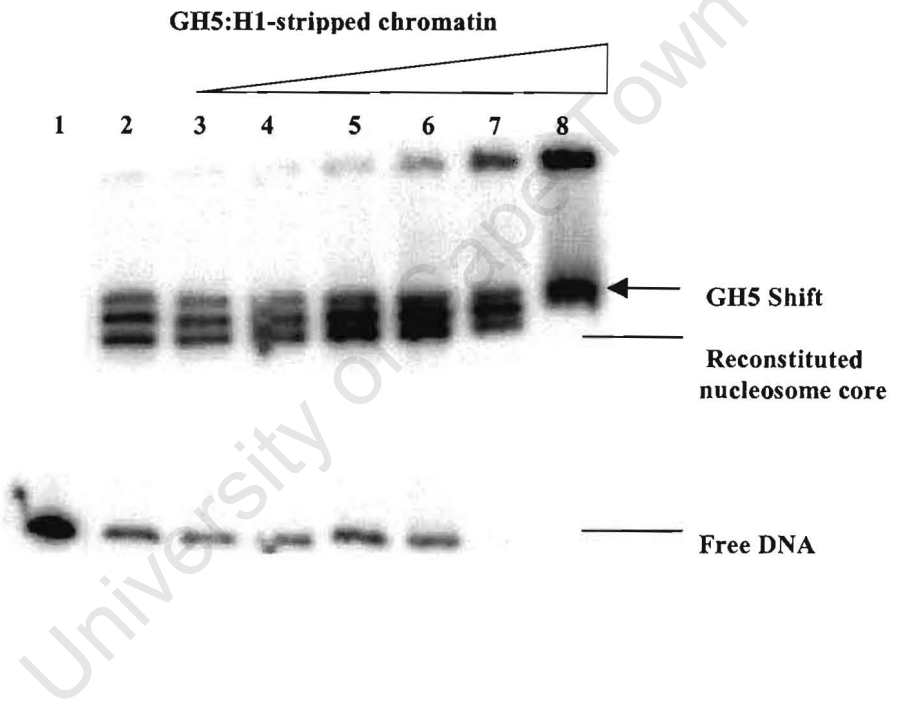
The structural assays would largely depend on the ability of the separate globular domains to associate with the nucleosomal DNA. The ability of the separate globular domains to bind to nucleosomes was determined by reconstitution of the purified GD1 and GD2 with nucleosomes by the high-salt exchange method. GD1 and GD2 were added to the reconstitution mix at the 50 mM NaCl dilution step. The reconstitution of GH5, which is a globular domain known to bind to the nucleosome (van Holde, 1989), was used as positive control in these experiments.

Following the incubation, the reconstitution samples were electrophoresed on native PAGE gels and the phosphorimage of the dried gel analysed. The GH5 reconstitution is shown in Figure 4.6.A. The molar ratios shown are approximate ratios. The titrations were done to determine the optimal input ratio, since we were not sure of the protein concentration which had been determined spectrophotometrically and by direct comparison with protein standards on a Coomassie-stained gel. From lane 5 to lane 8, a significant amount of aggregation is shown. Judging from the intensities of the three nucleosomal bands, it appears possible that GH5 binds more strongly to the central core (middle band). The intensities of the middle and bottom bands decrease with the increasing molar ratio, which suggests that these cores form aggregates at lower GH5 ratios compared to the complex that resolves as the middle band. Most of the free DNA is bound at the ratio of 1:16 (GH5: H1 stripped long chromatin). With the increasing ratio, the three core bands appear to move closer together, which may imply that the structures have become more similar. Alternatively, GH5 could induce re-positioning of the core, thus making the structures more alike. The change in the mobility of these structures may be an indication of GH5 binding.

The same conditions were used for the reconstitutions of both GD1 (Figure 4.6 B and C) and GD2. Although both proteins bound to free DNA (Figure 4.6 B and C, lane 8), neither of the proteins appeared to cause a change in the mobility of the three reconstituted core bands, observed with GH5 (Figure 4.6 A, lane 8).

The affinity of the proposed Hho1p globular domains for nucleosomal DNA might differ from that of GH5. The core structures formed upon reconstitution of GD1 and GD2 might also not be similar. The reconstitution of GD1 and GD2 with nucleosomal templates have previously been shown in another laboratory (Ali, 2000). In that study, a different reconstitution technique was employed, and GD1 reconstitution was only reported at low salt concentrations, which suggested that the interactions and affinities of the two globular domains differed.

