This is to certify that the

dissertation entitled

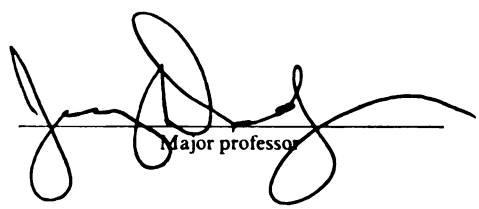
Ectopic Expression of Chicken
HMG14A and HMG17 Chromatin
Binding Proteins

presented by

Natalie Brown

has been accepted towards fulfillment
of the requirements for

PH.D. degree in Microbiology



Major professor

Date 2-17-95

ECTOPIC EXPRESSION OF CHICKEN HMG14A AND HMG17 CHROMATIN BINDING PROTEINS

By

Natalie S. Brown

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Microbiology

1995

ABSTRACT

ECTOPIC EXPRESSION OF CHICKEN
HMG14a AND HMG17 CHROMATIN
BINDING PROTEINS

By

Natalie S. Brown

The high mobility group (HMG) chromatin binding proteins HMG14 and HMG17 are abundant and highly conserved eukaryotic proteins. Although their function remains unknown, these proteins are known to bind to nucleosomes preferentially in regions of actively transcribing chromatin. To better understand these proteins and the role they play within the cell, we have ectopically expressed wild type and mutant chicken HMG14a and HMG17 cDNAs in QT6 (quail fibroblast cell line) cells.

Wild type and mutant cDNAs were introduced into several eukaryotic expression vectors and transfected, transiently and stably, into the QT6 cell line. Stable expression of the wild type HMG17 cDNA resulted in 3 to 5 fold increases in the level of detectable protein. The overproduction of this protein did not effect the cell viability, growth rate, or cause any gross morphological changes. Chromatin structural studies of the cell lines overexpressing HMG17 demonstrated moderately increased binding of HMG17 to individual nucleosomes yet there was no effect seen on the gross chromatin structure.

The production of wild type and mutant HMG14a fusion proteins containing an immunologically distinct C-terminus allowed us to distinguish between the endogenous and exogenously expressed proteins. Using a tetracycline regulated expression system designed to increase the amount of exogenous protein expressed in cells, we transiently expressed mutant, wild type, and fusion HMG14a clones in QT6 cells. We demonstrated that the cells can express these constructs with no difference seen in transfection efficiency, growth characteristics, or level of expression between the mutants and the wild type.

Using immunofluorescent confocal microscopy, we examined the *in-situ* location of exogenous wild type and mutant HMG14a fusion proteins in our transfected cell lines. We demonstrated that these proteins do localize exclusively to the nucleus. The staining patterns seen in these cell lines suggest the HMG14a proteins are in fact binding to the dispersed chromatin of the interphase cell as well as the condensed chromosomes in dividing cells. There was no detectable difference seen between the wild type and mutant proteins used in these studies.

Dedications

To Bill and Manitou

ACKNOWLEDGMENTS

I am extremely grateful to my advisor, Dr. Jerry Dodgson for his advice, help, support and patience during my graduate work. I feel it a privilege to have worked with him and know I have benefitted greatly from having him as a role model. I am especially grateful to the members of my committee, Drs. Richard Schwartz, Michelle Fluck, Larry Snyder and Birgit Zipser for their guidance and encouragement throughout the years.

I would like to thank the many people who have helped me with their friendship and expertise including my fellow labmates, Huei Min Lin, Yi Li, Kyoung Eun Kim, Don Salter, Bill Payne, Lyman Crittenden, and Wynne Lewis, Dr. Lee Velicer, who always had an encouraging word, Dr. Henry Hunt, who greatly helped me with research advice, and my fellow graduate students, especially Larry Martin, Nicco Yu, Sue Dahger, Michelle Anderson, and Brenda Kenney.

I feel it was a great privilege to work among the excellent faculty and support staff of the Department of Microbiology and wish them all continued success.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

APPENDIX II

INTRODUCTION

The high mobility group (HMG) chromatin binding proteins HMG14 and HMG17 are abundant and highly conserved eukaryotic nuclear proteins. Their high level of evolutionary conservation and wide distribution suggest they play an important role in chromatin structure. Although the exact nature of their cellular function remains uncertain, these proteins are known to bind to nucleosomes with specific stoichiometery. A variety of *in-vitro* experiments suggest that HMG14 and HMG17 are preferentially associated with regions of active or potentially active chromatin. The apparent exclusive association of these HMG proteins with transcribable sequences suggests that they may be involved in some aspect of transcriptional regulation, yet they have been found not to function directly as transcriptional activators. Much of the investigation of these proteins has focused on their physical structure and direct interactions with nucleosomes. With the recent availability of cloned genes for the chicken HMG14 and HMG17 and suitable expression vectors, we chose to study the expression of these proteins in a tissue culture system.

In this thesis, we demonstrate stable and transient expression of wild type, mutant, and fusion chicken HMG14a and HMG17 in a

quail fibroblast cell line. In stable transfections we have elevated the level of HMG17 to 3-5 times that of the normal cell. This was achieved with little or no effect on the cell phenotype, growth characteristics, or chromatin structure, suggesting these cells can tolerate at least this much excess protein. We have also shown that these proteins localize to the nucleus with *in-situ* immunofluorescent studies and appear to bind to the dispersed chromatin of the interphase cell.

Chapter 1

Literature Review: The HMG chromosomal proteins and chromatin structure

Chromatin Structure and Active Genes

The complex of DNA and protein in the nucleus of eukaryotic cells is termed chromatin. The dynamic structure of chromatin is evident in the mitotic processes of chromosome condensation and relaxation and the cellular mechanistics of DNA replication and transcription.

The basic repeating unit of chromatin is the nucleosome. The core nucleosome is composed of a 146 base pair duplex DNA wrapped twice around around an octamer of histone proteins. This octamer is composed of two each of the four core histones, H2A, H2B, H3, and H4. The nucleosomes are linked together by the DNA duplex and this linker DNA varies in length, producing the heterogeneity seen among nucleosomes from different cells and at different times in the developmental process. A fifth histone, H1, is associated with this linker DNA and is known to bind to the DNA strand as it enters and exits the core nucleosome.

The nucleosome is also involved in the higher orders of structure that compress the genomic DNA into the compact form found in the nucleus. The nucleosomal fiber, which appears as "beads-on-a-string" under an electron microscope and has a diameter of 10 nm is coiled into a 30 nm fiber helped by the addition of histone H1 which acts as a crosslinker between nucleosomes. This 30 nm fiber is then subject to further levels of condensation depending on the mitotic state of the cell, among other things, which is brought about by the

binding of specific nonhistone proteins which form the scaffold of the chromosome. An illustration of the various levels of chromatin condensation is shown in figure 1. The condensed structure of chromatin is neccessary to compact the large amount of DNA present in the nucleus. The human genome of 3×10^9 base pairs would extend over one meter if completely extended, yet is packaged into the nucleus with an average diameter of 10 um.

The dynamic state of chromatin is evidenced by the process of gene transcription. Highly condensed chromatin is too compact to allow the cellular transcriptional machinery to access the DNA. Therefore, the regions of DNA which are actively being transcribed or are in a potentially active state, approximated to be 10 - 20 % of the total genome, must be in a more accessible or less compact structure. This decondensed structure of active genes can be confirmed experimentally with the use of non-specific endonucleases such as DNaseI. It has been shown that DNaseI will preferentially digest active or potentially active gene sequences and that this increased sensitivity is a result of changes in chromatin and nucleosomal structure specific to these regions in the chromatin (1). These changes in the structural organization of active genes can also be seen by digesting chromatin with micrococcal nuclease which produces characteristic patterns of digestion that distinguishs bulk chromatin from the coding regions of

Figure 1 - **Progressing orders of chromatin condensation**

An illustration of the progressing orders of chromatin condensation giving rise to the highly condensed metaphase chromosome (83).

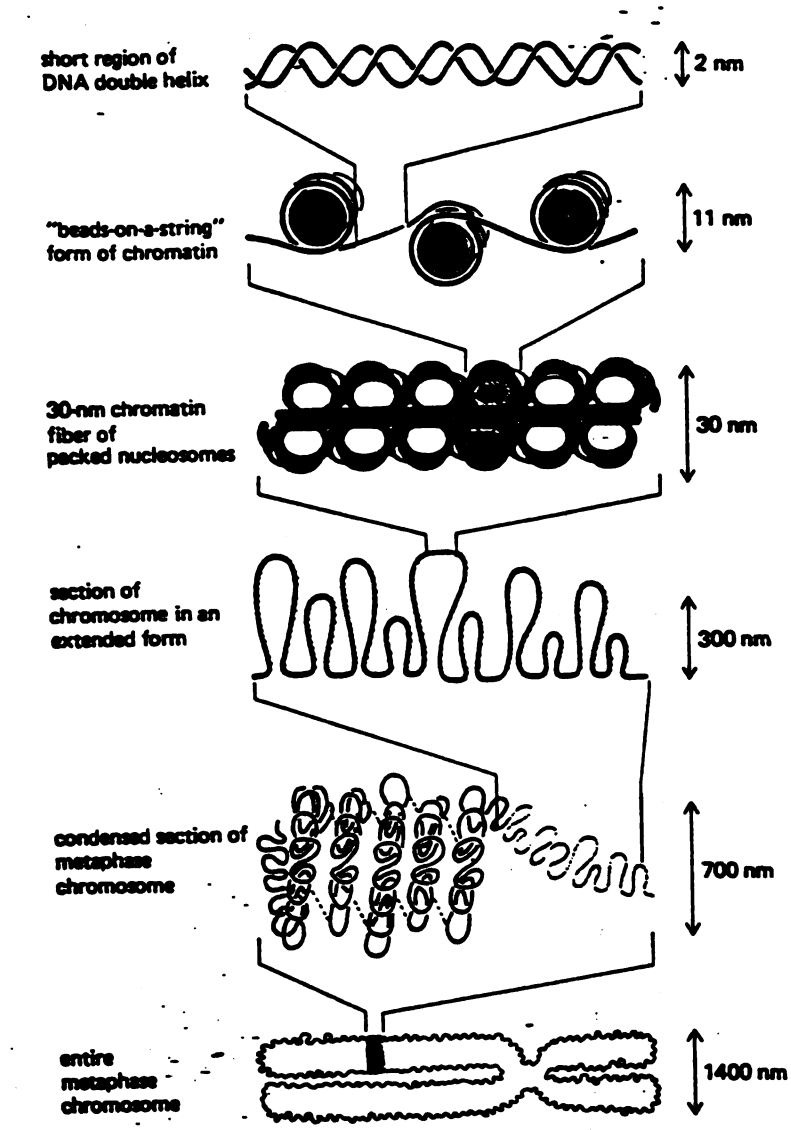

short region of DNA double helix
2 nm
"beads-on-a-string" form of chromatin
11 nm
30-nm chromatin fiber of packed nucleosomes
30 nm
section of chromosome in an extended form
300 nm
condensed section of metaphase chromosome
700 nm
entire metaphase chromosome
1400 nm

Figure 1

active genes (2). An illustration of the digestion of chromatin with DNaseI is shown in figure 2. The changes in chromatin structure associated with active genes are thus illustrated by the increased susceptability of these regions to nucleases and are due to a decondensing of the chromatin at these locations. This allows the nuclease molecules increased physical access to the DNA. These changes are coincident with modification of the histone proteins, possible substitution of the histones with variants, loss of histone H1, and the presence of nonhistone chromatin binding proteins such as the high mobility group (HMG) proteins on the nucleosomes.

Analysis of the core histone proteins found in chromatin fractions enriched for active sequences shows these proteins to be highly acetylated on lysines near their amino terminus. Although this modification produces subtle changes in the nucleosome conformation (5), it does not appear to be sufficient to generate the active chromatin conformation or to increase the rate of transcription (3,4). Acetylation at the histone tails may reduce the helical periodicity of the linker DNA and by loosening interactions at the periphery of the nucleosome, help to dissociate the histone H1, which is found in reduced amounts in active chromatin (6).

The two small, basic chromatin binding proteins, HMG14 and HMG17, are the most abundant nonhistone chromatin binding proteins found in the nucleus of higher eukaryotes and have

Figure 2 - **Digestion of chromatin with endonuclease**

An illustration of the digestion of chromatin with pancreatic DNaseI (83). Endonucleases such as DNase I cut chromatin first at nuclease hypersensitive sites characteristic to a particular gene. This is followed by selective degredation of the DNA sequence of actively transcribing and potentially active genes. The distinct "open" conformation in these regions of chromatin allows for this selective digestion. HMG14/17 are selectively associated with these regions of active and potentially active chromatin.

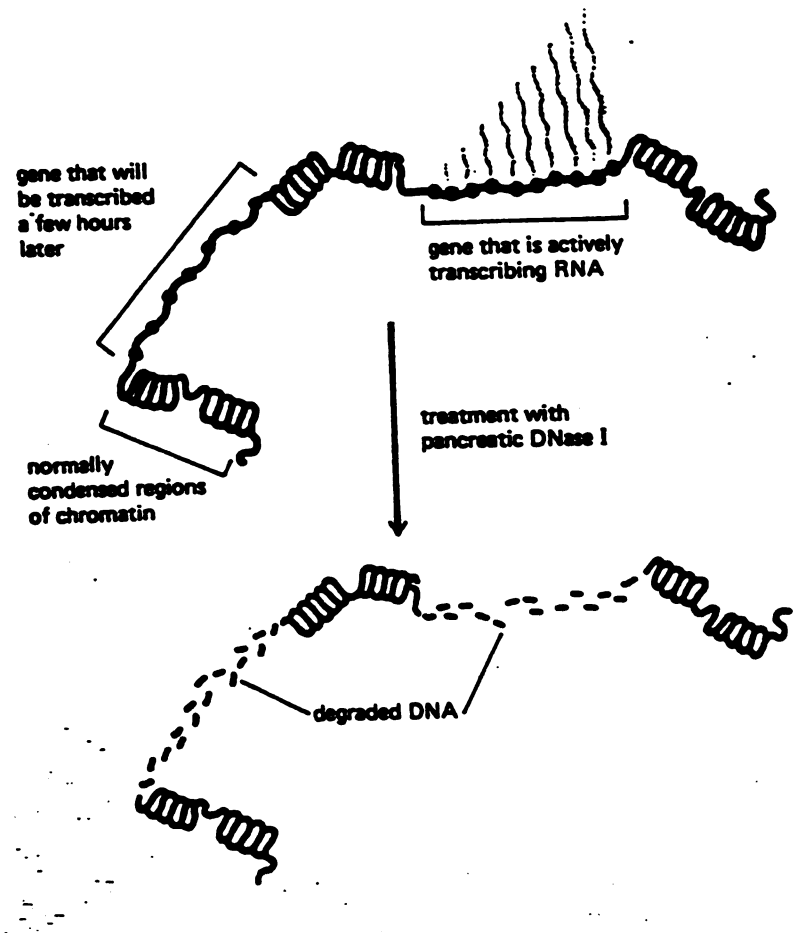

gene that will be transcribed a few hours later
gene that is actively transcribing RNA
treatment with pancreatic DNase I
normally condensed regions of chromatin
degraded DNA

Figure 2

been directly associated with active chromatin. Weisbrod et al showed that the DNaseI sensitivity of active gene sequences is lost upon extraction of the chromatin with 0.35M NaCl, which releases a fraction containing the HMG14 and HMG17 proteins. Upon reconstition with the 0.35 M NaCl fraction or with a subfraction highly enriched for HMG14 and HMG17, this sensitivity is restored (7,8,9). Others have shown that mononucleosome fractions enriched for active sequences are also enriched for HMG14 and HMG17 (12). Numerous other lines of evidence support the association of these HMG proteins with regions of active chromatin. HMG14 and HMG17 proteins immobilized on agarose columns can selectively bind to and retain actively transcribed nucleosomes (10). HMG17-specific antibodies have been used to selectively isolate oligonucleosomes from active chromatin (11), and microinjection of antibodies to HMG17 in somatic cells has been shown to inhibit transcription (13). Although the correlation between active gene structures and the HMG14 and HMG17 proteins appears well established, the actual function of these proteins and the role they play in the generation and/or maintanance of active chromatin remains unknown.

The HMG chromosomal proteins

The high mobility group (HMG) chromatin binding proteins are the most abundant and ubiquitous nonhistone proteins found in the nucleus of eukaryotic cells. These proteins are about one tenth as abundant as the histone proteins at approximately 10^6 molecules per cell. The HMG proteins have been highly conserved throughout evolution suggesting an important role in cellular function. Despite extensive investigation, their roles remain virtually unknown.

The HMG proteins were first identified by Johns and co -workers in their attempts to isolate histone H1 from calf thymus chromatin. In addition to histone H1, the nuclear preps contained a number of additional chromatin proteins with distinct charge and size properties. It was found that these nonhistone proteins could be removed from chromatin with 0.35 M NaCl extraction and further purified as a 2% TCA (trichloroacetic acid) soluble fraction (14). When this 0.35 M NaCl extractable, 2% TCA soluble fraction was run on polyacrylamide gels, up to 16 bands were identified, all of which ran with a high mobility due to their small size and highly charged nature. This group of proteins was thus named the high mobility group proteins and designated HMG1, HMG2 and so on according to their position in the gels. Later, it was determined that most of these bands were

proteolytic degradation products. HMG1, HMG2, HMG14, and HMG17 are now recognized as the major HMG proteins in vertebrate tissue, with HMGI and HMGY found in at least some species. As other groups continued to find these proteins in other organisms, it became apparent that they constituted a family of abundant nuclear proteins with distinct properties and a probable important role in cellular function.

The working definition of the HMG family of proteins was initially derived from the properties of the calf thymus proteins, these properties for the HMG proteins are: (14)

1. chromatin proteins
2. extractable from chromatin in 0.35 M NaCl
3. not precipitated from the above extract in 2% TCA
4. high in basic amino acids, approximately 25% or more
5. high in acidic amino acids, approximately 20-30%
6. relatively high in proline, 7% or more
7. soluble in 5% perchloric acid

This definition is somewhat problematic today as proteins isolated from lower eukaryotes, which are probably functionally related, may be excluded based upon differences in one or more of these properties. Isolation and characterization of the cDNAs for the HMG proteins from a variety of organisms has allowed this defintion to be re-evaluated and now HMG proteins are classified based primarily upon sequence similarities as well as their

physical properties.

HMG-like proteins from lower eukaryotes

HMG-like proteins have been isolated from a number of lower eukaryotic species (15). These include *Drosophila melanogaster, Ceratitis capitata, Saccharomyces cerevisiae, Neurospora crassa, Dictyostelium discoideum, Physarum polycephalum, Aspergillus nidulans, Tetrahymena pyriformis,* and *Tetrahymena thermophila*. The HMG-like proteins isolated from these organisms have physical characteristics similar to the HMGs of vertibrates such as, solubility in 5% perchloric acid, gel mobility and amino acid composition but few have been analyzed further. Protein sequence determination of the HMG-like proteins NHP6a, NHP6b and ACP2 found in *Saccharomyces cerevisiae* show that they are homologues to HMGI, yet lack some characteristic internal repeats found in these proteins. Deletion of the ACP2 locus from *S. cerevisiae* has been shown to be haploid lethal (15). The LGI protein from *T. thermophila* has also been shown to have partial homolgy to HMGI and is thought to share a common ancestor with the mammilian HMGI protein. Little else is known of these proteins and their function remains obscure.

The HMGI/Y family

The HMGI family is composed of HMGI and it's isoform HMGY.

These proteins are not detectable in terminally differentiated, non -proliferating cells but are enriched in rapidly growing, transformed or malignant tissues (15). These proteins resemble the HMG14 and HMG17 proteins and although very little is known of their function, it is thought they may replace HMG14 and HMG17 and influence cell growth and/or gene regulation. The full length cDNAs have been isolated for human HMGY and HMGI and for HMGY in the mouse (15). The two proteins are isoforms resulting from differential processing of the RNA. The HMGI and HMGY mRNAs are expressed preferentially in malignant, non-differentiated and neoplastic tissues. Induction of differentiation in murine teratocarcinoma cells results in down-regulation of mRNA expression (15). Thus it appears that the level of HMGI and HMGY expression is coupled to the proliferative state of the cell, yet it is not known whether this regulation is a consequence of or a prerequisite for differentiation.

The HMG1 and HMG2 family

The HMG1 and HMG2 proteins constitute the most abundant HMG family, at approximately one molecule per 3000 base pairs of DNA. HMG1 and HMG2 have molecular weights of 27 and 28 Kd respectively. These proteins can bind both single and double stranded DNA yet show a preference for binding to single stranded DNA (15). In addition, they show a preference for

binding cruciform DNA *in-vitro* though the presence of this form of DNA *in-vivo* is debatable. Experimental results on their cellular role are somewhat conflicting and their function remains unknown. Several lines of evidence suggest they may play a role in chromosomal replication (15). Antibodies to HMG1 and HMG2 microinjected into nuclei inhibit DNA synthesis and the proteins can stimulate DNA polymerases (15). Antibodies to HMG1 were shown not to inhibit RNA synthesis by RNA polymerase II in somatic cells, yet microinjection into amphibian oocyte nuclei showed retraction of the transcription loops, suggesting that these proteins stimulate transcription (17). These proteins have also been shown to increase the rate of binding of some transcription factors to their DNA recognition sequences (15). In contrast to HMG14 and HMG17, these proteins do not bind selectively to regions of active chromatin. Recently, it has been observed that HMG1 and HMG2 can bend DNA efficiently into small circles and may facilitate cooperative interactions between cis-acting factors by promoting DNA flexibility. The ability to promote highly compact forms of DNA may indicate a general role for these HMG proteins in chromatin structure (19).

HMG1 and HMG2 proteins have been isolated and sequenced from a variety of sources including rat, bovine, human, pig, and Chinese hamster ovary cells. Alignment of the various

HMG1 protein sequences show them to be 97-99% homologous to the human proteins. HMG1 is 214 amino acids in length and can be divided into three domains (15). The first two N-terminal regions, up to amino acid 164, constitute an internal repeat separated by 10 amino acids from residue 80-89. These domains are thought to have a globular structure and are rich in charged amino acids with a net positive charge of +20. This region is most likely involved in binding to the DNA. The C -terminal portion of the protein is unstructured with the last 29 residues being highly acidic (15). This region may be involved with binding to positively charged chromatin proteins such as the histones. The distibution of charges along the HMG1 molecule is thus highly asymmetric. A number of unrelated eukaryotic DNA binding proteins contain regions homologous to the N-terminal repeats found in HMG1. This motif has been named the HMG box and is involved in the DNA binding ability of these various proteins.

cDNA clones have been isolated for HMG1 from several tissues from various species including human, bovine, rat, pig and Chinese hamster ovary cells (CHO). No cDNAs for HMG2 have yet been isolated. All mammalian cells contain three mRNA species for HMG1 as detected by northern analysis. The sizes of these transcripts are 2.4 kb, 1.4 kb and 1 kb, with the 2.4 kb band being the predominant species (15). This heterogeneity may be due to differential processing of a

primary transcript. One of these species may be the transcript for HMG2 as the primary sequence of the proteins predicts similarities in the transcript sequences. The various cDNAs for HMG1 are highly conserved showing a 93% sequence identity among them in the open reading frame (15). Portions of the 3' untranslated sequence show very high levels of homology with the first 300 bp homologous at about the 90% level (15). This suggests that this region may be important for message processing. Southern analysis suggests there may be multiple copies of HMG1 and HMG2 genes in the genome of the cells and tissues examined. It is probable some of these represent psuedogenes. No genomic clone has yet been isolated for HMG1 or HMG2.

The level and cellular distribution of HMG1 and HMG2 proteins have been investigated in mammilian tissues (17). These proteins are present in the nucleus and the cytoplasm and the subcellular distribution as well as the amount of protein are highly tissue specific. In general, in slow growing or differentiated cells, HMG1 and HMG2 are found predominantly in the cytoplasm. In actively dividing cells, these proteins are found in the nucleus. In addition, the levels of HMG1 and HMG2 are inversely correlated with the level of histone H1° (17). These results are consistant with a role of HMG1 and HMG2 in replication or transcription.

The HMG14 and HMG17 family

HMG 14 and HMG17 are the smallest HMG proteins with molecular weights of 14 and 11 kd. They are the only known DNA binding proteins with a higher affinity for nucleosomes than naked DNA. The HMG14 and HMG17 family of proteins are the best characterized HMGs and, although their physical properties and specific interactions with the nucleosome are well understood, the function of these proteins remains controversial.

Characteristics of the HMG14 and HMG17 proteins

HMG14 and HMG17 proteins have been isolated and analyzed from a variety of species including human, calf, mouse, and chicken. Trout chromatin protein H6 is closely related to HMG14 and HMG17 and is considered a member of this family. Analysis of the protein sequences of these HMG proteins shows them to be highly conserved with stretches of amino acids remaining invariant in all species tested. The HMG17 proteins appear to have evolved slower than HMG14 since their sequence is more highly conserved. The sequence homolgy among the various HMG17 proteins is over 91%, while the sequence homolgy among the HMG14 proteins ranges from 49-94% (15). HMG14 and HMG17 share several highly conserved elements and have over 30% of their amino acid sequence in common; yet they are quite distinct protein families, the

similarity between any members of the two groups being less than 52% (15). Trout H6 is approximately 65% similar to the HMG17 proteins and 55% similar to the HMG14 proteins (15). The major chicken HMG14 protein, called HMG14a, is also intermediate in similarity between the the HMG14 and HMG17 subfamilies. The high level of conservation seen among the HMG proteins, representing an evolutionary span of more than 400 million years from trout to human, suggests probable functional significance for those regions showing the highest evolutionary constraints.

The chicken HMG14 and HMG17 proteins are 104 and 89 amino acids in length, respectively. They can be divided into several domains based upon sequence homology, charge and function. The N-terminal regions of the molecules are basic and contain the DNA binding domains which have a high positive charge. The acidic C-termini have a negative net charge and may mediate interactions with positively charged proteins such as the histones. The C-termini of two HMG molecules bound to the same nucleosome may also interact with each other to stabilize the particle and/or facilitate a cooperative mode of binding. The homology of the proteins vary along the polypeptide chain. The first four amino acids of HMG14 and HMG17 are PKRK (appendix I) and are absolutely conserved among species. The location of the DNA binding domain varies slightly between species. In the

chicken proteins, the HMG14a DNA binding domain is located from residue 16-40 and the HMG17 from 21-45. In the calf proteins, DNA binding has been localized to amino acids 17 -60 in HMG14 and 15-40 in HMG17 (21). Regardless of its location, the DNA binding domains of these proteins contain two stretches of virtually invariant residues, P(K,Q)RRSARLSA and KPKKA. An alignment of the known protein sequences of HMG14 and HMG17 is shown in figure 3.

Analysis of the primary protein structures of HMG14 and HMG17 also shows conservation in the distribution of charged residues along the polypeptide chains. The first 20 amino acids of HMG17 contain 12 charged residues with a net charge of +2. Similarly, the first 15 amino acids of HMG14 contain 7-8 charged residues with a net charge of +1. The central region of these molecules, which contains the DNA binding domain, is highly positively charged with residues 20-64 in HMG17 and 15-79 in HMG14 having net positive charges of +16 and +15 respectively. The DNA binding domains in this central region, approximately 25 amino acids long, have basic/acidic ratios of 9/1. The C-terminal 26 amino acids of HMG17 contains 9 charged residues with a net charge of -3. In HMG14, the C-terminal 23 amino acids contain 18 charged residues with an overall charge of -3. This high number of charged residues in the C-terminus is characteristic of the other HMG protein families as well.

Figure 3 - **Alignment of HMG14 and HMG17 protein sequences**

Alignment of HMG14 and HMG17 protein sequences from various organisms (15). Overlined sequences indicate regions of invariant amino acids or those which are conservatively substituted among the proteins.

```
                 ----    ---       ---- ----------------------
human   HMG14    PKRKVSSAEG--------AKEE-PKRRSARLSAKPPAKVEAKPKKK
calf    HMG14    PKRKVSSAEG-------AAKEE-PKRRSARLSAKPAPAKVETKPKK
mouse   HMG14    PKRKVS-ADG-------AAKAE-PKRRSARLSAKPAPAKVDAKPKK
chicken HMG14a   PKRKAP-AEGE-------AKEE-PKRRSARLSAKPAPPKPEPKPKK
chicken HMG14b   PKRKVAASRG--------GREEVPKRRSARLSAKPVPDKAEPKAKA
human   HMG17    PKRK---AEGDAKGDKAKVKDE-PQRRSARLSAKPAPPKPEPKPKK
calf    HMG17    PKRK---AEGDAKGDKAKVKDE-PQRRSARLSAKPAPPKPEPKPKK
mouse   HMG17    PKRK---AEGDAKGDKTKVKDE-PQRRSARLSAKPAPPKPEPKPKK
chicken HMG17    PKRK---AEGDTKGDKAKVKDE-PQRRSARLSAKPAPPKPEPKPKK
trout   H6       PKRKSA-----TKG------DE-PARRSARLSARPVP-KPAAKPKK

                ----       - -  --  -----          --      -----
                AAAK------DKSSDKKVQTKGKRGAKGKQ-AEVANQETKEDLPAENG
                AAGK------DKSSDKKVQTKGKRGAKGKQ-AEVANQETKEDLPAENG
                AAGK------DKASDKKVQIKGKRGAKGKQ-ADVADQQTT-ELPAENG
                AAPKKEKAANDKKEDKKAATKGKKGAKGKD--ETKQEDAKEENHSENG
                LAAK------DKSENKKAQSKGKKGPKGKQTEETNQEQIKDNLPAENG
                APAK-------KGE--KVP-KGKKG--------KADAGKEGNNPAENG
                APAK-------KGE--KVP-KGKKG--------KADAGKDGNNPAENG
                APAK-------KGE--KVP-KGKKG--------KADAGKDANNPAENG
                AAPK-------KSE--KVP-KGKKG--------KADAGKEGNNPAENG
                AAAP-----------KKAV-KGKKA------------------AENG

                     ----              ---
                     ETKTEESPASDEAGEK-EAKSD
                     ETKNEESPASDEAEEK-EAKSD
                     ETENQ-SPASEE--EK-EAKSD
                     DTKTNEAPAAEASDDK-EAKSE
                     ETKSEETPASDAAVEKEEVKSE
                     DAKTDQAQK--AEGAG-DAK--
                     DAKTNQAEK--AEGAG-DAK--
                     DAKTDQAQK--AEGAG-DAK--
                     DAKTNQAEK--AEGAG-DAK--
                     DAKAEAKVQAAGDGAG-NAK--
```

Figure 3

The specific amino acid contents of the chicken HMG14a and HMG17 proteins is listed in figure 4. Interestingly, all HMG14 and HMG17 proteins have an unusually high number of proline residues. In chicken HMG17, there are 8 proline in a stretch of 17 residues in the DNA binding domain. Likewise, in chicken HMG14a, there are 7 prolines out of 17 residues. Proline residues will produce a bend in a polypeptide chain, normally disrupting or preventing regions of secondary and tertiary structure from forming in the molecule. The high level of proline in these HMG proteins most likely contributes to the lack of higher order structure seen in these polypeptides.

Using circular dichroism and NMR studies, it was shown that these proteins do not form any secondary or tertiary structure over a wide range of salt conditions (21,22). This suggests that the conformation of these proteins is controlled by the binding to other factors in the chromatin, particulary DNA. However, analysis of the HMG14 and HMG17 protein sequences by the Chou-Fasman secondary structure prediction model suggests the acidic C-termini could be in a weak alpha helical conformation with the acidic residues positioned on a single face of the helix (21,22). This lends support to the suggestion that this region participates in protein-protein interactions primarily through the acidic face of the helix as is seen

Figure 4A - **Translated coding region of chicken HMG14a cDNA**

The translated coding region of chicken HMG14a is shown. The total number and percent content of each amino acid is listed. The predicted molecular weight is 11,225 daltons. See appendix I for amino acid abbreviation list.

```
CCC AAA AGA AAG GCT CCA GCT GAA GGC GAG GCG AAG GAG GAG CCA
P   K   R   K   A   P   A   E   G   E   A   K   E   E   P 15

AAG AGA AGG TCG GCC AGA CTA TCT GCT AAA CCT GCT CCG CCT AAA
K   R   R   S   A   R   L   S   A   K   P   A   P   P   K 30

CCG GAG CCA AAG CCC AAA AAG GCA GCA CCT AAG AAA GAA AAG GCA
P   E   P   K   P   K   K   A   A   P   K   K   E   K   A 45

GCA AAC GAT AAA AAG GAA GAC AAA AAG GCA GCA ACA AAA GGG AAG
A   N   D   K   K   E   D   K   K   A   A   T   K   G   K 60

AAA GGA GCC AAA GGC AAA GAC GAA ACT AAA CAA GAG GAT GCA AAA
K   G   A   K   G   K   D   E   T   K   Q   E   D   A   K 75

GAA GAA AAC CAC TCT GAA AAT GGA GAT ACC AAA ACT AAT GAG GCA
E   E   N   H   S   E   N   G   D   T   K   T   N   E   A 90

CCA GCT GCT GAA GCA TCT GAT GAT AAG GAA GCC AAG TCC GAG
P   A   A   E   A   S   D   D   K   E   A   K   S   E   104
```

Amino acid	Total number	Total percent
A	19	18.1
C	0	0.0
D	7	6.7
E	16	15.2
F	0	0.0
G	5	4.8
H	1	1.0
I	0	0.0
K	26	24.8
L	1	1.0
M	0	0.0
P	11	10.5
Q	1	1.0
R	4	3.8
S	5	4.8
T	4	3.8
V	0	0.0
W	0	0.0
Y	0	0.0

Figure 4A (con.)

