Biochemistry of a mutant, cancer-causing DNA licensing protein MCM4

Honors Thesis

Presented to the College of Agriculture and Life Sciences,

Genetics Honors Program

of Cornell University

in Partial Fulfillment of the Requirements for the

Research Honors Program

by

Dian Yang

May, 2011

Research Advisor: Dr. John C. Schimenti

Table of Content

Abstract

Biochemistry of a mutant, cancer-causing DNA licensing protein MCM4

Dian Yang

Under the supervision of John C. Schimenti

Department of Biomedical Sciences

Chromosomal instability is a hallmark of cancer cells and impaired DNA replication is a major cause of chromosome instability. This thesis investigates the mutant allele *Chaos3* (chromosome aberrations occurring spontaneously 3) isolated in a N-ethyl-N-nitrosourea (ENU) mutagenesis screen of mice for chromosome instability. This mutation causes high levels of chromosome instability. Over 80% of mutant female homozygous mice develop mammary adenocarcinomas with a mean latency of 12 months.

The *Chaos3* mutation occurs in the *Mcm4* gene (minichromosome maintenance 4) and causes a single amino acid change (F345I) in a highly conserved region. *Mcm4* encodes the MCM4 subunit of the hetero-hexameric MCM2-7 complex. MCM2-7 is a DNA replicative helicase. During DNA replication, MCM2-7 is loaded on the replication origins as double hexamers.

The goal of this study was to characterize the biochemical consequence of the point mutation, focusing on protein interactions measured by co-immunoprecipitation. Here, we report that the *Chaos3* mutation causes a dramatic decrease of MCM4-MCM6 interaction and a slight decrease of MCM4-MCM7 interaction. This finding suggests that the loss of interaction might cause structural instability of the replicative machinery,

leading to increased number of stalled replication forks and chromosome segregation defects. Another finding was identifying an interaction between MCM4 and HSP70.

Key Words

DNA Replication, Breast Cancer, MCMs, *Chaos3* Mutation, Protein Interactions

Introduction

This thesis investigates the biochemical consequences of a mutant, cancer-causing, DNA licensing protein MCM4, called MCM4^{Chaos3}.

In a forward genetic screen in mice identifying genes or alleles that predispose to cancer, N-ethyl-N-nitrosourea (ENU) was used to introduce point mutations in the mouse genome. To screen for tumorigenesis, some challenges are that first, cancer is believed to be caused by accumulation of genetic and epigenetic alterations (Chan, 2008; Thompson, 2008; Thoma, 2011; Vaiserman, 2011). Therefore, we need to raise a large number of mice until the onset of tumors. Second, the penetrance of cancer is not 100%. Because of the challenges of performing a screen to directly detect the incidence of cancer, it was decided to use chromosome instability, a hallmark of cancer cells, as a biomarker for cancer instead. The level of chromosome instability was assessed by quantifying spontaneous micronuclei levels in erythrocytes because spontaneous micronuclei are caused by chromosome instability and are easier to be measured in erythrocytes which don't have nuclei. Spontaneous micronuclei were measured via propidium iodide staining and flow cytometry. By this assay, the Schimenti lab screened a few thousand mice and found some mutants exhibiting high level of chromosome instability (Shima, 2007).

A mutation called Chaos3 (**Ch**romosomal **a**berration **o**ccurring **s**pontaneously 3) was isolated from this screen. In terms of the level of micronuclei, heterozygotes showed a mild elevation of micronuclei while homozygotes showed a dramatic elevation compared to wild-type mice. As for tumorigenesis, the heterozygotes are grossly normal

while over 80% of homozygous female mice develop mammary adeneocarcinomas with a mean latency of 12 months (Shima, 2007).

Previous work mapped the *Chaos3* mutation to a 1.3-Mb region on mouse chromosome 16, which contains 11 RefSeq genes. By sequencing of these candidate genes, a single point mutation was identified in the essential and highly conserved *Mcm4 (**m**ini-**c**hromosome **m**aintenance)* gene. The point mutation causes a T-A transversion at nucleotide 1033 of the coding region of *Mcm4*. This point mutation encodes a single amino acid change from phenylalanine to isoleucine at residue 345 (Shima, 2007).

The *Mcm4* gene, encoding the MCM4 subunit of the MCM2-7 complex, is a highly conserved gene in all eukaryotic organisms (Tye, 1999; Bochman, 2008). The MCM 2-7 complex is the presumptive DNA helicase, a ring-shaped hetero-hexamer of six distinct but structurally related proteins (Blow, 2005; Moyer, 2006; Bochman, 2008). In mammalian cells, MCMs are mainly found in the form of subcomplexes rather than heterohexamers (Prokhorova, 2000). The genes encoding the MCM2-7 complex were identified in a genetic screen in budding yeast for DNA replication mutants in the 1980s (Maine, 1984). As shown in Figure 1, from late M to early G1 phase of cell cycle, DNA replication origins are bound by the hexameric origin recognition complex (ORC) (Gilbert, 2001; Forsburg, 2004; Lei, 2005). Initiation factors CDC6 and CDT1, which are necessary for further recruiting the MCM2- 7 complex to replication origins, were recruited by ORC (Randell, 2006). All six MCM subunits colocalize to origins of replication during pre-replicative complex (pre-RC) formation. During S phase, MCM2-7 double hexamers are activated to unwind the DNA by phosphorylation, an event catalyzed by cyclin-dependent kinases (CDKs) and Dbf4-dependent kinase (DDK) Cdc7.

With the help of CDK, DDK, Cdc7 as well as other replication factors of currently unknown functions, the replication forks are assembled and bidirectional DNA replication begins (Bell, 2002).

Even though many down-stream effects of this *Chaos3* mutation have been discovered, the direct biochemical consequence of this mutation is still not clear. Based on the structural data from archaeal *M. thermoautotrophicum* and the high conservation of MCM protein, the site where the point mutation resides is predicted to be a critical region for interaction with other MCMs (Fletcher, 2003; Shima, 2007). The goal of my research project is to identify the biochemical consequences of the *Chaos3* point mutation on the MCM4 protein, especially focusing on protein interactions. The three kinds of interaction I am studying are those within one MCM hexamer, those between two hexamers, and those with other MCM binding proteins. MCM4 over-expression and co-immunoprecipitation were used to study these interactions. Mass spectrometry was used to identify unknown MCM binding proteins.

In this study, I found a dramatic decrease of $MCM4^{Chaos3}$-MCM6 interaction and a slight decrease of $MCM4^{Chaos3}$-MCM7 interaction in the *Chaos3* cells based on Western blotting results. The MCM4-MCM7 interaction is more critical for the survival of mice, while the loss of MCM4-MCM6 interaction leads to the structural instability of the MCM2-7 helicase and might cause increased number of stalled replication forks and chromosome segregation problems, resulting in chromosome instability (Kawabata, 2011). Besides, I also characterize the interaction between Hsp70 protein 1B and MCM4. Further experiments are needed to characterize the role of Hsp70 in DNA replication and its influence on MCM complex.

