INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Université d'Ottawa · University of Ottawa

The Role of Xist and CHD-1 in Gene Silencing

Helen H. Tai

Thesis submitted to
the Department of Biochemistry, Microbiology, and Immunology in partial of fulfilment of the requirements for the degree of Doctor of Philosophy

University of Ottawa
Ottawa, Ontario, Canada
2000

Acknowledgements

My thesis project was supported by a grant from the Medical Research Council of Canada. I also received fellowships from Medical Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada. My thesis work was carried out in the laboratory of Dr. Michael W. McBurney. I would like to thank Dr. McBurney for giving me the autonomy to work on some of my own ideas. Drs. John Bell, Douglas Gray, Alexander MacKenzie, David Picketts and Mario Chevrette were members of my thesis advisory committee. Their sound advice was much appreciated.

I would like to thank Dr. Verne Chapman of the Roswell Park Cancer Institute in Buffalo, NY, for giving us unpublished information from his lab about the Hae III polymorphism used in Chapter I. Dr Hunt Willard donated the Xist probe used for the northerns presented in Chapter I.

Much appreciation also goes to Drs. Robert Perry and David Stokes of the Fox Chase Cancer Centre, in Philadelphia, PA. They generously donated to me the full-length CHD-1 cDNA as well as the polyclonal antibody against CHD-1 used in this thesis.

Dr. Micheline Paulin-Lavasseur consulted me about doing the nuclear matrix extractions presented in Chapter II. Also my father, Dr. George Tai, designed the statistical analysis used to analyze nuclear matrix localization. Their expert advice was greated appreciated.

Dr. Karen Colwill isolated the N-CoRβ partial cDNA in Dr. Tony Pawson's

laboratory at the University of Toronto. I would like to thank them for letting me use the N-CoRβ partial cDNA for the studies presented in this thesis.

Over the years I spent at the Cancer Research Group many people helped me though my doctoral degree. I couldn't have finished this degree without them. First of all I would like to thank the technicians in the lab, Jane Craig, Karen Jardine, and Xiao-Feng Yang. Their attention to the many details involved in running the lab and their supprt for me was much appreciated. Karen also made many of the vectors that I used to make plasmid constucts for my thesis project. I would also like to thank Beth Mason, Dr. McBurney's secretary, for all the help she provided during the preparation of my thesis.

I would like to thank Richa Sharan, a graduate student in the lab, for introducing me to chromodomains. Another graduate student in the lab, Sandra Milligan, provided me with good advice and protocols forimmunofluorescence. I'd also like to thank Ken Garson, a post-doc in the lab, for many insightful discussions concerning research and for his dry sense of humor. Ilona Skeranc, Leslie Sutherland, and Debbie Blake were former members of the lab who provided me as much technical advice as moral support.

Other members of the Cancer Research Group were also instrumental to the success of my thesis project. I would especially like to thank Dr. Margit Geisterfer, a post-doc in Dr. Bell's lab; she provided much assistance and advice on my yeast two-hybrid screen. One of the clones I identified in my yeast two-hybrid screen was mKIAA00164, which Margit recognized as a protein that also

interacts with Clk1. Margit also cloned the full-length mKIAA0164, which is presented in Figure II.11 of this thesis. Margit's keen observation is what eventually led us to discover an interaction between N-CoRβ and CHD-1 presented in Chapter II. Margit also shared with me many of the ups and downs (mostly ups) of life as a Mommy-scientist.

Scott Kelso, a technician in John Bell's lab, did all the sequencing presented in this thesis. I would like to thank Scott for his care and consideration with his work. I would also like to thank David Stojdl, a graduate student in John Bell's lab, for sharing his many thoughts on research and life in general; and for spurring me on to go ahead with the yeast two-hybrid screen. I'd like to thank Alan Cheng and Ninan Abraham (a fellow leftist), former graduates in the Bell lab, for sharing many interesting discussions.

Paola Blanchette and Maria Tsirigotis are graduate students in Doug Gray's lab, who were much support to me during immunoprecipitations and westerns. I'd also like to thank them for many laughs and roundhouse kicks. Sherry Thurig, Josée Coulombe, and Yvonne Avis are current and former member of the Gray lab who I would also like to thank for their support in the lab.

Audri Brewster, another leftist, supplied many though-provoking discussions I will miss. I would like to thank Audri for her comraderie.

There are many other members of the Cancer Research Group who I have not named here but who also helped me through my doctoral degree. The open atmosphere at the Cancer Research Group has made my years of graduate

studies easier and a lot more fun. I have made many friends at the Cancer Researhc Group who have been a tremendous support to me and I would like to take this opportuinity to thank them. I would also like to thank all those who entertained my daughter and made the lab a playground for her.

I could not have finished my doctoral degree without Diane Farrell, my daughter's caregiver. I would like to thank Diane for giving me the peace of mind I needed to focus on my thesis work.

I would especially like to thank my family. First of all my parents, George and Cecilia Tai, who have always encouraged and supported me in all my academic pursuits. I would also like to thank my good little girl, Mia Parenteau, who was my buddy at the lab on the weekends. Her enthusiasm for the lab and good behavior allowed Mommy to get her "speriments" done. Most of all I'd like to thank my partner and best friend, Bill Parenteau, who makes my life so much fun. Without his unwavering support of me, I would not have been able to finish this degree.

Finally, I would like to dedicate this thesis to the late Dr. Joe Connolly, my friend and former graduate advisor at the University of Toronto. Joe showed us all how fun in the lab. His enthusiasm for science was an inspiration to me during my doctoral thesis work.

Abstract

Gene silencing is a stable and heritable form of repression of gene expression that is thought to be mediated by higher order chromatin structures. This thesis consists of two separate studies on gene silencing. The first chapter deals with X chromosome inactivation which is a particular case of gene silencing where almost all the genes one X chromosome are inactivated early in the development of female mammals. The Xist gene is a candidate gene for controlling X chromosome inactivation because the Xist gene resides on the X chromosome near the XIC (X inactivation center) and is expressed from the inactive X in female somatic cells. Chapter I describes an investigation of the expression of Xist in diploid P10 female embryonal carcinoma cells that have two active X chromosomes. The transcripts expressed from Xist alleles in P10 cells could be distinguished by restriction enzyme digestion of their cDNAs. It was found that both alleles Xist are expressed; thus, Xist transcripts are expressed prior to X chromosome inactivation and are not always derived from the inactive X chromosome. Therefore, Xist expression *per se* cannot be a sufficient signal to inactivate an X chromosome.

The second chapter is a study of CHD-1, a mouse protein with two chromodomains, a helicase/ATPase domain, and a DNA binding domain. The chromodomain connects CHD-1 to gene silencing because the chromodomain, also present in the *Drosophila* proteins HP-1 and Polycomb, is required for the gene silencing function of these proteins. Therefore, CHD-1 was further examined in Chapter II. In part A of Chapter II, the finding that CHD-1 was associated with the nuclear matrix, a protein structure in the nucleus, is reported. Part B of Chapter II describes a study investigating proteins that interact with the chromodomain of CHD-1 in a yeast two hybrid screen. One of the protein identified as interacting with CHD-1 was N-CoR which is a protein that is associated with histone deacetylation. Since deacetylated histones are associated with chromatin containing silenced genes, it was concluded that CHD-1 can function in gene silencing as part of a complex with histone deacetylase.

List of Tables

List of Figures and Illustrations

General Introduction

Only a subset of genes are expressed in a cell at any particular time and the rest of the genome is in a silenced state. Silenced genes remain stably repressed throughout the life of the cell and through cell division. The heritable nature of gene silencing is a defining characteristic (104). Silenced genes are also associated with a condensed chromatin structure called heterochromatin whereas the decondensed structure in euchromatin structure is associated with active genes (104). In addition, regions of silenced DNA are often methylated and the surrounding chromatin is deacetylated; in contrast to active genes which are associated with demethylated DNA and acetylated chromatin (115) (12).

This thesis examines two types of gene silencing: X chromosome inactivation and the maintenance of developmentally regulated gene expression. X chromosome inactivation involves the inactivation of almost all the genes on a single X chromosome. In female mammals one X chromosome inactivates early in the developing embryo. Once the X chromosome is inactivated is remains silent in daughter cells after cell division. In addition, in female cells there exists both an active X chromosome and an inactive X. Therefore, the inactive X is exposed to the same transcriptional regulatory proteins as the active X but remains stably inactive (70).

The maintenance of gene expression patterns in the developing embryo also involves gene silencing. In Drosophila, a specific pattern of gene expression is established in each segment. The Polycomb group of genes were required for maintaining this pattern of expression, though these protein do not contribute the

establishment of the expression pattern (52). One protein in the Polycomb complex is the Polycomb protein, Pc. Pc shares with another Drosophila protein, HP-1, a protein domain called the chromodomain (95). HP-1 is heterochromatin protein involved in position effect variegation (38). Position effect variegation occurs when a gene translocates to a region of heterchromatin and be comes silenced. Therefore, the chromodomain is domain associated with gene silencing. The second chapter of this thesis examines a mammalian chromodomain protein, CHD-1.

Although, the two chapters of this thesis are separate studies, they share the common theme of gene silencing.

Chapter I:

Xist and X chromosome inactivation in female embryonal carcinoma cells

Introduction

The first chapter of this thesis examines X chromosome inactivation. In mammals, males (XY) inherit only the maternal X chromosome whereas females (XX) inherit X chromosomes from both parents. Although female mammals have twice the number of X-linked genes, the level of expression of these genes is the same in males and females since the genes on one of the X chromosomes are inactivated in female somatic cells (68). The inactive X chromosome remains transcriptionally inert through many cell divisions and may be maintained in this state by DNA methylation of genes on the inactive X (for reviews see (43), (44)). Hence, X chromosome inactivation in female mammals is an example of gene silencing.

In early female embryos, both X chromosomes are active until the early blastocyst stage when X inactivation is first detected (39) (60). X chromosome inactivation takes place in all cells of the female mouse embryo between days 5.5 to 8.5 post-coitum (100) (111). Inactivation is thought to initiate at a unique site on the X chromosome called the X inactivation center (Xic/XIC; (19) 1 (101), for review see (22)). Once X inactivation has been initiated it spreads along the X chromosome in *cis* (43) (70). The inactivation of either the maternally or paternally-inherited X chromosome is random in cells of in the inner cell mass that give rise to the embryo proper. However, the paternal X chromosome is preferentially inactivated in cells of the primitive endoderm and trophectoderm (111) (111) (120). The timing of X chromosome inactivation in the embryo seems to coincide with tissue differentiation (82).

The human XIC was localized using naturally occurring human X chromosome mutations combined with somatic cell hybrid technology (21). The XIC was defined as the minimal segment of the X chromosome that could still induce X chromosome inactivation (21). Therefore, genes that affect X inactivation could be isolated from the XIC. Brown et al. cloned one such gene derived from the human XIC, XIST (21) (19). They found that XIST transcripts were expressed from the inactive X but not the active X chromosome in female somatic cells, which is opposite to most X-linked genes. An XIST homologue in mouse, Xist, was also cloned (15) (17). Murine Xist is also expressed exclusively from the inactive X chromosome in female somatic tissues (18) (17) (15) and in male germ cells that also inactivate their X chromosomes (76) (103) (105). Furthermore, the Xist transcript was found in the nucleus associated with the condensed inactive X chromosome (20) (18). Hence, it was proposed that the Xist gene has a role in controlling X chromosome inactivation. The XIST/Xist mRNA is 15 kb in mouse and 17 kb in humans but contains no long open reading frame with no protein coding potential (18) (20). Regions of homology between the mouse and human genes are highest in repeated sequences with low protein coding potential. It was hypothesized that Xist/XIST either functions as an RNA molecule or as a sterile transcript where transcription through that region of the genome results in a change in DNA structure that favors inactivation (18) (20). Furthermore, Xist expression was shown to precede X chromosome inactivation in the female mouse embryo (58). Therefore the timing of Xist expression is also consistent with the idea that Xist may play a role in

initiation of X inactivation.

Embryonal carcinoma (EC) cells resemble cells of the early embryo. Female EC cells with two X chromosomes may bear two active X chromosomes or one active and one inactive X chromosome (69) (74) (71). We examined Xist expression in the P10 line of female EC cells that have two active X chromosomes. . P10 cells can be induced to differentiate with retinoic acid. This differentiation is accompanied by inactivation of one X chromosome in each cell (74) (96). Xist expression was examined in P10 cells in the study described in Chapter I of this thesis. These results were also published in a paper in 1994 (110). Northern blots show that Xist is expressed prior to X chromosome inactivation, which confirms the finding published by other using female embyos (58). In addition, expression of Xist increases in differentiating cells undergoing X chromosome inactivation.

P10 cells were made using embryos derived from a cross between a female *Mus musculus* mouse and a feral Danish male mouse. Therefore, P10 female embryonic carcinoma cells are heterozygous for many genes including Xist. The two Xist alleles in P10 cells contained Restriction Length Polymorphisms (RFLP) which were used to identify which allele of Xist was expressed. Chapter I of this thesis as well as Tai et al. (110) describe experiments I did using Reverse-transcription polymerase chain reaction (RT-PCR) to monitor expression of the Xist. The RT-PCR products were digested with Hae III or Hind II whose recognition sites were present in one allele of Xist but not the other. The results indicate that Xist expression prior to X chromosome inactivation bi-allelic

(expressed from both X chromosomes). This was a novel finding at the time I published this study (110) that challenged existing models for Xist function. Subsequent studies from other labs (9) (92) confirmed my results. Instead of RT-PCR used in my work, these newer studies employed allele-specific probes for RNA-FISH to allow for the detection, in single cell, of Xist transcripts derived both alleles.

Previous to the publication of my studies and others which showed bi-allelic expression; it was the induction of Xist expression that triggered the inactivation of the X chromosome from which the Xist transcript derived. Bi-llelic expression ox Xist in cells with two active X chromosomes indicate that the control of X chromosome is more complicated. Work done by Sheardown et al (107) and Panning et al. (92) show that Xist transcripts derived from the future active X chromosomes are selectively degraded. On the other hand, the Xist transcript from the inactive X is stabilized and becomes physically associated to the inactive X. Therefore, Xist expression is controlled by degradation of the Xist transcript.

There has been some controversy concerning how Xist transcript degradation is regulated. One study shows the presence of transcripts containing different 5' ends using RT-PCR (107). This result was thought to indicate the presence of early (Po) and late (P1/2) promoters driving Xist transcript expression. Transcripts driven by the Po promoter were found to be less stable than transcripts driven by P1/2. They hypothesize that switching from Po to P1/2 is what triggers Xist transcript degradation. However, others have found that there

is an antisense transcript that overlaps with the Xist gene, Tsix. The 3' end of the Tsix gene is the same as the 5' end of Xist (118). These authors claim that is evidence against the Po promoter. At present, it is still unclear how Xist degradation is achieved.

Materials and Methods

Cell lines and culture conditions. P10 cells were maintained in alpha medium (Gibco) supplemented with 10% fetal bovine serum (Hyclone) and cultured on a layer of irradiated STO fibroblast feeder cells (74). P10 (X^pO) and P10 (OX^m) are clones derived from P10 cells in which the maternal or paternal X chromosome has been lost. These cells are maintained in the same media supplemented with leukemia inhibitory factor (LIF, (122)) and on culture dishes without STO feeder fibroblasts. Cultures of P10 cells can become overgrown with cells that have lost an X chromosome. To ensure that a culture was comprised of populations in which each cell had two X chromosomes, P10 cells were plated at low density and colonies grown up from single cells. The alleles of the X-linked phosphoglycerate kinase gene (Pgk-1) in P10 cells differ at a single nucleotide in their coding regions (14) which results in two PGK-1 isoforms that migrate differently on cellulose acetate electrophoresis (26). Colonies were cut in half and one half was used to assay for the presence of two Pgk-1 alleles by cellulose acetate electrophoresis and the other half passaged to 24-well plates. Colonies of cells that had both Pgk-1 alleles were expanded for use.

C86 (XX) cells are female embryonal carcinoma cells that contain one active and one inactive X chromosome (71) and were maintained in alpha medium with 10% fetal bovine serum.

To induce differentiation, P10 cells are plated at a density of about 3×10^6 cells/150mm culture dish without STO feeder cells. 24 h later retinoic acid (RA)

was added at a concentration of 10nM or 100nM. The media was replaced every 2 days.

Northern blot analysis. Total RNA was prepared by the guanidinium isothiocyanate procedure (27) and was electrophoresed in 0.7% agarose gels containing 1.1M formaldehyde. The RNA was electroblotted to sheets of Hybond N (Amersham) using a sandwich apparatus run at 100V for 4 h in 1X TAE (10mM Tris, pH7.5, 5mM sodium acetate, 0.5mM EDTA) with cooling. Blots were hybridized to [^{32}P]-labeled DNA probes. Hybridizations were done at 42°C in a buffer containing 50% formamide, 5X Denhardt's solution, 5X SSPE (1X SSPE = 10mM sodium phosphate, 0.15M NaCl, 1mM EDTA, pH 7.4), 1% SDS and 100mg/ml denatured salmon sperm DNA. Washes were performed twice in 2X SSC/0.1% SDS at room temperature for 15 min and twice in 0.2X SSC/0.1% SDS at 65°C for 15 min.

Probe DNAs. A 600bp Xba I fragment of the Xist cDNA 3' end was used as a probe (15). The Pgk-1 probe was a 1.7 Eco RI fragment of the mouse Pgk-1 pseudogene called B13 (2). The Endo A probe was a 2.2kb Eco R I fragment of the processed pseudogene (116). The Oct-3 probe was a 220 bp Ava III-Eco RI fragment from the cDNA (89). DNA was labelled with [a-^{32}P]dCTP by random primed replication.

Genomic DNA preparation. Cells were extracted in 10mM Tris, pH8.0, 10mM

EDTA, 10mM NaCl, 0.5% SDS with 50mg/ml Proteinase K at 55°C overnight. DNA was then extracted gently once with phenol and twice with chloroform:isoamyalcohol (24:1).

Polymerase Chain Reaction (PCR) conditions. PCR amplification of genomic DNA was with 2.5units/100µl Taq polymerase (BRL) in BRL Taq buffer (with 2.0µM $MgCl_2$ and 1mg/ml BSA, 200µM dNTPs) and 0.25mmol primers for 30 cycles (95°C for 40 sec, 55°C for 1 min, 72°C for 2 min) in the thermal cycler (MJ Research).

Reverse Transcription - Polymerase Chain Reaction (RT-PCR). Total RNA was treated with RNase-free DNase (0.02U/mg RNA, BRL) for 10min at room temperature before reverse transcription. 2mg of total RNA was reversed transcribed with Moloney murine leukemia virus reverse transcriptase (200U/ml, BRL) for 1hr at 37°C using 10pmol of specific primers. For each sample of RNA a control experiment was carried out in parallel omitting the reverse transcriptase (-RT). One fiftieth of the cDNA or -RT was used for PCR amplification with Taq polymerase (BRL). PCR conditions were the same as above. RT-PCR products were analyzed on a 1% agarose gel. Gels were blotted by capillary transfer and probed with labelled PCR product.

Sequencing. Genomic DNA amplified using PCR with primers containing Bam

HI restriction enzyme recognition sites was subcloned into Bluescript (Stratagene). Double stranded sequencing was done using dideoxy chain termination and [^{35}S]dATP (Amersham). The sequencing reactions were done using a Pharmacia kit with universal primer.

Restriction Fragment Length Polymorphism (RFLP) analysis. PCR or RT-PCR products were digested with restriction enzymes Hae III or Hind III. Digested DNA was electrophoresed on 5% polyacrylamide mini-gels. Gels were electroblotted in a sandwich apparatus at 100V for 1h in 1X TAE (10mM Tris, pH7.5, 5mM sodium acetate, 0.5mM EDTA) with cooling. Blots were hybridized with labelled PCR product.

Results

A) P10 cells express Xist

Northern blots of total RNA from female somatic tissue was probed with the 3' end of the Xist cDNA and showed a smear of hybridizing material along with two bands of approximately 14.3 and 11.8kb (Figure I.I.1a, lane 2). No hybridization to RNA from male somatic tissue was evident. These results are similar to those of Borsani et al. (15) and Brockdorff et al. (17) and are consistent with Xist expression from the inactive X chromosome. We examined Xist expression in P10 female embryonal carcinoma cells with two active X chromosomes. The two Xist transcripts were present in these P10 cells (Figure I.1a, lanes 3 and 13). Because both X chromosomes are active in P10 cells, this result indicates that expression was not strictly associated with the inactive X chromosome.

Reverse Transcription-Polymerase Chain Reaction (RT-PCR) confirmed the results obtained with northern blots. Figure I.2a shows RT-PCR amplified Xist transcripts from P10 cells (Figure I.2a, lane 1). Because the amplified region was present in the last Xist exon, we carried out the reaction with (+) and without (-) reverse transcription to ensure that the 397bp signal derived from cDNA and not contaminating genomic DNA. RT-PCR of Xist transcripts was also done with primers that spanned the spliced region of the Xist gene from the 3' end of the first exon to the 5' end of the last exon (Figure I.2b). The 601bp PCR product from P10 cells with two active X chromosomes (Figure I.2b, lane 1) indicated that these cells expressed spliced transcripts of the Xist gene.

C86 (XX) is female embryonal carcinoma cell line with one active X chromosome and one inactive X chromosome (71). This cell expresses Xist (Figure I.1a, lane 8; Figure I.2a lane 6); Figure I.2b, lane 6) whereas the male embryonal carcinoma cell line, P19 (XY), did not (Figure I.1a, lane 11). Irradiated STO fibroblasts, which were used as feeder cells in P10 cell cultures did not express Xist (Figure I.1a, lane 12).

B) Xist expression is increased with X chromosome inactivation

P10 cells differentiate when exposed to retinoic acid (RA). This differentiation is accompanied by inactivation of one of the X chromosomes in each cell 3-4 days following RA treatment (96). Figure I.1a shows that Xist transcript levels increased after RA treatment. The biggest increase in level of Xist occured between 1 and 3 days after RA (Figure I.1a, lanes 4 and 5) when X inactivation occurs. The experiment was repeated with a 10 fold higher dose of RA (Figure I.1a , lane 14). When Xist expression was normalized to loading controls (Figure I.1b), transcripts were found to increase 6.5 fold after 10nM RA treatment (Figure I.1a, lane 3 versus lane 5) and 16 fold after 100nM RA treatment (Figure I.1a, lane 13 versus lane 14) reflecting the more efficient induction of differentiation induced by the higher dose of RA (see below). Hybridization to the 14.3kb transcript was 2-5 fold more intense than the 11.8kb band in differentiated cells (Figure I.1a, lanes 5-7), whereas in undifferentiated cells with two active X chromosomes (lane 3) almost equal amounts of the two transcripts were present.

To confirm that P10 cells were differentiated with RA in these experiments the expression of two genes, Endo A and Oct-3, was analyzed using northern blots. Endo A encodes a cytokeratin and is expressed in visceral and parietal endoderm but not in undifferentiated embryonal carcinoma cells (116). Endo A expression in P10 cells increased with RA induced differentiation (Figure I.1c). Oct-3 is a transcription factor expressed in embryonal carcinoma cells but down-regulated after RA induced differentiation (89). Figure I.1d shows that Oct-3 is expressed in all the embryonal carcinoma cell lines analyzed but not in adult tissue or STO fibroblasts. In P10 cells the level of Oct-3 transcripts is diminished after RA induced differentiation but only at the higher dose of RA (Figure I.1, lane 14). These data show that P10 cells had characteristics of undifferentiated embryonal carcinoma cells and differentiated into an endoderm-like cells after RA treatment (96).

C) Both alleles of Xist are expressed

We further examined the expression of Xist in P10 cells to determine whether the Xist transcript was derived from one X chromosome or both. To do this the paternal and maternal alleles of Xist had to be distinguished. The P10 cell line was derived from an embryo obtained from a cross between a male mouse carrying an X chromosome originating from a feral Danish animal (26) and a female *Mus musculus*. Thus, a number of allelic differences exist between maternal and paternal X-linked genes in P10 cells. To characterize polymorphisms we made use of clones of P10 cells that had lost an X

chromosome: P10 (X^pO) and P10 (OX^m) are clones that have retained the paternal or the maternal X chromosome respectively. Genomic DNA from P10 (X^pO) and P10 (OX^m) cells amplified with primers exA and exB (Table I.1, Figure I.3a) yielded a 397bp fragment that was digested with Hae III. This revealed a polymorphism allowing us to distinguish paternal and maternal Xist alleles (V. Chapman, personal communication and Figure I.3a). The maternal Xist allele yielded Hae III fragments of 72, 140, 185 bp (Figure I.3b , P10 (OX^m) lane C) while the paternal allele resulted in two bands of 72 and 325 (Figure I.3b, P10 (X^pO) lane C). Sequencing of the two alleles revealed a 15bp deletion in the paternal allele from P10 (X^pO) that encompassed the Hae III site that digests the 325bp fragment into 140bp and 185bp fragments (Figure I.3a).

The Hae III RFLP was then used to analyze RT-PCR products to see if the expression of Xist prior to X chromosome inactivation was derived from both X chromosomes. P10 cells used in our experiments are clonal cell lines derived from a single cell (74). RNA was isolated from three P10 (XX) subclones and reverse transcribed as described in Figure I.2. These cDNAs were amplified with primers exA and exB. RNA preparations that were not reverse transcribed were also subjected to PCR with the same primers but no PCR product was present indicating that there was no genomic DNA contaminating these RNA preparations. The amplified RT-PCR products were then digested with Hae III, electrophoresed, blotted and hybridized for Xist sequences. Four bands of 397, 325, 185, 140 bp were in the RT-PCR products from each of the P10 (XX) cell

lines (Figure I.3c, -RA samples). The predicted band of 72bp was not detected likely because it failed to adhere to the membrane. The band at 397bp represented undigested DNA. The band at 325 bp derived from cDNA of the paternal allele (which lacks the Hae III site). Bands at 185 and 140 are the digested fragments from the cDNA of the maternal Xist allele. The results indicate that both Xist alleles were expressed in undifferentiated P10 cells.

A second Xist polymorphism was investigated. A Hind III site is present in exon 5 of the maternal but not the paternal Xist allele ((22), see Figure I.4a). cDNA from P10 (XX) cells was amplified with primers spA and spB and the 601bp product digested with Hind III. C86 (XX) cells yielded 3 bands of 320, 204, 77 bp (Figure I.4b). Undifferentiated P10 (XX) cells yielded four Hind III fragments of 397, 320, 204, 77 bp (Figure I.4b, -RA). The paternal allele of P10 (XX) cells lacks the Hind III site that cut the 397 bp fragment into 320 bp and 77 bp (Figure I.4a). The presence of the four bands indicates that both alleles of the spliced Xist transcripts were present in undifferentiated P10 cells.

In the experiment shown in Figure I.4b, the 320 and 77 bp bands from P10 cells were less intense than expected. In other experiments we noticed that the relative abundance of the 320 and 77 bp DNAs varied slightly but never exceeded the intensity of the bands at 397 and 204 bp. We suspect that there is a slight bias in amplification of the cDNA products from the two alleles and that the higher intensity of the 397 band is not indicative of preferential expression of this allele. No difference in allele abundance was detected with the Hae III

polymorphism shown in Figure I.3c.

Both Hind III and Hae III polymorphisms of Xist cDNAs indicate that both alleles were expressed in undifferentiated P10 (XX) cells. Both Xist alleles were also expressed in cultures of RA treated P10 (XX) cells (Figures 3c and 4b, +RA). Since each cell in these differentiated cultures could have randomly inactivated either X chromosome (74), the population would be expected to consist of equal numbers of cells expressing each Xist allele.

D) Xist is not expressed in P10 cells that have lost an X chromosome

P10 cells that had spontaneously lost an X chromosome were examined for Xist expression. RNA from P10 ($X^{?}O$) and P10 ($OX^{?}$) were probed for Xist. Figure I.1a shows that P10 ($X^{?}O$) and P10 ($OX^{?}$) cells do not contain Xist transcripts. P10 ($X^{?}O$) and P10 ($OX^{?}$) cDNA did not give amplified products with primers for Xist (data not shown). Thus, P10 cells with two active X chromosomes express Xist but that expression is lost in cells that have lost one or the other of the X chromosomes.

Table and Figures

Table I.1: Oligonucletide primers

	Nucleotide	primer location	sequence(5'-3') number*
exA	TTAGGATCCAGACTGCCCTGAGAAATAATC	exon 6	11723
exB	CCTTCTTGGTCTTGGGGATAG	exon 6	12004
exC	CCCGGATCCTGCATTAGTGACCTAGTTT	exon 6	13484
spA	CCCGGATCCCTCTGGCTGTTTAGACTACA	exon 1	9473
spB	TTAGGATCCGTAATATTTGGTAGATGGCA	exon 6	10083
spC	GGCATGTTGATCCTCGGGTC	exon 6	10186

*5' nucleotide numbers refer to Brockdorff et al., 1992 (19)

underlined sequences indicate Bam HI restriction site extension

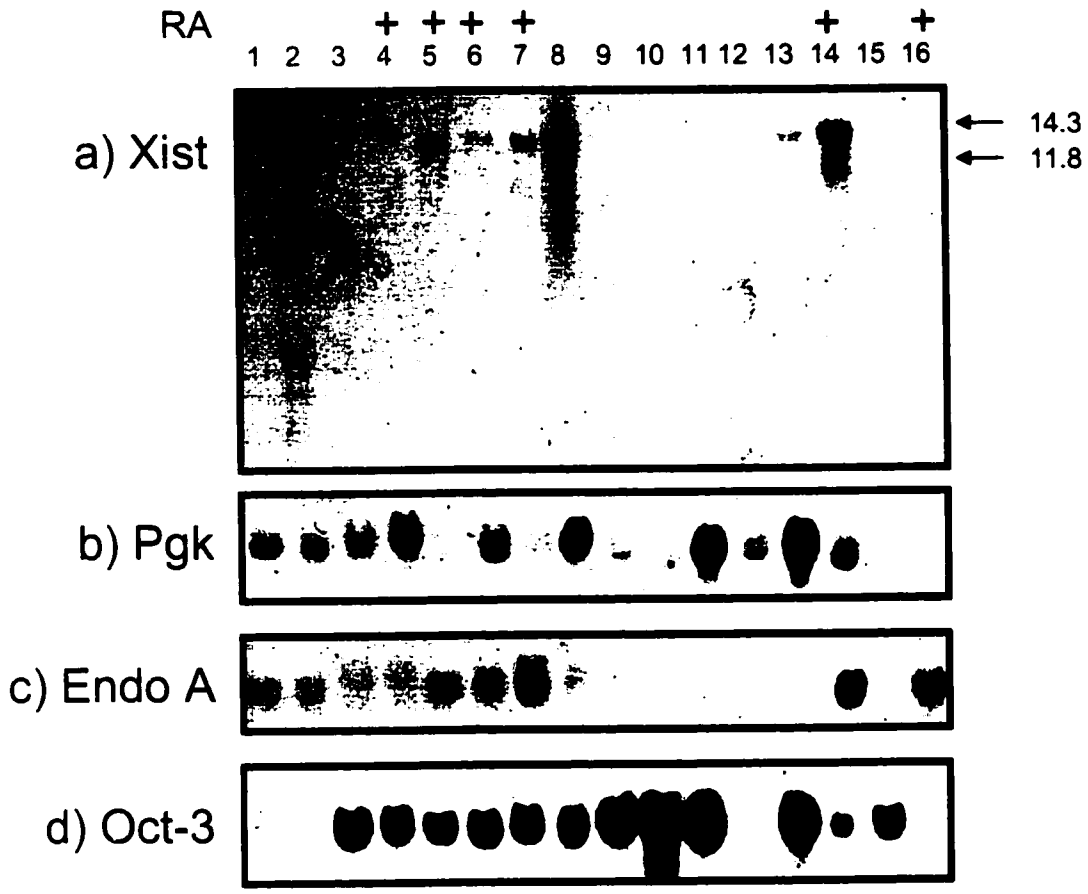
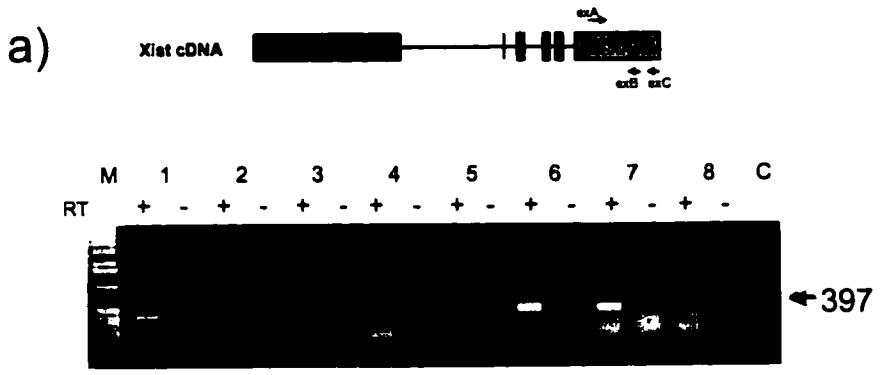

RA
+ + + + + +
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
a) Xist
14.3
11.8
b) Pgk
c) Endo A
d) Oct-3

Figure I.1. Northern blots of RNA showing Xist expression as two large transcripts from P10 cells. Each lane contains 20mg of total RNA from **1)** male and **2)** female C3H mouse adult kidney; **3)** P10 female embryonal carcinoma (EC) cells; P10 cells incubated in 10nM retinoic acid for **4)** 1 day, **5)** 3 days, **6)** 5 days, **7)** 7 days; **8)** C86 (XX) female EC cells; **9)** P10 (X^pO) EC cells; **10)** P10 (OX^m) EC cells; **11)** P19 (XY) EC cells; **12)** STO feeder fibroblasts; **13)** P10 cells; **14)** P10 cells incubated in 100nM retinoic acid for 11 days; **15)** P10 (X^pO);**16)** P10 (X^pO) incubated in 100nM retinoic acid for 11 days. Blots were probed with a) Xist , b) Pgk-1, c) Endo-A and d) Oct-3. The arrows in panel a indicate the two large Xist transcripts.


a)
Xist cDNA
exA
exB
exC
M 1 2 3 4 5 6 7 8 C
RT + - + - + - + - + - + - + - + -
397

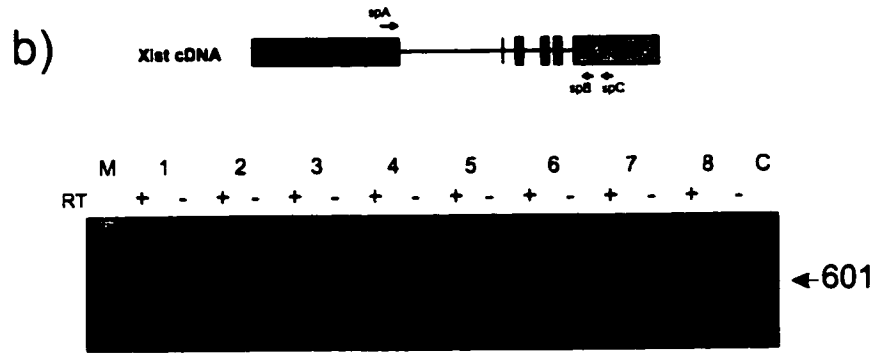
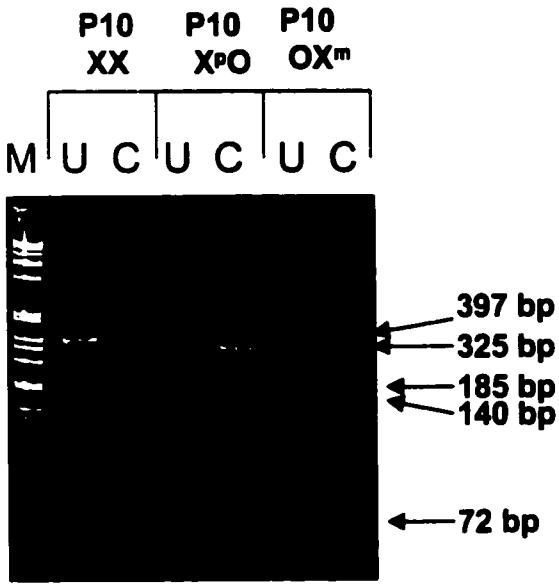
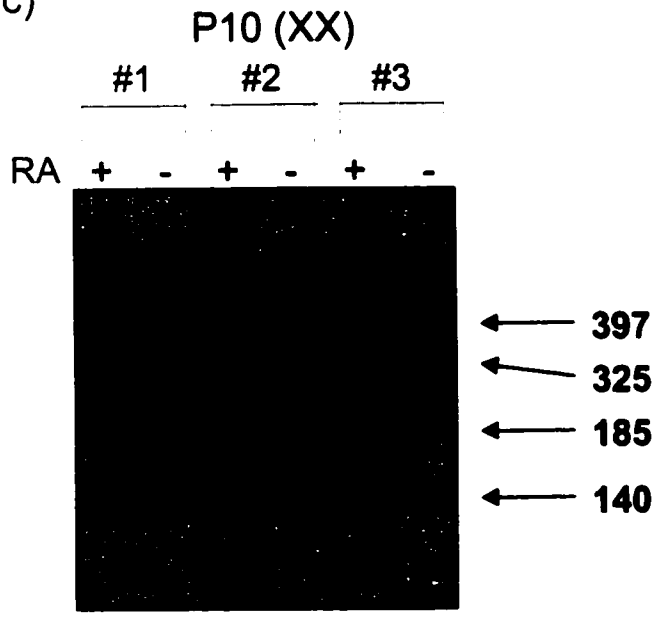

b)
Xist cDNA
spA
spB
spC
M 1 2 3 4 5 6 7 8 C
RT + - + - + - + - + - + - + - + -
601

