Spring 4-2006

Identification of CHD2 Nuclear Localization Signals

Alisha Gail Coffey

University of Tennessee-Knoxville

Follow this and additional works at: https://trace.tennessee.edu/utk_chanhonoproj

Recommended Citation


This is brought to you for free and open access by the University of Tennessee Honors Program at Trace: Tennessee Research and Creative Exchange. It has been accepted for inclusion in University of Tennessee Honors Thesis Projects by an authorized administrator of Trace: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Identification of CHD2 Nuclear Localization Signals

A Seniors Honors Project
In Partial Fulfillment of
Bachelor of Science with University Honors in
Biological Sciences: Biochemistry and Cellular and Molecular Biology
The University of Tennessee, Knoxville

Alisha G. Coffey
April 2006

Senior Project Advisor: Dr. Sundaresan Venkatachalam
Assistant Professor
Biochemistry and Cellular and Molecular Biology
ABSTRACT

According to the American Cancer Society, cancer is a group of diseases characterized by uncontrolled growth and spread of abnormal cells. The genetic alterations that cause such uncontrolled growth can disrupt the proper functioning of affected organ systems and, ultimately, can result in death. It has been suggested in recent studies that chromatin remodeling machinery may play an important role in the prevention of these diseases (1). The CHD2 gene (Chromo-domain Helicase DNA binding protein 2) is a protein thought to play an essential role in the process of chromatin remodeling, and this gene is present in a chromosomal region that is frequently lost in various cancers. Thus, in an effort to understand the role of CHD2 deficiency in cancer formation, we have generated a mutant mouse model that is deficient for Chd2 expression. Preliminary data indicate that the Chd2 deficient mice are susceptible to cancer formation. Our latest research involves the identification of the region of the CHD2 protein that allows it to enter the nucleus, where chromatin remodeling takes place. Computational analysis of the protein indicates the presence of four putative nuclear localization signals (NLS) that may be responsible for this entry into the nucleus. Three of these signals are located in the C-terminus end of the protein while the fourth is located near the N-terminus of the protein. We generated a CHD2 expression clone that contains only the N-terminus signal, and an HA-tag has been attached. The expression of this protein fragment was analyzed by immuno-fluorescence and found to localize within the nucleus with only the first putative NLS.
BACKGROUND

Chromatin Structure and Remodeling

Gene regulation is an essential ability in living cells, and one way this is achieved in eukaryotic cells is with the chromatin structure of DNA. The DNA found in a eukaryotic chromosome is found in association with histones and other small proteins in a complex referred to as chromatin. Five types of histones are present in chromatin, and they are designated H1, H2A, H2B, H3, and H4 (2).

Figure 1. The packaging of DNA in a eukaryotic cell from “beads on a string” to chromosome form.

Chromatin contains repeating structural units often described as “beads on a string” known as nucleosomes which contain approximately 1.8 turns of DNA around a core particle of histone proteins. The core is composed of an octamer of histones: two each of H2A, H2B, H3, and H4 histones. About 146 base pairs are tightly bound to this
core while another 20 base pairs are associated with the H1 histone (2). The DNA is then condensed further and tightly packed into chromosome form (see Figure 1).

DNA remains associated with these proteins through electrostatic forces between the negatively charged phosphate groups in the DNA backbone and the positively charged amino acids, such as lysine and arginine, found in the amino terminal tail of the histone proteins. These proteins can be covalently modified through the addition of acetyl, methyl, or phosphate groups which, in turn, affect the affinity of the histones for the DNA (3-4). These modifications are important to biological processes including DNA replication, gene expression, chromatin assembly, condensation, and cell division. Histone variants as well as non-histone chromosomal proteins such as high mobility group proteins can also be incorporated to affect both chromatin structure and activity (5). These characteristics are involved in chromatin remodeling which allow certain factors to interact with the DNA and can be used in transcription, replication, repair, recombination, and nucleosome assembly (6-17). The exact mechanism of chromatin remodeling is not known, and this is an active and ongoing field of research.

Chromodomain Helicase DNA Binding Proteins

One group of proteins thought to be important in this process of chromatin remodeling is the chromodomain helicase DNA binding (CHD) proteins. This family of proteins was characterized in the late 1990’s (18-20). The presence of at least eight members of the CHD family in mammals suggests that these proteins might have diverse functional roles in higher eukaryotes (21, 22). Several conserved domains among the proteins point to roles in chromatin remodeling and transcriptional control. Common
domains among the proteins include the chromodomain (Chromatin organization modifier), SNF2-related ATP-dependent helicase domain, DNA binding domains, and the C-terminal helicase domain.

