# Aberrant DNA Replication at an Ectopic Chromosomal Site in Human Cells

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

By

XIAOMI CHEN

B.S., Hebei University of Science and Technology, 1986

_____

2011

Wright State University

WRIGHT STATE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

March 2011

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY SUPERVISION BY Xiaomi Chen ENTITLED Aberrant DNA Replication at an Ectopic Chromosomal Site in Human Cells BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Doctor of Philosophy.

_____

Michael Leffak, Ph.D.

Dissertation Director

_____

Gerald Alter, Ph.D.

Director, Biomedical Sciences

Ph.D. Program

_____

Andrew Hsu, Ph.D.

Dean, School of Graduate Studies

Committee on

Final Examination

_____

Steven Berberich, Ph.D.

_____

Paula Bubulya, Ph.D.

_____

John Paietta, Ph.D.

_____

Courtney Sulentic, Ph.D.

ABSTRACT

Xiaomi Chen. Ph.D., Biomedical Sciences Program, Wright State University, 2011. Aberrant DNA Replication at an Ectopic Chromosomal Site in Human Cells

Aberrant DNA replication, including over-replication or under-replication may lead to life-threatening mutation or even cause human diseases. This thesis focused on three issues related to abnormal DNA replication in human chromosomes including: I) to define the function of DNA unwinding element (DUE) and DNA unwinding element-binding protein (DUE-B) to maintain an active c-myc replicator; II) to determine the role of trans-acting factors in defining a replication origin on human chromosomes; III) to investigate the mechanism by which hairpins affect DNA replication and instability of $(CTG)_n$•$(CAG)_n$ trinucleotide repeat tracts in human cells.

Our laboratory previously demonstrated that both DUE and DUE-B, are essential in c-myc DNA replication initiation. In part I, I have shown that the increased binding of DUE-B and Cdc45 correlated with the decrease of the DUE helical stability and increased origin activity for the chimeric c-myc/SCA10 replicators. However, tethered binding of DUE-B on a mutant c-myc replicator with DUE deletion could not confer the DNA replication activity.

In part II, I explored the induction of DNA replication origin via trans-acting factors. My data suggested that tethered binding of transcription factor, E2F1, could induce replication activity likely by changing the chromatin structure. Tethered binding of pre-RC component, Orc2 and Cdt1 also can induce replication origin activity while Mcm7 and Cdc45 could not. Unlike episomal

systems, our system revealed that the induction of replication origin activity on human chromosome also required the essential cis-acting elements including the DUE.

Although it is widely accepted that unstable triplet nucleotide repeat (TNR) caused instability, it still remains elusive how and when the hairpins form during DNA replication. In part III, engineered zinc finger nucleases (ZFNs) and small pool PCR (spPCR) were applied to probe the hairpin formation in vivo in human cells. In our system, it was demonstrated that $(CTG)_n$•$(CAG)_n$ repeat tracts could form hairpins on either lagging strand or leading strand template, and the formation of hairpins is DNA replication associated.

# TABLE OF CONTENTS


| | Page |
|---|---|
| INTRODUCTION……. | 1 |
| Diversified DNA replication origins……..…….. | 1 |
| Simple replicon model in bacteria……..…… | 1 |
| Replication origins in unicellular eukaryotic cells……. | 2 |
| Replication origin in mammals…….……..… | 5 |
| c-myc origin……..…..…… | 5 |
| Conservative trans-acting factors of DNA replication origin……..... | 8 |
| Modulation of DNA replication activity in mammalian cells……...…. | 15 |
| Cis-acting elements and origin activity……..…..…….. | 15 |
| The role of trans-acting factors in replication origin activity…….…… | 15 |
| Transcription factor……..…….…….……. | 15 |
| Directed binding of components of pre-RC ……..……..……. | 17 |
| Aberrant DNA replication associate with human genomic instability…… | 18 |
| Triple nucleotide-repeat (TNR) instability is associated with human genetic diseases……..……...……. | 18 |
| Replication activity affects $(CTG)_n$•$(CAG)_n$ instability……..... | 18 |

DNA replication associated $(CTG)_n \bullet (CAG)_n$ instability in vivo in human cells……. 24

Significance……..…….……..…….…… 25

MATERIALS AND METHODS……. 26

Construction of plasmids……..……..……. 26

Plasmids expressing $GAL4^{DBD}$ fusion proteins ……..……... 26

Plasmids for cell line construction ……..……..…….. 26

Plasmids expressing ZFN and ZFP…….…….. 27

DNA sequencing……...…..…….……..…. 27

Construction of site-specific integration cell lines……..…… 27

Cell culture ……..……. 28

Transfection ……..……..……. 28

Flow cytometry……..……..…… 29

Nascent DNA abundance analysis (NDAA)…… 29

Polymerase Chain Reaction (PCR)……..……..…….. 30

Standard PCR……..……...……. 30

Quantitative Real-time PCR……...…… 30

Small Pool PCR (spPCR)……..……..……. 31

Western Blotting……..……..…… 33

Chromatin immunoprecipitation (ChIP)…….…….. 33

Isolation of cross-linked nuclei……... 33

Prepare cross-linked chromatin…….. 34

Immunoprecipitation……. 34

Design of zinc finger nucleases (ZFNs) ……. 35

Purification of ZFP and ZFN by immunoprecipitation ……..…… 36

DNA substrates used for testing ZFN activity in vitro……..……. 37

Assay of enzymatic activity of ZFNs in vitro……..... 37

Assay of enzymatic activity of ZFNs in vivo……... 37

Cytotoxic analysis of ZFNs with trypan blue staining…….…….…..... 38

RESULTS…….. 39

Part I. Function of DUE in an active chromosomal replicator……... 39

Binding of DUE-B on c-myc/SCA10 chimerical DUEs ……..…. 40

Making site specific cell line with GAL4/DUE substitution…….…. 48

Construction GAL4$^{DBD}$-DUE-B fusion protein expression plasmid… 54

Binding of GAL4$^{DBD}$–DUE-B on endogenous c-myc and ectopic GAL4SubDUE mutant origin…… 54

Tethered binding of DUE-B did not recruit Cdc45 in GAL4SubDUE mutant origin…..... 55

Tethered binding of DUE-B could not recover replication activity..... 55

SUMMARY I……. 63

Part II. Induction of replication activity by trans-acting factors on an ectopic chromosomal site…….. 64

Expression GAL4$^{DBD}$ fusion protein in HeLa cells ..... 64

Binding of trans-acting factors on the ectopic c-myc replicator in HeLa cells…….. 65

Binding of endogenous trans-acting factors at the ectopic mutant c-myc replicators in HeLa cells…… 66

Binding of exogenous GAL4$^{DBD}$ fusion protein on endogenous c-myc origin in HeLa cells.....…. 75

Tethered binding of E2F1 induced replication activity..... 78

GAL4$^{DBD}$-E2F1 binding altered chromatin structure..... 78

E2F1 binding recruited pre-RC components, Orc2, Mcm7, Cdt1 and pre-IC component, DUE-B and Cdc45.....…. 78

E2F1 binding stimulated DNA replication activity..... 79

Pre-RC components..... 86

Tethered Orc2 binding recruited Cdt1, Mcm7, DUE-B and Cdc45, and recovered the DNA replication activity.....….. 86

Tethered Cdt1 binding recruited Orc2, Mcm7, DUE-B and Cdc45,

and recovered the DNA replication activity..... 86

Tethered binding of Mcm7 could not recruit Orc2, DUE-B or Cdc45 and failed recover DNA replication activity 96

Pre-IC component..... 101

Tether binding of Cdc45 could not recruit Orc2, Mcm7, DUE-B and failed to recover DNA replication activity…….. 101

Summary II ……..…….. 106

Part III. Probing the formation of hairpins by $(CTG)_n$•$(CAG)_n$ repeat tracts during DNA replication in human cells in vivo…… 107

Assembly of three-zinc finger nuclease targeting on TNR repeat tract..... 107

Expression of ZFNs and ZFPs in HeLa cell..... 108

Function of $ZFN_{CTG}$ and $ZFN_{CAG}$ on TNR in vitro…….. 118

Function of $ZFN_{CTG}$ and $ZFN_{CAG}$ on TNR in vivo…….. 121

$(CTG)_{12}$•$(CAG)_{12}$ was able to be cleaved by double nuclease, $ZFN_{CTG}$+$ZFN_{CAG}$..... 121

$(CTG)_{12}$•$(CAG)_{12}$ could not form hairpins capable to be cleaved by $ZFN_{CAG}$ or $ZFN_{CTG}$ individually……. 122

$(CTG)_{102}$•$(CAG)_{102}$ could form hairpins which were cleaved by

either $ZFN_{CTG}$ or $ZFN_{CAG}$ individually..... 125

$(CTG)_{102}$•$(CAG)_{102}$ could be cleaved by double nuclease, $ZFN_{CAG}$+$ZFN_{CTG}$ together..... 125

Hairpin forming is DNA replication associated..... 130

Summary III..... 133

DISCUSSION..... 136

1. DUE is an essential functional element for an active replicator..... 136

2. Binding of DUE-B on chromatin is ORC dependent..... 138

3. Transcription factor, E2F1, induced DNA replication origin activity… 140

4. Binding of Orc2 and replication activity..... 143

5. Tethered Cdt1 could recruit ORC ..... 145

6. Tethered binding of Mcm7 and DNA replication activity…… 147

7. Tethering of Cdc45 could not induce replication origin activity 149

8. DNA sequence binding specificity and cell toxicity of zinc finger nuclease……..... 150

9. Hairpins form on both leading strand and lagging strand templates within the $(CTG)_n$•$(CAG)_n$ in human cells…… 151

REFERENCES……. 153

# LIST OF FIGURES


| Figure | | Page |
|---|---|---|
| 1. | Replication origin in budding yeast…… | 3 |
| 2. | c-myc replicator …… | 6 |
| 3. | Cis- and trans-acting elements in c-myc replicator……. | 11 |
| 4. | Model of DNA replication initiation in human cells…….. | 13 |
| 5. | Microbial model of instability of $(CTG)_n$•$(CAG)_n$..... | 22 |
| 6. | DUE is an essential element in the c-myc replicator…….. | 42 |
| 7. | Binding of DUE-B and Cdc45 on the ectopic SCA10DUE/mycDUE ……. | 44 |
| 8. | SCA10 DUEs can functionally replace c-myc DUE…… | 46 |
| 9. | Making HeLa406h/pFRT.mycGAL4SUBDUE cell line…….. | 50 |
| 10. | Analytical PCR and qPCR to identify single copy site-specific integration cell lines…… | 52 |
| 11. | Expression of functional $GAL4^{DBD}$-DUE-B…….. | 57 |
| 12. | Tethered binding of $GAL4^{DBD}$-DUE-B does not recruit Cdc45 .. | 59 |
| 13. | Tethered $GAL4^{DBD}$-DUE-B could not recover origin activity….. | 61 |

14. Diagram of expression plasmids of GAL4$^{DBD}$ fusion protein… 67

15. Binding of endogenous pre-RC components and pre-IC components on the ectopic c-myc replicator …….. 69

16. Binding of GAL4$^{DBD}$-E2F1 on the ectopic locus…….. 71

17. Replication activity of c-myc replicator and its mutants……. 73

18. Binding of GAL4$^{DBD}$ fusion proteins on endogenous c-myc replication origin……... 76

19. Tethered binding of E2F1 and histone acetylation ……. 80

20. Tethered binding of GAL4$^{DBD}$-E2F1 recruited pre-RC and pre-IC components to acceptor site…….. 82

21. Tethered binding of E2F1-GAL4$^{DBD}$ recovered replication origin activity……. 84

22. Tethered Orc2 binding recruited Mcm7, DUE-B and Cdc45….. 88

23. Tethered Orc2 binding recovered the DNA replication activity. 90

24. Tethered binding of Cdt1 recruited Orc2, Mcm7, DUE-B and Cdc45…….. 92

25. Tethered Cdt1 binding recovered the DNA replication activity.. 94

26. Tethered binding of Mcm7-GAL4$^{DBD}$ did not recruit Orc2, DUE-B and Cdc45……. 97

27. Tethered binding of Mcm7-GAL4$^{DBD}$ did not recover the DNA replication activity……. 99

28. Tethered binding of Cdc45-GAL4$^{DBD}$ could not recruit Orc2, Mcm7 and DUE-B……..…….. 102

29. Tethered binding of Cdc45-GAL4$^{DBD}$ could not recover the DNA replication activity…… 104

30. Design ZFNs targeting $(CTG)_n$•$(CAG)_n$ formed hairpins…… 110

31. Assembly of zinc fingers…….. 112

32. Expression of ZFPs and ZFNs in HeLa cells…… 114

33. Toxicity of ZFNs…… 116

34. Function analysis of ZFNs in vitro……. 119

35. In vivo effect of ZFN on $(CTG)_{12}$•$(CAG)_{12}$ cell lines……... 123

36. In vivo effect of ZFN on $(CTG)_{102}$•$(CAG)_{102}$ cell lines…… 126

37. In vivo effect of ZFN on $(CTG)_{102}$•$(CAG)_{102}$ cell lines…… 128

38. Formation of hairpins is replication dependent…… 131

39. Model of forming hairpins during DNA replication……... 134

## LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| ACS | ARS Consensus sequence |
| ARS | Autonomous Replication Sequence |
| ATP | Adenosine triphosphase |
| AT rich | Adenine-Thymine rich |
| bp | Base Pair |
| CDK | cyclin-dependent kinase |
| cDNA | Complementary DNA |
| ChIP | Chromatin Immunoprecipitation |
| CMV | Cytomegalovirus |
| CREB | cAMP Response Element Binding Protein |
| DDK | Dbf4 dependent kinase |
| DHFR | Dihydrofolate Reductase |
| DM1 | Myotonic Dystrophy Type 1 |
| DUE | DNA Unwinding Element |
| DUE-B | DUE Binding Protein |
| E. coli | Escherichia coli |
| FRAX | Fragile X Syndrome |
| FRT | FLP recombinase target site |
| $GAL4^{DBD}$ | GAL4 DNA Binding Domain |

| | |
|---|---|
| GC-rich | Guanine-Cytosine rich |
| GCV | Ganciclovir |
| GINS | Go-Ichi-Ni-San meaning 5-1-2-3, after the four related subunits of the complex Sld5, Psf1, Psf2 and Psf3 |
| HD | Huntington Disease |
| Hyg | Hygromicin |
| kb | Kilo Base |
| kDa | Kilo Daltons |
| MCM | Minimum Chromosomal Maintenance |
| NDAA | Nascent DNA Abundance Assay |
| Neo | Neomycin |
| NT | Nucleotide |
| ORC | Origin Recognition Complex |
| PCR | Polymerase Chain Reaction |
| Pre-IC | Pre-initiation Complex |
| Pre-RC | Pre-replication Complex |
| qPCR | Quantitative PCR |
| RPA | Replication Protein A |
| RSV | Rous Sarcoma Virus |
| SCA10 | Spinocerebellar Ataxia Type 10 |

| | |
|---|---|
| spPCR | Small Pool PCR |
| STS | Sequence Tagged Site |
| TK | Thymidine Kinase |
| TNR | Triple Nucleotide Repeat |
| WT | Wild Type |
| ZiFiT | Zinc Finger Targeter |
| ZFN | Zinc Finger Nuclease |
| ZFP | Zinc Finger Protein |

ACKNOWLEDGMENTS

I would like to thank Dr. Michael Leffak for providing help, guidance, and an excellent training environment, all of which has been invaluable in my development as a scientist. I would like also to recognize Drs. Malloy Ghosh and Guoqi Liu, for their assistance and collaborations, and for performing much of the background work that allowed me to propose and carry out many of the experiments in this dissertation. Lastly, the comments, criticisms, and ideas offered by all the other members of the Leffak laboratory, Shere´ Myers, Michael Kemp, Asif Chowdhury, Lubna Abu-Niaaj, Janhong Yao, Joanna Barthelemy and Yanzhe Gao, with whom I have worked are greatly appreciated.

**COPYRIGHT**

**XIAOMI CHEN**

**2011**

# INTRODUCTION

DNA replication is one of the most fundamental aspects of cellular metabolism in humans. During each cell division cycle, the human genome duplicates itself only once (Machida et al., 2005). This delicate process is precisely regulated by cis-acting elements and trans-acting factors. Cis-acting elements refer to replication origins, DNA sequences, from which DNA starts to copy itself. Trans-acting factors are a group of proteins, which participate in the activation of DNA replication. Normal DNA replication is critical to maintain genomic integrity in healthy life while aberrant replication may lead to life-threatening mutation or even cause human diseases (Mirkin, 2006; Nasheuer, 2010).

## Diversified DNA replication origins

### *Simple replicon model in bacteria*

The replicon model proposed by Jacob and Brenner (Jacob and Brenner, 1963) hypothesized that a specific *cis*-acting element was activated by a *trans*-acting factor (protein) to initiate DNA replication. This model suggested that DNA sequences ultimately determine replication start sites, and that protein factors provided the triggering mechanism for starting the synthesis of new DNA

molecules.

***Replication origins in unicellular eukaryotic cells***

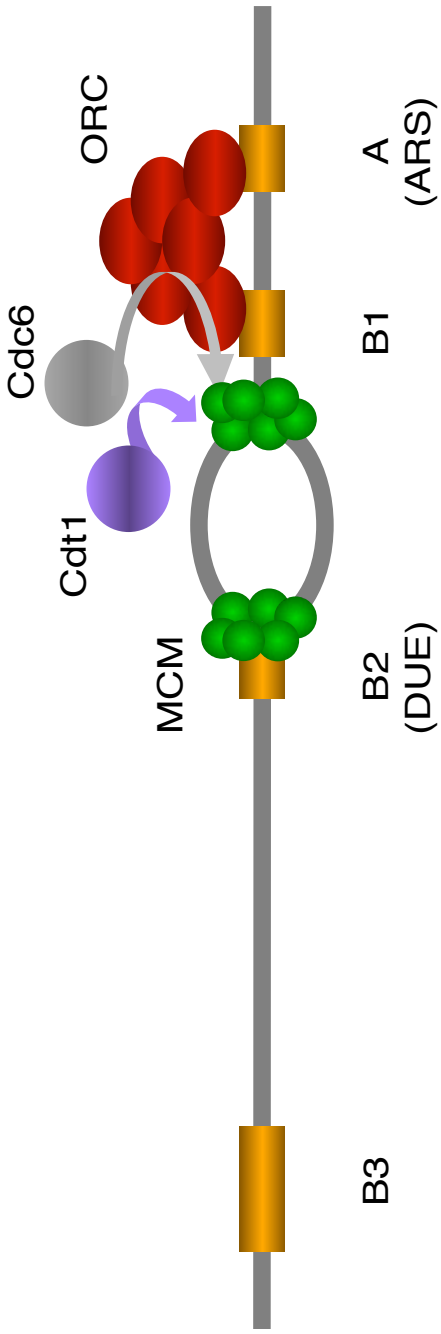
Replication in eukaryotic cells, even in the budding yeast *Saccharomyces cerevisiae,* is much more complicated than that in bacteria (Dubey et al., 1996). In general, replication origins in *S. cerevisiae* are defined by specific DNA sequence composed of multiple necessary sequence elements. The best characterized example of a yeast origin is ARS1 (Marahrens and Stillman, 1992) (Figure 1). ARS1 is ~125 bp long and is composed of 4 necessary sequence elements: the ARS consensus sequence (ACS), B1, B2, and B3 elements. The ACS sequence is an 11 bp sequence that is conserved with at least a 9/11 bp match in all *S. cerevisiae* origins. The origin recognition complex (ORC) binds to the ACS and B1 elements and recruits additional replication factors including Cdc6, Cdt1, MCM and Cdc45 (Bryant et al., 2001; Cocker et al., 1996; Gavin et al., 1995; Hardy and Pautz, 1996; Rowley et al., 1995; Santocanale and Diffley, 1996). The B2 element functions as a DNA unwinding element and may also serve other roles. The B3 element binds the transcription factor Abf1 and maintains a specific chromatin conformation at the origin. New DNA synthesis begins at a specific nucleotide between the B1 and B2 elements (Diffley, 1994).

Unlike in budding yeast, replication origins in fission yeast are far different (Clyne and Kelly, 1995; Dai et al., 2005; Dubey et al., 2010; Zhu et al., 1992). Fission yeast origins tend to be larger in size (~450 bp) and lack consensus

**Figure 1. Replication origin in budding yeast**

A DNA replication origin in budding yeast consists of several essential elements. Element A contains an asymmetric A:T-rich “ARS consensus sequence” that is required for origin activity. Element B1 facilitates A in binding ScORC (S. cerevisiae origin recognition complex*)*. Element B2 is a weak ScORC binding site. Element B3 (~22 bp), binds transcription factor Abf-1.

In order to initiate DNA replication, a pre-replication complex (pre-RC) must form on the origin, which requires several protein factors involving in the DNA replication initiation. ORC: origin recognition complex; Cdc6 and Cdt1: replication origin licensing factors recruited by ORC. MCM: putative DNA helicase to unwind DNA duplex.




Pre-replicative Complex (pre-RC)

sequences (Dubey et al., 1996). AT-richness is the only common feature found in most of fission yeast origins (Gomez and Antequera, 1999; Kim and Huberman, 1999; Okuno et al., 1999), which facilitates binding of the spOrc4 subunit of the origin recognition complex (ORC) through its AT hook (Chuang and Kelly, 1999; Lee et al., 2001)

***Replication origin in mammals***

Although recent high-throughput methods have identified more replication origins in mammals (Karnani et al., 2010; Lucas et al., 2007), a limited number of mammalian DNA replication origins have been well characterized. In general, this small number of replication origins can be grouped into two categories, one group of origins presents as a large initiation zone (~55 kb) containing multiple inefficient replication starting sites. The Chinese hamster dihydrofolate reductase (DHFR) replication origin (Burhans et al., 1991; Hamlin et al., 1992; Heintz et al., 1983) is an example of this group. Another group of mammalian replication origins tends to be a short DNA fragment with high efficiency like the well characterized human replication origins, c-myc (Berberich et al., 1995; Trivedi et al., 1998; Waltz et al., 1996), lamin B2 (Biamonti et al., 1992; Falaschi et al., 2007), β-globin (Aladjem et al., 1998; Wang et al., 2004) and DBF4 (Romero and Lee, 2008).

***c-myc origin***

Our laboratory first mapped that DNA replication initiated in a 12 kb region containing 5' portion of the c-myc gene in human cells (Leffak and James, 1989) (Figure 2). Later work demonstrated that there are multiple potential initiation

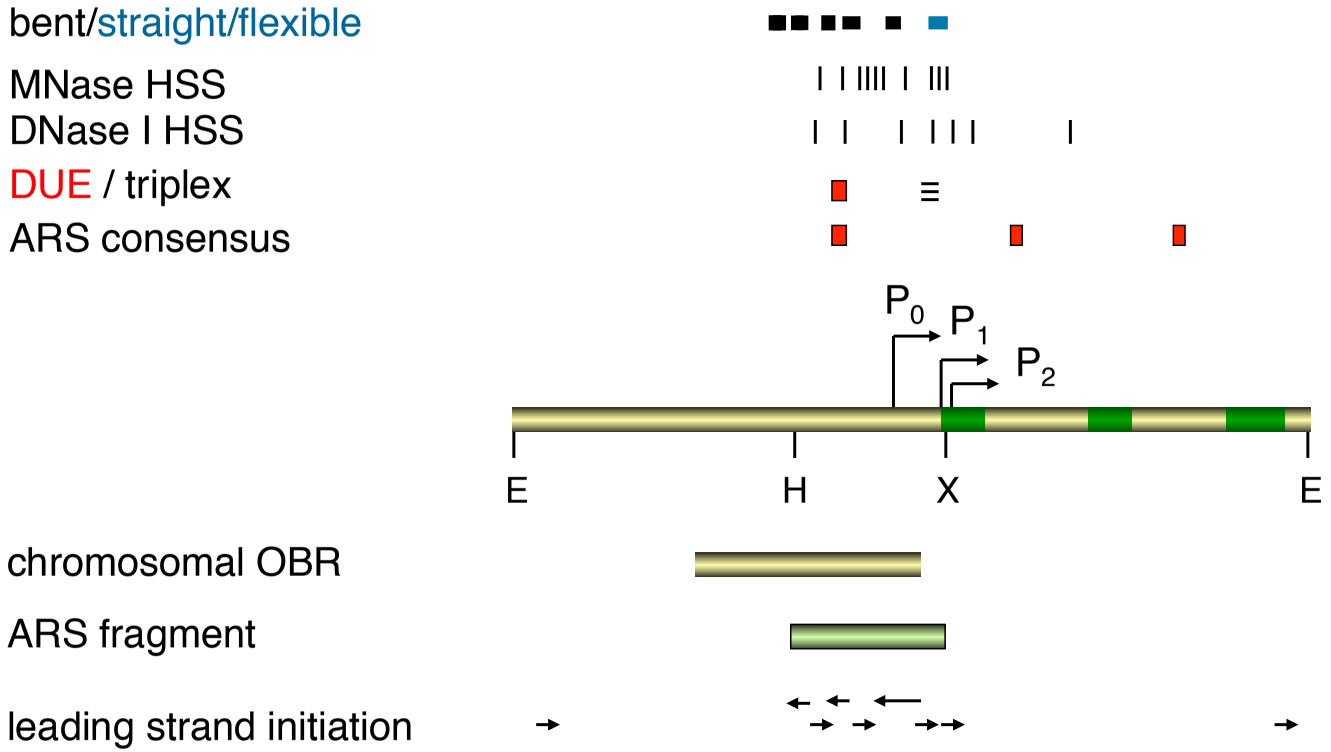
**Figure 2. c-myc replicator**

The c-myc replication origin was mapped in the 5' promoter region of c-myc oncogene locus in 12 kb EcoRI to EcoRI DNA fragment. DNA replication initiates from 2.4 kb core region (HindIII to XhoI fragment). Green bars are exons of the c-myc gene. E: EcoR I; H: HindIII; X: XhoI. $P_1$, $P_2$ and $P_0$ indicate the sites of three promoters. Arrows around $P_1$, $P_2$ and $P_0$ indicate the transcription directions.

Structure features are shown above the diagram of c-myc locus including bent/straight/flexible sequence, DNA unwinding element, triplex forming sequences, micrococcol nuclease/DNase I nuclease hypersensitive sites and three 9/11 matches to the ARS consensus sequence.

Biological features are shown below the diagram of c-myc replicator including the chromosomal origin bi-direction replication fragment, ARS (autonomously replicating sequence) fragment and leading strand synthesis sites and directions.

# Human c-myc locus



sites scattered throughout the c-myc locus (Trivedi et al., 1998), among which a 2.4 kb region immediately upstream of the c-myc gene was a replicator when moved to an ectopic chromosomal site (Liu et al., 2003; Malott and Leffak, 1999). Mutational analyses of the c-myc replicator demonstrated that multiple functional elements including numerous transcription factor-binding sites, a promoter region, a DNA unwinding element and triplex forming sequences are essential to maintain c-myc origin activity at an ectopic chromosomal site (Figure 2)(Liu et al., 2003). Recently, the binding sites of pre-RC, including ORC, MCM and pre-IC components, DUE-B and Cdc45, are detected through ChIP assay (Figure 3) (Chowdhury et al., 2010; Ghosh et al., 2006).

**Conservative trans-acting factors of DNA replication origin**

Figure 4 illustrates the sequential binding of trans-acting factors in eukaryotic cells during DNA replication initiation. The origin recognition complex (ORC) is a primary eukaryotic replication initiation factor required for the replication of chromosomal DNA. ORC was first identified from yeast and comprises six interacting subunits (Orc1-6) that define the sites along chromosomes from which replication can initiate (Bell et al., 1993; Diffley and Cocker, 1992). Homologs of the ORC subunits have been identified in all eukaryotic replication model systems (Romanowski et al., 1996; Takahara et al., 1996). It is believed that ORC binding on chromatin is the first step of DNA replication initiation (Bell and Dutta, 2002).

Upon loading on chromatin, ORC recruits two other important factors,

Cdc6 and Cdt1. Together, Cdc6 and Cdt1 function with ORC to load the mini-chromosome maintenance (MCM) double hexamers onto the replication origins to form the pre-replication complex (pre-RC)(DePamphilis et al., 2006). MCM, a putative replicative helicase, is composed of six subunits termed the MCM complex (Romanowski et al., 1996). The six subunits (Mcm2-7) are related to one another and have been conserved throughout eukaryotes and archaea (Chong et al., 2000). The MCM complex is required not only for the initiation of DNA replication but also for the continued elongation phase of DNA synthesis (Aparicio et al., 2006; Gambus et al., 2006; Moyer et al., 2006). Upon the loading of MCM, pre-RC formation is completed.

DNA unwinding element binding protein (DUE-B) (Casper et al., 2005; Kemp et al., 2007) is a recently discovered essential pre-IC component and binds on chromatin after recruitment of MCM in an ORC dependent manner. Our preliminary data demonstrated that DUE-B functioned as a human ortholog of yeast Sld3 required for recruitment of Cdc45 during the formation of pre-IC in yeast (Kamimura et al., 2001; Yabuuchi et al., 2006). Most recently, immunoprecipitation (IP) and ChIP assay demonstrate that DUE-B is closely associated with Cdc45, a component of pre-initiation complex (Chowdhury et al., 2010; Fu and Walter, 2010). Meanwhile, three other groups (Balestrini et al., 2010; Fu and Walter, 2010; Kumagai et al., 2010; Sansam et al., 2010) also claimed the finding of the metazoan orthologs of yeast Sld3. The first group found GEMC1, a TopBP1-interacting protein required for chromosomal DNA replication, which initiates chromosomal DNA replication in multicellular

organisms by mediating TopBP1- and Cdk2-dependent recruitment of Cdc45 onto replication origins (Balestrini et al., 2010). The second group identified a novel protein, Treslin, which collaborates with TopBP1 in triggering the initiation of DNA replication by Cdk2-mediated loading of Cdc45 onto replication origins (Kumagai et al., 2010). The third group identified ticrr (for TopBP1-interacting, checkpoint, and replication regulator), which acts in association with TopBP1 and plays an essential role in pre-IC formation (Sansam et al., 2010). Currently, more work needs to be done to in order to confirm which protein is the Sld3 ortholog in metazoan.

Recruited to the pre-RC in a cyclin-dependent kinase (CDK)- and Dbf4 dependent kinase (DDK)- dependent manner, Cdc45 is an essential protein for the initiation of DNA replication interacting with both components of the pre-RC as well as DNA polymerases (Bell and Dutta, 2002). Cdc45 may serve as a bridge between the proteins that select origin sites with the actual replication machinery. Several lines of evidence demonstrated that Cdc45, MCM and GINS forms CMG complex to activate the helicase activity of MCM (Ilves et al., 2010; Moyer et al., 2006). Binding of Cdc45 was mediated by Sld3 in yeast (Kamimura et al., 2001; Kanemaki and Labib, 2006; Nakajima and Masukata, 2002; Pospiech et al., 2010; Takayama et al., 2003; Tanaka et al., 2007; Yabuuchi et al., 2006; Yamada et al., 2004), indicating the formation of pre-initiation complex (pre-IC) (Lei and Tye, 2001).

**Figure 3. Cis- and trans-acting elements in c-myc replicator**

Origin activity of c-myc mutants at an ectopic chromosomal site was measured by relative nascent DNA abundance (vertical black cylinders). Wt, nascent strand abundance in cells containing wild-type c-myc DNA (pFRT.myc); Δ1 to Δ12, nascent strand abundance in cells containing mutant c-myc DNA (406h/pFRT.myc. (Δ1 to Δ12) (Liu et al., 2003).

Shown in upper box are approximate locations of different transcription factor-binding sites in the c-myc origin core DNA aligned with the positions of origin deletions (Δ1 to Δ12) (Liu et al., 2003). p, PuF; m, c-Myb; n, NF1; s, SP1; a, AP-2, e, Ets; f, far upstream element binding protein. $P_0$ and $P_1$, c-myc promoters.

Shown under the 2.4 kb c-myc core (blue bar) are the approximate binding sites of pre-RC components, ORC, Cdc6 and MCM (Ghosh et al., 2006) and pre-IC components, DUE-B and Cdc45 (Chowdhury et al., 2010).

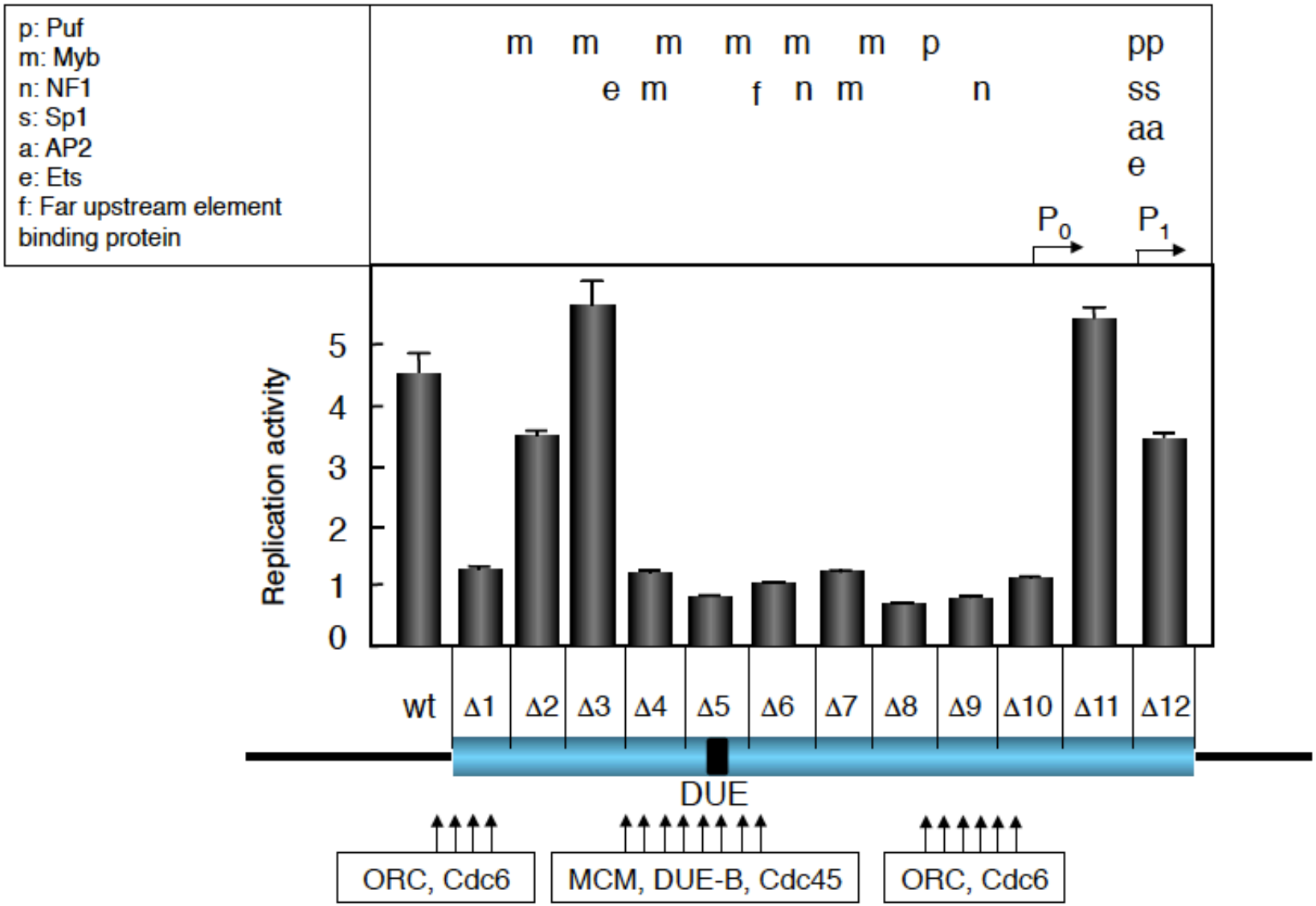
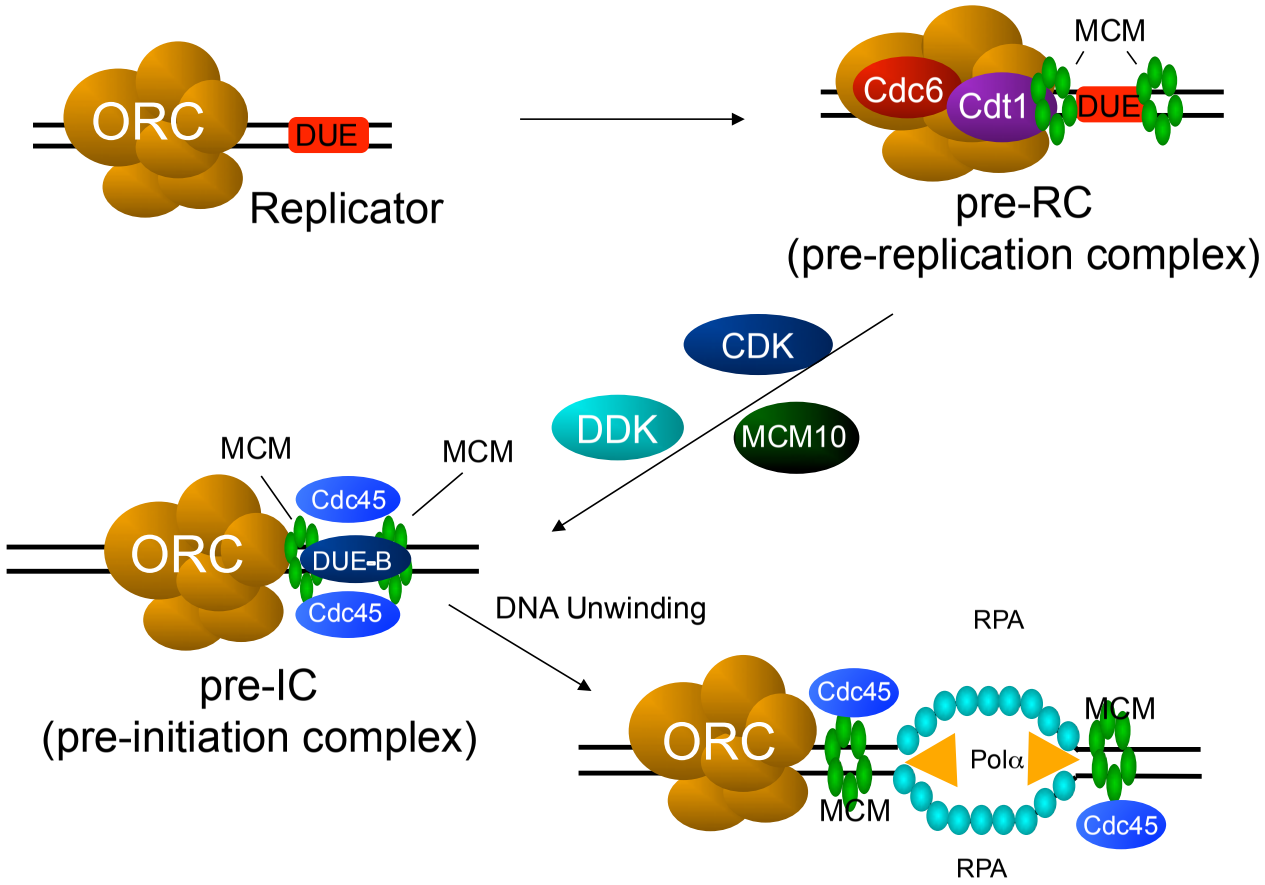

p: Puf
m: Myb
n: NF1
s: Sp1
a: AP2
e: Ets
f: Far upstream element binding protein
m m m m m m p pp
e m f n m n ss
aa
e
P0
P1
Replication activity
0
1
2
3
4
5
wt
Δ1
Δ2
Δ3
Δ4
Δ5
Δ6
Δ7
Δ8
Δ9
Δ10
Δ11
Δ12
DUE
ORC, Cdc6
MCM, DUE-B, Cdc45
ORC, Cdc6

**Figure 4. Model of DNA replication initiation in human cells**

Origin recognition complex (ORC), a heterohexamer (Orc1-6), binding on the replication origin is the first step. Cdc6 and Cdt1 are further recruited by ORC to chromatin. Double heterohexamer MCM (2-7) is later recruited to chromatin site near DUE, which marks the end of the assembly of pre-replication complex (pre-RC).

Catalyzed by Dbf4 dependent kinase (DDK) and cyclin-dependent kinase (CDK), GINS (from the Japanese go-ichi-ni-san meaning 5-1-2-3, after the four related subunits of the complex Sld5, Psf1, Psf2 and Psf3), DUE-B, Cdc45 and TopBP1 are further recruited to pre-RC to form pre-initiation complex (pre-IC). MCM10 is believed to be involved in this step. Cdc45 activates MCM helicase activity leading to the unwinding of DNA double helix. Replication protein A (RPA) coats single stranded DNA, and the primase, DNA polymerase alpha (Pol $\alpha$), binds on RPA-coated single stranded DNA, and the synthesis of new DNA molecules was initiated.




DNA Replication Initiation

## Modulation of DNA replication activity in mammalian cells

### Cis-acting elements and origin activity

Cis-acting elements play an essential role in defining mammalian DNA replication origins. Several groups have demonstrated that DNA replication activity is eliminated by deletion of essential elements in lamin B2 (Paixao et al., 2004), β-globin (Aladjem et al., 1998), DHFR (Altman and Fanning, 2004), and c-myc origins (Figure 3) (Liu et al., 2003). Even though DUEs (AT rich sequences) were often found within DNA replication origins (Gilbert, 2004), consensus DNA sequence from different mammalian DNA replication origins have not been found. Moreover, the DUE can be functionally substituted by some heterogeneous DUE with the equivalent low helical stability in ori-beta of DHFR (Altman and Fanning, 2001) and c-myc origins (Liu et al., 2007).