Figure I.2. RT-PCR analysis shows Xist expression in P10 cells. Total RNA used for RT-PCR was from **1)** P10 EC cells; P10 cells incubated in retinoic acid for **2)** 1 day, **3)** 3 days, **4)** 5 days, **5)** 7 days; **6)** C86 (XX) EC cells; **7)** female and **8)** male adult kidneys; **C)** negative control (no DNA). a) Primer exC was used for cDNA synthesis followed by PCR with primers exA and exB. A 397bp product was amplified from exon 6 of the Xist gene. b) Primer spC was used for cDNA synthesis followed by amplification with primers spA and spB. The 601bp amplified product derived from Xist RNA from which all 5 introns had been spliced. The RT-PCR products were run on a 1% agarose gel. Lanes labeled "+" contained samples subjected to amplification following reverse transcription while the lanes labeled "-" contained RNA that had not been reverse transcribed as a control for contaminating DNA. M= marker (1kb ladder from BRL).

b)


P10 XX
P10 XPO
P10 OXm
M U C U C U C
397 bp
325 bp
185 bp
140 bp
72 bp

c)


P10 (XX)
#1
#2
#3
RA + - + - + -
397
325
185
140

a) Exon 6

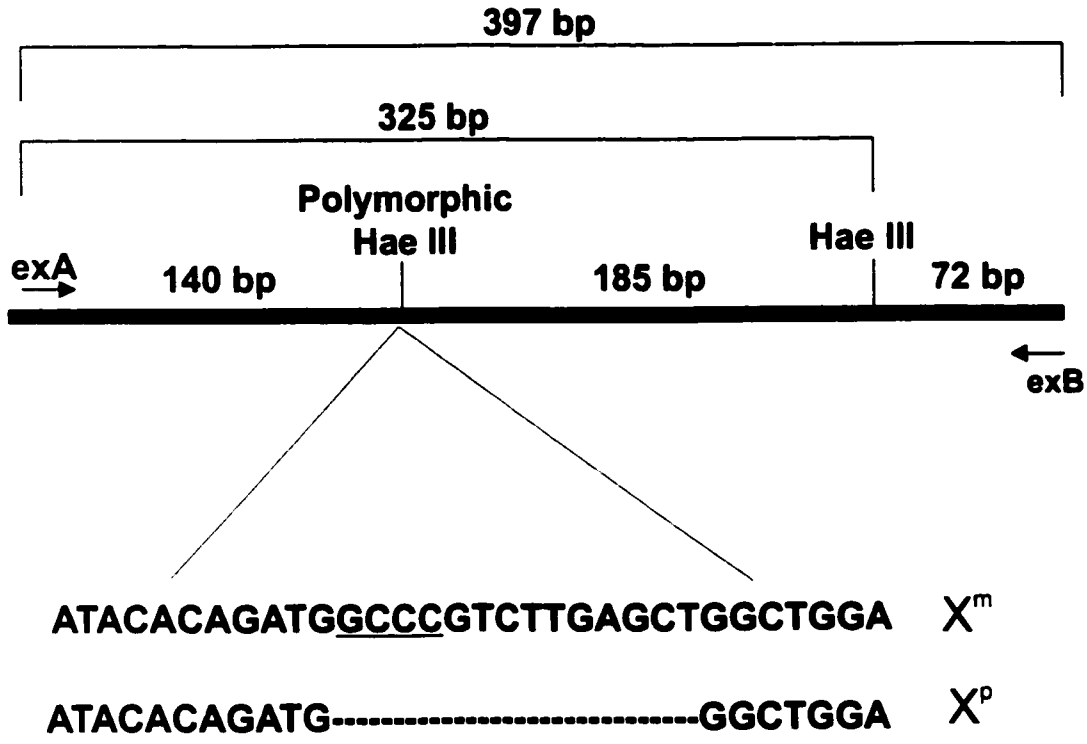
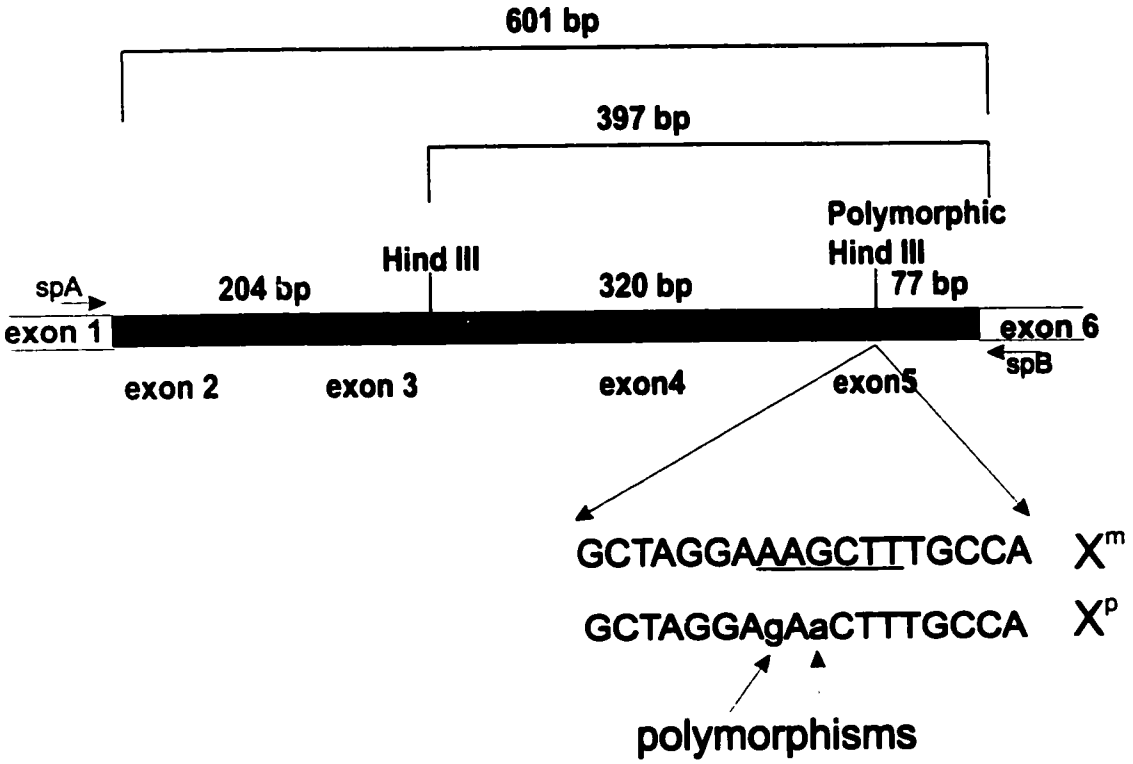

397 bp
325 bp
Polymorphic
Hae III
Hae III
exA
140 bp
185 bp
72 bp
exB
ATACACAGATGGCCCGTCTTGAGCTGGCTGGA
Xm
ATACACAGATG----------GGCTGGA
Xp

Figure I.3. Hae III restriction fragment length polymorphism in Xist cDNAs indicate that P10 cells express Xist from both X chromosomes in P10 cells. a) Schematic diagram of the 397bp region from Xist exon 6 amplified by PCR using exA and exB primers (Table I.1). Genomic DNA from P10 ($X^{s}O$) was used to amplify the paternal allele (X^{s}) and genomic DNA from P10 (OX^{m}) was used to amplify the maternal allele (X^{m}). The sequences of these Xist alleles around the 15 base pair deletion in exon 6 of the paternal allele are shown. The deletion includes a Hae III restriction enzyme site which is underlined in the sequence of the maternal allele. The sizes if the Hae III fragments of the maternal allele are shown above the line. b) Characterization of Hae III Xist polymorphism. Genomic DNA from P10 (XX), P10 ($X^{s}O$), or P10 (OX^{m}) was amplified using primers exA and exB and the products digested with Hae III and run on a 5% polyacrylamide gel. U=uncut, C=cut with Hae III, M=marker. The labels on the right indicate the expected fragments. c) RNA from three P10 (XX) subclones were prepared from undifferentiated (-) cells or from cultures treated for 11 days with 100 nM RA (+). This RNA was reverse transcribed and amplified as described above, digested with Hae III, run on a 5% polyacrylamide gel, electroblotted and probed with the Xist cDNA. Distinct bands of 325bp, 185bp, and 140bp indicate the presence of both Xist transcripts in each RNA sample.

a) Spliced region



b)

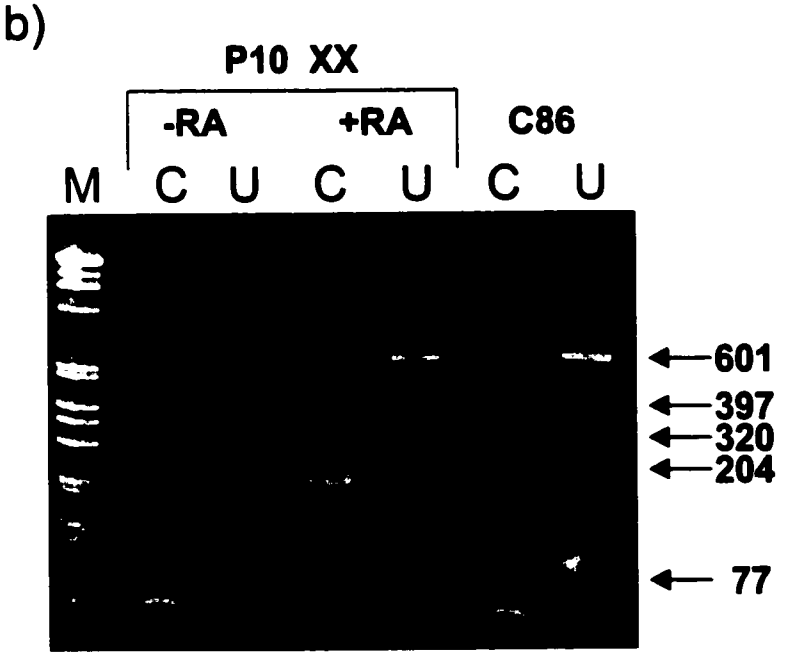


Figure I.4. Hind III restriction fragment length polymorphism indicates expression of Xist from both X chromosomes in P10 cells. a)Schematic diagram of the 601bp region of Xist cDNA amplified using spA and spB primers (Table I.1). The 601bp product is spliced without introns. The boundaries between exons is shown. The sequence around the polymorphic Hind III site is shown (16). The sizes of the Hind III fragments from the maternal allele are shown above the line. b) Total RNA was isolated from undifferentiated P10 (XX) cells (-RA), P10 (XX) cells treated for 11 days with 100nM RA (+RA), or C86 (XX) cells. This RNA was reverse transcribed and amplified as described above. The DNA products were digested with Hind III, run on a 5% polyacrylamide gel and visualized following ethidium bromide staining. The expected fragments are indicated on the right. The presence of bands of 397bp, 320bp, and 77bp in samples of cDNA from P10 cells indicates the expression of both Xist alleles in these cultures. U=uncut, C=cut with Hind III, M=marker.

Discussion

The results presented here show that Xist transcripts are present in P10 female embryonal carcinoma cells with two active X chromosomes and that the Xist genes on both X chromosomes are expressed. Studies from other labs with somatic tissues and male germ cells demonstrated that Xist transcripts were only detected in cells carrying an inactivated X chromosome. The results of Chapter I, on the other hand, show that Xist expression is not exclusively derived from inactive X chromosomes in embryonic cells. The implication is that the activation of Xist expression alone does not induce X chromosome inactivation.

However the expression of both Xist alleles detected in P10 cell cultures could have derived from subpopulations of cells that spontaneously differentiated and inactivated one X chromosome. We think this interpretation is unlikely. The level of Xist transcript in undifferentiated P10 cells is at least 10-20% of that present in C86 (XX) cultures with an inactive X chromosome in each cell. Cultures of P10 cells do not contain any cytologically detectable inactive X chromosomes (96) and the population of cells in these cultures expressing a differentiated cell marker (Endo A) is less than 5%. In addition, our experiments were performed with freshly cloned populations of cells that had been in culture less than 3 weeks. Thus, it seems unlikely that the Xist expression observed could derive from a differentiated subpopulation within the P10 cultures.

In female mouse embryos X inactivation first occurs in the trophectoderm

and the paternal X chromosome is preferentially inactivated (93) (111). Xist expression in embryos was first detected from the paternal allele before X inactivation had occured (57). Thus, Xist expression appeared to precede X inactivation and its paternal-specific expression suggests a role for Xist in realizing preferential paternal inactivation in trophectoderm. P10 cells are thought to be derived from the embryonic ectoderm region of the embryo in which either the maternal or paternal X chromosome can inactivate. If the resemblance of P10 cells to embryonic ectoderm holds, our results suggest that both Xist alleles are initially active in each cell of the embryonic ectoderm of female embryos prior to X inactivation. The abundance of the Xist transcripts increased in differentiating P10 cells at the time when these cells were undergoing inactivation of one X chromosome. Whether this increased expression is causally involved with X inactivation is not yet clear

Clones of P10 cells that had spontaneously lost an X chromosome, P10 ($X^{p}O$) and P10 (OX^{m}), fail to express Xist from the remaining allele suggesting that Xist expression is regulated by the ratio of X to autosomal chromosomes. In *Drosophila* and in nematodes the X:autosome ratio determines sexual characteristics and dosage compensation of X-linked genes (6) (50). The X:autosome ratio in *Drosophila* appears to be determined in early embryos by dimerization of transcription factors of the basic helix-loop-helix (bHLH) type (94). By analogy, the loss of Xist expression in P10 ($X^{p}O$) and P10 (OX^{m}) cells may be dependent on the reduced proportion of bHLH factors encoded by X-linked genes.

In subsequent studies in other labs Xist expression was confirmed to be expressed in embryonic cells from both X chromosomes prior to X inactivation in experiments in another lab (92). There is selective degradation of Xist transcripts derived from the X chromosome that will remain active (92) (107). The stabilized Xist transcript derived from the inactive X coats one of the X chromosome in cis and this intimate association of the Xist transcript and the X chromosome is part of the mechanism for X chromosome inactivation (28) (92) (107). This is also further evidence that Xist encodes a functional RNA molecule.

The Xist transcription initiation site further characterized using RTPCR (107). Evidence from these studies indicated the presence of two promoters, early expression of Xist is initiated from the P_0 promoter and later expression is stimulated from the P_1/P_2 promoter. The switch from the P_0 promoter to the P_1/P_2 occurred on only one X chromosome leading to the stabilization of the Xist transcript derived from that chromosome. However, there is some controversy concerning the existence of the P_0 promoter as another gene, Tsix, was found that overlaps Xist in the antisense orientation (65). The 3' end of the Tsix transcript coincides with the P_0 promoter (118). The Tsix transcript also has no open reading frame and is thought to function as a RNA molecule like Xist (65).

Deletions of the Tsix gene indicate that it is involved in the choice of which X chromosome inactivate (65) (35). At present, it is unclear if Tsix has a role in the degradation of the Xist transcript. Further studies are needed to elucidate the mechanism of Xist transcript degradation.

Observation made by others have confirms the results presented in the first chapter of this thesis that Xist is expressed prior to X chromosome inactivation. In addition, evidence presented in Chapter I that Xist expression is bi-allelic has also been confirmed.

Chapter II:

A study of the CHD-1 protein

Introduction

A) Chromodomain proteins and gene silencing

The maintenance of gene expression patterns established during development is another type of gene silencing. The silencing of homeobox genes in *Drosophila* is an example of this type of gene silencing. After segment-specific expression of homeotic genes is initiated early in *Drosophila* development, the repression of homeotic gene expression in the inappropriate segments is maintained by the Polycomb group of genes. Mutations in the Pc protein result in the loss of repression of homeotic gene expression and ectopic expression of posterior phenotypes was observed inappropriately in anterior segments as would be predicted (54). The repression of homeotic genes is classified as gene silencing since the regions of repressed chromatin established by the Polycomb group are stable throughout development and are passed on to subsequent cell generations (98).

The chromodomain was first described as a domain of homology between the proteins Polycomb (Pc) and HP-1 (95). The Pc protein is one of a complex of proteins called the Polycomb group (54). HP-1 is another protein with a chromodomain and it is also involved in gene silencing in a separate phenomenon, position effect variegation in *Drosophila* (38). Position effect variegation occurs when a gene is transposed to heterochromatin containing inactive genes. The translocated gene is usually an eye color gene whose gene product is easily measured. The repressive nature of heterochromatin in this

region spreads to the translocated eye color gene. The spreading of heterochromatin, however, varies giving rise to cells that either express eye color or not. The result is the development of eyes with variegated color. A lot of heterochromatin spreading gives rise to flies with eyes that have very few patches of color; conversely, little heterochromatin spreading results in eyes with many patches of color. This phenomenon is known to be regulated by a number of proteins which when mutated either increase or decrease the spread of heterochromatin and hence the amount of eye color. For example, loss-of-function mutations in the *Su(var) 2-5* gene encoding the protein HP-1, result in more eye color and gain-of-function mutations result in less eye color (38). The behavior of HP-1 is consistent with it being a repressor of gene expression.

The chromodomain is a stretch of 37 amino acids (a.a.) from a.a. 26 to 62 in Pc and a.a. 24 to 60 in HP-1 (95). 65% (24/37) of amino acids are identical and 7/37 are conservative changes between the two proteins. Chromodomain mutations in Pc and HP-1 result in proteins that do not mediate gene silencing indicating the importance of this domain. The presence of a chromodomain in two proteins that were involved in two different gene silencing phenomena not only indicates a role of the chromodomain in gene silencing but also that the mechanism of silencing is similar. The chromodomain has subsequently been found to be present in a variety of genes in plant and animal organisms (108) which further underlines the importance of the chromodomain.

B) CHD-1, a mouse protein with two chromodomains

Since chromodomain proteins in *Drosophila* were associated with gene silencing it was of interest to find out if mammalian chromodomain proteins were also involved in gene silencing. This thesis describes a study the mammalian chromodomain protein, CHD-1. The CHD-1 protein is classified in a group of chromodomain proteins, those with two chromodomains near the amino-terminus (1). CHD-1 was subsequently found to be part of a family of CHD proteins that share the signature domain structure of two **C**hromodomains, a **H**elicase domain, and a **D**NA binding.

The chromodomain of other proteins indicated a role in protein-protein interactions. Pc and HP-1 are localized to distinct regions on polytene chromosomes. In an experiment where the chromodomains of Pc and HP-1 were exchanged it was found that the chromodomain determines the location of binding on polytene chromosomes, i.e. the Pc bound to the HP-1 polytene chromosome site if it had the HP-1 chromodomain (78) (99). The interpretation is that the chromodomain determined the location of the protein on the polytene chromosome through specific protein interactions.

The NMR structure of the chromodomain of the MoMOD1 protein, a mouse homolog of HP-1, supports the hypothesis that the chromodomain is involved in protein interactions. The chromodomain consists of an amino-terminal three-stranded anti-parallel ß sheet folding against a carboxy-terminal α helix (7). This structure is similar to the DNA binding domains from the archebacterial

proteins Sac 7d and Sso 7d. The particular charge distribution for the binding to the major groove of DNA is not present in the chromodomain so the proposed function of the chromodomain is not to bind DNA but rather to bind other proteins.

The studies described above support a role for the chromodomain in protein-protein interactions. Identification of proteins that interact with the chromodomain would not only confirm the protein-binding property, but the function of the chromodomain might be elucidated if the interacting proteins have established functions. This is because the chromodomain would likely have a role in the function of the proteins it interacts with. Therefore a search for proteins that interact with a chromodomain was done as part of this thesis.

C) CHD-1 has a helicase/ATPase domain

CHD-1 also has a helicase/ATPase domain that is homologous to the SWI2/SNF2 protein in yeast. The SWI2/SNF2 protein is part of the SWI/SNF complex that also includes SWI1, SNF5, SNF6, and SWI3 (32). This complex enhances transcription by a number of sequence-specific activators such as GAL4 (97), which might suggest that CHD-1 could also be a gene activator rather than a silencer. However, CHD-1 homology to SWI2/SNF2 is limited to the helicase/ATPase domain and this domain is also present in other proteins that are not transcriptional activators. In fact, another helicase/ATPase containing protein in yeast, MOT1, was found to be a transcriptional repressor (4). Therefore, having a helicase domain does not rule out the idea that CHD-1

is a gene silencing protein.

Instead of transcriptional activation the presence of the helicase/ATPase domain in CHD-1 may indicate that the protein remodels chromatin. Chromatin remodelling or nucleosome remodelling is the shifting of histones in the nucleosome that can result in the exposure of regulatory DNA sequences to transcription factors. Studies of the SWI/SNF complex showed that it is a chromatin remodelling complex that displaces histones from the GAL4 DNA binding site (91). Furthermore, the complex was shown to displace histones by sliding them along DNA in cis (121). Chromatin remodelling by the SWI/SNF complex is dependent on the SWI2/SNF2 helicase/ATPase domain that is homologous to CHD-1. Therefore the helicase/ATPase domain of CHD may mediate chromatin remodelling as well. This is supported by the finding that the chromodomain and the helicase/ATPase domain of CHD-1 bind to chromatin (59).

D) CHD-4 is in a chromatin remodelling complex with histone deacetylase

Further evidence that CHD-1 could be involved in chromatin remodelling came from studies showing that another CHD family member, CHD-4, is in a complex that could remodel chromatin (113) (126). These studies also demonstrate that histone deacetylase, HDAC, is in the complex along with CHD-4. This complex is called the NURD complex (see below). Studies on histone acetylation revealed that histone acetylases and deacetylases provide another level of control of gene expression. Chromatin containing deacetylated

histones is associated with unexpressed genes and, conversely, acetylated histones are associated with chromatin containing active genes (115). The association of CHD-4 with histone deacetylase points to a role for CHD proteins in the gene silencing mediated by histone deacetylation.

Further evidence that histone acetylation is associated with transcriptional activation is the discovery that Gcn5, a transcriptional co-activator in yeast, is homologous to the histone acetyltransferase in *Tetrahymena*, HAT A (23). Therefore histone acetylation is probably part of the mechanism of transcription activation by Gcn5. Since the discovery that Gcn5 has histone acetylase activity other transcriptional activators have also been found to have histone acetyltransferase activity - (P/CAF (124), p300/CBP (8) (85), TAFII250 (81)) further indicating the importance of histone acetylation in gene activation.

Since transcriptional activators have histone acetylase activity, do repressors have histone deacetylase activity? In fact, mammalian histone deacetylases, mRpd3 (mouse) and HDAC 1 and 2 (human) have been isolated and were all found to have similarity to a yeast transcriptional repressor protein, Rpd3 (112) (56) (117). The mammalian histone deacetylases were found to function in large complexes of proteins. The first such complex, discovered in mouse, contained mRPD3 protein and the mSin3 protein (47). Sin3 was originally identified in yeast as a protein that could repress the same set of genes as the Rpd3 protein and both proteins were also found together in a large complex in yeast (56). The mSin3 protein is thought to "tether" the histone deacetylase to transcriptional repressors. This was concluded from

experiments that used column purification, immunoprecipitation and yeast two hybrid assays to show that the mSin3 protein bound both histone deacetylase and transcriptional repressor proteins such as Mad (86) (63) (47), N-CoR (86) (48) and SMRT (86). These studies also showed that transcriptional repression could be inhibited by histone deacetylase inhibitors, trichostatin A and trapoxin, implying that association of transcriptional repressor with mSin3 and histone deacetylase is a critical step in repression. The repressor complexes containing mSin3 are referred to as Sin complexes (see table below adapted from Ayer (5)).

A number of other complexes containing histone deacetylase have subsequently been discovered and one of these, the NURD complex, which contains the CHD-4 protein, was mentioned above. Some proteins such as HDAC1/2 and RbAp48 (a protein that was originally found to be associated with the retinoblastoma protein, Rb) were found in the NURD complex as well as Sin complexes. The NURD complex is also thought be a transcriptional repressor complex since it contains the HDAC's. Notably absent from the NURD complex, however, is the mSin3 protein; and absent from the Sin complex is CHD-4 (See table below adapted from Ayer (5)). This suggests that there are functional differences between the complexes.

CHD-4, like CHD-1, has a helicase/ATPase domain similar to that of the SWI2/SNF2 protein which led investigators to ask if the NURD complex could remodel chromatin like the SWI/SNF complexes. Chromatin remodeling stimulated by the NURD complex was found and was dependent on histone

deacetylase activity. Another unique feature of the NURD complex is that it can deacetylate histones in oligonucleosomes *in vitro*. The Sin complex, on the other hand, could only deacetylate the core histone octamer or mononucleosomes. The ability of the NURD complex to deacetylate oligonucleosomes was dependent on the chromatin remodeling activity of the helicase/ATPase domain. Therefore, it was suggested that CHD-4 is responsible for remodeling chromatin which would expose histones to the histone deacetylase. Since CHD-1 and CHD-4 are similar proteins, a similar function for CHD-1 is proposed; i.e., CHD-1 associates with other repressor proteins such as histone deacetylase to form a transcriptional repressor complex that remodels chromatin.

The Sin complex		The NURD complex	
protein	function	protein	function
HDAC1/HDAC2	Histone deacetylation	HDAC1/HDAC2	Histone deacetylation
RbAp46/48	HDAC targeting	RbAp48	HDAC targeting
SAP30a	adaptor	CHD-4	Nucleosome remodelling
SAP30b	adaptor	MTA1/MTA2	Metastatic potential
SAP18	?	p32	?

E) CHD-1 binds to AT-rich DNA like that at nuclear matrix attachment regions

The final signature domain of CHD-1 is the DNA binding domain which gel shifted AT-rich DNA (109). AT-rich DNA is characteristic of DNA that is attached to the nuclear matrix which is referred to as MAR (Matrix Attachment Region) DNA (13) (16) (79). Since CHD-1 bound to MAR-like DNA it is reasonable to propose that CHD-1 could be nuclear matrix protein. The nuclear matrix is operationally defined as the proteinaceous structure in the nucleus that remains after high-salt and detergent extraction and Dnase digestion (11) (31) (49). The nuclear matrix is thought to organize chromatin into loop domains and MAR DNA is responsible for attaching chromatin to the nuclear matrix base of the loop (79) (34).

Many proteins in the nuclear matrix have been identified and they include proteins involved in DNA replication (46), transcription (42) (83), and RNA processing (80) (123). Of particular interest are the nuclear matrix proteins that bound to MAR DNA. MAR binding proteins that have been identified include lamins which are proteins in the nuclear lamina (67); HET/SAFB, a protein that represses the hsp27 promoter (88); hnRNP-U/SAF-A, a protein in hnRNP particles that are involved in packaging and processing RNA (40); SATB1 which binds to the base of chromatin loops (34); and ARBP (attachment region binding protein)/ MeCP2 (methyl-CpG binding protein), a transcriptional repressor that binds to methylated DNA and is associated with histone

deacetylase (87).

Histone deacetylase (HDAC1) itself was also shown to be associated with the nuclear matrix in a study that examined proteins cross-linked to the nuclear matrix (106). HDAC1 binding to the nuclear matrix supported the argument that CHD proteins are part of the nuclear matrix since, as discussed in the previous section, the CHD protein, CHD-4, was also shown to be associated with HDAC1. Therefore, the function of the CHD family proteins could include association with the nuclear matrix. Furthermore, studies of the human SWI/SNF complex showed that the SWI2/SNF2 homologues, hbrm and BRG-1, were also found in the nuclear matrix (102). This finding also supports the argument that CHD proteins are part of the nuclear matrix since the CHD family proteins also have homology to SWI2/SNF2 over the helicase/ATPase domain. However, it is not known if the helicase/ATPase domain of hbrm and BRG-1 are involved in the nuclear matrix association of these proteins. Since CHD-1 is a member of the CHD family and it binds MAR-like DNA it is proposed in this thesis that CHD-1 is a nuclear matrix protein.

F) The study of CHD-1

The study of CHD-1 described in this chapter starts with a characterization of the cellular localization of CHD-1 protein carrying deletions of the signature domains - the chromodomains, the helicase/ATPase domain and the DNA binding domain. The function of all the domains appeared to be in the nucleus. In the next section the CHD-1 protein was found in the nuclear matrix fraction of

a cell extraction. This concurred with studies by others (109) that showed the CHD-1 protein bound to MARs-like AT-rich DNA. The smallest mutant CHD-1 protein found in the nuclear matrix contained the two chromodomains and the helicase/ATPase domain.

The final part of this thesis examined proteins that interacted with the chromodomain of CHD-1 using a yeast two-hybrid screen with the chromodomain as a bait. Proteins with homologies to HUB1, KIAA1237, KIAA0164 and N-CoR were identified as proteins that interacted with the chromodomain of CHD-1. The interaction with N-CoR was of particular interest since N-CoR associates with histone deacetylase in the Sin complex and the CHD-4 protein also interacted with histone deacetylase in the NURD complex. CHD-1 was also shown to immunopreciptitate mSin3B indicating that it is part of a Sin complex.

Materials and Methods

Cell lines and culture conditions

P19 cell culture: P19 was a mouse male embryonal cell line (73). Cells were maintained in alpha medium with a 6.7% donor bovine serum and 3.3% fetal bovine serum. P19 cells were induced to differentiate with $3X10^{-7}$ M retinoic acid (RA) (53). Briefly, P19 cells were aggregated and cultured in suspension in bacterial-grade plastic Petri dishes (Falcon) for five days in the presence of RA before plating on tissue culture grade plastic dishes (Falcon). Cells were allowed to grow for another 48 h before harvesting. Differentiation was also induced with 1% (v/v) dimethyl sulfoxide (DMSO) (72). Cells were also aggregated and cultured in suspension in bacterial-grade plastic Petri dishes for five days in DMSO before plating on tissue grade tissue culture dishes. Cells differentiated with DMSO were also harvested after another 48 h.

293T cell culture: 293T is a human kidney epithelial cell line that was immortalized with the adenovirus T antigen. Cells were maintained in alpha medium with a 6.7% donor bovine serum and 3.3% fetal bovine serum.

Northern blot analysis. Total RNA and northern blots were prepared the same as in Chapter 1. Probe cDNA used in this chapter are the XhoI / Eco RI fragment of the CHD-1 cDNA (bp 31-769) and the fragment of the N-CoR cDNA from bp 339-1574 of the cDNA. The probe for N-CoR was the fragment contained in the cDNA library used for the yeast two hybrid screen.

SDS-PAGE and immunoblots. P19 or 293T cells were lysed in 2X SDS sample buffer (120mM Tris-Cl pH 6.8, 10% (w/v) SDS, 2% dithiothreitol (w/v), 40% gycerol, 0.002% bromophenol blue). Protein concentrations were determined using a protein assay (BioRad detergent compatible protein assay). 40 μg of protein lysate was loaded onto an SDS-polyacrylamide gel (SDS-PAGE) containing 10 % polyacrylamide gel (29:1 acrylamide:bisacrylamide) in 0.1% SDS and 0.375 M Tris-Cl pH8.8 (62). Samples were electrophoresed at 10 mAmp constant current overnight or 40mAmps constant current for 2-3h.

Proteins separated on SDS-PAGE were transferred to nitrocellulose membranes (Hybond C+, Amersham) by electrophoresis in a buffer containing 25mM Tris-Cl pH8.3, 192 mM glycine and 20% (v/v) methanol (114) at 20 V over night. The membranes were blocked for 30 minutes in 2% skim milk powder in TBST (10mM Tris-Cl pH 7.4, 150mM NaCl, 0.05% (v/v)Tween) before probing with the primary antibody.

The primary antibodies used in this section were as follows (dilutions are indicated in brackets): rabbit anti-CHD-1 polyclonal (1/2000) (109), rabbit anti-topoisomerase II polyclonal (1/500) (Topogen), mouse anti-myc 9E10 monoclonal (1/200 of the culture supernatant), anti-his tag monoclonal (1/1000) (Qiagen) and the anti-mSin3B polyclonal (1/500) (Santa Cruz). Primary antibodies were diluted in 2% skim milk powder in TBST. Before adding to blots. The blot was washed 3 times for 5 min in TBST before application of the goat anti-rabbit or anti-mouse secondary antibody conjugated to horseradish peroxidase (1/5000). A chemiluminescent substrate, lumiglo (Kirkegaard and

Perry Laboratories) was applied to the blot and immediately afterward the blot was exposed to X-ray film for visualiztion.

CHD-1 deletion mutants. The CHD-1 cDNA, pCHD-1, obtained from Robert Perry (36) was cloned into the bluescript vector (Stratagene). The base pair (bp) numbering for the CHD-1 cDNA started at the adenosine of the initiator ATG. A plasmid with the PgK-1 promoter (75) driving the expression of a myc-tagged CHD-1 fusion protein was constructed for use in transfections of mammalian cells. The CHD-1 cDNA in pCHD-1 was digested with Sal I at a unique site near the 5' end of the CHD-1 cDNA (bp 241) and Xba I at site in the multiple cloning site after the termination codon of cDNA (bp 5136). The excised CHD-1 fragment was ligated in-frame with the myc epitiope tag in an expression vector with the Pgk-1 promoter and polyadenylation signal to create the plasmid pHT8.

Digestion of pCHD-1 with restriction enzymes truncated cDNA's that were cloned in frame with the myc tag in the expression vector. The plasmids containing truncated CHD-1 cDNA's were as follows: pHT1 contained CHD-1 582-1203 (Bgl II to Pvu II); pHT14 contained CHD-1 241-2321 (Sal I to Eco RI) ; pHT16 contained CHD-1 241-3193 (Sal I to Eco RV); pHT15 contained CHD-1 bp 241-1596 (Sal I to Spe I) and pHT12 contained CHD-1 bp 1507-5136 (Spe I to the termination codon). All CHD-1 sequences were myc epitope tagged on the 5' end.

Transfections. All transfections were done using Fugene (Boehringer mannheim) according to manufacturer's instructions. P19 were plated at a density of $1X10^5$ cells/35mm tissue culture dish and $5X10^5$ cells/60 mm tissue culture dish. 293T cells were plated at $5X10^4$ cells/35 mm tissue culture dish and $3X10^5$ cells/60mm tissue culture dish. For 35 mm dishes 3μg of plasmid DNA was mixed with 6 μl Fugene and for 60 mm dishes 9 μg of plasmid DNA was mixed with 18 μl of Fugene. The Fugene/DNA mixture was added to cells and incubated for 48 h.

Immunofluorescent staining. $1X10^5$ cells were plated on gelatin-coated coverslips in 35 mm dishes. The next day cells were fixed in 3% paraformaldehde for 15 min. Fixed cells were incubated with a primary antibody diluted in 0.3% Triton X-100 in PBS (phosphate buffered saline) for 1 h at room temperature. The primary antibodies used for immunofluorescence were anti-myc 9E10 (1/5), anti-his tag and anti- topoisomerase II. Cells were washed 3x in PBS before the secondary antibody was applied. The secondary antibody against rabbit was conjugated to the fluorescent marker FITC (green) and the secondary antibody against mouse was conjugated to the cy3 (red) fluorescent markers. The DNA in the nuclei were stained with Hoechst. The cells were observed under the microscope and the immunofluorescent image was digitized using the Matrox Intellicam system.