Materials and Methods

DNA constructs:

3X *Flag* Epitope tag was amplified by PCR from *pBICEP-CMV-2* using the primer pair 5'-AAACTCGAGGACTAGACCATGACGGT-3' (Forward) and 5'-AAAACCGGTCACTCGTCATCCTTGTA-3' (Reverse). As shown in **fig. 2**, the resultant was sub-cloned into pcDNA4/TO/Myc/His/Mcm4^{WT} and pcDNA4/TO/Myc/His/ *Mcm4^{Chaos3}* using *Age*I and *Xho*I site into the plasmid to replace *Myc* epitope tag. pcDNA4/TO/Myc/His/Mcm4^{WT} and pcDNA4/TO/Myc/His/Mcm4^{Chaos3} were provided by Chen-Hua Chuang in the Schimenti lab. Inserts and vectors were purified by gel extraction kit (Invitrogen) and ligation was done using T4 ligase (NEB) at room temperature for 2 hours. Ligation product was transformed into competent cells (Top10) and plated out on Luria Bertani-ampicillin (LB-Amp) plates. After 16h, several colonies were picked and confirmed by PCR amplification. Finally, the plasmids were prepared by plasmid mini-prep kit (Invitrogen) and sequenced at Cornell Life Science Core Facility to validate the sequence of the constructs. Thus, four plasmid constructs were made, pcDNA4/TO/His/Mcm4^{WT}/3XFlag, pcDNA4/TO/His/Mcm4^{Chaos3}/3XFlag, pcDNA4/TO /His/Mcm4^{WT}/Myc, and pcDNA4/TO/His/ Mcm4^{Chaos3} /Myc.

Cell culture:

HEK-293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% Fetal Bovine Serum, 1mM sodium pyruvate and 100U/ml penicillin and streptomycin.

Mcm4$^{WT/WT}$ and *Mcm4*$^{Chaos3/Chaos3}$ cells were isolated as primary MEFs and were made into immortalized cell lines according to 3T3-L1A Cell Line Protocol by Chen-Hua Chuang in the Schimenti lab. *Mcm4*$^{WT/WT}$ and *Mcm4*$^{Chaos3/Chaos3}$ cells were maintained in regular cell culture.

Transfection:

Approximately 1x10^6 HEK cells were seeded in 60 mm cell culture dishes 24 hours before transfection. 2 hours before transfection, the medium was changed to DMEM medium without penicillin and streptomycin. Cells were transfected with 600 ng each plasmid using Lipofectamine 2000 (Invitrogen) according to the protocol provided by Lipofectamine 2000 kit. 48 hours after transfection, cells were harvested by RIPA as indicated in *Cell Lysis* section. In the MCM4 over-expression and co-IP experiment, four groups of co-transfection were done: 1. *Mcm4*^{WT}*-Flag* and *Mcm4*^{WT}*-Myc* 2. *Mcm4*^{Chaos3}*-Flag* and *Mcm4*^{Chaos3}*-Myc* 3. *Mcm4*^{WT}*-Flag* and *Mcm4*^{Chaos3}*-Myc* 4.Negative Control: *Mcm4*^{Chaos3}*-Myc* and Vector alone. In the MCM-HSP70 reciprocal IP experiment, five groups of transfection were done in HEK-293 cells: 1. *Mcm4*^{WT}*-Flag, 2. LacZ-Myc, 3. Mcm4*^{WT}*-Myc, 4. Mcm4*^{Chaos3}*-Myc 5. Mcm6*^{WT}*-Myc*. (All the plasmids used for transfection were prepared by Invitrogen Maxi-Prep Kit.)

Cell Lysis by Radio-Immunoprecipitation Assay (RIPA) buffer:

48h after transfection, the plates were wash twice by PBS briefly and gently. Whole cell lysates were harversted by adding 600ul RIPA buffer (50mM Tris-HCl pH 8.0, 150mM NaCl, 0.1% SDS, 1% NP-40, 1mM EDTA, 0.5% sodium deoxycholate, 50mM NaF) with proteinase inhibitor cocktail Tablet (Roche) to each 60 mm plate. The plates were put in 4 C for 10 minutes and cells lysates were collected. Lysates were incubated on ice for an hour and were vortexed every 5 minutes. After an hour, lysates were centrifuged at 10,000 xg for 10 mins at 4 C and the supernatant was kept for immunoprecipitation.

Immunoprecipitation:

Lysates were prepared using 600ul RIPA lysis buffer supplemented with protease inhibitor cocktail Tablet (Roche) per 60 mm plate. 30ul supernatants were saved as input sample and the rest were used for immunoprecipitation. Lysates were incubated with anti-FLAG monoclonal antibody (Sigma) at a dilution of 1: 300 overnight at 4 C and then precipitated with 100ul Protein A Agarose beads (Millipore) for 1h at room temperature. Precipitates were washed three times with RIPA lysis buffer. Bound proteins were eluted from the beads with 50ul 2X Laemmli buffer (SDS 2% (w/v), 1M Tris, pH 6.8, 50 mM, bromphenol blue 0.2 mg/mL, DTT 0.1 M DTT, H_2O 10% (v/v), Glycerol) at 95 C for 5 min before immunoblotting. For input samples, 30ul 2X Laemmli buffer was added to each. For the reciprocal endogenous immunoprecipitation, anti-MCM6 (Santa Cruz, 1: 250) and anti-MCM7 (cell signaling, 1:250) were used separately. For MCM4 – HSP70

immunoprecipitation experiment, lysates were incubated with anti-MYC (Sigma, 1: 350 dilution) and anti-HSP70 antibody (cell-signaling, 1: 250 dilution) separately.

Immunoblotting:

Whole cell lysate samples and elutes from protein A beads were separated by 7.5% SDS-PAGE (Bio-Rad), electrotransferred to nitrocellulose membranes (Bio-Rad). Membranes were washed twice in TBS (50 mM Tris.HCl, pH 7.4 and 150 mM NaCl.) containing 1% Tween 20 (TBST) for 3min and then blocked in TBST and 5% nonfat dry milk. Membranes were then incubated in TBST and 3% nonfat dry milk at 4 C overnight with the following primary antibodies: Rabbit polyclonal anti-MCM4 ab4459 (1:2000, abcam); Rabbit polyclonal anti-MCM6 NB100-78262 (1:1000, Novus); Mouse monoclonal anti-MCM7 ab2360 (1:1500, abacam); Mouse monoclonal anti-c-Myc M4439 (1:2000, Sigma); Rabbit polyclonal anti-FLAG #2368 (1:2000, Cell Signaling); Rat monoclonal anti-HSP 70 #4873 (1:1000, Cell Signaling); Mouse anti-GAPDH (1:5000, Advanced Immunochemicals); Mouse anti-β-Actin (1:5000, Sigma). Membranes were washed four times with TBST for 5 minutes each and incubated in TBST and 3% nonfat milk for 1.5 h at room temperature with the following secondary antibodies: HRP Goat anti-Rabbit IgG #31460 (1:3000, Pierce); HRP Goat anti-Mouse IgG #31430 (1:3000, Thermo Scientific); HRP Goat anti-Rat IgG # 31470 (1:5000, Pierce). After 4 washes with TBST, peroxidase activity was revealed by Pierce ECL kit. Quantifications were done using ImageJ software.

Colloidal Commassie Blue Staining and Mass Spectrometry Analysis:

Protocols were obtained from the Cornell Life Science Core Facility. The protein samples were prepared as mentioned before and separated by SDS-polyacrylamide gel. The SDS-polyacrylamide gel was fixed in the Fix Buffer for 1 hour. (50% acetic acid, 40% ethanol, 10% ddwater.) Then the fixed gel was stored in ddwater. Then gel was sent to the Life Sciences Core Facility, stained with colloidal commassie blue and 2 bands of protein was cut off for mass spectrometry analysis.

Results

The Mcm4^{Chaos3} mutation caused a dramatic decrease of MCM4-MCM6 interaction and a slight decrease of MCM4-MCM7 interaction.

The *Chaos3* mutation causes a single amino acid change at residue 345 from Phe to Ile in the highly conserved domain of MCM4 protein. This point mutation can cause a pan-reduction of all MCM2-7 components at both mRNA and protein levels as well as a reduction of dormant origins at cellular level (Chuang, 2010; Kawabata, 2011). It can also cause genomic instability and extreme cancer susceptibility in homozygous mutant mice (Shima, 2007). However, the direct biochemical consequences of this mutation at protein level are still unknown.