## The role of trans-acting factors in replication origin activity

### Transcription factors

The interplay between transcription and DNA replication in mammalian cells is well documented. Transcription of DHFR and neighboring genes prevents replication initiation within the transcribed regions, but this block is suppressed upon treatment of the cells with a transcription inhibitor. It was also demonstrated that transcription termination signals at the 3' end of the DHFR gene block transcription elongation into the intergenic region, and deletion of these signals

causes a loss of replication initiation in the entire intergenic region (Kalejta et al., 1998; Mesner et al., 2003). Using a Tet-on inducible transcription system, our laboratory (Ghosh et al., 2004) demonstrated that basal level transcription favored c-myc replication origin activity, while induced transcription, on the contrary, repressed DNA replication locally in human cells. Another group (Danis et al., 2004) reported that a site-specific DNA replication origin can be induced by the assembly of an active transcription domain by tethering $GAL4^{DBD}$-VP16 to a plasmid promoter region in *Xenopus* egg extract. This activation of the origin is due to the local acetylation of histones. This result agrees with the observation of Ghosh et al., which demonstrated that binding of transcription factor cAMP response element-binding protein (CREB) itself activated an inactive c-myc DNA replication origin but did not activate origin activity at an non-origin site in human cells in vivo (Ghosh et al., 2004).

Previously our laboratory found the histone deacetylase inhibitor trichostatin A (TSA) altered the pattern of DNA replication origin activity in human cells via modulation of chromatin structure (Kemp et al., 2005), which is consistent with the finding that chromatin regulates origin activity in Drosophila follicle cells (Aggarwal and Calvi, 2004). These results suggest that nucleosome acetylation and other epigenetic changes are important modulators of origin activity in metazoan.

Most recently, the role of chromatin acetylation in activation of replication origin activity in human cells has been supported by several new findings. A group reported that HBO1, an H4-specific histone acetylase, directly interacts

with Cdt1 and enhances Cdt1-dependent rereplication (Miotto and Struhl, 2010). The second group claimed that histone acetylation promoted MCM loading via enhanced chromatin accessibility, which was regulated positively by Cdt1 and HBO1 in G1 and repressed by Geminin-HDAC11 association with Cdt1 in S-phase (Wong et al., 2010). Taken together, the current data tend to support the influence of chromatin structure on modulation of replication origin activity (Gerbi and Bielinsky, 2002).

**Directed binding of components of pre-RC**

It has been demonstrated that pre-RC components, initiator Orc2 or Cdc6, fused to the DNA binding domain of the yeast GAL4 protein could stimulate extra chromosomal replication of plasmids containing $GAL4^{DBD}$ binding sites in human cells. Other pre-RC components were also recruited to the same sites in human cells (Takeda et al., 2005). This experiment strongly indicated that artificial recruitment of eukaryotic cellular replication initiation factors to a DNA sequence was able to create a functional origin of replication on plasmids in human cells. Because mammalian chromosomes possess much more complicated structures than plasmid chromatin, and may have many differences from plasmid in DNA replication initiation. As described below, I used the GAL4 tethering system, and found that Orc2 and Cdt1 could activate replication activity at an inactive chromosomal site adjacent to an inactive replication origin, but could not activate replication origin activity on the same chromosomal site without inactive c-myc replication origin. Moreover, Mcm7 and Cdc45 could not recover DNA replication

origin activity at an inactive chromosomal site adjacent to a DNA fragment containing AT-rich DNA unwinding elements using the same system.

## Aberrant DNA replication associates with human genomic instability

### *Triple nucleotide-repeat (TNR) instability is associated with human genetic diseases*

Simple DNA repeat expansions can cause nearly 30 human hereditary diseases including fragile X syndrome (FRAX) (Richards et al., 1993), Huntington's disease (HD) (MacMillan et al., 1993) and myotonic dystrophy type 1 (DM1) (Sinden et al., 2002). As shown in Table 1 (Sinden et al., 2002), $(CTG)_n$•$(CAG)_n$ repeat tract expansion is associated with about a dozen human neurodegenerative and neuromuscular diseases. Most of these diseases shared the following features: patients exhibit an earlier age of onset and increased disease severity in successive generations (genetic anticipation) (Mirkin, 2007); the length of the expanded repeat tract is critical in the pathogenic mechanism and is one of the most important determinants of disease severity; expanded repeats in many diseases show both somatic and germline instability, usually with a bias toward expansion (Pearson et al., 2005).

## Replication activity affects $(CTG)_n$•$(CAG)_n$ instability

Several lines of evidence from bacterial (Bowater et al., 1997; Fieck et al., 1992; Hebert and Wells, 2005; Pan, 2006), yeast (Miret et al., 1997; Pelletier et al., 2003; Yang and Freudenreich, 2010), and transgenic mouse systems

(Monckton et al., 1997; van den Broek et al., 2007) suggest multiple mechanisms leading to $(CTG)_n$•$(CAG)_n$ instability including DNA replication, DNA repair and recombination. In plasmid-based model systems (Cleary et al., 2002), the role of DNA replication in $(CTG)_n$•$(CAG)_n$ instability has been suggested by the observation that the proximity and direction relative to a neighboring DNA replication origin can influence $(CTG)_n$•$(CAG)_n$ instability (Cleary et al., 2002). The effects of replication polarity and origin proximity, and the association of $(CTG)_n$•$(CAG)_n$ instability with cell proliferation are readily detected in population assays and strongly support the role of replication in this process.

In microbial and yeast model systems, it has been suggested that replication affects $(CTG)_n$•$(CAG)_n$ instability due to the formation of unusual DNA structures, including hairpins which cause DNA replication forks to pause, collapse, or break (Darlow and Leach, 1995; Henricksen et al., 2000; Hou et al., 2009; Mirkin, 2007). As shown in Figure 5, hairpins can form within the portion of single stranded $(CTG)_n$•$(CAG)_n$ slipped out from normal double helix structure. The effect of hairpins on $(CTG)_n$•$(CAG)_n$ instability will depend on their locations, i.e. hairpins on template strand may cause contraction, and hairpins on nascent DNA strand will lead to more expansion. It is believed that hairpins may form on lagging strand template or leading strand template and newly synthesized lagging strand DNA or newly synthesized leading strand DNA. However, it is still unclear about the location on which $(CTG)_n$•$(CAG)_n$ hairpins form in bacteria (Pearson et al., 1998; Wells et al., 2005) and yeast system (Freudenreich et al., 1997; Pelletier et al., 2003). Little is known about the mechanism of

$(CTG)_n$•$(CAG)_n$ instability in human cells. Most recently, we revealed the formation of hairpin in human cells during DNA replication (Liu et al., 2010).

## Table 1. Human diseases caused by expanded $(CTG)_n \bullet (CAG)_n$

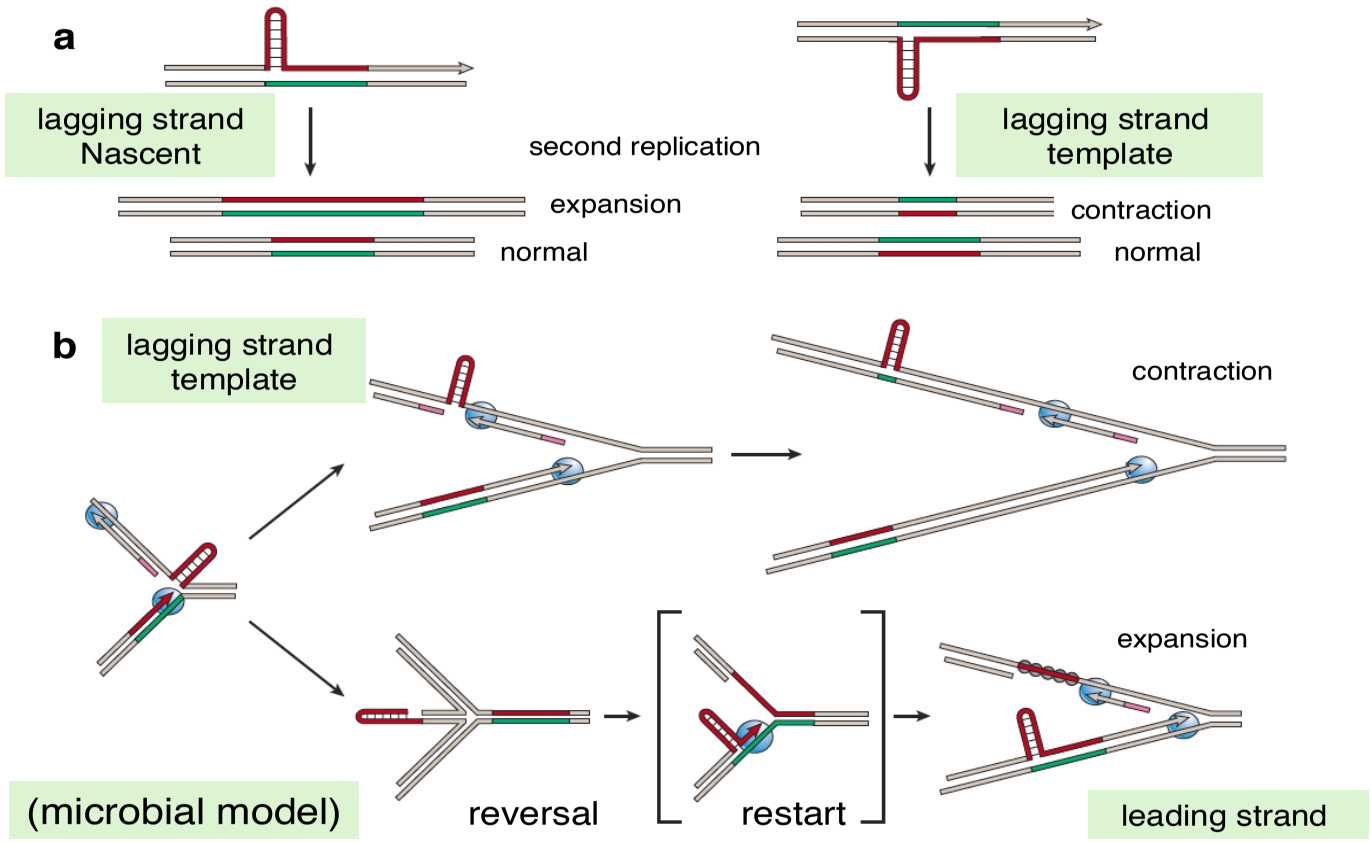
| Disease | Gene | Locus | Repeat[a] | Repeat length | | | Protein/possible biological effect of expansion |
|---|---|---|---|---|---|---|---|
| | | | | Normal | Pre-mutation | Disease | |
| Spinobulbar muscular atrophy (SBMA) (Kennedy's disease) | AR | Xq13–21 | $(CAG)_n$ | 14–32 | ? | 40–55 | Androgen receptor (AR)/polyglutamine tract expansion |
| Myotonic dystrophy type 1 (DM1) | DMPK | 19q13 | $(CTG)_n$ | 5–37 | 50–80 | 80–1000 | Myotonic dystrophy protein kinase/altered mRNA processing and transport, altered gene expression due to chromatin changes |
| Huntington disease | HD | 4p16·3 | $(CAG)_n$ | 10–34 | 36–39 | 40–121 | Huntingtin/polyglutamine expansion |
| Spinocerebellar ataxia 1 | SCA1 | 6p23 | $(CAG)_n$ | 6–44 | – | 39–82 (pure) | Ataxin-1/polyglutamine expansion |
| Spinocerebellar ataxia 2 | SCA2 | 12q24·1 | $(CAG)_n$ | 14–31 | – | 34–59 (pure) | Ataxin-2/polyglutamine expansion |
| Spinocerebellar ataxia 3 (Machado Joseph disease) | SCA3 | 14q32·1 | $(CAG)_n$ | 13–44 | NA | 55–84 | Ataxin-3/polyglutamine expansion |
| Spinocerebellar ataxia 6 | SCA6 | 19p13 | $(CAG)_n$ | 4–18 | NA | 21–33 | $\alpha_{1A}$-voltage-dependent calcium channel subunit/polyglutamine expansion |
| Spinocerebellar ataxia 7 | SCA7 | 13p12–13 | $(CAG)_n$ | 4–34 | NA | 37–306 | Ataxin-7/polyglutamine expansion |
| Spinocerebellar ataxia 8 | SCA8 | 13q21 | $(CTG)_n$ | 15–27? | 110–200 | | 3′ to SCA8gene/alters gene expression of adjacent genes? |
| Spinocerebellar ataxia 12 | SCA12 | 5q31–33 | $(CAG)_n$ | 7–28 | ? | 66–78 | 5′ to SCA12 gene/alters gene expression? |
| Dentatorubral-pallidoluysian | DRPLA | 12p13·31 | $(CAG)_n$ | 7–25 | ? | 49–75 | Polyglutamine expansion |

Adapted from Sinden et al. (2002). J Biosci

**Figure 5. Microbial model of instability of $(CTG)_n \bullet (CAG)_n$**

a. Hairpin associated with instability of $(CTG)_n \bullet (CAG)_n$. Formation of a hairpin on the nascent strand DNA results in repeat expansions after two rounds of replication (left panel), whereas the presence of the same structure on the template strand results in repeat contractions (right panel).

b. Replication fork stalling mediated instability of $(CTG)_n \bullet (CAG)_n$. When a hairpin forms on the lagging strand template two consequences may occur. If polymerase goes through hairpin smoothly contraction in lagging strand nascent DNA occurs. Alternatively, hairpins cause fork stalling and generate a chicken-foot structure, which mediates the hairpin formation in a nascent leading strand DNA resulting in expansion.




(Adapted from Mirkin, Nature, 2007)

**DNA replication associated trinucleotide repeat $(CTG)_n$•$(CAG)_n$, (TNR), instability in vivo in human cells.**

During recent years, our laboratory has developed a novel human cell culture system, in which a replicator could be connected with microsatellite, a repeating sequence in DNA, including $(ATTCT)_n$• $(AGAAT)_n$ or $(CTG)_n$•$(CAG)_n$ with different length of repeats and different orientations relative to the direction of replication. This system allowed allowing us to carry out systematic studies of DNA replication associated microsatellite instability using HeLa cells (Liu et al., 2007; Liu et al., 2010; Liu et al., 2003). Our data demonstrate that $(CTG)_n$•$(CAG)_n$ instability is determined by repeat length, adjacent replication origin activity, $(CTG)_n$•$(CAG)_n$ orientation, and tissue culturing time indicating our system could mimic the TNR instability similar to a human disease scenario (Liu et al., 2010).

In this thesis, I used synthetic zinc finger nuclease technique (Benabdallah et al., 2010; Doyon et al., 2010; Hockemeyer et al., 2009; Katada and Komiyama, 2009; Lombardo et al., 2007; Mani et al., 2005; Orlando et al., 2010; Radecke et al., 2010) to probe the formation of hairpins in $(CTG)_n$•$(CAG)_n$ repeat tract in human cells in vivo (Liu et al., 2010). Our data showed that $(CTG)_n$•$(CAG)_n$ repeat tract formed hairpins during DNA replication on either leading strand template or lagging strand template, which demonstrated that the complexity of the instability of $(CTG)_n$•$(CAG)_n$ on human chromosome, may not be mimicked in other systems.

**Significance**

Since DNA replication is one of the most fundamental metabolic activities in living cells, it is worthwhile to know whether and how trans-acting factors can induce replication origin activity. There are multiple mechanisms serving as safety guards to prevent aberrant replication, which is extremely important for keeping the fidelity of genetic materials from one generation to the next. Although human cells have developed a sound self-protection machinery, numerous human diseases have been associated with the malfunction of DNA replication. It is no doubt that determining the replication origin activity in some circumstances, e.g. induction of trans-acting factors, may be important in gene therapy in diseases caused by the failure of DNA replication from the right orientation.

$(CTG)_n$•$(CAG)_n$ instability is associated with more than two dozen human neurodegenerative and neuromuscular diseases which constitute a significant health burden at the individual and national levels (Sinden et al., 2002). Thus, it is also significant to know the mechanism by which aberrant DNA replication causes $(CTG)_n$•$(CAG)_n$ instability. Unfortunately, results based on the current bacterial and yeast model systems have not given a consistent mechanistic explanation for $(CTG)_n$•$(CAG)_n$ instability in human cells. Our results will be extremely helpful to clarify the mechanisms of $(CTG)_n$•$(CAG)_n$ instability driven by the forming of hairpins during DNA replication in human cell lines.

# MATERIALS AND METHODS

## Construction of plasmids

### *Plasmids expressing GAL4$^{DBD}$ fusion proteins*

Mcm7 was cloned from human cDNA library by PCR. Orc2 and Cdt1 were subcloned from plasmids provided by the Dutta laboratory (Orc2)(Takeda et al., 2005) and Nishitani laboratory (Cdt1)(Nishitani et al., 2000). Orc2 and Mcm7 were fused to GAL4 DNA binding domain (GAL4$^{DBD}$, amino acids 1-147) at the N-terminus, and subcloned into pEGFP-N1 (Clontech). GAL4$^{DBD}$-Cdt1 was expressed in pcDNA3.1 vector (Invitrogen). Cdc45 was subcloned from plasmid provided by Kukimoto laboratory and fused to GAL4$^{DBD}$ at the N-terminus in plasmid pEGFP-N1. Cdc45 was fused with GAL4 DNA binding domain, Cdc45-GAL4$^{DBD}$. Plasmid pBS-GAL4-(E2F1) is a generous gift from Dr. Flemington, in which GAL4$^{DBD}$ was fused with full length E2F1 at its N-terminus (Flemington et al., 1992; Flemington et al., 1993). GAL4$^{DBD}$-DUE-B is constructed by fusion of GAL4$^{DBD}$ with human DUE-B at its N-terminus and subloned into plasmid pEGFP-N1 at *NheI* and *Sac II* restriction cutting sites.

### *Plasmids for cell line construction*

pFRT.GAL4 and pFRT.myc-GAL4subDUE are pFRT.myc (Malott and Leffak, 1999) derivative plasmids, which were constructed by replacement either 2.4 kb c-myc replicator or 200 bp DNA fragment containing c-myc DNA unwinding element (DUE) by five GAL4 DNA binding sites (5×GAL4), GAL4$^{DBD}$

binding sequences. pOG44 is a FLP recombinase expression vector purchased from Invitrogen.

***Plasmids expressing ZFN and ZFP***

The single zinc finger plasmids pc3XB-ZF72, pcXB-ZF83, and a mammalian expression vector, pST1374, were purchased from Addgene Inc. (Cambridge, MA). Proteins with three zinc fingers (ZFPs) were assembled using standard molecular cloning methods. $ZFP_{CTG,}$ containing three ZF72 zinc fingers (Addgene plasmid 13206), and $ZFP_{CAG}$, containing three ZF83 zinc fingers (Addgene plasmid 13217), were designed with Zinc Finger Targeter (ZiFiT) Version 2.0. ZFPs cDNAs were subcloned into the mammalian expression vector pST1374 (Addgene plasmid 13426) between the coding sequences of the FLAG tag and Fok I DNA cleavage domain to generate zinc finger nucleases, $pZFN_{CTG}$ and $pZFN_{CAG}$ (ZFNs). ZFP expression plasmids, $pZFP_{CTG}$ and $pZFP_{CAG}$ are derivative plasmids from $pZFN_{CTG}$ and $pZFN_{CAG}$, respectively, by adding a translational stop codon at BamH I restriction site between ZFP and Fok I DNA sequences.

**DNA sequencing**

The fragments of interest in all plasmids constructed in this thesis were sequenced by Retrogen Inc. or Lark Technologies.

**Construction of site-specific integration cell lines**

$4\times10^5$ cells of HeLa cell line 406h, an acceptor cell line harboring the FLP

recombination targeting (FRT) site, were seeded the day before transfection. Cells were co-transfected with total 1μg DNA of pOG44 and pFRT.GAL4subDUE in a 24-well plate. When transfection, the optimal molar ratio between pOG44 and pFRT.GAL4subDUE is about 9:1. Twenty-four hr after transfection, cells were trypsinized and transfered into plates with 15 cm diameter in Dulbecco's Modified Eagle Medium (DMEM) media with 500 μg/ml G418 for 14 days till single colonies formed. Colonies were screened with 5.5 μg/ml ganciclovir (GCV) from G418 resistant colonies, which took another 14 days. Single colonies with GCV and G418 resistance were picked up by pipette tips and kept growing in DMEM for further identification of site-specific integration cell lines through diagnostic PCR and real-time quantitative PCR.

**Cell culture**

HeLa 406h and derivative cell lines were cultured in DMEM with 10% new born calf serum in a 5% $CO_2$ atmosphere. As an exception, 0.5% serum was applied in order to arrest and maintain cells in a nonproliferating state by serum starvation. When necessary, 50 μg/ml hygromycin, 500 μg/ml G418, and 5.5 μg/ml ganciclovir (GCV) were applied.

**Transfection**

Lipofectamine 2000 (Invitrogen) was employed for plasmid DNA transfection as recommended by the manufacturer. Cell lines, HeLa/pFRT.GAL4 (HeLa/GAL4, this work), HeLa/pFRT.myc3'Δ1420 (HeLa/1420) and

HeLa/pFRT.myc3'GAL4 (HeLa/1420-GAL4) (Ghosh et al., 2004) were transfected by $GAL4^{DBD}$ fusion protein expressing plasmids. Sixteen hr later cells were splitted into new plates with 40% confluency. Twenty-four hr later, cells were harvested. $4\times10^7$cells were used for nascent DNA isolation by denaturing gel electrophoresis as described previously (Liu et al., 2003). $1\times10^7$ cells were used for Western blot and $1\times10^8$ cells were used for chromatin immunoprecipitation as described previously (Ghosh et al., 2004). All experiments were repeated at least three times. To keep the consistency of each transfection, GFP plasmid was mixed and co-transfected with $GAL4^{DBD}$ fusion protein expressing plasmids to monitor the transfection efficiency of each individual experiment (about 65-70%).

**Flow cytometry**

Cells were trypsinized, washed in cold PBS (pH 7.4), and fixed overnight at –20°C in 70% ethanol. Permiblized cells were resuspended in PBS (pH 7.4), treated with 100 U of RNase A for 20 min at 37°C, and stained with 50 μg of propidium iodide/ml (Sigma). Transfection efficiency was detected using a Becton Dickinson FACScan flow cytometer.

**Nascent DNA abundance analysis (NDAA)**

HeLa cells were plated in tissue culture plates with 15 cm diameter to be ~70% confluent at time of harvest. Cells were harvested by trypsinization, washed once with ice cold PBS (pH 7.4), then resuspended in a minimal volume

of 10% glycerol (in PBS, pH 7.4). 1.25% alkaline agarose gel (50 mM NaOH, 1 mM EDTA) was prepared and stored in alkaline buffer (50 mM NaOH, 1 mM EDTA) at 4°C. 50-60 µl of the cell suspension was loaded into individual wells of the alkaline gel and allowed to lyse for 10 min at 4°C before electrophoresis. The gel was electrophoresed overnight at 30 to 40 V. The gel slice containing 0.6 to 2 kb nascent DNA was excised, neutralized and purified with a Gel Extraction Kit (Qiagen).

**Polymerase chain reaction (PCR)**

***Standard PCR***

In general, standard PCR were performed on Perkin Elmer Geneamp PCR system 2400 thermocycler. Hotstar polymerase is a product of Qiagen, and dNTP is a product of New England Biolabs. Three holds of PCR program were carried out. Hold I: Hotstar polymerase was activated under 95$^{o}$C for 15 min, Hold II: 94$^{o}$C, denaturing 30 sec; 55~65 $^{o}$C, annealing 30 sec; 72 $^{o}$C, polymerization 30-90 sec, total 40 cycles. Hold III: 72 $^{o}$C: 7 min. PCR reaction were carried put in 50 µl volume containing 10mM Tris-HCl, pH 8.0, 50 mM KCl, 1.5mM MgCl, 1.25 units Taq polymerase and 0.2 mM each dATP, dCTP, dTTP, dGTP, 25 pmoles each upper and lower primers and template DNA.

***Quantitative real-time PCR***

Quantitative, real-time PCR (qPCR) analyses were performed on the ABI Prism 5700 and Prism 7000 sequence detection systems. PCR reagent, Master Mix of Sybgreen PCR kit, is a product of Applied Biosystems. DNA copy number

was quantitated by absolute copy numbers using standard curves generated for each primer set with known amounts of sheared, genomic DNA. The sizes of all PCR products from different primer pairs are about 80-100 bp long. Each specific primer pair only amplify a unique band. PCR reactions were carried out in 25 μl volume and following the next method: stage I, 50°C: 2 min; stage II, 95°C: 10 min; stage III (40 cycles): 90°C, 15 sec, 60°C, 1 min; stage IV: 95°C: 15 sec, 60°C, 25 sec, 90°C, 15 sec. Data were collected at stage III, step 2. For nascent DNA abundance assay, use β-globin STS-54.8 (Kamath and Leffak, 2001) as an internal normalizer. For ChIP assay, pre-immune IgG (pre-immune serum) was used as a negative control.

***Small pool PCR (spPCR)***

Small pool PCR was performed on Perkin Elmer Geneamp PCR system 9600. About 150 pg genomic DNA, equal to 200 molecules were applied as template in small pool PCR in 40-50 μl volume reaction. Amplification conditions were 95° C (15 min); 25-35 cycles of 94° C (30 sec), 65° C (30 sec), 72° C (90 sec); and 72° C (7 min). PCR products were resolved in 7.5% polyacrylamide gels. 1×TBE buffer was used in elctrophoresis system.

All PCR primers used in this thesis are listed in Table 2.

## Table 2. List of PCR primers

| Primer | Forward primer sequence | Reverse primer sequence |
|---|---|---|
| **QPCR** | | |
| Myc 5'-far | GCCACGGCTCAAGCAATC | GCCTCATGGCTGTAATCCTATCTACT |
| A | GGCTCCTGCGGGAAGG | CCTGACGGTGTCTGATCACTTAGA |
| C | GGGAAAGACGCTTTGCAGC | TTTGCCGCAAACGCG |
| E | CGCGCCCATTAATACCCTT | AGGGCCGCGCTTTGA |
| pVD | TCCCTGGCTCCCCTCCT | CATAAGTGCGGCGACGATA |
| pVU | CCAAAGCGGTCGGACAGT | GCGCTATATGCGTTGATGCA |
| Hyg | TGCTCCGCATTGGTCTTGA | TGCGCCCAAGCTGCAT |
| TK | AGCAAGAAGCCACGGAAGTC | GTTGCGTGGTGGTGGTTTTC |
| 54.8 | AAACCCTGCTTATCTTAAACCAACC | ACTCTGCCCTGCCTTTTATGC |
| | | |
| **Diagnostic PCR** | | |
| 1 | TGACTGGAGCGAGGCGAT | |
| 2 | | GCAGTAGCGTGGGCATTT |
| 3 | | TCAGAGCAGCCGATTGTCT |
| **Small pool PCR** | | |
| U2876 | GGAAGGAGCTGACTGGGTTGAAGGCTCT CT | |
| L2661 | | CTGGACGGCTGAGGACCCCCGAGCTGTGCT |
| **Standard PCR** | | |
| Mcm7 | ATATGAATTCCCACCATGGCACTGAAGGA CTACGCGCTA | CCGAACTAGTGACAAAAGTGATCCGTGTCC |
| DUE-B/GAL4 | CCGAATTCCGGAAGACTCTCCTCCGAA | TTTAGATCTGTACAGTCAACTGTCTTTGACCT |
| Cdc45 | GTCAGAATTCCCATCCGGAACTTGGAGG CCTCCACCATGTTCGTGTCCGATTTCCGC AAAG | CGTAACTAGTGGACAGGAGGGAAATAAGTG |
| Orc2 | CGTACCCGGGATGAGTAAACCAGAATTA AAG | CGTAACTAGTAGCCTCCTCTTCTTCCTTTTCC |
| Cdt1 | ACCGTCTAGACCACCATGGCAAAGCTACT GTCTTCTATCGAA | TTTTGCGGCCGCCGATACAGTCAACTGTCT |
| T7 | AAT ACG ACT CAC TAT AGG | |
| BGH | | TAGAAGGCACAGTCGAGG |

**Western Blotting**

Protein expression in cell lysates was determined by SDS-PAGE and Western blot analysis. In brief, M-PER mammalian protein extraction reagent was major lysis buffer (Thermo Scientific) and proteinase inhibitor (Sigma-Aldrich) was applied at 20 μg/ml. After electrophoresis, proteins were transfered to PVDF membrane (Immobilon-P transfer Millipore) through Genie Blotter (Idea Scientific), 45 min, constant current, 0.85 mA. Anti-actin antibody was purchased from Oncogene. Anti-acetylated histone H4 antibody is a product of Upstate Biotechnology. Polyclonal antibodies against human Orc2 and Mcm7 used in some Western blots were obtained from Dr. Aloys Schepers. Anti-DUE-B antibody was raised in rabbits against recombinant human DUE-B expressed in bacteria. Anti-E2F1 antibody and anti-Cdc45 antibody were purchased from Santa Cruz Biotechnology. Anti-$GAL4^{DBD}$ antibody was purchased from Abcam. Anti-FLAG antibody was purchased from Sigma-Aldrich.

**Chromatin immunoprecipitation (ChIP)**

***Isolation of cross-linked nuclei***

Cells with 80% confluence from 6×150 mm plates were harvested, washed with cold PBS (pH 7.4) with protein inhibitor, and resuspend the cells in cold hypotonic buffer A (10 mM HePES, 10 mMKCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM DTT with protease inhibitor). The cells were lysed with Triton X-100 with final concentration 0.04% on ice for 10 min. The nuclei were pelleted

with spinning for 5 min at 4000 g, 4 $^{o}$C and wash with Buffer A once. The nuclei were resuspend in 1ml Buffer A, and mixed with 9 ml pre-warmed Buffer A supplemented with 1.1% formaldehyde and incubated in 37 $^{o}$C for 10 min. The cross-linked nuclei were washed twice with PBS plus 0.5% NP-40 and resuspend in 800 μl low salt buffer (LSB,10 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 1 mM EDTA, 0.1% (v/v) NP40).

***Prepare cross-linked chromatin***

200 μl of 10% sarcosyl were added into cross-linked chromatin. After 5 min ice bath, the sample were layered on a 10 ml sucrose cushion (9 ml LSB, 1ml 1M sucrose) and spun for 5 min at 4000 g at 4$^{o}$ C. The pellet was resuspend in 10:1 TE buffer (10 mM Tris, 1 mM EDTA, pH, 8.0). Chromatin was sheared with sonication (75% duty cycle, 10 pulses, 50 sec, interval 10 sec). Spinning the sonicated sample at 12,000 g, 10 mins, 4 oC could pellet other non-chromatin residual. Upon the required resolution of a specific experiment, the sizes of sheared DNA could be variable. In order to obtain DNA fragment less than 500 bp, micrococcal nuclease was applied at 0.05 units micrococcal nuclease /100 μg chromatin with 1.2 mM $CaCl_2$ buffer, and digested for 10 mins at 37 $^{o}$C. Reaction was stopped by adding 15 mM ethylene glycol tetraacetic acid(EGTA). The final concentration of EGTA is 0.15 mM.

***Immunoprecipitation***

Before immunoprecipitation**,** chromatin was quantified either by spectrophotometer or quantitative gel. When setting immunoprecipitation, 5 μg specific antibody (refer to antibodies used for Western blot) or pre-immune IgG

was mixed with 250 μg chromatin in 1X NET buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 10 mM $Na_2S_2O_5$) in 1 ml tube and rotated at room temperature. Two hr later 120 μl salmon sperm DNA/protein A agarose (Upstate Cell Signaling Solution) was added to each tube and continued to rotate for 2 hr at room temperature. The agarose beads: antibody: chromatin complex was percipited and washed with four times with RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, 1 mM EDTA, 10 mM $Na_2S_2O_5$), LiCl buffer (10 mM Tris-HCl, pH 7.4, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA, 10 mM $Na_2S_2O_5$), and TE buffer, sequentially. To reverse cross-linked chromatin, 20 μl 5M NaCl was added and incubated at 65 $^oC$ for 4 hr. ChIP DNA Clean & Concentrator (Zymo Research) kit was used to purify DNA for real-time quantitative PCR analysis.

**Design of zinc finger nucleases (ZFNs)**

Synthetic $ZFP_{CTG}$ or $ZFP_{CAG}$, specifically binding to $(CTG)_n$ or $(CAG)_n$ repeat sequence formed hairpins, were designed with web-based software, Zinc Finger Targeter (ZiFiT) Version 2.0 (Wright et al., 2006). In general, the single zinc finger plasmids could be purchased from Addgene Inc.. ZFPs or ZFNs with multiple zinc fingers (ZF) were assembled as described in Figure 31. Different finger numbers were checked by PCR amplification using T7 and BGH primers (Figure 31). ZFPs with three-zinc finger derivatives were subcloned into mammalian expression vector pST1374 between FLAG tag and Fok I DNA cleavage domain. ZFPs without Fok I DNA cleavage domains were generated by

addition of a translational stop codon in BamH I restriction site between ZFP binding domain and Fok I nuclease domain. All the ZFPs and ZFNs were FLAG tagged which would allow them to be detected through FLAG antibody (Figure 32).

**Purification of ZFP and ZFN by immunoprecipitation**

In order to test the enzymatic activity of ZFNs in vitro, $ZFN_{CTG}$ or $ZFN_{CAG}$ were expressed in HeLa cells and purified by immunoprecipitation using EZview red ANTI-FLAG M2 affinity gel (Sigma-Aldrich). In brief, $1\times10^7$ cell lysate was prepared in the following lysis buffer: 50 mM Tris HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 % Triton X-100. Clarify cell lysates by centrifugation for 10 minutes 8,200 ×g at 2~8 °C. Cell lysate supernatant (200 ~1,000 µl) was added to 50-100 µl anti-FLAG M2 affinity gel in the final volume of 1 ml. Samples was gently shaked on a roller shaker for overnight at 4 °C, and the complex of protein:(anti-FLAG M2 affinity gel) was percipited by centrifuge, and kept on ice and washed by 500 µl TBS. After briefly vortex the complex was incubated at 4 °C for 5 min and pelleted by centrifuge. The supernant was removed completely and 100 µl of 3X FLAG peptide elution solution (150 ng/µl, Sigma-Aldrich) was added to each sample resin. After Incubation for 30 min at 4 °C, the sample was centrifuge at 8, 200 ×g for 30 seconds. The supernatants containing the interested proteins are transferred into fresh tubes and stored at either 4 °C (for immediate use) or – 20 °C (long-term storage).

**DNA substrates used for testing ZFN activity in vitro**

DNA substrates were direct PCR products amplified using linear $(CTG)_{102}$•$(CAG)_{102}$ DNA as template.

**Assay of enzymatic activity of ZFNs in vitro**

ZFNs and ZFPs expressed from HeLa cells were used to treat DNA substrates. All reactions were carried out in 50 μl volume with 1 X Fok I restriction buffer system and incubated in 37$^{o}$C for different time courses. At defined time points, reactions were stopped by freezing with ice, and DNA samples were purified with Qiagen kit. The cleavage effects were examined by electrophoresis on 7.5% polyacrylamide gel. Fujifilm LAS-300 image system was used to record the results.

**Assay of enzymatic activity of ZFNs in vivo**

Single ZFN plasmid ($ZFN_{CTG}$ or $ZFN_{CAG}$) or double ZFNs plasmid ($ZFN_{CTG}$ plus $ZFN_{CAG}$) and ZFP plasmids were used to transfect HeLa cell lines harboring normal $(CTG)_{12}$•$(CAG)_{12}$ repeats, and expanded repeat, $(CTG)_{102}$•$(CAG)_{102}$ repeats, in which TNR locates at the downstream of c-myc replicator with two orientations. Cell lines HeLa/myc-$(CTG)_{12}$•$(CAG)_{12}$ or HeLa/myc-$(CTG)_{102}$•$(CAG)_{102}$ indicate CTG repeat tract locate on lagging strand template, while cell lines HeLa/myc-$(CAG)_{12}$•$(CTG)_{12}$ or HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ indicate CAG repeat tract locate on lagging strand template. Small pool PCR (spPCR) was

performed to examine the length change of TNR in the above HeLa cells treated with ZFNs and ZFPs.

**Cytotoxic analysis of ZFNs with trypan blue staining**

In a test tube, 0.5 ml of a suitable cell suspension (approximate concentration of 1 x $10^5$ ~ 2 x $10^5$ cells per ml) was mixed with 0.1 ml of 0.4% trypan blue stain. After thoroughly mixing, the sample was allowed to stand for 5 min at room temperature. A hemocytometer was filled with about 20 μl of solution with treated cells. A microscope was used to visualize the stained non-viable cells and non-stained viable cells. The results were recorded using digital camera on a computer imaging system.

# RESULTS

## PART I. FUNCTION OF DUE IN AN ACTIVE CHROMOSOMAL REPLICATOR

DNA unwinding elements (DUE) refer to AT-rich DNA fragments with low helical stability (DePamphilis, 1993). Our laboratory showed DUE is an essential element for a functional c-myc replicator on an ectopic chromosomal site (Figure 6)(Liu et al., 2003). Using the c-myc DUE as bait, our laboratory carried out yeast one-hybrid assay and identified DUE-B as an essential protein for human DNA replication initiation (Casper et al., 2005). Mutagenesis analysis of the c-myc replicator demonstrated that deletion of DUE abolished DUE-B binding and eliminated the c-myc replicator origin activity at an ectopic chromosomal site. Furthermore, substitutions with SCA10 DUEs, AT-rich DNA fragment containing $(ATTCT)_n$ repeat at the ATN10 locus, with certain length and low helical stability could restore replication activity (Figure 7). We hypothesize that DUE-B binding is an essential step to activate origin activity in human cells. In this project, I examined the relationship between binding of DUE-B to DUE and DNA replication origin activity at an ectopic chromosomal site.

**Binding of DUE-B on c-myc/SCA10 chimerical DUEs**

Previously, our lab showed that different c-myc/SCA10 chimerical DUEs show different capacities for restoring DNA replication activity at an ectopic chromosomal site. When c-myc DUE was replaced by $(ATTCT)_8$ or $(ATTCT)_{13}$, origin activity was not recovered. In contrast, $(ATTCT)_{27}$ or $(ATTCT)_{48}$ could functionally replace the c-myc DUE and restore DNA replication activity (Figure 7)(Liu et al., 2007). In the current experiment, we further examined the binding of Orc2, Mcm7, DUE-B, and Cdc45 at the chimerical origins in several isogenic cell lines. As shown by Figure 7, there are no significant changes of the binding of Orc2 and Mcm7 among these cell lines regardless of the difference in origin activity. However, the binding of DUE-B and Cdc45 shows significant increase between the active chimerical origins (WT, sub27, 48) and inactive chimerical origins (sub8, 13) (Figure 8). Interestingly, the binding of DUE-B on the active origins is also increased with the increase of the length of $(ATTCT)_n$. $(ATTCT)_{27, 48}$ binds much more DUE-B than $(ATTCT)_{8,13}$ and wt c-myc do. In addition, although they showed a similar level of origin activity, $(ATTCT)_{48}$ attracted more DUE-B than that in the $(ATTCT)_{27}$ cell line. Surprisingly, the binding of Cdc45 showed a similar pattern with DUE-B, suggesting that binding of Cdc45 is DUE-B associated and correlated with replication origin activity. This phenomenon was also observed at the endogenous SCA10 locus, in which the expanded $(ATTCT)_n$ ($n>1000$) attracted more DUE-B and Cdc45 binding while binding of Orc2 and Mcm7 remains unchanged (Chowdhury et al., 2010). Taken together, these data suggest that functional DUEs can recruit efficient DUE-B and Cdc45 binding

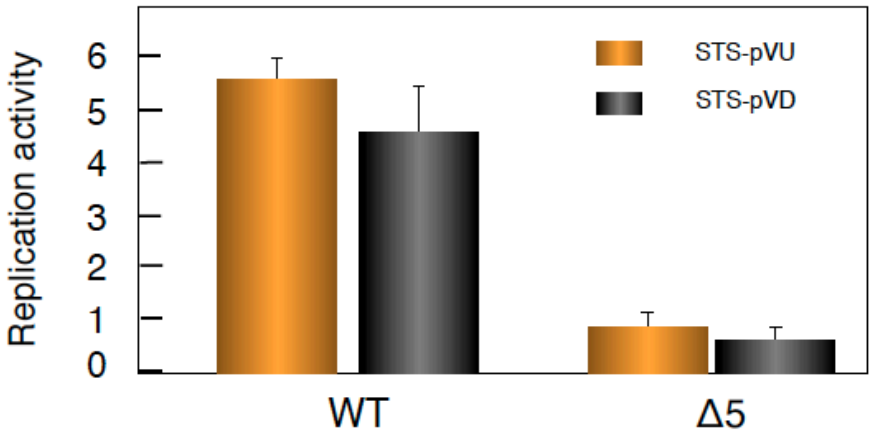
which determines the origin activity. The finding that Cdc45 co-localized with DUE-B on chromatin indicated a possible interaction between DUE-B and Cdc45 during DNA replication initiation. Indeed, this interaction between DUE-B and Cdc45 has been revealed in in vitro system (Chowdhury et al., 2010). Still, the role of DUE for DUE-B to recruit Cdc45 in human cells in vivo remains unknown.

**Figure 6. DUE is an essential element in the c-myc replicator**

**a**. Origin activity of the c-myc mutants at an ectopic chromosomal site measured by relative nascent DNA abundance (vertical cylinders) at sequence tagged sites (STS) pVU (orange) and STS-pVD (black). wt, nascent strand abundance in cells containing wild-type c-myc replicator (pFRT.myc); Δ5, nascent strand abundance in cells containing 200 bp deletion including DUE of c-myc replicator (406h/pFRT.myc. Δ5 ). Similar results were obtained with DUE substitution mutants (Liu et al., 2003). See materials and methods for details. Error bars represent standard deviation from three repetitive experiments.

**b**. Map of wt c-myc replicator and Δ5 on the ectopic chromosomal site. Locations of STS-pVU and STS-pVD were marked by horizontal bars.