Nuclear matrix extraction. Two methods for extracting the nuclear matrix were used. The first was described by Cockerill and Garrard (29). Briefly, cells were homogenized in RSB (10 mM NaCl, 3 mM $MgCl_2$, 10 mM Tris-Cl, 0.5 mM PMSF, pH 7.4) with a Dounce homogenizer and nuclei were purified by centrifugation through a cushion of RSB-2 M sucrose at 24,000 rpm for 30 min in an SW27 rotor. The purified nuclei pellet was digested with DNase I then extracted with 2M NaCl, 10 mM EDTA, 10 mM Tris-Cl, and 0.25 mg/ml BSA, pH 7.4 and centrifuged at 4500 x g for 15 min at 4°C. The supernant was the 2 M NaCl fraction and the pellet was the insoluble nuclear matrix fraction. The protease inhibitors 200 μg/ml PMSF, 2 μg/ml aprotinin, 5 μg/ml leupeptin were added to all buffers.

In the other method (25) for nuclear matrix extraction, cells were left attached to the tissue culture dish or gelatin-coated coverslips and extraction buffers were added directly into the dish. Briefly cells were permeated with 0.5% Triton X-100 in STM pH 6.2 (0.25 M sucrose, 10 mM Tris-Cl, 2 mM $MgCl_2.6H_2O$, 25 mM KCl) for 25 min, washed in STM pH7.4 then digested with 10 μg/ml DNase I in STM pH 7.4 for 30 min all at 4°C. Cells were then extracted on ice with LMB (0.05 mM $MgCl_2.6H_2O$, 10 mM Tris-Cl pH 7.4), then 2 M NaCl in LMB, followed by 4 M NaCl in LMB. The cells were then digested with 50μg/ml DNase in LMB on ice and again with 50μg/ml Dnase I and 50μg/ml RNase A in LMB). Finally cells were washed in STM pH 7.4 and the residual protein structure that was left on the culture dish or coverslip was the nuclear matrix.

Statistical analysis. The binomial test was done to compare counted data. Counted data were transformed to proportions (p) by dividing the positives counted (number of immunofluorescent cells) by the total number counted (n) (number of immunofluorescent cells + the number of non-immunofluorescent cells). The hypothesis tested in this binomial test was H_o: μ_1-μ_2=0, that is, there are no differences between p_1 and p_2, the two proportions compared.

The binomial test was done to test if $p_1 - p_2 = 0$. The standard error of the mean (s) for each proportion is calculated as follows:

$$s = pq/n ,$$

p is proportion immunofluorescent, q is 1-p, n is total number counted.

To find if $p_1 - p_2$ differed significantly from 0, a value for t was calculated as (See Table II.1). The calculated t values were compared to the t value in the t-distribution at the 95% confidence limit which is 1.6449 (from the table of hypothetical t values). If the t value calculated is lower that the1.6449 then the hypothesis tested, $p_1 - p_2$=0, is true. If the t value calculated is greater than 1.6449 the hypothesis is false and there are significant differences between p_1 and p_2.

Yeast strains and culture. The *Saccharomyces cerevisiae* yeast strain Y190 was used. It has the genotype MATa gal4 his3 trp1-901 ade2-101 ura3-52 leu2-3,-112 URA3::GAL-lacZ, LYS2::GAL(UAS)-HIS3 cyh.r. This genotype indicated that the yeast carried a genomic insertion of the GAL4 promoter driving expression of the ß-galactosidase gene and the gene that permitted cells

to grow in media lacking histidine, HIS3. For routine growth YPD was used (1% yeast extract, 2% peptone, 2% glucose). Synthetic dropout media (SD = 16.7% yeast nitrogen base, 0.5% $(NH_4)_2SO_4$, 2% dextrose, 0.058% amino acid mix) was made up with different amino acid mixes The basic amino acid mix contained 10% adenine SO_4, 5.5% arginine, 5.5% aspartic acid, 5.5% histidine, 5.5% isoleucine, 5.5% leucine, 5.5% lysine, 5.5% methionine, 8.5% phenylalalnine, 5.5% tryptophan, 5.5% tyrosine, 5.5% uracil, 26.5% valine. SD media made up with amino acid mixes lacking leucine, tryptophan, or histidine were used to select for yeast expressing nutritional marker genes LEU, TRP and HIS, respectively.

Yeast two hybrid screen. The CHD-1 fragment from pHT1 (b.p. 583-1203) was used as a target to screen for interacting proteins in a yeast two hybrid screen. The CHD-1 fragment from the pHT1 was excised with Sty I and BamHI and ligated in-frame with the GAL4 DNA binding domain in the NcoI/ Bam HI site of pAS2 (contained TRP gene) (37) (45) to create the plasmid pHT17. pHT17 was cotransformed with a mouse embryo Matchmaker two-hybrid library (Clontech, Palo Alto, CA) into the Y190 strain of *Saccharomyces cerevisiae*. The library cDNA's were cloned into the pGAD10 plasmid that contained the LEU gene. The transformants were plated on SD media containing 2% agar in 150mm petri dishes. The SD media was lacking in leucine, tryptophan and histidine. 30 mM 3-amino-1,2,4-triazole (3-AT)was added to media to provide additional selection for yeast that carried the HIS gene.

The yeast colonies formed were further screened for ß-galactosidase expression using the x-gal agarose overlay assay. Low melting temperature agarose was dissolved in a potassium phosphate buffer (0.5M potassium phosphate pH7.0, 6% dimethyl formamide, 0.1% SDS) at 60-70°C. The solution was cooled to 40°C and 0.1 mg/ml of x-gal and 0.05% (v/v) ß-mercaptoethanol was added. The x-gal agarose mixture was added to cover the surface of the SD agar media where the yeast colonies had formed. After the x-gal-agarose solution solidified, plates were incubated at 25°C overnight in the dark. The blue colored product of x-gal formed in yeast colonies that expressed ß-galatosidase. These colonies were picked through the top of the x-gal-agarose layer and restreaked on 100 mm SD media agar plates lacking leucine,tryptophan and histidine.

The streaked yeast were re-tested for ß-galactosidase expression using another procedure with the x-gal substrate. This time some of the yeast was lifted off the plates using nitrocellulose filters. The yeast were lysed by immersion of filters in liquid N_2 briefly. The filter with the yeast side up was placed on a whatman filter saturated with a solution of 1mg/ml x-gal in Z buffer (0.1 M sodium phosphate pH 7.0, 10 mM KCl, 1 mM $MgSO_4$). The yeast were monitored for the formation of the blue-colored product. The time taken for the blue colored product to form was proportional to the amount of ß-galactosidase present.

Recovery of plasmids from yeast. The QIAprep Spin miniprep kit (Qiagen) was used to isolate plasmids from yeast with the following modifications of the manufacturer's instructions: yeast grown in SD liquid media lacking leucine, tryptophan, histidine and containing 30mM 3-AT were pelleted and resuspended in yeast lysis solution (2% Triton X-100, 1% SDS, 100mM, 10 mM Tris-Cl pH 8.0, 1 mM EDTA) with 10μg/ml of RNase A instead of P1 (50 mM Tris-Cl pH 8.0, 10 mM EDTA, 100 μg/ml Rnase A) .

The low levels of plasmid isolated from yeast were transformed into bacteria using electroporation at a voltage of 2.5kV and a capacitance of 25μF. The transformed bacteria were grown on LB (1% Bacto tryptone, 0.5% yeast extract, 0.5% NaCl) agar plates containing ampicillin (100μg/ml). Colonies were lifted to nitrocellulose filters and colony hybridization with a radioactively labelled probe containing DNA from the activation domain of pGAD10 was done to identify plasmids derived from the cDNA library. Hybridizing colonies were picked and grown up for minipreps using the QIAprep Spin miniprep kit according to manufacturer's instructions.

Sequencing. The purified pGAD plasmids containing cDNA's from the mouse embryo library were sequenced by dye terminator cycle sequencing with an ABI, Inc., automated sequencer. The cycle sequencing was primed with an oligonucleotide with sequence from the activation domain of pGAD10 (TACCACTACAATGGATG).

The BLAST algorithm (3) was used to search for sequence homology in

Genbank (10). Sequence alignments of homologous sequences were also done using the DNASTAR software (DNASTAR).

Polymerase Chain reaction (PCR) cloning of N-CoRβ 5' sequence. The N-CoRβ cDNA present in the clone from the T-cell yeast two hybrid library (6.12) was incomplete (Fig.11.13). To clone more of the N-CoRβ cDNA reverse transcription PCR was used. The region targeted for amplification was the 5' end from the initiator ATG to bp 339 from the ATG of the N-CoRβ cDNA. The cDNA generated by reverse transcription was primed with the oligonucleotide sequence, 5'CGATCGACACGATCC (N-CoRβ 569A). This oilgonucleotide was the antisense sequence from a region 569 bp from the initiator ATG of the N-CoRβ cDNA. PCR amplification of the cDNA was primed with the oligonucleotides N-CoRβ ATG (5'ATGTCAAGTTCAGGTTATCC) on the 5' end and N-CoR 552A (5'GCTCTGTATCAGCTCTT) on the 3' end. The resulting PCR product was a 534 bp sequence from the 5' end of N-CoRβ that starting from at the initiator ATG.

Tagging N-CoR 5' with the his tag. The 552 bp N-CoRβ 5' sequence generated was tagged with the his tag which was a sequence containing six histidines for binding to Ni-NTA (Ni^{2+}-nitrilotriacetic acid) cross-linked to agarose beads (66) as well as the Met-Arg-Gly-Ser for recognition with a monoclonal anti-his tag antibody. PCR was used to make an in-frame fusion of the his tag with the N-CoRβ 5' sequence. The oligonucleotide used as the 5' primer for

PCR was

5'CCGCGTCCATACC**ATGAGAGGATCGCATCACCATCACCATCAC**ATGTCA

AGTTCAGGTTATCCTCCCAACCAGGGGGCG. The sequence in bold encodes the peptide Met-Arg-Gly-Ser-His-His-His-His-His-His which is the his tag. The underlined sequence anneals to the N-CoRβ 5' sequence. The 3' primer for PCR is the N-CoRβ 552A oligonucleotide described above. The PCR product of this reaction was a his-tagged N-CoRβ 5' sequence.

Construction of the his-tagged N-CoR expression plasmid. To create the his-tagged N-CoRβ expression plasmid, the first step was to excise the N-CoRβ partial cDNA from the T-cell yeast two hybrid library (6.12) using Xho I, and to ligate it into an expression vector containing the Pgk-1 promoter and polyadenylation signal (75) to make an intermediary plasmid. The intermediary plasmid was digested with Sma I, which was 5' of the N-CoRβ insert, and Pst I within the N-CoRβ sequence. The next step was to blunt the 5' end of the his-tagged N-CoR PCR product described in the previous section and digest the 3' end with Pst I. The his-tagged N-CoR PCR product could now be ligated with the N-CoRβ DNA in the intermediary plasmid to create a continuous 1567 bp sequence of N-CoRβ that his-tagged at the initiator Met which we termed his-tagged N-CoRß.

***In vitro* protein binding assay.** 293T cells grown on 60 mm tissue culture dishes were transfected with the his-tagged N-CoRß plasmid described in the

section above either alone, his-tagged N-CoRß + pHT1 or pHT1 alone. The cells were lysed in 100 μl his-tag lysis buffer (HLB) (50mM Tis-Cl pH8.0, 150 mM NaCl, 1% NP-40) with protease inhibitors (200 μg/ml PMSF, 2 μg/ml aprotinin, 5 μg/ml leupeptin). Protein concentrations were quantified using a protein assay (BioRad) and 100 μg of protein from the lysate were added to 50 μl of a 50% slurry of Ni-NTA cross linked agarose beads (Ni beads). The Ni beads and protein lysate were incubated overnight at 4°C with rotation. The next day the Ni beads were pelleted with a brief spin and were washed four times with HLB containing proteases and 10 mM imidazole. Proteins were eluted in 100μl of his tag elution buffer (HEB) (50 mM sodium phosphate buffer pH 6.0, 100 mM KCl, 2 % glycerol, 0.2 % NP-40, 250 mM imidazole) with rotation for 1 min at 23°C. Eluted proteins were run on SDS-PAGE, immunoblotted and probed with the anti-myc 9E10 antibody for analysis of protein interaction.

Co-immunopreciptiation. P19 cells grown on 100 mm tissue culture dishes were lysed with 300 μl of lysis buffer (10 mM Tris pH 7.5, 150 mM NaCl, 5mM EDTA, 1% Triton X-100) with protease inhibitors (200 μg/ml PMSF, 2 μg/ml aprotinin, 5 μg/ml leupeptin). Protein concentrations were quantified with a protein assay (BioRad) and 100 μg of protein from the lysate was added to the anti-CHD-1 anti--*-body (final dilution was 1/1000). The protein lysate was incubated with the anti-CHD-1 antibody overnight with rotation at 4°C. The next day 30 μl of a 50% slurry of IgG beads (gammabind, Pharmacia) equilibrated in

lysis buffer were added to the protein lysate antibody mixture and incubated at 4°C with rotation for 30 min. The beads were pelleted with a brief centrifugation and were washed four times with lysis buffer containing protease inhibitors. Proteins were eluted from beads using 2X SDS sample buffer with incubation at 100°C for 5 min. Eluted protein was run on SDS-PAGE, immunoblotted and probed with the anti-mSin3B antibody for analysis of binding between CHD-1 and mSin3B.

Results

A) CHD-1 expression

1) Expression of the endogenous CHD-1 protein in P19 cells

CHD-1 was previously found to be expressed in a variety of cell lines (36). The Northern blot in Figure II.1 using CHD-1 as a probe shows that CHD-1 was also expressed in P19 cells as well as in P19 cells differentiated with DMSO or with Retinoic Acid (RA). Stokes and Perry observed a number of different transcripts that could be detected with different parts of the CHD-1 cDNA as probes (109). Figure II.1 is a northern blot of RNA from P19 cells and P19 cells differentiated with RA and DMSO. The blot was probed with the CHD-1 cDNA and only one band at about 5 kb that corresponds to the full length CHD-1 transcript is detected. It is possible that the smaller alternatively spliced transcripts are degraded or expressed at low levels in P19 cells. It was observed that the level of CHD-1 transcripts decreased with differentiation.

David Stokes and Robert Perry generously provided the anti-CHD-1 rabbit polyclonal antibody used in this study. The anti-CHD-1 antibody was generated using the DNA binding domain as the immunogen for this antibody. The immunoblot using this anti-CHD-1 antibody shows that the 201kD CHD-1 protein was present in P19 cells and differentiated P19 cells (Fig. II.2). The decrease in expression of CHD-1 with differentiation that was noted on the northern blot (Fig II.1) is not apparent on the immunoblot. The CHD-1 antibody was also found to detect smaller protein bands. Only the full-length transcript of CHD-1 was detected in P19 cells so these smaller bands are most likely due

to degradation of the 201 kD protein. Degradation of the CHD-1 protein was also been observed by others in plasmacytoma cells (109).

2) Deletion mutants of CHD-1 are all nuclear

The CHD-1 protein is part of a family of proteins that shares the following domain structure: two **C**hromodomains, **H**elicase/ATPase domain, **D**NA binding domain. The domains are defined through sequence homologies and the actual function of these domains is as yet undetermined. To further explore the function of the signature domains of CHD-1, CHD-1 mutants were engineered with deletions of increasingly larger regions from the carboxy-terminus to delete progressively the DNA binding domain and helicase/ATPase domain of CHD-1 (Fig. II.3). The CHD-1 constructs were myc-tagged so that they could be detected with the anti-myc 9E10 monoclonal antibody, 9E10. In order to myc tag CHD-1, 241 bp of the 5' end was removed and replaced with the myc_6 tag. The largest myc-tagged clone, pHT8 (CHD-1 bp 241-5136), contained the full length cDNA of CHD-1 minus the 241 bp on the N-terminus. The rest of the constructs carry deletions that removed the DNA binding domain (CHD-1 bp241-2312, pHT14 and CHD-1 bp 241-3193, pHT16), and the helicase/ATPase domain (CHD-1 bp 241-1596, pHT15). The mutant CHD-1 in pHT1 (CHD-1 bp 583-1203) is the smallest CHD-1 fragment containing the first chromodomain but not the second. PHT12 (CHD-1 bp 1597-5136) is unlike the other mutant proteins in that the deletion is from the amino-terminus and removes the two chromodomains, and the helicase/ATPase domain.

The myc-tagged CHD-1 deletion mutants were cloned into expression vectors and transfected into 293T cells. 48 hours later, the protein from transfected cell lysates was analyzed on a western blot. All cell lysates contained a myc-tagged protein band of the predicted size (Figure II.4). The lane showing expression of the largest myc-tagged protein (pHT8) contains protein bands at smaller sizes in addition to the band at the predicted size of 200 kD. This is likely due to degradation of the protein as was found for the endogenous protein. Myc-tagged puromycin N-acetyltransferase (Fig. II.8g) was used as a positive control for the anti-myc 9E10 antibody.

Cellular localization of the mutant proteins was also examined in the transfected cells using the anti-myc 9E10 antibody for immunofluorescence. A control transfection was done with the myc-tagged puromycin N-acetyltransferase gene (myc-tagged puromycin). Puromycin N-acetyltransferase is a cytoplasmic protein (61) and Figure II.5g shows that myc-tagged puromycin remains cytoplasmic and that the myc tag on its own does not confer nuclear localization.

Immunofluorescence of the endogenous CHD-1 protein shows that the protein has punctate nuclear staining that is excluded from the nucleolus (109). Immunofluorescence of cells transfected with the myc-tagged CHD-1 mutants (Fig. II5a-f) also showed a nuclear localization that was excluded from the nucleolus. However, the immunofluorescence was not observed to be as punctate in appearance as the endogenous CHD-1 (109). All the deletion constructs had similar nuclear distributions (Fig. II.5).

B) Nuclear Matrix localization of CHD-1

1) Extraction of the CHD-1 protein in the nuclear matrix of P19 cells

The nuclear matrix is defined as a salt, nuclease and detergent resistant protein structure of the nucleus (41). DNA sequences that bind to the nuclear matrix have been characterized as being AT-rich DNA and are referred to as MARs or SARs. Bacterially synthesized CHD-1 was found to bind to such AT-rich sequences in a mobility shift assay (109). It was also found, however, that CHD-1 was soluble in the 2 M NaCl high salt extract unlike nuclear matrix protein which are insoluble in 2 M NaCl. Those experiments had been done on S194 mouse plasmacytoma cells and studies have shown that the composition of the nuclear matrix changes with cell type (41) (51). Therefore, CHD-1 could still be a nuclear matrix protein in another cell line.

The same nuclear matrix extraction (29) was done in this study with P19 cells to determine if CHD-1 is located in the nuclear matrix. The cell extracts from the cytoplasmic fraction, the soluble 2 M NaCl fraction and final insoluble nuclear matrix fraction were electrophoresed on a 6%SDS PAGE and immunoblotted (Fig. II.6). The blot was probed with the anti-CHD-1 antibody, which was described above, and the 201kD band (Fig. II.2) was detected. The blot was also reprobed with anti- topoisomerase IIa (170kD, anti-topo IIa) which is known to be a constituent of the nuclear matrix (29). Unlike S194 plasmacytoma cells, the endogenous CHD-1 of P19 cells, like topoisomerase IIa, was found in the insoluble nuclear matrix fraction and was not soluble in 2 M NaCl fraction (Fig.II.6). Anti-tubulin was also used to probe this blot. Tubulin

is a cytoplasmic protein and was not found in the nuclear matrix as shown in Figure II.6.

2) Analysis of nuclear matrix localization of deletion mutants of CHD-1

Since all the myc-tagged CHD-1 deletion mutant proteins were found localized in the nucleus, it was of interest to see if they, too, were located in the nuclear matrix. The previous method was initially described for cells growing in suspension. Another nuclear matrix extraction procedure could be done on cells that are attached to culture dishes like P19 cells (25). Like other nuclear matrix isolation procedures, extractions are done to remove non-nuclear matrix proteins with high salt, detergent and nucleases except that extractions are done on attached cells without removing them from the culture dish, and the procedure does not involve mechanical homogenization or centrifugation. At the end of the extractions the insoluble nuclear matrix is all that is left on the coverslip.

To see if myc-tagged CHD-1 mutant proteins are located in the nuclear matrix, cells were grown attached to gelatin-coated coverslips in duplicate. Transfection mixes containing each plasmid (pHT1, pHT8, pHT12, pHT15, pHTT14, and pHT16) were made and were split to transfect each duplicate coversplip. The nuclear matrix was extracted using the method of Chaly et al. (25) isolated nuclear matrices on coverslips were stained for immunofluorescene with the anti-myc 9E10 antibody to detect the presence of myc-tagged CHD-1 mutant proteins in the nuclear matrix. The advantage of

using immunfluorescence to detect the presence of myc-tagged proteins in the nuclear matrix is that detection was possible even if the transfection efficiency was poor and only a small number of cells are transfected. Some of the plasmids containing myc-tagged CHD-1 mutant proteins were expressed poorly in transfected P19 cells and myc-tagged protein levels were too low to allow for the detection on immunoblots. However, detection of the myc-tagged protein was possible with immunfluorescence.

Figure II.7 shows that the nuclear matrices isolated on coverslips are stained with anti-topo IIa, which confirms that nuclear matrix proteins are left on coverslips. The Hoechst stain was used to detect the presence of residual DNA in the nuclear matrix. There should be little DNA because of nuclease digestion steps in the nuclear matrix extraction procedure. Figure II.7b shows that the Hoechst stain is negative for nuclear matrices, which indicated that the DNA was extracted in a soluble non-nuclear matrix. To confirm that non-nuclear matrix proteins are not present in the insoluble nuclear matrix fraction, cells were transfected with the negative control protein, myc-puromycin. Around 20 % of intact cells were stained with anti-myc 9E10 but no anti-myc 9E10 staining was detected in the nuclear matrices prepared from transfected cultures (data not shown).

P19 cells were grown on coverslips in duplicate and transfected with the same transfection mix containing plasmids carrying genes encoding the myc-tagged CHD-1 deletion mutants. These cells were extracted for the nuclear matrix using the method described above (25). The residual nuclear matrix on

coverslips was probed with anti-myc 9E10 followed by a secondary anti-mouse cy3 for immunofluorescence. Another set of P19 cells on coverslips were also transfected with the same plasmids at the same time. These cells were not extracted and intact cells were stained for immunofluorescence with the anti-myc 9E10 antibody. The number cells or nuclear matrices staining with anti-myc were counted under the microscope (Appendix I). The proportion (p_1) of cells or the proportion (p_2) of nuclear matrices stained with the anti-myc 9E10 antibody within the total population were calculated for each of the recombinant myc-tagged CHD-1 proteins that were transfected (Table II.1).

The method used to analyze the nuclear matrix association of CHD mutant proteins is summarized in Figure II.9a. If myc-tagged CHD-1 mutant proteins are associated with the nuclear matrix, the proportion of cells (p_1) and nuclear matrices (p_2) staining with the anti-myc 9E10 antibody should be equal ($p_1 = p_2$). On the other hand if the myc-tagged CHD-1 mutant protein is not associated with the nuclear matrix the protein will be lost in soluble fractions and the proportion of cells (p_1) staining for immunofluorescence will be greater than the proportion of nuclear matrices (p_2) staining for immunofluorescence ($p_1 \neq p_2$). Hence, for nuclear matrix proteins $p_1 = p_2$ or $p_1 - p_2 = 0$. A statistical test was done to determine if the transfected myc-tagged CHD-1 mutant proteins are significantly associated with the nuclear matrix by testing if $p_1 - p_2$ values are significantly different from 0.

Table II.1 summarizes the statistical analysis of the immunofluorescence data. The binomial test was done to test if $p_1 - p_2 = 0$. A t value for this

difference was calculated as indicated in Table II.1 and was compared to the t value at the 95% confidence limit (1.6449). The t values lower than 1.6449 indicate that no significant difference is found between p_1 and p_2; that is the proportion of cells and nuclear matrices staining with the anti-myc 9E10 are the same which is the resulted expected for nuclear matrix proteins. On the other hand, if t values are greater than 1.6449 then significant differences were found between $p_1$and p_2 and the protein is classified as a non-nuclear matrix (Table II.1).

Cells transfected with pHT8 express the largest myc-tagged protein (CHD-1 aa 81-1711) (Fig. II.8a). For pHT8 transfected cells the proportion of immunofluorescent cells (p_1) that were stained with anti-myc 9E10 was 0.043 compared to the proportion of immunofluorescent nuclear matrices (p_2) which was 0.046. The value for p_1- $p_2$was -0.0030. The t-value calculated at 0.46 is lower than the t value at the 95% confidence limit, 1.6449; therefore, the hypothesis, p_1- p_2 = 0, is accepted as indicated in Table II.1and Figure II.9b. This is interpreted as showing that the myc-tagged CHD-1 aa 81-1711 (pHT8) is associated with the nuclear matrix (Fig. II.8a). For transfection with pHT14 and pHT16 p_1- p_2 values were also not found to be significantly different from 0 (Table II.1, Fig. II9b) with calculated t values of 0.103 and 0.755 for pHT14 and pHT16, respectively. Hence, the myc-tagged CHD-1 aa 81-769 (pHT14) (Fig. II.8b) and myc-tagged CHD-1 aa 81-769 (pHT16) (Figure II.8c) are also associated with the nuclear matrix.

P19 cells transfected with pHT1, on the other hand, show a significant

difference between p_1 and p_2 values with a calculated t-value of 12.38 that is higher than 1.6449. This is evidence that the myc-tagged CHD-1 protein encoded in pHT1 (aa 195-401) is not a nuclear matrix associated protein. The t values calculated with data from pHT12 and pHT15 transfected cells were 2.598 and 2.759 respectively, and are also higher than t 1.6449 indicating that the myc-tagged CHD-1 aa 81-532 and aa 533-1711 are also not associated with the nuclear matrix.

Figure II.9b is a graphical illustration of the results of the binomial test described above. A histogram of $p_1 - p_2$ values for each transfection are presented. From this graph it can be seen that for cells transfected with pHT8, pHT14, and pHT16, the $p_1 - p_2$ values are not significantly different from 0. Whereas with cells transfected with pHT1 the difference of $p_1 - p_2$ is much larger than 0. The $p_1 - p_2$ values for cells transfected with pHT15 and 12 are also significantly higher than 0.

From the analysis of the nuclear matrices extracted from cells carrying CHD-1 mutant proteins, the smallest region of CHD-1 that is associated with the nuclear matrix is aa 81-769. This region contains sequences that encoded the two chromodomains and the helicase/ATPase domain. However, the chromodomain was not found to be sufficient on its own to target the protein to the nuclear matrix (data from pHT1 transfected cells). The helicase ATPase domain was also found to be insufficient for targeting the protein to the nuclear matrix since the CHD-1 proteins that encoded in pHT12 (aa 533-1711) and pHT15 (aa 81-532) contain the helicase/ATPase domain but are not associated

with the nuclear matrix. The putative nuclear matrix targeting signal is discussed further in the Discussion.

C) Proteins interacting with CHD-1

1) The yeast two-hybrid screen

Though CHD-1 was found to be a nuclear matrix protein in this study, the function of this protein in the nuclear matrix is still unclear. The biological function of CHD-1 could be elucidated through identifying proteins that interacted with it. The implication is that CHD-1 participates in the biological functions of the proteins it interacts with. For this reason a search for proteins that interacted with CHD-1 was done using the yeast two-hybrid system (Fig.10). The advantage of the yeast two-hybrid system is that it screens for *in vivo* interactions between proteins. The yeast two-hybrid screen selects cDNA's that encode proteins ("prey") that bind to a "bait" protein. The cDNA expression libraries are constructed so that "prey" protein are fused to the GAL4 activation domain (Clontech). The "bait" is fused with the GAL4 DNA binding domain. The interaction between the "bait" protein and the "prey" protein recreates the GAL4 transcription factor at a genomic GAL4 promoter - the DNA binding domain fused to the "bait" binds to the GAL4 promoter and the "bait" binds to the "prey" which is fused to the acidic activation domain promoter. This arrangement results in the activation of transcription at the genomic GAL4 promoter. The GAL 4 promoter was engineered so that the activation of transcription results in the expression of the ß-galactosidase gene and the His

nutritional selection gene. His$^-$ gene expression is selected for in media lacking histidine. The interactions could be confirmed by an assay for ß-galactosidase enzyme activity in yeast cell lysates.

2) The chromodomain of CHD-1 was used as the bait for the screen

The bait used for this screen was aa195-401 fragment of CHD-1 (pHT1) containing the chromodomain of CHD-1. Figure II.10 is a schematic drawing of the yeast two-hybrid screening strategy. To create the chromodomain bait, the chromodomain cDNA from CHD-1 (bp 583-1203) was fused in-frame with the GAL4 DNA binding domain in the vector pAS. This plasmid is named pHT17. The pAS2 vector contains the Trp$^-$ nutritional selection gene, TRP, so that yeast transformed with pas can be selected for in media lacking tryptophan. The cDNA library used for yeast two hybrid screening was constructed from mouse 17-day embryo RNA (Clontech). The library contains $3X10^6$ independent cDNA clones. It was created from oligo dT and random primed cDNA's which were cloned into the pGAD10 vector to create fusion proteins with the GAL4 activation domain. The pGAD10 vector also contains the nutritional selection gene, LEU, which allows cells transformed with pGAD 10 to grow in media lacking leucine.

Before the screen was carried out control experiments were done to rule out false positive ß-galactosidase expression induced by the pHT17 (chromodomain bait) alone. First, the yeast were transformed with pHT17 alone and plated on SD media lacking tryptophan. The colonies formed on these plates were tested

for ß-galactosidase expression using the x-gal substrate. Formation of a blue product in this assay indicates that transcription of ß-galactosidase was activated. The results of this experiment confirm that the chromodomain bait is not able to induce ß-galactosidase expression by itself (Table II.2). The same experiment was done with yeast transformed with pHT17 and the empty pGAD10 plasmid. This control is also negative for ß-galactosidase expression. The results of the controls confirm that the chromodomain bait could be used for the yeast two-hybrid screen.

Plasmid	ß-Galactosidase expression
pAS	negative
pGAD10	negative
pHT17	negative
PHT17 + pGAD10	negative

Table II.3. The chromodomain "bait" does not stimulate the GAL4 promoter on its own. The empty plasmid vectors, pAS and pGAD10 and the chromodomain bait plasmid, pHT17, were transformed into yeast and the activation of the GAL 4 promoter in transformed yeast was tested by assaying ß-galactosidase expression using x-gal as the substrate. The lack of blue product formation meant that expression was not stimulated and this is indicated in the table as "negative".

3) Isolation of clones that encoded protein interacting with the chromodomain

For our yeast two-hybrid screen, $1X10^{10}$ cells were transformed with 100 mg of pHT17 and 100mg of pGAD10 library DNA. The transformants were plated on agar plates containing SD media that was lacking leucine, tryptophan and histidine. 30mM aminotriazole was used to further suppress growth of yeast that did not express the His nutritional selection gene. $2.2X10^{5}$ colonies formed on the plates indicating a plating efficiency of about $2.2X10^{3}$ transformants/mg of library DNA. Of these colonies 103 expressed ß-galactosidase in an assay using x-gal as the substrate. These x-gal positive clones were picked and restreaked on agar plates containing SD media that lacked leucine, tryptophan and histidine. 65 colonies were positive on the secondary ß-galactosidase screen and were picked and grown in liquid SD media lacking leucine, tryptophan and lysine for plasmid isolation.

The plasmids recovered from the 65 yeast clones were a mixture of pHT17 and pGAD10 plasmids containing cDNA's from the yeast two-hybrid library. The plasmids have to be separated in order to identify the cDNA contained in pGAD10 . To separate mixtures of plasmids isolated from yeast, each plasmid preparation from yeast was electroporated into *E. coli* and selected in ampicillin for bacteria containing plasmid. An *E. coli.* bacterium is typically only transformed with one plasmid so this step served to separate pHT17 from pGAD10 plasmids containing cDNA clones form the yeast two-hybrid library. In addition, plasmids isolated from yeast were low in copy number so transformation of plasmids into *E. coli* also amplified plasmid quantities. All 65 yeast plasmid preparations transformed in *E. coli* resulted in colony formation on LB ampicillin plates. To identify which *E. coli* colonies contain pGAD10 plasmids with cDNA's from the yeast two-hybrid library, the transformants were subjected to colony hyrbridization with a probe containing the activation domain contained in the pGAD10 vector. All 65 *E. coli* transformations produced colonies that hybridized to the activation domain probe. Colonies that hybridized to the activation domains were picked and grown for plasmid isolation. Isolated plasmids were digested with Eco RI to release the cDNA insert from the multiple cloning site of pGAD10. 17 of the 65 plasmids isolated contained inserts that were over 100 bp. The insert-containing plasmids were sequenced using the ABI sequencing machine with a primer from the pGAD10 activation domain 5' of the cDNA insert. Successful sequences were obtained for 15 of the plasmids,

and most sequences were over 500 bp in length.

4) Identification of proteins interacting with the chromodomain

The sequences obtained from the ABI machine were used to search for sequence homologies in Genbank (10) using BLAST (3). The results of the BLAST search of Genbank are summarized in Table II.3. The plasmids are labelled as indicated in the table in the column entitled "Clone #". The "size insert" column lists the sizes of the cDNA inserts estimated from agarose gel electrophoresis. The "bases sequenced" column lists the lengths of the sequence obtained from the ABI sequencing machine for the cDNA inserts. The "Genbank" columns indicate the identity and size of the cDNA in Genbank that had homology to the sequence of the cDNA insert. The column entitled "Homology", lists the regions of the cDNA insert that were found to have homology to the Genbank cDNA sequence and the percentage of nucleotides in the pGAD10 cDNA insert that were identical to the nucleotides in the corresponding region of the cDNA in Genbank. There are a number of cDNA clones from the yeast two-hybrid library that do not have any homology with sequences in Genbank and they are clones E5, A41, A19, B47, E120, and E33.

The cDNA of the clones A113, F11, A25 and C47 were found to be homologous with 17-b-hydroxysteroid dehydrogenase, tRNA, IGF-II (exon 6) and PTD004 (direct submission), respectively. However, it was found that the homology was at the nucleotide level only. Proteins expressed by the cDNA

inserts in pGAD10 contain an amino-terminal activation domain and a carboxy terminus equivalent to the protein encoded by the cDNA insert. Therefore, the cDNA is translated in the same codon reading frame as the activation domain and for A113, F11, A25 and C47 this resulted in a protein that had an amino acid sequence that was different from the protein encoded by the Genbank cDNA. However, sequencing errors may have occurred, therefore, additional sequencing is being done to confirm the reading frame of the cDNA in A113, F11, A25 and C47.

The sequences of the cDNA of the clones A91, B18, B41, and D18 indicate that they are homologous to KIAA1237, glutathione-S-tranferase, HUB1 and KIAA0164, respectively. In addition, the cDNA of these clones also appear to encode the same amino acid sequence as the homologous cDNA found in Genbank. KIAA1237 is a human cDNA cloned as part of a screen for large proteins from the human brain (84). 83% of the nucleotides in A91 cDNA are identical with the human KIAA1237 cDNA from bp 523-1246. There is little information about KIAA1237 so the function of this region, or of the entire KIAA1237 protein, is unknown. Preliminary sequence data indicates that another cDNA present in A85, may also contain a cDNA with homology to KIAA1237. The incomplete sequence of the cDNA in A85 shows that it too has homology to KIAA1237 but further sequence information is needed to confirm this result (data not shown).

The nucleotide sequence for cDNA clone B80 was 96% identical with

glutathione S-transferase (24). However, glutathione S-transferase is ruled a false positive since glutathione-S-transferase is known to be a cytoplasmic protein and the chromodomain bait, pHT17, was shown to be a nuclear protein. Thus, it is unlikely that these proteins would interact in mouse cells.

The cDNA of B41 was found to be homologous to human HUB1 (90). The nucleotide sequence of the B41 cDNA is 84% identical to a region from bp 780-1285 of the HUB1 cDNA. HUB1 encodes a novel Kruppel-type zinc finger protein and it is also known that HUB1 represses expression mediated by the human T cell leukemia virus type I long terminating repeat. The B41 cDNA clone may be a mouse homolog of this protein.