Based on the fact that Phe345 in MCM4 is predicted to reside at the protein-protein interface region of the different MCM subunits of the MCM2-7 complex and that MCM has several subcomplex forms in cells (Remus, 2009; Gambus, 2011; **Fig. 3.**), we

hypothesized that the *Chaos3* mutation might change the interaction between the MCM4 and other proteins, which may affect the stability or firing of the Pre-RC (Shima, 2007). Three kinds of interactions might be disrupted by the *Chaos3* mutation. They are the interactions with other MCM monomers, those between two MCM hexamers, and those with other MCM binding proteins. To study these interactions, MCM4 over-expression and co-IP were done to compare the interaction between WT and mutant MCM4.

A problem for studying the interaction between two MCM hexamers is that the two hexamers are identical, thus we cannot distinguish the corresponding MCM subunits of both MCM2-7 complex in our co-IP and western-blot analysis. To distinguish the two MCM complexes, epitope tags, FLAG and MYC were introduced to tag both MCM4 and $MCM4^{Chaos3}$ separately. 4 groups of co-transfections were done in HEK-293 cells and anti-FLAG antibody was used to precipitate all FLAG-tagged MCM4 proteins.

As shown in **Fig. 4**, there was no change in interaction between the two $MCM4^{Chaos3}$ subunits of double hexamers. This observation is inconsistent with our original hypothesis that the mutation might disrupt inter-hexamer interactions. However, the *Chaos3* mutation caused a dramatic qualitative decrease (92%-96% loss) in MCM4-MCM6 interactions and a relatively slight decrease (35-40% loss) of MCM4-MCM7 interaction, as quantified by ImageJ software.

We further investigated the intra-hexamer interaction of MCM complex, focusing on MCM4, MCM6 and MCM7 by performing a reciprocal co-IP experiment. In the reciprocal experiment, anti-MCM6 and anti-MCM7 antibodies were used separately to precipitate endogenous MCM complexes in $Mcm4^{WT/WT}$ and $Mcm4^{Chaos3/Chaos3}$ mouse embryonic fibroblasts. The results in **Fig. 5** were consistent with the MCM4 over-

expression and co-IP experiment. When we pull down the MCM complex using anti-MCM6 antibody, $MCM4^{WT}$, but not $MCM4^{Chaos3}$, could be precipitated. We didn't pull down MCM7. One explanation is that because MCM7 indirectly interacts with MCM6 via MCM4, the loss of interaction between $MCM4^{Chaos3}$ and MCM6 will lead to a failure in precipitating MCM7. The results were similar in the co-IP MCM7 experiment. When we pulled down the MCM complex using anti-MCM7 antibody, approximately 40% more $MCM4^{WT}$ than $MCM4^{Chaos3}$ could be precipitated as measured by ImageJ software. MCM6 was only precipitated when $MCM4^{WT}$ was present, further indicating the loss of interaction between $MCM4^{Chaos3}$ and MCM6. To sum up, the *Chaos3* mutation caused a slight decrease of interaction between $MCM4^{Chaos3}$ and MCM7, and a dramatic decrease of interaction between $MCM4^{Chaos3}$ and MCM6.

The slight decrease in the MCM4-MCM7 interaction was validated by colloidal commassie blue staining and mass spectrometry analysis (**Fig. 6**). By commassie staining the SDS polyacrylamide gel, we observed a decreased interaction of a 75-80 kD protein with $MCM4^{Chaos3}$, which is consistent with the molecular weight of MCM7. Mass spectrometry analysis showed that the main component of that band on the gel is MCM7, which validated our findings in previous results from Western blotting.

We suggest a model for the biochemical consequences of the F345I mutation in the MCM4 subunit on protein interactions between MCM monomers. As is shown in **Fig. 7**, the F345I mutation causes a complete loss of $MCM4^{Chaos3}$-MCM6 interaction and a slight decrease in $MCM4^{Chaos3}$-MCM7 interactions. It should be noted that only a tiny amount of MCM2, MCM3 and MCM5 was pulled down in the co-IP experiments (Data not shown). Based on our reciprocal IP results and the different forms of MCM sub-

complexes in the cell (Fig.1), our model for MCM in the whole cell lysate suggests that MCM proteins mainly exist as sub-complexes in mouse cells, MCM4, 6, 7, MCM3, 5 and MCM2. This point will be further explained in the Discussion section.

MCM4 interacts with hsp70 protein, but the interaction is not MCM4-specific.

We became interested in a highly abundant 70 kD MCM binding protein (**Fig. 6**, the protein in the lower band of the blue-stained SDS-PAGE). Mass spectrometry results showed that this protein was heat shock 70KD protein 1-B. The well-characterized heat shock protein 70 is a highly conserved family of molecular chaperons. Hsp70 has several functions: it could maintain protein precursor proteins for translocation, bind to incoming polypeptides on the trans side of mitochondrial and ER membranes and plays a key role in protein folding (Hartl, 1992; Beckmann, 1990; Bochkareva, 1988). Hsp70 and Hsp90 are important for regulating protein folding and degradation (Hohfeld, 2001; Cyr, 2002).

Reciprocal co-IP experiments were done to study the interaction between MCM4 and Hsp70 protein 1B. Five groups of transfection were done: *$Mcm4^{WT}$-Flag, LacZ-Myc, $Mcm4^{WT}$-Myc, $Mcm4^{Chaos3}$-Myc and $Mcm6^{WT}$-Myc*. *$Mcm4^{WT}$-Flag* was used for control of the precipitation of MYC-tagged protein. *LacZ-Myc* was used for control of the specificity of Hsp70 1-B binding.

In the reciprocal co-IP experiments, Anti-MYC and Anti-HSP70 antibody were used separately. In **Fig. 8** of representative western blot results, we could observe that both $MCM4^{WT}$ and $MCM4^{Chaos3}$ have similar interactions with HSP70 protein in the reciprocal IP experiments, however, the *Lac-Z encoded* β-gal could also bind HSP70

protein, suggesting that the interaction between MCM4 and HSP70 was not specific. One thing we need to mention is the high background of HSP70 in the control. When we immunoprecipitate using anti-MYC antibody in lysates from cells transfected with only *Mcm4-Flag,* we observed a weak signal, indicating that the HSP70 protein was so abundant that the beads itself could pull down HSP70. Even though the background is strong, we can still believe the interaction between MCM4, LacZ and HSP70, because the signals of the experimental samples were much stronger than that of the control. In future, the protein A beads should be blocked by BSA before the precipitation to minimize the influence of background.

Discussion

The biological meanings of the loss of MCM4-MCM6 interaction

Based on the results generated from the reciprocal co-IP experiments, the *Chaos3* mutation disrupted the MCM4-MCM6 interaction dramatically but only slightly influenced the MCM4-MCM7 interaction.

In both co-immunoprecipitation systems, distinct amounts of MCM6 and 7 could be pulled down, however, only trace amounts of MCM2, 3 or 5 could be co-precipitated with either WT or *Chaos3* MCM4. This can be explained by the fact that the MCM complex mainly exists in the form of several sub-complexes (**Fig. 3**), and the combinations of MCM4, MCM6 and MCM7 exist in almost all kinds of the subcomplexes as well as hexamers, except the MCM3, 5 sub-complex (Ma, 2010;

Prokhorova, 2000). This is why we precipitated a distinct amount of MCM4, MCM6 and MCM7 and very tiny amount of MCM2, MCM3 and MCM5. This result is consistent with previous biochemical studies that MCM4, 6, 7 sub-complexes is a dominant form in glycerol gradient sedimentation analysis (Ma, 2010; Bochman, 2008; Lee, 2000). Functional studies showed that MCM4, 6, 7 sub-complexes formed the core of helicase activity (You, 2002, 2005; Ishimi, 1997). In our MCM4 over-expression and co-IP experiment, the interaction between two $MCM4^{Chaos3}$ subunits is the same as that of $MCM4^{WT}$. This, together with the loss of interaction between $MCM4^{Chaos3}$ and MCM6, leads us to postulate the assumption that the *Chaos3* mutation causes the MCM4,6,7 double-trimer to become a MCM4, 7 double dimer.