A



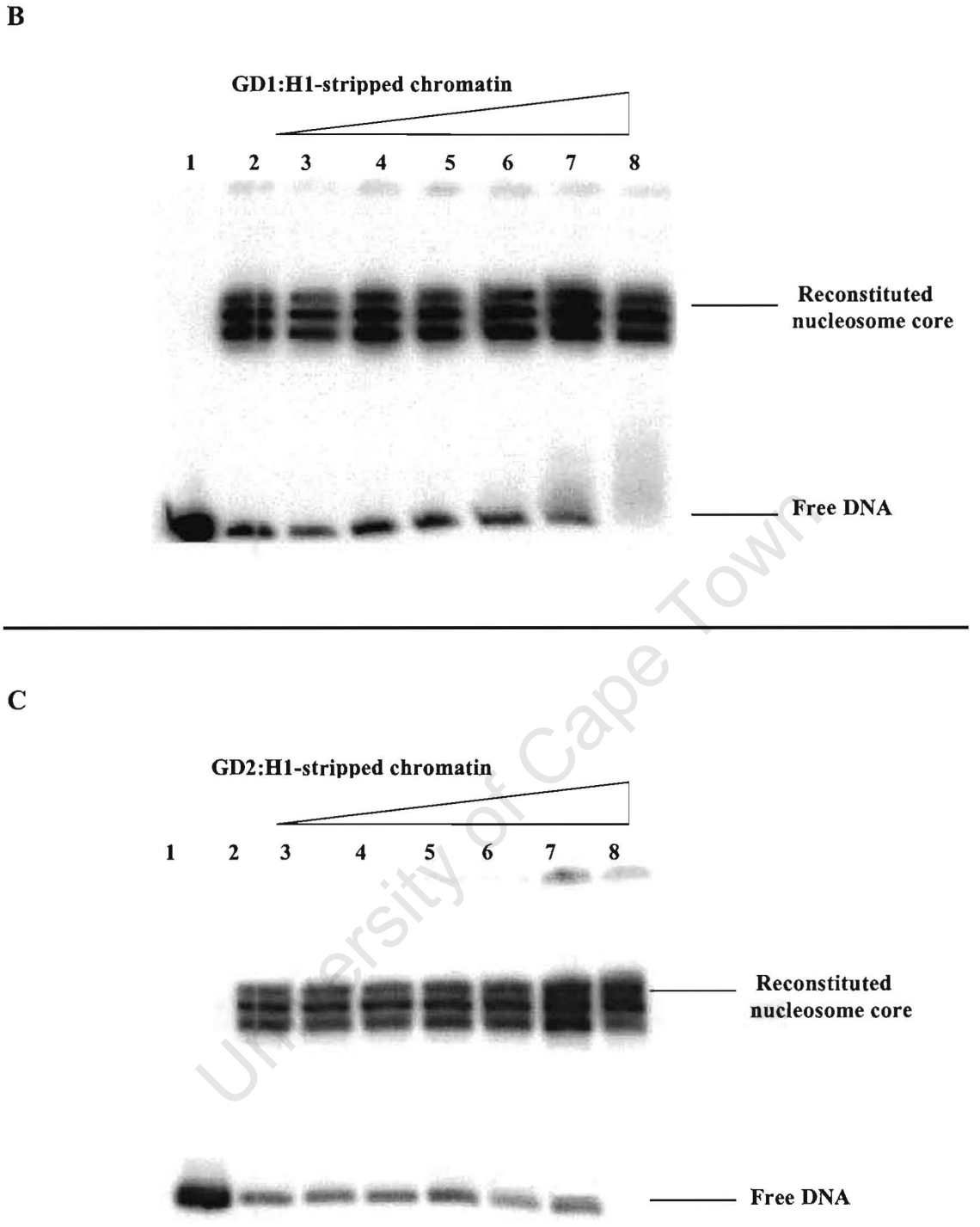


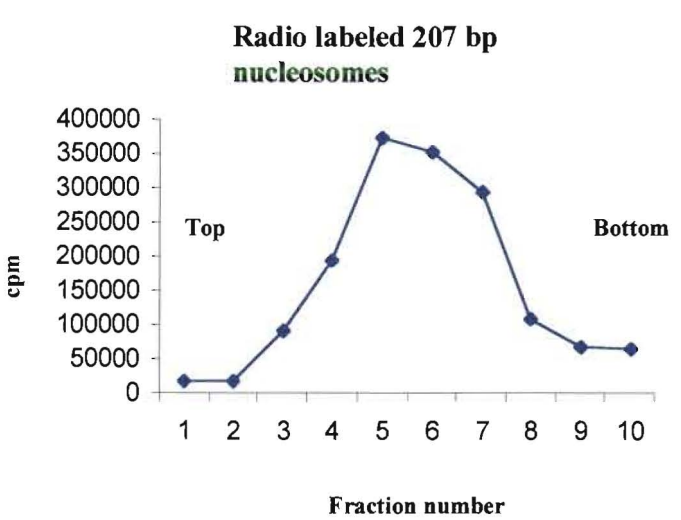
Figure 4.6 Binding of GH5, GD1 and GD2 to a 207-bp 5S DNA nucleosome core.
An aliquot containing 207-bp 5S DNA was reconstituted into nucleosome cores using H1-stripped long chromatin as histone source. The increase in the ratio of input linker histone peptide to total nucleosome (radiolabeled 207-bp core and cold, bulk nucleosome cores) is indicated above the Figure.
Panel A, GH5 reconstitution; **Panel B**, GD1 reconstitution and **Panel C**, GD2 reconstitution. An 80 µℓ aliquot of each reconstitution mix was electrophoresed on a 8% native polyacrylamide gel. A phosphorimage of the gel is shown. **A** and **B**; lane **1**, free 5S radiolabeled fragment; lane **2**, control for the octamer reconstitution containing no recombinant protein; lanes **3-8** increasing approximate molar ratios of linker peptide; 1:1, 1:2, 1:4, 1:8, 1:16, 1:32 molar ratios.

4.2.4 Preparation of mononucleosomes and initial magnetic bead analysis.

4.2.4.1 Sucrose gradients.

In order to perform the magnetic bead assays, labeled, pure mononucleosomes were prepared. This was achieved by separating the various species of the core nucleosome reconstitutions using ultracentrifugation through 5-20% isokinetic sucrose gradients. Fractions collected from these gradients were analysed by scintillation counting (Figure 4.7 A) for the radiolabeled nucleosomes, or by dot blotting (Figure 4.7 B) with chemiluminescent detection for the biotin-labeled nucleosomes. The pooled fractions were dialysed to remove the sucrose and concentrated.

A



B

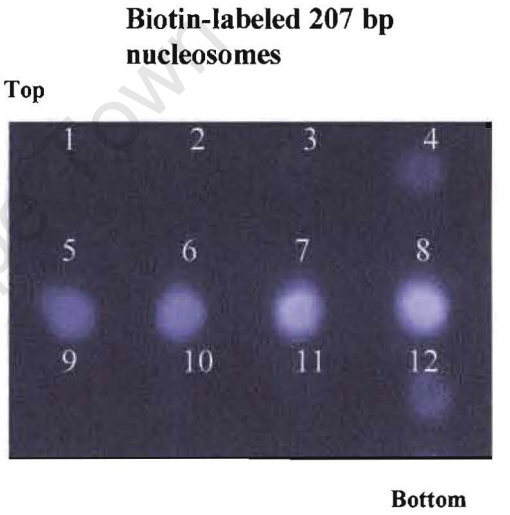


Figure 4.7 Generation of radio- and biotin-labeled mononucleosmes.
Nucleosomal cores were purified from a reconstitution mixture of the separate biotin- and radio-labeled 207-bp reconstitutions by ultracentrifugation and fractions collected from the separate gradients. The distribution of the nucleosome cores on the gradient were determined by scintillation counting in the case of ^{32}P-labeled cores (**A**), fractions 4-7 were pooled; or by chemiluminescent detection of the biotin-labeled cores (**B**), fractions 5-8 were pooled.