Figure 4B - **Translated coding region of chicken HMG17 cDNA**

The translated coding region of chicken HMG17 is shown. The total number and percent content of each amino acid is listed. See appendix I for amino acid abbreviation list

```
CCG AAG AGA AAG GCT GAA GGA GAT ACC AAG GGC GAT AAG GCC AAA
P   K   R   K   A   E   G   D   T   K   G   D   K   A   K 15

GTT AAG GAT GAG CCA CAA CGG AGA TCG GCA AGG TTA TCT GCT AAA
V   K   D   E   P   Q   R   R   S   A   R   L   S   A   K 30

CCT GCC CCT CCG AAG CCA GAG CCT AAA CCT AAA AAG GCA GCT CCA
P   A   P   P   K   P   E   P   K   P   K   K   A   A   P 45

AAG AAG AGT GAG AAG GTG CCC AAG GGA AAG AAG GGG AAA GCT GAT
K   K   S   E   K   V   P   K   G   K   K   G   K   A   D 60

GCT GGC AAG GAG GGA AAC AAC CCT GCA GAA AAT GGA GAT GCC AAA
A   G   K   E   G   N   N   P   A   E   N   G   D   A   K 75

ACA GAC CAG GCA CAG AAA GCC GAA GGT GCT GGT GAT GCC AAG
T   D   Q   A   Q   K   A   E   G   A   G   D   A   K    89
```

Amino Acid	Total number	Total percent
A	15	16.9
C	0	0.0
D	7	7.9
E	7	7.9
F	0	0.0
G	9	10.1
H	0	0.0
I	0	0.0
K	22	24.7
L	1	1.1
M	0	0.0
N	3	3.4
P	11	12.4
Q	3	3.4
R	4	4.5
S	3	3.4
T	2	2.2
V	2	2.2
W	0	0.0
Y	0	0.0

Figure 4B (con.)

with other classes of proteins such as transcription factors.

Association of HMG14 and HMG17 with active chromatin

HMG14 and HMG17 are the only DNA binding proteins with a higher affinity for nucleosomal DNA than naked DNA. In early experiments to determine the stoichiometry of HMG binding to the nucleosome, nucleosomal core particles were salt stripped with 0.35M NaCl to release the HMG proteins, then titrated with HMG14 or HMG17 and electrophoresed on native gels (22,23,24). Gel shift patterns clearly show two more slowly migrating bands indicating two potential binding sites per nucleosome for HMG14 or HMG17. These binding sites were located on the inner face of the DNA as it emerges from the nucleosome core. There was no difference seen when using HMG14 or HMG17 alone or in combination in these reconstitution experiments. This suggests that these two proteins bind with equal specificity to the nucleosomal core particle. Recently, it was shown that a 30 amino acid long peptide corresponding to the DNA binding region of HMG17 can bind with equal affinity and specificity to that of the intact protein (32). Recent analysis of the specific binding of the HMG14 and HMG17 to nucleosome cores and H1 depleted chromatosomes confirmed the stoichiometry of two nucleosomes per particle

and found the path of HMG14 on the surface of the nucleosome is indistingiushable from that of HMG17 (26). Using hydroxy radical cleavage mapping experiments, it was shown that these HMG proteins protect DNA 25 base pairs out from the core particle on either end (26). The sites occupied by HMG14 and HMG17 on the ends of the core particle are distinct from the sites occupied by the linker histones (26). A model of the binding of HMG14 and HMG17 to the nucleosome is shown in figure 5. Other studies of the specific binding of these HMG proteins to the nucleosome further characterize this interaction. Using circular dichroism and thermal denaturation, it was shown that the binding of HMG17 produces an overall stabilization and condensation of the core particle (26). Neutron scattering experiments demonstrated that the binding of HMG14 produced a slight increase in the radius of gyration of the DNA and that there are fewer nucleosomes per repeat in HMG14 containing fibers (27,28). Binding of HMG14 or HMG17 to the nucleosome core does not induce any detectable rearrangment of the core histones but may promote a slightly more compact core histone structure (27). Crosslinking studies show the C-terminus of HMG17 interacts primarily with the core histone H2A (29,30).

Although every nucleosome has two potential binding sites for HMG14 and HMG17, the limited amount of these proteins in

Figure 5 - **Binding of HMG14/17 on nucleosome core**

Location of HMG14 and HMG17 on the nucleosome core (25). The model illustrates binding of the HMG proteins to the nucleosome core. The nucleosome core potentially contains two HMG molecules which bind to the DNA as it enters and exits the core structure. The acidic tails of the HMG molecules are free to interact with core histones or each other.

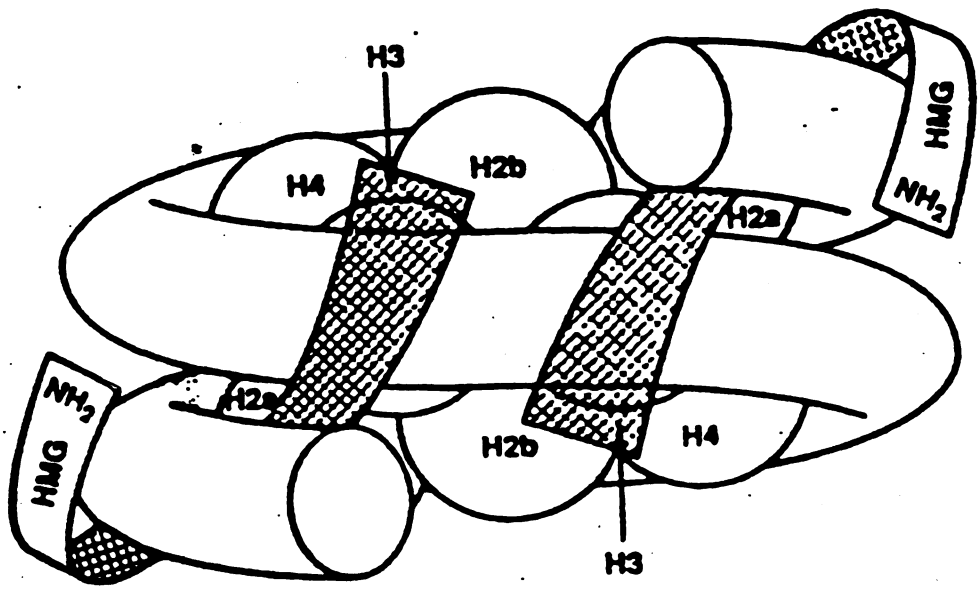


Figure 5

the nucleus restricts this binding to a subset. It is estimated that 1 out of every 10 nucleosomes will have HMG bound *in vivo*. This coincides with the estimated 10-20% of the chromatin which is actively transcribing or in a potentially active state in the cell. Actively transcribing genes are preferentially digested by DNaseI, a non-specific endonuclease. This DNase sensitivity of actively transcribing genes is dependent upon their association with the HMG14 and HMG17 proteins. The association of HMG14 and HMG17 with active chromatin was first shown by Weisbrod and Weintraub in 1979 (8). In erythrocytes, the globin gene is active and preferentially digested with DNase while the ovalbumin gene is inactive and is not sensitive to this enzyme. In erythrocyte chromatin depleted with 0.35 M NaCl, the globin gene looses this preferential sensitivity while the ovalbumin gene remains unaffected. Reconstition of chromatin or mononucleosomes with the 0.35 M NaCl fraction or purified HMG14 and/or HMG17 restores the DNase sensitivity of the globin gene while the ovalbumin gene remains unaffected (7,8). The cellular source of the HMG14 or HMG17 used did not alter the specificity of the reconstitution. Full sensitivity to DNaseI is restored to active genes at about 1 mole HMG14 or HMG17 per 10-20 nucleosomes, a ratio equivalent to the amount of these HMG proteins found in the nucleus (7). In a similar manner, HMG14 and HMG17 were also shown to

specifically confer DNase sensitivity to nuclear RNA (nRNA) sequences as well as bulk active sequences (7,9). Nucleosome fractions containing active genes were found to be enriched for HMG14 and HMG17 and nucleosomes with bound HMG14 and HMG17 were found to be enriched in active, DNase sensitive sequences (12,32). HMG14 and HMG17 immobilized by crosslinking to agarose specifically retained actively transcribing nucleosomes when exposed to bulk digested chromatin (10). Quantitative analysis in chicken erythrocytes of the distribution of chromosomal proteins on transcribed chromatin shows a 1.5-2.5 higher density of HMG14 and HMG17 and a 2 fold lower density of histone H1 and H5 as compared to inactive sequences (36). Antibodies to HMG17 were also used to specifically isolate nucleosomes containing active genes (11,33). In oligonucleosomes isolated with antibodies to HMG17, the distribution of HMG17 with respect to the coding region of specific genes was analyzed in detail. It was found that HMG17 is bound only downstream of the transcriptional start site (34). This correlates with the observation that the coding regions of active genes are most sensitive to digestion by DNaseI, while regions upstream are only moderately sensitive. All of these results suggest that HMG14 and HMG17 bind specifically to regions of chromatin containing active genes and indicate that active nucleosomes possess some feature that allows HMG14 and HMG17 to

specifically distinguish them from the inactive nucleosomes. It is known that in active chromatin the core histones are hyperacetylated and there is a marked decrease in the amount of histone H1 bound. These changes and their consequences on the fine structure of nucleosomes may be involved in the specificity of binding of the HMG14 and HMG17 proteins. Indeed, HMG14 and HMG17 have been shown to partially inhibit histone deacetylase in active chromatin (51).

Due to this well established association of HMG14 and HMG17 with active gene sequences, the influence of these proteins on the process of transcription was analyzed. Microinjection of antibodies to HMG17 into fibroblasts inhibited transcription as shown by significant reductions in nuclear incorporation of tritiated uridine (13). This suggests that the HMG proteins are required for transcriptional activity in the cell. The ability of these HMG proteins to function as direct transcriptional activators was studied in *Saccharomyces cerevisiae* cells (35). These cells express fusion proteins comprised of the *lexA* DNA binding domain fused to the HMG14 or HMG17 acidic C-terminus, which has features reminiscent of some transcriptional activators. This construct was unable to stimulate transcription from reporter constructs containing the *lexA* operator sequences upstream of the *B*-galactosidase gene. Thus, in this system at least, HMG14 and HMG17

do not function as direct transcriptional activators.

Modifications of the HMG14 and HMG17 proteins

HMG14 and HMG17 are known to undergo several post -transcriptional modifications including methylation, acetylation, phosphorylation, poly(ADP) ribosylation and glycosylation. The functional significance,if any, of these modifications is unknown. These HMG proteins can be poly(ADP) ribosylated (38); in the chicken proteins the putative sites for this modification are glutamine 81 in HMG14a and glutamine 70 in HMG17. While the biological significance of this modification are unknown, the ribosylation sites are located in the C-terminus of the molecules and may be involved in interactions of the HMG proteins with the nucleosome and/or the core histones. Studies with trout H6 protein show there is a preferential localization of ribosylated HMG proteins in DNase sensitive chromatin (37). HMG14 and HMG17 can be glycosylated, and *in-vivo* labeling experiments show these proteins can incorporate fucose, galactose, mannose and *N*-acetylglucosamine. Most of these oligosaccharide linkages are of the *N*-glycosidic type (38). The function of these modifications are unknown, yet it has been shown that gylcosylated HMG proteins bind preferentially to the nuclear protein matrix in mammalian cells (39). The nuclear matrix is the site of both RNA transcription and DNA

replication which is consistent with a role of these proteins in actively transcribing chromatin.

Phosphorylation of HMG14 in mammilian cells can occur at two sites, serine 6 and serine 24. These residues are located in two highly conserved regions of the protein. Phosphorylation at these sites was shown to be hormonally regulated by thyroid stimulating hormone *in-vivo* and can be mimicked by forskolin and cAMP analogues *in-vitro* (40). HMG17 is also phosphorylated but does not show this regulation. Phosphorylation of HMG14 reduces its ability to interact with nucleosomes by diminishing interactions with DNA. This could be an important regulatory mechanism for HMG14 binding to nucleosomes and illustrates an important biological difference between HMG14 and HMG17.

Tissue specificity and developmental regulation of HMG14 and HMG17

Antibody injection experiments have localized HMG14 and HMG17 to the nucleus in human fibroblast cells (41). There appears to be no functional difference in the HMG proteins produced in different tissues as HMG14 or HMG17 from chicken brain cells can reconstitute DNase sensitivity on erythrocyte chromatin (7).

The cell cycle regulation of HMG17 mRNA has been analyzed in HeLa cells (43). The level of HMG17 mRNA in these cells is high and does not correspond to the amount of protein present

in the cell. The level of HMG17 mRNA is much higher than the level of the message for actin which is a much more abundant protein in the cell. This suggests that the HMG17 is rapidly turned over, however this conclusion is disputed by other evidence. HMG17 mRNA is present throughout the cell cycle with a sharp rise at the beginning of S-phase, followed by another rise near the end of S-phase. This rise in the rate of transcription of the gene is not coupled to DNA synthesis as inhibition of DNA sythesis does not affect HMG17 mRNA synthesis (43). Interestingly, the level of HMG17 protein present in HeLa cells does not change noticeably with fluctuations in the mRNA levels (15).

Isolation of HMG14 and HMG17 proteins from various organs of the chicken show that in organs with a higher proportion of replicating cells, the amount of HMG17 is much higher than HMG14 while in transcriptionally active organs with a very small proportion of replicating cells, the levels of HMG14 and HMG17 are low and roughly equal (42). The relative levels of HMG14 and HMG17 proteins varies in the different cells tested. In chicken erythroctyes and transformed quail fibroblasts, HMG14a is more abundant than HMG17, while in many human cell lines, HMG17 is more abundant than HMG14 with slight variations in the ratio depending upon growth conditions (50). The expression of these two proteins does not seem to be coordinately regulated, as changes in the

level of one does not effect the level of the other. In a chicken lymphoid cell line, that has the single copy gene for HMG17 inactivated at both allelles, there is no HMG17 protein present in the cell yet the level of HMG14a does not seem to be affected (Yi Li, personal communication). However, recent results from our lab suggest that HMG14 protein levels may be down regulated in transformed quail cell lines ectopically expressing high levels of chicken HMG17 (unpublished results).

There is evidence that HMG14 and HMG17 mRNA synthesis may be coupled to the differentiation state of the cell. In myogenesis, myoblasts differentiate into myotubes. During mouse myogenesis, the HMG14 and HMG17 mRNA levels and protein synthesis rate decrease to 20% of that seen in myoblasts, while the levels of the protein in the cell remains unaltered. These fluctuations were not coupled to DNA synthesis (44). Similar results were shown in rat cells, with HMG14 and HMG17 mRNA and protein synthesis rates in proliferating myoblasts significantly higher than in nondividing myotubes (45). To investigate the importance of the downregulation of the HMGs in myogenesis, exogenous human HMG14 expressed in mouse myoblasts was found to block the differentiation of these cells to myotubes (48). This suggests that proper regulated expression of the HMG proteins may be required for some aspects of cellular differentiation.

This is consistent with a role of the HMG14 and HMG17 proteins in gene regulation. During erythropoiesis in chicken, the levels of the HMG proteins and their mRNAs also change as the cells develop from embryonic to adult stages (46,47). Embryonic erythroid cells have significantly higher levels of HMG14 and HMG17 mRNAs than adult cells. In addition, HMG14a, the major chicken HMG14 protein, is expressed primarily in embryonic and developing cells while HMG14b is preferentially expressed in definitive cells. The developmental regulation of the HMG14 and HMG17 proteins was also analyzed in osteoblast cells that can be induced to differentiate into osteocytes and in promyelotic leukemia cells that can be induced to a post-proliferative state (49). The results were consistent with the other reports in that the HMG proteins were downregulated as the cells progressed into their differentiated states. It is not clear whether the switch in HMG expression levels is a prerequisite or a consequence of differentiation.

HMG14 and HMG17 cDNAs

cDNAs for HMG14 and HMG17 have been isolated and sequenced from the chicken, human, and mouse. Two types of HMG14 have been isolated from the chicken, HMG14a, which is the major chicken HMG14, and HMG14b, a slightly smaller protein, more closely related to the human HMG14. The

overall structure of the cDNAs for the HMG14 and HMG17 proteins are similar with clear sequence distinctions between the subfamilies.

All the cDNAs for HMG14 and HMG17 contain a short 5' untranslated sequence with no sequence similarity among the clones. The coding region of these cDNAs have stretches of high similarity. Among the HMG14 cDNAs, the sequence similarity in this region ranges from 61% to 77%, while the HMG17 cDNAs show a higher degree of evolutionary conservation with sequence similarites ranging from 85% to 93%. The 3'-untranslated region of all the HMG14 and HMG17 cDNAs is unusually long, A+T rich, and is highly conserved in certain sections among and between the subfamilies. In the HMG14 subfamily, the first 250 nucleotides adjacent to the coding region shows similarities ranging from 48% to 86%, with the more distal segment being less conserved. In the HMG17 subfamily, the first 150 nucleotides adjacent to the coding region show sequence similarities ranging from 85% to 95%, with the similarity being less pronounced in the distal portion of the clones. Between the HMG14 and HMG17 subfamilies, this portion of the 3'-untranslated region is also highly conserved. The functional significance of the high level of conservation seen in the 3'-untranslated portion of these cDNAs is unknown, but this region may be important for gene regulation, message

stability, or processing. A dot matrix analysis comparison of the human HMG14 and HMG17 and the chicken HMG14a and HMG17 cDNAs is shown in figure 6 (53).

Chicken HMG14a cDNA

The cDNA for chicken HMG14a was isolated by screening a chicken liver cDNA library with synthetic oligonucleotide pools whose sequence was derived from the partial amino acid sequence of chicken HMG14 (53). The sequence of the chicken HMG14a cDNA is 900 bp in length. It contains 104 bp of 5'-untranslated sequence and a long 3'-untranslated sequence of 445 bp. The coding region codes for a protein of 104 amino acids. The complete cDNA sequence for this clone is shown in figure 7.

Chicken HMG17 cDNA

The cDNA for chicken HMG17 was isolated by sceening a chicken liver cDNA library with synthetic oligonucleotide pools derived from the complete amino acid sequence of chicken HMG17 (53). The cDNA is 1360 bp in length. It contains 177 bp of 5'-untranslated sequence and a 843 bp 3'-untranslated region. The coding region predicts a protein of 89 amino acids. The complete cDNA sequence for chicken HMG17 is shown in figure 8.

Figure 6 - **Dot matrix comparison of HMG14 and HMG17 cDNA sequences**

Dot matrix comparison of HMG14 and HMG17 cDNA sequences (53). (A) Comparison of human HMG17 cDNA (X-axis) to chicken HMG17 cDNA (Y-axis). (B) Comparison of human HMG14 cDNA (X axis) to chicken HMG14a cDNA (Y-axis). (C) Comparison of chicken HMG14a cDNA (X-axis) to chicken HMG17 cDNA (Y-axis). A window of 10 nt residues was used with a 50% identity required for a positive result.

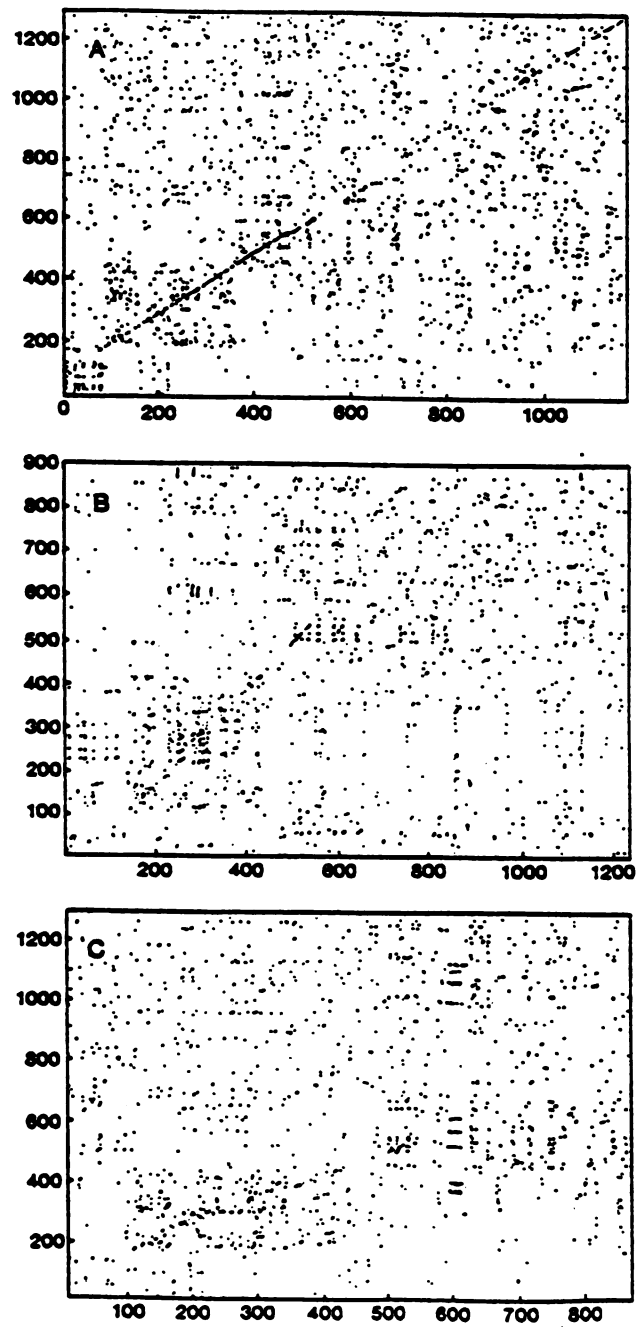

A
1200
1000
800
600
400
200
0
200
400
600
800
1000
B
900
800
700
600
500
400
300
200
100
200
400
600
800
1000
1200
C
1200
1000
800
600
400
200
100
200
300
400
500
600
700
800

Figure 6

Figure 7 - **Nucleotide sequence of chicken HMG14a cDNA**

Nucleotide sequence of chicken HMG14a cDNA. The complete sequence of the chicken HMG14a cDNA is shown with the translated amino acid sequence beneath. Seven bases of *Eco*RI linkers are at each end of the clone.

```
   GAATTCCGTC CCCTTCCTCA GGACGCTCGA AAACAGTTTC TCGGCGGTTC CCTTCCTATT    60
   TTTTACACCT CTCCCGATCT CTCTATTTGC AGTCAACTAT TAAGGTGCAA CTATGCCCAA   120
                                                              ProLy
   AAGAAAGGCT CCAGCTGAAG GCGAGGCGAA GGAGGAGCCA AAGAGAAGGT CGGCCAGACT   180
 2 sArgLysAla ProAlaGluG lyGluAlaLy sGluGluPro LysArgArgS erAlaArgLe
   ATCTGCTAAA CCTGCTCCGC CTAAACCGGA GCCAAAGCCC AAAAAGGCAG CACCTAAGAA   240
22 uSerAlaLys ProAlaProP roLysProGl uProLysPro LysLysAlaA laProLysLy
   AGAAAAGGCA GCAAACGATA AAAAGGAAGA CAAAAAGGCA GCAACAAAAG GGAAGAAAGG   300
42 sGluLysAla AlaAsnAspL ysLysGluAs pLysLysAla AlaThrLysG lyLysLysGl
   AGCCAAAGGC AAAGACGAAA CTAAACAAGA GGATGCAAAA GAAGAAAACC ACTCTGAAAA   360
62 yAlaLysGly LysAspGluT hrLysGlnGl uAspAlaLys GluGluAsnH isSerGluAs
   TGGAGATACC AAAACTAATG AGGCACCAGC TGCTGAAGCA TCTGATGATA AGGAAGCCAA   420
82 nGlyAspThr LysThrAsnG luAlaProAl aAlaGluAla SerAspAspL ysGluAlaLy
   GTCCGAGTAA TGTTAACCCT GCCCTATATC TCCATCATTT GGTATCCGTA CCTCCATGCT   480
102 sSerGlu***
   GTATTGTTAA CAGAGAGGAA TATTTTTATC AACTATTTTA TAAATGCAGG TTTTTTTAGC   540
   ATGAATTTAA TTATGGAACA TCTTCATCTC GGTTACTTGG GAATTAAATC CCTAACAAAC   600
   AAAACAAAAC AAAACAAAAA AAAATCATTG TTTTAAATTT GTGATTGTAA TAGTTTGTAT   660
   GGTACATGGA AAGAATAAGT GGTGGTAGCT TTTGACTTCT GTCAGTGTGT CCCTTTTTGT   720
   GTAAGTCATG CTTACAGACT TCAGATTTTA ATTTTACCCT TGTATGTGTT GTATGGTTTC   780
   TTAAAGTGGG GAGGTCTCAA AACAGATAAC TGTGTTAAAC ATTCCAGTGG TTCTGTGGGT   840
   TGCTTTTATA AAGAAGGTGA GCTATTTTCA TGAAAAAAAA AAAAAAAAAA AACGGAATTC   900
```

Figure 7

Figure 8 - **Nucleotide sequence of chicken HMG17 cDNA**

Nucleotide sequence of chicken HMG17 cDNA. The complete sequence of the chicken HMG17 cDNA is shown with the translated amino acid sequence beneath. Seven bases of EcoRI linkers are at each end of the clone.

```
   GAATTCCGCA GCCAGCGCAG CGAGCCGGCC GCCAGCCCCG CCGCGCCGCC CCGCTCTCCC    60
   CCTCGGCCCT CCCCCGCTTC TCGCCGCCAC CGAGCGAGCC CGGCTGCCCG CCCCCGCCCG   120
   CCCCCTCCGC TCGCTCTCTC CCTCCTCGCA CAACACACGC ACGCGCCGCC CGGAGCTATG   180
   CCGAAGAGAA AGGCTGAAGG AGATACCAAG GGCGATAAGG CCAAAGTTAA GGATGAGCCA   240
 1 ProLysArgL ysAlaGluGl yAspThrLys GlyAspLysA laLysValLy sAspGluPro
   CAACGGAGAT CGGCAAGGTT ATCTGCTAAA CCTGCCCCTC CGAAGCCAGA GCCTAAACCT   300
21 GlnArgArgS erAlaArgLe uSerAlaLys ProAlaProP roLysProGl uProLysPro
   AAAAAGGCAG CTCCAAAGAA GAGTGAGAAG GTGCCCAAGG GAAAGAAGGG GAAAGCTGAT   360
41 LysLysAlaA laProLysLy sSerGluLys ValProLysG lyLysLysGl yLysAlaAsp
   GCTGGCAAGG AGGGAAACAA CCCTGCAGAA AATGGAGATG CCAAAACAGA CCAGGCACAG   420
61 AlaGlyLysG luGlyAsnAs nProAlaGlu AsnGlyAspA laLysThrAs pGlnAlaGln
   AAAGCCGAAG GTGCTGGTGA TGCCAAGTAA AATGTGTGAA TTTTTGATAA CTGTGTACTT   480
81 LysAlaGluG lyAlaGlyAs pAlaLys***
   CTGGTGACTG TACAGTTTGA AATACTATTT TTTATCAAGT TTTATAACAA TGCAGAATTT   540
   TGTTTTACTT TTTTTAAGCT ATGTTGTTAG CACACAGACC GCTTTGTTGT TGTGTTTTGA   600
   GGGGGGGGCA GTGGGACAAA CGTCACTTAA TCTGTTTCTT GGAACCTAAA TTTTAAAAGT   660
   TTACCCCTTC CCAGTTTTTT AGAAGGACTC TTCCTAAATG GAGCAGGAAG GGATTCCTTC   720
   GTGCTGCACA CCTCTTCCGT TTTGTGGACC GCATCAGAGT GAACGGAAGC TCCCGAGATG   780
   CCTGTTGCCA ACTTCAGAAC TGCAGTTTGC AGTGCCCTCT GCGTTTCCTT TCATGCCCTC   840
   CCTTTTTGCC TAGAGCCTAT CACTCCGAAA TACAGCAGAC ATGGCATGTT GGGACTCACC   900
   ACTCTAAATG CATTGTCAGG TGATCTGGAC TTCTGGTGTC TAATTTGGGA TATAATAGCT   960
   CTAAAAGGAG CTGCATTTCC TCTTTCATAT TGTAGATCTA CAGATTAAGG AATCTGCAGT  1020
   TTTTAATTTT TCCTCGCAAA GCTAGGGTAG ATTTGTGAAG AGTTGTTAAA CAACATGCTA  1080
   AATGTGAAAG TGTCCGCCCT CACTCTAAAC ATTTCCCTCT ACAAGTATAC AAAAATGAAG  1140
   ATTTGTCGGT TTTATAGCAA CCTTTATGTT TGGGTAGTCC ATGAAGGGAG GGGAGTTTGA  1200
   CAGTTGTTGT AAAATGTTGC AGATTGTAGC CCATGTCCTG CCTAAATTAC CATGATTGTT  1260
   TATGAAAAGT ACCTTTAATA AAGCTGGATA CGGTTTGGCT TGGAAAAAAA AAAAAAAAAA  1320
   AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AACGGAATTC                        1360
```

Figure 8

HMG14 and HMG17 chromosomal genes

The chromosomal genes for HMG14 and HMG17 have been cloned and sequenced in the human and chicken (53-59). Southern analysis shows that the human genome contains 35 -50 HMG17 gene copies and 60-90 HMG14 gene copies (60,61). Most of these are believed to be retropseudogenes making HMG14 and HMG17 the largest known human retropseudogene families. The structure of the human genes are similar to the chicken genes which are described below. Comparison of the genomic structure of the HMG14 and HMG17 genes suggests these genes evolved from a common ancestor. The functional human gene for HMG17 has been mapped to band 1p36.1 and the many pseudogenes are disbursed over several chromosomes (55). Interestingly, the functional human gene for HMG14 maps to chromosome region 21q22.3 which is the region associated with the pathogenesis of Down syndrome (63,64). Down syndrome is characterized by extra copies of this region on the chromosome. RNA and protein analysis from mouse trisomy 16 embyros, a mouse model for Down syndrome, shows elevated levels of HMG14 present. It is not known if elevated levels of this chromosomal protein contribute to the etiology of the disease.

Chicken HMG14a chromosomal gene

The chicken genome contains one gene each for HMG14a and

HMG14b. HMG14a is the major HMG14 expressed in avian cells with HMG14b, the mammalian homologue, being expressed at low levels. It is not known if there is a homologue for HMG14a in mammalian cells.

The chromosomal gene for HMG14a was isolated by screening a chicken library with the HMG14a cDNA (56). Positive clones were picked and analyzed by restriction enzyme and sequence analysis. The HMG14a gene spans approximately 10 kb and is single copy in the chicken genome. The gene contains 7 exons and 6 introns, the exons being labeled exon 0 through exon 6. Exon 0 contains about 126 bp of 5'non-coding sequence. Exon 1 contains the rest of the 5' untranslated region, the ATG translational initiation codon and the first 4 codons of the protein sequence. Exons 2,3,and 4 are small at 30, 30 and 51 bp, respectively, and exon 5 is somewhat larger at 144 bp. These exons contain all coding region sequence. Exon 6 is the longest exon and contains the last 15 codons of the coding sequence, the termination codon, TAA, and the entire 3'-untranslated region of the gene. The exon-intron boundries have been mapped for this gene and the splice acceptor and splice donor sites are similar to those found in other nuclear genes (56).

The HMG14a promoter region has been sequenced and was found not to contain the consensus TATAA or CCAAT elements

found in many eukaryotic gene promoters; however a number of putative SP1 binding sites were located therein. The promoter region is very G+C rich at 76% and contains 7 HpaII sites (CCGG) (56). These features are characteristic of housekeeping genes which are transcribed constituitively in most cells. The 3'-untranslated region of the HMG14a gene contains 4 copies of the canonical polyadenylation signal AATAAA. Primer extension analysis and S1 mapping verify that the HMG14a gene is transcribed into multiple mRNAs arising from two or more transcriptional start sites with alternative splicing and utilization of two or more polyadenylation signals (59).

The chicken HMG14b gene is similar in structure to that of HMG14a. The promoter region contains no consensus elements other than the SP1 binding sites and is very G+C rich. Also, like the HMG14a gene, it contains multiple polyadenylation sites (56). Aside from very different protein coding sequences, the HMG14b gene differs from HMG14a in that exons 2 and 3 are fused. These exons code for the most highly conserved portion of the HMG14 and HMG17 protein.