From the whole cell lysate IP result, we can see almost a complete loss of MCM4-MCM6 interaction and a slight decrease of MCM4-MCM7 interaction. Even though the interaction between MCM4-MCM7 seems more important in DNA replication, it has long been proposed that the intact ring structure of MCM complex that could encircle single and/or double stranded DNA is important for the unwinding function (Bochman, 2008). Two hypotheses for compensating this complete loss of interaction are that: (1) It is possible that when MCMs are recruited on the chromosomal replication origins, some MCM binding proteins might cause some conformational change of MCM proteins to stabilize the decreased MCM4-MCM6 interactions on the chromatin. To test this possibility, MCM fractionation and co-IP using chromatin-bound fraction of MCMs rather than the whole cell lysate may help uncover the remaining MCM4-MCM6 interactions. (2) Another possibility is that the MCM4-MCM7 interaction might be sufficient for helicase activity. MCM4 and MCM7 can form an active helicase assembly

even without MCM6 in mammalian cells (Kanter, 2008). So even though there is a slight decrease of MCM4-MCM7 interaction, it seems that the decrease is so minor that it doesn't influence the function of MCM complexes.

In terms of MCM function, *in vivo* studies showed that *Mcm4*$^{Chaos3/Chaos3}$ mice are on average viable for at least 1 year on average until the onset of cancer (Shima, 2007). *In-vitro* helicase assay showed that the dramatic loss of MCM4-MCM6 interaction and the slight loss of MCM4-MCM7 interaction didn't result in obvious defects in helicase activity (Kawabata, 2011). On the contrary, the helicase with MCM4^{Chaos3} could unwind DNA substrate a little bit faster than did the wild-type helicase (Kawabata, 2011). What's more, the Shima group also discovered that the *Chaos3* mutation might cause an increased number of stalled replication forks and chromosome segregation problems in late M phase, resulting in chromosome instability. These stalled replication forks can escape the DNA damage checkpoints responses and cause chromosome segregation problems later in cell division (Kawabata, 2011). However, the exact biochemical components as well as pathways involved in how the disrupted MCM4-MCM6 interaction finally lead to these outcomes is still unknown. These interesting phenotypes suggested that we need to explore other possible biochemical roles of the loss of MCM4-MCM6 interaction in order to fully understand downstream effects of this loss of interaction and how this structural instability finally leads to severe genomic instability.

Characterizing the role of HSP 70 protein in DNA replication

For the other project studying the MCM-Hsp70 interaction and the role of Hsp70 protein in DNA replication, even though this experiment showed clearly that the

interaction between MCM4 and Hsp70 seems not to be MCM-specific, there are several limitations and questions about this experiment. Due to the several roles of Hsp70 in protein quality control, the timing and cellular condition for certain interactions are more important to characterizing the exact effect of Hsp70 on MCM2-7 complex during DNA replication.

There are several hypothesized roles that need to be verified in future experiments. First of all, it is suggested that molecular chaperons might be required for helping the conformational change of MCM helicase proteins during DNA replication (Bochman, 2008). It is possible that HSP70 protein might maintain some MCM4-MCM6 interaction to compensate for the structural instability caused by $MCM4^{Chaos3}$. To test this possibility, we should focus on the chromatin-bound fraction of MCM complex to study the role of HSP70 in DNA replication rather than use whole cell lysate sample in which a lot of nascent MCM proteins may need molecular chaperon just for protein folding. Another possible role of HSP70 is in protein quality control. The HSP70 may either help the protein fold properly and maintain some interactions or bind the $MCM4^{Chaos3}$ protein and trigger ubiquitination. However, in our over-expression and reciprocal IP experiment, the excess amount of expressed proteins might induce stress in cells, which might induce molecular chaperon to help protein folding or degradation. Thus, this over-expression creates a non-physiological environment and might cover the real situation in $Mcm4^{Chaos3/Chaos3}$ cells. It will be better to conduct an experiment in $Mcm4^{WT/WT}$ and $Mcm4^{Chaos3/Chaos3}$ cells studying whether there is an increased level of HSP70 protein in $Mcm4^{Chaos3/Chaos3}$ cells.

Precise DNA replication is important for maintaining genome integrity and stability, which requires the cooperation of all the replication components. Defects in any of these factors may cause genomic instability, finally leading to cancer. In the *Chaos3* cancer mouse model, the F345I mutation in the helicase MCM4 subunit causes a dramatic structural instability of the helicase. This might cause an increased number of stalled replication forks and chromosome segregation problems in late M phase, resulting in chromosome instability. The accumulation of chromosomal instability finally leads to tumorigenesis in female $Mcm4^{Chaos3/Chaos3}$ mice. The significance of this study of the viable Chaos3 allele is that it can not only help us uncover the role and mechanism of DNA replication machinery in maintaining genome integrity, but also opens a door for us to explore the potential role of other components of the DNA replication machinery in chromosomal instability and cancer predispositions.

Literature

Bell, S. P., and Dutta, A. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71:333–374.

Bell, S. P., and Dutta, A. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71:333–374.

Beckmann, R. P., Mizzen L. E., and Welch W. J. 1990. Interaction of Hsp 70 with newly synthesized proteins: implications for protein folding and assembly. *Science.* 248(4957):850-4.

Berretta, R. and Moscato, P. 2010. Cancer Biomarker Discovery: The Entropic Hallmark. *PLoS One.* 2010; 5(8): e12262.

Bochkareva, E.S., Lissin, N.M., and Girshovich, A.S. 1988. Transient association of newly synthesized unfolded proteins with the heat-shock GroEL protein. *Nature.* 1988 Nov 17;336(6196):254-7.

Bochman, M and Schwacha. 2008. A The Mcm Complex: Unwinding the Mechanism of a Replicative Helicase, *Microbiology and Molecular Biology Reviews*, Dec. 2009, p. 652–683

Bochman, M. L., and A. Schwacha. 2007. Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and 5 define a slow ATP-dependent step. *J. Biol. Chem.* 282:33795–33804.

Bochman, M. L., and A. Schwacha. 2008. The Mcm2-7 complex has in vitro helicase activity. *Mol. Cell* 31:287–293.

Blow, JJ and Dutta, A. 2005. Preventing re-replication of chromosomal DNA. *Nature Reviews Molecular Cell Biology* **6**, 476-486

Chan, T.A, Glockner, S., Yi, J.M., Chen, W., Van, Neste. L, et al. (2008) Convergence of mutation and epigenetic alterations identifies common genes in cancer that predict for poor prognosis. PLoS Med 5(5):e114.

Chuang C-H, Wallace MD, Abratte C, Southard T, Schimenti JC. 2010. Incremental Genetic Perturbations to MCM2-7 Expression and Subcellular Distribution Reveal Exquisite Sensitivity of Mice to DNA Replication Stress. *PLoS Genet* 6(9)

Cyr DM, Höhfeld J, Patterson C. 2002 Protein quality control: U-box-containing E3 ubiquitin ligases join the fold. *Trends Biochem Sci*. (7):368-75.

Fletcher, R.J. et al. The structure and function of MCM from archaeal M. thermoautotrophicum. *Nat. Struct. Biol.* 10, 160–167 (2003).

Forsburg, S.L. Eukaryotic MCM proteins: beyond replication initiation. Microbiol. Mol. Biol. Rev. 68, 109–131 (2004).