4.2.4.2 Non-specific binding to streptavidin magnetic beads.

We tested non-specific binding of the nucleosomes to streptavidin-coated magnetic beads where the 5S fragment was both radio-labeled at the 5' end and biotin-labeled at the 3' end, or only radiolabeled at the 5' end.

Equimolar quantities of each of the nucleosomal species we incubated separately with streptavidin magnetic beads. An aliquot of this incubation was used to determine the total radioactivity. Nucleosomes bound to the streptavidin magnetic beads via the biotin moiety were separated from the unbound nucleosomes by application of the magnetic particle separator. The supernatant from this "pull-down" was collected as the unbound fraction. The magnet was removed and the bond between the biotin and streptavidin magnetic beads broken with guanidinium hydrochloride. This eluent fraction was an indication of the bound fraction of nucleosomes.

The results of the assay are shown in Figure 4.8. In the graph, bars 1 and 2 indicate the total radioactivity for both radio- and biotin-labeled or only radiolabeled nucleosomes, respectively. Bars 3 and 4 indicate the radioactivity in the supernatant after the application of the magnetic bead separator and subsequent pull-down of bound nucleosomes. Bars 5 and 6 indicate the radioactivity in the supernatant after release of the bound nucleosomes from the streptavidin-coated magnetic beads.

From this figure, it is observed that there is virtually no non-specific binding by the nucleosomes lacking the biotin moiety to the streptavidin beads. The radioactivity in the supernatant following the pull-down is much lower in the sample that contained the nucleosomes which were radio- and biotin-labeled as compared to the nucleosome that lacked the biotin moiety. Confirmation of this is shown with the release of the bound nucleosomes, where the radio- and biotin-labeled sample has a much higher reading, indicating binding of the biotin moiety and non-binding of the nucleosomes lacking this moiety.

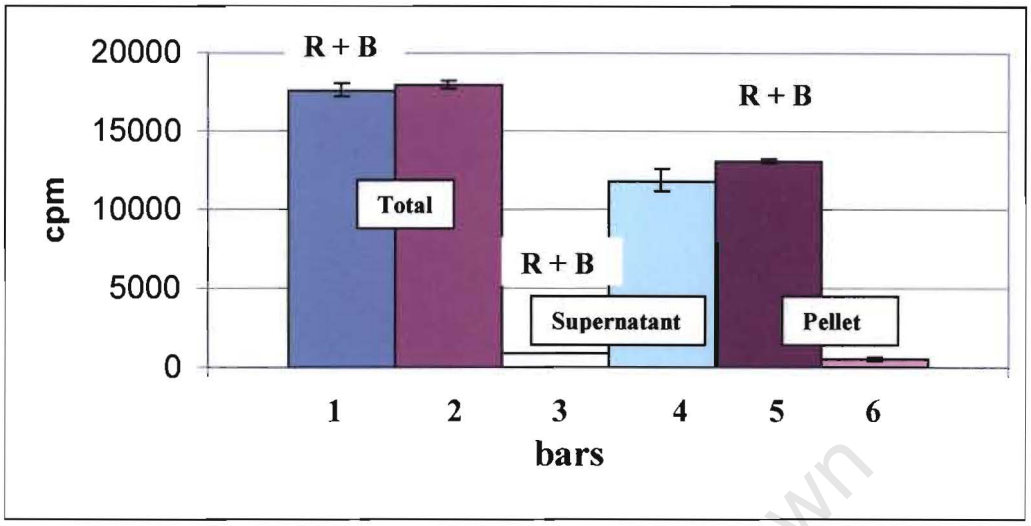


Figure 4.8 Lack of non-specific binding to streptavidin-coated magnetic beads. Radioactivity of the different samples was determined by scintillation counting. Bars **1**, **3** and **5** represent the radioactivity (cpm) present in the aliquots of the nucleosome core samples that were both radiolabeled and biotin-labeled. Bars **2**, **4** and **6** represent aliquots from ^{32}P-labeled nucleosomes lacking the biotin end-label (see section 4.2.4.2). Bars **1** and **2**, represent the total input count (total); bars **3** and **4** represent the cpm present in the supernatant after pull-done via the magnetic bead separator (supernatant); bars **5** and **6**, represent the radio label released from beads after the magnet was removed (pellet).

4.3 Summary.

Having successfully produced the proteins, Hho1p, Hho1-FXap, GD1 and GD2 (Chapter 3), the reagents for the nucleosomal binding and magnetic bead assays are available. H1-stripped long chromatin was isolated in such a way as to produce donor chromatin where the DNA was of suitable length, lacked native linker histones, and the histones are free of detectable degradation. A quantification protocol has been established that showed that the biotin labeling of the 207-bp template DNA was highly efficient.

Initial binding assays with both GD1 and GD2 were unsuccessful, despite the positive control, GH5, showing binding to the nucleosomal cores. The binding assays may be improved by electrophoresis on a polyacrylamide gel, which could allow improved resolution of the species formed by reconstitution with the recombinant proteins.

Further attempts at the optimizing GD1 and GD2 reconstitutions may include the lowering the salt concentration in the reconstitution mixes, especially for GD1, which was shown to bind to H1-stripped chromatin only at low salt concentrations (Ali, 2001). The lack of binding by GD2 could firstly be examined in terms of structure by analyzing the structural integrity of the purified protein by circular dichroism and also repeating the binding assays at different NaCl concentrations.

Chapter 5: Discussion.