**a**


Replication activity
6
5
4
3
2
1
0
STS-pVU
STS-pVD
WT
Δ5

**b**

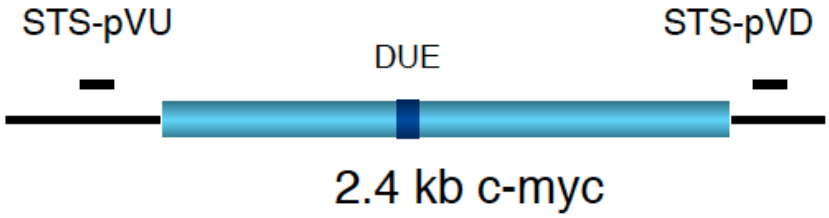

STS-pVU
DUE
STS-pVD
2.4 kb c-myc

**Figure 7. Binding of DUE-B and Cdc45 on the ectopic SCA10DUE/mycDUE**

**a**. Binding of Orc2, Mcm3, DUE-B, and Cdc45 at the ectopic c-*myc* replicators assayed by ChIP at STS-C (in FLP'd cell lines containing the wild-type (WT; 406h.myc) c-myc replicator, the c-myc replicator minus the DUE (ΔDUE)), or the c-myc replicator in which the DUE has been replaced by the SCA10 DUE containing $(ATTCT)_8$, $(ATTCT)_{13}$, $(ATTCT)_{27}$, or $(ATTCT)_{48}$ repeats. Protein binding at the ectopic c-myc locus was normalized using protein binding at the endogenous c-myc locus as an internal standard, i.e. enrichment at the ectopic locus is shown as the ratio of DNA copies immunoprecipitated at the ectopic locus versus DNA copies immunoprecipitated at the endogenous c-myc locus divided by ¼ (number of genomic ectopic copies/number of genomic germ line copies).

**b.** Structure of ectopic c-myc replicator and the position of STS-C (vertical bar) used for comparison of nascent DNA abundance assay.

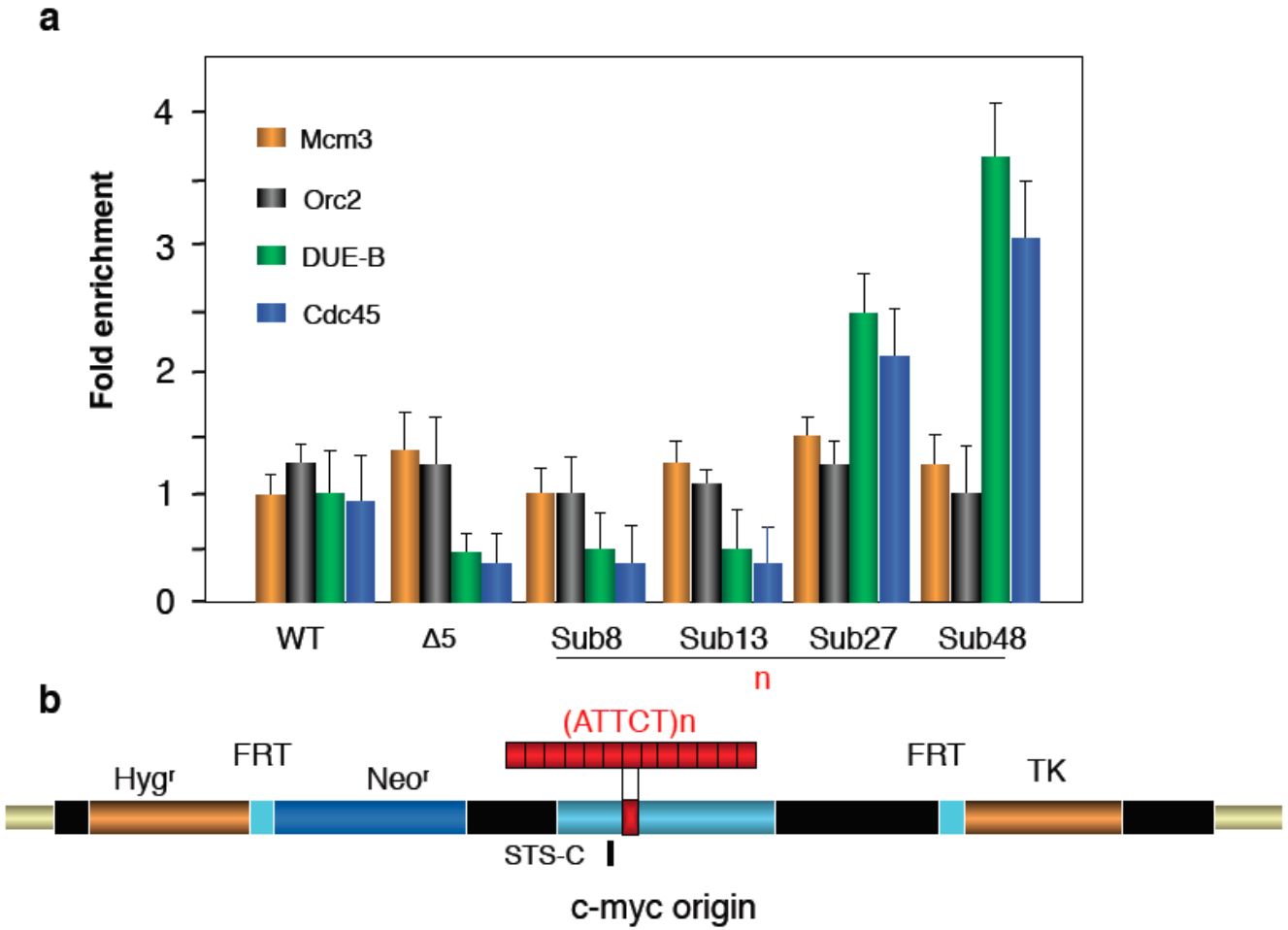
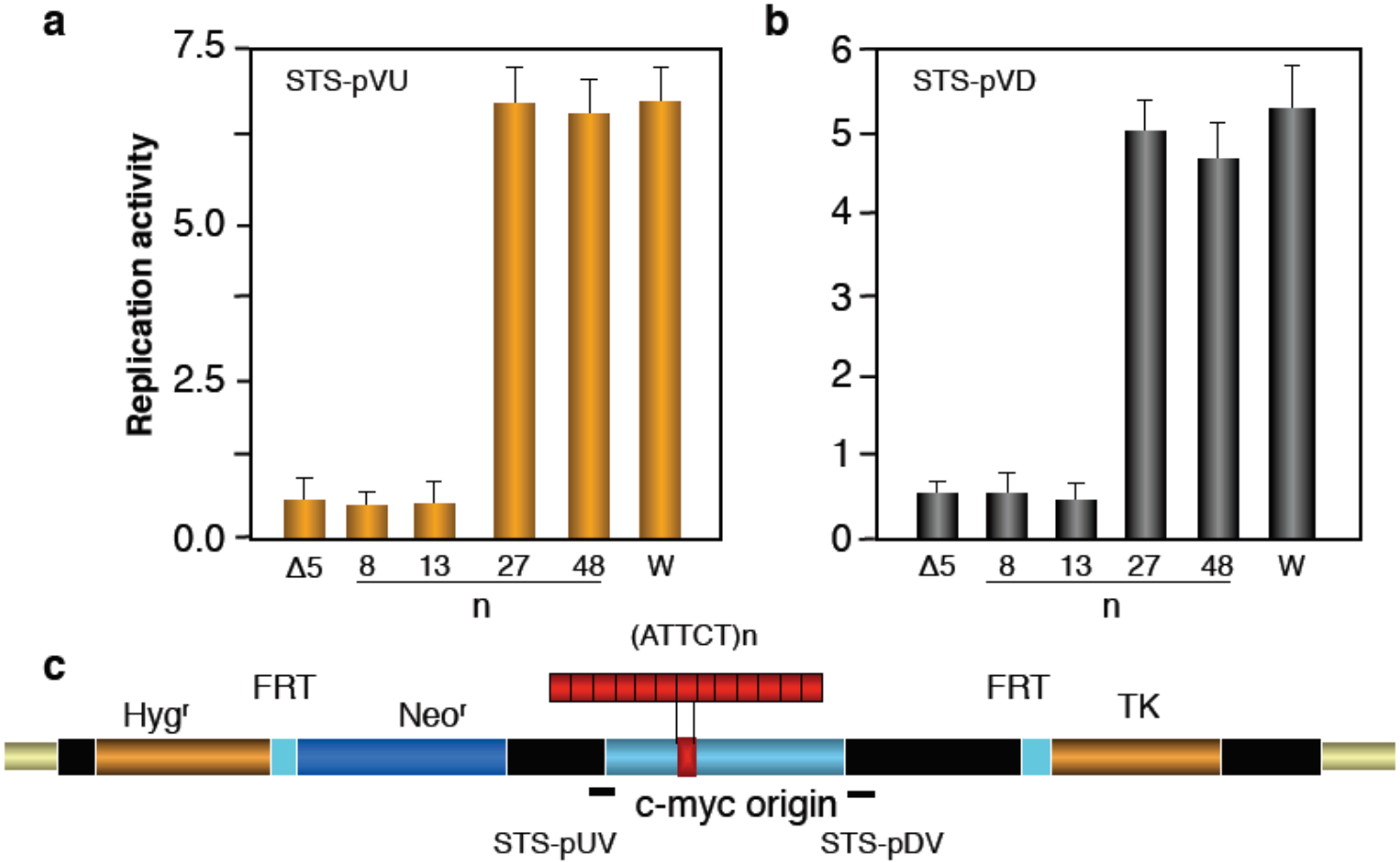

a
Fold enrichment
0
1
2
3
4
Mcm3
Orc2
DUE-B
Cdc45
WT
Δ5
Sub8
Sub13
Sub27
Sub48
n
b
(ATTCT)n
Hygr
FRT
Neor
FRT
TK
STS-C
c-myc origin

**Figure 8. SCA10 DUEs can functionally replace c-myc DUE (Liu et al., 2007)**

**a and b.** Expanded $(ATTCT)_n$ tracts function as DNA unwinding elements in chimerical c-myc/$(ATTCT)_n$ replicators. Nascent DNA was isolated from asynchronously growing wild type c-myc (W), Δ5Sub8, Δ5Sub13, Δ5Sub27, and Δ5Sub48 cell lines and quantitated as described in materials and methods.

**c.** Position of STS-pVU, STS-pVD used for qPCR quantitation of nascent DNA abundance assay.


a
b
c
Replication activity
STS-pVU
STS-pVD
7.5
5.0
2.5
0.0
6
5
4
3
2
1
0
Δ5
8
13
27
48
W
n
(ATTCT)n
Hygr
FRT
Neor
FRT
TK
c-myc origin
STS-pUV
STS-pDV

**Making a site specific cell line with GAL4SUBDUE substitution**

We hypothesize that both DUE DNA sequence and DUE-B binding are two critical aspects of the c-myc replication origin activity in human chromosomes. To test this hypothesis, we constructed a clonal HeLa cell line containing an ectopic mutant c-myc replicator in which the c-myc DUE was replaced by five yeast GAL4 DNA binding sequences. Plasmid pFRT.mycGAL4SubDUE was constructed by substitution of DUE by 5×GAL4 binding sequence. Clonal cell line, HeLa/GAL4SubDUE was generated by integrating the plasmid pFRT.mycGAL4SubDUE into a known ectopic chromosomal site in HeLa/406h. Figure 9 illustrates the process of making the cell line HeLa/GAL4SUBDUE. After screening cells with the drug resistance of neomycin, hygromycin and ganciclovir, single colonies were further characterized by analytical PCR and quantitative PCR (Figure 10 a, b). Diagnostic PCR (Figure 10a) using primer 1 and 3 or primer 1 and 2 can distinguish the acceptor cell line (Figure 10c) and integration cell line (Figure 10 d). As shown in Figure 10a, primer 1 and 3 can amplify a band from genomic DNA isolated from 406h acceptor cell line. However, after integration occurred primer 1 and 3 are 7 kb far apart which limited the PCR amplification. Primer 2 was designed to confirm the integration event. Since primer 2 located at donor plasmid, together with primer 1, a new band will be amplified from integration cell line, GAL4SubDUE, but not 406h cell line. Although we can identify the colonal cell line with integration at FRT site we could not rule out the possibility that other random integration may

occur at non-FRT sites. qPCR using sequence tagged at sites, Hyg, pVU, pVD, and TK, was applied to identify the target cell line with only one copy of plasmid integrating at the FRT site (Figure 10 b). Hyg and TK are sequences tagged sites on acceptor cell line, but STS-pVD and STS-pVU are from donor plasmid. When only one copy of donor plasmid integrated on the FRT site, the copy number of STS-hyg, STS-TK, STS-pVD, and STS-pVU should be equal. Otherwise, when non-FRT site integration occur there will be more copies of STS-pVU and STS-pVD compared to STS-Hyg and STS-TK. Using this method we isolated a cell line with only one copy of pFRT.GAL4SubDUE integrated at the FRT site (Figure 10b).

**Figure 9. Making HeLa406h/pFRT.mycGAL4SUBDUE cell line**

HeLa/406h cell is the HeLa acceptor cell line containing a single chromosomal copy of the plasmid pHyg.FRT.TK, representing an unoccupied chromosomal acceptor site. Hyg, hygromycin resistance gene; FRT, FLP recombinase target site; TK, herpes simplex virus thymidine kinase ($GCV^s$) gene. The Hyg-FRT-TK fusion protein renders cells resistant to hygromycin and sensitive to ganciclovir. pFRT.mycGAL4SubDUE contains the mutant origin, mycGAL4SubDUE. Neo, G418 resistance gene with promoter replaced by FRT. pOG44 is an FLP recombinase expression plasmid, cotransfected with pFRT.mycGAL4SubDUE. Accurately targeted cells are resistant to hygromycin, G418, and ganciclovir.

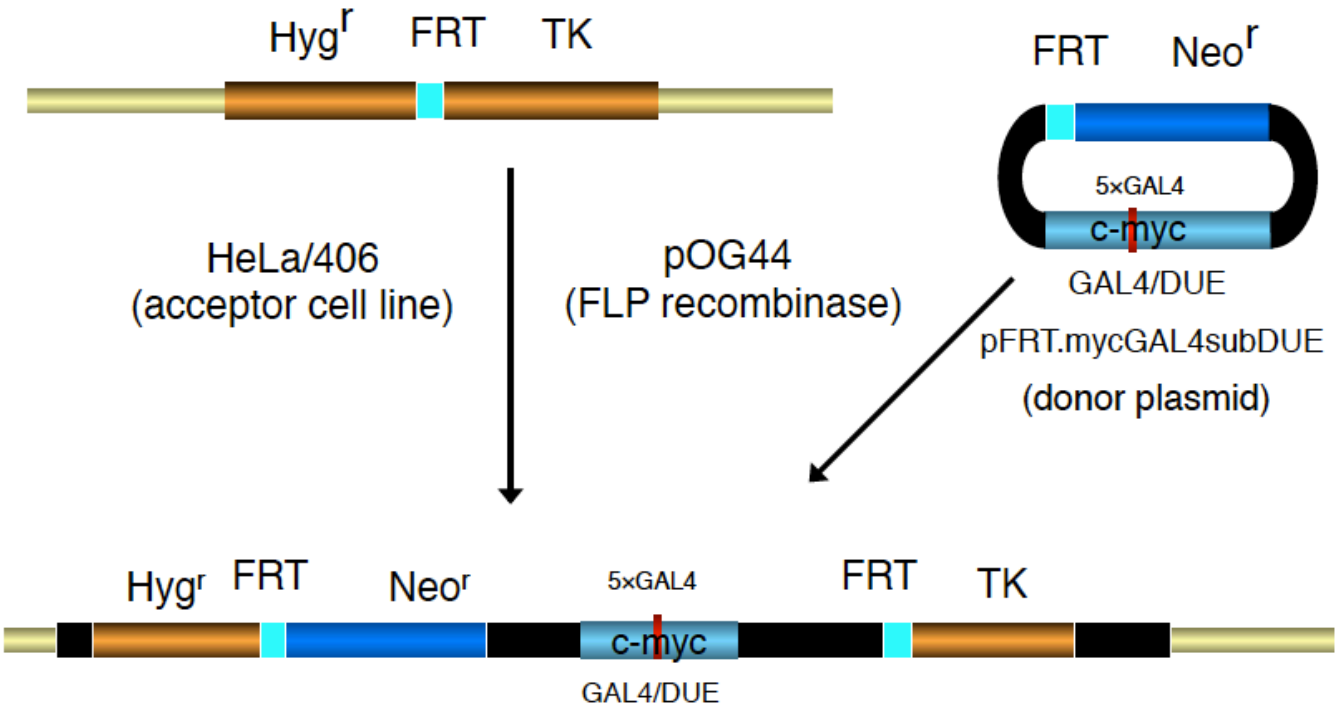

Hyg^r FRT TK
FRT Neo^r
5×GAL4
c-myc
GAL4/DUE
pFRT.mycGAL4subDUE
(donor plasmid)
HeLa/406
(acceptor cell line)
pOG44
(FLP recombinase)
Hyg^r FRT Neo^r 5×GAL4 FRT TK
c-myc
GAL4/DUE

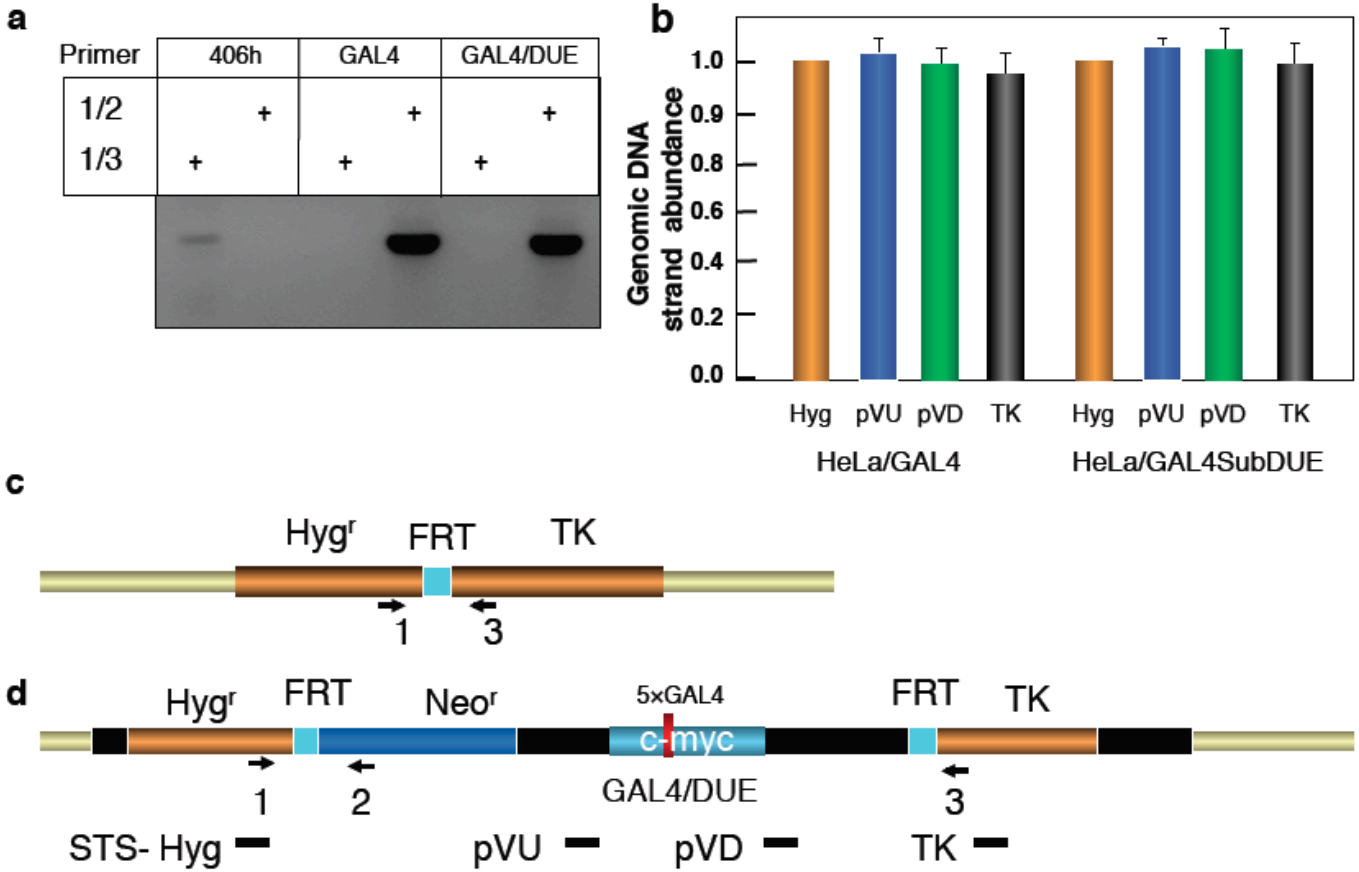
**Figure 10. Analytical PCR and qPCR to identify single copy site-specific integration cell lines**

**a**. Diagnostic PCR using primers 1 and 2 or 1 and 3 with genomic DNA from 406h acceptor cells with an unoccupied FRT acceptor site and genomic DNA from clonal cell lines containing the wild type GAL4 fragment (pFRT.GAL4) (GAL4, positive control), a plasmid in which the entire c-myc replicator was replaced by 5×GAL4 DNA binding sequences, or pFRT.mycGAL4SubDUE, a plasmid in which only c-myc DUE was been replaced by the 5×GAL4 DNA binding sequences, at the acceptor site.

**b.** qPCR quantitation of genomic DNA copy number at sequence-tagged sites Hyg (orange), pVU (blue), pML (green), and TK (black) are indicated. Error bars indicate the standard deviation of 3 experiments.

**c**. Position of primer 1 and 3 on the acceptor cell line, 406h.

**d**. Position of primer 1, 2, 3 and STS-Hyg, -TK, -pVU and -pVD on the integrated cell lines.


a
Primer
406h
GAL4
GAL4/DUE
1/2
1/3
+
b
Genomic DNA strand abundance
1.0
0.9
0.8
0.6
0.4
0.2
0.0
Hyg
pVU
pVD
TK
HeLa/GAL4
HeLa/GAL4SubDUE
c
Hygr
FRT
TK
1
3
d
Hygr
FRT
Neor
5×GAL4
c-myc
FRT
TK
GAL4/DUE
STS- Hyg
pVU
pVD
TK

**Construction GAL4$^{DBD}$-DUE-B fusion protein expression plasmid**

I constructed expression the plasmid pGAL4$^{DBD}$-DUE-B by fusing GAL4$^{DBD}$ at the N-terminus of DUE-B (Figure 11). Upon transfection into HeLa/GAL4SubDUE cells, HeLa406h/DUESubGAL4, the expression of pGAL4$^{DBD}$-DUE-B fusion protein was detected by Western blot (Figure 11). The non-transfected cell lines only showed the endogenous DUE-B protein (24 kD) while transfected cells showed the GAL4$^{DBD}$-DUE-B fusion protein (42 kD) and the endogenous DUE-B proteins (24 kD).

**Binding of GAL4$^{DBD}$-DUE-B on endogenous c-myc and ectopic GAL4SubDUE mutant origin**

Using GAL4 antibody and DUE-B antibody, I detected the binding of GAL4$^{DBD}$-DUE-B fusion protein on the endogenous c-myc replication origin. As shown in Figure 11, GAL4$^{DBD}$-DUE-B was enriched 5-6 fold at STS-myc-C site compared to STS-5' myc-Far (5'F), indicating the fusion protein is able to recognize the c-myc DUE. Meanwhile, we tested the binding of GAL4$^{DBD}$-DUE-B on the ectopic GAL4SubDUE mutant origin upon the transfection of GAL4$^{DBD}$-DUE-B into GAL4SubDUE cell line. As shown in Figure12, the enrichment of GAL4$^{DBD}$-DUE-B was increased about ~7 times at the ectopic GAL4SUBDUE mutant origin, which confirmed the effective tethered binding of GAL4$^{DBD}$-DUE-B on c-mycGAL4SubDUE origin.

**Tethered binding of DUE-B did not recruit Cdc45 in GAL4SubDUE mutant origin.**

Previously, I found that active c-myc/SCA10 chimerical origins bind increased DUE-B and Cdc45, encouraging me to further investigate the binding of Cdc45 on the DUE-B tethered site. As predicted, DUE-B was enriched 5-6 times at both STS-pVU and –pVD compare to non-transfected cells (Figure 12a). These results demonstrated $GAL^{DBD}$-DUE-B could bind to GAL4 site in GAL4SubDUE substitution cell lines.

However, binding of Cdc45 was not observed at two sequence tagged sites, STS-pVU and STS-pVD (Figure 12 b) compared with $GAL4^{DBD}$-DUE-B non-transfected control of the same cell lines. The failed binding of Cdc45 may suggest that DUE is a non-redundant functional element and other essential factors (e.g. pre-RC proteins) recruit Cdc45 to the c-myc DNA replication origin.

**Tethered binding of DUE-B could not recover replication activity**

Previous work demonstrated that origin activity of chimerical c-myc/SCA10 origins correlated with the elevated binding of DUE-B and Cdc45 (Chowdhury et al., 2010). Here we carried out the nascent DNA abundance assay (NDAA) to detect the replication activity upon transfection of $GAL4^{DBD}$-DUE-B fusion protein in HeLa/GAL4SubDUE cells.

Accordingly, the tethered binding of DUE-B failed to confer DNA replication origin activity. Compared with non-transfected cells, nascent DNA

abundance at STS-pVU and –pVD in GAL4$^{DBD}$-DUE-B transfected cells remains at the same level, indicating tethered binding of GAL4$^{DBD}$-DUE-B could not recover DNA replication activity regardless of a significant increase in binding of DUE-B. This negative result indicates an essential role of Cdc45 binding on an active DUE, or the requirement of binding other essential factors, like GINS (Pospiech et al., 2010), on an active DUE during the DNA replication initiation.

**Figure 11. Expression of functional GAL4$^{DBD}$-DUE-B**

**a.** Western Blot of GAL4$^{DBD}$-DUE-B. pGAL4$^{DBD}$-DUE-B was transiently expressed in the indicated cell lines. Western blot was carried out by using anti-DUE-B antibody. Endogenous DUE-B served as a loading control.

**b.** Diagram of GAL4$^{DBD}$-DUE-B.

**c.** Binding of GAL4$^{DBD}$-DUE-B on endogenous c-myc replication origin, anti-GAL4$^{DBD}$ antibody was used in ChIP assay. See material and methods for detailed ChIP assay. Error bars indicate standard deviations of three experiments. Position and distances between STSs are indicated. STS-Myc-Far (5'F) was used as a normalizer.

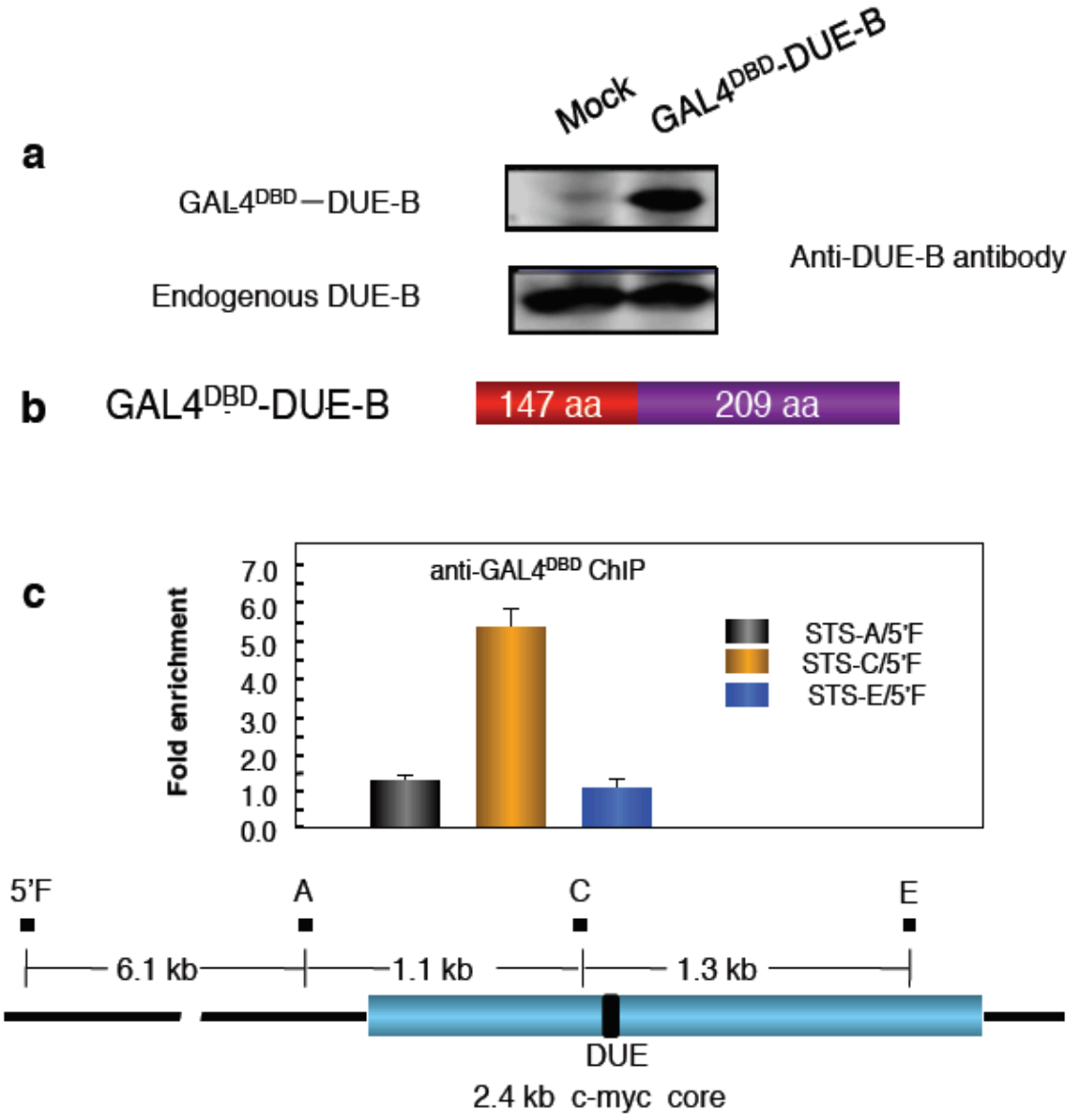

a
Mock
GAL4DBD-DUE-B
GAL4DBD—DUE-B
Endogenous DUE-B
Anti-DUE-B antibody
b
GAL4DBD-DUE-B
147 aa
209 aa
c
anti-GAL4DBD ChIP
Fold enrichment
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
STS-A/5'F
STS-C/5'F
STS-E/5'F
5'F
A
C
E
6.1 kb
1.1 kb
1.3 kb
DUE
2.4 kb c-myc core

**Figure 12. Tethered binding of GAL4$^{DBD}$-DUE-B does not recruit Cdc45**

**a**. DUE-B binding at the ectopic GAL4SubDUE.

**b**. Binding of Cdc45.

**c**. Position of STS-pVU and STS–pVD used in qPCR of ChIP assay.

See materials and methods for detailed ChIP assay. Error bars indicate standard deviations from three experiments.

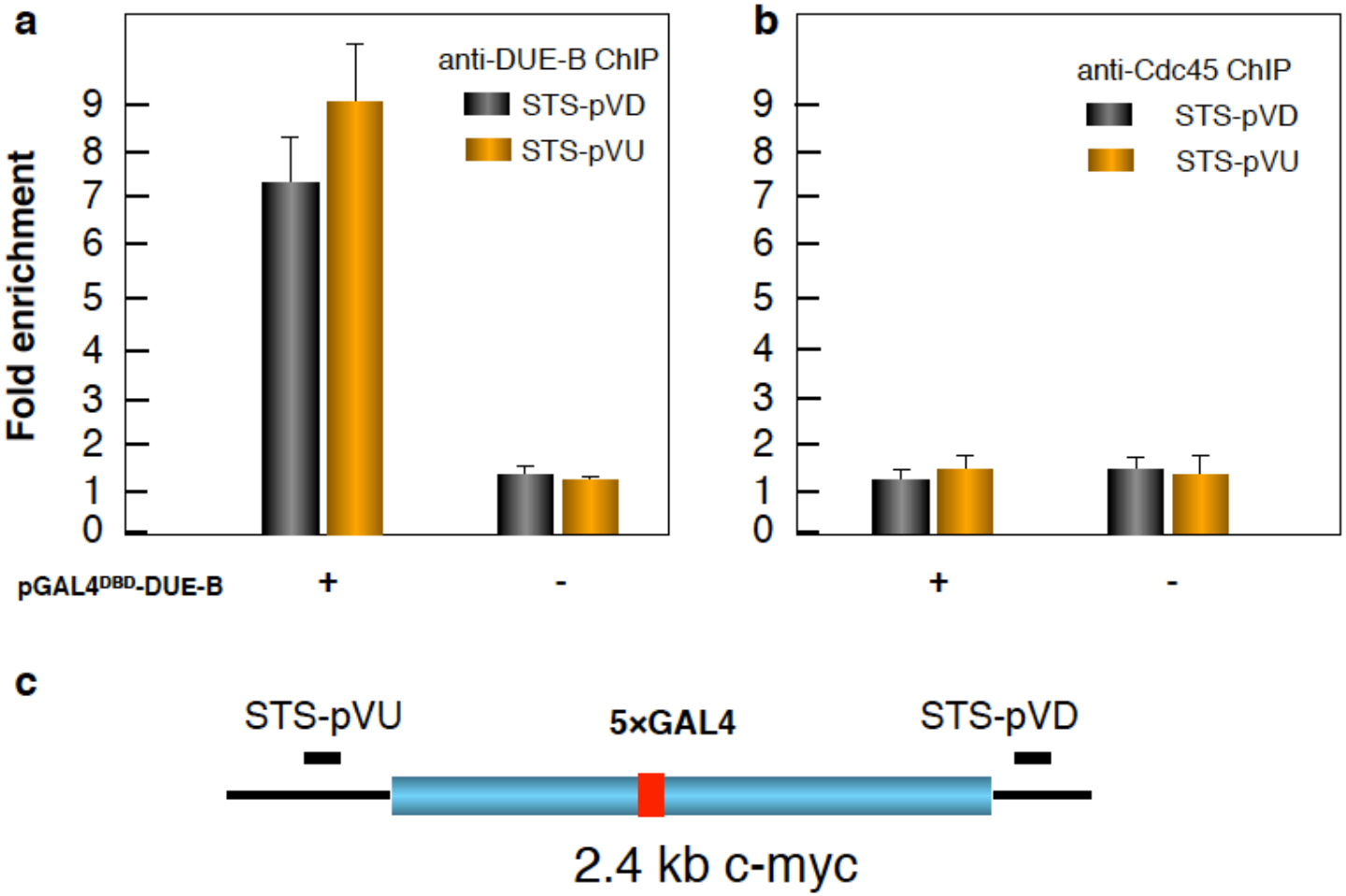
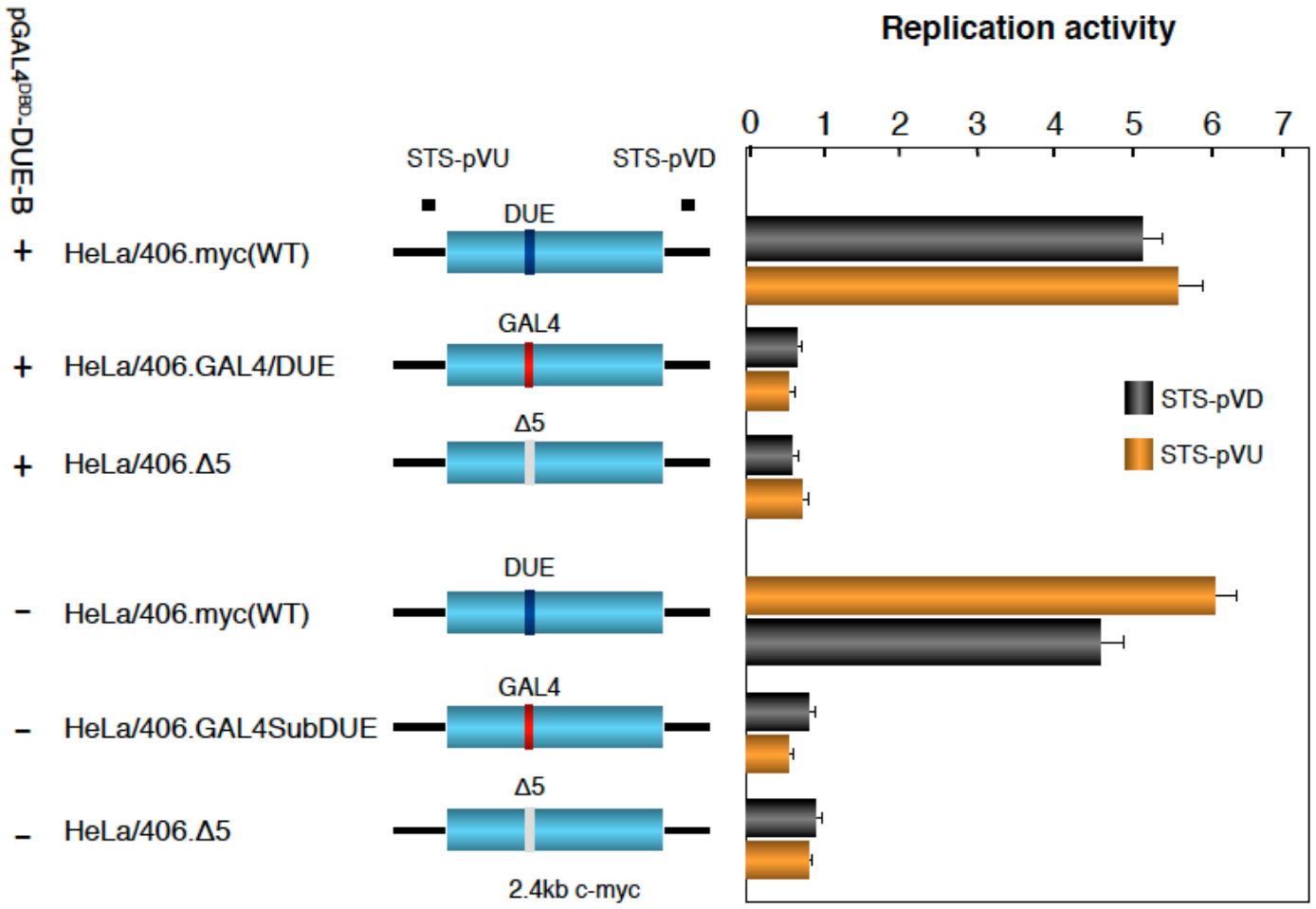

a
anti-DUE-B ChIP
STS-pVD
STS-pVU
Fold enrichment
0
1
2
3
4
5
6
7
8
9
pGAL4DBD-DUE-B
+
-
b
anti-Cdc45 ChIP
STS-pVD
STS-pVU
0
1
2
3
4
5
6
7
8
9
+
-
c
STS-pVU
5×GAL4
STS-pVD
2.4 kb c-myc

**Figure 13. Tethered GAL4$^{DBD}$-DUE-B could not recover origin activity**

DNA replication origin activity, as reflected by the abundance of short (0.6- to 2 kb) nascent DNA, was measured by qPCR. Relative nascent strand abundance by qPCR at sequence-tagged sites STS-pVU and STS-pVD in 406h cells containing the wild-type c-myc origin fragment (WT) or pFRT.myc GAL4SubDUE (GAL4SubDUE) and pFRT.mycΔ5 (Δ5) at the acceptor site. See materials and methods for detailed NDAA. Error bars indicate standard deviations from three experiments.


pGAL4DBD-DUE-B
Replication activity
0 1 2 3 4 5 6 7
STS-pVU
STS-pVD
DUE
GAL4
Δ5
+ HeLa/406.myc(WT)
+ HeLa/406.GAL4/DUE
+ HeLa/406.Δ5
− HeLa/406.myc(WT)
− HeLa/406.GAL4SubDUE
− HeLa/406.Δ5
STS-pVD
STS-pVU
2.4kb c-myc

## SUMMARY I

In this project, I demonstrated that replication origin activity correlates with increased binding of DUE-B and Cdc45 in chimerical c-myc/SCA10 replicators at an ectopic chromosomal site.

A cell line harboring a mutant c-myc replicator was constructed by replacing c-myc DUE by $GAL4^{DBD}$ binding sequence (GAL4). The replication activity of this mutant replicator was examined by nascent DNA abundance assay. My data revealed that tethered binding of DUE-B neither promoted the binding of Cdc45 nor restored the replication activity, suggesting that DUE-B binding is not sufficient to activate a replication activity at the ectopic chromosomal site. The DUE fragment itself plays an essential role apart from providing the binding site for pre-RC or pre-IC factors during DNA replication initiation.

## PART II. INDUCTION OF REPLICATION ACTIVITY BY TRANS-ACTING FACTORS ON AN ECTOPIC CHROMOSOMAL SITE

### Construction of GAL4$^{DBD}$ fusion protein expression plasmids

Figure14 lists all GAL4$^{DBD}$ fusion proteins used in this section. Except GAL4 $^{DBD}$-E2F1plasmid, all other four GAL4$^{DBD}$ fusion protein expression plasmids were constructed in this project. The Mcm7 gene was cloned from a human cDNA library by PCR. Orc2, Cdt1 and Cdc45 were amplified via PCR from plasmids provided by other researchers and subcloned into plasmids with a CMV promoter.

### Expression GAL4$^{DBD}$ fusion protein in HeLa cells

The expression levels of GAL4$^{DBD}$ fusion proteins were detected by Western blot using proper antibodies. E2F1, Orc2, Mcm7, and Cdc45 were detected using antibodies against E2F1, Orc2, Mcm7 and Cdc45 respectively. As shown in the Figure 21, 23, 27 and 29, both endogenous proteins and GAL4$^{DBD}$ fusion protein were detected. We used GAL4$^{DBD}$ antibody to detect the expression of Cdt1-GAL4 $^{DBD}$ in transfected cells indicated in Figure 25. These experimental results provide evidences that all the GAL4$^{DBD}$ proteins can be successfully expressed in our cell lines. In order to keep the experimental results comparable, GFP expression plasmids were co-transfected to monitor the transfection efficiency. In

general, the transfection efficiency is about 65-70% according to FACS data in my experiments.