One cDNA clone out of this screen, D18, had 87% cDNA nucleotide sequence identity with the human cDNA, KIAA0164. This is of particular interest since KIAA0164 was also found by others to interact with CLK1, a protein involved in splicing RNA, in another yeast two hybrid screen using a T-cell leukemia cell line library (30). Colwill et al. refer to their clone as 5.7 and the cDNA in the 5.7 and D18 clones may be mouse homologs of the human KIAA0164. The insert contained in D18 plasmid is smaller than that of the 5.7 plasmid. The cDNA sequences of D18 and 5.7 show that the D18 cDNA is identical to a corresponding region within the 5.7 cDNA (Figure II.11). Therefore, it is concluded that the D18 and 5.7 contained sequences from the same gene.

KIAA0164 was originally cloned from a large screen of DNA binding proteins from HepG2 cells (85). The exact function of KIAA0164 is as yet unknown but

it did have Ser and Arg repeats in the amino-terminus. Ser/Arg domains (SR) are associated with proteins involved in mRNA splicing such as CLK1 indicating that KIAA0164 may function in splicing. The cDNA of the 5.7 cDNA clone that interacts with CLK1 does not encode a corresponding Ser/Arg rich region; however, the 5.7 cDNA is missing sequences from the 5' end of the mouse KIAA0164 cDNA. The 5' end of the mouse KIAA0164 cDNA was cloned using Rapid Amplification of 5' cDNA ends (RACE) and was found to contain a sequence that encodes for the Ser/Arg rich region of the protein (Margit Geisterfer and John Bell, personal communication). The complete mouse and human KIAA0164 cDNA nucleotide sequence is presented in Figure II.11.

5) Proteins that interacted with both CLK1 and CHD-1

To compare the interactions of KIAA0164 with CLK or the chromodomain of CHD-1, a yeast two-hybrid assay was done as described Figure II.10. Yeast were transformed with plasmids containing either CLK1 or the chromodomain (pHT17) fused to the GAL4 DNA binding domain. Plasmids with GAL 4 DNA binding domain fusion protein were co-transformed with either 5.7 or D18 (which contained protein fused to the GAL 4 activation domain) and plated on to agar plates containing SD media lacking leucine, tryptophan and histidine to select for protein interactions. The strength of the protein-protein interaction in the co-transformed yeast is proportional to the amount of ß -galactosidase

expressed as measured using x-gal as the substrate. Large amounts of ß-galactosidase will result in rapid formation of the blue colored product; therefore, the time taken for blue colored x-gal product to form is a relative measure of the amount of ß-galactosidase expression. Table II.4 summarizes the results of the ß-galactosidase assay. The time taken for the formation of the blue product after x-addition is indicated with "++++" for 5 min , "+++" for 15 min, "++" for 30 min, "+" for 1h, "+/-" greater than 1 hour and "-" no blue product. CLK1 interacts strongly with the of 5.7 clone (++++) but there is no interaction with the clone of D18 (-). On the other hand, pHT17 interacted with both 5.7 (+) and D18 (+) which confirmed that CLK1 and CHD-1 bind to a common protein. Since CLK1 did not show an interaction with D18, it is concluded that CLK1 interacts with a region of the cDNA that is outside of the region corresponding to the D18 cDNA.

In addition to the 5.7 cDNA, a number of other cDNAs from the yeast two-hybrid T-cell library encode protein that interact with CLK1. Do these other proteins also interact with the chromodomain of CHD-1? To test this possibility pHT17 was co-transformed with plasmids from the yeast two-hybrid T-cell library that were isolated based on their interaction with CLK1. The co-transformed yeast were tested for expression of ß-galactosidase as described above and the data are presented in Table II.4. Interestingly, a number of the cDNA's tested encode proteins that interact with the chromodomain as well as CLK1. These cDNA's are identified as N-CoR (6.12), c-myc (2.8), and X16 (1.6). On the other

hand, cDNA's encoding ASF/SF2 and mammalian tra-2 proteins interact with CLK1 but do not interact with the chromodomain of CHD-1.

Since a number of proteins interact with both the chromodomain of CHD-1 and CLK1, it is possible that CHD-1 and CLK1 interact with each other. To test this, chromodomain of CHD-1 was cloned in frame with the GAL4 activation domain to turn it into a prey protein (Table II.4). The CLK1 protein fused with the GAL4 DNA binding domain was co-transformed with the chromodomain fused to the GAL 4 activation domain. ß -galactosidase assays were performed on co-transformed yeast to test for protein interactions. This assay shows that the chromodomain of CHD-1 does not interact with CLK1. However, the possibility that CLK1 interacts with another part of CHD-1 outside of the chromodomain cannot be ruled out.

The chromodomain fused to the GAL4 activation domain was co-transformed in yeast with pHT17 (the chromodomain fused to the Gal4 DNA binding domain. This was done to test for interaction of the chromodomain with another chromodomain. The ß-galactosidase assay shows that there is a chromodomain-chromodomain protein interaction. The interaction between the Gal4 DNA binding domain fused chromodomain and the GAL4 activation domain fused chromodomain is not a strong one as it took 1 h for the blue x-gal product to form. Chromodomain-chromodomain interactions were also found in a study done with *in vitro* translated protein (33).

D) The interaction of CHD-1 with N-CoR

1) Differences with the published N-CoR sequence

The 6.12 cDNA encoded a protein with homology to N-CoR that interacts with both CLK1 and CHD-1 (Table II.4). N-CoR is of particular interest since N-CoR is part of a repressor complex with histone deacetylase (See also Introduction). The hypothesis that CHD-1 is also associated with histone deacetylase is supported by the finding that another CHD protein , human CHD-4 also associates with histone deacetylase. Therefore, the cDNA of 6.12 was examined more closely.

N-CoR is a co-repressor of nuclear hormone receptor transcription (52). There are two regions of N-CoR that repress transcription in isolation from the rest of the protein. They are referred to as repressor domains I and II (52). The cDNA contained in the 6.12 plasmid corresponds to part of the repressor domain I of N-CoR. Repressor domain I stretches at the initiator methionine to aa 312 of the N-CoR protein. The protein encoded by the cDNA in 6.12 corresponded to aa 114 to 522 of N-CoR and is therefore missing the first 113 aa of repressor domain I. To complete repressor domain I of N-CoR, RT-PCR of brain RNA was used to clone the first 113 aa of repressor domain I missing from the clone in 6.12. The RT-PCR product encoding the first 113 aa of N-CoR was ligated with the cDNA contained in 6.12 at a natural Pst I site. This creates a cDNA that encodes an N-CoR protein that starts at the initiator Met and goes to aa 522 and includes all of repressor domain I (initiator Met to aa 312) (Fig. II.12). This

construct was referred to as N-CoRβ.

Initial sequencing of N-CoRβ indicate that there are several differences with the published sequence of N-CoR which are pointed out in Figure II.12. 7.9 % of the nucleotides of N-CoRß are different from the published N-CoR sequence. 76.4% of nucleotide substitutions occur at the third nucleotide of the codon. The amino acid sequence for N-CoRß is only 2.9 % different from that of N-CoR. The greatest number of nucleotide substitutions are in sequences encoding Repressor domain I of N-CoR (12.6% nucleotide substitutions versus 0.008% for the rest of the DNA). This indicates that N-CoR and N-CoRβ are highly related proteins that function similarly.

2) Cellular localization of the his-N-CoRß.

The N-CoRß protein was his tagged on the amino-terminus . The PCR product encoding his-N-CoRß was ligated into a mammalian expression vector and transfected into 293T cells grown on coverslips. These cells were then fixed and stained for immunofluorescence using the anti-his antibody followed by a secondary anti mouse conjugated to cy3. The results presented in Figure II.13 show that his-N-CoRß is a nuclear protein with a punctate pattern of staining. In addition, the nuclear staining is excluded from the nucleolus. The pattern of staining of the his-N-CoRß protein is similar to that for CHD-1 which supports the yeast two-hybrid assay result that CHD-1 and N-CoR are associated.

Figure II.14 is a northern blot of P19 and differentiated P19 total RNA. This blot was probed with the cDNA insert of the 6.12 clone (See Figure II.12). There is one large transcript running at about the size of the full length N-CoR transcript (c. 7.8 kb). There are no differences in expression of the N-CoR transcript over the time course of differentiation for either DMSO of RA differentiated P19 cells.

3) Confirmation of the interaction between N-CoRß and CHD-1

The yeast two-hybrid assay is prone to false positive interactions so an *in vitro* protein binding assay was done with mammalian cell lysates to confirm the interaction between N-CoRß and CHD-1. The cell lysates used were from 293T cells that were co-transfected with the his-N-CoRß expression plasmid and the pHT1 plasmid. The pHT1 plasmid was discussed previously (Fig. II.3); it contains a myc-tagged chromodomain in an expression vector. The pHT1 plasmid was chosen for this experiment since it contains the identical CHD-1 fragment (bp 583-1203) that was fused with the GAL4 DNA binding domain in the bait in the yeast two-hybrid screen (pHT17).

The *in vitro* protein binding assay uses the his tag on his-N-CoRß to bind Ni^{2+}-nitrilotriacetic acid which is cross-linked to agarose beads (66) (Ni beads). His-tagged proteins are easily isolated when bound to Ni beads since Ni beads can be pelleted by centrifugation. Only proteins with a his tag bind to the Ni beads. Proteins in the lysate that bind to his-N-CoRß are co-precipitated with the beads (Fig.II.15b). Hence the bead pellet should contain his-N-CoRß as well as protein

that bind to his-N-CoRß. The prediction is that the Ni bead pellet of cells co-transfected with his-N-CoRß and pHT1 should contain both the his-N-CoRß protein and the myc-tagged chromodomain protein if the myc-tagged chromodomain protein is bound to the his-N-CoRß.

Figure II.15 is a schematic diagram of the *in vitro* protein binding assay. The results of this assay are presented in Figure II.16. The immunoblot probed with the anti-his antibody is presented in the top box of Figure II.16. Lane 1 contains protein bound to Ni beads from lysates of cells transfected with his-N-CoRß expression vector alone. A band that corresponds to the size of the his-N-CoRß protein is detected in this lane demonstrating that the his-N-CoRß protein bound to Ni agarose beads as predicted. Lane 2 was loaded with the total protein lysate from cells transfected with his-N-CoRß only. The his-N-CoRß band present in this lane represents the amount of his-N-CoRß protein in 40 µg of cell lysate. Lanes 3 and 4 contain protein from cells co-transfected with both his-N-CoRß and pHT1. The his-N-CoR protein is present in both lane 3 which contains protein that bound to Ni beads and lane 4 which contains the total protein lysate , showing that co-transfection with myc-tagged chromodomain does not affect the ability of the his-N-CoR protein to bind to Ni beads. The last two lanes, 5 and 6, contain protein from cells transfected with the myc-tagged chromodomain (pHT1) alone. Since his-N-CoRß was not transfected into these cells there is no band for his-N-CoRß protein in either of these lanes.

The immunoblot was also probed anti-myc 9E10. Lanes 1 and 2 were loaded

with protein from cells transfected with his-N-CoRß alone so there is no anti-myc band in these lanes. Lanes 5 and 6 were loaded with protein from cells transfected with myc-tagged chromodomain alone (pHT1). Lane 6 contains protein from the total cell lysate and the myc-tagged chromodomain protein is present in this lane. On the other hand, in lane 5 which contains protein bound to the Ni beads and the myc-tagged chromodomain protein is not present in this lane, confirming that the Ni beads do not bind to the myc-tagged chromodomain protein. Lanes 3 and 4 contain protein from cells that were co-transfected with his-N-CoRß and the myc-tagged chromodomain protein (pHT1). Lane 4 was loaded with protein from the total protein lysate of the co-transfected cells and the myc-tagged chromodomain protein is detected with anti-myc 9E10 in this lane. The immunoblot had been probed with anti-his as discussed above and was already shown to contain the his-N-CoR protein. Therefore the lysate from co-transfected cells contains both the myc-tagged chromodomain and the his-N-CoRß protein. Lane 3 was loaded with protein from the co-transfected cell lysate that bound to Ni beads. This lane also contained an anti-myc 9E10 immunoreactive band that corresponds in size to the myc-tagged chromodomain (pHT1). Since the myc-tagged chromodomain protein does not bind to the Ni beads on its own (lane 5) then the band in lane 3 must be the myc-tagged chromodomain that bound to his-N-CoRß. This is confirmation that the myc-tagged chromodomain protein interacted with the his-N-CoRß.

The final box on Figure II.15 was the same immunoblot described above

probed with anti-tubulin. This was done as a loading control for the lysate lane and also to ensure that the Ni agarose was not binding non-specific protein. The blot shows that tubulin is present only in the lanes with total protein lysate and not in the lanes with protein that bound to Ni beads. This verifies that the Ni agarose cannot bind protein without a his tag.

4) The endogenous CHD-1 protein co-immunoprecipitates with mSin3B

As discussed in the Introduction to this chapter, N-CoR binds to the mSin3 protein in a large transcriptional repressor complex. If CHD-1 is found to bind to another member of this large complex then this would provide further evidence that CHD-1 has a role in the transcriptional repression mediated by this complex. Hence, co-immunoprecipitation was done with the anti-CHD-1 antibody to determine whether CHD-1 was present in the mSin3 complex. Figure II.17 is a western blot of that co-immunoprecipitation experiment. The anti-CHD-1 antibody was used to immunoprecipitate the endogenous CHD-1 protein and co-immunoprecipitate endogenous proteins that are bound to CHD-1. Immunoprecipitated proteins were bound to IgG beads and to identify co-immunoprecipitated proteins, proteins were eluted from the IgG beads and run on SDS-PAGE followed by immunoblotting.

The immunoblot was initially probed with anti-CHD-1 to confirm that the endogenous CHD-1 protein is immunoprecipitated with the anti-CHD-1 antibody (Fig. II.17a). Lane 3 was loaded with the total protein lysate from P19 cells,

which contain the full length CHD-1 band at 201 kD. Smaller bands are also detected that are attributed to degradation of the 201 kD CHD-1 protein as is typically observed for this protein. Lane 2 was loaded with protein that bound to the IgG beads alone. There are no bands in this lane indicating that the IgG beads do not bind to CHD-1 protein in the absence of antibody. Lane 1 was loaded with immunoprecipitated protein and a band corresponding to the size of the 201 kD CHD-1 protein is detected in the immunoprecipitate. Degradation of the 201 kD CHD-1 protein is observed in this lane as well. In addition the IgG heavy chain band is present as indicated on the figure. This result confirms that the CHD-1 is immunoprecipitated by the anti-CHD-1 antibody.

To see if CHD-1 could bind to mSin3 an immunoblot with anti-CHD-1 co-immunoprecipitated proteins was probed with the anti-mSin3B antibody. There are two mSin3 proteins in mouse, mSin3A and B and both these proteins bind to N-CoR. Only the mSin3B protein is examined in this experiment. The mSin 3B protein will only be present in the co-immunoprecipitate if it interacts with CHD-1. Lane 3 was the total protein lysate of P19 cells. The single band in this lane corresponds to the 175kD mSin3B protein. Lane 2 was protein that bound to beads alone. There are no bands here indicating that mSin3B does not non-specifically bind to IgG beads alone. Lane 1 was the lane containing co-immunoprecipitated protein. There was a band corresponding to the 175kD mSin3B in this lane. There are also bands of smaller sizes located just below the band for full length mSin3B. These bands are thought to be degradation

products of the IgG heavy chain.

Tables, Figure and Illustrations

treatments	n	IF counts	p	$p_1 - p_2$	t	Ho: $\mu_1 - \mu_2 = 0$
pHT8 cells	1000	43	0.0430	-0.0030	0.460	accept
pHT8 nuclear matrix	1000	46	0.0460			
pHT1 cells	1000	109	0.1090	0.093	12.38	reject
pHT1 nuclear matrix	1000	16	0.0160			
pHT12 cells	1900	41	0.0216	0.0086	2.508	reject
pHT12 nuclear matrix	1000	13	0.0130			
pHT15 cells	1300	17	0.0185	0.0095	2.759	reject
pHT15 nuclear matrix	1000	9	0.0090			
pHT14 cells	1800	20	0.0111	0.0003	0.103	accept
pHT14 nuclear matrix	1200	13	0.0108			
pHT16 cells	1700	10	0.0059	-0.0016	0.755	accept
pHT16 nuclear matrix	1200	9	0.0075			

Table II.1. Analysis of the nuclear matrix count data. The columns were labelled as follows:
treatments - the P19 cells were transfected with plasmids indicated (e.g. pHT8) and were intact cells (cells) or cells extracted for the nuclear matrix (nuclear matrix)
n - the total number of cells or nuclear matrices that were counted under the microscope (immunofourescent + non-immunofluorescent) .
IF counts - the number of immunofluorescent cells or nuclear matrices within n.
p - proportion immunfluorescent cells or nuclear matrices = IF counts divided by n.
$p_1 - p_2$ = p for cells (p_1) - p for nuclear matrix (p_2). If $p_1 = p_2$, then $p_1 - p_2 = 0$.
t - the t value of the binomial test which was done to see if $p_1 - p_2$ values deviated significanltly from 0. An example of a t value calculation for the pHT8 transfection follows:

$$t = \frac{p_1 - p_2}{\sqrt{[(n1 \times s1) + (n2 \times s2)] / (n1 + n2)}}$$

$$t = \frac{0.046 - 0.043}{\sqrt{[(1000 \times 0.000044) + (1000 \times 0.000041)] / (1000+1000)}} = 0.46$$

p_1 is the proportion of immunofluorescent cells for the pHT8 transfection
p_2 is the proportion of immunofluorescent nuclear matrices for the pHT8 transfection.
n_1 is the total number of cells counted
n_2 is the total number of nuclear matrices counted
$s_1 = p_1 q_1 / n_1$, where $q_1 = 1 - p_1$
$s_2 = p_2 q_2 / n_2$, where $q_2 = 1 - p_2$

Ho: $\mu_1 - \mu_2 = 0$ - The hypothesis as stated is that the difference between p_1 and p_2 equals 0 or $p_1 - p_2 = 0$. The results of hypothesis testing are in this column. The critical t-value with degrees of freedon >1000 (degrees of freedom = $n_1 + n_2 - 2$) at the 95% confidence limit for a one-tail test is 1.6449. If the calculated t values in the previous column are greater that 1.6449 then the hypothesis is rejected ($p_1 \neq p_2$); if the calculated t values are lower that 1.6449 the the hypothesis is accepted ($p_1 = p_2$).

A41	1300	914	No homology				+
A19	1000	612	No homology				+
B47	2700	680	No homology				+
E120	250	250	No ho mology				++
E33	800	800	No homology				+

1) Yeast two-hybrid clone			2) Genbank		3) Homology		4)
a) clone #	b) Size insert (bp)	c) bases seq (bp)	a) Identity of Genbank clone	b)Size (bp)	a) Alignment	b) % identical nucleotides	ß-gal (xgal)
A91	1700	736	KIAA1237 (human)	5101	i) A91: 1-736 ii) KIAA1237: 523-1246	83% of A91 609/736	+/-
B80	800	570	Glutathione S-transferase (GST) (mouse)	901	i) B80: 7-570 ii) GST:1-560	96%of B80 sequence 543/570	+
B41	800	513	HUB1 (human)	3721	i) B41: 1-513 ii) HUB1: 780-1285	84% of B41 sequence 463/513	++
D18	670	310	KIAA0164 (human)	5101	i) D18: 1-310 ii) KIAA0164: 2163-2472	87%of D18 sequence 272/310	+
A113	1000 650 550	871	17-β-hydroxysteroid dehydrogenase	1741	i)A113:1-129 ii) steroid dehydro-genase: 1609-1738	14.8% of A113 sequence 129/871	+
F111	1000 650	932	tRNA for Phe Val Leu, 12S ribosomal RNA and 16S ribosomal RNA	3121	i) F111: 1-344 ii) RNA: 1268-1612	36.9 %o f F111 sequence 344/932	++
A25	300	850	IGF-II exon 6	3301	i)A25:1-618 ii) IGF-II: 2552-3167	72.7% of A25 sequence	+/-
C47	1000	619	PTD004 (human)	1741	i)C47:1-619 ii)PTD004: 433-1047	91.2% of C47 sequence 565/619	+
E5	1000 400	740	No homology				

Table II.2. Identification of the clones from the mouse embryonic yeast two-hybrid library that interacted with the chromodomain. The columns in the table are labelled as follows:

1) Yeast two-hybrid clones - clones from the mouse embryo library that interacted with the chromodomain.

a) clone # - assigned plasmid numbers for the clones isolated from the mouse embryonic yeast two hybrid library.

b) size insert (bp) - the size in bp of the cDNA inserts in plasmids isolated from the yeast two hybrid library.

c) bases seq (bp) – length of the sequence obtained the ABI sequencer in base pairs

2) Genbank - the cDNA sequence of the clones isolated from the yeast two-hybrid were used to query Genbank for nucleotide sequences showing similarities using the BLAST algorithm (BLASTN version 2.0.10 [August,1999])

a) Identity of Genbank clone - Species origin of the sequence is indicated in brackets. "No homology" indicated that there were no matching sequences in Genbank.

b) Size (bp) - Size of full-length cDNA sequence of the Genbank clone.

3) Homology - analysis of sequence homologies between clones from the yeast two-hybrid and the clones identified in Genbank.

a) Alignment - cDNA sequences from the yeast two-hybrid library clones and Genbank clones were aligned using BLAST and regions where the sequences matched were identified. The number reported is the nucleotide number of the cDNA sequence. i) Region of alignment in the yeast two-hybrid cDNA clone. The nucleotides were numbered starting at the first nucleotide of the open reading frame. ii) Region of alignment in the Genbank cDNA clone. The first nucleotide was that of the cDNA sequence submitted to Genbank.

b) % of identical nucleotides - the number of nucleotides that were found to be identical between the Genbank and yeast two-hybrid cDNA sequences is expressed as a % total number of nucleotides sequenced for the yeast two-hybrid cDNA

4) ß-gal (x-gal) - the strength of interaction between the chromodomain bait and the yeast two-hybrid clones was measured by the amount of ß-galactosidase activity present in a yeast two-hybrid assay (see also Fig. II.9). X-gal was used as the substrate for this assay and the time taken for the blue-colored product to form was proportional to the relative amount of ß-galactosidase activity. The weakest interactions were indicated with "+/-" which mean that the product took greater than 1 hour to form. A "+" indicated that the product was detected a by 45 min, "++" indicates 30 min, "+++" indicates 15 min, and "++++" indicates 5 min which is the strongest interaction.

Prey: yeast two-hybrid clone #	Homologous sequence identified in Genbank	Bait: pHT17 (chromodomain) (ß-galactosidase formed)	Bait: clk-1 (ß-galactosidase formed)	Bait: pAS (ß-galactosidase formed)
D18	KIAA0164	+	-	-
5.7	KIAA0164	+	+++	-
6.12	NCoR	+++	+++	-
2.8	c-myc	++++	+	-
3.9	SAF-B	++	+++	-
6.3	ASF/SF2	+/-	+++	-
1.6	X16 (9.4,10.4)	++	++++	-
6.9	mammalian tra2	-	+++	-

Table II.4 Clk1 and the chromodomain of CHD-1 interact with common proteins. Yeast two-hybrid assay with yeast transformed with the prey and baits indicated above the columns. Transformed yeast were streaked on agar plates containing SD media lacking leucine, tryptophan and histidine. Colony lifts were done to transfer yeast colonies to filters. Yeast were lysed on filter with liq. N_2 and a ß-galactosidase assay was done using X-gal as the substrate. The time taken for the blue-colored product to form was proportional to the relative amount of ß-galactosidase activity. The weakest interactions were indicated with "+/-" which mean that the product took greater than 1 hour to form. A "+" indicated that the product was detected a by 45 min, "++" indicates 30 min, "+++" indicates 15 min, and "++++" indicates 5 min which is the strongest interaction. A "-" indicates that there was no product formed. The negative control pAS is the vector containing the DNA-

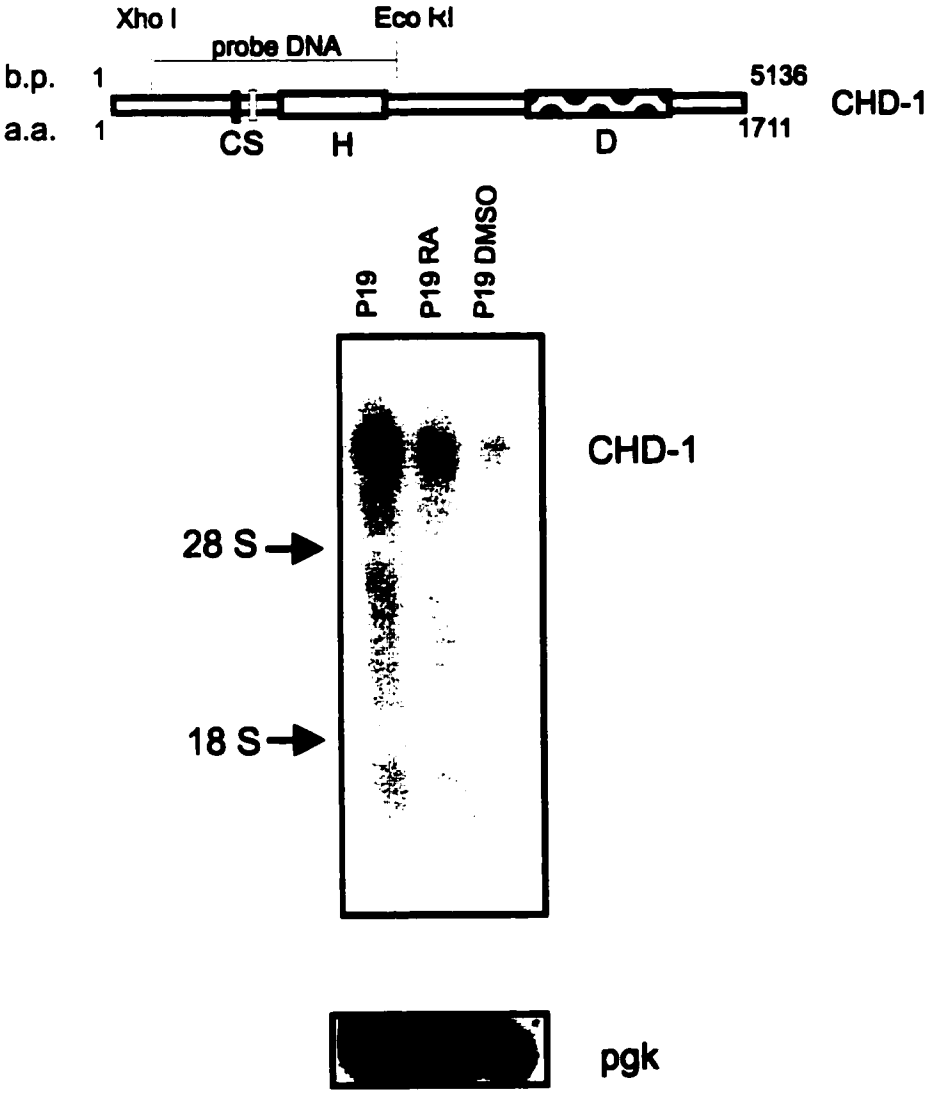

Xho I
Eco kI
probe DNA
b.p. 1
5136
CHD-1
a.a. 1
1711
CS
H
D
P19
P19 RA
P19 DMSO
CHD-1
28 S
18 S
pgk

Figure II.1. CHD-1 mRNA is expressed in P19 cells and differentiated P19 cells. Total RNA from P19 cells (P19), and P19 cells differentiated for 10 days in 10^{-7} M RA (P19 RA) or 10% DMSO (P19 DMSO) were separated on a 0.7% formaldehyde agarose gel and blotted to a nylon membrane (Hybond N). The upper box is the blot was probed with the Xho I to Eco RI fragment of the CHD-1 cDNA as indicated on top of the box. The location of the 28S and 18S ribosomal RNA bands are indicated to the left of the box. The lower box is the same blot was probed with Pgk-1 as a loading control.

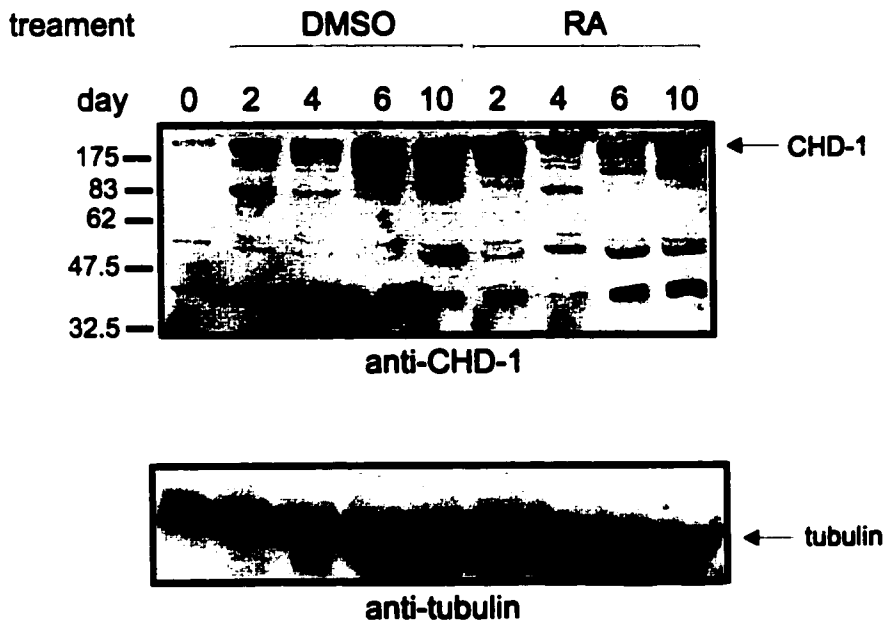

treament
DMSO
RA
day
0
2
4
6
10
2
4
6
10
175
83
62
47.5
32.5
CHD-1
anti-CHD-1
tubulin
anti-tubulin

Figure II.2. CHD-1 protein is expressed in P19 cells and differentiated P19 cells. P19 cells were differentiated by treating them with either 10% DMSO or 10^{-7} M RA as indicated at the top of the figure. The number of days cells received the treatment is also indicated. Lysates from P19 cells treated with either DMSO or RA were loaded onto a 10% SDS PAGE and blotted to nitrocellulose. Top panel: The blot probed with the anti-CHD-1 antibody(Stokes and Perry, 1994). The band corresponding to 201kD CHD-1 band is indicated on the right. The location of molecular weight size markers in kiloDaltons (kD) are indicated on the left of the panel . Bottom panel: The same blot was reprobed with an antibody against tubulin as a loading control.

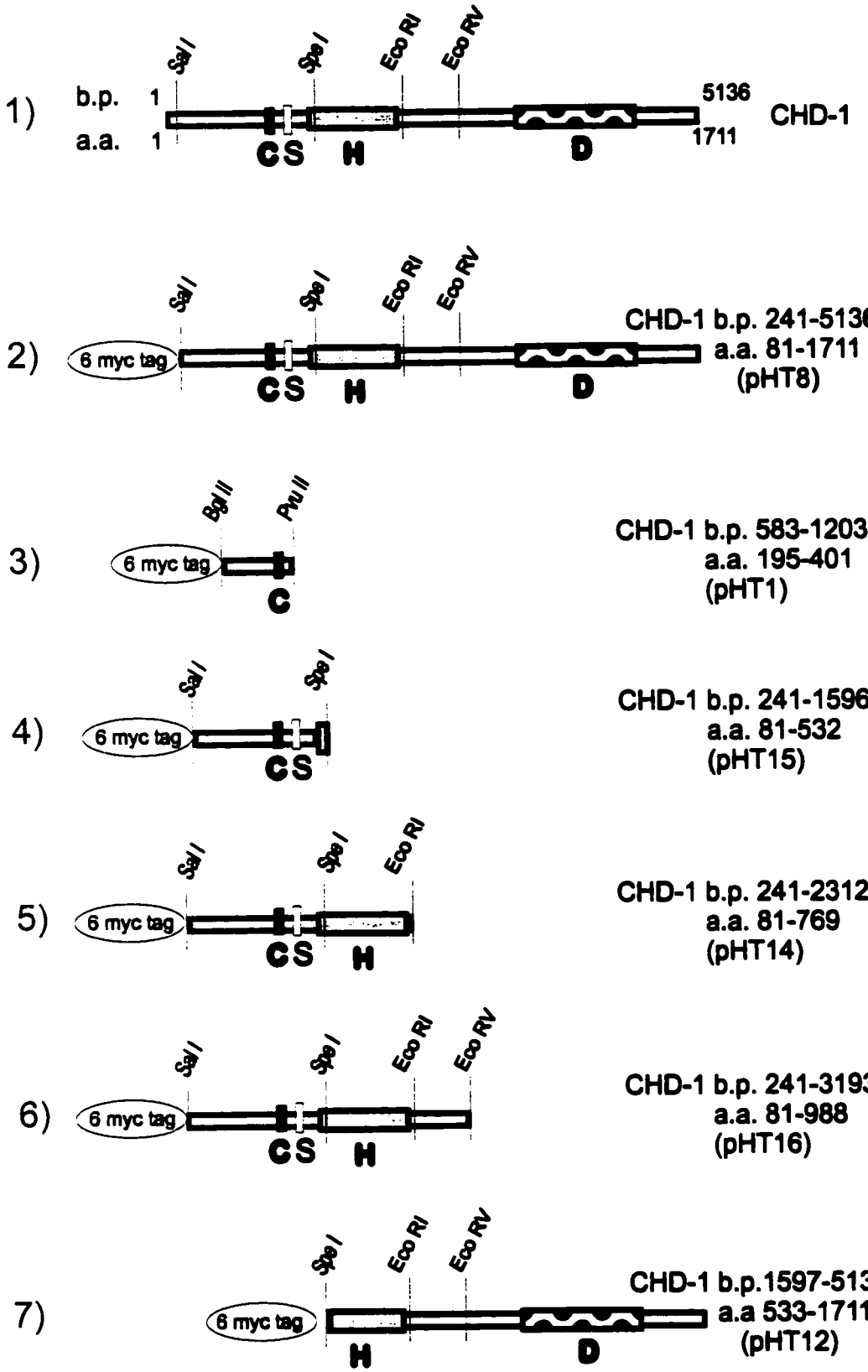

1)
b.p. 1
a.a. 1
Sal I
Spe I
Eco RI
Eco RV
5136
1711
CS
H
D
CHD-1
2)
6 myc tag
CHD-1 b.p. 241-5136
a.a. 81-1711
(pHT8)
3)
Bgl II
Pvu II
CHD-1 b.p. 583-1203
a.a. 195-401
(pHT1)
4)
CHD-1 b.p. 241-1596
a.a. 81-532
(pHT15)
5)
CHD-1 b.p. 241-2312
a.a. 81-769
(pHT14)
6)
CHD-1 b.p. 241-3193
a.a. 81-988
(pHT16)
7)
CHD-1 b.p.1597-5136
a.a 533-1711
(pHT12)

Figure II.3. Schematic diagram of the CHD-1 cDNA deletion mutants. The black rectangle represents the first chromodomain (C); the white rectangle represents the second chromodomain (S); the stippled rectangle represents the helicase domain (H) and the spotted rectangle represents the DNA-binding domain (D). Restriction enzyme sites are indicated for Sal I, Spe I, Eco RI, and Eco RV. The names of the plasmids containing deleted CHD-1 sequences are indicated in brackets. The size of the deleted CHD-1 fragments are also indicated next to the schematic in b.p. and a.a. 1) Full length cDNA of CHD-1 which is 5136 base pairs (b.p.) with a protein sequence that is 1711 amino acids (a.a.). 2) pHT8: myc-tagged CHD-1 from b.p. 241-5136. 3) pHT1: myc tagged CHD-1 b.p. 583-1203. 4) pHT15: myc-tagged CHD-1 b.p. 241-1596. 5) pHT14: myc-tagged CHD-1 b.p. 241-2312. 6) pHT16: myc-tagged CHD-1 b.p. 241-3193. 7) pHT12: myc-tagged CHD-1 b.p. 1597-5136. See also "Material and Methods" section for details on construction of plasmids.