5.1 Introduction.

The aim of the study was to develop the reagents that are required for the detailed investigation of the structural functionality of the globular domains of the yeast linker histone Hho1p. The *Saccharomyces cerevisiae* Hho1p has an unique structure, as does the linker histone in *Tetrahymena*, where the central globular domain is absent in the macronuclear H1 of the heterokaryon (Hayashi *et al.*, 1987). Unlike the tri-partite structure of the canonical linker histones, Hho1p contains two homologous globular domains separated by a lysine-rich region and does not contain a C-terminal tail (Landsman, 1996). There are also differences between the chromatin structure of yeast and that of many higher eukaryotes; especially in the nucleosomal repeat length, which in yeast is unusually short at ~165 bp (Thomas & Furber, 1976; Lohr *et al.*, 1977). Yeast histones are also highly acetylated (Davie *et al.*, 1981; Waterberg, 2000).

This implies that the yeast genome may have little need for a linker histone, since one of the functions of the linker histones C-terminal tail is binding to the linker region and facilitating the folding and stabilization of higher order chromatin structures. The yeast linker has no C-terminal tail, and the linker DNA between adjacent nucleosomes is unusually short. Yet, despite this, Hho1p displays properties that are characteristic of linker histones. These include binding to nucleosomes and producing the ~168 bp kinetic pause upon micrococcal nuclease digestion (Patterton *et al.*, 1997).

The assay under development will shed some light on the structural role of this linker histone, and at least indicate if both of the proposed globular domains function like canonical linker histone globular domains.

5.2 Protein over-expression in *E. coli.*

The expressed GD1 and GD2 proteins were required to investigate the DNA binding properties of these proteins. Studying the ability of these globular domains to associate with nucleosomal DNA is essential to our understanding of how a linker

histone containing two putative globular domains can bind to a nucleosome core. However, the two separate globular domains may associate with DNA differently compared to the complete Hho1p protein which includes the lysine-rich linker region.

5.2.1 Expression of linker histones in *E. coli.*

It was previously shown that chicken histone H5 was poorly expressed in *E. coli* (Gerchman *et al.,* 1994), and due to the high sequence homology between the globular domains of yeast Hho1p and chicken H5, there was a concern that similar low expression levels would be observed with the expression of the yeast Hho1p domains in *E. coli*. It was reported (Gerchman *et al.,* 1994) that the addition of a three amino acid N-terminal motif improved the expression level of the globular domain of chicken H5. The sequence encoding this motif was incorporated into the 5' ends of both GD1 and GD2. The induction time-course studies showed that, in contrast to the chicken H5, this modification was not required for the efficient expression of GD1 or GD2 in *E. coli* (data not shown).

5.2.2 Interference of the hexa histidine-tag with nucleosomal binding.

It was a concern that the N-terminal hexa histidine-tag needed for the purification of the expressed proteins could influence the binding of these proteins to DNA. Both GD1 and GD2 were therefore also cloned into *E. coli* expression vectors which expressed a C-terminal histidine-tag that could be removed from the purified protein by protease cleavage.

There was a notable difference in both the expression and purification behavior of the two globular domains. The final protein yield of GD1, subsequent to the first Ni-agarose purification step, was vastly greater than the yield of GD2. This suggested some instability in the case of GD2, since both domains were expressed with identical promoters. It is also possible that the protein conformation of GD2 was less stable than that of GD1, and that GD2 was actively targeted by the degradation machinery.

Initial structural integrity studies were performed by means of circular dichroism measurements (CD) on GD1. These spectra showed the characteristic profile for a predominately α-helical protein, and it would be interesting to see if the spectrum for

GD2 is that of an unfolded, unstructured protein. Such a result will explain the low expression levels and high degree of degradation during the purification steps.

5.3 Feasibility of the magnetic bead assay approach.

The interpretation of the results of the bead "pull-down" assay will require that one clearly define the limitations of the approach. For this purpose appropriate controls had to be designed.

We designed two controls, the first of which was a modified yeast linker histone, which contained a protease site within the peptide region linking the two putative globular domains of yeast Hho1p. The purpose of the cleavable protease site was to test the ability of the radiolabeled nucleosome core, expected to be immobilised on the streptavidin-coated bead if Hho1p cross-bound to two nucleosomes cores, to appear in the supernatant following cleavage of this modified yeast Hho1p. This would be an important control to indicate that an intact H1 protein containing two globular domains was required to bind the radiolabeled nucleosome core to the bead. This control construct was successfully expressed, and purified using the same methods used for the globular domains.

The second control construct will consist of two chicken H5 linker histone globular domains (GH5), known to bind to nucleosomal DNA, which will be connected by a peptide consisting of the lysine-rich linker area of the yeast Hho1p. This control will be important in establishing if the general experimental approach is feasible, and that two linked proteins, known to bind to nucleosomal DNA, are able to cross-bind to two nucleosome cores under the experimental conditions used. The cloning of the GH5-linker-GH5 construct was problematic, probably due to the distally, directly repeated sequence of the GH5 domains. Initial attempts involved recombination-deficient strains of *E. coli*. Work is currently underway to test the stability of expression constructs in additional *E. coli* and in yeast strains.

5.4 Binding assays and magnetic bead technology.

Numerous studies have emphasized the difficulty of obtaining a clear bandshift with a linker histone globular domain. Although some investigators have reported obtaining

atypical "downshifts" (Hamiche *et al.,* 1996), it is generally agreed that the small globular domain does not alter the shape, size or charge of a ternary complex to the extent that produces a clear change in the electrophoretic character of the complex. Globular domain-dependent gel shifts previously reported may be a function of the DNA sequence and reconstituted fragment length.

Although I have shown that incubation of a radiolabeled nucleosome core with chicken GH5 produced a change in the mobility of the nucleosomal structures electrophoresed on a non-denaturing gel, there was no clear single or group of bands that appeared with increasing GH5 concentration that could confidently be ascribed to the binding of GH5. Similarly, no shifted complex was electrophoretically observed following incubation of GD1 or GD2 with radiolabeled nucleosome cores.