The chromodomain portion can be used to self associate and interacts with the heterochromatic regions at centromeres, telomeres and polytene chromosomes (23, 24). It is thought that the helicase regions play a role in the unwinding of DNA and in destabilizing protein-DNA interactions (21, 13). Both CHD3 and CHD4 have provided information on this family of proteins as they were isolated from the nucleosome remodeling and histone deacetylation complex (NuRD) in Hela cells which is a complex that has been implicated in playing a major role in the organization of higher order chromatin complexes and gene repression (25, 27, 28).

**CHD1/CHD2 Sub-family**

CHD1 and CHD2 belong to a sub-family within the CHD proteins. They share a high degree of homology within their chromodomains and their Myb-related DNA-binding domains, and unlike other CHD proteins, they also contain an HMG-1 domain that is present in high mobility group (HMG) proteins (29). The Myb DNA binding domain (DBD) is well conserved, and several proteins with Myb-related DBD’s have been implicated in the regulation of cell proliferation, differentiation, and apoptosis (30). HGM-1 domains (A+T hooks) bind to DNA and chromatin and induce short and long range changes in the structure of their binding sites (31). These domains are thought to be involved in transcription, replication, recombination, and repair, and they are known to
play a role in preventing cancer formation due to these activities and their effect on cellular proliferation (32-34).

CHD1 has recently been shown to associate with active chromatin by binding to histone 3 (H3) methylated on lysine 4 (H3K4) which has been known to activate transcription (34, 35). The H3K4 binding motif (KWxxLxY) found in CHD1 is also seen in CHD2 which has approximately 60% similarity with CHD1 at the protein level (34, 35). This motif is not seen in CHD3 and CHD4 which have been implicated in transcriptional repression (36).

**CHD2 Structure and Function**

Few studies have focused on the role of CHD2; thus, its biological role is not yet known. The human gene encoding CHD2 is mapped to chromosome 15q26.2 which is a region that has been associated with rare genetic disorders that lead to growth retardation, cardiac defects and early post natal lethality (37,38). The 99% similarity found between the chromodomain and the Myb DNA binding domain of CHD2 with that of CHD1 suggests that the protein could be involved in the transcriptional regulation of target genes. The presence of the Myb DNA binding domain specifically suggests that CHD2 is involved in the regulation of genes with Myb-DNA binding sequences. Furthermore, deregulation of Myb protein has been seen in human and murine lymphoid cancers (39,40). This leads to the hypothesis that CHD2 may be an important protein involved in preventing cancer. This is supported by its possible role in DNA repair due to the presence of the HGM-1 domain, as well as a role in genomic stability and transcriptional
regulation as indicated by the helicase domains, chromodomains, and Myb DNA binding domain. A representation of the protein motifs found in CHD2 can be seen in Figure 2.

**CHD2 Nuclear Localization Signals**

In addition to these domains, the CHD2 protein contains four sites that could possibly be used as bipartite nuclear localization signals (NLS). These NLS’s are clusters of residues that signal the active transport of nuclear proteins into the nucleus by binding to molecules such as importins and karyopherins that recognize these distinct targeting signals. NLS’s with a bipartite motif are comprised of two sets of basic residues separated by 9-12 residues (41). Three of these sites are found near the C-terminus of the protein while the fourth is found close to the N-terminus of the protein. The goal of our current experiments has been to determine which sequence is used as the authentic nuclear localization signal. The nuclear localization is necessary to the proper
functioning of the protein because it must first enter the nucleus in order to interact with the chromatin.