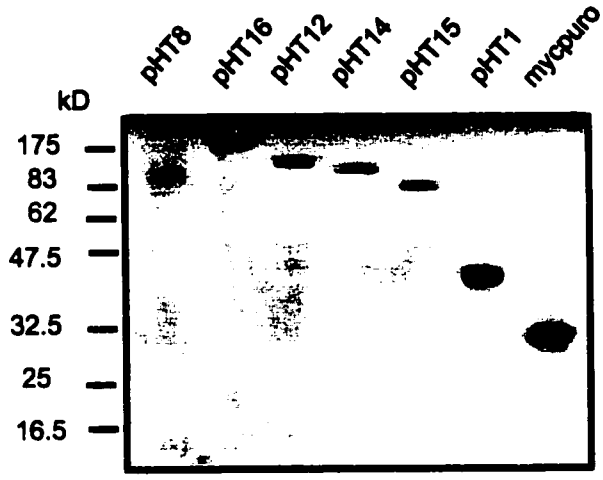
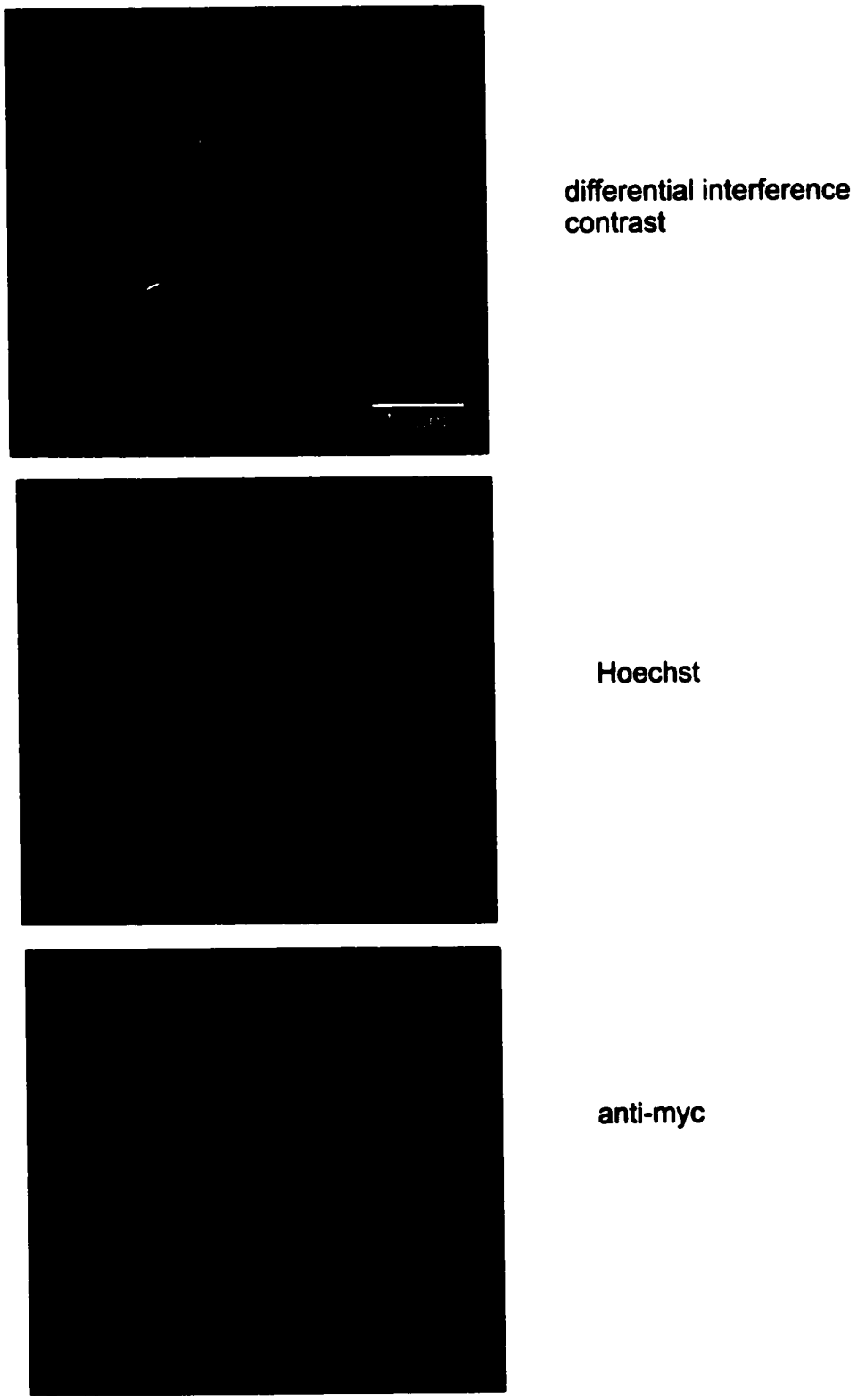

pHT8
pHT16
pHT12
pHT14
pHT15
pHT1
mycpuro
kD
175
83
62
47.5
32.5
25
16.5

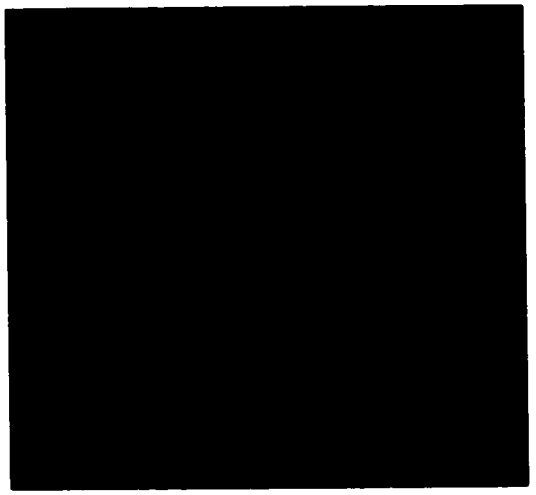
Figure II.4. 293T cells transfected with myc-tagged CHD-1 mutants express tagged proteins. Plasmids used to transfect cells are indicated at the top of the figure (see Figure II.3). The myc-puro lane contained protein from cells transfected with myc-tagged puromycin (myc-puro). The plasmid contained the puromycin antibiotic-resistance gene fused in-frame with the myc-tag. This protein served as a positive control for the anti-myc 9E10 antibody. Protein from transfected cells were lysed with 2X SDS loading buffer and loaded onto a 10% SDS PAGE. The gel was then blotted to nitrocellulose and probed with anti-myc 9E10. The migration of molecular weight size markers in kiloDaltons (kD) are indicated on the left of the figure.

g)

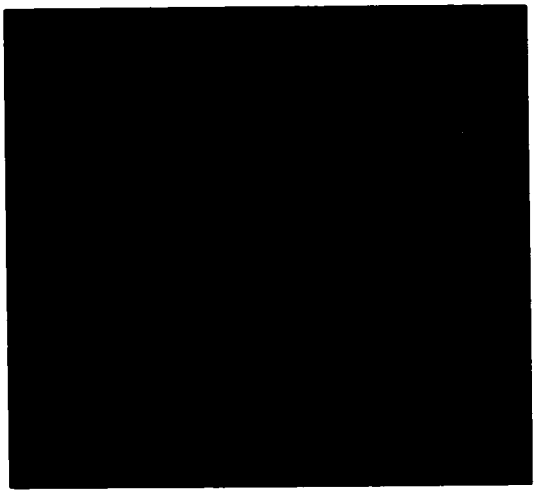

Cells
differential interference contrast
Hoechst
anti-myc
myc-tagged puromycin

f) Cells

differential interference contrast



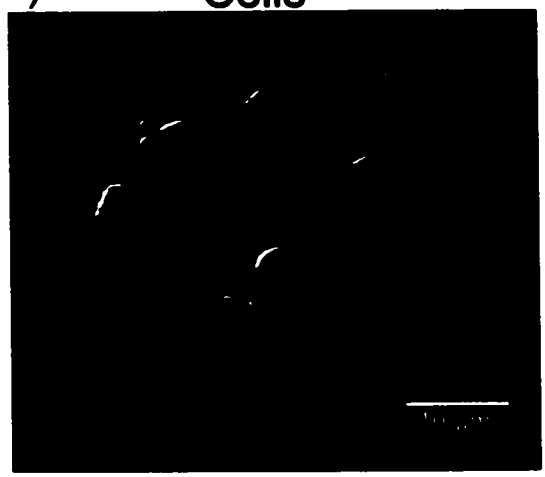
Hoechst



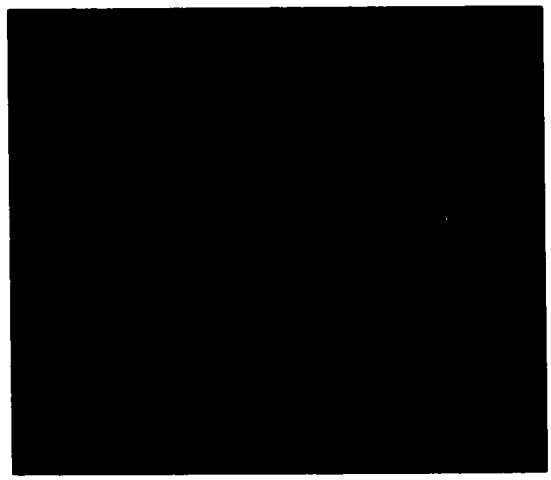
anti-myc

myc-tagged CHD-1 b.p.1597-5136
(pHT12)

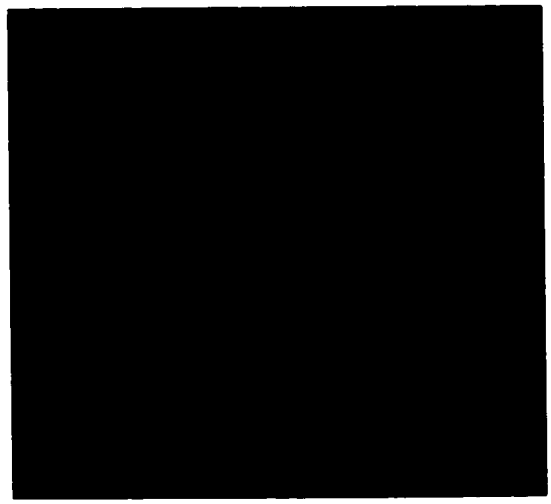
e) Cells



differential interference contrast



Hoechst

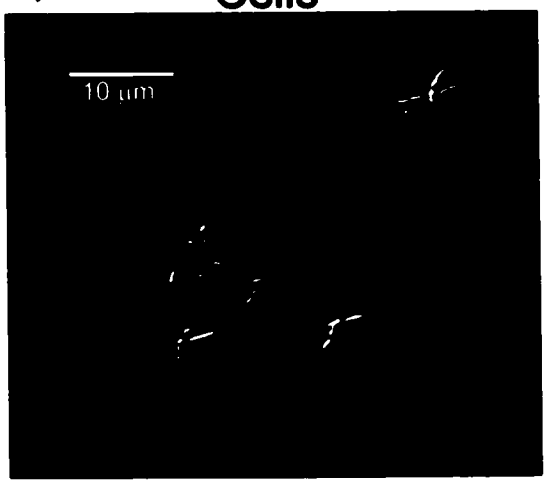


anti-myc

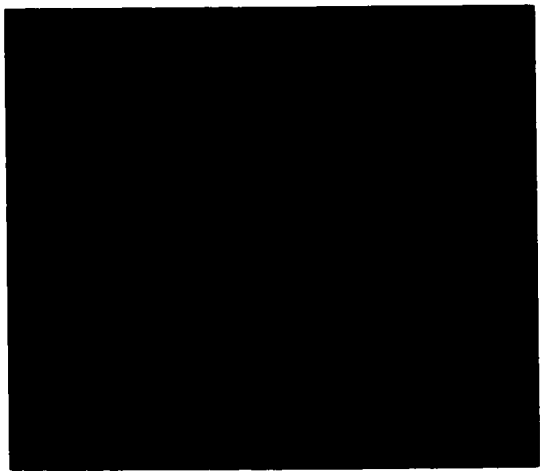
myc-tagged CHD-1 b.p.241-3193 (pHT16)

d)

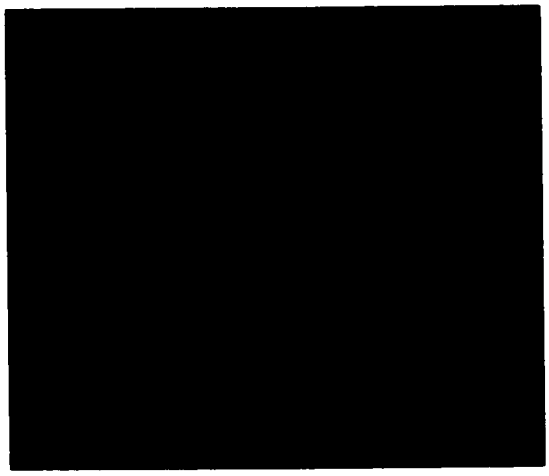
Cells



differential interference contrast



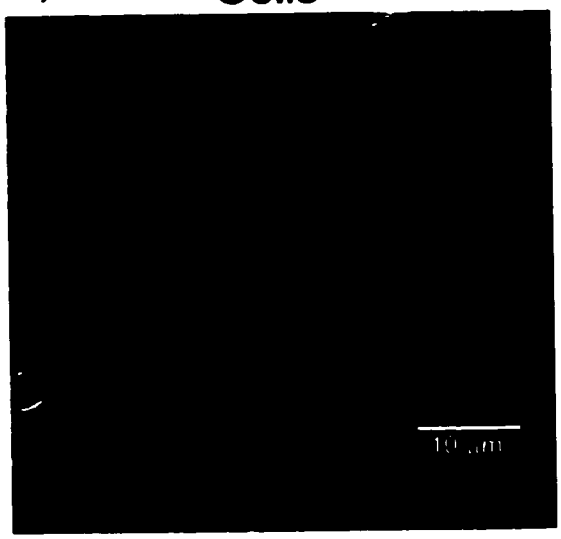
Hoechst



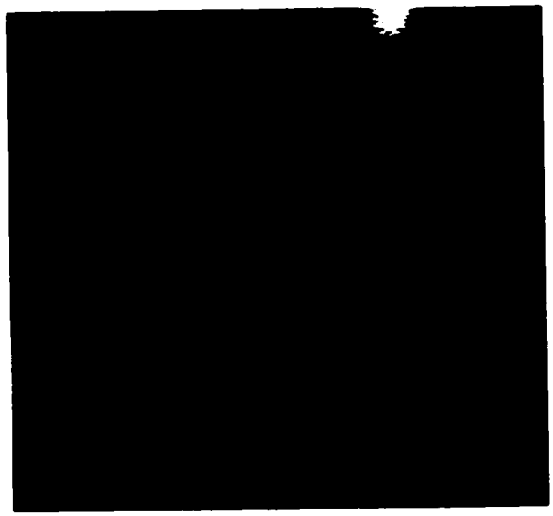
anti-myc

myc-tagged CHD-1 b.p. 241-2312 (pHT14)

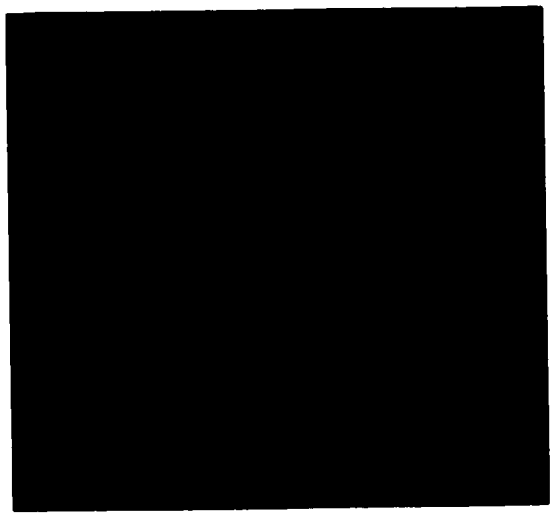
c) Cells



differential interference contrast



Hoechst

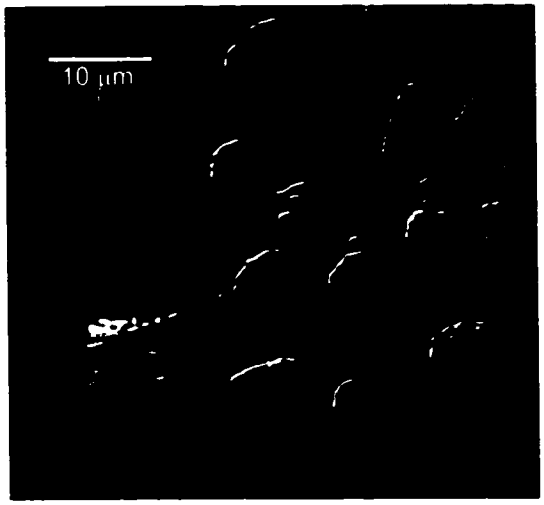


anti-myc

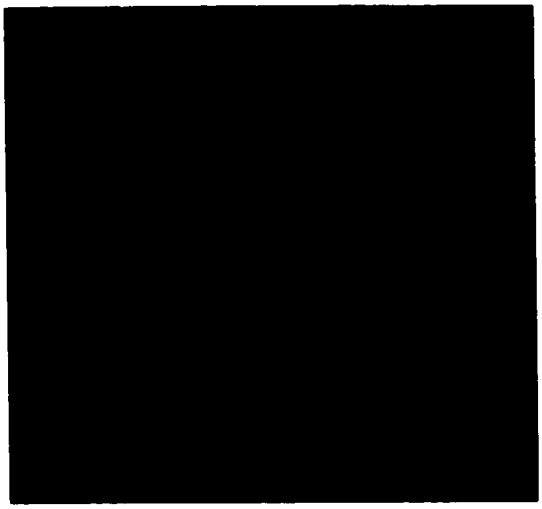
myc-tagged CHD-1 b.p. 241-1596 (pHT15)

b)

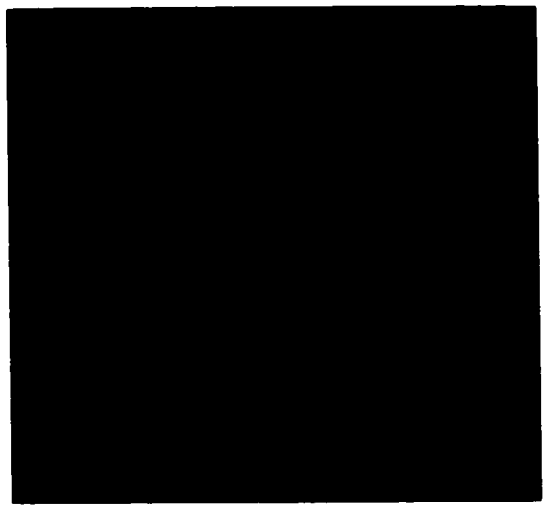
Cells



differential interference contrast



Hoechst

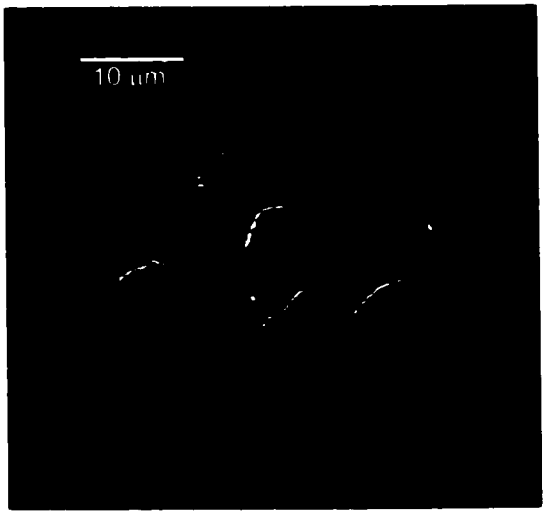


anti-myc

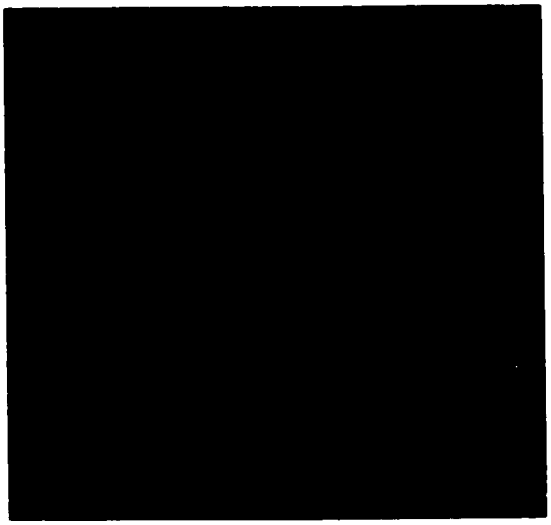
myc-tagged CHD-1 b.p.583-1203 (pHT1)

a)

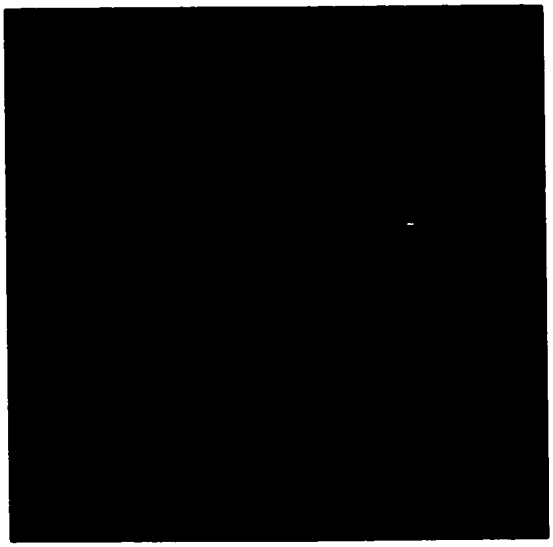
Cells



differential interference contrast



Hoechst



anti-myc

myc-tagged CHD-1 b.p.241-5136 (pHT8)

Figure II.5a. P19 cells transfected with myc-tagged CHD-1 mutants show nuclear localization of tagged proteins. P19 cells were grown on coverslips and transfected with a) myc-tagged CHD-1 b.p. 241-5136 (pHT8); b) myc-tagged CHD-1 b.p. 583-1203 (pHT1); c) myc-tagged CHD-1 b.p. 241-1596 (pHT15); d) myc-tagged CHD-1 b.p. 241-2312 (pHT14); e) myc-tagged CHD-1 b.p. 241-3193 (pHT16); f) myc-tagged CHD-1 b.p. 1597-5136 (pHT12); g) myc-tagged puromycin. Transfected cells were fixed with paraformaldehyde and stained with anti-myc (9E10) followed by a secondary anti-mouse cy3-conjugated antibody for visualization. The top panel is the differential interference contrast image of the field. The middle panel is the Hoechst DNA stain of the same field. The bottom panel is the immunofluorescence staining with the anti-myc (9E10) followed by anti-mouse cy3. Bar = 10μm.

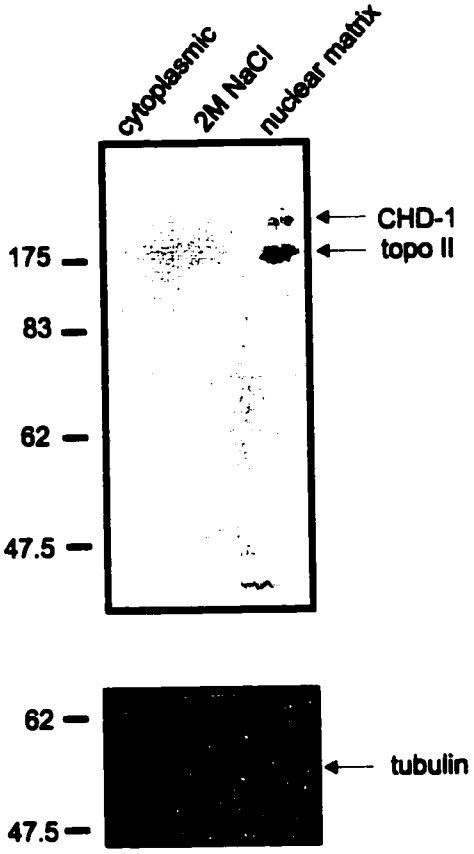
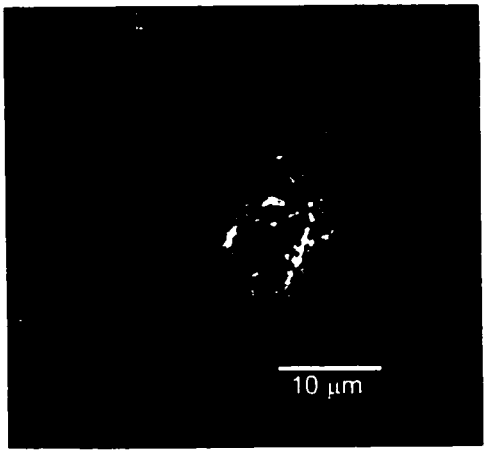
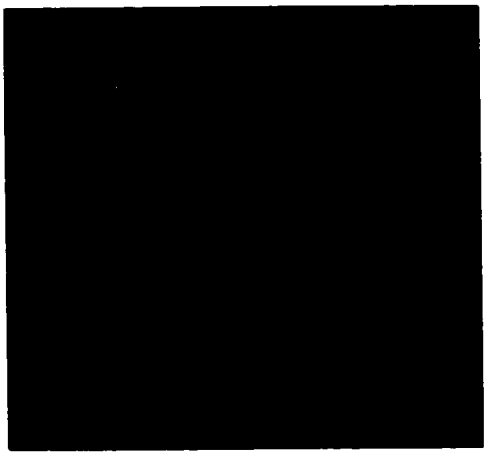

cytoplasmic
2M NaCl
nuclear matrix
CHD-1
topo II
175
83
62
47.5
62
47.5
tubulin

Figure II.6. CHD-1 protein is in the nuclear matrix fraction of P19 cells. A nuclear matrix extraction was done on P19 cells (29). Cytoplasmic refers to the cytoplasmic fraction of this extraction, 2M NaCl refer to the proteins that were soluble in 2 M NaCl and the nuclear matrix contains the proteins that are not soluble in 2 M NaCl. Proteins from these three fractions were run on an SDS-PAGE gels and blotted. Top panel: the blot was probed with the anti-CHD-1 and the anti-topoisomerase II antibody. The locations of the endogenous CHD-1 and topoisomerase II (topo II) protein are indicated on the right side of the panel. Bottom panel: the same blot was re-probed with the anti-tubulin antibody. The location of the tubulin protein is indicated on the right side of the panel. Numbers on the left indicate the location and size of the protein molecular weight marker in kiloDaltons.

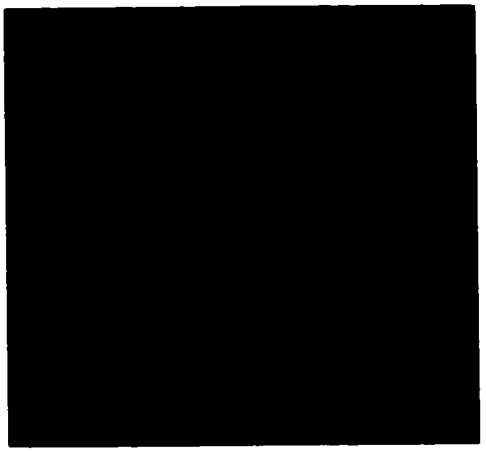
Cells



differential interference contrast



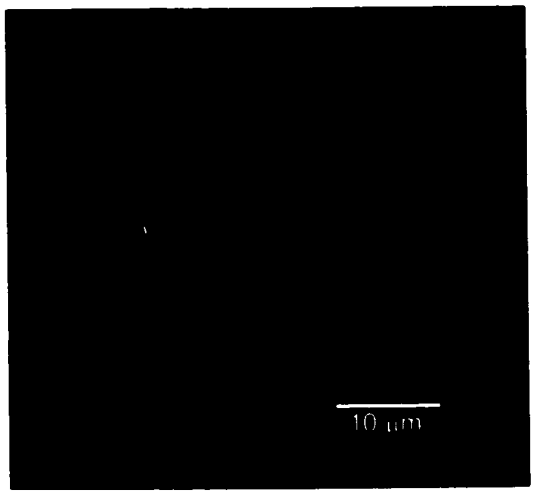
Hoechst



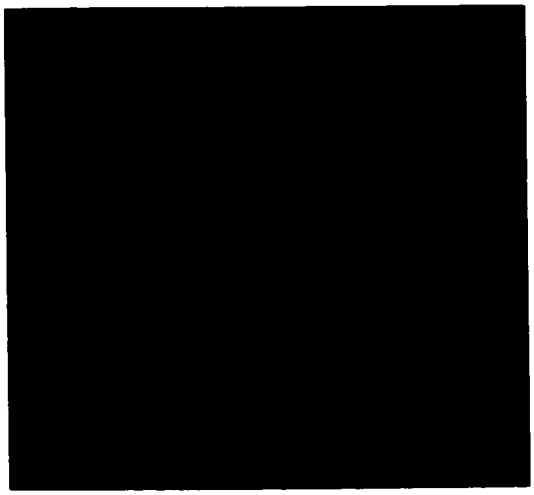
anti-Topo II

Figure II.7a. Anti-Topoisomerase II (topo II) immunofluorescence of P19 cells. P19 cells grown on coverslips were fixed with paraformaldehyde and stained with the anti-topo II antibody followed by an anti-rabbit FITC-conjugated secondary antibody. The top panel is the differential contrast image of the field. The middle panel is the Hoechst DNA stain of the same field. The bottom panel is the immunofluorescence staining with the anti-topo II antibody followed by anti-rabbit FITC. Bar = 10μm.

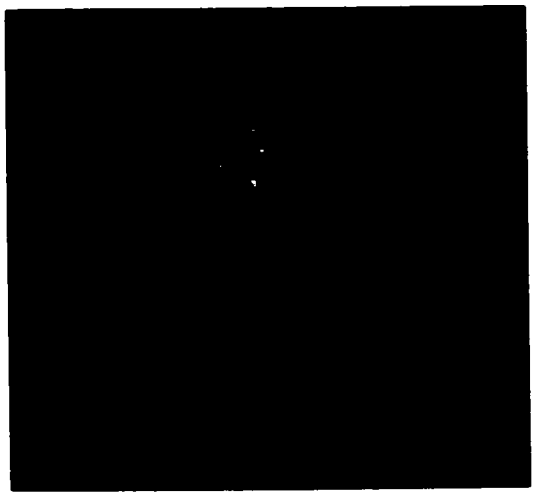
Nuclear Matrix



differential interference contrast



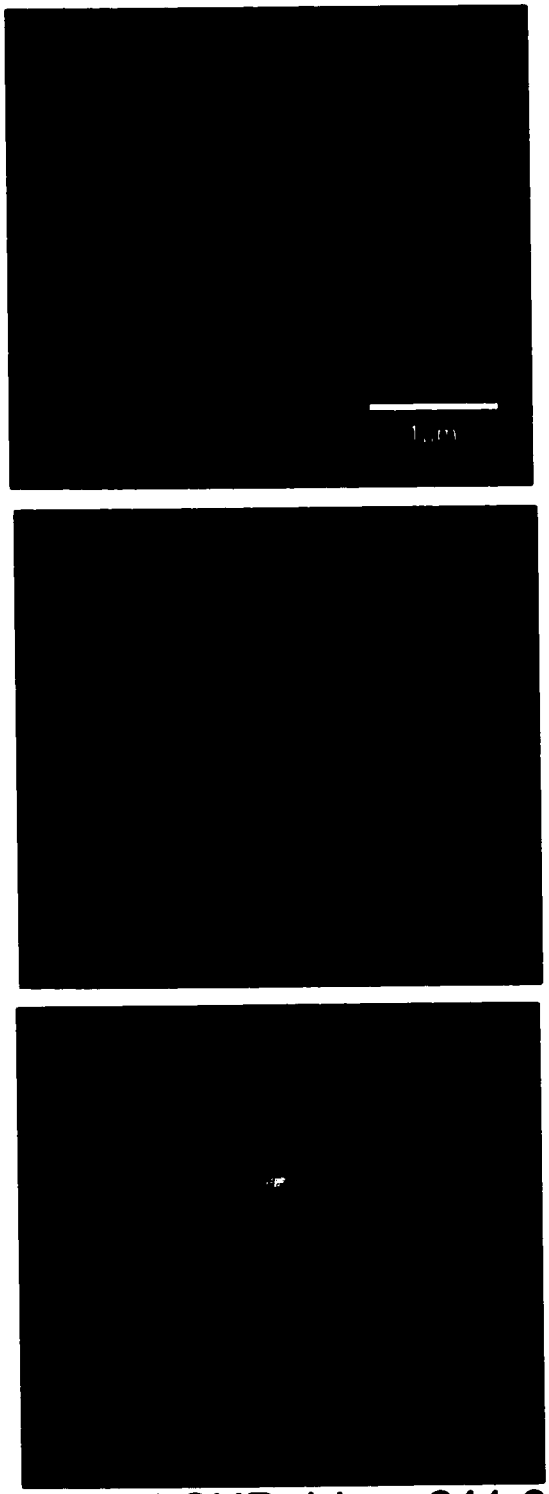
Hoechst



anti-Topo II

Figure II.7b. Anti-Topoisomerase II (Topo II) immunofluorescence of nuclear matrices from P19 cells. A nuclear matrix extraction was done on P19 cells grown on coverslips (25). The residual nuclear matrix left on coverslips was fixed in paraformaldehyde and stained with the anti-topo II antibody followed by an anti-rabbit FITC-conjugated secondary antibody. The top panel is the differential contrast image of the field. The middle panel is the Hoechst DNA stain of the same field. The bottom panel is the immunofluorescence staining with the anti-topo II antibody followed by anti-rabbit FITC. Bar = 10μm.

c) Nuclear Matrix



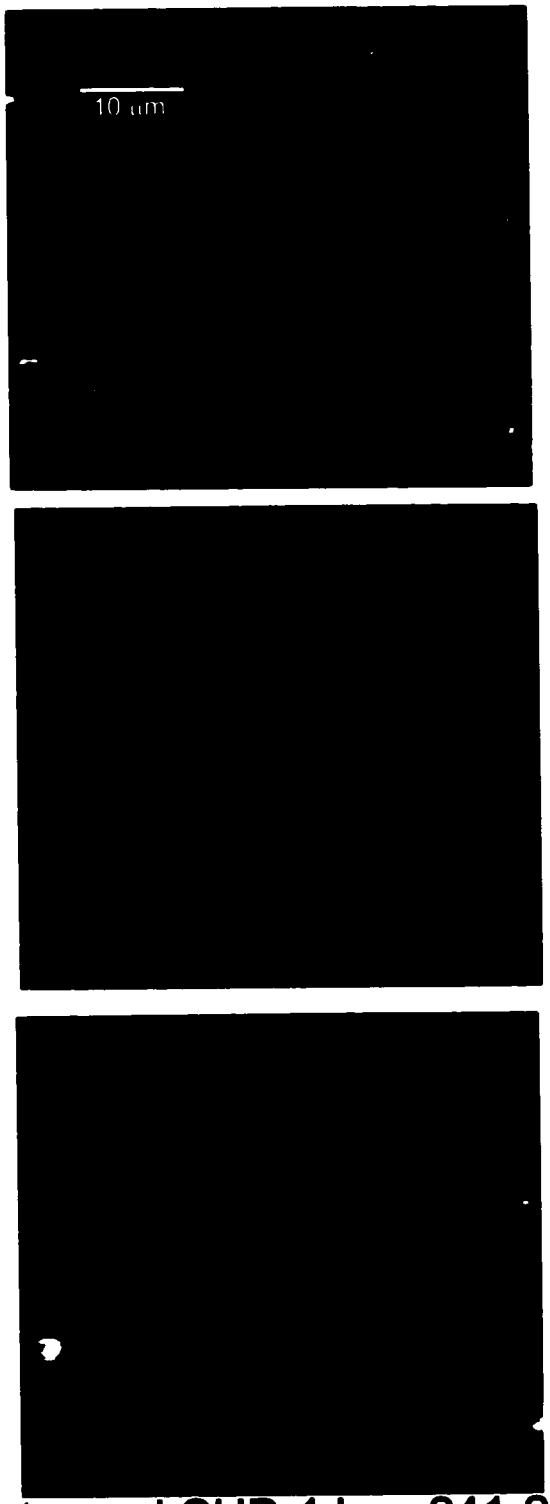
differential interference contrast

Hoechst

anti-myc

myc-tagged CHD-1 b.p. 241-3193 (pHT16)

b) Nuclear Matrix



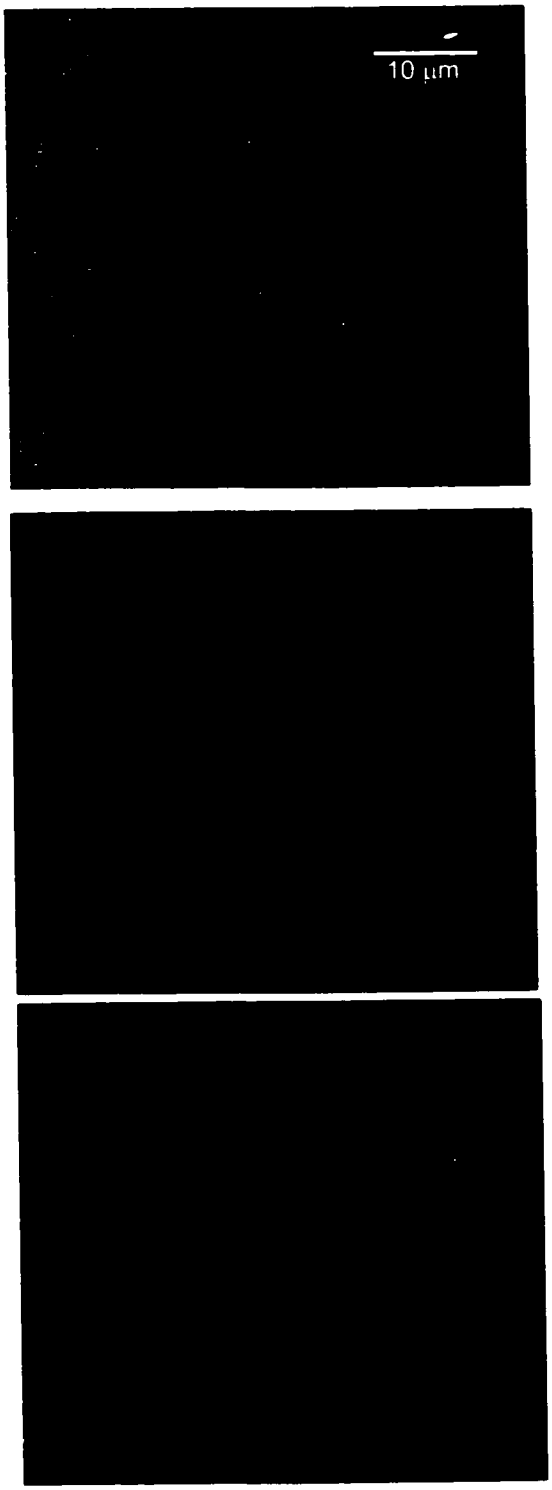
differential interference contrast

Hoechst

anti-myc

myc-tagged CHD-1 b.p. 241-2312 (pHT14)

a) Nuclear Matrix



differential interference contrast

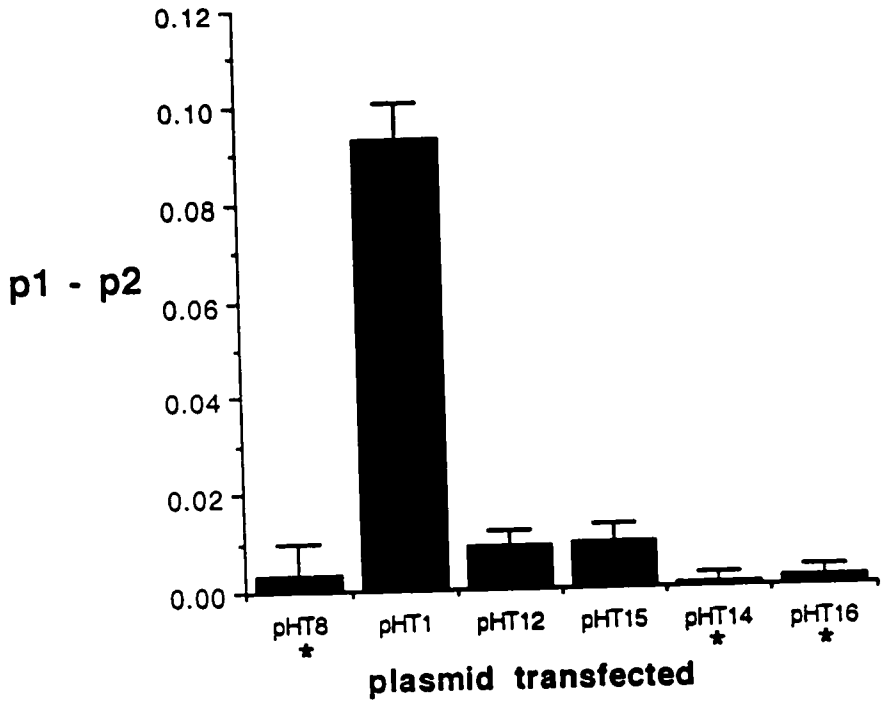
Hoechst

anti-myc

myc-tagged CHD-1 b.p. 241-5136 (pHT8)

Figure II.8a. Anti-myc (9E10) immunofluorescence of nuclear matrices prepared from P19 cells transfected with a) myc-tagged CHD-1 b.p. 241-5 136 (pHT8); b) myc-tagged CHD-1 b.p. 241-2312 (pHT14); c) myc-tagged CHD-1 b.p. 241-3193 (pHT16). The top panel is the differential interference contrast image of the field. The middle panel is the Hoechst DNA stain of the same field. The bottom panel is the same field probed with anti-myc (9E10) and visualized using cy3 conjugated secondary antibody against mouse. Bar = 10μm.