Chicken HMG17 chromosomal gene

The chicken genome contains one single copy gene for HMG17. The chromosomal gene for HMG17 was isolated by

screening a chicken genomic library with the chicken cDNA for HMG17 (56). Positive clones were analyzed by restriction enzyme analysis and DNA sequencing. The HMG17 gene spans less than 4 kb in the genome, much less than the gene for HMG14a. The gene contains 6 exons and 5 introns. Exon 1 contains all of the 5' untranslated sequences, the ATG translational initiation codon and the first 4 amino acid codons for the protein. Exons 2,3,4 and 5 are small, at 45 bp, 30 bp, 51 bp and 96 bp, respectively, and contain protein coding sequences. Exon 6 is much larger than the other exons and contains the last 33 codons of protein sequence, the TAA termination codon and all of the long 3' untranslated region of the gene. The extensive similarities in genome structure of the HMG14a and HMG17 genes suggest that they evolved from a common ancestor. The exon-intron boundries have been mapped for this gene and the splice donor and splice acceptor sites are similar to those found in other nuclear genes.

The promoter region for the HMG17 gene has been sequenced and was found to contain many common promoter elements, unlike the HMG14a gene. The promoter region of HMG17 contains two TATAA boxes 31 and 41 bp upstream of the initiation site. A CCAAT element is found 61 bp upstream and the G+C content of the region is high at 75% with 4 putative SP1 binding sites. The existence and location of

these promoter elements is typical of many eukaryotic genes. The 3' untranslated region of the HMG17 gene contains one copy of the canonical polyadenylation signal AATAAA 27 bp upstream of the polyadenylation site.

Chapter II

Materials and Methods

Chapter II

Materials and Methods

1. Subcloning HMG14a and HMG17 wild type and mutant cDNAs into expression vectors

HMG14a and HMG17 cDNAs were originally cloned into the *EcoRI* site of plasmid pCla12 in both the sense and antisense orientations. pCla12 is a 2 kb adapter plasmid designed to aid in cloning fragments into expression vector TFANEO (70). The polylinker of pCla12 is flanked by *ClaI* sites, allowing the cDNAs to be excised with *ClaI* and ligated into the *ClaI* site of expression vector TFANEO (71). TFANEO is a 7.2 kb expression vector used in our experiments. The expression cassette of TFANEO consists of two LTRs derived from the Schmidt-Rupin avian Rous sarcoma virus that provide transcriptional and polyadenylation signals. A unique *ClaI* restriction site lies between the two LTRs. This vector also contains the ampicillin resistance gene and the neomycin gene driven by a chicken *B*-actin promoter for selection by G418 in tissue culture. All ligations were carried out at 15° overnight using T4 DNA ligase and buffer from New England Biolabs (NEB). When neccessary, blunt ends were created by filling in the ends of the fragment and vector with the Klenow fragment of DNA polymerase I (NEB) using standard procedures (72). Free

ends of fragments and vectors were dephoshporylated with calf alkaline phosphatase from Boehringer Mannheim (BM) using standard procedures (72). Ligation reactions were transformed to *E.Coli DH5* by standard procedures and plated onto Luria -Bertani medium (LB) plates containing ampicillin for plasmid selection (72). Colonies were picked and plasmid DNA was isolated by alkaline lysis and tested by restriction enzyme digests for presence of insert and orientation. Most of our restriction enzymes were purchased from NEB or BM. The orientation of the HMG cDNAs in TFANEO was determined by restriction enzyme analysis as follows. Within the pCla12 polylinker, there is a *BamHI* site located 3' to the 5' *ClaI* site and 5' to the *EcoRI* site, such that the *BamHI* site is always 5' to the cDNA insert when the insert is cloned in the sense orientation in pCla12. Additionally, there are two *BamHI* sites in TFANEO located 5' to the *ClaI* cloning site in this vector. These sites are located 600 and 1000 bp upstream of the *ClaI* site. When digesting TFANEO containing an HMG cDNA insert, a 400 bp *BamHI* band is always present and a 600 bp band will be present when the insert is in the sense orientation. If the insert is in the antisense orientation, there will be a 400 bp *BamHI* band along with a band of 600 bp plus the size of the cDNA insert. This cloning strategy is illustrated in figure 1.

Figure 1 - **TFANEO subcloning strategy**

The subcloning strategy to move cDNA inserts from plasmid pCla12 to plasmid TFANEO is shown. This is described in detail in the text.

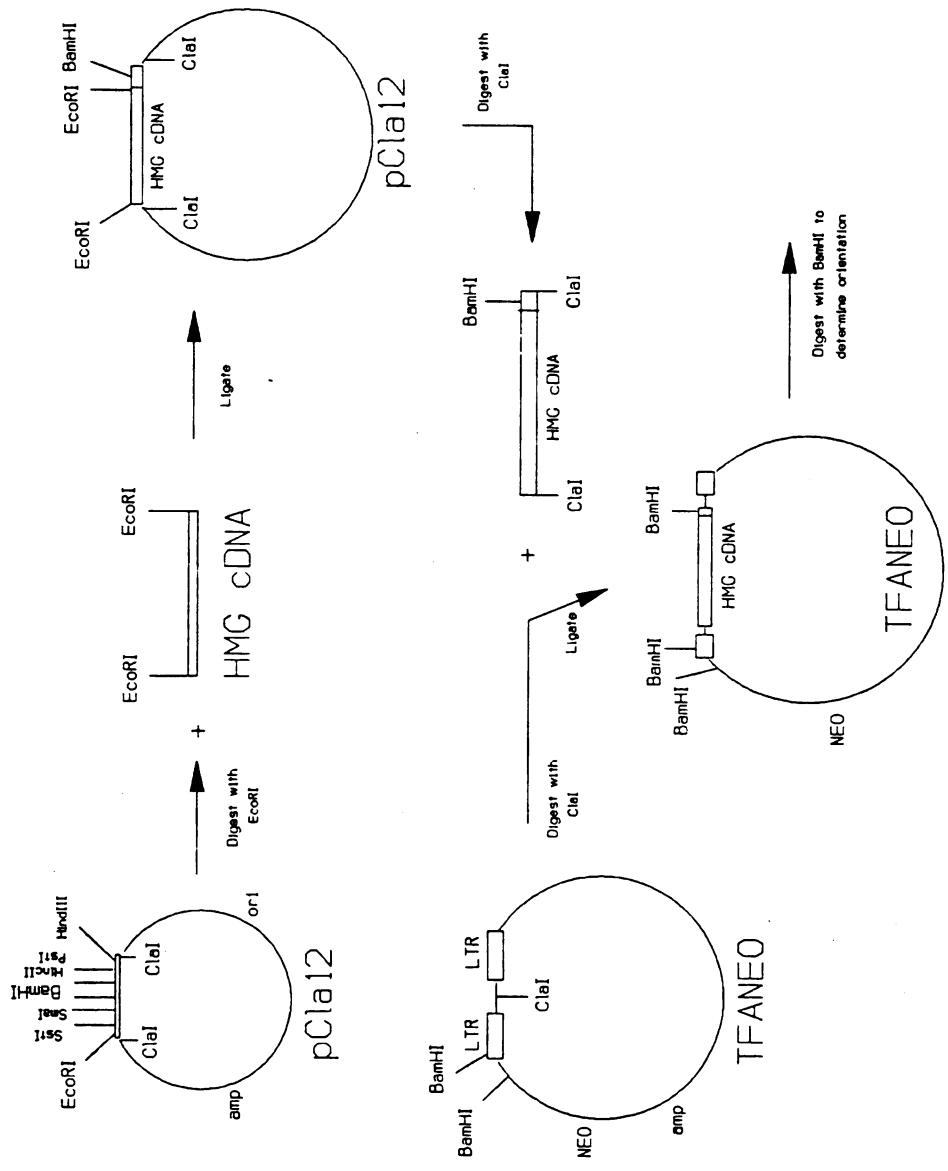

EcoRI
SstI
SmaI
BamHI
PstI
HindIII
ClaI
ClaI
ori
amp
pCla12
Digest with EcoRI
EcoRI
EcoRI
HMG cDNA
+
Ligate
EcoRI BamHI
ClaI
HMG cDNA
ClaI
EcoRI
pCla12
Digest with ClaI
BamHI
ClaI
HMG cDNA
ClaI
+
Digest with ClaI
Ligate
BamHI
LTR
LTR
ClaI
BamHI
NEO
amp
TFANEO
BamHI
HMG cDNA
BamHI
BamHI
NEO
TFANEO
Digest with BamHI to determine orientation

Figure 1

The wild type HMG14a and 17 cDNAs were cloned into TFANEO as well as several mutant cDNAs. The mutant constructs are described in detail in chapter 3. C14Ala11 is the HMG14a cDNA with a site-specfic silent mutation in the third position of codon 11 which codes for alanine. This generates a new *StyI* restriction site. C14ADPR and C17ADPR are the HMG cDNAs containing mutations at the putative ADP-riboyslyation sites, Glu 81 to Arg in HMG14a (generating a new *XbaI* site) and Glu 70 to Ala in HMG17 (generating new *BbvI* and *FnuHI* sites). HMG14-RS contains a 194 bp *RsaI* deletion in the conserved 3'-untranslated region. HMG17-RS contains a similar 178 bp deletion between the *RsaI* and *DraI* sites in it's 3'-untranslated region. A number of site-specific mutations in the DNA binding domain of each cDNA were also prepared and cloned into TFANEO. These are discussed in detail in chapter 3. The coding region only from each of the cDNAs was isolated and blunt\ end ligated into TFANEO vector cut with *EcoRI* and filled in with Klenow. The HMG14a coding region is contained in a 340 bp *HincII* fragment, and the HMG17 coding region in a 313 bp *BstUI/RsaI* fragment. The polyadenylation site contained in the HMG 17 cDNA was deleted in one clone to eliminate any interference with the downstream polyadenylation site in the vector LTR. This was deleted by digesting pC17cw with *AccI*, cutting out a 259 bp fragment containing the

polyA consensus site and religating the plasmid.

The HMG14a and HMG17 wild type cDNAs were also cloned into the expression vector pRC/CMV (Invitrogen). This vector drives expression of the inserted gene from a human CMV promoter and uses a SV40 polyA site. pRC/CMV also contains ampicillin and neomycin resistance genes. Due to a limited polylinker in this vector, the cDNAs were cut out of pCla12 with *EcoRI* and ligated into the *HindIII* site of pRC/CMV using *EcoRI-HindIII* adapters (NEB). Orientation was determined by restriction enzyme analysis using various restriction enzyme sites in the vector and insert.

HMG14a wild type cDNA and coding region were cloned into the regulatable eukaryotic expression vector pUHD10-3 (73). This vector contains a minimal CMV promoter fused downstream to multiple *E.Coli* tetracycline operator sequences such that transcription is activated upon binding of the *TetR* activator protein to the operator sequences. The *TetR* is present as a VP16-*TetR* activator expressed from a second vector, pUHD15-1 (73). The presence of tetracycline in the tissue culture media represses transcription by complexing with the VP16-*TetR* and inhibiting its interaction with the promoter operator sequences.

The full length HMG14a cDNA was ligated into the *EcoRI* site in the polylinker of pUHD10-3, while the coding region

was blunt end ligated into a Klenow filled *EcoRI* site. Orientation was determined by restriction enzyme analysis.

HMG14aFLAG and HMG17FLAG contain an additional 24 bp (8 amino acid codons) at the 3' end of their respective coding region. The translated 24 bp sequence is specifically recognized by a monoclonal antibody, M2 (IBI-Kodak). The cDNA *EcoRI* PCR fragment containing this fusion was ligated into the *EcoRI* site of pUHD10 -3. Site-specific mutants in the DNA binding domain of HMG14aFLAG were prepared by PCR using the wild type HMG14aFLAG cloned into pCla12 as template.

pUHD10-3/neo was prepared by excising the 2.4 Kb *BamHI/HindIII* fragment containing the neomycin resistance gene and chicken *B* -actin promoter from TFANEO, filling in the ends with Klenow and blunt end ligating this into pUHD10-3 cut with *HindIII* and filled in with Klenow.

2. Production of site-specific mutations in HMG14a and HMG17 cDNAs

a. Oligonucleotide directed in-vitro mutagenesis

A series of site specific mutations were prepared in the conserved DNA binding domain of the HMG cDNAs. Specific mutations, oligonucleotides used and significance are discussed in chapter 3. Mutants were prepared using Muta -gene M13 in-vitro mutagenesis kit (Bio-Rad).

Oligonucleotides containing internal single and double mismatches at the site of mutation were synthesized (Macromolecular Structure Facility, MSU). 200 pmol of individual oligonucleotides were phosphorylated with 4.5 units of T4 polynucleotide kinase (NEB) at 65° for 10 minutes. The wild type HMG14a and HMG17 cDNAs were cloned into the *EcoRI* site of M13 plasmid vector.These clones were transformed into *E.Coli* strain CJ236 which is a *dut,ung* mutant, allowing occasional incorporation of uracil in the place of thymine in the DNA synthesized in this bacterium. The uracil containing single stranded plasmid DNA was isolated using procedures recommended by the manufacturer. 3 pmol of the phosphorylated oligos were annealed to 200 ng of the uracil containing DNA by mixing with annealling buffer and heating to 70°, then slowly cooling to room temperature. 5 units of T4 DNA ligase and 1 unit of T4 DNA polymerase were added to this reaction and synthesis of the complementary strand was performed by incubation on ice for 5 minutes, at 25° for 5 minutes and then at 37° for 90 minutes. The reaction was stopped by the addition of 10 mM EDTA. This synthesis reaction was then transformed into wild type *E.Coli* strain MV1190, which only replicates the non-uracil containing DNA. Viable plaques were picked from transformation plates. Single stranded phage DNA was isolated from several plaques per reaction, cleaned

through a glass bead column (US Bioclean) and sequenced using standard di-deoxy sequencing procedures. The devised mutation was found in an average of 50% of the colonies picked.

b. **Site-specific mutagenesis with PCR**

Several mutations were prepared in the DNA binding domain of the HMG14a cDNA using the technique of Ho et.al. (74). Specific oligonucleotide primers used are described in chapter 3. Primers with internal mismatches at the site of mutation were synthesized (Macromolecular Structure Facility, MSU) in the sense and antisense orientation. Perfect match external primers were prepared complementary to sequences 200 to 900 base pairs on either side of the mutation site. DNA fragments were synthesized using standard PCR methods. The HMG14a cDNA cloned into pCla12 (pC14cw) was used as template with one of two pairs of primers, one external and one internal per pair. These PCR products, each with the mutated region at one end, were then combined in a second PCR reaction as template using the external primers to synthesize a fragment containing the cDNA with a internal mutation. This fragment was then digested with restriction enzyme *EcoRI*, releasing the cDNA carrying the mutation, and this fragment was then ligated into pCla12 for further subcloning into expression vectors

followed by sequencing using standard di-deoxy methods.

3. **Cell lines and cell culture**

QT6 is a chemically transformed quail fibroblast cell line (67). These adherant cells were grown in DMEM supplemented with 4% fetal calf serum, 1% chicken serum, 1% DMSO, 1X fungizone, and 1X Gentamycin. All tissue culture media materials were purchased from Gibco-BRL. The cells were grown at 37° and 5% CO_2.

4. **Transfections of tissue culture cells**

Stable and transient transfections of the QT6 cell line were carried out using standard calcium phosphate transfection procedures with no glycerol shock (72). Cells were grown in 10 cm tissue culture plates to 50-60% confluency, or approximately 10^6 cells per plate, and transfected with 10 ug DNA. 24 hours following transfection, cells were exposed to media containing 1 mg/ml G418 for stable selection. Single colonies arising after 10 days were isolated and frozen down in QT6 media supplemented with 20% fetal calf serum and 10% DMSO for later analysis. Transient transfections were performed as above with the cells being isolated at 24 hours post -transfection. For transfections using the tetracycline regulatable vector pUHD15-1, media was supplemented with

1.0-4.0 ug/ml tetracycline (Sigma).

5. **Protein Isolation**

HMG proteins were isolated from tissue culture cell lines as follows: Plates of cells were washed with cold PBS (Phosphate buffered saline (72)) containing the protease inhibitors aprotinin, 1%, (Sigma) and 0.1mM PMSF (Sigma). 5 ml of cold PBS with protease inhibitors was added to plates and cells were scraped free with rubber policeman and transfered to a 15 ml centrifuge tube. Cells were counted then pelleted at 1500 rpm at 4° for 2 minutes, then resuspended in lysis buffer containing 200 ul of 10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 1% aprotinin and 0.1 mM PMSF. 5% perchloric acid (14.4 ul) was added to the cells and this was incubated on ice for 10 minutes after vortexing.The cell extract was then spun out at 15,000 rpm at 4° for 10 minutes. The supernatant, which contains the HMG proteins in solution, was collected and made 0.3N in HCl (5.2 ul of conc. HCl). 7 volumes of acetone were used to precipitate the protein at -20° overnight. HMG samples were centrifuged at 4° for 10 minutes and the protein pellet was taken up in 50 ul water or SDS sample loading buffer (see below). The amount of protein was quantitated using the Pierce BCA protein reagent assay or approximated by counting the cell number prior to

isolation. The HMG protein samples were then run on SDS-PAGE gels (see below).

6. **Electrophoresis of HMG proteins on SDS-PAGE gels**

HMG protein samples were taken up in sample loading buffer containing 50 mM Tris-HCl, pH 6.8, 10% 2-mercaptoethanol, 2% SDS, 0.1% bromophenol blue, and 20% glycerol. Approximately 1 ug of 5% PCA extracted protein was loaded per lane. Electrophoresis was carried out on 15% or 18% SDS-PAGE gels with a 5% stacker gel at 150 V for 1.5-2 hours (72). The proteins were electro-transfered to PVDF membrane (Amersham) overnight at 4° at 50 V. Following transfer, the gels were stained with coomassie blue to visualize untransfered protein. Histone H1 is isolated along with the HMG proteins and due to the large amounts isolated, is not all transfered to the membrane. The coomassie blue stained histone H1 remaining on the gel is used as a loading control. The membranes are then subjected to western blotting techniques to visualize the HMG proteins.

7. **Western blotting**

Anti-sera recognizing a highly conserved portion of the DNA binding domain of both HMG14 and HMG17 was obtained from Michael Bustin and used in most western blotting (66).

Anti-FLAG M2 antibody was purchased from Kodak-IBI and was used for detection of FLAG-containing HMG constructs. For the HMG14/17 specific antibody, following transfer, the blots were washed twice for 15 minutes in PBS/0.1% Tween 20, and blocked for 1 hour in PBS/Tween 20/5% dehydrated milk at room temperature. The blots were then incubated for 1 hour with 5 ul of 100 mg/ml anti-sera in 50 ml of blocking solution. The blots were washed twice for 15 minutes in PBS/Tween 20 and incubated with 16.5 ul of anti-rabbit HRP conjugated secondary antibody (Bio-Rad) for 45 minutes at room temperature. This was followed by washing again twice for 15 minutes, each wash in PBS/Tween 20. When using the anti-FLAG M2 antibody, the blots were washed were washed twice for 15 minutes in TBS (0.05M Tris Base, 0.015M NaCl, pH 7.4), and blocked for 15 minutes in TBS/3% dehydrated milk at 37°. The blots were then incubated for 1 hour with anti-FLAG antibody at a concentration of 10 ug/ml in TBS. The blots were washed twice for 10 minutes in TBS/0.05% Tween 20 and incubated with 16.5 ul of anti-mouse HRP conjugated secondary antibody in 50 ml of TBS for 1 hour at room temperature. This was folwed by washing again three times at room temperature for 15 minutes, each wash TBS. Chemiluminescent detection of the HMG proteins was done using the Amersham ECL kit or Dupont NEN Renaissance western blot reagents. The chemiluminescence treated blots were then exposed to Amersham

hybond MP film for 5-30 seconds and the film was developed. Quantification of bands on the exposed film was done using the optical imaging software, AMBIS. The relative levels of HMG protein contained in each cell line were expressed as the ratio of HMG14/ HMG17 or as individual bands normalized to the control cell line histone H1 protein level.

8. **Production of HMG14a/HMG17-FLAG fusions by PCR**

To aid in distinguishing between the exogenous and endogenous HMG proteins in the cells, a fusion construct was prepared by addition of 24 bp to the carboxy terminus of the HMG14a and HMG17 coding region. The 24 bp sequence codes for 8 amino acids recognized with high specificity by antibodies and other conjugated reagents sold by Kodak scientific products. This fusion construct was produced using a modified version of the PCR mutagenesis scheme of Ho et al. Oligonucleotide primers were synthesized (Macromolecular Structure Facility, MSU), one complementary to the last 16 bp of the HMG14a/17 coding sequence with the 24 bp of FLAG sequence added at the 3' end, and one with the Flag sequences at the 5' end followed by the translational stop codon and the following 13 bp of noncoding sequence of the cDNA. The FLAG sequence and specific primers used are described in chapter 3. Perfect match primers used in the previous PCR mutagenesis reactions that are complementary to

sequences just outside of the cDNA (in the pCla12 cloning vector) were used in pairs (one inside plus one outside) in a PCR reaction resulting in amplification of two fragments, each containing the FLAG sequence at the 3' or 5' end. These two fragments were then combined in another PCR reaction using the two outside primers, resulting in a fragment which contained the full length cDNA with the FLAG sequences fused internally to the 3'end of the coding region. The PCR reaction was then extracted with phenol/chloroform followed by ether extraction to get rid of any residual phenol. The HMG14/17-FLAG cDNA was digested from the PCR product with *EcoRI*, the fragment was gel purified and ligated into the *EcoRI* site in the pUHD10-3 vector. The sequence of these constructs were verified by dye terminator DNA sequencing provided by the MSU sequencing facility.

9. **Isolation of RNA**

Cells were grown in 10 cm tissue culture plates until near confluencey and the RNA was harvested. All materials used were treated with DEPC to get rid of RNase or were RNase free from the manufacturer. 2 ml of denaturing solution containing 4M guanidine thiocyanate, 42mM sodium citrate, 0.83% N-lauryl sarcosine and 0.2 mM 2-mercaptoethanol, was added to each plate and gently swirled until cells were lysed. This was transfered to centrifuge tubes and the plates were washed

with another 2 ml, which was added to the same tubes. To these lysed cells was added 0.4 ml of 2 M sodium acetate, pH 4.0. An equal volume of water saturated, pH 5.0, phenol/chloroform/isoamyl alcohol mixture was added and the aqueous phase was separated by centrifuging at 4° for 20 minutes. The RNA in aqueous solution was precipitated with isopropanol overnight at -20°, and pelleted by centrifugation for 30 minutes at 4°. The RNA was resuspended in 2 ml denaturing buffer and precipitated as above. The final RNA pellet was washed with 70% ethanol and taken up in water and stored at -20° .

10. **Northern analysis**

20-30 ug of RNA per lane was run on 1.2% formaldehyde gels using standard procedures (72). Ethidium bromide staining of the gels was used to visualize the RNA before transfer to estimate the quality of the RNA according to the amount of degradation seen. The RNA gels were transferred by capillary blot to MSI magnagraph membrane and the blots were baked for two hours at 80° in a vacuum oven. Radioactive DNA probes were made using the Stratagene random primer kit with incorporation of 32P-dATP (3,000 Ci/mmol; Dupont NEN). Specific DNA fragments used for probes are discussed in the text. Prehybridization and hybridization were carried out in buffer containing 0.263 M Na_2HPO4, pH 7.4,

7% SDS, 1% BSA, and 1 mM EDTA, at 65° overnight. Blots were washed twice in 2X SSC, 0.1% SDS for 15 minutes at 65° , followed by washes in 1X SSC, 0.1% SDS at 65° . The blots were then exposed to film for varying lengths of time and developed.

11. **Micrococcal nuclease digestions**

Nuclei were isolated from cells, digested with micrococcal nuclease, and the genomic DNA was analyzed by electrophoresis using the following procedure. Cells were grown in 10 cm plates until they were near confluencey. The plates were washed with cold PBS then 5 ml of cold PBS was added to each plate and the adherant cells were scraped free with a rubber policeman and transferred to 15 ml centrifuge tubes. The cells were pelleted by centrifuging at 1500 rpm for 2.5 minutes at 4° . The cell pellet was taken up in 10 ml in cold RSB (10 mM Tris-HCl, pH 7.6, 10 mM NaCl, 3 mM MgCl2) and repelleted as above. The cell pellet was then taken up in 1.5 ml RSB, transferred to a microfuge tube and set on ice for 10 minutes, allowing the cells to swell and burst, releasing the intact nuclei. Detergent isolation of nuclei was not used as it led to clumping of the nuclei. The lysed cells were then Dounced 10 times using a glass homogenizer, spun down briefly for 15-30 seconds, and rinsed twice with cold RSB. The quality of the nuclei was checked at this point

by adding 1 ul of trypan blue to 10 ul of nuclei suspension and observing the nuclei with a microscope. The nuclei should be intact with very little cell debri remaining.

At this point the DNA equivalent of the nuclear sample was determined by adding 50 ul of the final RSB wash to 950 ul of 2 N NaCl, vortexing, and setting at room temperature for 10 minutes to allow nuclear lysis. The absorbance at A260 was measured and the DNA equivalent was estimated using 1 OD = 30 ug/ml. The nuclear preps were then taken up at a concentration of 0.5 mg/ml in micrococcal nuclease digestion buffer containing 15 mM Tris -HCl, pH 7.5, 15 mM NaCl, 60 mM KCl, 15 mM 2-mercaptoethanol, 0.15 mM neutralized spermine, 0.5 mM spermidine, 0.34 M sucrose, 1 mM CaCl2, and 0.1 mM PMSF. One nuclear prep from two confluent plates of cells should give 200-500 ug DNA equivalent. Micrococcal nuclease (Sigma) stock is made up at 100 units/ul, and diluted to 1 unit/ul for use in reactions. Each nuclei prep was digested at two concentrations of micrococcal nuclease, 50 units/mg and 300 units/mg by aliquoting each prep into tubes containing 100 -200 ug total DNA equivalent, adding the appropriate amount of micrococcal nuclease, and digesting at room temperature for 20 minutes. Similar levels of digestion will occur by incubation at 37° for 10 minutes. The reaction is stopped by addition of an equal amount of stop buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 15 mM EDTA

and 0.3% SDS. The digested prep is then incubated with 50 ug/ml RNase for 1 hour at 37° , followed by incubation with 50 ug/ml proteinase K for 2 hours to overnight at 37°. The genomic DNA is then extracted with phenol/chloroform in the presence of 1/10 volume 3M NaOAc. The DNA is then precipitated with 100% ethanol, followed by a wash with 70% ethanol, and taken up in 25-50 ul water. The concentration of the DNA is determined by spectrophotometry (1 OD = 50 ug/ml) and 15 ug per lane is loaded and run on a 1.8% agarose or 4% acrylamide gel.

12. **Isolation of nucleoprotein complexes and chromatin fractionation**

Isolation and analysis of mono and di-nucleosome complexes containing DNA, histones and HMG proteins bound was done by a modified version of Albright et al (24). Nuclei were isolated and the DNA equivalent was determined as above. The nuclei were resuspended in micrococcal nuclease digestion buffer at a concentration of 1 mg/ml DNA equivalent in microfuge tubes. Samples were digested with micrococcal nuclease at 150 units/mg DNA by prewarming the nuclei at 37° for 5 minutes, then adding micrococcal nuclease and incubating at 37° for 10 minutes. The reaction was stopped by rapidly cooling the samples on ice and then spun at 15,000 rpm at 4° for 10 minutes. The supernatant or S1 fraction,

which should contain primarily transcriptionally inactive gene sequences, is removed and the pellet is taken up in 50 ul of 2 mM EDTA by first adding 10 ul and slowly dissolving the pellet, then adding 40 ul and gently stirring. This was incubated on ice for 10 minutes and then spun out at 15,000 rpm at 4° for 10 minutes. The supernatant or S2 fraction, which should contain primarily active or potentially active gene sequences, is added to an equal volume of glycerol and is ready to load on a 3.5% acrylamide, 0.5% agarose gel containing 30% gylcerol. The gel is prepared as follows: to a 250 ml erlenmyer flask add 56.8 g glycerol, 4 ml water, 26.2 ml solution B (20% acrylamide, 1% bisacrylamide) and 2.4 ml solution A (400 mM Tris-HCl pH 8.0, 200 mM sodium acetate, 20 mM EDTA, pH adjusted to 8.5 with acetic acid), mix, and set at 55°. In another flask add 750 ml agarose and 37.5 ml water, melt and cool at 55°. Mix the two solutions together, add 120 ul TEMED, 120 ul 20% ammonium persulfate and pour immediately into a 16.5 cm by 16.5 cm vertical gel casting unit using combs with 2 mm by 7 mm wells. When the gel polymerizes, assemble the running aparatus, overlay the wells with 1.6% solution A, 30% glycerol, fill the upper and lower buffer wells with 1.6% solution A and equilibrate at 4°. Immediately before loading samples, rinse the overlay solution from the wells and load the samples. The running

buffer must be rapidly recirculated between the upper and lower buffer chambers during the gel run to avoid pH gradients from forming. Run the gel at 4° overnight at 9 mA. The gels are then stained in running buffer containing ethidium bromide and the nucleoprotein complexes are visualized under UV light and photographed. The various electroporetic forms of nucleosomes are distinguished according to Albright et al (24).

Chapter III

Design and production of chicken HMG14a and HMG17 mutants

Chapter III

INTRODUCTION

The HMG14a and HMG17 chromosomal proteins have remained highly conserved throughout evolution. This suggests that these proteins play an important role in cellular function. Specific residues which are invariant or regions which are the most highly conserved are likely to play a vital role in protein function. Mutation or deletion of invariant residues or conserved regions may alter the protein's ability to function and/or produce a detectable phenotype which would help to better understand the cellular role of the wild type protein. To investigate this possibility, a series of mutations were produced at specific locations in the cDNAs of chicken HMG14a and HMG17.

The DNA binding domain of HMG14 and HMG17 is highly conserved overall and contains a number of invariant residues. Two sequences within this region, P(K,Q)RRSARLSA and KPKKA have remained virtually invariant in all the HMG14 and HMG17 proteins analyzed, from trout to human, representing an evolutionary span of more than 400 million years. A number of these residues have been targeted for site-specific mutagenesis in the cDNAs of the chicken HMG14a

and HMG17.

The region of highest sequence homology among the HMG14 and HMG17 family resides in the 3' untranslated region of the mRNA. Portions of the 3' untranslated region in other eukaryotic nuclear transcripts have been shown to play a role in regulation of transcription, message stability, and mRNA processing (76). Dot matrix analysis comparing the chicken and human HMG14 and HMG17 cDNAs shows a portion of the 3' untranslated region to be highly conserved (56). Deletion mutants covering this region were prepared in the HMG14a and HMG17 cDNAs.

The HMG proteins undergo a number of postranscriptional modifications and the effect of these on protein function is unknown at this time. HMG14 and HMG17 can be poly(ADP) ribosylated, which may be involved in crosslinking of the HMGs to other chromosomal proteins such as the histones (14). Site-specific mutation of the sites of ribosylation on the HMG14a and HMG17 proteins have been prepared to test this hypothesis.

RESULTS

1. Production of site-specific mutations in the DNA binding domain of chicken HMG14a and HMG17

The DNA binding domain of the chicken HMG 14a and HMG17 proteins is located at the amino terminus of the polypeptide and covers amino acids 12 - 40 in HMG14a and 21 - 45 in HMG17. This domain begins, in both proteins, with the invariant sequence P(K,Q)RRSARLSA, with the second amino acid being K in HMG14a and Q in HMG17. This sequence is followed by a 9 amino acid sequence of KPAPPKPEP which is identical in these two chicken proteins but slightly variable among HMG14 and HMG17 from other species. The end of the DNA binding domain is comprised of the invariant sequence KPKKA. The DNA binding domain of these proteins is positively charged overall due to the high number of basic residues, primarily lysine and arginine. Alteration of these residues to uncharged or acidic amino acids will change the charge properties of this region and likely the proteins ability to bind to DNA. The DNA binding region of these HMG proteins also contains an unusually high number of proline residues, 8/29 in HMG14a and 7/24 in HMG17. In addition to being conserved in the HMG proteins, proline residues produce a unique bend conformation in polypeptide chains such that alteration of these sites to small alanine or lysine residues may effect significant changes in the final structure of the

protein.

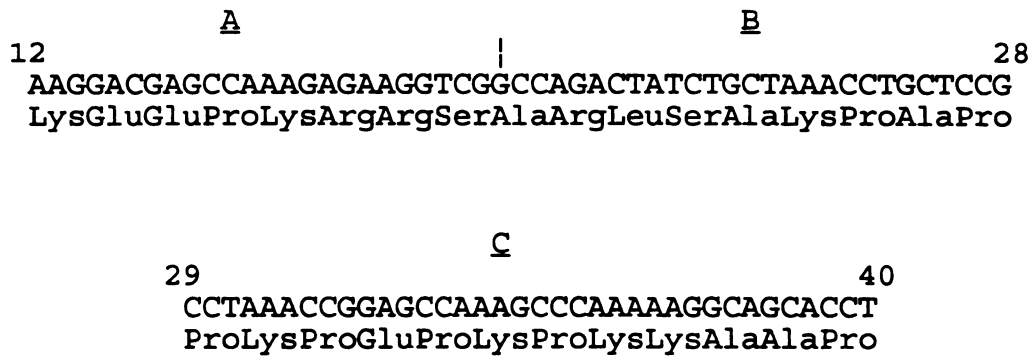
A number of the invariant residues within the DNA binding domain of chicken HMG14a and HMG17 were altered by site -specific mutagenesis of the cDNAs. For ease in presentation, the DNA binding domains of each cDNA can be divided into regions A,B and C in HMG14a and A and B in HMG17. The DNA and amino acid sequence of these proteins and the sites chosen for mutagenesis are illustrated in figures 1 and 2.