Gambus A, Khoudoli GA, Jones RC, Blow JJ. 2011. MCM2-7 form double hexamers at licensed origins in xenopus egg extract. *J Biol Chem*.

Hartl, F.U. and Martin, J. 1992 PROTEIN FOLDING IN THE CELL: The Role of Molecular Chaperones Hsp70 and Hsp60. Annu. Rev. Biophys. Biomol. Struct. 1992.21:293-322

Höhfeld J, Cyr DM, Patterson C. 2001. From the cradle to the grave: molecular chaperones that may choose between folding and degradation. *EMBO Rep*. (10):885-90

Ishimi Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem*. 272(39):24508-13

Kawabata T, Luebben SW, Yamaguchi S, Ilves I, Matise I, Buske T, Botchan MR, Shima N. 2011. Stalled Fork Rescue via Dormant Replication Origins in Unchallenged S Phase Promotes Proper Chromosome Segregation and Tumor Suppression. *Mol Cell*. 41(5):543-53

Maine, G.T., Sinha, P. & Tye, B.K. Mutants of S. cerevisiae defective in the maintenance of minichromosomes. Genetics 106, 365–385 (1984).

Moyer, S.E., Lewis, P.W. & Botchan, M.R. Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc. Natl. Acad. Sci. USA 103, 10236–10241 (2006).

Lee JK, Hurwitz J. 2000. Isolation and characterization of various complexes of the minichromosome maintenance proteins of Schizosaccharomyces pombe. *J Biol Chem*. 275(25):18871-8

Lei, M. The MCM complex: its role in DNA replication and implications for cancer therapy. Curr. Cancer Drug Targets 5, 365–380 (2005).

Prokhorova, T, A. and Blow, J. J. Sequential MCM/P1 Subcomplex Assembly Is Required to Form a Heterohexamer with Replication Licensing Activity. *The Journal of Biological Chemistry,* Vol. 275, No. 4, Issue of January 28, pp. 2491–2498, 2000

Shima, N, Alcaraz, A, Liachko, I, Buske, TR, Andrews, CA, Catherine AA, Munroe, RJ, Hartford, SA, Tye BK & Schimenti, JC. (2007) A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* 39: 93–98.

Shima, N, Buske, TR, Schimenti, JC. 2007. Genetic screen for chromosomeinstability in mice: Mcm4 and breast cancer. *Cell Cycle* 6: 1135–1140.

Thoma, C. R., Toso, A., Meraldi, P., and Krek, W. 2011. Mechanisms of aneuploidy and its suppression by tumour suppressor proteins. *Swiss Med Wkly*. 2011;141:w13170

Thompson, A., Brennan, K., Cox, A., Gee, J., Harcourt, D., et al. 2008. Evaluation of the current knowledge limitations in breast cancer research: a gap analysis. *Breast Cancer Research* 10: R26.

Tye, BK (1999) MCM proteins in DNA replication. *Annu Rev Biochem* 68: 649–686.

Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, Diffley JF. 2009. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*. 2009 Nov 13;139(4):719-30.

Randell, J.C., Bowers, J.L., Rodriguez, H.K. and Bell, S.P. 2006. Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol. Cell* 21:29–39.

Vaiserman A. M., Voitenko V. P., and Mekhova L. V. 2011. Epigenetic epidemiology of age-related diseases. *Ontogenez*. 42(1):30-50.

You, Z., Y. Ishimi, H. Masai, and F. Hanaoka. 2002. Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex. *J. Biol. Chem*. 277:42471–42479

You, Z., Masai, H. 2005. DNA binding and helicase actions of mouse MCM4/6/7 helicase. *Nucleic Acids Res*. 33(9):3033-47.

Acknowledgments

I would like to thank Dr. John Schimenti for the opportunity to work in his lab and for all of the patience, encouragement and support that he has given me over the past two years. In addition, I greatly appreciate the wisdom and patience of my student mentors, Chen-Hua Chuang, a very talented graduate student in the Schimenti lab, and the help of all the members of the Schimenti lab. I thank Dr. Jerry Feigenson, Dr. Volker Vogt and Dr. Laurel Southard for giving me a lot advice for undergraduate research and thank all the professors for reading my thesis and giving me very helpful suggestions.

I am also going to thank my parents for supporting me to study at Cornell University and finish my bachelor's degree and thank Dr. Zhangliang Chen for launching the CAU-Cornell Transfer program. I also thank Xin Li for continuous encouragement. Finally, I thank Cornell Hughes Scholars Program and the Stem Cell program for funding this research project.

Figures and Legends

Figure 1.

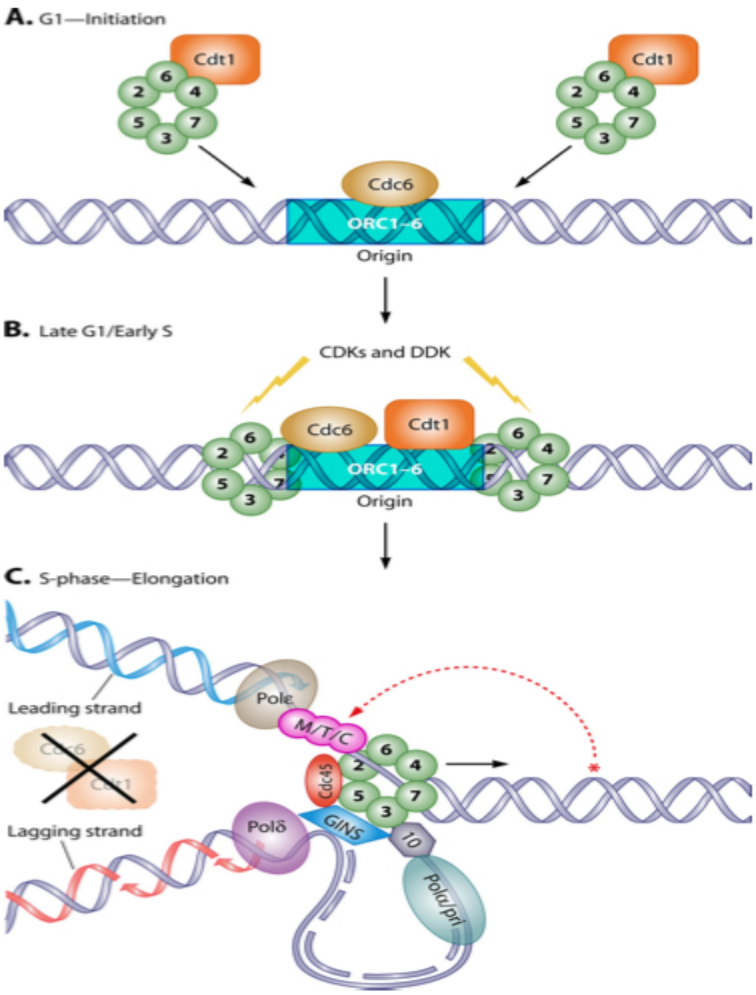


Fig. 1. DNA replication initiation (Adapted from Bochman, et al. 2009)

(A) During early G1 phase of cell cycle, DNA replication origins are bound by the origin recognition complex. **(B)** Initiation factors CDC6 and CDT1, which are necessary for further recruiting the MCM2- 7 complex to replication origins, were recruited by ORC. All six MCM subunits colocalize to origins of replication during pre-replicative complex (pre-RC) formation. **(C)** During S phase, MCM2-7 double hexamers are activated to unwind the DNA by CDK and DDK. With the help of Cdc7 as well as other replication factors, the replication forks are assembled and bidirectional DNA replication begins (Bell, 2002; Randell, 2006; Gilbert, 2001; Forsburg, 2004; Lei, 2005).

Figure 2.