Graph of the difference between the proportions cells nuclear matrices that were stained with anti-myc 9E10.



p1 = proporiton of cells that were immunofluorescent when stained with anti-myc 9E10

p2 = proportion of nuclear matrices that were immunofluorescent when stained with anti-myc 9E10

* p1 - p2 = 0, proportion of immunofluorescent cells and nuclear matrices are the same

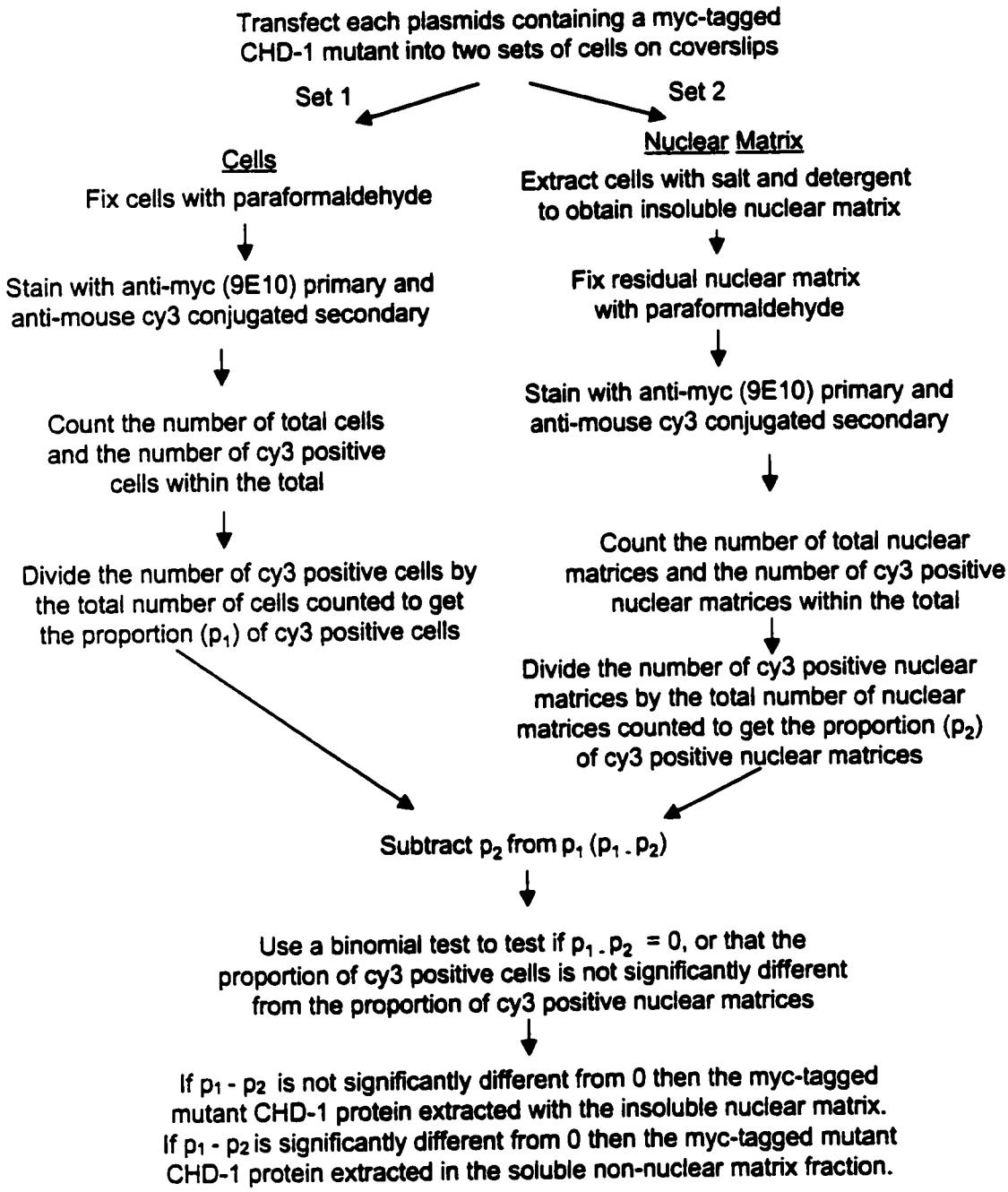

Transfect each plasmids containing a myc-tagged CHD-1 mutant into two sets of cells on coverslips
Set 1
Set 2
Cells
Fix cells with paraformaldehyde
Stain with anti-myc (9E10) primary and anti-mouse cy3 conjugated secondary
Count the number of total cells and the number of cy3 positive cells within the total
Divide the number of cy3 positive cells by the total number of cells counted to get the proportion (p_1) of cy3 positive cells
Nuclear Matrix
Extract cells with salt and detergent to obtain insoluble nuclear matrix
Fix residual nuclear matrix with paraformaldehyde
Stain with anti-myc (9E10) primary and anti-mouse cy3 conjugated secondary
Count the number of total nuclear matrices and the number of cy3 positive nuclear matrices within the total
Divide the number of cy3 positive nuclear matrices by the total number of nuclear matrices counted to get the proportion (p_2) of cy3 positive nuclear matrices
Subtract p_2 from p_1 ($p_1 - p_2$)
Use a binomial test to test if $p_1 - p_2 = 0$, or that the proportion of cy3 positive cells is not significantly different from the proportion of cy3 positive nuclear matrices
If $p_1 - p_2$ is not significantly different from 0 then the myc-tagged mutant CHD-1 protein extracted with the insoluble nuclear matrix.
If $p_1 - p_2$ is significantly different from 0 then the myc-tagged mutant CHD-1 protein extracted in the soluble non-nuclear matrix fraction.

Figure II.9a) The method used for analyzing nuclear matrix localization of mutant CHD-1 proteins. b) A graph of $p_1 - p_2$ values (see Table II.1). p_1 and p_2 are the proportions of cells and nuclear matrices, respectively, that were immunofluorescent when stained anti-myc 9E10. The y-axis are values of $p_1 - p_2$. The error bars represent the standard error of the mean (SEM) for $p_1 - p_2$ values which was calculated as follows:

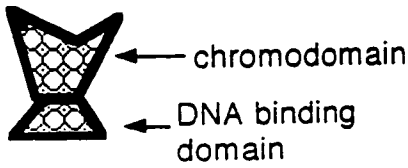
$$SEM = \sqrt{[(n1 \times s1) + (n2 \times s2)] / (n1 + n2)}$$

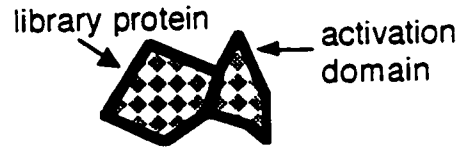
See Table II.1 for values of n_1, n_2, s_1, s_2.

The x-axis indicates the plasmid that was transfected into cells. The cells transfected with pHT8, 14, 16 have proportions of immunofluorescent cells and nuclear matrices that are not significantly different as indicated (*).

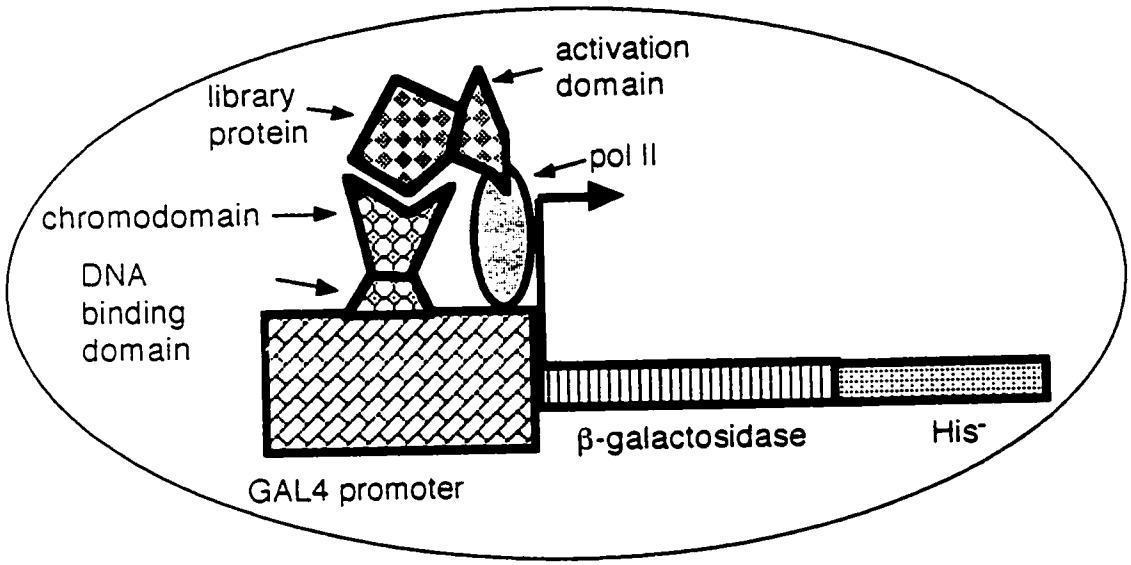
Plasmid containing the chromodomain "bait" (PHT17) fused to the DNA binding domain also had a Trp$^-$ resistance gene.

Plasmid containing the cDNA library fused to the activation domain also had a Leu$^-$ resistance gene.





Yeast transformed with both plasmids were selectable in Leu$^-$Trp$^-$ media. The yeast cell line used for the yeast two-hybrid screen had a genomic DNA insert with the GAL4 promoter driving the β-galactosidase gene and the His$^-$ resistance gene. The GAL4 promoter bound to the DNA binding domain but could not stimulate transciption without the activation domain. Likewise, the activation domain could not stimulate transcription unless it was at a promoter which brings it close to RNA pol II.

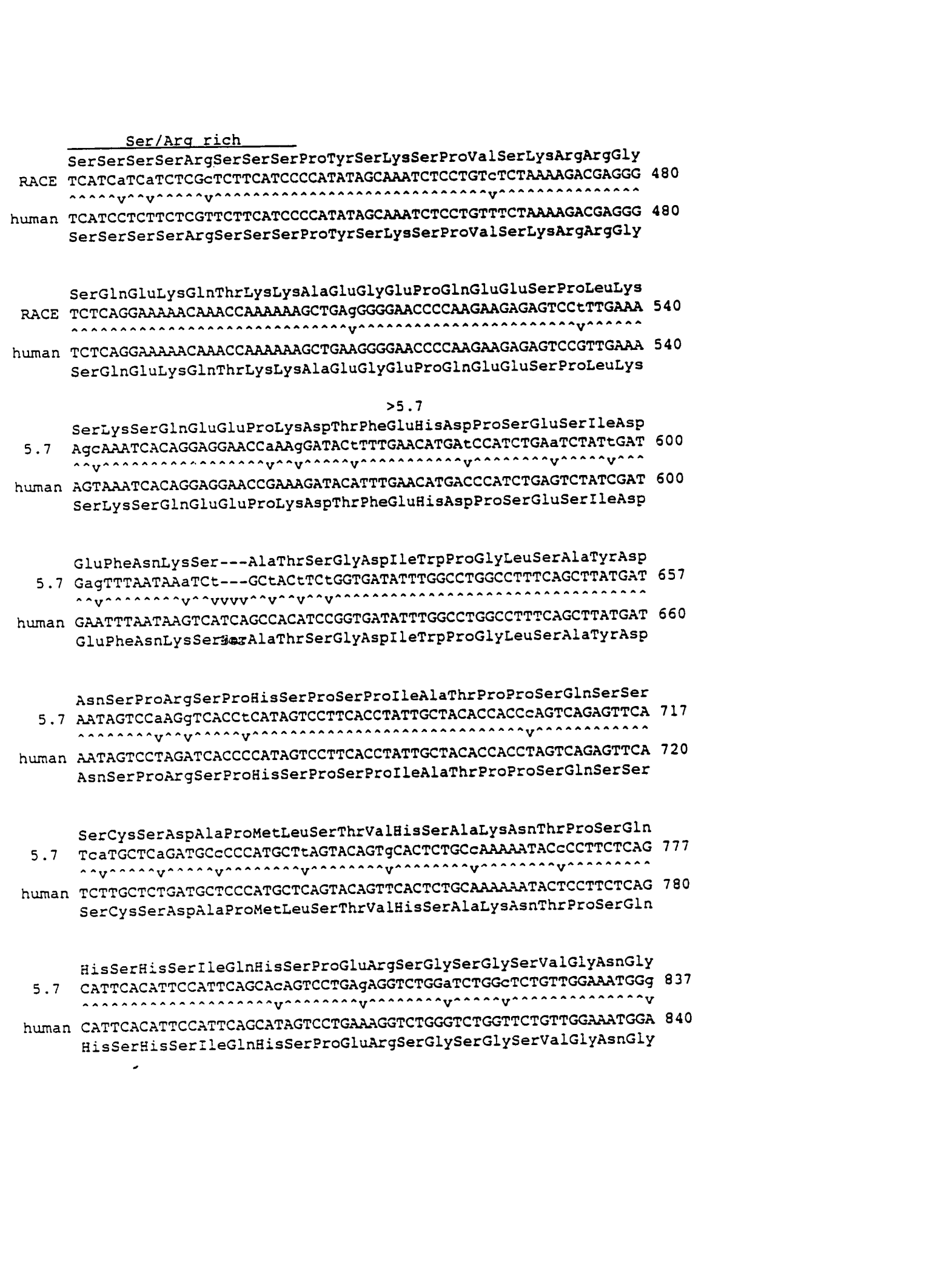


A positive interaction between the chromodomain and the library protein recreates a transcription factor with both a DNA binding domain and activation domain which activates RNA pol II. This results in transcription of the β-galactosidase and the His$^-$ resistance gene. Yeast containing interacting proteins were selected in Leu$^-$Trp$^-$His$^-$ media. This could be confirmed by testing the yeast for β-galactosidase expression using an enzyme activity assay.

Figure II.10 Schematic diagram of the yeast two-hybrid screen for proteins that interacted with the chromodomain.

```
      ThrProLysSerLysLysTyrPheLeuHisAspAspArgAspAspGlyValAspTyrTrp
 5.7  ACCCCAAAGAGCAAGAAGTACTTCTTGCAcGAtGACAGAGATGATGGTGTGGATTAcTGG 2574
      ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^v^^v^^^^^^^^^^^^^^^^^^^^^v^^^
human ACCCCAAAGAGCAAGAAGTACTTCTTGCATGACGACAGAGATGATGGTGTGGATTATTGG 2577
      ThrProLysSerLysLysTyrPheLeuHisAspAspArgAspAspGlyValAspTyrTrp

      AlaLysArgGlyArgGlyArgGlyThrPheGlnArgGlyArgGlyArgPheAsnPheLys
 5.7  GCCAAAAGAGGAAGAGGTCGTGGTACTTTTCAACGTGGCAGAGGGCGCTTTAAtTTCAAA 2634
      ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^v^^^^^^
human GCCAAAAGAGGAAGAGGTCGTGGTACTTTTCAACGTGGCAGAGGGCGCTTTAACTTCAAA 2637
      AlaLysArgGlyArgGlyArgGlyThrPheGlnArgGlyArgGlyArgPheAsnPheLys

      LysSerGlySerSerProLysTrpThrHisAspLysTyrGlnGlyAspGlyIleValGlu
 5.7  AAATCAGGTAGCAGTCCaAAATGGACTCATGACAAATACCAAGGaGATGGGATTGTTGAA 2694
      ^^^^^^^^^^^^^^^^^^v^^^^^^^^^^^^^^^^^^^^^^^^^^v^^^^^^^^^^^^^^^
human AAATCAGGTAGCAGTCCTAAATGGACTCATGACAAATACCAAGGGGATGGGATTGTTGAA 2697
      LysSerGlySerSerProLysTrpThrHisAspLysTyrGlnGlyAspGlyIleValGlu

      AspAspGluGluThrMetGluAsnAsnGluGluLysLysAspArgArgLysGluGluLys
 5.7  GATGAcGAAGAGACCATGGAAAATAATGAAGAgAAGAAGGACAGACGCAAaGAAGAAAAG 2754
      ^^^^^v^^^^^^^^^^^^^^^^^^^^^^^^^^^v^^^^^^^^^^^^^^^^^v^^^^^^^^^
human GATGAAGAAGAGACCATGGAAAATAATGAAGAAAAGAAGGACAGACGCAAGGAAGAAAAG 2757
      AspAspGluGluThrMetGluAsnAsnGluGluLysLysAspArgArgLysGluGluLys

      GluEnd
 5.7  GAATAg 2760
      ^^^^^v
human GAATAA 2763
      GluEnd
```

```
                                                  D18
          ArgLysGluArgSerLysGluArgGlyAspSerLysGlySerArgGluSerSerGlySer
 5.7      cGtAAAGAgAGAAGTAAAGAgCGGGGAGATTCCAAGGGCTCCAGGGAgTCCAGTGGgTCc 2154
          v^v^^^^^v^^^^^^^^^^^v^^^^^^^^^^^^^^^^^^^^^^^^^^v^^^^^^^^v^^v
human     AGAAAAGAAAGAAGTAAAGAACGGGGAGATTCCAAGGGCTCCAGGGAATCCAGTGGATCA 2160
          ArgLysGluArgSerLysGluArgGlyAspSerLysGlySerArgGluSerSerGlySer

                                                  D18
          ArgLysGlnGluLysThrProLysAspTyrLysGluTyrLysProTyrLysAspAspSer
 5.7      AGAAAGCAaGAAAAAACaCCAAAgGATTACAAGGAgTAtAAAcCTTACAAAGATGACAGc 2214
          ^^^^^^^^v^^^^^^^^v^^^^^v^^^^^^^^^^^v^^v^^^v^^^^^^^^^^^^^^^v
human     AGAAAGCAGGAAAAAACTCCAAAAGATTACAAGGAATACAAATCTTACAAAGATGACAGT 2220
          ArgLysGlnGluLysThrProLysAspTyrLysGluTyrLysSerTyrLysAspAspSer

          LysHisLysGlyArgGluArgAspHisSerArgSerSerSerSerSerAlaSerProSer
 5.7      AAACATAAAggtAGAGAGCggGAcCATTCcCGgTCaTCgTCCTCgTCAGCATCgCCgTCc 2274
          ^^^^^^^^^vvv^^^^^^^vv^^v^^^^^v^^v^^v^^v^^^^^v^^^^^^^^v^^v^^v
human     AAACATAAA---AGAGAGCAAGATCATTCTCGATCTTCATCCTCTTCAGCATCACCTTCT 2277
          LysHisLys---ArgGluGlnAspHisSerArgSerSerSerSerSerAlaSerProSer

          SerProSerSerArgGluGluLysGluSerLysLysGluArgGluGluGluPheLysThr
 5.7      TcgCCCAGTTCTCGAGAAGAAAAaGAGAGTAAGAAaGAgAGAGAAGAAGAATTTAAAACT 2334
          ^^v^^^^^^^^^^^^^^^^^^^^v^^^^^^^^^^^v^^v^^^^^^^^^^^^^^^^^^^^
human     TCTCCCAGTTCTCGAGAAGAAAAGGAGAGTAAGAAGGAAAGAGAAGAAGAATTTAAAACT 2337
          SerProSerSerArgGluGluLysGluSerLysLysGluArgGluGluGluPheLysThr

          HisHisGluMetLysAspTyrSerGlyPheAlaGlyValSerArgProArgGlyThrPhe
 5.7      CACCAcGAgATGAAAGAcTACTCAGGaTTTGCAGGAGTTAGCaGACCACGAGGAACCTTT 2394
          ^^^^^v^^v^^^^^^^^v^^^^^^^^v^^^^^^^^^^^^^^^v^^^^^^^^^^^^^^^^
human     CACCATGAAATGAAAGAATACTCAGGCTTTGCAGGAGTTAGCCGACCACGAGGAACCTTT 2397
          HisHisGluMetLysGluTyrSerGlyPheAlaGlyValSerArgProArgGlyThrPhe

          PheArgIleArgGlyArgGlyArgAlaArgGlyValPheAlaGlyThrAsnThrGlyPro
 5.7      TTTCGAATTAGAGGCAGAGGAAGAGCCAGAGGAGTTTTTGCTGGGACAAATACTGGTCCA 2454
          ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
human     TTTCGAATTAGAGGCAGAGGAAGAGCCAGAGGAGTTTTTGCTGGGACAAATACTGGTCCA 2457
          PheArgIleArgGlyArgGlyArgAlaArgGlyValPheAlaGlyThrAsnThrGlyPro

          AsnAsnSerAsnThrThrPheGlnLysArgProLysGluGluGluTrpAspProGluTyr
 5.7      AACAACTCAAATACcACTTTTCAAAAGAGACCGAAGGAAGAGGAATGGGATCCAGAATAT 2514
          ^^^^^^^^^^^^^^v^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
human     AACAACTCAAATACTACTTTTCAAAAGAGACCGAAGGAAGAGGAATGGGATCCAGAATAT 2517
          AsnAsnSerAsnThrThrPheGlnLysArgProLysGluGluGluTrpAspProGluTyr
```


Ser/Arg rich
SerSerSerSerArgSerSerSerProTyrSerLysSerProValSerLysArgArgGly
RACE TCATCaTCaTCTCGcTCTTCATCCCCATATAGCAAATCTCCTGTcTCTAAAAGACGAGGG 480
human TCATCCTCTTCTCGTTCTTCATCCCCATATAGCAAATCTCCTGTTTCTAAAAGACGAGGG 480
SerSerSerSerArgSerSerSerProTyrSerLysSerProValSerLysArgArgGly
SerGlnGluLysGlnThrLysLysAlaGluGlyGluProGlnGluGluSerProLeuLys
RACE TCTCAGGAAAAACAAACCAAAAAAGCTGAgGGGGAACCCCAAGAAGAGAGTCCtTTGAAA 540
human TCTCAGGAAAAACAAACCAAAAAAGCTGAAGGGGAACCCCAAGAAGAGAGTCCGTTGAAA 540
SerGlnGluLysGlnThrLysLysAlaGluGlyGluProGlnGluGluSerProLeuLys
>5.7
SerLysSerGlnGluGluProLysAspThrPheGluHisAspProSerGluSerIleAsp
5.7 AgcAAATCACAGGAGGAACCaAAgGATACtTTTGAACATGAtCCATCTGAaTCTATtGAT 600
human AGTAAATCACAGGAGGAACCGAAAGATACATTTGAACATGACCCATCTGAGTCTATCGAT 600
SerLysSerGlnGluGluProLysAspThrPheGluHisAspProSerGluSerIleAsp
GluPheAsnLysSer---AlaThrSerGlyAspIleTrpProGlyLeuSerAlaTyrAsp
5.7 GagTTTAATAAaTCt---GCtACtTCtGGTGATATTTGGCCTGGCCTTTCAGCTTATGAT 657
human GAATTTAATAAGTCATCAGCCACATCCGGTGATATTTGGCCTGGCCTTTCAGCTTATGAT 660
GluPheAsnLysSerSerAlaThrSerGlyAspIleTrpProGlyLeuSerAlaTyrAsp
AsnSerProArgSerProHisSerProSerProIleAlaThrProProSerGlnSerSer
5.7 AATAGTCCaAGgTCACCtCATAGTCCTTCACCTATTGCTACACCACCcAGTCAGAGTTCA 717
human AATAGTCCTAGATCACCCCATAGTCCTTCACCTATTGCTACACCACCTAGTCAGAGTTCA 720
AsnSerProArgSerProHisSerProSerProIleAlaThrProProSerGlnSerSer
SerCysSerAspAlaProMetLeuSerThrValHisSerAlaLysAsnThrProSerGln
5.7 TcaTGCTCaGATGCcCCCATGCTtAGTACAGTgCACTCTGCcAAAAATACcCCTTCTCAG 777
human TCTTGCTCTGATGCTCCCATGCTCAGTACAGTTCACTCTGCAAAAAATACTCCTTCTCAG 780
SerCysSerAspAlaProMetLeuSerThrValHisSerAlaLysAsnThrProSerGln
HisSerHisSerIleGlnHisSerProGluArgSerGlySerGlySerValGlyAsnGly
5.7 CATTCACATTCCATTCAGCAcAGTCCTGAgAGGTCTGGaTCTGGcTCTGTTGGAAATGGg 837
human CATTCACATTCCATTCAGCATAGTCCTGAAAGGTCTGGGTCTGGTTCTGTTGGAAATGGA 840
HisSerHisSerIleGlnHisSerProGluArgSerGlySerGlySerValGlyAsnGly

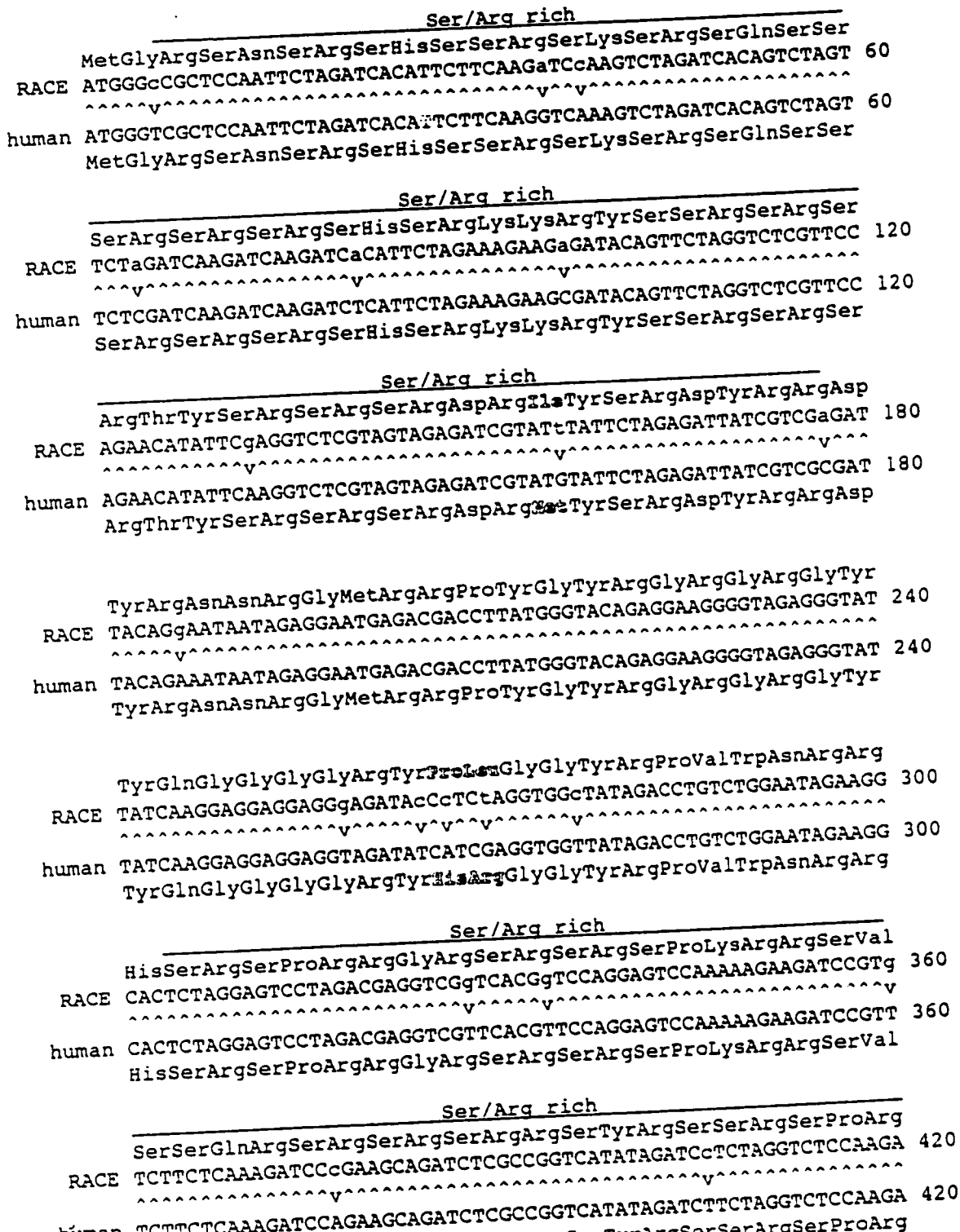

Ser/Arg rich
MetGlyArgSerAsnSerArgSerHisSerSerArgSerLysSerArgSerGlnSerSer
RACE ATGGGcCGCTCCAATTCTAGATCACATTCTTCAAGaTCcAAGTCTAGATCACAGTCTAGT 60
human ATGGGTCGCTCCAATTCTAGATCACATTCTTCAAGGTCAAAGTCTAGATCACAGTCTAGT 60
MetGlyArgSerAsnSerArgSerHisSerSerArgSerLysSerArgSerGlnSerSer
Ser/Arg rich
SerArgSerArgSerArgSerHisSerArgLysLysArgTyrSerSerArgSerArgSer
RACE TCTaGATCAAGATCAAGATCaCATTCTAGAAAGAAGaGATACAGTTCTAGGTCTCGTTCC 120
human TCTCGATCAAGATCAAGATCTCATTCTAGAAAGAAGCGATACAGTTCTAGGTCTCGTTCC 120
SerArgSerArgSerArgSerHisSerArgLysLysArgTyrSerSerArgSerArgSer
Ser/Arg rich
ArgThrTyrSerArgSerArgSerArgAspArgIleTyrSerArgAspTyrArgArgAsp
RACE AGAACATATTCgAGGTCTCGTAGTAGAGATCGTATtTATTCTAGAGATTATCGTCGaGAT 180
human AGAACATATTCAAGGTCTCGTAGTAGAGATCGTATGTATTCTAGAGATTATCGTCGCGAT 180
ArgThrTyrSerArgSerArgSerArgAspArgIleTyrSerArgAspTyrArgArgAsp
TyrArgAsnAsnArgGlyMetArgArgProTyrGlyTyrArgGlyArgGlyArgGlyTyr
RACE TACAGgAATAATAGAGGAATGAGACGACCTTATGGGTACAGAGGAAGGGGTAGAGGGTAT 240
human TACAGAAATAATAGAGGAATGAGACGACCTTATGGGTACAGAGGAAGGGGTAGAGGGTAT 240
TyrArgAsnAsnArgGlyMetArgArgProTyrGlyTyrArgGlyArgGlyArgGlyTyr
TyrGlnGlyGlyGlyGlyArgTyrProLeuGlyGlyTyrArgProValTrpAsnArgArg
RACE TATCAAGGAGGAGGAGGgAGATAcCcTCtAGGTGGcTATAGACCTGTCTGGAATAGAAGG 300
human TATCAAGGAGGAGGAGGTAGATATCATCGAGGTGGTTATAGACCTGTCTGGAATAGAAGG 300
TyrGlnGlyGlyGlyGlyArgTyrHisArgGlyGlyTyrArgProValTrpAsnArgArg
Ser/Arg rich
HisSerArgSerProArgArgGlyArgSerArgSerArgSerProLysArgArgSerVal
RACE CACTCTAGGAGTCCTAGACGAGGTCGgTCACGgTCCAGGAGTCCAAAAAGAAGATCCGTg 360
human CACTCTAGGAGTCCTAGACGAGGTCGTTCACGTTCCAGGAGTCCAAAAAGAAGATCCGTT 360
HisSerArgSerProArgArgGlyArgSerArgSerArgSerProLysArgArgSerVal
Ser/Arg rich
SerSerGlnArgSerArgSerArgSerArgArgSerTyrArgSerSerArgSerProArg
RACE TCTTCTCAAAGATCCcGAAGCAGATCTCGCCGGTCATATAGATCCcTCTAGGTCTCCAAGA 420
human TCTTCTCAAAGATCCAGAAGCAGATCTCGCCGGTCATATAGATCTTCTAGGTCTCCAAGA 420
SerSerGlnArgSerArgSerArgSerArgArgSerTyrArgSerSerArgSerProArg

Figure II.11. The alignment of nucleotide and protein sequences of the human and mouse KIAA0164. The human KIAA0164 cDNA nucleotide sequence from Genbank is labelled "human" and protein sequence for human KIAA0164 is below it. The nucleotide sequence of the mouse KIAA0164 cDNA is located above the nucleotide sequence of the human cDNA.