The majority of site-specific mutations were prepared by oligonucleotide directed *in-vitro* mutagenesis (81). Individual mutagenic oligonucleotides used are listed in figure 3. The cDNAs were cloned into an M13 phage vector and mutagenesis was performed in this vector. The M13 plasmids containing the cDNAs were transformed into *E. Coli* strain CJ236 which is a *dut, ung* mutant strain. The *dut* mutation inactivates the enzyme dUTPase and results in high intracellular levels of dUTP. The *ung* mutation inactivates uracil N-glycosylase which allows the incorporated uracil to remain in the DNA strand. This results in the transformed plasmid DNA containing a number of uracil residues in place of thymine. The uracil-containing single-stranded DNA was isolated and then used as a template for synthesis *in-vitro* of the complementary strand using a mutagenic oligonucleotide as a primer. The resulting double-stranded DNA was then transformed into *E. coli* strain MV1190, which is wild type

Figure 1 - **DNA and amino acid sequence of chicken HMG14a and HMG17 DNA binding domain**

The DNA and amino acid sequence of the DNA binding domain of HMG14a and HMG17 is shown. For presentation purposes, the DNA binding domains are divided into subregions as shown. Site-specific mutations produced in these regions are labeled according to the subdomain (A, B, or C) in which they are located.

HMG14a



HMG17

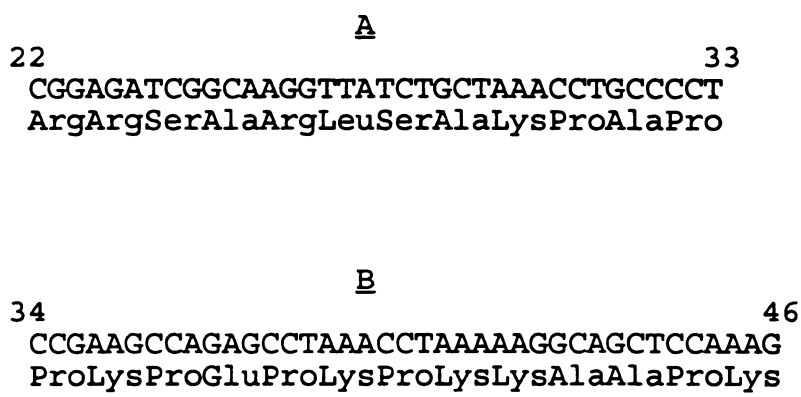


Figure 1

Figure 2 - **Mutations produced in the HMG14a and HMG17 cDNAs**

The individual site-specific mutations made in the HMG14a and HMG17 cDNAs are listed. Each mutant is labeled according to the subregion in the DNA binding domain in which it is located (A, B, or C). The change in codon sequence is shown along with the resulant change in amino acid and location.

Mutant	Codon change	Amino acid change/ Amino acid number
14A1	CCA - CTA	PRO - LEU / 15
14A4	AGG - GGG	ARG - GLY / 18
14B1	AGA - GGA	ARG - GLY / 21
14B3	CCT - GCT	PRO - ALA / 26
14C1	AAA - AAC	LYS - ASN / 30
14C2	AAG - GAG	LYS - GLU / 34
14C3	CCC - CTC	PRO - LEU / 35
14C4	AAA - ACC	LYS - THR / 36
14C5	AAG - ACT	LYS - THR / 37
17A2	AGG - AGC	ARG - SER / 26
17A4	CCT - GCT	PRO - ALA / 31
17B3	CCT - GCT	PRO - ALA / 40
17B6	AAG - AAC	LYS - ASN / 46

Figure 2

Figure 3 - **Oligonucleotides used for site-specific mutagenesis in HMG14a and HMG17 cDNAs.**

For each site-specific mutant produced in HMG14a or HMG17 cDNA, the individual oligonucleotides used are listed. The mismatched nucleotide(s) are indicated by a number or star. Mixed oligonucleotide pools were used in some instances to increase the number and variety of mutants isolated from an individual *in-vitro* mutagenesis reaction.

Mutant	Oligo	Oligonucleotide sequence 5' - 3'
14A1,14A4	14SE-AS	GCCGACC1TC12TTT3GCTCCTC
		1=T:C in 3:1 ratio 2=C:G in 3:1 ratio 3=G:A in 3:1 ratio
	14SE-S	CAAGGAGGAGC1AAA23GA3GGTC
		1=C:T in 3:1 ratio 2=G:C in 3:1 ratio 3=A:G in 3:1 ratio
14B1,14B3	14EB-AS	TAGGCG1AGCAG12TTAGCAGATAGTCC
		1=G:C in 3:1 ratio 2=T:G in 3:1 ratio
	14EB-S	GGCC1GACTATCTGCTAA23CTGCT3CG
		1=A:G in 3:1 ratio 2=A:C in 3:1 ratio 3=C:G in 3:1 ratio

Figure 3

14C1	14C1	CTTTGGCTCCGGGTTAGGC *	* = T-G
14C2	14C2	CTTTTTGGGCTCTGGCTC *	* = T-C
14C3	14C3-S	CAAAGCTCAAAAAGGCAGC *	* = C-T
	14C3-AS	GCTGCCTTTTTGAGCTTTG *	* = G-A
14C4	14C4	AGGTGCTGCCTTGGTGGGCTTTGGC **	* = T-G
14C5	14C5	CTTAGGTGCTGCAGTTTTGGGCTTT **	* = T-G,C-A
17A2	17A2	AGCAGATAAGCTTGCCGA *	* = C-G
17A4	17A4	CGGAGGGGCAGCTTTAGCAGA *	* = G-C
17B3	17B3	CGAGCCTTTTTAGCTTTAGGCTC *	* = G-C
17B6	17B6	CCTTCTCACTCTTGTTTGGAGC *	* = C-G

Figure 3 (con.)

for *dut* and *ung*. This results in the uracil-containing strand being degraded while the non-uracil DNA strand containing the desired mutation is preferentially replicated. An illustration of this *in-vitro* mutagenesis procedure is shown in figure 4. The mutated plasmid DNA was then isolated and sequenced by standard di-deoxy procedures. The primers used for sequencing are listed in figure 5. Following sequence verification of the mutation, the HMG cDNA was digested from M13 and cloned into the appropriate expression vector for use in tissue culture experiments.

Several of the site-specific mutations produced in the HMG14a cDNA were prepared using a polymerase chain reaction (PCR) based strategy (74). In this method, complementary oligonucleotide primers containing internal mismatches at the site of the desired mutation were used in combination with oligonucleotide primers outside of the region to be mutated. Two initial amplifications were made using pC14cw (pCla12 plasmid containing HMG14a cDNA in the sense orientation) as the template DNA, one using outside sense primer A' with the internal antisense mutagenic primer, and one using outside anti-sense primer D' with the internal sense mutagenic primer. These reactions produce two fragments that overlap at the site of mutation by the length of the mutagenic oligonucleotide and that contain the desired mutation. A second amplification is performed using these two

Figure 4 - ***in-vitro* mutagenesis reaction scheme**

The procedure used for the site-specific *in-vitro* mutagenesis reaction is shown. The individual steps involved are described in detail in the text.

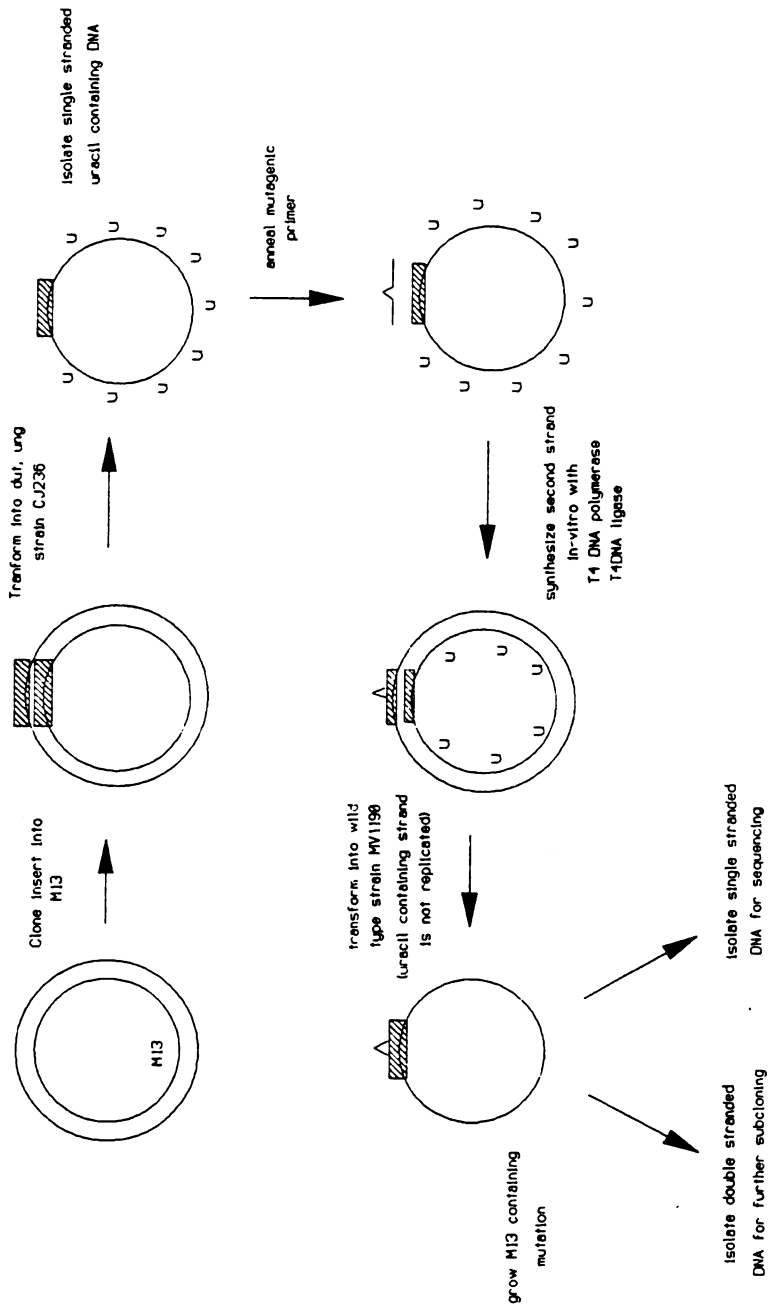

Clone insert into M13
M13
Tranform into dut, ung strain CJ236
Isolate single stranded uracil containing DNA
anneal mutagenic primer
synthesize second strand in-vitro with T4 DNA polymerase T4DNA ligase
transform into wild type strain MV1190 (uracil containing strand is not replicated)
grow M13 containing mutation
Isolate single stranded DNA for sequencing
Isolate double stranded DNA for further subcloning

Figure 4

Figure 5 - **Oligonucleotides used for sequencing mutants in the HMG14a and HMG17 cDNAs**

The oligonucleotide primers used for sequence verification of site-specific mutations in the HMG14a and HMG17 cDNAs are shown and their locations are indicated.

Oligo name	Sequence 5' - 3'	Location of 3' end
14seqpr1(AS)	TGGCTCCTTTCTTCCCTT	nt 287 in cDNA
14seqpr2(AS)	GCTGGTGCCTCATTAGTTTTGG	nt 368 in cDNA
14seqpr3(AS)	GCATCCTCTTTGTTTAG	nt 320 in cDNA
A' (S)	AGGCGTATCACGAGGC	nt 2791 in pCla12 60 bp 5' of polylinker
D' (AS)	TCTGACTTGAGCGTCG	nt 1035 in pCla12 135 bp 3' of polylinker

Figure 5

PCR generated fragments as template DNA and the two outside primers, A' and D'. This produces a DNA fragment that includes the entire cDNA with an internal mutation as well as sequences outside of the convenient cloning sites, EcoRI and ClaI. This fragment is then digested with EcoRI, releasing the mutated cDNA and two small fragments from either end of sizes 60 and 130 bp, which can be visualized on an agarose gel to verify digestion. This procedure eliminates the need to digest the PCR product close to the ends of the fragment which can produce uncohesive ends. An illustration of this PCR based mutagenesis strategy is shown in figure 6. The mutated cDNA was then subcloned back into pCla12 for sequencing verification of the mutation. The mutated cDNA was then subcloned into the appropriate expression vector for use in tissue culture experiments.

2. **Production of deletion mutations in HMG14a and HMG17 cDNAs**

Among the HMG14 and HMG17 cDNAs studied from various species, the 3' untranslated region contains a region which is the most highly conserved portion of the cDNA. A dot matrix analysis of the chicken and human HMG14 and HMG17 cDNAs located this region in the chicken HMG14a and HMG17 at nt 500 - 550 (59). Although the significance of this conservation is unknown, these sequences may be important for

Figure 6 - **PCR mutagenesis reaction scheme**

A schematic of the procedure used for making site-specific mutations in the HMG14a and HMG17 cDNAs is shown. The individual steps involved are described in detail in the text.

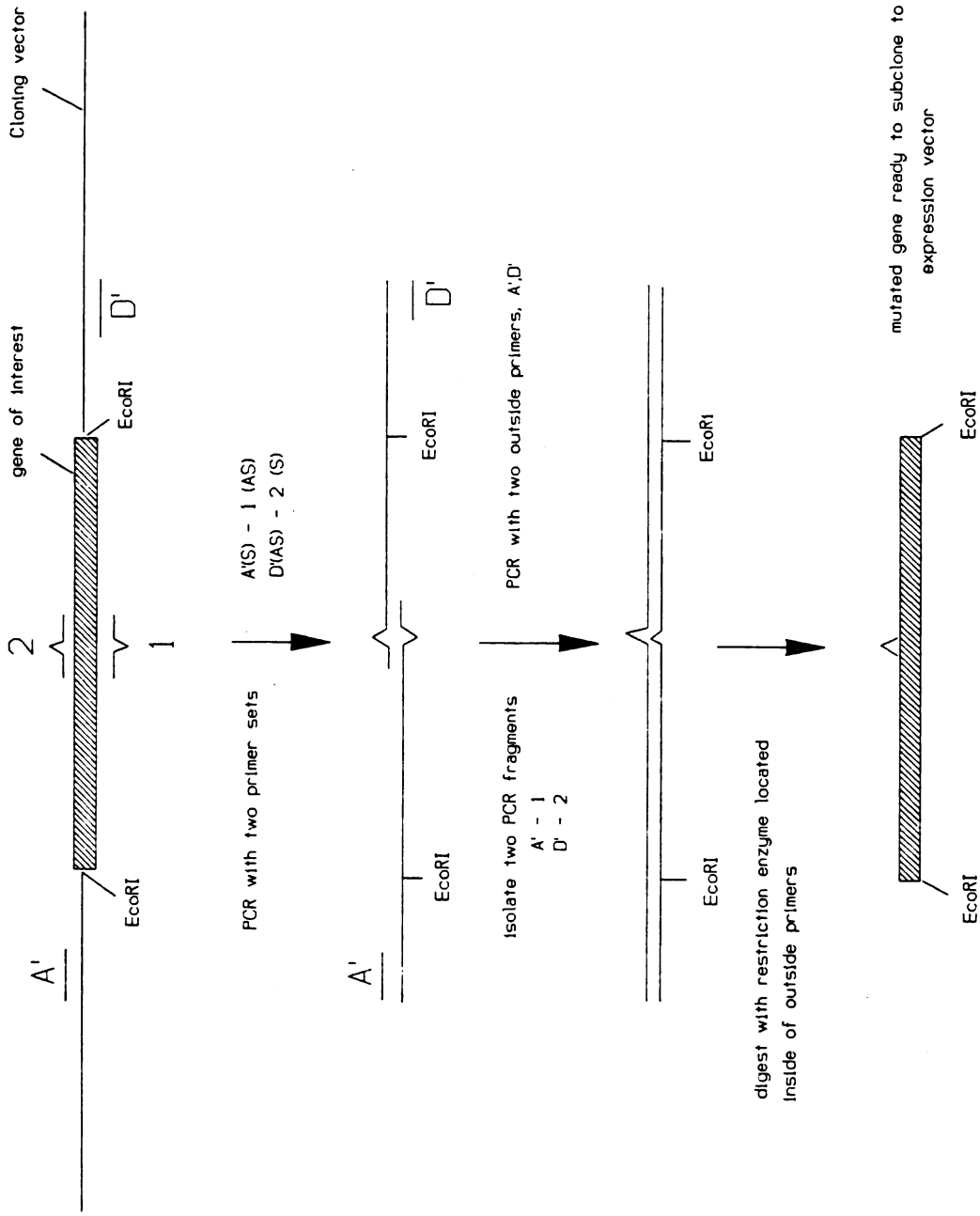

Cloning vector
gene of interest
A'
D'
EcoRI
EcoRI
2
1
PCR with two primer sets
A'(S) - 1 (AS)
D'(AS) - 2 (S)
A'
D'
EcoRI
EcoRI
Isolate two PCR fragments
A' - 1
D' - 2
PCR with two outside primers, A',D'
EcoRI
EcoRI
digest with restriction enzyme located
inside of outside primers
mutated gene ready to subclone to
expression vector
EcoRI
EcoRI

Figure 6

regulation of transcription, proper mRNA polyadenylation and/or mRNA structure and stability. To investigate these possibilities, this region was deleted in both the HMG14a and HMG17 cDNAs.

In HMG14a, this conserved 3'untranslated region was deleted by digesting pC14cw with the restriction enzyme RsaI, which cuts at postions 470 and 664 in the HMG14a cDNA. The internal 196 bp fragment was removed and the plasmid was ligated back together. The result is removal of the highly conserved region and some surrounding sequences, leaving 276 bp of 3' untranslated sequence. In HMG17, this region was deleted by digesting pC17cw with enzymes RsaI and DraI, which cut at positions 477 and 655, respectively, in the cDNA, removing the 178 bp fragment and religating the plasmid back together. This leaves 725 bp of 3' untranslated sequence in the cDNA. These two deletion mutants were then subcloned into expression vectors for use in tissue culture experiments. The location of the 3' untranslated region deletions are shown in figure 7.

3. Production of mutations in HMG14a and HMG17 altering the site of poly(ADP) ribosylation

The HMG14 and HMG17 proteins undergo a number of posttrancriptional modifications. Among these is poly(ADP) ribosylation. The effects of this modification on the

Figure 7A - **DNA sequence and location of mutations in the HMG14a cDNA**

The complete cDNA sequence of chicken HMG14a. The 3' untranslated region deletion defined by RsaI at nt477 and nt 664 is underlined. The location of the nt mutation designed to alter the poly(ADP) riboslyation site is starred. The mutation is G-A and changes the amino acid from Glu(GAA) - Arg(AGA).

HMG14a cDNA

GAATTCCGTCCCCTTCCTCAGGACGCTCGAAAACAGTTTCTCGGCGGTTCCCTTCCTATT
1

TTTTACACCTCTCCCGATCTCTCTATTTGCAGTCAACTATTAAGGTGCAACTATGCCCAA
61

AAGAAAGGCTCCAGCTGAAGGCGAGGCGAAGGAGGAGCCAAAGAGAAGGTCGGCCAGACT
121

ATCTGCTAAACCTGCTCCGCCTAAACCGGAGCCAAAGCCCAAAAAGGCAGCACCTAAGAA
181

AGAAAAGGCAGCAAACGATAAAAAGGAAGACAAAAAGGCAGCAACAAAAGGGAAGAAAGG
241

AGCCAAAGGCAAAGACGAAACTAAACAAGAGGATGCAAAAGAAGAAAACCACTCTGAAAA
301

*

TGGAGATACCAAAACTAATGAGGCACCAGCTGCTGAAGCATCTGATGATAAGGAAGCCAA
361

GTCCGAGTAATGTTAACCCTGCCCTATATCTCCATCATTTGGTATCCGTACCTCCATGCT
421

GTATTGTTAACAGAGAGGAATATTTTTATCAACTATTTTATAAATGCAGGTTTTTTTTAGC
481

ATGAATTTAATTATGGAACATCTTCATCTCGGTTACTTGGGAATTAAATCCCTAACAAAC
541

AAAACAAAACAAAACAAAAAAAAATCATTGTTTTAAATTTGTGATTGTAATAGTTTGTAT
601

GGTACATGGAAAGAATAAGTGGTGGTAGCTTTTGACTTCTGTCAGTGTGTCCCTTTTTGT
661

GTAAGTCATGCTTACAGACTTCAGATTTTAATTTTACCCTTGTATGTGTTGTATGGTTTC
721

TTAAAGTGGGGAGGTCTCAAAACAGATAACTGTGTTAAACATTCCAGTGGTTCTGTGGGT
781

TGCTTTTATAAAGAAGGTGAGCTATTTTCATGAAAAAAAAAAAAAAAAAAACGGAATTC*
841 900

Figure 7A

Figure 7B - **DNA sequence and location of mutations in the HMG17 cDNA**

The 3' untranslated region deletion defined by RsaI at nt 477 and DraI at nt 655 is underlined. The location of the nt mutation designed to alter the poly(ADP) riboslyation site is starred. The mutation is A-C and changes the amino acid from Glu(GAA) to Ala(GCA).

HMG17 cDNA

GAATTCCGCAGCCAGCGCAGCGAGCCGGCCGCCAGCCCCGCCGCGCCGCCCCGCTCTCCC
1

CCTCGGCCCTCCCCCGCTTCTCGCCGCCACCGAGCGAGCCCGGCTGCCCGCCCCCGCCCG
61

CCCCCTCCGCTCGCTCTCTCCCTCCTCGCACAACACACGCACGCGCCGCCCGGAGCTATG
121

CCGAAGAGAAAGGCTGAAGGAGATACCAAGGGCGATAAGGCCAAAGTTAAGGATGAGCCA
181

CAACGGAGATCGGCAAGGTTATCTGCTAAACCTGCCCCTCCGAAGCCAGAGCCTAAACCT
241

AAAAAGGCAGCTCCAAAGAAGAGTGAGAAGGTGCCCAAGGGAAAGAAGGGGAAAGCTGAT
301

*

GCTGGCAAGGAGGGAAACAACCCTGCAGAAAATGGAGATGCCAAAACAGACCAGGCACAG
361

AAAGCCGAAGGTGCTGGTGATGCCAAGTAAAATGTGTGAATTTTTGATAACTGTGTACTT
421

CTGGTGACTGTACAGTTTGAAATACTATTTTTTATCAAGTTTTATAACAATGCAGAATTT
481

TGTTTTACTTTTTTTAAGCTATGTTGTTAGCACACAGACCGCTTTGTTGTTGTGTTTTGA
541

GGGGGGGGCAGTGGGACAAACGTCACTTAATCTGTTTCTTGGAACCTAAATTTTAAAAGT
601

TTACCCCTTCCCAGTTTTTTAGAAGGACTCTTCCTAAATGGAGCAGGAAGGGATTCCTTC
661

GTGCTGCACACCTCTTCCGTTTTGTGGACCGCATCAGAGTGAACGGAAGCTCCCGAGATG
721

CCTGTTGCCAACTTCAGAACTGCAGTTTGCAGTGCCCTCTGCGTTTCCTTTCATGCCCTC
781

Figure 7B

CCTTTTTGCCTAGAGCCTATCACTCCGAAATACAGCAGACATGGCATGTTGGGACTCACC
841

ACTCTAAATGCATTGTCAGGTGATCTGGACTTCTGGTGTCTAATTTGGGATATAATAGCT
901

CTAAAAGGAGCTGCATTTCCTCTTTCATATTGTAGATCTACAGATTAAGGAATCTGCAGT
961

TTTTAATTTTTCCTCGCAAAGCTAGGGTAGATTTGTGAAGAGTTGTTAAACAACATGCTA
1021

AATGTGAAAGTGTCCGCCCTCACTCTAAACATTTCCCTCTACAAGTATACAAAAATGAAG
1081

ATTTGTCGGTTTTATAGCAACCTTTATGTTTGGGTAGTCCATGAAGGGAGGGGAGTTTGA
1141

CAGTTGTTGTAAAATGTTGCAGATTGTAGCCCATGTCCTGCCTAAATTACCATGATTGTT
1201

TATGAAAAGTACCTTTAATAAAGCTGGATACGGTTTGGCTTGGAAAAAAAAAAAAAAAAA
1261

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACGGAATTC
1321 1360

Figure 7B (con.)

function of these HMG proteins is unknown but the putative site of this modification has been localized to glutamic acid residues 81 in HMG14a and 70 in HMG17. Poly(ADP) ribosylation has been implicated in crosslinking of other chromosomal proteins such as the histones and modification of these residues may lead to alterations in the proteins ability to functionally interact with the nucleosomes. These residues have been altered by oligonucleotide directed *in-vitro* mutagenesis by others in our lab. In HMG14a, this glutamic acid residue (GAA) was changed to arginine (AGA), and in HMG17 to alanine (GCA). The locations of these mutations are shown in figure 5.

DISCUSSION

For the production of site-specific mutations in the HMG14a/17 cDNAs, the in-vitro mutagenesis procedure proved to be the most expedient and efficient. Once a stock of single -stranded uracil containing DNA was isolated for each clone in M13, the procedure can be completed in a matter of hours using standard reagents and techniques. The frequency of obtaining the desired mutation was 50% or higher. For each mutagenesis reaction performed, 4 colonies were picked and their DNA was sequenced over a region spanning approximately 200-300 bp within the cDNA and including the mutated region. Consistantly, at least 2 of these clones contained the

desired mutation. Undesired mutations, presumably produced from mistakes made during replication by the DNA polymerase, were noticed at a very low but detectable level. Out of approximately 34 clones sequenced, 3 mismatches were detected in the region sequenced, 2 in one clone and 1 in another.

Several mutations were produced using the PCR based strategy. This procedure is slightly more time consuming as it requires two consectutive PCR reactions and the reaction conditions must be optimized for each set of mutagenic primers used. In addition, the amplified bands frequently required gel isolation/purification prior to the second amplification and prior to subcloning with the final DNA fragment. Subcloning with PCR generated fragments is often problematic, in part due to the periodic addition of extra nucleotides on the fragment ends. This problem can be avoided by restriction enzyme digestion of the final product (provided there are convenient sites available) designed to cut off the ends and produce a fragment with the appropriate ends for subcloning. This was accomplished with our clones by using the EcoRI sites located internal to the external primer binding sites. The location of the external primers in pCla12 were chosen such that digestion with EcoRI releases two end fragmnets of detectable size on a mini gel for verification of digestion. The restriction enzyme digest was often done directly in the final PCR reaction mix. For efficient

subcloning of the resultant fragment, it is neccessary to thoroughly extract the final reaction mix with phenol/chloroform followed by an ether extraction to completly rid the DNA fragment of excess nucleotides and primers, Taq polymerase (which may add nucleotides to fragments ends) and all traces of phenol (which can prohibit subsequent ligation). The efficiency of this PCR based mutagenesis strategy was 100% as detected by sequencing or restriction enzyme digest of the final subcloned product. Although only a few clones were sequenced, no undesired mutations were detected in the region sequenced.

Chapter IV

Expression of chicken HMG14a and HMG17 proteins in tissue culture cells

Chapter IV

INTRODUCTION

Chicken chromosomal binding proteins HMG14a and HMG17 bind to and interact with nucleosomes in regions of active or potentially active chromatin. While much research has focused on the specific physical interactions involved between the HMG proteins and the nucleosome, very little is known of the cellular function of this binding. We have chosen to address this question by ectopically expressing wild type and mutant chicken HMG14a and HMG17 proteins in abnormal amounts in tissue culture cells and determining the resultant effects on the cellullar phenotype and the structure of the cellular chromatin.

Each nucleosome contains two potential binding sites for the HMG14/17 proteins yet the amount of these proteins present in the cell is regulated such that only about one out of every ten nucleosomes actually has HMG14/17 bound. Overexpression of wild type and mutant HMG14/17 proteins may cause the exogenously expressed proteins to occupy new sites in cellular chromatin and, in the case of mutant proteins, to outcompete the endogenous HMG for binding sites on the nucleosome. Any effects of this aberrant binding or other influences the exogenously expressed protein may have on the cell phenotype and/or its chromatin structure can then be observed.

As with many cellular proteins, the level of HMG14/17 present in the cell may be regulated. Bustin et al. attempted to transiently express exogenous human HMG14 and HMG17 in COS cells and reported that while relatively high levels of mRNA could be detected, HMG14/17 protein levels were limited to 2 -4 times that seen in untransfected cells (65). It is unclear if this level of exogenous HMG14/17 expression is a result of cellular regulation or, rather, of the limitations of the system used. We have undertaken a series of experiments designed to over express chicken HMG14a and HMG17 proteins for the first time in transformed quail fibroblast cells.

RESULTS

Transfection of QT6 tissue culture cells

QT6 cells are a chemically transformed quail fibroblast cell line used in our expression experiments (67). This cell line was chosen for these experiments due to our familiarity with its properties, and more important, its ease of transfection. There are very few immortalized, easily transfectable avian cell lines available.

Transfections of QT6 cells were performed using standard calcium phosphate transfection procedures. The transfection efficiency using these cells is consistently high. Routinely, 10 ug of DNA per 10 cm tissue culture plate with 10^6 cells gave 500-1000 stable colonies following selection. In some instances, it was neccessary to reduce the amount of DNA transfected to ensure isolation of single colonies. Single colonies were isolated by picking well isolated colonies from the transfected plates.

Transfection using expression vector TFANEO

TFANEO is an expression vector designed by Federspiel et al (71). The wild type and mutant cDNAs for chicken HMG14a and HMG17 were cloned into a *ClaI* restriction site located between two long terminal repeat (LTR) sequences derived from the Schimdt-Rupin avian Rous sarcoma retrovirus. The upstream LTR provides promoter and enhancer functions while the

downstream LTR polyadenylates the resultant transcript. This vector also contains a neomycin gene driven by the chicken *B* -actin promoter for selection by G418 in culture. A map of TFANEO is shown in figure 1.

The wild type cDNAs for HMG14a and HMG17, deletion mutants 14-R and 17-R, and site-specific mutants 14A1, 14A4, 14C1 and 17B6 were cloned into TFANEO, as described in chapter II and III, and transfected into QT6 cells. Stable pooled and single colonies were isolated from these transfections. Northern and western analysis were used to determine the level of HMG14a or HMG17 mRNA or protein ectopically expressed in these clones.

Northern analysis of HMG14/17 transfected cell lines

Figure 2 shows a northern analysis of total mRNA isolated from stable colonies transfected with the wild type HMG14a cDNA cloned into TFANEO in the sense and antisense orientation, and a 3'- untranslated region deletion mutant. Portions of the 3'untranslated region of the HMG14/17 proteins are very highly conserved suggesting they play an important role in gene function. It is possible that sequences within this region are important for mRNA processing, message stability, or lack thereof. Clones labeled 14-RS are HMG14a cDNA with a 194 bp deletion in the conserved portion of the 3'-untranslated region. The results in figure 2 show the presence of two transcripts in the

Figure 1 - **Map of expression vector TFANEO**

Expression vector TFANEO is 6.8 Kb in length. Transcriptional controls are located in the expression cassette consisting of two LTR sequences derived from the Schmidt-Rupin Avian Rous sarcoma retrovirus. HMG14/17 cDNAs were cloned into the single *ClaI* restriction site located between the LTRs. A neomycin gene driven by the chicken *B*-actin promoter allows for selection by G418 in tissue culture. The arrows indicate the direction of transcription.