The reconstitutions were performed by salt exchange to a final NaCl concentration of 25 mM. The electrophoretic separations were performed in 45 mM Tris acetate (pH 8.3), 1 mM EDTA at room temperature. A similar absence of clear nucleosomal association of GD1 and GD2 under similar ionic strengths was recently reported (Ali, 2001). It therefore seems that an investigation of the binding of isolated GD1 and GD2 to nucleosome cores may require gentle hydrodynamic approaches, where a range of ionic conditions can be investigated. Such a study may be preceded by studying the binding of GD1 and GD2 to synthetic four-way junctions, to which linker histones may bind more tightly, and which has previously been used to investigate linker histone binding (Varga-Weisz *et al.*, 1993).

The initial test of the specificity of the binding of biotin-labeled reconstituted nucleosome cores to streptavidin-coated magnetic beads, demonstrated that the biotin moiety was essential for a detectable "pull-down" signal. This result indicated that very little non-specific interactions between the streptavidin-coated bead and the reconstituted nucleosomes occurred, suggesting that the magnetic bead approach, in principle, is feasible.

5.5 Future Work.

The model being proposed for the binding of the yeast linker histone to the nucleosome involves the simultaneous binding of the globular domains to two

adjacent nucleosomes, with the lysine-rich linker region interacting with the linker DNA connecting these nucleosomes.

An alternative model was proposed (Ali, 2001) where GD1 binds close to the nucleosomal dyad, contacting the DNA both in the vicinity of the dyad axis and at the point where the DNA enters/exits the nucleosome core. GD2 was proposed to bind in a similar fashion to the symmetrically opposed site on the other side of the dyad axis. In this model, there was no interaction proposed between the lysine-rich linker region connecting the globular domains and DNA. It is not clear in this model why the peptide linking the two globular domains should display an amino acid composition typical of that of metazoan H1 C-terminal tails. Crosslinking studies have clearly shown that the C-terminal tail is associated with the internucleosomal linker DNA (Mirzabekov *et al.,* 1989). For this reason, it appears more likely that the lysine-rich region in yeast H1 similarly interacts with the linker DNA.

It is for this reason also unlikely that the yeast linker histone would display simultaneous binding to nucleosomes situated on two adjacent chromatin fibers (interchromosomal binding) as opposed to binding to adjacent nucleosomes (intrachromosomal binding). Interchromosomal binding by the yeast linker histone would leave the lysine-rich region between the globular domains exposed to solvent and unshielded. Intrachromosomal binding (as shown in Figure 4.1) would utilise the lysine-rich region by binding this region to the linker DNA between the adjacent nucleosomes, contributing to the higher-order chromatin structures in the yeast genome.

Recently, Jackson and colleagues (Downs *et al.,* 2003) established a role for the yeast linker in DNA repair in the homologous recombination pathway. It is plausible that the binding of the yeast linker histone to the nucleosome would inhibit various processes that require access to the DNA molecule. The mildly repressive role of Hho1p on homologous recombination is therefore more likely an ancillary effect, and not the central molecular role of this protein in the yeast nucleus. The study by Jackson and co-workers therefore does not impact on our hypothesis for the binding mode of the yeast linker via the globular domains.

Additional work in the immediate future will involve the magnetic bead pull-down assays. It is expected that the results from the completion of the pull-down assay will contribute to our understanding of the binding of yeast H1 in particular, and that of the H1 molecules in higher organisms in general.

References.

Ali, T. (2001) Characterisation of Hho1p, the putative *Saccharomyces cerevisiae.* Department of Biochemistry. University of Cambridge. Cambridge.

Allan,J., Hartman, P.G., Crane-Robinson, C. and Aviles, F.X. (1980) The structure of histone H1 and its location in chromatin. *Nature.* **288,** 675-679.

Allan, J., Cowling, G.J., Harborne,N., Cattini, P., Craigie, R., and Gould, H. (1981) Regulation of the higher-order structure of chromatin by histones H1 and H5. *Journal of.Cellular Biology.* **90**, 279-288

Allan, J., Mitchell, T., Harborne, N., Bohm, L. and Crane-Robinson, C. (1986) Roles of H1 domains in determining higher order chromatin structure and H1 location. *Journal of Molecular Biology.* **187**, 591-601.

An, W., van Holde, K. and Zlatanova, J. (1998) Linker histone protection of chromatosomes reconstituted on 5S rDNA *Xenopus borealis* : a reinvestigation. *Nucleic Acids Research.* **26**, 4042-4046.

Arents, G., Burlingame, R.W., Wang, B.W., Love, W.E. and Moudrianakis, E.N. (1991) The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National.Academy.Science.USA.* **88**, 10148-10152.

Arents, G and Moudrianakis, E.N. (1995) The histone fold – a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proceedings of the National.Academy.Science.USA.* **92**, 11170-11174.

Ausubel, F., Brent., Kingston, R.E., Moore, D.D., Seidman., Smith, J.A and Struhl, K. (1995) Short Protocols in Molecular Biology. Third edition. Jhon Wiley & Sons. Inc.

Aviles, J., Chapman, G.E., Kneale, G.G., Crane-Robinson, C., Bradbury, E.M. (1978) The conformation of histone H5. *European.Journal of.Biochemistry.* **88**, 363-371.

Baneyx, F. (1999) Recombinant protein expression in Escherichia coli. *Current. Opinion in.Biotechnoogy.,* **10,** 411-21.

Bednar, J., Horowitz, R.A., Dubochet, J. and Woodcock, C.L. (1995) Chromatin conformation and salt induced compaction: three dimensional structure information from cryoelectron microscopy. *Journal of. Cellular. Biology.* **131**, 1365-1376.

Bradford, M.M. (1976) A rapid and sensitive method foer the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry.* **72**, 248-254.

Buckle, R.S., Maman, J.D., Allan, J. (1992) Site-directed Mutagenesis studies on the Binding of the Globular Domain of Linker Histone H5 to the Nucleosome. *Journal of.Molecular.Biology.* **223,** 651-659.