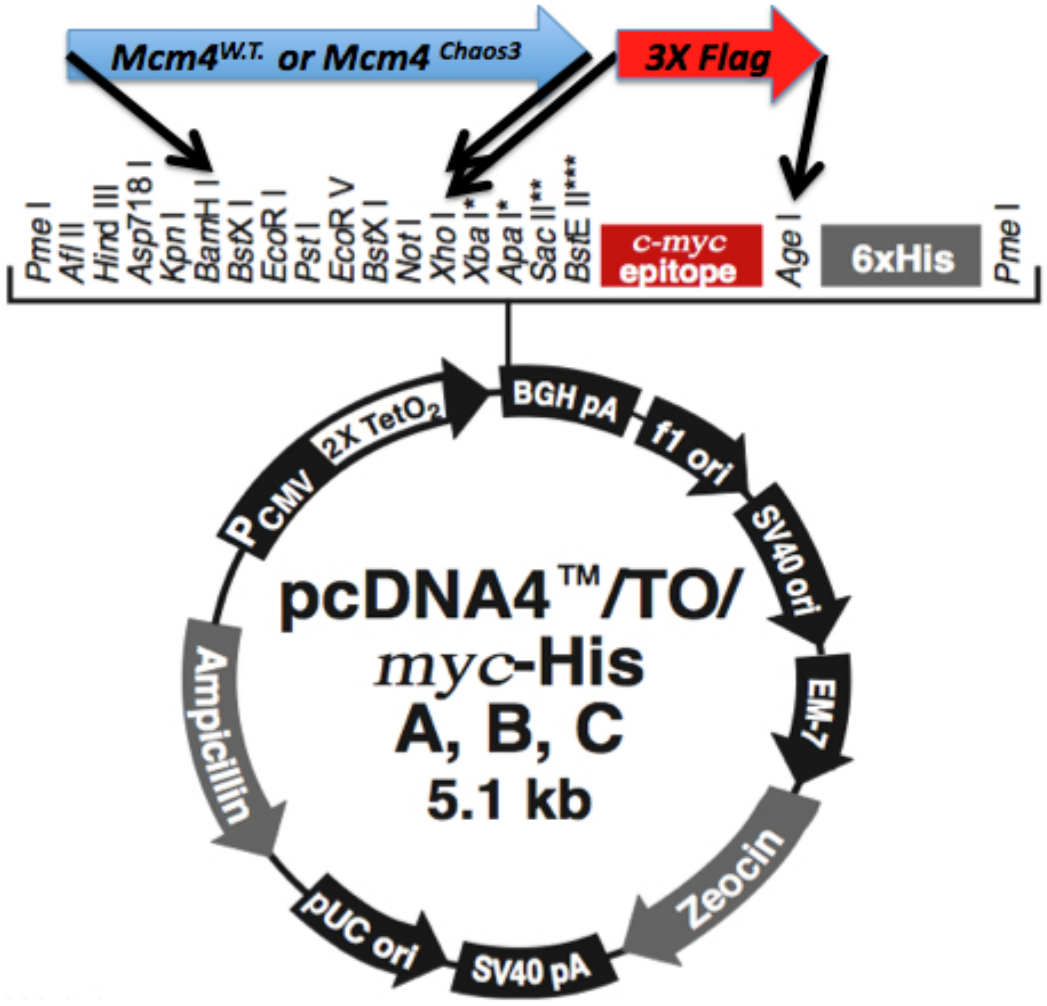


Fig. 2. The expression constructs. 3X *Flag* was amplified by PCR from *pBICEP-CMV-2* using the primer pair 5'-AAACTCGAGGACTAGACCATGACGGT-3' (Forward) and 5'-AAAACCGGTCACTCGTCATCCTTGTA-3' (Reverse). The resultant was sub-cloned into pcDNA4/TO/myc-His/Mcm4^{WT} and pcDNA4/TO/myc-His/Mcm4^{Chaos3} using *Age*I and *Xho*I site into the plasmid to replace *Myc* epitope. Both pcDNA4/TO/myc-His/Mcm4^{WT} and pcDNA4/TO/myc-His/Mcm4^{Chaos3} were provided by Chen-Hua Chuang in the Schimenti lab and *Mcm4^{WT}*and *Mcm4^{Chaos3}* have already been sub-cloned into *pcDNA4/TO/myc-His* using *BamH1 and XhoI*. Finally, four plasmid constructs were made, pcDNA4/TO/His/Mcm4^{WT}/3XFlag, pcDNA4/TO/His/Mcm4^{Chaos3} /3XFlag, pcDNA4/TO/His/Mcm4^{WT}/Myc, and pcDNA4/TO/His/ Mcm4^{Chaos3} /Myc.

Figure 3.