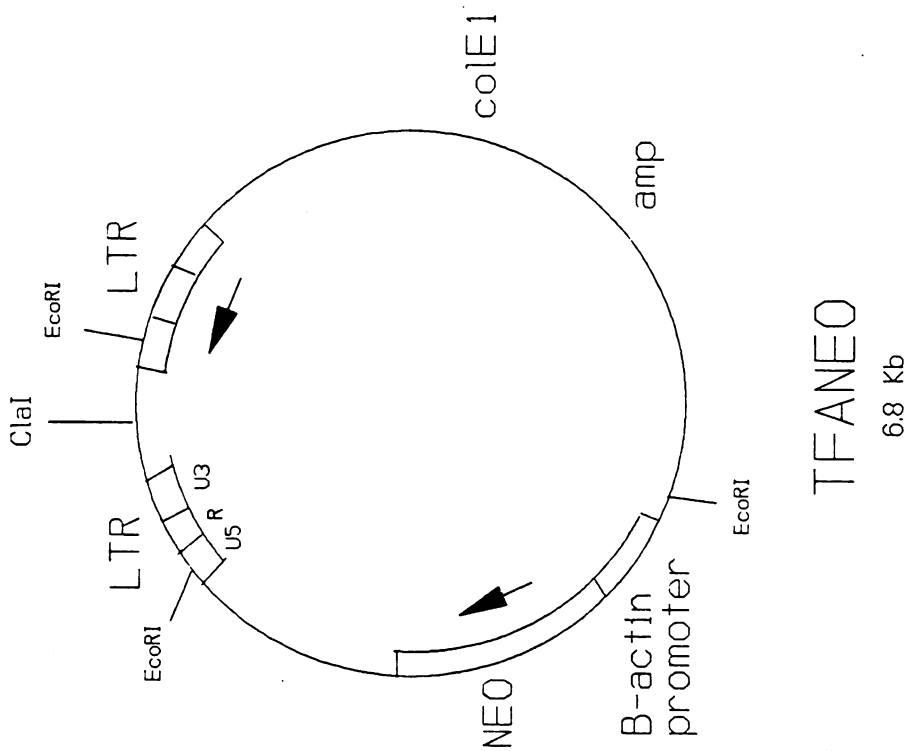

colE1
amp
TFANEO
6.8 Kb
EcoRI
B-actin
promoter
NEO
LTR
EcoRI
U5 R U3
LTR
EcoRI
ClaI

Figure 1

Figure 2 - **Northern analysis of total RNA isolated from QT6 cells transfected with HMG14a cDNA constructs**

Total RNA was isolated from stable cell lines transfected with TFANEO containing various chicken HMG14a cDNA clones. 30 ug of total RNA was loaded per lane. The blot was probed with a random labeled full length HMG14a cDNA. Two transcripts are seen, the upper band is presumably exogenous HMG14a mRNA and the lower, the endogenous quail HMG14. It is possible the lower band represents an additional alternatively polyadenylated exogenous transcript. Transcript sizes were estimated by comparison to a simultaneously run DNA standard. QT6 is the untransfected parental cell line. VC3 is QT6 transfected with TFANEO alone. 14S and 14AS are single colonies stably transfected with TFANEO containing the wild type chicken HMG14a cDNAs cloned in the sense and antisense orientation. 14-R are single colonies stably transfected with TFANEO containing HMG14a 3'-untranslated region deletions.

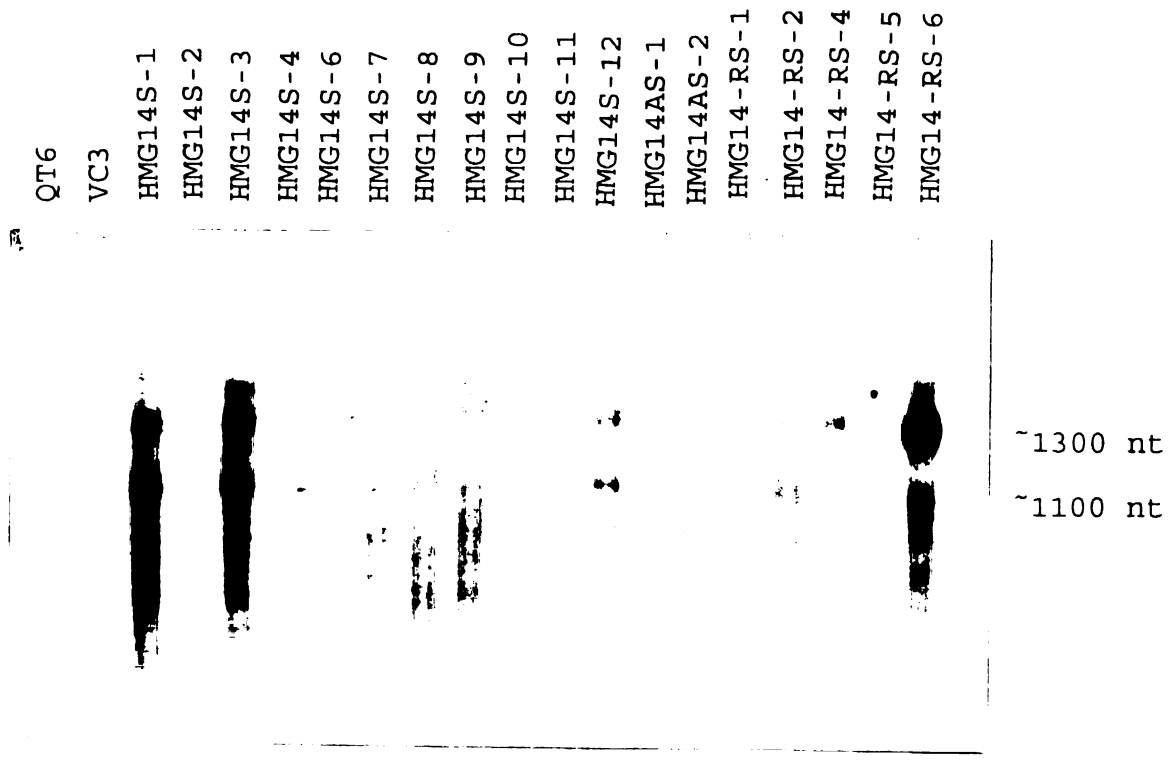

QT6
VC3
HMG14S-1
HMG14S-2
HMG14S-3
HMG14S-4
HMG14S-6
HMG14S-7
HMG14S-8
HMG14S-9
HMG14S-10
HMG14S-11
HMG14S-12
HMG14AS-1
HMG14AS-2
HMG14-RS-1
HMG14-RS-2
HMG14-RS-4
HMG14-RS-5
HMG14-RS-6
~1300 nt
~1100 nt

Figure 2

transfected cell lines. Presumably, the lower band represents endogenous quail HMG14 message and the upper band, the exogenously transfected mRNA. Ethidium bromide staining of the RNA gel shows all lanes contain equivalent amounts of good quality, undegraded rRNA. Clones 14-RS-6 and 14S-1 and 14S-2 express the highest amount of exogenous mRNA. The control cell lines QT6 and VC3, representing untransfected QT6 cells and QT6 cells transfected with the TFANEO vector only, show barely detectable levels of quail HMG14 specific mRNA. This suggests that quail HMG14 is expressed at very low levels in these cells or that the endogenous quail HMG14 message is sufficiently different in sequence from the chicken HMG14a so as to not hybridize efficiently with the chicken HMG14a cDNA probe used in these experiments. Most of the 14S and 14AS single colonies screened did not show elevated levels of exogenous HMG14a mRNA. Approximately 15 additional wild type and deletion mutant HMG14a clones were screened for mRNA expression in this manner (data not shown). The 3 overexpressing clones shown in figure 2 remain the highest expressors of exogenous chicken HMG14a mRNA seen in the cells tested. Interestingly, the 3'untranslated deletion mutant 14-RS-6 showed the highest levels of expression. This suggests that the deleted region may be involved in the regulation of mRNA expression levels.

Figure 3 shows a similar experiment screening mRNA

Figure 3 - **Northern analysis of total RNA isolated from QT6 cells transfected with HMG17 cDNA constructs**

Total RNA was isolated from stable cell lines transfected with TFANEO containing various chicken HMG17 cDNA clones. 30 ug of total RNA was loaded per lane. The blot was probed with full length HMG17 cDNA. Two transcripts are seen, the upper band is presumably exogenous HMG17 mRNA and the lower, the endogenous quail HMG17. It is possible that the lower band represents an additional alternatively polyadenylated exogenous transcript. Transcript sizes were estimated by comparison to a simultaneously run DNA standard. QT6 is the untransfected parental cell line. VC3 is QT6 transfected with TFANEO alone. 17S and 17AS are single colonies derived from stable transfections of TFANEO containing wild type chicken HMG17 cDNA cloned in the sense and antisense orientation. 17-R are single colonies from transfections of TFANEO containing HMG17 cDNA 3'-untranslated region deletions.

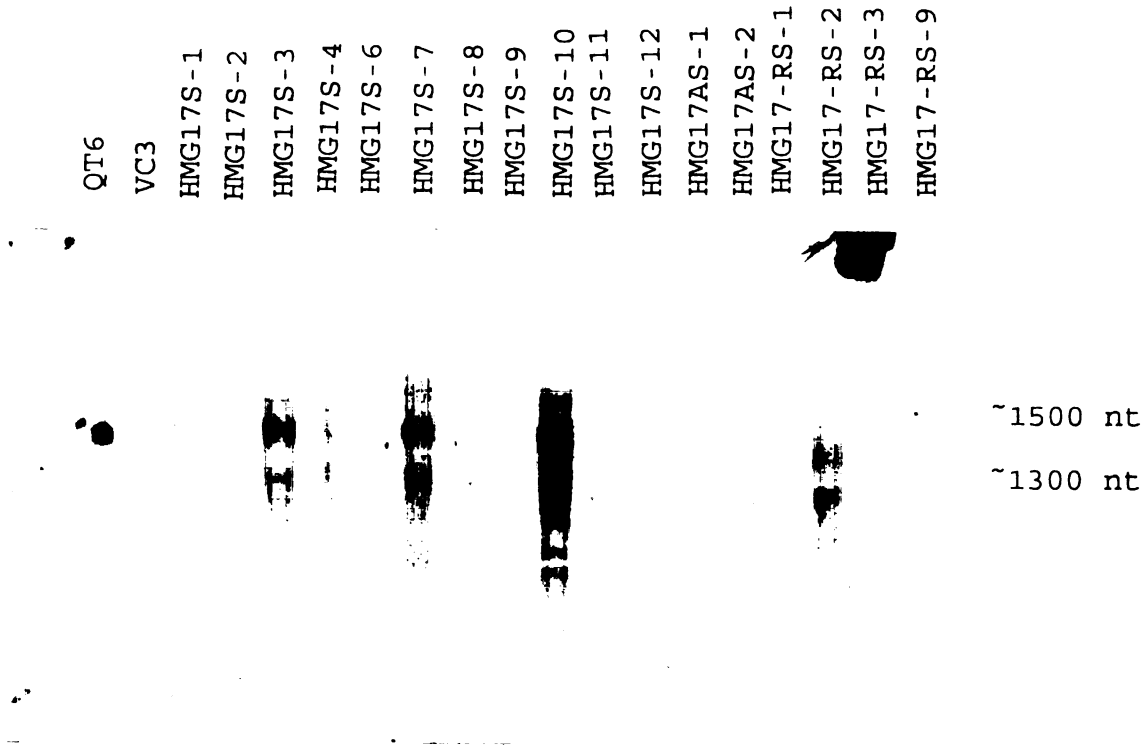

QT6
VC3
HMG17S-1
HMG17S-2
HMG17S-3
HMG17S-4
HMG17S-6
HMG17S-7
HMG17S-8
HMG17S-9
HMG17S-10
HMG17S-11
HMG17S-12
HMG17AS-1
HMG17AS-2
HMG17-RS-1
HMG17-RS-2
HMG17-RS-3
HMG17-RS-9
~1500 nt
~1300 nt

Figure 3

expression levels in quail cells stably transfected with chicken HMG17 cDNA clones. 17S 1-12 and 17AS 1,2 represent single colonies tranfected with wild type HMG17 cDNAs cloned into expression vector TFANEO in the sense and antisense orientation. 17-RS 1,2 are HMG17 cDNA deletion mutants with a 178 bp deletion in the conserved portion of the 3'untranslated region. QT6 and VC3 are the control cell lines described above. As with the HMG14a transfected cell lines, most of the single colonies do not show overexpression of the exogenous message. Clones 17S 3,7 and 10 were the highest overexpressers seen in all cell lines screened, including an additional 10 clones tested but not shown.

The fact that many stable colonies did not show overexpression of the exogenously transfected cDNAs may be explained in several ways. Very high levels of HMG14/17 mRNA may be lethal or deleterious to the QT6 cells, thereby naturally selecting only those clones which express low or moderate HMG14/17 mRNA levels. Alternatively, some of the integrated vectors may have the HMG14/17 cDNA disrupted in the quail cell genome. Most likely, we believe that the amount of total HMG14/17 mRNA present in the cell is regulated at the transcriptional level by the cell, by sequences present in the 3'-untranslated region or by rapid degradation of excessively transcribed HMG14/17 mRNA.

Western analysis of HMG14/17 transfected cell lines.

The amount of mRNA transcribed from a particular gene does not neccessarily reflect the amount of protein translated and present in the cell. To address the question of whether the cell lines shown to overexpress HMG14 or HMG17 mRNA also overexpress the HMG protein, western analysis was performed with a variety of cell lines.

HMG proteins were isolated from the various cell lines by a 5% perchloric acid (PCA) extraction of the nuclei. This isolation procedure has been shown to selectively extract the HMG proteins and histone H1. An antibody that recognizes both HMG14 and HMG17 was used to detect these proteins on the blots (66). The antibody was kindly provided by Michael Bustin. The antisera was elicited in rabbits using a synthetic peptide corresponding to the highly conserved 30 amino acid stretch of the human HMG17 DNA binding domain. This antibody has been shown to recognize both HMG14 and HMG17 specifically, with a slightly higher affinity for HMG17.

Figure 4 shows a western blot of protein samples from several of the stable cell lines described above. In all lanes bands corresponding to the HMG14 and HMG17 proteins are present and run at the appropriate mobility. To control for the total amount of protein loaded per lane, the gels were stained with coomassie blue and the lanes were normalized to

Figure 4 - **Western blot of QT6 transfected cell lines**

HMG14/17 proteins were selectively isolated from cells with a 5% PCA extraction procedure. Approximately 1 ug of protein was loaded per lane. Proteins were run on 15% SDS-PAGE gels and electroblotted to PVDF membrane. HMG14/17 proteins were detected with an antibody specific to the conserved DNA binding domain of these proteins. The upper band is HMG14 and the lower band is HMG17. QT6 is the untransfected parental cell line. 14-RS-6,8 are two single colonies derived from transfection with HMG14a cDNA 3'-untranslated region deletion mutants. 17S-7,10 are two single colonies derived from transfection of wild type HMG17 cDNA. Clones 17S-7 and 17S-10 show 5 and 3 times overexpression of HMG17 as compared to the control cell lines.

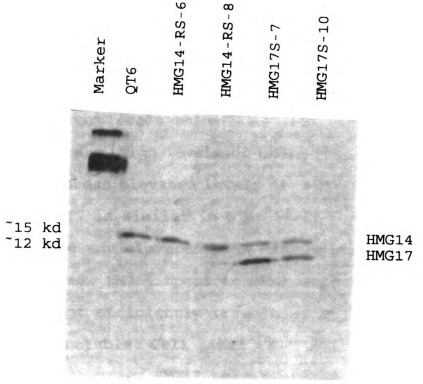

Marker
QT6
HMG14-RS-6
HMG14-RS-8
HMG17S-7
HMG17S-10
~15 kd
~12 kd
HMG14
HMG17

Figure 4

the slower migrating histone H1 band. The HMG14/17 bands seen on the radiographs were optically quantitated using AMBIS imaging software. The QT6 control cell lane in figure 4 clearly shows that more HMG14 is expressed in these cells than HMG17. We see this result consistently with our western blots, however published reports from M.Bustin et al. have shown that in human cell lines, HMG17 is expressed at slighlty higher levels than HMG14 (65). Cell line 14-RS-6, shown to express high levels of HMG14 specific mRNA does not appear to express elevated levels of HMG14 protein. The ratio of HMG14/HMG17 is similar in QT6, 14-RS-6 and 14-RS-8, which does not show any elevated mRNA expression. This suggests that the excess HMG14 specific mRNA expressed in cell line 14-RS-6 is not efficiently translated or the translation product is unstable. Cell lines 17S-7 and 17S-10, which were shown to overexpress HMG17 specific mRNA clearly show significantly higher levels of HMG17 protein as compared to the control cells. The level of total HMG17 expressed in these cells is consistently 3-5 times that seen in the control QT6 cells. Interestingly, 17S-10 displayed slightly higher levels of HMG17 mRNA than 17S-7 yet 17S-7 expresses slightly higher levels of HMG17 protein. In these cell lines it appears that this overexpression of HMG17 protein may be at the expense of HMG14. The amount of total HMG14 + HMG17 protein, normalized to histone H1, is roughly equal in the HMG17 overexpressing lines and the control

cells. This suggests that the cell may be regulating the total amount of HMG14/17 protein allowed in the cell.

Figures 5-7 show representative western blots from a variety of stable cell lines transfected with wild type or mutant HMG14/17 cDNAs. The wild type and deletion mutant clones have been described above. 17B-6 and 14A-1 are site specific mutants in the DNA binding domains of the respective HMG cDNAs. These mutant clones are described in detail in chapter III. Figure 5 shows clones 14S-2 and 14S-3 do not overexpress HMG14 protein. 14S-3 was shown in northern blots to express elevated levels of HMG14 specific mRNA while clone 14S-2 did not. This again illustrates the lack of correlation with HMG14 clones between mRNA and protein expression levels. Figures 5 and 6 show protein expression screening from a total of 20 single colonies derived from a stable transfection of mutant 17B-6. There are no obvious overexpressors of HMG17 protein. Figure 6 also confirms the overexpression of HMG17 protein in clones 17S-7 and 17S-10. Figure 7 illustrates HMG14/17 protein levels seen in 7 clones derived from mutant 14A-1. None of these clones express excess amounts of HMG14. In addition to the samples shown in western blots shown here, another 29 clones were screened in this manner for HMG14/17 protein expression levels. These include clones derived from site-specific mutants 14A-1, 14A-4, 14C-1, 17B-6 and wild type cDNAs. Of

Figure 5 - **Western blot of QT6 transfected cell lines**

HMG14/17 proteins were isolated from stable cells lines transfected with various HMG14/17 cDNA clones. Approximately 1 ug of protein was loaded per lane on a 15% SDS-PAGE gel. The proteins were detected with an antibody specific to the conserved DNA binding domain of these proteins. The upper band is HMG14 and the lower is HMG17. QT6 is the untransfected parental cell line. 14S-2,3 are two single colonies derived from transfection of wild type HMG14a cDNA. 17B6 clones are single colonies derived from transfection of a HMG17 cDNA containing the B6 site-specific mutation in the conserved DNA binding domain. There are no apparent overexpressors of HMG14 or HMG17 in this panel.

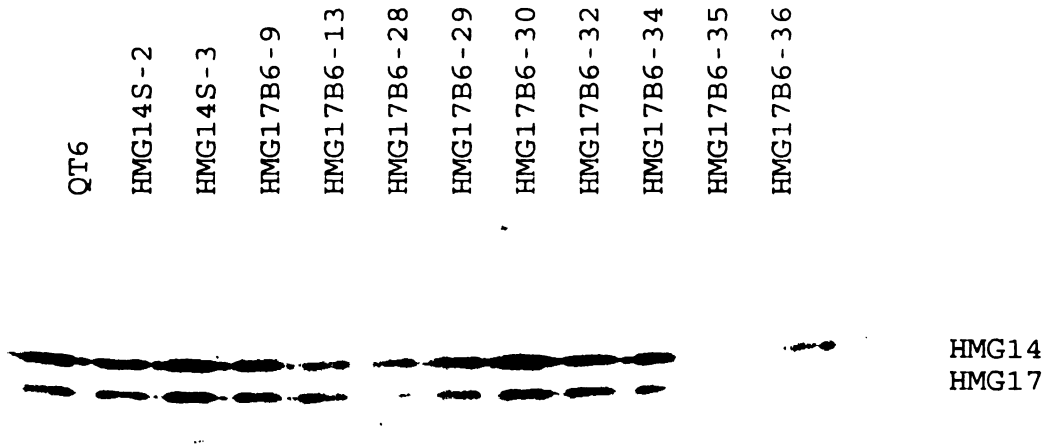

QT6
HMG14S-2
HMG14S-3
HMG17B6-9
HMG17B6-13
HMG17B6-28
HMG17B6-29
HMG17B6-30
HMG17B6-32
HMG17B6-34
HMG17B6-35
HMG17B6-36
HMG14
HMG17

Figure 5

Figure 6 - **Western blot of QT6 transfected cell lines**

HMG14/17 proteins were isolated from stable cell lines transfected with chicken HMG17 cDNA clones. Approximately 1 ug of protein was loaded per lane on a 15% SDS-PAGE gel. The proteins were detected with an antibody specific to the conserved DNA binding domain of these proteins. The upper band is HMG14 and the lower band is HMG17. QT6 is the untransfected parental cell line. 17S-7,10 are two single colonies derived from transfection of a wild type HMG17 cDNA. 17B6 clones are single colonies derived from a transfection of HMG17 cDNA containing the B6 site-specific mutation in the conserved DNA binding domain. Clones 17S-7 shows overexpression of HMG17 5 times and 17S-10 2 times that of the control cells. None of the HMG17B6 mutants show overexpression of HMG17.

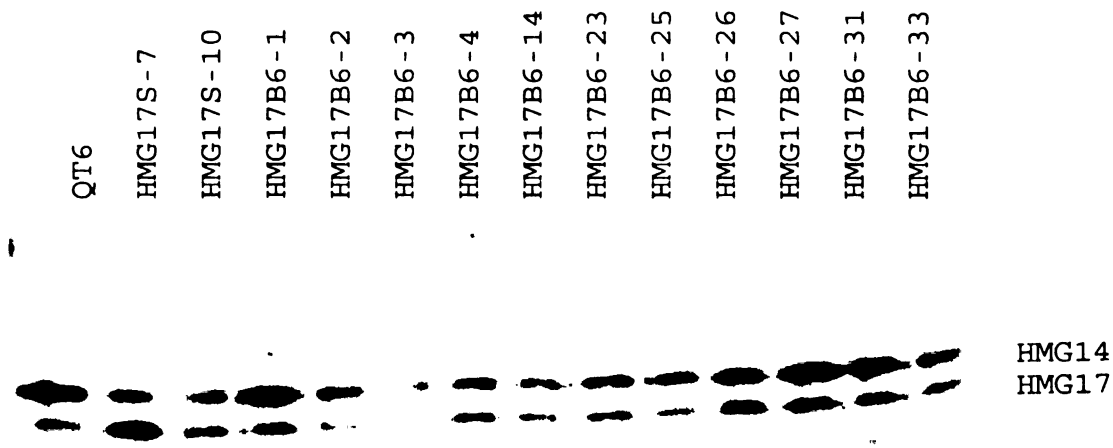

QT6
HMG17S-7
HMG17S-10
HMG17B6-1
HMG17B6-2
HMG17B6-3
HMG17B6-4
HMG17B6-14
HMG17B6-23
HMG17B6-25
HMG17B6-26
HMG17B6-27
HMG17B6-31
HMG17B6-33
HMG14
HMG17

Figure 6

Figure 7 - **Western blot of QT6 transfected cell lines**

HMG14/17 proteins were isolated from stable cell lines transfected with chicken HMG14a cDNA clones. Approximately 1 ug of protein was loaded per lane on a 15% SDS-PAGE gel. The proteins were detected with an antibody specific to the conserved DNA binding domain of these proteins. The upper band is HMG14 and the lower band is HMG17. The bands visible below HMG17 are probably HMG14/17 degredation products. The band visible between HMG14 and HMG17 may be HMG14b which is expressed at low levels in avian cells. QT6 is the untransfected parental cell line. Clones 14-RS-6,8 are two single colonies derived from a transfection of HMG14a cDNA with a 3'-untranslated region deletion. 14A1 clones are single colonies derived from a transfection of HMG14a cDNA containing the A1 site-specific mutation in the conserved DNA binding domain. There are no apparent HMG14 overexpressors in this panel.

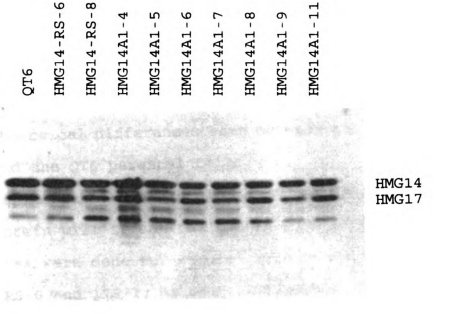

QT6
HMG14-RS-6
HMG14-RS-8
HMG14A1-4
HMG14A1-5
HMG14A1-6
HMG14A1-7
HMG14A1-8
HMG14A1-9
HMG14A1-11
HMG14
HMG17

Figure 7

all the clones screened, clones 17S-7 and 17S-10 were the only ones that showed significantly overexpressed levels of HMG14 or HMG17 protein.

Growth rate and morphology of HMG14/17 transfected cell lines

To examine the possible morphological effects on the QT6 cells due to overexpression of HMG14 or HMG17 mRNA or protein, various stable cell lines were microscopically photographed and/or examined visually. There was no apparent gross morphological differences seen between the cells lines examined and the QT6 parental cells.

It is possible that overexpression of the HMG14 or HMG17 mRNA or protein will affect the growth rate of the cells. Growth curves were done for several cell lines including clones 14-RS-6 and 17S-7. As seen in figure 8, 14-RS-6 grew at a significantly slower rate than the control cells. The doubling time was approximately 48 hours for the 14-RS-6 clone compared to less than 24 hours for QT6. As shown above, clone 14-RS-6 expresses high levels of HMG14a mRNA but not HMG14a protein. The decrease in growth rate seen with this clone may be influenced by this overexpression but it also could be a chance effect of the intregration site of the exogenous DNA in the host cell chromosome. Clone 17S-7, which overexpresses HMG17, showed a doubling time similar to QT6.

Figure 8 - **Growth curve for control and transfected QT6 cells.**

A 6 day growth curve for untransfected QT6 cells and 3 cell lines transfected with TFANEO driven constructs. All time points were done in duplicate. Plates were seeded with equal numbers of cells on day 0 and duplicate plates were counted for cell number at each time point. QT6 is the untransfected parental cell line. VC3 (vector control 3) is QT6 transfected with TFANEO only. 14-RS-6 is a transfected cell line with the HMG14a cDNA containing a 3'-untranslated region deletion. 17S-7 is a cell line transfected with the HMG17 cDNA in the sense direction.

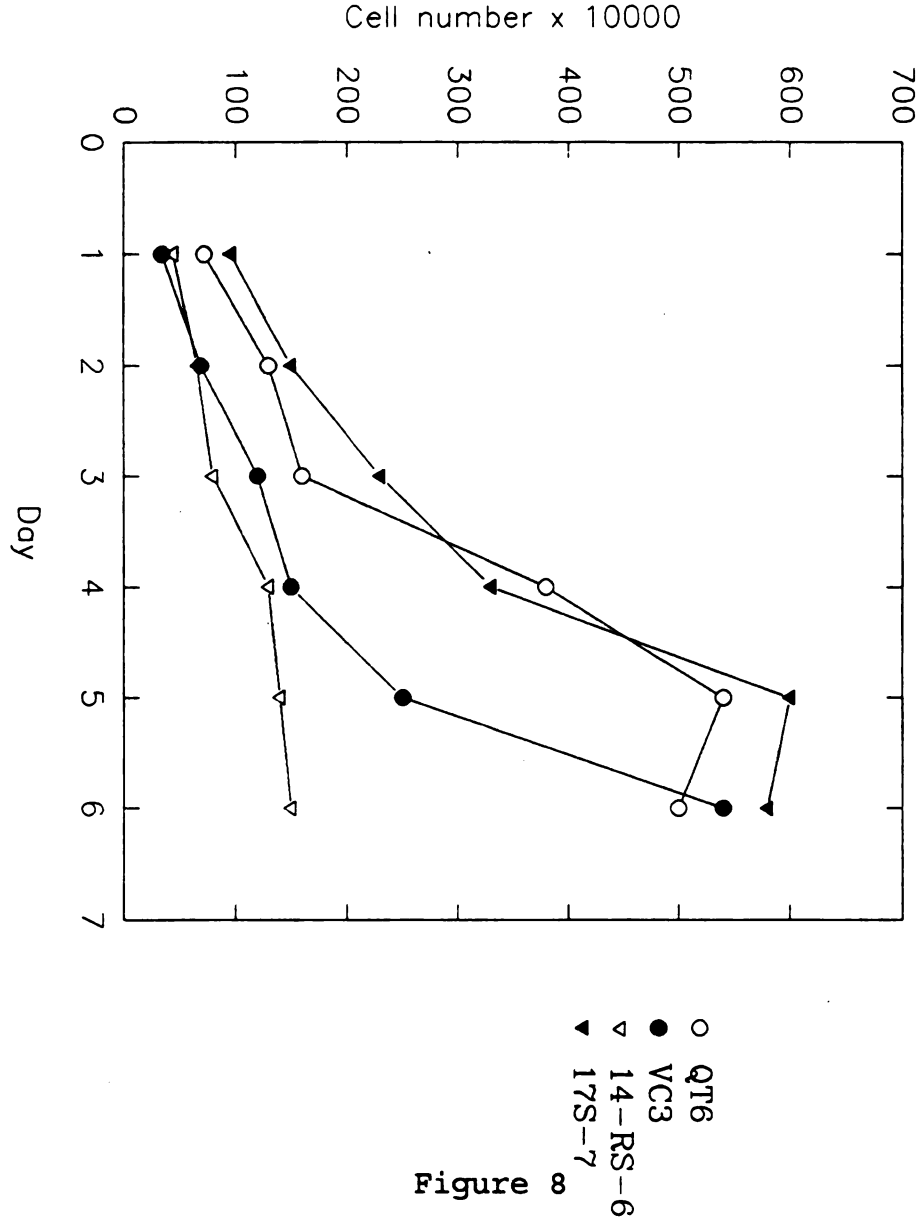

Cell number x 10000
0
100
200
300
400
500
600
700
Day
0
1
2
3
4
5
6
7
QT6
VC3
14-RS-6
17S-7

Figure 8

Discussion

The primary objective of transfecting QT6 cells with a vector expressing the HMG14/17 chromatin binding proteins is to achieve a level of over expression sufficient to induce detectable changes in the cell. The retroviral LTRs in expression vector p779NGneo contain transcriptional promoter and enhancer elements which should constitutively transcribe high levels of mRNA. Despite screening numerous transfected clones for mRNA expression, very few actually demonstrated elevated levels of exogenously expressed mRNA. Two wild type HMG14a clones, 14S-1 and 14S-3, show moderate HMG14 mRNA over expression. Likewise, two wild type HMG17 clones, 17S-7 and 17S-10 show HMG17 mRNA overexpression. Interestingly, clone 14-RS-6 demonstrated the highest levels of exogenously expressed message. This clone contains a deletion in the highly conserved portion of the 3'-untranslated region. Most likely, this region exerts an effect on mRNA stability or RNA processing. However, this region may influence expression in other ways. Sequences in the 3'-untranslated region as well as in the coding region of several genes have been found to influence transcription (68). The DNA or nascent RNA from these regions may have the ability to form secondary structures which can cause the RNA polymerase to pause, sometimes resulting in premature termination. This termination has been shown to occur in T-rich regions of untranscribed sequences, such as is seen in the HMG14/17 3'

-untranslated regions (68). It is also possible that DNA or RNA binding proteins can bind to this secondary structure and influence its interaction with the enzyme. These mechanisms could result in a level of post-initiation transcriptional control. It is possible that the highly conserved portion of the 3'-untranslated region deleted in clone 14-RS-6 functions in such a way. If so, then deletion of this region may remove a negative regulatory element, resulting in increased levels of detectable mRNA as seen in the results.

Although clone 14-RS-6 expresses high levels of HMG14 mRNA, there is no apparent increase in protein expression as compared to control cell lines. This result is consistent with evidence demonstrating that sequences in the 3'untranslated region of some transcripts can positively (and negatively) influence translational efficiency (69, 85). The deleted sequences in the HMG14a cDNA 3'-untranslated region have been highly conserved throughout evolution and therefore may be involved in this type of post-transcriptional control. Deletion of these sequences may lead to an RNA which is inactive or repressed in translation.

Two clones were shown to overexpress HMG17. Clones 17S-7 and 17S-10 show elevated levels of HMG17 3-5 times that seen in the control cells. This maximum level of stable expression is consistent with the only published account of exogenous expression of these HMGs, that being in transient

transfections as shown by M. Bustin et al (65). This group expressed human HMG14 and HMG17 under the control of the SV40 promoter in COS cells. They found that while HMG14/17 mRNA expression was increased 17-50 fold, the protein levels increased only 3 fold. However, unlike those of Bustin's group, our results show that with substantial increases in HMG17, there is a decrease in the amount of HMG14. These results suggest that there may be a limit on the total amount of HMG14/17 protein in the cell, or at least occupying sites on chromatin. HMG14/17 not within the chromatin structure could be rapidly degraded. Our inability at this point to isolate clones which express higher levels of HMG14/17 may indicate that such excesses are deleterious to the cell. Alternatively, excess HMG14/17 protein may just be rapidly degraded or may repress its own translation by some unknown mechanism.

To establish whether QT6 cells can transiently tolerate greater excesses of these HMG proteins, we next expressed the HMG14/17 cDNAs from a regulated promoter. The results from these experiments are described below.