Bussey, H., Storms, R.K., Ahmed, A., Albarmann, K., Allen, E., Ansorge, W., Araujo, R., Aparicio, A., Brarrell, B., Badcock, K., Benes, V., Botsiein, D., Chrucher, C., Foster, F., Davies, K., Davies, R.W., Diedrich, F.S., Delus, H., DiPaolo, T., Hani, J., *et al.* (1997) *Nature.* **387**, 303-305.

Carruthers, L.M. and Hansen, J.C. (2000) The core histone N-termini function independently of linker histones during chromatin condensation. *Journal of Bioligical Chemistry.* **275**, 37825-37290.

Cerf, C., Lippens, G., Muydermans, S., Segers, A., Ramakrishnan, V., Wodak, S.J., Hallenga, K. and Wyns, L. (1993) Homo-and Heteronuclear two-dimensional NMR studies of the globular domain of histone H1: Sequential assignment and secondary structure. *Biochemistry.* **32**, 11345-11351.

Certa, U., Colavito-Shepanski, M. and Grunstein, M. (1984) Yeast may not contain H1: the only known "histone H1-like"protein in *Saccharomyces cerevisiae* is a mitochondrial protein. *Nucleic Acids Research.* **12**, 7975-7985.

Clauser, K.R., Baker, P., Burlingame, A.L. (1999) Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical Chemistry.* **71**, 2871-82.

Clore, G.M., Gronenborn, A.M., Nilges, M., Sukumararan, K.D., Zarbock, J. (1987) The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO Journal.* **6**, 1833-1842.

Cole, R.D. (1989) *Methods in Enzymology.* **170**, 524-532.

Clark, D.J., Hill, C.S., Martin, S.R. and Thomas, J.O. (1988) A-Helix in the carboxy terminal domains of histones H1 and H5. *EMBO Journal.* **7**, 69-75.

Clark, D.J., O'Neill, L.P. and Turner, B.M. (1993) Selective use of H4 acetylation sites in the yeast *Saccharomyces cerevisiae. Journal of Biochemistry.* **294**, 557-561.

Crane-Robinson, C. (1997) Where is the globular domain of linker histone located on the nucleosome? *Trends Biochem Sci.* **22**, 75-77.

Davie, J.R., Saunders, C.A., Walsh, J.M. and Weber, S.C. (1981) Histone modifications in yeast *S.cerevisiae. Nucleic Acids Research.* **9**,3205-3216.

DeLange, R.J., Farnbrough, D.M., Smith, E.L. and Bonner, J. (1969) Calf and pea histone IV: complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. *Journal of Biological Chemistry.* **244**, 5669-5679.

Dong, F., Hansen, J.C. and Van Holde, K.E. (1990) DNA and protein determinats of nucleosome positioning on sea urchin 5S rRNA sequences *in vitro*. *Proceedings of the National Academy of Science USA*. **87**, 5724-5728.

Draves, P.H., Lowary, P.T., Widom, J. (1992) Co-operative binding of the globular domain of histone H5 to DNA. *Journal of Molecular Biology*. **225**, 1105-1121.

Faus, A., Luger, K., Tan, S., Richmond, T.J. (1996) Mapping nucleosome position at sigle base-pair resolution by using site-directed hydroxyl radicals. *Proceedings of the National Academy of Science USA*. **93**, 1370-1375.

Finch, J.T., Lutter, L.C., Rhodes, D., Brown, A.S., Rushton, B., Levitt, M., and Klug, A. (1977) Structure of nucleosome core particles of chromatin. *Nature*. **269**, 29-36.

Gerchman, S.E., Graziano, V. and Ramakrishnan, V. (1994) Expression of Chicken linker histone in *E.coli* : Sources of Problems and Methods for Overcoming Some of the Difficulties. *Protein Expression and Purification*. **5**, 242-251.

Goffeau, A., Barrell, B.G., Bussey,H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S.G. (1996) Life with 6000 genes. *Science*. **274**, 563-7.

Gou, D., Mant, C.T., Taneja, A.K., Parker, J.M.R., and Hodges, R.S. (1986) Prediction of peptide retention times in reversed-phase high-performance liquid chromatography.I. Determination of retention coefficients of amino acid residues of model synthetic peptides. *Journal of Chromatography*. **359,** 499.

Goytisola, F.A., Gerchman, S.E., Yu, X., Rees, C., Graziano, V et al. (1996) Identification of two DNA-binding sites on the globular domain of histone H5. *EMBO* Journal. **15**, 3421-3429.

Graziano, V., Gerchman, S.E., Schneider, D.K. and Ramakrishnan, V. (1994) Histone H1 is located in the interior of the chromatin 30-nm filament. *Nature*. **368**, 351-354.

Green, N.M. (1975) Advances in Protein Chemistry, Vol 29. Academic Press, San Diego, 85-133.

Grunstein, M. (1990). Histone function in transcription. *AnnualReview in Cellular Biology*. **6,** 643-678.

Grunstein, M. (1997) Molecular model for telomeric heterochromatin in yeast. *Current Opinion In Cell Biology*, **9**, 383-387.

Guschin, D., Chandler, S. and Wolffe, A.P. (1998). Asymmetric linker histone association directs the asymmetric rearrangement of core histone interactions in a positioned nucleosome containing a thyroid hormone response element. *Biochemistry*. 37, 8629-8636.

Hamiche, A., Schultz, P., Ramakrishnan, V., Oudet, P. and Purnell, A. (1996) Linker Histone-dependent DNA Structure in Linear Mononucleosomes. *Journal of Molecular Biology*. **257**, 30-42.

Hartman, P.G., Chapman, G.E., Moss, T., and Bradbury, E.M. (1977) Studies on the role and mode of operation of the very-lysine- rich histone H1 in eukaryote chromatin. *European Journal of Biochemistry*. **77**, 45-51.

Hayashi, T., Hayashi, H., Iwai, K. (1987) Tetrahymena histone H1. Isolation and amino acid sequence lacking the central hydrophobic domain conserved in other H1 histones. *Journal of Bicohemistry*. (Tokyo) **102**, 369-376.

Hayes, J.J. (1996) Site-directed cleavage by a linker histone-Fe(II) EDTA conjugate:localization of a globular domain binding site within a nucleosome. *Biochemistry*. **35**, 11931-11937.