The complete mouse cDNA was created by ligating the 5' cDNA generated by RACE to the cDNA insert of the plasmid 5.7. The sequence of the 5' RACE product is labelled "RACE". The sequence of the cDNA insert of 5.7 is labelled 5.7 and an arrow above the sequence indicates where the cDNA sequence of 5.7 started.

The sequence for the cDNA insert in the plasmid D18 was identical to a corresponding sequence of the cDNA insert of 5.7 that is indicated with (.....) above the nucleotide sequence of the cDNA insert of 5.7.

The differences in nucleotide sequence between the mouse KIAA0164 and human KIAA0164 are indicated with a " v " beneath the differing nucleotides of the 5.7 clone. Consensus nucleotides are indicated with the symbol, " ^ " . Protein sequence differences between mouse and human are indicated by amino acids that are in the outline style (Met). The numbers on the right side refer the number of nucleotides in the human KIAA0164 and mouse KIAA0164 starting from the initiator ATG. The line above the sequence indicates the location of Ser/Arg rich regions in the human protein. Three-letter amino acid codes are used.

```
==================================6.12==================================
      Gln Met Arg Gln Leu Ser Val Ile Pro Pro Met Met Phe Asp Ala Glu Gln Arg
NCoRß CAA ATG CGT CAG CTT TCT GTG ATT CCA CCT ATG ATG TTT GAT GCA GAA CAA AGA 1242
NCoR  CAA ATG CGT CAG CTT TCT GTG ATT CCA CCT ATG ATG TTT GAT GCA GAA CAA AGA 1239
      Gln Met Arg Gln Leu Ser Val Ile Pro Pro Met Met Phe Asp Ala Glu Gln Arg

==================================6.12==================================
      Arg Val Lys Phe Ile Asn Met Asn Gly Leu Met Glu Asp Pro Met Lys Val Tyr
NCoRß AGG GTC AAA TTC ATC AAT ATG AAT GGG CTG ATG GAG GAT CCA ATG AAG GTT TAT 1296
 NCoR AGG GTC AAA TTC ATC AAT ATG AAT GGG CTG ATG GAG GAT CCA ATG AAG GTT TAT 1293
      Arg Val Lys Phe Ile Asn Met Asn Gly Leu Met Glu Asp Pro Met Lys Val Tyr

==================================6.12==================================
      Lys Asp Arg Gln Phe Met Asn Val Trp Thr Asp His Glu Lys Glu Ile Phe Lys
NCoRß AAA GAC AGA CAG TTT ATG AAT GTT TGG ACT GAC CAT GAA AAG GAG ATC TTT AAG 1350
NCoR  AAA GAC AGA CAG TTT ATG AAT GTT TGG ACT GAC CAT GAA AAG GAG ATC TTT AAG 1347
      Lys Asp Arg Gln Phe Met Asn Val Trp Thr Asp His Glu Lys Glu Ile Phe Lys

==================================6.12==================================
      Asp Lys Phe Ile Gln His Pro Lys Asn Phe Gly Leu Ile Ala Ser Tyr Leu Glu
NCoRß GAC AAG TTT ATC CAG CAT CCA AAA AAC TTT GGA CTA ATT GCA TCC TAT TTG GAA 1404
NCoR  GAC AAG TTT ATC CAG CAT CCA AAA AAC TTT GGA CTA ATT GCA TCC TAT TTG GAA 1401
      Asp Lys Phe Ile Gln His Pro Lys Asn Phe Gly Leu Ile Ala Ser Tyr Leu Glu

==================================6.12==================================
      Arg Lys Ser Val Pro Asp Cys Val Leu Tyr Tyr Tyr Leu Thr Lys Lys Asn Glu
NCoRß AGG AAG AGT GTT CCT GAT TGT GTT TTA TAT TAC TAT TTA ACC AAG AAA AAT GAG 1458
NCoR  AGG AAG AGT GTT CCT GAT TGT GTT TTA TAT TAC TAT TTA ACC AAG AAA AAT GAG 1455
      Arg Lys Ser Val Pro Asp Cys Val Leu Tyr Tyr Tyr Leu Thr Lys Lys Asn Glu

==================================6.12==================================
      Asn Tyr Lys Ala Leu Val Arg Arg Asn Tyr Gly Lys Arg Arg Gly Arg Asn Gln
NCoRß AAT TAT AAG GCC CTC GTG AGA AGG AAT TAT GGA AAA CGC AGA GGC AGA AAT CAG 1512
 NCoR AAT TAT AAG GCC CTC GTG AGA AGG AAT TAT GGA AAA CGC AGA GGC AGA AAT CAG 1509
      Asn Tyr Lys Ala Leu Val Arg Arg Asn Tyr Gly Lys Arg Arg Gly Arg Asn Gln

==================================6.12==================================
      Gln Ile Ala Arg Pro Ser Gln Glu Glu Lys Val Glu Glu Lys Glu Glu Asp Lys
NCoRß CAG ATT GCC CGT CCC TCA CAA GAA GAA AAA GTA GAA GAA AAG GAA GAG GAT AAA G 1567
NCoR  CAG ATT GCC CGT CCC TCA CAA GAA GAA AAA GTA GAA GAA AAG GAA GAG GAT AAA G 1564
      Gln Ile Ala Arg Pro Ser Gln Glu Glu Lys Val Glu Glu Lys Glu Glu Asp Lys
```

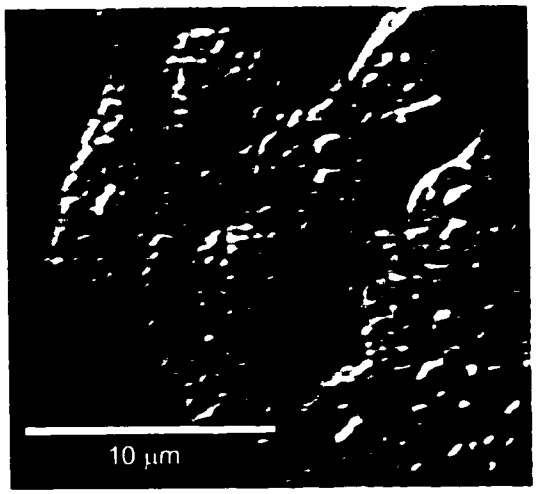
```
      ================================6.12=================================
      Lys Ile Phe Glu Gly Leu Gly Pro Lys Val Glu Leu Pro Leu Tyr Asn Gln Pro
NCoRß AAA ATA TTT GAA GGT CTT GGC CCA AAA GTT GAA CTG CCa CTC TAT AAC CAG CCA 810
 NCoR AAA ATA TTT GAA GGT CTT GGC CCA AAA GTT GAA CTG CCg CTC TAC AAC CAG CCG 807
      Lys Ile Phe Glu Gly Leu Gly Pro Lys Val Glu Leu Pro Leu Tyr Asn Gln Pro
      ----------------------Repressor Domain I----------------------


      ================================6.12=================================
      Ser Asp Thr Lys Val Tyr His Glu Asn Ile Lys Thr Asn Gln Val Met Arg Lys
NCoRß TCA GAT ACC AAG GTG TAC CAt GAG AAC ATC AAG ACA AAC CAG GTG ATG AGG AAA 864
NCoR  TCA GAT ACC AAG GTG TAC CAc GAG AAC ATC AAG ACA AAC CAG GTG ATG AGG AAA 861
      Ser Asp Thr Lys Val Tyr His Glu Asn Ile Lys Thr Asn Gln Val Met Arg Lys
      ----------------------Repressor Domain I----------------------


      ================================6.12=================================
      Lys Leu Ile Leu Phe Phe Lys Arg Arg Asn His Ala Arg Lys Gln Arg Glu Gln
NCoRß AAA CTC ATT TTA TTT TTT AAA AGA AGA AAT CAT GCA AGA AAA CAA AGG GAA CAA 918
NCoR  AAA CTC ATT TTA TTT TTT AAA AGA AGA AAT CAT GCA AGA AAA CAA AGG GAA CAA 915
      Lys Leu Ile Leu Phe Phe Lys Arg Arg Asn His Ala Arg Lys Gln Arg Glu Gln
      ----------------------Repressor Domain I----------------------


      ================================6.12=================================
      Lys Ile Cys Gln Arg Tyr Asp Gln Leu Met Glu Ala Trp Glu Lys Lys Val Asp
NCoRß AAA ATC TGC CAG CGC TAT GAT CAG CTC ATG GAA GCG TGG GAG AAA AAA GTG GAC 972
NCoR  AAA ATC TGC CAA CGT TAT GAT CAG CTC ATG GAA GCA TGG GAG AAA AAA GTG GAC 969
      Lys Ile Cys Gln Arg Tyr Asp Gln Leu Met Glu Ala Trp Glu Lys Lys Val Asp
      ----Repressor DomainI----


      ================================6.12=================================
      Arg Ile Glu Asn Asn Pro Arg Arg Lys Ala Lys Glu Ser Lys Pro Arg Glu Tyr
NCoRß AGA ATA GAA AAT AAT CCc CGG AGG AAA GCa AAA GAA AGC AAA cCA AGG GAA TAC 1026
NCoR  AGA ATA GAA AAT AAT CCt CGG AGG AAA GCt AAA GAA AGC AAA aCA AGG GAA TAC 1023
      Arg Ile Glu Asn Asn Pro Arg Arg Lys Ala Lys Glu Ser Lys Thr Arg Glu Tyr


      ================================6.12=================================
      Tyr Glu Lys Gln Phe Pro Glu Ile Arg Lys Gln Arg Glu Gln Gln Glu Arg Phe
NCoRß TAT GAg AAG CAG TTT CCA GAA ATT CGA AAA CAA AGA GAA CAG CAA GAA AGA TTT 1080
NCoR  TAT GAa AAG CAG TTT CCA GAA ATT CGA AAA CAA AGA GAA CAG CAA GAA AGA TTT 1077
      Tyr Glu Lys Gln Phe Pro Glu Ile Arg Lys Gln Arg Glu Gln Gln Glu Arg Phe


      ================================6.12=================================
      Gln Arg Val Gly Gln Arg Gly Ala Gly Leu Ser Ala Thr Ile Ala Arg Ser Glu
NCoRß CAG CGA GTT GGT CAG AGG GGA GCT GGT CTT TCA GCC ACC ATT GCT AGG AGT GAG 1134
NCoR  CAG CGA GTT GGT CAG AGG GGA GCT GGT CTT TCA GCC ACC ATT GCT AGG AGT GAG 1131
      Gln Arg Val Gly Gln Arg Gly Ala Gly Leu Ser Ala Thr Ile Ala Arg Ser Glu


      ================================6.12=================================
      His Glu Ile Ser Glu Ile Ile Asp Gly Leu Ser Glu Gln Glu Asn Asn Glu Lys
NCoRß CAT GAG ATT TCT GAA ATT ATT GAT GGT CTT TCT GAA CAG GAG AAT AAT GAG AAG 1188
NCoR  CAT GAG ATT TCT GAA ATT ATT GAT GGT CTT TCT GAA CAG GAG AAT AAT GAG AAG 1185
      His Glu Ile Ser Glu Ile Ile Asp Gly Leu Ser Glu Gln Glu Asn Asn Glu Lys
```

```
      ..............................RT-PCR.................................
      ==================================6.12=================================
      Val Leu Pro Leu Val His Ser Leu Pro Glu Gly Leu Arg Ser Ser Ala Arg Ala
NCoRß GTt tTa CCT TTa GTt CAC tCG CTG CCA GAA GGc tTG AGG TCg TCT GCa gAT GCT 432
NCoR  GTc cTc CCT TTg GTg CAC aCG CTG CCA GAA GGa cTG AGG TCt TCT GCc aAT GCT 429
      Val Leu Pro Leu Val His Ser Leu Pro Glu Gly Leu Arg Ser Ser Ala Asn Ala
      ------------------------Repressor Domain I-------------------------

      ..............................RT-PCR.................................
      ==================================6.12=================================
      Lys Lys Asp Ser Ala Phe Gly Ser Lys His Glu Ala Pro Ser Ser Pro Leu Ala
NCoRß AAG AAG GAT tCa GCA TTT GGA agC AAA CAT GAA GCT CCa TCC TCT CCt TTg gCT 486
NCoR  AAG AAG GAT cCg GCA TTT GGA gtC AAA CAT GAA GCT CCt TCC TCT CCc TTc tCT 483
      Lys Lys Asp Pro Ala Phe Gly Val Lys His Glu Ala Pro Ser Ser Pro Leu Ser
      ------------------------Repressor Domain I-----------------------

      ..............................RT-PCR.................................
      ==================================6.12=================================
      Gly Gln Pro Cys Gly Asp Asp Gln Asn Ala Ser Pro Ser Lys Leu Ser Lys Glu
NCoRß GGG CAa CCA TGT GGA GAT GAc CAA AAT GCT TCA CCT TCA AAG CTT TCA AAG GAg 540
NCoR  GGG CAg CCA TGC GGA GAT GAt CAG AAT GCC TCA CCT TCA AAA CTG TCA AAG GAa 537
      Gly Gln Pro Cys Gly Asp Asp Gln Asn Ala Ser Pro Ser Lys Leu Ser Lys Glu
      ------------------------Repressor Domain I-----------------------

      ....RT-PCR.....
      ==================================6.12=================================
      Glu Leu Ile Gln Ser Met Asp Arg Val Asp Arg Glu Ile Ala Lys Val Glu Gln
NCoRß GAG TTa ATA CAG AGt ATG GAc CGG GTa GAc CGA GAg ATT GCa AAA GTA GAg CAG 594
NCoR  GAG CTg ATA CAG AGc ATG GAt CGT GTc GAt CGA GAa ATT GCg AAA GTA GAa CAG 591
      Glu Leu Ile Gln Ser Met Asp Arg Val Asp Arg Glu Ile Ala Lys Val Glu Gln
      ------------------------Repressor Domain I-----------------------

      ==================================6.12=================================
      Gln Ile Leu Lys Leu Lys Lys Lys Gln Gln Gln Leu Glu Glu Glu Ala Ala Lys
NcoRß CAG ATC CTT AAA tTG AAA AAG AAa CAa CAA CAG CTa GAA GAA GAA GCa GCT Aag 648
NCoR  CAG ATC CTT AAA cTG AAA AAG AAg CAg CAA CAG CTc GAA GAA GAA GCt GCT Aaa 645
      Gln Ile Leu Lys Leu Lys Lys Lys Gln Gln Gln Leu Glu Glu Glu Ala Ala Lys
      ------------------------Repressor Domain I-----------------------

      ==================================6.12=================================
      Pro Pro Glu Pro Glu Lys Pro Val Ser Pro Pro Pro Val Glu Gln Lys His Arg
NCoRß CCC CCt GAa CCa GAG AAG CCT GTG TCC CCT CCT CCT GTG GAa CAa AAa CAC CGt 702
NCoR  CCC CCa GAg CCt GAG AAG CCT GTG TCC CCT CCT CCC GTG GAg CAg AAg CAC Cga 699
      Pro Pro Glu Pro Glu Lys Pro Val Ser Pro Pro Pro Val Glu Gln Lys His Arg
      ------------------------Repressor Domain I-----------------------

      ==================================6.12=================================
      Ser Ile Val Gln Ile Ile Tyr Asp Glu Asn Arg Lys Lys Ala Glu Glu Ala His
NcoRß AGT ATT GTC CAA ATT ATT TAT GAt GAG AAT CGG AAA AAA GCA GAA GAA GCT CAT 756
NCoR  AGT ATT GTC CAA ATC ATT TAT GAc GAG AAT CGG AAA AAA GCA GAA GAA GCT CAT 753
      Ser Ile Val Gln Ile Ile Tyr Asp Glu Asn Arg Lys Lys Ala Glu Glu Ala His
      ------------------------Repressor Domain I-----------------------
```

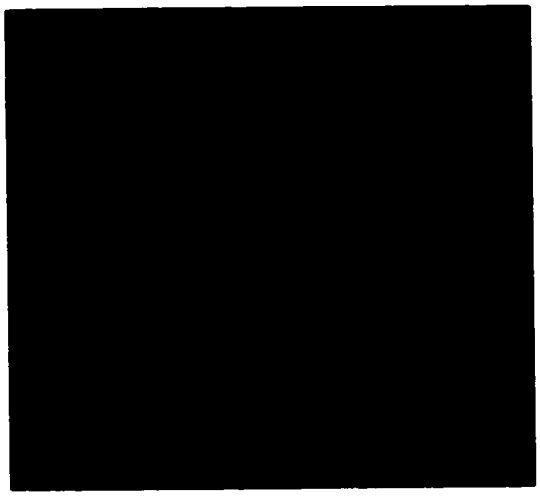
```
.....RT-PCR.....
     Met Ser Ser Ser Gly Tyr Pro Pro Asn Gln Gly Ala Phe Ser Thr Glu Gln Ser
NCoRß ATG TCA AGT TCA GGT TAT CCT CCC AAC CAa GGa GCt TTC AGC ACA GAG CAa AGT 54
NCoR  ATG TCA AGT TCA GGT TAT CCT CCC AAC CAg GGg GCg TTC AGC ACA GAG CAg AGT 54
     Met Ser Ser Ser Gly Tyr Pro Pro Asn Gln Gly Ala Phe Ser Thr Glu Gln Ser
     ----------Repressor Domain I----------

.....RT-PCR.....
     Arg Tyr Pro Ser His Ser Val Gln Tyr Thr Phe Pro Ser Thr Arg His Gln Gln
NCoRß CGt TAt CCT TCa CAt TCt GTc CAG TAt ACC TTT CCC AGt aCC CGA CAC CAG CAG 108
NCoR  CGc TAc CCT TCg CAc TCg GTt CAG TAc ACC TTT CCC AGc gCC CGT CAC CAG CAG 108
     Arg Tyr Pro Ser His Ser Val Gln Tyr Thr Phe Pro Ser Ala Arg His Gln Gln
     ----------Repressor Domain I----------

.....RT-PCR.....
     Glu Phe Ala Val Pro Asp Tyr Arg Ser Ser His Ile Glu Val Ser Gln Ala Ser
NCoRß GAA TTT GCA GTt CCT GAC TAC CGc TCT TCT CAT aTT GAA GTt AGc CAG GCa TCA 162
 NCoR GAA TTT GCA GTc CCT GAC TAC CGt TCT TCT CAT cTT GAA GTc AGt CAG GCg TCA 162
     Glu Phe Ala Val Pro Asp Tyr Arg Ser Ser His Leu Glu Val Ser Gln Ala Ser
     ----------Repressor Domain I----------

.....RT-PCR.....
     Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln Leu Arg Arg Arg Pro Ser Leu
NCoRß CAG CTt TTA CAG CAG CAG CAG CAG CAG CAG cag CTT CGa AGA CGa CCT TCC tTG 216
NCoR  CAG CTc TTG CAG CAG CAG CAG CAG CAG CAG --- CTT CGc AGA CGg CCT TCC cTG 213
     Gln Leu Leu Gln Gln Gln Gln Gln Gln Gln --- Leu Arg Arg Arg Pro Ser Leu
     ----------Repressor Domain I----------

.....RT-PCR.....
     Leu Ser Glu Phe His Pro Gly Ser Asp Arg Pro Gln Glu Arg Arg Thr Gly Tyr
NCoRß CTT TCa GAA TTT CAC CCG GGT TCt GAC AGG CCt CAa GAA AGG AGA AcT GGA TAt 270
NCoR  CTT TCc GAG TTT CAC CCG GGT TCc GAC AGG CCc CAg GAA AGG AGA AgT GGA TAc 267
     Leu Ser Glu Phe His Pro Gly Ser Asp Arg Pro Gln Glu Arg Arg Ser Gly Tyr
     ----------Repressor Domain I----------

.....RT-PCR.....
     Glu Gln Phe His Ser Gly Pro Ser Pro Val Asp His Asp Ser Leu Glu Ser Lys
NCoRß GAa CAG TTt CAC tCa GGA CCc TCa CCG GTG GAt CAT GAt TCC TTG GAG TCC AAG 324
NCoR  GAg CAG TTc CAC cCg GGC CCt TCc CCG GTG GAc CAT GAc TCG CTG GAG TCC AAG 321
     Glu Gln Phe His Pro Gly Pro Ser Pro Val Asp His Asp Ser Leu Glu Ser Lys
     ----------Repressor Domain I----------

.....RT-PCR.....
     ==========6.12==========
     Arg Pro Arg Leu Glu Pro Val Ser Asp Ala His Phe Gln Arg Val Ser Ala Ala
NCoRß CGC CCa CGC CTG GAG Cca GTT TCt GAT gCC CAt TTC CAG CGt gTt AGT GCT GCg 378
NCoR  CGG CCt CGC CTG GAG Cag GTT TCc GAC tCC CAc TTC CAG CGc aTc AGT GCT GCc 375
     Arg Pro Arg Leu Glu Gln Val Ser Asp Ser His Phe Gln Arg Ile Ser Ala Ala
     ----------Repressor Domain I----------
```

Figure II.12. The alignment of nucleotide and protein sequences of mouse N-CoRß from b.p. 1-1567 and mouse N-CoR from b.p. 1-1564. The partial nucleotide sequence of N-CoRß from b.p. 1-1567 is in rows labelled "N-CoRß" and the protein sequence for N-CoRß is located above it. The partial nucleotide sequence of N-CoR (b.p. 1-1564) is located below that of N-CoRß for comparison (the rest of the nucleotide sequence of N-CoR from b.p. 1564 -7780 b.p. is not shown). The nucleotides that differ between N-CoRß and N-CoR are bold lower-case letters (gatc). The consensus nucleotides are upper case letters (GATC). The differences in protein sequence are indicated by amino acids in the outline style (Met). The numbers on the right indicate the number of nucleotides in N-CoRß and N-CoR. The region of N-CoR that comprised Repressor domain I (~~~~) is indicated beneath the protein sequence for N-CoR. The double dashed lines (====) over the N-CoRß cDNA sequence is the sequence in the plasmid 6.12. The part the N-CoRß cDNA that was cloned by RT-PCR is indicated with dots over the sequence of N-CoRß (.....). The underlined N-CoRß sequence is a Pst I site present only in N-CoRß. This is also the site where the RT-PCR sequence was ligated to the cDNA sequence of 6.12

Cells



differential interference contrast



anti-his

his-tagged NCoRβ $_{1-1574}$

Figure II.13. His-NCoRβ transfected protein is nuclear. His-NCoRβ was transfected into 293T cells on coverslips and cells were fixed 48 hours later and stained with the anti-his antibody followed by the cy3 conjugated anti-mouse secondary antibody. The top panel is the differential contrast image of the field. The bottom panel is the immunofluorescence staining with the anti-his tag antibody followed by anti-mouse cy3. The immunofluorescence staining pattern of His-NCoRβ is similar to that for CHD-1. Bar = 10 μm.

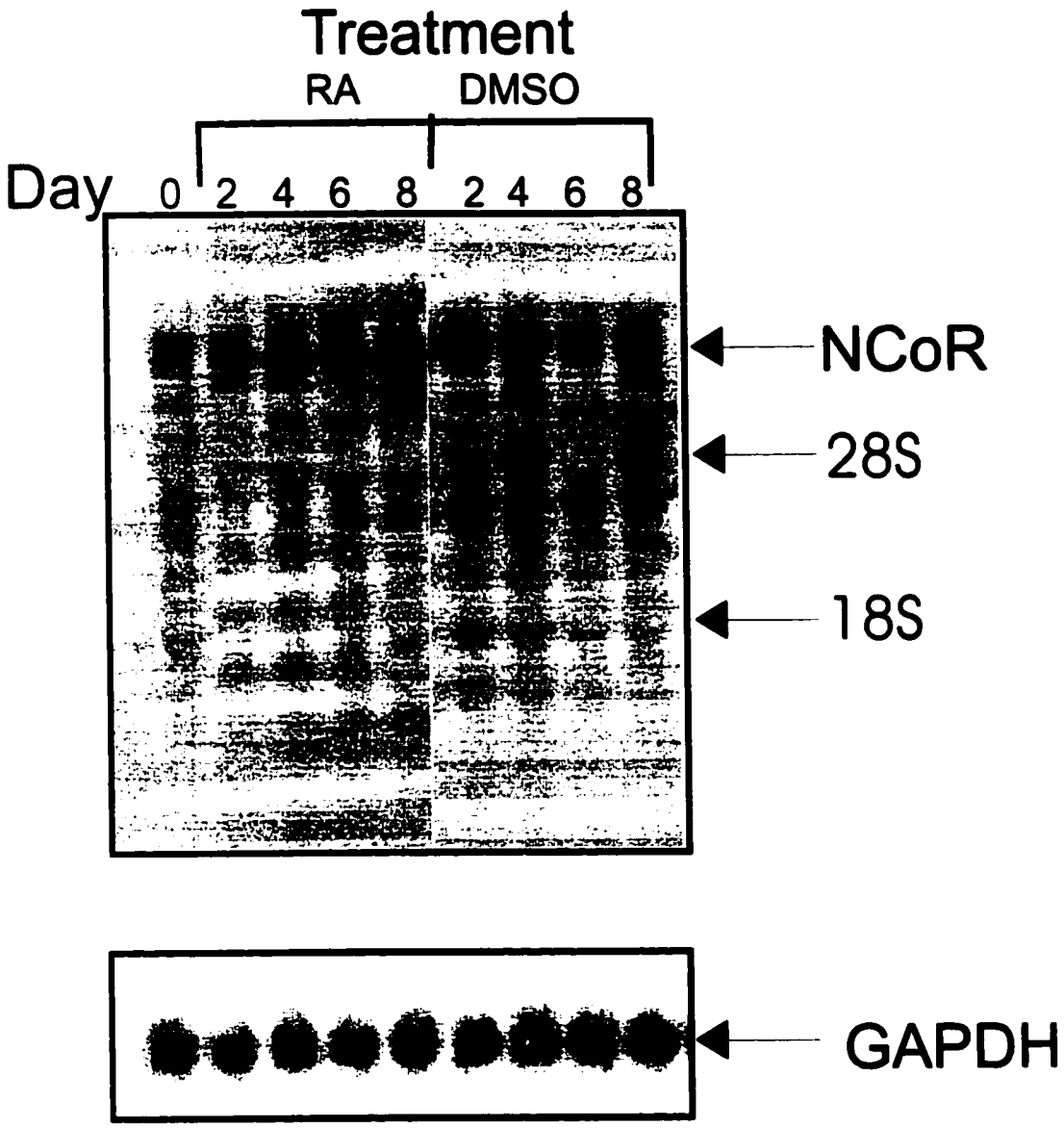
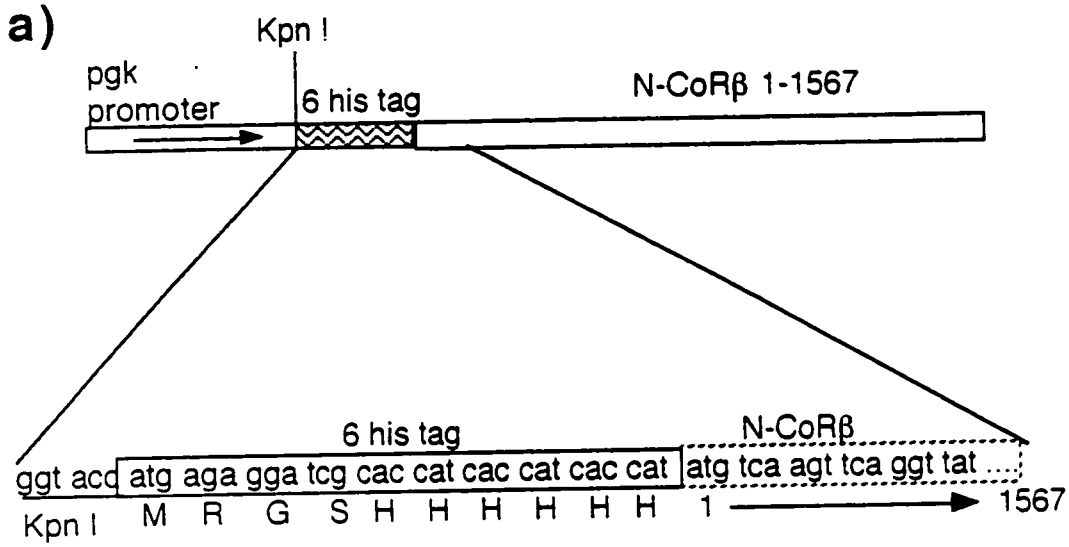

Treatment
RA
DMSO
Day
0 2 4 6 8 2 4 6 8
NCoR
28S
18S
GAPDH

Figure II.14. N-CoR is expressed in P19 cells and differentiated P19 cells. Northern blot of P19 cells and P19 cells differentiated with DMSO and retinoic acid (RA). The number of days that cells were treated with RA or DMSO is indicated at the top of the figure. The location of the full-length N-CoR mRNA is indicated by an arrow on the right. The arrows on the left indicate the location of 28S and 18S rRNA bands. The top box is the northern blot probed with N-CoRβ b.p. 1-336. The bottom box is the same northern probed with GAPDH as a loading control.


a)
Kpn I
pgk
promoter
6 his tag
N-CoRβ 1-1567
6 his tag
N-CoRβ
ggt acc atg aga gga tcg cac cat cac cat cac cat atg tca agt tca ggt tat ....
Kpn I M R G S H H H H H H 1
1567

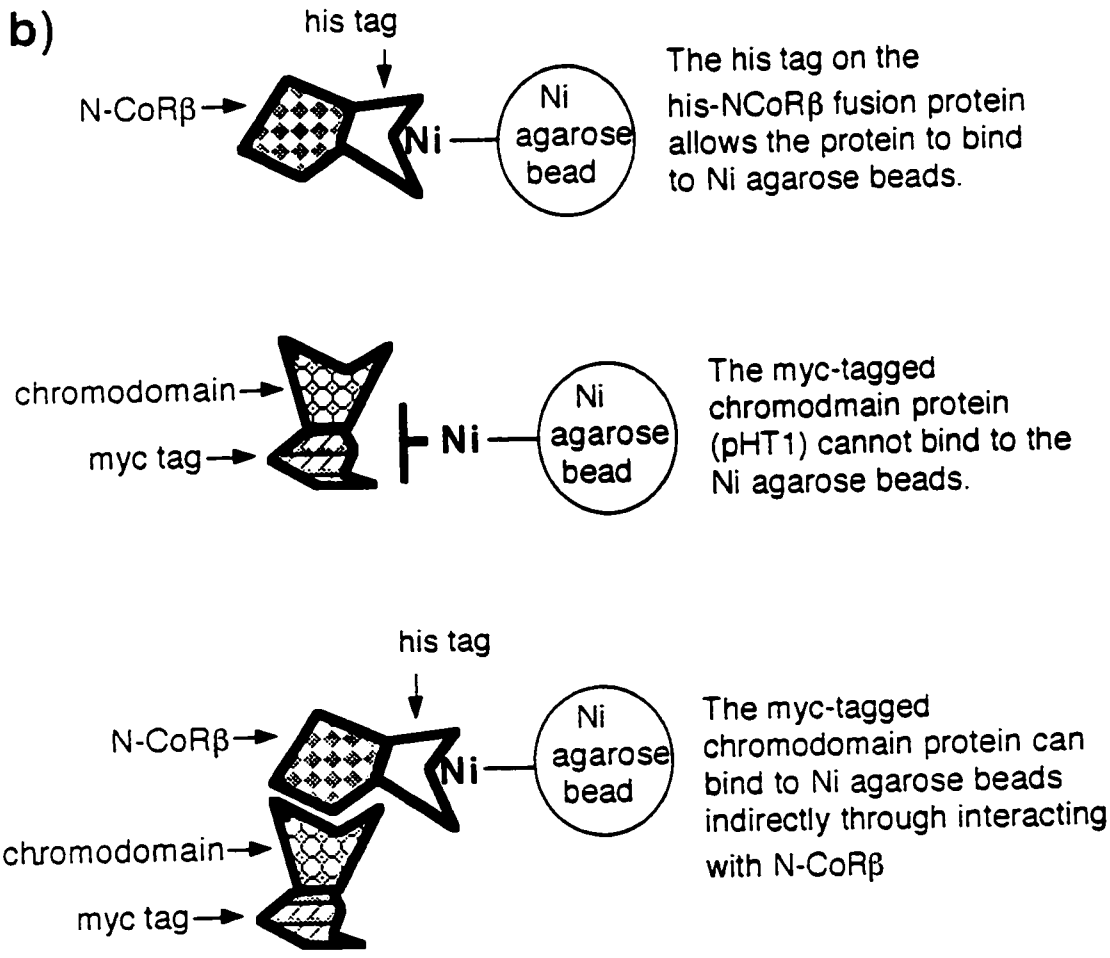

b)
his tag
N-CoRβ
Ni
Ni agarose bead
The his tag on the his-NCoRβ fusion protein allows the protein to bind to Ni agarose beads.
chromodomain
myc tag
Ni
Ni agarose bead
The myc-tagged chromodmain protein (pHT1) cannot bind to the Ni agarose beads.
his tag
N-CoRβ
Ni
Ni agarose bead
chromodomain
myc tag
The myc-tagged chromodomain protein can bind to Ni agarose beads indirectly through interacting with N-CoRβ

Figure II.15 a). Schematic diagram of the his-tagged N-CoRß. N-CoRß sequences in 6.12 contained only part of Repressor domain I. RT-PCR was used to amplify the 5' end of N-CoRß from the initiator Met to the internal Pst I site at bp 415. This site was used to clone the RTPCR product in frame with the N-CoRß sequence in 6.12. This resulted in a sequence of N-CoRß that went from bp 1-1567. The 5' primer used for RTPCR had a Kpn I site on the 5' end. This site was used to clone the N-CoRß into an expression vector driven by the pgk promoter. b) The in vitro protein binding assay using his-N-CoRß and Ni agarose beads.