Results

Transfection of QT6 cells with tetracycline regulated expression system

The HMG14a wild type cDNA and the HMG14a coding region only were cloned into expression vector pUHD10-3 as described in the legend to figure 9 (73). pUHD10-3 contains a minimal human cytolomegalovirus (CMV) promoter fused downstream of multiple *E.Coli* tetracycline operator sequences such that transcription is activated upon binding of a hybrid activator protein to the operator sequences. An SV40 polyadenylation cassette placed downstream of the polylinker region polyadenylates the transcribed message. The activator protein is a hybrid construct containing the VP16 activation domain (73) fused to the DNA binding domain of the tetracycline repressor, *TetR*. The hybrid activator is expressed from a second vector, pUHD15-1 (73). The presence of tetracycline in the media represses transcription by complexing with the VP16-*TetR* and blocking its interaction with the promoter/operator sequences. A map of pUHD10-3 and pUHD15-1 is shown in figure 9. For selection of stable transfectants, these two vectors were co -transfected with the vector TFANEO which provides resistance to G418 in tissue culture. To increase the chance of all three vectors being taken up by the transfected cells, ten fold more pUHD10-3 and pUHD15-1 DNA than TFANEO was used in the stable transfections (10 ug each of pUHD10-3 and pUHD15-1 and 1 ug of TFANEO were used per plate in most

Figure 9 - **Map of vectors used in tetracycline regulatable expression system**

A. Maps of vectors pUHD10-3 and pUHD15-1 (73). pUHD10-3 contains a minimal human CMV promoter fused downstream to multiple *E.Coli* tetracycline operator sequences. The HMG cDNA constructs were cloned into the *EcoRI* restriction site in the polylinker. pUHD15-1 contains the CMV promoter and enhancer sequences which drive expression of the VP16-*TetR* activator protein sequences. Both vectors contain the ampicillin resistance gene and the colE1 orgin of replication.

B. Maps of vector pUHD10-3 containing the full length HMG14a cDNA and the HMG14a coding region portion of the cDNA. The full length 900 bp *EcoRI* HMG14a cDNA fragment from plasmid pC14cw was ligated into the *EcoRI* site in the polylinker of pUHD10-3 by standard techniques (72). The coding region of the HMG14a cDNA is contained within a 340 bp *HincII* fragment which was isolated from plasmid pC14cw. This fragment was blunt end ligated into pUHD10-3 which had been digested with *EcoRI* and end filled with Klenow using standard techniques (72).

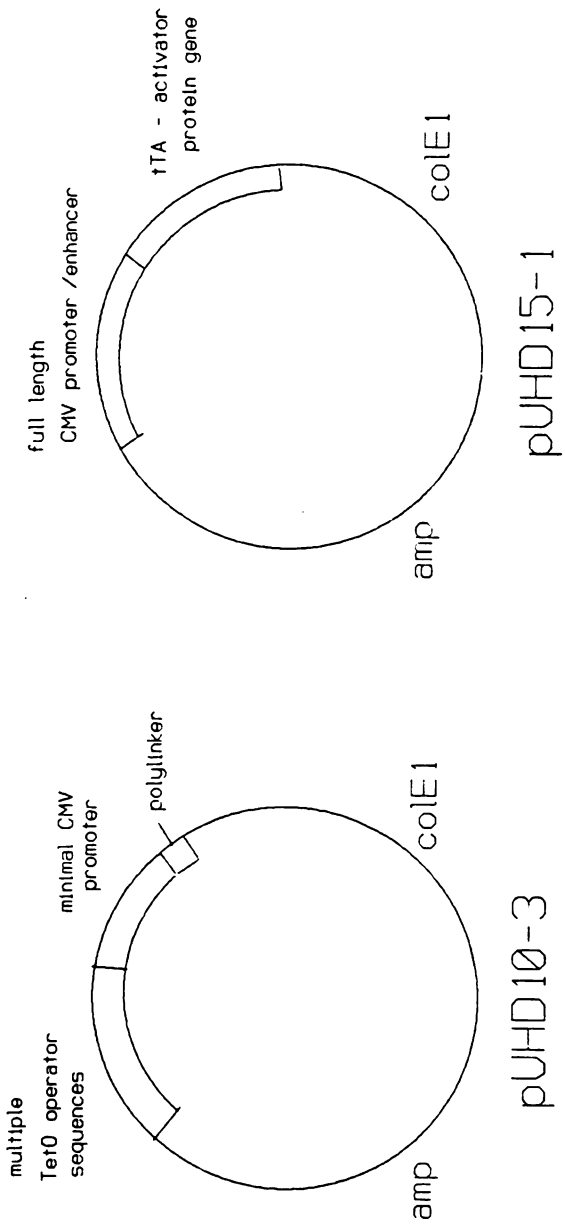

full length
CMV promoter /enhancer
tTA - activator
protein gene
colE1
amp
pUHD15-1
multiple
TetO operator
sequences
minimal CMV
promoter
polylinker
colE1
amp
pUHD10-3

Figure 9A

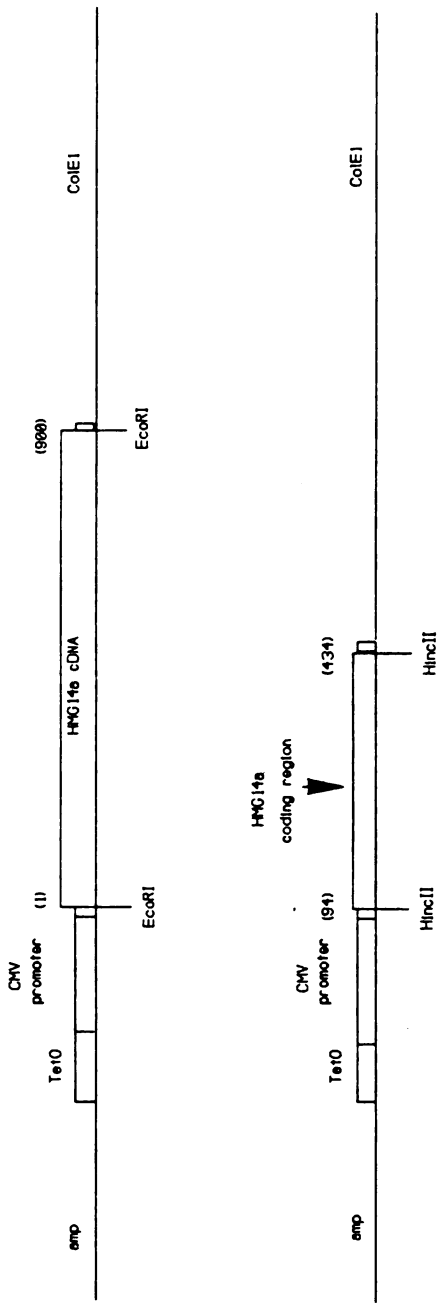

ColE1
(900)
EcoRI
HMG14a cDNA
(1)
EcoRI
CMV promoter
Tet0
amp
ColE1
(434)
HincII
HMG14a coding region
(94)
HincII
CMV promoter
Tet0
amp

Figure 9B

transfections). The various HMG14a clones were transiently and stably transfected into QT6 cells as described in earlier in this chapter. Protein was isolated from the transfectants and western analysis was used to determine the amount of HMG14a expressed in these clones. Protein isolation, gel conditions, loading controls, and antibody detection were done as previously described in this chapter.

Western analysis of HMG14a transfected cell lines

Initial transient transfections were done to establish whether the tetracycline regulatable vector system would express HMG14a in our QT6 tissue culture system. Figure 10 shows a western analysis of two such transfection experiments. Lane 1 is a mock transfection with no DNA, and lane 2 is a co-transfection of pUHD15-1 and pUHD10-3 containing the HMG14a coding region only. The transfections were done in the absence of tetracycline which should allow maximal expression of the HMG clone. The HMG14 band in lane 2 is at least 3.5 times greater than the corresponding endogenous band in the mock transfection. This suggests that the HMG14a clone is being expressed at easily detectable levels using the 2 vector expression system in our QT6 cell line. This is despite the fact that in this, as in any other transient system, many cells may not have been successfully transfected with either or both DNAs necessary for

Figure 10 - **Western analysis of protein isolated from cell lines transiently transfected with the coding region of HMG14a**

HMG14/17 proteins were isolated from cells 24 hours post-transfection with a 5% PCA extraction as described in chapter II. Approximately 1 ug of protein was loaded per lane on a 15% SDS-PAGE gel. The HMG proteins were detected with an antibody specific to the conserved DNA domain of these proteins as described in chapter II. The upper band is HMG14 and the lower is HMG17. Transfections were done in the absence of tetracycline which should allow maximal transcription of the HMG14-containing clone. Lane 1 represents the mock-transfected QT6 and lane 2 represents QT6 transfected with 10 ug of pUHD10-3/HMG14a coding region and 10 ug of pUHD15-1. In lane 2, HMG14 is being expressed at least 3.5 times that seen in the control cells. The levels of expression were optically quantitated using AMBIS imaging software. The level of HMG14 protein expression determined in lane 2 is a minimal value due to the intensity of the band which is out of the linear analyzing range of the AMBIS software.

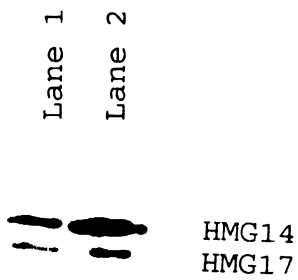


Figure 10

regulatable HMG14 expression.

Single and pooled colonies were isolated from stable transfections of HMG14a wild type and coding region only clones in pUHD10-3 co-transfected with pUHD15-1 and TFANEO. Figures 11 and 12 show western blots of HMG protein isolated from various cell lines derived from these transfections. The HMG protein samples were isolated with a 5% PCA extract of total cells. Single colonies were grown in the presence of tetracycline and split to 4 plates each. The media was changed to tetracycline free media and protein was isolated from each clone at 0, 6, 12 and 24 hours. The time course was done to determine the kinetics of protein expression over the first 24 hours of promoter induction. Surprisingly, there was no increase in HMG14 expression seen in any of the single colonies shown as well as in 22 other single colonies and 2 pooled colonies not shown. Protein isolated from clones after 6 and 14 days in tetracycline free media showed similar results (data not shown). The lack of exogenous HMG14a overexpression seen with the stable cell lines tested could be due to some type of cellular regulation of the HMG protein levels, or the absence of one or more of the 3 vectors used in the transfection.

The stable transfectants were isolated in the presence of G418 in the tissue culture media; thus, the vector TFANEO must be present in all the stable cell lines. To determine

Figure 11 - **Western analysis of protein isolated from cell lines stably transfected with the coding region of HMG14a**

HMG14/17 proteins were isolated from cells with a 5% PCA extraction as described in chapter II. Approximately 0.5 ug protein was loaded per lane on a 15% SDS-PAGE gel. The HMG proteins were detected with an antibody specific to the conserved DNA domain of these proteins. Stable colonies were grown in the presence of tetracycline and protein samples were isolated 0, 6, 12, and 24 hours after the media was changed to tetracycline free. Clones HMG14H-6 and HMG14H-7 are two single colonies isolated from a stable co-transfection of the coding region of the HMG14a cDNA cloned into pUHD10-3 (H signifies the *HincII* restriction fragment of the HMG14a cDNA which contains the coding region) along with pUHD15-1 and TFANEO. No perceptible increase in HMG14 expression was seen over the first 24 hours in either clone.

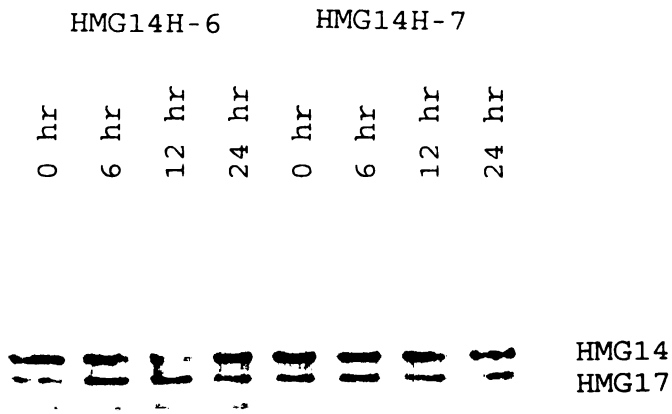

HMG14H-6
HMG14H-7
0 hr
6 hr
12 hr
24 hr
0 hr
6 hr
12 hr
24 hr
HMG14
HMG17

Figure 11

Figure 12 - **Western analysis of protein isolated from cell lines stably transfected with the HMG14a cDNA**

HMG14/17 proteins were isolated and detected as described in figure 11. Protein samples were taken at 0, 6, 12, and 24 hours after the media was changed to tetracycline free. Clones HMG14E-7, HMG14E-8, HMG14E-9 and HMG14E-10 are single colonies isolated from a stable transfection of the full length HMG14a cDNA cloned into pUHD10-3 (E signifies the *EcoRI* restriction fragment containing the full length HMG14a cDNA) along with pUHD15-1 and TFANEO. No perceptible increase in HMG14 expression was seen over the first 24 hours in any of these clones.

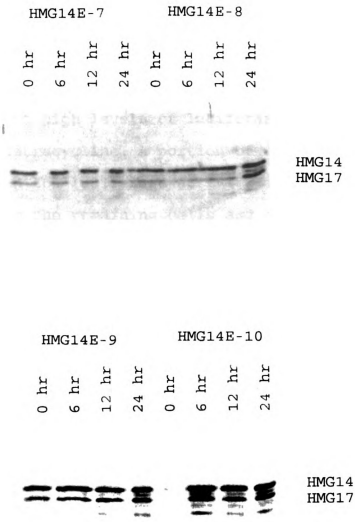

HMG14E-7
HMG14E-8
0 hr
6 hr
12 hr
24 hr
0 hr
6 hr
12 hr
24 hr
HMG14
HMG17
HMG14E-9
HMG14E-10
0 hr
6 hr
12 hr
24 hr
0 hr
6 hr
12 hr
24 hr
HMG14
HMG17

Figure 12

whether the VP16-*TetR* fusion activator protein was being expressed from vector pUHD15-1 in these cell lines, pUHD10-3 containing a luciferase gene cloned in place of the HMG14a clone was transiently transfected into pools of the stable cell lines (73). Expression of the fusion activator protein should lead to high levels of luciferase activity in the absence of tetracycline. A portion of the transfected cells were analyzed for luciferase activity and protein was isolated from the remaining cells and analyzed by western blot for HMG14a expression. The results of this experiment are shown in figure 13. As can be seen from the western blot, there is no apparent increase in HMG14a expression in any of the pools tested. However, the data from the luciferase assay clearly show that the luciferase gene is induced in tetracycline free media indicating that the activator protein is, in fact, expressed and is functional. Thus, the tetracycline regulatable system appears to be operative in the clones tested, yet there is no apparent increased expression of the HMG14a protein in these clones.

Discussion

It is possible that the cell will tolerate only a limited amount of HMG protein expressed in the cell. Possible mechanisms for this regulation have been discussed in the previous discussion section. The use of a regulatable expression vector to express exogenous HMG14a may be one way

Figure 13 - **Western analysis and luciferase assay of protein isolated from stably transfected clones transiently transfected with pUHD10-3 containing the luciferase gene.**

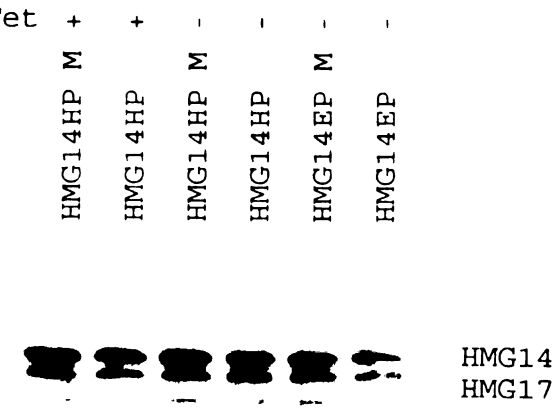
A. Western analysis

HMG14/17 protein was isolated and detected as previously described from pooled colonies of stable transfectants. Tetracycline (4 ug/ml) was present or absent in the media as indicated by +/- tet. HMG14E represents the full length HMG14a cDNA cloned into pUHD10-3. HMG14H represents the coding region only portion of the HMG14a cDNA cloned into pUHD10-3. P indicates pooled stable transfectants.The pools for each construct were transiently transfected with the luciferase gene cloned into pUHD10-3. Mock represents the stable colonies mock transfected with no luciferase DNA. There was no increase in HMG14 expression in any of the clones tested upon tetracycline removal.

B. Luciferase assay

Samples from each of the transient transfections were analyzed for luciferase activity. Elevated levels of luciferase activity were seen in transfections carried out in the absence of tetracycline.

A.



B.

Clone	Tet	Mock	Luciferase activity
HMG14HP	+	Y	0.00
HMG14HP	+	N	293.60
HMG14HP	-	Y	0.00
HMG14HP	-	N	1767.00
HMG14EP	+	Y	0.00
HMG14EP	+	N	500.10
HMG14EP	-	Y	3.16
HMG14EP	-	N	2378.00

Figure 13

to overcome potential feedback regulation of the HMG protein levels. A sharp induction of expression may transiently elevate the amount of HMG protein before the cell can adjust to the normal levels. However, our experiments suggest that either this is not the case or if feedback regulation of excess HMG expression exists, it is most likely a rapid process. Bujard et al. showed that expression of luciferase from vector pUHD10-3 in the presence of pUHD15-1 in tetracycline free media reached maximum levels at 12-24 hours (73). Our time course experiments showed that HMG14a expressed from pUHD10-3 did not show elevated levels of expression from 0-24 hours nor after 6-14 days. It remains a formal possibility that none of the pUHD10-3/HMG14a DNA was taken up by the cells along with pUHD15-1 and TFANEO during the stable transfections. However, a number of pooled colonies were tested for protein expression and the chance that this particular DNA would be preferentially excluded in all the transfected cells seems highly unlikely. It is also possible that all the transfecting pUHD10-3/HMG14a DNA copies disrupted in the HMG14a cDNA sequence (or in nearby flanking sequences required for expression) during integration into the QT6 genome. However, this again seems rather unlikely.

Due to the fact that we can not distinguish between the endogenous and exogenously expressed HMG14 protein with the antibody used in these experiments, there remains the

possibility that the protein detected is in fact the exogenous chicken HMG14a and that the endogenous quail protein is being repressed. Alternatively, our HMG14a cDNA clones may contain unknown elements which make them unrecognizable or inefficient in translation. To address these last possibilities, HMG14a cDNA clones were prepared which contain an additional 24 bp at the 3' end of the coding region that, when translated, code for eight amino acid residues recognized with high specificity by a commercial antibody. These clones and their expression are discussed in chapter V.

Chapter V

Expression of HMG14aFLAG in tissue culture cells

Chapter V

INTRODUCTION

Detection of total HMG14 and HMG17 protein expressed in our stably transfected tissue culture cell lines with an antibody specific to the DNA binding domain of these proteins does not allow us to distinguish between the endogenous quail HMG14 and the exogenously transfected chicken HMG14a. Our previous results show that very few transfected clones express elevated amounts of the transfected protein and we cannot be sure that this additional expression is, in fact, due to the exogenously transfected HMG cDNA, or rather, due to aberrant endogenous HMG expression. While it is possible that the inability to isolate numerous highly expressing HMG14/17 clones is due to an unknown cellular regulatory mechanism which limits the amount of HMG14/17 protein allowed in the cell, we cannot be sure the exogenous HMG14/17 clones are translated efficiently or remain stable once translated. To date, none of the point mutations made in HMG14 or HMG17 has resulted in an obvious change in electrophoretic mobility.

To address these concerns, we prepared HMG14a wild type and mutant cDNA clones that contain an additional 24 bp at the 3' end of the coding region such that an additional 8 amino acids are fused to the C-terminus of the translated protein. These eight amino acid residues are recognized with high affinity by a commercially available antibody, FLAG M2 (80).

This allows us to specifically detect the exogenously transfected HMG14 and therefore address the issues raised above.

The C-terminus of the HMG14a protein is highly acidic. It is possible that the residues located here (or at least their negative charge) directly participate in the protein's ability to function in the cell. Though no function has been assigned to this region of the protein, it is thought that it may be involved in binding to histones or to other HMG molecules on the nucleosomal surface. The 8 amino acid FLAG sequence added to the end of the coding region is also highly acidic and its overall charge properties are similar to the HMG14 C-terminus. It is hoped that the additional sequences will not disrupt the protein's ability to bind to the nucleosome.

RESULTS

Transfection and expression system

Wild type and mutant HMG14aFLAG clones were stably and transiently transfected into QT6 cells as previously described. The FLAG constructs were cloned into the expression vector pUHD10-3 and expression in culture was regulated by the presence or absence of tetracycline in the media as previously described in chapter IV. Protein isolation and western blotting were done as previously described except for the use of the anti-FLAG M2 antibody in addition to the anti-HMG antibody.

Construction of the HMG14aFLAG cDNA clones

HMG14a cDNA cloned into plasmid pCla12 (pC14cw) was used as a template in a PCR scheme to prepare the wild type HMG14aFLAG construct. The basic PCR scheme used was the same as for the site-specific mutagenesis described previously using complementary internal primers containing the additional FLAG sequences instead of primers containing single base mismatches. The internal FLAG primers used are listed in figure 1. The PCR scheme is illustrated in chapter III, figure 6. The two outside primers used, A' and D' are described in chapter III, figure 5. For production of the site-specific mutant HMG14aFLAG constructs, the wild type HMG14aFLAG in pCla12 was used as a template and the internal

Figure 1 - **Oligonucleotide primers for PCR production of HMG14FLAG constructs**

Oligo name	Sequence 5' - 3'
Wild type	
FLAG1-AS	CTTGTCATCGTCGTCCTTGTAGTCCTCGGACTTGGCTTCC **********
FLAG2-S	GGACGACGATGACAAGTAATGTTAACCCTGCC **********
	* = FLAG sequence
Mutant	
14C3-S	CAAAGCTCAAAAAGGCAGC
14C3-AS	GCTGCCTTTTTGAGCTTTG
14A2-S	GAGCCAAACAGAAGG
14A2-AS	CCTTCTGTTTGGCTC
14B1-S	GTCGGCCGGACTATCTGC
14B1-AS	GCAGATAGTCCGGCCGAC

mutagenic primers used are listed in figure 1. Portions of the coding regions of the resultant FLAG constructs were verified by DNA sequence analysis (Michigan State University sequence facility). The wild type and mutant HMG14aFLAG cDNA fragments produced in the PCR scheme were digested with *EcoRI* (see chapter III, figure 6) and cloned into the *EcoRI* site in the pUHD10-3 polylinker for use in transfection experiments. Figure 2 shows an illustration of the complete HMG14aFLAG cDNA clone as well as the DNA and amino acid sequence of the C-terminal region of HMG14a with and without FLAG sequences.

Transient expression of HMG14aFLAG constructs

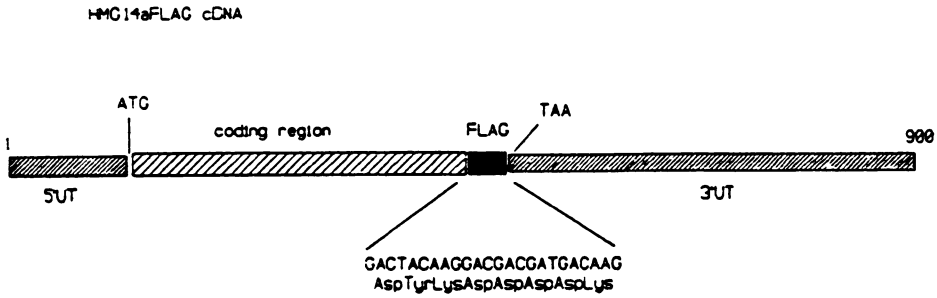
Initial transient transfections of the HMG14aFLAG in expression vector pUHD10-3 were done to establish whether the protein would be expressed and could be detected with the anti-FLAG M2 antibody specific to the 8 amino acids added to the 3' end of the HMG14a coding region. 10 ug of the wild type HMG14aFLAG/pUHD10-3 DNA and 10 ug of vector pUHD15-1 (expressing the transcriptional transactivator protein) were transfected onto plates of QT6 cells at approximately 60% confluency. Duplicate transfections were performed in the presence or absence of tetracycline (4.0 ug/ml) in the media. The HMG14aFLAG protein should be expressed in the absence of tetracycline. HMG protein was isolated with a 5% PCA

Figure 2 - **Schematic of HMG14aFLAG construct and comparison of C-terminus of HMG14a with and without FLAG sequences**

A. Schematic illustration of HMG14a wild type cDNA with fused FLAG sequences at the 3' end of the coding region. The DNA and amino acid sequence of FLAG is shown.

B. Comparison of the C-terminal residues of HMG14a coding region with and without FLAG sequences added starting at residue #72 (out of 104 in HMG14a). Acidic and charged residues are labeled. (* = acidic residue, + = basic residue) Amino acid abbreviations are listed in appendix.

A.



B.

	HMG14a	EDAKEENHSENGDTKTNEAPAAEASDDKEAKSE
	HMG14aFLAG	EDAKEENHSENGDTKTNEAPAAEASDDKEAKSEDYKDDDDK
		** +*** + * * + * * **+* + ** +*****+

Figure 2

extraction at 48 hours and the protein was electrophoresed on 15% or 18% SDS-PAGE gels. Detection of FLAG containing proteins was done using anti-FLAG M2 antibody purchased from Kodak-IBI. This antibody recognizes the FLAG protein sequence with high affinity when it is located at the C-terminus of a polypeptide (for N-terminal FLAG constructs, another antibody, anti-FLAG M1 can be used). Figure 3 shows a western blot of a transient transfection with the cells grown in the presence or absence of tetracycline. As expected, in the transfection containing tetracycline in the tissue culture media, no HMG14aFLAG was detected (lane 2). In the transfection containing no tetracycline in the media, the HMG14aFLAG protein is clearly expressed and is detected as a single band migrating at the expected mobility of approximately 14 kd (lane 2). This result demonstrates that the HMG14aFLAG clone is expressed and translated efficiently. Because the cDNA clone is the same as used in all previous experiments, with the exception of the extra FLAG sequences at the 3' end, this result demonstrates that the cDNA clone is able to be expressed and translated effectively and that the lack of overexpression seen previously is not due to any problems in the construction of the cDNA clones or in their ability to be translated. In addition, these results demonstrate that the regulated tetracycline expression system is functional and that expression of the exogenous gene is

Figure 3 - **Detection of HMG14aFLAG by Western blot**

QT6 cells were transiently transfected with plasmid constructs as listed in the figure. Cultures were grown in the presence (4.0 ug/ml) or absence of tetracycline in the media. Protein was isolated at 24 hours post-transfection with a 5% PCA extraction as described in chapter II. Protein was electrophoresed on 18% SDS-PAGE gels, blotted and detected with anti-FLAGM2 antibody (80). HMG14aFLAG protein is clearly detected in lane 3 while none of the control lanes showed evidence of expression.

	1	2	3	4	5	6	7	8	9
pUHD10-3/ HMG14aFLAG-S	-	+	+	-	-	-	-	-	-
pUHD10-3/ HMG14a-AS	-	-	-	-	-	-	+	+	-
pUHD10-3/ HMG14a-S	-	-	-	-	+	+	-	-	-
pUHD15-1	-	+	+	+	+	+	+	+	+
Tet	-	+	-	+	-	+	-	+	-

Figure 3

tightly controlled.

HMG proteins were isolated from the transfected cells with a 5% PCA total cell extraction. For the exogenously transfected protein to exert any effects on the chromatin structure within the cell, it is necessary that these proteins be located in the nucleus and not solely in the cytoplasm. To demonstrate the cellular location of the transfected protein, transient transfections were performed as above with each plate being divided into nuclear, cytoplasmic and total cell fractions. In addition to the wild type HMG14aFLAG, a site-specific mutant in HMG14a, HMG14aC3 -FLAG was also used in these transfections. The HMG14aC3 mutant contains a single base pair mutation at codon 35 in the cDNA sequence resulting in a proline to lysine change in the protein sequence. 5% PCA extractions were done on each fraction and the isolated protein was run out on 18% SDS-PAGE gels for analysis. Figure 4 shows the results of the western blots using the anti-FLAG M2 and anti-HMG antibodies to detect the protein. In figure 4A, the anti-FLAG M2 antibody detects HMG14aFLAG protein only in the nuclear and total fractions in the absence of tetracycline. In figure 4B, the anti-HMG antibody detects total HMG14/17 in only the nuclear and total fractions. These results therefore indicate that the wild type HMG14aFLAG and the mutant HMG14aC3FLAG are being expressed and the protein is localized to the nucleus.

Figure 4 - **Detection of HMG14a and HMG14aFLAG proteins in cell fractions from transiently transfected cells**

A. QT6 cells were transiently transfected with plasmid constructs as listed in the figure. Cultures were grown in the presence (4.0 ug/ml) or absence of tetracycline in the media. Protein was isolated with a 5% PCA extraction from nuclear, cytoplasmic and total fractions of transfected cells at 24 hours post-transfection. The fractions were prepared as follows. Nuclei were isolated by cellular lysis as described in chapter II and the cell lysate was taken as the cytoplasmic fraction. Nuclear pellets were washed 3-5 times in lysis buffer to rid nuclei of lingering cytoplasmic proteins and washes were added back to cytoplasmic fractions. Nuclear intactness was determined by microscopy. Total fractions refer to a 5% PCA extraction of whole cells. Protein was loaded (approximately 10^6 cells/lane) and electrophoresed on 18% SDS-PAGE gels. Gels were blotted and protein was detected with anti-FLAGM2 antibody. HMG14aFLAG wild type and mutant proteins were detected only in the nuclear and total cell fractions of the appropriate transfections.

B. Western blots from figure 4A were stripped and reprobed with anti-HMG antibody. The lanes correspond to those in figure 4A. HMG14/17 was detected in all nuclear and cytoplasmic fractions.

	1	2	3	4	5	6	7	8	9
pUHD10-3/ HMG14aFLAG-S	-	-	-	+	+	+	+	+	+
pUHD15-1	+	+	+	+	+	+	+	+	+
Cell fraction	N	C	T	N	C	T	N	C	T
Tet	+	+	+	+	+	+	-	-	-

	10	11	12	13	14	15	16	17	18
pUHD10-3/ HMG14C3FLAG	-	-	-	+	+	+	+	+	+
pUHD15-1	+	+	+	+	+	+	+	+	+
Cell fraction	N	C	T	N	C	T	N	C	T
Tet	+	+	+	+	+	+	-	-	-

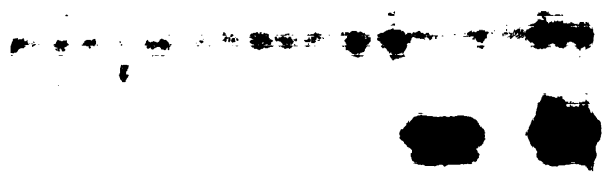


Figure 4A

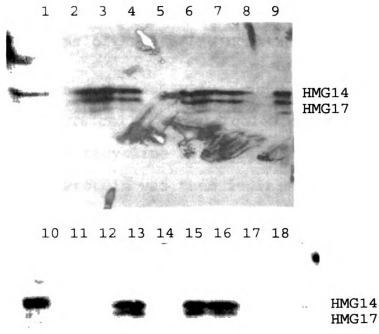


Figure 4B

In addition, this demonstrates that there is no significant leakage from nuclei isolated by our standard procedure.

Stable transfections of HMG14aFLAG constructs

The HMG14aFLAG and HMG14aC3FLAG constructs were stably transfected into QT6 cells. Single colonies were grown up in the presence of tetracycline to repress expression of the FLAG constructs. Protein was then isolated from duplicate plates of cells after 48 hours of growth with or without tetracycline in the media. 25 single colonies of HMG14aFLAG and 8 colonies of HMG14aC3FLAG transfectants were tested by western analysis for expression of the respective FLAG containing proteins. Results of several of these experiments are shown in figures 5 and 6. As can be seen in the blots probed with anti-FLAG M2 antibody, there was no FLAG containing protein detected in any of the single colonies tested. In most lanes, two series of bands were detected at high mobilities inconsistant with the expected size of the FLAG proteins. These bands are present in the untransfected QT6 cell lines as well and therefore represent cross reactivity with unknown proteins present in the QT6 cell line. All the blots were reprobed with the anti-HMG antibody showing endogenous HMG14/17 protein was detectable at normal levels as demonstrated in figure 6.

The results with the transient transfections demonstrate

Figure 5 - **Western blot of protein isolated from stably transfected cell lines**

Single colonies were isolated from QT6 cells stably transfected with TFANEO, pUHD15-1 and pUHD10-3 containing HMG14FLAG (designated F) or HMG14a cDNA (designated E). Cells were grown in the presence or absence of tetracycline for 24 hours and protein was isolated, electrophoresed, and blotted as previously described. Blots were probed with anti-FLAGM2 antibody. There was no detection of HMG14FLAG in any of the cell lines tested. Several low mobility bands were detected in all lanes that do not correspond to known proteins. M refers to a FLAG control protein purchased from Kodak-IBI (80) and serves as a positive control for the anti-FLAGM2 antibody.