Hayes, J.J. and Wolffe, A.P. (1993) Preferential and asymmetric interaction of linker histones with 5S DNA in the nucleosome. *Proceedings of the National Academy of Science USA*, **90**: 6415-6419.

Hayes, J.J. and Lee, K.M. (1997) In vitro reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. *Methods in Enzymology.* **12**, 2-9.

Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M. and Grunstein, M. (1995) Histone H3 and H4 N termini interact with SIR3 and SIR4 proteins: A molecular model for the formation of heterochromatin in yeast. *Cell.* **80**, 583-592.

Hereford, L, and Rosbash, M. (1977) Number and distribution of polyadenylated RNA sequences in yeast. *Cel.,* **10**, 453-462.

Hewish, D.R. and Burgoyne, L.A. (1973) Chromatin sub-structure: the digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem.Biophys.Res.Comm.* **52**, 504-510.

Higuchi, R., Krummel, B., and Saiki, R.K. (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Research.* **16**, 7351-7367.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene.* **77**, 51-59

Inoue et al., (1990) *Gene.* **96**, 23-28

Kasinsky, H.E., Lewis, J.D., Dacks, J.B. and Ausio, J. (2001) Origin of H1 linker histones. *FASEb Journal.* **15**, 34-42.

Khochbin, S. (2001) Histone H1 diversity: bridging regulatory signals to linker histone function. *Gene.* **271**, 1-12.

Klug, A., Rhodes, D., Smith, J.T., and Thomas, J.O. (1980) A low resolution structure for the histone core of the nucleosome. *Nature.* **287**, 509-516.

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

Kornberg, RD, and Lorch, Y. (1995) Interplay between chromatin structure and transcription. *Current.Opinion in Cellular Biology.* **7,** 371-375.

Laemmli, U.K.(1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T7. *Nature.* **227**, 680-685.

Lambert, S., Muyldermans, S., Baldwin, J., Kilner, J., Ibel, K. and Wijns, L. (1991). Neutron scattering studies of chroamtosomes. *Bioc. Biophy. Res. Comm.* **179** (2), 810-816.

Landsman, D. (1996) Histone H1 in *Saccharomyces cerevisiae* – a double mystery solved. *Trends in Biochemical Science.* **21**, 287-288.

Leuda, H.S., Zlatanova, J. and van Holde, K. (1994) On the location of linker DNA in the chromatin fiber – studies with immobilized and soluble micrococcal nuclease. *Journal of Molecular Biology.* **235**, 871-880.

Lohr, D., Kovacic, R.T. and Van Holde, K.E. (1977) Quantitative analysis of the digestion of yeast chromatin by staphylococcal nuclease. *Biochemistry.* **16**, 463-471.

Lohr, D., Hereford, L. (1979) Yeast chromatin uniformily digested by DNase I. *Proceedings of the National Academy of Science USA.* **76**,4285-4288.

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 A resolution.*Nature*.**389**, 251-260.

Mant, C.T. and Hodges, R.S (1991) High-Performance Liqouid Chromatography of peptides and proteins: Separation, Analysis, and Conformation. CRC Press, Inc. Florida.

McGhee, J.D. and Felsenfeld, G. (1980) Nucleosome structure. *Annual Review in. Biochemistry.* **49**, 1151-1156.

McGhee, J.D. and Felsenfeld, G. (1983) Another potential artifact in the study of nucleosome phasing by chromatin digestion with micrococcal nuclease. *Cell.* **32**: 1205-1212.

Meersseman, G., Pennings, S., Bradbury, E.M. (1991) Chromatosome positioning on assembled long chromatin. Linker histones affect nucleosome placement on 5S rDNA. *Journal of Molecular Biology.* **211**, 479-491.

Mirabekov, A.D., Pruss, D.V. and Ebralidse, K,K. (1989) Chromatin superstore-dependent Crosslinking with DNA of the histone H5 residues Thr1, His25 and His62. *Journal of Molecular Biology.* **211,** 479-491.

Muyldermans, S. and Travers, A.A. (1994) DNA sequence organization in chromatosomes. *Journal of Molecular Biology.* **235**: 855-870.

Noll, M. (1974) Subunit structure of chromatin. *Nucleic Acids Research.* **1**,1573-1578.

Noll, M. and Kronberg, R.D. (1977) Action of micrococcal nuclease on chromatin and the location of histone H1. *Journal of Molecular Biology,* **109**, 393-404.

Panetta, G., Buttinelli, M., Flaus, A., Richmond, T, J. and Rhodes, D. (1998) Differential nucleosome positioning on *Xenopus* oocyte and somatic 5S RNA genes determines both TFIIIA and H1 binding: a mechanism for selective H1 repression. *Journal of Molecular Biology.* **282**, 683-697.

Patterton, H.G., Landel, C.C., Landsman, C., Peterson, C.L and Simpson, R.T. (1998) The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry.* **273,** 7268-7276.

Patterton, H.G., Graves, S. (2000) DNAssist, a C++ Program for Editing and Analysis of Nucleic Acid and Protein Sequences on PC-Compatible Computers Running Windows 95, 98, NT 4.0 or 2000. *BioTechniques*, **28**, 6-11.

Pennings, S., Meersseman, G., Bradbury, E.M (1994) Linker histones H1 and H5 prevent mobility of positioned nucleosomes. *Proceedings of the National Academy of Science USA.* **91**, 10275-10279.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N. and Wolffe, A.P. (1996) An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres. *Science.* **257**, 614-617.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. and Sweet, R.M. (1993) Crystal structure of globular domain of histone H5 and its implications for the nucleosomal binding. *Nature.* **362**, 219-223.

Rhodes, D. (1985) Structural analysis of a triple complex between ths histone octamer, a Xenopus gene for 5S RNA and transcription factor IIIA. *EMBO Journal.* **4**, 3473-3482.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. (1984) Structure of the nucleosome core particle at 7A resolution. *Nature.* **311,** 532-537.