**Binding of trans-acting factors on the ectopic c-myc replicator in HeLa cells**

Our previous work demonstrated that 2.4 bp fragment of c-myc replication origin worked as a replicator at an ectopic chromosomal site in HeLa cells (Malott and Leffak, 1999). Using the cell line HeLa406h/pFRT.myc that harbored a 2.4 kb c-myc replicator, we checked the binding of trans-acting factor E2F1, pre-RC components Orc2, Cdt1, Mcm7, as well as pre-IC components, DUE-B and Cdc45, at the ectopic c-myc replicator (Figure 15a). We chose two sequence tagged sites, STS-pVU and STS-pVD, to test the enrichment of targeted proteins (Figure 15a). As shown in Figure 15, Orc2 (WT, Figure 15c) was enriched 4-5 fold. Mcm7 (WT, Figure 15d) and DUE-B (WT, Figure 15e) were enriched about 5-6 fold. The pre-IC component, Cdc45 (WT, Figure 15f), was enriched about 3-4 fold. Taken together, these results demonstrated that pre-RC and pre-IC assembled at the ectopic c-myc replicator, indicating that the ectopic c-myc replicator could act in the same way as the endogenous c-myc replication origin. Interestingly, as shown in Figure 16, E2F1 did not have pronounced enrichment on 2.4 kb c-myc replicator (WT) as well as in other cell lines (GAL4, 1420, and 1420-GAL4) although there is a predicted consensus E2F1 binding sequence, $TTT^{G}/_{C}{}^{G}/_{C}GCCGC$ located 800 bp downstream of STS-pVU.

**Binding of endogenous trans-acting factors at the ectopic mutant c-myc replicators in HeLa cells**

Using ChIP assay, I checked the binding of $GAL4^{DBD}$, Orc2, Mcm7, DUE-B, and Cdc45 on mutant c-myc replicators in HeLa cell lines, GAL4, Δ1420, and Δ1420-GAL4. As shown in Figure 15, unlike wild type c-myc replicator (Figure 15 c, d, e, f, WT), I found no binding of Orc2 (15c), Mcm7 (15d), DUE-B (15e), and Cdc45 (15f) in mutation cell lines, GAL4, Δ1420, and Δ1420-GAL4. These results strongly demonstrated that endogenous Orc2, Mcm7, DUE-B, and Cdc45 did not bind to the ectopic GAL4, Δ1420 and Δ1420-GAL4 fragments. Most likely, the loss of replication origin activity of Δ1420 fragment (Figure 17) maybe due to the failure of assembly of pre-RC although the DNA unwinding element (DUE) still remains in the Δ1420 fragment (Figure 17).

**Figure 14. Diagram of expression plasmids of GAL4$^{DBD}$ fusion protein**

GAL4 DNA binding domain (1-147 amino acids, GAL4 $^{DBD}$) was fused in either C-terminus of proteins to be tethered, Orc2, Mcm7, Cdc45, or N-terminus of E2F1 and Cdt1. Expression of GAL4-E2F1 was driven by RSV promoter. Expression of all other fusion proteins was driven by CMV promoter. Vectors used for making expression plasmids were listed at the left of the GAL4-fusion protein expression cassettes. $P_{cmv}$, cytomegalovirus (CMV) promoter/enhancer. $P_{RSV}$, Rous sarcoma virus promoter/enhancer. pA, simian virus 40 (SV40) poly(A) signal.

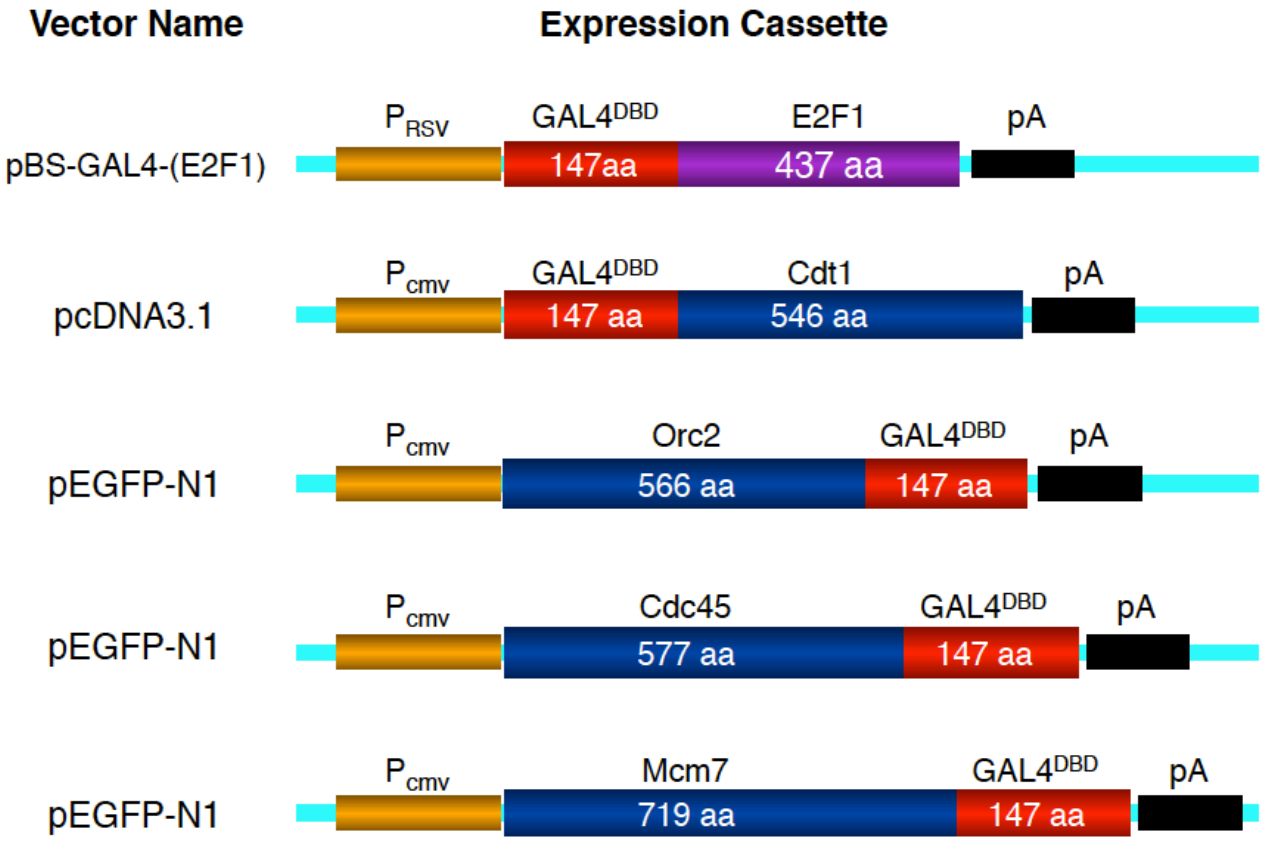

Vector Name
Expression Cassette
pBS-GAL4-(E2F1)
P RSV
GAL4 DBD
147aa
E2F1
437 aa
pA
pcDNA3.1
P cmv
GAL4 DBD
147 aa
Cdt1
546 aa
pA
pEGFP-N1
P cmv
Orc2
566 aa
GAL4 DBD
147 aa
pA
pEGFP-N1
P cmv
Cdc45
577 aa
GAL4 DBD
147 aa
pA
pEGFP-N1
P cmv
Mcm7
719 aa
GAL4 DBD
147 aa
pA

**Figure 15. Binding of endogenous pre-RC components and pre-IC components on the ectopic c-myc replicator**

**a. Isogenic cell lines used for analysis of trans-acting replication factors.** All cell lines were constructed using FLP-recombinase mediated site-specific integration into the single FLP recombinase target in HeLa/406h cells. Δ1420: the 5' 930 bp of the c-myc replicator. Δ1420–GAL4: the 5' 930 bp of the c-myc replicator is flanked by 5×GAL4 binding sites. GAL4: five GAL4 binding sites were integrated without flanking c-myc DNA. c-myc (WT): the 2.4 kb human c-myc replicator core. DUE: DNA unwinding element.

**b-f**. ChIP assay. Formaldehyde cross-linked chromatin was isolated and GAL4, Orc2 and Mcm7, DUE-B and Cdc45 antibodies were used to immunoprecipitate chromatin. Reverse cross-linked DNA was used as template in real time PCR to detect the enrichment of DNA fragment at STS-pVU (orange) and STS-pVD (black) after (b). $GAL4^{DBD}$, (c). Orc2, (d). Mcm7, (e). DUE-B and (f). Cdc45 pull down. ChIP data were normalized by pre-immune IgG from same species. Error bars, standard deviations from three repetitive experiments.

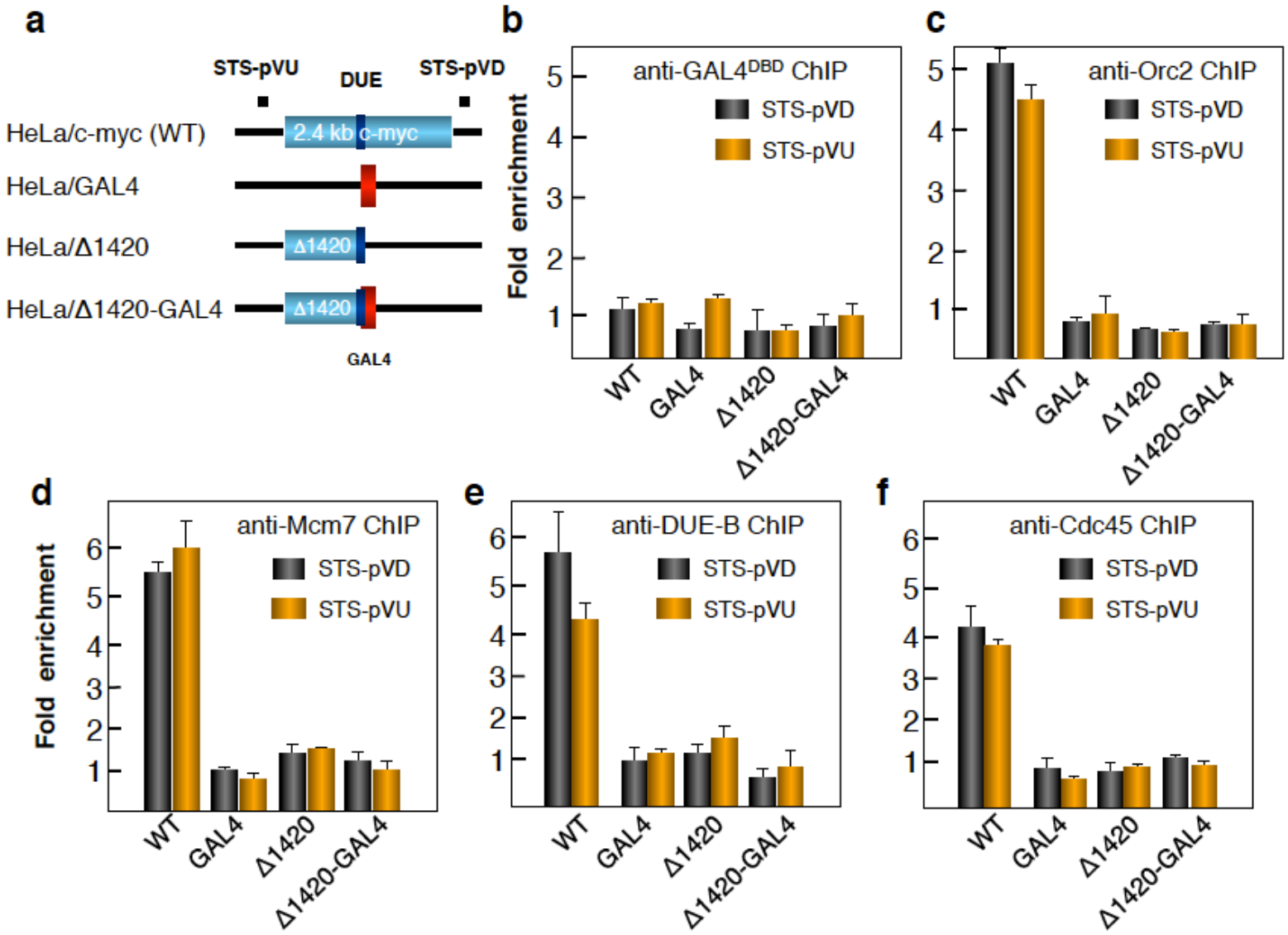
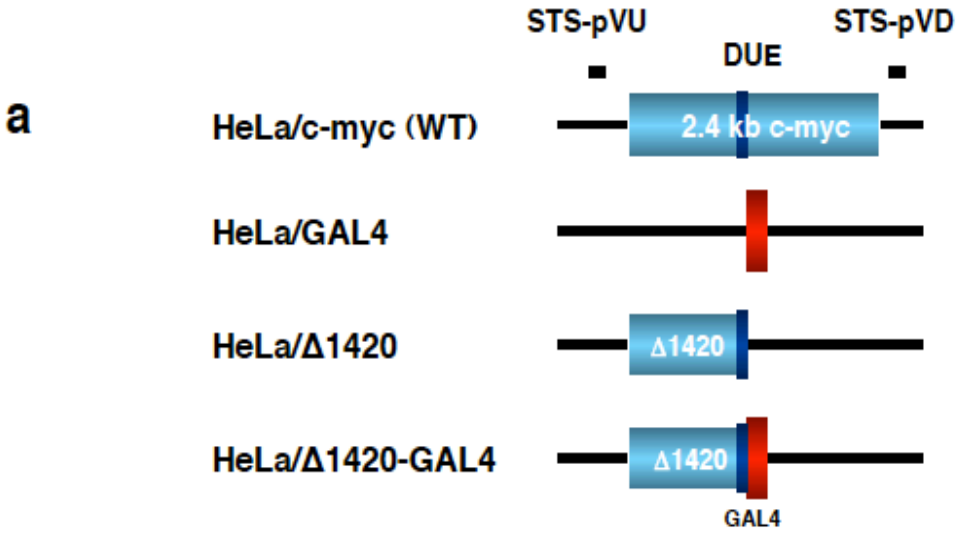

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
WT
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU

**Figure 16. Binding of GAL4$^{DBD}$-E2F1 on the ectopic locus**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b**. ChIP assay. Formaldehyde cross-linked chromatin was isolated and polyclonal antibody against human E2F1 were used to immunoprecipitate chromatin. Reverse cross-linked DNA was used as template in real time qPCR to detect the enrichment of DNA fragment at STS-pVU (orange) and STS-pVD (black) after E2F1 (left) pull down. ChIP data were normalized by pre-immune IgG from the same species. Error bars, standard deviations from three repetitive experiments.


a
STS-pVU
STS-pVD
DUE
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4

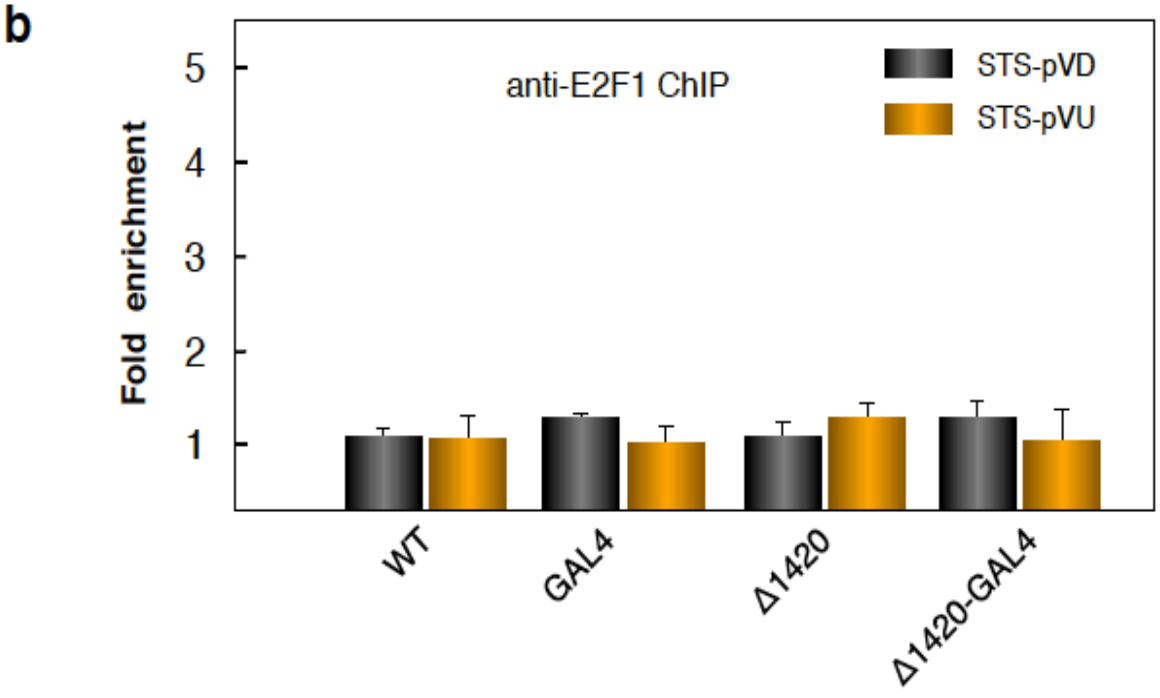
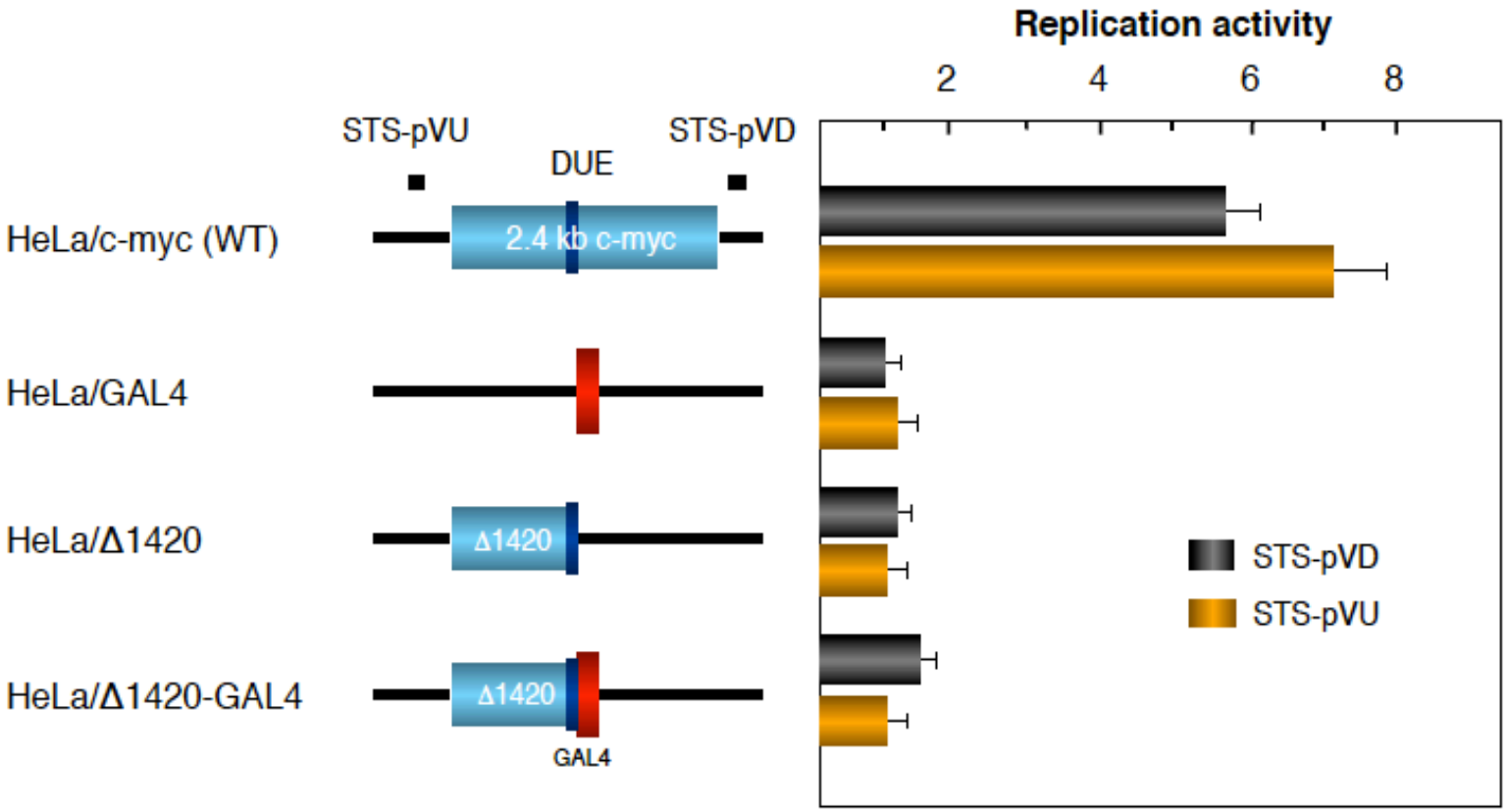

b
anti-E2F1 ChIP
STS-pVD
STS-pVU
Fold enrichment
5
4
3
2
1
WT
GAL4
Δ1420
Δ1420-GAL4

**Figure 17. Replication activity of c-myc replicator and its mutants**

DNA replication activity tested in cell lines HeLa/c-myc, HeLa/Gal4, HeLa/Δ1420 and HeLa/Δ1420-GAL4 by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8 from another replication origin, β-globin, was selected as a internal normalizer to compare different nascent DNA preparations (Kamath and Leffak, 2001). Error bars, standard deviations from three repetitive experiments.


Replication activity
2
4
6
8
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
STS-pVD
STS-pVU

**Binding of exogenous GAL4$^{DBD}$ fusion protein on endogenous c-myc origin in HeLa cells**

By fusing GAL4$^{DBD}$ on either C-terminus or N-terminus of pre-RC or pre-IC components and the E2F1 transcription factor we constructed GAL4$^{DBD}$ fusion proteins. In order to ensure the fusion of the GAL4$^{DBD}$ proteins on either C- or N-terminus of the replication proteins did not affect their normal biological function we carried out ChIP experiments by checking the binding of GAL4$^{DBD}$ fusion proteins to the endogenous c-myc replication origin. As shown in the Figure 18, using GAL4$^{DBD}$ antibody, Orc2 is enriched 4 and 5 fold at STS-A site and STS-E, respectively. Cdt1 was enriched 3 and 5.5 fold at STS-A and STS-C, respectively. On the contrary, Mcm7 and Cdc45 were enriched 3.5 fold or 4.5 fold at STS-C, respectively. These results agree with previous data, suggesting that GAL4$^{DBD}$ fusion did not affect the involvement of GAL4$^{DBD}$ fusion proteins in the assembly of pre-RC or pre-IC on the endogenous c-myc replication origin (Bieda et al., 2006; Ghosh et al., 2006; Kinoshita and Johnson, 2004; Roussel et al., 1994; Weinmann et al., 2001).

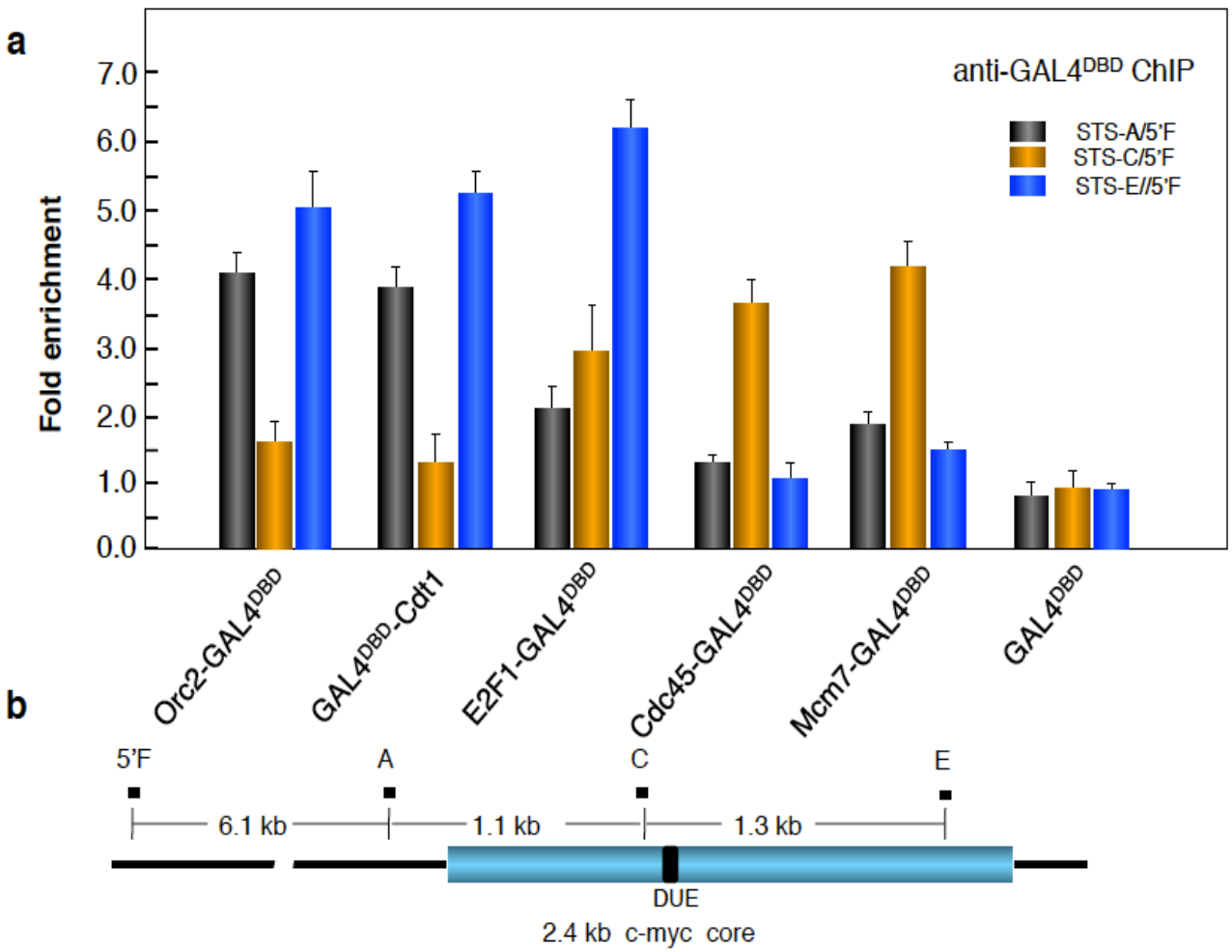
The enrichment of E2F1 reached the highest level, 6 fold, at STS-E, and was also detected at STS-A, 2 fold, and STS-C, 3 fold, respectively. Again, the enrichment of E2F1 on endogenous c-myc origin demonstrates that E2F1-GAL4$^{DBD}$ is able to bind to c-myc promoter region as the wild type E2F1 does, as previously reported (Bieda et al., 2006; Ghosh et al., 2006; Kinoshita and Johnson, 2004; Roussel et al., 1994; Weinmann et al., 2001).

**Figure 18. Binding of GAL4$^{DBD}$ fusion proteins on endogenous c-myc replication origin**

HeLa406h/GAL4 cell line was transfected, respectively, by Orc2-GAL4$^{DBD}$, GAL4$^{DBD}$-Cdt1, GAL4$^{DBD}$-E2F1, Mcm7-GAL4$^{DBD}$, DUE-B-GAL4$^{DBD}$ and Cdc45-GAL4$^{DBD}$. GAL4DBD antibody was used to detect the enrichment of GAL4$^{DBD}$ fusion protein at the endogenous c-myc replication origin. Error bars, standard deviations from three repetitive experiments.

a. ChIP assay to test enrichment of GAL4$^{DBD}$, Orc2, Cdt1, E2F1, Mcm7, DUE-B and Cdc45 on endogenous c-myc origin. See materials and methods for detailed ChIP assay.

b. Position of sequence tagged site, STS-5'F, -A, C, E.


a
anti-GAL4DBD ChIP
STS-A/5'F
STS-C/5'F
STS-E//5'F
Fold enrichment
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Orc2-GAL4DBD
GAL4DBD-Cdt1
E2F1-GAL4DBD
Cdc45-GAL4DBD
Mcm7-GAL4DBD
GAL4DBD
b
5'F
A
C
E
6.1 kb
1.1 kb
1.3 kb
DUE
2.4 kb c-myc core

## Tethered binding of E2F1 induced replication activity

### *GAL4$^{DBD}$-E2F1 binding altered chromatin structure*

Previously, our laboratory demonstrated that tethered binding of CREB modulated the local chromatin structure and conferred replication origin activity on the inactive replication origin. To test if tethered E2F1 also functions the same as CREB, I analyzed the acetylation status of histone 4 (H4) at the ectopic chromosomal site in HeLa/406h derivative cell lines using ChIP assay. My results demonstrated that tethered binding of E2F1 significantly led to the increase of the acetylation of histone H4 (Figure19). In HeLa/Δ1420-GAL4 cell line, tethered E2F1 binding led to 8 fold enrichment of acetyl-H4 at STS-pVU and STS-pVD, respectively, but no pronounced enrichment was seen in the HeLa/Δ1420 cell line, indicating that the binding of E2F1 significantly altered the chromatin structure flanking GAL4 DNA binding sequence. Meanwhile, binding of E2F1-GAL4$^{DBD}$ also led to significant increase of acetyl-H4, 6 fold, at the acceptor site in the HeLa/GAL4 cell line, suggesting a change of chromatin structure upon binding of E2F1-GAL4$^{DBD}$ is not c-myc DNA sequence dependent.

### *E2F1 binding recruited pre-RC components, Orc2, Mcm7, Cdt1 and pre-IC component, DUE-B and Cdc45*

Amazingly, ChIP further showed that the tethered binding of GAL4$^{DBD}$-E2F1 on Δ1420-GAL4 recruited Orc2 (Figure 20c), Mcm7 (Figure 20d), DUE-B (Figure 20e), and Cdc45 (Figure 20f). On the contrary, although GAL4$^{DBD}$-E2F1 also bound on GAL4 in HeLa/GAL4 cell line, enrichment of Mcm7 (Figure 20d), DUE-

B (Figure 20e) and Cdc45 (Figure 20f) was not detected regardless of slight recruitment of Orc2 (Figure 20c), suggesting that recruitment of pre-RC and pre-IC component by ORC may require essential DNA sequence which exists in either wild type c-myc or Δ1420 fragment but not in GAL4 DNA binding sequences.

***E2F1 binding stimulated DNA replication activity***

Using NDAA, I checked the effect of tethered $GAL4^{DBD}$-E2F1 on replication activity in three cell lines, HeLa/GAL4, HeLa/Δ1420 and HeLa/Δ1420-GAL4. As shown in Figure 21, tethered binding of E2F1 on GAL4 DNA sequence recovered the DNA replication activity in HeLa/Δ1420-GAL4 cell line but not in HeLa/GAL4 or HeLa/Δ1420 cell lines (Figure 21), suggesting that both the binding of $GAL4^{DBD}$-E2F1 on chromosome and the Δ1420 sequences are essential to confer the replication origin activity.

**Figure 19. Tethered binding of E2F1 induced histone acetylation of local chromatin**

ChIP assay of the enrichment of acetylation of histone 4 at the acceptor site. Refer to materials and methods for detailed ChIP assay except that the antibody used for pull down is rabbit polyclonal antibody against acetylated human histone H4. See materials and Methods for ChIP assay. Error bars, standard deviations from three repetitive experiments.

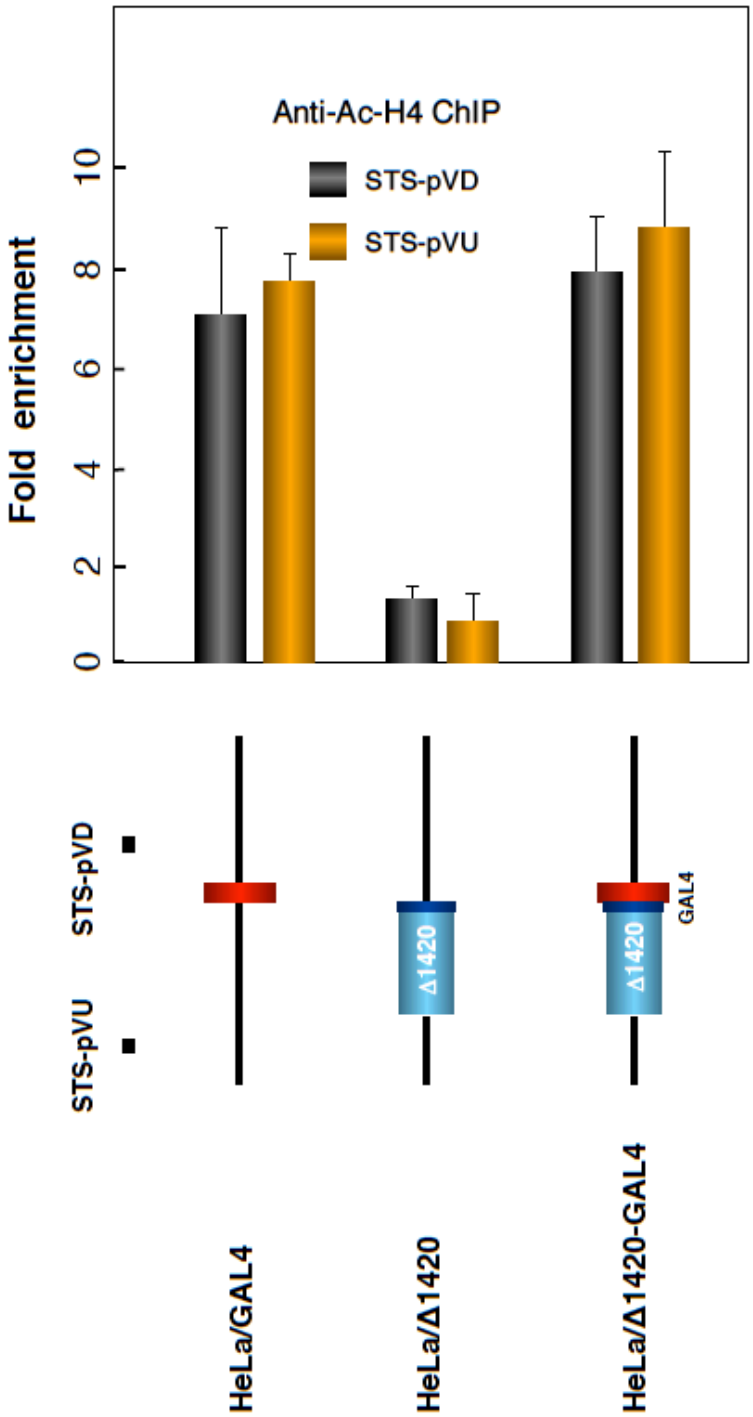

Anti-Ac-H4 ChIP
STS-pVD
STS-pVU
Fold enrichment
0
2
4
6
8
10
STS-pVU
STS-pVD
Δ1420
Δ1420
GAL4
HeLa/GAL4
HeLa/Δ1420
HeLa/Δ1420-GAL4

**Figure 20. Tethered binding of GAL4$^{DBD}$-E2F1 recruited pre-RC and pre-IC components to acceptor site**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b-f**. Recruitment of GAL4$^{DBD}$ fusion protein, pre-RC components, Orc2, Mcm7 and pre-IC components, DUE-B and Cdc45, were detected by ChIP. See materials and methods for detailed ChIP assay. Error bars, standard deviations from three repetitive experiments.

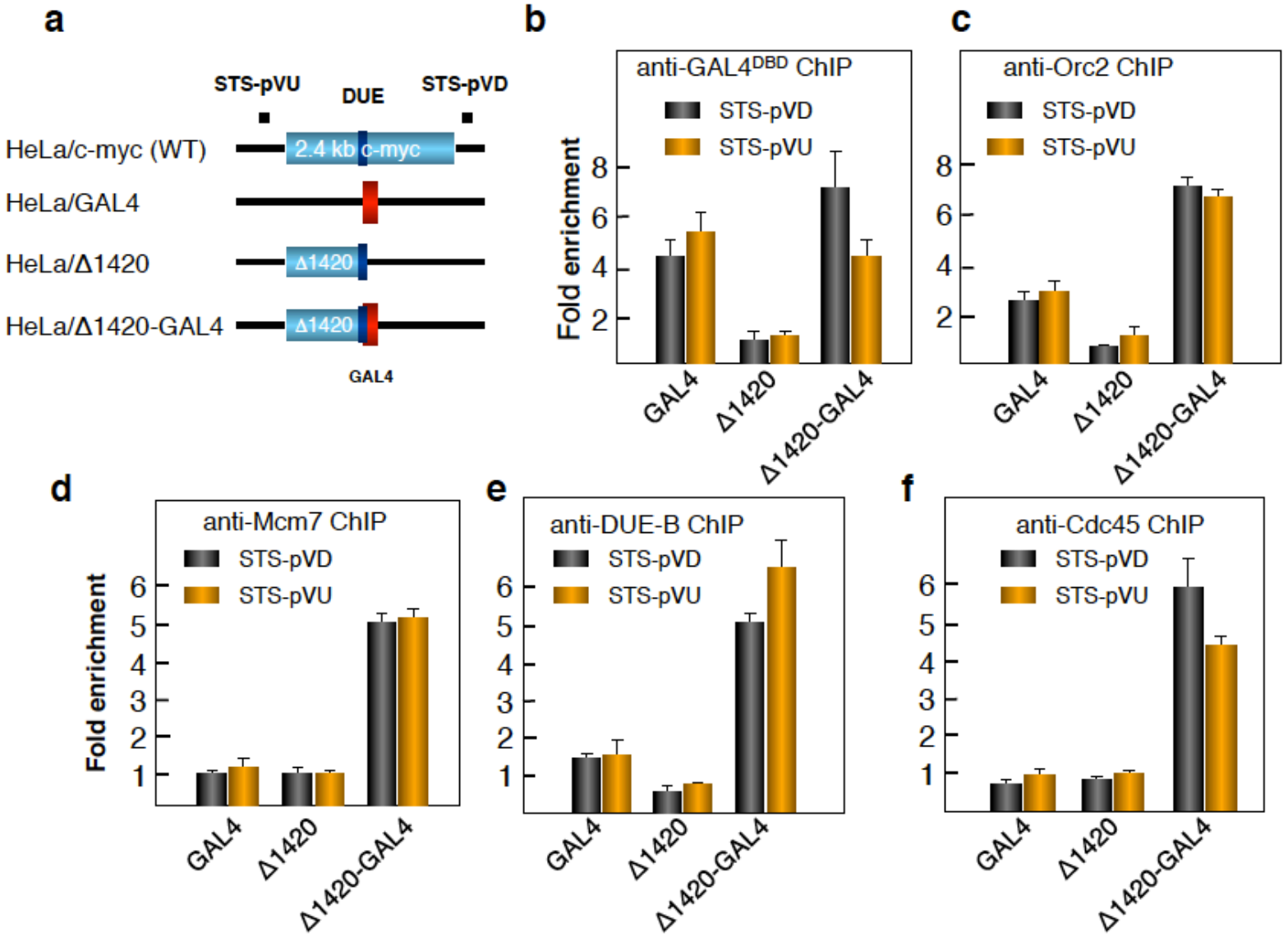
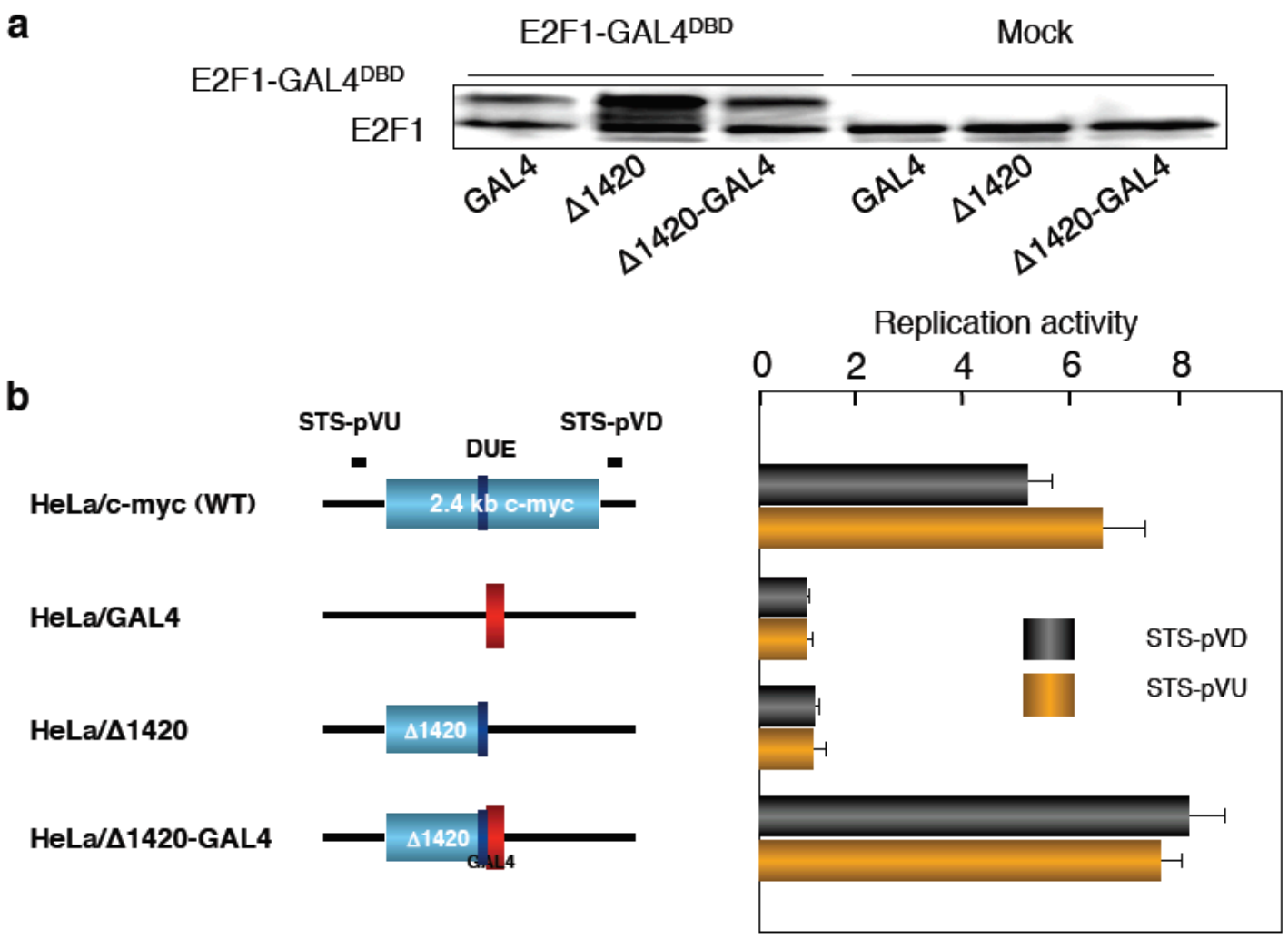

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4

**Figure 21. Tethered binding of E2F1-GAL4$^{DBD}$ recovered replication origin activity**

**a**. Western blot detection of the expression of E2F1 in HeLa cell lines. Mouse monoclonal antibody against E2F1 was used.