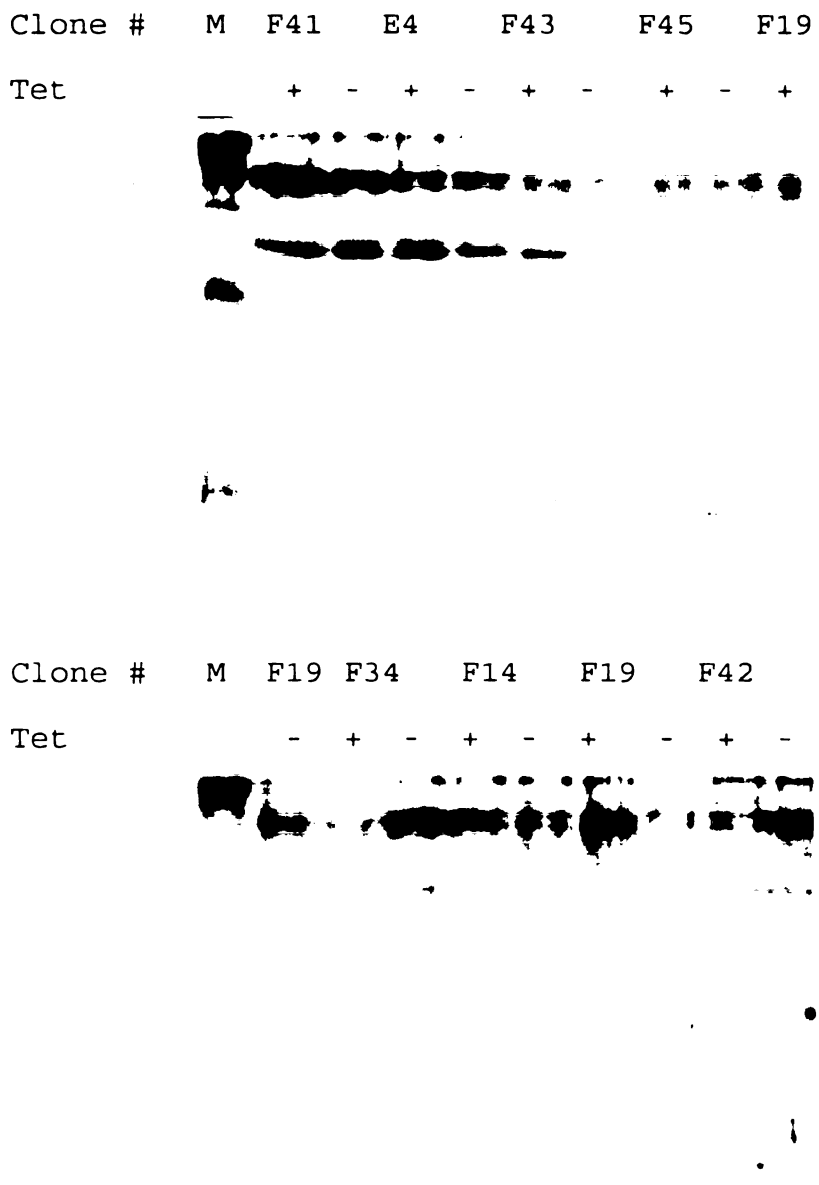


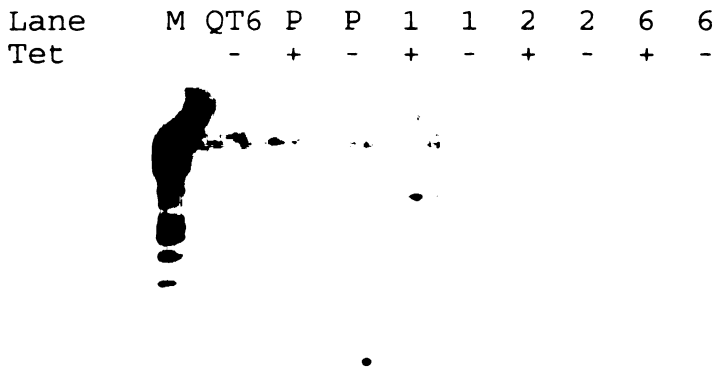
Figure 5

Figure 6 - **Western blot of protein isolated from stably transfected cell lines**

A. Protein was isolated from QT6 cells stably transfected with plasmids TFANEO, pUHD15-1, and HMG14C3FLAG/pUHD10-3. Cells were grown in the presence or absence of tetracycline and the protein was isolated, electrophoresed, and blotted as previously described. The protein was detected with anti-FLAGM2 antibody. There was no HMG14FLAG detected in any of the cell lines tested. Lane M refers to a FLAG control protein purchased from Kodak-IBI (80) and serves as a positive control for the anti-FLAGM2 antibody. QT6 refers to untransfected control cell line. P refers to a pooled colony derived from the transfection and 1,2, and 6 refer to individual single colonies.

B. Duplicate gel run simultaneously with gel in 6A and detected with anti-HMG antibody. Normal amounts of HMG14/17 are seen in all lanes.

A.



B.

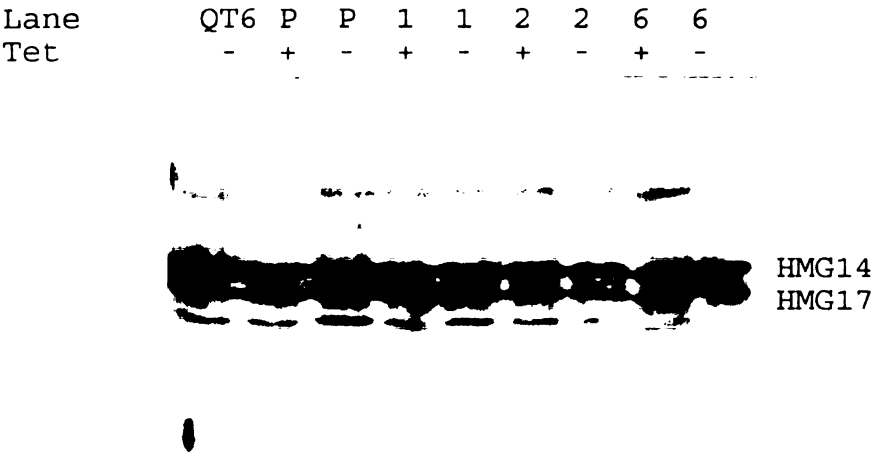


Figure 6

that the constructs used in these experiments are able to be expressed and detected by our antibodies. The inability to isolate stable colonies that express high levels of FLAG containing protein could be explained in several ways. The cell may only be able to tolerate a limited amount of HMG14 protein such that additionally expressed HMG14 is rapidly degraded or there may be limited sites available on the nucleosomes for HMG14 and additional non-bound protein is rapidly degraded. If this is the case, the exogenous chicken derived fusion protein may be preferentially degraded over the endogenous quail protein. Alternatively, perphaps there is a copy number effect, such that in the copy number may be low to moderate in stables, but very high in transients. Finally, since the plasmid containing the HMG14FLAG constructs is unselected in the cell, over the time required to grow the single colonies to to the point where there are enough cells from which to isolate protein, about 3-4 weeks, the plasmid could be lost or may become methylated or otherwise modified such that expression is severly limited or repressed altogether.

PCR detection of plasmid sequences in transfected cell lines

To determine whether the transfected HMG14aFLAG plasmid sequences are present and uninterrupted in the stably transfected cell lines, PCR was used to detect diagnostic

Figure 7 - **PCR strategy for detecting HMG14aFLAG/pUHD10-3 sequences in stably transfected cell lines**

A. Schematic of plasmid HMG14aFLAG/pUHD10-3 integrated in the cell genome. pUHD10-3 promoter (460 bp) contains the hCMV minimal promoter with heptemerized Tet operator sequences fused upstream (73). Genomic DNA was isolated from stable cell lines, digested with *HindIII*, and amplified with the primer pairs shown. Primer 1 (sense) binds to sequences within the Tet operator and was paired with primers 2,3 and 4 (anti-sense) that bind to sequences within the HMG14a cDNA coding region.

B. Sequences of primers used to amplify portions of HMG14aFLAG/pUHD10-3.

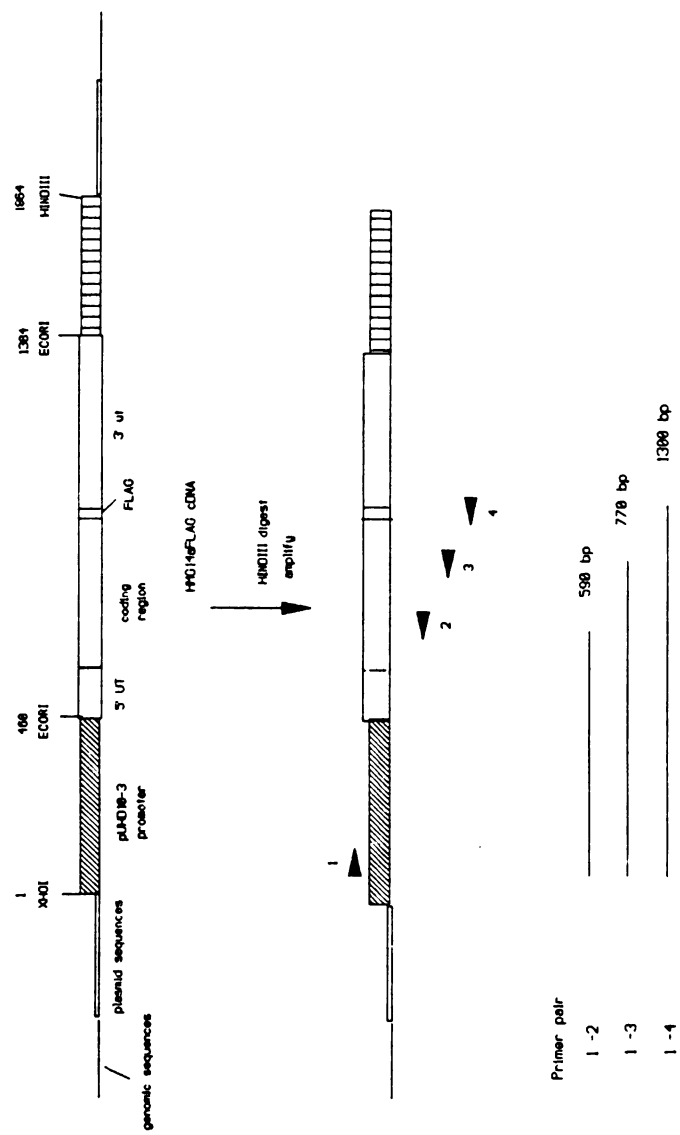

1
XHOI
plasmid sequences
genomic sequences
pJHD1B-3
promoter
100
ECORI
5' UT
coding
region
FLAG
3' ut
1364
ECORI
1964
HINDIII
HMD14dFLAG cDNA
HINDIII digest
amplify
1
2
3
4
Primer pair
1 -2
1 -3
1 -4
590 bp
770 bp
1300 bp

Figure 7A

Primer		Sequence 5' - 3'
1	(Tet operator)	GTGAAAGTCGAGTTTACCACTCC
2	(14C3-AS)	GCTGCCTTTTTGAGCTTTG
3	(14FLAG-AS)	CTTGTCATCGTCGTCCTTGTAGTCCTCGGACTTGGCTTCC
4	(14SEQPR2)	CCAAAACTAATGAGGCACCAGCTGC

Figure 7B

fragments in genomic DNA isolated from several of the colonies tested. The PCR strategy and primers used are illustrated in figure 7. Primer pairs used included a sense primer hybridizing to Tet operator sequences in the 5' end of the HMG14aFLAG/pUHD10-3 and several antisense primers hybridizing to sequences internal in the HMG14a coding region. Primer pairs were designed to test for the intactnesss of the promoter region and most of the FLAG cDNA sequence. Amplification products were electrophoresed on agarose gels and primer pair #3 generated DNA fragments corresponding to the expected size (770 bp, see figure 7A) were in 4 of the 6 clones tested. These results indicated that at least some of the stable colonies contained the HMG14aFLAG plasmid intact within the quail genome (presumably at more than one copy per haploid genome).

Western analysis of HMG14aFLAG stable clones

One of the previously unscreened single colonies which tested positive using PCR for the presence of the HMG14FLAG containing plasmid, HMG14C3FLAG-18, was analyzed for protein expression along with QT6 and an untested pooled colony, HMG14C3P. Cells were grown and protein was isolated as previously described. Protein samples were electrophoresed on 18% SDS-PAGE gels, blotted and probed with anti-HMG and anti-FLAGM2 antibodies. Figure 8 shows the results of this experiment. There was no HMG14FLAG protein detected in either

Figure 8 - **Western blots of protein isolated from stably transfected cell lines**

A. Protein was isolated from the previously untested cell line HMG14C3FLAG-18 and from control QT6 cells. Cells were grown in the presence or absence of tetracycline and the protein was isolated and blotted as previously described. Protein was detected with anti-FLAG M2 antibody. Lane M refers to a FLAG control protein purchased from Kodak-IBI (80) and serves as a positive control for the anti-FLAG M2 antibody. There was no HMG14aFLAG protein detected.

B. Duplicate gel to that in A above, using anti-HMG antibody. Normal levels of HMG14/17 are seen in all lanes.

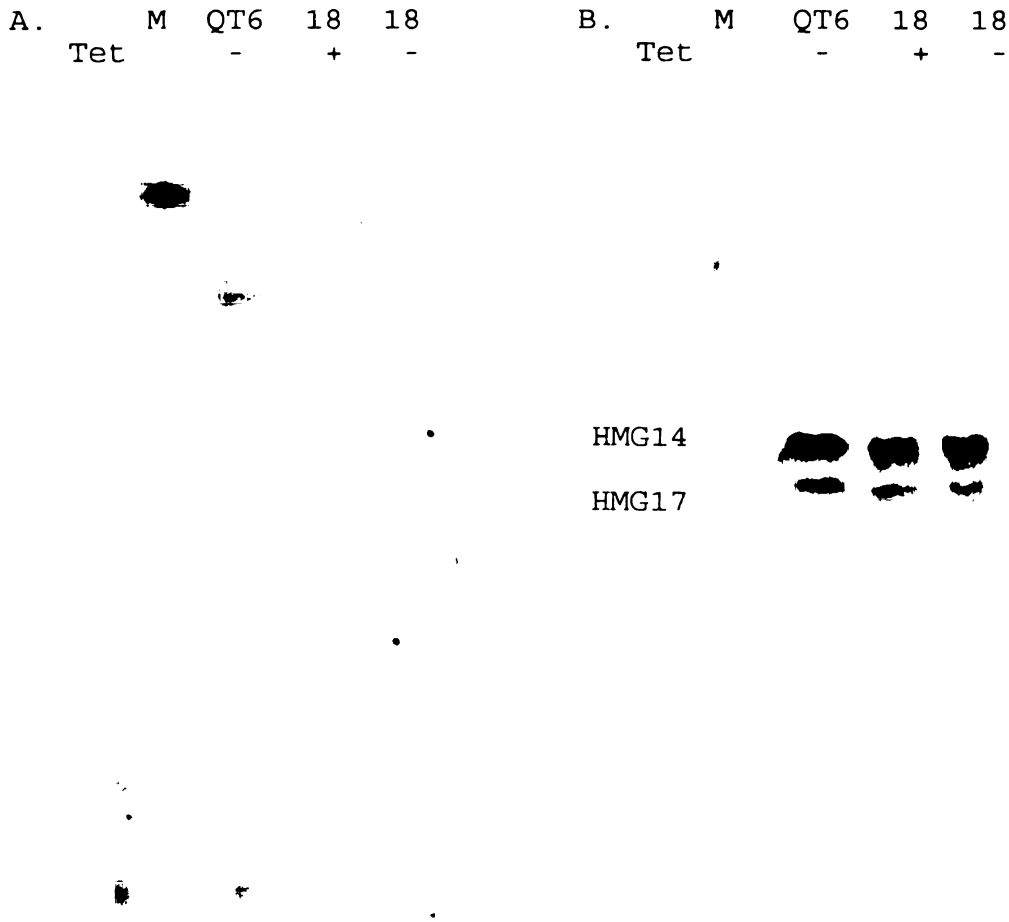


Figure 8

of the cell lines tested. The anti-HMG antibody detected normal amounts of HMG14 and HMG17, while the anti-FLAGM2 antibody detected only the previously seen high molecular weight bands. The repeated inability to isolate stable colonies that express significant levels of exogenously transfected HMG14a protein is therefore not due to loss or disruption of the HMG14-containing plasmid in the stable lines.

Discussion

The inability to isolate stably transfected QT6 cells that express high levels of the exogenous HMG14FLAG proteins from plasmid pUHD10-3 is consistent with our results using plasmid vector TFANEO. Previous experiments indicate that the HMG14aFLAG constructs are reproducibly expressed and the protein is detectable in transient transfection experiments. In at least 60% of our stable cell lines, the plasmid sequences are present and uninterrupted in the QT6 genome. In transient and stable cell lines the tetracycline regulated expression system has been shown to tightly control expression of HMG14a, HMG14aFLAG and luciferase gene constructs. The absence of detectable HMG14aFLAG in the stable cell lines tested is therefore most likely due to one or more cellular regulatory mechanisms functioning to repress expression or eliminate excess HMG14 protein.

The question of cellular regulation of HMG protein levels

raises several possibilities. It may be that only HMG14 bound to chromatin is stable. It is known that the HMG14 proteins are found bound only to a subset of nucleosomes *in-vivo*. If the limited binding sites are occupied by endogenous quail protein, and the exogenous chicken HMG14aFLAG does not compete efficiently for these sites, then the cell may rapidly degrade the unbound protein. Although HMG14 proteins show a high level of evolutionary similarity, the chicken HMG14a proteins may be sufficiently different from the (as yet uncloned) quail HMG14, particularily with the added FLAG tail, as to be relatively unstable or unable to efficiently compete for the limited binding sites available on chromatin. It may also be that the chicken HMG14a mRNA levels are feedback or otherwise regulated by the quail cell.

The transfected HMG14a proteins were reproducibly detected in transient transfections, yet were not detectable in cell lines derived from stable transfections. The expression of stably integrated plasmids may be repressed by a number of mechanisms. Expression can be influenced by the location of integration in the genome, the presence or absence of regulatory elements controlling gene expression over large distances, or presence of endogenous repressor activity (84). Alternatively, the exogenous HMG genes contain high G + C levels and once integrated, could be silenced by *de-novo* CpG methylation (85). Finally, the lack of detectable protein may

be due to a copy number difference. The endogenous quail HMG14a is known to be present at 2 copies per cell genome. The HMG14aFLAG is clearly detectable in transients where the copy number is probably greater than 100 copies per cell. In stable transfectants, the copy number is presumably much lower, at 1-10 copies per cell genome, and thus may account for a greatly reduced or undetectable expression level.

Chapter VI

Chromatin structural studies of cell lines ectopically expressing chicken HMG14a and HMG17

Chapter VI

INTRODUCTION

The HMG14 and HMG17 chromatin binding proteins have been highly conserved throughout evolution suggesting they provide a vital cellular function. The location and specificity of their binding suggests a role in the regulation of gene expression. These proteins bind to the subset of chromatin known to be in an active or potentially active transcriptional state. The amount of HMG14 and HMG17 present in the cell is regulated and limits binding to 1 out of every 10 nucleosomes, with up to two molecules bound per nucleosome. Despite a great deal of research characterizing the specific molecular interactions of HMG14/17 with the chromatin, their function remains obscure.

To address the question of the cellular function of these proteins, we have chosen to ectopically overexpress chicken HMG14a and HMG17 in tissue culture and measure the changes, if any, in cell characteristics, including chromatin structure. Overexpression of mutant and wild type HMG14a and HMG17 proteins may cause the exogenous proteins to occupy new sites in the cellular chromatin and, in the case of mutants, to outcompete the endogenous HMG for binding sites. Aberrant binding of the mutant or wild type HMG proteins may then produce changes in chromatin structure. Analysis of dominant phenotypes produced in these cell lines will help to

establish the role these proteins play in chromatin structure and regulation of gene expression.

Our studies focus on the bulk chromatin structure and make use of endonucleases that cleave at predictable locations in chromatin. Parameters such as nucleosome repeat length, spacing, sensitivity to nucleases, and changes in the structure or distribution of active and inactive sequences are determined in cell lines shown to ectopically over-express HMG14a and HMG17.

RESULTS

Micrococcal nuclease digestion of chromatin

Micrococcal nuclease is an endonuclease that digests DNA at locations not tightly complexed with proteins. In chromatin, this enzyme cuts in the linker region between nucleosomes resulting in the release of individual nucleosomes with the length of linker DNA present dependent upon the extent of digestion. DNA extracted from a mild micrococcal nuclease digestion of chromatin and run out on an electrophoretic gel results in a ladder pattern with bands representing mononucleosomes, dinucleosomes, trinucleosomes and so on. The core nucleosome contains approximately 140-160 bp of DNA wrapped tighty around the octomer of core histones. The length of the linker DNA between nucleosomes is variable, usually 60-140 bp, depending upon the organism, cell type and developmental stage (75). The repeat length of nucleosomes is the total length of the DNA associated with that nucleosome, i.e. the total of the core plus linker DNA. The spacing of nucleosomes is usually consistent within a given cell and depends on the length of linker DNA between nucleosomes. It has been suggested that HMG14/17 are required for the proper spacing of nucleosomes (82). Digestion of chromatin with micrococcal nuclease and analysis of the resultant DNA will thus demonstrate the repeat length and nucleosome spacing properties characteristic of a particular cell line.

In addition, the overall sensitivity of the chromatin to digestion with micrococcal nuclease will be characteristic of a particular cell line. Regions of transcriptionally active sequences demonstrate an increased sensitivity to nucleases such as DNaseI. That the HMG14/17 proteins are also associated with these regions of active chromatin suggests they may influence this property.

Figure 1 shows micrococcal nuclease digestions of chromatin from five experimental cell lines. Nuclei were isolated, digested with two concentrations of micrococcal nuclease, the DNA was extracted and then it was electrophoresed on a 1.8% agarose gel. QT6 is the untransfected control cell line, 14-RS-6 and 14-RS-8 are two single colony clones from QT6 transfected with the HMG14a cDNA containing a 3'-untranslated region deletion. 14-RS-6 is known to overexpress HMG14a mRNA but not protein. 17S-7 and 17S-10 are two single colony clones from QT6 transfected with wild type chicken HMG17 cDNA. These two clones express elevated levels of HMG17 specific mRNA as well as 3-5 times more HMG17 protein than the control cells. There is no apparent difference seen in the repeat length or spacing properties of the nucleosomes from these different cell lines. In addition, the overall nuclease sensitivity appears to be similar in the cells tested. This result is somewhat surprising, particularly in cell lines 17S-7,10 which overexpress HMG17. This

Figure 1 - **Micrococcal nuclease digestion of chromatin**

DNA extracted from chromatin was digested with 50 and 300 units of micrococcal nuclease per mg nuclei DNA equivalent. The DNA was electrophoresed on a 1.8% agarose gel and visualized with ethidium bromide. QT6 is the untransfected parental cell line. 14-RS-6 and 14-RS-8 are two single colony clones from a QT6 transfection of chicken HMG14a cDNA with a deletion in the 3'- untranslated region. 17S-7 and 17S-10 are two single colony clones from a QT6 transfection of chicken HMG17 wild type cDNA. There is no apparent difference seen in the nucleosomal repeat length, nucleosome spacing or overall nuclease sensitivity among these clones.

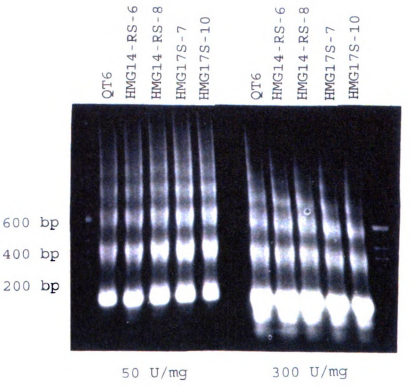

QT6
HMG14-RS-6
HMG14-RS-8
HMG17S-7
HMG17S-10
QT6
HMG14-RS-6
HMG14-RS-8
HMG17S-7
HMG17S-10
600 bp
400 bp
200 bp
50 U/mg
300 U/mg

Figure 1

may suggest that the extra HMG17 present in the cell is either not binding to additional sites on the chromatin, and/or that this binding has no effect at this level of analysis. Additionally, the level of overexpression of the HMG17 protein may not be high enough to exert any changes which can be seen in this type of analysis. Alternatively, it may be that these HMG proteins have no influence over the chromatin structural properties tested in this assay.

Characterization of nucleoprotein complexes present in cell lines

To determine whether the excess HMG17 protein expressed in cell lines 17S-7 and 17S-10 is binding to sites on the nucleosomes other than those already bound by the endogenous proteins, we assayed the subfractions of nucleosomes present in the various cell lines. In methods developed by Albright et al; (24), mononucleosomes can be electrophoretically seperated into five classes characterized by the presence of a variety of small proteins bound, including the HMG14/17 proteins (24). Forms MI - MV are defined as follows:

	HMG14/17 bound/particle	H1 bound/particle
MI	0	0
MII	1	0
MIIIA	0	1
MIIIB	2	0
MIV	1	1
MV	2	1

By comparing the distribution of these heterogeneous nucleosomal fractions between cell lines, we can determine if the excess HMG17 produced in our cell lines is binding to additional sites on the individual nucleosomes. Nuclei were isolated from the various cell lines, treated briefly with micrococcal nuclease and the resultant nucleoprotein complexes were isolated and run on acrylamide/ agarose/ glycerol gels.

Figure 2 shows the results of this nucleoprotein assay testing the five cell lines described above in the micrococcal nuclease digestion assay. Cell line 17S-7, and to a lesser extent 17S-10, show a definite increase in band MII, which represents one particle of HMG17 bound per nucleosome. This indicates that these cell lines, which overexpress HMG17 3-5 times that of the control QT6 cells, do have an increase in binding of HMG17 on their nucleosomes. Apparently, this excess binding is not sufficient to be able to alter the chromatin structure as assayed in the micrococcal nuclease digest above.

DISCUSSION

The HMG14 and HMG17 chromatin binding proteins bind to the nucleosome in regions of structurally distinct chromatin, therefore, it may be expected that overexpression of these proteins in the cell would lead to structural changes in the chromatin. With this hypothesis in mind, we tested various

Figure 2 - **Analysis of nucleoprotein complexes**

Chromatin was digested briefly with micrococcal nuclease and the nucleoprotein complexes were fractionated and run out on acrylamide/agarose/glycerol gels. The cell lines are as described in figure 1. The content of HMG14/17 in each nucleosomal class MI -MV are desribed in the text. Clones 17S-7 and 17S-10 show an increase in band MII as compared to control cells, indicating the excess HMG17 expressed in these cells lines is taking up additional binding sites on the nucleosomes.

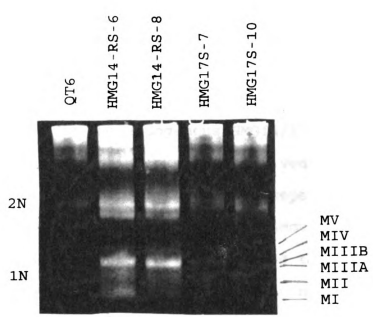

QT6
HMG14-RS-6
HMG14-RS-8
HMG17S-7
HMG17S-10
2N
1N
MV
MIV
MIIIB
MIIIA
MII
MI

Figure 2

cells lines shown to ectopically overexpress chicken HMG17. It is known that the levels of total HMG14 and HMG17 are regulated in the cell such that about 1 out of 10 nucleosomes will have HMG14/17 bound. The HMG bound nucleosomes are preferentially localized in regions of active or potentially active chromatin. To determine if the excess HMG17 produced in our cell lines, HMG17S-7 and HMG17S-10, will bind to nucleosomes and possibly take up additional sites not bound by the endogenous HMG14/17, we analyzed the individual nucleoprotein complexes derived from mononucleosomes. The complexes can be separated on electrophoretic gels according to the amount of HMG14/17 bound per particle. Our results indicate that the exogenous transfected HMG17 does, in fact, bind to nucleosomes resulting in an increase in the electrophoretic band corresponding to one HMG bound per particle. However, the changes observed were not dramatic. The increased binding of the HMG17 molecule to the nucleosomes might be expected to produce detectable alterations in the gross chromatin structure. However, our analysis of micrococcal nuclease digested chromatin isolated from these cell lines showed no detectable change in the nucleosome repeat length, nucleosomal spacing, or overall sensitivity to the endonuclease. This result suggests several possibilities. First, the level of expression of the exogenous HMG17 may not

be sufficient to induce changes in the gross chromatin structure as assayed at this level. Alternatively, the chromatin structural properties associated with the binding of the HMG14/17 molecules, such as increased sensitivity to micrococcal nuclease (associated with regions of active chromatin and bound HMG) may not be caused by the direct binding of these proteins. It may be that these proteins are required, at most, to maintain this micrococcal nuclease sensitive structure rather than to generate it, such that excess HMG17 bound would not be sufficient to produce detectable changes in the chromatin structure. Finally, though highly unlikely due to the high conservation of the HMG14/17 proteins, the exogenous chicken HMG17 may be sufficiently different from the endogenous quail proteins such that they will bind to the nucleosomes yet not influence the chromatin structure in the predicted fashion.

Chapter VII

Immunofluorescence of cells expressing HMG14aFLAG

Chapter VII

INTRODUCTION

We have demonstrated the successful transient transfection of wild type and mutant HMG14aFLAG fusion clones and the specific detection of their protein product in the QT6 cell line. Here we examine the *in-situ* localization of these proteins using indirect immunofluorescence to determine more specifically the subcellular localization of these proteins.

The subcellular distribution of the HMG14/17 proteins has not been studied in detail although these proteins are known to bind to nucleosomes and have been shown to be associated *in-vitro* with regions of actively transcribing chromatin. We have previously shown, by crude cellular fractionation and Western blot, that these proteins localize preferentially to the nucleus of the transfected cells. Here we examine the cellular distribution of transiently transfected wild type and mutant HMG14aFLAG using indirect immunofluorescence in the QT6 cell line.

RESULTS

HMG14aFLAG constructs and transfection system

The construction of clones for the wild type HMG14aFLAG and the site-specific mutant, HMG14C3FLAG, are described in chapter V. These FLAG constructs were cloned into the expression vector pUHD10-3 and co-transfected into QT6 cells with plasmid pUHD15-1 as previously described in chapter IV. The expression of the cloned genes in culture is regulated by the presence or absence of tetracycline in the tissue culture media. The transfected cells were grown on coverslips in tissue culture plates for 48 hours before fixation.

Immunofluorescence and microscopy

Transfected cells on coverslips were fixed for staining by two methods which gave similar results. In method I, the cells were extracted for 5 minutes in microtubule stabilizing buffer (100mM PIPES, 2 mM EGTA, 4% w/v PEG 8000, 0.025 % sodium azide, pH 6.8)) with 0.5% Triton X-100. In method II, there was no fixation used. All cells were then fixed in 4% paraformaldehyde for 30 minutes, washed in acetone for 5 minutes, washed briefly in distilled water and held in PBS (phoshate buffered saline, pH 7.5) overnight at 4°. The cells were then stained with the primary antibody, anti-FLAGM2 (described in chapter V) diluted 1/100 in TBS (Tris buffered saline, 0.05 M Tris base, 0.015 M NaCl, pH

7.5), 0.5% Bovine serum albumin, and 0.05% Tween 20. This incubation was performed by dropping the primary antibody solution on the coverslips in a closed moist container and gently shaking at room temperature for 4 hours. Coverslips were then washed by repeated dipping in TBS/0.05% Tween20, followed by washing in PBS. The coverslips were then incubated for 4 hours at room temperature in a 1/50 dilution of fluorescein conjugated anti-rabbit IgG (Dako Co., Carpinteria, Ca.) in a moist container covered with aluminium foil to prevent exposure to light. The coverslips were then washed in PBS and dried briefly. A drop of embedding medium was applied to the coverslips, these were then placed upside down on the slide, sealed with clear nail polish and stored at -70°. The slides were analyzed under the Odyssey Real Time Laser scanning confocal microscope (Noran Instruments Inc., Madison, Wi.). Serial optical sections were captured using 10X, 25X , 60X and 100X Nikon oil immersion objectives. Laser excitation and primary barrier filters for fluorescein were 488 and 550 nm respectively. Images were analyzed and processed using IMAGE I software (Universal Image Co., Westchester, Penn.).

Subcellular distribution of HMG14aFLAG proteins in cells

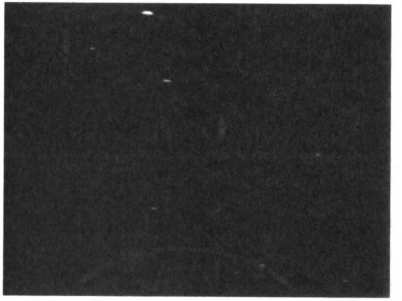
Figure 1 shows the immunofluorescence staining pattern of QT6 cells transiently transfected with wild type HMG14aFLAG

Figure 1 - **Immunofluorescent staining of cells transfected with HMG14aFLAG**

Panel A - QT6 cells transfected with pUHD10-3/HMG14aFLAG and pUHD15-1 as previously described and grown for 48 hours in the presence of tetracycline. Cells were fixed and stained as described in the text. There is no cellular or nuclear staining in this 10X magnification control slide.

Panel B - Identical to panel A but cells were grown in the absence of tetracycline to allow expression of the HMG14aFLAG gene construct. There is clear staining of nuclei with an estimated transfection efficiency of 10-20%. 10X magnification.