Roth, S.Y. and Allis C.D. (1992) Chromatin condensation: Does histone H1 dephosphorylation play a role? *Trends Biochem. Sci*, **17**, 93-98.

Simpson, R.T. (1978) Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry.* **17** (25), 5524-5531.

Simpson, R.T., Stafford, D.W. (1983) Structural features of a phased nucleosome core particle. *Proceedings of the National Academy of Science USA.* **80**, 51-55.

Smith, B.J., Harris, M.R., Sigournay, C.M., Mayes, E.L., Bustin, M. (1984) A survey of H1o-and H5-like protein structure and distribution in higher and lower eukaryotes. *European Journal of Biochemistry..* **138**, 309-17.

Srebreva, L., Zlatanova, J., Miloshev, G., Tsanev, R. (1987) Immunological evidence for the existence of H1-like histone in yeast. *European Journal of Biochemistry.* 165, 449-54.

Staynov, D.Z. and Crane-Robinson, C. (1988) Footprinting of linker histones H1 and H5 on the nucleosome. *EMBO Journal.* **7**, 3685-3691.

Stein, A. and Mitchell, M. (1988) Generation of different nucleosome spacing periodicities in vitro :Possible origin of cell type specificity. *Journal of Molecular Biology.* **203**, 1029-1043.

Strahl, B.D., Allis, C.D. (2000) The language of covalent histone modifications. *Nature.* **403**, 41-5.

Sulkowski, E. (1989) The saga of IMAC and MIT. *Bioessays.* **10**, 170-5.

Sun, J.M., Wiaderkiewicz, R and Ruiz-Carrillo, A.. (1989) Histone H5 in the control of DNA synthesis and cell proliferation. *Science.***245**, 68-71.

Thoma, F., Koller, T., Klug, A. (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *Journal of Cellular Biology.* **83**, 203-227.

Thoma, F., Losa, R. and Koller, T. (1983) Involvement of the domains of histone H1 and H5 in the structural organization of soluble chromatin. *Journal of Molecular Biology.* **167**, 619-640.

Thomas, J.O. and Kornberg, R.D. (1975) An octamer of histones in chromatin free in solution. *Proceedings of the National Academy of Science USA.* **72**, 2626-2630.

Thomas, J.O and Furber, V. (1976) Yeast chromatin structure. *FEBS Letters.* **66**, 274-280.

Thomas, J.O.(1984) The higher order structure of chromatin and histone H1. *J.Cell. Sci. Suppl.*,**1**, 1-20.

Thomas, J.O.and Wilson, C.M. (1986) Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5. *EMBO Journa.*, **5**, 3531-3537.

Thomas, J.O., Rees, C.and Finch, J.T. (1992) Co-operative binding of the globular domains of histone H1 and H5 to DNA. *Nucliec Acids Research.* **20**, 187-194

Travers, A.A. and Muyldermans, S.V. (1996) A DNA sequence for positioning chromatosomes. *Journal of Molecular Biology.* **257**, 486-491.

Travers, A. (1999) The location of the linker histone on the nucleosome. *TIBS.* **24**, 4-7.

Ura, K., Haye, J.J.and Wolffe, A.P. (1995) A positive role for nucleosome mobility in the transcription activity of chromatin: restriction by linker histones. *EMBO* Journal. **14**, 3752-3765.

Ushinsky, S.C., Bussey, H., Ahmed, A.A., Wang, Y., Williams, B.A.and Storms, R.K. (1997) Histone H1 in Saccharomyces cerevisiae. *Yeast.* **13**, 151-161.

Van Holde, K.E. (1989) Chromatin, 1st edition, A. Rich, ed. (new York: Springer-Verlag).

Varga-Weisz, P., Van Holde, K. and Zlatanova, J. (1993) Preferential binding of H1 to fourway-junction DNA. *Journal Biology Chemistry.* **268**, 20699-20700.

Waterberg, J.H. (2000) Steady-state levels of histone acetylation in *Saccharomyces cerevisiae. Journal of Biological Chemistry.* **275**, 13002-13011.

Wolffe, A.P. (1994) Transcription: in tune with the histones. *Cell.***77**, 13-16.

Wolffe, A. (1995) Chromatin structure and function. Second edition. Academic Press. London.

Wolffe, A.P., Khochbin, S. and Dimitrov, S. (1996) What do linker histones do in chromatin? *BioEssays.* **19**(3), 249-255.

Woodcock, C.L., Frado, L.L. and Rattner, J.B. (1984) The higher-order structure of chromatin : evidence for a helical ribbon arrangement. *Jouranl of Cellular Biology* **99**, 42-52.

Woodcock, C.L. and Dimitrov, S. (2001) Higher-order structure of chromatin and chromosomes. *Current Opinion in Genetic Development.* **11**, 130-5.

Wu, M., Allis, L.D., Richman, R., Cook, R.G. and Govorsky, M.A. (1986) An interesting sequence in an unusual histone H1 gene of *Tetrahymena thermophilia. Proceedings of the National Academy of Science USA.* **83**, 8674-8678.

Yaneva, J., Zlatanova, J., Paneva, E., Srebreva, L. and Tsanev, R. (1990) Interaction of histones H1 and H1 (0) with superhelical and linear DNA. *FEBS Letters.* **263(2),** 225-228.

Zentgraf, H. and Franke, W.W. (1984) Differences of supra nucleosomal organization in different kinds of chromatin : cell type-specific globular subunits containing different numbers of nucleosomes. *Journal of Cellular Biology.* **99**, 272-286.

Zhou, Y-B., Gerchman, S.E., Ramamkrishnan, V., Travers, A.and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature.* 395, 402-405.

Zlatanova, J. (1990). Histone H1 and the regulation of transcription of eukaryotic genes. *TrendsBiochem Science.* 15, 273-276.

Zlatanova, J. and Yaneva, J. (1991). DNA sequence specific interactions of histone H1. *Mol.Biol.Rep.* 15(1), 53-56.

Zlatanova, J. and Van Holde, K. (1996) *Progress in Nucelic Acid research and Molecular Biology*. **52**, 217-259.