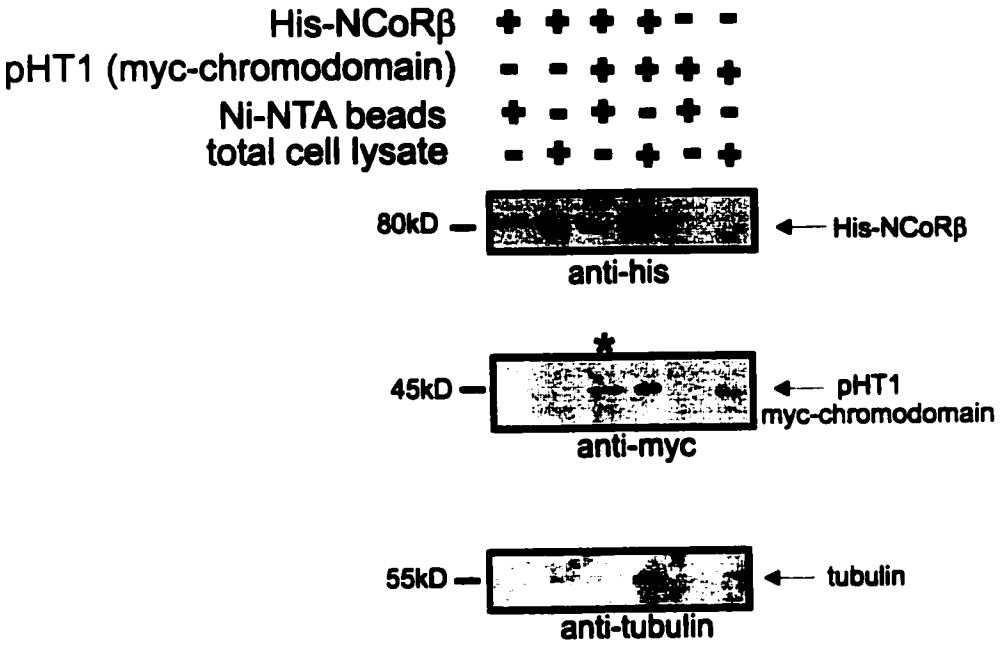
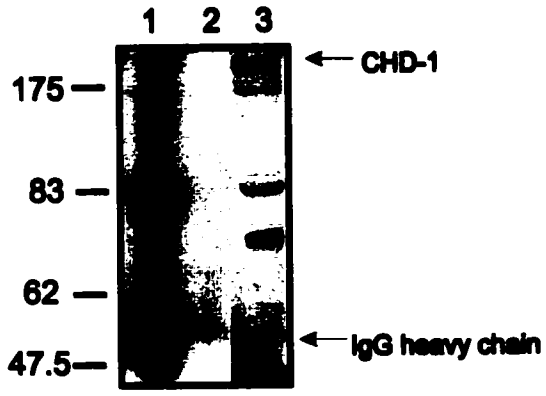

His-NCoRβ + + + + - -
pHT1 (myc-chromodomain) - - + + + +
Ni-NTA beads + - + - + -
total cell lysate - + - + - +
80kD
His-NCoRβ
anti-his
*
45kD
pHT1
myc-chromodomain
anti-myc
55kD
tubulin
anti-tubulin

Figure II.16. The chromodomain binds to N-CoRß *in vitro*. 293T cells were transfected with plasmids containing his-tagged N-CoRß (His-NcoRß) (lanes 1 and 2), myc-tagged chromodomain of CHD-1 (pHT1) (lanes 5 and 6), or both (lanes 3 and 4). Transfected cells were lysed and combined with Ni-NTA agarose beads, which bind to the his-tagged N-CoRß protein. Proteins binding to Ni-NTA agarose beads were then washed and eluted and run on 10% SDS-PAGE. "+" in the row beside His-NCoRß indicates that lane directly below the "+" was loaded with protein from cells transfected with His-NCoRß. "-" indicates that the lane was loaded with protein that was not transfected with His-NCoRß. The same is true for "+'s" and "-'s" beside myc-chromodomain (pHT1). The "+'s" next to Ni-NTA beads indicate that the lane below was loaded with protein that had been eluted from Ni-NTA agarose beads. The "+'s" next to whole cell lysate indicate that the lane below was loaded with protein from the whole cell lysate. The gel was blotted and probed with anti-his (top box), anti-myc (9E10) (middle box), and anti-tubulin (bottom box).

a)


1
2
3
CHD-1
175
83
62
47.5
IgG heavy chain

b)

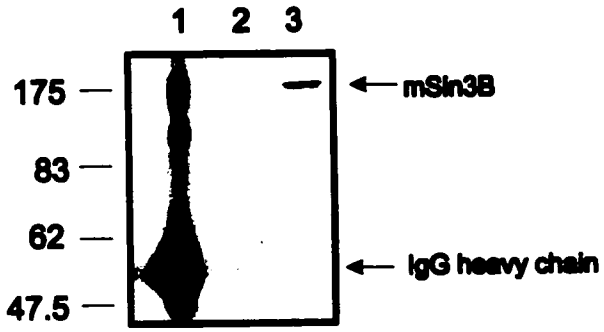

1
2
3
175
83
62
47.5
mSin3B
IgG heavy chain

Figure II.17. The endogenous CHD-1 and mSin3B proteins are associated in P19 cells. P19 cells were lysed and the anti-CHD-1 antibody was added to immunoprecipitate the endogenous CHD-1 protein and co-immunoprecipitate CHD-1 associated proteins. IgG cross-linked beads were used to bind precipitated proteins which were then run on 10% SDS-PAGE and blotted.

a) The immunoblot probed with anti-CHD-1. 1) Immunoprecipitate with anti-CHD-1; 2) IgG beads alone; 3) whole cell lysate of P19. The arrows indicates the location of the CHD-1 protein and also the heavy chain of IgG. The numbers on the left indicate the sizes and migration in kiloDaltons (kD) of the protein molecular weight markers.

b) The immunoblot probed with anti-mSin3B. 1, 2, 3 lanes are same as above. The arrows indicates the location of the mSin3B protein and the IgG heavy chain.

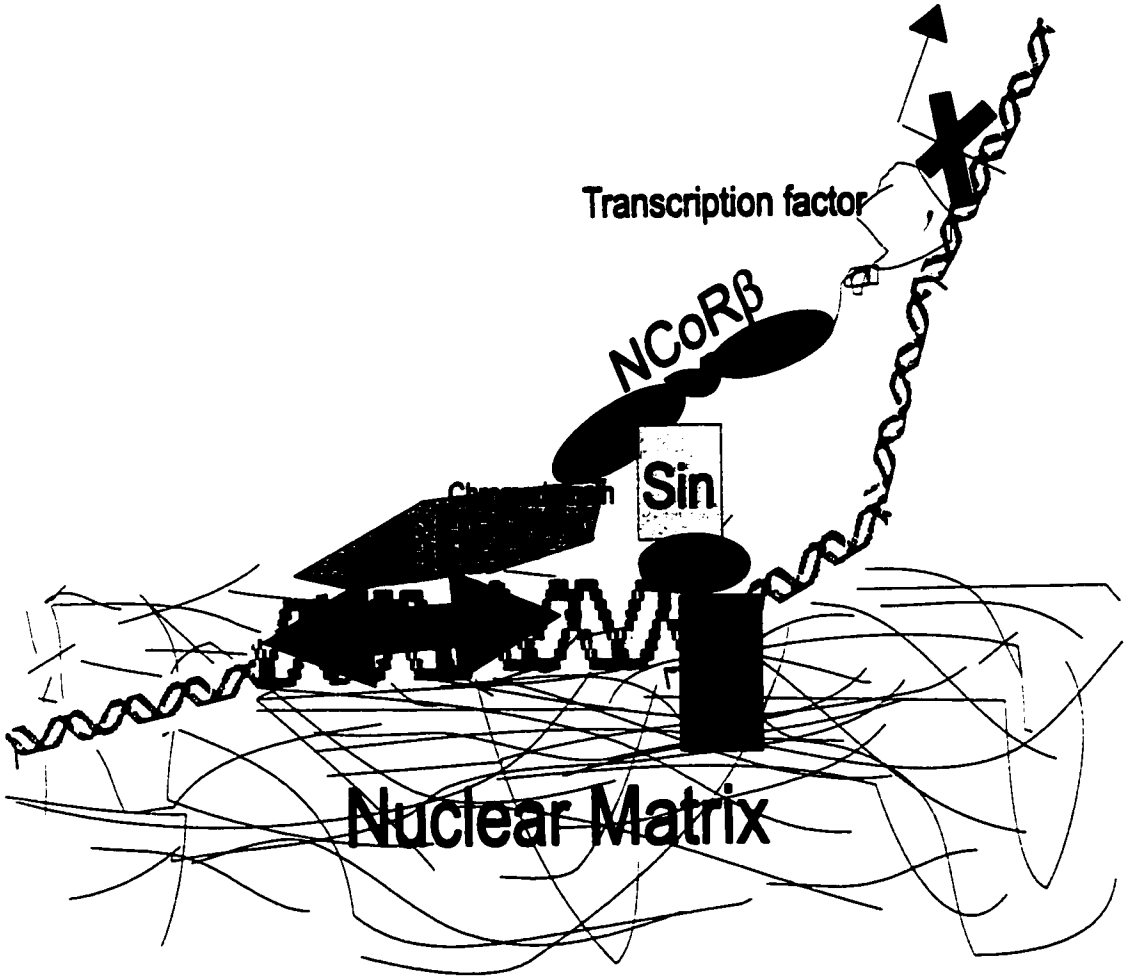

Transcription factor
NCoRβ
Sin
Nuclear Matrix

Figure II.18. Schematic diagram of a model for the function of CHD-1.

Discussion

Two novel discoveries are described in this chapter: A) the CHD-1 protein is associated with the nuclear matrix, B) the CHD-1 protein is associated with the Sin repressor complex.

A) The CHD-1 protein is associated with the nuclear matrix.

1) The endogenous CHD-1 protein in P19 cells

The results of Chapter II show that the endogenous CHD-1 protein is associated with the nuclear matrix in P19 cells, which is different from plasmacytoma cells where the CHD-1 protein was found in the soluble fraction rather than in the insoluble nuclear matrix fraction (109). However, these results do not necessarily contradict each other, since the composition of nuclear matrix proteins can differ between different cell types (41) (51). Nevertheless, our result concurs with the finding that CHD-1 binds to AT-rich DNA (109) which is similar to DNA in nuclear matrix attachment regions (MAR) (13) (16) (79). Hence, it is proposed that CHD-1 is a MAR DNA binding protein in the nuclear matrix.

MAR DNA binding proteins in the nuclear matrix have a number of functions. Recently, MeCP2 was identified as a MAR binding protein in the nuclear matrix (119). MeCP2 is a protein that binds to methylated CpG residues that are found in silenced genes (12), indicating that gene silencing is also part of the function of the nuclear matrix. In addition, HDAC 1(human histone deacetylase I) was

shown to bind to the nuclear matrix (106) and MeCp2 also associates with HDAC (87) which further indicates that the nuclear matrix can function in gene silencing. Therefore, the association of CHD-1 with the nuclear matrix is consistent with a gene silencing function for this protein.

2) Some mutant CHD-1 proteins containing deletions are also associated with the nuclear matrix

CHD-1 contained the following signature domains: two chromodomains, a helicase/ATPase domain, and a DNA binding domain. These domains were deleted progressively to generate mutant proteins that contained deletions of the DNA-binding domain, helicase/ATPase domain and the second chromodomain. One other mutant was constructed with a deletion of the amino-terminal end containing the chromodomains and part of the helicase/ATPase domain. The mutant CHD-1 proteins were all found to be localized to the nucleus. The amino acid sequence of CHD-1 indicates that there were seven possible sequences in CHD-1 that could function as nuclear localization signals (NLS) (55) and they are found at a.a. 138 (AKRKKH), 351 (KKK), 694 (RRVKK), 925 (RAKKK), 1104 (RKRPKKR), 1330 (KRRKTRAKK), and 1498 (KHAIKKR).

Are any of these mutant CHD-1 proteins also associated with the nuclear matrix ? To answer this question, the nuclear matrix was prepared from cells carrying transfected mutant CHD-1 proteins. The results showed that not all mutant proteins are associated with the nuclear matrix. The smallest mutant

protein showing nuclear matrix localizaton was the fragment of CHD-1 from aa 81-769 suggesting the presence of a nuclear matrix targeting signal within this sequence. The protein sequence within CHD-1 aa 81-769 contains the two chromodomains and the helicase/ATPase domain. However, the mutant CHD-1 protein aa 195-401 (pHT1) which encodes the first chromodomain but is missing the second chromodomain and the helicase domain, was not found to be associated with the nuclear matrix. In addition, the CHD-1 sequence contained in the plasmid, pHT15, encodes the second chromodomain and motifs I and II of the helicase/ATPase domain (a.a. 81-532) and this protein was also not found in the nuclear matrix. Furthermore, the rest of the helicase/ATPase domain, motifs III- VII, is present in CHD-1 aa 533-1711 which is contained in plasmid pHT12; and this protein was not found in the nuclear matrix, either. Hence, targeting the CHD-1 protein to the nuclear matrix may require that two or more regions within CHD-1 aa 81-769 mediate nuclear matrix attachment. Alternatively, an intact helicase/ATPase domain (motifs I - VI) may be required for nuclear matrix association.

Other proteins containing a homologous helicase/ATPase domain, BRG-1 and hbrm, were also found to be associated with the nuclear matrix (102) which supports the hypothesis that the helicase ATPase domain of CHD-1 is involved in targeting the protein to the nuclear matrix. However, the study of BRG-1 and hbrm did not specify if the helicase/ATPase domains of these proteins were required for nuclear matrix association.

Nuclear matrix targeting signals have been defined in the YY1 (77) and AML/CBF-α (125) protein sequences. The sequence encoding the nuclear matrix targeting signal of YY1 is different from that of AML/CBF-α . In addition, neither of these nuclear matrix targeting signals were found in CHD-1 aa 81-769 which is the region thought to contain the putative nuclear matrix targeting signal. Therefore, these nuclear matrix proteins are probably attached to different protein in the nuclear matrix rather than to a common structural target.

B) The CHD-1 protein is associated with the mSin repressor complex

1) The search for proteins that interact with the chromodomain of CHD-1

A function for CHD-1 might also emerge by identifying proteins associated with it. It was reasoned that if CHD-1 interacted with a particular protein then it was likely that CHD-1 has a role in the biological function of the interacting protein. Hence, a yeast two-hybrid screen was done to find proteins that interacted with the chromodomain of CHD-1.

The chromodomain was chosen as a target for interacting proteins because it is a domain associated with gene silencing in *Drosophila* (69). Hence, it is also thought that CHD-1 functions in gene silencing. The chromodomains of Pc and HP-1 have characteristics of a protein-protein interaction domain and can target the protein to sites chromatin. Therefore, it is predicted that the proteins that bind to the chromodomain of CHD-1 would be involved in gene silencing.

The chromodomain was found to interact with proteins encoded by cDNA

from a mouse embryo yeast two-hybrid library which confirmed that it could mediate protein interactions. Furthermore, the chromodomain of CHD-1 also interacted a protein with near identity with N-CoR, N-CoRβ. N-CoR is a co-repressor of transcription as the high degree of homology between the partial cDNA of N-CoRb and N-CoR indicates that N-CoRβ may also function in repression of gene expression. The cloning of the full-length cDNA of N-CoRβ will help confirm a function in repression. The association of CHD-1 with N-CoRβ and mSin3B indicates that CHD-1 is also involved in gene repression.

2) CHD-1 is associated with the Sin complex

N-CoR was of particular interest concerning CHD-1 because it was a co-repressor of nuclear hormone receptor mediated transcription which concurs with hypothesis that CHD-1 functions in gene silencing. Furthermore, N-CoR also bound the Sin complex which contains histone deacetylase. Since, deacetylated histones are also found in chromatin around silenced genes, an association between CHD-1 and the Sin complex is another indication that CHD-1 functions in gene silencing.

The interaction between the N-CoR and the chromodomain of CHD-1 was confirmed using an *in vitro* protein binding assay with lysates from cells containing tagged N-CoRβ and CHD-1 proteins. Additional experiments showed

that the endogenous CHD-1 protein was co-immunoprecipitated with the endogenous mSin3B protein a member of the Sin. This confirmed that CHD-1 was associated wit the Sin complex. Experiments in progress will determine if histone deacetylase also co-immunoprecipitates with CHD-1.

In addition, the expression of the transfected his-tagged N-CoRβ protein shows a punctuated nuclear staining like that found for CHD-1 confirming that the two proteins are co-localized.

The chromo shadow domain, which is similar to the chromodomain, is only contained in proteins with a chromodomain. The chromo shadow domain of human HP-1 and MoMOD were found to be associated with TIFa which can also repress transcription mediated by nuclear receptors (64). This is further evidence of the association of chromodomain proteins with repression of gene expression mediated by nuclear hormone receptors.

C) A model for the function of CHD-1

The CHD-4 protein, a- CHD-1 related protein, was also found to be associated with histone deacetylase in the NURD complex. Furthermore, chromatin remodelling activity of the NURD complex is attributed to CHD-4 which indicates that CHD-1 could also remodel chromatin.

A hypothetical model for the function of CHD-1 is presented in Figure II.18. A transcriptional repressor protein, such as N-CoR, recruits the Sin repressor complex to sites in chromatin to induce gene silencing. The CHD-1 protein as

part of the Sin complex anchors the chromatin to the nuclear matrix and also remodels chromatin to expose histone tails to the histone deacetylase in the complex. The deacetylation of histones promotes the formation of silenced chromatin. The association of the Sin complex with the nuclear matrix is also supported by studies that showed that HDAC 1 could be cross-linked to the nuclear matrix. Another protein associated with gene silencing MeCP2, which binds to methylated DNA that contains silenced genes, was also shown to be a nuclear matrix protein indicating further that the nuclear matrix is a site for gene silencing proteins. In addition, MeCp2 was also associated with histone deacetylase (87); therefore methylated DNA containing silenced genes is also a target for histone deacetylase which increases the evidence that histone deacetylation is part of a gene silencing mechanism. Finally it is suggested that CHD-1 is also associated with gene silencing through functioning in protein complexes that contain histone deacetylase.

References

1. Aasland R and Stewart AF (1995). The chromo shadow domain, a second chromo domain in heterochromatin- binding protein 1, HP1. Nucleic.Acids.Res. 23: 3168-3174.

2. Adra CN, Ellis NA, and McBurney MW (1988). The family of mouse phosphoglycerate kinase genes and pseudogenes. Somat Cell Mol Genet 14: 69-81.

3. Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ (1990). Basic local alignment search tool. J Mol.Biol. 215: 403-410.

4. Auble DT, Hansen KE, Mueller CG, Lane WS, Thorner J, and Hahn S (1994). Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. Genes Dev. 8: 1920-1934.

5. Ayer DE (1999). Histone deacetylases: transcriptional repression with SINers and NuRDs. Trends Cell Biol. 9: 193-198.

6. Baker BS and Belote JM (1983). Sex determination and dosage compensation in Drosophila melanogaster. Annu Rev Genet 17: 345-393.

7. Ball LJ, Murzina NV, Broadhurst RW, Raine AR, Archer SJ, Stott FJ, Murzin AG, Singh PB, Domaille PJ, and Laue ED (1997). Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. EMBO J 16: 2473-2481.

8. Bannister AJ and Kouzarides T (1996). The CBP co-activator is a histone acetyltransferase. Nature 384: 641-643.

9. Beard C, Li E, and Jaenisch R (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. Genes Dev 9: 2325-2334.

10. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, and Wheeler DL (GenBank. Nucleic Acids Res.2000.Jan.1.;28.(1.):15.-18. 28: 15-18.

11. Berezney R and Coffey DS (1974). Identification of a nuclear protein matrix. Biochem Biophys.Res.Commun. 60: 1410-1417.

12. Bird AP and Wolffe AP (1999). Methylation-induced repression--belts, braces, and chromatin. Cell 99: 451-454.

13. Bode J, Schlake T, Rios-Ramirez M, Mielke C, Stengert M, Kay V, and

Klehr-Wirth D (1995). Scaffold/matrix-attached regions: structural properties creating transcriptionally active loci. Int.Rev.Cytol. 162A:389-454: 389-454.

14. Boer PH, Potten H, Adra CN, Jardine K, Mullhofer G, and McBurney MW (1990). Polymorphisms in the coding and noncoding regions of murine Pgk-1 alleles. Biochem Genet 28: 299-308.

15. Borsani G, Tonlorenzi R, Simmler MC, Dandolo L, Arnaud D, Capra V, Grompe M, Pizzuti A, Muzny D, and Lawrence C (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature 351: 325-329.

16. Boulikas T (1993). Nature of DNA sequences at the attachment regions of genes to the nuclear matrix. J Cell Biochem 52: 14-22.

17. Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, and Rastan S (1991). Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature 351: 329-331.

18. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, and Rastan S (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 71: 515-526.

19. Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, and Willard HF (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349: 38-44.

20. Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, and Willard HF (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell 71: 527-542.

21. Brown CJ, Lafreniere RG, Powers VE, Sebastio G, Ballabio A, Pettigrew AL, Ledbetter DH, Levy E, Craig IW, and Willard HF (1991). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature 349: 82-84.

22. Brown S (1991). Xist and the mapping of the X chromosome inactivation center. Bioessays 13: 607-612.

23. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, and Allis CD (1996). Tetrahymena histone acetyltransferase A: a homolog

to yeast Gcn5p linking histone acetylation to gene activation. Cell 84: 843-851.

24. Buetler TM and Eaton DL (1992). Complementary DNA cloning, messenger RNA expression, and induction of alpha-class glutathione S-transferases in mouse tissues. Cancer Res. 52: 314-318.

25. Chaly N, Little JE, and Brown DL (1985). Localization of nuclear antigens during preparation of nuclear matrices in situ. Can.J Biochem.Cell Biol. 63: 644-653.

26. Chapman VM, Kratzer PG, and Quarantillo BA (1983). Electrophoretic variation for X chromosome-linked hypoxanthine phosphoribosyl transferase (HPRT) in wild-derived mice. Genetics 103: 785-795.

27. Chirgwin JM, Przybyla AE, MacDonald RJ, and Rutter WJ (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294-5299.

28. Clemson CM, McNeil JA, Willard HF, and Lawrence JB (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J Cell Biol 132: 259-275.

29. Cockerill PN and Garrard WT (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell 44: 273-282.

30. Colwill K, Pawson T, Andrews B, Prasad J, Manley JL, Bell JC, and Duncan PI (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J 15: 265-275.

31. Comings DE and Okada TA (1976). Nuclear proteins. III. The fibrillar nature of the nuclear matrix. Exp.Cell Res. 103: 341-360.

32. Cote J, Quinn J, Workman JL, and Peterson CL (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. Science 265: 53-60.

33. Cowell IG and Austin CA (1997). Self-association of chromo domain peptides. Biochim.Biophys.Acta 1337: 198-206.

34. de B, I, Cai S, and Kohwi-Shigematsu T (1998). The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. J Cell Biol. 141: 335-348.

35. Debrand E, Chureau C, Arnaud D, Avner P, and Heard E (1999).

Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. Mol Cell Biol 19: 8513-8525.

36. Delmas V, Stokes DG, and Perry RP (1993). A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. Proc.Natl.Acad.Sci U.S.A. 90: 2414-2418.

37. Durfee T, Becherer K, Chen PL, Yeh SH, Yang Y, Kilburn AE, Lee WH, and Elledge SJ (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes Dev. 7: 555-569.

38. Eissenberg JC, Morris GD, Reuter G, and Hartnett T (1992). The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. Genetics 131: 345-352.

39. Epstein CJ, Smith S, Travis B, and Tucker G (1978). Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. Nature 274: 500-503.

40. Fackelmayer FO, Dahm K, Renz A, Ramsperger U, and Richter A (1994). Nucleic-acid-binding properties of hnRNP-U/SAF-A, a nuclear-matrix protein which binds DNA and RNA in vivo and in vitro. Eur.J Biochem 221: 749-757.

41. Fey EG and Penman S (1988). Nuclear matrix proteins reflect cell type of origin in cultured human cells. Proc.Natl.Acad.Sci U.S.A. 85: 121-125.

42. Forrester WC, van Genderen C, Jenuwein T, and Grosschedl R (1994). Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. Science 265: 1221-1225.

43. Gartler SM and Riggs AD (1983). Mammalian X-chromosome inactivation. Annu Rev Genet 17: 155-190.

44. Grant SG and Chapman VM (1988). Mechanisms of X-chromosome regulation. Annu Rev Genet 22: 199-233.

45. Harper JW, Adami GR, Wei N, Keyomarsi K, and Elledge SJ (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75: 805-816.

46. Hassan AB, Errington RJ, White NS, Jackson DA, and Cook PR (1994). Replication and transcription sites are colocalized in human cells. J Cell Sci 107: 425-434.

47. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, and Ayer DE (1997).

Histone deacetylase activity is required for full transcriptional repression by mSin3A. Cell 89: 341-347.

48. Heinzel T, Lavinsky RM, Mullen TM, Soderstrom M, Laherty CD, Torchia J, Yang WM, Brard G, Ngo SD, Davie JR, Seto E, Eisenman RN, Rose DW, Glass CK, and Rosenfeld MG (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression [see comments]. Nature 387: 43-48.

49. Herman R, Weymouth L, and Penman S (1978). Heterogeneous nuclear RNA-protein fibers in chromatin-depleted nuclei. J Cell Biol. 78: 663-674.

50. Hodgkin J (1987). Sex determination and dosage compensation in Caenorhabditis elegans. Annu Rev Genet 21: 133-154.

51. Holth LT, Chadee DN, Spencer VA, Samuel SK, Safneck JR, and Davie JR (1998). Chromatin, nuclear matrix and the cytoskeleton: role of cell structure in neoplastic transformation (review). Int.J Oncol 13: 827-837.

52. Horlein AJ, Naar AM, Heinzel T, Torchia J, Gloss B, Kurokawa R, Ryan A, Kamei Y, Soderstrom M, and Glass CK (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor [see comments]. Nature 377: 397-404.

53. Jones-Villeneuve EM, Rudnicki MA, Harris JF, and McBurney MW (1983). Retinoic acid-induced neural differentiation of embryonal carcinoma cells. Mol.Cell Biol. 3: 2271-2279.

54. Jurgens G (1985). A group of genes controlling the spatial expressinof the bithorax complex in Drosophila. Nature 316: 153-155.

55. Kalderon D, Richardson WD, Markham AF, and Smith AE (1984). Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature 311: 33-38.

56. Kasten MM, Dorland S, and Stillman DJ (1997). A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators. Mol.Cell Biol. 17: 4852-4858.

57. Kay GF, Barton SC, Surani MA, and Rastan S (1994). Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell 77: 639-650.

58. Kay GF, Penny GD, Patel D, Ashworth A, Brockdorff N, and Rastan S (1993). Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. Cell 72: 171-182.

59. Kelley DE, Stokes DG, and Perry RP (1999). CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin. Chromosoma 108: 10-25.

60. Kratzer PG and Gartler SM (1978). HGPRT activity changes in preimplantation mouse embryos. Nature 274: 503-504.

61. Lacalle RA, Pulido D, Vara J, Zalacain M, and Jimenez A (1989). Molecular analysis of the pac gene encoding a puromycin N-acetyl transferase from Streptomyces alboniger. Gene 79: 375-380.

62. Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

63. Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, and Eisenman RN (1997). Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell 89: 349-356.

64. Le Douarin B, Nielsen AL, Garnier JM, Ichinose H, Jeanmougin F, Losson R, and Chambon P (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J 15: 6701-6715.

65. Lee JT, Davidow LS, and Warshawsky D (1999). Tsix, a gene antisense to Xist at the X-inactivation centre [see comments]. Nat Genet 21: 400-404.

66. Lu T, Van Dyke M, and Sawadogo M (1993). Protein-protein interaction studies using immobilized oligohistidine fusion proteins. Anal.Biochem 213: 318-322.

67. Luderus ME, den Blaauwen JL, de Smit OJ, Compton DA, and van Driel R (1994). Binding of matrix attachment regions to lamin polymers involves single-stranded regions and the minor groove. Mol.Cell Biol. 14: 6297-6305.

68. Lyon M (1961). Gene action in the X chromosome o fhte mouse (Mus musculus). Nature 190: 372-373.

69. Martin GR, Epstein CJ, Travis B, Tucker G, Yatziv S, Martin DWJ, Clift S, and Cohen S (1978). X-chromosome inactivation during differentiation of female teratocarcinoma stem cells in vitro. Nature 271: 329-333.

70. McBurney MW (1988). X chromosome inactivation: a hypothesis. Bioessays 9: 85-88.

71. McBurney MW and Adamson ED (1976). Studies on the activity of the X

chromosomes in female teratocarcinoma cells in culture. Cell 9: 57-70.

72. McBurney MW, Jones-Villeneuve EM, Edwards MK, and Anderson PJ (1982). Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. Nature 299: 165-167.

73. McBurney MW and Rogers BJ (1982). Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Dev.Biol. 89: 503-508.

74. McBurney MW and Strutt BJ (1980). Genetic activity of X chromosomes in pluripotent female teratocarcinoma cells and their differentiated progeny. Cell 21: 357-364.

75. McBurney MW, Sutherland LC, Adra CN, Leclair B, Rudnicki MA, and Jardine K (1991). The mouse Pgk-1 gene promoter contains an upstream activator sequence. Nucleic Acids Res. 19: 5755-5761.

76. McCarrey JR and Dilworth DD (1992). Expression of Xist in mouse germ cells correlates with X-chromosome inactivation. Nat Genet 2: 200-203.

77. McNeil S, Guo B, Stein JL, Lian JB, Bushmeyer S, Seto E, Atchison ML, Penman S, van Wijnen AJ, and Stein GS (1998). Targeting of the YY1 transcription factor to the nucleolus and the nuclear matrix in situ: the C-terminus is a principal determinant for nuclear trafficking. J Cell Biochem 68: 500-510.

78. Messmer S, Franke A, and Paro R (1992). Analysis of the functional role of the Polycomb chromo domain in Drosophila melanogaster. Genes Dev. 6: 1241-1254.

79. Mirkovitch J, Mirault ME, and Laemmli UK (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell 39: 223-232.

80. Misteli T, Caceres JF, and Spector DL (1997). The dynamics of a pre-mRNA splicing factor in living cells. Nature 387: 523-527.

81. Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y, and Allis CD (1996). The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell 87: 1261-1270.

82. Monk M and Harper MI (1979). Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. Nature 281: 311-313.

83. Mortillaro MJ, Blencowe BJ, Wei X, Nakayasu H, Du L, Warren SL, Sharp PA, and Berezney R (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. Proc.Natl.Acad.Sci U.S.A. 93: 8253-8257.

84. Nagase T, Ishikawa K, Kikuno R, Hirosawa M, Nomura N, and Ohara O (1999). Prediction of the coding sequences of unidentified human genes. XV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro [In Process Citation]. DNA Res. 6: 337-345.

85. Nagase T, Seki N, Ishikawa K, Tanaka A, and Nomura N (1996). Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by analysis of cDNA clones from human cell line KG-1. DNA Res. 3: 17-24.

86. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, and Evans RM (1997). Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell 89: 373-380.

87. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, and Bird A (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex [see comments]. Nature 393: 386-389.

88. Oesterreich S, Lee AV, Sullivan TM, Samuel SK, Davie JR, and Fuqua SA (1997). Novel nuclear matrix protein HET binds to and influences activity of the HSP27 promoter in human breast cancer cells. J Cell Biochem 67: 275-286.

89. Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, and Hamada H (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell 60: 461-472.

90. Okumura K, Sakaguchi G, Naito K, Tamura T, and Igarashi H (1997). HUB1, a novel Kruppel type zinc finger protein, represses the human T cell leukemia virus type I long terminal repeat-mediated expression. Nucleic Acids Res. 25: 5025-5032.

91. Owen-Hughes T, Utley RT, Cote J, Peterson CL, and Workman JL (1996). Persistent site-specific remodeling of a nucleosome array by transient action of the SWI/SNF complex. Science 273: 513-516.

92. Panning B, Dausman J, and Jaenisch R (1997). X chromosome inactivation is mediated by Xist RNA stabilization. Cell 90: 907-916.

93. Papaioannou VE and West JD (1981). Relationship between the parental origin of the X chromosomes, embryonic cell lineage and X chromosome expression in mice. Genet Res 37: 183-197.

94. Parkhurst SM, Bopp D, and Ish-Horowicz D (1990). X:A ratio, the primary sex-determining signal in Drosophila, is transduced by helix-loop-helix proteins [published erratum appears in Cell 1991 Mar 8;64(5):following 1046]. Cell 63: 1179-1191.

95. Paro R and Hogness DS (1991). The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. Proc.Natl.Acad.Sci U.S.A. 88: 263-267.

96. Paterno GD and McBurney MW (1985). X chromosome inactivation during induced differentiation of a female mouse embryonal carcinoma cell line. J Cell Sci 75: 149-163.

97. Peterson CL and Herskowitz I (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell 68: 573-583.

98. Pirrotta V (1997). PcG complexes and chromatin silencing. Curr.Opin.Genet.Dev. 7: 249-258.

99. Platero JS, Hartnett T, and Eissenberg JC (1995). Functional analysis of the chromo domain of HP1. EMBO J 14: 3977-3986.

100. Rastan S (1982). Timing of X-chromosome inactivation in postimplantation mouse embryos. J Embryol Exp Morphol 71: 11-24.

101. Rastan S (1983). Non-random X-chromosome inactivation in mouse X-autosome translocation embryos--location of the inactivation centre. J Embryol Exp Morphol 78: 1-22.

102. Reyes JC, Muchardt C, and Yaniv M (1997). Components of the human SWI/SNF complex are enriched in active chromatin and are associated with the nuclear matrix. J Cell Biol. 137: 263-274.

103. Richler C, Soreq H, and Wahrman J (1992). X inactivation in mammalian testis is correlated with inactive X-specific transcription. Nat Genet 2: 192-195.

104. Rivier DH and Rine J (1992). Silencing: the establishment and inheritance of stable, repressed transcription states. Curr.Opin.Genet.Dev. 2: 286-292.

105. Salido EC, Yen PH, Mohandas TK, and Shapiro LJ (1992). Expression of

the X-inactivation-associated gene XIST during spermatogenesis. Nat Genet 2: 196-199.

106. Samuel SK, Spencer VA, Bajno L, Sun JM, Holth LT, Oesterreich S, and Davie JR (1998). In situ cross-linking by cisplatin of nuclear matrix-bound transcription factors to nuclear DNA of human breast cancer cells. Cancer Res. 58: 3004-3008.

107. Sheardown SA, Duthie SM, Johnston CM, Newall AE, Formstone EJ, Arkell RM, Nesterova TB, Alghisi GC, Rastan S, and Brockdorff N (1997). Stabilization of Xist RNA mediates initiation of X chromosome inactivation [see comments]. Cell 91: 99-107.

108. Singh PB, Miller JR, Pearce J, Kothary R, Burton RD, Paro R, James TC, and Gaunt SJ (1991). A sequence motif found in a Drosophila heterochromatin protein is conserved in animals and plants. Nucleic Acids Res. 19: 789-794.

109. Stokes DG and Perry RP (1995). DNA-binding and chromatin localization properties of CHD1. Mol.Cell Biol. 15: 2745-2753.

110. Tai HH, Gordon J, and McBurney MW (1994). Xist is expressed in female embryonal carcinoma cells with two active X chromosomes. Somat Cell Mol Genet 20: 171-182.

111. Takagi N and Sasaki M (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature 256: 640-642.

112. Taunton J, Hassig CA, and Schreiber SL (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p [see comments]. Science 272: 408-411.

113. Tong JK, Hassig CA, Schnitzler GR, Kingston RE, and Schreiber SL (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature 395: 917-921.

114. Towbin H, Staehelin T, and Gordon J (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc.Natl.Acad.Sci U.S.A. 76: 4350-4354.

115. Turner BM (1993). Decoding the nucleosome. Cell 75: 5-8.

116. Vasseur M, Duprey P, Brulet P, and Jacob F (1985). One gene and one pseudogene for the cytokeratin endo A. Proc Natl Acad Sci U S A 82: 1155-1159.

117. Vidal M and Gaber RF (1991). RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in Saccharomyces cerevisiae. Mol.Cell Biol. 11: 6317-6327.

118. Warshawsky D, Stavropoulos N, and Lee JT (1999). Further examination of the Xist promoter-switch hypothesis in X inactivation: evidence against the existence and function of a P(0) promoter. Proc Natl Acad Sci U S A 96: 14424-14429.

119. Weitzel JM, Buhrmester H, and Stratling WH (1997). Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. Mol.Cell Biol. 17: 5656-5666.

120. West JD, Frels WI, Chapman VM, and Papaioannou VE (1977). Preferential expression of the maternally derived X chromosome in the mouse yolk sac. Cell 12: 873-882.

121. Whitehouse I, Flaus A, Cairns BR, White MF, Workman JL, and Owen-Hughes T (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. Nature 400: 784-787.

122. Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, and Gough NM (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336: 684-687.

123. Xing Y, Johnson CV, Moen PTJ, McNeil JA, and Lawrence J (1995). Nonrandom gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains. J Cell Biol. 131: 1635-1647.

124. Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, and Nakatani Y (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature 382: 319-324.

125. Zeng C, van Wijnen AJ, Stein JL, Meyers S, Sun W, Shopland L, Lawrence JB, Penman S, Lian JB, Stein GS, and Hiebert SW (1997). Identification of a nuclear matrix targeting signal in the leukemia and bone-related AML/CBF-alpha transcription factors. Proc.Natl.Acad.Sci U.S.A. 94: 6746-6751.

126. Zhang Y, LeRoy G, Seelig HP, Lane WS, and Reinberg D (1998). The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell 95: 279-289.