A



B

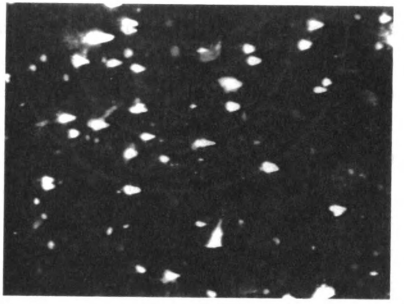


Figure 1

grown in the presence (panel A) and the absence (panel B) of tetracycline. Cells grown in the presence of tetracycline show no cellular or nuclear staining, as is expected with repression of expression of the HMG14aFLAG gene. Cells grown in the absence of tetracycline show highly fluorescent nuclei with barely detectable background staining of the cytoplasm. Transfection efficiency was visually estimated at 10-20%. These results clearly show that the transfected HMG14aFLAG gene is being expressed and the protein is localized preferentially to the nucleus. Additionally there appears to be very tight control of expression by the tetracycline regulated vector system.

Figure 2 shows a negative control transfection of QT6 cells transfected with plasmid pUHD15-1 which expresses the activator protein only. Background staining shows a faint outline of the cells and clearly unstained nuclei.

Upon higher magnification, two distinct staining patterns appear within the nuclei of the transfected cells. Each slide analyzed showed a mixture of the two staining patterns with no clear preference overall for either one. Figure 3, panel A, shows highly fluorescent nuclei with clear exclusion of staining in the nucleoli. The staining pattern within the nucleus is not uniform, but gives a characteristic mottled appearance, presumably due to the HMG14aFLAG proteins binding to the disbursed chromatin. Panel B shows a similar overall

Figure 2 - **Negative transfection control**

QT6 cells transfected with pUHD15-1 only and treated and stained identically to all other transfection plates. There is a faint nonspecific background staining with no specific cellular or nuclear staining evident.

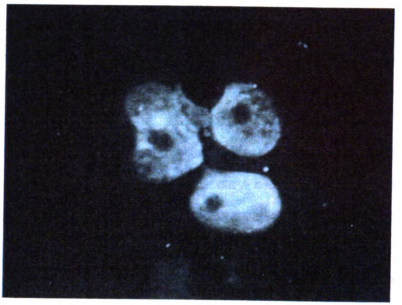
Figure 2

Figure 3 - **Immunofluorescent staining of cells transfected with HMG14aFLAG**

Panel A - QT6 cells transfected with pUHD10-3/HMG14aFLAG and pUHD15-1 as previously described and grown in the absence of tetracycline. Cells were fixed and stained as described in the text. This 100X magnification plate shows the specific nuclear staining with preferential exclusion from the nucleoli. Notice the uneven or mottled appearance of the nuclear material.

Panel B - Photograph taken from the same slide as panel A. In these nuclei the staining is most intense inside the nucleoli.

A



B

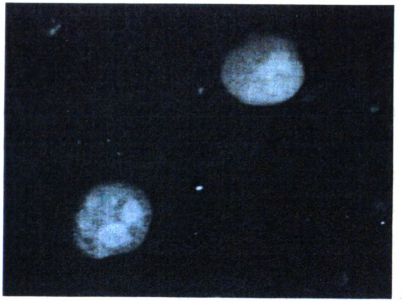


Figure 3

nuclear staining appearance with the multiple nucleoli preferentially stained. This pattern of exclusion or preference of staining of the nucleoli in these cells, taken from a single slide, seems unlikely to be due to antibody accessability or differences in sample preparation. Alternatively, this difference may be due to the state of the individual cell within the cell cycle and may represent the distribution of these HMG14 proteins accordingly. Bhullar et al (84) showed that the high mobility group proteins of cultured trout cells (analagous to the mammalian proteins HMG1/2) preferentially stained nucleoli in interphase cells. More investigation is warranted to determine the significance of these specific patterns.

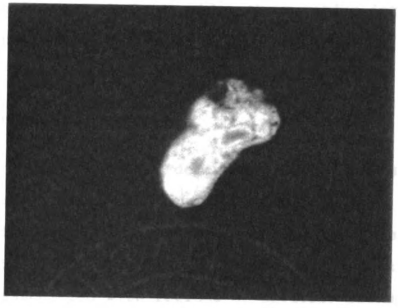
Cells transfected with the site-specific mutant HMG14C3FLAG (14C3 is a mutation in the DNA binding domain of HMG14 that changes Pro 35 - Leu) showed similar transfection efficiencies and staining patterns as seen above with the wild type. Figure 4, panel A, shows a typical nuclear staining pattern, in this case the nucleoli can be seen to preferentially exclude the stain. Panel B was included for interest for the unusual staining pattern seen in the nuclei in the lower right. This cell may be undergoing division with the HMG14 proteins intensely staining the condensed chromosomes. It appears that the mutant, 14C3, does not present any differences overall in staining patterns from the

Figure 4 - **Immunofluorescent staining of cells transfected with HMG14C3FLAG**

Panel A - QT6 cells transfected with pUHD10-3/HMG14C3FLAG and pUHD15-1 as previously described and grown in the absence of tetracycline in the media. Cells were fixed and stained as described in the text. The nuclear staining pattern of this mutant HMG14 appears identical to the wild type. In this panel there is preferential exclusion of staining from the nucleoli. 100X magnification.

Panel B - Photograph from the same slide as panel A. Notice the staining pattern in lower right nucleus. This pattern may be due to intense staining of condensed chromosomes prior to cell division.

A



B

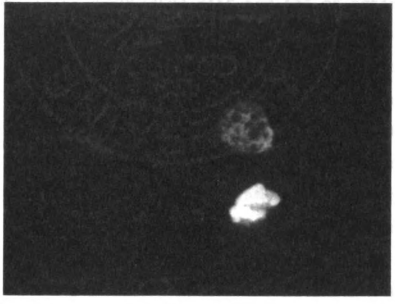


Figure 4

wild type HMG14aFLAG. Presumably, the mutation introduced to this clone does not affect binding to the chromatin or at least nuclear localization.

Discussion

The data presented here demonstrate that our HMG14FLAG clones are transfectable into QT6 cells at a high efficiency and localize to the nucleus of these cells. It has long been established that the HMG14/17 proteins bind to nucleosomes and are associated with regions of actively transcribing chromatin *in-vitro*. It is hypothesized that these proteins, upon binding to nucleosomes, help to generate or maintain the DNAse sensitivity associated with regions of active or potentially active genes. This evidence that HMG14 is, in fact, localized to the nucleus *in-situ* lends support to the proposed role that the HMG proteins play in the cell.

Upon high magnification, the staining pattern of the HMG14FLAG proteins within the nucleus gives a nonuniform and mottled appearance. Presumably, this may be due to the binding of these HMG proteins to the dispersed chromatin of the interphase chromosomes.

The interesting patterns of nuclear staining seen in this study, namely preferentially localized to or excluded from the nucleolus, warrant further investigation. Fibroblasts, such as the QT6 cell line, are known to have numerous

nucleoli. The nucleoli are distinct, nonmembraned regions within the nucleus where the rapid transcription of ribosomal RNA genes takes place. The presence or absence of HMG14/17 proteins within these regions may be influenced by the state of the cell cycle and, if indeed these proteins are required on the chromatin of actively transcribing genes, their presence in the nucleolus may be indicative of rapid or elevated transcription of ribosomal RNA as it is required by the cell.

Chapter VIII

Summary

SUMMARY

The aim of the research project reported here was to help develop techniques and data that will further our understanding of the role of the HMG14/17 proteins in the cell. Toward that end, we chose to express wild type and mutant HMG14/17 proteins from cloned chicken cDNAs in tissue culture and to observe morphological and phenotypic changes that may result from this expression. We have successfully expressed a variety of HMG14/17 clones in transient transfection systems, several wild type HMG17 in stable transfections, and analyzed, via chromatin structural studies, cell growth properties, and microscopy, the resultant cellular characteristics.

In the quail fibroblastic cell line, QT6, we transiently expressed exogenous wild type, site-specific mutants and HMG14 fusion proteins in clearly detectable amounts. These results were achieved using two different expression vectors, TFANEO and the tetracycline regulated pUHD10-3. Expression levels were determined by Northern analysis and Western analysis using two different antibodies for detection of protein. Transfection efficiencies were reliably high, suggesting that, at least in transients, the cells can tolerate high levels of exogenously expressed HMG14/17 proteins. However, since only a fraction of the cells were

successfully transfected and since the avidities of the anti-FLAG and anti-HMG antibodies may differ, it is difficult to estimate the exogenous/endogenous HMG ratio *in-vivo*. The growth rates of the transfected cell lines did not differ noticably from control transfections, though this characteristic is difficult to quantify in these types of experiments. The ectopic HMG14/17 expressed in these cell lines was shown to localize to the nucleus, as does the endogenous protein. There appeared to be no discernable difference in transfection efficiency or detectability between the wild type, mutant, or fusion HMG14/17 proteins expressed. Additionally, none of the site-specific mutants tested showed electrophoretic variability distinct from the wild type protein.

We have isolated a limited number of cell lines stably expressing wild type HMG17 driven by the vector TFANEO. These cell lines show HMG17 expression to be elevated 3-5 times that of the untransfected cells. This overexpression appears to cause no gross morphological changes or alterations in cell growth rate. Chromatin structural studies on two such stable cell lines indicate that the exogenous protein does bind to additional unoccupied sites on the chromatin, but does not induce any detectable gross structural changes. Techniques employed to demonstrate this include micrococcal nuclease digestion and electrophoretic

analysis of total chromatin, and electrophoretic analysis of characteristic nucleoprotein complexes present in the chromatin of these cell lines. Presumably, the lack of gross phenotypic alterations in the chromatin structure of these stably transfected cell lines may be due to the moderate level of overexpression of the exogenous clones.

With the exception of these two stable cell lines expressing only moderately higher levels of exogenous HMG17, we were unable to reproducibly isolate stable cell lines expressing detectably increased levels of HMG14/17. Following careful evaluation of our transfection systems, we conclude that this effect is most likely due to cellular mechanisms which act to limit the expression of exogenous HMG14/17 and/or the total amount of HMG14/17 within the cell. The molecular mechanisms of such a limitation and whether it affects both endogenous and exogenous HMg14/17 remains unclear. Unfortunately, stable cell lines present the best model for further analysis of the effects of exogenously expressed HMG14/17 proteins in the cell. The lack of stable cell lines expressing high and/or regulatable levels of HMG14/17 has thus limited our ability to analyze the role of HMG14/17 in the cell using these techniques.

In further attempts to characterize the activity of the HMG14/17 proteins in the cell, we used immunoflorescent confocal microscopy to visualize HMG14 fusion proteins

within transiently transfected cells. These experiments demonstrated, as shown previously, that the transfected gene products localize to the nucleus of the cells. Within the nucleus, various distinct staining patterns can be seen. In many cells, the HMG14 fusion protein can be seen throughout the nucleus but preferentially excluded from the presumed nucleolus. In other cells, this pattern is distinctly reversed, with the HMG14 fusion protein seen to be preferentially localized within the nucleolus. These two unique staining patterns may be related to the state of the individual cell with regards to its cell cycle and the distribution of the HMG14 proteins within the nucleus may be regulated according to the cells replicative needs. These results are provocative enough to warrant further investigation.

Overall, we have demonstrated that wild type, mutant and fusion chicken HMG14/17 cDNA clones can be successfully transfected and expressed in QT6 cells. The expressed protein is localized to the nucleus and as shown in some stable cell lines, binds to sites on chromatin previously unoccupied by the endogenous protein. The distribution of the exogenous HMG14 proteins within the nucleus is shown to be non-uniform with regards to the nucleolus. The specific role that these HMG14/17 proteins play in the cell and, in particular, in the regulation of gene expression remains unclear, and it is

hoped that the methods and results developed here will contribute to a better understanding of these proteins.

LIST OF REFERENCES

LIST OF REFERENCES

1. Weintraub, H., Groudine, M. (1976) Science 193, 848-856

2. Levy, A., Noll, M. (1981) Nature 289, 198-203

3. Pogo, B.G.T., Allfrey, V.G., Mirsky, V.G. (1966) Proc. Natl. Acad. Sci. U.S.A. 59, 2239-2243

4. Mathis, D.J., Oudet, P., Wasylyk, B., Chambon, P. (1978) Nuc. Acids Res. 5, 3523-3547

5. Oliva, R., Bazett-Jones, D.P., Dixon, G. (1990) Nuc. Acids Res. 18, 2739-2747

6. Wolfe, A. (1992) Chromatin Structure and Function. Academic Press, New York.

7. Weisbrod, S., Groudine, M., Weintraub, H. (1980) Cell 19, 289-301

8. Weisbrod, S., Weintraub, H. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 630-634

9. Gazit, B., Panet, A., Cedar, H. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 1787-1790

10. Weisbrod, S., Weintraub, H. (1981) Cell 23, 391-400

11. Dorbic, T., Wittig, B. (1986) Nuc. Acids Res. 14, 3363-3376

12. Albanese, I., Weintraub, H. (1980) Nucl. Acids Res. 8, 2787-2805

13. Einck, L., Bustin, M. (1983) Proc. Natl. Acad. Sci. 80, 6735-6739

14. Johns, E.W. (1982) The HMG Chromosomal Proteins, Academic Press, New York

15. Bustin, M., Lehn, D.A., Landsman, D. (1990) Biochem. Biophys. Acta 1049, 231-243

16. Lewin, B. (1990) Genes IV, Oxford University Press, Cambridge, U.K.

17. Einck, L., Bustin, M. (1985) Exp. Cell Res. 156, 295-310

18. Mosevitsky, M., Novitskaya, V., Iogannsen, M., Zabezhinsky, M. (1989) Eur. J Biochem 185, 303-310

19. Paull, T., Haykinson, M., Johnson, R. (1993) Genes and Development 7, 1521-1534

20. Bustin, M., Becerra, P., Crippa, M., Lehn, D., Pash, J., Shiloach, J. (1991) Nuc. acids Res. 19, 3115-3121

21. Abercrombie, B., Kneale, G., Crane-Robinson, C., Bradbury, E.M., Goodwin, G., Walker, J., Johns, E.W. (1978) Eur. J. Biochem 84, 173-177

22. Cary, P.D., King, D.S., Crane-Robinson, C., Bradbury E.M., Rabbani, A., Goodwin, G., Johns, E.W. (1980) Eur. J Biochem 112, 577-580

23. Sandeen, G., Wood, W., Felsenfeld, G. (1980) Nuc. Acids Res. 8, 3757-3778

24. Albright, S.C., Wiseman, J.M., Lange, R.A., Garrard, W.T. (1980) J. Biol. Chem. 255, 3673-3684

25. Mardian, J., Paton, A., Bunick, G., Olins, D. (1980) Science 209, 1534-1536

26. Alfonso, P.J., Crippa, M.P., Hayes, J.J., Bustin, M. (1994) J. Mol. Biol. 236, 189-198

27. Sasi, R., Huvos, P.E., Fasman, G.D. (1982) J. Biol. Chem. 257, 11448-11454

28. Uberbacher, E.C., Mardian, J.K., Rossi, R.M., Olins, D.E., Bunick, G.J. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 5258-5262

29. Graziano, V., Ramakrishan, V. (1990) J. Mol. Biol. 214, 897-910

30. Shick, V.V., Belyavsky, V., Mirzabekoz, A.D. (1985) J. Mol. Biol. 185, 329-339

31. Cook, G.R., Yau, P., Yasuda, H., Traut, R.R., Bradbury, E.M. (1986) J. Biol. Chem. 261, 16185-16190

32. Crippa, M.P., Alfonso, P.J., Bustin, M. (1992) J. Mol. Biol. 228, 442-449

33. Brotherton, T.W., Ginder, G.D. (1986) Biochemistry 25, 3447-3454

34. Druckman, S., Mendelson, E., Landsman, D., Bustin, M. (1986) Exp. Cell Res. 166, 486-496

35. Dorbic, T., Wittig, B. (1987) EMBO 6, 2393-2399

36. Landsman, D., Bustin, M. (1991) Mol. Cell. Biol. 11, 4483-4489

37. Postinkov, Y.V., Shick, V.V., Belyavsky, V., Khrapko, K., Brodolin, K., Nikolskaya, T.A., Mirzabekov, A.D. (1991) Nuc. Acids Res. 19, 717-725

38. Levy-Wilson, B. (1981) Arch. Biochem. Biophy. 208, 528-534

39. Reeves, R., Chang, D., Chung, S. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 6704-6708

40. Reeves, R., Chang, D. (1983) J. Biol. Chem. 258, 679-687

41. Spaulding, S.W., Fucile, N., Bofinger, D.P., Sheflin, L.G. (1991) Mol. Endo. 5, 42-50

42. Einck, L., Soares, N., Bustin, M. (1984) Exp. Cell Res. 152, 287-301

43. Pedrini, M., Grunicke, H., Csordas, A. (1992) Electrophoresis 13, 397-399

44. Bustin, M., Soares, N., Landsman, D., Srikantha,T., Collins, J.M. (1987) Nuc. Acids Res. 15, 3549-3561

45. Pash, J.M., Bhorjee, J.S., Patterson, B.M., Bustin, M. (1990) J. Biol. Chem. 265, 4197-4199

46. Begum, N., Pash, J.M., Bhorjee, J.S. (1990) J. Biol. Chem. 265, 11936-11941

47. Crippa, M.P., Nickol, J.M., Bustin, M. (1991) J. Biol. Chem. 266, 2712-2714

48. Crippa, M.P., Nickol, J.M., Bustin, M. (1991) J. Mol. Biol. 217, 75-84

49. Pash, J.M., Alfonso, P.J., Bustin, M. (1993) J. Biol. Chem. 268, 13632-13638

50. Shakoori, A.R., Owen, T.A., Shaloub, V., Stein, J.L., Bustin, M., Stein, G.S., Lian, J.B. (1993) J. Cell. Biochem. 51, 479-487

51. Crippa, M.P., Pash, J.M., Gerwin, B.I., Smithgall, T.E., Glazer, R.I., Bustin, M. (1990) Cancer Res. 50, 2022-2026

52. Reeves, R., Candido, E.P.M. (1980) Nuc. Acids Res. 8, 1947-1963

53. Landsman, D., McBride, O.W., Soares, N., Crippa, M.P., Srikantha, T., Bustin, M. (1989) J. Biol. Chem. 264, 3421-3427

54. Landsman, D., McBride, O.W., Bustin, M. (1989) Nuc. Acids Res. 17, 2301-2314

55. Popescu, N., Landsman, D., Bustin, M. (1990) Hum. Genetics 85, 376-378

56. Browne, D. Ph.D. Dissertation (1990) Michigan State University, Dept. Microbiology

57. Landsman, D., Srikantha, T., Bustin, M. (1988) J. Biol. Chem. 263, 3917-3923

58. Srikantha, T., Landsman, D., Bustin, M. (1987) J. Mol. Biol. 197, 405-413

59. Browne, D., Dodgson, J. (1993) Gene 124, 199-206

60. Landsman, D., Srikantha, T., Westerman, R., Bustin, M. (1986) J. Biol. Chem. 261, 16082-16086

61. Landsman, D., Soares, N., Gonzalez, F.J., Bustin, M. (1986) J. Biol. Chem. 261, 7479-7484

62. Srikantha, T., Landsman, D., Bustin, M. (1990) J. Mol. Biol. 211, 49-61

63. Pash, J., Popescu, N., Matocha, M., Rapoport, S., Bustin, M. Proc. Natl. Acad. Sci. U.S.A. 87, 3836-3840

64. Pash, J., Smithgall, T., Bustin, M. (1991) Exp. Cell. Res. 193, 232-235

65. Giri, C., Landsman, D., Soares, N., Bustin, M. (1987) J. Biol. Chem. 262, 9839-9843

66. Bustin, M., Crippa, M.P., Pash, J.M. (1990) J. Biol. Chem. 265, 20077-20080

67. Moscovici, C., Moscovici, M.G., Jimenez, H. (1977) Cell 11, 95-103

68. Kerppola, T.K., Kane, C.M. (1991) FASEB Journal 5, 2833-2842

69. Jackson, R.J., Standart, N. (1990) Cell 62, 15-24

70. Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J., Sutrave, P. (1978) J. Virology 61, 3004-3012

71. Federspeil, M.J., Crittenden, L.B., Hughes, S.J. (1989) Virology 173, 167-177

72. Sambrook, J., Fritsch, E.F., Maniatus, T. (1989) Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

73. Gossen, M., Bujard, H. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 5547-5551

74. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R. (1989) Gene 49-53

75. van Holde, K.E. (1989) Chromatin, pp. 291-302, Springer-Verlag Inc., New York

76. Manley, J.L., Proudfoot, N.J., Platt, T. (1989) Genes and Development 3, 2218-2244

77. Brush, D., Dodgson, J.D., Choi, O., Stevens, P.W., Engel, J.D. (1985) Mol. Cell. Biol. 5, 1307-1317

78. Stein, G.S., Stein, J.L., Marzluff, W.F. (1984) Histone Genes, Jonh Wiley and Sons, Inc.

79. Dodgson, J.D., Yamamoto, M., Engel, J.D. (1987) Nuc. Acids Research 15, 6294

80. Kodak IBI (1993) Anti-FLAG M2 antibody technical article Product # IBI3010

81. Bio-Rad Inc. (1990) MUTA-GENE in vitro mutagenesis kit manual, Product # 170-3571

82. Tremethick, D.J., Drew, H.R. (1993) J. Biol. Chem 268, 11389-11393

83. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D. (1989) Molecular Biology of the Cell, Garland Publishing, Inc., N.Y.

84. Bhullar, B.S., Hewitt, J., Candido, E.P.M. (1981) Journal Biol. Chem. 256, 8801-8806

85. Goodwin, E.B., Okkema, P.G., Evans, T.C., Kimble, J. (1993) Cell 75, 329-339

APPENDIX I

APPENDIX I

Table of amino acids and their abbreviations

Amino Acid	Abbreviation	Characteristics
Alanine	A	nuetral,hydrophobic
Cysteine	C	nuetral,polar
Aspartic acid	D	acidic
Glutamic acid	E	acidic
Phenylalanine	F	nuetral,hydrophobic
Glycine	G	nuetral,polar
Histidine	H	basic
Isoluecine	I	neutral,hydrophobic
Lysine	K	basic
Luecine	L	neutral,hydrophobic
Methionine	M	neutral,hydrophobic
Asparagine	N	neutral,polar
Proline	P	neutral,hydrophobic
Glutamine	Q	neutral,polar
Arginine	R	basic
Serine	S	neutral,polar
Threonine	T	neutral,polar
Valine	V	neutral,hydrophobic
Tryptophan	W	neutral,hydrophobic
Tyrosine	Y	neutral,polar

APPENDIX II

Production of cell lines expressing chicken histone variants H3.3B and H3.2

APPENDIX II

INTRODUCTION

Chicken histones and their variants

The highly conserved major core histone proteins, H2A, H2B, H3 and H4, are the most abundant chromatin binding proteins found in the nucleus of eukaryotic cells (78). These histones comprise a set of proteins whose basic function is to bind nuclear DNA and compact it into the specific condensed structure seen in eukaryotic chromatin. The basic repeating structural unit of chromatin, the nucleosome, is comprised of approximately 144 base pairs of DNA wrapped around an octomer of 2 each of the 4 core histones, H2A, H2B, H3 and H4 (78). This nucleosomal structure is further compacted by the binding of the linker histone H1, which binds to the linker DNA between nucleosomes (78). While the cellular role of these abundant core histones is well known, there exist a number of histone variants whose functions remain obscure. These histone variants can differ from the core proteins in amino acid sequence, degree of chemical modification, and/or regulation of expression with respect to the cell cycle, tissue, or developmental program (77). The fact that the expression of these variants is regulated suggests they have functional significance, probably at the level of chromatin structure. It is thought that the presence of

histone variants in chromatin may influence tissue-specific gene expression during development (78).

Histone variants can be divided into several classes according to their pattern of expression (77). Replication variants are those whose expression is tightly coupled with the S phase of the cell and DNA replication. Replacement variants are expressed at a constitutive level throughout the cell cycle and tissue specific variants are found in certain tissues at unique stages in development. Certain histone variants have been found to play important and unique roles in cell function such as the chicken red blood cell specific H5 and the spermatocyte specific H1 variant H'. Due to the wide species distribution and high level of evolutionary conservation of other histone variants, it is reasonable to assume they also have unique roles in cell function and development.

The genes for the major non-variant histone types have been isolated and examined from various species and have been found to share certain characteristics in most cases. These include:

- The genes are closely linked to one another
- There are multiple copies of genes for each histone
- The genes do not contain introns
- The transcripts are not polyadenylated

- Expression is tightly coupled to DNA synthesis

While most histone genes share these properties, there are some histone variants found that display certain different characteristics from those mentioned above. Our lab has previously cloned and sequenced two different genes which code for the same chicken histone replacement variant H3.3, namely H3.3A and H3.3B (77). These genes are unique in that they:

- Are not linked to each other or the other histone genes
- Their mRNA is polyadenylated
- Their genes do contain introns
- They are expressed in the absence of DNA replication (i.e. replacement variants)

In the chicken genome, these two H3.3 replacement variants are members of the H3 subfamily of genes. This subfamily includes at least these two H3.3 genes as well as about nine replication variant H3.2 genes. The apparent cellular requirement for the presence of replacement and replication variants within this subfamily is not clear but the high evolutionary conservation of both types suggests a functional significance. Analogous H3 histones to H3.3 have been found in mammals, *Drosophila*, and *Tetrahymena* (78).

To investigate the cellular role of the replacement variant histone H3.3B, the cDNAs encoding H3.3B and, for comparison,

the replication variant H3.2, were cloned into an expression vector and transfected into tissue culture cells. The results from these experiments are discussed below.

RESULTS

Expression of the chicken histone replacement variant H3.3B and replication variant H3.2 at abnormal times and/or in abnormal amounts in the cell may lead to changes in cellular phenotypes which can be observed. These observations may then lead to a better understanding of the individual role these variants play in the cell.

Replacement variant H3.3B

The gene encoding the chicken replacement variant H3.3B was isolated (along with H3.3A) by hybridization to heterologous sea urchin and *Drosophila* histone coding region probes (77). The H3.3B gene was later used to isolate corresponding H3.3 cDNA clones (79). The 1128 bp cDNA for H3.3B contains 405 bp of coding region (135 AA predicted protein), about 100 bp of 5' noncoding sequences and a long 585 bp 3' untranslated region (79). The coding region of this cDNA was isolated and cloned into plasmid pCla12. Most of the noncoding sequences were eliminated for our experiments as they most likely contribute to the unique regulatory properties of this gene and we wish to avoid that regulation for the purpose of

expressing the H3.3 protein at abnormal times and/or in abnormal amounts in the cell. The H3.3 cDNA was digested with *EcoRV* at the 5' end, resulting in 45 bp of 5' noncoding sequences remaining and was digested with *ApoI* at the 3' end, resulting in 32 bp of 3' untranslated remaining with the coding region and elimination of the downstream polyadenylation site. This truncated H3.3 cDNA in pCla12 was then subcloned into expression vector TFANEO (see chapter IV, figure 1) for transfection into QT6 tissue culture cells.

Replication variant H3.2

511 bp of the 1 kb chicken H3.2 gene containing primarily coding region was cloned into pCla12 and then TFANEO for use in our experiments. The 1 kb *EcoRI* fragment containing the H3.2 gene was digested with *BspDI* giving the 511 bp fragment. The 511 bp fragment contains approximately 70 bp of 5' noncoding sequences, 325 bp of coding region, and 116 bp of 3' untranslated sequence. Previous work by our lab and others determined that S phase dependent regulatory elements of this gene were controlled primarily by sequences 5' (130 bp of 5'noncoding sequence) and 3' (containing a stem loop region) to the coding region (79). As with H3.3B, we wished to eliminate these regions in our clone. The coding region portion of H3.2 cloned into expression vector TFANEO was then used in our transfection experiments.

Southern analysis of QT6 transfected with histone variants

Genomic DNA was isolated from stable colonies derived from transfections of QT6 cells with H3.3B and H3.2 in the expression vector TFANEO. The DNA was digested with *EcoRI* and subjected to Southern analysis. Figure 1 shows the results of two of these southern blots. Cell lines transfected with H3.3B were probed with a 500 bp *EcoRI* cDNA fragment isolated from pCla12/H3.3B. Cell lines transfected with H3.2 were probed with a 500 bp *ClaI* fragment isolated from pCla12/H3.2. In both blots, bands migrating at the expected mobility were detected confirming that the transfected cDNAs were integrated and non-interupted in the quail cell genome. Additional bands are due to cross hybridization with endogenous quail H3 histone genes.

Northern analysis of transfected cell lines

To determine the levels of expression of the transfected histone genes, RNA was isolated from the stable cell lines and was subjected to northern analysis. 30 ug of total RNA was loaded per lane onto agarose/formamide gels and these were then blotted to nitrocellulose membranes. The blots were probed with either cDNA fragments specific to the transfected construct (as in the Southern analysis) or with an *EcoRI* fragment containing LTR sequences from TFANEO. Sequences in the TFANEO LTR expression cassette should be transcribed in

Figure 1 - **Southern analysis of stable cells lines transfected with H3.3B and H3.2 cDNAs**

Genomic DNA was isolated from QT6 cells transfected with H3.3B and H3.2 cDNAs in expression vector TFANEO. The DNA was digested with *EcoRI* and 10 ug was run per lane on 1.5% agarose gels. The gels were blotted to nitrocellulose membrane and probed with the random primed H3.3B or H3.2 sequences. Cell lines tested were: untransfected QT6, QT6 transfected with TFANEO vector only (VC3), and cell lines transfected with H3.3B and H3.2 cloned in the sense and antisense direction in vector TFANEO. Panel A shows a band at the expected size of ~500 bp in the H3.3B sense lane. Panel B shows a fragment of the expected size of ~550 in the H3.2 sense and antisense lanes. This confirms that in these cell lines, the transfected DNA has integrated and is uninterrupted in the quail cell genome.

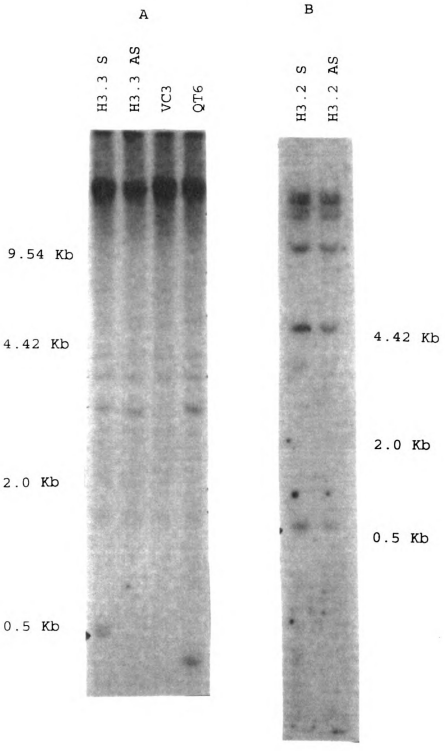

A
B
H3.3 S
H3.3 AS
VC3
QT6
H3.2 S
H3.2 AS
9.54 Kb
4.42 Kb
2.0 Kb
0.5 Kb
4.42 Kb
2.0 Kb
0.5 Kb

Figure 1

Figure 2 - **Northern analysis of cell lines transfected with H3.3B and H3.2 cDNAs**

30 ug of total RNA isolated from cell lines transfected with the H3.3B and H3.2 cDNAs in vector TFANEO was run out on 1.2% agarose/formamide gels. The gels were blotted to nitrocellulose and probed with a *EcoRI* LTR fragment from TFANEO. Cell lines tested include: untransfected QT6 control, QT6 transfected with vector alone, and cell lines transfected with H3.3B and H3.2 in the sense and antisense direction in TFANEO. Bands visible in the lanes representing H3.3B and H3.2 cloned in the sense direction demonstrate that these cell lines are expressing high levels of exogenous message.

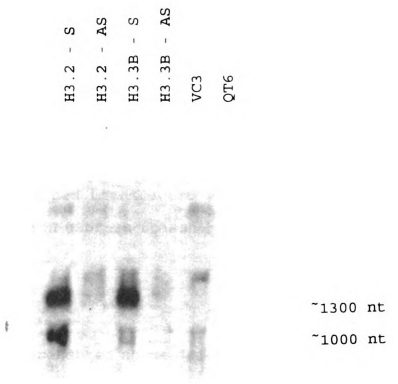


Figure 2

the 5' end of the exogenous histone variant mRNAs and should be unique in this cell line. Figure 2 shows a northern analysis of various cell lines probed with the LTR fragment. Cell lines tested include the parental QT6, VC3 which is QT6 transfected with the vector alone, and H3.3B and H3.2 in the sense and antisense direction in TFANEO. The results show that the cell lines transfected with H3.3B and H3.2 in the sense orientation express reasonably high levels of exogenous message.

Discussion

The results presented above demonstrate that the cell lines we transfected with histone variants H3.3B and H3.2 have the transfected cDNAs integrated into the cellular genome and express significant levels of exogenous mRNA. Microscopic examination of these cell lines shows no gross morphological alterations and their growth rate is similar to the parental QT6. It thus appears that these cells can tolerate high levels of exogenously transfected histone H3.3B and H3.2 message with no obvious defects. Further analysis of these cell lines is required to determine if the exogenous histone mRNA is being transcribed efficiently into protein. In addition, studies of the chromatin structure in these cell lines may uncover changes brought about by the unregulated expression of these histone variants.