**b**. DNA replication activity tested by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8, a site from another human replication origin, β-globin locus with low nascent DNA abundance (Kamath and Leffak, 2001), was selected as a internal normalizer to compare different nascent DNA preparations. Error bars, standard deviations from three repetitive experiments.


a
E2F1-GAL4DBD
Mock
E2F1-GAL4DBD
E2F1
GAL4
Δ1420
Δ1420-GAL4
GAL4
Δ1420
Δ1420-GAL4
b
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
Replication activity
0
2
4
6
8
STS-pVD
STS-pVU

**Pre-RC components**

***Tethered Orc2 binding recruited Cdt1, Mcm7, DUE-B and Cdc45, and recovered the DNA replication activity.***

As shown in Figure 4, binding of ORC on chromosome is the first step to form pre-RC during DNA replication initiation. In this experiment, we first tested the effect of tethered binding of Orc2. As expected, we found the tethered binding of Orc2 (Figure 22c) on Δ1420-GAL4, recruited Mcm7 (Figure 22d), DUE-B Figure 22e), and Cdc45 (Figure 22f) in HeLa/Δ1420-GAL4 cells and not on GAL4 in HeLa/GAL4 cells.

Furthermore, as shown in Figure 23, tethering of Orc2-GAL4$^{DBD}$ in HeLa/Δ1420-GAL4 cells recovered the replication activity, which was not found in HeLa/GAL4 and HeLa/Δ1420 cell lines, suggesting that the DNA sequences in Δ1420 fragment played an essential role for ORC to recruit other pre-RC and pre-IC components in order to activate DNA replication origin activity.

***Tethered Cdt1 binding recruited Orc2, Mcm7, DUE-B and Cdc45, and recovered the DNA replication activity.***

Cdt1 is a licensing factor involving in recruiting explicative DNA helicase Mcm2-7 proteins into the pre-explicative complex together with Cdc6 (Jee et al., 2010). Once origins have fired, Cdt1 is either exported out of the nucleus or degraded, thereby preventing another round of replication. Higher eukaryotes have evolved another redundant mechanism, in which an inhibitor called geminin

restrains Cdt1 activity (Ballabeni et al., 2004; Saxena and Dutta, 2005). It was believed that the binding of Cdt1 is ORC dependent (Chen et al., 2007). Most recently, it was shown that human Cdt1 associated with the Cdc7 kinase recruits Cdc45 to chromatin and the amount of Cdt1 bound to chromatin is regulated by Cdc7 (Ballabeni et al., 2009).

In the current experiment, I directly tested the effect of tethered binding of Cdt1 directly on a mutant c-myc replicator in human cells. As shown in Figure 24, tethered Cdt1 binding recruited Orc2 (Figure 24a, c), Mcm7 (Figure 24a, d), DUE-B (Figure 24a, e), and Cdc45 (Figure 24a, f). Most importantly, $GAL4^{DBD}$-Cdt1 fusion protein recovered the DNA replication activity in HeLa/Δ1420-GAL4 and not in HeLa/GAL4 or HeLa/Δ1420 cell lines (Figure 25b). Western blot result showed the successful expression of Cdt1-$GAL4^{DBD}$ fusion protein (Figure 25a).

**Figure 22. Tethered Orc2 binding recruited Mcm7, DUE-B and Cdc45.**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b-f**. Recruitment of $GAL4^{DBD}$ fusion protein, Orc2-$GAL4^{DBD}$ and endogenous Mcm7, DUE-B and Cdc45-GAL4, were detected by ChIP. See materials and methods for detailed ChIP assay. Error bars, standard deviations of three repetitive experiments.

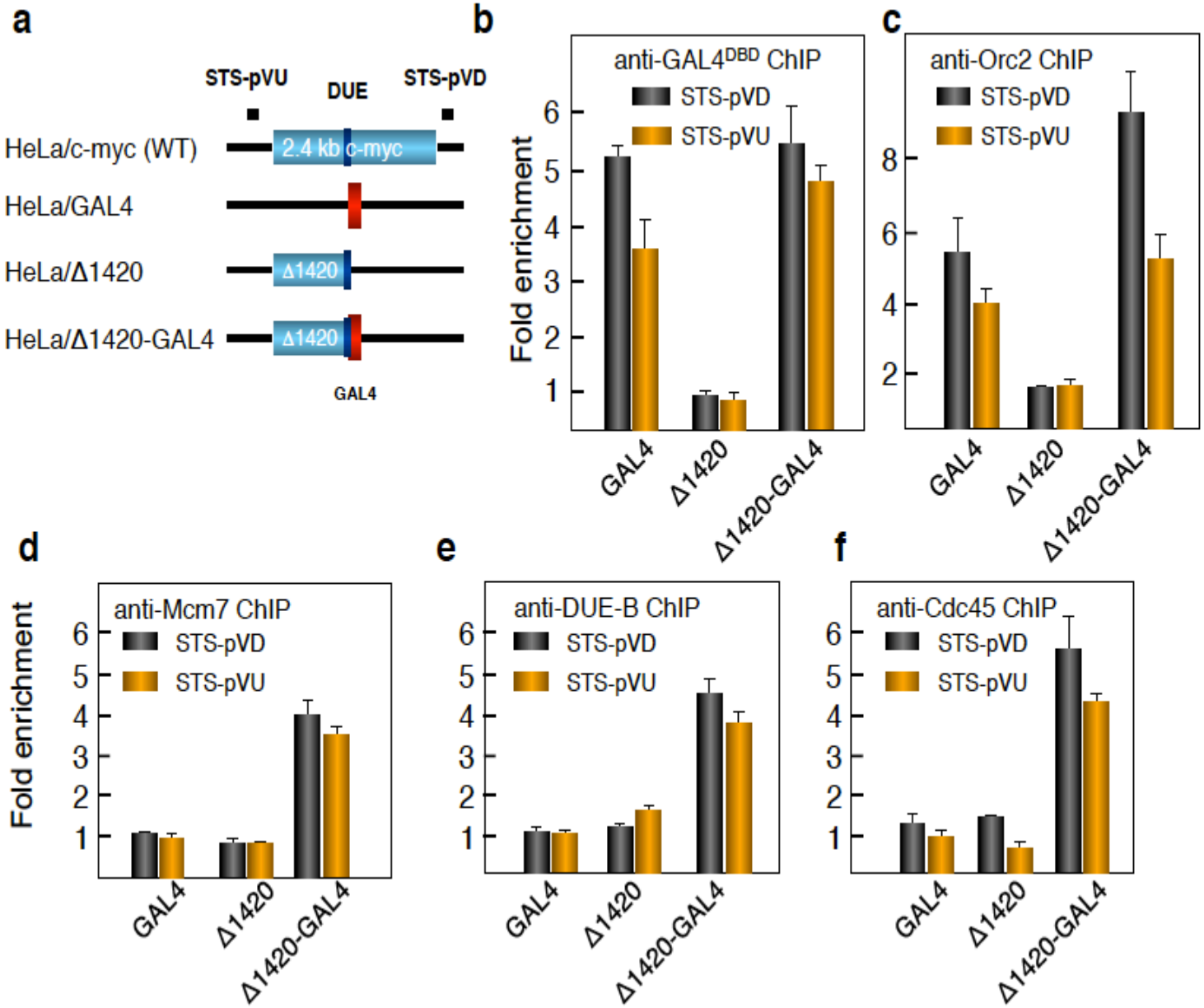

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4

**Figure 23. Tethered Orc2 binding recovered the DNA replication activity.**

**a.** Western blot of Orc2-GAL4$^{DBD}$ protein. Rabbit polyclonal antibody against human Orc2 was applied to detect the expression of Orc2-GAL4$^{DBD}$ and endogenous Orc2 was used as loading control.

**b**. DNA replication activity tested by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8 from another replication origin, β-globin, was selected as a internal normalizer to compare different nascent DNA preparations. Error bars, standard deviations of three repetitive experiments.

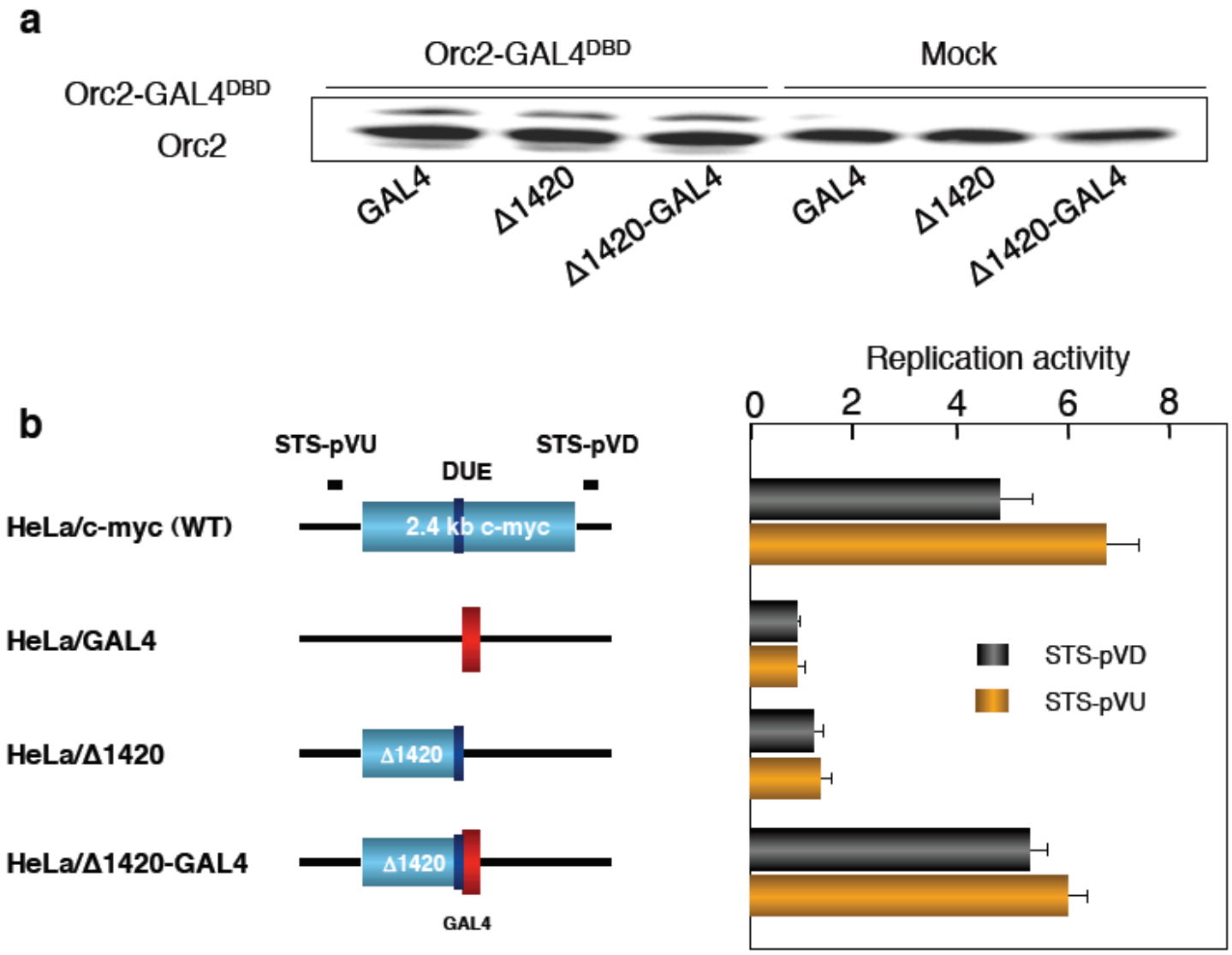

a
Orc2-GAL4DBD
Orc2
Orc2-GAL4DBD
Mock
GAL4
Δ1420
Δ1420-GAL4
GAL4
Δ1420
Δ1420-GAL4
b
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
Replication activity
0
2
4
6
8
STS-pVD
STS-pVU

**Figure 24. Tethered binding of Cdt1 recruited Orc2, Mcm7, DUE-B and Cdc45.**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b-f**. Recruitment of $GAL4^{DBD}$–Cdt1 fusion proteins and endogenous Orc2, Mcm7, DUE-B and Cdc45, were detected by ChIP. See materials and methods for detailed ChIP assay. Error bars, standard deviations of three repetitive experiments.

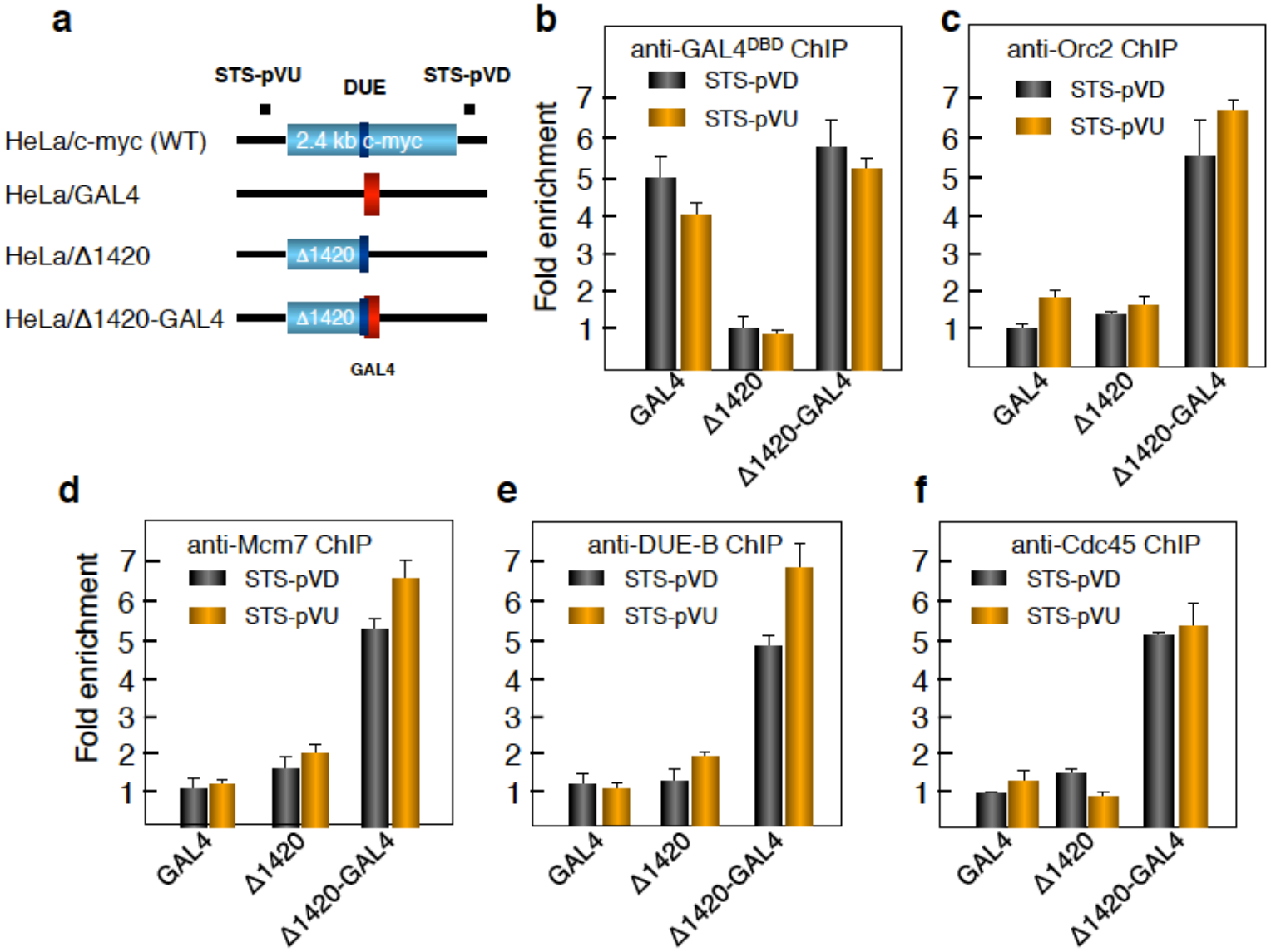
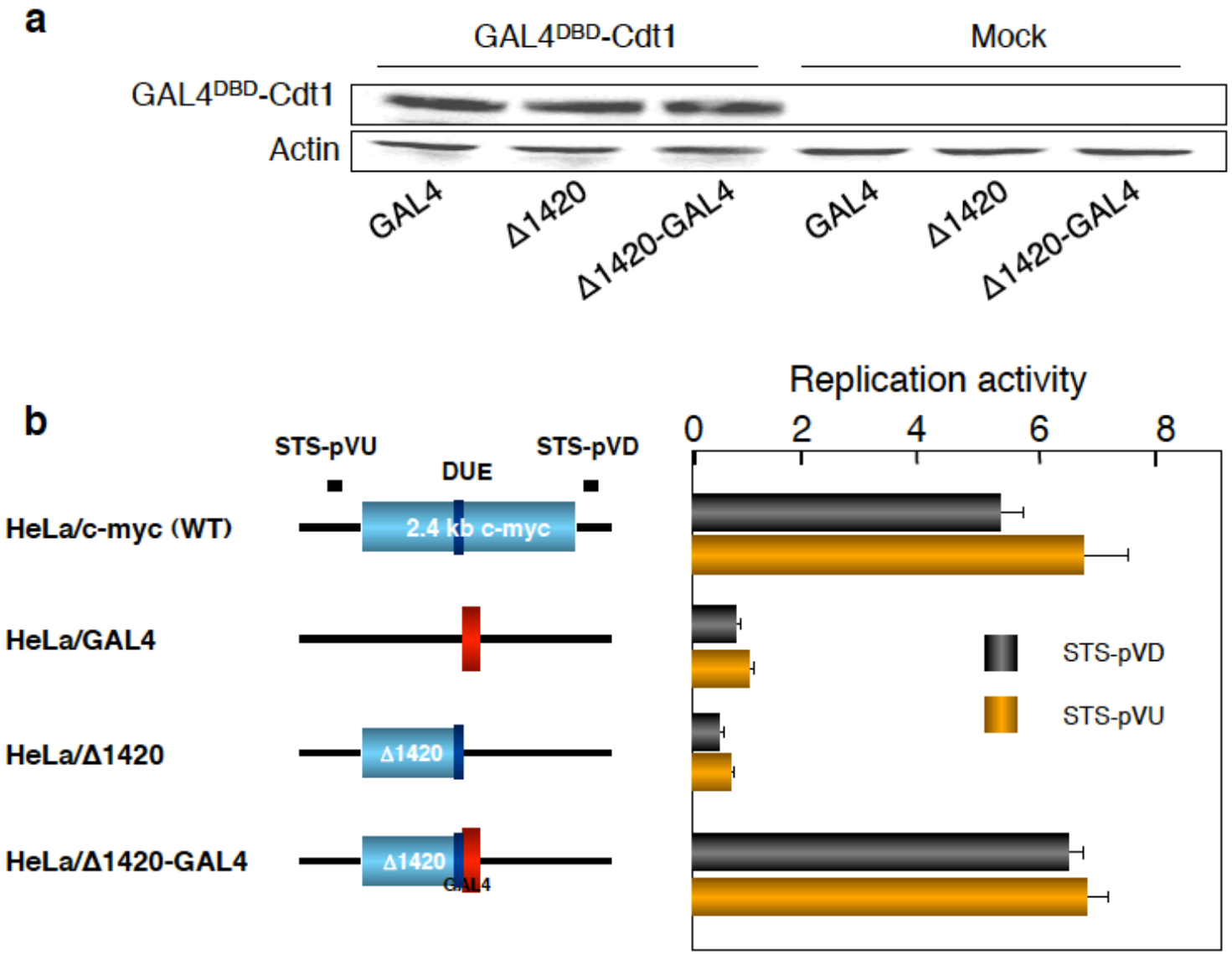

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4

**Figure 25. Tethered Cdt1 binding recovered the DNA replication activity.**

**a.** Western blot of $GAL4^{DBD}$–Cdt1protein. Mouse polyclonal antibody against $GAL4^{DBD}$ was applied to detect the expression of $GAL4^{DBD}$–Cdt1protein and endogenous actin was used as loading control.

**b**. DNA replication activity tested by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8 from another replication origin, β-globin, was selected as a internal normalizer to compare different nascent DNA preparations. Error bars, standard deviations of three repetitive experiments.


a
GAL4DBD-Cdt1
Mock
GAL4DBD-Cdt1
Actin
GAL4
Δ1420
Δ1420-GAL4
GAL4
Δ1420
Δ1420-GAL4
b
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
Replication activity
0
2
4
6
8
STS-pVD
STS-pVU

***Tethered binding of Mcm7 could not recruit Orc2, DUE-B or Cdc45 and failed to recover DNA replication activity.***

MCM complex is recruited by ORC during the formation of pre-RC, and functions as a helicase (Ilves et al., 2010; Kaplan and Bruck, 2010; Pospiech et al., 2010). ORC binding and the following recruitment of Cdc6 and Cdt1 is to prepare for the binding of MCM. The above results from Cdt1 and Orc2 seems to indicate that any single component of pre-RC or pre-IC might be able to recover DNA replication activity. To test this idea, I further tested the effect of tethered binding of Mcm7-GAL4$^{DBD}$ fusion protein. However, ChIP assay demonstrated that tethered binding of Mcm7 (Figure 26a,b, d) did not recruit Orc2 (Figure 26c), DUE-B (Figure 26e), or Cdc45 (Figure 26f). Meanwhile, as shown in Figure 27 b, tethered binding of Mcm7 failed to recover DNA replication activity in three cell lines based on NDAA. The Western blot (Figure 27a) ruled out the possibility that the failed recovery of replication activity is due to the poor expression of Mcm7-GAL4$^{DBD}$.

**Figure 26. Tethered binding of Mcm7-GAL4$^{DBD}$ did not recruit Orc2, DUE-B and Cdc45.**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b-f**. Recruitment of GAL4$^{DBD}$-Mcm7 fusion proteins and endogenous Orc2, DUE-B and Cdc45, were detected by ChIP. See materials and methods for detailed ChIP assay. Error bars, standard deviations of three repetitive experiments.

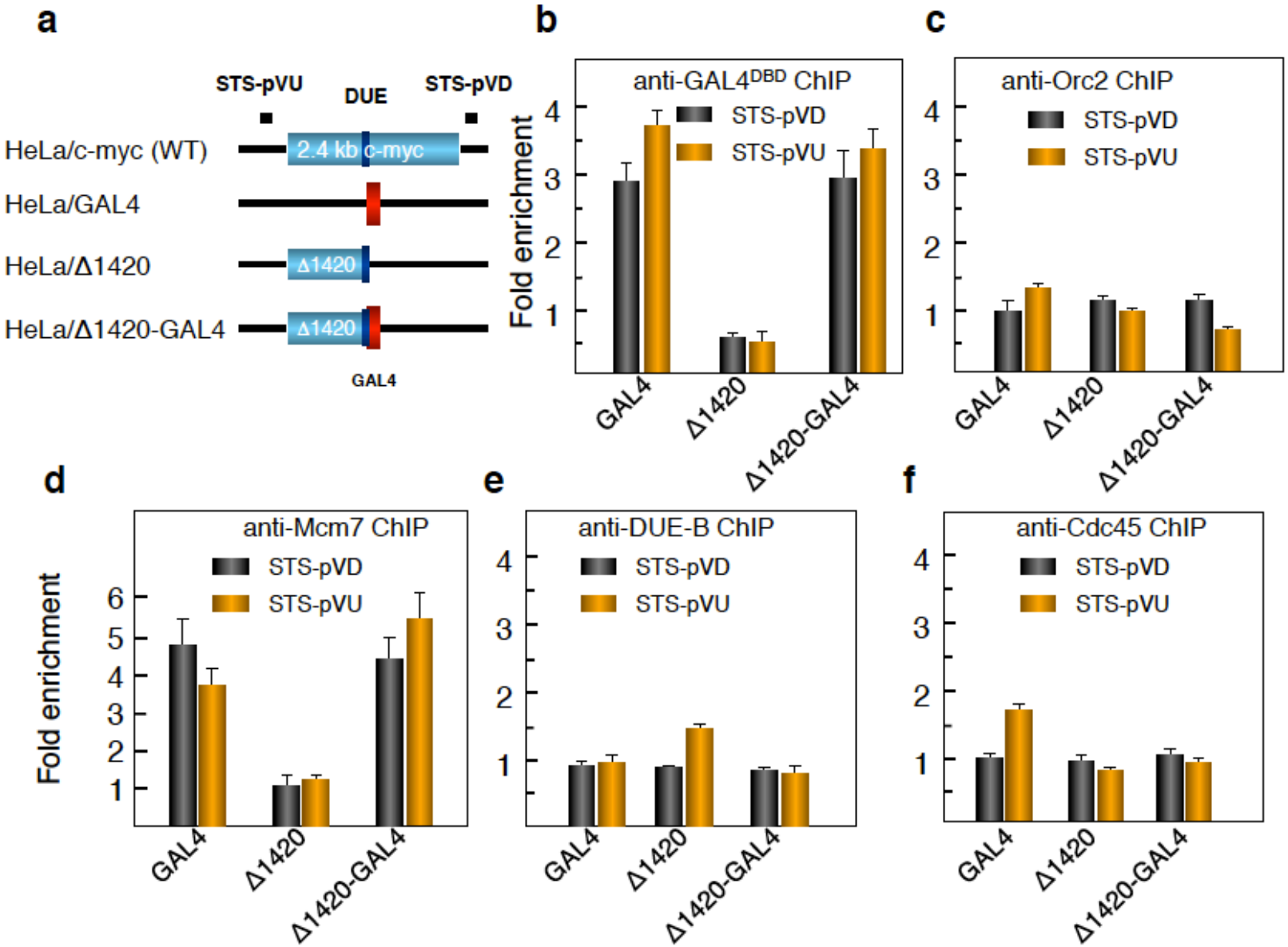
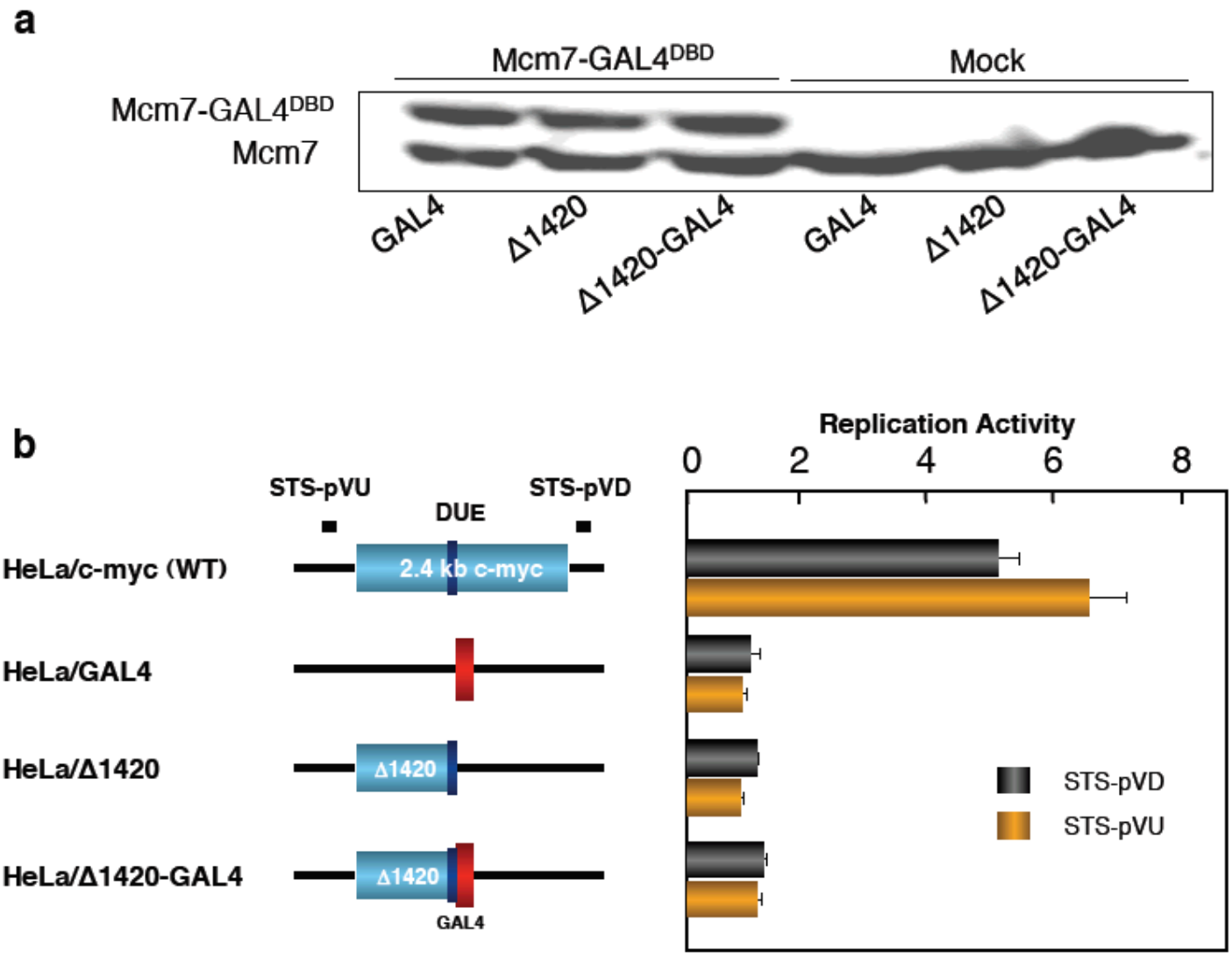

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4

**Figure 27. Tethered binding of Mcm7-GAL4$^{DBD}$ did not recover the DNA replication activity.**

**a.** Western blot of Mcm7-GAL4$^{DBD}$ protein. Rabbit polyclonal antibody against Mcm7 was applied to detect the expression of Mcm7-GAL4$^{DBD}$ protein and endogenous Mcm7 was used as a loading control.

**b**. DNA replication activity tested by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8 from another replication origin, β-globin, was selected as a internal normalizer to compare different nascent DNA preparations. Error bars, standard deviations of three repetitive experiments.


a
Mcm7-GAL4DBD
Mock
Mcm7-GAL4DBD
Mcm7
GAL4
Δ1420
Δ1420-GAL4
GAL4
Δ1420
Δ1420-GAL4
b
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
Replication Activity
0
2
4
6
8
STS-pVD
STS-pVU

**Pre-IC component**

***Tethered binding of Cdc45 could not recruit Orc2, Mcm7, DUE-B and failed to recover DNA replication activity.***

As shown in Figure 4, as a co-factor of MCM, Cdc45 function to activate the helicase activity of MCM upon binding on pre-RC, and the binding of Cdc45 indicates formation of pre-initiation complex (pre-IC). I tested the effect of tethered binding of Cdc45-GAL4$^{DBD}$ on replication activation on the mutant c-myc replicators. As shown in Figure 28, the ChIP data clearly demonstrated that there are no Orc2 (Figure 28a, b, c), Mcm7 (Figure 28d), and DUE-B (Figure 28e) recruitments upon tethered binding of Cdc45 (Figure 28f) on GAL4 binding site. Consequently, Cdc45-GAL4$^{DBD}$ binding in a chromosomal context did not confer a DNA replication origin activity on the ectopic chromosomal site in the three HeLa cell lines (Figure 29b).

**Figure 28. Tethered binding of Cdc45-GAL4$^{DBD}$ could not recruit Orc2, Mcm7 and DUE-B.**

**a.** Isogenic cell lines used for analysis of trans-acting replication factors (refer to Figure 15 legend)**.**

**b-f**. Recruitment of GAL4$^{DBD}$-Cdc45 fusion proteins and endogenous Orc2, Mcm7, DUE-B, was detected by ChIP. See materials and methods for detailed ChIP assay. Error bars, standard deviations of three repetitive experiments.

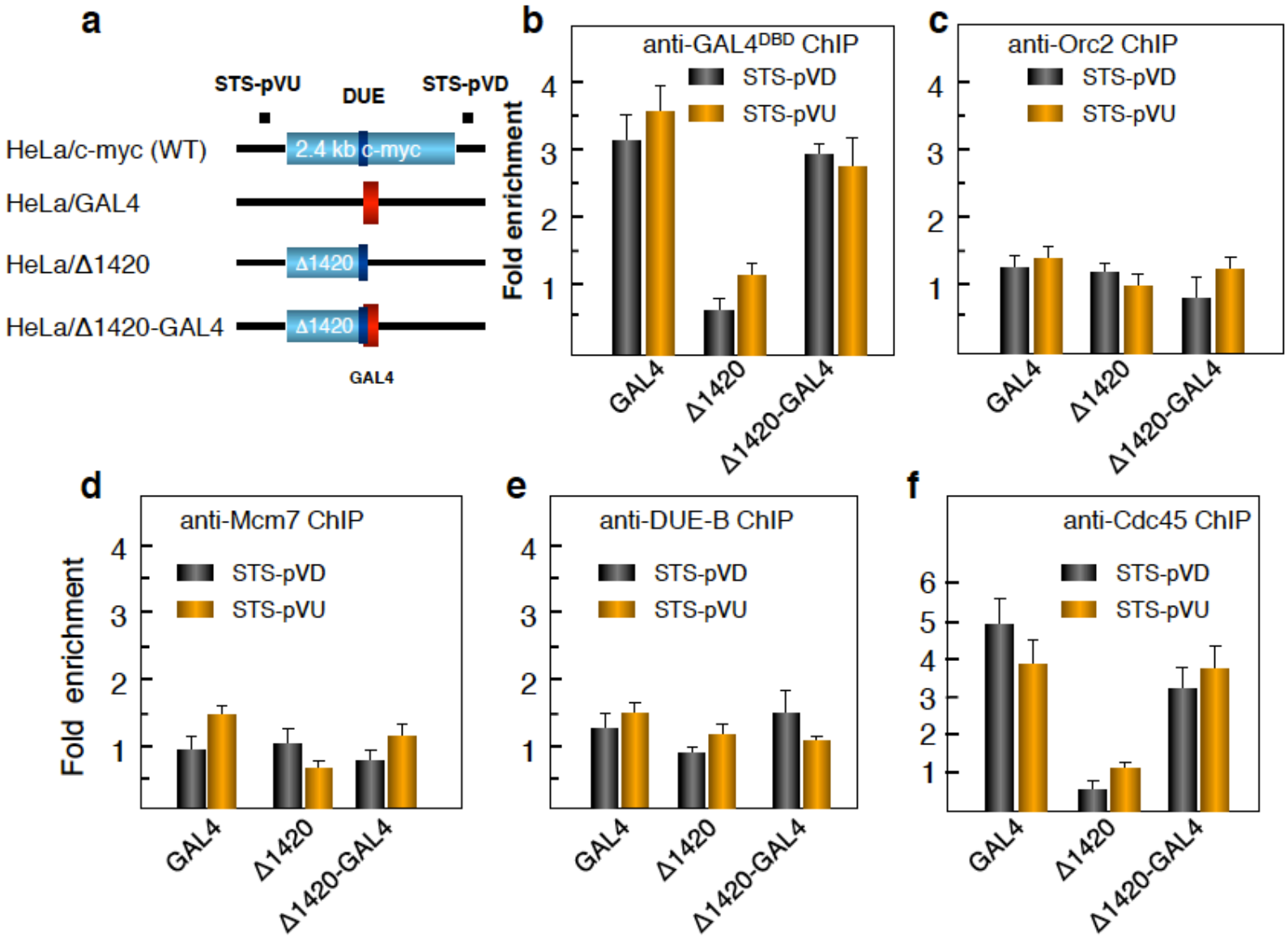
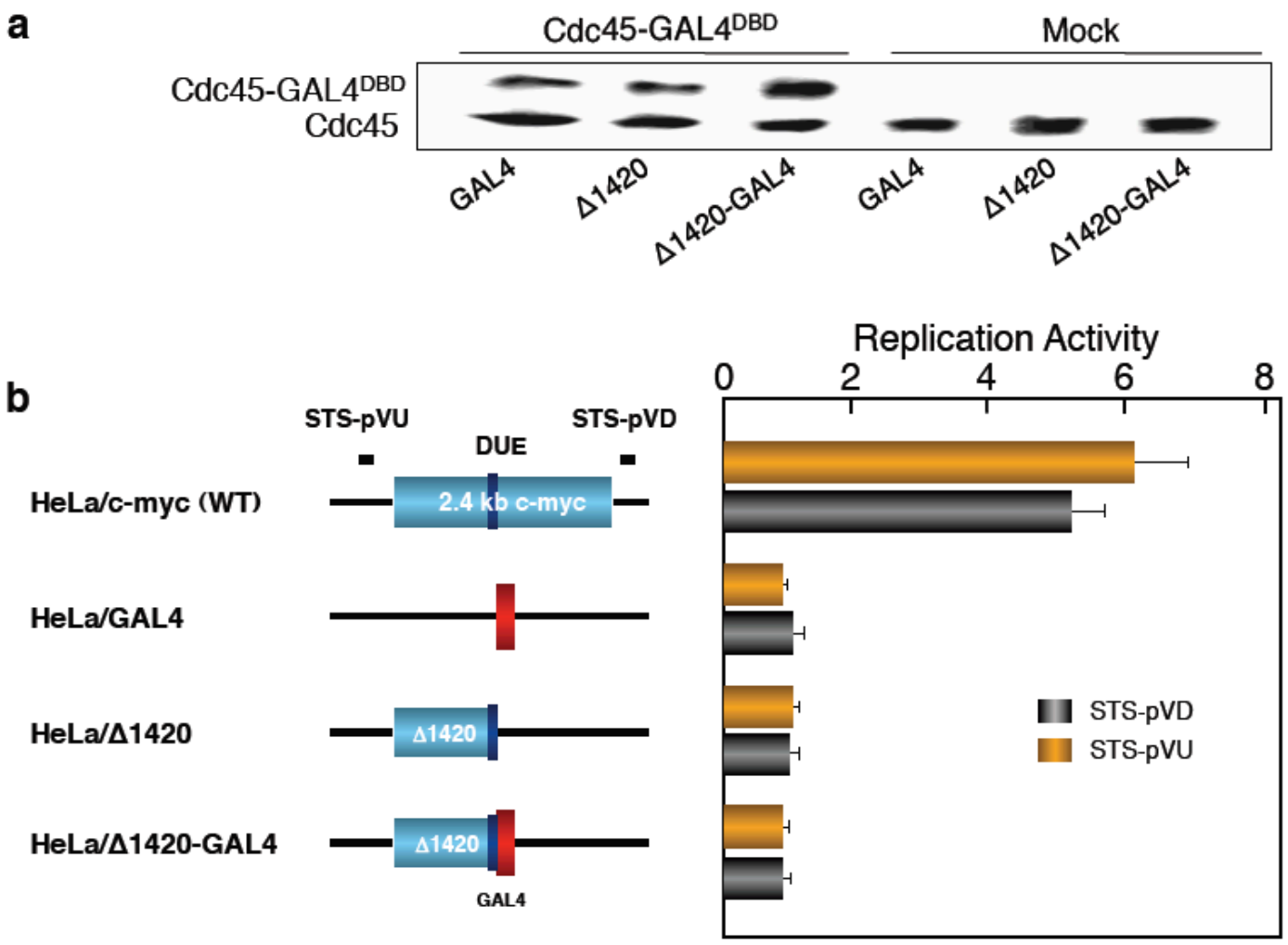

a
STS-pVU
DUE
STS-pVD
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
b
anti-GAL4DBD ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
c
anti-Orc2 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
d
anti-Mcm7 ChIP
STS-pVD
STS-pVU
Fold enrichment
GAL4
Δ1420
Δ1420-GAL4
e
anti-DUE-B ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4
f
anti-Cdc45 ChIP
STS-pVD
STS-pVU
GAL4
Δ1420
Δ1420-GAL4

**Figure 29. Tethered binding of Cdc45-GAL4$^{DBD}$ could not recover the DNA replication activity.**

**a.** Western blot of Cdc45-GAL4$^{DBD}$. Rabbit polyclonal antibody against GAL4$^{DBD}$ was applied to detect the expression of Cdc45-GAL4$^{DBD}$ and endogenous Cdc45 was used as loading control.

**b**. DNA replication activity tested by relative nascent DNA abundance at STS-pVD (black) and STS-pVU (orange). STS-54.8 from another replication origin, β-globin, was selected as a internal normalizer to compare different nascent DNA preparations. Error bars, standard deviations of three repetitive experiments.


a
Cdc45-GAL4DBD
Mock
Cdc45-GAL4DBD
Cdc45
GAL4
Δ1420
Δ1420-GAL4
GAL4
Δ1420
Δ1420-GAL4
b
STS-pVU
STS-pVD
DUE
HeLa/c-myc (WT)
2.4 kb c-myc
HeLa/GAL4
HeLa/Δ1420
Δ1420
HeLa/Δ1420-GAL4
Δ1420
GAL4
Replication Activity
0
2
4
6
8
STS-pVD
STS-pVU

## SUMMARY II

In this project I tested the tethered binding of trans-acting factors and DNA replication origin activity in human chromosome.