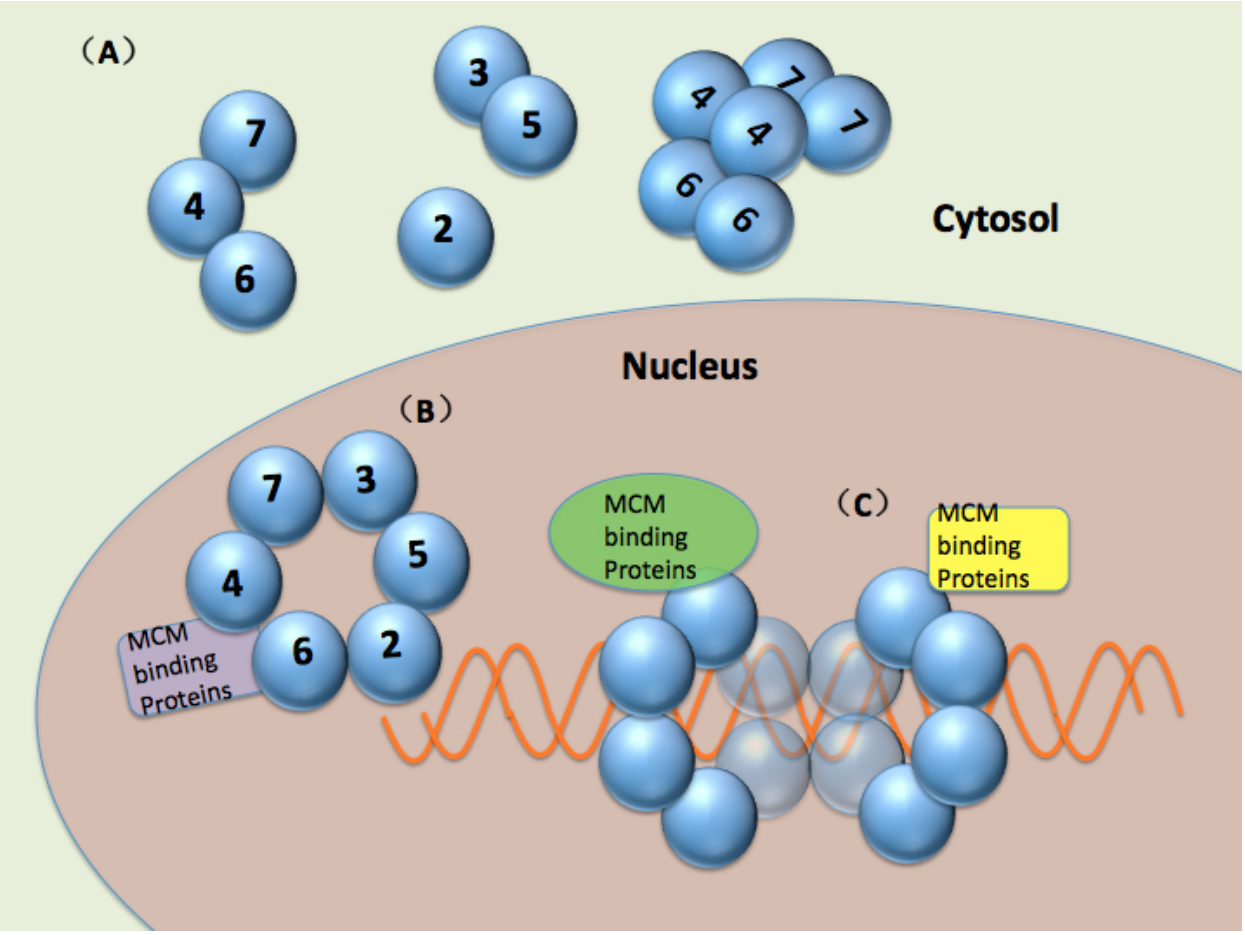


Fig. 3. MCM complex exists as several forms in cells: (A) In the cytosol, MCM mainly exists as several kinds sub-complexes, such as MCM3,5, MCM4,6,7 and MCM4,6,7 double trimer. (B) In the nucleus, MCM exists as a hexamer and bound by some MCM binding proteins. (C) At replication origins, MCM complexes are recruited as double hexamers to form the Pre-Replication Complex (Pre-RC). Some MCM binding proteins are also involved in this replication initiation process. (Ma, 2010; Bochman, 2008; Lee, 2000)

Figure 4.

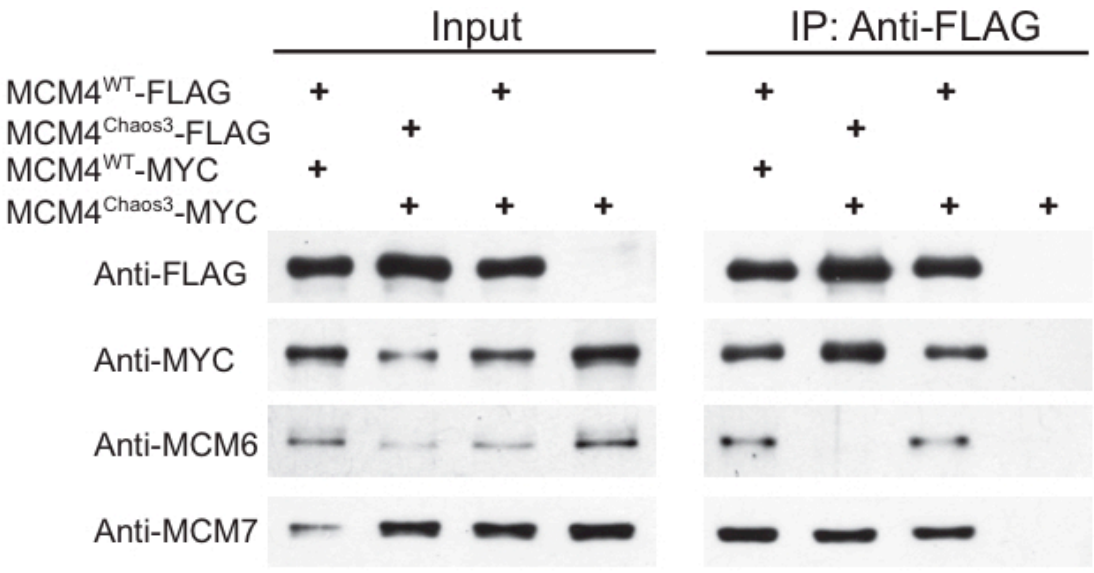


Fig. 4. MCM4 over-expression in HEK and Co-IP results Western blot result of the MCM4 over-expression in HEK and Co-IP experiment. The figure is a representative of western blot. Flag-tag proteins were immunoprecipitated from whole cell lysates prepared from HEK cells transiently expressing indicted plasmid. Both Input and IP product were electrophoresed by SDS-PAGE, transferred to nitrocellulose membrane and probed with anti-FLAG, anti-MYC, anti-MCM6 and anti-MCM7.

Figure 5.

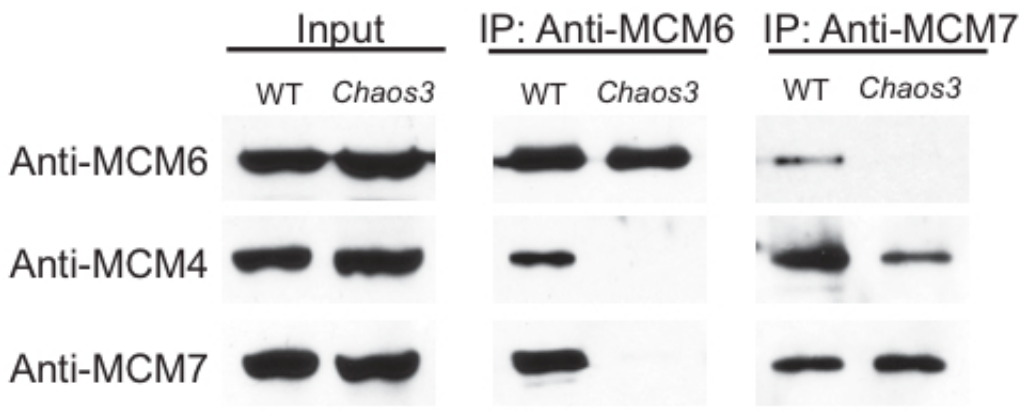
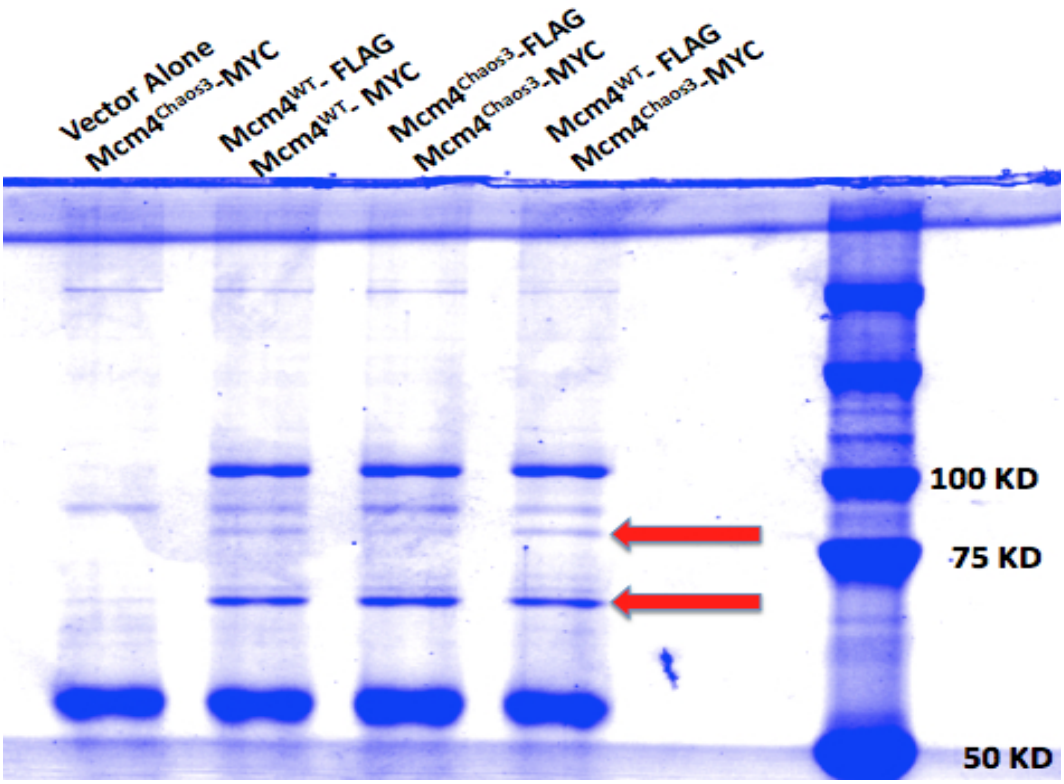


Fig. 5. Reciprocal Co-IP results in *Mcm4*$^{WT/WT}$ and *Mcm4*$^{Chaos3/Chaos3}$ cells Representative Western blots of the reciprocal co-IP experiments in *Mcm4*$^{WT/WT}$ and *Mcm4*$^{Chaos3/Chaos3}$ cells. The figure is a representative of 3 repeats. MCM6 and MCM7 were immunoprecipitated from whole cell lysates prepared from *Mcm4*$^{WT/WT}$ and *Mcm4*$^{Chaos3/Chaos3}$ immortalized MEF separately. Both Input and IP product were electrophoresed by SDS-PAGE, transferred to nitrocellulose membrane and probed with anti-MCM4, anti-MCM6 and anti-MCM7.