**Generation of CHD2 Deficient Mice**

In order to further study the role of CHD2 in tumor suppression in a mammalian model, we generated a mutant mouse model for Chd2 using the Baygenomics genetrap ES cell resource (42, 43). One of the ES cell clones that had been characterized to have a genetrap insertion within the *Chd2* gene was represented in the Baygenomics ES cell library. The *Chd2* trapped ES cells were obtained from Baygenomics and analyzed by PCR to confirm Chd2 disruption by using primers that were specific for *Chd2* and the gene-trap sequences. Sequencing of the PCR product indicated that the gene trap has indeed integrated within intron 27 (1563 base pairs from the beginning of the intron) of the *Chd2* gene. The insertion of the gene-trap results in the loss of 595 amino acids at the C-terminus and results in the formation of a Chd2-β-gal-neomycin fusion gene. Figure 2 shows the insertion site of the gene-trap vector within the *Chd2* gene and the RFLP fragments generated due to the insertion. The ES cells confirmed for Chd2 deletion were used for blastocyst injections using the micro-injection services at the University of Massachusetts Medical School, Worcester and the subsequent generation of chimeric founder mice. Out of the seven high degree chimeras obtained from the blastocyst injections, analysis of the first three litters from all founder males indicated that two of the chimeric founders produced agouti germ line litters for Chd2 deletion. The colonies from the two founders were expanded further for the analysis of the mutant offspring.
Figure 3. Genotype analysis of Chd2 mutant animals and cell lines. Schematic representation of wild type and trapped Chd2 alleles (top panel) is shown above. The expected sizes of the genomic DNA fragments obtained from Xba1 digestion are indicated along with the probe (dark rectangle) used for the southern blot analysis. Southern blot analysis of genomic DNA isolated from representative mouse embryonic fibroblasts (MEF) and mouse tails is shown in the middle panel along with the genotypes and the sizes of DNA fragments.

Analysis of the mutant Chd2-fusion protein localization

At the protein level, the insertion of the gene trap was determined to be downstream of the HMG-1 domain that leads to the loss of 585 amino-acids at the C-terminus of the Chd2 protein. Analysis of the Chd2 protein indicated that the insertion of the gene trap also results in the loss of 3 out of the 4 putative nuclear localization signals. This further indicated that the disruption of Chd2 will lead to the absence of the protein in the nucleus. To confirm the absence of the Chd2 protein in the nucleus, we used
antibodies against the β-galactosidase and found that the Chd2-β-gal-neomycin fusion gene product was localized only in the cytoplasm Figure 3).

Figure 4. Absence of Chd2-β-gal-neomycin fusion protein in the nuclei of Chd2 mutant MEFs. FITC-conjugated antibodies specific for β-galactosidase were used for detecting the fusion protein. Please note the exposure times were the same for both the cell lines that were grown on the same chamber slide.

Although this procedure confirmed that our mutant gene was not entering the nucleus, we did not know if this was because the putative NLS at the N-terminal end was not functioning as a nuclear localization signal in the absence of the other three NLS’s or if the size of the Chd2-β-gal-neomycin fusion gene was serving to block the gene from entering the nucleus. Our first step in testing the authenticity of this NLS site was to make a clone of the first 447 amino acids which contains only the first NLS sequence in addition to a hemagglutinin (HA) tag, transfecting these clones into a mammalian cell line, and analyzing the clone with immuno-fluorescence. We will also attach a FLAG-tag to the clone as well as further clones of varying sizes, and western blot analysis will be performed. This combined information will then be used to determine which NLS is (are) the authentic sequence(s) used by the protein for entry into the cell. This will provide more information on the CHD2 gene, and it will allow further research to be
conducted on the putative sequences that are not actually used as nuclear localization signals.

MATERIALS AND METHODS

*Generation of Chd2 gene fragments in mammalian expression vector*

To obtain clones of the Chd2 protein, the cDNA sequence of human CHD2 was blasted with mouse Chd2 and found to have approximately 94% homology. A complete clone of Chd2 was not available; however, a Chd2 clone, M2021, available from Open Biosystems was found to show 99.9% homology to the first 520 base pairs (bp) of the human CHD2. Another clone, M5007, contained 3430bp of Chd2 with the first 80bp of this section overlapping with the last 80bp of the M2021 clone. The M2021 clone was available in the PT7T3D-PacI bacterial vector, and the M5007 had been cloned in the pSport6 mammalian vector. In order to begin our experiments, we first set out to place both segments in the same mammalian expression vector, pSport6, in order to express the peptide in mammalian cells. First, M2021 was cut from PT7T3D-PacI with restriction enzymes XhoI and NotI. Then, M5007 was cut from pSport6 using SalI and NotI. Digestion with XhoI and SalI results in compatible ends. M2021 was then ligated into the empty pSport6 vector and transformed into JM109 competent cells, and the resulting clone was labeled pSP2021. Ligation was confirmed by performing a restriction digest on the miniprep colonies with EcoRV and NotI.
Generation of N-terminal tagged peptide fragment of Chd2