I found that tethered binding of E2F1 modulated chromatin structure and induced the replication origin activity on human chromosome in HeLa/Δ1420-GAL4 cells but not in HeLa/Δ1420 cells and HeLa/GAL4 cells indicating that nucleosome acetylation favored DNA replication origin activation. I also found that tethered binding of Orc2 recruited Mcm7, DUE-B, and Cdc45, and activated DNA replication activity on the human chromosome in HeLa/Δ1420-GAL4 cells but not in the HeLa/Δ1420 cells and HeLa/GAL4 cells. Surprisingly, I found that tethered binding of Cdt1 recruited Orc2, Mcm7, DUE-B, and Cdc45 and activated DNA replication activity on the human chromosome in HeLa/Δ1420-GAL4 cells but not in the HeLa/GAL4 cells and HeLa/Δ1420 cells.

On the contrary, tethered binding of Mcm7 could not recruit Orc2, DUE-B, and Cdc45 and failed to confer a replication activity in HeLa/Δ1420-GAL4 cells, HeLa/Δ1420 cells and HeLa/GAL4 cells. Similarly, tethered binding of Cdc45 could not recruit Orc2, Mcm7, DUE-B and Cdc45 and failed to confer a replication activity in HeLa/Δ1420-GAL4 cells, HeLa/Δ1420 cells, and HeLa/GAL4 cells.

Taken together, tethered binding experiments demonstrated 1) the binding of MCM, DUE-B and Cdc45 is ORC dependent and 2) DNA sequence in the mutant c-myc replicator (Δ1420) is essential for induced replication origin activity on human chromosome.

# PART III. PROBING THE FORMATION OF HAIRPINS BY $(CTG)_N \bullet (CAG)_N$ REPEAT TRACTS DURING DNA REPLICATION IN HUMAN CELLS IN VIVO

Previous studies in prokaryotes and lower eukaryotes implied that $(CTG)_n \bullet (CAG)_n$ repeat tracts formed hairpin structures, which led to expansion and contraction (Lenzmeier and Freudenreich, 2003; Ohshima and Wells, 1997; Pearson et al., 1998). We designed zinc finger nucleases with DNA sequence binding specificity to detect the possibility to form hairpins by $(CTG)_n \bullet (CAG)_n$ repeat tracts in human chromosome.

**Assembly of three-zinc finger nuclease targeting on TNR repeat tract**

Based on the database (Zinc Finger Consortium) (Sander et al., 2010; Sander et al., 2007), we designed two three-zinc finger nucleases, named $ZFN_{CAG}$ and $ZFN_{CTG}$, in which the zinc finger protein targeting on CAG repeat ($ZFP_{CAG}$) and zinc finger protein targeting on CTG repeat ($ZFP_{CTG}$) domains were designed to recognize repeating sequence of GCA and GCT respectively (Figure 30a,b, c). The assembled ZFPs with three zinc fingers (Figure 32a) were confirmed by PCR (Figure 31b) and then subcloned into a mammalian cell expression vector, pST1374, between Xba I and BamH I restriction sites (Figure

32a). Each ZFN contains a nuclear localization sequence (NLS), a FLAG tag (FLAG) and a nuclease cleavage domain of Fok I, a type II restriction enzyme, which cleaves DNA in a sequence nonspecific manner. The molecular weight of ZFN, $ZFN_{CAG}$ or $ZFN_{CTG}$, is about 36 kDa (Figure 32b). ZFPs were constructed by introducing a protein translational stop codon at BamH I restriction site. The predicted molecular weight of each ZFP is about 17 kDa (Figure 32b). Unlike ZFN, a ZFP contains a DNA binding domain without DNA cleavage activity due to the omission of the Fok I cleavage domain.

**Expression of ZFNs and ZFPs in HeLa cells**

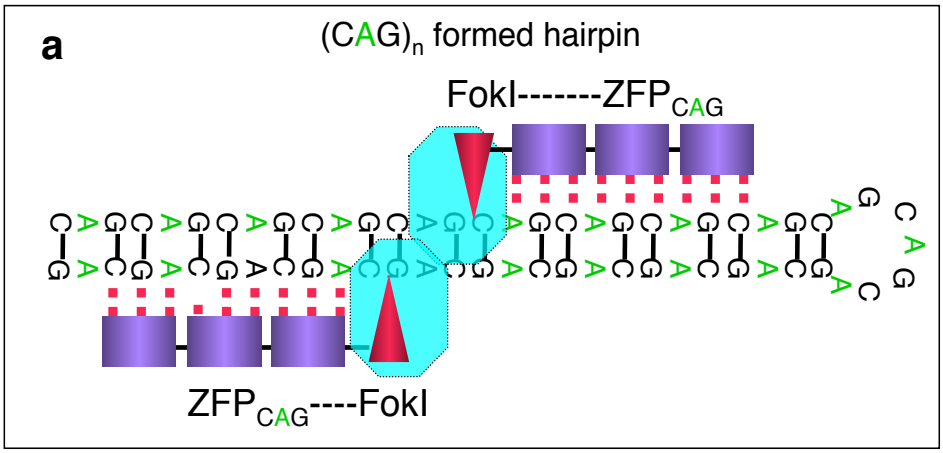
About 10 μg ZFN or ZFP expression plasmids were used to transfect $1\times10^7$ HeLa cells. After 24 hr, cells were harvested and whole cell extract was used for Western blot. As shown in Figure 32b, $ZFN_{CAG}$ or $ZFN_{CTG}$ was detected by anti-FLAG antibody, indicating the expression of $ZFN_{CAG}$ or $ZFN_{CTG}$ in HeLa cells. However, I noticed that ZFNs expressed at very low level in HeLa cells. Possibly, ZFNs are toxic to human cells due to potential binding to TNR formed hairpins along the human genome leading to DNA damage.

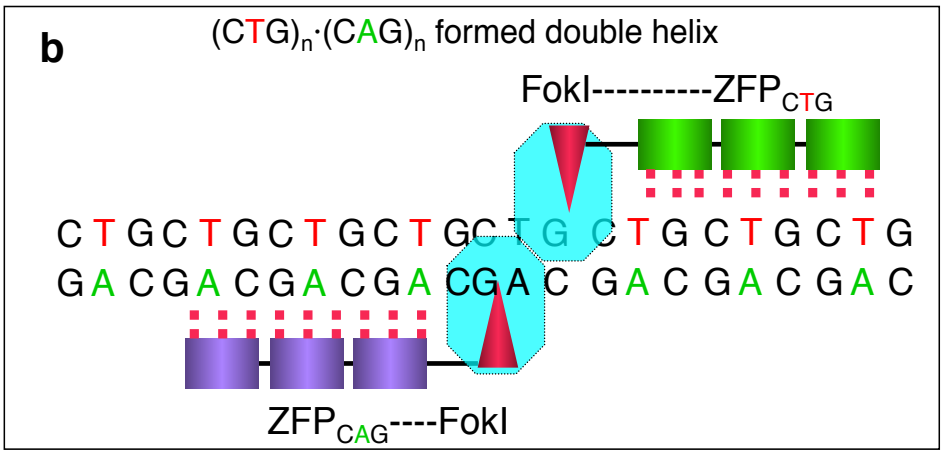
Experiments were carried out to detect the cytoxicity of ZFNs and ZFPs using trypan blue staining. As shown in Figure 33, transfection with single ZFN caused about 15% increase of cells death compare to transfection with single ZFP. On the contrary, transfection with equal amount of DNA containing double ZFNs caused about 27% increase of cell death compare to transfection with

double ZFP (~1.0% cell death), suggesting ZFN is the major reason that leads to cell death.

**Figure 30. Design of ZFNs targeting hairpins formed by $(CTG)_n$•$(CAG)_n$**

Zinc-finger nucleases (ZFNs) were designed to target $(CAG)_n$ formed hairpins (**a**), $(CTG)_n$ formed hairpin (**c**) and double stranded fragment of $(CTG)_n$•$(CAG)_n$ repeat tract (**b**). Each ZFN contains two parts, ZFP (purple or green) contains three zinc finger protein providing DNA binding ability with DNA sequence specificity. Nuclease (red arrow) is nuclease cleavage domain of restriction enzyme, Fok I. $ZFP_{CAG}$-Fok I and $ZFP_{CTG}$-Fok I would form heterodimer targeting double stranded fragment of $(CTG)_n$•$(CAG)_n$ repeats tract. But, $ZFP_{CAG}$-Fok I or $ZFP_{CTG}$-Fok I would form homodimers targeting on $(CAG)_n$ or $(CTG)_n$ formed hairpins respectively.


a
(CAG)n formed hairpin
FokI-------ZFPCAG
ZFPCAG----FokI


b
(CTG)n·(CAG)n formed double helix
FokI----------ZFPCTG
C T G C T G C T G C T G C T G C T G C T G C T G
G A C G A C G A C G A C G A C G A C G A C G A C
ZFPCAG----FokI

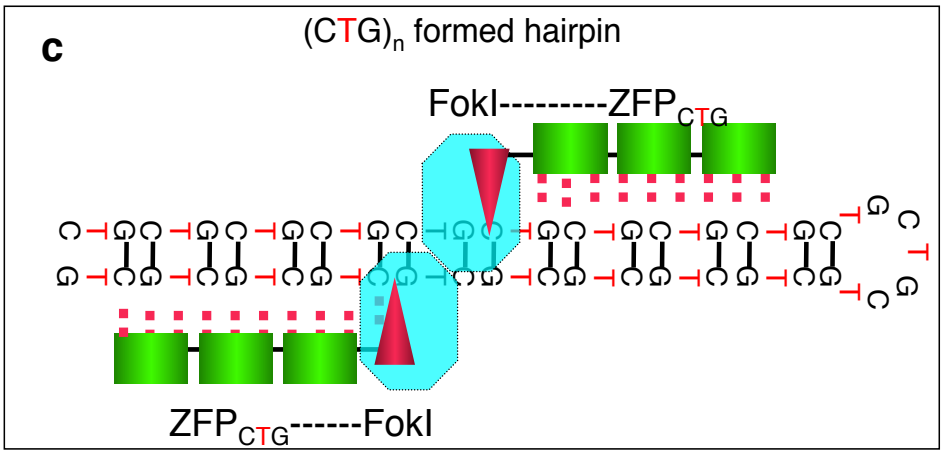
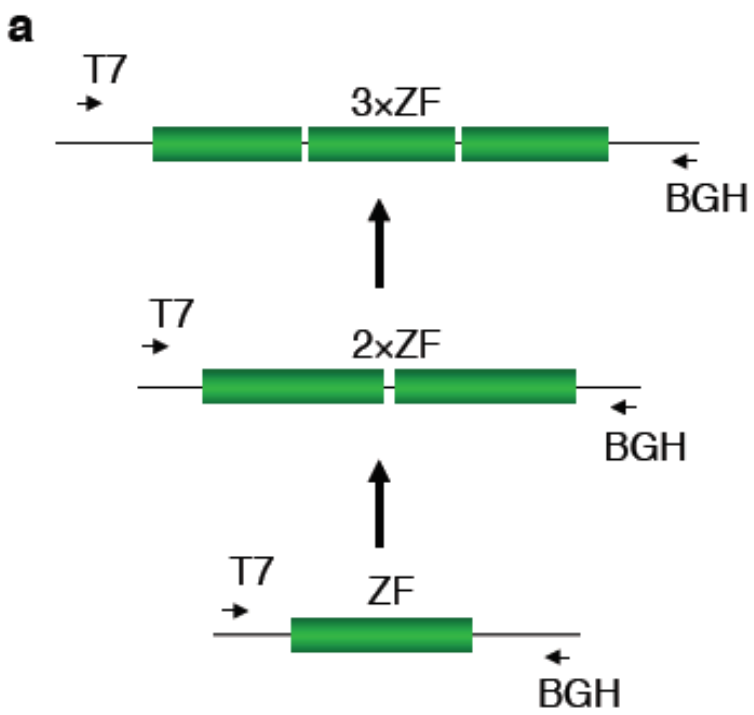

c
(CTG)n formed hairpin
FokI---------ZFPCTG
ZFPCTG------FokI

**Figure 31. Assembly of zinc fingers**

**a**. Starting from one zinc finger containing plasmid, the second and third zinc finger protein sequence were inserted sequentially to generate a three-zinc finger plasmid. T7 and BGH primer pair (arrows, panel **a**) was used to check the size and number of zinc fingers.

**b**. Electrophoresis of PCR product amplified by T7 and BGH primers. From left to right, M represents DNA marker; lane1 and 6: one-zinc finger; lane 2 and 5: two zinc-finger; lane 3 and 4: three-zinc finger.


a
T7
3xZF
BGH
T7
2xZF
BGH
T7
ZF
BGH

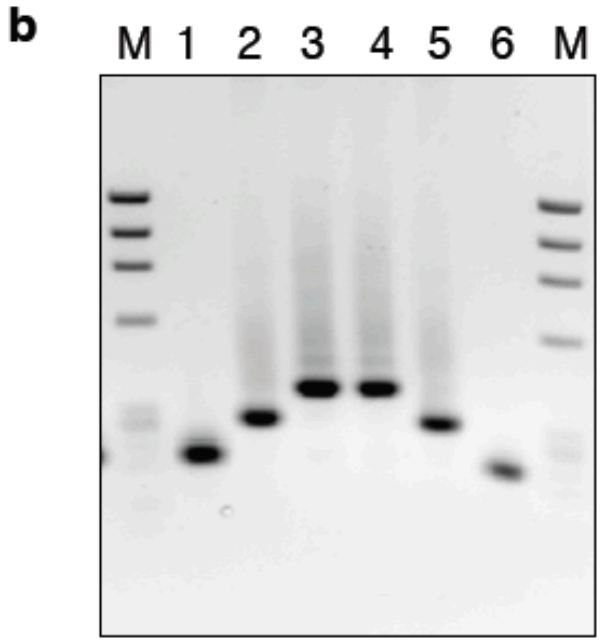

b
M 1 2 3 4 5 6 M

**Figure 32. Expression of ZFPs and ZFNs in HeLa cells**

**a**. Strategy to make ZFP and ZFN expression plasmids. Three zinc fingers were subcloned into mammalian expression vector pST1374 between FLAG tag and Fok I nuclease cleavage domain generating ZFN. ZFPs were made by introducing a translational stop codon between the third zinc finger and Fok I nuclease cleavage domain. See material and methods for detailed methods to construct the ZFP plasmids .

**b**. Expression of $ZFP_{CTG}$, $ZFN_{CTG}$, $ZFP_{CAG}$ and $ZFN_{CAG}$ in HeLa cells. FLAG antibody was used in Western blot and actin was a loading control.

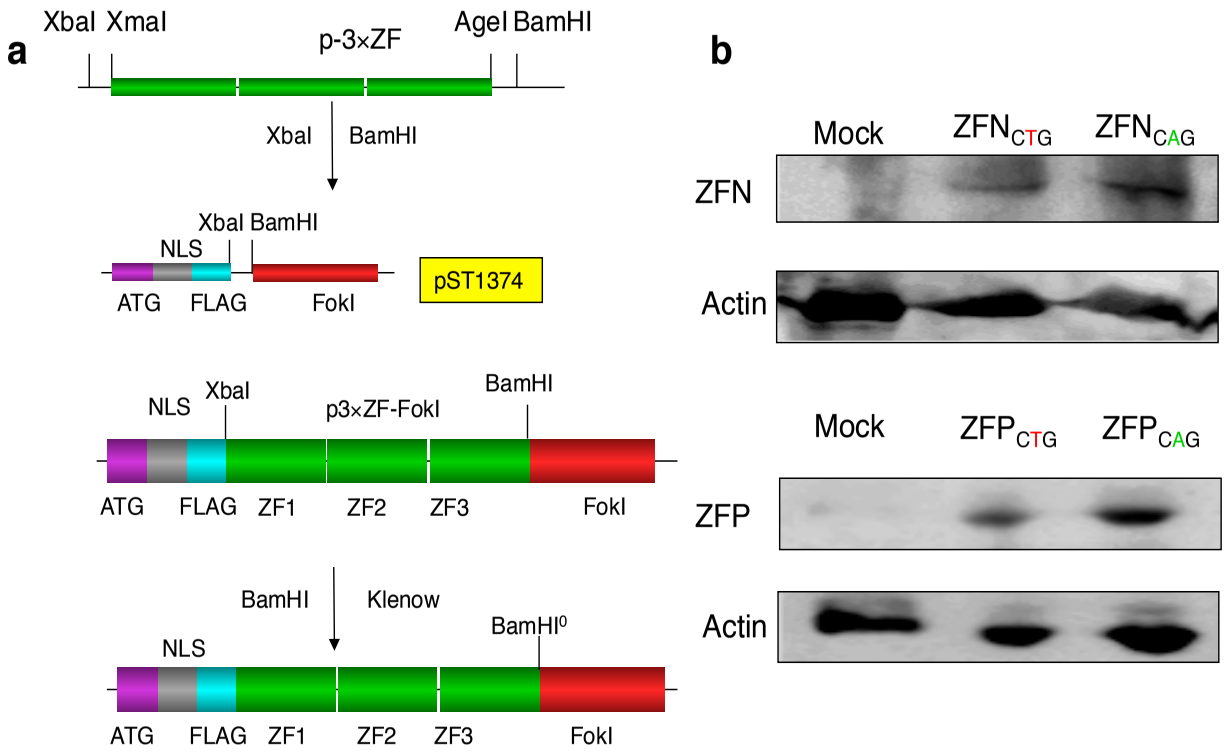
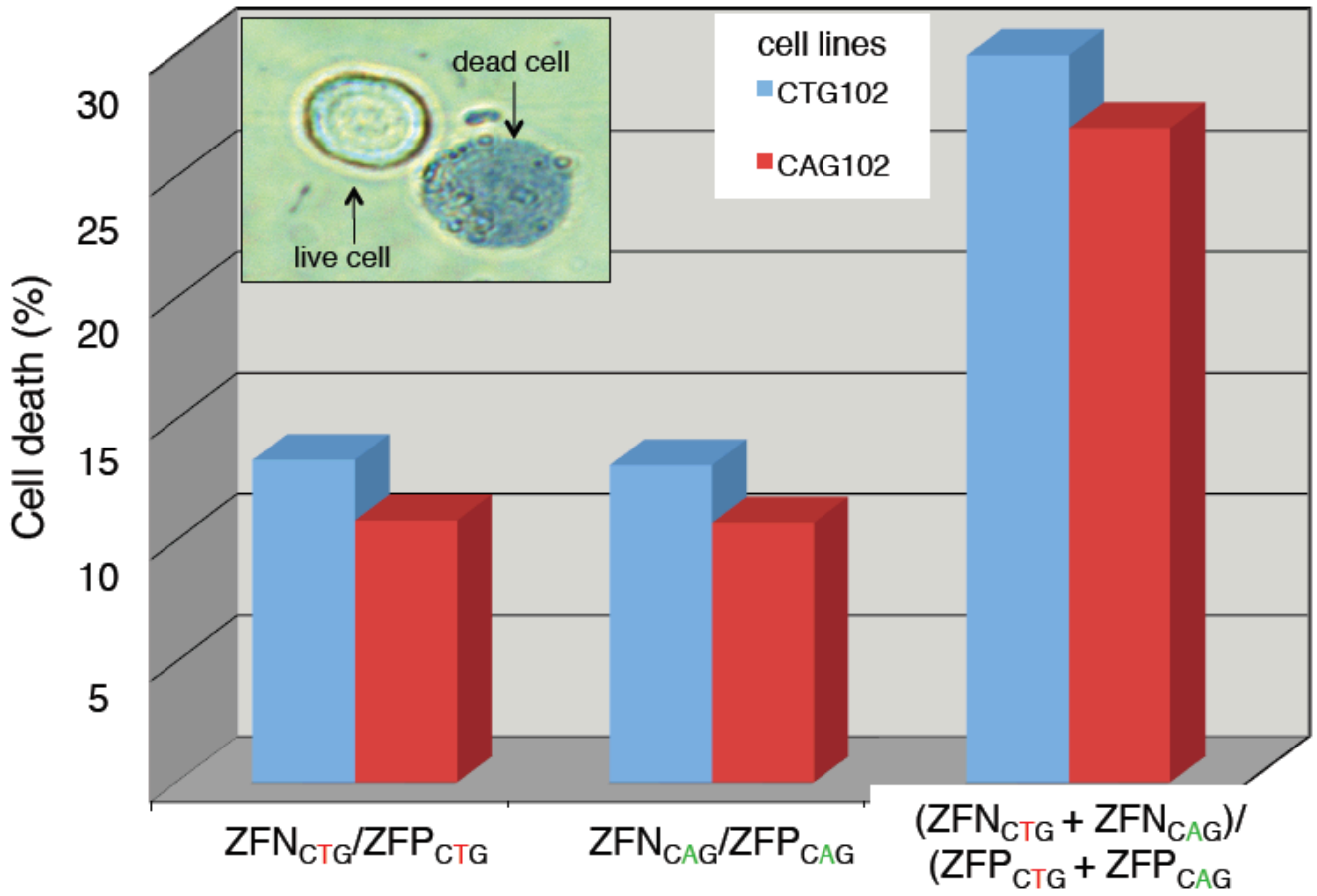

a
XbaI XmaI
p-3×ZF
AgeI BamHI
XbaI
BamHI
XbaI BamHI
NLS
ATG
FLAG
FokI
pST1374
NLS
XbaI
p3×ZF-FokI
BamHI
ATG
FLAG
ZF1
ZF2
ZF3
FokI
BamHI
Klenow
BamHI⁰
NLS
ATG
FLAG
ZF1
ZF2
ZF3
FokI
b
Mock
ZFN$_{CTG}$
ZFN$_{CAG}$
ZFN
Actin
Mock
ZFP$_{CTG}$
ZFP$_{CAG}$
ZFP
Actin

**Figure 33. Toxicity of ZFNs**

Trypan blue staining was applied to test the cell viability (shown in the upper left corner in this figure). Blue cell: Stained non-viable; Clear cell: viable cells excluded the staining.

Cell death was quantitated after transfection with plasmids expressing ZFN and ZFP. Y-axis: increased cell death after treatment with ZFN normalized by cell death treatment with ZFP. X-axis: treatment by either single nuclease ($ZFN_{CTG}$ or $ZFN_{CAG}$) or double nucleases ($ZFN_{CTG}$ and $ZFN_{CAG}$). Blue bar: $(CTG)_{102}$ cell line; red bar: $(CAG)_{102}$ cell line.


Cell death (%)
30
25
20
15
10
5
dead cell
live cell
cell lines
CTG102
CAG102
ZFN$_{CTG}$/ZFP$_{CTG}$
ZFN$_{CAG}$/ZFP$_{CAG}$
(ZFN$_{CTG}$ + ZFN$_{CAG}$)/
(ZFP$_{CTG}$ + ZFP$_{CAG}$

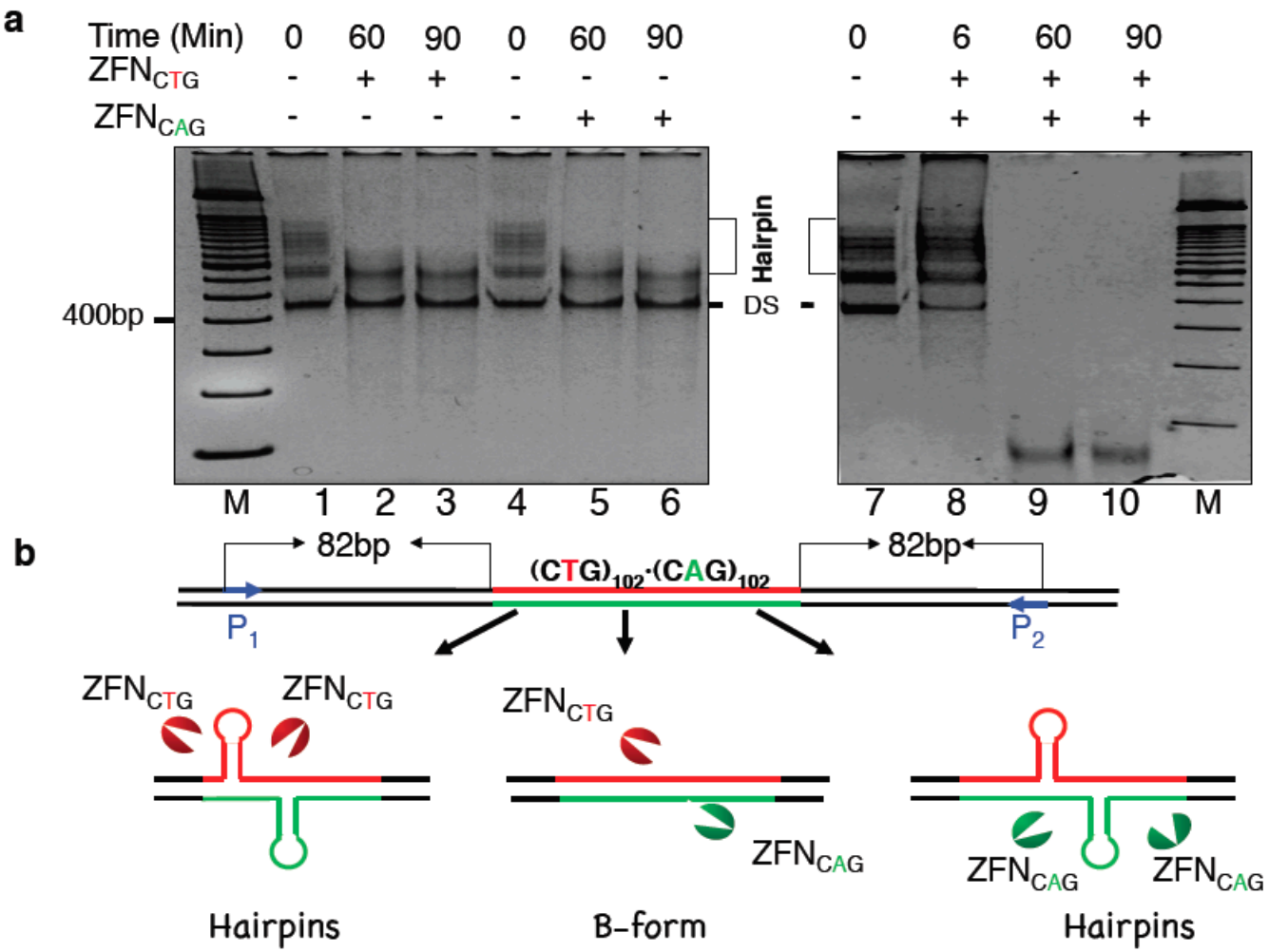
**Function of $ZFN_{CTG}$ and $ZFN_{CAG}$ on TNR in vitro**

Like Fok I alone, ZFNs dimerize through their Fok I cleavage domains only when bound to DNA, and this dimerization is necessary for catalytic activity (Carroll et al., 2006). Using immunoprecipitation of purified fusion protein FLAG-ZFNs, I carried out in vitro experiments in which PCR products containing $(CAG)_{102}$•$(CTG)_{102}$ were used as substrates. Previous experiments from our laboratory (Liu, unpublished) and others (Chastain et al., 1995; Pearson et al., 1998) suggested that PCR products containing $(CAG)_{102}$•$(CTG)_{102}$ are a mixture of several forms of structures including linear double strand DNA and secondary structures including slipped DNA with hairpins on one or both strands, which showed different migration rate on polyacrylamide gel. In general, the slipped DNA migrates slower than the linear DNA does. In my experiment, ZFNs were used to treat above substrates in vitro. As shown in Figure 34, some of the hairpin forms of $(CAG)_{102}$•$(CTG)_{102}$ (slipped DNA, bands with lower mobility) were cleaved by either ZFN separately (Figure 34, lane 1-3, 4-6). Single digestion did not cleave double strand DNA (DS, band with higher mobility). However, double strand DNA was cleaved faster than the slipped strand (hairpin) forms of $(CAG)_{102}$•$(CTG)_{102}$ by double ZFN digestion after 1 hr and 1.5 hr (Figure 34, lane 9, 10) indicating the cleavage occurred inside the TNR region. Taken together, in vitro, our engineered ZFNs targeted TNR formed hairpins in pure DNA by forming homodimers while ZFNs targeted double stranded TNR by forming heterodimer.

**Figure 34. Functional analysis of ZFNs in vitro**

**a**. Single digestion with time course; PCR product of $(CTG)_{102}$•$(CTG)_{102}$ was treated by single ZFN immunoprecipitates, $ZFN_{CTG}$ (lane 1-3) or $ZFN_{CAG}$ (lane 4-6). Double digestion with time course; PCR product of $(CTG)_{102}$•$(CTG)_{102}$ was treated with 1:1 mixture of single ZFN immunoprecipitates, $ZFN_{ctg}$ plus $ZFN_{cag}$ (lane 7-10). DS: double stranded DNA. M: DNA marker.

**b**. Diagram of possible cleavage.


a
Time (Min) 0 60 90 0 60 90 0 6 60 90
ZFN$_{CTG}$ - + + - - - - + + +
ZFN$_{CAG}$ - - - - + + - + + +
400bp
Hairpin
DS
M 1 2 3 4 5 6 7 8 9 10 M
b
82bp
(CTG)$_{102}$·(CAG)$_{102}$
82bp
P$_1$
P$_2$
ZFN$_{CTG}$ ZFN$_{CTG}$
ZFN$_{CTG}$
ZFN$_{CAG}$
ZFN$_{CAG}$ ZFN$_{CAG}$
Hairpins
B-form
Hairpins

## Function of $ZFN_{CTG}$ and $ZFN_{CAG}$ on TNR in vivo

### $(CTG)_{12}$•$(CAG)_{12}$ was able to be cleaved by double nuclease, $ZFN_{CTG}$+$ZFN_{CAG}$

Previously, our lab constructed several FRT acceptor cell lines in which different number of $(CAG)_n$•$(CTG)_n$ was connected with the c-myc replicator at the 3' with two orientations. c-myc-$(CTG)_n$•$(CAG)_n$ or c-myc-$(CAG)_n$•$(CTG)_n$ refer to CTG repeat tract or CAG repeat tract on the lagging strand template when DNA replication starts from c-myc origin (Liu et al., 2010). We carried out the ZFN in vivo experiment using four cell lines, two of them harbored normal number of repeat tract, $(CTG)_{12}$•$(CAG)_{12}$ repeat tract, the other two cell lines contain expanded number of repeat tracts, $(CTG)_{102}$•$(CAG)_{102}$.

HeLa/c-myc-$(CTG)_{12}$•$(CAG)_{12}$, or HeLa/c-myc-$(CAG)_{12}$•$(CTG)_{12}$ cell lines were transfected with double nucleases, $ZFN_{CTG}$+$ZFN_{CAG}$. About 250 pg genomic DNA was used as small pool PCR template. PCR primer $P_1$ and $P_2$ (Figure 35b) were used to amplify the fragment containing the full length repeat tract, $(CTG)_{12}$•$(CAG)_{12}$ and flanking sequences. The size of wild type fragment without contraction or expansion should be about 200 bp. As shown in Figure 35a, lane 17-28, except the wild type PCR product other smaller bands were also detected, indicating that $ZFN_{CTG}$+$ZFN_{CAG}$ could cleave the $(CTG)_{12}$•$(CAG)_{12}$ containing fragment, which demonstrates that the synthesized ZFNs, $ZFN_{CTG}$, $ZFN_{CAG}$ possess the normal enzyme function in vivo.

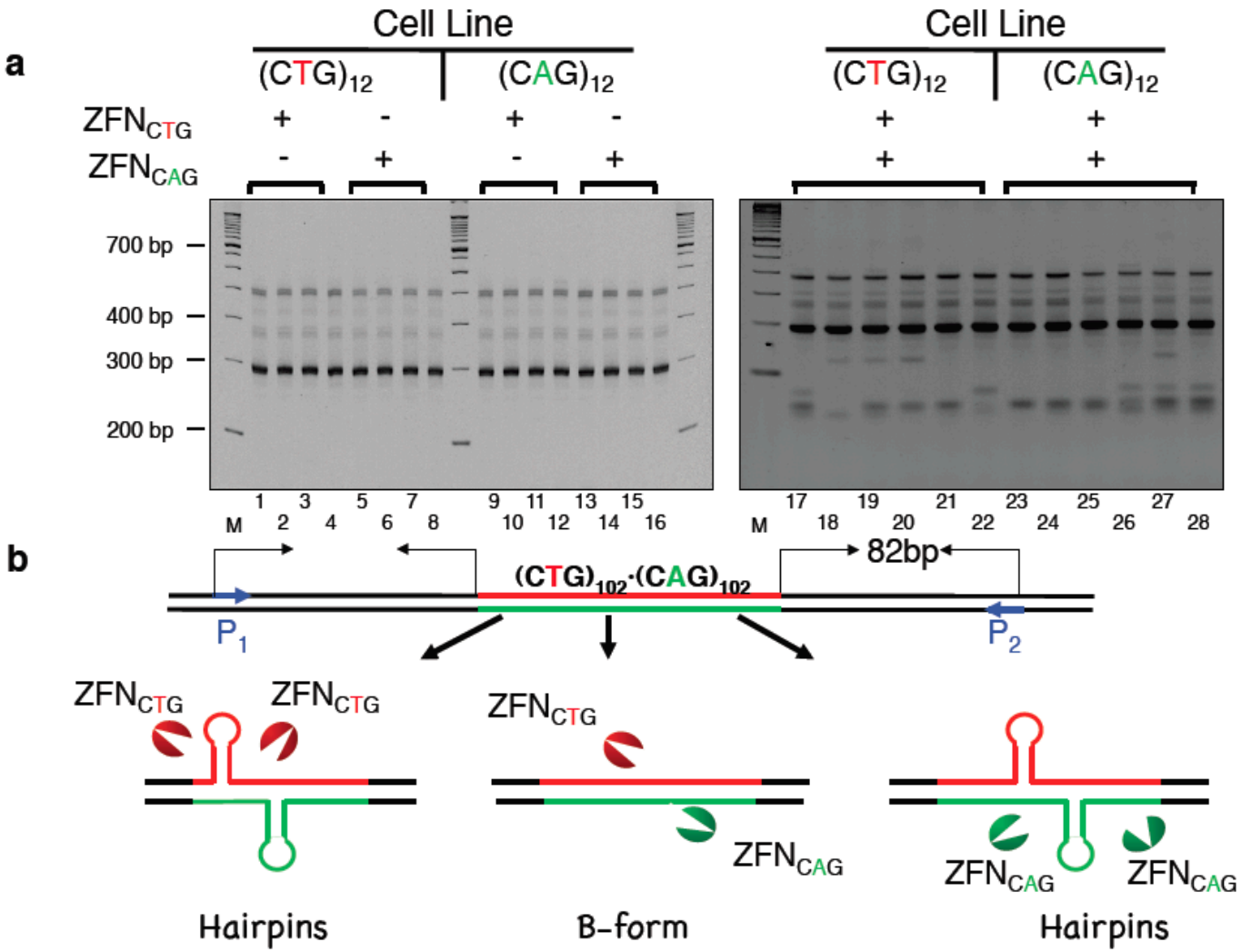
**$(CTG)_{12}$•$(CAG)_{12}$ could not form hairpins capable to be cleaved by $ZFN_{CAG}$ or $ZFN_{CTG}$ individually**

To detect whether $(CTG)_{12}$•$(CAG)_{12}$ repeat tract form hairpins which could be bound and cleaved by ZFNs, I transfected HeLa/c-myc-$(CTG)_{12}$•$(CAG)_{12}$ or HeLa/c-myc-$(CAG)_{12}$•$(CTG)_{12}$ individually with single nuclease, $ZFN_{CAG}$ or $ZFN_{CTG}$. As shown in Figure 35a, neither $ZFN_{CTG}$ (Figure 35a lane 1-8) nor $ZFN_{CAG}$ (Figure 35a lane 9-16) treated cells could generate deleted bands with the size smaller than wild type size indicating that neither $ZFN_{CAG}$ nor $ZFN_{CTG}$ could cut DNA at the $(CTG)_{12}$•$(CAG)_{12}$ repeat tract region. This result demonstrated that $(CTG)_{12}$•$(CAG)_{12}$ repeat tract region was not able to form hairpins suitable for functional ZFN binding and subsequent cleavage. This result also showed that single ZFN treatment did not cause cleavage in $(CAG)_{12}$•$(CTG)_{12}$ TNR or $(CAG)_{12}$•$(CTG)_{12}$ double strand DNA (DS-DNA). Because the predicted minimum binding site for a dimeric ZFN is eight trinucleotide repeats, the stems of the TNR hairpins in these cells should be too short to allow nuclease binding and dimerization, whereas the B-form $(CTG)_{12}$•$(CAG)_{12}$ TNRs could theoretically bind dimeric ZFNs. Indeed, the individual ZFNs did not cleave the $(CTG)_{12}$•$(CAG)_{12}$ cell TNRs, but the coexpressed ZFN pair could digest the TNRs. Presumably because there are fewer possible binding configurations that would accommodate the dimeric nuclease, digestion was less efficient on the shorter repeat tracts.

**Figure 35. In vivo effect of ZFN on $(CTG)_{12}$•$(CAG)_{12}$ cell lines**

**a.** HeLa/myc-$(CAG)_{12}$•$(CTG)_{12}$ or HeLa/myc-$(CTG)_{12}$•$(CAG)_{12}$ cells were treated with mock, $ZFP_{CAG}$ (lane 1-4) or $ZFP_{CTG}$ (lane 9-12) , single $ZFN_{CAG}$ (lane 5-8) or $ZFN_{CTG}$ (lane 13-16) **,** and double ZFN, $ZFN_{CAG}$ plus $ZFN_{CTG}$ (lane 17-28) . Genomic DNA was isolated from treated cell lines, HeLa/myc-$(CAG)_{12}$•$(CTG)_{12}$ or HeLa/myc $(CTG)_{12}$•$(CAG)_{12}$ , and used as template in small pool PCR using $P_1$ and $P_2$. PCR products were resolved on 7.5% polyacrylamide gel. DNA marker is 100 bp DNA ladder.

**b.** Diagram of possible cleavage.


a
Cell Line
(CTG)12
(CAG)12
Cell Line
(CTG)12
(CAG)12
ZFNCTG
ZFNCAG
700 bp
400 bp
300 bp
200 bp
M
b
(CTG)102·(CAG)102
82bp
P1
P2
ZFNCTG
ZFNCTG
ZFNCTG
ZFNCAG
ZFNCAG
ZFNCAG
Hairpins
B-form
Hairpins

**$(CTG)_{102}$•$(CAG)_{102}$ could form hairpins which were cleaved by either $ZFN_{CTG}$ or $ZFN_{CAG}$ individually**

We further applied the $ZFN_{CAG}$ or $ZFN_{CTG}$ in HeLa/c-myc-$(CTG)_{102}$•$(CAG)_{102}$ or HeLa/c-myc-$(CAG)_{102}$•$(CTG)_{102}$ cell lines (Figure 36). Surprisingly, single nuclease, $ZFN_{CTG}$ (Figure 36a lane 9-16, and lane 25-32) or $ZFN_{CAG}$ (Figure 36a lane 37-40, and lane 45-48) cut either HeLa/c-myc-$(CTG)_{102}$•$(CAG)_{102}$ or HeLa/c-myc-$(CAG)_{102}$•$(CTG)_{102}$ cell lines DNA regardless of the orientation of TNR. Since either ZFN would cut both CTG or CAG orientation and I did not know if the template or the nascent DNA was being cut; and I could not say that both arms of the fork form hairpins. However, since the CTG orientation gave almost entirely contractions with long-term culture (Liu et al., 2010), the hairpins cut by $ZFN_{CTG}$ should be on the lagging strand template and the hairpins cut by $ZFN_{CAG}$ should be on the leading strand template.

**$(CTG)_{102}$•$(CAG)_{102}$ could be cleaved by double nuclease, $ZFN_{CAG}$+$ZFN_{CTG}$ together**

Equal amount of plasmid DNA expressing each nuclease, $ZFN_{CAG}$ and $ZFN_{CTG}$, had been used to transfect the above cell lines. As shown in Figure 37a, in both cell lines, HeLa/c-myc-$(CTG)_{102}$•$(CAG)_{102}$ (Figure 37a, lane 9-12), or HeLa/c-myc-$(CAG)_{102}$•$(CTG)_{102}$ (Figure 37a, lane 13-16) cell lines, the wild type band shown in control (Figure 37a, lane 1-8) was eliminated in spPCR products amplified using double nuclease treated cells. Contracted bands are major PCR products in Figure 37a.

**Figure 36. In vivo effect of ZFN on $(CTG)_{102}$•$(CAG)_{102}$ cell lines**

**a.** HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc-$(CTG)_{102}$•$(CAG)_{102}$ cells were treated with mock, $ZFP_{CTG}$ (lane 1-8, 18-24), or $ZFP_{CAG}$ (lane 33-36 or 41-44), $ZFN_{CAG}$ (lane 9-16, 25-32) or $ZFN_{CTG}$ (lane 37-40, 45-48. Genomic DNA was isolated from treated cell lines, HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc $(CTG)_{102}$•$(CAG)_{102}$, was used as template in small pool PCR using $P_1$ and $P_2$. PCR products were resolved on 7.5% polyacrylamide gel. DNA marker is 100 bp ladder.

**b.** Diagram of possible cleavage.

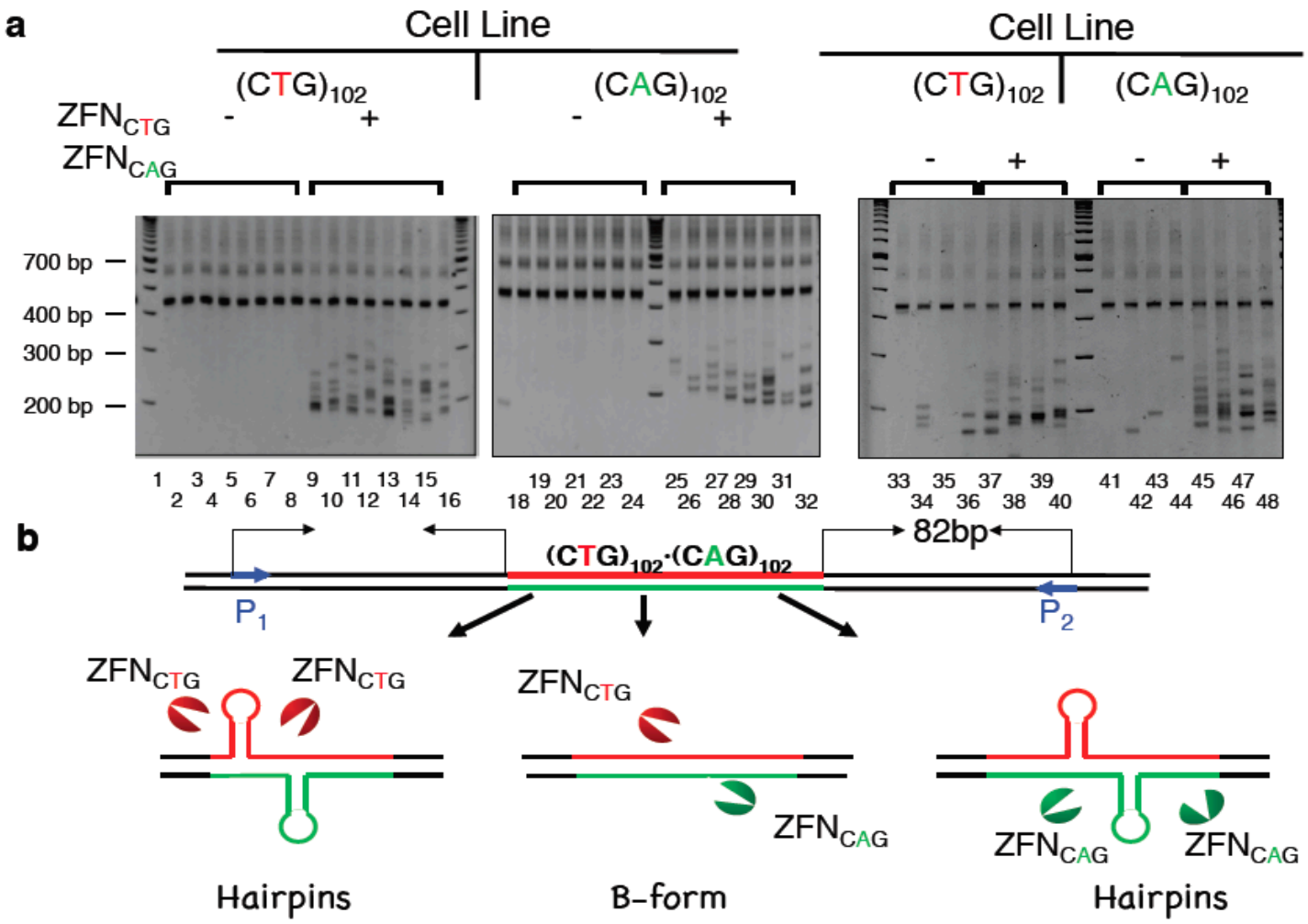
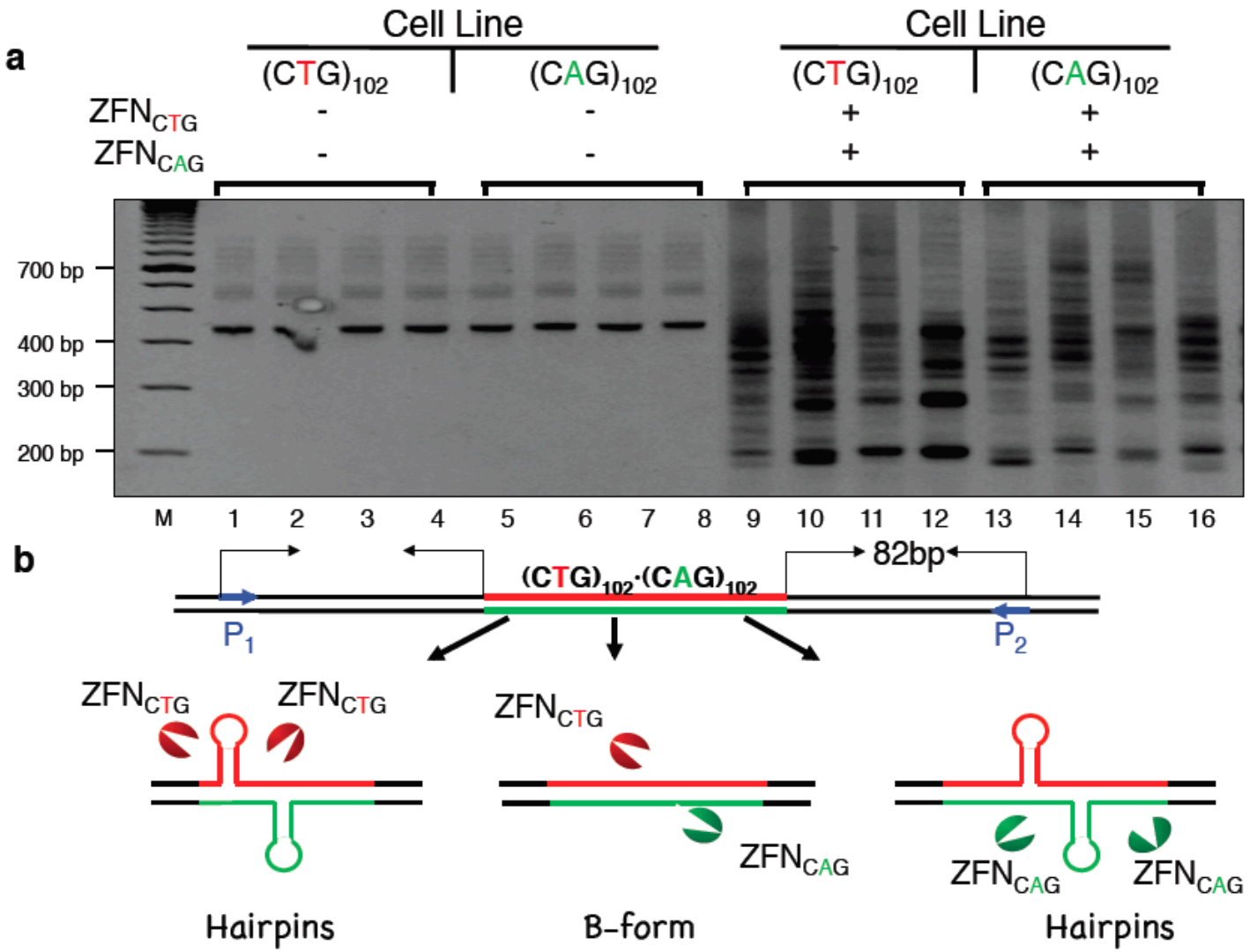

a
Cell Line
(CTG)102
(CAG)102
ZFNCTG
ZFNCAG
-
+
700 bp
400 bp
300 bp
200 bp
1 3 5 7 9 11 13 15
2 4 6 8 10 12 14 16
19 21 23
18 20 22 24
25 27 29 31
26 28 30 32
33 35 37 39
34 36 38 40
41 43 45 47
42 44 46 48
b
(CTG)102·(CAG)102
82bp
P1
P2
ZFNCTG
ZFNCAG
Hairpins
B-form
Hairpins

**Figure 37. In vivo effect of ZFN on $(CTG)_{102}$•$(CAG)_{102}$ cell lines**

**a**. HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc $(CTG)0_{12}$•$(CAG)_{102}$ cells were treated with $ZFP_{CAG}$ plus $ZFP_{CTG}$ (lane 1-8) and double ZFNs, $ZFN_{CAG}$ plus $ZFN_{CTG}$ (lane 9-16). Genomic DNA was isolated from cell lines, HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc-$(CTG)_{102}$•$(CAG)_{102}$, was used as template in small pool PCR using $P_1$ and $P_2$. PCR products were resolved on nondenaturing 7.5% polyacrylamide gel. DNA marker is 100 bp ladder.

**b**. Diagram of possible cleavage.


a
Cell Line
Cell Line
(CTG)102
(CAG)102
(CTG)102
(CAG)102
ZFNCTG
ZFNCAG
700 bp
400 bp
300 bp
200 bp
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
b
(CTG)102·(CAG)102
82bp
P1
P2
ZFNCTG
ZFNCTG
ZFNCTG
ZFNCAG
ZFNCAG
ZFNCAG
Hairpins
B-form
Hairpins

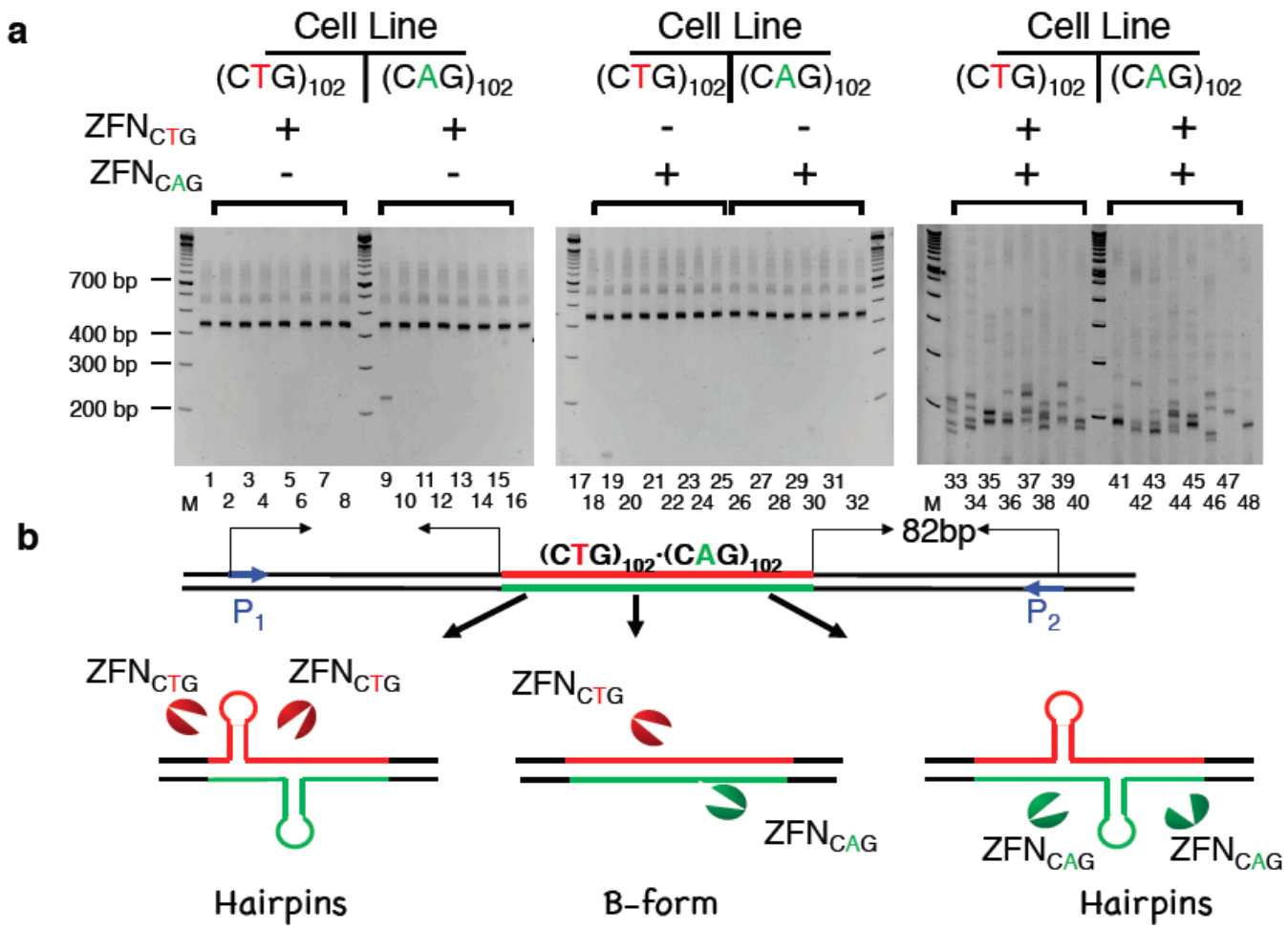
**Hairpin forming is DNA replication associated**

Another experiment was carried out to investigate whether the ZFN effect is replication initiation relevant by inhibition of the DNA replication activity by serum depletion. Cell starvation was performed as described (Yang et al., 2003) with the following modification. In brief, HeLa cells with 50% confluence were grown in DMEM media containing 0.5% serum for three days. Transfection with ZFN plasmid was carried out twice every 72 hr and maintained in DMEM media with 0.5% serum. 72 hr later after the second transfection, genomic DNA was isolated and used as a template for spPCR. Under this condition, there was no cell division found in HeLa cells, suggesting that DNA replication was inhibited. As shown in Figure 38a, there was no significant cleavage of hairpins detected when cells were maintained in serum depleted DMEM media (Figure 38a, lane 1-32), indicating that cleavage of hairpins by single ZFN was associated with DNA replication. However, double ZFN could cleave double strand portion (Figure 38a, lane 33-48) regardless of the lack of the replication-associated hairpins within the TNR repeat tract under the serum starvation.

**Figure 38. Formation of hairpins is replication dependent**

**a**. HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc $(CTG)_{102}$•$(CAG)_{102}$ cells were treated with $ZFN_{CTG}$ (lane 1-8, lane 9-16) or $ZFN_{CAG}$ (lane17-32) and double ZFNs, $ZFN_{CTG}$ plus $ZFN_{CAG}$ (lane 33-48). Cells were treated with ZFNs under serum deprivation condition (0.5% serum v/v). Genomic DNA was isolated from cell lines, HeLa/myc-$(CAG)_{102}$•$(CTG)_{102}$ or HeLa/myc-$(CTG)_{102}$•$(CAG)_{102}$, was used as template in small pool PCR using $P_1$ and $P_2$. PCR products were resolved on native 7.5% polyacrylamide gel. DNA marker is 100 bp ladder.

**b**. Diagram of possible cleavage.


a
Cell Line
(CTG)102
(CAG)102
ZFNCTG
ZFNCAG
700 bp
400 bp
300 bp
200 bp
M
b
(CTG)102·(CAG)102
82bp
P1
P2
ZFNCTG
ZFNCAG
Hairpins
B-form
Hairpins

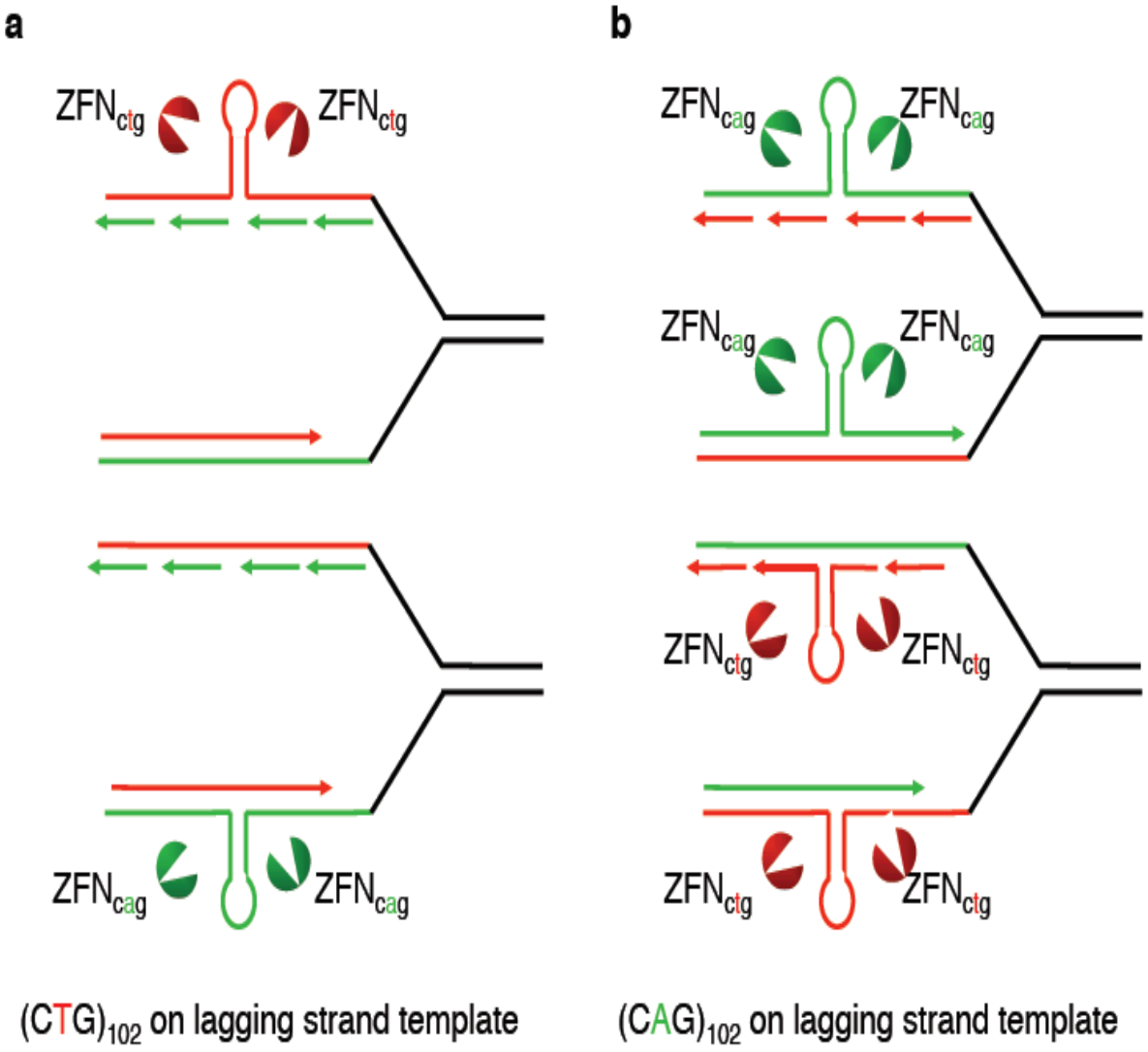
## SUMMARY III

Previous data demonstrated that contraction was a dominant event when $(CTG)_n$ is located on lagging strand template and both expansion and contraction would occurred when $(CAG)_n$ is located on lagging strand template (Mirkin, 2006, 2007). The current data combined with previous results (Liu et al., 2010)tend to support the following model (Figure 39). On one hand, hairpins form when $(CTG)_n$ is on template strands in myc-$(CTG)_n$ orientation (Figure 39a), which potentially leads to contraction. On the other hand, hairpins may form on templates and nascent strands in myc-$(CAG)_n$ orientation, which might lead to both contraction and expansion (Figure 39b).

**Figure 39. Model of forming hairpins during DNA replication**

**a**. In $(CTG)_{102}$ cell line, hairpin formed on lagging strand template or leading strand template.

**b**. In $(CAG)_{102}$ cell line, hairpin formed on both template and nascent strands.


a
ZFNctg
ZFNctg
ZFNcag
ZFNcag
(CTG)102 on lagging strand template
b
ZFNcag
ZFNcag
ZFNcag
ZFNcag
ZFNctg
ZFNctg
ZFNctg
ZFNctg
(CAG)102 on lagging strand template

# DISCUSSION

## 1. DUE is an essential functional element for an active replicator

DUEs have been widely found among different DNA replication origins from prokaryotic cells to eukaryotic cells (Benbow et al., 1992; Dobbs et al., 1994; Huang and Kowalski, 1996; Ishimi and Matsumoto, 1994; Li et al., 1995; Liu et al., 2003; Matsumoto and Ishimi, 1994). In yeast, it has been shown that two essential components are substantial in DNA replication initiation regardless of the size difference among different replication origins. One is the binding site for origin recognition complex (ORC), another is an easily unwound sequence, DUE, facilitating to open the DNA double helix (Dobbs et al., 1994; Wilmes and Bell, 2002).

Saccharomyces cerevisiae replication origins (autonomous replication sequences, ARS) are 100 to 150 bp in length. Each ARS contains two genetically identifiable elements, A and B1, as well as a DUE that usually contains the genetically identifiable B2 element (Celniker et al., 1984; Fangman et al., 1983; Stinchcomb et al., 1979). Replication origin in fission yeast Schizosaccharomyces pombe usually contain two or more regions that are required for full ARS activity, and these regions consist of asymmetric A:T-rich sequences with A residues clustered on one strand and T residues on the other, facilitating the binding of ORC in Saccharomyces cerevisiae (spORC) (Lygerou and Nurse, 1999; Takahashi and Masukata, 2001).

Our laboratory and others have repetitively demonstrated that DUEs are essential elements of mammalian DNA replication origins. Previously, we demonstrated that deletion of DUE inactivated c-myc replication origin and DUE-B binding and origin activity (Ghosh et al., 2006; Liu et al., 2007; Liu et al., 2003)(Figure 6, 7). But, the lost origin activity can be recovered by the replacement by DUEs with certain length of $(ATTCT)_n$•$(AGAAT)_n$ repeat with compatible low helical instability (Liu et al., 2007)(Figure 7). Using SCA10 DUE substitution cell lines, I tested the binding of DUE-B and Cdc45 on c-myc/SCA10 chimerical replicators. The functional SCA10 DUEs bind significantly increased DUE-B and Cdc45 compared with the inactive chimerical c-myc/SCA10 replicator. However, binding of ORC and MCM remains at the same level among inactive and active chimerical replicators (Chowdhury et al., 2010). These results were consistent with the finding that the replication origin activity, Cdc45, RPA and DUE-B binding are induced by the expansion of $(ATTCT)_n$•$(AGAAT)_n$ repeat, the SCA10 DUE, at the SCA10/ATX10 locus (Chowdhury et al., 2010). Taken together, our data favor the hypothesis that DUEs function as a site for essential proteins, including DUE-B, to bind during the pre-replication complex (pre-RC) assembly (Wilmes and Bell, 2002).

If binding of initiation protein is the sole role of DUEs, any DNA sequences capable of attracting DUE-B binding may function as a DUE to recover inactive replication activity. To test this hypothesis, we constructed a HeLa/mycGAL4SubDUE cell line and made a $GAL4^{DBD}$-DUE-B fusion protein. Unfortunately, GAL4 sequence could not functionally replace DUE, and the

tethered binding of DUE-B failed to recover replication activity, indicating another unidentified essential protein or proteins may functionally interact with DUE. It was reported that FBP [far upstream element (FUSE)-binding protein] binds on DUE with DNA sequence specificity, and loss of FBP function led to cell cycle arrest (He et al., 2000). Most recently, a novel essential protein, Treslin (Fu and Walter, 2010; Kumagai et al., 2010), was identified in both frog and human and functions as the Sld3 homolog, similar to DUE-B, although the binding of Treslin on DUE remains unclear. Nonetheless, the non-redundant role of DUE makes DUE the necessity in a functional replication origin in human cells.

**2. Binding of DUE-B on chromatin is ORC dependent**

DUE-B was discovered by yeast one-hybrid assay using DUE sequence from c-myc replicator as a bait (Casper et al., 2005). Based on the finding that DUE-B was an essential DNA replication protein in human cells, it is a reasonable speculation that DUE-B binds the c-myc replicator in a DUE DNA sequence specific. However, work from our laboratory demonstrated that DUE-B binds on not only active c-myc replicator but also active lamin B2 (Ghosh et al., 2006) and active SCA10 replication origins (Chowdhury et al., 2010; Liu et al., 2007), on which we did not find c-myc DUE consensus sequences. Thus, it seems that DUE-B binds on DUE with DNA sequence flexibility although we do not know how DUE-B recognizes DUE structure.

In one set of experiments (Figure 7a), using chimeric myc/SCA10 replicators, we demonstrated that origin activity correlates with the increase of

DUE-B binding, suggesting that increased binding of DUE-B activated replication origin activity (Chowdhury et al., 2010). In another set of experiments, we found a wild type c-myc DUE harbored in Δ1420 or HeLa/Δ1420-GAL4 cell lines could not bind DUE-B (Figure 15e). Coincidently, binding of ORC (Figure 15c), MCM (Figure 15 d), and Cdc45 (Figure 15f) were also not found. However, It is further demonstrated that DUE-B binds to the native DUE upon the tethered binding of Orc2 (Figure 22c,e), suggesting that DUE-B binds to the active replication origin with ORC binding dependency. This finding was further supported by tethered E2F1 (Figure 20c,e) and tethered Cdt1 (Figure 24c,e) experiments in which Orc2 and DUE-B were recruited and origin activity was induced. On the contrary, ORC binding was not detected when tethering Mcm7 and Cdc45 were on the acceptor site, nor was DUE-B. Taken together, DUE-B binding on chromatin correlated with replication origin activity and is ORC dependent. In addition, tethered binding of Cdt1 was able to reversely recruit Orc2, DUE-B, Mcm7 and Cdc45. Based on these observations I hypothesize that the chromatin structure at the inactive mutant c-myc origin plays an essential role to maintain the replication origin activity. Tethered binding of any of the pre-RC components including Orc2 (Takeda et al., 2005), Cdt1(this work), Cdc6 (Takeda et al., 2005), or transcription factors, E2F1 (this work) and CREB (Ghosh et al., 2004), may alter chromatin structure, favoring the binding of other pre-RC or pre-IC components and resulting in restoring replication origin activity.

### 3. Transcription binding factor, E2F1, induced DNA replication origin activity

The facts that replication origins were often mapped in promoter regions, and actively transcribed regions in genome usually replicate earlier than non-transcribed regions suggest a correlation between transcription and DNA replication. Our lab previously showed that tethered binding of $GAL4^{DBD}$-CREB binding to an ectopic inactive c-myc replication origin recovered the mutant replication origin activity in HeLa cells (Ghosh et al., 2004), which was consistent with a previous report (Danis et al., 2004) demonstrating that tethered binding of $GAL4^{DBD}$-VP16 conferred a replication origin activity on a plasmid in Xenopus extract (Danis et al., 2004). However, data from Takeda et al. suggested that tethered binding of VP16 failed to confer a replication origin on plasmid in human cells (Takeda et al., 2005). In the current experiment, we tested the effects of tethered binding of full length of E2F1 transcription factor to a mutant c-myc replication origin. Our results demonstrated that tethered E2F1 modulated the change of chromatin structure by increasing the H4 acetylation significantly on both HeLa/Δ1420-GAL4 and HeLa/GAL4 only cell lines. However, pre-RC assembly and induced origin activity were only detected on the HeLa/Δ1420-GAL4 cell line, again suggesting the left 1 kb c-myc mutant replicator plays an essential role to restore the replication activity. This result is in agreement with our previous CREB data (Ghosh et al., 2004) and results from others (Danis et al., 2004).

It has been repetitively shown that histone acetylation of chromatin favor DNA replication (Burgess et al., 2010; Davie and Hendzel, 1994; Glozak and Seto, 2009; Han et al., 2007; Hasan and Hottiger, 2002; Yang and Freudenreich, 2010). Most recently, Unnikrishnan et al. and others (Aggarwal and Calvi, 2004; Chadha and Blow, 2010; Unnikrishnan et al., 2010) demonstrate that dynamic changes in histone acetylation regulate origins of DNA replication in budding yeast, and this acetylation is required for efficient origin activation during S-phase. In the current experiment, tethered E2F1 restored origin activity, which coincides with the increased histone acetylation, suggesting histone acetylation plays an important role in the induction of replication activity. Recently, it has been shown that chromatin acetylation by HBO1 favored the recruitment of Cdt1, a replication licensing factor (Chadha and Blow, 2010; Iizuka et al., 2009; Miotto and Struhl, 2008, 2010; Wong et al., 2010). In order to distinguish the role of general acetylation of chromatin and Cdt1 associated acetylation, we have initiated work to investigate the role of tethered binding of HBO1 or Tip60 (Sapountzi et al., 2006; Taubert et al., 2004) to Δ1420-GAL4. Since Tip60 is a general acetylase and not associated with Cdt1, if tethered binding of HBO1 but not tethered Tip60 could restore replication activity we will conclude that it is Cdt1 associated acetylation event induced the activation of origin activity. Otherwise, it will indicate that global histone acetylation activity is sufficient to activate the origin activity at the acceptor site.

E2F1 has been linked to replication origin activity since several DNA replication genes including Cdc6, Cdt1, geminin, Orc6, Mcm7 are regulated by

E2F1(Ahlander et al., 2008; Esposito et al., 2009; Karakaidos et al., 2004; Mallik et al., 2008; Ohtani et al., 1996; Yoshida and Inoue, 2004). Overexpression of E2F1 will up-regulate the expression of them and favor DNA replication. In the current experiment, the recovery of a mutant replication activity in HeLa/Δ1420-GAL4 cells, not in HeLa/Δ1420, argued against the possibility that increased expression of replication proteins upregulated by the overexpression of E2F1 conferred the replication origin activity in HeLa/Δ1420-GAL4 cells. Taken together, the origin activity at HeLa/Δ1420-GAL4 cells was not due to the overcome of limitation of DNA replication proteins.

E2F1 associated with DNA replication origins through the interaction with Nbs1 (Maser et al., 2001). E2F1 interacts with Nbs1 at the origin-proximal E2F site in c-myc replication origin (Maser et al., 2001). The interaction is enhanced in active replication origins suggesting that E2F1 binding on replication origin favors its activity. Although we did not test the interaction between Nbs1 and E2F1 in our cell lines, the finding that E2F1 binding on HeLa/Δ1420-GAL4 cell lines restore the replication activity demonstrated the role of E2F1 binding in active replication origin.

There is a well studied E2F1 binding site TTTSSCCGGC located outside 2.4 kb c-myc replicator in endogenous promoter region (Roussel et al., 1994). Using ChIP assay we also were able to detect the binding of E2F1 on endogenous c-myc replication origin (Figure 18 a, b). However, we failed to find the enrichment of E2F1 binding at STS-A, -C sites flanking a predicted E2F1/Ets binding site, TTTSSCCGGC, in HeLa/Δ1420 and HeLa/Δ1420-GAL4 as well as in

2.4 kb replicator. The fact that E2F1 binds on endogenous c-myc replication origin (Figure 18 a, b) but not on 2.4 kb ectopic replicator and mutant replicator Δ1420 (Figure 16 a, b) suggests that the binding activity of E2F1 on the predicted E2F1 binding site has been lost at the ectopic 2.4 kb c-myc and Δ1420. Our previous data showed that deletion of 200 bp DNA fragment including the predicted E2F1 binding site did not affect the replication activity suggesting that there is no actual E2F1 binding at the predicted E2F1 binding site.

It was found that E2F1 is important factor for proper localization of ORC within the chorion gene cluster during embryogenesis in *Drosophila melanogaster* (Cayirlioglu et al., 2001; Royzman et al., 1999). Since this effect did not require E2F1 transcriptional activity, it is clear that E2F1 acts directly at or near origins of replication. In good agreement with this, tethered binding of E2F1 helped to recruit ORC, MCM, DUE-B, and Cdc45 suggesting that the tethered binding of E2F1 was able to re-localize ORC in HeLa/Δ1420-GAL4 cells as well. The fact of no E2F1 binding on WT 2.4 kb c-myc replicator raised a possibility that other transcription factors, i.e., CREB and HBO1, may also facilitate the localization of ORC via change of chromatin structure locally.

**4. Binding of Orc2 and replication origin activity**

Assembly of the six subunits of ORC (ORC1-6) on chromatin is the primary step during DNA replication initiation. Although ORC binding in budding yeast is sequence specific, however, ORC binding on metazoan replication origin lacks DNA sequence specificity (Houchens et al., 2008; Remus et al., 2004;

Vashee et al., 2003; Vashee et al., 2001), suggesting that DNA replication may initiate at any DNA sequences upon ORC binding, which is further supported by the finding that tethered binding of GAL4-Orc2 to GAL4 binding sequence on a plasmid conferred a replication origin activity on plasmid in human cells regardless of DNA sequence specificity (Takeda et al., 2005).

Our previous study and current work demonstrated that replication origin activity was induced upon tethered binding CREB or E2F1, which is correlated with the chromatin structure alteration and Orc2 recruitment. Based on the results from our laboratory and others (Takeda et al., 2005), we speculate that origin activity will be conferred by tethering Orc2 to an inactive chromosomal site with DNA sequence preference in human cells. Indeed, we observed that tethered binding of Orc2 conferred a replication origin in HeLa/Δ1420-GAL4 cell line but not in HeLa/GAL4 cell lines. Partially, our result is consistent with the finding that tethered binding of Orc2 to GAL4 could confer a replication origin activity even on human chromosome. However, we also found tethered binding of Orc2 on HeLa/GAL4 failed to show replication origin activity on human chromosome suggesting that ORC binding on chromosome is not sufficient to confer an origin activity.

The above differences may reflect the different requirements to initiate DNA replication on episomal plasmid and on human chromosomes. In fact, our laboratory and others have demonstrated that a fragment may initiate DNA replication on plasmid but not on a chromosomal context (McWhinney and Leffak, 1990; McWhinney et al., 1995). One explanation will be that human

chromosome normally forms more complicated nucleosomal structure than that on plasmid, and DUEs facilitate DNA replication initiation by lowering the free energy to unwind duplex DNA.

We hypothesize that binding of ORC is essential but not sufficient to activate a replication origin activity. There are several lines of evidence from our experiments that strongly support this hypothesis. First, previous data (Ghosh et al., 2006; Liu et al., 2003) demonstrated that deletion of a fragment including DUE, named Δ5, and another deletion, named Δ7, did not change the ORC binding on c-myc replicators at an ectopic chromosomal site although the replication origin activity was abolished from these two cell lines. Second, although the binding of Orc2 on inactive and active SCA10 origins are comparable, the inactive SCA10 origin were activated after the SCA10 DUE expanded, which lowered the helical stability and attracted elevated binding of DUE-B and Cdc45 (Chowdhury et al., 2010; Liu et al., 2007). Third, the fact that the binding of Orc2 at the ectopic chimeric c-myc/SCA10DUE origins behaved similarly with endogenous SCA10 origins, suggesting that ORC binding is essential but not sufficient to confer an active replication origin unless ORC binds at a site flanking a functional DUE.

**5. Tethered Cdt1 could recruit ORC**

Cdt1 functions as licensing factor recruited by ORC during the assembly of pre-replication complex (Asano et al., 2007; Maiorano et al., 2000; Rialland et al., 2002). It is believed that Cdc6 and Cdt1 are recruited by ORC independently

(Drury and Diffley, 2009; Speck and Stillman, 2007; Tsuyama et al., 2005), which are critical steps to load MCM (Evrin et al., 2009; Remus et al., 2009). A recent study (Takeda et al., 2005) showed that tethered binding of Cdc6 recruited ORC reversely and triggered an origin activity on plasmid in human cells. Here, we tested the idea whether tethered binding of Cdt1 may also confer an origin activity on a human chromosome. As we expected, tethered binding of Cdt1 recovered replication origin activity in HeLa/Δ1420-GAL4 cell line. Surprisingly, binding of Cdt1 also reversely recruited Orc2, as well as, Mcm7, DUE-B, and Cdc45.

In order to maintain genome stability, the replication licensing mechanism must be strictly regulated to guarantee the entire genome duplicates once per cell cycle. One regulatory mechanism is to control the licensing process by targeting Cdt1. Inhibition of Cdt1 activity occurs by multiple mechanisms, including ubiquitin-dependent proteolysis and binding to its potent inhibitor Geminin (Arias and Walter, 2007; Tada et al., 2001; Wohlschlegel et al., 2000). It is believed that the expression of Cdt1 was tightly regulated in human cells to prevent rereplication (De Marco et al., 2009; Kulartz and Knippers, 2004; Liontos et al., 2007; Lovejoy et al., 2006). Overexpression of Cdt1 lead to rereplication in Xenopus egg extract and amplification in *Drosophila melanogaster* and checkpoint activation in human cells (Hall et al., 2008; Jin et al., 2006; Kulartz and Knippers, 2004). In our experiment, exogenous Cdt1 was expressed in HeLa cells and we speculated that the replication origin activity may be due to re-replication. To test this idea the genomic DNA was isolated from Cdt1 transfected

cells and small pool PCR and quantitative PCR were carried out to test the possibility of rereplication, amplification and damage events. However, the spPCR, qPCR, and FACS data suggest that there were no detectable re-replication events occurring. The negative Western blot results with phosphorylated Chk1ser$^{345}$ further supported the idea that overexpression of Cdt1 in our experimental system did not cause DNA damage and activation of the checkpoint (Liu et al., unpublished data).

Based on the current data, we proposed the following model, in which Cdt1 may function as a cofactor of an unknown protein to form a complex, which is essential to facilitate ORC loading on chromatin. Tethered binding of Cdt1 to chromatin favored the recruitment of ORC to chromatin and formed pre-RC, which further triggered firing of the replication origins.

**6. Tethered binding of Mcm7 and DNA replication activity**

During the assembly of pre-RC, six subunits, Mcm2-7, form double hexamer, which is believed to function as a DNA helicase to open double strand DNA, and keep moving with replication fork (Aparicio et al., 1997; Drury and Diffley, 2009; Evrin et al., 2009; Francis et al., 2009; Labib et al., 2000; Liu et al., 2009; Lutzmann and Mechali, 2009; Remus et al., 2009; Tercero et al., 2000).

Binding of MCM is the last step in the assembly of pre-RC. Since tethered Cdt1 could confer replication origin activity we are curious about the virtual function of the earlier steps during the formation of pre-RC. In other word, if the early steps during the formation of pre-RC are solely to prepare for the

recruitment of MCM, the tethered binding of MCM may skip the earlier steps, or like Cdt1 (this work) and Cdc6 (Takeda et al., 2005) reversely recruit ORC and other pre-RC components ahead of MCM binding. To test this idea, I examined the effect of tethered binding of Mcm7. My data demonstrated that the tethered Mcm7 to GAL4 sites flanking a mutant c-myc origin could not recruit Orc2, DUE-B, and Cdc45 suggesting that tethered binding of Mcm7 could not form a functional pre-RC and pre-IC. As a consequence, tethered binding of Mcm7 could not recover DNA replication origin activity in HeLa/Δ1420-GAL4 cells. The fact that the Mcm7-GAL4$^{DBD}$ fusion protein is capable to bind endogenous c-myc origin argues against the possibility that the failure of recovery of origin activity is due to the malfunction of Mcm7-GAL4$^{DBD}$ fusion.

During nascent DNA synthesis, as a helicase, MCM complex moves with replication fork in order to keep unwinding duplex DNA (Bochman and Schwacha, 2009). In our experiment, the GAL4$^{DBD}$ mediated tethering of MCM complex, to some extent, may lock the movement of MCM complex, which may inhibit the replication fork moving regardless of assembly of pre-RC or pre-IC. However, our ChIP data argued that the failure to recover the DNA replication origin activity after tethering of Mcm7 is due to the incapability of pre-RC and pre-IC formation rather than replication fork blockage. Taken together, tethered binding of Mcm7 did not trigger assembly of pre-RC and could not recruit co-factor Cdc45 to form pre-IC. Consequently, tethered binding of Mcm7 was not able to induce replication origin activity.

**7. Tethering of Cdc45 could not induce replication origin activity**

As a co-factor of MCM complex, Cdc45 and GINS form a MCG complex with MCM during DNA replication initiation (Ilves et al., 2010; MacNeill, 2010; Pospiech et al., 2010). In yeast, Cdc45 is believed to be a limiting factor for replication timing, and recycled from early-fired origins to late-fired origins (Aparicio et al., 1999; Aparicio et al., 1997). We designed Cdc45 tethering experiment to check whether the replication origin activity of a mutant c-myc replicator can be recovered by elevated level of Cdc45 and whether the binding of Cdc45 can recruit other earlier bound pre-RC or pre-IC components during DNA replication initiation. My data demonstrated that tethered binding of Cdc45 could not recruit ORC, MCM and DUE-B, and could not confer origin activity in HeLa/Δ1420-GAL4 cell line. The fact that Cdc45-GAL4$^{DBD}$ can be functionally involved in the assembly of pre-IC in endogenous c-myc replication origin indicated that the GAL4$^{DBD}$ fused Cdc45 could function as wild type Cdc45, and the failure of recovery of DNA replication origin activity is not due to the defect of the function of Cdc45-GAL4$^{DBD}$.

Meanwhile, like tethered binding of Mcm7, one might argue that the tethered binding of Cdc45-GAL4$^{DBD}$ would block the movement of Cdc45-GAL4$^{DBD}$ along with the replication fork (Akman and MacNeill, 2009; Aparicio et al., 1997; MacNeill, 2010; Tercero et al., 2000), which abolished the nascent DNA synthesis. Obviously, the Cdc45 ChIP data ruled out this possibility since tethered Cdc45 binding did not recruit Orc2, DUE-B and Mcm7. Thus, there is no pre-RC assembly upon Cdc45-GAL4$^{DBD}$ tethering on GAL4 site.

## 8. DNA sequence binding specificity and cell toxicity of zinc finger nucleases

Zinc finger is a popular protein-DNA recognition motif for transcription factors to bind specific DNA sequence. CTCF, a transcription factor, is a ZFP binding on flanking sequence of DM1 $(CTG)_n$ repeat tract (Cho et al., 2005; Filippova et al., 2001). Our initial intention is to design a tool detecting hairpin formation during DNA replication in vivo. Synthetic ZFN came to our mind due to its sequence specific binding and subsequent cleavage of DNA upon dimerization. In general, ZFP portion determines the binding specificity and nuclease portion provides the cleavage function upon dimerization of cleavage domains.

ZFNs were normally designed based on B-form double strand DNA sequences (Figure 30b), and several lines of evidence have demonstrated their high efficiency in the editing of genomic DNA. However, the target DNA molecules in my experiment are hairpins, non-B DNA (Sinden et al., 2002; Sinden et al., 2007), and imperfect duplex DNA (Figure 30a, c). Thus, we wondered if the ZFNs we designed could bind to $(CTG)_n$ or $(CAG)_n$ formed hairpin efficiently. Meanwhile, due to the high similarity between $(CTG)_n$ repeat tract and $(CAG)_n$ repeat tract, we were concerned that the $ZFN_{CTG}$ or $ZFN_{CAG}$ may not be specific enough to distinguish $(CTG)_n$ and $(CAG)_n$ sequences. Our in vitro experiment indicated that these two concerns are not necessary. First, $ZFN_{CAG}$ only cut $(CAG)_n$ formed hairpins (Liu et al., 2010), and $ZFN_{CTG}$ only cleaved $(CTG)_n$ formed hairpins (Liu et al., 2010). Together, $ZFN_{CTG}$ or $ZFN_{CAG}$

could cleave $(CTG)_n$ and $(CAG)_n$ formed double stranded DNA. The cleavage itself has indicated the binding specificity of DNA binding domain, ZFP, of ZFNs because cleavage events occurred upon the binding of ZFP portion of two ZFN molecules on each strand of a duplex DNA. The formation of a ZFN heterodimer provides the chance for the nuclease domain of ZFN to cut double stranded DNA.

Off target cleavage is one of the major concerns of ZFN technique (Durai et al., 2005; Radecke et al., 2010; Wu et al., 2007). In general, non-sequence specific cleavage causes cell toxicity. In my experiment, I indeed found single ZFN treatment caused 10% increase of cell death compared to with treatment with single ZFP; double ZFN treatment, however, killed about 27% of the cells (Figure 34). Fortunately, one of the spPCR primers, primer 2 (Figure 35 b), is from the sequence of plasmid vector, thus our primer pair only amplified the DNA fragment in which we are interested. All the off-target cleavages will not affect our experiment results.

**9. Hairpins form on both leading strand and lagging strand templates within the $(CTG)_n$•$(CAG)_n$ in human cells**

A previous report (Delagoutte et al., 2008) proposed “template push” model based on the observation of more contraction seen in cultured cells. In their model, they believed that hairpins formed on the leading strand template, which slowed the speed of DNA polymerase passing across the 5’-CAG leading strand template and created a threat to helicase-polymerase coupling. To prevent uncoupling, the TNR template is pushed out and by-passed. Hairpins do not

cause the block, but appear to occur as a consequence of polymerase pass-over (Delagoutte et al., 2008).

Since the significant differences between in vitro and in vivo system, we designed our experiment to explore the formation of hairpins associated with DNA replication in human cells. Surprisingly, our current in vivo results cannot be explained by the template push model (Delagoutte et al., 2008). In our system, since the c-myc replicator was at the 5' of the $(CTG)_n$ or $(CAG)_n$ repeat tract, the replication direction is well defined. In other words, the top strand DNA will work as lagging strand template and the bottom strand will function as leading strand template (Figure 39). Meanwhile, the two ZFNs, $ZFN_{CAG}$ and $ZFN_{CTG}$, will target $(CAG)_n$ formed hairpins and $(CTG)_n$ formed hairpins, respectively. My results demonstrated that a similar hairpin cleavage fragment pattern was observed in two cell lines, HeLa/$(CAG)_n$ and HeLa/$(CAG)_n$, regardless of the ZFNs used, suggesting that hairpins formed in both leading strand nascent/leading strand template and lagging strand nascent/lagging strand template (Figure 39). Specifically, ZFN excision of a hairpin is predicted to generate gapped linear DNA consistent with the results of in vitro ZFN digestion (Figure 33a). During DNA replication, at replication fork, excision of nascent strand hairpins by ZFN would not be detectable after gap filling, whereas excision of template strand hairpins would result in contractions. Therefore, the contractions induced by either $ZFN_{CTG}$ or $ZFN_{CAG}$ treatment of HeLa/$(CTG)_{102}$ or HeLa/$(CAG)_{102}$ cells support the conclusion that both the leading and lagging template strands can form hairpins.

# REFERENCES


Aggarwal, B.D., and Calvi, B.R. (2004). Chromatin regulates origin activity in Drosophila follicle cells. Nature *430*, 372-376.

Ahlander, J., Chen, X.B., and Bosco, G. (2008). The N-terminal domain of the Drosophila retinoblastoma protein Rbf1 interacts with ORC and associates with chromatin in an E2F independent manner. PLoS One *3*, e2831.

Akman, G., and MacNeill, S.A. (2009). MCM-GINS and MCM-MCM interactions in vivo visualised by bimolecular fluorescence complementation in fission yeast. BMC Cell Biol *10*, 12.

Aladjem, M.I., Rodewald, L.W., Kolman, J.L., and Wahl, G.M. (1998). Genetic dissection of a mammalian replicator in the human beta-globin locus. Science *281*, 1005-1009.

Altman, A.L., and Fanning, E. (2001). The Chinese hamster dihydrofolate reductase replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequence elements for activity. Mol Cell Biol *21*, 1098-1110.

Altman, A.L., and Fanning, E. (2004). Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. Mol Cell Biol *24*, 4138-4150.

Aparicio, O.M., Stout, A.M., and Bell, S.P. (1999). Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. Proc Natl Acad Sci U S A *96*, 9130-9135.

Aparicio, O.M., Weinstein, D.M., and Bell, S.P. (1997). Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM proteins and Cdc45p during S phase. Cell *91*, 59-69.

Aparicio, T., Ibarra, A., and Mendez, J. (2006). Cdc45-MCM-GINS, a new power player for DNA replication. Cell Div *1*, 18.

Arias, E.E., and Walter, J.C. (2007). Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. Genes Dev *21*, 497-518.

Asano, T., Makise, M., Takehara, M., and Mizushima, T. (2007). Interaction between ORC and Cdt1p of Saccharomyces cerevisiae. FEMS Yeast Res *7*, 1256-1262.

Balestrini, A., Cosentino, C., Errico, A., Garner, E., and Costanzo, V. (2010). GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. Nat Cell Biol *12*, 484-491.

Ballabeni, A., Melixetian, M., Zamponi, R., Masiero, L., Marinoni, F., and Helin, K. (2004). Human geminin promotes pre-RC formation and DNA replication by stabilizing CDT1 in mitosis. EMBO J *23*, 3122-3132.

Ballabeni, A., Zamponi, R., Caprara, G., Melixetian, M., Bossi, S., Masiero, L., and Helin, K. (2009). Human CDT1 associates with CDC7 and recruits CDC45 to chromatin during S phase. J Biol Chem *284*, 3028-3036.

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

Bell, S.P., Kobayashi, R., and Stillman, B. (1993). Yeast origin recognition complex functions in transcription silencing and DNA replication. Science *262*, 1844-1849.

Benabdallah, B.F., Allard, E., Yao, S., Friedman, G., Gregory, P.D., Eliopoulos, N., Fradette, J., Spees, J.L., Haddad, E., Holmes, M.C.*, et al.* (2010). Targeted gene addition

to human mesenchymal stromal cells as a cell-based plasma-soluble protein delivery platform. Cytotherapy *12*, 394-399.

Benbow, R.M., Zhao, J., and Larson, D.D. (1992). On the nature of origins of DNA replication in eukaryotes. Bioessays *14*, 661-670.

Berberich, S., Trivedi, A., Daniel, D.C., Johnson, E.M., and Leffak, M. (1995). In vitro replication of plasmids containing human c-myc DNA. J Mol Biol *245*, 92-109.

Biamonti, G., Perini, G., Weighardt, F., Riva, S., Giacca, M., Norio, P., Zentilin, L., Diviacco, S., Dimitrova, D., and Falaschi, A. (1992). A human DNA replication origin: localization and transcriptional characterization. Chromosoma *102*, S24-31.

Bieda, M., Xu, X., Singer, M.A., Green, R., and Farnham, P.J. (2006). Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. Genome Res *16*, 595-605.

Bochman, M.L., and Schwacha, A. (2009). The Mcm complex: unwinding the mechanism of a replicative helicase. Microbiol Mol Biol Rev *73*, 652-683.

Bowater, R.P., Jaworski, A., Larson, J.E., Parniewski, P., and Wells, R.D. (1997). Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in Escherichia coli. Nucleic Acids Res *25*, 2861-2868.

Bryant, J.A., Moore, K., and Aves, S.J. (2001). Origins and complexes: the initiation of DNA replication. J Exp Bot *52*, 193-202.

Burgess, R.J., Zhou, H., Han, J., and Zhang, Z. (2010). A role for Gcn5 in replication-coupled nucleosome assembly. Mol Cell *37*, 469-480.

Burhans, W.C., Vassilev, L.T., Wu, J., Sogo, J.M., Nallaseth, F.S., and DePamphilis, M.L. (1991). Emetine allows identification of origins of mammalian DNA replication by

imbalanced DNA synthesis, not through conservative nucleosome segregation. Embo J *10*, 4351-4360.

Carroll, D., Morton, J.J., Beumer, K.J., and Segal, D.J. (2006). Design, construction and in vitro testing of zinc finger nucleases. Nat Protoc *1*, 1329-1341.

Casper, J.M., Kemp, M.G., Ghosh, M., Randall, G.M., Vaillant, A., and Leffak, M. (2005). The c-myc DNA-unwinding element-binding protein modulates the assembly of DNA replication complexes in vitro. J Biol Chem *280*, 13071-13083.

Cayirlioglu, P., Bonnette, P.C., Dickson, M.R., and Duronio, R.J. (2001). Drosophila E2f2 promotes the conversion from genomic DNA replication to gene amplification in ovarian follicle cells. Development *128*, 5085-5098.

Celniker, S.E., Sweder, K., Srienc, F., Bailey, J.E., and Campbell, J.L. (1984). Deletion mutations affecting autonomously replicating sequence ARS1 of Saccharomyces cerevisiae. Mol Cell Biol *4*, 2455-2466.

Chadha, G.S., and Blow, J.J. (2010). Histone acetylation by HBO1 tightens replication licensing. Mol Cell *37*, 5-6.

Chastain, P.D., 2nd, Eichler, E.E., Kang, S., Nelson, D.L., Levene, S.D., and Sinden, R.R. (1995). Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. Biochemistry *34*, 16125-16131.

Chen, S., de Vries, M.A., and Bell, S.P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. Genes Dev *21*, 2897-2907.

Cho, D.H., Thienes, C.P., Mahoney, S.E., Analau, E., Filippova, G.N., and Tapscott, S.J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol Cell *20*, 483-489.

Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M., and Stillman, B. (2000). A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. Proc Natl Acad Sci U S A *97*, 1530-1535.

Chowdhury, A., Liu, G., Kemp, M., Chen, X., Katrangi, N., Myers, S., Ghosh, M., Yao, J., Gao, Y., Bubulya, P.*, et al.* (2010). The DNA unwinding element binding protein DUE-B interacts with Cdc45 in preinitiation complex formation. Mol Cell Biol *30*, 1495-1507.

Chuang, R.Y., and Kelly, T.J. (1999). The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. Proc Natl Acad Sci U S A *96*, 2656-2661.

Cleary, J.D., Nichol, K., Wang, Y.H., and Pearson, C.E. (2002). Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat Genet *31*, 37-46.

Clyne, R.K., and Kelly, T.J. (1995). Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe. Embo J *14*, 6348-6357.

Cocker, J.H., Piatti, S., Santocanale, C., Nasmyth, K., and Diffley, J.F. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. Nature *379*, 180-182.

Dai, J., Chuang, R.Y., and Kelly, T.J. (2005). DNA replication origins in the Schizosaccharomyces pombe genome. Proc Natl Acad Sci U S A *102*, 337-342.

Danis, E., Brodolin, K., Menut, S., Maiorano, D., Girard-Reydet, C., and Mechali, M. (2004). Specification of a DNA replication origin by a transcription complex. Nat Cell Biol *6*, 721-730.

Darlow, J.M., and Leach, D.R. (1995). The effects of trinucleotide repeats found in human inherited disorders on palindrome inviability in Escherichia coli suggest hairpin folding preferences in vivo. Genetics *141*, 825-832.

Davie, J.R., and Hendzel, M.J. (1994). Multiple functions of dynamic histone acetylation. J Cell Biochem *55*, 98-105.

De Marco, V., Gillespie, P.J., Li, A., Karantzelis, N., Christodoulou, E., Klompmaker, R., van Gerwen, S., Fish, A., Petoukhov, M.V., Iliou, M.S.*, et al.* (2009). Quaternary structure of the human Cdt1-Geminin complex regulates DNA replication licensing. Proc Natl Acad Sci U S A *106*, 19807-19812.

Delagoutte, E., Goellner, G.M., Guo, J., Baldacci, G., and McMurray, C.T. (2008). Single stranded DNA binding protein in vitro eliminates the orientation dependent impediment to polymerase passage on CAG/CTG repeats. J Biol Chem.

DePamphilis, M.L. (1993). Origins of DNA replication that function in eukaryotic cells. Curr Opin Cell Biol *5*, 434-441.

DePamphilis, M.L., Blow, J.J., Ghosh, S., Saha, T., Noguchi, K., and Vassilev, A. (2006). Regulating the licensing of DNA replication origins in metazoa. Curr Opin Cell Biol *18*, 231-239.

Diffley, J.F. (1994). Eukaryotic DNA replication. Curr Opin Cell Biol *6*, 368-372.

Diffley, J.F., and Cocker, J.H. (1992). Protein-DNA interactions at a yeast replication origin. Nature *357*, 169-172.

Dobbs, D.L., Shaiu, W.L., and Benbow, R.M. (1994). Modular sequence elements associated with origin regions in eukaryotic chromosomal DNA. Nucleic Acids Res *22*, 2479-2489.

Doyon, Y., Choi, V.M., Xia, D.F., Vo, T.D., Gregory, P.D., and Holmes, M.C. (2010). Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nat Methods *7*, 459-460.

Drury, L.S., and Diffley, J.F. (2009). Factors affecting the diversity of DNA replication licensing control in eukaryotes. Curr Biol *19*, 530-535.

Dubey, D.D., Kim, S.M., Todorov, I.T., and Huberman, J.A. (1996). Large, complex modular structure of a fission yeast DNA replication origin. Curr Biol *6*, 467-473.

Dubey, D.D., Srivastava, V.K., Pratihar, A.S., and Yadava, M.P. (2010). High density of weak replication origins in a 75-kb region of chromosome 2 of fission yeast. Genes Cells *15*, 1-12.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., and Chandrasegaran, S. (2005). Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res *33*, 5978-5990.

Esposito, F., Pierantoni, G.M., Battista, S., Melillo, R.M., Scala, S., Chieffi, P., Fedele, M., and Fusco, A. (2009). Interaction between HMGA1 and retinoblastoma protein is required for adipocyte differentiation. J Biol Chem *284*, 25993-26004.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc Natl Acad Sci U S A *106*, 20240-20245.

Falaschi, A., Abdurashidova, G., Sandoval, O., Radulescu, S., Biamonti, G., and Riva, S. (2007). Molecular and structural transactions at human DNA replication origins. Cell Cycle *6*, 1705-1712.

Fangman, W.L., Hice, R.H., and Chlebowicz-Sledziewska, E. (1983). ARS replication during the yeast S phase. Cell *32*, 831-838.

Fieck, A., Wyborski, D.L., and Short, J.M. (1992). Modifications of the E.coli Lac repressor for expression in eukaryotic cells: effects of nuclear signal sequences on protein activity and nuclear accumulation. Nucleic Acids Res *20*, 1785-1791.

Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet *28*, 335-343.

Flemington, E.K., Borras, A.M., Lytle, J.P., and Speck, S.H. (1992). Characterization of the Epstein-Barr virus BZLF1 protein transactivation domain. J Virol *66*, 922-929.

Flemington, E.K., Speck, S.H., and Kaelin, W.G., Jr. (1993). E2F-1-mediated transactivation is inhibited by complex formation with the retinoblastoma susceptibility gene product. Proc Natl Acad Sci U S A *90*, 6914-6918.

Francis, L.I., Randell, J.C., Takara, T.J., Uchima, L., and Bell, S.P. (2009). Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. Genes Dev *23*, 643-654.

Freudenreich, C.H., Stavenhagen, J.B., and Zakian, V.A. (1997). Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. Mol Cell Biol *17*, 2090-2098.

Fu, Y.V., and Walter, J.C. (2010). DNA replication: metazoan Sld3 steps forward. Curr Biol *20*, R515-517.

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat Cell Biol *8*, 358-366.

Gavin, K.A., Hidaka, M., and Stillman, B. (1995). Conserved initiator proteins in eukaryotes. Science *270*, 1667-1671.

Gerbi, S.A., and Bielinsky, A.K. (2002). DNA replication and chromatin. Curr Opin Genet Dev *12*, 243-248.

Ghosh, M., Kemp, M., Liu, G., Ritzi, M., Schepers, A., and Leffak, M. (2006). Differential binding of replication proteins across the human c-myc replicator. Mol Cell Biol *26*, 5270-5283.

Ghosh, M., Liu, G., Randall, G., Bevington, J., and Leffak, M. (2004). Transcription factor binding and induced transcription alter chromosomal c-myc replicator activity. Mol Cell Biol *24*, 10193-10207.

Gilbert, D.M. (2004). In search of the holy replicator. Nat Rev Mol Cell Biol *5*, 848-855.

Glozak, M.A., and Seto, E. (2009). Acetylation/deacetylation modulates the stability of DNA replication licensing factor Cdt1. J Biol Chem *284*, 11446-11453.

Gomez, M., and Antequera, F. (1999). Organization of DNA replication origins in the fission yeast genome. Embo J *18*, 5683-5690.

Hall, J.R., Lee, H.O., Bunker, B.D., Dorn, E.S., Rogers, G.C., Duronio, R.J., and Cook, J.G. (2008). Cdt1 and Cdc6 are destabilized by rereplication-induced DNA damage. J Biol Chem *283*, 25356-25363.

Hamlin, J.L., Dijkwel, P.A., and Vaughn, J.P. (1992). Initiation of replication in the Chinese hamster dihydrofolate reductase domain. Chromosoma *102*, S17-23.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M., and Zhang, Z. (2007). Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science *315*, 653-655.

Hardy, C.F., and Pautz, A. (1996). A novel role for Cdc5p in DNA replication. Mol Cell Biol *16*, 6775-6782.

Hasan, S., and Hottiger, M.O. (2002). Histone acetyl transferases: a role in DNA repair and DNA replication. J Mol Med *80*, 463-474.

He, L., Liu, J., Collins, I., Sanford, S., O'Connell, B., Benham, C.J., and Levens, D. (2000). Loss of FBP function arrests cellular proliferation and extinguishes c-myc expression. EMBO J *19*, 1034-1044.

Hebert, M.L., and Wells, R.D. (2005). Roles of double-strand breaks, nicks, and gaps in stimulating deletions of CTG.CAG repeats by intramolecular DNA repair. J Mol Biol *353*, 961-979.

Heintz, N.H., Milbrandt, J.D., Greisen, K.S., and Hamlin, J.L. (1983). Cloning of the initiation region of a mammalian chromosomal replicon. Nature *302*, 439-441.

Henricksen, L.A., Tom, S., Liu, Y., and Bambara, R.A. (2000). Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. J Biol Chem *275*, 16420-16427.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B.*, et al.* (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol *27*, 851-857.

Hou, C., Chan, N.L., Gu, L., and Li, G.M. (2009). Incision-dependent and error-free repair of (CAG)(n)/(CTG)(n) hairpins in human cell extracts. Nat Struct Mol Biol *16*, 869-875.

Houchens, C.R., Lu, W., Chuang, R.Y., Frattini, M.G., Fuller, A., Simancek, P., and Kelly, T.J. (2008). Multiple mechanisms contribute to Schizosaccharomyces pombe origin recognition complex-DNA interactions. J Biol Chem *283*, 30216-30224.

Huang, R.Y., and Kowalski, D. (1996). Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome. Nucleic Acids Res *24*, 816-823.

Iizuka, M., Takahashi, Y., Mizzen, C.A., Cook, R.G., Fujita, M., Allis, C.D., Frierson, H.F., Jr., Fukusato, T., and Smith, M.M. (2009). Histone acetyltransferase Hbo1: catalytic activity, cellular abundance, and links to primary cancers. Gene *436*, 108-114.

Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Mol Cell *37*, 247-258.

Ishimi, Y., and Matsumoto, K. (1994). Loading of a DNA helicase on the DNA unwinding element in the yeast replication origin: mechanism of DNA replication in a model system. Biochemistry *33*, 2733-2740.

Jacob, F., and Brenner, S. (1963). [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon.]. C R Hebd Seances Acad Sci *256*, 298-300.

Jee, J., Mizuno, T., Kamada, K., Tochio, H., Chiba, Y., Yanagi, K., Yasuda, G., Hiroaki, H., Hanaoka, F., and Shirakawa, M. (2010). Structure and mutagenesis studies of the C-terminal region of licensing factor Cdt1 enable the identification of key residues for binding to replicative helicase Mcm proteins. J Biol Chem *285*, 15931-15940.

Jin, J., Arias, E.E., Chen, J., Harper, J.W., and Walter, J.C. (2006). A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. Mol Cell *23*, 709-721.

Kalejta, R.F., Li, X., Mesner, L.D., Dijkwel, P.A., Lin, H.B., and Hamlin, J.L. (1998). Distal sequences, but not ori-beta/OBR-1, are essential for initiation of DNA replication in the Chinese hamster DHFR origin. Mol Cell *2*, 797-806.

Kamath, S., and Leffak, M. (2001). Multiple sites of replication initiation in the human beta-globin gene locus. Nucleic Acids Res *29*, 809-817.

Kamimura, Y., Tak, Y.S., Sugino, A., and Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae. Embo J *20*, 2097-2107.

Kanemaki, M., and Labib, K. (2006). Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. EMBO J *25*, 1753-1763.

Kaplan, D.L., and Bruck, I. (2010). Methods to study how replication fork helicases unwind DNA. Methods Mol Biol *587*, 127-135.

Karakaidos, P., Taraviras, S., Vassiliou, L.V., Zacharatos, P., Kastrinakis, N.G., Kougiou, D., Kouloukoussa, M., Nishitani, H., Papavassiliou, A.G., Lygerou, Z*., et al.* (2004). Overexpression of the replication licensing regulators hCdt1 and hCdc6 characterizes a subset of non-small-cell lung carcinomas: synergistic effect with mutant p53 on tumor growth and chromosomal instability--evidence of E2F-1 transcriptional control over hCdt1. Am J Pathol *165*, 1351-1365.

Karnani, N., Taylor, C.M., Malhotra, A., and Dutta, A. (2010). Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. Mol Biol Cell *21*, 393-404.

Katada, H., and Komiyama, M. (2009). Artificial restriction DNA cutters as new tools for gene manipulation. Chembiochem *10*, 1279-1288.

Kemp, M., Bae, B., Yu, J.P., Ghosh, M., Leffak, M., and Nair, S.K. (2007). Structure and Function of the c-myc DNA-unwinding Element-binding Protein DUE-B. J Biol Chem *282*, 10441-10448.

Kemp, M.G., Ghosh, M., Liu, G., and Leffak, M. (2005). The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells. Nucleic Acids Res *33*, 325-336.

Kim, S.M., and Huberman, J.A. (1999). Influence of a replication enhancer on the hierarchy of origin efficiencies within a cluster of DNA replication origins. J Mol Biol *288*, 867-882.

Kinoshita, Y., and Johnson, E.M. (2004). Site-specific loading of an MCM protein complex in a DNA replication initiation zone upstream of the c-MYC gene in the HeLa cell cycle. J Biol Chem *279*, 35879-35889.

Kulartz, M., and Knippers, R. (2004). The replicative regulator protein geminin on chromatin in the HeLa cell cycle. J Biol Chem *279*, 41686-41694.

Kumagai, A., Shevchenko, A., and Dunphy, W.G. (2010). Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. Cell *140*, 349-359.

Labib, K., Tercero, J.A., and Diffley, J.F. (2000). Uninterrupted MCM2-7 function required for DNA replication fork progression. Science *288*, 1643-1647.

Lee, J.K., Moon, K.Y., Jiang, Y., and Hurwitz, J. (2001). The Schizosaccharomyces pombe origin recognition complex interacts with multiple AT-rich regions of the replication origin DNA by means of the AT-hook domains of the spOrc4 protein. Proc Natl Acad Sci U S A *98*, 13589-13594.

Leffak, M., and James, C.D. (1989). Opposite replication polarity of the germ line c-myc gene in HeLa cells compared with that of two Burkitt lymphoma cell lines. Mol Cell Biol *9*, 586-593.

Lei, M., and Tye, B.K. (2001). Initiating DNA synthesis: from recruiting to activating the MCM complex. J Cell Sci *114*, 1447-1454.

Lenzmeier, B.A., and Freudenreich, C.H. (2003). Trinucleotide repeat instability: a hairpin curve at the crossroads of replication, recombination, and repair. Cytogenet Genome Res *100*, 7-24.

Li, L., Li, B.L., Hock, M., Wang, E., and Folk, W.R. (1995). Sequences flanking the pentanucleotide T-antigen binding sites in the polyomavirus core origin help determine selectivity of DNA replication. J Virol *69*, 7570-7578.

Liontos, M., Koutsami, M., Sideridou, M., Evangelou, K., Kletsas, D., Levy, B., Kotsinas, A., Nahum, O., Zoumpourlis, V., Kouloukoussa, M.*, et al.* (2007). Deregulated overexpression of hCdt1 and hCdc6 promotes malignant behavior. Cancer Res *67*, 10899-10909.

Liu, G., Bissler, J.J., Sinden, R.R., and Leffak, M. (2007). Unstable spinocerebellar ataxia type 10 (ATTCT*(AGAAT) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells. Mol Cell Biol *27*, 7828-7838.

Liu, G., Chen, X., Bissler, J.J., Sinden, R.R., and Leffak, M. (2010). Replication-dependent instability at (CTG)*(CAG) repeat hairpins in human cells. Nat Chem Biol.

Liu, G., Malott, M., and Leffak, M. (2003). Multiple functional elements comprise a Mammalian chromosomal replicator. Mol Cell Biol *23*, 1832-1842.

Liu, P., Slater, D.M., Lenburg, M., Nevis, K., Cook, J.G., and Vaziri, C. (2009). Replication licensing promotes cyclin D1 expression and G1 progression in untransformed human cells. Cell Cycle *8*, 125-136.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D.*, et al.* (2007). Gene editing in human

stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol *25*, 1298-1306.

Lovejoy, C.A., Lock, K., Yenamandra, A., and Cortez, D. (2006). DDB1 maintains genome integrity through regulation of Cdt1. Mol Cell Biol *26*, 7977-7990.

Lucas, I., Palakodeti, A., Jiang, Y., Young, D.J., Jiang, N., Fernald, A.A., and Le Beau, M.M. (2007). High-throughput mapping of origins of replication in human cells. EMBO Rep *8*, 770-777.

Lutzmann, M., and Mechali, M. (2009). How to load a replicative helicase onto chromatin: a more and more complex matter during evolution. Cell Cycle *8*, 1309-1313.

Lygerou, Z., and Nurse, P. (1999). The fission yeast origin recognition complex is constitutively associated with chromatin and is differentially modified through the cell cycle. J Cell Sci *112 ( Pt 21)*, 3703-3712.

Machida, Y.J., Hamlin, J.L., and Dutta, A. (2005). Right place, right time, and only once: replication initiation in metazoans. Cell *123*, 13-24.

MacMillan, J.C., Morrison, P.J., Nevin, N.C., Shaw, D.J., Harper, P.S., Quarrell, O.W., and Snell, R.G. (1993). Identification of an expanded CAG repeat in the Huntington's disease gene (IT15) in a family reported to have benign hereditary chorea. J Med Genet *30*, 1012-1013.

MacNeill, S.A. (2010). Structure and function of the GINS complex, a key component of the eukaryotic replisome. Biochem J *425*, 489-500.

Maiorano, D., Moreau, J., and Mechali, M. (2000). XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis. Nature *404*, 622-625.

Mallik, I., Davila, M., Tapia, T., Schanen, B., and Chakrabarti, R. (2008). Androgen regulates Cdc6 transcription through interactions between androgen receptor and E2F transcription factor in prostate cancer cells. Biochim Biophys Acta *1783*, 1737-1744.

Malott, M., and Leffak, M. (1999). Activity of the c-myc replicator at an ectopic chromosomal location. Mol Cell Biol *19*, 5685-5695.

Mani, M., Kandavelou, K., Dy, F.J., Durai, S., and Chandrasegaran, S. (2005). Design, engineering, and characterization of zinc finger nucleases. Biochem Biophys Res Commun *335*, 447-457.

Marahrens, Y., and Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science *255*, 817-823.

Maser, R.S., Mirzoeva, O.K., Wells, J., Olivares, H., Williams, B.R., Zinkel, R.A., Farnham, P.J., and Petrini, J.H. (2001). Mre11 complex and DNA replication: linkage to E2F and sites of DNA synthesis. Mol Cell Biol *21*, 6006-6016.

Matsumoto, K., and Ishimi, Y. (1994). Single-stranded-DNA-binding protein-dependent DNA unwinding of the yeast ARS1 region. Mol Cell Biol *14*, 4624-4632.

McWhinney, C., and Leffak, M. (1990). Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-myc gene. Nucleic Acids Res *18*, 1233-1242.

McWhinney, C., Waltz, S.E., and Leffak, M. (1995). Cis-acting effects of sequences within 2.4-kb upstream of the human c-myc gene on autonomous plasmid replication in HeLa cells. DNA Cell Biol *14*, 565-579.

Mesner, L.D., Li, X., Dijkwel, P.A., and Hamlin, J.L. (2003). The dihydrofolate reductase origin of replication does not contain any nonredundant genetic elements required for origin activity. Mol Cell Biol *23*, 804-814.

Miotto, B., and Struhl, K. (2008). HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1. Genes Dev *22*, 2633-2638.

Miotto, B., and Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin. Mol Cell *37*, 57-66.

Miret, J.J., Pessoa-Brandao, L., and Lahue, R.S. (1997). Instability of CAG and CTG trinucleotide repeats in Saccharomyces cerevisiae. Mol Cell Biol *17*, 3382-3387.

Mirkin, S.M. (2006). DNA structures, repeat expansions and human hereditary disorders. Curr Opin Struct Biol *16*, 351-358.

Mirkin, S.M. (2007). Expandable DNA repeats and human disease. Nature *447*, 932-940.

Monckton, D.G., Coolbaugh, M.I., Ashizawa, K.T., Siciliano, M.J., and Caskey, C.T. (1997). Hypermutable myotonic dystrophy CTG repeats in transgenic mice. Nat Genet *15*, 193-196.

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

Nakajima, R., and Masukata, H. (2002). SpSld3 is required for loading and maintenance of SpCdc45 on chromatin in DNA replication in fission yeast. Mol Biol Cell *13*, 1462-1472.

Nasheuer, H.P. (2010). Genome stability and human diseases. Preface. Subcell Biochem *50*, vii-viii.

Nishitani, H., Lygerou, Z., Nishimoto, T., and Nurse, P. (2000). The Cdt1 protein is required to license DNA for replication in fission yeast. Nature *404*, 625-628.

Ohshima, K., and Wells, R.D. (1997). Hairpin formation during DNA synthesis primer realignment in vitro in triplet repeat sequences from human hereditary disease genes. J Biol Chem *272*, 16798-16806.

Ohtani, K., DeGregori, J., Leone, G., Herendeen, D.R., Kelly, T.J., and Nevins, J.R. (1996). Expression of the HsOrc1 gene, a human ORC1 homolog, is regulated by cell proliferation via the E2F transcription factor. Mol Cell Biol *16*, 6977-6984.

Okuno, Y., Satoh, H., Sekiguchi, M., and Masukata, H. (1999). Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. Mol Cell Biol *19*, 6699-6709.

Orlando, S.J., Santiago, Y., Dekelver, R.C., Freyvert, Y., Boydston, E.A., Moehle, E.A., Choi, V.M., Gopalan, S.M., Lou, J.F., Li, J.*, et al.* (2010). Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. Nucleic Acids Res.

Paixao, S., Colaluca, I.N., Cubells, M., Peverali, F.A., Destro, A., Giadrossi, S., Giacca, M., Falaschi, A., Riva, S., and Biamonti, G. (2004). Modular structure of the human lamin B2 replicator. Mol Cell Biol *24*, 2958-2967.

Pan, X.F. (2006). Mechanism of trinucleotide repeats instabilities: the necessities of repeat non-B secondary structure formation and the roles of cellular trans-acting factors. Yi Chuan Xue Bao *33*, 1-11.

Pearson, C.E., Nichol Edamura, K., and Cleary, J.D. (2005). Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet *6*, 729-742.

Pearson, C.E., Wang, Y.H., Griffith, J.D., and Sinden, R.R. (1998). Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)n. (CAG)n repeats from the myotonic dystrophy locus. Nucleic Acids Res *26*, 816-823.

Pelletier, R., Krasilnikova, M.M., Samadashwily, G.M., Lahue, R., and Mirkin, S.M. (2003). Replication and expansion of trinucleotide repeats in yeast. Mol Cell Biol *23*, 1349-1357.

Pospiech, H., Grosse, F., and Pisani, F.M. (2010). The initiation step of eukaryotic DNA replication. Subcell Biochem *50*, 79-104.

Radecke, S., Radecke, F., Cathomen, T., and Schwarz, K. (2010). Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. Mol Ther *18*, 743-753.

Remus, D., Beall, E.L., and Botchan, M.R. (2004). DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. Embo J *23*, 897-907.

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. Cell *139*, 719-730.

Rialland, M., Sola, F., and Santocanale, C. (2002). Essential role of human CDT1 in DNA replication and chromatin licensing. J Cell Sci *115*, 1435-1440.

Richards, R.I., Holman, K., Yu, S., and Sutherland, G.R. (1993). Fragile X syndrome unstable element, p(CCG)n, and other simple tandem repeat sequences are binding sites for specific nuclear proteins. Hum Mol Genet *2*, 1429-1435.

Romanowski, P., Madine, M.A., Rowles, A., Blow, J.J., and Laskey, R.A. (1996). The Xenopus origin recognition complex is essential for DNA replication and MCM binding to chromatin. Curr Biol *6*, 1416-1425.

Romero, J., and Lee, H. (2008). Asymmetric bidirectional replication at the human DBF4 origin. Nat Struct Mol Biol.

Roussel, M.F., Davis, J.N., Cleveland, J.L., Ghysdael, J., and Hiebert, S.W. (1994). Dual control of myc expression through a single DNA binding site targeted by ets family proteins and E2F-1. Oncogene *9*, 405-415.

Rowley, A., Cocker, J.H., Harwood, J., and Diffley, J.F. (1995). Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. Embo J *14*, 2631-2641.

Royzman, I., Austin, R.J., Bosco, G., Bell, S.P., and Orr-Weaver, T.L. (1999). ORC localization in Drosophila follicle cells and the effects of mutations in dE2F and dDP. Genes Dev *13*, 827-840.

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. (2010). ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res *38 Suppl*, W462-468.

Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. (2007). Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Res *35*, W599-605.

Sansam, C.L., Cruz, N.M., Danielian, P.S., Amsterdam, A., Lau, M.L., Hopkins, N., and Lees, J.A. (2010). A vertebrate gene, ticrr, is an essential checkpoint and replication regulator. Genes Dev *24*, 183-194.

Santocanale, C., and Diffley, J.F. (1996). ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. Embo J *15*, 6671-6679.

Sapountzi, V., Logan, I.R., and Robson, C.N. (2006). Cellular functions of TIP60. Int J Biochem Cell Biol *38*, 1496-1509.

Saxena, S., and Dutta, A. (2005). Geminin-Cdt1 balance is critical for genetic stability. Mutat Res *569*, 111-121.

Sinden, R.R., Potaman, V.N., Oussatcheva, E.A., Pearson, C.E., Lyubchenko, Y.L., and Shlyakhtenko, L.S. (2002). Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. J Biosci *27*, 53-65.

Sinden, R.R., Pytlos-Sinden, M.J., and Potaman, V.N. (2007). Slipped strand DNA structures. Front Biosci *12*, 4788-4799.

Speck, C., and Stillman, B. (2007). Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. J Biol Chem *282*, 11705-11714.

Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979). Isolation and characterisation of a yeast chromosomal replicator. Nature *282*, 39-43.

Tada, S., Li, A., Maiorano, D., Mechali, M., and Blow, J.J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. Nat Cell Biol *3*, 107-113.

Takahara, K., Bong, M., Brevard, R., Eddy, R.L., Haley, L.L., Sait, S.J., Shows, T.B., Hoffman, G.G., and Greenspan, D.S. (1996). Mouse and human homologues of the yeast origin of replication recognition complex subunit ORC2 and chromosomal localization of the cognate human gene ORC2L. Genomics *31*, 119-122.

Takahashi, T., and Masukata, H. (2001). Interaction of fission yeast ORC with essential adenine/thymine stretches in replication origins. Genes Cells *6*, 837-849.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. Genes Dev *17*, 1153-1165.

Takeda, D.Y., Shibata, Y., Parvin, J.D., and Dutta, A. (2005). Recruitment of ORC or CDC6 to DNA is sufficient to create an artificial origin of replication in mammalian cells. Genes Dev *19*, 2827-2836.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., and Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. Nature *445*, 328-332.

Taubert, S., Gorrini, C., Frank, S.R., Parisi, T., Fuchs, M., Chan, H.M., Livingston, D.M., and Amati, B. (2004). E2F-dependent histone acetylation and recruitment of the Tip60 acetyltransferase complex to chromatin in late G1. Mol Cell Biol *24*, 4546-4556.

Tercero, J.A., Labib, K., and Diffley, J.F. (2000). DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. EMBO J *19*, 2082-2093.

Trivedi, A., Waltz, S.E., Kamath, S., and Leffak, M. (1998). Multiple initiations in the c-myc replication origin independent of chromosomal location. DNA Cell Biol *17*, 885-896.

Tsuyama, T., Tada, S., Watanabe, S., Seki, M., and Enomoto, T. (2005). Licensing for DNA replication requires a strict sequential assembly of Cdc6 and Cdt1 onto chromatin in Xenopus egg extracts. Nucleic Acids Res *33*, 765-775.

Unnikrishnan, A., Gafken, P.R., and Tsukiyama, T. (2010). Dynamic changes in histone acetylation regulate origins of DNA replication. Nat Struct Mol Biol *17*, 430-437.

van den Broek, W.J., Wansink, D.G., and Wieringa, B. (2007). Somatic CTG*CAG repeat instability in a mouse model for myotonic dystrophy type 1 is associated with changes in cell nuclearity and DNA ploidy. BMC Mol Biol *8*, 61.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T.J., and Walter, J.C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes Dev *17*, 1894-1908.

Vashee, S., Simancek, P., Challberg, M.D., and Kelly, T.J. (2001). Assembly of the human origin recognition complex. J Biol Chem *276*, 26666-26673.

Waltz, S.E., Trivedi, A.A., and Leffak, M. (1996). DNA replication initiates non-randomly at multiple sites near the c-myc gene in HeLa cells. Nucleic Acids Res *24*, 1887-1894.

Wang, L., Lin, C.M., Brooks, S., Cimbora, D., Groudine, M., and Aladjem, M.I. (2004). The human beta-globin replication initiation region consists of two modular independent replicators. Mol Cell Biol *24*, 3373-3386.

Weinmann, A.S., Bartley, S.M., Zhang, T., Zhang, M.Q., and Farnham, P.J. (2001). Use of chromatin immunoprecipitation to clone novel E2F target promoters. Mol Cell Biol *21*, 6820-6832.

Wells, R.D., Dere, R., Hebert, M.L., Napierala, M., and Son, L.S. (2005). Advances in mechanisms of genetic instability related to hereditary neurological diseases. Nucleic Acids Res *33*, 3785-3798.

Wilmes, G.M., and Bell, S.P. (2002). The B2 element of the Saccharomyces cerevisiae ARS1 origin of replication requires specific sequences to facilitate pre-RC formation. Proc Natl Acad Sci U S A *99*, 101-106.

Wohlschlegel, J.A., Dwyer, B.T., Dhar, S.K., Cvetic, C., Walter, J.C., and Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science *290*, 2309-2312.

Wong, P.G., Glozak, M.A., Cao, T.V., Vaziri, C., Seto, E., and Alexandrow, M. (2010). Chromatin unfolding by Cdt1 regulates MCM loading via opposing functions of HBO1 and HDAC11-geminin. Cell Cycle *9*, 4351-4363.

Wright, D.A., Thibodeau-Beganny, S., Sander, J.D., Winfrey, R.J., Hirsh, A.S., Eichtinger, M., Fu, F., Porteus, M.H., Dobbs, D., Voytas, D.F.*, et al.* (2006). Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. Nat Protoc *1*, 1637-1652.

Wu, J., Kandavelou, K., and Chandrasegaran, S. (2007). Custom-designed zinc finger nucleases: what is next? Cell Mol Life Sci *64*, 2933-2944.

Yabuuchi, H., Yamada, Y., Uchida, T., Sunathvanichkul, T., Nakagawa, T., and Masukata, H. (2006). Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. Embo J *25*, 4663-4674.

Yamada, Y., Nakagawa, T., and Masukata, H. (2004). A novel intermediate in initiation complex assembly for fission yeast DNA replication. Mol Biol Cell *15*, 3740-3750.

Yang, J.H., and Freudenreich, C.H. (2010). The Rtt109 histone acetyltransferase facilitates error-free replication to prevent CAG/CTG repeat contractions. DNA Repair (Amst) *9*, 414-420.

Yang, Z., Lau, R., Marcadier, J.L., Chitayat, D., and Pearson, C.E. (2003). Replication inhibitors modulate instability of an expanded trinucleotide repeat at the myotonic dystrophy type 1 disease locus in human cells. Am J Hum Genet *73*, 1092-1105.

Yoshida, K., and Inoue, I. (2004). Regulation of Geminin and Cdt1 expression by E2F transcription factors. Oncogene *23*, 3802-3812.

Zhu, J., Brun, C., Kurooka, H., Yanagida, M., and Huberman, J.A. (1992). Identification and characterization of a complex chromosomal replication origin in Schizosaccharomyces pombe. Chromosoma *102*, S7-16.