Figure 6.

(A) Colloidal Coomassie Blue Staining Result



(B) Mass Spectrometry Result (just show the top 3 components of each band)

(1) Upper band:

hit_num	prot_description	prot_score	molar %
1	**MCM 7 isoform 1**	12408	50.78
2	**MCM 4**	1955	5.79
3	**keratin 1**	1240	8.91

(2) Lower band

hit_num	Prot_description	prot_score	molar %
1	**hsp 70kD prot 1B**	13459	58.90
2	**hsp 70kD prot 1**	5744	3.77
3	**hsp 70kDa prot 6**	3437	1.14

Fig. 6. Colloidal Coomassie Blue Staining and mass spectrometry results The figure is a representative of 2 repeats. (A) Samples from the MCM4 over-expression and co-IP using anti-FLAG antibody was separated by SDS-PAGE and then the gel was stained by colloidal coomassie blue. (B) The two bands indicated on the gel were cut and analyzed by mass spectrometry and the results are shown above.

Figure 7.

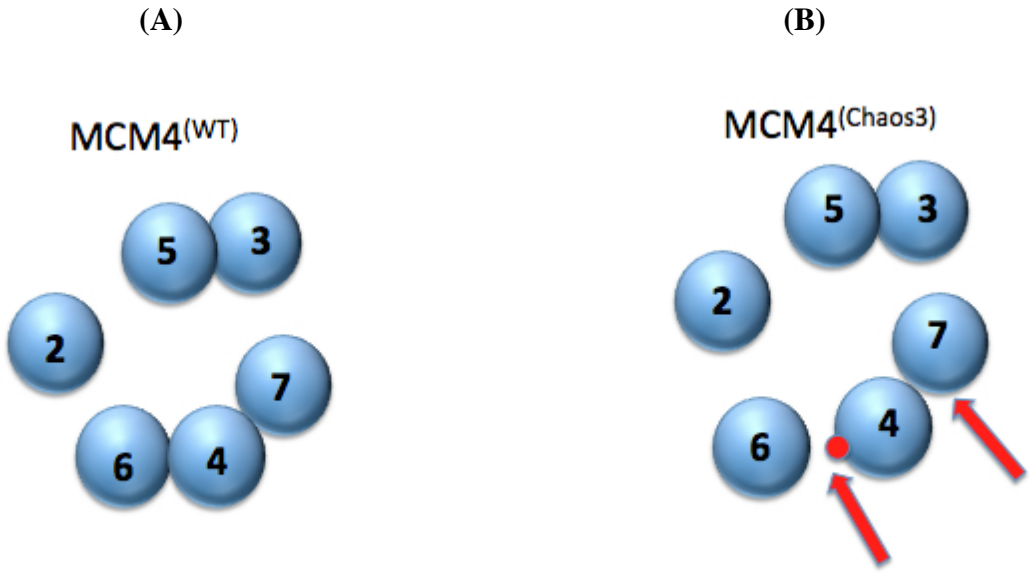


Fig. 7. A Model for the biochemical consequences of MCM4^{Chaos3} In the whole cell lysate, the MCM complex mainly exists in the form of several sub-complexes(Fig.1), however the combinations of MCM4, MCM6 and MCM7 exist in almost all kinds of the sub-complex as well as hexamers. This explains why we get very strong signal for MCM4, MCM6 and MCM7 and very weak signal for MCM2, MCM3 and MCM5. So in our model for MCM in the whole cell lysate, MCM proteins mainly exist as sub-complexes, MCM4,6,7, MCM3,5 and MCM2. (A) Wild-type MCM4 has interaction with both MCM6 and MCM7. (B) MCM4-Chaos3 causes a dramatic loss of MCM4-MCM6 interaction and a slight decrease of MCM4-MCM7 interaction.

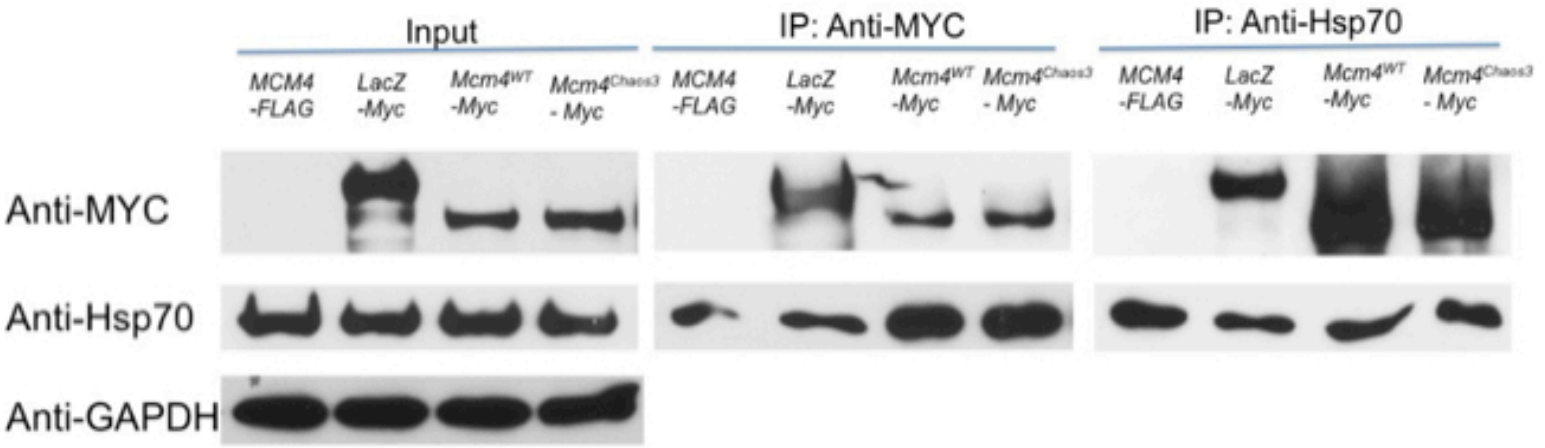


Fig. 8. The reciprocal IP experiment indicates the interaction between MCM4 and HSP70, but the interaction is not MCM4 specific. Western blot results of the co-IP experiment in HEK-293 cells after transient transfection of certain MYC tagged proteins as indicated in the figure. MYC-tagged proteins and HSP70 were immunoprecipitated from whole cell lysates separately. Both input and IP sample were separated by SDS-PAGE, transferred to nitrocellulose membranes and incubated with anti-MYC, anti-HSP70 and anti-GAPDH. When we precipitate using anti-MYC antibody in lysates from cells transfected with only *Mcm4-Flag*, we observed a weak signal, indicating that the HSP70 protein was so abundant that the beads itself could pull down HSP70. Even though the background is strong, we can still believe the interaction between MCM4, LacZ and Hsp70, because the signals of the experimental samples were much stronger than that of the control. In future, the protein A beads should be blocked by BSA before the precipitation to minimize the influence of background.