In order to perform analysis with immuno-fluorescence and western blotting, we attached an HA-tag to the M5007 clone using tag specific primers. The primers HAX and Pvu2f were used for this process, and the resulting PCR product, 5007-HA tag, was approximately 950bp. This product was then PCR purified and digested with XbaI overnight. The clone pSp2021 was also digested overnight with XbaI and with 1:10 diluted PvuII for 90 minutes. The digested 5007-HA and pSp2021 were then ligated and transformed into JM109 competent cells, resulting in the HA-tagged N-terminal peptide clone herein referred to as lklHA. Colonies were picked, and the resulting minipreps were digested with XbaI to confirm the success of the ligation. Further confirmation and determination of the orientation of the Chd2 clone within the vector was gained through sequencing. A schematic representation of the N-terminal HA-tagged peptide is shown in figure 5.

Figure 5. Structural motifs seen in the N-terminal HA-tagged peptide fragment in comparison to the full length CHD2 protein.
**Expression of HA-tagged N-terminal peptide in mammalian cells**

The 1kHA clone was then transfected into a mammalian cell line (U2OS) via lipofectamine-mediated DNA transfer, and the cells were grown on glass chamber slides. After 24 hours, the cells were fixed with 1:1 acetone and methanol (10 min), washed with Tris buffered saline and blocked for one hour in 10% antibody dilution buffer (ADB) from the stock solution containing 3% BSA, 10% goat serum, 0.05% Triton X-100 in PBS pH 7.4. The primary anti-HA antibody (1:1000 dilution in ADB; rabbit) was then added, and the slides were incubated in a humidified chamber for four hours at room temperature. The slides were then washed three times with ADB buffer and incubated for one hour at room temperature with fluorescein isothiocyanate (FITC)-conjugated goat anti mouse antibody (1:500 dilution). The slides were washed with PBS containing 0.2% photoflo for ten minutes and PBS containing 0.05% triton X-100 for ten minutes and, this was followed by two washes with water and 0.2% photoflo for five minutes each. Finally, the DNA was counterstained with 4'-6-Diamidino-2-phenylindole (DAPI-0.1ug/ml sigma), and the slides were mounted with anti-fade reagent overnight. Images of the cells were captured using a Zeiss fluorescent microscope.

**RESULTS**

Analysis of the expression patterns of HA-tagged N-terminal region of Chd2 indicated that the N-terminal NLS was sufficient for the expression of the Chd2 protein fragment within the nucleus. The resulting pictures from the immuno-fluorescent work are shown in figure 6A. The presence of the HA-tagged peptide stains positive for the FITC label and appears as a green fluorescent signal while the DAPI staining of the
nuclei of the cells appears blue. Based on the presence of the green fluorescence seen only in the nucleus, we can conclude that the 1kHA peptide can actually enter the nucleus. Therefore the first NLS for the CHD2 protein is authentic and sufficient for nuclear localization. This is an interesting and surprising result considering the fact that CHD2 shares a high homology with CHD1 that has a single NLS signal at the C-terminus. Despite this result, the mutant Chd2 protein that retains this NLS signal and an additional peptide of ~1300 amino acids comprised of the β-galactosidase-neomycin fusion gene does not localize to the nucleus (Figure 6B). One probable reason for the localization of the mutant protein in the cytoplasm (and not within the nucleus) may be the presence of the β-gal-neomycin fusion peptide. We propose that the presence of the bacterial- β-galactosidase peptide sequence may prevent its nuclear localization.

Figure 6. (A.) Nuclear localization of 1K-HA Chd2 fragment in mammalian U20S cells. (B.) Cytoplasmic localization of the Chd2 mutant protein in mouse embryonic fibroblasts.
FUTURE EXPERIMENTS

In the future, we hope to confirm the results of the HA tag immuno-fluorescence work by adding a Flag tag to the CHD2 clones. In addition, nuclear and cytoplasmic fractions will be analyzed with western-blot assays. We will also express additional regions of the Chd2 protein by cloning a larger region of the gene and identify interacting proteins using mass spectrometry.

Acknowledgements:

I would like to sincerely thank Dr. Sundaresan Venkatachalam for his guidance and assistance with this project. I also greatly appreciate the time and effort that Dr. Prabakaran Nagarajan and Sangeetha Rajagopalan have put into the CHD2 project. Lastly, I would like to acknowledge Michelle Chi, Nicole Mitchell, Lesley Starnes, Guy Wiles, and John York for their assistance with this project.

References:


