Characterisation of the avian TopBP1 protein and its functions

MELITI SKOUTERI

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF SUSSEX

APRIL 2017

"There is nothing over which a free man ponders less than death; his wisdom is, to meditate not on death but on life."

DECLARATION

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:.....

II

UNIVERSITY OF SUSSEX

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

MELITI SKOUTERI

Characterisation of the avian TopBP1 protein and its functions

One of the proteins that lie at the heart of the DNA Damage Response (DDR) is Topoisomerase II-binding protein I (TopBP1). TopBP1 was initially identified and has been extensively studied in the yeast model organisms. However, the lack of readily available tools, including genetically defined mutant cell lines, has rendered the characterisation of TopBP1 in higher eukaryotes more challenging.

Sequence information obtained from the characterisation of the *gallus gallus* TopBP1 mRNA revealed a different splicing pattern at the 5'end to the one reported in the Genome Browser. Our assembled TopBP1 mRNA sequence containing a novel open reading frame (ORF) enabled the creation of a conditional knockout cell line of TopBP1 in DT40, which has been impossible with the use of the annotated cDNA sequence. Thus the avian TopBP1 ORF identified herein contained the necessary function(s) to sustain viability of DT40 cells in the absence of the endogenous protein. Additionally, the establishment of an isogenic set of stable cell lines from the chicken B cell line DT40 by targeted deletion of the *TopBP1* alleles revealed a gene dosage-dependent reduction of the TopBP1 protein levels and functions. This work establishes a novel gene-dosage system that can be used for the knock in of point mutations within the endogenous *TopBP1* locus. Using this system, a novel characterisation of knock-in point mutants of the ATR Activation Domain (AAD) of TopBP1 was carried out, providing *in vivo* evidence of its DDR function(s). Finally, a stably integrated overexpression system (SIOS) capable of producing increased amounts of a protein of interest has been established in DT40 cells. SIOS represents an easy to use versatile system for various experimental purposes in the field of DT40.

The work presented in this thesis represents a novel characterisation of the avian TopBP1 mRNA and the TopBP1 protein and its functions. This is crucial to gain insight into the mechanistics of the DDR network and the genetic instability characterising cancer development.

III

Acknowledgments

I would like to thank Professor Tony Carr for giving me the opportunity to pursue research in his laboratory as well as Dr Helfrid Hochegger for his interest in my project.

I would also like to express my gratitude to my colleagues. A big thank you to Carr group members, past and present!

Above all I would like to thank my family for their continuous support and unconditional love and for believing in me when my own faith waivered. I wouldn't have made it without you.

Many thanks to Professor Simon Morley and Dr. Felicity Watts for their help and support during the past four years.

Special thanks to Kostas Nestoras for his guidance and support;
You have been a great mentor.

IV

TABLE OF CONTENTS

V

Figures

Chapter 1

Introduction

1.1 The cell division cycle; from its discovery to its understanding

1.1.1 Cell division cycle; definition and a historical perspective

The cell division cycle is a complex series of events that involve the duplication of the genetic material and other cellular compartments of a single parental cell and the subsequent division and formation of two daughter cells. It is the process by which a multicellular organism arises from a single fertilized oocyte as well as the process by which stem cells in our bodies acquire differentiation fates to replace mature tissue like hair, skin and blood cells when these become lost, damaged or scarce (Alberts 2002, Dalton 2015). It is thus central to the understanding of multicellular life.

The concepts of the "cell" and the "cell theory" have been proposed as early as the 18th and 19th centuries respectively but it was not until the improvement of microscopy at the turn of the 20th century that enabled scientists to study the cytology of cell division (Nurse, Masui et al. 1998, Alberts 2002). In particular, the pioneer developmental biologist Theodor Boveri (1862-1915) introduced the concepts of cell division and for the first time identified the chromosomes as the carriers of the heritable material (Boveri, 1902). Not surprisingly, with his elegant experiments and groundbreaking concepts Boveri has set the ground for the cell cycle research of the new century (Balmain 2001).

Half a century later, technological advances have made it possible for scientists to go beyond the mere observations of the cellular division and the cell cycle. Early in the 1950's the structure of the DNA had been proposed by James Watson and Francis Crick in one of the most famous papers of all time (Watson and Crick 1953). A few years later and in accordance with the Watson-Crick model, Meselson and Stahl shed light on the molecular mechanism of DNA synthesis; within a single cell cycle the genetic material doubles and each daughter cell receives one copy composed of a parental strand (subunit) and a newly synthesized strand (Meselson and Stahl 1958).

The unprecedented breakthrough in unravelling the molecular basis of the cell cycle, however, came in the late 1960's when three scientist, two geneticists and a biochemist, although working from different angles and using different model systems, made monumental contribution into understanding the regulation of the cell cycle in eukaryotic cells. Working with *S. pombe* and *S. cerevisiae*, respectively, Paul Nurse and Lee Hartwell created mutants that deregulated or halted the progression through the cell

cycle and by cloning the genes they identified a network of proteins responsible for cell cycle control. Most importantly, the first members of the cyclin dependent kinase (CDKs) family of enzymes, namely spCdc2 and scCdc28, were identified by Nurse and Hartwell respectively, in genetic studies looking for *cdc* (cell division cycle) mutants, defective in cell cycle progression (Nurse 1997). Future studies by numerous researchers have discovered human homologues of these genes by complementation analysis of yeast mutants (Malumbres and Barbacid 2005). Tim Hunt on the other hand used sea urchin embryos and biochemistry to identify yet another key player, the *cyclin*. (Evans, Rosenthal et al. 1983, Hartwell and Weinert 1989, Nurse 1997, Nurse, Masui et al. 1998). Soon after, the biochemical connection between the two families of enzymes was established and the regulatory circuit of the cell cycle has just started to be delineated. For their pioneering research, Paul Nurse, Lee Hartwell and Tim Hunt were awarded the 2001 Nobel Prize in Physiology and Medicine.

1.1.2 The eukaryotic cell cycle and its regulation

The cell cycle of eukaryotic cells (**Figure 1.1**) is comprised of two main phases, synthesis (S) and mitosis (M). The S phase is when the genetic material gets duplicated through the process of DNA replication. Mitosis is the nuclear division, which involves the segregation of sister chromatids into separate nuclei and is followed by cytokinesis, the cellular division and the formation of two daughter cells. Cells spend only 5% of their life in the M phase, with the rest of their cell cycle spent in interphase, the period between two sequential mitoses. The main event of the interphase is the duplication of the DNA during S phase and the two key phases of the eukaryotic cell cycle, S and M, are separated by two gap phases, G1 and G2. During G1 the cell grows in size and prepares for DNA replication and during G2 the cell continues to grow and prepares for the onset of mitosis in human cells (Cooper, 2000). At the onset of the G1 phase there is a specific point known as the “restriction point” and once the cell is past this point for example in response to mitogen stimulation, it becomes committed to the next cell cycle. If however it is favourable for the cell to remain non dividing, it enters the G0 phase which is a state of quiescence and can be reversible or irreversible (Pardee 1974).

Passage from one stage to the next is a highly regulated process and the mechanisms of regulation are highly conserved across species, from yeast to higher eukaryotes. Yeast has only one CDK responsible for both S and M phase entry, Cdc2 in

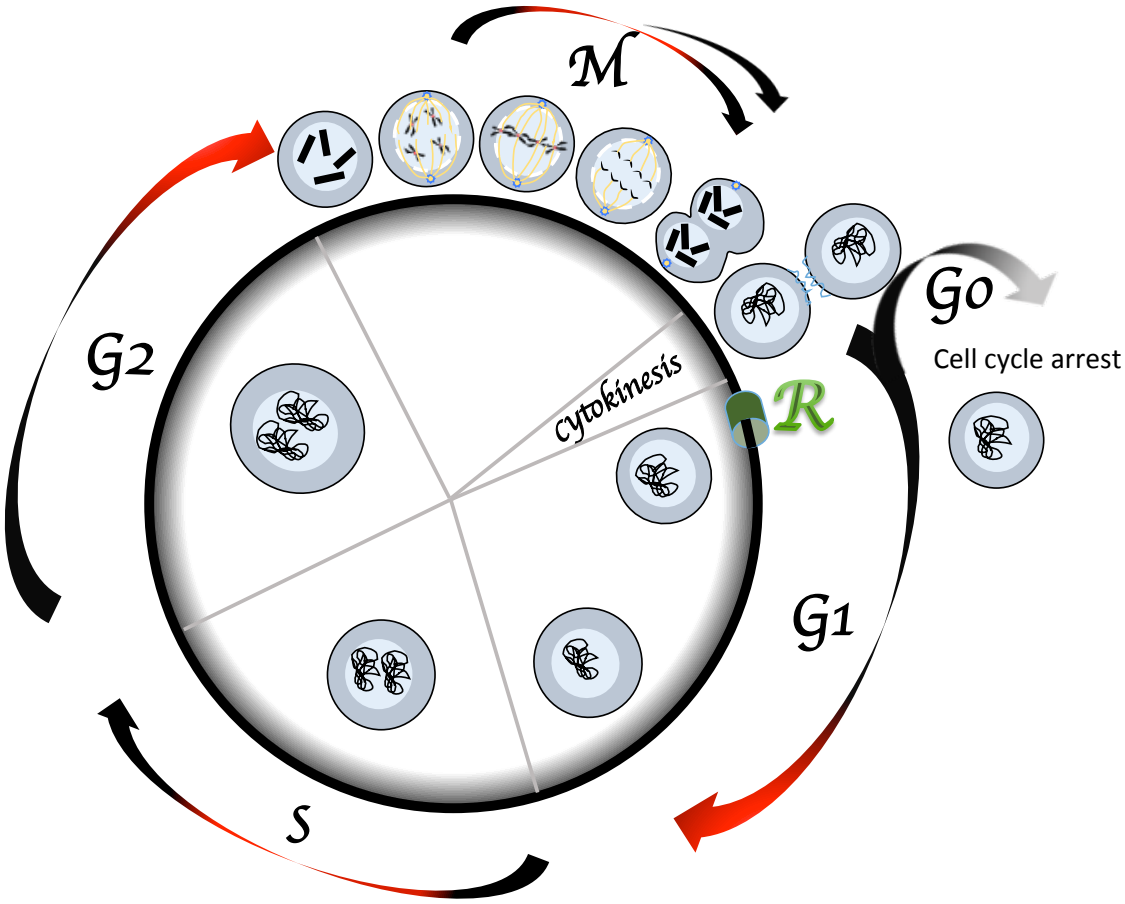


Figure 1.1: Diagrammatic representation of the standard eukaryotic cell cycle. It consists of four phases: G1, S, G2, and M phases. G1, S and G2 are collectively referred to as Interphase. There are three main phase transitions in the cell cycle: G1/S to initiate DNA replication, G2/M to enter mitosis, and metaphase/anaphase to exit mitosis (not shown). R in green is the restriction point at which the cell becomes committed to the cell cycle. Red coloured segments in the arrows represent the checkpoints, namely the G1/S, intra-S, G2/M and S/M, which are discussed in later sections.

S. pombe and Cdc28 in *S. cerevisiae* (Bueno, Richardson et al. 1991, Forsburg and Nurse 1991, Hayles, Fisher et al. 1994, Mondesert, McGowan et al. 1996). At the heart of the human cell cycle regulation machinery are also the CDKs, which encompass a family of 20 members in human cells. Of these, CDK1, 2, 3, 4, 6 and 7 are serine/threonine kinases functioning in cell cycle regulation. For full enzymatic activity they all (except CDK7) require the provision of additional peptide sequences through association with cyclins. CDK-cyclin complexes phosphorylate specific protein targets on K/R-S/T-P-X-K/R consensus sites in a temporally specific manner and this is what allows the proper progression through the cell cycle (Masumoto, Muramatsu et al. 2002, Malumbres and Barbacid 2005, Malumbres 2014). The protein levels of CDKs remain fairly constant throughout the division cycle but their respective activities oscillate cyclically (Hochegger, Takeda et al. 2008). CDK activity is largely dependent on its association with cyclins and biochemical experiments have shown that binding of cyclins to CDK results to a 40,000 fold increase to its kinase activity (Lees 1995).

Additionally, regulation of CDK activity occurs via phosphorylation/dephosphorylation and binding to CDK inhibitors (CKIs) (Nurse 1997). Phosphorylation of the tyrosine 15 (Y15) and additionally, only in mammals, of the adjacent threonine 14 (T14) residues within the active site inhibits CDK activity. For example, Wee1 and Myt1, whose activity remains high throughout interphase, phosphorylate $CDK1^{Cdc2}$ thus preventing premature entry into mitosis. At the G2/M transition however, this inhibitory phosphorylation is removed by the Cdc25 (A/B/C) phosphatase, favouring an active $CDK1^{Cdc2}$ and entry into mitosis (Russell and Nurse 1986, Russell and Nurse 1987, Lees 1995, Malumbres and Barbacid 2005, Calonge and O'Connell 2008). Cdc25 is activated by phosphorylation by the Polo kinase as well as by the CDK1-cyclin B complex (also called the mitosis promoting factor; MPF) itself, which also phosphorylates and inhibits Wee1. This creates a positive feedback loop that quickly leads to the activation of all M-CDK complexes in the cell (Alberts 2002, Toyoshima-Morimoto, Taniguchi et al. 2002). Finally, full activation of the CDK catalytic subunit requires CAK-mediated phosphorylation of the conserved T160 residue within the CDK activation loop (Lees 1995).

Inactivation of CDKs is also promoted through their binding to CKIs. CKIs bind CDK-cyclin complexes and inhibit their activity. One of the most potent inhibitors is $p21^{Cip1}$ which inhibits the complexes of both CDK1 and CDK2 by either direct binding or by preventing CAK-mediated phosphorylation of T160 (Lees 1995).

The cell cycle regulation machinery described above not only controls the ordered sequence of events within the cell cycle but also cross talks to a network of checkpoint factors which monitor the cell for the presence of incomplete DNA synthesis or unrepaired DNA damage. One of the end receivers of this signalling mechanism is of course the cell cycle machinery again; in the event of damage to the genetic material, cell cycle progression is halted to allow cells time to overcome the damage. These mechanisms contribute to the successful completion of the two central events of the cell cycle, DNA replication (S) and chromosome segregation (M), and ensure fidelity of genetic transmission to future generations.

1.1.3 DNA Replication

Replication initiation

A fundamental aspect of genome stability is the faithful replication of DNA prior to each cell division. Chromosome duplication in eukaryotes occurs during the S phase of the cell cycle and is initiated by multiple origins. Chromosome duplication initiated at origins is mediated by a number of essential and non-essential proteins that collaborate in a spatially and temporally regulated manner to perform the basic steps of DNA replication: pre-replicative complex (pre-RC) formation, maturation and activation of the pre-RC to form the pre-initiation complex (pre-IC) and progression of the fully functional replisome, known as the replisome progression complex (RPC) (Bryant and Aves 2011, Wu, Liu et al. 2014) (**Figure 1.2**).

Pre-RC formation

The first step of replication initiation is the binding of the 6-subunit origin recognition complex (ORC, Orc1-6) to those origins that will fire during the S phase, thus marking them as active origins.

During the G1 phase cell division control protein 6 (Cdc6) and chromatin licensing and DNA replication factor 1 (Cdt1), associate with the origin bound ORC independently of each other (Maiorano, Moreau et al. 2000). They are both essential for origin licensing (Coleman, Carpenter et al. 1996, Rialland, Sola et al. 2002). Cdc6 is an ATPase and is thought to bind ORC (Speck, Chen et al. 2005), whereas Cdt1 on the other hand is a coiled-coil domain protein that associates with the C terminus of Cdc6 and together they cooperatively promote the recruitment and subsequent loading of the

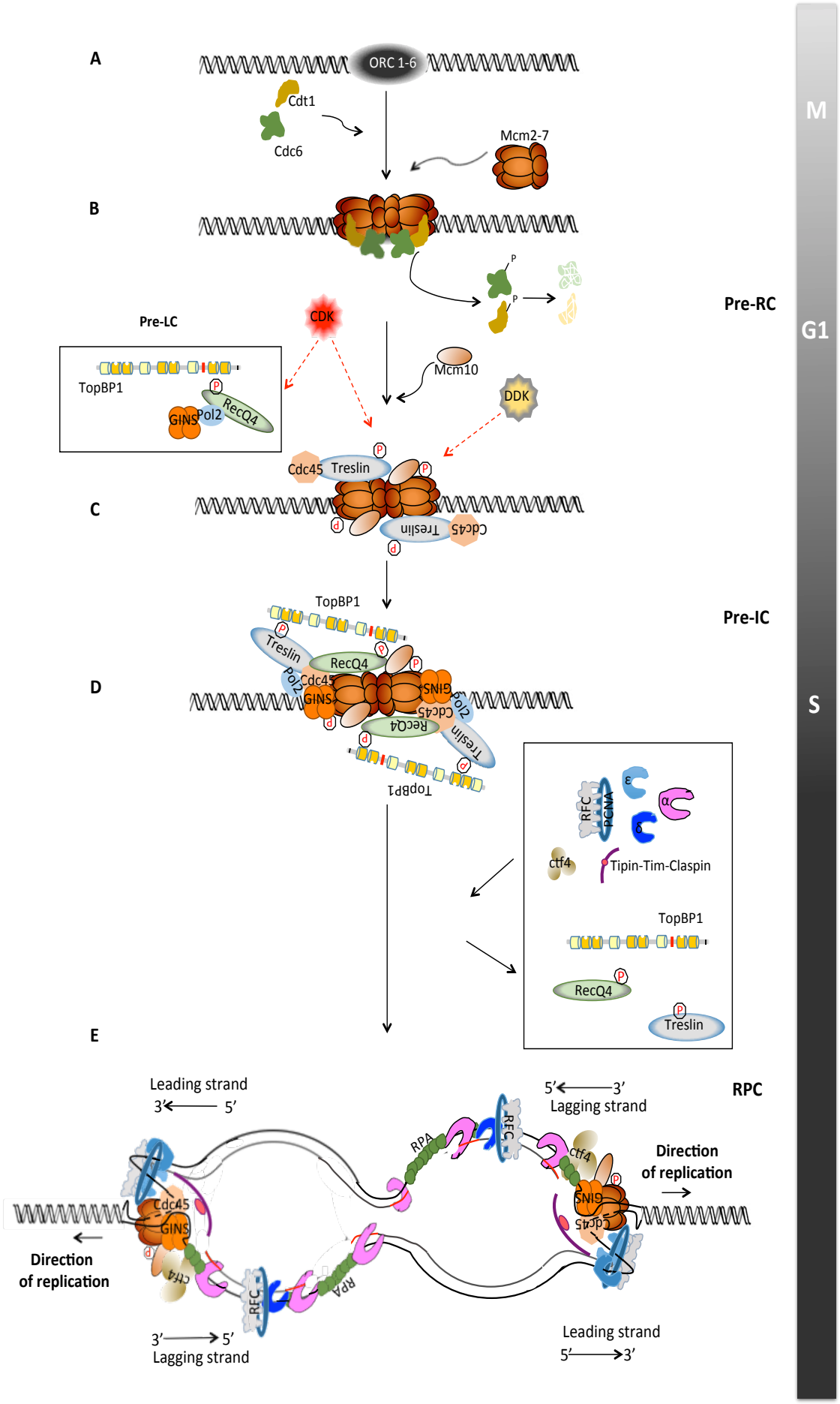

A
ORC 1-6
Cdt1
Cdc6
Mcm2-7
B
Pre-RC
Pre-LC
CDK
TopBP1
RecQ4
GINS
Pol2
Mcm10
DDK
C
Cdc45
Treslin
D
Pre-IC
E
RPC
Leading strand
3′ 5′
5′ 3′
Lagging strand
Direction of replication
RFC
PCNA
ctf4
Tipin-Tim-Claspin
RPA
Cdc45
GINS
M
G1
S

Figure 1.4: Initiation of DNA replication in eukaryotes. Origin licensing: **(A)** During late M-early G1 ORC binds the origin of replication and **(B)** together with Cdc6 and Cdt1 help recruit the MCM complex, forming the pre-RC. Soon after Cdc6 and Cdt1 dissociate and get destroyed. **(C)** At the start of S, CDK and DDK-mediated recruitment of Cdc45 and GINS form the CMG replicative helicase. Cdc45 is found associated with Treslin and binds the origins whereas RecQL4 is found in complex with GINS, pol2 and TopBP1 in a complex known as the pre-LC. Mcm10 is also recruited via an unknown mechanism. **(D)** All factors recruited in C define the pre-IC, a macromolecular complex at the origin that still has not fired replication. **(E)** Recruitment of additional factors such as the polymerases (α, ε and δ), RFC, PCNA, Ctf4 and Tipin-Tim-Claspin (and others not shown) results to the formation of the RPC and causes the necessary biochemical changes that will fire replication, thus defining the S phase. Treslin, RecQL4 and TopBP1 do not travel with the fork. (For details see the text)

hexameric minichromosome maintenance (MCM, Mcm2-7) complex onto double-stranded DNA as an inactive double hexamer (Nishitani, Lygerou et al. 2000). Given their importance in recruiting MCM, Cdc6 and Cdt1 proteins are both regulated by the cell cycle machinery. Cdc6 is subject to CDK-mediated phosphorylation, which takes it out of the nucleus in S phase as well as targeting to APC-mediated ubiquitylation and subsequent proteolysis during M phase. Geminin on the other hand, associates with and inhibits Cdt1 in human cells and this inhibition translates to an inability to promote pre-RC formation (Wohlschlegel, Dwyer et al. 2000). In fact, Geminin is present in high levels after entry into S phase but it gets destroyed via an APC-dependent pathway in mitosis. This is thought to provide yet another mechanism for the prevention of endoreduplication of the genetic material within a single cell cycle (Bell and Dutta 2002). Overall, ORC, Cdc6, Cdt1 and the inactive double hexamer of MCM altogether form the pre-RC and soon after its formation, CDC6 and Cdt1 dissociate, defining the next step of replication, replication initiation (Blow and Laskey 1986, Blow 1993, Bell and Dutta 2002, Arias and Walter 2007, Araki 2011).

Pre-IC formation

MCM is loaded around double-stranded DNA as an inactive double hexamer and it is only at the start of S phase that it associates with Cdc45 and GINS to form the CMG replicative helicase holocomplex (Evrin, Clarke et al. 2009, Gambus, Khoudoli et al. 2011). Maturation of the pre-RC and CMG activation is mediated via the action of two essential kinases, CDK and DDK (Dbf4 and Drf1-dependent kinase). Studies have shown that phosphorylation of $Sld2^{RecQL4}$ and $Sld3^{Treslin/ticrr}$ by $Cdc28^{CDK}$ in budding yeast is the minimal requirement for replication initiation (Zegerman and Diffley 2007). The substrate of DDK, on the other hand, appears to be the MCM complex (Jones and Petermann 2012). In particular, DDK phosphorylates MCM2, MCM4 and MCM6 (Jiang, McDonald et al. 1999, Masai, Taniyama et al. 2006, Sheu and Stillman 2006, Chuang, Teixeira et al. 2009). Phosphorylation of $Sld2^{RecQL4}$ and $Sld3^{Treslin/ticrr}$ by CDK create binding sites for the C-terminal and N-terminal BRCT domains of $Dpb11^{TopBP1}$, respectively (Zegerman and Diffley 2007, Labib 2010).

The role of TopBP1 in pre-IC formation

An essential function of $TopBP1^{Dpb11/Rad4}$ that is conserved among all the homologs is contributing to the initiation of DNA replication. Its role during initiation

involves bringing Polε and GINS to the origin-loaded Cdc45-Mcm2-7 complex via scaffolding Sld2^{RecQL4} and Sld3$^{Treslin/ticrr}$ (Zegerman and Diffley 2007).

Work in *S. cerevisiae* has shown that the interaction between Sld2^{RecQL4} and Dbp11 is CDK-dependent. Phosphorylation of a number of canonical motifs within Sld2^{RecQL4} is first required but this does not play a direct role in complex formation, but rather regulates phosphorylation of another residue, T84. This phosphorylated Sld2^{RecQL4} is then competent to interact with the BRCT3 and BRCT4 of Dpb11^{TopBP1} (see Figure 1.2) (Masumoto, Muramatsu et al. 2002). CDK-phosphorylated Sld2^{RecQL4} forms a fragile complex with GINS (Sld5, Psf1-3), the replicative polymerase ε (the Pol2 catalytic subunit) and Dpb11^{TopBP1} *in vitro,* in a complex known as the pre-loading complex or pre-LC (Araki 2011).

Furthermore, CDK needs to phosphorylate Sld3$^{Treslin/ticrr}$ and this is required for its interaction with Dpb11^{TopBP1} (Zegerman and Diffley 2007, Tanaka and Araki 2010). In addition, Sld3$^{Treslin/ticrr}$ itself interacts with Cdc45 throughout the cell cycle and DDK phosphorylation of Sld3$^{Treslin/ticrr}$ promotes binding of Dpb11 to Cdc45-associated Sld3$^{Treslin/ticrr}$ (Tercero, Labib et al. 2000, Kamimura, Tak et al. 2001, Yabuuchi, Yamada et al. 2006). A more recent finding from the Diffley laboratory suggests that Sld3 recognises DDK-phosphorylated peptides on Mcm4 and Mcm6, and this provides another explanation of how Cdc45 associates with the inactive origin-loaded MCM (Deegan and Diffley 2016, Deegan, Yeeles et al. 2016). Overall, Sld2^{RecQL4} and Sld3$^{Treslin/ticrr}$ are the minimal set of CDK targets required for promoting replication by bringing GINS-Polε and Cdc45 together (Zegerman and Diffley 2007). GINS and Cdc45 can then interact with Mcm2-7, forming the active helicase complex CMG. Once CMG forms and departs from the origin, the Dpb11^{TopBP1}-Sld2^{RecQL4}-Sld3$^{Treslin/ticrr}$ complex dissociates and does not travel with the fork (Bruck et al. 2011).

Similarly to Dpb11^{TopBP1}, in fission yeast, CDK-mediated phosphorylation of Sld2^{RecQL4} and Sld3$^{Treslin/ticrr}$ allows their interaction with Rad4^{TopBP1} BRCTs3-4 and 1-2, respectively. Both of these interactions are required for replication initiation and like Dpb11^{TopBP1}, Rad4^{TopBP1} does not travel with the fork (Fukuura, Nagao et al. 2011).

These interactions seem to be conserved in higher eukaryotes though the exact mechanisms are less understood. Studies in Xenopus suggest that RecQL4^{Sld2} does not get phosphorylated by CDK, nevertheless it co-immunoprecipitates with TopBP1 and it is the C-terminus of TopBP1 that mediates the interaction (Doi, Nagasaki et al. 2006). Treslin/Ticcr on the other hand, the Sld3$^{Treslin/ticrr}$ homolog, gets phosphorylated on two

serine residues by CDK2-Cyclin E in Xenopus (Kumagai, Shevchenko et al. 2011) and evidence from human cells suggests that it contacts BRCTs 1-2 of TopBP1 (Boos, Sanchez-Pulido et al. 2011).

Two additional interactions of TopBP1 at origins of replication described in Xenopus extracts are with GEMC1 (GEMinin Coiled-coil containing protein 1) and DUE-B (DNA unwinding element binding protein), both suggested to promote replication initiation by mediating recruitment of Cdc45 on replication origins (Balestrini, Cosentino et al. 2010, Chowdhury, Liu et al. 2010). Whether these interactions are relevant to *H. Sapiens* remains to be identified.

RPC and DNA synthesis

The exact molecular mechanisms of the transition from a stationary MCM double hexamer to a processive CMG translocase are still the subject of intense research. Several recent studies, however, suggest that DDK mediated phosphorylation of the MCM ring results to a conformational change that promotes ATP hydrolysis thus allowing duplex unwinding. All current models favour the idea that the active CMG translocates along single stranded DNA in a process that excludes the lagging strand from the central channel of the MCM holoenzyme ring (strand exclusion model) (Simon, Sannino et al. 2016, Sun, Yuan et al. 2016, Yuan, Bai et al. 2016). An additional factor required for initiation of replication is Mcm10, which is essential for the activation of the CMG helicase and the unwinding of the duplex, at least in budding yeast but the exact underlying mechanism has not yet been elucidated (van Deursen, Sengupta et al. 2012). All these interactions are key in bringing together all the players required for the replication progression complex (RPC) assembly and origin firing. And although the main aspects of the initiation reaction have been elucidated there is still a lot to learn about the precise molecular mechanisms and perhaps about the involvement of novel proteins in some of these steps (Labib 2010).

Unwinding of the double helix by CMG is followed by DNA synthesis which is bidirectional and occurs in a 5'-3' direction with respect to the newly synthesized strand. Due to the antiparallel nature of the duplex, replication is semi-discontinuous. One strand (the leading strand) is synthesized continuously and in the same direction as fork unwinding whereas the other strand (the lagging strand) is synthesized discontinuously (i.e. requires repeated priming and synthesis events) which leads to Okazaki fragments being formed (Sakabe and Okazaki 1966). In addition, because of

the fact that DNA strands are intertwined, DNA unwinding of the double helix results in topological stress building up ahead of the fork which is relieved by topoisomerases (Recolin, van der Laan et al. 2014). The actual reaction of the bulk of DNA synthesis in eukaryotes is catalysed by the replicative polymerases ε and δ. Experiments in yeast model systems assign Pol ε to the leading strand and Pol δ to the lagging (O'Donnell and Li 2016). There is however evidence of Pol δ synthesizing both strands under normal circumstances (Johnson, Klassen et al. 2015) or after replication restart following a pause site (Miyabe, Mizuno et al. 2015).

Interestingly, although both replicative polymerases are capable of dNTP incorporation hence DNA synthesis, they are not able to initiate *de novo* DNA synthesis. A third polymerase, known as polymerase α-primase, is instead required to initiate replication once in the leading strand and for each Okazaki fragment on the lagging strand (Kunkel and Burgers 2014, O'Donnell and Li 2016). The current favoured model of the architecture of the eukaryotic replisome suggests that the single stranded DNA exiting from the N terminus of the MCM U-turns to reach Pol ε sitting at the C-terminus side of the CMG helicase. This positions Pol ε ahead of the replicative helicase, suggesting that it might be involved in nucleosome disruption (Foltman, Evrin et al. 2013, O'Donnell and Li 2016). Another interesting feature of polymerases is that their "semi-closed hand" structure (see figure 1.2) does not allow them to continuously and stably associate with the DNA. To increase their processivity, polymerases associate with sliding clamps. The eukaryotic sliding clamp is a homotrimeric ring shaped molecule known as PCNA, which in addition to strengthening the interaction between the polymerases and the DNA it also has regulatory roles at the fork. PCNA itself is loaded onto DNA by the Replication Factor C (RFC, RFC1-5), which recognises primer-template junctions and uses energy from ATP to open the PCNA ring and load it on the DNA so as to encircle the duplex and also contact the polymerase (Leman and Noguchi 2013).

Other factors travelling with the replication fork include Ctf4 (also called And1) and Tipin-Tim-Claspin. Ctf4 is a homotrimeric protein that functions as a scaffold, bridging the CMG helicase to polymerase α (Gambus, van Deursen et al. 2009), although recent work suggests that it might associate with more partners, e.g. GINS (Villa, Simon et al. 2016). The Tipin-Tim-Claspin$^{\text{Csm3-Tof1-Mrc1}}$ component of the RPC on the other hand, has more specialized roles in checkpoint signalling following replication fork pausing (Hodgson, Calzada et al. 2007, Yao and O'Donnell 2016) but has also been

suggested to maintain normal replication fork rates in human cells (Petermann, Helleday et al. 2008).

It has emerged that the duplication of the genetic material as part of the S phase of the cell cycle is a highly orchestrated event. The recruitment of specialized factors, the biochemical changes triggered and, in parallel, the regulatory effects of the cell cycle machinery on the complexes recruited at the origins, all set the scene for the eventual assembly of the eukaryotic RPC and the firing of replication. Completion of S phase thus results in the generation of two copies of the genetic material that will be segregated to the daughter cells in mitosis.

1.1.4 Nuclear and cellular division

Following duplication of the hereditary material, cells go through the other important phase of their cell cycle, chromosome segregation and cellular division. In fact progression to the M phase generally occurs only if cells have successfully completed DNA replication. If errors have been incorporated during synthesis or if the DNA has been damaged in any way checkpoint pathways will be activated that will halt the progression of the cell cycle and allow time for repair.

Chromosome segregation happens in M phase and involves the partitioning of each genetic copy made in S phase into independent nuclei. The most striking characteristic of a mitotic cell compared to an interphase cell is the condensation of chromatin. So although an interphase cell has its chromatin diffused and cohesin keeps the sister strands held together, during mitosis a remarkable condensation takes place such that individual chromosomes become visible (Figure 1.1). And this is what allowed researchers, like Walther Flemming (1843-1905) to correctly deduce the events of chromosome movements during mitosis even before the improvement in microscopy (Paweletz 2001). And this condensation is also what gave mitosis its actual name; mitosis in Greek means “threads” which refers to the threadlike appearance of chromatin.

An increase in the levels of CDK1-Cyclin B (MPF) signals entry into mitosis. Mitosis is divided into five phases based on the physical state of the DNA and the spindle, the organelle organising and separating the chromosomes during mitosis. These are prophase, prometaphase, metaphase, anaphase and telophase. During prophase, there is transition from an amorphous DNA mass, where sister strands are attached to each

other via cohesin, to two distinguishable rod-shaped arms joined at the centromere, in a process known as resolution (Hagstrom and Meyer 2003).

Chromosome condensation continues until metaphase when chromosomes are fully compacted (Hagstrom and Meyer 2003). During the next stage of mitosis, prometaphase, MPF phosphorylates microtubule-associated proteins leading to the formation of the spindle apparatus. The spindle apparatus consists of two pairs of centrioles that move to opposite poles of the cells and catalyse the nucleation of microtubules. Microtubules polymerise to a length sufficient to mediate binding to the kinetochore of the sister chromatids. The kinetochore is a structure that forms at the centromere of each sister chromatid during prophase and provides an attachment site for microtubules. Necessary for the assembly of the spindle in metazoan cells is the nuclear memabrane breakdown, which is mediated through phosphorylation and subsequent depolymerisation of the nuclear lamins. In addition to nuclear membrane breakdown, no vesicular transport happens inside a mitotic cell as MPF promotes the breakdown of the endoplasmic reticulum and the Golgi apparatus (Lodish et al 2000).

At the end of prometaphase sister chromatids are arranged in such a way so that their kinetochores are attached to microtubules of opposite poles, in a process known as bi-orientation (Hagstrom and Meyer 2003). Next, during metaphase, chromosomes become fully condensed and they align along the centre of the cell (at the metaphase plate). Tension builds up but chromosomes can withstand spindle forces due to their condensed physical state and cohesin acting like glue at the centromeres (Hagstrom and Meyer 2003). At this point the cell activates the spindle assembly checkpoint and it is only when all the kinetochores are biorientated on the spindle that the checkpoint is released and the cell progresses to anaphase. Anaphase is characterised by the separation of sister chromatids which involves the cleavage of cohesin by an enzyme called separase. Sister chromatids start moving towards opposite spindle poles in a process that involves shortening of the microtubules as well as the action of motor proteins that move along the microtubules (Peters 2002).

When chromosomes reach the poles at telophase they start decondensing and also lamins get dephosphorylated and the nuclear lamina reforms using membrane vesicles from the parent cell's old nuclear membrane. By the end of telophase two daughter nuclei, each with a full set of chromosomes in the interphase conformation have formed (Lodish et al. 2002). The final step of the cell cycle is the division of the

two daughter cells, which is mediated by the contractile ring (Hauf, Waizenegger et al. 2001).

Mitosis is a crucial step within the living cycle of cells as it ensures that the progeny will receive an identical set of chromosomes. Together with DNA replication they comprise the key processes underpinning the cell cycle, hence life. DNA replication synthesizes (copies) the molecule of DNA, which is the carrier of the hereditary information and mitosis passes a complete copy of this information to the progeny. This cycle of life, no matter how simple it might seem from the above description, it is in reality a very complex system. And perhaps the only way of making sense of a composite biological system is to think of it in an organised, but still complex, way.

1.2 The DNA Damage Response (DDR)

1.2.1 DNA damage

DNA is a macromolecule and it is perhaps not surprising that it is subject to physical and chemical damage. Every cell in the human body can experience tens of thousands of DNA lesions per day (Jackson and Bartek 2009). DNA lesions can be generated by intracellular DNA damage or result from exogenous insult. DNA lesions generated by intracellular DNA damage include depurination due to loss of bases, base transitions due to deamination, SAM-induced methylation as well as oxidation generated by physiological by-products of oxidative respiration (Ciccia and Elledge 2010). Another threat to the integrity of the DNA double helix is impediments to the movement of a replication fork causing replication stress. In addition to sites of damaged DNA, replication stress can also be induced by nucleotide depletion, clashes between the replicative and transcriptional machineries, specific secondary DNA structures (e.g. G-quadruplex), tight DNA-protein complexes, triplet repeats, telomeric repeats and tRNA genes. Replication fork barriers (RFBs) can either block the polymerization step or interfere with duplex opening and, depending on the nature of the barrier, different cellular responses will be mounted (Lambert and Carr 2005). DNA damage can also result due to exogenous insult. Ultra violet (UV) light from sunlight causes chemical changes to the DNA helix and can induce approximately 100,000 lesions (e.g. pyrimidine dimers) per hour in a single cell. In addition, cosmic ionising radiation (IR) as well as IR used in medical treatments can generate a variety of DNA

lesions, with the most toxic being single-strand and double-strand breaks (SSBs and DSBs, respectively). Other DNA damaging sources include chemotherapeutic agents, pollution, cigarette smoke as well as certain chemicals in foods such as heterocyclic amines in overcooked meat. Overall, the effects of these agents on the double helix include base lesions, base alterations, strand breaks, nicks and crosslinks (Jackson and Bartek 2009, Ciccia and Elledge 2010).

1.2.2 The DDR; a genome maintenance network

The cell cycle regulation machinery not only controls the ordered sequence of events within the cell cycle but also cross talks to a network of factors which monitor the cell for the presence of incomplete DNA synthesis or unrepaired DNA damage. Inherent to this network is the ability of some cell cycle proteins themselves to perform genome maintenance functions. The maintenance of genome integrity is key to cell and organismal survival as DNA encodes the information to direct the production of proteins required for correct development and proliferation. To counteract DNA damage thus diminishing the risk of mutations that would endanger both the survival of the organism as well as the accurate transmission of the genetic material to future generations, cells have evolved sophisticated response mechanisms, collectively known as the DNA Damage Response (DDR) (Polo and Jackson 2011). The DDR involves a complex interplay between DNA damage signal transduction pathways that act to promote the repair or bypass of the damage. At the same time DDR collaborates with the cell cycle and apoptotic machineries of the cell. The presence of damage temporarily arrests the progression of the cell cycle to allow time for repair (Gorgoulis, Vassiliou et al. 2005). During this window of arrest and repair, higher eukaryotes have evolved to consider several cell fate options such as apoptosis, senescence or even activation of the immune system (Ciccia and Elledge 2010). And this is because of the different priorities of multicellular systems; elimination rather than propagation of an "unhealthy" cell is more favourable for the organism (Wahl and Carr 2001).

1.2.2.1 **Checkpoint signaling**

The concept of the checkpoint was first described in studies using Ataxia Telangiectasia (AT) patient-derived cell lines. Two studies were published describing the defect of AT cells in inhibiting DNA synthesis (i.e. delaying S phase) after exposure to radiation compared to control cells. So despite the lack of advanced technological

means researchers realized that the phenotypes observed could not be explained by defects in the repair process *per se* but rather by a "*process playing a broader role in development*" that was actually compromised (Houldsworth and Lavin 1980, Painter and Young 1980). Almost a decade later, Hartwell and Weinert described it as a "*control mechanism enforcing dependency in the cell cycle" (Hartwell and Weinert 1989). In* other words, an alarm system in place within the cells that ensures that a proceeding event within the cell cycle begins only when and if the previous event has been successfully completed. Their idea was born from the observation that although wild type yeast cells irradiated in S phase delayed the cell cycle to allow time for repair before mitosis, cells defective in particular genes (Sc *RAD9,* homologue of the human 53BP1, in particular) failed to delay. Imposing an artificial delay by chemical means, however, would largely rescue the observed phenotypes (Weinert and Hartwell 1988). With the power of yeast genetics, a battery of genes have been cloned by complementation and many proteins have been identified as "stop buttons" or "checkpoints" within the reproductive cycle of a cell and have been shown to be conserved in *Homo sapiens* (Carr 1996, O'Connell, Walworth et al. 2000). These will be discussed in the following sections.

DNA damage checkpoints are subdivided into categories and named after the cell cycle stage at which they operate (**Figure 1.3)**. Thus there is a G1-S checkpoint responding to damage during G1, an intra-S checkpoint dealing with DNA damage during replication, a replication checkpoint responding to replication stress, a G2-M checkpoint dealing with damage during G2 (and lesions that have remained unrepaired from previous stages) and an S-M checkpoint that delays the onset of mitosis until DNA replication has been completed (Enoch and Nurse 1990, Kastan and Bartek 2004, Smith, Tho et al. 2010). These checkpoints are comprised of specialized macromolecular complexes, which upon DNA damage are capable of catalysing specific reactions to communicate the problem and help resolve it. They should not be viewed as solid machineries that have evolved to perform one function. In reality they are very complicated networks of dozens if not hundreds of proteins that are extremely dynamic in nature and pleiotropic in their functions and this explains why they are under stringent control. And it is the type of damage that dictates the specific mechanisms that will be activated and the cell-cycle checkpoint machineries will orchestrate the fine-tuning of the response.

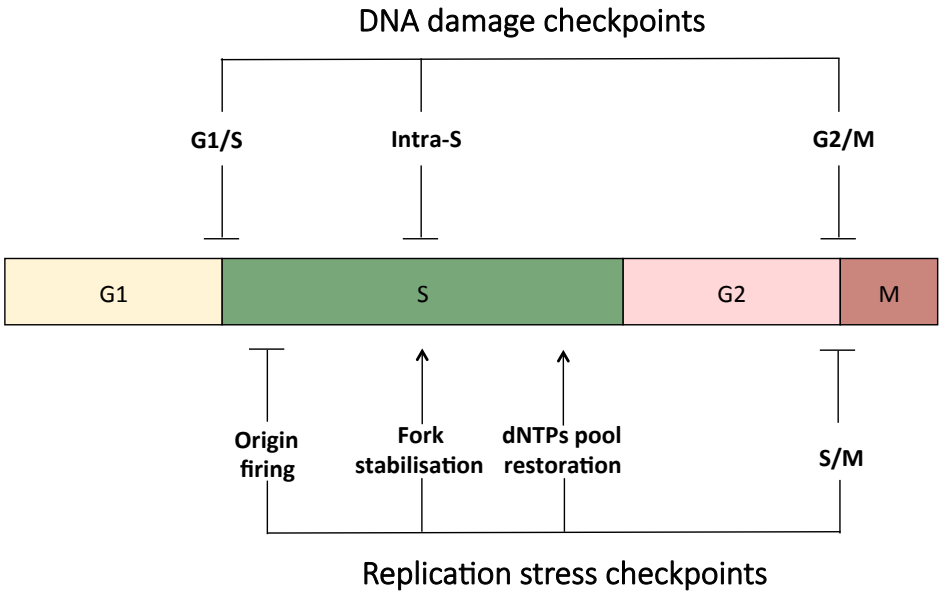


Figure 1.3: DNA structure-dependent checkpoints. Cells possess points along the length of the cell cycle where they can activate the checkpoint pathways, which act to halt the progression to the next phase until damage is repaired. In the presence of DNA damage, cell cycle halts at the G1-S, S or G2-M, depending on which phase exactly the cell was traversing at the time damage occurred. Replication stress-associated damage or other factors that cause replication fork stalling activate the replication checkpoint which has many functions including the inhibition of origin firing, stabilisation of stalled forks, restoration of the dNTP pools and inhibition of mitosis (S-M checkpoint). Adapted from Smith Gillespie et al 2010.

PIKKs at the heart of signalling

Life on earth would certainly not be as we know it if cells have not evolved defense mechanisms of keeping their precious code of life –the DNA– intact. Mechanisms which are themselves produced from what they evolved to protect; a series of nitrogen-based molecules which "carry" the code of life just like a computer software. A computer software however, more advanced than any other ever been created. And just like a defect in a computer will show as a pop up window to its user who will embark on strategies to try and repair it, defects on the DNA can somehow alert the cell and activate a series of processes that will attempt to restore the problem. Although it is not known exactly how DNA damage on the duplex is physically detected, it is well established that DNA damage leads to a quick and robust activation of a protein-signaling cascade that regulates several aspects of nuclear and cellular physiology to provide an environment conducive for successful maintenance of genomic integrity. The existence of such a cascade is beneficial as it allows a single input (damage) to be translated to many different outputs in a short period of time. At the heart of the signal transduction network of the DDR checkpoints lie two protein kinases, ataxia telangiectasia mutated (ATM) and ATM-and-Rad3-related (ATR).

Both kinases are conserved down to the yeasts *S. pombe* and *S. cerevisiae* as $Rad3^{ATR}$/$Tel1^{ATM}$ and $Mec1^{ATR}$/$Tel1^{ATM}$, respectively (Ciccia and Elledge 2010). The ATM gene was cloned in 1995 and took its name after being identified as the gene responsible for the AT phenotypes (A-T mutated) (Savitsky, Bar-Shira et al. 1995). Soon after, the full sequence of the yeast Rad3 was identified and the human ATR cDNA cloned by virtue of sequence similarities (Bentley, Holtzman et al. 1996). ATM and ATR are large kinases that belong to the phosphoinositide 3-kinase-like protein kinases (PIKKs) family of proteins together with mechanistic target of rapamycin (mTOR), suppressor of morphogenesis in genitalia (SMG1), DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and transformation/transcription domain-associated protein (TRRAP) (**Figure 1.4**). All members of this family share sequence similarity to the 300 amino acid kinase domain of phosphatidylinositol-3 kinases (PI3Ks), a classical family of inositol lipid kinases (Hunter 1995, Bosotti, Isacchi et al. 2000, Foster, Traer et al. 2003). They nonetheless lack lipid kinase activity; instead they are all highly conserved Ser/Thr-kinases and contain a kinase domain KD (although in TRRAP is not active) located at the C-terminus, flanked by the FRAP-ATM-TRRAP

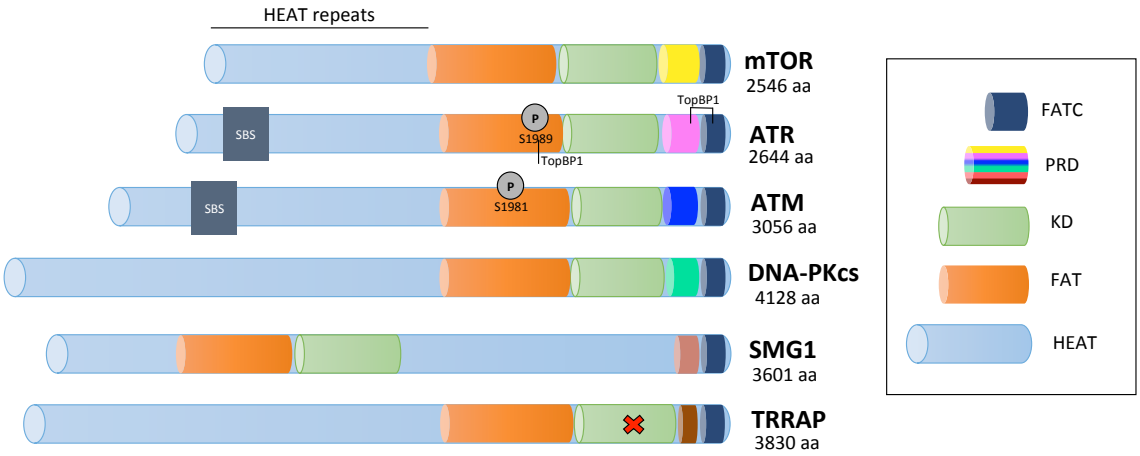


Figure 1.4: Domain architecture of PIKKs. The kinase domain (except TRRAP which lacks kinase activity) is located at the C-terminus and is flanked by the FAT and FATC domains. Note that the KD of TRRAP is inactive as designated by the red cross. Between the FAT and FATC domain lies the PRD domain, which is specific to each PIKK. The substrate-binding site SBS is shown for ATR and ATM and also regions important for the interaction of TopBP1 with ATR are depicted. Adapted from (Kastan and Bartek 2004, Derheimer and Kastan 2010).

(FAT) and FAT-C-terminal (FATC) domains. Sitting between the KD and FATC is the so-called PIK-kinase regulatory domain (PRD), which differs among the different PIKKs and is thought to be important for their regulation (Cimprich and Cortez 2008). The N-terminus on the other hand, does not seem to harbor specific domains but instead contains HEAT (Huntingtin Elongation Factor 3 Alpha subunit and TOR1) repeats, each of which is a pair of antiparallel α-helices linked by a flexible "intra-unit" loop (Perry and Kleckner 2003). Despite their sequence similarity, PIKKs have quite diverse biological functions (Lempiainen and Halazonetis 2009, Lovejoy and Cortez 2009). SMG1 regulates elimination of mRNA molecules containing premature termination codons, mTOR functions in nutrient signaling, metabolism and cell growth and TRRAP is involved in transcription (Lovejoy and Cortez 2009). **DNA-PKcs** on the other hand has a role in DNA repair and will be discussed later. So that leaves only **ATM** and **ATR** having the key role in DNA-damage-dependent checkpoint signaling.

ATM and ATR kinases, once activated, can phosphorylate substrates on serine and threonine residues followed by a glutamine (SQ/TQ). In particular, they have a preference in phosphorylating SQ/TQs found in clusters, known as SQ/TQ cluster domains (SCDs). Protein phosphorylation events therefore play a key role in signal transmission and amplification (Nam and Cortez 2011). In the following paragraphs I will attempt to illustrate the present-day understanding of the molecular aspects of this signaling cascade.

The ATM signaling cascade

The ATM kinase can be activated in the G1, S and G2 phases of the cell cycle. This became evident long before the identification of the ATM gene. Researchers observed that contrary to wild-type cells, AT cells failed to delay the cell cycle following irradiation in G1 and G2 and they also exhibited IR-insensitive DNA synthesis (Houldsworth and Lavin 1980, Painter and Young 1980, Hong, Gatti et al. 1994, Khanna, Beamish et al. 1995). But how do these cellular phenotypes translate to a human disorder? AT is a rare autosomal recessive disorder characterized by radio sensitivity, early onset debilitating cerebellar ataxia, dilation of blood vessels (telangiectasia), neurodegeneration, susceptibility to bronchopulmonary disease and predisposition to cancer, particularly lymphoid tumours. This broad spectrum of severe symptoms arises from the many roles of ATM in the DDR and other cellular processes

(Jeggo, Carr et al. 1998, Lavin 2008).

ATM activation

ATM is primarily activated in response to DSBs, a lesion generated when each strand of the duplex is broken close in time and space such that the DNA ends fail to be kept juxtaposed. DSBs represent one of the most toxic lesion on the DNA. Such dangerous breaks can be caused by ultraviolet light (UV), mutagenic chemicals, radiomimetic drugs and reactive oxygen species (ROS) (produced by IR or physiological processes of cellular metabolism) as well as be generated at broken or collapsed replication forks. The potential toxicity of DSBs lies in the susceptibility of the free damaged DNA ends towards pathogenic chromosomal rearrangements, which can result to genomic instability, a hallmark of cancer (Hoeijmakers 2001, Jackson 2002, Tsai and Lieber 2010).

Initial studies on its biochemical properties revealed that activation of the ATM kinase, which is present throughout the cell cycle, is one of the earliest events observed following DNA damage (Canman, Lim et al. 1998). In its inactive state ATM non-covalently homodimerises, with each monomer inserting its KD into the FAT region of its partner (**Figure 1.5**). Intermolecular phosphorylation of S1981 (a hallmark of activated human ATM) within the FAT domain from the pairing molecule is thought to be a necessary but not sufficient step in the monomerisation and subsequent activation of the kinase (Guleria and Chandna 2016). An engineered S1981A non-phosphorylatable form of ATM was unable to rescue the checkpoint defect of AT cells and also displayed a dominant inhibitory effect over activation of endogenous ATM when transfected in wild type human cells (Bakkenist and Kastan 2003). Phosphorylation of S1981 is also thought to stabilize ATM at the break site (So, Davis et al. 2009).

Subsequent studies (Kozlov, Graham et al. 2006, Kozlov, Graham et al. 2011) have revealed that activation of ATM is more complex as it involves autophosphorylation of more residues (S367, S1893, S2996). The importance of all these sites for the activity of the kinase is however a subject of debate.

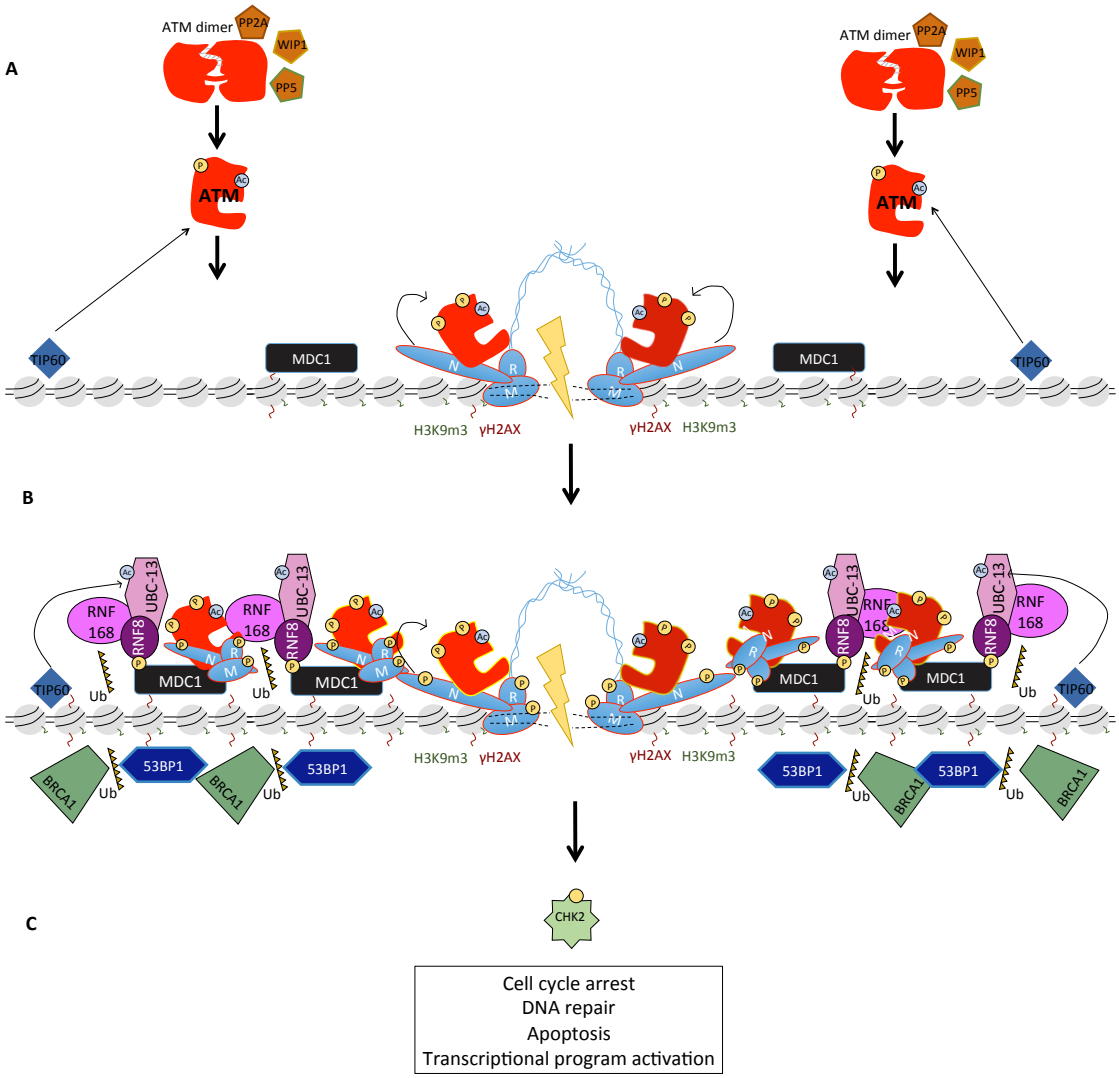


Figure 1.5: The ATM signalling cascade at DSBs. A) In an unstressed cell ATM exists in an inactive dimer as well as dephosphorylated through the action of phosphatases. **B)** In response to DSB formation, however, ATM undergoes autophosphorylation and TIP60-mediated acetylation and converts from inactive dimers to active monomers that bind the MRN sensor bridging the ends of the break. Full activation of ATM following recruitment allows the phosphorylation and activation of many factors. Phosphorylation of H2AX to form γ-H2AX at nearby nucleosomes allows recruitment of MDC1 that promotes a feed-forward loop spreading activated ATM and γ-H2AX to long distances flanking the site of break. Activated MDC1 recruits RNF8 ubiquitin ligase, which promotes the recruitment of yet another one RNF168. This sets the scene for a ubiquitination cascade that helps recruit RAP80-associated BRCA1 and 53BP1. **C)** Checkpoint signalling through the ATM cascade results to activation of the CHK2 effector kinase, which halts the cell cycle, promotes repair as well as activation of gene expression and apoptotic pathways. Adapted from (Marechal and Zou 2013).

Dephosphorylation and acetylation might also contribute to the activation of ATM. In fact, PP2A and WIP1 phosphatases interact with ATM and negatively regulate its phosphorylation status perhaps in an effort from the cell to prevent unwanted activation of the kinase (Ali, Zhang et al. 2004). In addition, the histone acetyl-transferase TIP60 modifies K3016 and this also promotes ATM activation (Derheimer and Kastan 2010). In fact this modification precedes the autophosphorylation of ATM thus possibly having a predominant role in ATM monomerisation and activation (Sun, Xu et al. 2007). TIP60 itself is regulated by the c-Abl kinase which catalyses the phosphorylation-mediated association of TIP60 with H3K9m3 and subsequent acetylation of ATM (Kaidi and Jackson 2013), thus suggesting that activation of ATM happens in a chromatin-proximal manner. Such post-translational modifications of ATM provide very dynamic mechanisms of not only contributing to the activation of the kinase but also in allowing recovery from DNA damage and return of the cell to its physiological state.

The exact biochemical mosaic of ATM activation as well as the initial trigger that will convert the quiescent dimers of ATM into an "avid phosphorylation machine" are still a matter of ongoing debate (Shiloh and Ziv 2013). Nevertheless, it is accepted that the other key step in the activation of this kinase is its (re)localization to the break site. Although ATM seems to possess a weak DNA binding activity (Smith, Cary et al. 1999) its localization *in vivo* is most of the times dependent on the Mre11-Rad50-Nbs1 (MRN) complex. Initial evidence for the role of MRN in ATM activation came from the observation that mutation of its components leads to human disease. The most favourable model suggests that MRN is the most proximal sensor of DSBs, possibly mediated through its intrinsic DNA binding ability. The recruitment of MRN makes the break lesion accessible to ATM and also promotes end-processing and repair to occur (Bhatti, Kozlov et al. 2011). Via the Mre11-Rad50 subunits, MRN binds DNA as a heterotetramer bridging in this way the broken ends together (Paull and Lee 2005). Its subsequent interaction with ATM allows the kinase to stably associate with the DSB.

Another key determinant for the activation of ATM is catalyzed by the C-terminal 20 amino acids of Nbs1 by a mechanism, which seems to be important for the activation of most PIKKs by their co-factors (Falck, Coates et al. 2005). Finally, as well as being an upstream regulator of ATM, MRN is also one of its downstream factors. ATM subjects all MRN subunits to phosphorylation and although the functional roles of these phosphorylation events are not entirely understood, it is clear that they regulate

many aspects of the damage-induced cascade, not only of ATM, but of ATR as well. For instance, Nbs1 gets phosphorylated on S278 and S343 by ATM, with the latter residue also being a target of the ATR kinase. Phosphorylation of these residues is required for S-phase checkpoint control and for activation of the Chk2 effector kinase (Paull 2015).

The mediators take over

The recruitment and interaction of ATM and MRN at the site of damage and the activation of ATM through the action of the mechanisms described above sets the scene for the propagation of the signal to more damage response protein factors on the chromatin flanking the break.

Activated ATM phosphorylates histone H2AX on S139 forming γ-H2AX (Burma, Chen et al. 2001). Although phosphorylation of H2AX is not required for the activation of ATM substrates like CHK2 and p53, it is important for the recruitment of DNA repair proteins (Fernandez-Capetillo, Chen et al. 2002). In fact, γ-H2AX provides a high-affinity binding platform for MDC1, which orchestrates the recruitment of essentially all of the downstream factors. As well as binding γ-H2AX via its BRCT domain, MDC1 also binds ATM via its FHA domain and also makes contact with Nbs1, thus helping concentrate ATM at the break and sustain signal amplification (Lou, Minter-Dykhouse et al. 2006, Wu, Luo et al. 2008). The direct binding of MDC1 to γ-H2AX *per se* is also thought to protect the histone variant from dephosphorylation, which again contributes to a sustained response (Stucki, Clapperton et al. 2005). As a result, γ-H2AX near the break gets continuously phosphorylated by ATM leading to the formation of a positive feedback loop that recruits more MDC1 molecules. This helps maintain the construction and the shape of the DSB focus and the processes that take place within it (Savic, Yin et al. 2009).

ATM also phosphorylates MDC1 within its various TQXF clusters and this phosphorylation is critical for the recruitment of the ubiquitin ligase RNF8 (Kolas, Chapman et al. 2007). RNF8 forms a complex with an E3 ubiquitin conjugating enzyme called UBC-13 and together catalyze polyubiquitination of γ-H2AX (Plans, Scheper et al. 2006, Mailand, Bekker-Jensen et al. 2007). Important for the ability of UBC-13 to catalyze ubiquitination is its acetylation by TIP60 (Ikura, Tashiro et al. 2007). This creates a favorable chromatin environment for the recruitment of another ubiquitin ligase RNF168 (Kolas, Chapman et al. 2007, Doil, Mailand et al. 2009). In fact,

RNF168 is capable of binding to the RNF8-created ubiquitin chains on γ-H2AX, in a reaction that is dependent on RNF8. Upon binding it further boosts the formation of K63-linked ubiquitin conjugates to a level that is favorable for recruitment of BRCAI and 53BP1 (Doil, Mailand et al. 2009). Another factor to join the party is RAP80, which has an intrinsic binding affinity for ubiquitin chains and once localized to the site is able to recruit many factors including BRCA1 (Sobhian, Shao et al. 2007, Wang and Elledge 2007). The recruitment of 53BP1 on the other hand is mediated by methylation marks on the chromatin, in particular H4-K20m2 and H3-K79m3. Evidence suggests that these histone residues are constitutively methylated in human cells so it is more likely that the processing of the DSB region results to a change of the conformational status of chromatin, allowing these methylation marks to become exposed and accessible (Huyen, Zgheib et al. 2004, Botuyan, Lee et al. 2006).

53BP1 and BRCA1 are both substrates of ATM (Cortez, Wang et al. 1999, Xia, Morales et al. 2001) and are able to not only influence the outcomes of the response by help recruiting other factors but also dynamically interact with factors already bound on the lesion. In fact, 53BP1 interacts with both ATM and Rad50, which is thought to sustain the signal at low levels of MRN (Lee, Goodarzi et al. 2010).

Interestingly, other molecular factors get recruited and activated by ATM. Factors involved in DSB repair like RAD9, DNA-PK, CtIP and the already bound Nbs1 and Rad50, proteins involved in chromatin relaxation, remodeling of nucleosomes and regulation of transcription. All of these proteins have their own substrates and this allows the signal to diverge (Lavin and Khanna 1999, Khanna, Lavin et al. 2001, Lavin 2008, Shiloh and Ziv 2013, Awasthi, Foiani et al. 2015). One of the best-characterized kinase substrates of ATM is the effector kinase CHK2. CHK2 is a protein expressed throughout the cell-cycle that remains stable but inactive unless damage occurs (Lukas, Bartkova et al. 2001). ATM phosphorylates CHK2 on T68 within the SQ-TQ-rich motif at the N-terminus. The phosphorylated SQ-TQ-rich motif of one CHK2 monomer is then recognized by the FHA domain of another, leading to homodimerisation and intermolecular activation (Ahn, Li et al. 2002). Activated CHK2 is then released from the site of the lesion to perform pan-nuclear functions by phosphorylating numerous substrates (Lukas, Falck et al. 2003). The tumor suppressor p53 and its regulator MDMX, the cell-cycle phosphatase CDC25, the tumor suppressor BRCA1 and the transcriptional regulator E2F1 are few of the substrates of CHK2 which are also ATM substrates. The concerted action of these and other proteins determine the choice of

repair machineries recruited and also contribute to the decision-making between the activation of cell survival and cell death programs (Smith, Tho et al. 2010).

The initial discovery of ATM and its position in the DDR was a milestone step towards uncovering a very complex signaling network that for sure researchers did not expect. Over 700 ATM substrates have been identified and the list might grow even bigger (Matsuoka, Ballif et al. 2007). A single lesion on the DNA can activate the apical kinase ATM and within minutes an orchestrated network of hundreds of proteins that all collaborate in an astonishingly orderly way to pursue unique physiological functions that determine the fate of the entire cell. The key determinants within this network are phosphorylation-dephosphorylation events, which modulate the activity, localization, stability, interplay and turnover of all these factors as well as chromatin dynamics to fine-tune the outcomes of this massive signaling network (Shiloh and Ziv 2013).

Interestingly, the fission and budding yeast homologs of ATM, Tel1, are involved in telomere maintenance and play a minor role in the checkpoint response (Greenwell, Kronmal et al. 1995).

Cell cycle functions of the ATM cascade

Damage within the G1 phase of the cell cycle leads to activation of the G1-S checkpoint, which is executed in two phases. The first phase involves the ATM-mediated activation of the effector kinase CHK2. Activated CHK2 phosphorylates and inhibits CDC25A phosphatase and in this way stabilizes the inhibitory phosphorylations of T14 and Y15 on CDK2. This blocks the formation of an active CDK2-CyclinE complex and prevents G1 to S transition (Mailand, Falck et al. 2000). The second phase of the G1 checkpoint is responsible for a more delayed and sustained response to the damage and relies on an ATM- and CHK2-dependent activation of the tumor suppressor p53 (Kastan, Onyekwere et al. 1991, Kastan, Zhan et al. 1992, Chehab, Malikzay et al. 2000). p53 is a transcription factor capable of activating numerous genes required for blocking the G1 to S transition, repairing the damage as well as genes involved in apoptotic pathways (Hirao, Cheung et al. 2002, Takai, Naka et al. 2002).

The molecular mechanisms underlying the operation of the S-phase checkpoint are much less understood. In general, activation of this checkpoint is thought to stabilize stalled replication forks and inhibit origin firing until damage has been repaired (Lambert and Carr 2005). S-phase checkpoint-mediated CHK2 activation is important for cell cycle arrest via the same mechanisms described for G1 arrest. In addition to

that, other factors that get phosphorylated by ATM include BRCA1, FANCD2, SMC1 and Nbs1. Although all of these factors are required for a successful S phase checkpoint, their exact effect on inhibiting replication is not yet clear. Nonetheless, the function of all these S-phase checkpoint regulators is better understood in the context of damage repair (Derheimer and Kastan 2010, Guleria and Chandna 2016).

Arrest at the G2 phase is achieved via CHK2- and NEK11-mediated phosphorylation of CDC25C. Phosphorylated CDC25C associates with 14-3-3-σ, which targets it for nuclear export (Peng, Graves et al. 1997, Melixetian, Klein et al. 2009). BRCA1 also has a role in inhibiting the CDK-Cyclin complex by regulating the expression of WEE1 and 14-3-3-σ (Yarden, Pardo-Reoyo et al. 2002). Interestingly, the G2-M-specific function of BRCA1 can be abrogated by mutation of S1423, which is a target of ATM (Xu, Kim et al. 2001). Finally, a p53-mediated expression of genes, such as p21 and 14-3-3-σ, also contributes to the delay of damaged cells in the G2 phase of their cell cycle (Flatt, Tang et al. 2000).

Repair function of the ATM cascade

Several repair pathways can process and correct sites of DSBs on the DNA, such as non-homologous end joining (NHEJ), homologous recombination (HR) and single-strand annealing (SSA). The key determinant in the choice of pathway is the extent of resection, with NHEJ not requiring any, MMEJ requiring 5-25 resected nucleotides and HR and SSA requiring more extensive tracts (Hartlerode and Scully 2009).

DSBs occurring before DNA replication are repaired by end joining pathways due to the lack of a homologous sequence (provided by the sister chromatid in S and G2) (Marechal and Zou 2013). NHEJ occurs throughout the cell cycle and is initiated with the binding of the Ku70/Ku80 heterodimer to the exposed ends of the break and the recruitment of the catalytic subunit of DNA-PK (DNA-PKcs). The DNA-PKcs/Ku70/Ku80 complex recruits ligase IV, which joins the ends together, thus healing the lesion (Blier, Griffith et al. 1993, Nick McElhinny, Snowden et al. 2000, Sonoda, Hochegger et al. 2006). DNA ends that are not compatible for ligation can be processed by nucleases, such as Artemis and APLF (Mahaney, Meek et al. 2009). This classic NHEJ pathway is thought to proceed with fast kinetics and to be ATM-independent, although several substrates of DNA-PKcs are phosphorylated by ATM (Ciccia and Elledge 2010). A more direct role of ATM and Artemis in DSB repair has been proposed by Jeggo and Lobrich laboratories, with both factors being responsible

for the slow repair of a subset of breaks in both the G1 and the G2 phases of the cell cycle. In particular, ATM and Artemis channeled the G1 subset of breaks to NHEJ repair whereas they promoted an HR-mediated repair of the G2 subset of breaks (Beucher, Birraux et al. 2009). In S and G2, CDK activation allows extensive generation of ssDNA through the MRN-mediated recruitment of CtIP, Exo1 and Dna2. This promotes a switch from ATM to ATR at DSBs and repair via homology search and will be discussed later (Ciccia and Elledge 2010, Marechal and Zou 2013).

The role of the ATM cascade in cell fate decision – live or die?

ATM-mediated activation of the tumor suppressor p53 activates the expression of genes that impose cell-cycle arrest but at the same time drives the expression of genes that promote programmed cell death (Shiloh and Ziv 2013). p53 mechanism of action has been described to entail wave-like changes or pulses of its protein levels following damage in human cell lines (Geva-Zatorsky, Rosenfeld et al. 2006). This pulsatile nature of p53 protein works like a molecular timer that allows cell cycle arrest at low thresholds and induction of apoptosis at higher thresholds. A key determinant in this model proposed by Zhang and colleagues is the severity of DNA damage. Low levels of damage go down the repair route whereas at high levels of damage repair is suppressed and apoptosis begins. This mechanism makes the p53 network a flexible and dynamic network that can very elegantly decide cell fates based on the integrity status of the duplex (Zhang, Liu et al. 2009). Post-translational modifications also play a role in the regulation of this network. Just like phosphorylation of p53 on S15 drives cell cycle arrest, phosphorylation of S46 is thought to promote expression of pro-apoptotic genes (Oda, Arakawa et al. 2000). Loss of p53 is a driver to uncontrolled growth and genomic instability, both hallmarks of carcinogenesis (Carr 2000).

The ATR signaling cascade

Although ATM is predominantly activated by DSBs, a relatively rare type of damage, ATR is primarily activated by ssDNA. This structure is generated during the physiological process of DNA replication but also, in higher amounts, at compromised replication forks (Lonn and Lonn 1988). Different amounts of ssDNA are thought to allow cells to sense the severity of the problem thus mounting various degrees of checkpoint activation (Recolin, van der Laan et al. 2014). ssDNA is also formed at sites of DNA damage through DSB processing by nucleases, and also at sites of ongoing

repair through the processing of bulky lesions. Thus the ability of ATR to sense ssDNA gives it a broad function and renders it a key response mechanism to a plethora of DNA metabolic processes (Lambert and Carr 2005, Cimprich and Cortez 2008, Marechal and Zou 2013).

Attempts from many labs to biochemically characterize ATR have all failed to purify an active form of the kinase. A key step in the activation of the kinase is its localization to sites of ssDNA where it is required. Early studies in *S. pombe* have identified $Rad3^{ATR}$ and $Rad26^{ATRIP}$ to exist as a soluble complex in cells, having a checkpoint function upstream of other Rad proteins (Edwards, Bentley et al. 1999). Subsequent studies in human cells have shown that the two partners are mutually dependent to each other as reduction of the protein levels of one leads to a decrease in the levels of the other and a subsequent checkpoint defect (Cortez, Guntuku et al. 2001). Perhaps the strongest evidence of the mutual dependence of ATR^{Rad3} and $ATRIP^{Rad26}$ comes from the fact that loss of either of the two produces the same phenotypes at an organismal and cellular level (Cimprich and Cortez 2008). Additionally, the stability of the ATR-ATRIP complex has been suggested to involve phosphorylation by the NEK1 kinase (Liu, Ho et al. 2013). So unlike ATM, which exists as inactive homodimers that break apart after damage, ATR is always found in a stoichiometric heterocomplex with its obligate co-factor ATRIP (ATR-interacting protein), even in untreated cells (Recolin, van der Laan et al. 2014). ATR itself is unable to sense ssDNA regions and it is instead ATRIP that acts as the sensor. ATRIP can localize ATR to sites of damage by binding directly not to ssDNA itself but to the tripartite complex of RPA, which always coats exposed DNA strands (Zou and Elledge 2003, Dart, Adams et al. 2004, Fanning, Klimovich et al. 2006).

RPA is composed of the subunits RPA70, RPA32 and RPA14 and binds ssDNA with a very high affinity (K_d ~10^{-9}-10^{-10} M) through oligonucleotide/oligosaccharide-binding (OB) fold domains (Kim, Paulus et al. 1994). RPA is a complex subjected to extensive post-translational modifications in a cell-cycle regulated manner. In response to replication stress ATR phosphorylates RPA32 on S33 whereas at DSBs S33 is followed by phosphorylation on S4/S8 by DNA-PK (Shiotani, Nguyen et al. 2013). Treatment of cells with chemical agents that cause replication stress in a checkpoint defective background causes exhaustion of RPA pools and leads to replication catastrophe, marked by phosphorylation of RPA32 on T21 and S4/S8 (Toledo, Altmeyer et al. 2013). The RPA70 subunit on the other hand is key for the DDR

through its ability to recruit important factors such as ATR-ATRIP, 9-1-1, MRN and PRP19 (Marechal and Zou 2015).

Crucial for ATR activation is the ability of ATRIP to recognize RPA70 via an acidic alpha helix within its checkpoint recruitment domain (CRD) (Ball, Ehrhardt et al. 2007). Mutating the ATRIP binding surface on RPA70 causes a significant defect in ATR-mediated Chk1 phosphorylation (Xu, Vaithiyalingam et al. 2008). The importance of ATRIP in ATR recruitment to RPA-ssDNA substrates in yeast model systems is evident from the fact that, in *S. cerevisiae*, *ddc2Δ*$^{ATRIP\Delta}$ has the same phenotype as a *mec1Δ*$^{ATR\Delta}$ and mutating RPA leads to a reduction of Ddc2^{ATRIP} foci formed following DNA damage (Zou and Elledge 2003). Another interesting observation coming from budding yeast is that the C-terminus of Mec1^{ATR} is also contacting RPA in a Ddc2^{ATRIP}-dependent way (Nakada, Hirano et al. 2005).

As with ATM, localization of ATR-ATRIP to RPA-ssDNA nucleofilaments is by itself not sufficient to activate the signaling response fully. In addition to RPA-ssDNA, the junction of ssDNA and dsDNA is also an important structure for ATR activation. These junctions are recognized by the PCNA-like 9-1-1 complex (*sp*/*h* Rad9-Rad1-Hus1, *sc* Ddc1-Rad17-Mec3), which is loaded onto chromatin independently of ATR and in an ATP-dependent manner (Parrilla-Castellar, Arlander et al. 2004, MacDougall, Byun et al. 2007). In particular, 9-1-1 recognizes the DNA end that is proximal to the RPA-coated ssDNA and its loading is mediated by the damage-specific clamp loader Rad17-RFC (RFC2-5; i.e. Rfc1 replaced by Rad17 for checkpoint control) in a manner analogous to PCNA loading by RFC during replication (Caspari, Dahlen et al. 2000, Bermudez, Lindsey-Boltz et al. 2003, Ellison and Stillman 2003). An important determinant during this process seems to be the RPA heterotrimer. Firstly, the defined polarity by which RPA wraps around the ssDNA (with the C-terminus of RPA1 facing the 3'-end of the ssDNA), directs loading of the 9-1-1 complex to 5'junctions (Bochkarev, Pfuetzner et al. 1997, Majka, Binz et al. 2006). Secondly, pre-mRNA processing factor (PRP19)-mediated ubiquitination of RPA is thought to promote ATRIP tethering on chromatin (Marechal, Li et al. 2014).

The independent recruitment of the two RPA-dependent complexes, ATR-ATRIP and the 9-1-1 complex, at the site of the damage promotes the in *trans* phosphorylation of ATR on T1989 in the FAT domain. This autophosphorylation event, although it only enables the basal kinase activity of ATR, has been described as a molecular switch towards robust checkpoint activation, owing to its ability to promote

interaction of the ATR kinase with its main activator –at least in higher eukaryotes– TopBP1 (Topoisomerase II binding protein I; spRad4, scDpb11) (Furuya, Poitelea et al. 2004, Liu, Shiotani et al. 2011). The importance of the 9-1-1 complex could be envisaged since its discovery as mutations within *the rad9* gene caused sensitivity to IR and UV, thus suggesting a role in the checkpoint response (Murray, Carr et al. 1991). *rad9* knockout mouse embryonic stem cells show spontaneous chromosomal aberrations, are sensitive to DNA damaging agents and have a partially defective G2 checkpoint whereas homozygous mutant embryos die within the early days of gestation (Hopkins, Auerbach et al. 2004). An interesting function of 9-1-1 has been described in yeasts. In *S. cerevisiae*, it has been shown that Ddc1-Rad17-Mec3$^{9\text{-}1\text{-}1}$ can directly stimulate the Mec1^{ATR} kinase activity *in vitro* (Majka, Binz et al. 2006). Using *in vivo* budding yeast LacO systems, Bonnila *et al* (2008) showed that artificial co-localization of Mec1^{ATR}-Ddc2^{ATRIP} and Ddc1-Rad17-Mec3$^{9\text{-}1\text{-}1}$ onto chromatin was sufficient to activate the checkpoint even in the absence of damage (Bonilla, Melo et al. 2008). Another *in vivo* study conducted by Navadgi-Patil and Burgers (2009) showed that residues W352 and W544 at the C-terminus of Ddc1^{Rad9} were sufficient to activate Mec1^{ATR} in both the G1 and G2 phases of the cell cycle (Navadgi-Patil and Burgers 2009). Whether the 9-1-1 complex can activate ATR-ATRIP activity in other organisms is still unclear. Instead, a conserved function of the 9-1-1 complex in checkpoint activation is its interaction with TopBP1 (Cimprich and Cortez 2008). Apart from its essential function in replication initiation, TopBP1 also functions in checkpoint activation.

TopBP1; the cornerstone of the ATR signaling cascade

The recruitment of the ATR-ATRIP complex is considered an independent event from the recruitment of 9-1-1 and TopBP1 (Melo, Cohen et al. 2001). This requirement for ATR activation has reasonably led Cimprich and Cortez (2008) to suggest that it might serve a functional importance. By having two independent sensor complexes being required for mounting a response, like a molecular version of the "two-man rule", cells ensure that the checkpoint will less likely be activated accidentally or by default. Alternatively, the functional importance of this independent recruitment might lie in an increased probability of ATR-ATRIP and 9-1-1 to co-localize to sites of perturbed replication as the length of the ssDNA tracts become longer (Cimprich and Cortez 2008). Or it could be that this mechanism provides flexibility for the activation of

different cellular responses when each is recruited alone (or with different partners), compared to when they co-localise. Indeed, ATM-mediated phosphorylation of Rad9 on S272 acts to promote damage repair in S and G2 (Shin, Yuan et al. 2012).

The TopBP1 protein and its homologs

TopBP1 was first identified in two independent screens in *S. pombe*. In the first screen it was identified as a radiation sensitive mutant *rad4* (Schupbach 1971). Interestingly, unlike most *rad* mutants it also displayed a temperature sensitive lethal phenotype indicating a function in an additional cellular pathway (Garcia, Furuya et al. 2005). More than a decade later, Hirano *et al* (1986) identified the same gene, *cut5*, in a *cut* mutant screen searching for mutations that block mitosis but allow cytokinesis such that the septum bisects the nucleus (hence the name "cells ultimately torn", *cut*) (Hirano, Funahashi et al. 1986). Saka *et al* (1993) cloned and sequenced *cut5* and showed that it was identical to *rad4*, which had in the meantime been cloned by Fenech *et al* (1991) and was already thought to play a role in DNA metabolism as judged by its sequence similarity to Xrcc1 (Fenech, Carr et al. 1991, Lehmann 1993). More importantly, Saka *et al* (1993) demonstrated that the function of the *rad4/cut5* was required both for the S-
M checkpoint and for the initiation of replication (Saka and Yanagida 1993). The essential role of this protein in replication initiation was strengthened by the identification of the *S. cerevisiae* homolog, $Dpb11^{TopBP1}$, as a high-copy-number suppressor of temperature sensitive mutants in the catalytic subunit as well as the DPB2 subunit of polymerase ε (Araki, Leem et al. 1995). Finally, the mammalian homolog of Rad4/Dpb11, namely TopBP1, was identified in a two-hybrid system that searched for interactors of the Topoisomerase IIβ protein and was shown to be required for cell survival (Yamane, Kawabata et al. 1997). Since then, homologs of this protein have been found in a plethora of species.

TopBP1 and its homologs are multiple BRCT domain proteins and a schematic representation of Rad4, Dpb11, *Xenopus* TopBP1 (XTopBP1) and human TopBP1 (hTopBP1) is shown in **Figure 1.6**. The BRCT domain is a folding unit that is approximately 95 residues long and consists of a four-stranded parallel β-sheet surrounded by three α-helices: two α-helices (α1 and α3) on one side and one α-helix (α2) on the other side of the β-sheet. The BRCT domains are named after the breast

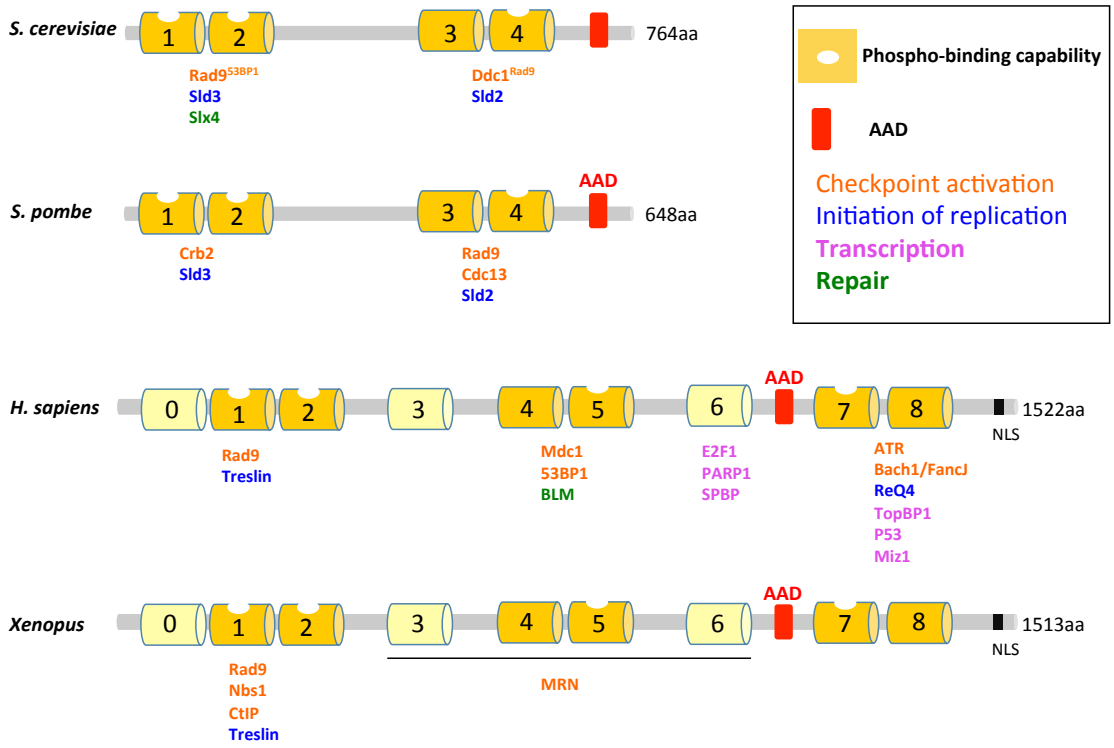


Figure 1.6: **Diagrammatic representation of TopBP1 and its homologs in yeasts, *Xenopus* and human**. BRCT domains are represented by boxes and known interactors are colour-coded according to their functionality. The position of the ATR activating domain within the C-terminus is also shown in red.

cancer associated protein 1 (BRCA1) C-terminus as the conserved globular domain they contain was first identified in BRCA1 (Lindsay A. Matthews and Alba Guarne, 2013, (Watts and Brissett 2010). The BRCT domain is very common among proteins involved in the DDR and mainly occurs as a singleton (single BRCT) or tandem pair (double BRCT) and often contains a phosphate-binding pocket (Manke, Lowery et al. 2003). Two types of single BRCT domains (sGroup I and sGroup II) and double BRCT domains (dGroup I and dGroup II) have been identified through phylogeny analysis and these four groups seem to differ in their phosphate-binding pockets. Evolution studies in eukaryotic organisms suggest that the phosphate-binding pocket changed from a DNA-binding type to a protein-binding site between the two types of single BRCT domains. Later in evolution a tandem duplication event of the protein-binding type domain gave birth to double BRCT domain. In all members of the first type of double BRCT domains, the phosphate-binding pocket is observed in BRCT a, but not in BRCT b and, unlike singletons, the main function of this pocket is binding to the phosphate of phosphopeptides. It is believed that the evolution of eukaryotic BRCT domains is associated with the evolution of the DNA damage response system. So in the early stages of evolution where the DNA damage response network of cells was less complicated, the BRCT-containing damage/repair proteins were targeted to the exact site of the damage directly via their BRCT domain(s). Thus, the BRCT domains had DNA-binding motifs. But as the DDR network became more sophisticated, more protein factors were targeted to the damaged sites, providing a more rapid and efficient way of mounting a response. And this may explain why the singletons changed their function from binding DNA to binding protein, such that they can transmit the signals generated by damage sensors to the repair machinery and other cellular processes (cell cycle checkpoint, apoptosis, transcription etc.) (Sheng, Zhao et al. 2011).

TopBP1 contains both Group I and Group II BRCT domains and its BRCT domains are essential for its ability to recruit client proteins and mediate protein-protein interactions. But the number of BRCT domains varies among the different species. The *S. pombe* Rad4^{TopBP1} and the *S. cerevisiae* Dpb11^{TopBP1} both contain four BRCT domains that come as two tandem repeats. The *C. elegans* and *D. melanogaster* homologs have 6 and 7 BRCT domains, respectively. The *Xenopus* and human TopBP1were originally shown to contain eight BRCT domains with BRCTs 1-2, 4-5 and 7-8 occurring as double domains whereas BRCT3 as a singleton (Garcia, Furuya et al. 2005, Wardlaw, Carr et al. 2014). More recently, an extra BRCT domain, BRCT0,

has been identified at the extreme N-terminus (Rappas, Oliver et al. 2011). Interestingly, studies on the crystal structure of TopBP1 have shown that not all the domains have the same phospo-binding ability; only BRCTs 1, 2, 5 and 7 seem to bind phosphorylated peptides. Interestingly, although all homologs function within the same cellular pathways, it overall seems that the similarity between the homologs is weak. TopBP1 shares 73% similarity (60% identity) with *Xenopus* XTopBP1. The similarity shared with *S. pombe* $Rad4^{TopBP1}$ and $Dpb11^{TopBP1}$ is concentrated over the regions containing the tandem BRCTs 1-2 and 4-5, which are conserved in all orthologs. $Rad4^{TopBP1}$ and $Dpb11^{TopBP1}$ share 38% similarity (24% identity) with each other (Garcia, Furuya et al. 2005, Wardlaw, Carr et al. 2014).

Perhaps the most interesting feature of the architecture of TopBP1 is that it contains an ATR Activation domain (AAD). In fission and budding yeasts this domain is located at the extreme C-terminus of the sequence whereas in *Xenopus* and humans it is found between the sixth and seventh BRCT domain (Kumagai, Lee et al. 2006, Mordes, Glick et al. 2008, Mordes, Nam et al. 2008, Lin, Wardlaw et al. 2012).

TopBP1 activates ATR

One of the biggest challenges in the DDR field has been the elucidation of the molecular mechanisms underlying ATR activation (**Figure 1.7**). The mode of activation of the ATR kinase seems to slightly differ between yeasts and mammals, possibly due to alterations in the protein's functions that have occurred in the course of evolution. The first evidence for the possible role of TopBP1 in the checkpoint response came from the observation that TopBP1 localised to sites of damage and replication stress (Makiniemi, Hillukkala et al. 2001). We now know that $TopBP1^{Rad4/Dpb11}$ is a key regulator of the ATR kinase activity, though the precise role it plays in different organisms and the extent to which it regulates the ATR pathway may slightly vary.

Perhaps the most interesting role of TopBP1 within the ATR signaling pathway is its ability to directly activate the ATR kinase. The AAD contacts a region in the C-terminus of ATR, located between the kinase and FATC domains, namely the PIKK Regulatory Domain (PRD). Mutating a conserved residue, K2598, within the PRD abolishes TopBP1-dependent activation of ATR. However, even in the absence of TopBP1-dependent ATR activation, ATR still possesses some basal kinase activity sufficient to promote the initial in *trans* phosphorylation of ATR on T1989. Thus PRD is not required for the basal activity of ATR but is required for its full activity both *in*

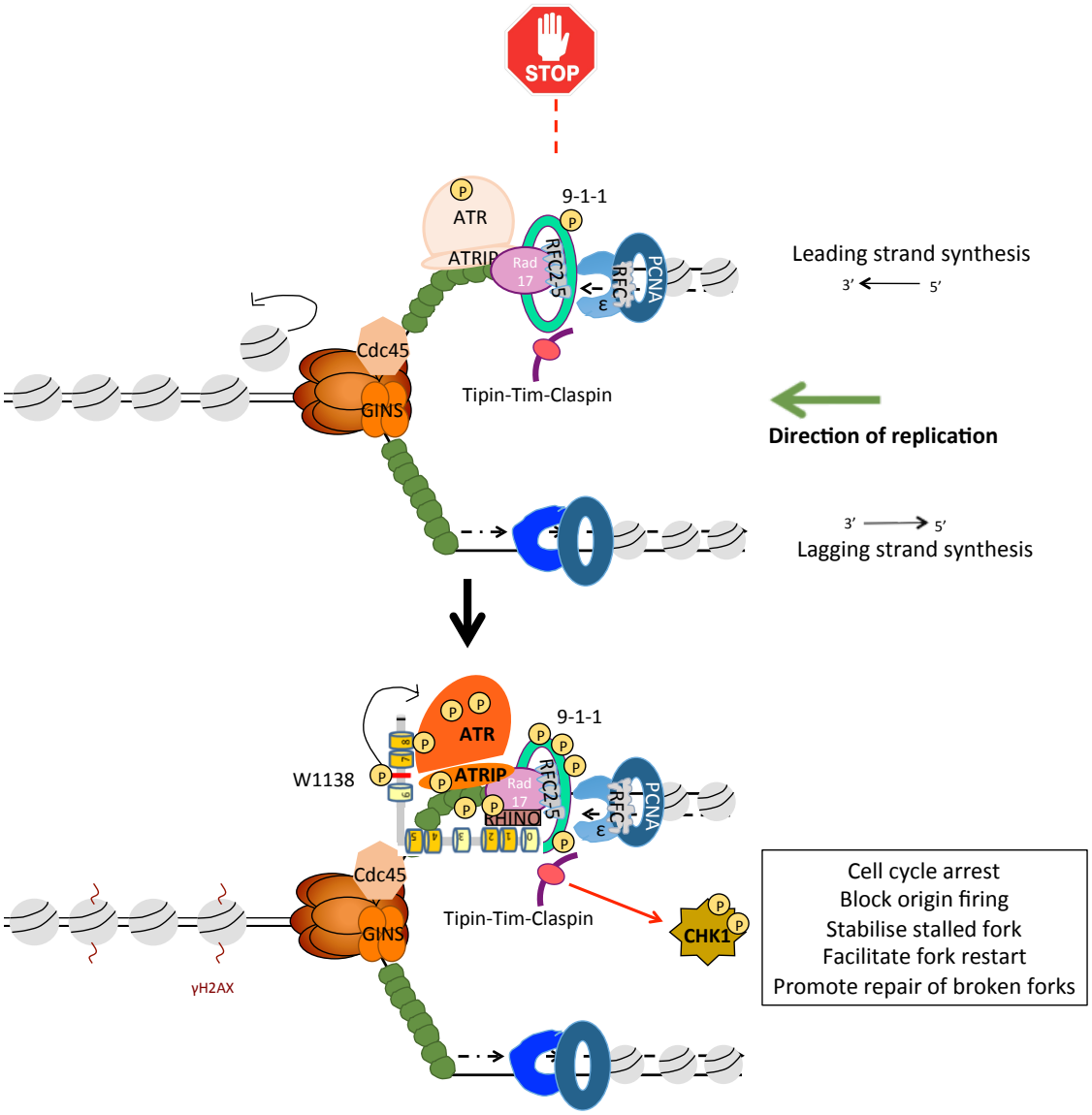


Figure 1.7: Model for the activation of the ATR checkpoint at perturbed replication forks. A) When a replication fork encounters a barrier, the polymerase and helicase activities may uncouple leading to the formation of long tracts of ssDNA in front of the stalled fork, which rapidly gets decorated with RPA. ssDNA formed at sites of replication stress recruits ATR-ATRIP and the 9-1-1 complex independently. **B)** The 9-1-1 complex in turn associates with and recruits the master ATR activator TopBP1. Rhino is also mediating the interaction between the 9-1-1 complex and TopBP1. ATR autophosphorylates and phosphorylates ATRIP which contacts the N-terminus of TopBP1. So TopBP1 acts like a bridge between the two binding partners, ATR and ATRIP. ATR itself is also interacting with TopBP1 via its autophosphorylation site and TopBP1 BRCTs 7-8. TopBP1 activates ATR through an interaction between its AAD and ATR's PRD. ATR then phosphorylates TopBP1, leading to further stimulation of ATR's kinase activity by TopBP1. Thus the interaction between ATR and TopBP1 provides a possible mechanism of signal auto-amplification. ATR is now competent to phosphorylate its downstream targets, thus fully activating the checkpoint. Whether TopBP1 can activate ATR at other stages of the cell cycle or after DNA damage awaits further characterization.

vitro and *in vivo*. Mutation of this domain in $ATR^{flox/-}$ cell lines causes defects in checkpoint activation and loss of viability (Mordes, Glick et al. 2008). Kumagai *et al* (2006) incubated recombinant TopBP1 with human and *Xenopus* ATR and showed that it can stimulate its kinase activity. The region of TopBP1 responsible for the activation was narrowed down to a region between the 6^{th} and 7^{th} BRCT domains (AAD) and the critical residue that mediates the interaction, W1138, was also identified. Mutation of this residue abolished this ectopic ATR activation and also caused a checkpoint defect in aphidicolin-treated egg extracts. What is more, it appears that the TopBP1/ATR/ATRIP interaction is transient and weak because TopBP1 separates from the kinase complex upon gel filtration and ATR returns to its basal kinase activity (Kumagai, Lee et al. 2006). The crucial function of the AAD in activating ATR is also evident from the fact that fusion of this domain to H2B histone or PCNA was sufficient to activate ATR in 9-1-1 knockout DT40 cells (Delacroix, Wagner et al. 2007) or Nbs1 knockout DT40 cells (Kobayashi, Hayashi et al. 2013). In addition, directly tethering TopBP1 to DNA is sufficient to induce ATR-mediated phosphorylation of Chk1 both *in vitro* and *in vivo*. Of interest is the synergistic activation of Chk1 when the mediator protein Claspin is also tethered to the DNA with TopBP1 (Lindsey-Boltz and Sancar 2011). A recent study in mice conducted by Zhou *et al* (2013) showed that the TopBP1 AAD is essential for embryonic development; a knock-in point mutant (W1147R) that ablates mouse TopBP1-AAD function causes early embryonic lethality, similar to an ATR knock out. In addition, AAD inactivation impaired cell proliferation, promoted premature senescence and compromised Chk1 activation following UV irradiation, highlighting the crucial role of TopBP1 as an ATR activator (Zhou, Liu et al. 2013). Finally, it has been reported that the TopBP1 AAD is capable of activating ATR and initiating checkpoint signaling even in the absence of damage. Persistent cell cycle arrest caused senescence (Toledo et al 2008).

Necessary for TopBP1's ability to activate ATR is yet another interaction, this time with ATR's binding partner, ATRIP. Mutating the domain of ATRIP that is involved in the interaction leads to HU sensitivity, defect in the G2/M checkpoint and reduced cell viability (Mordes, Glick et al. 2008). In their attempt to reconstitute the ATR checkpoint pathway, Choi *et al* (2010) showed that it is the N-terminus of TopBP1 that interacts with ATRIP but the details of this interaction are not yet known (Choi, Lindsey-Boltz et al. 2010).

A recently characterized protein called ETAA1 has also been suggested to have an ATR activating capability via a domain that harbors sequence similarity to the TopBP1 AAD (Bass, Luzwick et al. 2016, Haahr, Hoffmann et al. 2016). Authors suggest that ETAA1 represents a TopBP1-independent mechanism of activating ATR that is confined to stalled as opposed to collapsed replication forks. It is thus interpreted as a pathway parallel to TopBP1 that gives cells the flexibility to respond to replication stress in different ways according to the functional status of the perturbed chromatin region. However, the fact that depletion of ETAA1 has no discernable effect on CHK1 phosphorylation raises the question of whether this protein is only responsible for a small subset of ATR targets, such as RPA. Alternatively, it might be a mechanism specific to some cell lines or not conserved among all organisms, at least not to the same extent (Bass, Luzwick et al. 2016, Feng, Zhao et al. 2016, Haahr, Hoffmann et al. 2016, Lee, Zhou et al. 2016).

In the yeast model systems, the mode of ATR activation seems to involve more players. A functionally conserved AAD has been found in scDpb11^{TopBP1}. Despite the fact that the yeast TopBP1 has half the number of BRCT domains present in higher eukaryotes, studies in the budding yeast model have shown that Dpb11^{TopBP1} is able to bind and activate Mec1^{ATR}. Mordes *et al* (2008) showed that Dpb11^{TopBP1} interacts with Mec1^{ATR}-Ddc2^{ATRIP} and activates Mec1^{ATR} and similarly to higher eukaryotes this activation is dependent on Ddc2^{ATRIP}. This interaction is mediated via an AAD located within the C-terminus of Dpb11 and in particular the key residues that act as contact points are W700 and Y735. This AAD, however, does not have any sequence similarity with the mammalian AAD (Mordes, Nam et al. 2008, Navadgi-Patil and Burgers 2009). Ddc1^{Rad9}, as already mentioned, is also able to activate Mec1^{ATR} and this acts synergistically with Dpb11^{TopBP1} (Navadgi-Patil and Burgers 2008).

Interestingly, a third mode of activation of the Mec1^{ATR} kinase has been recently reported by Kumar and Burgers (2013). In particular, it has been shown that the unstructured N-terminal domain on the Dna2 nuclease is capable of stimulating Mec1 and that mutation of W128 and Y130 abrogates this function. Thus Dna2 shows partial redundancy for the replication checkpoint with checkpoint initiators Ddc1–Mec3–Rad17^{9-1-1} and Dpb11^{TopBP1}. A triple mutant that eliminates the checkpoint functions of all three initiators abrogates the Mec1^{ATR}-dependent checkpoint (Kumar and Burgers 2013). Finally, in an attempt to unravel the mechanistics of the Rad3^{ATR}/Rad4^{TopBP1} interaction in fission yeast, Lin et al (2012) identified and characterized an AAD within

Rad4^{TopBP1} that is crucial for activation of Rad3^{ATR} *in vivo*. AAD-defective strains were found to be DNA damage-sensitive during G1/S phases but not during G2, suggesting that this Rad4^{TopBP1} AAD is specific to G1/S but not evident in G2. Interestingly, it was suggested that *S. pombe* Rad4^{TopBP1} acts in a chromatin-dependent pathway to amplify the levels of activated Rad3^{ATR} in order to gain a full checkpoint response. This is predominant during the G1/S phases of the cell cycle where resection hence ssDNA, which is required for Rad3^{ATR} activation, is limited (Lin, Wardlaw et al. 2012).

Nonetheless, in all organisms studied to date, ATR activation is not exclusively mediated by a single protein. It is tempting to speculate that the existence of more than one AAD-containing proteins may dictate different modes of signaling through the ATR pathway in accordance with the morphology of the chromatin at the lesion.

Role of TopBP1 in ATM-mediated activation of ATR at DSBs

The initial study that identified the human TopBP1 homologue as a protein required for cell survival, has also characterized TopBP1 as a substrate of both the ATR and the ATM kinases (Yamane, Wu et al. 2002). Few years later, TopBP1 was proposed to have a role in a pathway that connects ATM to ATR at sites of DSBs (**Figure 1.8**). Yoo *et al* (2007) found that *Xenopus* ATM-catalyzed phosphorylation of TopBP1 on S1131 in the AAD is necessary for activation of ATR-ATRIP in response to DSBs, but not to replication blockages. The TopBP1 mutant unable to be phosphorylated on this residue displayed a defect in Chk1 phosphorylation following DNA damage but not replication stress. This favors a model whereby phosphorylation of this residue by ATM may mediate a handover from ATM to ATR checkpoint at DSBs by increasing the ability of TopBP1 to stimulate the kinase activity of ATR via the AAD (Yoo, Kumagai et al. 2007). The ATM-dependent phosphorylation on S1131 shows that there is a role of TopBP1 at DSBs and a quite intriguing question is whether TopBP1 can directly activate the ATM kinase by an as yet unidentified ATM activation domain, discussed later.

TopBP1 as a molecular scaffold of the ATR checkpoint

In addition to turning on the kinase activity of ATR, TopBP1 also acts as a molecular landing pad for checkpoint proteins. It brings a number of factors into close proximity at the site of damage thus promoting the propagation of the checkpoint signaling.

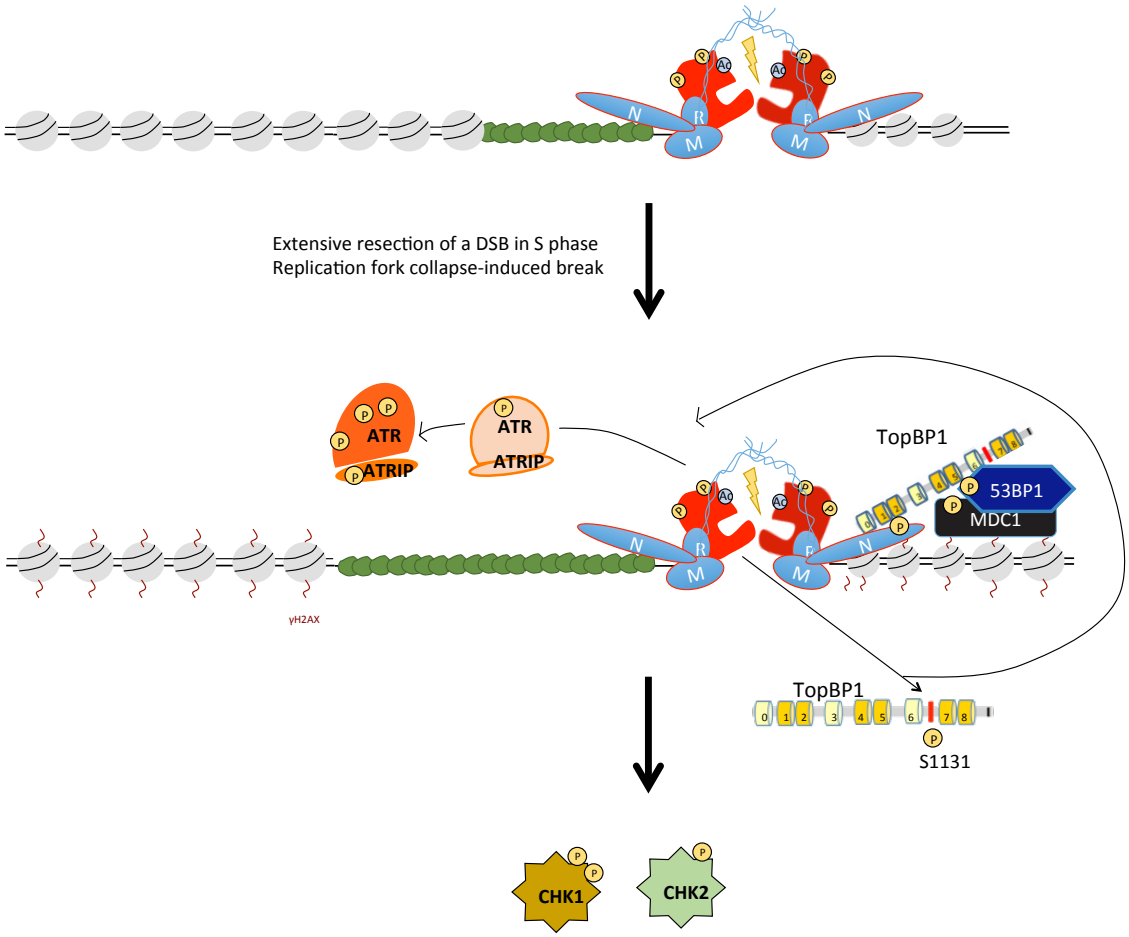


Figure 1.8: Handover from ATM to ATR of extensively resected DSBs. A) Activation of ATM at the site of DNA damage. **B)** Extensive resection of DSB leads to the formation of long tracts of RPA-coated ssDNA and the subsequent activation of the ATR kinase at the junction distal to the break. TopBP1 has a role to play in this handover from ATM to ATR as phosphorylation of its residue S1131 by ATM has been shown to activate ATR at regions of damage (see page 39 for details).

Efforts to unravel its mechanisms of recruitment revealed that TopBP1 is recruited to the lesion through interaction with phosphorylated Rad9. The $Rad4^{TopBP1}$-Rad9 interaction was first described in fission yeast and it was shown to occur between the $Rad4^{TopBP1}$ BRCTs3-4 and two damage induced phosphorylation events on Rad9, T412 and S423 (Furuya, Poitelea et al. 2004, Taricani and Wang 2006). This was similarly observed in *S. cerevisiae* where phosphorylation of a threonine residue on the tail of $Ddc1^{Rad9}$ by $Mec1^{ATR}$ allows its interaction with $Dpb11^{TopBP1}$ and recruitment of the mediator $Rad9^{53BP1}$ and the effector kinase $Rad53^{Chk2}$ (Puddu, Granata et al. 2008, Pfander and Diffley 2011). More recently, Germann *et al* showed that the interaction of $Dpb11^{TopBP1}$ with $Ddc1^{Rad9}$ is important for Dpb11 foci formation in response to damage in all phases of the cell cycle (Germann, Oestergaard et al. 2011).

In human cells, the interaction of TopBP1 with Rad9 seems to be constitutive. The important residue on Rad9 that contacts TopBP1 is S387 and it is constitutively phosphorylated by CK2. This suggests that the interaction between the two proteins may be independent of damage formation in human cells (St Onge, Besley et al. 2003). Consistent with this idea TopBP1 and Rad9 co-immunoprecipitate even in untreated cells (Greer, Besley et al. 2003). Rad9 S341 has also been found important for Rad9 tethering with TopBP1, hence for amplification of the checkpoint signal (Ueda, Takeishi et al. 2012, Ohashi, Takeishi et al. 2014). Rappas *et al* (2011) showed that it is the BRCTs 1-2 pair of TopBP1 that contacts Rad9 and mutational analysis of the phospho-binding pocket of each of these domains revealed that both domains provide binding surfaces. More specifically, BRCT1 is the primary binding site for S387 and BRCT2 plays a minor role in S387 binding but might also be involved in contacting another phosphorylated residue on Rad9, such as S341 (Rappas, Oliver et al. 2011). Interestingly, the ability of TopBP1 to bind Rad9 has switched from the second BRCT pair in yeast to the first one in humans.

Although the Rad9-TopBP1 interaction provides a mechanism of how TopBP1 is tethered to the lesion, the fact that they are recruited independently of each other leaves us with the question of what is the initial sensing mechanism of TopBP1. Using *Xenopus* extracts, Acevedo et al (2016) demonstrated that TopBP1 associates with sites of replication stress through a direct interaction with ssDNA-RPA via BRCT2 (BRCTs4-5 also interacted weakly) (Acevedo, Yan et al. 2016). TopBP1 has also been shown to bind damaged DNA *in vitro* through its C-terminus and this was sufficient to activate ATR (Yamane and Tsuruo 1999, Choi, Lindsey-Boltz et al. 2009). Another

protein able to associate with both 9-1-1 and TopBP1 is RHINO (Rad9, Rad1, Hus1 interacting nuclear orphan, encoded by RHNO1). It is not yet known what the binding surfaces that mediate these protein-protein interactions are and their functional importance in the ATR checkpoint is also unknown. One possibility is that it acts like a bridge between the checkpoint clamp and TopBP1. Alternatively, it might promote prolonged retention of TopBP1 to the site of the lesion or even enhance the ability of TopBP1 to activate ATR, thus providing another possible mechanism for signal amplification (Cotta-Ramusino, McDonald et al. 2011).

The C-terminal BRCT domains present in higher eukaryotes, BRCTs 7-8, were also identified as functional players in the activation of the checkpoint since their deletion led to a checkpoint defect in *Xenopus* extracts (Yan, Lindsay et al. 2006). Subsequent studies have made it clear that BRCT pairs 1-2 and 7-8 are important during replication stress (Cescutti, Negrini et al. 2010, Yan and Willis 2013). The 7^{th} and 8^{th} BRCT domains of TopBP1 bind to the Fanconi Anemia helicase FancJ and this is also a phospho-specific interaction that occurs after replication stress and is required for the extension of RPA-coated ssDNA and hence the activation of the checkpoint (Gong, Kim et al. 2010, Leung, Gong et al. 2011). Finally, BRCTs7-8 have the ability to bind ATR on its T1989 residue that gets autophosphorylated when ATR is recruited to the RPA-rich site of the damage. This positive feedback interaction promotes the full activation of ATR via TopBP1 and the subsequent activation of the checkpoint apparatus (Liu, Shiotani et al. 2011).

The mediators take over for propagation of the ATR signaling to transducers

The 9-1-1-mediated recruitment of TopBP1 is a catalytic event for the activation of the ATR kinase and for the molecular assembly of the cascades' mediators and transducers. Once activated by the localisation of 9-1-1 and TopBP1 (and/or ETAA1) to the lesion, ATR is fully active and ready to phosphorylate its substrates leading to the recruitment of the mediators and the propagation of the checkpoint signaling.

The mediator of the ATR signaling cascade at sites of replication stress is $Claspin^{Mrc1}$. It was initially identified in *Xenopus* extracts by the Dunphy lab as a protein essential for Chk1 phosphorylation and replication stress-dependent cell cycle arrest (Kumagai and Dunphy 2000). Claspin is a key component of the replication fork and in response to replication stress is required to bring ATR into close proximity with its effector kinase Chk1. Binding of Claspin to Chk1 in *Xenopus* requires

phosphorylation of Claspin residues S864 and S895 but the kinase responsible for this has not been identified yet (Kumagai and Dunphy 2003). Crucial for the retention of Chk1 phosphorylation is the binding of Claspin to Rad17, which has been already phosphorylated by ATR (Bao, Tibbetts et al. 2001, Wang, Zou et al. 2006). Another fork-associated protein that is also promoting ATR-mediated phosphorylation of Chk1 is Timeless/Tipin. Claspin together with Timeless/Tipin form the Fork Protection Complex (FPC) that apart from supporting Chk1 phosphorylation at sites of perturbed replication it also associates with the fork during unperturbed S phase (Unsal-Kacmaz, Chastain et al. 2007).

The yeast homolog, $Mrc1^{Claspin}$, was similarly identified as a protein required for the recruitment of the effector kinases, $Cds1^{Chk2}$ (*S. pombe*) and $Rad53^{Chk2}$ (*S. cerevisiae*). This recruitment required $Mrc1^{Claspin}$ phosphorylation and was confined to S phase (Alcasabas, Osborn et al. 2001, Tanaka and Russell 2001). The *S. pombe* and *S. cerevisiae* $Mrc1^{Claspin}$ also form the FPC with $Tof1/Swi1^{Timeless}$ and $Swi3/Csm3^{Tipin}$ (Unsal-Kacmaz, Chastain et al. 2007). It is interesting to note here that the functions of the Chk1 and Chk2 kinases have swapped over the course of evolution from fungi to higher eukaryotes. In particular, although $Cds1^{Chk2}$ responds to replication stress in fission yeast, its mammalian homolog Chk2 responds mainly to DNA damage whereas Chk1 is the replication stress effector. In budding yeast, however, it seems that $Rad53^{Chk2}$ is responsible for the response to both types of damage, with budding yeast Chk1 providing a relatively minor backup role (Rhind and Russell 2000).

The Chk1 transducer can then be recruited and activated by ATR. In mammals, this activation involves the phosphorylation of two SQ sites, S317 and S345, both required for Chk1 activation and used as surrogate markers of ATR activation (Guo, Kumagai et al. 2000, Liu, Guntuku et al. 2000). These two SQ sites are conserved in the fission yeast but only S345 is required to stimulate Chk1 (Lopez-Girona, Tanaka et al. 2001). The ATR-mediated phosphorylation is also important for inducing a conformational change that relieves the intra-molecular auto-inhibitory state at which Chk1 molecules exist in undamaged conditions. Activated Chk1 is then released from the chromatin such that the phosphorylated Chk1 pools can only be detected in soluble fractions of cellular lysates (Smits 2006, Smits, Reaper et al. 2006).

Overall, ATR activated after replication stress at disturbed forks phosphorylates various substrates to coordinate cell cycle arrest, maintenance of replication fork stability, origin firing and restart of broken forks.

TopBP1 scaffolds the mediators at sites of damage

In addition to activating ATR and providing docking sites for the replication stress proteins, TopBP1 also helps recruit mediator proteins to sites of damage.

Initial evidence for the interaction of TopBP1 with damage mediators came from Saka *et al* (1997) who showed that $Crb2^{53BP1}$ is recruited to damage-associated $Rad4^{TopBP1}$. $Crb2^{53BP1}$ interacts with $Rad4^{TopBP1}$ via BRCTs1-2 (Saka, Esashi et al. 1997). Recruitment of $Crb2^{53BP1}$ to sites of damage and activation of the $Rad3^{ATR}$ checkpoint was also found to depend on interaction of $Rad4^{TopBP1}$ with Rad9 (Furuya, Poitelea et al. 2004, Taricani and Wang 2006). Three CDK phosphorylation sites on $Crb2^{53BP1}$ have recently been identified to be required for contacting $Rad4^{TopBP1}$. These sites are phosphorylated sequentially with the phosphorylation of two canonical sites, T215 and T235, bringing $Crb2^{53BP1}$ into close proximity to the $Rad4^{TopBP1}$-associated $Cdc2^{CDK}$, which is then able to phosphorylate the non-canonical T187. This leads to a strong interaction between BRCTs1-2 and two molecules of Crb2, an interaction required for the recruitment of the effector kinase and the activation of the DNA damage checkpoint in G2 (Qu, Rappas et al. 2013). This interaction is also conserved in the budding yeast where the same BRCT domains of $Dpb11^{TopBP1}$ bind to two phosphorylated residues on $Rad9^{53BP1}$. This is also required for bringing the effector kinase to the lesion but whether there are two molecules of $Rad9^{53BP1}$ binding to $Dpb11^{TopBP1}$ is not yet known (Pfander and Diffley 2011). Conserved throughout evolution, 53BP1 is also interacting with BRCT4-5 of TopBP1 in human cells and this seems to be required for full activation of the checkpoint during the G1 phase of the cell cycle. Whether TopBP1 activates the G1 checkpoint through direct activation of ATM, ATR or ATM-mediated activation of ATR is not known (Cescutti, Negrini et al. 2010).

BRCTs 4-5 are also important for the recruitment of TopBP1 to sites of damage and replication-associated stress and it was Yamane *et al* (2002) who first showed that TopBP1 focus formation relies on BRCT5 (Yamane, Wu et al. 2002). BRCT5 was later found to interact with the MDC1 mediator through the SDT repeat of MDC1 (Wang, Gong et al. 2011).

TopBP1 can also be recruited to sites of damage in an MRN-dependent manner and this has been shown in *Xenopus*. The TopBP1-Rad9 interaction, however, was still required to activate the ATR cascade effector kinase Chk1. The fragment of XTopBP1 required for the interaction was narrowed down to BRCTs 3-6 (Duursma, Driscoll et al. 2013). XTopBP1 has also been reported to bind MRN at sites of DSBs through

interaction with the Nbs1 subunit of MRN (Yoo, Kumagai et al. 2009). A more recent study suggests that it is the first pair of the BRCT domains of TopBP1 that contacts Nbs1 and that MDC1 mediates this interaction (Choi and Yoo 2016). Interestingly, Shiotani et al (2013) showed that Nbs1 mediates ATR activation on RPA-ssDNA in a TopBP1-dependent but Rad17-independent manner. Authors suggested a quite interesting model of a bi-phasic nature of DSB processing resulting to distinct molecular requirements. This idea has been around for many years and supported by *in vitro* studies (Shiotani and Zou 2009). Initial processing of the break exposes limited amount of RPA-ssDNA, decorated by ATR-ATRIP and juxtaposed to the Rad17/9-1-1-decorated ss-dsDNA junction. In this context, RPA is phosphorylated in a Rad17-dependent manner. But as resection proceeds and long tracts of RPA-ssDNA form, the Rad17/9-1-1 complexes become out of reach. Now, Nbs1 recognizes the distal portion of RPA-ssDNA and ATR-ATRIP is activated by recruitment of TopBP1 in an Nbs1-dependent but Rad17-independent manner (Yoo, Kumagai et al. 2009, Shiotani, Nguyen et al. 2013). In line with this model, depletion of both Nbs1 and Rad17 from *Xenopus* egg extracts abolishes phosphorylation of CHK1 in response to the polymerase inhibitor aphidicolin (Lee and Dunphy 2013). Finally, CtIP is another documented interactor of XTopBP1 and this is a damage-stimulated and MRN-dependent interaction, possibly involved in an ATM to ATR hand-over as well (Ramirez-Lugo, Yoo et al. 2011). Future studies will hopefully shed light on the details of this interaction and reveal whether it is occurring in organisms other than *Xenopus* too.

Other mediators of the ATR pathway that do not interact with TopBP1 also exist. One of the best-characterized substrates of ATR is H2AX, which gets phosphorylated by the kinase at sites of defective replication to yield γ-H2AX. This favors the recruitment of MDC1 and interacting TopBP1, the propagation of γ-H2AX to flanking chromatin and subsequently the recruitment of other ATR substrates (Wang, Gong et al. 2011). Alternatively, it might promote recruitment of ATM to the chromatin surrounding the stressed fork (Ward and Chen 2001).

ATR functions during unperturbed replication

Unlike ATM, ATR is essential for viability of proliferating cells. Germ-line inactivation of ATR or Chk1 results in early embryonic lethality in mice and genetic inactivation causes cell death in human cells, whereas ATM- and CHK2-knock out mice are viable (Brown and Baltimore 2000, de Klein, Muijtjens et al. 2000, Liu, Guntuku et

al. 2000, Takai, Tominaga et al. 2000, Brown and Baltimore 2003). Furthermore, ablation of ATR and CHK1 causes high levels of genomic instability in S phase (Casper, Nghiem et al. 2002, Syljuasen, Sorensen et al. 2005). Surprisingly, CHK1 can be genetically deleted in DT40 lymphoma cells although CHK1-deficient DT40 show severe defects in both the DNA damage and the replication checkpoints (Zachos, Rainey et al. 2003). This points to fundamental differences in the normal physiological functions of the ATM and ATR checkpoint pathways. ATR responds to ssDNA and this is a much more physiologically relevant DNA structure compared to DSBs, which activate ATM (Cimprich and Cortez 2008).

The essential function of ATR seems to be the stabilisation of replication forks even in undamaged cells. ATR phosphorylates various factors at the replication fork and at origins of replication and although the precise molecular significance of most of these phosphorylation events is not entirely understood, it is well accepted that they promote genomic stability during S phase (Cimprich and Cortez 2008, Flynn and Zou 2011). For example, ATR, Claspin, Rad9 and Hus1 were found to be required for proper regulation of CHK1 and CDC25 in a normal S phase (Sorensen, Syljuasen et al. 2004). Similarly, in fission yeast, histone H2A is phosphorylated in S phase at difficult-to-replicate regions such as ribosomal DNA repeats (Rozenzhak, Mejia-Ramirez et al. 2010). Interestingly, this basal level of the activated checkpoint in yeasts increases DNA damage tolerance (Tsaponina and Chabes 2013). ATR also acts to sustain fork integrity in situations where a moving replication fork runs into an actively transcribed region (Bermejo, Capra et al. 2011). Finally, ATR acts to maintain stability of common fragile sites (Casper, Nghiem et al. 2002) and to control activation of the mitotic CDK at the centromeric regions (Kramer, Mailand et al. 2004).

Cell cycle functions of the ATR cascade

The levels of ATR and Chk1 increase in late G1 as part of the E2F-dependent S-phase promoting transcriptional programme that allows the expression of genes required for S phase (Kastan and Bartek 2004). Unlike CHK2, CHK1 is a quite labile protein that is restricted to S and G2 phases and it remains active even in unperturbed cell cycles (Lukas, Bartkova et al. 2001).

Replication stress results to an enhancement of the action of the CHK1 kinase, rather than *de novo* activation as is the case of CHK2, simply because CHK1 is already operating in unperturbed S. More specifically, Chk1 remains largely inactive through an

intra-molecular auto-inhibitory mechanism. Phosphorylation of CHK1 by ATR on residues S317 and S345, however, relieves this inhibition via an induction of a conformational change as well as via release from chromatin as already mentioned. This allows CHK1 to interact with other protein, such as 14-3-3 (Jiang, Pereira et al. 2003, Kastan and Bartek 2004, Smits 2006). One of the targets of CHK1 includes the CDC25 proteins, leading to S phase arrest via CDC25 phosphorylation and inhibition of CDK2/Cyclin E (Sanchez, Wong et al. 1997). A CHK1-mediated phosphorylation and activation of Wee1 kinase has also been described (Lee, Kumagai et al. 2001). Furthermore, CHK1 functions in the S-M checkpoint to prevent mitotic entry in cells with under-replicated or damaged DNA (Brown and Baltimore 2003). Finally, in response to persisting damage in G2, CHK1 promotes degradation of CDC25A and 14-3-3-medited cytoplasmic sequestration of CDC25C, promoting in this way the inactivation of CDK1 and imposing a G2/M arrest (Peng, Graves et al. 1997, Sanchez, Wong et al. 1997, Nghiem, Park et al. 2001, Mailand, Podtelejnikov et al. 2002).

ATR functions at stalled replication forks

To avoid activating late-firing origins in the presence of DNA damage, cells have evolved mechanisms to regulate the complexes required for replication initiation. Phosphorylation of Treslin^{Sld3} by Chk1 prevents its interaction with TopBP1 (Boos, Sanchez-Pulido et al. 2011) and a similar mechanism has been described in budding yeast (Zegerman and Diffley 2010). Similarly Rad53 phosphorylates Sld3 inhibiting its association with MCM in addition to Dpb11 and Cdc45 binding (Deegan, Yeeles et al. 2016).

ATR also physically functions on the chromatin to stabilize stalled forks. This function of ATR is in large part an enhancement of the already physiologically operating mechanism of the ATR-CHK1 arm. Perhaps the most vital function of ATR on stalled replication forks is promoting their stable association with the replisome. This is thought to prevent fork breakage and to assist in replication resumption (Dimitrova and Gilbert 2000). Interestingly, several observations have led to the suggestion that primase and polymerase activities at stalled replication forks are required for the actual activation of the ATR checkpoint. Synthesis and elongation of *de novo* DNA primers on the single-stranded templates of stalled forks promote the formation of such DNA structures required for the loading of the 9-1-1 and recruitment of TopBP1. This might explain why the checkpoint machinery functions to avoid the dissociation of

polymerases from stalled forks (Recolin, van der Laan et al. 2014).

Although ATR signaling through Chk1 acts to reduce the overall rates of replication, interestingly, ATR can also promote dormant origin firing to rescue forks stalled at DNA lesions. It does so via a mechanism that acts independently of Chk1 and involves phosphorylation of the MCM2 subunit on S108 (Cortez, Glick et al. 2004, Yoo, Shevchenko et al. 2004) and subsequent recruitment of Polo-like kinase 1 (PLK1) (Trenz, Errico et al. 2008). This seems to create a biphasic situation characterized by inhibition of origin firing on a global scale and replication synthesis on a local scale, near the stalled fork (Cimprich and Cortez 2008, McIntosh and Blow 2012).

The replicative machinery also has the ability to re-prime DNA synthesis downstream of the lesion, leaving behind single-stranded DNA gaps that can be filled by the so-called "DNA damage tolerance" (DDT) pathways (also known as "Post-replication repair, PRR). This involves the use of specialized low fidelity translesion synthesis (TLS) polymerases to bypass the lesion, for repair after the completion of DNA replication. The functional link between DDT and the ATR signalling pathway remains elusive. However, studies in yeasts suggest that the ATR pathway and its components are important for the recruitment of TLS polymerases as well as for the ubiquitination of PCNA, a catalytic event for the regulation of DDT (Ghosal and Chen 2013).

As part of its role in promoting replication fork stability of stalled replication forks, ATR recruits factors like the RecQ family of helicases, which help clear DNA structures that could be deleterious for the cells. Timeless/Tipin are also ATR substrates that accumulate at stalled forks and promote the Claspin-mediated activation of CHK1 as well as the prevention of continued unwinding of the DNA by the replicative helicase, hence the dissociation of the RFC from the site of synthesis. Finally, in fission yeast $Rad3^{ATR}$ phosphorylates Mus81 leading to its dissociation from chromatin and this helps avoid fork collapse or DSB formation induced by direct cleavage of the DNA by the Mus81/Eme1 endonuclease complex. Overall, the various interactions of ATR at stalled forks and their importance in fork stabilization are a matter of ongoing research (Paulsen and Cimprich 2007, Ciccia and Elledge 2010).

Early work suggested that failure to resume DNA synthesis leads to fork collapse (Cobb et al, 2003, 2005; Katou et al, 2003; Lucca et al, 2004). A stalled fork largely retains the ability to resume replication as it is stabilized by the checkpoint machinery, whereas a collapsed fork is not stabilized and requires the assistance of

restart mechanisms to resume. Whether failure to activate the checkpoint always leads to fork collapse and whether this collapse event is always characterized by the dissociation of replisome components is not yet clear (Lambert and Carr, 2013). Recombination-mediated repair pathways are among the mechanisms that promote continuity of DNA replication and faithful inheritance of the genetic code.

HR plays a pivotal role in the repair and restart of stalled replication forks (McGlynn and Lloyd 2002, Lambert and Carr 2013). HR is described as gene conversion in the sense that the sequence of a particular locus is replaced by copying the sequence of a homologous locus (Haber 2000). ATR signaling at collapsed replication forks regulates repair by HR as many recombination proteins are ATR substrates (Ciccia and Elledge 2010). A characteristic example of HR factors regulated by ATR are the BRCA1 and BRCA2 proteins, involved in end resection and loading of the Rad51 nucleoprotein concomitantly with RPA displacement (Prakash, Zhang et al. 2015). Furthermore, ATR is involved in the pathway of Nucleotide Excision Repair (NER), which removes bulky helix-distorting DNA lesions as part of an oligonucleotide fragment, creating in this way short stretches of ssDNA (Marteijn, Lans et al. 2014). Finally, the Fanconi Anemia pathway of repair is also targeted by ATR and this is partly a TopBP1-mediated regulation, discussed in the next section.

ATM and ATR; more than just independent DDR regulators

A particularly interesting aspect of the signaling network of the DDR is the cross-talk between the ATM and ATR kinases themselves. Although they sense different DNA structures and they respond to these in unique ways, ATM and ATR have interdependent activities as well, as it might have already emerged from previous sections (Cimprich and Cortez 2008). In particular, ATM and MRN are required for ATR activation in response to DSBs, suggesting the conversion of a DSB to a ssDNA checkpoint substrate (Myers and Cortez 2006). When cells are deficient for ATM, ATR is required for Chk2 phosphorylation signaling in response to IR-induced DSBs (Wang, Redpath et al. 2006). Similarly, ATM signaling is activated at collapsed replication forks where DSBs are formed (Cimprich and Cortez 2008). What is more, evidence of the interplay between ATM and ATR comes from the overlapping in substrate specificity. As well as having their unique substrates, they can also phosphorylate the same substrates (Cimprich and Cortez 2008).

Of particular interest in the current research study is the function of TopBP1 within the ATM and ATR signaling pathways. Its role in scaffolding the checkpoints and in acting as a catalyst for the activation of the ATR kinase on both broken chromatin and ssDNA renders TopBP1 an integral component of the checkpoint pathways. Unraveling its mechanisms of action and regulation will shed light into fundamental questions concerning the DDR.

Other roles of TopBP1

TopBP1 has also been reported to have a role in regulating DNA repair. Depletion of TopBP1 leads to an increase in inter-sister chromatid exchange and a reduced level of DSB-induced HR. Depletion of TopBP1 also causes sensitivity to DSB-forming agents (Morishima, Sakamoto et al. 2007). One explanation for the role of TopBP1 in suppressing sister-chromatid exchange is its recently identified interaction with the BLM helicase during S phase. BLM, which is part of the BLM/TOP3A/RMI1/RMI2 Dissolvasome complex, was shown to interact with the BRCT5 domain of TopBP1 and this interaction seems to be important for suppressing origin firing, sister chromatid exchanges (SCEs) and generally preventing the formation of aberrant DNA structures within recombinational repair (Wang, Chen et al. 2013, Blackford, Nieminuszczy et al. 2015). Phosphorylated S304 of BLM seems to be the point of contact for TopBP1 BRCT5 but the exact mechanism by which TopBP1 regulates the activity of BLM remains elusive (Wang, Chen et al. 2013). Furthermore, TopBP1 is involved in the regulation of the Fanconi-Anemia (FA) pathway, a group of proteins catalyzing the repair of interstrand crosslinks that would otherwise block the replication fork bi-directionally (Ciccia and Elledge 2010). A key player within this pathways is yet another helicase, FANCJ (also known as BACH1 or BRIP1), which is known to be phosphorylated after damage and replication stress (Peng, Litman et al. 2006). In particular, a phospho-specific interaction between FANCJ T1133 and TopBP1 BRCTs 7-8 was required for the loading of RPA onto chromatin following HU treatment suggesting a role in the activation of the replication stress checkpoint (Gong, Kim et al. 2010, Leung, Gong et al. 2011).

TopBP1 also emerges as a key regulator of other HR-mediated repair. Moudry *et al* (2016) found that depletion of TopBP1 sensitized cells to olaparib, a drug known to hyper sensitize cells already defective in *bona fide* HR factors. In an effort to understand the potential role of TopBP1 in HR, authors observed that TopBP1 promotes

the loading of Rad51. Although BRCTs 7-8 were found to be sufficient for this function, the exact mechanism remains unkown (Moudry, Watanabe et al. 2016). Other suggested interactions of TopBP1, which are of relevance to recombinational repair include the one with BRCA1 (Greenberg, Sobhian et al. 2006) and PARP1 (Wollmann, Schmidt et al. 2007). Such interactions, however, remain to be validated and their functional importance identified. Interestingly, TopBP1 has emerged as one of the minimal requirements for ATR-mediated repair of NER DNA substrates in an *in vitro* system developed by the Sancar laboratory (Lindsey-Boltz, Kemp et al. 2014). Finally, TopBP1 has been described as a key player in V(D)J recombination during B- and T-cell development in the immune system (Kim, Lee et al. 2014).

In yeast, the interaction of $Dpb11^{TopBP1}$ with the repair scaffold Slx4 provides another example of the various possibilities by which TopBP1 functions in DNA repair pathways. CDK-mediated phosphorylation of Slx4 on S486 enables binding to $Dpb11^{TopBP1}$ BRCTs 1-2 thus competing with the binding of the $Rad9^{53BP1}$ checkpoint adaptor. This was suggested to provide a mechanism of $Dpb11^{TopBP1}$ sequestration and dampening of the DNA damage signaling to allow room for repair and avoid persistent cell cycle arrest (Ohouo, Bastos de Oliveira et al. 2010, Ohouo, Bastos de Oliveira et al. 2013). The Slx4-Dpb11 interaction has more recently been described as being important in DNA end resection (Dibitetto, Ferrari et al. 2016) and resolution of repair intermediates (Gritenaite, Princz et al. 2014).

Furthermore, as an indirect way of exerting an effect on repair, TopBP1 seems to play a role in the recruitment of chromatin remodelers. A recent study from the Pfander laboratory has identified a novel interaction between the Fun30 nucleosome remodeler and BRCTs1-2 of $Dpb11^{TopBP1}$, which seems to be essential for the function of Fun30 in end resection (Bantele, Ferreira et al. 2017). Additionally, the BRG1 chromatin remodeler also involved in end resection has been suggested to be recruited in an Rb-dependent manner at DSBs, in what seems as a novel function of the Rb tumour suppressor at sites of damage. But localization of Rb depends on E2F1, which in turn interacts with TopBP1. Mutating the site of interaction of E2F1 with TopBP1 in mice impairs the recruitment of all these factors to the lesion and leads to checkpoint and repair defects (Velez-Cruz, Manickavinayaham et al. 2016).

In recent years, increasing evidence identifies TopBP1 as a key regulator of yet another cellular process, that of mitosis. It is possible for cells to proceed into mitosis with regions of their genome being under-replicated simply because of an inefficiency

of the G2-M checkpoint to sense and allow the repair of some lesions. This poses a problem for chromosome segregation and if such structures are not repaired within M, then they persist to the daughter cells in G1 as nuclear bodies of the 53BP1 marker of DNA damage. Interestingly, work from the Lisby and Oestergaard laboratories has revealed that TopBP1 marks such aberrant chromatin structures and its depletion causes an increase in anaphase bridges, probably caused by persisting HJs between the sister chromatids (Germann, Schramke et al. 2014). In fact, it was shown that TopBP1 promotes the resolution of HJs by Slx4 but also facilitates DNA synthesis at under-replicated or unreplicated regions in mitosis (Pedersen, Kruse et al. 2015), independently of translesion polymerases (Gallina, Christiansen et al. 2016).

TopBP1 has also been found to localize to mitotic centrosomes (Reini, Uitto et al. 2004, Bang, Kim et al. 2013) and centromeric UFBs (Broderick, Nieminuszczy et al. 2015). Centrosomes are crucial for the initiation of mitosis as it is the structure where activation of CDK1/CyclinB happens (Jackman, Lindon et al. 2003). At these regions TopBP1 co-localises with Topoisomerase IIα, thus presumably promoting the resolution of centromeric DNA catenations. Whether TopBP1 localises to centrosomes via its *in vitro* identified DNA-binding activity or via some protein-protein interaction remains unknown. Nonetheless, mutation of the phosphor-binding pocket of BRCT5 results to an increase in centromeric UFBs (Broderick, Nieminuszczy et al. 2015).

In addition to mitosis, TopBP1 is also emerging as a key regulator of transcription. Human TopBP1 has been shown to contain a transcription activation domain between residues 460 and 500 as well as two transcription repressor regions that map to BRCT2 and BRCT5 (Wright, Dornan et al. 2006). A number of interactions have also been identified concerning the role of TopBP1 in transcription control. Perhaps one of the most interesting interactions is the one with E2F-1, a member of the E2F family of transcription factors involved in the regulation of the cell cycle and, contrary to other members of its family, of apoptosis (Field, Tsai et al. 1996). The interaction of TopBP1 with E2F-1 inhibits the known E2F-1 activities but recruits E2F-1 to a BRCA1-containing repair complex, suggesting a direct role of E2F-1 in checkpoint/repair at stalled forks (Liu, Lin et al. 2003). The E2F-1-TopBP1 interaction has been found to occur at the G1-S boundary and after DNA damage. It presumably functions to prevent apoptosis during replication in the first case and to prevent replication of damaged genetic material in the second. For this interaction to occur, phosphorylated E2F-1 S31 interacts with BRCT6 of TopBP1 (Liu, Lin et al. 2003, Liu,

Luo et al. 2004). This interaction is also mediated by an AKT-dependent phosphorylation of S1159 on TopBP1, which leads to oligomerisation of TopBP1 via BRCTs7-8 and subsequent binding to E2F-1. Recent evidence suggests that oligomerisation of TopBP1 through BRCTs7-8 prevents its interaction with ATR and hence switches the role of TopBP1 from checkpoint activation to transcriptional regulation. AKT-mediated oligomerisation of TopBP1 seems to enable TopBP1 to interact with most of the transcription factors that are known to bind to it (Liu, Lin et al. 2003, Liu, Paik et al. 2006, Liu, Graves et al. 2013). Interestingly, TopBP1 has been reported to bind to the DNA-binding domain of p53 and repress its activity thus promoting growth and survival (Liu, Bellam et al. 2009). Finally, TopBP1 exerts an effect on the cell cycle machinery through its ability to interact with Miz1, a transcriptional activator of the p21 CDK inhibitor. In unperturbed cells, TopBP1 binds to and inhibits Miz1 whereas after UV irradiation it dissociates leading to cell cycle arrest (Herold, Wanzel et al. 2002). Perhaps more intriguing is the observation that depletion of TopBP1 in G1 led to an accumulation of p21 and p27, inactivation of the CDK2/Cyclin E complex and subsequent G1/S arrest. Co-depletion of p21, p27 and TopBP1 restored the levels of CDK2/Cyclin E but did not relieve the cell cycle block due to the additional role of TopBP1 in pre-IC assembly at origins (Jeon, Lee et al. 2007).

How to approach TopBP1 and what questions to ask

Conditional knock out of TopBP1 in mice results to early embryonic lethality and depletion in human cancer cell lines causes cellular apoptosis. TopBP1 depletion in mouse untransformed or human primary cells, however, leads to cellular senescence (Jeon, Ko et al. 2011). TopBP1 is also essential for neurogenesis as tissue specific deletion in the central nervous system (CNS) of mice leads to genomic instability and p53-mediated apoptosis in neuronal progenitors. Interestingly, authors observed that TopBP1 loss resulted to impaired neurogenesis more aggressively than ATR loss, pinpointing to a broader role in genome stability (Lee, Katyal et al. 2012).

TopBP1 has emerged as a central activator of the ATR kinase, a function that is of pivotal significance for the regulation of cellular proliferation under normal circumstances but also for the repair of potentially deleterious lesions on the genetic material. Like ATR and Chk1, TopBP1 is essential for progression of cells through the S phase, but to what extent this is attributed to its role in replication initiation or to its additional roles in cell cycle arrest, stabilization of stalled forks and restart/repair of broken forks, or indeed both remains enigmatic. The fact, nonetheless, that not that many mutations within *TopBP1* have been identified in cancer patients suggests that mutations affecting the protein's function, also impair the replicative potential of pre-cancerous lesions. TopBP1 thus functions as a crucial regulator of genome integrity. Its ability to dialogue with so many proteins downstream of the ATM and ATR kinases as well as with the kinases themselves renders it a central modulator in the fine-tuning of the numerous signaling routes encompassing the DDR and a key catalyst towards the repair of damaged chromatin and the suppression of tumorigenesis.

Interestingly signaling through the ATR-Chk1 arm can potentially lead to TopBP1-dependent repression of E2F-1- and p53-mediated apoptosis. So in addition to its checkpoint activation function, TopBP1 can also orchestrate a transcriptional regulatory programme (Akt-mediated) that suppresses checkpoint activation when growth-promoting signals are communicated to the cell. This function renders TopBP1 a global regulator of cellular proliferation in addition to its local activities at perturbed replication forks or sites of damage. It also renders TopBP1 an attractive target for cancer treatment. The progression of solid tumors is usually driven by a de-regulation of DDR pathways such as Rb, p53 and PI3K/Akt. And TopBP1 sits at the crossroad of all these oncogenic pathways, like a common modulator at the point of convergence. Could we perhaps target TopBP1 in cancer cells to activate apoptosis? Probably yes, as is the

case with other DDR factors. Chowdhury *et al* (2014) have identified Calcein as a molecule capable of blocking oligomerisation of TopBP1 hence inhibiting its binding to p53 and re-activating apoptosis in a spectrum of cancer cell lines over-expressing TopBP1 (Chowdhury, Lin et al. 2014).

Furthermore, though still contradictory, a number of studies have implicated aberrant expression of TopBP1 in breast cancer. Going *et al* (2007) analysed 12 samples of breast tissue from cosmetic breast reduction surgery; immunohistochemical analysis demonstrated that in a significant number of breast carcinomas TopBP1 was aberrantly expressed, as it was detected in the cytoplasm and nucleus of some tumors and exclusively in the cytoplasm of others. Liu *et al* on the other hand reported TopBP1 overexpression in the nucleus of primary breast cancer tissue-derived cells (Liu, Bellam et al. 2009). In another study, Forma *et al* (2013) examined the association between five single nucleotide polymorphisms located in the 3′UTR region of the TopBP1 gene and endometrial cancer risk and their results raise the possibility of TopBP1 being a susceptibility gene for endometrial cancer, though more studies are required to confirm this. Overall, whether TopBP1 can be used as a prognostic marker for breast cancer requires further investigation.

The question for now, I think, is not whether we can target TopBP1 to regulate its function but whether we truly know enough about this protein to be able to design the appropriate inhibitor molecules and predict with accuracy their effect firstly at a molecular level and then at the level of the cell/organism. Research in recent years has greatly expanded the information about TopBP1 and its multi-faceted nature in various cellular processes has been revealed. But we still lack answers to fundamental questions. For example; where is this protein localized throughout the cell cycle? How is its localization altered after DNA damage or replication stress? How is it regulated? How do the levels of TopBP1 affect its functions? What is the mechanism of activation of ATR by TopBP1? Can TopPB1 directly activate ATM? What is the ATM- and TopBP1-mediated mechanism of ATR activation at DSBs? Which of the *in vitro*-described interacting partners of TopBP1 happen *in vivo* as well? What are the *in vivo* physiological functions of TopBP1 protein-protein interactions? How are its pro- and anti-survival functions regulated? And many more.

Of course there are many unanswered questions. Experimental observations suggest levels of complexity beyond the skeletal framework of the ATM and ATR cascades. Especially critical will be the development of novel genetic systems that will

allow the analysis of separation of function point mutants of TopBP1 in genetically defined systems *in vivo*. Additionally, complete understanding of the underlying molecular mechanisms of action of TopBP1 or indeed any other protein will require high-resolution structural information.

TopBP1 is a multifunctional protein and unraveling in detail its plethora of effects on genome biology requires an understanding of every functional aspect of this protein in both stressed and unstressed cellular contexts. More importantly, a true mechanistic understanding of the antagonistic roles of TopBP1 in suppressing and supporting tumorigenesis will require new tools and persistent research to pull the puzzle apart.

Preface to Results

To better understand the roles of TopBP1 in DNA replication and the DDR, I have used reverse genetics to create genetically defined systems of TopBP1 in DT40 cells. The DT40 cells have rapidly gained a leading place in the field of cellular biology by providing an efficient and relatively easy model system to answer key biological questions by reverse genetics. DT40 is a transformed cell line derived from an avian leukosis virus (ALV)-induced bursal lymphoma. DT40 cells have a modal chromosome number of 80, with 11 autosomal chromosomes, the ZW sex chromosomes (thus making it a 'female') and 67 microchromosomes (Sonoda, Sasaki et al. 1998). This is two more chromosomes than generally found in the chicken (*Gallus gallus*) and is due to a trisomy of chromosome 2 and an additional microchromosome (Smith and Burt 1998, Sonoda, Sasaki et al. 1998). These pre-B lymphocytes have very high gene targeting frequency (around 80% efficiency), which makes entire gene deletions relatively easy. They also have a higher growth rate than mammalian cells, with the cell cycle lasting for about 8 hours. Importantly, DT40 cells have a very stable karyotype and have been successfully used to create isogenic mutants and assays by various laboratories working in the field of the DDR (Yamazoe, Sonoda et al. 2004, Hochegger, Dejsuphong et al. 2006). Finally, the DT40 lymphocytes are generally considered a cell line that does not express p53. This widespread idea is mainly based on a study conducted by Takao et al (1999) where it was shown that RT-PCR analysis using p53-specific primers produced a reaction product when RNA from chicken embryos, liver and testis were used but no product was generated from DT40 RNA (Takao, Kato et al. 1999). A few years later, however, a different group reported that Yin Yang 1 (YY1) negatively regulates p53 in DT40. In fact they showed that depletion of YY1 resulted to accumulation of p53 protein and proposed that YY1 exerts its function via direct physical interactions with both p53 and its regulator Mdm2 that promote ubiquitination and subsequent degradation of p53 (Sui, Affar el et al. 2004).

The work presented here describes my initial (unsuccessful) attempts to create a *TopBP1* knockout model system in DT40 cells. Part of this involved the development of a stably integrated overexpression system (SIOS) useful for the overexpression of the TopBP1 transgene that can be adapted to express any protein of interest. The next significant step towards achieving the experimental aim was the characterization of the *gallus gallus* TopBP1 RNA, which helped me identify previously unreported characteristics of its primary sequence and post-transcriptional control. This

characterization allowed the assembly of a novel TopBP1 cDNA that provided the essential function(s) missing from the annotated TopBP1 cDNA and which was necessary for the creation of a conditional *TopBP1* knockout model system in DT40 cells.

Furthermore, the creation of an isogenic set of stable cell lines with varying copies of TopBP1, allowed the study of the kinetics of the events induced by progressive loss of function of TopBP1. This work characterises the $TopBP1^{+/+/+}$, $TopBP1^{-/+/+}$ and $TopBP1^{-/-/+}$ cell systems for their replicative ability and cell cycle-dependent checkpoint proficiency. This data is necessary for the use of these systems as gene knock in platforms for the study of TOPBP1 mutants of interest. Additionally, to create an equivalent system for the study of human TopBP1, I generated a TopBP1 degron system in human RPE cells. This system allows the degradation of the endogenous protein and can be used to ectopically express mutants of interest.

Finally, application of the $TopBP1^{-/-/+}$ point mutation knock in system to the study of the TopBP1 AAD revealed various roles of this domain in checkpoint activation and signaling following replication stress and DNA damage.

Chapter 2

Materials and Methods

2.1 General Molecular Techniques

2.1.1 Gel Electrophoresis

Gels were poured at 0.8% agarose in 0.5x TBE containing 0.5µg/mg ethidium bromide (Sigma #E7637). Samples were mixed with loading dye and were loaded on the gel. The gel was run in 0.5× TBE at 100V and then visualised under UV light.

2.1.2 Nucleic Acid ethanol precipitation

2 volumes of 100% ethanol and 1/10 volume of 3M NaOAc were added to the DNA sample and the mixture was vortexed and placed on ice for 10 minutes. Samples were then centrifuged at 13000rpm for 15 minutes. The supernatant was removed and the samples were washed with 0.5ml 70% ethanol. The pellet was then air dried and then resuspended in the appropriate volume of 1× TE.

2.1.3 DNA Restriction Digests

Restriction digests were carried out using New England Biolabs (NEB) restriction enzymes or the Fermentas FastDigest™ enzymes according to the manufacturers recommended conditions. Restriction digested plasmid DNA fragments were gel purified using agarose gel electrophoresis and a Nucleospin clean up kit (Macherey Nagel, #740609.10).

2.1.4 Plasmid DNA Ligations

To set up DNA ligation mixtures, DNA concentrations of the insert and vector was measured using NanoDropTM 1000.

T4 Ligation

Restricted insert DNA was incubated with appropriate amount of restricted vector DNA and ligated using the T4 DNA Ligase (NEB, #M0202) and according to manufacturers' guidelines.

Infusion cloning

Restricted insert DNA was incubated with appropriate amount of restricted vector DNA using the In-Fusion® HD Cloning Plus kit (Clontech, #638910) according to the manufacturers' guidelines.

2.1.5 Oligos annealing

1µg of sense and anti-sense oligos were mixed to a final volume of 50µl with dH2O. The oligos were boiled for 5 minutes and then allowed to cool to room temperature slowly. The annealed oligos were stored at -20°C until use.

2.1.6 Removal of the 5' phosphate group from DNA ends

Antarctic Phosphatase (NEB, #M0289S) was used to remove the 5'-P group from DNA ends. To do this the linearized vector was mixed with Antarctic Phosphatase Reaction Buffer and Antarctic Phosphatase enzyme following the NEB guidelines and were incubated at 37°C for 30-60 minutes. The reaction was then stopped by incubating the samples at 80°C for 2 minutes.

2.1.7 TOPO cloning

PCR amplified products were cloned into the TOPO vector pCR® 4Blunt-TOPO (Amp/Kan).

First, PCR products were purified using the Gel Extraction method (See 2.1.13) and then cloned into the TOPO vector.

Reagents	Volume (µl)
Purified PCR Product	4
Salt	1
TOPO Vector	1
Total	6

The reaction mixture was incubated for 30-40 minutes at room temperature and then transformed into *E.coli* DH5α competent cells for amplification.

2.1.8 PCR for Molecular Cloning

For the amplification of plasmid DNA by the Polymerase Chain reaction method, KOD Hot Start DNA Polymerase (EMD Millipore) was used due to its high fidelity. The standard reaction was set up according to the manufacturers instructions. The annealing temperature was often adjusted between 50˚C and 58˚C depending on the primers used. PCR products were then purified using a Nucleospin clean up kit (Macherey Nagel, #740609.10)

2.1.9 PCR of DT40 genomic DNA

To PCR amplify genomic DNA LA Taq Polymerase (Clontech, #RR002A) was used, due to its high fidelity across long DNA templates. The standard reaction set up as well as the standard cycling conditions for the PCR reaction were as suggested by the manufacturer.

2.1.10 Site Directed Mutagenesis (SDM)

Site directed mutagenesis was used to insert DNA point mutations into a gene of interest, already subcloned in a plasmid. Overlapping forward and reverse primers of 20bp containing a point mutation were designed and used in the SDM PCR reaction. The PCR reaction was as follows: 50ng of plasmid DNA, 5µl of 10x PFU Turbo Buffer, 5µl of dNTPs at 2mM each, 1µl of PFU Turbo polymerase (Agilent, #600250), 0.25µM of primers and the volume made to 50µl with dH_2O. The standard cycling conditions for the SDM PCR were: 1x 94˚C for 3 minutes, 20x 94 for 30 seconds, 58˚C for 1 minute, 68˚C for 16.5 minutes. Followed by a final elongation step of 1 x 68˚C for 7 minutes. The annealing temperature may have been altered depending on the Tm of the primers and the elongation time was adjusted in accordance with the size of the plasmid. The template DNA was then digested using 1µl of Dpn1 (NEB) for 1 hour at 37˚C. The reaction was then cleaned up using a Nucleospin clean up kit (Macherey Nagel, #740609.10). All of the reaction was transformed into DH5α competent cells.

2.1.11 *E. coli* media

Luria-Bertani (LB)

10 g/l Tryptone

5 g/l Yeast Extract

10 g/l Sodium Chloride

Luria-Bertani Agar (LA/LB plates)

As LB plus:

12 g/l Agar

Drugs used for selection:

Antibiotic	Concentration	Stock
Kanamycin (Melford, #K0126)	35 μg/ml	35mg/ml
Ampicillin sodium salt (Sigma, #A95180)	100 μg/ml	100mg/ml

2.1.12 *E. coli* Transformation

Competent DH5α *E. coli* cells were thawed on ice. Plasmid DNA was mixed with the thawed cells and incubated on ice for 30 minutes. The DNA-cell mixture was heat shocked at 42˚C for 30 seconds and placed back on ice for 5 minutes. 1 ml of LB was added and transformed cells incubated for 60 minutes at 37˚C before being plated on to LB plates with the appropriate selection drug. Plates were then incubated at 37˚C overnight.

2.1.13 Extraction of Plasmid DNA from *E. coli*

For minipreps and maxipreps *E. coli* cells were inoculated in 5 ml or 100ml respectively of LB media containing Amplicillin at 100μg/ml or Kanamycin at 35μg/ml, and incubated over night at 37˚C. Cells were then pelleted at 4,600 rpm for 10 minutes at room temperature. The plasmid DNA was then extracted using a Qiagen Miniprep Kit (#27104) or a Midiprep Kit (#12145) according to the manufacturer's recommendations. The Plasmid DNA was resuspended in an appropriate volume of dH_2O and the concentration measured on a Nanodrop ND-1000 spectrophotometer.

2.1.14 Extraction of DNA from agarose gel

DNA sample of interest was run on 1% agarose gel (in 0.5% TBE) for 30min at 100V. The gel was then visualized using a UV transilluminator (350nm wavelength to avoid damage on DNA). The DNA fragment of interest was excised with a clean scalpel, transferred to a falcon tube and weighted. DNA was then isolated using the Nucleospin clean up kit (Macherey Nagel, #740609.10) in accordance with the manufacturer's instructions. The DNA was finally checked by gel electrophoresis.

2.1.15 Gateway cloning

Gateway technology is a quick and efficient cloning technique based on the bacteriophage λ recombination system in which the phage λ integrates into the bacterial

chromosome. Recombination takes place between specific sites called the *att* sites in the presence of specific enzymes, called the clonase enzymes, resulting in the formation of hybrid DNA sequences. These enzymes recognize the *att* sites and mediate their in between recombination reaction and then attach the ends together. The Gateway technology is based on the same principle with some modifications and consists of three major steps (see methods).

BP recombination reaction

The BP reaction was carried out for sub-cloning the 5' arm and the 3' arm into the donor vectors pDONOR P4-P1R and pDONOR P2-P3R, respectively (Invitrogen). The reaction was performed according to the manufacturers' instructions.

LR recombination reaction

The LR reaction was carried out for sub-cloning the 5' arm and the 3' arm entry clones and the Puromycin-containing entry clone into the pDEST DTA-MLS vector (Invitrogen). The reaction was performed according to the manufacturers' instructions.

2.1.16 DNA Sequencing

The Source Bioscience Sanger sequencing service in Nottingham, UK carried out DNA sequencing for all the work presented in the thesis.

2.1.17 RNA analysis

RNA purification

The TRI Reagent (Sigma Aldrich) was used to isolate RNA from DT40 cells, according to the manufacturer's instructions. The purity of the RNA was assessed on a denaturing MOPS/ buffer/formaldehyde gel by gel electrophoresis. To prepare 1% agarose/formaldehyde gel containing 0.5µg/ml ethidium bromide, 20 ml of 5xMOPS (0.2M MOPS pH7, 0.05M sodium acetate, 0.005M EDTA pH8) were mixed with 72 ml nuclease-free water and 1g of agarose (molecular biology grade). The mixture was heated to boiling in a microwave and let to cool down to 55°C before adding 17.6 ml of 37% formaldehyde and 5 µl of 10mg/ml ethidium bromide. The gel was then poured and allowed to solidify. The RNA samples were mixed with RNA sample buffer* at a 1:2 ratio, heated to 60°C for 5 minutes and cooled on ice for 2min. Then 2 µl of RNA loading buffer** was added and samples loaded on the gel. Gel electrophoresis was performed at 100V and gel visualized shortwave UV (254nm).

*RNA sample buffer: 3.5 ml 37% formaldehyde, 2ml 5x MOPS and 10 ml deionized formamide (Dowex XG8 mix-bed resin was added to formamide and stirred at room temperature for 45min and filtered through Whatman filter paper)

**RNA loading buffer: 50% glycerol, 1mM EDTA, 0.4% bromophenol blue, 1mg/ml ethidium bromide.

5' RACE

RACE was performed with the SMARTER RACE 5'/3' Kit (Clontech) and according to the manufacturer's instructions.

2.2 Cell Biology

2.2.1 Tissue culture

2.2.1.1 Culturing suspension cells (DT40)

DT40 Basic Cell Culture Conditions

DT40 cells were cultured in culture flasks, petri dishes or in 6/24- well plates, whereas microtiter plates were used for transfection or subcloning. Cells were passaged by diluting 1:10-1:20 into fresh media every 1-2 days to maintain the cells in exponential growth phase.

The optimum culture condition for the cells is 39.5◦C with 5%CO2 and the growth media used is:

RPMi complete Media:

RPMi 1640 [-L]	500ml
FCS	50ml
Chicken Serum	5ml
L-Glutamine (200mM)	5ml
Penicillin (10000U/ml)/Streptomycin (10mg/ml)	5ml
500mM β-Mercaptoethanol	0.6ml

Drugs used for selection:

The following drugs and chemicals were added to the appropriate media in order to select for the cells expressing the appropriate genetic marker (table 2-1).

Antibiotic	Concentration	Stock
G418	2mg/ml	50mg/ml
Histidinol	1mg/ml	100mg/ml
Hygromycin B	2.5mg/ml	100mg/ml
Puromycin	0.5µg/ml	0.5mg/ml
Blasticidin S	20µg/ml	10mg/ml

2.2.1.2 Culturing adherent cells (RPE)

RPE Basic Cell Culture Conditions

RPE cells were cultured in culture flasks whereas microtiter plates were used for

subcloning. Once the cells had reached confluence an aliquot of the cells was transferred to a new flask to allow continued growth of the culture. The media was first aspirated and cells were washed with 10mls of pre-warmed PBS. 1ml of pre-warmed 0.25% trypsin in PE was added to the flask (the volume of trypsin depends on the size of the flask/dish, 1ml is sufficient for a T75 flask) and the trypsin solution was distributed evenly over the surface of the flask/dish. The flask was then returned to the incubator for 2-3 minutes until the cells had detached form the plate. Once detached, 10mls of fresh pre- warmed media was added to the cells to inactivate the trypsin. Cells were then centrifuged at 1500rpm for 5min and resuspended in fresh 10ml media. A 1ml aliquot of that was added to a flask containing fresh media to continue growth.

The optimum culture condition for RPE cells is 37◦C with 5%CO2 and the growth media used is:

DMEM/F-12 (Dulbecco's Modified Eagle Medium/Ham's F-12 + L-glutamine)	500ml
FCS	50ml
Penicillin (10000U/ml)/Streptomycin (10mg/ml)	5ml

Drugs used for selection:

The following drugs and chemicals were added to the appropriate media in order to select for the cells expressing the appropriate genetic marker.

Antibiotic	Concentration	Stock
G418	1mg/ml	50mg/ml
Zeocin	500µg /ml	100mg/ml

2.2.1.3 Cryogenic preservation of cell lines

For long-term storage of cells, log phase healthy cells were pelleted and resuspended in 90% FCS/10% DMSO and divided into 500µl aliquots in 1.5ml cryovials. Initial freezing was carried out in a -70°C freezer to give a cooling rate of 1°C/minute. Once a temperature of -70°C was reached the cells were transferred to storage in liquid nitrogen vapour phase tanks at -180°C. To wake up the cells the vials were quickly warmed up to 37°C by placing in a container of warm water. Once thawed the cells were added to pre-warmed media. RPE cells were instead resuspended in 10ml prewarmed growth media, pelleted by centrifugation to remove the DMSO and added to a flask containing fresh media. The following day the cells were passaged or the media was changed depending

on the confluence of the cells.

2.2.1.4 Stable targeted Transfection of DT40 cells

Linearization of the plasmid for transfection

20µg of the targeting vector were linearized with the appropriate restriction enzyme prior to transfection to increase the chances of both the resistance gene and the gene of interest being expressed from the preferred locus. The linear DNA was ethanol precipitated and resuspended in 50µl dH_2O under the tissue culture hood to avoid contamination.

Transfection of DT40 cells by electroporation

Stable transfection was performed to insert a construct into a targeted locus in DT40. 20 µg of the targeting construct plasmid was linearized with the appropriate restriction enzyme and purified by ethanol precipitation. 5-10×10^6 DT40 cells were centrifuged at 1500 rpm for 5 minutes at room temperature and the cell pellet was resuspended in 0.5ml chilled PBS carefully. 30µg of linearized and purified plasmid was mixed with the cell suspension. All the mixture was transferred to an electroporation cuvette (Biorad, #1652088) and incubated on ice for 10 minutes. Cell electroporation was performed in the Gene Pulser Xcell total system at 550V and 25µF. The cuvette was incubated on ice for 10 minutes. The transfected cell suspension was transferred to a flask containing 20ml fresh pre-warmed RPMi media and incubated at 37^{o}C overnight. Selection was then made with the appropriate antibiotic and cells were plated on four 96-well plates.

2.2.1.5 Stable non-targeted Transfection of DT40 cells

Stable non-targeted transfection of DT40 cells was performed in the same way as the targeted transfection except that cell electroporation was performed at 250V and 950µF.

2.2.1.6 Expansion of drug resistant DT40 clones

Following the selection process clones were visible by eye in the 96 well plates. Clones were inspected under the microscope to ensure only single colonies were isolated. So single colonies were resuspended and then transferred to 800µl of fresh media (plus selection) in a 24 well plate. When the cells had reached confluence 500µl of the cells was added to 4.5mls of fresh media (plus selection) in a 6 well plate. The cells were transferred to T25 flasks while still being kept under selection. An aliquot of the newly

created clones was frozen down as soon as possibly to prevent *in vitro* evolution.

2.2.1.7 Excision of floxed-DNA sequences by induction of Mer-Cre-Mer

10^5 cells transfected with floxed vectors were cultured in 1ml of chickem medium containing 2μM 4-hydroxytamoxifen (4-HT) for 24 or 48h, depending on the experiment. Cell were then subcloned with limiting dilution for final concentration of 30, 100, 300 and 1000 cells per 96-well plate. 6-8 days ater subcloning, single clones can be observed as single colonies on the bottom of the plate. To expand single clones, 10μl of stable transfectants were transferred into 1ml of chicken medium. To assess for successful excision of the drug resistance gene cassettes, duplicates were made if the transfectants in selection drug-containing medium.

2.2.1.8 Stable Transfection of RPE cells using the Neon®Transfection System

Logarithmically growing cells were trypsinized and $1x10^6$ cells were transfected with the Neon®Transfection System according to the manufacturers' instructions.

2.2.2 DT40 Cell Biology techniques

2.2.2.1 Extraction of Genomic DNA from DT40 cells

DT40 cells were centrifuged at 1500 rpm for 5 minutes at room temperature and the cell pellet was resuspended in 0.5ml Tail Buffer (50mM Tris-HCl pH 8.0, 100mM EDTA, 100mM NaCl, 1% SDS) containing 0.5mg/ml Proteinase K.. The mixture was incubated at 55oC overnight. 300μl of 5M NaCl was added for protein denaturation and the mixture was vortexed for 1 minute. The mixture was centrifuged at 13,000 rpm for 5 minutes and 700μl of supernatant containing genomic DNA (gDNA) was transferred to 1.5ml fresh microfuge tube. DNA was precipitated with 400μl isopropanol, vortexed for 1 minute and centrifuged at 13,000 rpm for 5 minutes. The supernatant was discarded and 1 ml chilled 70% ethanol was added to the pellet. The tube was vortexed to wash the pellet and then centrifuged at 13,000 rpm for 5 minutes. The supernatant was discarded and the pellet was dried at room temperature for 10-15 minutes and resuspended in 50μl of TE buffer pH 8.0. The pellet was dissolved by 20 minutes incubation in a 65oC water bath.

2.2.2.2 Southern blot

Genomic DNA was extracted using a standard ethanol extraction process as described above. Concentration of DNA was measured using a Nanodrop and appropriate amount of DNA was digested in a final volume of 200µl at 37°C. Digested DNA was isopropanol precipitated and resuspended in 20µl of dH_2O and loading buffer (30% glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) Xylene cyanol FF). The samples were then run on a long agarose gel of appropriate concentration (depending on the size of the fragment) in 1× TBE at 50V. The gel was then incubated for 20 minutes in depurinating solution (0.25M HCl) in a shaker. Then the gel was washed in denaturing solution (1.5M NaCl and 0.5M NaOH) for 30 minutes on a shaker. Then gel was washed in neutralizing solution (1M Tris and 1.5M NaCl). The gel was then transferred to a membrane employing 10× SSC buffer (1.5M NaCl, 0.15M sodium citrate pH 7) and capillary force over night. The membrane was then washed in 2× SSC buffer for 5 minutes on a shaker. The membrane was air dried on a piece of filter paper and then the DNA was cross-linked to the membrane using UV light at 1200J/m2. The membrane was stored at 4°C.

A specific probe for Southern blot analysis was generated by PCR amplification of a ≈500bp fragment of the locus of interest and gel extraction. For Hybridising probe to the membrane, first the membrane was washed in dH2O for 5 minutes. Then 80ml of preheated 65°C hybridising solution I (6× SSC, 1x Denhardt [100x: 2% Ficoll 400, 300mM NaCl, 2% polyvinylpyrrolidone, 2% BSA], 1% sarcosyl, 0.1% BSA) was added to the hydrated membrane in a tube. The tube then was placed in hybridising oven for one hour at 65°C. Meanwhile, 1µl of 50ng/µl probe was added to 44µl dH2O. The solution was boiled in a water bath for 5 minutes and then placed on ice. In the radioactivity room, the labelling mix and 5µl of 32P-αdCTP were added to the DNA and the mixture was incubated at 37°C for 15 minutes. The labelled probe was then spun in a pre-spun G50 column at 3000rpm for 1 minute and incubated at 100°C for 5 minutes. Then the mixture was kept on ice. Then the probe was added to 20ml preheated 65°C hybridising solution II (6× SSC, 1x Denhardt, 1% sarcosyl, 200µl 10mg/ml salmon sperm DNA). Then hybridising solution I was replaced with hybridising solution II and the tube was put back in the oven at 65°C over night. The next day the membrane was washed with 50ml preheated 65°C wash buffer I (2× SCC, 1% SDS) in the oven for 10 minutes and then with 450ml of wash buffer I on a shaker

for 15 minutes. In the following step the membrane was washed twice, each time with 500ml of 42°C buffer II (0.1× SSC, 0.1% SDS) on a shaker for 15 minutes. The membrane was then air dried on tissue and wrapped in cling film and placed in a phosphoimager cassette over night. The membrane was scanned to obtain the Southern blotting results.

2.2.2.3 Growth curves and doubling time

1 x 10^5 cells were seeded in pre warmed growth media and counted after 24h. The culture was split 1/10 to the same volume, left to grow for another 24h and counted again. The procedure was repeated for the specified number of days. The doubling time *d* was calculated using the growth rate:

$$\boldsymbol{N(t) = N(0) \times e^{gr \times t}} \quad \gg \quad \boldsymbol{gr = \frac{ln\frac{N(t)}{N(0)}}{t}} \quad \text{,where}$$

N(t) = the number of cells at time t
N(0) = the number of cells at time 0
gr = growth rate (amount of doubling in a unit of time)
t = time (usually in hours)
Then the doubling time *d* can be derived as:

$$d = \frac{ln2}{gr}$$

Viable and dead cells were distinguished by trypan blue exclusion (Fisher, #15250061) in a Countess™ automated cell counter (Life Technologies) and values represent the mean of three independent experiments.

2.2.2.4 DNA damaging treatments of DT40 cells

To monitor the response of DT40 cells to replication stress, the ribonucleotide reductase inhibitor hydroxyurea (HU, Sigma) was added to the desired number of cells in culture, for the desired length of time and at the specified concentration depending on the experiment. For washing the drug off the culture, cells were spun down at 1500rpm for 5min, HU-containing media aspirated and cell pellet resuspended in fresh pre-warmed growth media.

To investigate the response to replication stall at damaged bases, MMS (Sigma) was added to cells in culture as in the case of HU.

As a method of inducing DNA damage cells were treated with either γ-irradiation or

UV-C light. Cells were irradiated with γ rays using an Alcyon II Cobalt-60 Teletherapy Unit. Dose rates varied from 1-10Gy. Cells were irradiated directly in the media in the culture flask. Control cells were handled in parallel but were not exposed to the radiation.

Cells were treated with UV-C light using a Stratalinker 1800 (Agilent Technologies, Stockport, UK) containing 254nm bulbs set to deliver 5000μJcm^{-2} (50Jm^{-2}). Prior to the UV-C exposure the DT40 cells were first washed with PBS and resuspended in a small volume of 1%FCS/PBS. The cells were returned to complete growth media following irradiation.

2.2.2.5 Colony Formation Assay

In general, serially diluted cells were plated in triplicate onto 6-well plates or single plates with 5 ml/well of DT40 growth medium with 1.5% (w/v) methylcellulose (Sigma). To determine sensitivity to MMS (Sigma) or HU (Sigma), serially-diluted cells were plated in MMS or HU-containing methylcellulose plates. To test sensitivity to UV or γ-irradiation, the appropriate number of cells was first resuspended in 1% FCS/PBS (for UV) or normal growth media (for γ-irradiation) and irradiated before being plate on methyl cellulose-containing media. A 60Co γ-ray source and UVC (254 nm wavelength) were used for the respective treatments. Colonies at the bottom of the plated were counted by eye 7-10 days after plating. Percentage survival was determined relative to the number of colonies formed in the untreated control. The experiment was repeated at least three independent times.

2.2.2.6 Cell Titer Blue Viability assay

To test the sensitivity of specified cell lines to the ribonucleotide reductase inhibitor HU, cells were treated with HU in 1 ml of medium using 24-well plates and incubated at 39.5°C for 48 h. To analyse sensitivity to MMS, 1 × 106 cells in PBS containing 1% FCS were exposed to MMS for 1h at 39.5°C and 10μl of exposed cells were transferred to 1 ml of medium using 24-well plates and incubated at 39.5°C for 48 h. To investigate sensitivity to γ-irradiation, cells were irradiated using a 60Co γ-ray source and diluted to 104 cells /ml in 24-well plates and incubated at 39.5°C for 48 h. For sensitivity to UV light, 1 × 106 cells were suspended in 0.5 ml of 1% FCS/PBS in 6-well plates and irradiated with UVC (254 nm wavelength) and 10 μl of irradiated cells were transfer to 1 ml of medium using 24-well plates and incubated at 39.5° C for 48 h. Then, we transferred 100μl of medium containing the cells to 96-well plates and measured the amount of ATP using Alamar Blue (Thermofisher) according to the manufacturer's instructions. Luminescence was measured by Fluoroskan Ascent FL (Thermo Fisher Scientific Inc.,

Pittsburgh, PA). Percentage viability was determined relative to an untreated sample prepared in parallel. The experiment was repeated at least three independent times.

2.2.2.7 Mitotic index

Cells were collected, cytospun onto microscopy slides and fixed with 3 % paraformaldehyde/PBS for 10min. Slides were then washed with PBS and a drop of DAPI mounting media (Vector laboratories) was added on top of the fixed cells. By microscopy a minimum of 200 nuclei was scored for each mitotic index measurement and the experiment was repeated at least three times. In some cases, indicated cells were incubated with nocodazole with or without prior γ-irradiation. A decrease in the number of mitotic cells was taken as indicative of activation of the G2/M checkpoint. Cells were imaged/counted on a wide-field DeltaVision Olympus IX70 microscope or on a Nikon E400 microscope.

2.2.2.8 Flow Cytometry

Fixing cells

500µl of mid logarithmically growing cells were washed with PBS and fixed in 1ml ice-cold 70% ethanol/PBS while vortexing. This minimises the formation of clumps and ensures uniform fixing of the cells. Fixed cells were stored at 4°C overnight or at -20°C for a couple of hours to several weeks before further analysis.

DNA content analysis

Fixed cells were centrifuged at 1500rpm for 5min, fixing solution aspirated and cell pellet washed twice with 3% BSA/PBS. Then cells were resuspended in 500µl of 3% BSA/PBS containing 250 µg/ml Ribonuclease A (RNase) and 10µg/ml propidium iodide (PI) (Sigma #81845). Samples were left for one hour at room temperature in the dark or at 4°C overnight before analysis on a BD FACSCanto machine (BD) using the FL-A setting.

*S phase analysis*To monitor actively replicating cells, cells were tested for their ability to incorporate the EdU analogue into their DNA. Cells in culture were treated as required and 30min before harvest 10µM EdU was added. Cells were washed and fixed as above and then the Click-iT® EdU Alexa Fluor® 488 Imaging Kit was used as a detection reagent to gain access to the DNA. The cells were then washed twice with 3% BSA/PBS and stained for DNA content analysis as explained above. Analysis was performed on FacsCanto machine using the FL-A setting.

2.3 Biochemical Techniques

2.3.1 Purification of antibody from rabbit serum

5ml of GST-fusion strain was grown at 37°C to confluence before added to 1 litre LB-Amp and grown to OD595 ~0.6. IPTG was added to final concentration of 0.5mM and cells were grown overnight at 20°C. Cells were pelleted by centrifugation and resuspended in 35ml PBS plus 1 protease inhibitor tablet (Roche) and AEBSF. Cell mixture was sonicated on ice and then cleared by centrifugation at 20k rpm for 1h. In the meantime, 1ml of Glutathione Sepharose was washed with water to remove the ethanol and then with PBS to equilibrate it. Cell lysate was then added and resin/lysate incubated for 1h at 4°C with rolling/gentle agitation. Resin was subsequently washed three times with 30ml of PBS and twice with 10ml of 0.2M sodium borate pH9. Dimethylpimelimidate (DMP) was then added to 20mM taking into account the volume of beads only. Resin/GST-fusion was gently mixed for 1h at room temperature to cross-link. Cross-linking was terminated with the addition of 0.2M Tris-HCl pH8. Resin was then washed with 5ml of 0.1M glycine-HCl pH2.5 and twice with 10ml PBS.

For antibody purification, 9ml of serum were mixed with 1ml of 10x PBS and incubated with 1ml GST-linked resin and incubated at 4°C rolling for 4h. Serum was then allowed to elute and column was washed with 10ml PBS. A second wash was performed with 10ml TrisHCl pH7.5 250mM NaCl and a third was with 10ml TrisHCl pH7.5 750mM NaCl. Antibody was eluted with triethylamine pH11.5 and equilibrated with TrisHCl pH5 to adjust the pH. The eluted serum was re-applied to the column (after column was washed with PBS) and this process was repeated three times.

2.3.2 Whole Cell Protein Extracts

Cells were collected, washed in PBS and resuspended in 100µl of 1x sample buffer (see below) and boiled for 5 minutes and spun for 5 minutes at 13000 rpm before loading on SDS gel.

<u>1x TCA Sample Buffer:</u>

1 volume 4x SDS sample buffer

1 volume 1 M Tris, pH 8

2 volumes dH_2O

2.5% β-mercaptoethanol

4x SDS Sample Buffer:

250 mM Tris-base, pH6.8

20% Glycerol

0.004 g/ml (w/v) Bromophenol blue

0.08 g/ml (w/v) SDS

2.3.3 SDS PAGE and Immunostaining of Proteins (Western Blot)

Whole cell protein extracts were separated by sodium dodecyl sulphatepolyacrylamide gel electrophoresis (SDS-PAGE).

Gels were run in a BIORAD Mini-POTEAN TetraCell or a C.B.S Double or Triple-wide electrophoresis system in 1x SDS running buffer (0.025M Tris Base, 0.25M Glycine, 0.1% SDS) at 80 volts constant through the stacking gel (Table 2-3) and 100-120 volts through the separating gel. A prestained Protein Marker (NEB, #P7708 or #14208S) was run alongside the samples.

Resolving gel:

Component volumes (ml) per gel mold volume of 5ml				
Component	**6%**	**8%**	**10%**	**12%**
H_2O	2.6	2.3	1.9	1.6
30% acrylamide mix	1	1.3	1.7	2
1M Tris (pH6.8)	1.3	1.3	1.3	1.3
10% SDS	0.05	0.05	0.05	0.05
10% APS	0.05	0.05	0.05	0.05
Temed	0.004	0.003	0.002	0.002

Stacking gel:

Component volumes (ml) per gel mold volume of 5ml	
Component	**5%**
H_2O	3.4
30% acrylamide mix	0.83
1M Tris (pH6.8)	0.63
10% SDS	0.05
10% APS	0.05
Temed	0.005

Proteins were then transferred from the gel to a Nitrocellulose membrane (GE Healthcare, Nitrocellulose, Hybond, #RPN3032D) via wet transfer at 300mA constant, in 1x transfer buffer (20mM Tris base, 750mM Glycine, 20% (v/v) Methanol, 0.025 % (v /v) SDS).

The membrane was stained with Ponceau-S solution (0.2% (w/v) Ponceau S, 3% (w/v) TCA) to confirm protein transfer and allow accurate cutting of the membrane for immunostaining with different antibodies.

The membrane was blocked with 3% milk PBST (Marvel dried skimmed milk in phosphate buffered saline (PBS), 0.1% Tween (Sigma #P7949)) or 5% TBST/BSA (for phosphor-specific antibodies) for 1 hr at room temperature. The primary antibody was added to the blocking solution and incubated with the membrane for 1 hr at room temperature or 4˚C overnight whilst being gently shaken. The primary antibody was then washed off by 3x 5 minute washes in PBST or TBST. It was then incubated with the appropriate secondary antibody in for 1 hr at room temperature whilst being gently shaken. The secondary antibody was washed off via 3x 5minute washes. The bound antibody was then detected by chemilluminescence (ECL Plus Western Lightning, Perkin Elmer, #NEL104001EA) and exposed to GE Healthcare Hyperfilm ECL (#GZ28906837). The film was developed with a Xograph Imaging Systems Compact X4. Quantification of western blots, was carried out using the 'Analyse -> Gels' function in ImageJ (NIH).

Specifications of antibodies used in this thesis and dilution factors

Anti-TopBP1 rabbit polyclonal (raised by Eurogentec) 1:250

Mouse anti-Chk1 (Santa Cruz Biotechnology) 1:1000

Rabbit anti pS345 Chk1 (Cell Signalling Technology) 1:1000

Mouse anti-pS139 (γH2A.X) clone JBW301 (Millipore) 1:1000

Mouse anti-Tubulin (Sigma, T5168) 1:10,000

Mouse anti-GADPH (abcam) 1:10,000

Rabbit anti-Mouse HRP Rabbit polyclonal (DakoCytomation P0260) 1:2500

Swine anti-Rabbit HRP Rabbit polyclonal (DakoCytomation P0217) 1:2500

Anti-TopBP1 rabbit polyclonal (Bethyl laboratories) 1 :1000

Anti-myc (Merck) 1 :5000

2.3.4 Oligonucleotides list

P	SEQUENCE (5'-3')
1	GGGGACAACTTTGTATAGAAAAGTTGATCACTAGTTGTGAAATCAGTGG
2	GGGGACTGCTTTTTTGTACAAACTTGCCTTTAACTGAGAAGCAAACC
3	GGGGACAGCTTTCTTGTACAAAGTGGTGAGAAAGTTCCGTGCTTTTAATTTG
4	GGGACAACTTTGTATAATAAAGTTGAAAACAGATGAGCTATGAGAAGG
5	GCGTCGAGATGCTAGCGAGGCTCACCTGGACTTCATATCCTTTTGG
6	ACGAAGTTATGTCGACGGATGGGAGAGAAGACTGGGAAATATTG
7	TAAGCAGGCGCGCCTAGTTCTGGGACAGTTTGCTACCC
8	TGCTTACTCGAGGGTACCTCTTTCTATGCATTTTATCCCTACCA
9	CAAGCTGGCGCGGCCGCATGAAAGGCAGCAAGGAGGTGTTCTT
10	TTTAAACTGACCCGGGTCAGTGCATTCTGGATCGCTTGA
11	GGGGACAGCTTTCTTGTACAAAGTGGACAGAAGATGAGAAGGTAGATG
12	GGGACAACTTTGTATAATAAAGTTGTTTCTGGGATAATCCAGTCC
13	GATTGCCAGATGTAGACTCAGATGC
14	CAGTGTCCTCTTCAGAGTCTGGATC
15	AGGCTAGAGCCACCGGATCCATGACTGAGTATAAACCAACCGTGAGAC
16	CCCCAGAGTCCCGCGAATTCTCAGGCTCCAGGTTTTCTTGTCATAC
17	TTGGGTACCGGGCCCCCCCTCGAGGTCGACATGAAAGGCAGCAAGGAGGTG
18	TCTGATATCAAGCTTATCGATACCGTCGACTGACCCCCAGCTCACTACAAC
19	GAATCGATAGCGATAATCTAGAGCGGCCGCCGTTCTGAGCACCCTCTTCC
20	AAGCTGGAGCTCCACCGCGGTGGCGGCCGCGTGCATTCTGGATCGCTTGAC
21	GGTACCCGCATGAAAGGCAGCAAGGAGGTGTT
22	GGTACCTCAATGCATCCGGCTCCTTTTTACTCTGCTCATTTCTCCCGGTGCTTTTC
23	GGGCTATCGAAACTTAATTAAAGAACCAGCTGTGGAATGTGTGTC
24	CTGACTTGACTGGTTAATTAAGGTACCTCTTTCTATGCATTTTATCCCTACCA
25	ACCTTTTTGGCAGCGATCGGAGCTCACCTGGACTTCATATCCTTTTGG
26	ATGTCGGGAGCCGCGATCGATAACTTCGTATATAATACCATATACGAAGTTATG TCGAC
27	TAA GCA GCT AGC CTTCCTATGTTTTGCTTTACCATTCAC
28	TGCTTAGTCGACGATTTCTGCAGTACAGAATATTAGG
29	TAAGCAGGCGCGCCCAGTACTCCCACACAAGCAAG
30	GACCACCTGAAATCTATAGTGATACACCTCGAGTGCTTA
31	GGGCTATCGAAACTTAATTAAAGAACCAGCTGTGGAATGTGTGTC
32	CTGACTTGACTGGTTAATTAAGACCACCTGAAATCTATAGTGATACAC
33	ACCTTTTTGGCAGCGATCGCTTCCTATGTTTTGCTTTACCATTCAC
34	GGATCCATGGCCAAGCCTTTGTCTCAAGAA
35	TCTGGTTATGTGTGGGAGGGCTAAGAATTC
36	CCCCCTCGAGGTCGACATCACTAGTTGTGAAATCAGTGG
37	TATCGATACCGTCGACCCTTTAACTGAGAAGCAAACC
38	TAATCTAGAGCGGCCGCTGAGAAAGTTCCGTGCTTTTAATTTG
39	ACCGCGGTGGCGGCCGCAAAACAGATGAGCTATGAGAAGG
40	GCAAGACTTGTCAGCAACTTCG
41	GGGCAGAACTGAACAAATCTGGC

42	GGTACCCACCCGCATGAAAGGCAGCAAGGAGGTGTT
43	GGTACCTCAGTGCATTCTGGATCGCTTGA
44	GGGCTATCGAAACTTAATTAAAGAACCAGCTGTGGAATGTGTGTC
45	CTGACTTGACTGGTTAATTAACACAGGAAACAGCTATGACCATG
46	CGATAACTTCGTATATGGTATTATATACGAAGTTATCGAT
47	CGATAACTTCGTATATAATACCATATACGAAGTTATCGAT
48	CACCAACCGTAAGACCTCCTAGAA
49	AAACTTCTAGGAGGTCTTACGGTTC
50	CACCGGCTTGCCCACCTTCCATGC
51	AAACGCATGGAAGGTGGGCAAGCCC
52	CACCGGACTGGATTATCACAAAAG
53	AAACCTTTTGTGATAATCCAGTCC
54	CACCGAGATGCGATTAGTGTACTCT
55	AAACAGAGTACACTAATCGCATCTC
56	CACCGTTTAATGTTTGGTAACTAAA
57	AAACTTTAGTTACCAAACATTAAAC
58	CTGGATTATCACAAAAAAGAAAAGCTCCTACAG
59	CTGTAGGAGCTTTTCTTTTTTGTGATAATCCAG
60	ACGACCTAGAGTTCATGTATACTCGCATC
61	GATGCGAGTATACATGAACTCTAGGTCGT
62	CTCGCATCTACCCTTCAGCTACCAAACATTAAATG
63	CATTTAATGTTTGGTAGCTGAAGGGTAGATGCGAG
64	GACCTAGAGTACACGTATACGGCGCCTCAGCGGCATCA
65	AAGGGTAGATGCGAGTATACTCAGAAGAACTCGTCAAGAAGGCGATAGAAGG
66	GAACTTGGGACTGGATTATCAC
67	CTATCACAGTCACATTCAGGCTTTC
68	CTGACATGGAAATGCATCGGGTATACCGTTCGTATAATGTATGCTATACGAAGT TATAGAACCAGCTGTGGAATGTGTG
69	ATGCCACACCTTTGCTTGTATACCGTTCGTATAGCATACATTATACGAAGTTATG GTTATCGCTATCGATTCACACAAAAAACCAAC
70	GATAGTTCTCCAGCCACTCCTG
71	GTTGCTGACTAATTGAGATGCATGC
72	TCACATGGAGGAATTCAATGCCACAC
73	AGCTCTCGAATTCAAAGGAGGTACCATGTCCAGAAATGACAAAGAACC
74	ACCTCGAGGTAGATATCGCGGTACCTTAGTGTACTCTAGGTCGTTTGATTT
75	CCTTCTTGACGAGTTCTTCTGAGAATTCG
76	GCTTACACGTAGCTTGCACACATAG
77	CTCGACACACCCGCCAGCGGCCGCTGCCAAGCTTCCGAGCTCTCGAATTC
78	CTTGGGACACATCTCTGGGACTGCATGCAGTACAGGACTACCTG
79	CAGGTAGTCCTGTACTGCATGCAGTCCCAGAGATGTGTCCCAAG
80	ACAACTTCTAGAAACCTCGAGGTAGATATCGCGGTACC

Chapter 3

Strategies employed for the generation of a TopBP1 knockout in the DT40 model system

Chapter 3: Strategies employed for the generation of a TopBP1 knockout in the DT40 model system

The recent years have witnessed an increase in the amount of information on the role of TopBP1 within different aspects of the cell cycle of eukaryotic cells. The use of RNA depletion approaches in human cell lines and also various *in vitro* systems have expanded our knowledge about this protein. Between the two, however, there seems to be a gap. The available tools in molecular biology have imposed constraints in the breadth and depth of experiments on human or other cell lines and also some of the interesting data obtained from *in vitro* studies has not yet been examined *in vivo*. The lack of a defined system, amenable to genetic manipulation, has thus left a lot of questions about the role of TopBP1 within the DDR largely unanswered. This chapter describes the different approaches followed towards the creation of a novel TopBP1 knockout system in the DT40 model system. Such a system would enable the functional analysis of TopBP1 by reverse genetics and would serve to bridge the aforementioned gap in the TopBP1 field.

3.1: Strategy 1; creation of the *TopBP1*$^{flox/flox/+}$/*Ova*$^{CMVTopBP1/+/+}$ DT40 cell line and initial attempts to create the knock out

The central aim of this current piece of research was to delete the entire *TopBP1* gene, located on the reverse strand of chromosome II of DT40 cells (chr2:43,231,685-43,253,592 21, UCSC Genome Browser), so as to create a *TopBP1*$^{-/-/-}$ model system. The reason why I decided to target the entire locus rather than a portion of it as is usually the preferred method, was that although a partial deletion would in most probability disrupt the transcription of the gene of interest, it does not eliminate the possibility of an alternatively spliced fragment, perhaps non-detectable by the antibody against the protein on a western blot, still being expressed. The chicken *TopBP1* locus consists of 27 exons and is predicted to produce a protein of 1512 amino acids in length. Briefly, the targeting strategy employed for the generation of a *TopBP1*$^{-/-/-}$ cell line involved the use of "recombineering" and homologous integration to complete creation of DT40 cells in which all endogenous *TopBP1* alleles would be deleted and one "transgenic" allele introduced at an ectopic locus, the *Ovalbumin (Ova)* locus also located on chromosome II. The transgenic allele was a wild-type copy of the *TopBP1* cDNA flanked by lox sites, thus easily excisable by Cre recombinase. Because of the essential nature of the TopBP1 protein, the only requirement in this targeting strategy was that the introduction of the wild-type cDNA should precede the deletion of the third endogenous *TopBP1* allele, as shown in **Figure 3.1**. DT40 cells are trisomic for chromosome II, but this did not pose any problem, except for the fact that an additional round of gene targeting had to be performed. The resulting *TopBP1*$^{-/-/-}$ /*Ova*$^{WTTopBP1cDNA/+/+}$ could be used for the introduction of transgenic mutant versions of *TopBP1*, resistant to the action of Cre recombinase, and the mutant phenotype revealed following loss of the wild-type copy.

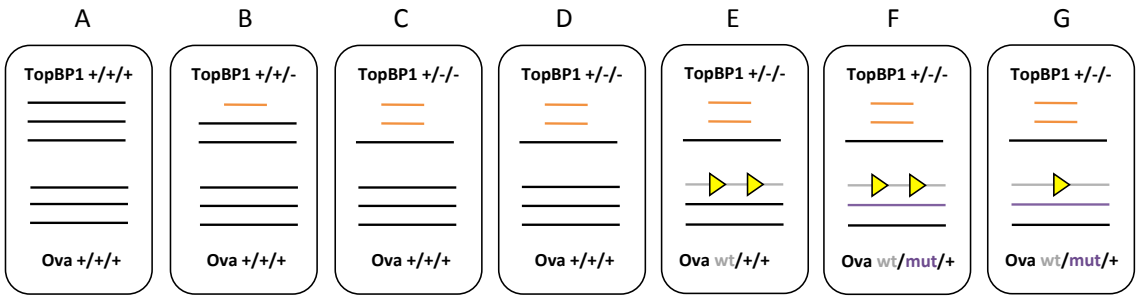


Figure 3.5: Diagrammatic representation of the TopBP1 knockout strategy. A) The *TopBP1* and *Ova* loci on the trisomic chromosome II of wild-type cells. **B**) Deletion of one *TopBP1* allele to create *TopBP1*$^{-/+/+}$. **C**) Deletion of the second *TopBP1* allele to create *TopBP1*$^{-/-/+}$. **D**) Integration of a wild-type copy of the *TopBP1* cDNA under the control of a constitutive promoter and flanked by lox sites in the *Ova* locus to create *TopBP1*$^{-/-/+}$/*Ova*$^{\text{WTTopBP1cDNA/+/+}}$. **E**) Deletion of the third *TopBP1* allele to create *TopBP1*$^{-/-/-}$/*Ova*$^{\text{WTTopBP1cDNA/+/+}}$. **F**) Integration of any mutant of interest of the *TopBP1* transgene into the second *Ova* allele. The mutant transgene is also under the control of a constitutive promoter but is not flanked by lox sites and also contains a different selection marker. **G**) Induction of Cre recombinase to flox the wild-type transgene and reveal the phenotype of the mutant transgene of the TopBP1 protein.

3.1.1 Deletion of two *TopBP1* alleles to create the *TopBP1*$^{flox/flox/+}$ cell line

3.1.1.1 Replacement of two *TopBP1* genomic alleles with selection markers to create the *TopBP1*$^{puro/his/+}$ cell line.

The method of gene targeting employed herein is based on the use of homologous recombination to delete the genomic fragment of interest with base-pair precision. This involves the assembly of homology arms that flank the region to be deleted within a targeting vector, which will be electroporated as a linearized DNA molecule to cells in culture. To enable successfully transfected cells to be identified, a selection cassette containing an antibiotic resistance marker under the expression of a viral promoter is cloned between the two homology arms. In essence, a successfully targeted clone of the cell population will have the gene of interest being replaced by the selectable marker cassette and will confer resistance to this particular antibiotic. To allow for recycling of selectable markers in future studies, such selection cassettes are flanked by lox sites so they can easily be removed by expression of the Cre recombinase. Overall, this so-called recombineering technique allows the introduction of targeting vectors that act as homologous substrates for the HR machinery of the cell to create genetic changes.

The first step of our gene targeting strategy was the deletion of two out of the three *TopBP1* alleles of DT40 cells. Construction of the appropriate gene-targeting vectors was performed using the method and vectors described by Iiizumi et al (2006) (and shared by the Hochegger laboratory) and the Multisite Gateway Cloning system from Invitrogen (see Materials and Methods). In particular, two sets of targeting constructs were created with homology arms flanking the gene of interest (designated as LA for left arm and RA for right arm), only differing in the selectable marker they contained. The selection cassettes were flanked by direct repeats of lox recombination sites making them amenable to Cre-mediated excision.

For generating these two targeting constructs (**Figure 3.2 A**), a 5' homology arm (4kb) and a 3' homology arm (2,003kb), flanking exons 1 to 27 of *TopBP1*, were selected. The sequence information was retrieved from the UCSC gallus gallus database. Different sets of primers (one pair for each arm) flanked by the *att* sites (sequence of *att* sites taken from Iiizumi et al (2006) were designed. The 5' homology arm (LA) was PCR amplified using primers P1 and P2 and the 3' homology arm (RA) was PCR amplified using primers P3 and P4. The LA and RA PCR products were then

gel purified and inserted into the pDONR P4-P1R and pDONR P2R-P3 donor vectors using BP recombination reaction of the Gateway technology *(Materials and Methods)*. The resulting LA and RA entry clones were confirmed by diagnostic enzyme digestion and Sanger sequencing. Subsequently, the LR recombination reaction was performed whereby the LA entry clone, RA entry clone, puromycin (or histidinol) entry clone and the pDEST DTA-MLS were all recombined together to produce the final gene targeting vector (Iiizumi, Nomura et al. 2006). In fact, two versions of the LA & RA-containing construct were obtained that contained different markers (Puromycin or Histidinol) for selection in DT40 cells. The final constructs were confirmed by PCR amplification of the arms, diagnostic restriction enzyme digestion and Sanger sequencing. For the sake of simplicity these constructs will be referred to as LARAPuro and LARAHis.

To generate a *TopBP1*$^{puro/+/+}$ cell line, the LARAPuro targeting vector was linearised by *PmeI* digestion (which linearizes the plasmid, cutting outside of the cloned sequences) and transfected by electroporation *(Materials and Methods)* into wild-type DT40 cells, kindly provided by the Hochegger laboratory. Out of 5 independent transfections, 62 puromycin (*puro*) resistant clones were obtained and of these only one was confirmed to be positive from Southern blot analysis with probe A, yielding 1.6% targeting efficiency. In more detail, genomic DNA was digested with *NcoI* restriction enzyme. In the wild-type locus probe A hybridizes to a 4.8kb fragment whereas in a successfully targeted locus the next *NcoI* restriction site is 10.5kb away (**Figure 3B C**). This positive clone was expanded and appropriately frozen down in aliquots as the *TopBP1*$^{puro/+/+}$ cell line.

The *TopBP1*$^{puro/+/+}$ clone was freshly woken up from liquid nitrogen stock and transfected with *PmeI*-linearised LARAHis targeting vector that contained the same combination of arms as LARAPuro (**Figure 3.2 A**). As a lower efficiency of successful targeting was expected, this time 10 rounds of transfections were performed and 150 histidinol (*his*) resistant clones screened by Southern blot analysis. Indeed, with two clones confirmed as positive, the targeting efficiency was at 1.3%. *NcoI* digestion of genomic DNA of successfully targeted clones creates a 3.7kb DNA fragment that is being recognized by probe A in addition to the already targeted and intact alleles (**Figure 3.2 D E**). The two positive clones were expanded and appropriately frozen down in aliquots as *TopBP1*$^{puro/his/+}$ clone1 and *TopBP1*$^{puro/his/+}$ clone2. Furthermore, to re-confirm the genotype of these positive clones, fresh genomic DNA was prepared and digested with *NcoI* as before or *SacI* enzymes. The probe A-hybridised fragments on the

blot were of the expected sizes as before (**Figure 3F left panel**). With *SacI* digestion of the wild-type locus a band of 5.9kb was expected whereas the *puro* and *his* targeted alleles were expected to give bands of 6.8kb and 8.5kb, respectively. The Southern blot results for *SacI* digests confirm the expected sizes (**Figure 3F right panel**).

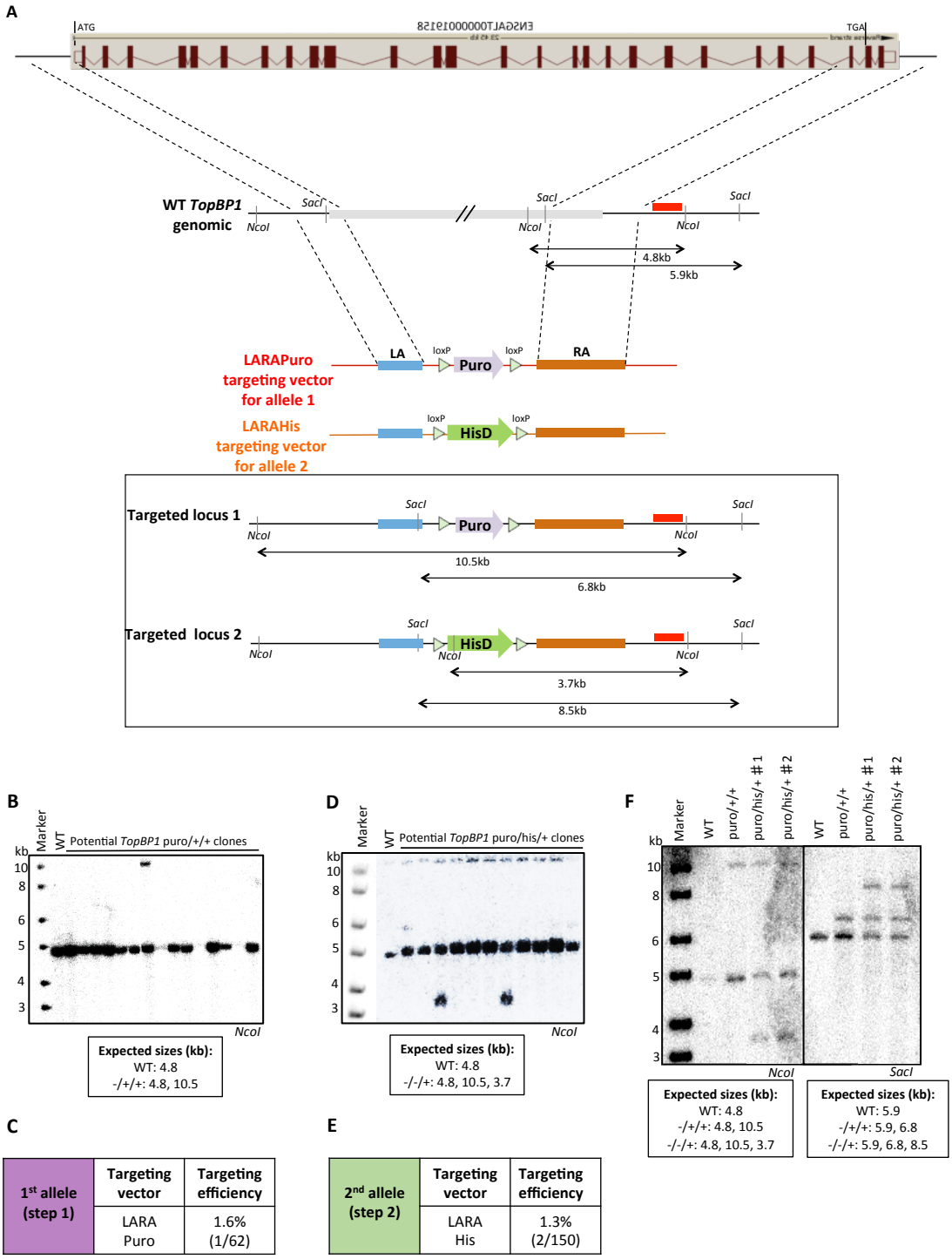


Figure 3.2: Targeted deletion of two *TopBP1* alleles and generation of the *TopBP1*$^{puro/+/+}$ and *TopBP1*$^{puro/his/+}$ cell lines. A) Schematic representation of the endogenous wild-type *TopBP1* locus, LARAPuro and LARAHis targeting vectors and successfully targeted allele. **B)** Southern blot analysis of wild type (WT) control (lane 2) and potential *TopBP1*$^{Puro/+/+}$ clones (10.5kb) digested with *NcoI*. **C)** Targeting efficiency of the transfections performed to obtain the TopBP1$^{puro/+/+}$ cell line. **D)** Southern blot analysis of wild type (WT) control (lane 2) and potential *TopBP1*$^{Puro/his/+}$ clones digested with *NcoI* (3.7kb). **E)** Targeting efficiency of the transfections performed to obtain the *TopBP1*$^{puro/his/+}$ cell line. **F)** Southern blot analysis of wild type (WT) control (lane 2) and *TopBP1*$^{puro/+/+}$ and *TopBP1*$^{puro/his/+}$ clone 1 and clone 2. *NcoI*-digested genomic DNA on the left panel and *SacI*-digested genomic DNA on the right panel. In all B, D, F digested genomic DNA was hybridized with the 3' external probe (red bar) shown in A.

3.1.1.2 Stable integration of Cre recombinase and removal of the selection cassettes from *TopBP1*$^{puro/his/+}$ cell line to create *TopBP1*$^{flox/flox/+}$

The next step of this strategy was to stably integrate the Cre recombinase enzyme in the genome to allow for excision of the selection marker cassettes or of any other lox-containing sequences valuable to future studies. To achieve this, the expression plasmid pANMerCreMer-*hygro* (provided from the Hochegger laboratory) encoding a tamoxifen-regulated chimeric Cre enzyme was linearised by *AhdI* digestion and introduced into the *TopBP1*$^{puro/his/+}$ clone 1 in a non-targeted way; i.e. randomly within the genome. The MerCreMer version of the Cre recombinase protein is inactive due to its retention in the endoplasmic reticulum in the absence of oestrogen derivatives (Zhang, Riesterer et al. 1996).

Transfected single clones were selected in the presence of hygromycin B and resistant clones were screened for inducible-Cre expression. Initially, Western blot analysis was used but due to the non-specific binding of the antibody (abcam #40011, data not shown), an alternative approach was applied: *TopBP1*$^{puro/his/+}$ cells were grown in 2μM 4-hydroxytamoxifen (4-HT) overnight to induce nuclear localization of the Cre enzyme. The selection cassettes integrated within the *TopBP1* loci of this cell line are flanked by direct repeats of lox recombination sites, easily excisable by induction of Cre recombinase. Cells were then serially diluted and plated on 96-well plates to isolate single colonies. To test for the success of the recombination reaction and the resulting deletion of the selection cassette, equal amount of 4-HT-treated clones were transferred to normal as well as puromycin- or histidinol-supplemented growth media and left to grow. Loss of the *puro* and *his* selection markers was expected to result in loss of viability when cells were grown in the presence of puromycin and histidinol, respectively. The untreated control culture served to assess the sensitivity of these cell lines to 4-HT but also as a healthy, untreated stock for future freezing down.

All clones screened had successfully floxed out the selection cassettes (**Figure 3.3 A**). To more directly confirm for successful integration of the Cre recombinase in the genome, two sets of primers were designed along the length of the pANMerCreMer-*hygro* expression vector and genomic DNA of *TopBP1*$^{flox/flox/+}$ clones 1-3 alongside the parental *TopBP1*$^{puro/his/+}$ were subjected to PCR amplification.

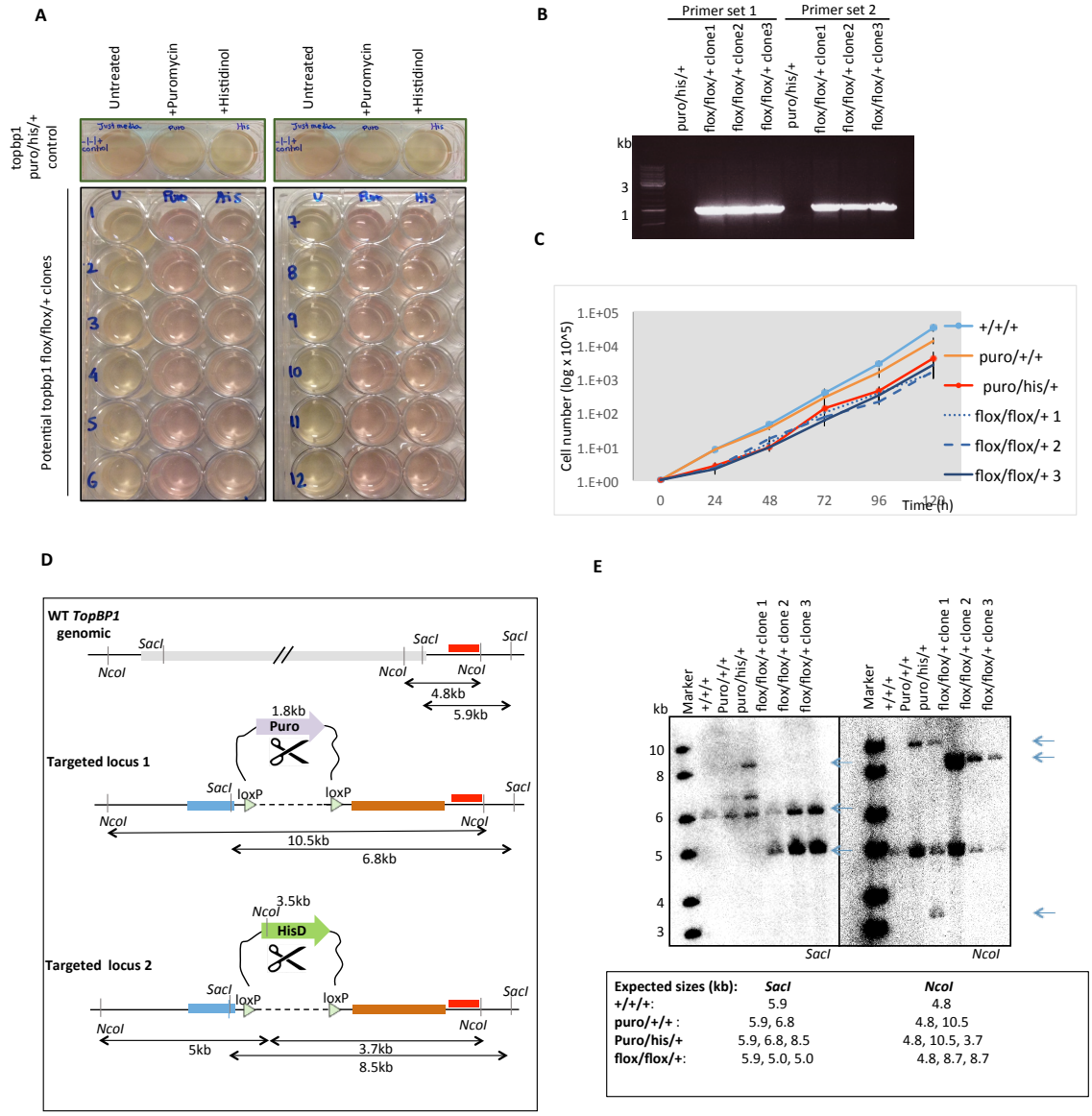


Figure 3.3: Stable non-targeted integration of the Cre recombinase and floxing of the selection markers from the *TopBP1*$^{puro/his/+}$ to create *TopBP1*$^{flox/flox/+}$ cell line. A) Following stable non-targeted integration of the pANMerCreMer-*hygro* vector into the *TopBP1*$^{puro/his/+}$ cell line, single isolated clones were tested for inducible Cre expression. 10µl of the single clones cultures were transferred to 6-wells containing normal media, 0.5µg/ml supplemented media or 1mg/ml histidinol suplemeted media. The top panel shows the parental cell line growing in all three conditions. The panels below show clones that have successfully induced Cre and have floxed the puro and his selection cassettes from the TopBP1 locus hence have become sensitive to puromycin and histidinol. **B)** PCR amplification across the pANMerCreMer-*hygro* vector from genomic DNA of the parental cell line (negative control) alongside three of the clones that have successfully integrated the Cre expression vector acccroding to A. **C)** Growth curves of wild type (*TopBP1*$^{+/+/+}$), *TopBP1*$^{puro/his/+}$ and *TopBP1*$^{flox/flox/+}$ clones 1, 2, 3. The error bars represent the standard deviation of the mean for three independent experiments. **D)** Diagrammatic representation of the wild-type, puro-targeted and his-targeted TopBP1 alleles showing the floxing of the selection cassettes and the expected sizes following *SacI* or *NcoI* digestion of genomic DNA. **E)** Southern blot analysis of *TopBP1*$^{+/+/+}$, *TopBP1*$^{puro/+/+}$, *TopBP1*$^{puro/his/+}$ and *TopBP1*$^{flox/flox/+}$ clones 1, 2, 3 following *SacI* or *NcoI* digestion of genomic DNA. In the left panel (*SacI),* diagnostic bands representing the puro (6.8kb) and his (8.5kb)-targeted *TopBP1* alleles as well as the puro-floxed (5kb) and his-floxed (5kb) *TopBP1* alleles are indicated. In the right panel (*NcoI),* diagnostic bands representing the puro (10.5kb) and his (3.7kb)-targeted *TopBP1* alleles as well as the puro-floxed (8.7kb) and his-floxed (8.7kb) *TopBP1* alleles are indicated. Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in D.

PCR products were run on agarose gel and results confirmed a single amplification band of the expected 1kb, as both primer sets amplified similar sized bands (**Figure 3.3 B)**. To chose which one of the *TopBP1*$^{flox/flox/+}$ clones 1-3 grows more similarly to the parental *TopBP1*$^{puro/his/+}$ cell line so that any potentially slow-growing clones would not be used for the subsequent steps the knockout strategy, a proliferation assay was performed to assess the proliferation rate of these clones. As shown in **Figure 3.3 C** they all proliferated with similar kinetics. Furthermore, to convincingly confirm the removal of the selection markers from the targeted *TopBP1* alleles, *TopBP1*$^{flox/flox/+}$ clones 1-3 were freshly woken-up, their DNA purified and digested with either *SacI* or *NcoI* enzymes as before. Successful removal of the *puro* and *his* cassettes (1.8kb- and 3.5kb-long, respectively) in *SacI*-digested samples was expected to shift the 6.8kb and 8.5kb bands down to 5kb and 5kb, respectively. Removal of the *puro* and *his* cassettes in *NcoI*-digested samples was expected to shift the 10.5kb and 3.7kb bands to 8.7kb and 8.7kb, respectively. The rest of the expected sizes have been explained previously (**Figure 3.3 B)**. Southern blot analysis with probe A did indeed confirm the expected results so the *TopBP1*$^{flox/flox/+}$ clones 1-3 were all expanded and frozen in aliquots for storage in liquid nitrogen. *TopBP1*$^{flox/flox/+}$ clone 1 only was used for subsequent experiments.

3.1.1.3 Stable integration of an ectopic transgenic copy of *TopBP1* to create *TopBP1*$^{flox/flox/+}$/*Ova*$^{CMVTopBP1/+/+}$

The next step of the knockout strategy was to introduce a wild-type copy of the *TopBP1* cDNA flanked by LoxM sites in the ovalbumin locus of chromosome II (Chromosome 2: 67,948,051-67,955,623 reverse strand, ENSGALT00000037195.1). The ovalbumin locus is repressed in the chicken B cells as this is a gene only expressed in oviduct cells in response to oestrogen (Buerstedde and Takeda 1991). Choosing a silent locus for the ectopic expression of the rescue construct is important for two reasons. Firstly, to avoid disrupting a housekeeping gene and secondly, to prevent potential effects on the expression of the transgenic allele by a local regulation of transcriptional activity. To generate a *TopBP1*$^{flox/flox/+}$ / *Ova*$^{CMVTopBP1/+/+}$ cell line, a targeting construct OVACMV-IRES had been designed by previous members of the Carr laboratory. To confirm the sequence of the construct before proceeding to the experiment, Sanger sequencing was performed. OVACMV-IRES contained the wild-

type cDNA of the TopBP1 protein under the control of the human cytomegalovirus (CMV) promoter (**Figure 3.4 A**).

The CMV promoter has been extensively used for the ectopic overexpression of transgenes and the substitution of endogenous genes in DT40 systems (Fukagawa, Mikami et al. 2001, Zimmermann, Ahrens et al. 2002, Johnston, Joglekar et al. 2010). Furthermore, transcription of the *TopBP1* transgene in our OVACMV-IRES targeting construct was coupled to transcription of the downstream neomycin selectable marker through the IRES (*Internal Ribosome Entry Site*), a sequence able to recruit the ribosome to the mRNA thus allowing cap-independent translation of the second gene. The IRES technology has also been exploited for the co-expression of genes in DT40 cells (Szuts, Simpson et al. 2006, Arakawa, Kudo et al. 2008). The entire CMV-*TopBP1*-IRES-*Neo* was flanked by direct repeats of LoxM recombination sites and at the extreme ends were the left and right ovalbumin homology arms, as shown in **Figure 3.4 A**.

The targeting vector was linearized by *PvuI* digestion and transfected into the *TopBP1*$^{flox/flox/+}$ cell line. 24h post-transfection, transfectants were microscopically observed and looked healthy. They were then selected with the standard concentration of 2mg G418 per ml of culture but no colonies were obtained in 12 transfections attempted. To try and assess if this was due to a potentially toxic concentration of the antibiotic, 16 transfections were performed and they were selected pairwise in variable concentrations of G418 (0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 and 2mg/ml) but still no colonies were observed. What is more, to eliminate the possibility of the cells being resistant to transfection, a wild-type cell line was also manipulated the same way but no colonies were obtained with OVACMV-IRES construct. As a positive control in the experiment, a plasmid containing a neomycin resistance gene under the control of the SV40 promoter successfully produced G418-resistant colonies at 2mg/ml G418 containing media. As an extra control, *TopBP1*$^{flox/flox/+}$ cells transfected with OVACMV-IRES and manipulated the same way but plated in G418 (-) media grew to confluence. From these experiments it was concluded that the IRES sequence did not work in our system and in the context of the CMV promoter. Thus a new strategy was devised for creating a *TopBP1*$^{flox/flox/+}$ cell line ectopically overexpressing wild-type TopBP1.

The new strategy was based on having the *TopBP1* transgene and the selection marker being independently transcribed from different promoters. A linear DNA fragment containing the CMV promoter, a poly A tail, the SV40 promoter, neomycin

gene and another polyA tail was synthesized from Genscript using the *EcoRV* cloning site of the pUC57 vector. To ensure that the expression system being developed would be functional, the nucleotide sequences of the aforementioned genetic elements were copied from widely used mammalian expression vectors. More specifically, the CMV promoter and the first polyA tail sequences were copied from the pCDNA3.1 vector (V790-20, Addgene) whereas the SV40-neomycin-polyA fragment was copied from pCI vector (E1841, Addgene). Unique cloning sites were created within the sequences of the different elements to allow for subcloning of the ovalbumin left arm (OVALA) at *NheI/SalI*, the ovalbumin right arm (OVARA) at *AscI/XhoI* and *TopBP1* cDNA at *NotI/XmaI*. The pairs of primers used for these three cloning steps were P5/P6, P7/P8 and P9/P10, respectively. Each cloning step was monitored by diagnostic digestion and after the final cloning step Sanger sequencing successfully confirmed the newly made OvaCMV expression vector (**Figure 3.4 A**).

OVACMV was linearized by *XhoI* digestion and transfected to the *TopBP1*$^{flox/flox/+}$ cell line. 4 independent transfections were performed and over 100 G418-resistant clones obtained at a selection of 2mg/ml G418. This confirmed that it was indeed the IRES sequence that was impairing the viability of the cells transfected with the OVACMV-IRES construct used before and allowed our gene knock out strategy to proceed. Of the 124 clones selected, 34 were initially screened by Southern blot analysis, as the *Ova* locus is well known for its high rate of integration. Genomic DNA from these 34 clones was purified and digested with *SphI* restriction enzyme for Southern blot analysis. A probe was also designed specific for the analysis of the *Ova* locus (probe B). *SphI* digestion of the wild-type *Ova* locus was expected to generate a 12.2kb band on the Southern blot after hybridization with probe B, which is indeed the case. Successful integration of the OVACMV expression vector was expected to produce a 7.6kb band. It is worth-mentioning that the extra band just above the 8kb denoted by an asterisk is a result of a polymorphism in one of the three *Ova* alleles.

Out of the 34 samples tested, 14 seem to have been successfully targeted according to **Figure 3.4 B**. To re-confirm this, fresh DNA samples from some positive (clones 1, 2, 30) and some negative (clones 20, 21) clones were digested with *BamHI* and analysed by Southern blot. The wild-type locus came up at the expected size of 18.4kb and the successful targeting events were visualized at the expected size of 10.7kb. The results obtained from the *BamHI* blot thus reconfirmed the ones obtained from the *SphI* blot. In addition, the polymorphism in one of the *Ova* alleles was not

generating any additional unexpected *BamHI* site so no extra band was obtained here (**Figure 3.4 C**). Finally, to have an idea of the efficiency of our rescue expression system, Western blot analysis was performed using total cell extracts of Southern-positive (left panel) and Southern-negative (right panel) clones, alongside the parental cell line serving as a control for the TopBP1 protein levels produced from the single remaining endogenous *TopBP1* allele. Surprisingly enough, no differences (in terms of protein expression levels) could be observed between the parental cell line or the clones negative for *Ova* targeting and the clones constitutively expressing the additional transgenic *TopBP1* allele from the *Ova* locus (**Figure 3.4 D**). To ensure that the transgene has been integrated intact with no mutations, PCR amplifications of slightly overlapping ~0.6kb-1kb fragments specifically from the integrated OVACMV fragment were performed and sent for Sanger sequencing. The specificity of the PCR reactions within the *TopBP1* cDNA region was achieved by designing primers that would not be able to anneal within the endogenous TopBP1 locus (i.e. an oligo perfectly annealing on the *TopBP1* transgene was comprised of the end of exon (*n*) and the start of exon (*n+1*) at the endogenous locus). Sanger sequencing of 4 independent clones confirmed that no mutations were incorporated at the *Ova* targeted locus (data not shown). These were expanded and frozen down in liquid nitrogen as $TopBP1^{flox/flox/+}$ /$Ova^{CMVTopBP1/+/+}$ clones 1, 2, 3 and 4.

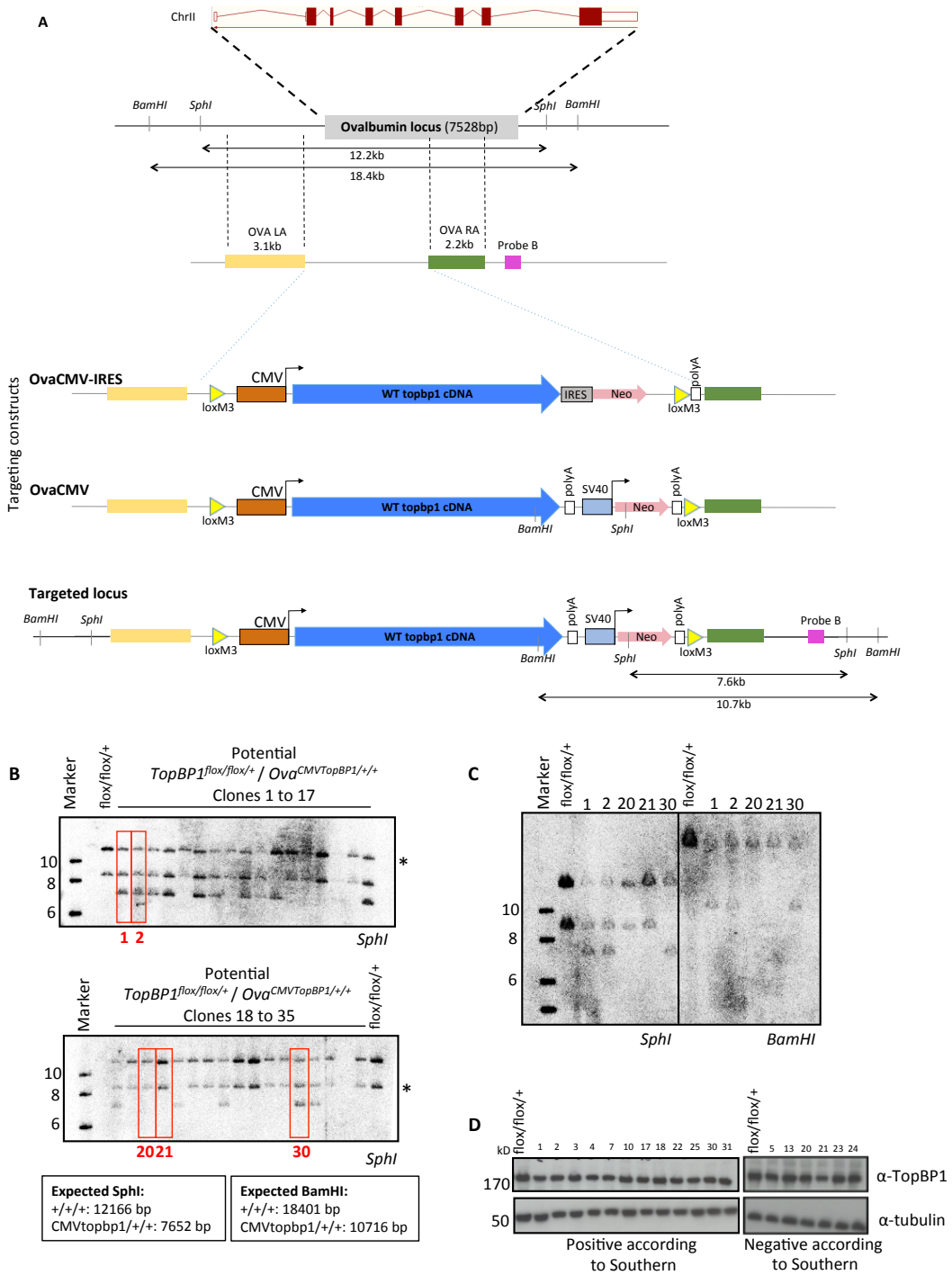


Figure 3.4: Stable integration at the *Ova* locus of the rescue wild-type *TopBP1* transgene under the control of the CMV constitutive promoter. **A)** Schematic representation of (from top to bottom) the ovalbumin (*Ova*) locus, the homology arms used, the targeting vectors attempted and the *OVA* locus successfully targeted with vector OVACMV. **B)** Southern blot analysis of *TopBP1*$^{flox/flox/+}$ parental cell line (control) and potential *TopBP1*$^{flox/flox/+}$ / *ova*$^{CMVTopBP1/+/+}$ clones following digestion of genomic DNA with *SphI*. Top panel depicts clones 1 to 17 and bottom panel clones 18 to 35. The diagnostic bands correspond to 12.2kb for an intact allele and 7.7kb for a successfully targeted allele. The asterisk corresponds to one of the three *Ova* alleles containing a polymorphism and producing an *SphI* diagnostic band of ~8.5kb. **C)** Clones denoted by red boxes in A were re-analysed by Southern blot analysis following digestion with *SphI* or *BamHI*. The diagnostic bands for *SphI* are same as in B, whereas for *BamHI* the wild-type diagnostic band is 18.4kb and the targeted band is 10.7kb. Digested genomic DNA was hybridized with the 3' external probe (purple bar) shown in A. **D)** Western blot analysis of indicated clones with an anti-TopBP1 (~180kDa) polyclonal antibody. Beta-tubulin was used as the loading control.

3.1.1.4 Attempting to knock out the third TopBP1 allele with LARAHis targeting construct

To indirectly but perhaps more effectively investigate the functionality of the ectopic expression, I decided to attempt knocking out the remaining third allele from the endogenous *TopBP1* locus. To do so the LARAHis vector was linearized and transfected as explained before. As deletion of the third allele could perhaps be a rare event and also to account for clonal variation, all four clones confirmed by sequencing in the previous step (clones 1, 2, 3 and 4) were used and a total of 32 knockout transfections performed. It is worth mentioning that since background activity of MerCreMer can lead to undesired excision of the *TopBP1* cDNA rescue construct from the ovalbumin locus during prolonged culture, we selected for cells retaining the transgene by culturing in media containing G418 prior to attempting the knockout transfection. Selection gave rise to 776 histidinol-resistant clones, which were subsequently screened for loss of the endogenous *TopBP1* locus. Screening all by Southern blot analysis would be quite time-consuming and expensive so a different method was followed. In fact, I took advantage of the floxable nature of the ectopic transgene and reasoned that a transfectant that would have successfully deleted the third endogenous copy of *TopBP1*, would display loss of viability following treatment with 4-HT. So equal amounts of all 776 cultures were transferred into fresh media with or without containing 4-HT and left to grow. Of the 776 clones, 126 were scored as 4-HT-sensitive (**Figure 3.5 B**) and these were frozen down and also subjected to Southern blot analysis. *SacI* digestion of wild-type DNA and probing with the *TopBP1* locus-specific probe A hybridized to a diagnostic 5.9kb band. The $TopBP1^{flox/flox/+}$ /$Ova^{CMVTopBP1/+/+}$ parental control generates one band at 5.9kb for the remaining intact allele and two overlapping bands at 5kb for the deleted/floxed alleles. Successful replacement of the intact allele with the LARAHis targeting vector was expected to shift the wild-type band to 8.5kb (**Figure 3.5 A**). However, as is evident from the representative Southern blot shown in Figure 3.6 C, the clones obtained were either genotypically the same as the parental cell line for the *TopBP1* locus or they had one of the already targeted alleles being retargeted and the wild-type copy still remaining intact (**Figure 3.5 A, C**). In fact, 60.3% of the 4-HT-sensitive clones apparently remained $TopBP1^{flox/flox/+}$ /$Ova^{CMVTopBP1/+/+}$ and 39.7% became $TopBP1^{flox/his/+}$/ $Ova^{CMVTopBP1/+/+}$ (**Figure 3.5 D**). Therefore, using the same targeting vector promotes a retargeting event mediated by the same homology arms. This retargeting event is more favorable than the deletion of the

third intact copy perhaps due to the much smaller size of the intervening sequence. The false positive result obtained from the 4-HT experiment is perhaps associated to an inherent sensitivity to this drug, even at low doses, for some of the clones.

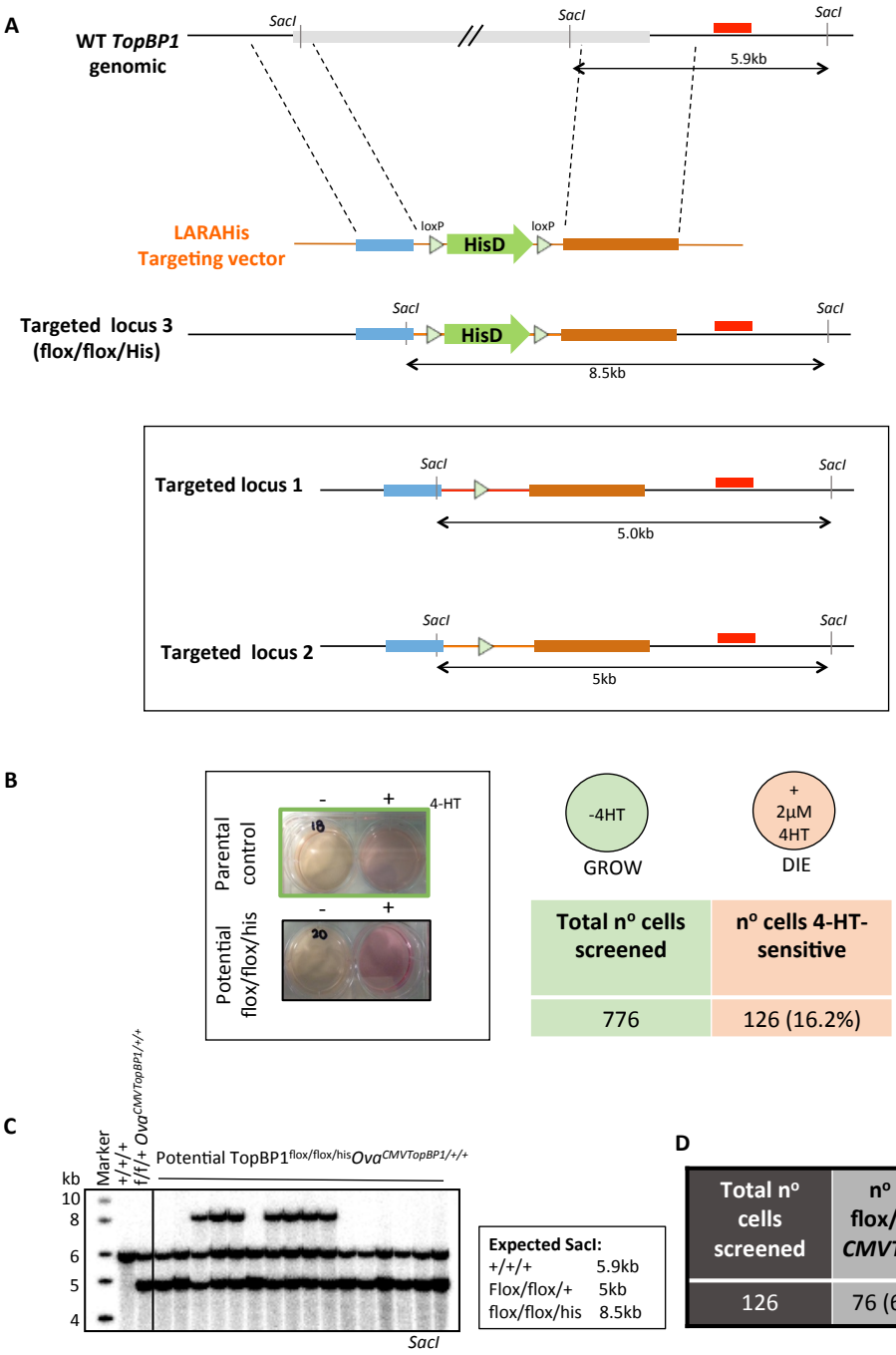


Figure 3.5: Attempt to knockout the *third* TopBP1 allele in *TopBP1*$^{flox/flox/+}$/*Ova*$^{CMVTopBP1/+/+}$ using the LARAHis targeting vector. A) Schematic representation of the wild-type *TopBP1* allele, LARAHis targeting vector and the successfully targeted locus. Inside the box at the bottom is a representation of the already targeted TopBP1 alleles from the previous steps. **B)** Screening for clones that have deleted all of the *TopBP1* alleles and hence should lose viability following induction of Cre recombinase. On the left panel is an example of a clone losing viability after 4-HT treatment (bottom) compared to the parental cell line (top). The right panel represents a table summarising the results obtained from this screening experiment. **C)** Southern blot analysis of the 4-HT-sensitive clones obtained in B. After *SacI* digestion of genomic DNA, a diagnostic band of 5.9kb represents the wild-type allele, a band of 5kb represents each of the floxed alleles and a band of 8kb represents a re-targeting event. Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in A. This blot is a representative figure and **D)** summarises all the results obtained from the Southern blots performed.

3.1.1.5 Attempting to knock out the third *TopBP1* allele with LARA2His targeting construct

To try and avoid the re-targeting event I changed the homology arms with arms homologous to internal regions of the *TopBP1* locus (now present only in the third copy). Using the Gateway recombination technology and donor vectors from Iiizumi *et al* (2006) again I attempted to assemble the LA2RA2His targeting construct (**Figure 3.6 A**). Numerous failures to PCR amplify the LA2, even trying different primer pairs, prompted us to assemble the LARA2His targeting construct instead, where the LA was exactly the same as the one used before (**Figure 3.6 A**). To do so the LA was amplified with primers P1 and P2 as before whereas primers P11 and P12 were used to amplify the RA2 from genomic DNA of the *TopBP1*$^{flox/flox/+}$ cell line. The resulting LARA2His plasmid was checked by PCR amplification of the subcloned arms, diagnostic enzyme digestion and Sanger sequencing.

Once verified, LARA2His was linearized by *PmeI* digestion and transfected into the *TopBP1*$^{flox/his/+}$ /*Ova*$^{CMVTopBP1/+/+}$ cell line. Transfectants were selected in the presence of histidinol and then were screened by the method of Cre induction as before. Of the 400 histidinol-resistant clones obtained from 22 transfections and screened for sensitivity to 4-HT, 341 were scored as resistant and 59 as sensitive (**Figure 3.6 B**). The clones displaying loss of viability following Cre induction were analysed for loss of the third endogenous *TopBP1* copy by Southern blot following *BglII* genomic DNA digestion. Probe A hybridized to a 7.5kb band whereas the already targeted alleles ran at 6kb. Deletion of the third copy and its replacement by the His cassette between the regions comprising the homology arms was expected to generate a 13.2kb diagnostic band following hybridization. However, an unexpected band at 9.5kb was observed that based on the relative intensities seemed to be a re-targeting of one of the already deleted copies. In fact, 22% of the clones screened were untargeted hence displaying same Southern blot profile as their parental counterparts, whereas the remaining 78% contained this band of unexpected size (**Figure 3.6 C, D**).

One possibility was that a novel *BglII* site was generated by some acquired mutation/polymorphism following the integration of the targeting vector. One way to avoid this potential novel site and screen for positive clones was to digest DNA with an alternative restriction enzyme. Unfortunately, all enzymes to the 3' region of the probe that could be used were also cutting within the area of the RA2 thus making the distinction between *TopBP1*$^{flox/flox/+}$ and *TopBP1*$^{flox/flox/his}$ impossible.

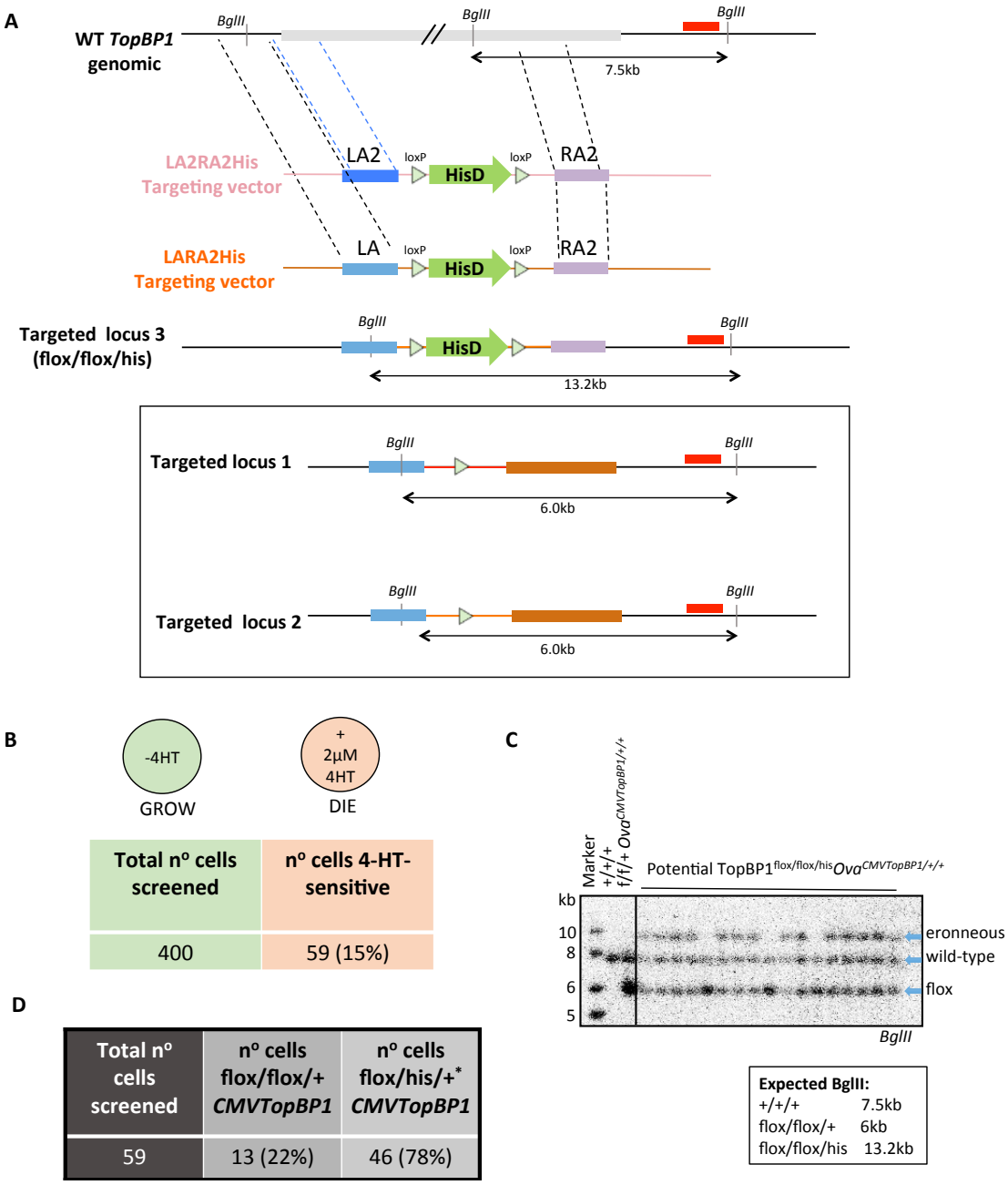


Figure 3.6: Attempt to knockout the third *TopBP1* allele in *TopBP1*$^{flox/flox/+}$/*Ova*$^{CMVTopBP1/+/+}$ using the LARA2His targeting vector. A) Schematic representation of the wild-type *TopBP1* allele, LARA2His targeting vector and the successfully targeted locus. Inside the box at the bottom is a representation of the already targeted TopBP1 alleles from the previous steps. **B)** Screening for clones that have deleted all of the *TopBP1* alleles and hence should lose viability following induction of Cre recombinase. The table summarises the results obtained from this screening experiment. **C)** Southern blot analysis of the 4-HT-sensitive clones obtained in B. After *BglII* digestion of genomic DNA, a diagnostic band of 7.5kb represents the wild-type allele, a band of 6kb represents each of the floxed alleles and a band of 9.5kb represents the product of erroneous recombination (see text for details). Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in A. This blot is a representative figure and **D)** summarises all the results obtained from the Southern blots performed. The asterisk represents the erroneous recombination event.

This was not considered a major problem as we reasoned that it is quite impossible for such a high percentage of clones to have all acquired the very same mutation/polymorphism and thus sought an alternative explanation.

Closer observation of the possible ways for a 9.5kb band to be generated revealed a potential erroneous recombination event (although a different scenario cannot be excluded). As depicted in **Figure 3.7 A**, the homology for this erroneous recombination event was mediated by the portion (shown in orange) of the LARAHis targeting vector surrounding the lox site that has remained in the second locus of *TopBP1* and that is homologous to the LARA2His targeting vector. This scenario requires resection of the linearized targeting vector beyond the RA2 region and thus implies that the product of the erroneous recombination would contain the RA instead of the RA2. So, in short, what this erroneous recombination event yields is a retargeted *TopBP1* allele (the one targeted with the His vector and floxed out) but instead of the presence of the RA2, the RA remains.

A second erroneous recombination event that could help explain why cells die following Cre induction would be at the region of the LA. In particular, it is possible that a recombination reaction between the LA on the targeting vector and the intact *TopBP1* allele rendered the protein produced from this allele partly- or non-functional with an impaired N-terminus. Alternatively, an inherent sensitivity to 4-HT can help explain the observed phenotypes as mentioned earlier. Finally, to prove this erroneous recombination event hypothesis a different probe, 5' to the LA was designed that would help distinguish the erroneous recombination product on the basis of size. Primers P13 and P14 were used to amplify the probe from wild-type DT40 genomic DNA. As for the Southern blot strategy, *EcoRV* digestion of genomic DNA of wild-type cells would yield a band of 9.8kb whereas the erroneous recombination product would be distinguished (21kb) from the already targeted alleles (17.5kb) (**Figure 3.7 B)**. *EcoRI* was also serving the same purpose leading to a diagnostic band of 10.5kb for wild-type, 11kb for intact targeted allele and 7.5kb for the product of the erroneous recombination. Unfortunately, the probe designed gave a lot of background non-specific binding to DT40 DNA, and although sample digestion was efficient for both enzymes, hybridization of the membrane could not deduce conclusive results (data not shown).

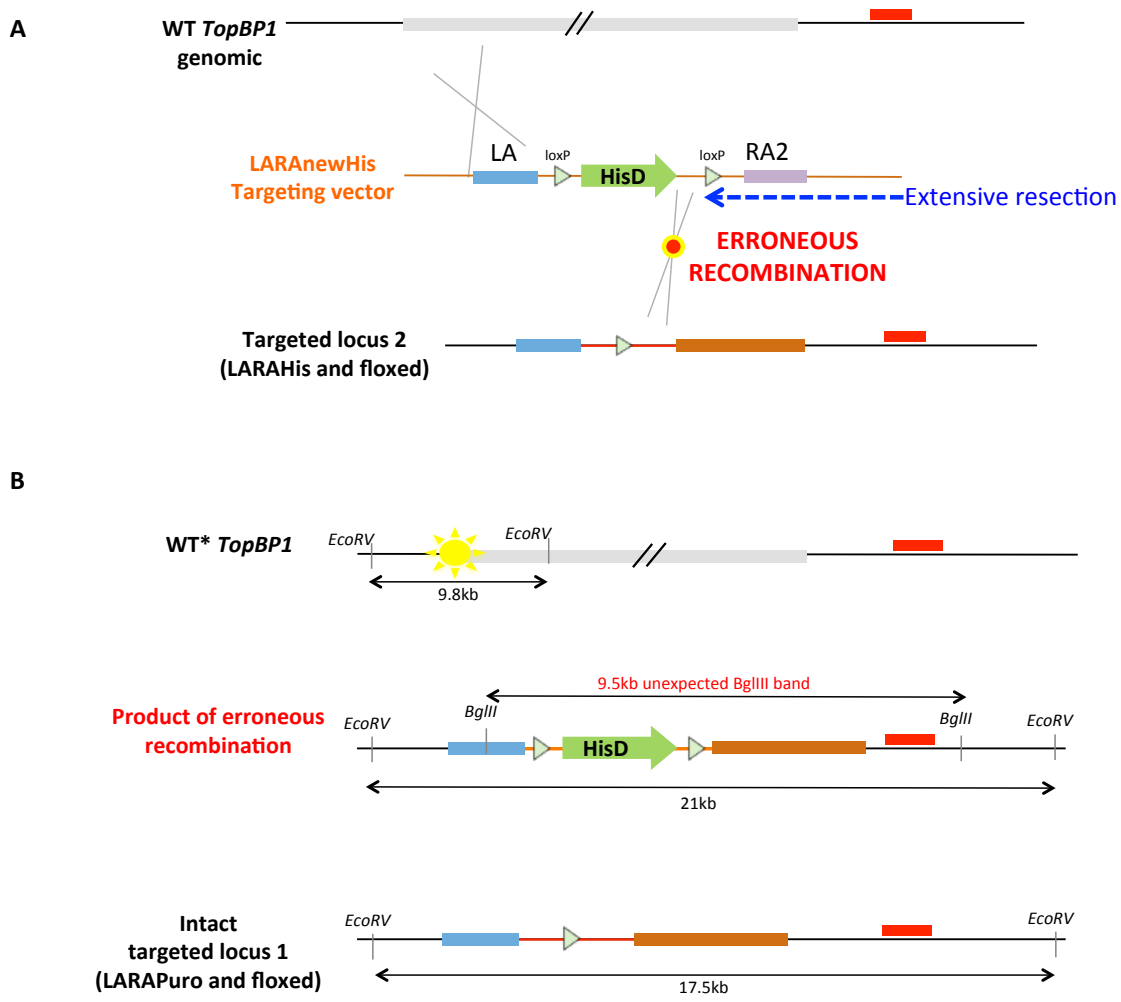


Figure 3.7: Erroneous recombination upon attempt to knockout the third *TopBP1* allele in *TopBP1*$^{flox/flox/+}$ /*Ova*$^{CMVTopBP1/+/+}$ using the LARA2His targeting vector. A) The predicted erroneous recombination involves two steps. In the first step, the LARA2His targeting vector recombines with the LA region of the intact *TopBP1* allele. In the second step, the LARA2His targeting vector erroneously recombines with the TopBP1 allele that has been already targeted with the LARAHis construct and the marker has been floxed. The homology is provided by the backbone sequences between the homology arms and the lox sites that have remained in the deleted allele but are also present in the LARA2His targeting vector. **B)** Schematic representation of the three *TopBP1* alleles following erroneous recombination. It is predicted that clones that displayed the 9.5kb unexpected band in figure 3.7 would be composed of a partly-functional full-length TopBP1 allele, an allele produced from the erroneous recombination event and the allele that was deleted with LARApuro and which remains intact during the erroneous recombination reaction.

Summary for chapter 3

In this chapter, I described the use of homology-directed repair to achieve complete deletion of two alleles of the *TopBP1* gene in DT40 cells. Next, I described the stable non-targeted integration of the MerCreMer enzyme under the control of a tamoxifen-inducible promoter, as a useful tool for future use of my model system. Then I went on to describe the removal of the selection cassettes from the targeted *TopBP1* alleles by induction of MerCreMer as well as the subsequent integration of the wild-type cDNA of *TopBP1* under the control of the CMV promoter at the *Ova* locus. Having confirmed the stable insertion of the rescue construct within the *Ova* locus, it was attempted to knockout the third endogenous allele of *TopBP1*. Two gene targeting constructs were tried and more than 1200 clones screened for loss of viability after treatment with 4-HT but no positive clone was obtained following Southern blot analysis. This prompted me to investigate whether the protein levels produced from the ectopic rescue construct were sufficient to sustain the viability of cells in the absence of the endogenous TopBP1. Also, the inability to obtain the knockout cell line despite the miscellaneous and intense approaches was informative and intriguing *per se*.

Chapter 4

Creation of a stably integrated overexpression system for the production of increased amounts of the ectopic *TopBP1* transgene and subsequent novel attempts to create the knock out

4.1: Creation of a stably integrated overexpression system for the production of increased amounts of the ectopic *TopBP1* transgene and subsequent novel attempts to create the knock out

The inability to obtain the knock out after screening more than a thousand clones raised the question of whether the levels being produced from the ectopic *TopBP1* copy at the *Ova* locus were enough to allow cells to let go of their final endogenous *TopBP1* allele. As discussed in the previous chapter, the cells harbouring the transgene at the *Ova* locus did not produce any apparent increase in the total amount of TopBP1 protein as judged by Western blot analysis of whole cell extracts using a polyclonal antibody raised against a C-terminal peptide of TopBP1. But since the ectopic transgene did not harbour any tag I was unable to specifically visualise the protein levels being produced from the *Ova* locus. Stable integration of more than one *TopBP1* transgenes under the control of the CMV promoter was the first method I employed to increase the levels of the protein. This time the transgene was also fused to a tag to enable separation of the TopBP1 protein pools, endogenous versus ectopic. The second method I embarked on involved a novel characterisation of different promoter regions driving expression of the *TopBP1* transgene from various loci in order to identify the best way of achieving maximal expression of the rescue construct and thus increase the probability of successfully achieving deletion of the remaining endogenous copy. These two methods together with novel ways of attempting the knockout are the focus of this subchapter.

4.1.1: Stable non-targeted transfection of the transgene under the control of CMV promoter to increase protein levels

The quickest way of achieving our goal of increasing the protein levels of TopBP1 was to stably integrate the *TopBP1* transgene randomly within the genome as a way of optimising overexpression and achieving sufficient complementation. Theoretically, the lack of homology arms and the high levels of energy used for electroporation allow the integration of more than one copy of the targeting vector in some, if not all, of the clones. To achieve non-targeted integration the OvaCMV expression vector described earlier was used to subclone the *TopBP1* cDNA at *NotI/XmaI* as before. For the reason explained in the introduction an array of three FLAG epitopes was introduced at the C-terminal end of the cDNA replacing the stop

codon. Subcloning of the FLAG-tags was achieved using annealed oligomers containing the octapeptides and ligation of the resulting linear molecule as an *XmaI/PmeI* fragment. The construct assembled called "CMVnon-targeted" contained no homology arms so by using relatively high electroporation energies that cause DNA breaks stable non-targeted integration into the genome could be achieved. Unlike for targeted integration, here it was not necessary to linearise the targeting construct.

The circular plasmid was transfected into the *TopBP1*$^{flox/flox/+}$ cell line and the G418-resistant clones obtained were screened by Western blot analysis with an anti-FLAG antibody. Almost 50% of the clones have successfully integrated the construct and it is interesting to note that some clones display higher expression levels than others, presumably due to more copies having integrated into the genome or different sites of integration affecting expression (**Figure 4.1 A**). All of the positive clones were expanded and frozen down as *TopBP1*$^{flox/flox/+}$ */Ova*$^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$. The only caveat with this system of non-targeted integration of the transgene was that induction of Cre by 4-HT treatment would not flox these randomly integrated copies of the *TopBP1* cDNA. Treating cells with the standard concentration of 2μM did not catalyse the deletion of the expression fragment flanked by the lox sites, as is evident from the persisting protein signal on the Western blot membrane. Doubling the concentration of 4–HT used did not make any difference whereas higher concentrations of 4-HT tested were toxic to the cells (**Figure 4.1 B**).

Despite this imperfection of the "CMVnon-targeted system, I decided to attempt the knockout targeting transfection to answer the question of whether more copies of the constitutively expressed and CMV-driven *TopBP1* cDNA would allow deletion of the endogenous locus. To avoid the issues of retargeting and erroneous recombination observed with the LARAHis and LARA2His targeting constructs, I had in the meantime designed a new *TopBP1* knockout vector containing arms homologous to internal regions within the gene and in respect to the LA and RA regions used before. It is worth mentioning that to be absolutely certain that the nucleotide sequence of the third intact copy does not differ from the consensus sequence on the genome database, I decided to design primers along the length of the new arm regions and PCR amplify them from the *TopBP1*$^{flox/flox/+}$ cell line. The ~0.7kb amplified fragments were gel extracted and sequenced and analysis of the sequencing data revealed no mismatches to the consensus sequence on the database. The new arm regions chosen, LA^{i}RA^{i} (I for internal) were therefore suitable for targeting of the third *TopBP1* copy. In addition to the construction

of internal arms, I also decided to use a different backbone vector (i.e. not the pDONOR vectors used before) so that the vector sequences present in the already targeted alleles would bear no homology to this newly assembled deletion construct. Finally, to avoid any erroneous inter-allelic recombination event lox sites were not included flanking the selection cassette.

To assemble the LA^{i}RA^{i}Puro construct, the *neomycin* gene within OvaCMV was replaced with a gene conferring resistance to puromycin using the flanking *EcoRI/BamHI* single cutters. The primers used were P15 and P16 and the substrate in the PCR reaction was a *puromycin*-containing expression vector kindly provided by the Hochegger laboratory. The resulting SV40-puro-pA fragment was excised by cutting with *EcoRV/XbaI* and subcloned into pBluescript SK+. Subsequently the LA^{i} and RA^{i} were amplified from genomic DNA with primers P17/P18 and P19/P20 and subcloned at *SalI* and *NotI*, respectively. The resulting LA^{i}RA^{i}Puro (**Figure 4.1 C**) was confirmed by diagnostic restriction enzyme digestion and Sanger sequencing.

Two of the *TopBP1*$^{flox/flox/+}$ / *Ova*$^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$ clones displaying high levels of expression of the transgene were transfected with *AhdI*-linearised LA^{i}RA^{i}Puro and selected in puromycin. This time the 4-HT screening method could not be used as explained earlier so all clones were subjected to genotyping by PCR (P40/P41). The primers used for the genotyping anneal to sequences in exon 19 and intron 22 respectively which serves two purposes; Firstly, amplification of the ectopic *TopBP1* transgene is impossible since the cDNA does not contain introns and secondly, it allows screening for knockout cells by negative selection, which means that no amplified band should be observed in a clone that has all endogenous *TopBP1* alleles deleted. None of the 250 clones tested had successfully deleted the *TopBP1* locus as is evident from representative results in **Figure 4.1 C**. A LA^{i}RA^{i}Puro successfully targeted clone should not produce an amplified product in the PCR reaction. Finally, representative results from Southern blot analysis of some of these clones confirmed that no successfully targeted clones were obtained. Following *BglII* digestion of wild-type cells the intact TopBP1 allele was represented by a diagnostic band of 7.5kb whereas the floxed alleles were both represented by overlapping bands at 6kb. A successfully targeted clone was expected to generate a diagnostic band of 9.2kb corresponding to the replacement of the third intact TopBP1 allele by the LA^{i}RA^{i}Puro targeting vector. As shown from the representative results in **Figure 4.1 D**, the obtained clones showed the exact same pattern as the parental cell line and have not lost the final endogenous allele

of *TopBP1*. At least no retargeting was observed with the use of the $LA^{i}RA^{i}$Puro targeting vector.

Overall, in this experiment we have managed to create the $TopBP1^{flox/flox/+}$ / $Ova^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$ cell line, visualise the protein produced from the exogenous construct and attempt the deletion of *TopBP1* with a newly made targeting construct specific for the third intact allele. The failure to target the endogenous wild-type copy prompted us to consider the possibility of the CMV promoter element not producing sufficient levels of the rescue TopBP1, not even when present in more than one copies within the genome. I thus sought novel ways of further increasing expression the exogenous protein expression.

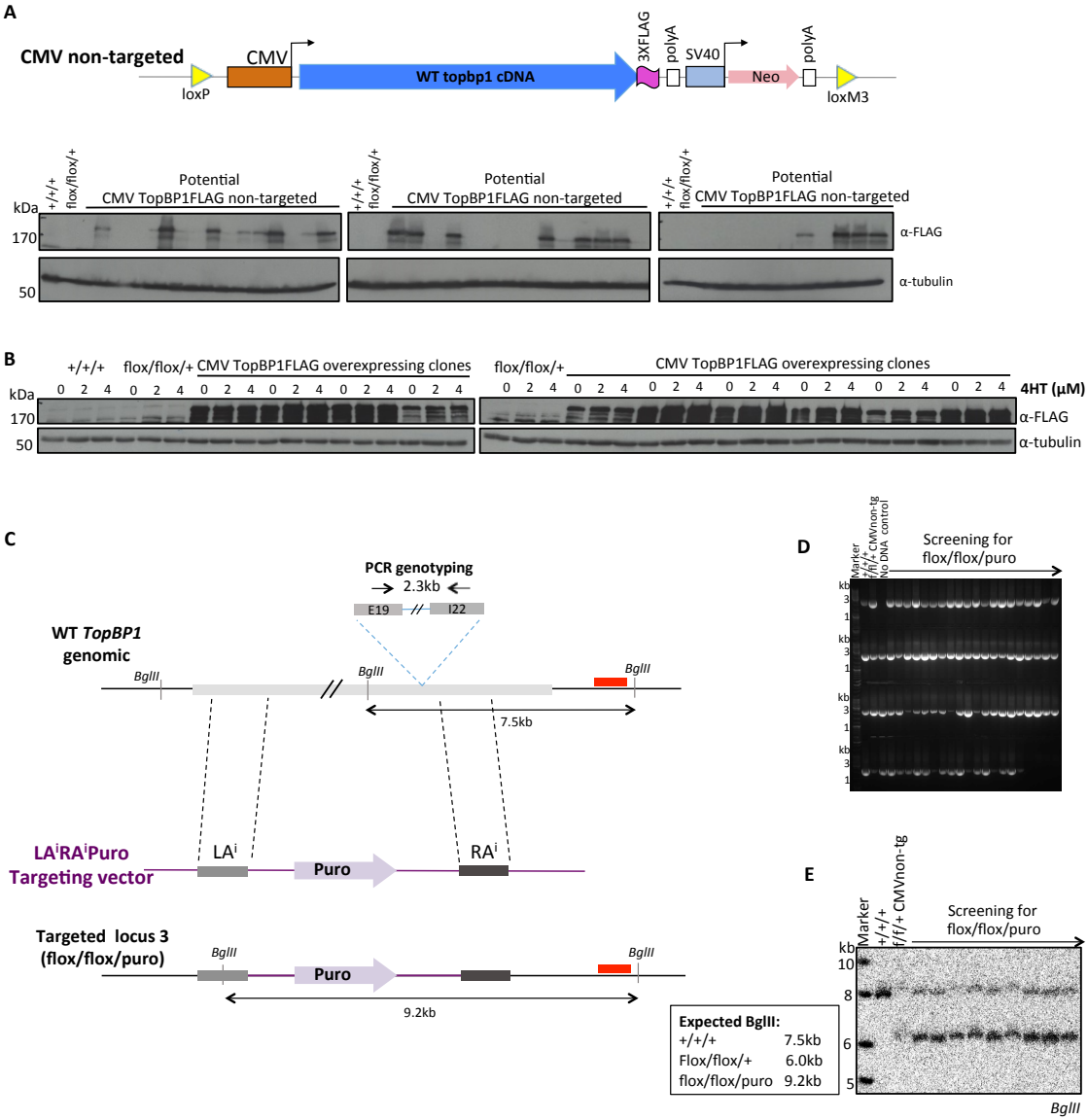


Figure 4.1: Creation of the *TopBP1*$^{flox/flox/+}$ /*CMV*$^{TopBP1FLAGnon\text{-}targeted}$ *cell line* and attempt to knockout the third endogenous copy with the LA^{i}RA^{i}Puro targeting vector. **A)** (top panel) Schematic representation of the "CMV non-targeted" targeting vector for stable non-targeted integration of the *TopBP1* transgene under the control of the constitutive CMV promoter. (bottom panel) Western blot analysis of clones transfected with the "CMV non-targeted" vector alongside the wild-type and parental cell lines in order to screen for the ones that display expression of the transgene. Cell lysates were blotted with an anti-FLAG antibody to control for transgenic TopBP1 expression. Blotting for beta-tubulin served as a loading control. **B)** Western blot analysis of lysates of the indicated cell lines following 4-HT treatment to screen for floxing of the ectopic non-targeted copy (copies) of the FLAG-tagged *TopBP1* transgene. Cells were treated with 0, 2 or 4μM 4-HT for 24h and then serially diluted to isolate single colonies. **C)** Diagrammatic representation of the wild-type *TopBP1* locus, LA^{i}RA^{i}Puro targeting vector and successfully targeted locus of *TopBP1*$^{flox/flox/+}$ /*Ova*$^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$ cells. **D)** PCR genotyping of wild-type DT40 and *TopBP1*$^{flox/flox/+}$ /*Ova*$^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$ parental control alongside puromycin-resistant clones obtained from transfection of the parental control with the LA^{i}RA^{i}Puro targeting vector shown in C. Successfully targeted clones would be negative for the 2.3kb amplified band, **E)** Southern blot analysis of *BglII*-digested DNA of wild-type DT40 and *TopBP1*$^{flox/flox/+}$ / *Ova*$^{CMVTopBP1FLAGnon\text{-}targeted/+/+}$ cells alongside puromycin-resistant clones obtained from transfection of the parental control with the LA^{i}RA^{i}Puro targeting vector shown in C. Diagnostic bands represent the intact (7.5kb) and floxed *TopBP1* alleles (6kb and 6kb). Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in C.

4.1.2: A comparison of viral promoter elements integrated at active or silent loci to enhance exogenous protein expression in the DT40 model system

The adequate choice of the promoter element and the characterization of the protein expression level is of significant importance towards the development of a suitable model system, tailored to meet the needs of the researcher and serve the purpose of the biological question. To our knowledge, no characterisation of promoter elements has been performed in the DT40 model system. In an effort to identify an expression system suitable for optimal exogenous protein expression in DT40 cells, I decided to compare the already used CMV promoter with two other promoter elements, CAG (CMV early enhancer and chicken beta actin) and CBA (chicken beta actin), in their ability to drive expression of the *TopBP1* transgene. The reason I chose these particular ubiquitous promoters was that they are generally described as being among the stronger constitutive promoters available in molecular biology (Powell, Rivera-Soto et al. 2015). Our work so far has indicated that the levels produced from the CMV element do not confer any overall increase in the total amount of TopBP1 produced inside the cells. Thus attempting the design of novel constructs that would increase the protein levels of our protein would possibly be of significant importance for sustaining viability of the *TopBP1* knockout.

To avoid the high variation in transgene expression resulting from uncontrolled copy number and chromosomal position effects when using non-targeted integration, I targeted the CAG, CBA and CMV expression systems in the same chromosomal position, the *Ova* locus, of the *TopBP1*$^{flox/flox/+}$ cell line thus creating three isogenic clones constitutively overexpressing the *TopBP1* transgene. This has enabled the direct comparison of constructs from within the same genomic context and allowed a systematic and quantitative assessment of the strengths of the promoters.

Furthermore, since *Ova* is a silent gene in the chicken B cells (Buerstedde and Takeda 1991), I wanted to investigate whether the local chromatin environment would have an effect on the levels of expression. Thus exogenous protein production from the *Ova* locus was compared to that of a transcriptionally active locus. The following paragraphs describe the creation of a stably integrated overexpression system (SIOS). SIOS is an easy to use and versatile system for constitutive, reversible exogenous protein production that provides a range of potential expression levels. This is a useful experimental tool for future DT40 experiments.

The three constructs assembled for the creation of the three expression systems at the *Ova* locus are shown in **Figure 4.2 A**. The design of the OvaCMV construct was explained in paragraph 3.1.1.3. To assemble the OvaCBA and OvaCAG constructs we purchased two expression vectors, pSF-CBA (OG262) and pSF-CAG (OG505). I then sequentially subcloned the *TopBP1* cDNA at the *KpnI* site, the SV40-neomycin-pA-loxM3-OvaRA fragment at the *PacI* site and the OvaLA-loxP at *PvuI*. Successfully cloned molecules were screened by diagnostic enzyme digestion after each cloning step and finally verified by Sanger sequencing. The primers used for these three steps were P21-P26. Additionally, the *TopBP1* transgene was fused to three FLAG epitopes included as an overhang in the reverse primer.

The OvaCMV targeting vector was linearized with *XhoI* as before whereas *ApaLI* was used to linearise OvaCBA and OvaCAG prior to electroporation into the *TopBP1*$^{flox/flox/+}$ cell line. Cells that successfully formed single colonies following selection in 2mg/ml G418 were subjected to Western blot analysis with an anti-FLAG antibody to visualize ectopic TopBP1-3XFLAG. The three targeting experiments and subsequent analysis were performed in parallel to eliminate variations in the experimental conditions. As shown in **Figure 4.2 B** most of the clones obtained were positive for the expression of the FLAG-tagged TopBP1 but to distinguish the clones that have targeted the construct specifically at the *Ova* locus, Southern blot analysis of *SphI*-digested genomic DNA was performed. Hybridisation with probe B revealed the overexpression clones positive for the *Ova* locus, evident from the 7.6kb band corresponding to a successfully targeted *Ova* allele (**Figure 4.2 D, E**, positives shown in red boxes).

Two clones of each of *TopBP1*$^{flox/flox/+}$*/Ova* $^{CMVTopBP1/+/+}$, *TopBP1*$^{flox/flox/+}$*/Ova* $^{CBATopBP1/+/+}$ *and TopBP1*$^{flox/flox/+}$*/Ova* $^{CAGTopBP1/+/+}$ were subjected to PCR amplification of ~0.7kb partly overlapping fragments and Sanger sequencing across the length of the transgene to ensure no mutations have been incorporated following integration. Interestingly, having the transgene fused to a protein tag allowed the characterization of the relative efficiencies of the three distinct promoters among the three isogenic counterparts. As depicted in **Figure 4.2 C**, the CAG promoter displayed the highest strength in relation to *TopBP1* expression. The average relative expression of the two CAG clones was 7.5AU, of the four CBA clones 2.8AU and of the two CMV clones 1.3AU. Hence the CAG promoter element showed more than 2.6-fold higher relative expression of the transgene than the CBA and 5.8-fold than the CMV.

Probing with the anti-TopBP1 antibody raised against a C-terminal peptide of the protein (Appendix 1) enabled a qualitative assessment of the three systems based on the fold-increase of the total protein levels over the amount of endogenous TopBP1 in the parental cell line. When compared to exogenous TopBP1, the CMV construct contributes less than 1-fold) of the total pool of TopBP1 protein inside the cells, whereas CBA and CAG contribute to 3- and 6-fold than the endogenous allele **(Figure 4.2 F, G)**. Thus, a single copy of the CBA promoter increases the TopBP1 levels to approximately the levels normally found in wild-type cells (where there are three copies of the endogenous gene) while the CAG promoter provides approximately twice the amount of TopBP1 than that found in wild-type cells.

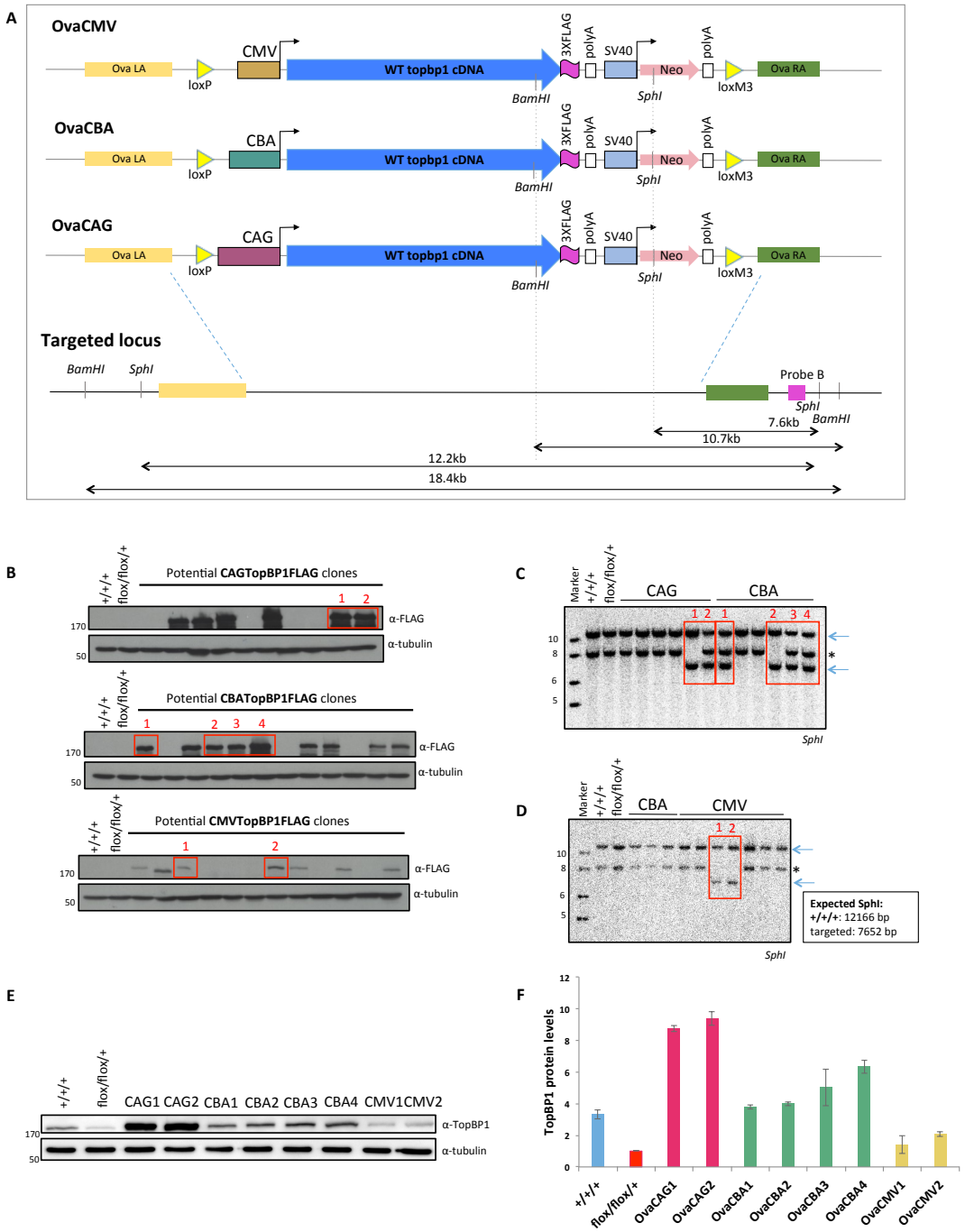


Figure 4.2: Relative promoter activities of *TopBP1* transgenic constructs stably integrated at *Ova*. **A)** Schematic representation of *TopBP1* expression constructs under the control of different promoters for stable integration into the *Ova* locus of DT40 cells. **B)** Western blot analysis of cell lysates from cells transfected with the expression constructs shown in A. Immunoblotting with an anti-FLAG antibody controls for the expression of the ectopic TopBP1-3xFLAG. Cell lysates of wild-type and parental cells were used as a negative control. The levels of beta-tubulin were used as a loading control. **C, D)** Clones showing expression of the *TopBP1* transgene in B were subjected to Southern blot analysis following *SphI* digestion of genomic DNA. The diagnostic bands corresponding to the wild-type allele (12.2kb) and the successfully targeted allele (7.6kb) are indicated. The asterisk corresponds to one of the three *Ova* alleles containing a polymorphism that generates an *SphI* restriction site. Digested genomic DNA was hybridized with the 3' external probe (purple bar) shown in A. **E)** Western blot analysis of cell lysates of cells overexpressing the *TopBP1* transgene under the control of the three constitutive promoters shown in A. Immunoblotting with an anti-TopBP1 antibody controls for the visualisation of the increase of TopBP1 protein levels over the levels produced from the endogenous promoter. Beta-tubulin serves as the loading control. **F)** Quantification of the blot in E.

Ova is a silent gene in the chicken B cells, which may affect the expression levels due to the local chromatin environment. To examine if this was the case we compared TopBP1 protein production from the CAG promoter integrated at the Ova locus to the levels of TopBP1 when expressed from the same promoter construct integrtated at a transcriptionally active locus. To achieve this, the TopBP1 transgene construct under the control of the CAG promoter was recreated with different homology arms. I chose to target a region downstream of the endogenous *TopBP1* locus on chromosome II (42,781,752-42,789,165), a region between the 3'end of the TopBP1 locus and the 5' start of the *CDV3* protein-coding gene.

To assemble the "euchromatinCAG" (euCAG) construct shown in **Figure 4.3 A**, the *TopBP1* cDNA was first subcloned at the *KpnI* site of pSF-CAG (OG505), as before. In the meantime, the euchromatin LA was PCR amplified from genomic DNA and subcloned in the OvaCMV using *NheI/SalI* sites and replacing in this way the existing OvaLA (P27 and P28). Following the same logic the euchromatin RA was amplified and subcloned in OvaCMV using *AscI/XhoI* sites and replacing in this way the existing OvaRA (P29 and P30). These cloning steps were confirmed by diagnostic enzyme digestion and Sanger sequencing. Subsequently, the SV40-neomycin-pA-loxM-euchromatinRA and the euchromatinLA-loxP fragments were PCR amplified from the aforementioned constructs and subcloned sequentially at the *PacI* and *PvuI* sites of pSF-CAG, respectively (P31 and P32 & P33 and P26). The final euCAG construct was verified by diagnostic enzyme digestion and Sanger sequencing.

The resulting construct was linearized by *ApaLI* digestion and transfected into the *TopBP1*$^{flox/flox/+}$ cell line. To be able to directly compare the relative efficiency of the CAG promoter element between the *Ova* locus and the euchromatic locus downstream of the endogenous *TopBP1* gene of the *TopBP1*$^{flox/flox/+}$ cell line, the OvaCAG targeting vector was transfected in parallel to the euCAG targeting vector (**Figure 4.3 B)**. Single clones were isolated following selection with G418 and were checked for expression of the transgene by Western blot analysis and immunoblotting with an anti-FLAG antibody. The majority of the *Ova*-targeted (**Figure 4.3 C**) and the euchromatic region-targeted (**Figure 4.3 D**) cells were positive for TopBP1-3xFLAG expression. But to test which of those had the transgenes integrated at the loci of interest, Southern blot analysis was performed. 6 out of the FLAG-positive clones had specifically targeted

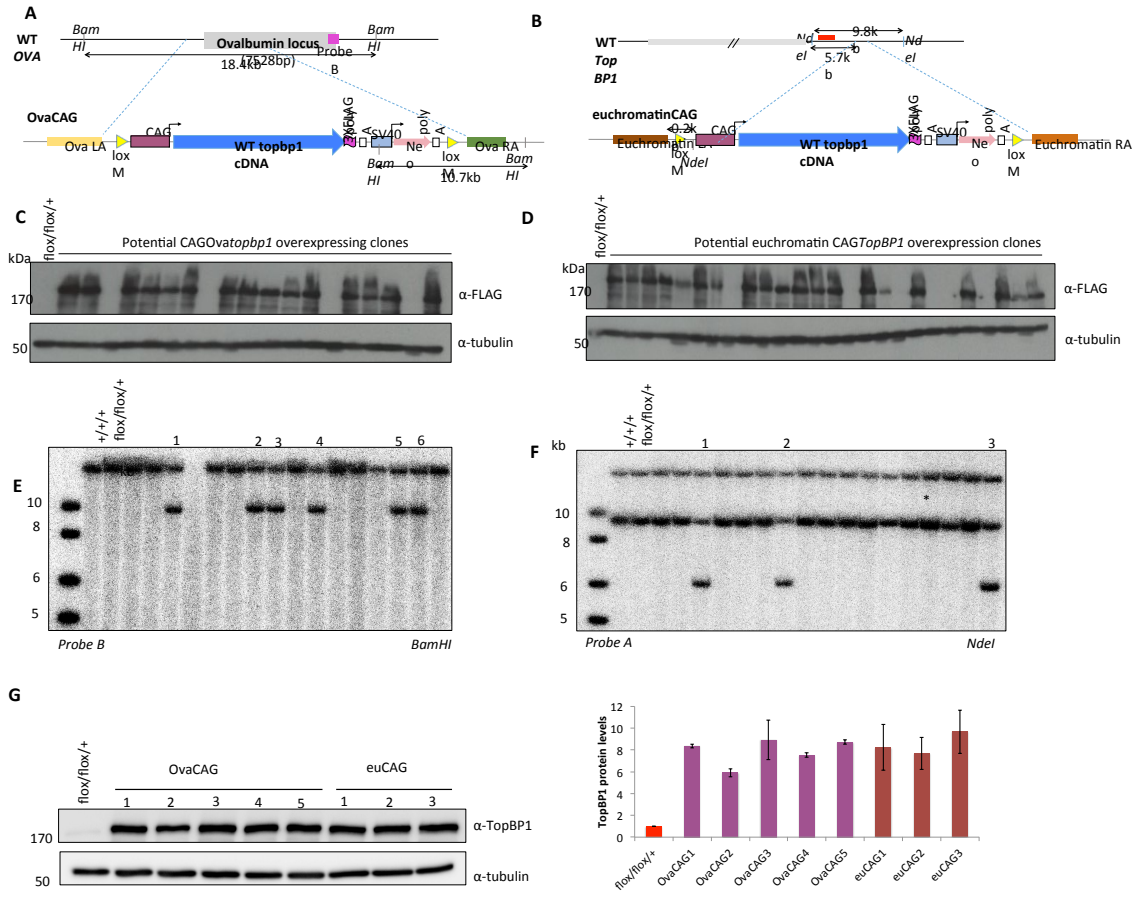


Figure 4.3: Relative strength of the CAG promoter driving expression of the *TopBP1* transgene when stably integrated at the *Ova* locus versus a euchromatic locus. A) Schematic representation of the targeting constructs used to stably integrate the CAG-*TopBP1* transgene at the *Ova* locus and **B)** A euchromatic locus nearby the endogenous *TopBP1* gene (right panel). **C)** Western blot analysis of cell lysates from cells transfected with the expression construct shown in A. Immunoblotting with an anti-FLAG antibody controls for the expression of the *Ova* TopBP1-3xFLAG. Cell lysates of wild-type and parental cells were used as a negative control. The levels of beta-tubulin were used as a loading control. **D)** Western blot analysis of cell lysates from cells transfected with the expression construct shown in B. Immunoblotting with an anti-FLAG antibody controls for the expression of the euchromatin TopBP1-3xFLAG. Cell lysates of wild-type and parental cells were used as a negative control. The levels of beta-tubulin were used as a loading control. Clones positive according to WB in C and D were subjected to Southern blot analysis in **E)** and **F)**. In E the diagnostic bands following *BamHI* digestion of genomic DNA correspond to the wild-type allele (18.4kb) and the successfully targeted allele (10.7kb). Digested genomic DNA was hybridized with the 3' external probe (purple bar) shown in A. In F the diagnostic bands following *NdeI* digestion of genomic DNA correspond to the wild-type allele (9.8kb) and the successfully targeted allele (6kb). The band at the top of the membrane results from a polymorphism (see text). Digested genomic DNA was hybridized with the 5' external probe (red bar) shown in B. **G)** Western blot analysis *Ova*-CAG*TopBP1* and euchromatin-CAG*TopBP1* lysates of cells positive according to E and F. Immunoblotting with an anti-TopBP1 antibody controls for the visualisation of the increase of TopBP1 protein levels over the levels produced from the endogenous promoter. Beta-tubulin serves as the loading control.

the *Ova* locus and 3 out of the 17 had successfully targeted the euchromatic locus. The expected sizes for the successfully targeted *Ova* locus has been explained already. For the euchromatic locus, *NdeI* digestion of genomic DNA and hybridization to probe A was expected to generate a 9.8kb band for wild-type cells and a 6kb band for clones that have successfully integrated the euCAG construct (**Figure 4.3 E, F**). The extra band above the 10kb is more likely a polymorphism within one of the three alleles on chromosome II as it was present in all the cell lines (**Figure 4.3 F**). To estimate the potential differences in expression between the two loci, total protein was prepared from five OvaCAG and three euCAG integrants, subject to Western blotting and probed with α-TopBP1. Some variation (up to three fold) was observed within the five OvaCAG clones and also within the three euCAG clones. In total, the data indicate that there is no major difference when comparing OvaCAG and euCAG, suggesting that the *Ova* locus in an appropriate site for integrating exogenous genes for expression. The main advantage of the *Ova* locus is its integration rate (**Figure 4.3 G**).

Furthermore, the CMV, CAG and CBA overexpression systems were characterised in terms of their stability to be used as reliable tools in future experiments. In fact, the behavior of these cell lines was followed for 8 days in culture to ensure that ectopically expressed TopBP1 did not reduce over time and that overexpression of the transgene was not toxic to the cells. The continued culturing of these cell lines did not affect the promoters' activity, as judged by Western blot analysis and immunoblotting against TopBP1 (**Figure 4.4 A**). I followed the cell cycle profiles by Flow Cytometry to establish of the overexpression of TopBP1 was influencing the progression of the cell cycle. No changes were observed over the course of the eight days (**Figure 4.4 B**). In addition, when compared to the parental strain, no growth defect was associated with overexpression of the *TopBP1* cDNA when cell numbers were followed over a 120 hour period (**Figure 4.4 C**). Finally, induction of Cre recombinase by 4-HT treatment for 24h and subsequent serial dilution to isolate single clones confirmed that the ectopic transgenes could successfully be floxed (**Figure 4.4 D**).

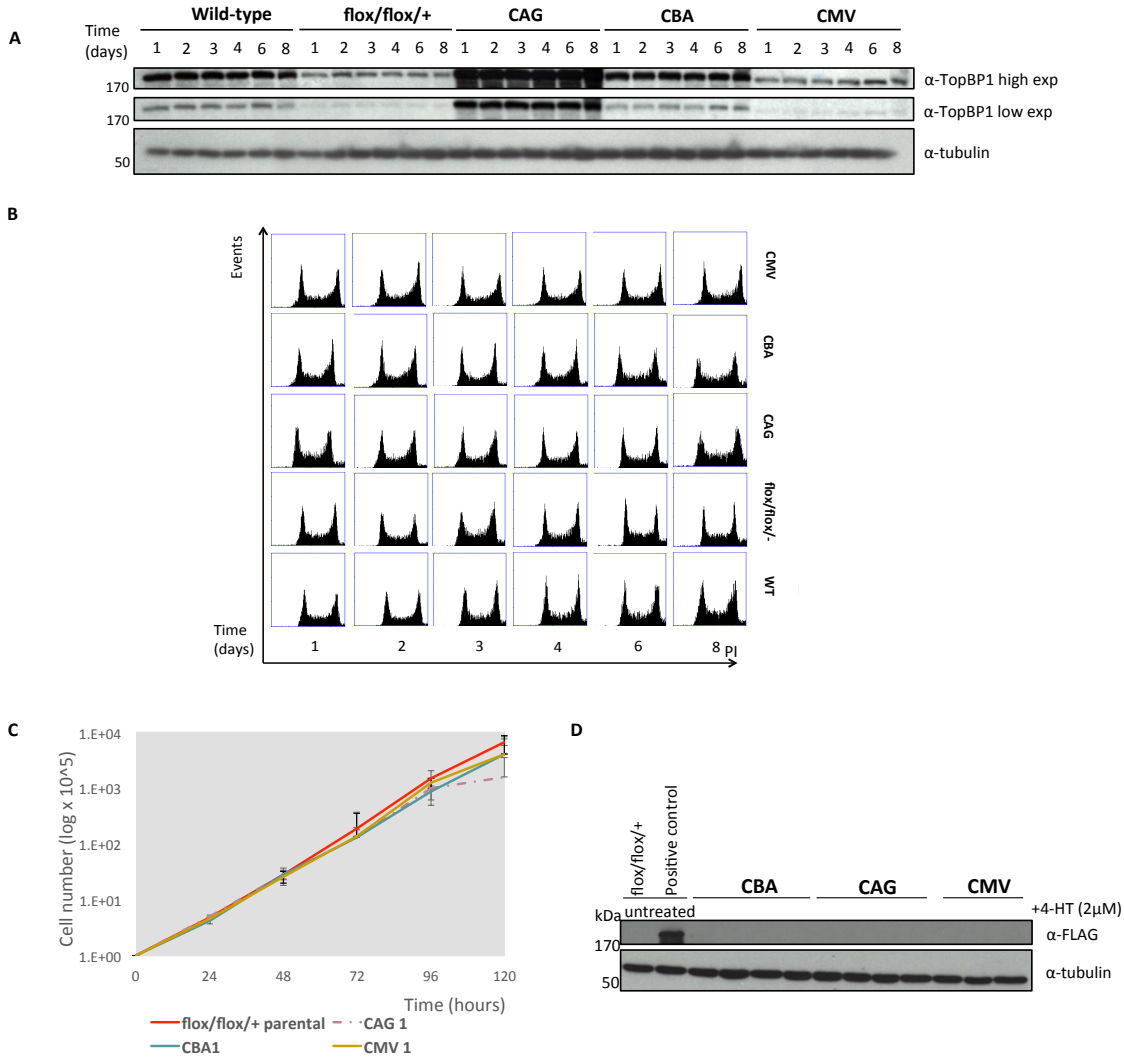


Figure 4.4: Stability of the SIOS system. A) The indicated cell lines were cultured for a period of seven days and samples were kept for Western blot analysis at the specified times. Immunoblotting with an anti-TopBP1 antibody controls for the expression of TopBP1. Cell lysates of wild-type and parental cells were used as a negative control. The levels of beta-tubulin were used as a loading control. **B)** Flow cytometry analysis of samples fixed and stained with propidium iodide. **C)** Growth curves for the indicated cell cultures. **D)** The indicated TopBP1 overexpressing cell lines were tested for their ability to flox the ectopic transgene following treatment with 4-HT (2μM) for 24h. The treated pools of cells were serially diluted and single clones were subsequently expanded and analysed by Western blotting. Immunoblotting with an anti-FLAG antibody controls for the expression of the TopBP1-3xFLAG transgene. Lysates of cells not containing any transgene or were used as a negative control. The levels of beta-tubulin were used as a loading control. The clones used in this experiment are CAG 2, CBA 4 and CMV 2 presented in figure 4.2.

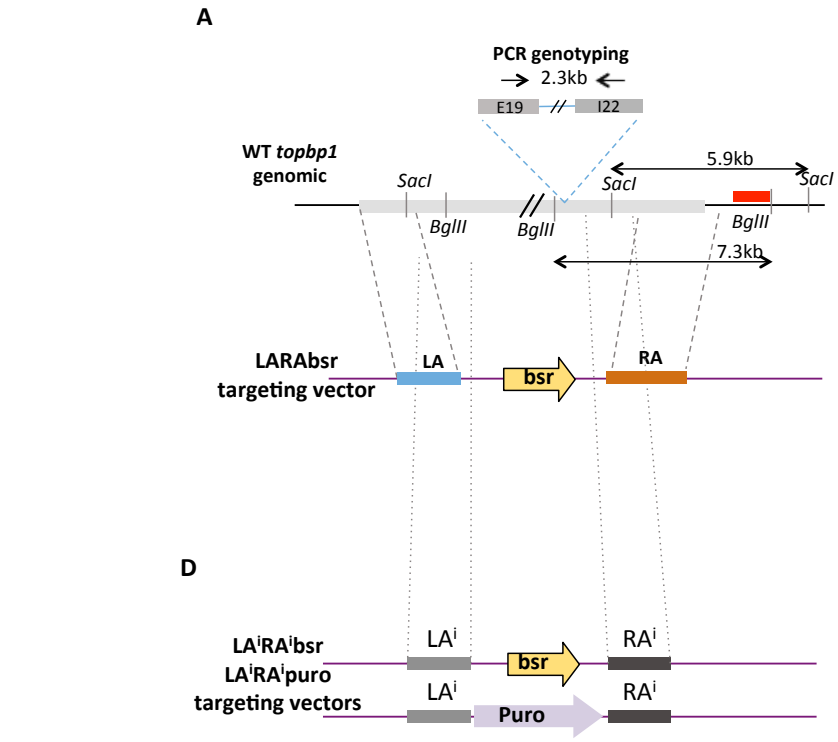
4.1.2.1 Attempting to knock out the third TopBP1 allele with novel methods

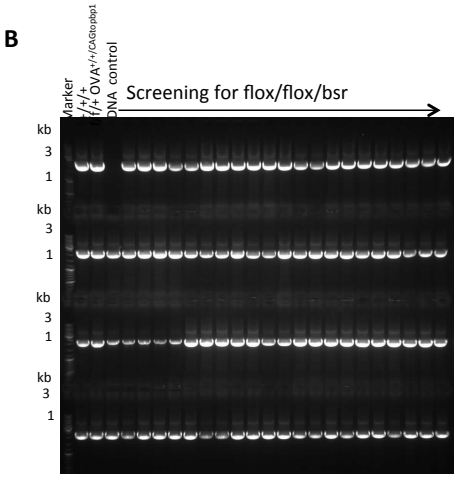
Having optimised the ectopic expression of the rescue transgene, the next step was to attempt knocking out the third intact copy of *TopBP1* in the *TopBP1*$^{flox/flox/+}$*/Ova* $^{CAGTopBP1/+/+}$ cell line (clones 1 and 2). To achieve this goal several strategies have been employed.

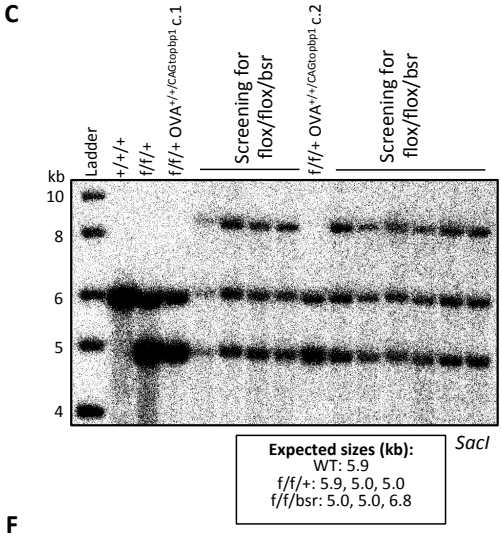
Firstly, a new knock out *TopBP1* targeting construct, LARAbsr (**Figure 4.5 A**), was assembled that contained the arms used for targeting the first two alleles, LA and RA, but a new selection marker cassette conferring resistance to blasticidin (*bsr*). This would prevent an erroneous recombination event promoted by the vector sequences between the arms and would help answer whether the optimised ectopic over-expression can help avoid re-targeting the already deleted *TopBP1* alleles. To assemble the LARAbsr construct, the pBSK+ containing the SV40-puro-pA fragment (described in 3.2.1) was digested with *EcoRI/BamHI* to replace the puromycin gene with *bsr*. Primers P34/P35 were used to amplify *bsr*. Subsequently the LA was amplified with primers P36/P37 and subcloned at *SalI* and the RA was amplified with primers P38/P39 and subcloned at *NotI*. Diagnostic enzyme digestion and Sanger sequencing confirmed the final sequence composition of the LARAbsr targeting vector. The vector was linearised by *AhdI* digestion and transfected into the *TopBP1*$^{flox/flox/+}$*/Ova* $^{CAGTopBP1/+/+}$ clones 1 and 2. The 150 clones that conferred resistance to blasticidin were genotyped by PCR (P40/P41). A positive clone should show no band after PCR as the region where the primers anneal would have been deleted. None of the clones screened had lost this region of PCR amplification meaning that none has deleted the locus (**Figure 4.5 B**). To visualize how the targeting construct has recombined with the *TopBP1* alleles, 10 of the clones screened by PCR were also subjected to Southern blot analysis following *SacI* digestion. This revealed that all of the clones had re-targeted one of the two deleted alleles (**Figure 4.5 C**) and from this it was concluded that using the LA and RA arms makes re-targeting a quite frequent event.

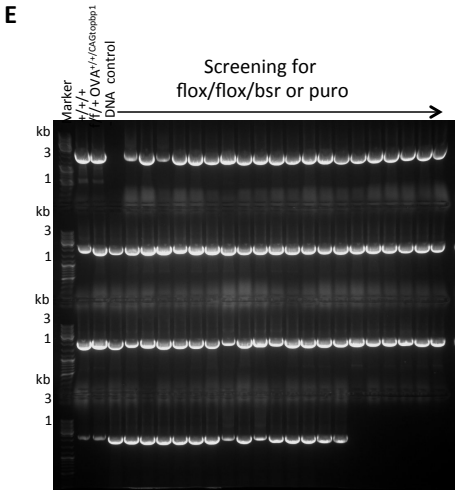
Secondly, it was decided that using the internal arms LA^{i} and RA^{i} was the optimal strategy for achieving the knockout and thus the previously described LA^{i}RA^{i}Puro construct was employed here as well. Additionally, the LA^{i}RA^{i}bsr construct was assembled which only differed in having the *bsr* gene in the place of *puro* at the *EcoRI/BamHI* site (**Figure 4.5 D**). The reason behind this was that by testing two different drugs I could eliminate the possibility of failing to obtain the knockout due to a potential low-level toxicity of one of the drugs to the cells. Both plasmids were

linearised by *AhdI* digestion and transfected into the *TopBP1*$^{flox/flox/+}$/*Ova* $^{CAGTopBP1/+/+}$ clones 1 and 2. The 90 clones that conferred resistance to blasticidin or puromycin were genotyped by PCR (P40/P41). Again all clones were negative for the deletion of the endogenous locus (**Figure 4.5 E**) and also negative when analysed by Southern blot. The only difference was that the use of internal arms eliminated the retargeting event and all clones analysed had exactly the same hybridisation pattern as the parental cells (**Figure 4.5 F**).


A
PCR genotyping
2.3kb
E19
I22
WT topbp1 genomic
SacI
BglII
SacI
BglII
5.9kb
SacI
BglII
7.3kb
LARAbsr targeting vector
LA
bsr
RA
D
LAiRAibsr
LAiRAipuro targeting vectors
LAi
bsr
RAi
LAi
Puro
RAi


B
Marker
+/+
f/+ OVA+/+/CAGtopbp1
DNA control
Screening for flox/flox/bsr
kb
3
1


C
Ladder
+/+/+
f/f/+
f/f/+ OVA+/+/CAGtopbp1 c.1
Screening for flox/flox/bsr
f/f/+ OVA+/+/CAGtopbp1 c.2
Screening for flox/flox/bsr
kb
10
8
6
5
4
SacI
Expected sizes (kb):
WT: 5.9
f/f/+: 5.9, 5.0, 5.0
f/f/bsr: 5.0, 5.0, 6.8


E
Marker
+/+/+
f/+ OVA+/+/CAGtopbp1
DNA control
Screening for flox/flox/bsr or puro
kb
3
1

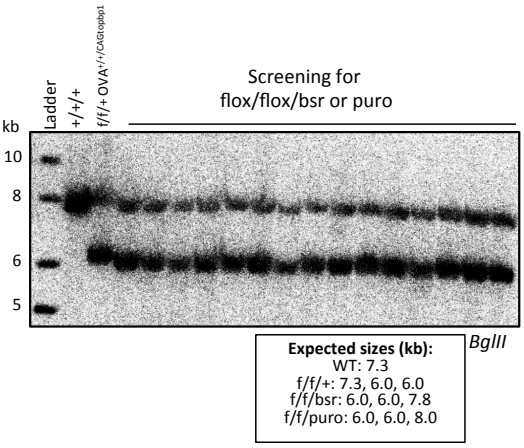

F
Ladder
+/+/+
f/f/+ OVA+/+/CAGtopbp1
Screening for flox/flox/bsr or puro
kb
10
8
6
5
BglII
Expected sizes (kb):
WT: 7.3
f/f/+: 7.3, 6.0, 6.0
f/f/bsr: 6.0, 6.0, 7.8
f/f/puro: 6.0, 6.0, 8.0

Figure 4.5: Attempt to knockout the third *TopBP1* allele in *TopBP1*$^{flox/flox/+}$/*Ova*^{CAGTopBP1} using the LARAbsr targeting vector. A) Diagrammatic representation of the wild-type *TopBP1* locus and LARAbsr targeting vector. **B)** PCR genotyping of wild-type DT40 and *TopBP1*$^{flox/flox/+}$/*Ova*$^{+/+/CAGTopBP1}$ parental control alongside bleomycin-resistant clones obtained from transfection of the parental control with the LARAbsr targeting vector shown in A. Successfully targeted clones would be negative for the 2.3kb amplified band. **C)** Southern blot analysis of *SacI*-digested DNA of wild-type DT40 and *TopBP1*$^{flox/flox/+}$/*Ova*$^{+/+/CAGTopBP1}$ cells alongside bleomycin-resistant clones obtained from transfection of the parental control with the LARAbsr targeting vector. Diagnostic bands represent the intact (5.9kb) and floxed *TopBP1* alleles (5kb and 5kb). Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in C. **D)** Diagrammatic representation of the LA^{i}RA^{i}bsr and LA^{i}RA^{i}Puro targeting vectors. **E)** PCR genotyping of wild-type DT40 and *TopBP1*$^{flox/flox/+}$/*Ova*$^{+/+/CAGTopBP1}$ parental control alongside bleomycin or puromycin-resistant clones obtained from transfection of the parental control with the targeting vector shown in D. Successfully targeted clones would be negative for the 2.3kb amplified band. **F)** Southern blot analysis of *SacI*-digested DNA of wild-type DT40 and *TopBP1*$^{flox/flox/+}$/*Ova*$^{+/+/CAGTopBP1}$ cells alongside bleomycin-resistant clones obtained from transfection of the parental control with the LARAbsr targeting vector. Diagnostic bands represent the intact (5.9kb) and floxed *TopBP1* alleles (5kb and 5kb). Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in C.

The third strategy devised was the creation of a cell line where the *TopBP1* transgene under the control of the strong CAG promoter was stably integrated into the genome in a non-targeted manner. This is similar to what was described in 3.2.1 section, although the much stronger CAG promoter was used here and also the transgene was not fused to an oligopeptide tag. To assemble the "CAG non-targeted untagged" construct (**Figure 4.6 A**), the *TopBP1* cDNA was first subcloned at the *KpnI* site of the original pSF-CAG (OG505) expression vector using primers P42/P43. Then the SV40-neomycin-pA-loxM fragment was subcloned at the *PacI* site (P44/P45) and a loxP site inserted at *PvuI* using annealed oligonucleotides (P46/P47). The resulting expression vector was confirmed by diagnostic digestion and Sanger sequencing. The "CAG non-targeted untagged" construct was then linearised by *ApaLI* digestion and transfected into the *TopBP1*$^{flox/flox/+}$ cell line using high voltage. More than 30 G418 resistant clones were obtained and 8 of them were checked by Western blot analysis using the anti-TopBP1 antibody. All but one of the clones were successfully showing overexpression of the TopBP1 protein (**Figure 4.6 B**). Two of these *TopBP1*$^{flox/flox/+}$ *CAG non-targeted* clones were subsequently transfected with the LA^{i}RA^{i}Puro construct for attempting once again to delete the endogenous copy of *TopBP1*. All of the 184 puromycin-resistant single clones obtained were genotyped by PCR amplification using the same primers as before. None was identified as a knockout clone (**Figure 4.6 C**).

Finally, the CRISPR (clustered regularly interspaced short palindromic repeats) and CRISPR-associated (Cas) 9 genome editing technology was employed to attempt the *TopBP1* knockout. The CRISPR-Cas9 system makes use of the mechanism of adaptive immunity present in some bacteria and archaea that enable these microorganisms to raise a response and eliminate invading genetic material. In short, these organisms have the ability to chop invading DNA into small fragments which they then integrate into their CRISPR loci as short repeats. These loci are transcribed producing transcripts, which are processed into CRISPR RNA molecules. These RNA oligos can then target cellular endonucleases to invading DNA molecules by means of sequence complementarity. Thus microbes have memory recording of their infections which allows them to defend themselves against pathogens (Jinek, Chylinski et al. 2012). And although the biology of microbes' immunity has been known for many years, it wasn't until 2012 that scientists including Feng Zhang and Jeniffer Doudna reported that a simplified version of CRISPR could be used for genome editing in mammalian and other cells (Lander 2016).

Here two constructs were assembled to specifically target the Cas9 enzyme to the LA^i and RA^i regions only present in the third intact *TopBP1* copy. The vector used for co-expression of the guide RNA and the Cas9 enzyme was pSpCas9(BB)-2A-Puro (PX459), initially described in Ran *et al* 2013 and available from addgene (Ran, Hsu et al. 2013). One guide RNA (gRNA) specific for the LA^i was synthesised by annealing primers P48/P49 and one gRNA specific for the RA^i was synthesised by annealing primers P50/P51 and subcloning at the *BbsI* site of PX459 (**Figure 4.6 D, E**). As the gRNAs were too small, bacterial clones could not be checked by diagnostic restriction digest and thus were confirmed by Sanger sequencing. The resulting PX459-gRNALA^i and PX459-gRNARA^i expression constructs were co-transfected with the LA^iRA^ibsr targeting construct. The thinking behind this experiment was the Cas9 enzyme would create DSBs within the genomic regions of the homology arms thus creating the need for repair. Having no other *TopBP1* alleles containing such regions so as to repair the damage cells would in theory use the repair template provided by LA^iRA^ibsr. It is worth mentioning that in order to render the LA^iRA^ibsr targeting vector (i.e. the repair template) resistant to cleavage by Cas9, one base within the ~20bp sequence corresponding to the gRNA specific for the LA^i was mutated (in a way so as to silently mutate the codon) by site-directed mutagenesis. The same was done for the RA^i region.

The transfection of the CRISPR constructs was attempted in both the *TopBP1*$^{flox/flox/+}$/*Ova* $^{CAGTopBP1/+/+}$ clones 1 and 2 and the *TopBP1*$^{flox/flox/+}$ *CAG non-targeted* clones. In total, 214 blasticidin-resistant clones were obtained which were all subjected to genotyping by PCR. None of the clones displayed loss of the *TopBP1* allele as is evident from the amplification of the 2.3kb band across exon 19 and intron 22 (**Figure 4.6 F**).

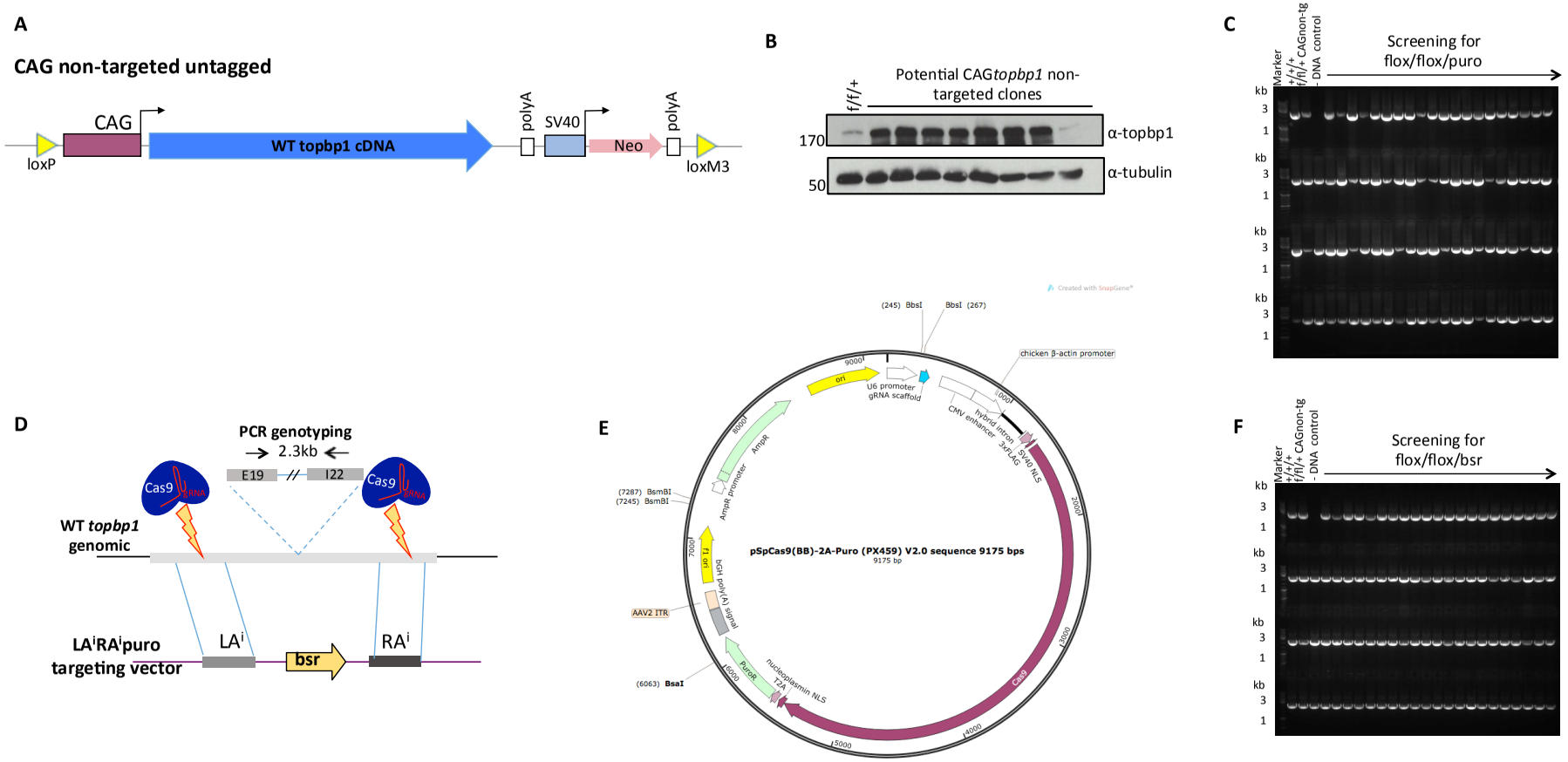

A
CAG non-targeted untagged
CAG
WT topbp1 cDNA
polyA
SV40
Neo
polyA
loxP
loxM3
B
f/f/+
Potential CAGtopbp1 non-targeted clones
170
α-topbp1
50
α-tubulin
C
Marker
+/+/+
f/fl/+ CAGnon-tg
- DNA control
Screening for flox/flox/puro
kb
3
1
D
PCR genotyping
2.3kb
E19
I22
Cas9
gRNA
WT topbp1 genomic
LAiRAipuro targeting vector
LAi
bsr
RAi
E
pSpCas9(BB)-2A-Puro (PX459) V2.0 sequence 9175 bps
9175 bp
ori
U6 promoter
gRNA scaffold
chicken β-actin promoter
CMV enhancer
hybrid intron
3xFLAG
SV40 NLS
AmpR promoter
AmpR
f1 ori
bGH poly(A) signal
AAV2 ITR
PuroR
nucleoplasmin NLS
T2A
Cas9
(245) BbsI
BbsI (267)
(7287) BsmBI
(7245) BsmBI
(6063) BsaI
F
Marker
+/+/+
f/fl/+ CAGnon-tg
- DNA control
Screening for flox/flox/bsr
kb
3
1

Figure 4.6: Attempt to knockout the third *TopBP1* allele in the newly created TopBP1$^{flox/flox/+}$/ *CAGTopBP1 untagged* non-targeted cell line using the LA^{i}RA^{i}Puro gene targeting vector alone or aided by the use of CRISPR gene editing technology. A) Schematic representation of the targeting construct used for stable non-targeted integration of the CAG-*TopBP1* transgene in the genome of TopBP1$^{flox/flox/+}$cells. **B)** Western blot analysis of cell lysates from cells transfected with the expression construct shown in A. Immunoblotting with an anti-TopBP1 antibody controls for the expression of the exogenous over the endogenous TopBP1 protein. Cell lysates of parental cells were used as a control. The levels of beta-tubulin were used as a loading control. **C)** PCR genotyping of wild-type DT40 and *TopBP1*$^{flox/flox/+}$ parental control alongside puromycin-resistant clones obtained from transfection of TopBP1$^{flox/flox/+}$/*CAGTopBP1untagged* non-targeted cells with the LA^{i}RA^{i}Puro targeting vector. Successfully targeted clones would be negative for the 2.3kb amplified band. **D)** Schematic representation of the CRISPR strategy used with gRNAs designed to recognise sequences within the LA^{i} and RA^{i} regions. **E)** Map of the PX459 CRISPR vector used to subclone the gRNAs at *BbsI* site. The resulting constructs were co-transfected with the LA^{i}RA^{i}bsr targeting construct in both the *TopBP1*$^{flox/flox/+}$/*Ova* $^{CAGTopBP1/+/+}$ and the *TopBP1*$^{flox/flox/+}$ *CAG non-targeted* cell lines. **F)** PCR genotyping of wild-type DT40 and parental controls alongside bleomycin-resistant clones obtained from transfection of the parental cell lines as explained in E. Successfully targeted clones would be negative for the 2.3kb amplified band.

Summary for chapter 4

To increase the TopBP1 protein levels and thus increase the probability of achieving the knock out of the third *TopBP1* allele, I first attempted integrating more copies of the CMV-*TopBP1* transgene within the genome but this proved fruitless in terms of achieving the knockout and also the presence of many copies rendered the floxing of the integrated transgenes impossible. The next set of experiments I performed involved a characterisation of the relative efficiencies of different constitutive promoter elements (CMV, CBA, CAG) in an isogenic background. With this characterisation I achieved the design of a stably integrated overexpression system (SIOS) that provides a range of potential expression levels of the *TopBP1* transgene. SIOS is also a stable system amenable to recombinase-mediated cassette exchange thus rendering it an easy to use versatile system for various experimental purposes in the field of DT40. Finally, I went on to show that various attempts to obtain the knockout cell line –including using several gene targeting constructs, targeted or non-targeted integrations of the rescue construct and CRISPR technology- were unsuccessful, despite the optimised exogenous expression of TopBP1. Overall the development of the *TopBP1*$^{-/-/+}$*MerCreMer* cell line and the SIOS were valuable tools for the subsequent work described in this thesis. Also, the inability to obtain the knockout cell line despite the miscellaneous and intense approaches was informative and intriguing *per se*.

Chapter 5

Novel characterization of the avian TopBP1 mRNA and successful generation of the TopBP1 knockout DT40 model system

Preface

In order to study the function of a gene and its protein product in an experimental system, it is important to eliminate the function of this gene and study the resulting phenotype. Depletion studies of the TopBP1 mRNA using siRNA or shRNA have produced most of the knowledge we now have about the role of the mammalian homologue in the initiation of DNA replication and the signalling and activation of DNA structure-dependent checkpoints. Despite being very informative, adapting RNAi techniques to silence gene expression *in vivo* in mammalian systems also has its disadvantages. Firstly, it does not take into account the existence of potential splice variants of the mRNA of interest, which might be resistant to the RNAi molecule introduced into the system. Secondly, such techniques are transient in nature thus only useful for short-term studies, which also poses the risk of the phenotype being lost due to the inherent instability of the RNAi molecule. Additional disadvantages of transient siRNA studies are the potential off-target effects and incomplete knockdown efficiency. So although RNAi techniques are useful for a short-term evaluation of gene expression, they can never replace knockout models. Knockout models can be powerfully used to create even an entire organism in which a specific gene has been deleted or deactivated in all its body cells. At a cellular level, gene knockouts allow the study of the specific phenotypes generated from loss of function of a gene, in an otherwise normal cell. Although being both labor and time consuming, the application of reverse genetics allows manipulation of the protein of interest at its source, hence the generation of stable experimental systems useful for allele-specific expression, complete loss of function, partial loss of function, mutational and complementation studies.

Chapter 5: Novel characterization of the avian TopBP1 mRNA and generation of a TopBP1 knockout DT40 model system

The failure of achieving a complete knockout of *TopBP1* in DT40 cells was interesting on its own right. Given that the tools that have been developed so far, including the gene knockout strategy and the SIOS system, were tailored to meet the needs of the experimental system and its purposes, I wondered whether the *TopBP1*

transgene expressed from the *Ova* locus, which corresponded to the cDNA sequence of TopBP1 available from the Genome Browser, was incorrect. In all genome browser databases, the *gallus gallus* TopBP1 protein sequence is classified as "predicted", a term used for entries without evidence at protein, transcript, or homology levels. A potential "mistake" in the protein sequence of TopBP1, as this is available on the genome database, could render the ectopic protein being produced from the OvaCAG overexpression system a partly or non-functional one. This would explain why the final *TopBP1* allele of $TopBP1^{flox/flox/+}/Ova^{CAGTopBP1/+/+}$ cells could not be deleted. Alternatively, it is possible that cells that had successfully deleted their intact *TopBP1* copy failed to sustain viability due to some essential function missing from the ectopically expressed *TopBP1* transgene. In either case, the ectopic transgene was unable to provide all the functions of TopBP1 required for survival of DT40 cells. To investigate this hypothesis I embarked on two experiments. I first created a cell line where the human homologue of TopBP1 was overexpressed from the *Ova* locus and attempted to knockout the remaining endogenous copy. Given that the human TopBP1 cDNA has been extensively used in the literature it was less likely that it would be incorrect and indeed it allowed the deletion of the endogenous *TopBP1* allele. The next experiment involved the isolation and characterisation of the *gallus gallus* TopBP1 mRNA sequence and successfully identified a fragment missing from the sequence available on the genome database. The newly identified TopBP1 sequence described herein allowed the creation of the DT40 TopBP1 knockout model system, $TopBP1^{flox/flox/puro}/Ova^{CAGnewTopBP1/+/+}$.

5.1.1: Creation of the DT40 *TopBP1* knockout model system by ectopic overexpression of the human TopBP1 cDNA

In an attempt to identify parts of the primary sequence of the avian TopBP1 that could be incorrect I performed multiple sequence alignments with the sequences of other mammalian species (Appendix 2) or indeed other organisms (data not shown). Alignment of the chicken and human TopBP1 protein sequences showed that they shared 68% identity (calculated using the ClustalW protein alignment software from alignment presented in Appendix 2). Close observation of the aligned sequences did not reveal any extended discrepancy except for an array of ~14amino acids (aa 519-533) that was absent from *gallus gallus*, partly absent from *mus musculus*, *gorilla gorilla* and

Pan troglodytes (chimpanzee) but present in *Homo sapiens*. One observation that was striking, however, was that the predicted chicken TopBP1 sequence initiated with an arginine (R) rather than a methionine (M) residue, which is the consensus start codon of an mRNA transcript translated by a ribosome in eukaryotes.

As a first way of getting an indication of whether the annotated chicken TopBP1 cDNA was incorrect, the human homologue of TopBP1 (hTopBP1) was stably integrated in the *Ova* locus of *TopBP1*$^{flox/flox/+}$ cells under the control of the CAG promoter and subsequently the knockout of the final endogenous allele attempted. To create the new targeting vector, the OvaCAG construct described in Chapter 4 was modified to replace the annotated chicken *TopBP1* cDNA with the human homologue. As no suitable restriction sites were available, the hTopBP1 cDNA (kindly provided by the Pearl lab) was subcloned at the *KpnI* site of the original pSF-CAG (OG505) expression vector with primers P73/P74. The SV40-neomycin-pA-loxM3-OvaRA fragment was then subcloned at the *PacI* site and the OvaLA-loxM3 at *PvuI*, as described before. The resulting OvaCAGhTopBP1 vector (**Figure 5.1 A**) was linearized by *AhdI* digestion prior to electroporation into the *TopBP1*$^{flox/flox/+}$ cell line. Ten of the clones that successfully formed single colonies following selection in 2mg/ml G418 were subjected to Southern blot analysis of *BamHI*-digested genomic DNA. Hybridisation with probe B showed that 8 out of the 10 clones examined had successfully integrated the OvaCAGhTopBP1 into the *Ova* locus, as evident from the 10.7kb targeted band (**Figure 5.1 B**). Two of these clones were frozen down in liquid nitrogen as *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ clones 1 and 2 and only clone 1 was used for future experiments. Western blot analysis of *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ clones 1 alongside the *TopBP1*$^{flox/flox/+}$ parental cell line confirmed the overexpression of the human transgene (**Figure 5.1 C**).

To attempt knocking out the third endogenous copy of *TopBP1*, LA^{i}RA^{i}Puro targeting vector was linearised by *AhdI* digestion and transfected into *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ clone 1. Transfectants were serially diluted and selected as described before. In total, 41 puromycin-resistant clones were obtained and subjected to genotyping by PCR with primers P40/P41 (**Figure 5.1 D**). One of the clones tested did not produce the 2.3kb band corresponding to amplification of the *TopBP1* locus between exon 19 and intron 22 and which was present in the *TopBP1*$^{flox/flox/+}$ and *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ controls. This was indicative of successful deletion of the remaining *TopBP1* allele (**Figure 5.1 E**). To conclusively

confirm this result, Southern blot analysis of *BglI* or *BglI/ClaI*-digested genomic DNA of the indicated cell lines was performed. Following *BglI* digestion and hybridisation with probe A, the intact *TopBP1* allele was represented by a diagnostic band of 5.7kb, which was expected to move to 6.2kb upon successful targeting. Additionally, *BglI/ClaI*-digestion and hybridisation with probe A generated a 5.7kb band corresponding to the wild-type *TopBP1* allele. This was expected to move to 5.3kb upon successful deletion of the remaining allele. Indeed, the Southern results confirmed the PCR result and the *TopBP1*$^{flox/flox/puro}$*/Ova*$^{CAGhTopBP1/+/+}$ cell line was appropriately frozen down (**Figure 5.1 D, F**, left panel). To my knowledge, this is the first time that a conditional knockout cell line of the *TopBP1* gene has been created.

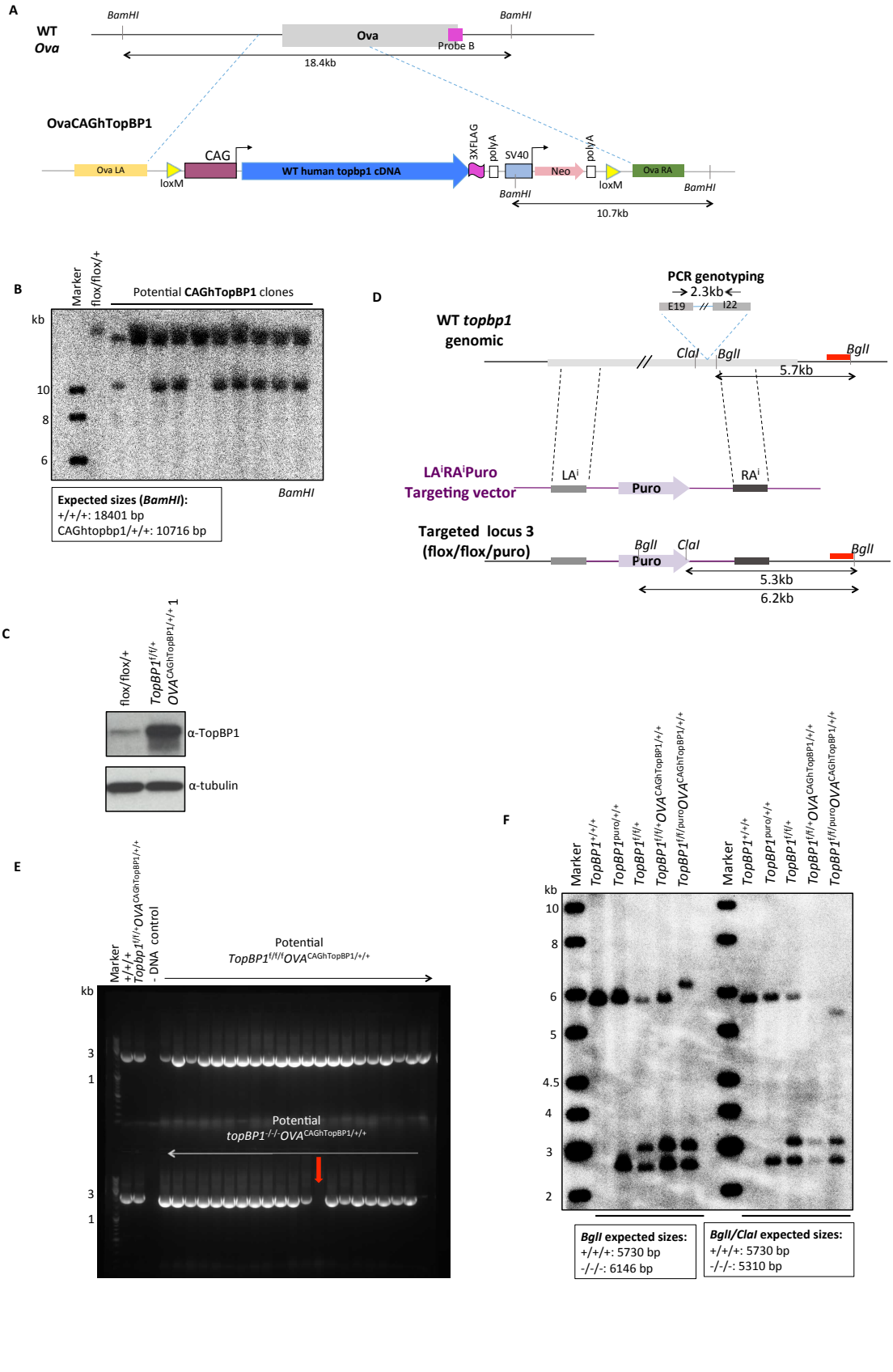

A
WT
Ova
BamHI
Ova
Probe B
BamHI
18.4kb
OvaCAGhTopBP1
Ova LA
loxM
CAG
WT human topbp1 cDNA
3XFLAG
polyA
SV40
Neo
polyA
loxM
Ova RA
BamHI
BamHI
10.7kb
B
Marker
flox/flox/+
Potential CAGhTopBP1 clones
kb
10
8
6
BamHI
Expected sizes (BamHI):
+/+/+: 18401 bp
CAGhtopbp1/+/+: 10716 bp
C
flox/flox/+
TopBP1f/f/+
OVACAGhTopBP1/+/+ 1
α-TopBP1
α-tubulin
D
PCR genotyping
2.3kb
E19
I22
WT topbp1
genomic
ClaI
BglI
BglI
5.7kb
LAiRAiPuro
Targeting vector
LAi
Puro
RAi
Targeted locus 3
(flox/flox/puro)
BglI
ClaI
Puro
BglI
5.3kb
6.2kb
E
Marker
+/+
TopBP1f/f/+OVACAGhTopBP1/+/+
- DNA control
Potential
TopBP1f/f/fOVACAGhTopBP1/+/+
kb
3
1
Potential
topBP1-/-/-OVACAGhTopBP1/+/+
3
1
F
Marker
TopBP1+/+/+
TopBP1puro/+/+
TopBP1f/f/+
TopBP1f/f/+OVACAGhTopBP1/+/+
TopBP1f/f/puroOVACAGhTopBP1/+/+
Marker
TopBP1+/+/+
TopBP1puro/+/+
TopBP1f/f/+
TopBP1f/f/+OVACAGhTopBP1/+/+
TopBP1f/f/puroOVACAGhTopBP1/+/+
kb
10
8
6
5
4.5
4
3
2
BglI expected sizes:
+/+/+: 5730 bp
-/-/-: 6146 bp
BglI/ClaI expected sizes:
+/+/+: 5730 bp
-/-/-: 5310 bp

Figure 5.1: Complete knockout of *TopBP1* by overexpression of the human TopBP1 cDNA. A) Schematic representation of the targeting constructs used to stably integrate the CAGhuman*TopBP1* transgene at the *Ova* locus. **B)** Southern blot analysis of potential *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ clones alongside the *TopBP1*$^{flox/flox/+}$ parental control. The diagnostic bands following *BamHI* digestion of genomic DNA correspond to the wild-type allele (18.4kb) and the successfully targeted allele (10.7kb). Digested genomic DNA was hybridized with the 3' external probe (purple bar) shown in A. **C)** Western blot analysis of cell lysate from one *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ clone positive according to B. Immunoblotting with an anti-TopBP1 antibody controls for the expression of total TopBP1. Cell lysates of the parental cells were used as a negative control. The levels of beta-tubulin were used as a loading control. **D)** Diagrammatic representation of the wild-type *TopBP1* locus, LA^{i}RA^{i}Puro targeting vector and successfully targeted locus of *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ cells. **E)** PCR genotyping of potential *TopBP1*$^{flox/flox/puro}$/*Ova*CAG$^{hTopBP1/+/+}$ clones obtained from transfection of the *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGhTopBP1/+/+}$ with the LA^{i}RA^{i}Puro targeting vector shown in D alongside wild-type and parental controls. Successfully targeted clones would be negative for the 2.3kb amplified band. **F)** Southern blot analysis of *BglI* or *BglI/ClaI*-digested DNA of the *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGhTopBP1/+/+}$ clone identified in E alongside the indicated controls. Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in D.

5.1.2: Novel characterisation of the DT40 TopBP1 mRNA and identification of a novel TopBP1 protein domain

The successful generation of a DT40 system genetically deficient of the *TopBP1* locus by overexpression of the human cDNA suggested that my hypothesis was right. That indeed the inability to create the knockout so far (after having established an optimal ectopic expression system and an efficient knockout strategy) was most likely due to some error or missing sequence in the annotated chicken TopBP1 cDNA. Using the exact same overexpression system (i.e. CAG SIOS) and only replacing the annotated chicken cDNA with the human one was enough to allow the creation of the *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGhTopBP1/+/+}$ conditional knockout.

The challenge now was to identify the region(s) of the chicken TopBP1 primary sequence that was missing or incorrect on the genome browser. To do so, it was necessary to isolate the TopBP1 RNA and identify its sequence, which would prove useful both for the creation of the final *TopBP1* conditional knockout (i.e. by overexpressing the correct chicken protein rather than the human homologue) as well as for the characterisation of the avian TopBP1 protein in general. The experiments performed towards achieving this goal are the focus of the following paragraphs.

In order to obtain the sequence of the TopBP1 cDNA, Rapid Amplification of cDNA Ends (RACE) was used. This is a method that allows the amplification of the 5' or 3'ends of a cDNA molecule of interest and its only requirement is that 23-28 nucleotides of sequence information are known in order to design gene-specific primers (GSPs) for the RACE reaction. I decided to use the SMARTer RACE 5'/3' kit (Clontech), which eliminates the need for adaptor ligation and allows the use of first-strand cDNA directly in RACE PCR. The other advantage of this technology is that it is optimised for using either polyA$^+$ or total RNA as a starting material and it also allows the construction of even long cDNAs.

As mentioned already it was thought bizarre that the TopBP1 protein sequence available on the database contained an arginine as the start codon. In addition to that, the first eight amino acids of TopBP1 were conserved across all the species examined but not in *gallus gallus* (**Figure 5.2 A**). So in order to characterise the 5' end of the avian TopBP1, total RNA was extracted from wild-type DT40 cells using the TRIzol method (described in methods and materials). To first assess total RNA template quality, which is a key determinant of successful RACE, the purified RNA was visualised on a denaturing formaldehyde agarose gel under UV light. Total eukaryotic

RNA of high quality is expected to give a theoretical 28S:18S ratio of approximately 2:1. As shown in **Figure 5.2 C**, the purified RNA was of good quality and so RACE analysis could be performed.

The purified total RNA was used for first strand cDNA synthesis according to the manufacturers' instructions and samples were kept at -20°C. For 5' RACE, seven "reverse" GSPs (F1-F7, **Figure 5.2 B**) were designed that would anneal in a 3' to 5' direction thus allow to obtain sequence information towards the start of the TopBP1 cDNA. There were two reasons why I decided to use more than one GSPs: First, since the GSPs were designed according to the database cDNA sequence of TopBP1, it was possible that any discrepancies with the RACE amplified TopBP1 cDNA would preclude annealing to the template. Second, some GSPs could be suboptimal for successful amplification due to their base composition or secondary structure, despite careful design. So 5' RACE reactions with GSPs F1-F7 were performed according to the manufacturers' instructions. The products obtained from the 5'RACE reaction of each GSP alongside a positive control of mouse heart total RNA were visualised by gel electrophoresis and ethidium bromide staining (**Figure 5.2 D**). Subsequently, the RACE fragments were characterised to make sure that the desired products have been amplified. This involved gel extracting the ethidium bromide-stained bands in Figure 5.2 D and subcloning them in the pRACE vector by In-Fusion cloning (vector provided with the Clontech kit). To obtain the maximum amount of sequence information, 15 different independent clones for each 5' RACE product were characterised by Sanger sequencing using primers suggested by the manufacturer. Analysis of the sequencing data of RACE products revealed some interesting information.

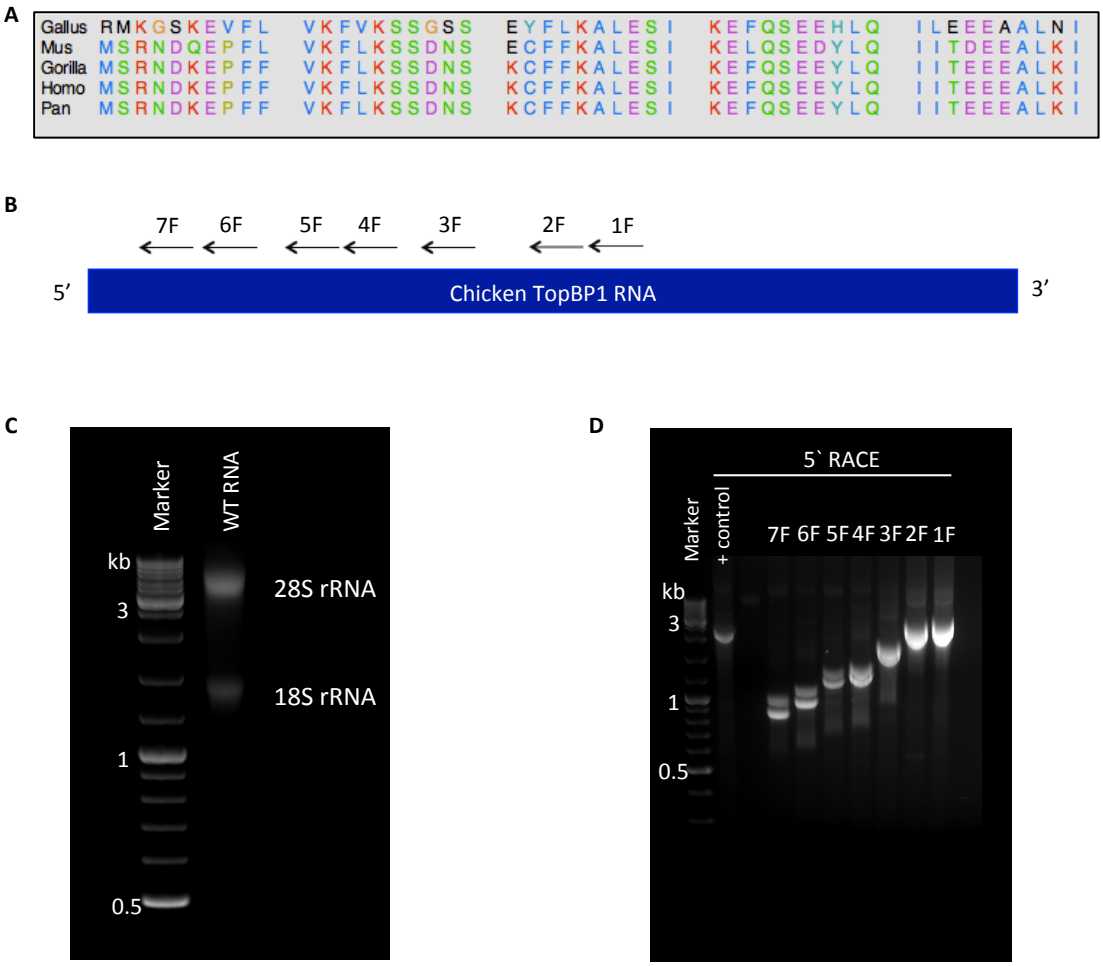


Figure 5.2: 5' RACE of TopBP1 cDNA. A) Alignment of the first 50 amino acids of TopBP1 across different species using the ClastalW alignment software. **B)** Schematic representation of the GSPs used for the 5' RACE reaction. **C)** Gel electrophoresis of a formaldehyde gel stained with ethidium bromide and visualized under UV light to assess the integrity of the total RNA purified from wild-type DT40 cells. **D)** 5' RACE products visualized by gel electrophoresis and ethidium bromide staining. The sample in lane 1 corresponding to mouse total heart RNA served as the positive control.

Identification of a new splicing pattern at the 5' region of the DT40 TopBP1 mRNA

Figure 5.3 A, shows a linear map of the intron/exon pattern of the 5' end of *TopBP1* as well as the start of the TopBP1 ORF ($ORF^{annotated}$), as these are available from the genome browser. The start codon for translation of TopBP1 according to the database genome data is located within the first exon (**Figure 5.3 A** top panel, red asterisk). Therefore the region upstream of the annotated initiation methionine represents the annotated 5' untranslated region (annotated 5' UTR). The 5' UTR of an mRNA molecule is the region that precedes the start codon and is important for translation of the mRNA to protein by the ribosome. Flanking exon 1 are the first and second introns.

Analysis of the RACE sequencing data showed that ~74% of the sequenced clones were missing a fragment of 128 bp (or 128 + 3 bp in some clones) within exon 1 (termed "new intron" in Figure 5.3 A). It is likely that this fragment corresponds to an intronic region as it is flanked by GT/AG, a sequence considered to be the splicing signal for the spliceosome at the intron/exon boundary. Thus, the current analysis defines a region within the exon 1 of the *TopBP1* locus that is intronic, thus splitting this first exon into two exons separated by this 128 bp (or 128 + 3 bp) identified intron. It should be noted that the 3' junction of this new intron was the same for all 70% of the clones and contained a consensus AG splice acceptor site. In contrast, the 5' end varied by 3bp. In other words, most of these clones contained a "new" intron 128 bp-long and a smaller subset had an extra codon at the 5' junction as part of this "new" intron, thus making their intron longer by 3bp. In both cases however, the 5' side of the newly identified intron contained a GT splice donor site (**Figure 5.3 B)**.

The new exon identified upstream of this aforementioned new intron contained a new open reading frame, termed ORF^{short}, which overlapped with the C-terminal end of the database intron 1. The fact that this region was present in the RACE product (thus in the cDNA copy of the TopBP1 mRNA) suggests that it part of the new exon in my RACE data. This data thus propose a different intron/exon junction between the annotated intron 1 and the annotated exon 1 but because of limited sequencing information it was unable for the splicing boundaries to be defined (**Figure 5.3 B**).

Interestingly, about ~10% of the ~74% clones containing ORF^{short}, were missing the oligonucleotide "GTTAAAGG" immediately upstream of the annotated methionine (hence called $ORF^{short\ processed}$), suggesting alternative processing within this region of a

population of TopBP1 RNA molecules. Thus, overall, ~64% of all RACE clones sequenced contained ORF^{short} and ~10% contained $ORF^{shortprocessed}$ (**Figure 5.3 B**).

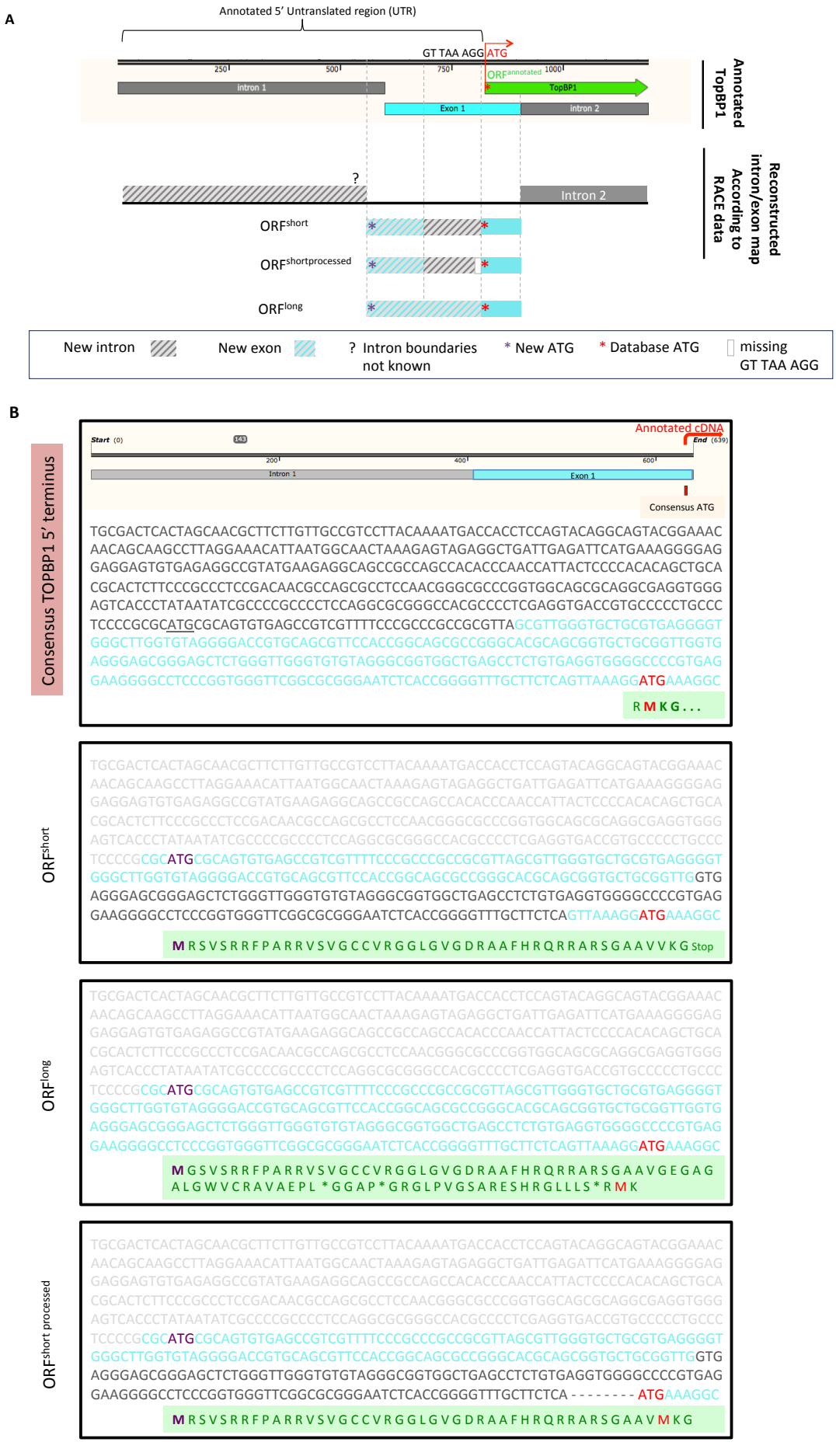
Furthermore, ~24% of the sequenced clones contained the entire sequence corresponding to the database exon 1 (ORF^{long}). This might be indicative of alternative splicing of a subset of TopBP1 transcripts at this particular region (**Figure 5.3 B**).

Two clones (~2%) contained sequences $ORF^{short'}$ and $ORF^{short''}$, which resembled sequence ORF^{short} but were shorter (ORFshort') or slightly different in sequence (ORFshort``). These were most likely a result of incorrect RACE amplification (as they occurred at a very low frequency) but were nonetheless included in the subsequent analysis (**Figure 5.3 B**).

Figure 5.3 C summarises the most abundant N-terminally extended ORFs identified in the present study (ORF^{short}, ORF^{long} and $ORF^{shortprocessed}$), aligned to the consensus TopBP1 ORF from the database ($ORF^{annotated}$) and **Figure 5.3 F** shows representative Chromas data for these ORFs. The differential splicing pattern described in the previous paragraphs concerning the region upstream of the annotated start codon of the TopBP1 transcript suggested that there could be a functional sequence element contained within the newly identified ORFs that was required for complete functionality of the TopBP1 protein. Therefore the results described so far do not propose the existence of a novel TopBP1 sequence absent from the genomic data available for the *TopBP1* locus but rather propose a different splicing pattern at the 5' end. The different arrangements of the splicing events at the 5' region of the genomic TopBP1 locus is shown in **Figure 5.3 D**.

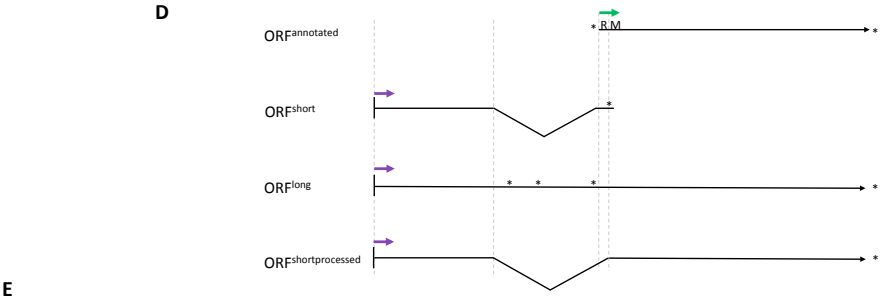
To graphically define the 3' start site of the RACE data across all sequenced clones, I counted how many transcripts start at each respective base within the 5' region of the TopBP1 genomic locus. The graph depicted in **Figure 5.3 E** shows that the majority of the RACE data read through the newly identified ATG (represented by nucleotides 14-16).

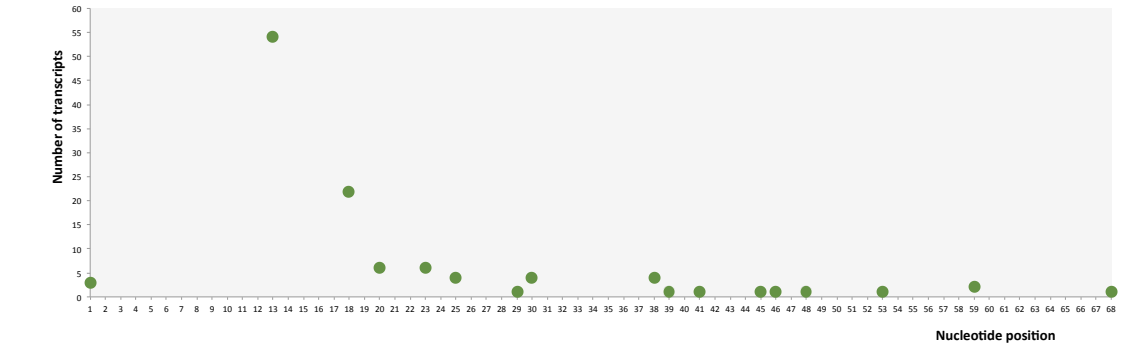
Finally, analysis of the sequencing data revealed that despite the confusion at the 5' end of the TopBP1 mRNA, the region downstream of the R and proceeding consensus M entirely aligned to the known primary sequence and no mismatches were identified. In fact, our sequencing data confirm the consensus amino acid sequence of TopBP1 (and the intron/exon boundaries therein at the nucleotide level) from amino acid 1 to amino acid 908 (data not shown).


A
Annotated 5' Untranslated region (UTR)
GT TAA AGG ATG
ORF annotated
TopBP1
intron 1
Exon 1
intron 2
Annotated TopBP1
Reconstructed intron/exon map According to RACE data
Intron 2
ORF short
ORF shortprocessed
ORF long
New intron
New exon
? Intron boundaries not known
* New ATG
* Database ATG
missing GT TAA AGG
B
Consensus TOPBP1 5' terminus
Annotated cDNA
Intron 1
Exon 1
Consensus ATG
TGCGACTCACTAGCAACGCTTCTTGTTGCCGTCCTTACAAAATGACCACCTCCAGTACAGGCAGTACGGAAAC
AACAGCAAGCCTTAGGAAACATTAATGGCAACTAAAGAGTAGAGGCTGATTGAGATTCATGAAAGGGGAG
GAGGAGTGTGAGAGGCCGTATGAAGAGGCAGCCGCCAGCCACACCCAACCATTACTCCCCACACAGCTGCA
CGCACTCTTCCCGCCCTCCGACAACGCCAGCGCCTCCAACGGGCGCCCGGTGGCAGCGCAGGCGAGGTGGG
AGTCACCCTATAATATCGCCCCGCCCCTCCAGGCGCGGGCCACGCCCCTCGAGGTGACCGTGCCCCCTGCCC
TCCCCGCGCATGCGCAGTGTGAGCCGTCGTTTTCCCGCCCGCCGCGTTAGCGTTGGGTGCTGCGTGAGGGGT
GGGCTTGGTGTAGGGGACCGTGCAGCGTTCCACCGGCAGCGCCGGGCACGCAGCGGTGCTGCGGTTGGTG
AGGGAGCGGGAGCTCTGGGTTGGGTGTGTAGGGCGGTGGCTGAGCCTCTGTGAGGTGGGGCCCCGTGAG
GAAGGGGCCTCCCGGTGGGTTCGGCGCGGGAATCTACCGGGGTTTGCTTCTCAGTTAAAGGATGAAAGGC
R M K G . . .
ORF short
TGCGACTCACTAGCAACGCTTCTTGTTGCCGTCCTTACAAAATGACCACCTCCAGTACAGGCAGTACGGAAAC
AACAGCAAGCCTTAGGAAACATTAATGGCAACTAAAGAGTAGAGGCTGATTGAGATTCATGAAAGGGGAG
GAGGAGTGTGAGAGGCCGTATGAAGAGGCAGCCGCCAGCCACACCCAACCATTACTCCCCACACAGCTGCA
CGCACTCTTCCCGCCCTCCGACAACGCCAGCGCCTCCAACGGGCGCCCGGTGGCAGCGCAGGCGAGGTGGG
AGTCACCCTATAATATCGCCCCGCCCCTCCAGGCGCGGGCCACGCCCCTCGAGGTGACCGTGCCCCCTGCCC
TCCCCGCGCATGCGCAGTGTGAGCCGTCGTTTTCCCGCCCGCCGCGTTAGCGTTGGGTGCTGCGTGAGGGGT
GGGCTTGGTGTAGGGGACCGTGCAGCGTTCCACCGGCAGCGCCGGGCACGCAGCGGTGCTGCGGTTGGTG
AGGGAGCGGGAGCTCTGGGTTGGGTGTGTAGGGCGGTGGCTGAGCCTCTGTGAGGTGGGGCCCCGTGAG
GAAGGGGCCTCCCGGTGGGTTCGGCGCGGGAATCTCACCGGGGTTTGCTTCTCAGTTAAAGGATGAAAGGC
M R S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V V K G Stop
ORF long
TGCGACTCACTAGCAACGCTTCTTGTTGCCGTCCTTACAAAATGACCACCTCCAGTACAGGCAGTACGGAAAC
AACAGCAAGCCTTAGGAAACATTAATGGCAACTAAAGAGTAGAGGCTGATTGAGATTCATGAAAGGGGAG
GAGGAGTGTGAGAGGCCGTATGAAGAGGCAGCCGCCAGCCACACCCAACCATTACTCCCCACACAGCTGCA
CGCACTCTTCCCGCCCTCCGACAACGCCAGCGCCTCCAACGGGCGCCCGGTGGCAGCGCAGGCGAGGTGGG
AGTCACCCTATAATATCGCCCCGCCCCTCCAGGCGCGGGCCACGCCCCTCGAGGTGACCGTGCCCCCTGCCC
TCCCCGCGCATGCGCAGTGTGAGCCGTCGTTTTCCCGCCCGCCGCGTTAGCGTTGGGTGCTGCGTGAGGGGT
GGGCTTGGTGTAGGGGACCGTGCAGCGTTCCACCGGCAGCGCCGGGCACGCAGCGGTGCTGCGGTTGGTG
AGGGAGCGGGAGCTCTGGGTTGGGTGTGTAGGGCGGTGGCTGAGCCTCTGTGAGGTGGGGCCCCGTGAG
GAAGGGGCCTCCCGGTGGGTTCGGCGCGGGAATCTCACCGGGGTTTGCTTCTCAGTTAAAGGATGAAAGGC
M G S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V G E G A G
A L G W V C R A V A E P L * G G A P * G R G L P V G S A R E S H R G L L L S * R M K
ORF short processed
TGCGACTCACTAGCAACGCTTCTTGTTGCCGTCCTTACAAAATGACCACCTCCAGTACAGGCAGTACGGAAAC
AACAGCAAGCCTTAGGAAACATTAATGGCAACTAAAGAGTAGAGGCTGATTGAGATTCATGAAAGGGGAG
GAGGAGTGTGAGAGGCCGTATGAAGAGGCAGCCGCCAGCCACACCCAACCATTACTCCCCACACAGCTGCA
CGCACTCTTCCCGCCCTCCGACAACGCCAGCGCCTCCAACGGGCGCCCGGTGGCAGCGCAGGCGAGGTGGG
AGTCACCCTATAATATCGCCCCGCCCCTCCAGGCGCGGGCCACGCCCCTCGAGGTGACCGTGCCCCCTGCCC
TCCCCGCGCATGCGCAGTGTGAGCCGTCGTTTTCCCGCCCGCCGCGTTAGCGTTGGGTGCTGCGTGAGGGGT
GGGCTTGGTGTAGGGGACCGTGCAGCGTTCCACCGGCAGCGCCGGGCACGCAGCGGTGCTGCGGTTGGTG
AGGGAGCGGGAGCTCTGGGTTGGGTGTGTAGGGCGGTGGCTGAGCCTCTGTGAGGTGGGGCCCCGTGAG
GAAGGGGCCTCCCGGTGGGTTCGGCGCGGGAATCTCACCGGGGTTTGCTTCTCA - - - - - - - - ATGAAAGGC
M R S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V M K G

C

Name	Amino acid sequence
ORF^{annotated}	----------R M TopBP1
ORF^{short}	M R S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V V ---------- K G *
ORF^{long}	M G S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V G E G A G A L G W V C R A V A E P L * G G A P G * R G L P V G S A R E S H R G L L L S * R M TopBP1
ORF^{shortprocessed}	M R S V S R R F P A R R V S V G C C V R G G L G V G D R A A F H R Q R R A R S G A A V ---------- M TopBP1





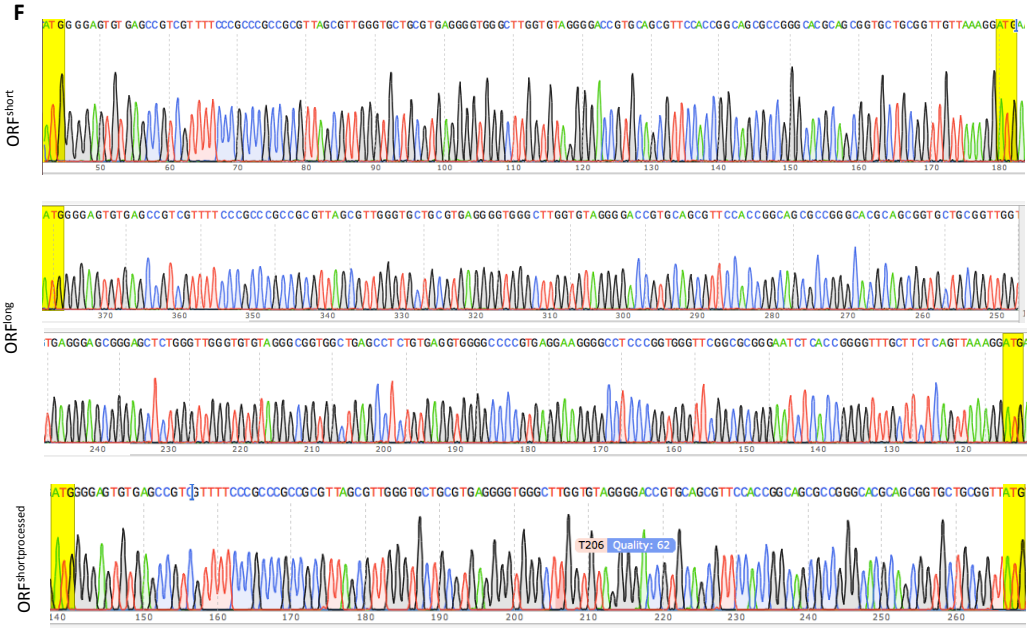


Figure 5.3: Analysis of 5' RACE data and identification of a new splicing pattern of the 5' region of the TopBP1 RNA. A) Schematic representation of the consensus *TopBP1* genomic region at the start of the *TopBP1* locus (top panel). The bottom panel represents the reconstructed intron/exon map based on the RACE data. **B)** Newly identified ORFs upstream of the annotated ATG. ORF^{short} was the most abundant among the 5' RACE data, followed by ORF^{long} and ORF^{shortprocessed}. Colour coding; cyan denotes exons, grey denotes introns and green denotes protein sequence. **C)** Amino acid sequence alignment of the consensus 5' TopBP1 sequence (ORF^{annotated}) and the extended versions of the TopBP1 protein (ORF^{short}, ORF^{long} and ORF^{shortprocessed}) identified in the present study. **D)** Diagrammatic representation of the splicing events in the 5' region of the TopBP1 gene locus identified in the present study. **E)** Graphic representation of the RACE sequencing start site. Base 1 is defined as the 3'-most nucleotide identified from the sequenced RACE clones. Newly identified ATG is represented by nucleotide positions 14-16. **F)** Chromas data of the sequenced RACE-identified ORFs.

5.1.3: Creation of the DT40 *TopBP1* knockout model system by ectopic overexpression of the 5'RACE-identified TopBP1 protein sequence

The newly identified 5' sequences were assembled on the rest of the TopBP1 cDNA sequence and new OvaCAG overexpression vectors were constructed. The resulting ovalbumin targeting constructs OvaCAGnewTopBP1_ORF^{short}, OvaCAGnewTopBP1_ORF^{long}, OvaCAGnewTopBP1_ORF^{shortprocessed}, OvaCAGnewTopBP1_ORF$^{short'}$and OvaCAGnewTopBP1_ORF$^{short''}$, (**Figure 5.4 B**) were confirmed by Sanger sequencing.

The constructs were linearized by *ApaLI* digestion prior to electroporation into the *TopBP1*$^{flox/flox/+}$ cell line. For each construct, ten of the clones that successfully formed single colonies following selection in 2mg/ml G418 were subjected to PCR analysis with primers P75/P76 to identify the ones that have specifically targeted the *Ova* locus. P75 annealed within the selection cassette whereas P76 annealed to a chromosomal region outside the OvaRA thus a successfully *Ova* targeted clone would produce a PCR product of ~3kb in length (**Figure 5.4 B**). Gel electrophoresis of the PCR products identified the targeted clones (**Figure 5.4 C**). To confirm this result, two PCR positive clones corresponding to each construct was subjected to Southern blot analysis of *BamHI*-digested genomic DNA. Hybridisation with probe B showed that all of the clones examined had successfully integrated the corresponding OvaCAGnewTopBP1 construct into the *Ova* locus, as evident from the 10.7kb targeted band (**Figure 5.4 D**). All the clones were frozen down in liquid nitrogen for long-term storage.

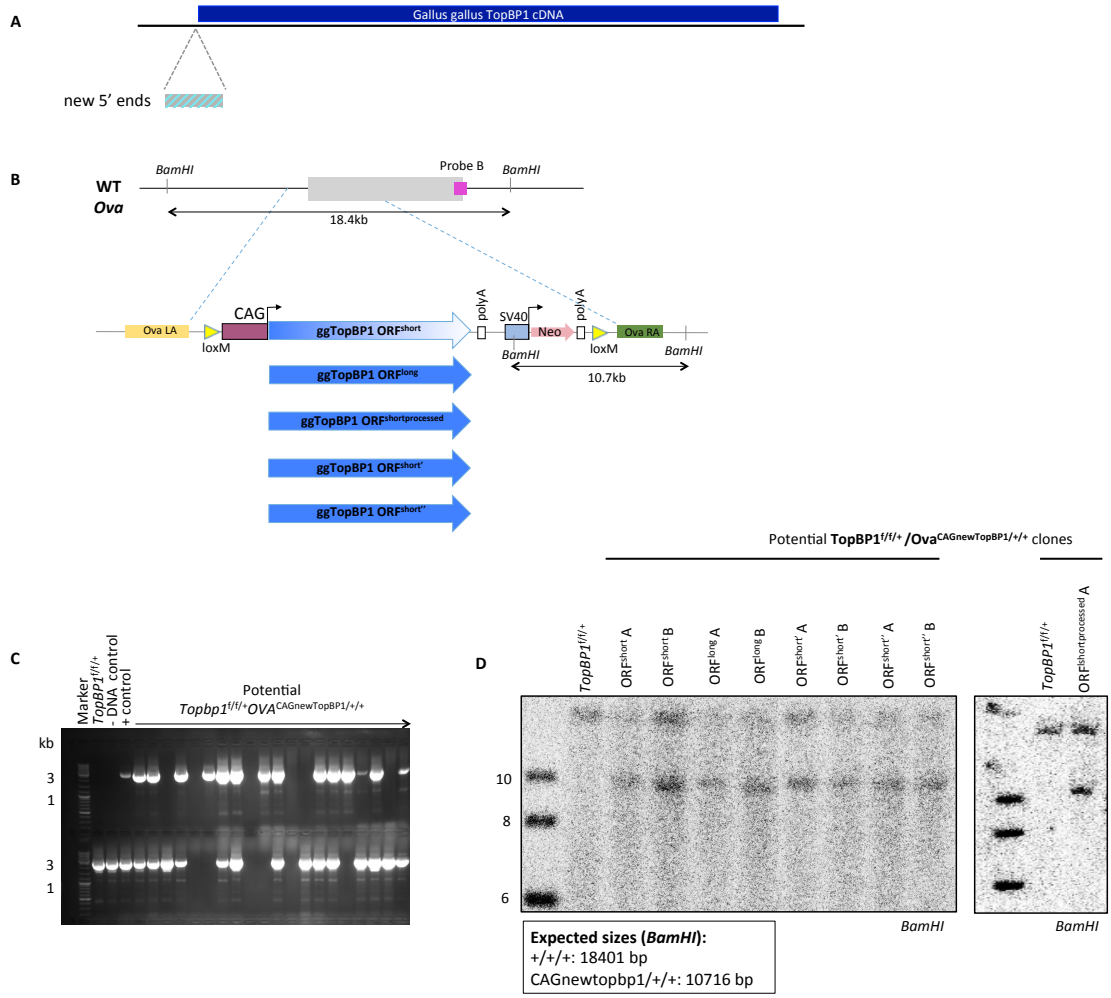


Figure 5.4: Stable integration of the 5' RACE-identified TopBP1 cDNAs in the ovalbumin locus of *TopBP1*$^{flox/flox/+}$ cells. A) Diagrammatic representation of the cloning strategy devised for the assembly of the new cDNA targeting constructs. **B)** Schematic representation of the targeting constructs used to stably integrate the newly assembled CAGnew*TopBP1* constructs at the *Ova* locus. **C)** PCR screening of clones that have successfully targeted the *Ova* locus. **D)** Southern blot analysis of potential *TopBP1*$^{flox/flox/+}$ *Ova* $^{CAGnewTopBP1/+/+}$ clones alongside the *TopBP1*$^{flox/flox/+}$ parental control. The diagnostic bands following *BamHI* digestion of genomic DNA correspond to the wild-type allele (18.4kb) and the successfully targeted allele (10.7kb). Digested genomic DNA was hybridized with the 3' external probe (purple bar) shown in B.

To knock out the third endogenous copy of *TopBP1*, LA^{i}RA^{i}Puro targeting vector was linearised by *AhdI* digestion and transfected into the respective *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGnewTopBP1/+/+}$ cell line (**Figure 5.5 A**). Transfectants were serially diluted and selected as described before. In total, 437 puromycin-resistant clones were obtained and subjected to genotyping by PCR with primers P40/P41. It should be noted that 137 potential *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshort/+/+}$, 35 potential *TopBP1*$^{flox/flox/}$puro/*Ova*CAGnewTopBP1$^{ORFlong/+/+}$, 141 potential *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshort'/+/+}$, 124 potential *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshort''/+/+}$ clones and 83 potential *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshortprocessed/+/+}$ clones were screened by PCR. Six of the clones tested did not produce the 2.3kb band corresponding to amplification of the *TopBP1* locus between exon 19 and intron 22 and which was present in the *TopBP1*$^{-/-/+}$ and *TopBP1*$^{flox/flox/+}$/*Ova*$^{CAGnewTopBP1/+/+}$ parental controls. This was indicative of successful deletion of the remaining *TopBP1* allele (**Figure 5.5 B**). In parallel, OvaCAG (i.e. the construct harbouring the annotated TopBP1 cDNA) was transfected into the same cells and all of the 95 clones obtained were negative, which served as a control (data not shown). To conclusively confirm this result, Southern blot analysis of *BglI* or *BglI/ClaI*-digested genomic DNA of the PCR-positive cell lines was performed. Following *BglI* digestion and hybridisation with probe A, the intact *TopBP1* allele was represented by a diagnostic band of 5.7kb, which was expected to move up to 6.2kb upon successful targeting. Additionally, *BglI/ClaI*-digestion and hybridisation with probe A generated a 5.7kb band corresponding to the wild-type *TopBP1* allele. This was expected to move down to 5.3kb upon successful deletion of the remaining allele. The Southern blot results confirmed that two of the six PCR-positive clones (clones 4 and 5) have successfully knocked out the remaining *TopBP1* chromosomal locus (**Figure 5.5 C**). These contained "ORF^{short}" or "ORF^{shortprocessed}" TopBP1 cDNA integrated at the *Ova* locus. Unsuccessful PCR reaction due to the lack of sufficient genomic DNA in the sample could explain the false positive result obtained for the other four clones (clones 1-3 and 6). Finally, *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshort/+/+}$ as well as *TopBP1*$^{flox/flox/puro}$/*Ova*$^{CAGnewTopBP1ORFshortprocessed/+/+}$ knockout clones were appropriately frozen down in liquid nitrogen. Finally, it is worth pointing out that the small number of clones obtained from transfection of the LA^{i}RA^{i}Puro targeting vector

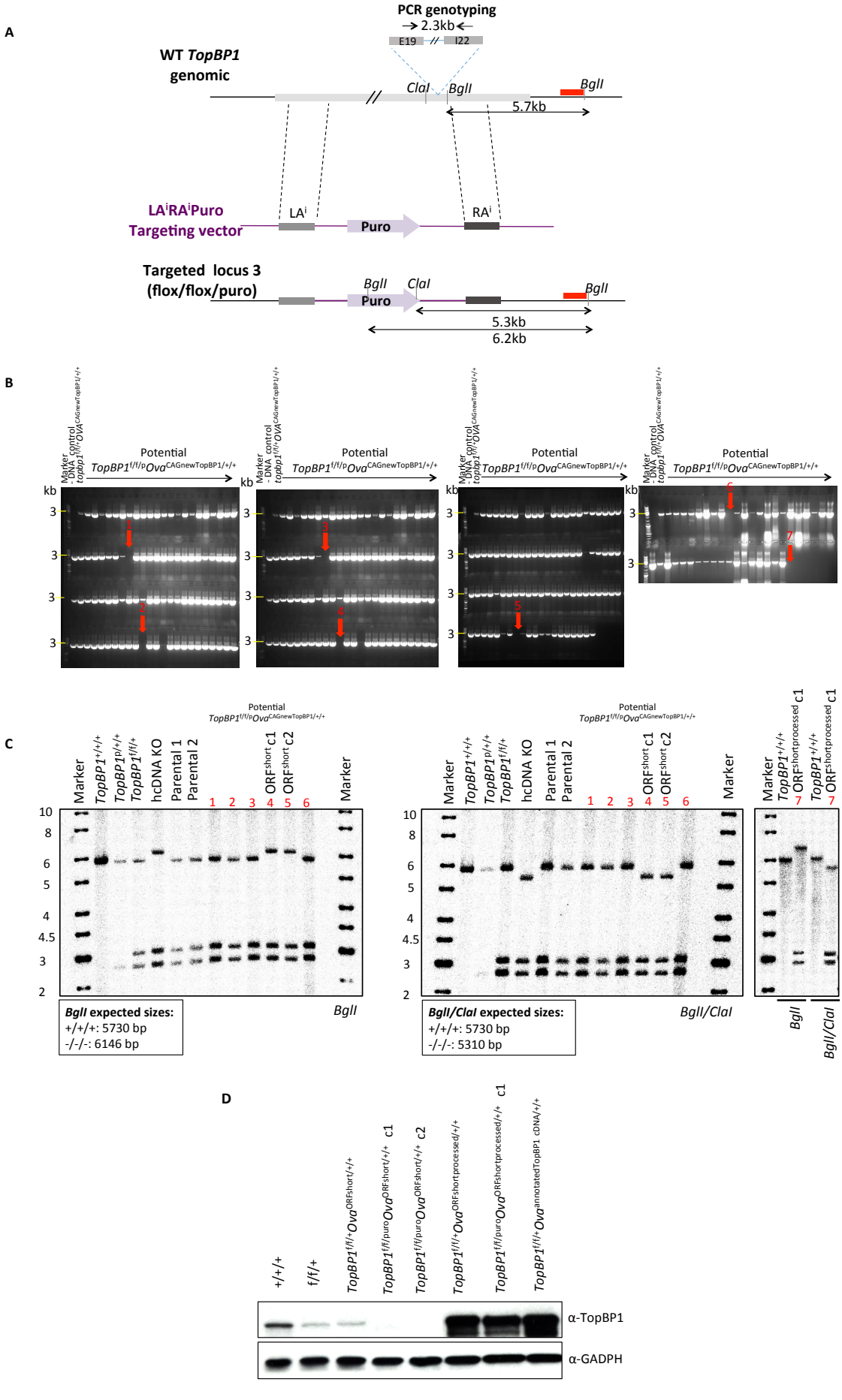

A
PCR genotyping
2.3kb
E19
I22
WT TopBP1 genomic
ClaI
BglI
BglI
5.7kb
LAiRAiPuro Targeting vector
LAi
Puro
RAi
Targeted locus 3 (flox/flox/puro)
BglI
ClaI
Puro
BglI
5.3kb
6.2kb
B
Marker
DNA control
topbp1f/f/pOVACAGnewTopBP1+/+
Potential
TopBP1f/f/pOvaCAGnewTopBP1/+/+
kb
3
1
2
3
4
5
6
7
C
Marker
TopBP1+/+
TopBP1p/+
TopBP1f/f/+
hcDNA KO
Parental 1
Parental 2
ORFshort c1
ORFshort c2
ORFshortprocessed c1
10
8
6
5
4
4.5
3
2
BglI expected sizes:
+/+/+: 5730 bp
-/-/-: 6146 bp
BglI
BglI/ClaI expected sizes:
+/+/+: 5730 bp
-/-/-: 5310 bp
BglI/ClaI
D
+/+/+
f/f/+
TopBP1f/f/+OvaORFshort/+/+
TopBP1f/f/puroOvaORFshort/+/+ c1
TopBP1f/f/puroOvaORFshort/+/+ c2
TopBP1f/f/+OvaORFshortprocessed/+/+
TopBP1f/f/puroOvaORFshortprocessed/+/+ c1
TopBP1f/f/+OvaannotatedTopBP1 cDNA/+/+
α-TopBP1
α-GADPH

Figure 5.5: Complete knockout of *TopBP1* by overexpression of the 5' RACE-identified avian protein sequences. A) Diagrammatic representation of the wild-type *TopBP1* locus, LA^{i}RA^{i}Puro targeting vector and successfully targeted locus of *TopBP1*$^{-/-/+}$/*Ova*$^{\text{CAGnewTopBP1/+/+}}$ cells. **B)** PCR genotyping of potential *TopBP1*$^{-/-/-}$/*OVA*$^{\text{CAGnewTopBP1/+/+}}$ clones obtained from transfection of the *TopBP1*$^{-/-/+}$/*OVA*$^{\text{CAGnewTopBP1/+/+}}$ with the LA^{i}RA^{i}Puro targeting vector shown in A alongside wild-type and parental controls. Successfully targeted clones would be negative for the 2.3kb amplified band. C) Southern blot analysis of *BglI* or *BglI/ClaI*-digested DNA of the six *TopBP1*$^{-/-/-}$/*OVA*$^{\text{CAGnewTopBP1/+/+}}$ clones identified in E alongside the indicated controls. Digested genomic DNA was hybridized with the 3' external probe (red bar) shown in A. D) Western blot analysis of *TopBP1*$^{\text{flox/flox/+}}$/*Ova*$^{\text{ORFshort/+/+}}$ and *TopBP1*$^{\text{flox/flox/+}}$/*Ova*$^{\text{ORFshortprocessed/+/+}}$ knockout. clones Immunoblotting with an anti-TopBP1 antibody controls for the expression of total TopBP1. Cell lysates of TopBP1$^{+/+/+}$, TopBP1$^{\text{flox/flox/+}}$ and *TopBP1*$^{\text{flox/flox/puro}}$/*Ova*$^{\text{CAGnewTopBP1ORFshort/+/+}}$ & *TopBP1*$^{\text{flox/flox/puro}}$/*Ova*$^{\text{CAGnewTopBP1ORFshortprocessed/+/+}}$ parental cells were used as a negative control. The levels of GADPH were used as a loading control.

into the *TopBP1*$^{flox/flox/+}$*OVA*$^{CAGnewTopBP1\mathbf{ORFlong}/+/+}$ cell line prevents me from drawing a conclusion concerning the ability of this sequence (sequence ORF^{long}, Figure 5.3 B) in providing the same function(s) required to achieve the knock out. Future experiments are required to assess the function of ORF^{long} by screening more cells for the knockout targeting.

Western blot analysis of the knockout clones alongside their parental controls reveals that only *TopBP1*$^{flox/flox/puro}$*/Ova*$^{CAGnewTopBP1ORFshortprocessed/+/+}$ cells show an overexpression pattern when immunoblotting with an anti-TopBP1 antibody. In contrast, *TopBP1*$^{flox/flox/puro}$*/Ova*$^{CAGnewTopBP1\mathbf{ORFshort}/+/+}$ cells show no TopBP1 corresponding band on the immunoblot (**Figure 5.5 D**). This is presumably due to the fact that ORF^{short} is out of frame with the annotated TopBP1. Closer look to Figure 5.3 B reveals that if translation starts at the novel methionine of ORF^{short} then this creates a stop codon at the start of the annotated TopBP1. This raises the possibility of the newly identified domain upstream of the annotated TopBP1 (ORF^{short} herein) being capable of providing the essential function required for DT40 viability on its own (without the rest of the TopBP1 protein sequence). Although the present data are a good indication towards this scenario, future work is required to prove it.

Overall, the 5' RACE results described herein and the new TopBP1 cDNA assembled provided the necessary function(s) of the protein required for achieving deletion of the endogenous *TopBP1* gene and sustaining viability of DT40 cells.

Summary of Chapter 5

This chapter describes the experiments performed to investigate the hypothesis of the consensus TopBP1 cDNA missing some function(s) necessary for allowing the knock out of the intact *TopBP1* allele of *TopBP1*$^{flox/flox/+}$ cells. An initial good indication in support of this hypothesis came from achieving the complete knock out of *TopBP1* by overexpressing the human cDNA using the SIOS system described before. The next challenge was to examine the primary sequence of the TopBP1 mRNA present in wild-type DT40 cells. This involved the isolation and characterisation of the 5' end of the TopBP1 cDNA using 5' RACE. Analysis of RACE sequencing results revealed different intron/exon boundaries and subsequently a different splicing pattern of the TopBP1 mRNA compared to the consensus. Additionally, an array of ~43 amino acid lying upstream of the annotated translational start site was identified in the majority of the sequenced RACE clones. Stable integration of this newly identified sequence of TopBP1 at the *Ova* locus using the SIOS system allowed the generation of the complete knock out of *TopBP1*. Thus the avian TopBP1 mRNA identified herein contained the necessary function(s) to sustain viability of DT40 cells in the absence of the endogenous protein.

Chapter 6

A novel knock in gene targeting system based on the gene dosage-dependent functions of TopBP1

Preface

Like most of the key proteins involved in the checkpoint signaling, TopBP1 was initially identified and has been extensively studied in *S. pombe* and *S. cerevisiae* yeast model organisms, where it is known as *Rad4/Cut5* and *Dpb11*, respectively. Depletion studies have also been used extensively to unravel the functions of TopBP1 in mammalian cells. However, the lack of readily available tools, including genetically defined mutant cell lines, has rendered the characterisation of TopBP1 in higher eukaryotes more challenging. In addition to gene knockout, gene knock inmodel systems

Chapter 6: TopBP1 functions are gene dosage-dependent

To unravel the functions of TopBP1 in cell proliferation, survival, DNA damage-dependent checkpoint activation and damage repair, it is important to create separation of function mutants using genetically defined model systems. In addition to the TopBP1 knockout model systems created (see Chapter 5), I wanted to generate a knock in system as well. The advantage of such a system is that it allows the study of a gene at its source by incorporating point mutations of interest at the endogenous locus. Thus, it is more physiologically relevant and together with the knockout system represent powerful tools for a structure-function analysis of the TopBP1 protein in DT40.

Theoretically, one could use the TopBP1 locus to knock in a point mutation(s) of interest to all three TopBP1 (+/+/+) alleles on chromosome II thus replacing the entire wild-type TopBP1 protein population by the mutant and studying the resulting phenotypes. Alternatively, one of the three endogenous alleles are knocked out (-/+/+) and the remaining is (are) used to knock in the point mutation(s) of interest. Interestingly, such a system can be used for recapitulating both homozygous and heterozygous mutants. A third scenario would be to delete two of the three TopBP1 alleles and use the third one as the knock in platform. Combining any of the +/+/+, -/+/+ or -/-/+ system with the SIOS system developed in Chapter 4, also allows the study of lethal knock in point mutations.

To create this *in vivo* TopBP1 gene knock in platform, an isogenic set of stable cell lines from the chicken B cell line DT40 was established by targeted deletion of the

TopBP1 alleles. This chapter describes the generation of the +/+/+, -/+/+ and -/-/+ novel knock in platforms as well as the study of the kinetics of the events induced by progressive loss of function of TopBP1. Thus, it provides a basic characterisation of the knock in platforms *per se,* prior to the integration of any point mutations.

6.1.1: Successive deletion of the *TopBP1* alleles in an isogenic set of stable cell lines results in successively decreasing levels of the TopBP1 protein

In chapter 3 I described the generation of the $TopBP1^{-/+/+}$ and $TopBP1^{-/-/+}$ from a $TopBP1^{+/+/+}$ wild-type DT40 parental strain. For these cell lines to only differ in their respective *TopBP1* status and be otherwise isogenic, pANMerCreMer was transfected in $TopBP1^{+/+/+}$ wild-type cells as well, in the same way as described before. The newly generated DT40 cell lines were expanded and appropriately frozen down in liquid nitrogen as $TopBP1^{+/+/+}$ (isogenic), $TopBP1^{flox/+/+}$ (isogenic) and $TopBP1^{flox/flox/+}$ (isogenic), but for the shake of simplicity will be referred to as $TopBP1^{+/+/+}$, $TopBP1^{-/+/+}$ and $TopBP1^{-/-/+}$ (or simply +/+/+, -/+/+ and -/-/+ in figures). All work described in the following paragraphs makes use of these three isogenic cell lines. A schematic representation of the three flavors of *TopBP1* alleles is shown in **Figure 6.1 A**.

Having established a stable and isogenic cell-based system with varying copies of the *TopBP1* gene, I wanted to investigate the effect, if any, of *TopBP1* gene dosage on the TopBP1 protein levels. Whole cell lysates were prepared from the three cell lines and analysed by Western blot with an anti-TopBP1 antibody raised against a C-terminal peptide of the protein (Appendix 1). Immunoblot analysis indicated that the successive deletion of the *TopBP1* alleles led to a concomitant decrease in the TopBP1 levels inside the cells. In fact, $TopBP1^{-/+/+}$ expressed ~67% and $TopBP1^{-/-/+}$ expressed ~37% of the levels found in the parental wild-type cells (**Figure 6.1 B, C**). Hence cells did not compensate for the loss of each *TopBP1* copy by increasing transcription from the remaining loci, as judged by the total TopBP1 protein levels in the three cell lines. This struck my attention and I decided to use these three isogenic cell lines as a novel system to study the effects of reducing TopBP1 levels on various aspects of cellular biology. **Figure 6.1 D** shows a diagrammatic representation of this novel TopBP1 system in the DT40 model organism, which is the focus of the work described in the following pages.

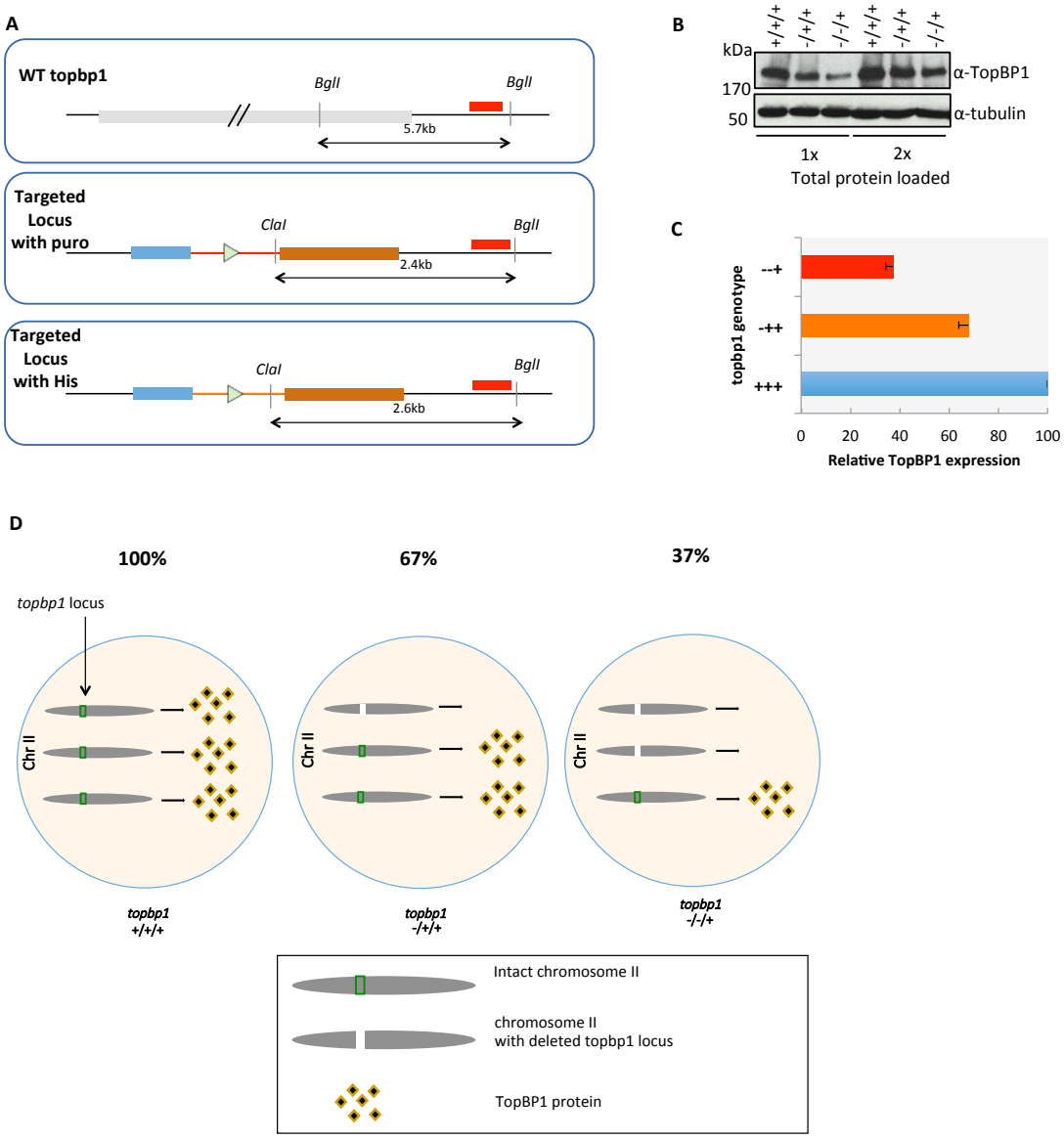


Figure 6.1: Deletion of the *TopBP1* alleles leads to a step-wise decrease in the TopBP1 protein levels. A) Schematic representation of the wild-type and deleted TopBP1 alleles on chromosome II of DT40. The *TopBP1*$^{+/+/+}$ cell line possesses three wild-type copies, the *TopBP1*$^{-/+/+}$ possesses two wild-type and one replaced by the puro cassette which has been floxed and *TopBP1*$^{-/-/+}$ possesses one wild-type, one puro-floxed and one his-floxed. **B)** Western blot analysis of cell lysates of the indicated genotypes. Immunolotting with an anti-TopBP1 antibody controls for the visualisation of the decrease of TopBP1 protein levels with the successive deletion of the *TopBP1* gene copies. Beta-tubulin serves as the loading control. **C)** Relative expression of the TopBP1 protein based on a quantification of the immunoblot in B. The levels found in the mutants are normalised to the levels found in wild-type cells. Also, the bar chart represents the average of the 1x and 2x panels that represent the total amount of protein loaded on each well. Error bar represent the standard deviation from the mean. **D)** A schematic representation of the three stable cell lines producing different levels of the TopBP1 protein from the *TopBP1* locus on chromosome II.

6.1.2: Successive deletion of the *TopBP1* alleles leads to altered replication properties

Being an essential protein for the initiation of replication, the first question I sought to address with the three-strain *TopBP1*-dosage system was whether the reduction of the TopBP1 levels impacted on the proliferative ability of cells. Although *TopBP1*$^{-/-/+}$ were viable, they multiplied more slowly than *TopBP1*$^{+/+/+}$ cells during exponential phase growth, whereas the *TopBP1*$^{-/+/+}$ cell line showed an intermediate rate of proliferation (**Figure 6.2 A**). This defect becomes more obvious if the proliferation rate is used to deduce the doubling time, the period of time required for the cell number of the respective strains in culture to double. Although it takes 8.2h for *TopBP1*$^{+/+/+}$ cells to double, *TopBP1*$^{-/+/+}$ double in 8.9h and *TopBP1*$^{-/-/+}$ in 10.8h. Thus to complete a full cell cycle *TopBP1*$^{-/+/+}$ cells need 0.7h and *TopBP1*$^{-/-/+}$ cells need 2.6h more than *TopBP1*$^{+/+/+}$ (**Figure 6.2 B**). This proliferation defect is also associated with an accumulation of more dead cells in the mutant cultures compared to the parental control, as determined by trypan blue exclusion and optical microscopy. Interestingly, dead cells in the *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cultures started to appear after the first day of seeding and gradually expanded as cells were kept in tissue culture for 72h (**Figure 6.2 C**). What is also worth noticing is that the differences observed in the growth rate and viability are bigger between the *TopBP1*$^{-/-/+}$ and *TopBP1*$^{-/+/+}$ than between the *TopBP1*$^{-/+/+}$ and the *TopBP1*$^{+/+/+}$, perhaps suggesting that the contribution of the deleted alleles to the observed phenotypes are not proportional.

As another way to assess their proliferation properties, the three cell lines were plated on semi-solid media and assessed for their ability to form colonies. The allele-dependent reduction in the levels of the TopBP1 protein impaired the clonogenic ability of mutant cells. In fact, although 80% of the *TopBP1*$^{+/+/+}$ cells seeded managed to form colonies in methylcellulose-containing growth media, 69% of *TopBP1*$^{-/+/+}$ and 42% of *TopBP1*$^{-/-/+}$ did (**Figure 6.2 D**). Finally, to determine the relative proportions of G1, S and G2/M phase cells and whether loss of *TopBP1* alleles affected the progression through the cell cycle, the three cell lines were subjected to Fluorescence-activated cell sorting (FACS) using the nucleic acid stain *propidium iodide (PI). Wild-type TopBP1*$^{+/+/+}$ cells contained 41%, 13% and 31% of G1, S and G2 phase cells, respectively. Interestingly, successive deletion of the *TopBP1* alleles led to a modest increase in the proportion of cells in S at the expense of G1 and G2/M as determined by DNA content (**Figure 6.2 E**). *Thus the overall slower multiplication of TopBP1*$^{-/+/+}$ and

to a greater extent slower multiplication of *TopBP1*$^{-/+/+}$ is due to a combination of slight alterations in cell cycle phases distribution and spontaneous cell death.

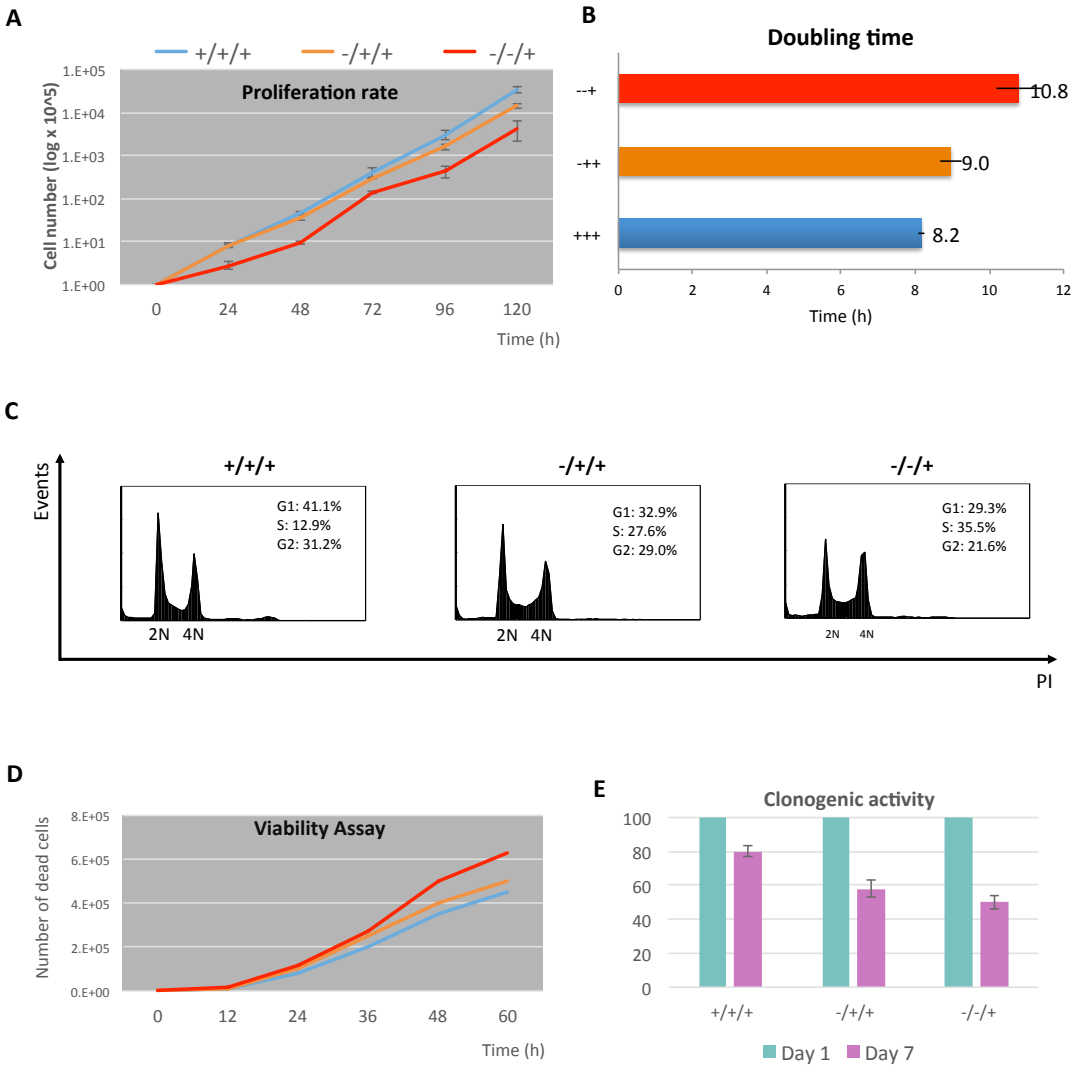


Figure 6.2: Proliferation properties of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells. A) Growth curves for the indicated cell cultures. **B)** Doubling time of the indicated cell lines calculated from the data in A. **C)** Accumulation of dead cells with time as determined by trypan blue staining and optical microscopy. **D)** Clonogenic ability of indicated cells grown in semi-solid media. Equal number of cells was seeded on day 1 and left to grow before colonies were visible and counted at the bottom of the plates. **E)** Flow cytometry analysis of asynchronous cultures fixed and stained with propidium iodide. All error bars show the standard deviation of the mean for three independent experiments.

6.1.3: Successive deletion of the *TopBP1* alleles renders cells more sensitive to killing with DNA damaging agents

To investigate the function of *TopBP1* gene-dosage on the sensitivity to DNA damaging agents, viability and survival assays were performed with a battery of genotoxic agents including UV, MMS, IR and HU. Such agents exert different effects on the DNA; UV causes structural alterations which distort base pairing, MMS is an alkylating agent causing DNA damage by stalling replication forks, IR directly causes DSBs and HU is a potent inhibitor of the ribonucleotide reductase holoenzyme thus leading to depletion of the dNTPs pools and replication stress. Firstly, the viability of $TopBP1^{-/+/+}$ and $TopBP1^{-/-/+}$ cells compared to their $TopBP1^{+/+/+}$ counterparts following treatment with the aforementioned damaging agents was measured using alamarBlue®.

Successive reduction of the *TopBP1* gene dosage was correlated with a stepwise increasing sensitivity to all the genotoxic agents tested (**Figure 6.3 A)**. More specifically, mutant strains seem to tolerate concentrations of HU up to 50μM comparable to wild-type cells but at increasing concentrations of HU the sensitivity of the mutants becomes apparent. It is worth noticing that the drop in the viability of the $TopBP1^{-/-/+}$ cells at 100μM HU is quite abrupt compared to $TopBP1^{-/+/+}$ cells whereas the drop in the viability of the latter compared to $TopBP1^{+/+/+}$ is quantitatively less severe. A similar pattern is observed following treatment of the three cell lines with ionizing radiation (IR) and ultraviolet light (UV), with $TopBP1^{-/-/+}$ cells exhibiting a more increased sensitivity to killing than $TopBP1^{-/+/+}$ cells, compared to the $TopBP1^{+/+/+}$. For viability following MMS treatment, the big error bars make it hard to get a clear picture of the precise pattern of the graphs but what is clear from the data is that the cell lines expressing less TopBP1 are significantly more sensitive than the $TopBP1^{+/+/+}$. Notably, in these assays sensitivity was measured using alamarBlue®, which assesses the ability of a cell population to metabolise resazurin but not its proliferative capacity. To measure cell survival and quantify the ability of individual cells to expand and form a viable population, colony formation assays were employed using semi-solid media. The data obtained from these survival experiments confirmed the data of the viability assays. As for the sensitivity to MMS, it is now more clear that deletion of one *TopBP1* allele renders DT40 cells sensitive to killing by MMS compared to $TopBP1^{+/+/+}$ but deletion of two out of the three alleles exacerbates this phenotype (**Figure 6.3 A**). Thus *TopBP1* gene dosage is correlated with the ability of DT40 cells to survive genotoxic stress.

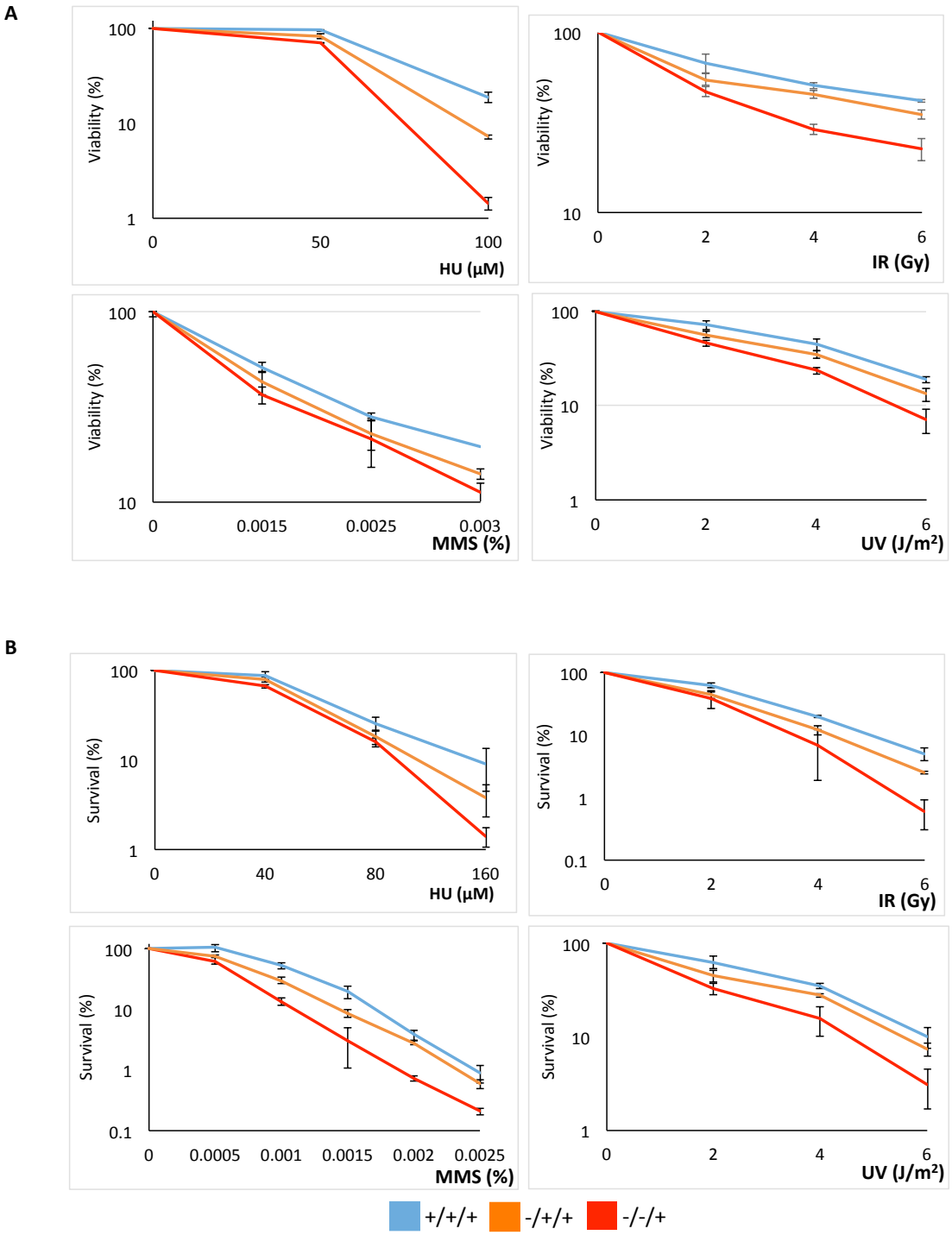


Figure 6.3: Successive deletion of TopBP1 alleles leads to successively increased sensitivity to damage caused by HU, IR, MMS and UV. A) Cell viability was measured after continuous 48h-treatment with increasing doses of HU, after 1h-treatment with MMS followed by 48h recovery time and after increasing doses of IR or UV followed by 48h recovery time. **B)** This method was compared to colony survival assays whereby cells were plated singularly in HU-containing semi-solid media or in drug-free semi-solid media following the indicated dosed of IR or UV or following 1h-treatment with the increasing doses of MMS. After 7-10 days from plating, white colonies were counted at the bottom of the plates. All error bars show the standard deviation of the mean for three independent experiments.

6.1.4: Successive deletion of the *TopBP1* alleles leads to defects in checkpoint signalling and activation

To try and understand the increasing sensitivity of DT40 cells as they become progressively deficient of their *TopBP1* alleles, the *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ isogenic cell lines were examined for their respective ability to activate the DNA structure-dependent checkpoints. HU was used to investigate the response of DT40 cells to replication stress and subsequent activation of the replication and S-M checkpoints. Additionally, I sought to gain insight into the kinetics of activation of the G2-M checkpoint following treatment with IR of this three-cell system. Previous evidence has shown that the G2/M and S/M checkpoints triggered by DNA damage or replication arrest in DT40 cells are both dependent on Chk1 (Zachos et al 2003 and Zachos et al 2005). Chk1 phosphorylation of S345 (pS345 Chk1) has also been described as a *bona fide* read-out of checkpoint activation in response to both DNA damage and replication arrest in DT40 cells (Wang et al 2009). To determine whether pS345 Chk1 could indeed be used as a marker of checkpoint activation in our system and with the available antibodies, wild-type DT40 cells were treated with 2mM HU for two hours and whole cell lysates were prepared at 1h and 2h time points for Western blot analysis. In parallel, wild-type DT40 cells were treated with 5Gy of IR and lysed 1h post-irradiation for Western blot analysis. As shown in **Figure 6.4 A and B**, Chk1 kinase was indeed phosphorylated after both HU-induced replication stress and IR-induced DNA damage. Consistent with Wang et al (2006), IR induced a more modest increase in pS345 Chk1.

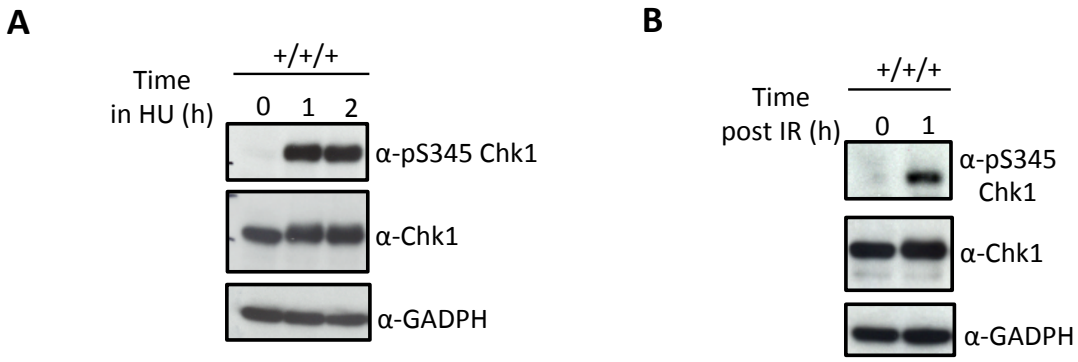


Figure 6.4: Phosphorylation of S345 Chk1 as a marker of checkpoint activation following replication stress or DNA damage. A) Whole cell lysates of wild-type DT40 cells were prepared after 1h and 2h of continuous HU treatment (2mM) and run on SDS-PAGE gel. **B)** Whole cell lysates of wild-type DT40 cells were prepared 1h after irradiation with 5Gy of IR and run on SDS-PAGE gel. In both A and B western blots were analysed using anti-S345 phospho-specific Chk1 antibody and an antibody against total Chk1, while GADPH served as the loading control. Experiments in A and B have been performed in parallel and only presented separately.

6.1.4.1: Response of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells to replication stress caused by hydroxyurea

To assess the kinetics of checkpoint activation in response to replication stress, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were grown in the continuous presence of 2mM HU for 6h and samples for Western blot and FACS analyses were kept at 0, 1, 2, 4 and 6h post addition of the drug. As shown in **Figure 6.5 A, B**, treatment of wild-type DT40 cells with HU led to activation of Chk1, evident from both increased pS345 and altered electrophoretic mobility. Additionally, an increase in the total TopBP1 levels and phosphorylation of S139 on γH2AX were observed. However, cells deleted for one or two of the *TopBP1* alleles showed an impaired checkpoint response compared to *TopBP1*$^{+/+/+}$ cells. In fact, although *TopBP1*$^{+/+/+}$ cells presented an increase in the levels of the TopBP1 protein as early as 1h (1.6-fold) post addition of HU and reaching a maximum at 2h (2.5-fold), *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells only slightly increased (~1.5-fold) TopBP1 above the basal levels at 1h. Interestingly, mutant cells were also impaired in their ability to phosphorylate S345 Chk1, with *TopBP1*$^{+/+/+}$ cells presenting a 23-fold increase over the untreated sample and *TopBP1*$^{-/+/+}$ & *TopBP1*$^{-/-/+}$ cells only below 5- and 4-fold, respectively. Phosphorylation of γH2AX S139 reached a maximum level at 2h for all three strains but it was more robust in *TopBP1*$^{+/+/+}$ (27-fold) than in cells deleted of their *TopBP1* alleles (12- and 18-fold, respectively).

Furthermore, cell cycle analysis by flow cytometry revealed that despite the defects of *TopBP1*$^{-/+/+}$ & *TopBP1*$^{-/-/+}$ cells in the strength and kinetics of checkpoint activation in response to HU, they were able to arrest the cell cycle by 6h similarly to their wild-type counterparts, albeit with slower kinetics. As shown in **Figure 6.5 C**, untreated control cultures contained cells distributed across all phases of the cell cycle. However, after treatment with 1mM HU for 2h, cells started accumulating in G1/early S and by 6h most cells were in the G1/S as judged by PI staining with a corresponding reduction in the proportion of cells in S and G2/M. What is more, *TopBP1*$^{-/-/+}$ and, to a lesser extent, *TopBP1*$^{-/+/+}$ cells contained more sub-G1 cells compared to *TopBP1*$^{+/+/+}$.

Taken together, data show that HU activates Chk1 in wild-type DT40 cells and leads to increased TopBP1 levels and increased pS345 Chk1 and pS139 γH2AX. Gene dosage reduction of *TopBP1* impairs the response to replication stress and although most cells successfully arrest at G1/S, they do so in an altered way and with slower kinetics than *TopBP1*$^{+/+/+}$ DT40 and also at the expense of some cell death. The

impairment of the checkpoint response is thus proportional to *TopBP1* gene dosage as well as to the extent of replication stress, with higher concentrations of HU exacerbating the observed phenotypes.

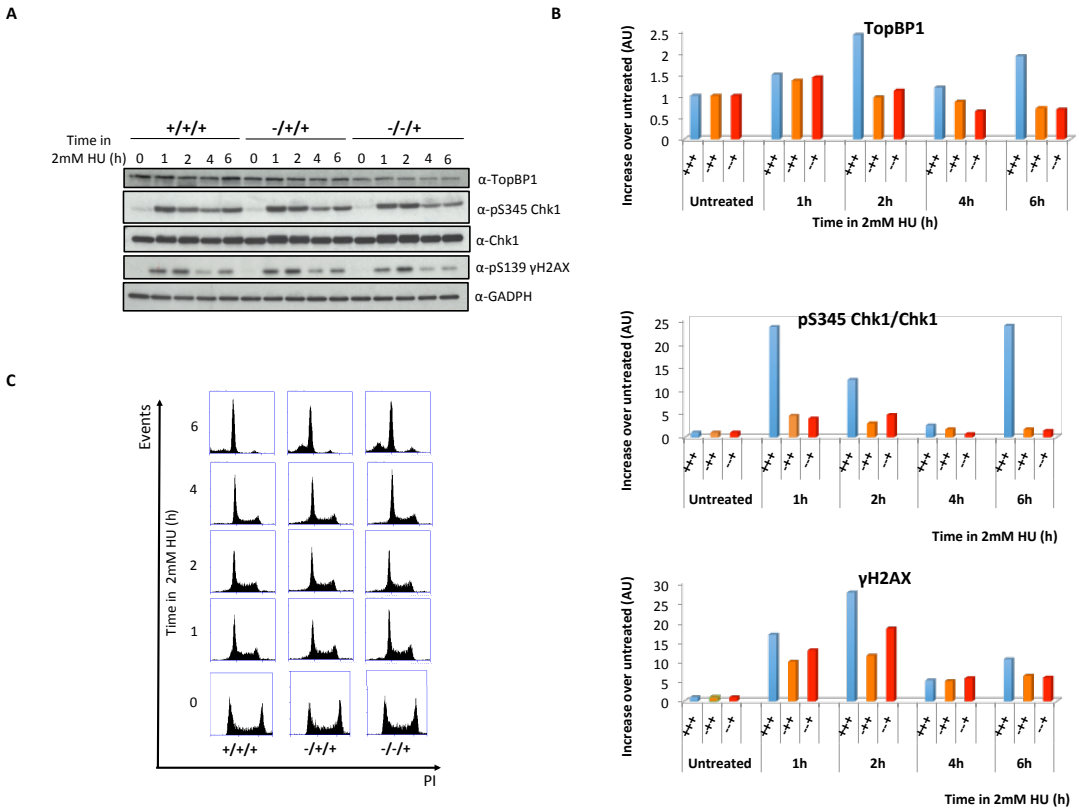


Figure 6.5: Gene dosage-dependent defect of checkpoint activation in response to continuous replication stress. The indicated strains were treated with 2mM HU and samples were kept at the specified times for analysis by Western blotting and FACS. **A)** Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phospho-specific S345 Chk1 antibody, total Chk1 and phospho-specific S139 γH2AX, while GADPH served as the loading control. **B)** Quantification of the blot in A, using the ImageJ software. **C)** Flow cytometry analysis of samples fixed and stained with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure.
Note that equivalent results have been obtained from repeat experiments.

6.1.4.2: Response of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells during recovery from replication stress caused by hydroxyurea

To investigate whether *TopBP1* gene dosage is also important for the recovery from replication stress, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were arrested in G1/S by treatment with 1mM HU for 12h and then were washed clear of the drug and followed through a period of 6h recovery time. Western blot analysis (**Figure 6.6 A, B**) showed that an increase in the levels of TopBP1 and pS345 Chk1 following HU block and release were dependent on *TopBP1* gene dosage. In fact, by the end of the 12h arrest in HU *TopBP1*$^{+/+/+}$ cells increased TopBP1 protein levels more than 2-fold compared to the basal (untreated) levels whereas *TopBP1*$^{-/+/+}$ showed a 1.5-fold increase and *TopBP1*$^{-/-/+}$ cells no increase at all. During the recovery period, TopBP1 levels started to steadily decline for *TopBP1*$^{+/+/+}$ cells before reaching the basal level at 6h post-release. On the contrary, the mutant cell lines displayed an increasing amount of TopBP1. More specifically, *TopBP1*$^{-/+/+}$ cells further increased their TopBP1 levels to 1.7- and 1.8-fold at 1h and 2h post-release respectively, before returning to near normal levels (0.9-fold) by 4h. Similarly, *TopBP1*$^{-/-/+}$ cells reached a 1.4-fold increase at 2h post-release and by 4h they presented TopBP1 protein below basal levels (0.7-fold), which further decreased (0.3-fold) by 6h.

Furthermore, phosphorylation of the replication checkpoint effector kinase Chk1 was also impaired in a gene dosage-dependent manner. The HU block resulted in a 8-fold increase of the pS345 Chk1 pool in *TopBP1*$^{+/+/+}$ cells and this phosphorylation persisted 4h post-release before starting to decline. Consistent with the phenotypes observed in the experiment in Figure 6.6, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were impaired in their ability to phosphorylate Chk1 on S345 to the extent of *TopBP1*$^{+/+/+}$ cells, following 12h incubation with HU. More specifically, only a 3.2- and 2.7-fold increase in the pS345 Chk1 population was observed in *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ mutants respectively by the end of the arrest. Contrary to *TopBP1*$^{+/+/+}$ DT40, pS345 Chk1 was not maintained elevated in the mutants. Both *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells reduced their pool of phosphorylated Chk1 in a stepwise manner before returning to the levels of undamaged cells by 6h post-release. This is markedly different from what was observed in *TopBP1*$^{+/+/+}$ cells, which displayed increased pS345 Chk1 even 6h post-release (4.1-fold over untreated), indicative of persisting checkpoint activation.

Additionally, *TopBP1*$^{-/-/+}$ cells were characterized by an increased phosphorylation of S139 γH2AX by the end of the HU block (8.1-fold over untreated) compared to *TopBP1*$^{+/+/+}$ (5.1-fold) and *TopBP1*$^{-/+/+}$ (6.5-fold). At later time points, *TopBP1*$^{+/+/+}$ cells decreased pS139 before phosphorylation starting returning to basal levels (2.8, 4.2, 4 and 2.5-fold over untreated at 1, 2, 4 and 6h post-release, respectively). *TopBP1*$^{-/+/+}$ cells followed a similar trend like wild-type cells (3.5, 6.8, 3.2 and 3.3-fold, respectively) but in marked contrast, *TopBP1*$^{-/-/+}$ cells presented higher levels of pS139 across all time points compared to the other cell lines (5.3, 7.7, 4.5 and 4.3-fold, respectively) (**Figure 6.6 A, B**).

Cell cycle analysis by flow cytometry revealed that the three cell lines block at the G1-S in response to HU in a gene dosage-dependent manner, with *TopBP1*$^{-/-/+}$ blocking less efficiently. The resumption of the cell cycle following release from the HU was found to be *TopBP1* gene dosage-dependent. In particular, a subset of *TopBP1*$^{-/-/+}$ -and to a lesser extent *TopBP1*$^{-/+/+}$- cells displayed 2N DNA content at 2 and 4h post HU release, when in fact *TopBP1*$^{+/+/+}$ cells have all synchronously moved to the next G2 phase. And by 6h, *TopBP1*$^{-/+/+}$ and to a greater extent *TopBP1*$^{-/-/+}$ cultures had more cells in the G1 phase compared to the *TopBP1*$^{+/+/+}$ wild-type control (**Figure 6.6 C**).

Taken together, the observed data suggest that successive deletion of the *TopBP1* alleles results in not only an impaired checkpoint response to replication stress but also in an impaired recovery from replication stress. Compared to *TopBP1*$^{+/+/+}$ DT40 cells, this recovery period in *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells is characterized by a defect in increasing the expression of the TopBP1 itself, a severe impairment in phosphorylating the effector kinase Chk1 on S345 and a premature restoration of this phosphorylation to basal levels, which is accompanied by a faster resumption of the cell cycle. At the same time, increased phosphorylation of S139 γH2AX was observed in a gene dosage-dependent manner.

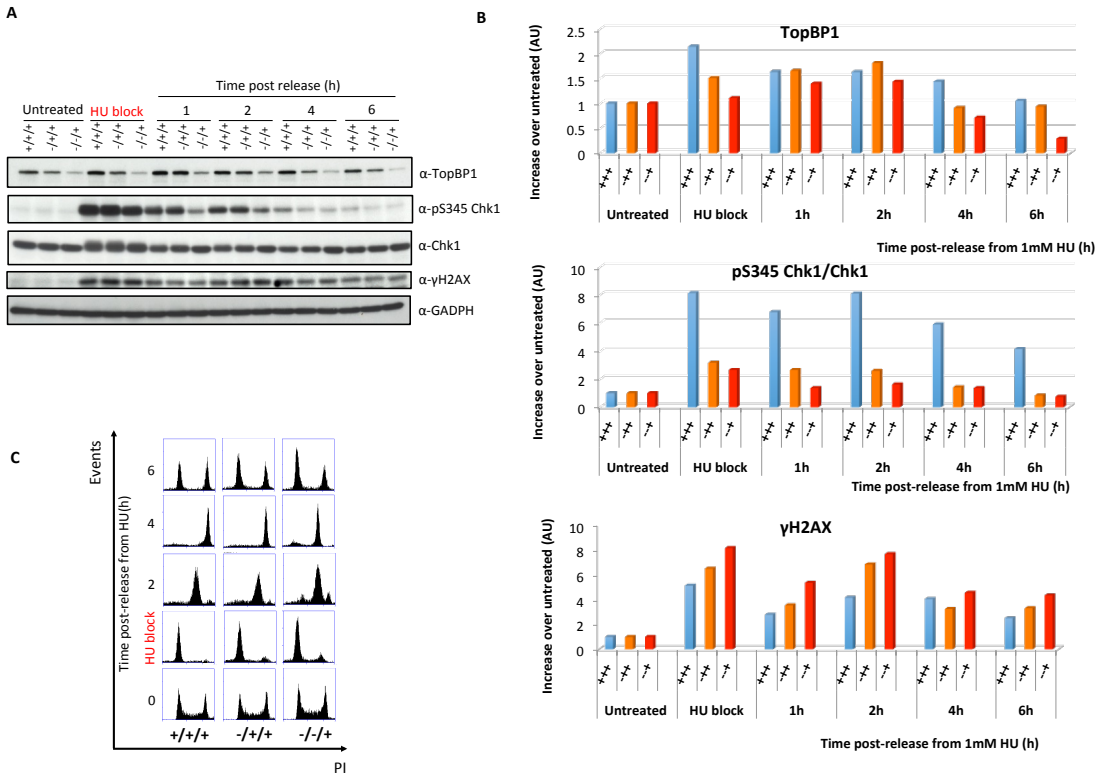


Figure 6.6: Gene dosage-dependent recovery from replication stress. The indicated strains were treated with 1mM HU for 12h before being transferred into normal media. Samples were kept at the specified times for analysis by Western blotting and FACS. **A)** Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phospho-specific S345 Chk1 antibody, total Chk1 and phospho-specific S139 γH2AX, while GADPH served as the loading control. **B)** Quantification of the blot in A, using the ImageJ software. **C)** Flow cytometry analysis of samples fixed and stained with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure.
Note that equivalent results have been obtained from repeat experiments.

6.1.4.3: Response of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells to DNA damage caused by ionizing radiation.

To determine whether *TopBP1* gene-dosage is important for DNA damage-induced G2/M arrest, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were exposed to 10Gy of IR and Western blot analysis was used to assess the activation of the G2/M checkpoint. Irradiation of the *TopBP1*$^{+/+/+}$cells led to progressively increasing levels of the TopBP1 protein reaching a 2.7-fold compared to untreated at 4h post-IR. At later time points the TopBP1 levels of *TopBP1*$^{+/+/+}$ cells returned nearly to the untreated levels (1.3- and 1.2-fold at 6h and 8h, respectively). In striking contrast *TopBP1*$^{-/+/+}$ cells completely failed to increase the levels of TopBP1 following irradiation and at time points beyond 2h post-IR TopBP1 dropped to levels below what is found in the untreated. Now the *TopBP1*$^{-/-/+}$ cells showed an inability to increase the levels of TopBP1 to the extent of *TopBP1*$^{+/+/+}$ but they presented a slight 1.4-fold increase at 4h post-IR. At later time points however they exhibited very low levels of the protein (0.4- and 0.3-fold at 6h and 8h, respectively). Similarly to TopBP1, the extent and kinetics of Chk1 phosphorylation at S345 displayed a different pattern in the mutant cells compared to the wild-type. In particular, *TopBP1*$^{+/+/+}$ displayed a 1.9-fold increase of pS345 Chk1 over untreated at 1h post-IR, which remained relatively stable until 8h post-IR (2.5-fold). The mutant cell lines, however, showed a *TopBP1* gene dosage-dependent increase of pS345 Chk1. Throughout the 8h timecourse, *TopBP1*$^{-/+/+}$ cells displayed more than 3-fold increase in the levels of pS345 indicating that a higher percentage of their Chk1 pool was phosphorylated on S345 compared to *TopBP1*$^{+/+/+}$. Even more pronounced was the 6.4-fold increase over untreated of pS345 Chk1 in the *TopBP1*$^{-/-/+}$ cells 1h post-IR, which gradually increased reaching a 10.2-fold increase at 8h. Examination of pS139 γH2AX revealed that *TopBP1*$^{-/-/+}$ cells showed a consistently decreased ability of phosphorylating this residue compared to *TopBP1*$^{+/+/+}$ across all the time points examined (**Figure 6.7 A, B**).

At the same time the effect of IR on cell cycle progression was determined by flow cytometry. Wild-type DT40 cells lack functional p53 (Takao et al, 1999) and 6-8h after irradiation accumulated predominantly in G2/M. In comparison, *TopBP1*$^{-/+/+}$ and to a greater extent *TopBP1*$^{-/-/+}$ cells exhibited a slower accumulation in G2/M, indicative of a *TopBP1* gene dosage-dependent checkpoint defect. In fact, 8h after irradiation *TopBP1*$^{-/-/+}$ cultures still contained a significant amount of S phase cells (**Figure 6.7 C**).

To further confirm that the G2/M defect of the mutant cell types is attributable to gene dosage reduction, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were incubated in medium containing nocodazole for 12h with or without prior irradiation (10Gy). The percentage of mitotic cells was then determined by DAPI staining of the nucleus. As shown in **Figure 6.7 D**, successive deletion of the *TopBP1* alleles leads to a successive defect in G2/M arrest following the nocodazole block. In fact, 88% of the *TopBP1*$^{+/+/+}$ cells successfully arrested in mitosis by the end of the nocodazole block, but only 82.5% of *TopBP1*$^{-/+/+}$ and 73.5% of *TopBP1*$^{-/-/+}$ managed to do so. Additionally, irradiation completely prevented nocodazole-treated *TopBP1*$^{+/+/+}$ cells from accumulating in mitosis (11.5%), indicating a successful arrest in the G2 phase. In contrast, 17% of *TopBP1*$^{-/+/+}$ and 28% of *TopBP1*$^{-/-/+}$ cells entered mitosis regardless of prior irradiation.

Taken together the above data suggest that TopBP1 is required in a gene dosage-dependent manner for proper functioning of the G2/M checkpoint in DT40 cells. Successive deletion of the *TopBP1* alleles results to a progressively defective DNA damage response characterized by an inability to increase TopBP1 protein levels, an abnormally elevated phosphorylation of Chk1 on S345 and γH2AX on S139 and a defect in arresting the cell cycle at the G2 phase following IR or combined IR and nocodazole trap.

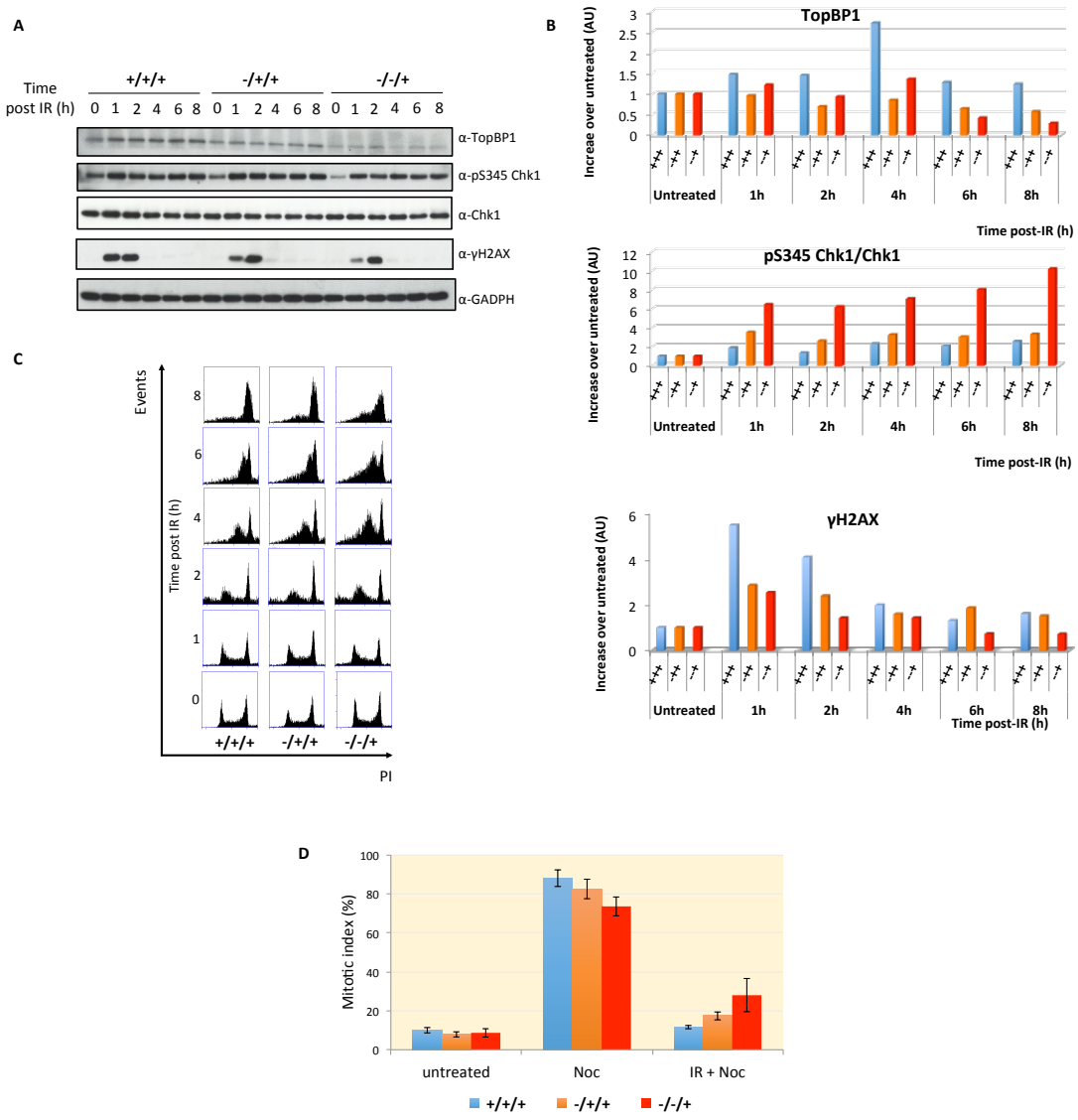


Figure 6.7: Gene dosage-dependent defect of checkpoint activation in response to IR. The indicated strains were irradiated with 10Gy of IR before being transferred into normal media. Samples were kept at the specified times for analysis by Western blotting and FACS. **A)** Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phospho-specific S345 Chk1 antibody, total Chk1 and phospho-specific S139 γH2AX, while GADPH served as the loading control. **B)** Quantification of the blot in A, using the ImageJ software. **C)** Flow cytometry analysis of samples fixed and stained with propidium iodide. **D)** Mitotic indices of the indicated cells incubated with nocodazole for 12h with or without prior irradiation (10Gy). A minimum of 200 nuclei was counted for each mitotic index measurement. Error bars indicate the standard deviation of the mean for two independent experiments. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

6.1.4.4: Milder defects of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells to lower doses of replication stress or DNA damage

To investigate whether the *TopBP1* gene dosage-dependent defects of the checkpoint response were correlated with the extent of genotoxic stress, cells were treated with lower doses of HU and IR and checkpoint activation assessed by Western blot analysis and FACS.

Blocking *TopBP1*$^{+/+/+}$ cells in early S phase with 0.25mM HU for 12h resulted in a 1.8-fold increase of the TopBP1 levels compared to the untreated sample. The protein levels further increased to 2.6-fold at 1h post-release from HU before returning to the basal levels at subsequent time points. *TopBP1*$^{-/+/+}$ cells showed a similar pattern reaching 1.7- and 2-fold increase of TopBP1 at 12h block and 1h post-release respectively before returning to basal levels of TopBP1 protein. In contrast, *TopBP1*$^{-/-/+}$ cells reached the peak of TopBP1 protein levels at the end of the HU block (1.9-fold over untreated) but failed to sustain it and returned back to basal earlier than the other two cell types cells. Compared to the HU block-release experiment described in Figure 6.7 where a concentration of HU used was 4-fold higher (1mM instead of 0.25mM used here), it is observed that *TopBP1*$^{-/-/+}$ cells have a milder defect in boosting TopBP1 protein levels when exposed to a milder HU block. In fact, when treated with 1mM HU *TopBP1*$^{-/-/+}$ cells only increased TopBP1 protein levels to 1.1-fold above untreated, whereas with 0.25mM they can achieve a 1.9-fold increase above untreated.

Furthermore, phosphorylation of S345 Chk1 of *TopBP1*$^{+/+/+}$ cells was markedly increased to 8.1-fold by the end of the HU block, whereas *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells displayed a *TopBP1* gene dosage-dependent reduced ability (4.3- and 3.2-fold, respectively) of catalysing this modification compared to wild-type. Following release into normal media, pS345 Chk1 started to gradually reduce for all cell types. By 8h post-release however, *TopBP1*$^{+/+/+}$ cells still possessed pS345 at 1.7-fold over untreated but for the mutants a gene dosage-dependent premature return to basal levels was observed. When comparing the 1mM versus 0.25mM HU-block release experiments it is noticed that there is a comparable gene dosage-dependent defect in sustaining checkpoint activation (as judged by pS345 Chk1) compared to *TopBP1*$^{+/+/+}$. Additionally, all three cell types displayed phosphorylation of γH2AX after treatment with HU for 12h. Interestingly, *TopBP1*$^{-/-/+}$ cells displayed higher amounts of S139 γH2AX phosphorylation than *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/+/+}$ both by the end of the HU block and at 1h post-release release period (**Figure 6.8 A, B**). This was consistent with

the observed phosphorylation pattern of S139 γH2AX across the three cell types following a 12h HU block (1mM) and 1h release from this harsh block, described in Figure 6.6. At later time points, however, the milder damage seems to have alleviated the observed phenotypes of the mutants (**Figure 6.8 A, B**).

Flow cytometry analysis showed that similarly to the stronger HU block, the majority of cells successfully arrested their cell cycles at early S phase and at 2h and 4h post-release they were synchronously going through S and G2 phases, respectively. Interestingly, at 6h post-release where cells have progressed through a next round of the cell cycle, *TopBP1*$^{-/-/+}$ cultures contained more G1 cells compared to *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/+/+}$. So although the milder HU treatment seems to have diluted out the phenotype of the *TopBP1*$^{-/+/+}$ cells, *TopBP1*$^{-/-/+}$ cells were still presented with a phenotype defective compared to the wild-type cells (**Figure 6.8 C**).

To determine whether *TopBP1* gene-dosage is important for DNA damage-induced G2/M arrest after treatment with low IR doses, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells were exposed to 5Gy of IR and Western blot analysis was used to assess the activation of the G2/M checkpoint. 1h post-IR, wild-type DT40 showed 3-fold increase of their total TopBP1 protein levels, which returned to near normal levels at later time points. In marked contrast, the mutant cell lines were unable to increase TopBP1 levels, in a similar manner as observed after irradiation with 10Gy of IR. Additionally, the IR dose used was too mild to induce pS345 Chk1 above the basal levels in *TopBP1*$^{+/+/+}$ cells, whereas 1h post-IR *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells displayed 1.6- and 1.8-fold increase of pS345 phosphorylation, respectively, compared to the untreated samples. This increased phosphorylation persisted until later time points. This is comparable to what was observed when the three cell types were irradiated with 10Gy of IR where the mutant cells exhibited higher levels of pS345 Chk1 than wild-type and *TopBP1*$^{-/-/+}$ cells displayed the highest among all. It is worth noticing however, that although the pattern of pS345 Chk1 induction is comparable between the 10Gy versus 4Gy irradiation experiments, the lower dose of IR causes a more modest phosphorylation of S345. Finally, 5Gy of IR were sufficient to increase pS139 γH2AX 12.2-fold above untreated in *TopBP1*$^{+/+/+}$ cells 1h post-IR. In contrast, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells showed a 3.2- and 1.8-fold increase over the basal levels. What is more, the levels of pS139 γH2AX followed a decreasing pattern between 1h and 6h post-IR for both *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/+/+}$ cells and at 8h they showed a slight increase again. *TopBP1*$^{-/-/+}$ cells however, increased pS139 γH2AX to 3.8-fold at

2h after irradiation, higher than the other cell types. At later time points their levels of pS139 decreased again to levels below wild-type (**Figure 6.8 D, E**). Therefore, the milder irradiation resulted in a limited amount pS139 γH2AX in cells harboring less TopBP1, which is comparable to what observed in Figure 6.7 where a harsher IR treatment was used.

In parallel flow cytometry was used to assess the effect of this lower dose of IR on the progression of the cell cycle of the respective cell types. In marked contrast to the higher IR experiment described earlier, here the mutant cell lines displayed similar cell cycle progression to wild-type. So by 6h *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ cells had all successfully blocked at the G2 phase of the cell cycle (**Figure 6.8 F**).

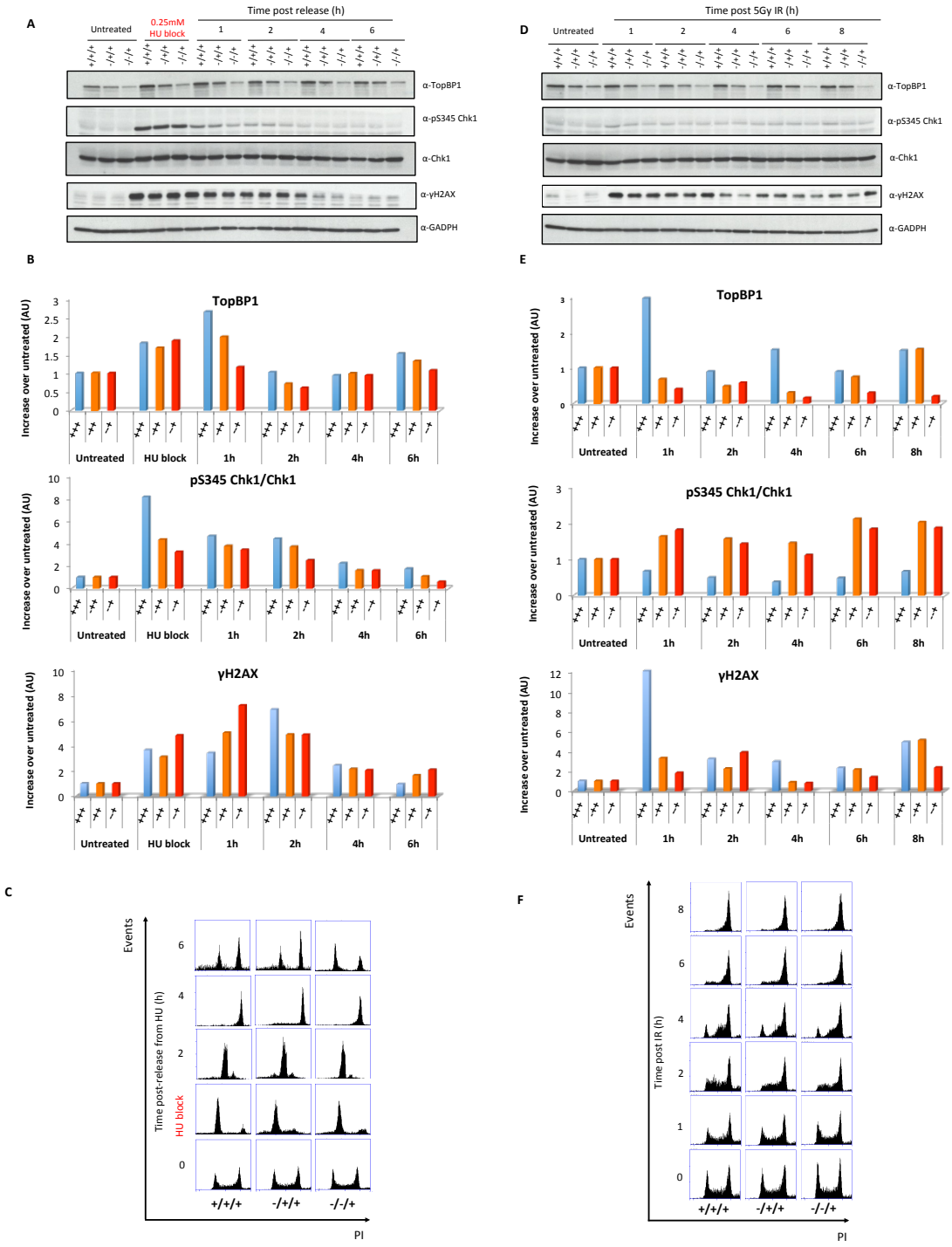


Figure 6.8: Gene dosage-dependent checkpoint defects at lower doses of replication stress or DNA damage. (left panel) The indicated strains were treated with 0.25mM HU for 12h before being transferred into normal media. (right panel) The indicated strains were irradiated with 5Gy of IR before being transferred into normal media. Samples were kept at the specified times for analysis by Western blotting and FACS. **A) D)** Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phospho-specific S345 Chk1 antibody, total Chk1 and phospho-specific S139 γH2AX, while GADPH served as the loading control. **B) E)** Quantification of the blot in A and D, using the ImageJ software. **C) F)** Flow cytometry analysis of samples fixed and stained with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

6.2: Creation of a novel system to study the effects of TopBP1 destruction in the human h-TERT RPE-1 cell line

To gain insight into the dosage function of TopBP1 in mammalian cells, I used the small molecule-assisted shutoff (SMASh) technology (Chung, Jacobs et al. 2015) combined with the mini-AID degron (Nishimura and Kanemaki 2014) to create a human cell line amenable to manipulation of the TopBP1 protein levels. This involved tagging both *TopBP1* alleles of the human h-TERT RPE-1 cell line with an array of the SMASh and min-AID destabilizing degrons that can be conditionally (and independently) activated to target the tagged protein for degradation. The combined use of two degrons provides a powerful means to effectively degrade the protein of interest in a minimally disruptive, inducible and reversible way. The TopBP1mAIDSMASh model developed herein provides a novel system to study the effects of TopBP1 protein dosage on checkpoint signaling and activation, much like the *TopBP1* gene dosage system developed in DT40. Additionally, the TopBP1mAIDSMASh system provides a time window between TopBP1 protein destruction where the phenotypes arising from TopBP1 absence on checkpoint functioning or indeed other cellular processes can be assessed.

6.2.1: Generation of the *TopBP1*$^{+/mAID\text{-}S}$ *rosa26*$^{+/osTIR1myc}$ and *TopBP1*$^{mAID\text{-}S/mAID\text{-}S}$ *rosa26* $^{+/osTIR1myc}$ RPE cell lines

A CRISPR/Cas-based tagging method was used to tag the endogenous *TopBP1* gene with the miniAID/SMASh array of degrons in the RPE cell line. Although these two degrons both enable the destruction of the tagged protein, they work in completely different ways and have different requirements. The SMASh tag system is composed of the Hepatitis C virus (HCV) protease NS3 followed by the degron tag. Fusion of the SMASh tag to the target protein occurs via an HCV NS3 protease-recognition site. After protein folding, the SMASh tag is removed from the protein by its internal protease activity and is also degraded owing to its internal degron activity. Addition, however, of a protease inhibitor called asunaprevir (ASV) blocks the auto-cleavage step of the SMASh tag leading to degradation of the synthesized copies of the tagged protein (Chung, Jacobs et al. 2015). The miniAID degron on the other hand, requires the introduction of a second component inside the cells at a safe harbor locus, like the *rosa26* locus used herein (Friedrich and Soriano 1991, Irion, Luche et al. 2007). This component is the auxin-responsive F box protein, TIR1, which forms a functional

SCF^{TIR1} (Skp1–Cullin–F-box) ubiquitin ligase with the endogenous components in human cells. The TIR1 used here is derived from *Oryza sativa* (OsTIR1), because previous evidence suggests that it works better than its homolog in *A. thaliana* at higher temperatures (Nishimura, Fukagawa et al. 2009). So in the presence of auxin, a protein tagged with miniAID is polyubiquitinated and targeted for ubiquitin-mediated proteosomal degradation. To render this a conditional rather than a constitutive event, the OsTIR1 transgene was placed under the control of tetracyclin responsive TRE promoter (TRE-OsTIR1) such that transcription is reversibly turned on or off in the presence of the antibiotic tetracyclin or its derivative doxycycline (Dox) (Gossen, Freundlieb et al. 1995). Overall, SMASh tag-mediated degradation of a miniAID-SMASh-tagged TopBP1 would require addition of ASV, whereas mini-AID-mediated degradation would require addition of auxin and Dox in the culture. Treatment of cells with all three drugs -ASV, auxin and Dox- would activate both degrons, presumably leading to a more robust degradation of the TopBP1 protein, compared to ach system alone. A schematic representation of the two degron systems is depicted in **Figure 6.9 A, B**.

As shown in **Figure 6.9 C**, the targeting strategy devised for the generation of the cell lines of interest involved C-terminally tagging the endogenous *TopBP1* gene with the miniAID-SMASh array. A first targeting step would yield $TopBP1^{+/mAID\text{-}S}$ heterozygous cells and a second targeting step would yield $TopBP1^{mAID\text{-}S/mAID\text{-}S}$ homozygous. Alternatively, owing to the previously reported efficiency of the CRISPR-Cas9 system, heterozygous and homozygous cells could be obtained from a single targeting step (and indeed that was the case eventually). The generation of these cells lines would resemble the DT40 gene dosage system described earlier in the sense that the protein product of each *TopBP1* allele could be selectively degraded. To fulfill the second requirement of the miniAID system, a myc-tagged OsTIR1 transgene was integrated at the safe harbour *rosa26* locus under the control of a TRE promoter. This targeting strategy was used to generate $TopBP1^{+/mAID\text{-}S}$ $rosa26^{+/osTIR1myc}$ and $TopBP1^{mAID\text{-}S/mAID\text{-}S}$ $rosa26^{+/osTIR1myc}$ RPE cell lines.

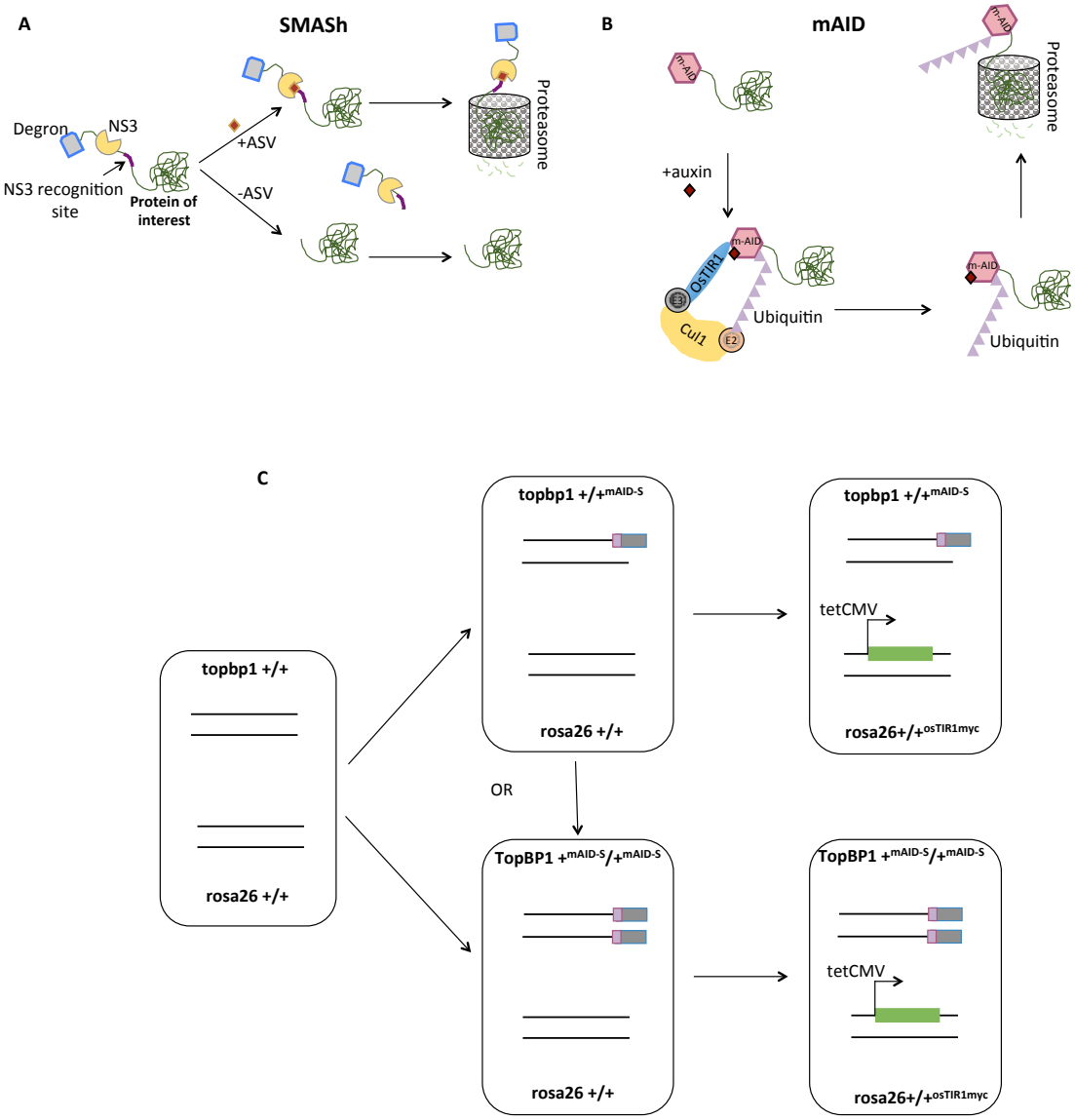


Figure 6.9: Schematics of the SMASh and miniAID systems and targeting strategy. A) Translation of a SMASh-tagged protein results to intramolecular cleavage by the NS3 protease, releasing the protein in its native form whereas the NS3 protease and degron are degraded by the cell (bottom panel). Addition of asunaprevir (ASV) however, inhibits NS3 and consequently the degron promotes proteasomal degradation of the tagged protein (top panel) (Adapted from Chung et al 2015). **B)** Expressed OsTIR1 forms a functional SCF^{OsTIR1} E3 ubiquitin ligase complex with the endogenous components of the pathway in human cells. In the presence of auxin, a miniAID-tagged protein is degraded via polyubiquitylation mediated proteosomal degradation. **C)** To generate an RPE cell line where both of the *TopBP1* alleles would be tagged with the combined miniAID-SMASh system, a miniAID-SMASh fusion would first be integrated using CRISPR at the C-terminal end of *TopBP1* removing the stop codon. To also satisfy the requirement of the miniAID system, an OsTIR1 transgene would be integrated under the control of an inducible promoter at the *rosa26* safe harbour locus using CRISPR.

To C-terminally tag the *TopBP1* alleles with the miniAID-SMASh degron tag array, a homology directed repair (HDR) template was co-transfected with the Cas9-expressing vector pX459, where the single-guide RNA (sgRNA) has been subcloned. More specifically, to assemble the HDR template (**Figure 6.10 A top panel**), a left homology arm (LHA) -corresponding to 500bp upstream of the stop codon of the genomic *TopBP1* locus- followed by a *BstZI7I* restriction site and a right homology arm (RHA) –corresponding to 500bp downstream of the stop codon- were synthesized in pBluescript SK+ as an *XhoI/SpeI* fragment. In parallel, three sgRNAs were selected using the Benchling software and the algorithm therein, capable of locating potential protospacer adjacent motif (PAM) and target sequences and ranking the associated sgRNAs based on their predicted on-target and off-target activity. In this case, sequences amenable to recognition and cleavage by the *Streptococcus pyogenes* (SP) Cas9 (SpCas9) with a PAM of NGG were selected. Each sgRNA was assembled by annealing a forward and a reverse primer and a total of three sgRNAs were made that all were specific for regions ~100bp around the stop codon of the endogenous *TopBP1* gene (primers used P52-57). The ends of the sgRNA oligonucleotides contained the appropriate sequences as described by Ran *et al* 2013 to enable their subcloning at the *BbsI* site of pX459 vector.

Additionally, PCR mutagenesis was performed to render the HDR template resistant to cleavage by the Cas9 enzyme with primers P58-63. This involved silently mutating two juxtaposed codons within the Cas9 target site/PAM site on the HDR template so as to avoid both the cleavage of the HDR prior to integration as well as the continuous cleavage of the genomic site following successful targeting. Overall, three versions of the pX459 vector, each with a different *TopBP1*-specific sgRNA and three versions of the HDR template (TopBP1taggingHDR 1-3), each silently mutated in a manner specific for each sgRNA were designed. The final cloning step for the design of the HDR 1-3 templates involved the subcloning of a miniAID-SMASh-T2A-Neomycin fragment, kindly provided by the Hochegger laboratory. This fragment was PCR amplified with primers P64 & P65 and subcloned at the unique *BstZI7I* restriction site of HDRs 1-3. Sanger sequencing analysis confirmed all the DNA molecules created.

To integrate the miniAID-SMASh double degron tag at the C-terminus of the endogenous *TopBP1* gene, each HDR template was co-transfected with its appropriate sgRNA-expressing Cas9 pX459 vector into wild-type h-TERT RPE-1 cells using the Neon transfection system (see Methods and Materials). To control for the efficiency of

the transfection, an EGFP tagged H2B-expressing vector was transfected in parallel. 24h post-transfection, control cells were examined by fluorescence microscopy for the EGFP signal and this confirmed successful transformation. Media was refreshed and cells left to recover for another 24h before being serially diluted and selected in the presence of 1mg/ml G418 in 96-well plates. Three weeks later single colonies started appearing and these were expanded in duplicates; one of the samples was kept for future freezing down and the other sample was used for genomic DNA extraction and PCR analysis. Successful tagging of the *TopBP1* locus would increase the 125bp PCR product obtained with primers P66 & P67 to 2200bp as shown in **Figure 6.10 A**.

Of the 18 clones screened by PCR, 6 contained both the wild-type and targeted amplification products hence were classified as heterozygous, whereas 12 contained only the targeted product and were classified as homozygous (**Figure 6.10 B**). Two heterozygous and three homozygous clones were confirmed by Sanger sequencing, appropriately frozen down in liquid nitrogen and also subjected to Western blot analysis of whole cell extracts. Wild-type TopBP1 protein has a predicted size of 170.7kDa whereas a tagged TopBP1 is expected to be 181.3kDa owing to the presence of the miniAID tag attached to its C-terminus (SMASh tag is auto cleaved) (**Figure 6.10 C**). Western blot analysis enabled the size-dependent separation of the tagged versus untagged TopBP1 protein populations and confirmed the creation of the *TopBP1*$^{+/mAID\text{-}S}$ (which is heterozygous hence expresses both the 170.7 and 181.3kDa TopBP1 populations) and *TopBP1*$^{mAID\text{-}S/mAID\text{-}S}$ (which is homozygous and only expresses the 181.3kDa tagged TopBP1) RPE cell lines (**Figure 6.10 D**).

To complete the design of the TopBP1 degradation system, a myc-tagged OsTIR1 transgene under the control of the TRE promoter was integrated at the *rosa26* locus of both *TopBP1*$^{+/mAID\text{-}S}$ and *TopBP1*$^{mAID\text{-}S/mAID\text{-}S}$ cells, using CRISPR technology **(Figure 6.10 E)**. Similarly as before, a pX459 vector expressing a *rosa26* locus-specific sgRNA was co-transfected with a targeting HDR vector harbouring arms homologous to sequences of the *rosa26* locus (LHA and RHA). Successful integration of the HDR vector would result in the replacement of the *rosa26* locus with a transgene containing OsTIR1-myc under the control of TRE promoter. In addition, to enable selection of the successfully transfected cells, the ectopic transgene also contained a *Sh ble* gene under the control of the SV40 promoter. Both the sgRNA-expressing Cas9 vector and the rosa26 HDR were kindly provided by the Hochegger laboratory. Transfectants were thus selected in the presence of 500μg/ml zeocin and once expanded, they were

subjected to Western blot analysis to check for the expression of OsTIR1-myc. In fact, parental cell lines alongside 21 bleomycin-resistant clones were treated with 1μg/ml Dox for 24h and whole cell extracts run on SDS PAGE gel.

Parental cell lines were negative for the expression of OsTIR1-myc whereas 19 out of the 21 clones showed expression of OsTIR1-myc showing successful integration of the transgene **(Figure 6.10 F**). However, to distinguish between a correct targeting of the *rosa26* locus and an erroneous targeting of the HDR somewhere else inside the genome, 4 of the *TopBP1*$^{+/mAID\text{-}S}$ and 7 of the *TopBP1*$^{mAID\text{-}S/\ mAID\text{-}S}$ that showed OsTIR1-myc expression in Figure 6.11 F were subjected to PCR analysis. The forward primer of the reaction annealed within the SV40 promoter in the HDR whereas the reverse primer was specific to a region of the *rosa26* locus located 3' to the RHA region (**Figure 6.10 E** bottom diagram (purple arrows)). In all of the clones tested, the transgene had indeed integrated at the *rosa26* locus as is evident from the 2.3kb band amplified from the genomic DNA of the clones but not that of the parental cell lines **(Figure 6.10 G)**. This series of targeting events has successfully generated the *TopBP1*$^{+/mAID\text{-}S}$ *rosa26*$^{+/osTIR1myc}$ and *TopBP1*$^{mAID\text{-}S/mAID\text{-}S}$ *rosa26* $^{+/osTIR1myc}$ RPE cell lines.

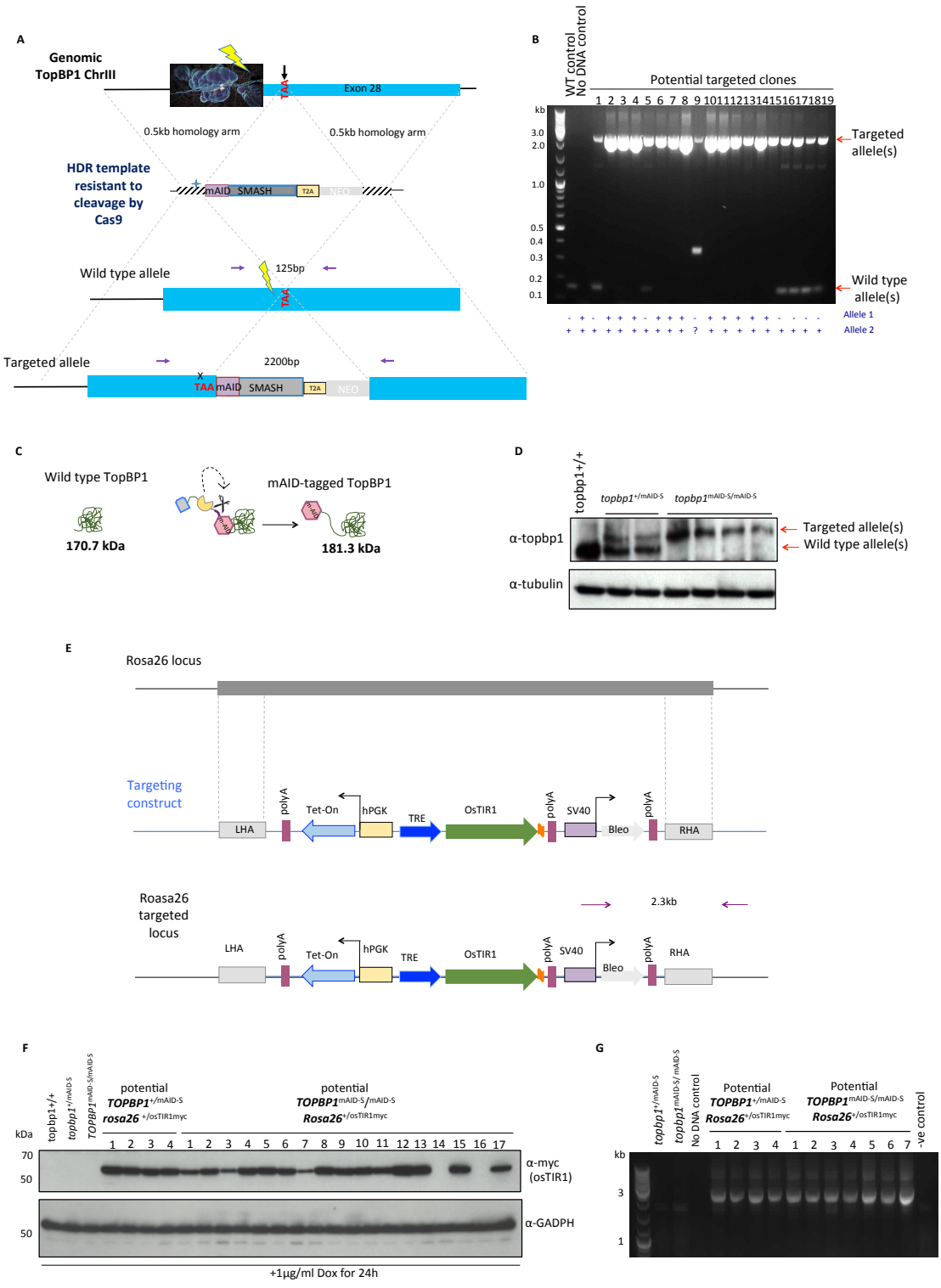

A
Genomic
TopBP1 ChrIII
Exon 28
0.5kb homology arm
0.5kb homology arm
HDR template
resistant to
cleavage by
Cas9
mAID
SMASH
T2A
NEO
Wild type allele
125bp
Targeted allele
2200bp
B
WT control
No DNA control
Potential targeted clones
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
kb
3.0
2.0
1.0
0.5
0.4
0.3
0.2
0.1
Targeted
allele(s)
Wild type
allele(s)
Allele 1
Allele 2
C
Wild type TopBP1
170.7 kDa
mAID-tagged TopBP1
181.3 kDa
D
topbp1+/+
topbp1+/mAID-S
topbp1mAID-S/mAID-S
α-topbp1
α-tubulin
Targeted allele(s)
Wild type allele(s)
E
Rosa26 locus
Targeting
construct
LHA
polyA
Tet-On
hPGK
TRE
OsTIR1
polyA
SV40
Bleo
polyA
RHA
Roasa26
targeted
locus
2.3kb
F
topbp1+/+
topbp1+/mAID-S
TOPBP1mAID-S/mAID-S
potential
TOPBP1+/mAID-S
rosa26+/osTIR1myc
potential
TOPBP1mAID-S/mAID-S
Rosa26+/osTIR1myc
kDa
1 2 3 4 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
70
50
α-myc
(osTIR1)
50
α-GADPH
+1μg/ml Dox for 24h
G
topbp1+/mAID-S
topbp1mAID-S/mAID-S
No DNA control
Potential
TOPBP1+/mAID-S
Rosa26+/osTIR1myc
Potential
TOPBP1mAID-S/mAID-S
Rosa26+/osTIR1myc
-ve control
1 2 3 4 1 2 3 4 5 6 7
kb
3
1

Figure 6.10: Construction of a conditional miniAID-SMASh TopBP1 system in human RPE cells. **A)** Schematic of the degron targeting construct, comprising the miniAID and SMASh sequences coupled to the neomycin marker via the T2A sequence and all flanked by 500bp arms homologous to the regions of the *TopBP1* gene immediately upstream and downstream the stop codon. Unsuccessful targeting would leave the TopBP1 allele intact whereas gene targeting would result in replacement of the stop codon with the targeting construct. PCR with the primers shown (purple arrows) could be used to distinguish successfully targeted clones as well as heterozygously- from homozygously-tagged based on the size of the PCR product (125bp VS 2200bp). T2A is a small "self-cleaving" peptide that was initially identified in the foot-and-mouth disease virus (FMDV) (Ryan, King et al. 1991). Subsequent studies have demonstrated that ribosomes skip the synthesis of the glycyl-prolyl peptide bond at the C-terminus of a 2A peptide, leading to the cleavage between a 2A peptide and its immediate downstream peptide (Donnelly, Luke et al. 2001). In molecular biology, T2A has proved a useful tool for the fusion of genes whose expressions are driven by a single promoter (Kim, Lee et al. 2011). **B)** G418-resistant were genotyped by PCR (as explained in A) for determining clones that had successfully performed gene targeting. Wild-type cells generating a 125bp PCR product -indicative of an untagged *TopBP1* C-terminus- served as the negative control. **C)** Expected sizes of untagged versus tagged TopBP1 protein as calculated with the ExPasy tool. **D)** Whole cell lysates of clones classified as heterozygous or homozygous in B were prepared and run on SDS-PAGE gel. Western blots were analysed using antibody against total TopBP1 while beta-tubulin served as the loading control. The tagged and untagged TopBP1 protein populations can be separated on the basis of size on an SDS gel; top band corresponding to the tagged and bottom band to the untagged on the anti-TopBP1 immunoblot. **E)** Schematic representation of the rosa26 locus, OsTIR1 containing targeting vector and successfully targeted rosa26 locus. The purple arrows indicate the primers used for PCR analysis in G. **F)** Whole cell lysates of clones obtained from transfection of the OsTIR1 construct in *TopBP1*$^{+/mAID\text{-}S}$ and *TopBP1*$^{mAID\text{-}S/mAID\text{-}S}$ cells (obtained from B and D) were prepared and run on SDS-PAGE gel. Western blots were analysed using antibody against the myc-tagged OsTIR1 while GADPH served as the loading control. To activate transcription of the transgene cells were pre-treated with 1μg/ml doxycycline 24h prior to harvesting. **G)** PCR analysis of the indicated clones classified as positive for the OsTIR1 integration in F to show that the transgene has indeed integrated at the rosa26 locus. Primers are depicted in E (purple arrows).

6.2.2: Activation of the miniAID-SMASh system leads to TopBP1 degradation and inability to incorporate EdU and DNA damage.

I next tested the efficacy of the miniAID-SMASh combined system as well of each system on its own in suppressing the protein levels of TopBP1 present inside the cells. *TopBP1*$^{+/+}$ (wild-type), *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ and *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells were exposed to ASV (3μM) alone, IAA (500μM) alone or a combination of both drugs in order to activate the SMASh or miniAID or both, respectively. It should also be noted that activation of the miniAID required pre-treatment of cells with Dox (1μg/ml) for a period of 2h prior to the addition of IAA (N.B. Dox was not washed off). This was necessary in order to activate expression of the ectopic OsTIR1 transgene required for the degrading mechanism of the miniAID system. Following exposure to the respective drugs, cells were followed for a period of 6h and whole cell extracts were analysed by Western blotting at 0, 1, 2, 4 and 6h time points (**Figure 6.11 A**). Exposure of wild-type cells to any of the drugs' combinations did not have an effect on the levels of the TopBP1 protein and also no OsTIR1-myc expression was observed, which served as a control in the experiment.

Activation of only the SMASh tag by exposure to ASV and subsequent stabilisation of the degron tag on TopBP1 did not result to any obvious degradation of the TopBP1 protein in the tagged cell lines. The miniAID system, on the other hand, when activated by exposure to Dox and 2h later addition of IAA, was found to be more efficient in promoting the degradation of TopBP1 within the time period of 6h studied. Interestingly, making use of both the miniAID and SMASh tag systems led to a similar depletion of TopBP1 within 6h after addition of ASV and IAA. A drawback of this miniAID-SMASh system is that, in the absence of any drugs, the levels of TopBP1 protein in the tagged cell lines are decreased compared to wild-type RPE cells. In fact, tagging of one of the *TopBP1* alleles with miniAID-SMASh causes a reduction of the TopBP1 levels of untreated cells whereas tagging of both alleles leads to a further reduction by default.

To summarize my results so far, the miniAID system is more potent than the SMASh system in promoting degradation of the tagged protein. Additionally, tagging of one or both of the endogenous *TopBP1* alleles with the miniAID-SMASh degron results to a slight reduction of the protein levels by default. Nonetheless, this double degron system can be activated to drastically reduce the levels of the tagged TopBP1 protein beyond the extent of the default situation.

Having established that the combined miniAID-SMASh degron leads to the depletion of the protein of interest, I then wanted to investigate the kinetics of this depletion since the 6h period examined so far was not sufficient to completely eliminate TopBP1 from the cells. FACS analysis was used this time to follow cells for a period of 24h following activation of the miniAID, SMASh or both systems. This was considered a more informative experiment than performing Western blot analysis, as disappearance of the TopBP1 band from the Western blot would not necessarily mean that TopBP1 is entirely depleted. Instead, factors like sensitivity of the antibody or exposure time could result to a false interpretation of the effect of degron activation on the levels of the protein. On the other hand, maximum activation of the miniAID-SMASh system and subsequent depletion of TopBP1 would prevent cells in culture from actively replicating their genomes as TopBP1 is essential for the initiation step of DNA replication. Pulse labelling cells with EdU, therefore, would allow the identification of the time point following addition of the drugs, that TopBP1 depletion would have reached low enough levels to render cells non-replicating hence unable to incorporate EdU.

TopBP1$^{+/+}$ (wild-type), *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ and *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells were exposed to ASV (3μM) alone, IAA (500μM) alone or a combination of both. Activation of the miniAID required pre-treatment of cells with Dox (1μg/ml) for a period of 2h prior to addition of IAA, as before. FACS data presented in **Figure 6.11 B** show that depletion of the TopBP1 protein expressed from one or two of the TopBP1 alleles leads to a proportional reduction of EdU positive cells, compared to *TopBP1*$^{+/+}$ (wild-type) cells. What is more, activation of the miniAID degron results in a more potent reduction of EdU positive cells compared to SMASh degron activation at 24h post drug addition. Activation of both degrons leads to an inability of *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells to incorporate EdU, whereas *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ cells show a partial inability and *TopBP1*$^{+/+}$ no defect at all. Overall, combined activation of miniAID and SMASh for 24h is sufficient for cells to stop actively replicating their DNA, without causing cell death, within the time frame studied. However, the effect of both degrons is similar to the effect of miniAID degron activation alone with respect to EdU incorporation.

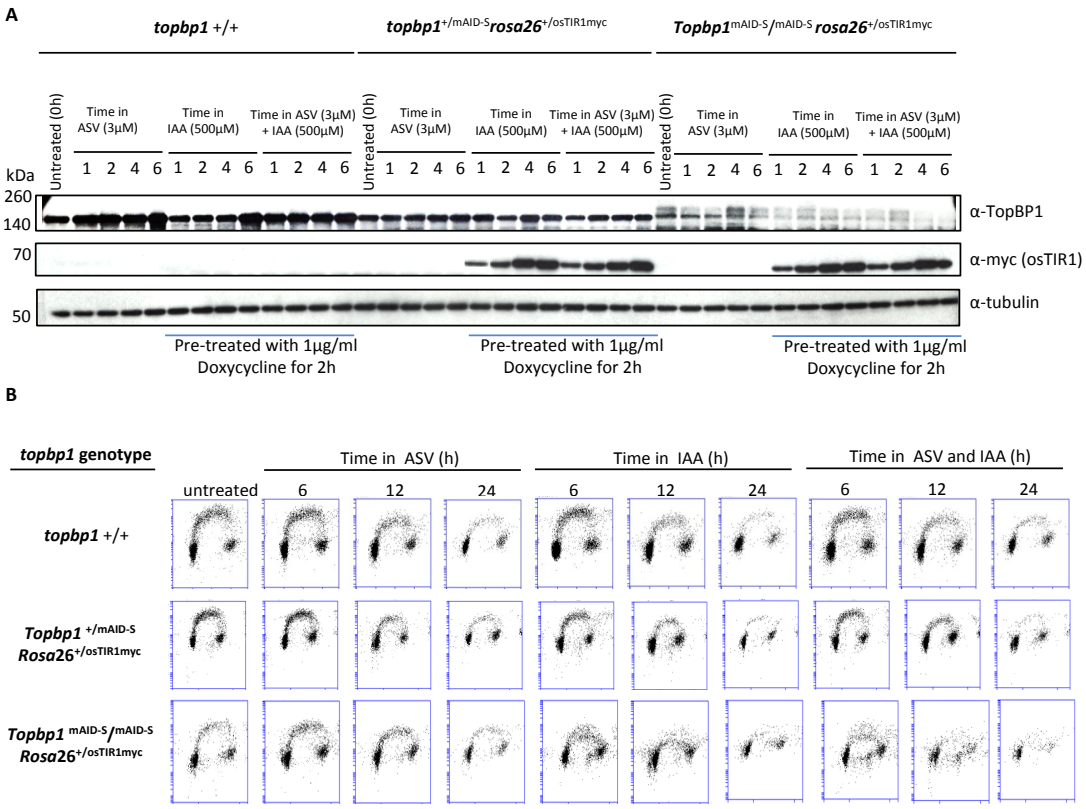


Figure 6.11: Kinetics of TopBP1 destruction following activation of the miniAID-SMASh system. **A)** Detection of TopBP1 and OsTIR1 using anti-TopBP1 and anti-myc antibodies of whole cell extracts of the indicated cell lines run on an SDS-PAGE gel. Cells were treated with either ASV (3μM) alone to activate the SMASh tag or Dox (1 μg/ml for 2h) and IAA (500 μM) to activate the miniAID tag or a combination of all to make use of the degrading ability of the combined miniAID-SMASh system. Cells were followed for a period of 6h and samples were collected at 0 (untreated), 1, 2, 4 and 6h post drug addition. Note that Dox was added 2h prior to the 0h time point and was not washed off. Beta-tubulin served as the loading control. **B)** Cell cycle distribution of the indicated cell lines at 0, 6, 12 and 24h post addition of ASV (3μM) alone or Dox (1 μg/ml fro 2h) and IAA (500 μM) or a combination of all, as measured by EdU incorporation and DNA content Flow cytometry analysis.

Having established that 24h of miniAID-SMASh activation is sufficient to prevent cells from replicating due to TopBP1 depletion, I then wanted to test what effect this depletion has on the checkpoint proficiency of the respective cell lines. To do so, *TopBP1*$^{+/+}$ (wild-type), *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ and *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells were exposed to ASV (3μM) and IAA (500μM) for a period of 36h. Activation of the miniAID required pre-treatment of cells with Dox (1μg/ml) for a period of 2h prior to addition of IAA, as before. About 2h before the final harvesting samples of each cell line were also exposed to HU (0.25mM) or IR (2Gy). At 6, 12, 24 and 36h post-drug addition samples were collected for Western blot analysis. As a negative control in this experiment cell lysates of untreated cells at 0 and 36h were also collected.

As depicted in **Figure 6.12**, exposure of wild-type cells to ASV and IAA had no effect on the levels of total TopBP1 inside the cells. On the other hand, *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ cells showed a reduction in TopBP1 12h after activation of the degrons, which became more potent at 24 and 36h post-degron activation. This suggests that the default (i.e. degrons not activated by drug addition) depletion of the miniAID-SMASh tagged protein reduced TopBP1 to levels that produce an unwanted phenotype. Furthermore, 6h following activation of the degron system TopBP1 was no longer detectable by the anti-TopBP1 antibody. By 36h, phosphorylation of S139 γH2AX was observed even in otherwise undamaged cells, suggesting that depletion of TopBP1 *per se* caused DNA damage.

The extent of this phosphorylation in *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ undamaged cells is comparable to samples treated with HU or IR. Interestingly, if we compare the three panels of blots from the three respective cell lines we observe that the extent of depletion of TopBP1 is proportional to the amount of γH2AX phosphorylation observed, with wild-type cells presenting hardly any, *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ cells presenting some and, strikingly, *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells showing increased phosphorylation of γH2AX following TopBP1 depletion. It also worth-mentioning that treatments of 0.25mM HU for 2h and 4Gy of IR were sufficient to make *TopBP1*$^{\text{mAID-S/mAID-S}}$ *rosa26*$^{+/\text{osTIR1myc}}$ cells that have been depleted of TopBP1 detach from the culture plate, something that was not observed in the *TopBP1*$^{+/+}$ (wild-type) and *TopBP1*$^{+/\text{mAID-S}}$*rosa26*$^{+/\text{osTIR1myc}}$ cells.

To sum up, activation of the miniAID-SMASh double degron efficiently depleted homozygously tagged cells of their TopBP1 protein pool leading to an inability

to actively replicate as well as to accumulation of DNA damage evident from increased S139 phosphorylation on γH2AX.

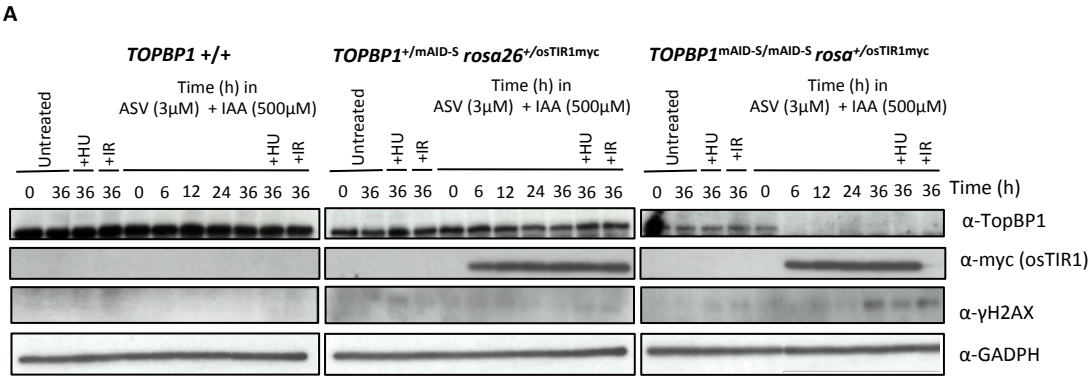


Figure 6.12: Depletion of TopBP1 in human cell leads to DNA damage. The indicated cell lines were pre-treated with Dox (1 μg/ml for 2h) before the addition of ASV (3μM) and IAA (500μM). Whole cell extracts were collected at 0, 6, 12, 24 and 36h post addition of the drugs and run on an SDS-PAGE gel. Untreated samples were also collected at the start (0h) and end (36h) of the experiment serving as a negative control. In addition, 2h prior to harvesting both untreated and ASV/IAA treated cells were either exposed to HU (0.25mM for 2h) or IR (4Gy) to activate the checkpoints. Antibodies were used to visualise TopBP1, Os-TIR1myc, pS139 γH2AX whereas GADPH served as the loading control.

Summary for chapter 6

This work describes the establishment of a TOPBP1 knock in gene targeting system, useful for the generation of mutant cell lines of interest in the DT40 model organism. In fact, depending on the researchers needs, all three, two or one of the endogenous TOPBP1 alleles can be mutated in *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ or *TopBP1*$^{-/-/+}$ cell lines, respectively, leading to the replacement of the total wild-type TopBP1 protein inside the cell by the mutant. The use of this system, however, requires the characterization of the events induced by the deletion of the alleles *per se* (in *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ systems) before proceeding to the characterization of the mutants. In the same way, using the wild-type system to knock in all three TopBP1 alleles requires characterization of the *TopBP1*$^{+/+/+}$ system prior to integration of the mutations of interest. Here I showed that successive deletion of the avian *TopBP1* alleles leads to a progressive reduction of the TopBP1 protein levels. Interestingly, this reduction is accompanied by progressively decreasing cellular proliferation and clonogenic ability, increasing sensitivity to DNA damaging agents, altered cell cycle profile as well as mild defects in the checkpoint response following replication stress or DNA damage. Such defects are associated with the extent of genotoxic stress imposed on the cells, with milder HU or IR treatments yielding milder defects in the observed phenotypes and harsher treatments exacerbating those defects.

Finally, to deplete TopBP1 in human cells in an allele-specific manner, I tagged one or both of the *TOPBP1* alleles of RPE cells with the miniAID-SMASh degron. Tagging of the *TopBP1* alleles with miniAID-SMASh leads to a reduction of the TopBP1 levels by default. This suggests that there is an inherent instability to the miniAID-SMASh system, mediated either by a default activation of miniAID due to leaky expression of the OsTIR1 transgene or by an incomplete NS3-mediated auto cleavage of the SMASh degron off TopBP1 or indeed by a combination both. Alternatively, tagging of the TopBP1 protein *per se* may have an effect on the stability of the protein. Nonetheless, activation of the degron tags by drug addition in cells that are heterozygous leads to depletion of the TopBP1 population produced from the tagged allele whereas applying this technology in homozygously tagged cells allowed the complete or near-complete depletion of the entire TopBP1 pool inside the cells. Such complete depletion of TopBP1 led to a progressive elimination of S phase cells from the culture and persisting damage on the chromatin (even in the absence of genotoxic stress, as judged by pS139 γH2AX).

The two systems described herein allowed a characterization of the effects of *TopBP1* gene dosage as well as TopBP1 protein elimination on the checkpoint proficiency of DT40 and RPE cells, respectively. More importantly, they represent useful tools for future structure-function studies of TopBP1 in DT40 and human cell lines.

Chapter 7

In vivo characterization of the avian TopBP1 ATR activation domain and its functions

Preface

In addition to gene knockouts, reverse genetics can be used for knock in studies. Apart from eliminating a gene and observing the phenotypes generated from loss of function of the protein product, it is often useful to specifically mutate the gene and investigate the structure and biological activity of the mutant protein. Generating separation of function mutants allows the identification of specific roles of our favourite protein within distinct cellular pathways and provides us with a powerful tool to understand the underlying molecular mechanisms in a genetically defined system *in vivo*.

7: *In vivo* characterization of the avian TopBP1 ATR activation domain and its function

TopBP1 performs vital roles for the correct functioning of the checkpoint in response to replication stress and DNA damage. Firstly, by acting as a molecular scaffold, TopBP1 bridges between the DNA damage sensors and the checkpoint mediators, which present the effector kinases to ATR/ATM. It thus performs an important skeletal function for the propagation of the checkpoint signalling. Additionally, TopBP1 harbours the ATR Activation Domain (AAD), which in higher eukaryotes is necessary for ATR activation. Replacing *Xenopus* TopBP1 with recombinant TopBP1 mutated within the AAD (W1138) completely abolished ATR activation in aphidicolin-treated egg extracts (Kumagai et al 2006). TopBP1 AAD was also described to be essential for early embryonic development in mice as mutation of the AAD arrested development at the blastocyst stage (Zhou et al 2013). Furthermore, TopBP1 has been described as having a functional role in a pathway that connects ATM to ATR at sites of DSBs. In fact, ATM-catalysed phosphorylation of *Xenopus* TopBP1 S1131 was shown to be necessary for activation of ATR-ATRIP in response to DSBs but not replication stress (Yoo et al 2007). This favours a model whereby phosphorylation of this residue by ATM may mediate a handover from ATM to ATR activation at DSBs by increasing the ability of TopBP1 to stimulate the kinase activity of ATR via the AAD.

The multifaceted nature of TopBP1 render it an integral part of the DDR, required both for the generation and sustaining of the checkpoint signal in response to DNA lesions but also for the correct transmission of the signal across the checkpoint

cascade. Interestingly, these functions of TopBP1 are conserved from yeasts to humans but the precise molecular mechanisms have evolved. So although the yeast TopBP1$^{Rad4/Dpb11}$ AAD is dispensable for ATR activation, the architecture of ATR activation in higher eukaryotes is thought to be largely dependent on TopBP1 AAD. In order to investigate the function of the TopBP1 AAD *in vivo*, a knock-in gene-targeting approach was designed to introduce point mutations of interest into the intact *TopBP1* genomic locus of *TopBP1*$^{-/-/+}$ DT40 cells. The use of the *gallus gallus*-derived DT40 cells would hopefully help us avoid the issue of embryonic lethality of the AAD mutants and hence allow a proper *in vivo* characterisation of the resulting phenotypes. The work presented here investigates the function of the DT40 TopBP1 AAD and in particular of residues S1132 and W1139 using a knock-in gene targeting strategy, based on the gene dosage knock in platform described in the previous chapter.

7.1.1: Generation of AAD mutant DT40 cells

The characterisation of the *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/+/+}$ and *TopBP1*$^{-/-/+}$ gene dosage systems (Chapter 6) allows their use as knock in gene targeting systems for the study of TopBP1. Here this system was used to knock in point mutations of interest within the intact *TopBP1* locus of *TopBP1*$^{-/-/+}$ cells. In this way, the population of the TopBP1 protein inside the cells would be replaced by a mutant population of interest (in a single gene targeting step). The advantage of mutating the endogenous *TopBP1* allele is that the gene will still be expressed under the control of the endogenous promoter and subjected to physiological cell cycle regulation or mRNA splicing. Such a system would enable the analysis of TopBP1 mutants under physiological conditions *in vivo* and would help avoid potential undesirable effects often associated with overexpressing the cDNA from a viral promoter.

The knock-in targeting strategy devised herein makes use of homologous recombination to integrate point mutations within the genomic locus of interest. Briefly, a selection marker cassette flanked by lox sites and long homology arms (where the mutation is incorporated) are synthesized in a plasmid. The arm regions are chosen in such a way so as to integrate the marker cassette within an intronic region and without affecting the amino acid frame downstream following successful integration. PCR analysis and Restriction Fragment Length Polymorphism (RFLP) are used to screen for transfected clones that have both integrated the marker cassette at the locus of interest and have incorporated the mutation. Having characterized the *TopBP1*$^{-/-/+}$ cell line in

terms of its replication properties, response to replication stress and DNA damage compared to its wild-type counterpart, I decided to use it as the parental cell line for the generation of the TopBP1 mutants. This would offer the advantage of having to perform only one targeting step to create a cell line where all the wild-type TopBP1 population would have been replaced by the mutant. It is also worth noticing that the *TopBP1*$^{-/-/+}$ cell line used herein contains the MerCreMer enzyme stably integrated into the genome (described in chapter 3).

The first biological question addressed concerned the physiological function of the *gallus gallus* TopBP1 AAD in checkpoint signaling. The TopBP1 AAD in chickens spans exons 19 and 20, with the core indispensable residue W1139 as well as the ATM-targeted residue S1132 both encoded within exon 20 (**Figure 7.1 A**). To change serine 1132 to alanine (S1132A), T3394 was mutated to G3394 and to change tryptophan 1139 to arginine (W1139R), T3415 was mutated to C3415 in the remaining intact *TopBP1* allele of *TopBP1*$^{-/-/+}$ cells. In parallel, I wanted to create cell lines that would be silently mutated for the exact same residues, which would serve as a control for future experiments. Thus C3396 was mutated to A3396 (S1132S) and G3417 was mutated to A3417 (W1139W). Overall, four knock-in targeting constructs were assembled, namely AAD-SAki and its control AAD-SSki as well as AAD-WRki and its control AAD-WWki (**Figure 7.1 B**). The cloning strategy involved two cloning steps. First, a 2kb left homology arm and a 2kb right homology arm separated by a *BstZI71* restriction site were synthesized as a *NotI/XhoI* fragment in pUC57 vector. The sequence of the arms was retrieved from the genome browser so it was the wild-type sequence of *TopBP1*. The unique *BstZI71* cutting site was used to subclone the lox-SV40-puromycin-polyA-lox fragment, which was amplified from the previously described LA^{i}RA^{i}Puro construct with primers P68/P69. For the second cloning step I made use of two unique restriction sites that flanked the Exon20 region of interest. In fact, this region was flanked by *EcoRV* and *SpeI* restriction sites that did not cut anywhere else in the insert or the pUC57 backbone. Thus, *EcoRV/SpeI*-flanked gene block fragments that were homologous to wild-type *TopBP1* but contained the mutation of interest (SA, SS, WR, WW) were synthesized from IDT and subcloned in the targeting vector. The resulting knock-in targeting constructs, AAD-Saki (S1132A), AADSSki (S1132S), AAD-WRki (W1139R) and AAD-WWki (W1139W) were verified by Sanger sequencing.

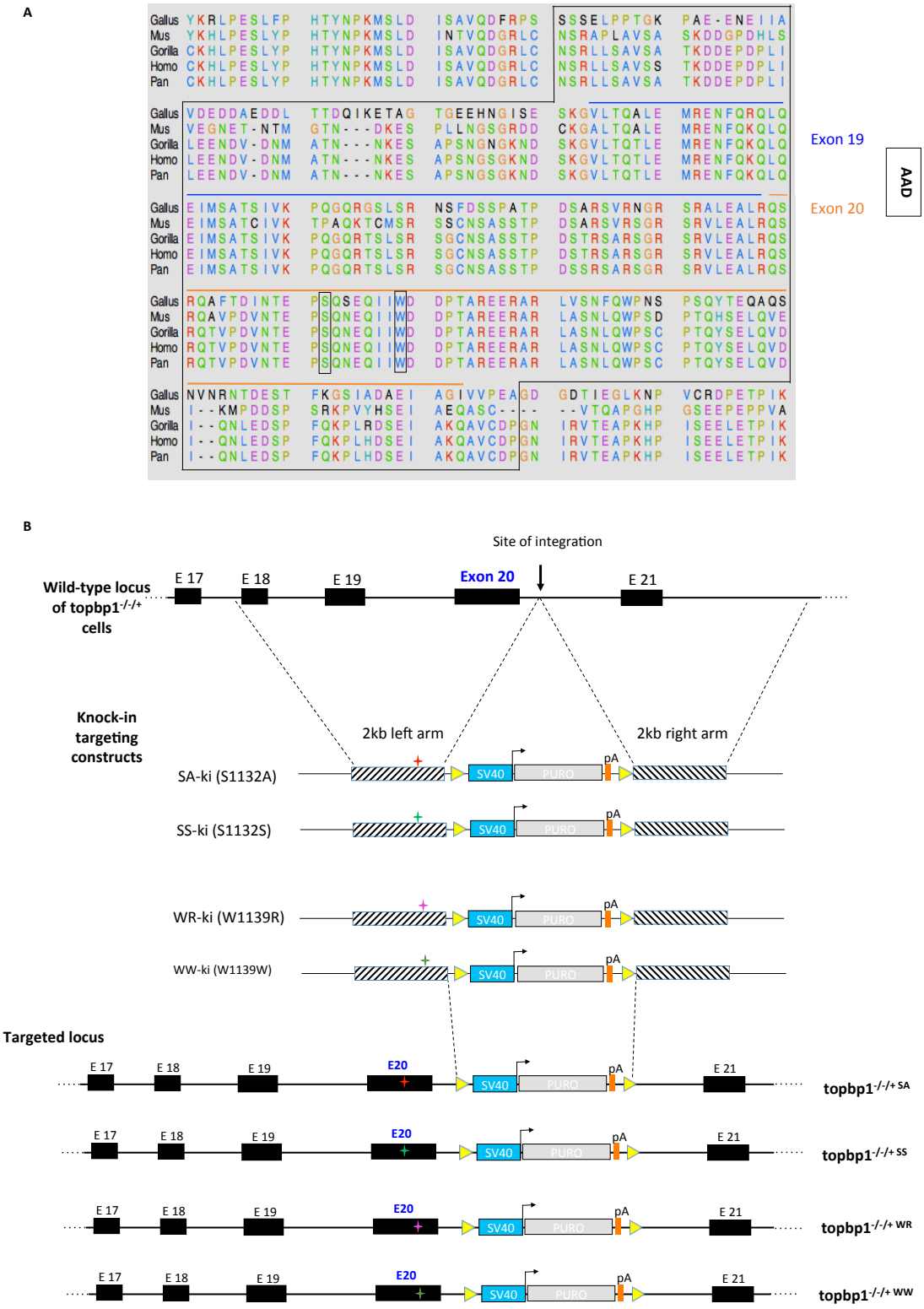


Figure 7.1: Knock-in targeting strategy for AAD mutants. A) Structure alignment of TopBP1 AAD obtained using the ClustalW sequence alignment software. Alignment of the AAD sequences from different species is shown, with the AAD domain being enclosed in the black box. The sequence encoded by exon 19 is highlighted by the blue line and that encoded by exon 20 by the yellow line. **B)** Schematic representation of the AAD genomic locus, the knock-in targeting constructs used and the *TopBP1* locus successfully targeted with the respective vectors. Coloured crosses represent the novel restriction sites generated after gene targeting to allow RFLP analysis of the obtained clones.

Knock-in targeting vectors were linearized by *AhdI* digestion and electroporated into *TopBP1*$^{-/-/+}$ cells. Transformants were selected in the presence of puromycin and 5-7 days later puromycin-resistant clones were expanded and screened by PCR analysis for successful knock-in targeting. More specifically, the screening strategy was a two-step process. First, PCR with primers P70 and P71 was performed to screen for cells that had successfully integrated the knock-in construct within intron 21 of the intact *TopBP1* locus of *TopBP1*$^{-/-/+}$ cells. P70 annealed to a genomic region 5' with respect to the left homology arm and P71 annealed within the selection marker cassette (**Figure7.2 A**). In this way the PCR product would be specific to the locus of interest avoiding the selection of false positive clones that have erroneously integrated the targeting construct. Of the 35 potential *TopBP1*$^{-/-/+SS}$ and of the 45 potential *TopBP1*$^{-/-/+SA}$ clones screened, 9 and 7 were classified as positive according to PCR analysis, respectively. Additionally, of the 28 potential *TopBP1*$^{-/-/+WW}$ and of the 55 potential *TopBP1*$^{-/-/+WR}$ clones screened, 4 and 10 were classified as positive according to PCR analysis, respectively (data not shown).

Next, the method of RFLP was employed whereby incorporation of the mutation would create a novel restriction site allowing in this way identification of successfully targeted clones. More specifically, successful targeting of the AAD-SAki construct and incorporation of the S1132A mutation would generate a novel *NaeI* site, whereas successful targeting of the AAD-WRki construct and incorporation of the W1139R mutation would generate a novel *NciI* site. Clones successfully targeted with the control constructs, AADSSki (S1132S) and AAD-WWki (W1139W), would not contain such restriction sites. Therefore in the second step of the screening strategy, the PCR product of clones classified as positive from step 1 was digested with the appropriate restriction enzyme to identify those clones that have incorporated the mutation of interest. As shown in **Figure 7.2 B**, all *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+WW}$ control clones did not present fragment size variation before and after digestion, as expected (compare top and bottom of the gel). In contrast, 3 of the 7 potential *TopBP1*$^{-/-/+SA}$ clones presented bands corresponding to *NaeI*-digested DNA fragments suggesting successful incorporation of the S1132A mutation. Additionally, 7 of the 10 potential *TopBP1*$^{-/-/+WR}$ clones presented bands corresponding to *NciI*-digested DNA fragments suggesting successful incorporation of the W1139R mutation. All clones classified as positive by the second step of the screening strategy were confirmed by Sanger sequencing and appropriately frozen down in liquid nitrogen.

Finally, to remove the selection marker cassette from intron 21, the stably integrated Cre recombinase was induced by treatment of cells with 2μM 4-HT. 24h later, treated cells were serially diluted and 5-7 days later single colonies were expanded and tested for their sensitivity to puromycin. All of the single clones isolated have lost resistance to puromycin suggesting successful removal of the selection cassette from the *TopBP1* locus (**Figure 7.2 C**). This was also confirmed by PCR analysis with primers P70 and P72 (data not shown).

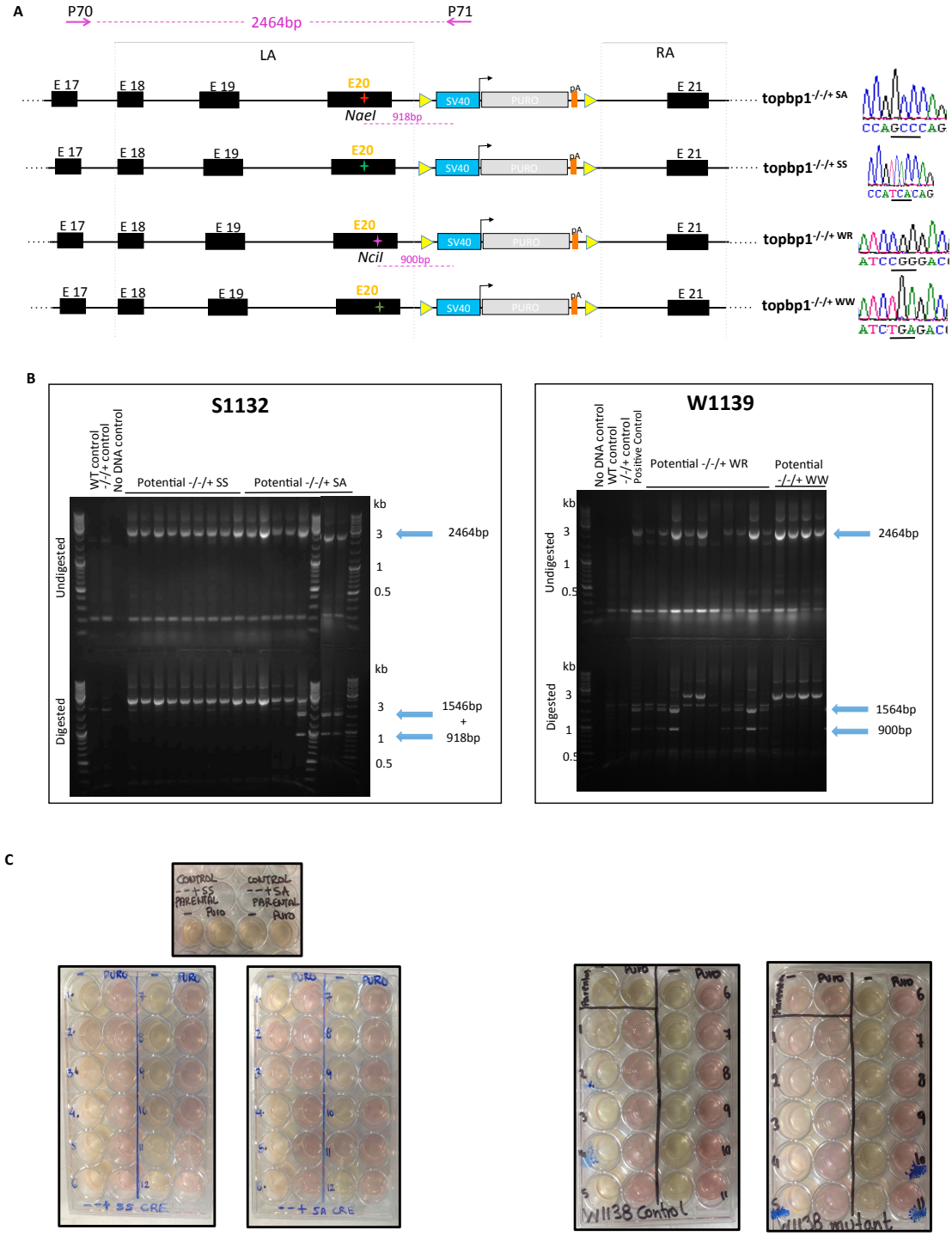


Figure 7.2: Generation of AAD mutant DT40 cells. A) Schematic representation of the AAD locus of *TopBP1*$^{-/-/+SS}$, *TopBP1*$^{-/-/+SA}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cell lines following successful knock-in targeting. On the right four panels show the sequencing data of the genomic DNA from the aforementioned cell lines, confirming the mutation of TCC to GCC (S1132A) and of TGG to CGG (W1139R) as well as the corresponding silent mutations in the control cell lines (see text for details).
B) RFLP analysis of puromycin resistant clones obtained following transfection of the knock-in targeting constructs into the TopBP1$^{-/-/+}$ cell line. PCR was performed with primers P70 and P71 shown in A and the PCR products were digested with *NaeI* (SA) or *NciI* (WR) to confirm successful integration of the point mutations. **C)** Positive clones obtained from B were incubated with 2μM 4-HT for 24h and serially diluted to obtain single clones that have successfully floxed the selection cassette and have become puromycin sensitive again. The parental cell lines served as the control on the experiment.

7.1.2: Inactivation of TopBP1 S1132 leads to a defective checkpoint response to replication stress caused by hydroxyurea

To gain insight into a potential functional role of the TopBP1 S1132 residue during replication stress, colony formation assays in HU were initially performed for *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cell lines. The data obtained from continuous exposure to HU using semi-solid media suggested that *TopBP1*$^{-/-/+SA}$ cells are more sensitive to replication stress caused by HU than their *TopBP1*$^{-/-/+SS}$ control counterparts. More specifically, the mutant strains seemed to tolerate concentrations of HU up to 20μM but their sensitivity became apparent at higher concentrations (40-80μM). The data showed that the S1132 residue of TopBP1 is important for the ability of DT40 cells to survive HU-caused genotoxic stress (**Figure 7.3 A**).

To further explore this increased HU sensitivity, *TopBP1*$^{-/-/+SA}$ cells were examined for their ability to activate the G1-S and S-M as well as the G2-M DNA damage checkpoints when treated with HU. To do so, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells were grown in the continuous presence of 1mM HU for 6h and samples for Western blot, FACS and microscopic analyses were taken at 0, 1, 2, 4 and 6h post addition of the drug. As shown in **Figure 7.3 B**, treatment of wild-type cells with 1mM HU led to a slight increase in the levels of TopBP1 at 6h after addition of HU. Additionally, HU addition led to phosphorylation of pS345 Chk1 after 1h to levels 22.9-fold over the untreated. This phosphorylation was maintained at similar levels until 2h and during the last two hours of the time course it further increased to 24.6- and 27-fold. Phosphorylation of S139 γH2AX increased to 23.9-fold 1h post HU addition and reached a maximum of 38.6-fold over untreated at 2h whereas at later time points it reduced to 19.1- and 22.4-fold. The control strain *TopBP1*$^{-/-/+SS}$ displayed an altered response similar to the response of *TopBP1*$^{-/-/+}$ compared to wild-type, as explained in Chapter 6. In fact, *TopBP1*$^{-/-/+SS}$ displayed a reduced ability of phosphorylating Chk1 with pS345 reaching levels of 13.8-, 16.3-, 23.4- and 21.6-fold over untreated at 1, 2, 4 and 6h, respectively. Phosphorylation of γH2AX was also impaired as *TopBP1*$^{-/-/+SS}$ cells failed to display the initial increase of pS139 observed in the wild-type cells. In fact, *TopBP1*$^{-/-/+SS}$ increased pS139 to only 17.7- and 12-fold over untreated 1h and 2h post HU addition, respectively. Checkpoint activation of *TopBP1*$^{-/-/+SS}$ persisted until 6h as indicative of Chk1 phosphorylation but the extent of pS345 was reduced compared to wild-type cells. Interestingly, the *TopBP1*$^{-/-/+SA}$ mutant presented an even more defective checkpoint response than its *TopBP1*$^{-/-/+SS}$ counterpart control. Mutation of the S1132

residue of TopBP1 interfered with the ability of DT40 cells to both phosphorylate S345 Chk1 and S139 γH2AX. *TopBP1*$^{-/-/+SA}$ cells phosphorylated S345 with the same kinetics as the control cells and following the same pattern but the extent of pS345 Chk1 was consistently reduced. In fact, they presented a 12.3- and 14.3-fold increase at 1 and 2h of HU exposure and reached a maximum of 17.6-fold after 4h (versus 23.4-fold observed in the control cells) before a slight drop to 17.1-fold at the end of the time course. Furthermore, consistent with a defective checkpoint response, *TopBP1*$^{-/-/+SA}$ mutants showed only an 8.2- and 9.5-fold increase over untreated of pS139 γH2AX when exposed to 1mM HU for 1 and 2h, respectively. Phosphorylation of γH2AX remained at low levels at 4h and reached a maximum value at 6h post HU addition (7.4- and 10.4 fold over untreated, respectively). As for the levels of the TopBP1, the mutant cells were able to increase protein levels over the untreated in a manner very much like the wild-type cells so in that sense they were not defective, which comes in contrast to their *TopBP1*$^{-/-/+SS}$ counterparts. However, it is worth noticing that the levels of the TopBP1 protein pool inside the mutant cells are by default reduced compared to *TopBP1*$^{-/-/+SS}$. In fact, quantification of the TopBP1 levels over background showed that *TopBP1*$^{-/-/+SA}$ mutants contain only 56% of the TopBP1 protein levels found in *TopBP1*$^{-/-/+SS}$, raising the question of whether the S1132A mutation influences the stability of the protein.

To assess whether treatment of the mutant cells with HU interfered with cellular integrity, samples from the 6h time course were fixed on glass slides, stained with DAPI and analyzed by microscopy. This experiment showed that HU treatment led to a progressive accumulation of fragmented nuclei –a hallmark of apoptosis- in the cultures of the mutant cells. In fact, 4h in HU were enough to kill 6.5% of the mutant cells when only 3.5% of *TopBP1*$^{-/-/+SS}$ contained fragmented nuclei. By 6h the extent of fragmented nuclei further increased to 4.5% and 8% for the *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$, respectively (**Figure 7.3 C**).

Cell cycle analysis by flow cytometry (**Figure 7.3 D**) revealed that despite the slight defects of the S132A mutants in the strength and kinetics of checkpoint activation in response to HU, they were able to arrest the cell cycle by 6h similarly to the wild-type and control counterparts. Importantly, to ensure that the phenotypes observed were not specific to a single *TopBP1*$^{-/-/+SA}$ mutant clone, four different isolates were analysed by FACS and two different isolates by Western blot.

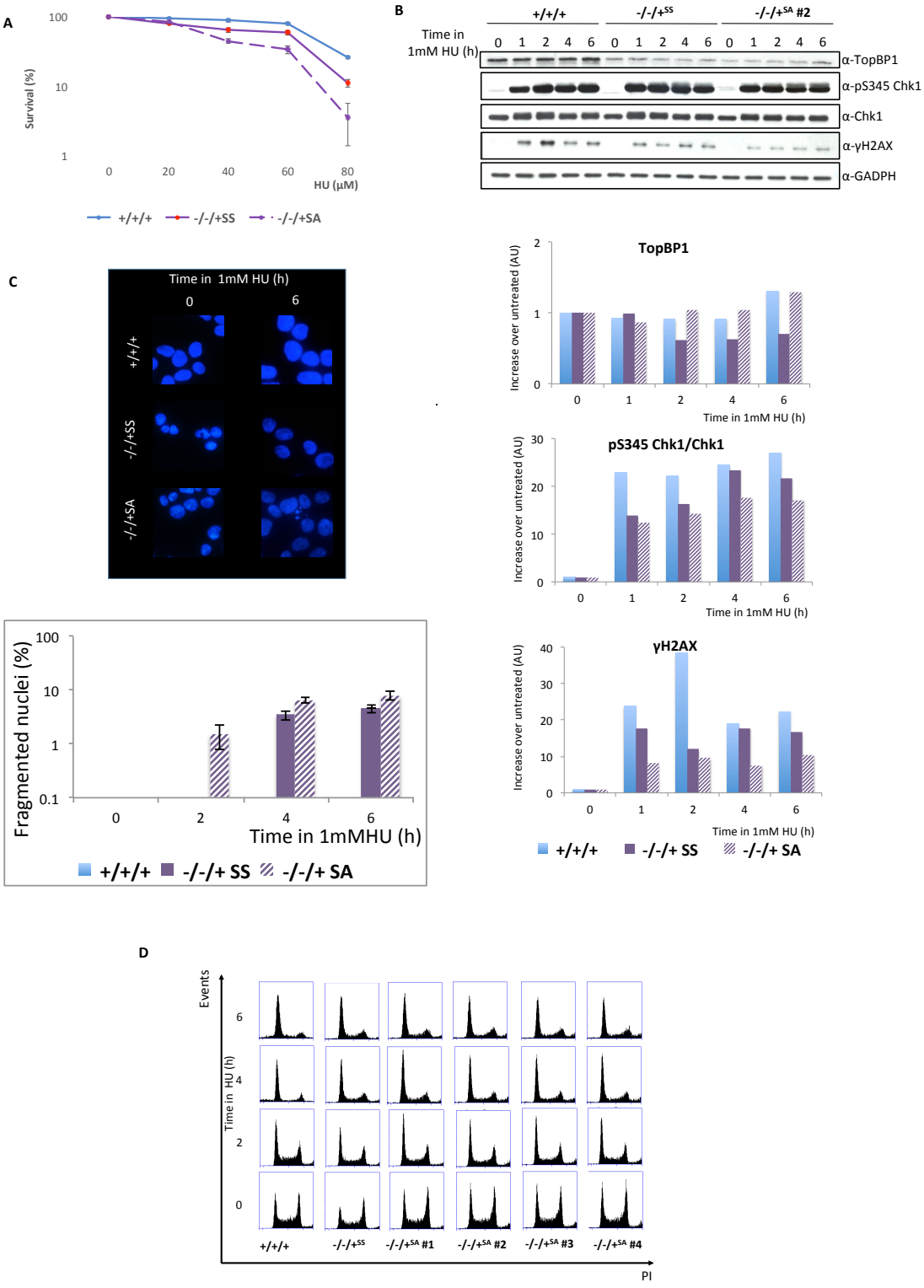

A
100
10
1
Survival (%)
0
20
40
60
80
HU (μM)
+/+/+
-/-/+SS
-/-/+SA
B
+/+/+
-/-/+SS
-/-/+SA #2
Time in 1mM HU (h)
0 1 2 4 6 0 1 2 4 6 0 1 2 4 6
α-TopBP1
α-pS345 Chk1
α-Chk1
α-γH2AX
α-GADPH
C
Time in 1mM HU (h)
0
6
+/+/+
-/-/+SS
-/-/+SA
TopBP1
Increase over untreated (AU)
Time in 1mM HU (h)
pS345 Chk1/Chk1
Increase over untreated (AU)
Time in 1mM HU (h)
γH2AX
Increase over untreated (AU)
Time in 1mM HU (h)
Fragmented nuclei (%)
100
10
1
0.1
0
2
4
6
Time in 1mMHU (h)
+/+/+
-/-/+ SS
-/-/+ SA
D
Events
Time in HU (h)
6
4
2
0
+/+/+
-/-/+SS
-/-/+SA #1
-/-/+SA #2
-/-/+SA #3
-/-/+SA #4
PI

Figure 7.3: Checkpoint activation defect of S1132A TopBP1 mutant in response to continuous replication stress. A) Colony survival assay of the indicated cell types plated singularly in HU-containing semi-solid media at the indicated concentrations. Error bars represent the standard deviation of the mean for three independent experiments. **B, C, D**) The indicated strains were treated with 1mM HU and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **B**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. Charts at the bottom represent quantification of the Western blots using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **C**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **D**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

To investigate whether the observed checkpoint response defects of the *TopBP1*$^{-/-/+SA}$ mutants were correlated with the extent of genotoxic stress, cells were treated with lower doses of HU and checkpoint activation assessed in the same way as before. Treating asynchronously growing *TopBP1*$^{-/-/+SA}$ cells with 0.25mM HU resulted to increased TopBP1 protein levels 2h post drug addition for both the wild-type and the *TopBP1*$^{-/-/+SS}$ controls (1.4-fold over untreated for both). At later time points TopBP1 levels returned to basal levels for wild-type cells whereas *TopBP1*$^{-/-/+SS}$ further increased TopBP1 to 1.6-fold over untreated at 4h and this increase was sustained until 6h. *TopBP1*$^{-/-/+SA}$ mutants, however, failed to increase their TopBP1 levels. Furthermore, although *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/-/+SS}$ presented a significant increase in the extent of S345 Chk1 phosphorylation as early as 1h post addition of HU (8.2- and 6.9-fold, respectively), *TopBP1*$^{-/-/+SA}$ mutants failed to do so. At later time points, wild-type cells sustained pS345 Chk1 at constant levels of around 5-fold over untreated. Similarly, *TopBP1*$^{-/-/+SS}$ maintained the S345 phosphorylated pool of Chk1 but the extent of phosphorylation was lower than wild-type at 2 and 4h (~4-fold over untreated) and slightly higher at 6h (5.6-fold). Interestingly, 2h post drug addition mutant cells displayed pS345 Chk1 of levels comparable to the control cell lines, but during later time points they sustained pS345 at lower but steady levels. Additionally, phosphorylation of S139 γH2AX followed the same pattern as in Figure 7.3. More specifically, although the lower HU concentration used here ameliorated the defects of the *TopBP1*$^{-/-/+SS}$ controls, *TopBP1*$^{-/-/+SA}$ mutants still displayed a markedly reduced ability to phosphorylate S139 γH2AX (**Figure 7.4 A**).

Microscopic analysis of DAPI-stained time course samples revealed that the milder HU treatment helped reduce the number of dead cells in the cultures of *TopBP1*$^{-/-/+SA}$ mutants to 3.5% at 6h (when *TopBP1*$^{-/-/+SS}$ controls had 1% of their cells displaying an apoptotic phenotype) (**Figure 7.4 B**). Furthermore, cell cycle analysis by flow cytometry revealed that despite the defects of *TopBP1*$^{-/-/+SA}$ mutant cells in the strength and kinetics of checkpoint activation in response to 0.25mM HU, they were able to arrest the cell cycle by 6h similarly to their *TopBP1*$^{-/-/+SS}$ counterparts, albeit with slower kinetics (**Figure 7.4C**).

Taken together, the above data suggest that the TopBP1 S1132 residue is required for proper functioning of the replication stress checkpoint. Mutation of this phosphorylatable serine to a non-phosphorylatable alanine results to a defective checkpoint response characterized by an inability to increase TopBP1 protein levels and

abnormally reduced phosphorylation of Chk1 on S345 and γH2AX on S139. And although mutant cells successfully arrested at G1/S, they did so with slightly slower kinetics and at the expense of some cell death, as evident from DAPI staining of the nucleus. Milder replication stress by using lower concentrations of HU did not ameliorate the defects of the mutant cells in checkpoint activation, although it did improve the response of *TopBP1*$^{-/-/+SS}$ controls as has been already described in Chapter 6.

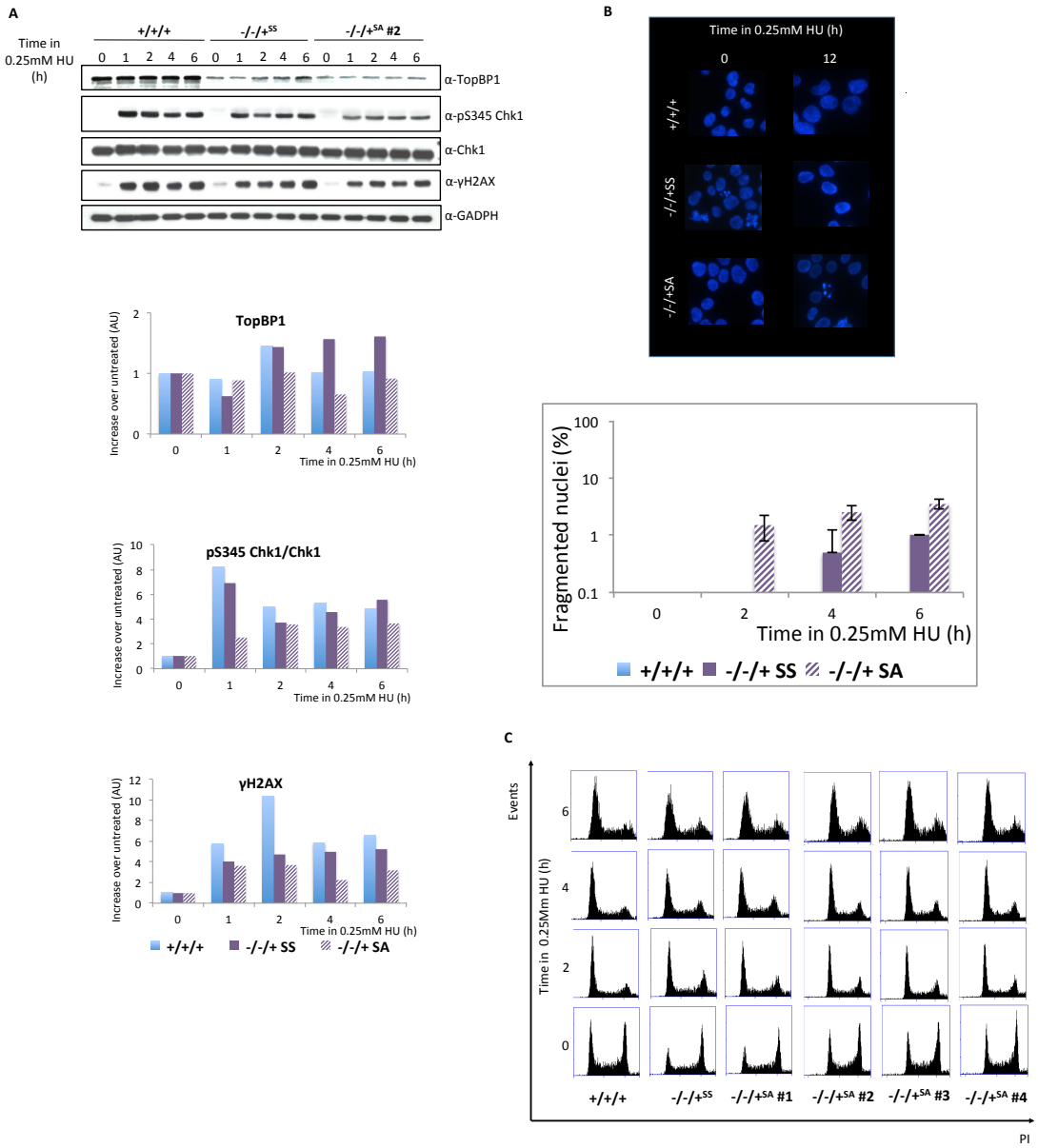


Figure 7.4: Milder checkpoint activation defect of S1132A TopBP1 mutant in response to lower HU-caused replication stress. A, B, C) The indicated strains were treated with 0.25mM HU and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **A**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. Charts at the bottom represent quantification of the Western blots using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **B**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **C**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

7.1.3: Inactivation of TopBP1 S1132 leads to a defective recovery from replication stress caused by hydroxyurea

To examine whether the S1132 residue of TopBP1 is also important for the recovery from replication stress, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells were arrested in G1/S by treatment with 0.25mM HU for 12h and then were washed clear of the drug and followed through a period of 6h recovery time. Western blot analysis (**Figure 7.5 A, B**) revealed that phosphorylation of S345 Chk1 and S139 γH2AX after exposure to HU as well as during release from the block was slightly impaired in a TopBP1 S1132-dependent manner. The 12h HU block resulted in a 9.2-fold increase of the pS345 Chk1 pool in *TopBP1*$^{+/+/+}$ cells and this phosphorylation persisted at elevated levels during the 6h recovery time. Consistent with the phenotypes observed in Figures 7.3 and 7.4, *TopBP1*$^{-/-/+SA}$ cells were impaired in their ability to phosphorylate Chk1 on S345 to the extent of the *TopBP1*$^{-/-/+SS}$ controls after 12h incubation with HU. Only a 5.3- and 4.7-fold increase in the pS345 Chk1 population was observed for the two *TopBP1*$^{-/-/+SA}$ independent clones tested when in fact *TopBP1*$^{-/-/+SS}$ controls displayed a 7.4-fold increase. From this analysis it is also observed that the mutant cells failed to maintain elevated levels of pS345 during the recovery period and so pS345 returned to near-basal levels. This is different from what was observed in wild-type cells, which displayed increased pS345 Chk1 even 6h post-release (6.8-fold over untreated versus 0.7- and 1.8-fold for the mutant clones), indicative of persisting checkpoint activation. Furthermore, *TopBP1*$^{-/-/+SS}$ cells were characterized by an increase in the population of phosphorylated γH2AX by the end of the HU block (22.9-fold over untreated), compared to *TopBP1*$^{+/+/+}$ (21-fold) and *TopBP1*$^{-/-/+SA}$ (13.3- and 11.7-fold for the two clones). During the release period, *TopBP1*$^{+/+/+}$ cells retained increased pS139 for two hours, before phosphorylation started to decline (18.6-, 17.3-, 5.7- and 4.2-fold at 1, 2, 4 and 6h post-release, respectively). *TopBP1*$^{-/-/+SS}$ control followed a similar trend like wild-type albeit with altered kinetics. In contrast, *TopBP1*$^{-/-/+SA}$ clones displayed a severe inability of increasing pS139 γH2AX. By the end of the HU block they managed to reach a 13.3- and 11.7-fold increase over the untreated levels but this quickly returned to basal levels at 4h post release.

At the same time analysis of the cell cycle profile by flow cytometry revealed that all the three cell types successfully halted the cell cycle in response to the HU

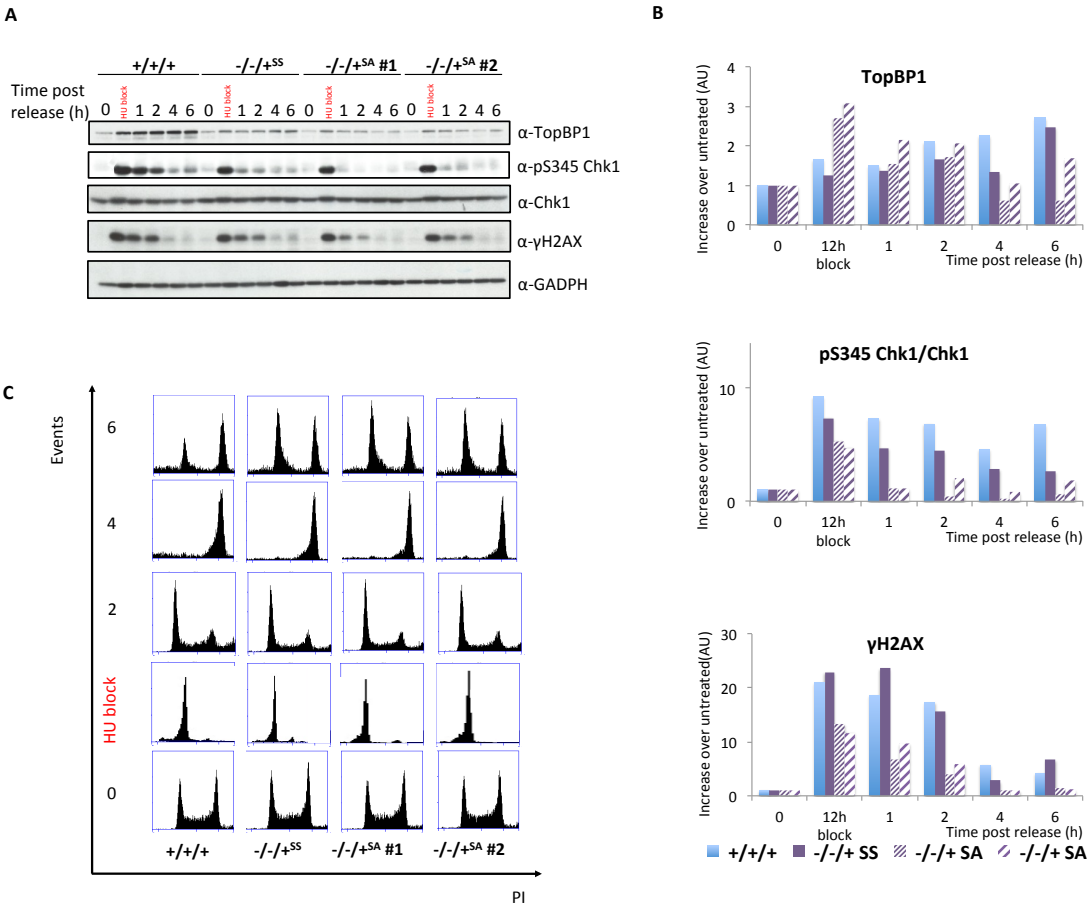


Figure 7.5: S1132 TopBP1-dependent recovery from replication stress. A, B, C) The indicated strains were treated with 0.25mM HU for 12h, washed clear of the drug and samples were kept at the specified times for analysis by Western blotting and FACS. **A**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. **B**) Quantification of the Western blots in A using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **C**) Flow cytometry analysis of samples fixed with propidium iodide.

treatment, suggesting that the slightly altered checkpoint response observed by Western blot analysis did not affect the cell cycle kinetics (**Figure 7.5 C**).

7.1.4: Inactivation of TopBP1 S1132 leads to a defective response to DNA damage caused by ionizing radiation and induces cell death.

To determine whether mutation of TopBP1 S1132 confers sensitivity to killing by IR, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells were exposed to 1 or 2Gy of IR and then plated onto semi-solid media to examine their colony forming ability, hence their replicative and survival potential following DNA damage. As shown, in **Figure 7.6 A**, the *TopBP1*$^{-/-/+SA}$ mutants displayed significantly reduced survival compared to *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/-/+SS}$ controls, which was exacerbated with increasing doses of radiation. It should be noted that the mutant cells did not survive higher doses of IR tested (4, 6 and 8Gy) despite the fact that 10-fold more cells were plated compared to wild-type.

To further explore a potential role of TopBP1 S1132 in damage-induced G2/M arrest, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells were exposed to 6Gy of IR and followed through a period of 6h recovery time. Analysis of whole cell extracts by Western blot at 0, 2, 4, 6 and 8h post-IR revealed a defective checkpoint activation response compared to control cell types. In fact, mutant cells were characterized by an inability to phosphorylate S345 Chk1 to the levels observed for their *TopBP1*$^{-/-/+SS}$ counterparts. So although *TopBP1*$^{-/-/+SS}$ showed 3.8-, 2-, 1.5- and 1.4-fold increase over untreated at 2, 4, 6 and 8h post-IR, respectively, *TopBP1*$^{-/-/+SA}$ showed only 2.4-, 1.2-, 0.8- and 1-fold increase at the same time points. Thus there was a consistently decreased ability of phosphorylating the effector kinase following DNA damage. Phosphorylation of γH2AX on S139 was also found to be dependent on TopBP1 S1132 as *TopBP1*$^{-/-/+SA}$ mutants displayed an inability to boost pS139 1h post IR to the extent of the *TopBP1*$^{-/-/+SS}$ controls (3.3/4.7- versus 9.1-fold over untreated). This reduced phosphorylation of γH2AX was consistent across all the time points tested suggesting an inherent defect in checkpoint activation of DT40 cells expressing only TopBP1 S1132A (**Figure 7.6 B**).

Interestingly, DAPI staining of the nuclei revealed that 6Gy of ionizing radiation sensitized *TopBP1*$^{-/-/+SA}$ mutants, as evident from the altered morphology of the nuclear chromatin. In fact, the number of fragmented nuclei in the *TopBP1*$^{-/-/+SA}$ cultures was almost two-fold higher at 4, 6 and 8h post-IR (9, 13.5 and 18.5%) compared to *TopBP1*$^{-/-/+SS}$ controls (4.5, 6.5 and 9%) (**Figure 7.6 C**).

Examination of the cell cycle phase distribution revealed that the *TopBP1*$^{-/-/+SA}$ mutants progressed towards a damage induced G2/M arrest with altered kinetics compared to both the wild-type and the *TopBP1*$^{-/-/+SS}$ controls. In fact, 8h following irradiation *TopBP1*$^{+/+/+}$ and the *TopBP1*$^{-/-/+SS}$ displayed a 4N DNA content indicative of a successful G2/M arrest. In contrast, although the majority of *TopBP1*$^{-/-/+SA}$ cells had reached a G2/M arrest, a subset still presented a DNA content corresponding to S phase (**Figure 7.6 D**).

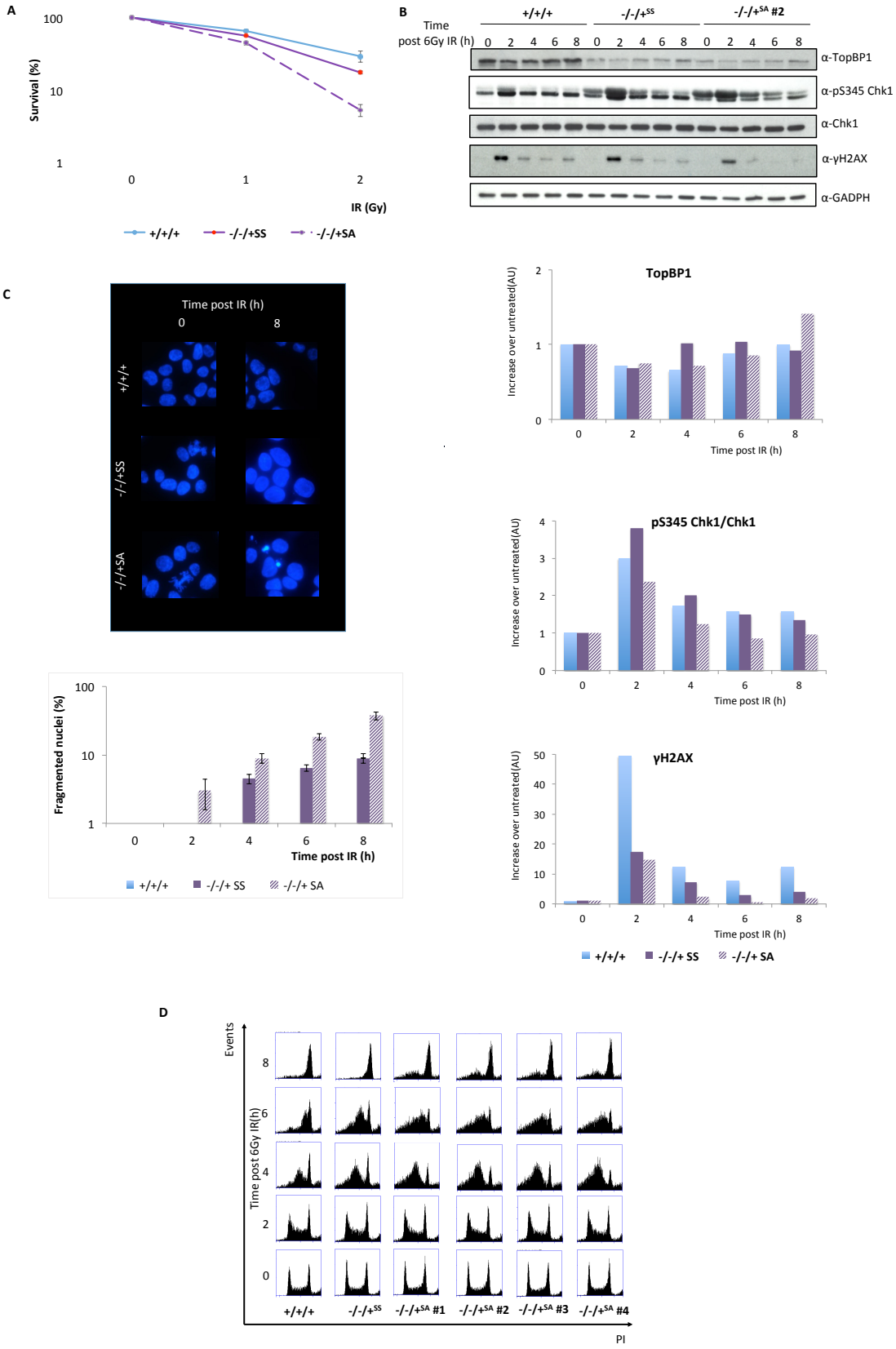

A
Survival (%)
100
10
1
0
1
2
IR (Gy)
+/+/+
-/-/+SS
-/-/+SA
B
Time post 6Gy IR (h)
+/+/+
-/-/+SS
-/-/+SA #2
0 2 4 6 8 0 2 4 6 8 0 2 4 6 8
α-TopBP1
α-pS345 Chk1
α-Chk1
α-γH2AX
α-GADPH
C
Time post IR (h)
0
8
+/+/+
-/-/+SS
-/-/+SA
TopBP1
Increase over untreated(AU)
Time post IR (h)
pS345 Chk1/Chk1
Increase over untreated(AU)
Time post IR (h)
γH2AX
Increase over untreated(AU)
Time post IR (h)
Fragmented nuclei (%)
100
10
1
0 2 4 6 8
Time post IR (h)
+/+/+
-/-/+ SS
-/-/+ SA
D
Events
Time post 6Gy IR(h)
8
6
4
2
0
+/+/+
-/-/+SS
-/-/+SA #1
-/-/+SA #2
-/-/+SA #3
-/-/+SA #4
PI

Figure 7.6: S1132 TopBP1-dependent checkpoint activation in response to IR. A) Colony survival assay of the indicated cell types plated singularly in semi-solid media following exposure to the indicated doses of IR. Error bars represent the standard deviation of the mean for three independent experiments. **B, C, D**) The indicated strains were irradiated with 6Gy of IR and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **B**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. Charts at the bottom represent quantification of the Western blots using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **C**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **D**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

7.1.5: Inactivation of TopBP1 S1132 leads to a milder defect in checkpoint activation in response to DNA damage caused by low doses of ionizing radiation and does not induce cell death.

To examine whether lower doses of ionizing radiation could alleviate the phenotypes of the S1132A mutants, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells were exposed to 3Gy of IR and followed through a period of 6h recovery time as before. Western blot analysis revealed a similar response as in Figure 7.6 B, with *TopBP1*$^{-/-/+SA}$ mutants being characterized by a reduced ability to phosphorylate the Chk1 effector kinase on S345 as well the γH2AX marker on S139 compared to the *TopBP1*$^{-/-/+SS}$ controls. In fact, 2h post-irradiation *TopBP1*$^{-/-/+SS}$ presented 3.1-fold increase of pS345 Chk1 and 25.7-fold increase of pS139 γH2AX over untreated. In contrast, the two independent *TopBP1*$^{-/-/+SA}$ mutant clones tested presented 1.6-/1.9-fold increase of pS345 Chk1 and 12.1-/15-fold increase of pS139 γH2AX over untreated at the same time point. This reduced ability of the mutant cell type to catalyze the Chk1 and γH2AX phosphorylation events was consistent across the entire 6h recovery period investigated (**Figure 7.7 A, B**).

Interestingly, microscopic analyses of DAPI-stained nuclei revealed that the low dose of IR used herein was insufficient to induce nuclear fragmentation in the *TopBP1*$^{-/-/+SA}$ cultures. So only 2.5% and 2% of the *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+SA}$ cells appeared to contain fragmented nuclei 8h post-IR. This suggests that although TopBP1 S1132A leads to a defective checkpoint response after both high (6Gy) and low (3Gy) doses of IR, it is the extent of the damage that will dictate the outcome in terms of cellular viability (**Figure 7.7 C**).

Cell cycle progression towards an IR-caused G2/M arrest was found to be slower for the *TopBP1*$^{-/-/+SA}$ cells than for the controls, as in Figure 7.6 D. At 8h following irradiation *TopBP1*$^{+/+/+}$ and the *TopBP1*$^{-/-/+SS}$ displayed 4N DNA content whereas a subset of *TopBP1*$^{-/-/+SA}$ cells still presented a DNA content corresponding to S phase (**Figure 7.7 D**).

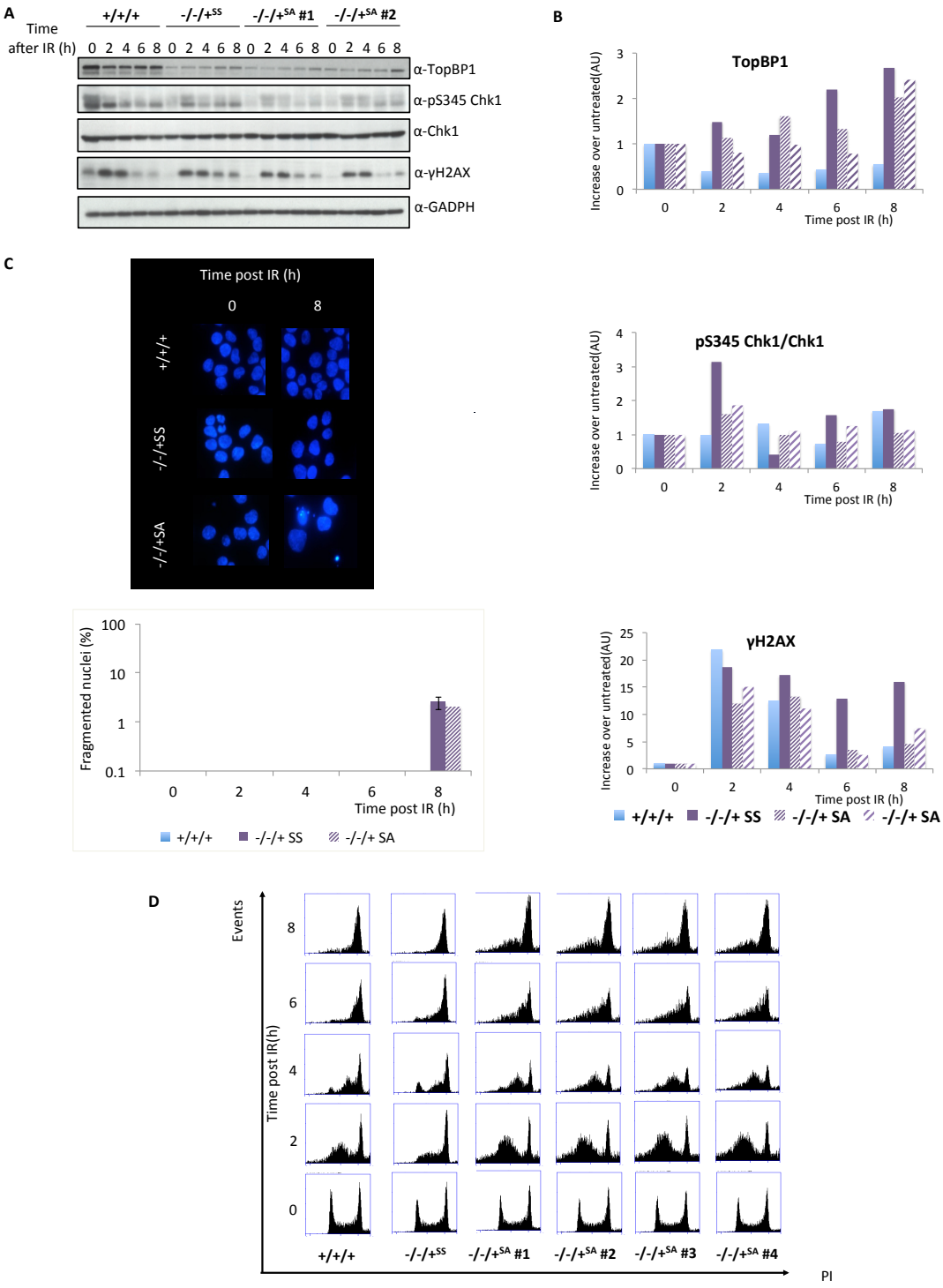


Figure 7.7: S1132 TopBP1-dependent checkpoint activation in response to low IR doses. A, B, C) The indicated strains were irradiated with 3Gy of IR and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **A**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. **B**) Quantification of the Western blots in A using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **C**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **D**) Flow cytometry analysis of samples fixed with propidium iodide.

7.1.6: The W1139 core AAD residue of TopBP1 is required for ATR pathway activation in response to replication stress caused by hydroxyurea.

To investigate the function of the W1139 core indispensable aromatic residue of the TopBP1 AAD in ATR pathway activation following replication stress, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cells were plated onto HU-containing semi-solid media at the indicated concentrations. Mutation of the W1139 residue to R was correlated with a markedly increased sensitivity to killing by HU even at concentrations as low as 20 and 40μM that do not affect the survival of the control cell lines. This suggested that the TopBP1 W1139 residue is critical for sustaining cell survival during continuous exposure to HU (**Figure 7.8 A**).

To more precisely understand the function of TopBP1 W1139 in the replication stress checkpoint response, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cells were grown in the continuous presence of 0.2mM HU for 6h and samples for Western blot, FACS and microscopy were kept at 0, 1, 2, 4 and 6h post addition of the drug. It is worth mentioning that higher concentrations of HU induced cell death within one cell cycle and the 0.2mM used herein was the optimal concentration that could both yield checkpoint activation in the control cell lines and not entirely kill the mutant cells within the time window studied. Interestingly, the mutant cells presented an inability of increasing the levels the total TopBP1. Instead the total TopBP1 protein levels were attenuated.

Immunostaining for phosphorylated Chk1 (pS345 Chk1) to detect substrates downstream of ATR activation showed an increase of pS345 Chk1 in *TopBP1*$^{+/+/+}$ (between 4.6- and 6.2-fold over untreated) and to a lesser extent *TopBP1*$^{-/-/+WW}$ (between 3.2- and 3.4-fold over untreated) controls across the entire period of 6h post HU addition. In marked contrast, *TopBP1*$^{-/-/+WR}$ mutants displayed a marginal increase of pS345 Chk1 only reaching 1.9- and 1.5-fold over untreated at 1h and 2h post HU addition, respectively. At later time points the pool of pS345 Chk1 in *TopBP1*$^{-/-/+WR}$ cells returned to levels below what is found in the untreated situation which contrasts with the increased and sustainable phosphorylation of the effector kinase observed in the control cell types. Notably, the mutant cells also seem to contain less of total Chk1 protein and their Chk1 pool is characterized by an altered electrophoretic mobility compared to the control cell lines (α-Chk1 immunostaining). Interestingly, mutant cells also presented a defective pattern of S139 γH2AX phosphorylation. More specifically, *TopBP1*$^{-/-/+WR}$ presented higher levels of pS139 (5.4-fold over untreated) compared to

their *TopBP1*$^{-/-/+WW}$ counterparts (3.4-fold over untreated) after 1h in the presence of HU. This hyper phosphorylation returned to abnormally low levels at later time points when in fact control cell types maintained pS139 at high levels (**Figure 7.8 B**). Finally, it should be noted that another independent *TopBP1*$^{-/-/+WR}$ clone was subjected to the same experiment and the same phenotypes were observed.

What is more, continuous exposure to 0.2mM HU affected the integrity of the *TopBP1*$^{-/-/+WR}$ cells. DAPI staining of fixed samples across the same time points revealed a progressive accumulation of cells with fragmented nuclei in the *TopBP1*$^{-/-/+WR}$ cultures. As early as 2h after drug addition 4.5% of the mutant cells displayed fragmented nuclei, which increased to 19.5% and 29% at 4 and 6h, respectively. Prolonged exposure to HU for 14h resulted to 69% of the *TopBP1*$^{-/-/+WR}$ cells presenting fragmented nuclei, a hallmark of apoptosis. This apoptotic phenotype was severe compared to the 6, 6.5 and 9% of fragmented nuclei observed in the *TopBP1*$^{-/-/+WW}$ controls at 4, 6 and 14h, respectively (**Figure 7.8 C**).

Analysis of the cell cycle by flow cytometry revealed a sub-G1 accumulation of a subset of the mutant cells at 4h post HU addition. Notably, by 6h the majority of the cells displayed a sub-G1 phenotype. All of the four independent *TopBP1*$^{-/-/+WR}$ clones tested displayed a similar cell cycle profile following HU treatment (**Figure 7.8 D**).

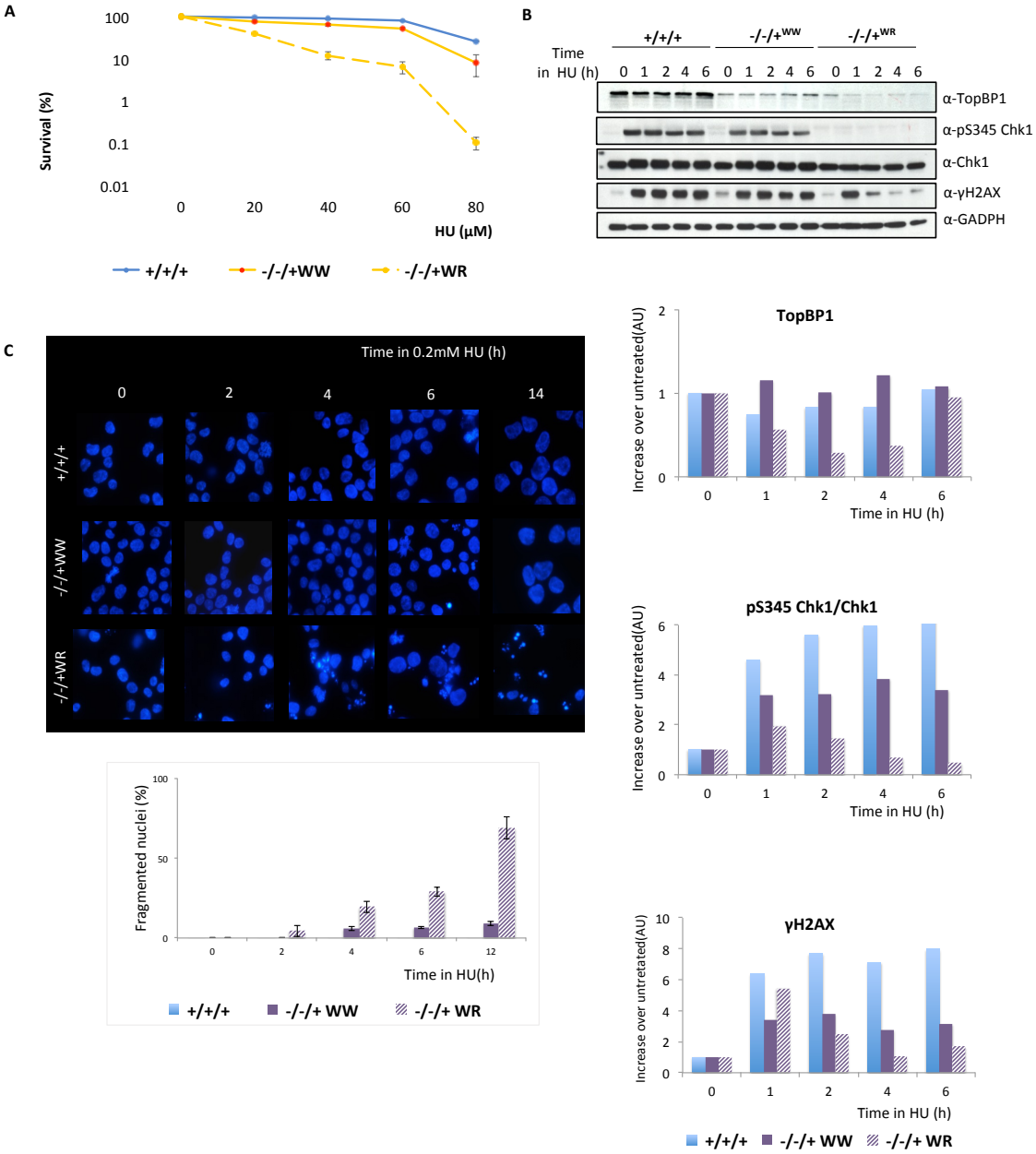

A
Survival (%)
100
10
1
0.1
0.01
0
20
40
60
80
HU (μM)
+/+/+
-/-/+WW
-/-/+WR
B
Time in HU (h)
+/+/+
-/-/+WW
-/-/+WR
0 1 2 4 6 0 1 2 4 6 0 1 2 4 6
α-TopBP1
α-pS345 Chk1
α-Chk1
α-γH2AX
α-GADPH
C
Time in 0.2mM HU (h)
0
2
4
6
14
+/+/+
-/-/+WW
-/-/+WR
Fragmented nuclei (%)
Time in HU(h)
+/+/+
-/-/+ WW
-/-/+ WR
TopBP1
Increase over untreated(AU)
Time in HU (h)
pS345 Chk1/Chk1
γH2AX
D
Events
Time in 0.2mM HU (h)
+/+/+
-/-/+WW
-/-/+WR #1
-/-/+WR #2
-/-/+WR #3
PI

Figure 7.8: W1139 TopBP1-dependent replication checkpoint activation in response to HU. A) Colony survival assay of the indicated cell types plated singularly in HU-containing semi-solid media at the indicated concentrations. Error bars represent the standard deviation of the mean for three independent experiments. **B, C, D**) The indicated strains were treated with 0.2mM HU and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **B**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. Charts at the bottom represent quantification of the Western blots using the ImageJ software. Experiments have been repeated three times and representative data is shown. **C**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **D**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

7.1.7: The W1139 core AAD residue of TopBP1 is required for ATR pathway activation and cell fate during recovery from replication stress caused by hydroxyurea.

To assess the function of the W1139 residue of TopBP1 during recovery from replication stress, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cells were arrested in G1/S by treatment with 0.125mM HU for 15h and then were washed clear of the drug and followed through a period of 6h recovery time. The longer HU block (15h herein compared to 12h block used in previous experiments) was applied due to the longer doubling time of the *TopBP1*$^{-/-/+WR}$ cells (~14.2h) compared to their *TopBP1*$^{-/-/+WW}$ counterparts (~11h). Western blot analysis at the indicated time points revealed a down regulation of the TopBP1 protein levels in *TopBP1*$^{-/-/+WR}$, which contrasts the increased TopBP1 levels observed in the *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/-/+WW}$ controls.

Furthermore, phosphorylation of the effector kinase Chk1 on S345 was attenuated in *TopBP1*$^{-/-/+WR}$ cells (pS345 Chk1 immunoblot). This attenuation of pS345 was coupled with a concomitant down regulation of the total Chk1 protein levels (Chk1 immunoblot) observed at the end of the HU block as well as post release. The percentage of pS345 Chk1 relatively to total Chk1, however, was higher in the mutant cells. In fact, the *TopBP1*$^{-/-/+WW}$ controls showed 8.2-fold increase of pS345 Chk1 over untreated by the end of the arrest, which was not too different from what was observed in the *TopBP1*$^{-/-/+WR}$ clones (10.6- and 7.3-fold). During the release from the HU block, however, the total Chk1 protein was reduced in the mutant cells that the percentage of Chk1 being phosphorylated on S345 increased to 28.6/12.6-, 18.6/30.1-, 52/47.9- and 26/8.1-fold for the *TopBP1*$^{-/-/+WW}$♯1 and *TopBP1*$^{-/-/+WW}$♯2 mutant clones, respectively, at 1, 2, 4 and 6h post release. In marked contrast, pS345 Chk1 in *TopBP1*$^{-/-/+WR}$ controls progressively returned to near basal levels during the recovery period. Another interesting observation is that the autophosphorylation of Chk1 picked up by the antibody against total Chk1 at the end of the HU block is not observed in the mutant cells. Finally, phosphorylation of pS139 γH2AX was defective in cells containing a mutated AAD. Both of the *TopBP1*$^{-/-/+WR}$ clones were unable to catalyse this phosphorylation event to the extent of the *TopBP1*$^{-/-/+WW}$ controls across all the time points examined (**Figure 7.9 A, B**).

The attenuation of the Chk1 protein levels observed in the above experiment as well as the observation that continuous exposure to 0.125mM HU for 15h leads to a substantial increase in nuclear fragmentation (Figure 7.8 C), prompted me to examine in

yet another way whether replication stress caused by HU induces apoptosis in DT40 cells mutated for the TopBP1 AAD core residue. **Figure 7.9 C** shows agarose gel electrophoresis analysis at the indicated time points of *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ blocked in 0.125mM HU and released. This experiment demonstrated the ladder pattern of DNA fragmentation in the cultures of the mutant cells. Overall, abolishment of TopBP1 W1139 leads to apoptosis following replication stress as evident from both inactivation of Chk1 and internucleosomal DNA fragmentation.

Finally, cell cycle analysis of the respective samples confirmed that the lower HU block applied here was sufficient to arrest the cells at G1/S, albeit less efficiently in the mutants. Interestingly, the *TopBP1*$^{-/-/+WR}$ were unable to recover from the HU block and gradually accumulated at the sub-G1 phase (**Figure 7.9 D**).

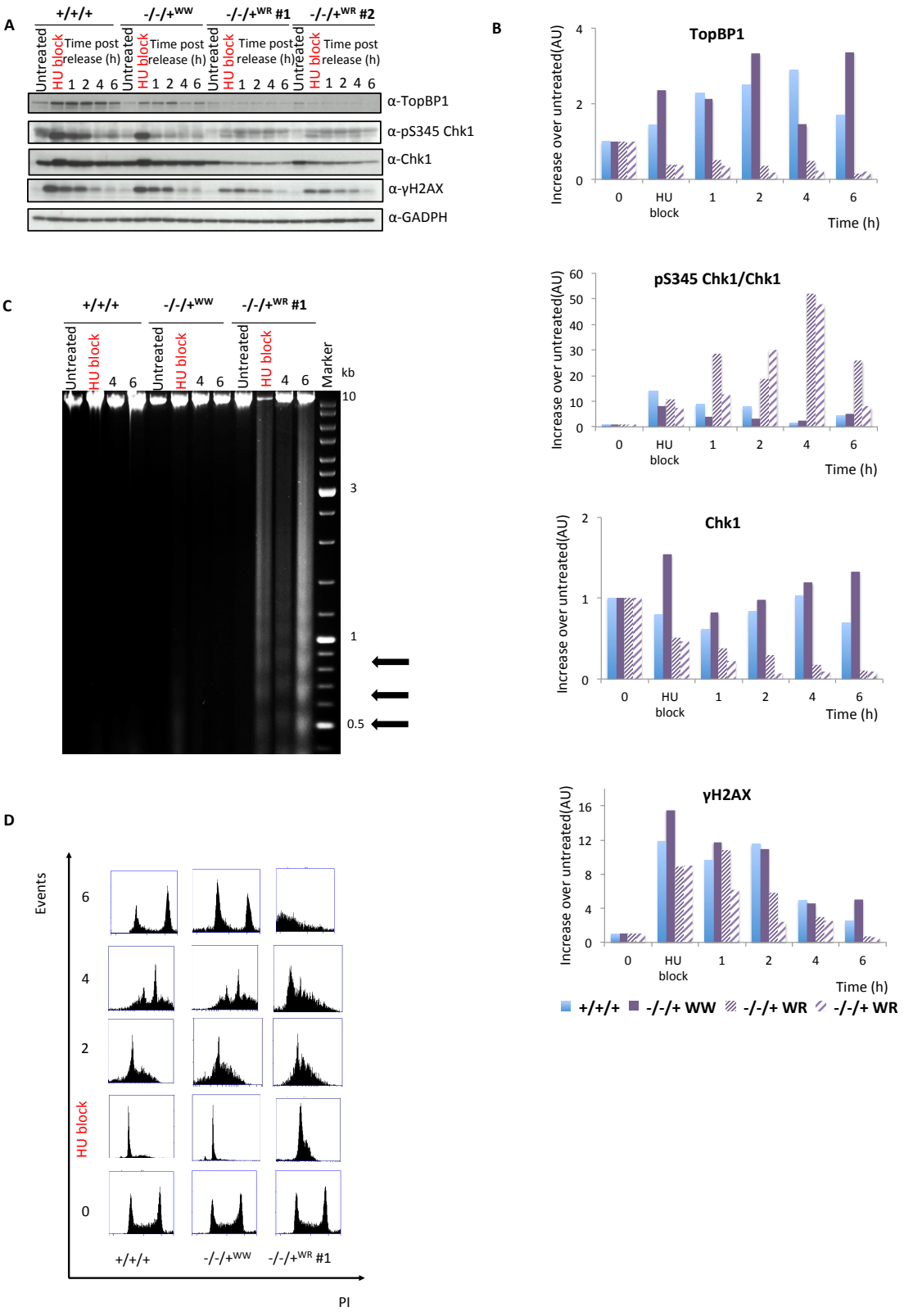

A
+/+/+
-/-/+WW
-/-/+WR #1
-/-/+WR #2
Untreated
HU block
Time post release (h)
1 2 4 6
α-TopBP1
α-pS345 Chk1
α-Chk1
α-γH2AX
α-GADPH
B
TopBP1
pS345 Chk1/Chk1
Chk1
γH2AX
Increase over untreated(AU)
HU block
Time (h)
C
Marker
kb
10
3
1
0.5
D
Events
PI
+/+/+ WW WR WR
Time (h)

Figure 7.9: W1139 TopBP1-dependent cell fate decision during recovery from replication stress. A, B, C) The indicated strains were treated with 0.125mM HU for 15h, washed clear of the drug and samples were kept at the specified times for analysis by Western blotting, FACS and genomic DNA. **A**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. **B**) Quantification of the Western blots in A using the ImageJ software. Experiments have been repeated at least twice and representative figures are shown. **C**) Internucleosomal DNA fragmentation in cells treated with 0.125mM HU for 15h and released in drug-free media. Genomic DNA was extracted and electrophoresed with ethidium bromide staining. In control cells there is lack of DNA fragmentation whereas mutants present a "ladder" of internucleosomal DNA fragments at ~200bp intervals (black arrows). Molecular weight standards are shown on the left. **D**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

7.1.8: Inactivation of TopBP1 W1139 leads to a mild defect in checkpoint activation in response to DNA damage caused by low doses of ionizing radiation

To investigate whether mutation of W1139 within TopBP1 AAD confers sensitivity to IR, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cells were plated onto semi-solid media immediately after being exposed to the indicated doses of IR. Quantitative assessment of the colony forming abilities of the respective cell lines revealed an inherent sensitivity of *TopBP1*$^{-/-/+WR}$ cells to killing my IR. This suggested that the W1139 residue of TopBP1 might also be required for checkpoint activation in response to DNA damage (**Figure 7.10 A**).

To mechanistically understand the role of W1139 in the DNA damage response, *TopBP1*$^{+/+/+}$, *TopBP1*$^{-/-/+WW}$ and *TopBP1*$^{-/-/+WR}$ cell lines were irradiated with 3Gy and followed through a period of 8h. It should be noted that higher doses of IR were not tolerated by the mutant cells. Western blot analysis showed that *TopBP1*$^{-/-/+WR}$ mutants were unable to phosphorylate S345 Chk1 to the extent of the *TopBP1*$^{-/-/+WW}$ controls. So although *TopBP1*$^{-/-/+WW}$ showed 6.8-fold increase of pS345 Chk1 2h post IR, the two mutant clones showed an attenuated response at 1.5- and 2.7-fold, respectively. This reduced ability of phosphorylating S345 was also observed at later time points. Contrary to the response observed following HU treatment, here there was no inactivation of the Chk1 protein as judged by immunostaining with an antibody against total Chk1. What is more, phosphorylation of pS139 γH2AX was severely impaired in the *TopBP1*$^{-/-/+WR}$ mutant clones as they hardly managed to increase pS139 above the basal levels when in fact the *TopBP1*$^{-/-/+WW}$ controls reached a 12.4-fold increase over untreated as early as 2h following irradiation (**Figure 7.10 B**).

DAPI staining of fixed cells and assessment of nuclear morphology showed that *TopBP1*$^{-/-/+WR}$ mutants were slightly sensitive to IR. In fact, 6h post IR 10.5% of the cells in the *TopBP1*$^{-/-/+WR}$ cultures presented fragmented nuclei and this rose to 20.5% by 8h. In contrast, only 4% of the *TopBP1*$^{-/-/+WW}$ controls presented the same nuclear morphology at 8h (**Figure 7.10 C**).

Cell cycle analysis of the same time points following irradiation with 3Gy revealed that despite their slower cell cycle, *TopBP1*$^{-/-/+WR}$ mutants reached the G2/M block with slightly faster kinetics than the *TopBP1*$^{+/+/+}$ and *TopBP1*$^{-/-/+WW}$ controls. (**Figure 7.10 D**).

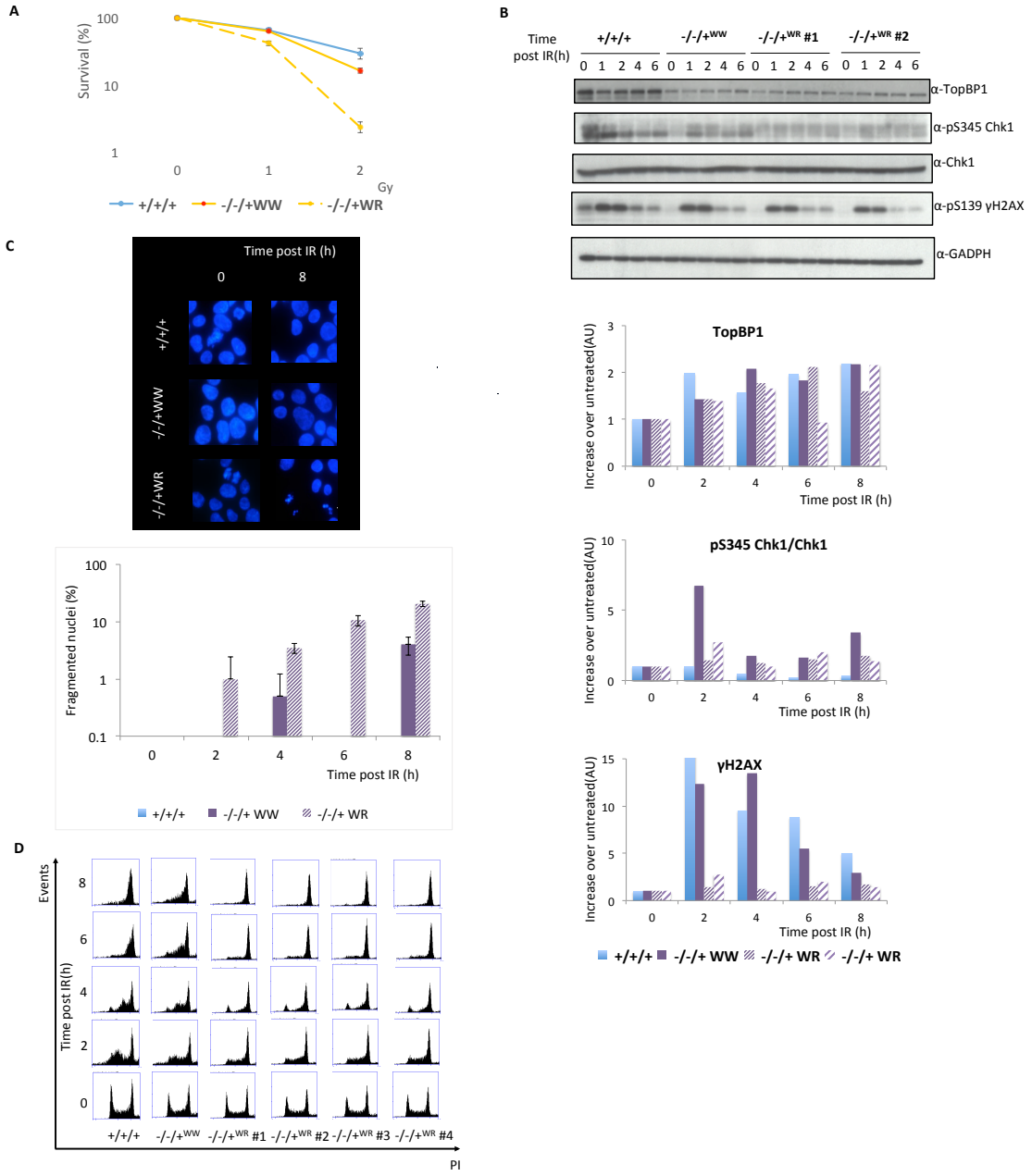


Figure 7.10: W1139 TopBP1 mutants are not sensitive to IR but still possess a defective checkpoint response. A) Colony survival assay of the indicated cell types plated singularly in semi-solid media following exposure to the indicated doses of IR. Error bars represent the standard deviation of the mean for three independent experiments. **B, C, D**) The indicated strains were irradiated with 3Gy of IR and samples were kept at the specified times for analysis by Western blotting, FACS and DAPI staining. **B**) Whole cell lysates of the indicated DT40 cell types were prepared and run on SDS-PAGE gel. Western blots were analysed using antibodies against total TopBP1, phosphor-specific S345 Chk1 antibody, total Chk1 and phosphor-specific S139 γH2AX, while GADPH served as the loading control. Charts at the bottom represent quantification of the Western blots using the ImageJ software. Experiments have been repeated three times and representative data is shown. **C**) Cells were fixed on microscopy slides and stained with DAPI to analyse nuclear morphology. 200 cells were scored in total and error bars represent the standard deviation of the mean for at least two independent experiments. **D**) Flow cytometry analysis of samples fixed with propidium iodide. A representative blot of three independent and consistent experiments is presented in this figure. Note that equivalent results have been obtained from repeat experiments.

Summary of Chapter 7

This chapter describes the use of a novel TopBP1 knock in gene targeting system to study the function of the TOPBP1 AAD in DT40. This involved the generation of the *TopBP1*$^{-/-/+SA}$ and *TopBP1*$^{-/-/+WR}$ mutant cell lines and their respective controls, *TopBP1*$^{-/-/+SS}$ and *TopBP1*$^{-/-/+WW}$. Creation of these mutant cell lines allowed the investigation of the function of the TopBP1 AAD in vertebrates, which proved to be impossible in the published mice system as inactivation of the AAD led to early embryonic lethality (Zhou *et al*, 2013). The current analysis revealed a role of the S1132 residue of TopBP1 in checkpoint activation following replication stress and DNA damage. Mutation of this residue to a non-phosphorylatable alanine led to defects in phosphorylation of the effector kinase Chk1 on S345 as well as defects in S139 γH2AX phosphorylation. Such defects were accompanied by nuclear fragmentation of a small percentage of the mutant cells in the *TopBP1*$^{-/-/+SA}$ cultures but did not affect genotoxic stress-dependent cell cycle arrest after HU or IR. On the contrary, abolishment of the core TopBP1 residue W1139 led to severe phenotypes in response to HU, including defective checkpoint signaling, Chk1 inactivation and nuclear fragmentation. TopBP1 W1139 is thus required for activation of the ATR-Chk1 pathway. Defective checkpoint signaling was also observed following irradiation of the *TopBP1*$^{-/-/+WW}$ cells but this did not affect cell fate. The work described herein provides novel *in vivo* evidence of the functions of the TopBP1AAD residues S1132 and W1139 in checkpoint activation and signaling in vertebrates.

Chapter 8

Discussion

8: Discussion

The work presented in this thesis describes the development of genetically defined systems that enable the study of the TopBP1 protein and its functions in the DT40 cells. In the following paragraphs I will briefly summarise the results presented in the preceding chapters and discuss their relevance to the literature.

Unsuccessful attempts to knock out TopBP1 in DT40 and establishment of SIOS

The aim of this project (Chapters 3 and 4) was to delete the three alleles of the *TopBP1* gene in DT40 cells, while sustaining cell viability by stably integrating a TopBP1 cDNA transgene under the control of a constitutive promoter at the *Ova* locus. With the use of homology-mediated genetic engineering I have achieved the deletion of two of the three *TopBP1* alleles (~23kb) located on chromosome II of DT40 cells and their respective replacement with selection marker cassettes. These cell lines were designated *TopBP1*$^{puro/+/+}$ and *TopBP1*$^{puro/his/+}$. Next, MerCreMer was stably integrated in the genome of these two cell lines in a non-targeted manner. Induction of MerCreMer nuclear localisation by treatment with 4-HT led to the generation of *TopBP1*$^{flox/+/+}$ and *TopBP1*$^{flox/flox/+}$ cell lines. Before attempting knocking out the third allele of *TopBP1*, the chicken TopBP1 cDNA under the control of the CMV promoter was stably integrated within the *Ova* locus. Subsequent attempts to complete deletion of the endogenous *TopBP1* copies with LARAHis and LARA2His were unsuccessful (more than a thousand cells have been screened in total). The use of the same homology arms present in LARAHis probed retargeting of the already targeted alleles. The use of LARA2His, on the other hand, revealed a potential erroneous recombination event. Finally, the other major caveat of this initial strategy was the insufficient levels of TopBP1 being produced from the OvaCMV expression construct, thus perhaps not creating the favourable environment for cells to delete their final endogenous *TopBP1* copy.

To specifically visualise the TopBP1 protein produced from the ectopic transgene and also stably integrate more copies of it within the genome (or at potentially more favourable loci), a 3xFLAG-tagged version of TopBP1 was subcloned in the OvaCMV construct and integrated within the genome of *TopBP1*$^{flox/flox/+}$ cells in a non-targeted manner. Additionally, a new gene-targeting construct (LA^{i}RA^{i}Puro) specifically tailored to help avoid retargeting was assembled. Attempts to knock out the

third *TopBP1* allele, however, were again unsuccessful. Overall, this work has established an efficient gene knock out strategy that was successfully used for the deletion of the two alleles of *TopBP1*. Additionally, it was shown that the use of the CMV promoter was inadequate for the purposes of our biological system. The next challenge was to optimise the levels of TopBP1 protein being produced from the ectopic rescue transgene. This involved a novel characterisation of different promoter regions driving expression of the *TopBP1* cDNA from the *Ova* (a locus not expressed in DT40 cells) versus a transcriptionally active locus (designated "euchromatic").

Exogenously expressing a protein of interest has proved to be a valuable tool during the study of diverse biological processes. A limiting factor for the development of artificial genetic systems is the availability of suitable promoter elements. Many constitutive promoters are used during the study of either loss of function (i.e. shRNA) or gain of function (i.e. cDNA expression) systems as well as systems designed to replace a gene of interest with a mutated copy upon deletion of the endogenous copies or to produce recombinant proteins. However, there are few examples of systematic comparative studies of the expression levels of commonly used promoter elements in specific cell types. Because different experiments have distinct requirements for the level of transgene expression, and because there is limited information available on the efficiency of different promoter systems, the choice of appropriate promoter is often ambiguous: choices are often based on technical convenience or the assumption that, since a particular promoter worked in one cell line, it will work similarly in another.

A characterization of promoter strengths has been performed by Qin et al 2010 who tested a group of six commonly used mammalian constitutive promoters for their ability to drive expression of a GFP reporter in as a panel of eight mammalian cell types. They observed that expression from the CMV promoter was the most variable, being strong in some cell types but approximately 7–8 fold lower in others. (Qin, Zhang et al. 2010). This finding was consistent with the observation from many other groups that the CMV promoter becomes silenced in some cells. A second important finding by Qin *et al.* (2010) was that the EF1A and CAG promoters were similarly highly expressed, but were more reliable across all the cell lines tested, varying by less than a factor of two (Teschendorf, Warrington et al. 2002, Brooks, Harkins et al. 2004, Meilinger, Fellinger et al. 2009, Qin, Zhang et al. 2010).

To our knowledge, no charaterisation of promoter elements has been performed in the DT40 model system. To identify an expression system suitable for optimal

exogenous protein expression in DT40 cells, the already used CMV (cytomegalovirus) promoter was compared to the CAG (CMV early enhancer and chicken beta actin) and CBA (chicken beta actin) promoter elements, in their respective ability to drive expression of the *TopBP1* transgene in DT40. The reason I chose these particular ubiquitous promoters was that they are generally described as being among the stronger constitutive promoters available in molecular biology (Powell, Rivera-Soto et al. 2015). Additionally, the CBA promoter has been successfully used for the overexpression of an *ATR* transgene integrated within the *OVA* locus and the subsequent disruption of the endogenous *ATR* locus of DT40 cells (Eykelenboom, Harte et al. 2013).

I demonstrate that CMV is a relatively weak promoter in DT40, producing approximately the same quantity of an exogenous TopBP1 reporter protein as is produced from a single copy of the native *TopBP1* gene. CBA shows several fold higher activity while CAG showed the highest level of activity in my assay, resulting in approximately 7-8 fold more protein than a single endogenous TopBP1 allele. I di observe variation between individual clones for a specific promoter construct but this did not exceed three fold. Importantly, the level of protein produced from each of the three promoters was not altered following growth for 8 days. This data suggests that integration of these promoters in DT40 does not result in silencing over this timeframe and thus these are suitable tools for protein expression studies.

The *Ova* locus is commonly used for protein expression in DT40 cells because it is transcriptionally silent in tis cell line and it shows a high targeting efficiency compared to multiple other loci studied (Buerstedde and Takeda 1991) however, it was not clear if the silent nature of the locus would affect the efficiency of the integrated promoters. Here I compared the levels of TopBP1 produced from a CAG-driven transgene at the *Ova* locus (OvaCAG) with the same transgene and promoter integrated at a euchromatic site (eCAG). Although I did observe that individual clones showed variation within each of the two loci (not exceeding three fold), I did not see any major difference between the OvaCAG and eCAG clones. From this I conclude that the *Ova* locus is an appropriate choice for integration.

Taken together, this work has established the creation of SIOS, a system useful for the overexpression of the *TopBP1* transgene or any protein of interest. Being a stable and versatile system, SIOS can indeed prove useful for various experimental purposes in the field of DT40, including protein production for biochemical studies and generation of knock out cell lines of genes of interest. SIOS is also amenable to

recombinase-mediated cassette exchange, which allows the integration of any transgene in a one-step reaction. This makes SIOS a rapid tool to examine cellular phenotypes of site-specific vertebrate mutants of a gene of interest and examine, in detail, the biological and biochemical outcomes with efficiencies equivalent to yeast genetics.

Despite the establishment of SIOS, however, subsequent attempts to create the *TopBP1* knockout using targeted or non-targeted integration of the OvaCAG expression construct as well as CRISPR technology were unsuccessful.

Characterisation of the TopBP1 mRNA and assembly of a new TopBP1 cDNA capable of sustaining viability of TopBP1 knock out DT40 cells.

The inability to obtain the *TopBP1* knockout in DT40 cells raised the question of whether the amino acid sequence of the TopBP1 protein available from the Genome Browser was erroneous or incomplete. The protein sequence of avian TopBP1 on the genome browser (XP_015137236.1) is classified as "uncharacterised" and "predicted by automated computational analysis", which suggests there has not been an *in vivo* characterisation of the *gallus gallus* TopBP1 protein or its RNA, at least not published. Instead, the sequence was generated by a prediction program, Gnomon. The main problem with computational annotation is that it relies on sequence similarities of ORFs, with putatively functional short ORFs (of less than 80 amino acids) being assigned lower quantitative conservation scores than protein-coding long ORFs. And because the human genome contains millions of small ORFs, cut-offs for 50-100bp-long ORFs are used by annotation programs, discarding in this way any small ORFs for which no experimental evidence exists (Couso and Patraquim 2017).

Protein sequence-alignment analysis of the avian TopBP1 with the TopBP1 homologues of other species did not reveal any obvious misalignment that could help identify if the "predicted" primary sequence on the database was incomplete or incorrect. The fact, however, that the consensus *gallus gallus* TopBP1 protein sequence started with an arginine instead of a methionine combined with the misalignment of the first ten amino acids with those of the TopBP1 homologues, led me to hypothesise that perhaps the 5'end of the protein was incorrect or incomplete in some way. Before embarking on characterising the TopBP1 mRNA, I decided to use the human cDNA transgene as the rescue construct in my knock out system. Overexpression of the human TopBP1 cDNA using the OvaCAG SIOS system allowed deletion of the final copy of $TopBP1^{flox/flox/+}/OVA^{CAGhTopBP1/+/+}$ cells, thus strengthening my hypothesis that the primary sequence of the available chicken cDNA was not providing the function(s) required to maintain viability of DT40 cells.

To precisely address my hypothesis, however, I isolated the total RNA of wild-type DT40 cells and performed 5' RACE with primers specific to the TopBP1 cDNA. Analysis of the sequencing data of the 5' RACE products revealed some interesting observations concerning the primary sequence and splicing pattern of the TopBP1 mRNA in DT40 cells. First and foremost, it was reassuring that the sequencing data confirmed the nucleotide and amino acid sequence of the first 908 amino acids of the

TopBP1 protein, with amino acid 1 being defined as the database annotated M (ATG). Interestingly, however, the region upstream of the consensus M, as this was identified by the 5' RACE, presented some interesting features. Firstly, the 5' RACE analysis defined a region within the database exon 1 of the *TopBP1* locus as an intron. A smaller subset of the sequenced RACE products contained the entire region annotated as exon 1 on the genome database, suggesting that the region upstream of the consensus start site is subjected to alternative splicing or some other processing activity that results to the production of two versions of TopBP1 mRNAs; one with a shorter and one with a longer 5' end. Furthermore, a small region encompassing the end of the database intron 1 was sequenced in all the 5' RACE products, suggesting a different intron/exon boundary between the first intron and the downstream exon than the consensus. Unfortunately, limited sequence information precluded the identification of this boundary. Finally, the current analysis identified a new open reading frame (ORF) within this 5' end sequence, which -according to the database- spans across the end of intron 1 and the first 90bp (or 87bp) of exon 1. This ORF^{short} was present in ~74% while a smaller subset of those (~10%) containing the $ORF^{shortprocessed}$, which was identical to ORF^{short} but lacked "GTTAAAGG" upstream of the annotated M. In addition, ~24% of the RACE products contained a longer ORF (ORF^{long}), which corresponds to the OR^{short} region but also contains all downstream sequences until the annotated ATG. Finally, an $ORF^{short'}$ being composed of the first 20 aa of the ORF^{short} and an $ORF^{short''}$ resembling $ORF^{short'}$ were identified in two of the sequenced clones.

Assembly of these newly identified ORFs at the 5' end of the consensus cDNA and the OvaCAG SIOS, stable integration within the *Ova* locus and subsequent attempt to knock out the last endogenous copy of *TopBP1* showed that the ORF^{short} and $ORF^{shortprocessed}$ sequences provided the function(s) necessary for the viability of DT40 cell devoid of endogenous TopBP1. Insufficient number of clones obtained from $TopBP1^{flox/flox/+}/OVA^{CAGnewTopBP1ORFlong/+/+}$ knockout targeting does not allow us to draw conclusions about the functionality of the ORF^{long}.

Closer examination of this newly identified domain reveals that it is a sequence rich in arginine residues, the most positive among the 20 amino acids. In fact, this peptide is 23.9% R-rich, 15.2% G-rich, 15.2% V-rich and 13% A-rich (**Figure 8.1 A**). A very similar amino acid composition is also observed in the region upstream of the ATG of the human homologue. Structure prediction of this sequence using the Phyre[2] software reveals a primarily alpha-helical secondary structure with small intervening

disordered sequences (**Figure 8.1 B**). One possibility is that this ORF provides a missing function(s) to the TopBP1 protein itself or alternatively complements a function that was partially missing from the 5' end of the protein. Indeed, one of the most common functions of arginine-rich cationic peptides is the stabilisation of macromolecular structures through the establishment of appropriate interactions (e.g. salt bridges and hydrogen bonds) (Chandana and Venkatesh 2016). Glycine-arginine-rich (GAR) domains are also frequent targets for arginine methylation, a modification that occurs frequently on nucleic acid binding proteins containing a GAR motif (Boisvert, Chenard et al. 2005). In fact, GAR domains are particularly common in proteins involved in RNA processing and trafficking (Godin and Varani 2007) therefore it could be related to the known role of TopBP1 as a transcriptional regulator. Interestingly, Mre11 contains a C-terminal GAR domain subjected to methylation. This methylation regulates the exonuclease activity of Mre11 as well as its association with sites of DNA damage, thus playing a significant role in the outcome of the checkpoint response (Boisvert, Dery et al. 2005, Boisvert, Hendzel et al. 2005).

The other possibility is that this newly identified ORF is important for post-transcriptional control of the TopBP1 RNA. This region may adopt some secondary structure(s) that ultimately dictates the function and fate of the TopBP1 RNA. Using the RNAstructure Web Server available from the Mathews lab, several potential structures that this region can adopt have been computed (**Appendix 3**). Interestingly, this region contains several arrays of guanine bases separated by one or more bases, with potential to form G-quadruplex structures (putative quadruplex sequence, PQS). In fact, bioinformatic analysis has revealed that RNA PQSs are highly enriched within the 5' UTR of human genes compared to the entire transcriptome (Huppert, Bugaut et al. 2008). Increasing evidence also suggests the PQSs are involved in the regulation of translation (Bugaut and Balasubramanian 2012).

Another interesting feature of this domain within $ORF^{shortprocessed}$ is that its methionine is not in frame with the consensus methionine of TopBP1 such that translation seems to terminate at the start of the annotated TopBP1 transcript. Given that Western blot analysis with an anti-TopBP1 antibody shows no transgene overexpression, it is probable that the reading frame of the protein has been changed. Thus it seems that there is some kind of translation initiation regulation within this region, given that 64% of the RACE clones contained ORF^{short} and 10% contained $ORF^{shortprocessed}$.

One mechanism for the regulation of translation as well as gene expression involves a combination of tRNA abundance with codon usage bias. In general, codons recognized by abundant tRNAs are more efficiently translated (Stoecklin and Diederichs 2014). Thus a potentially lower abundance of tRNA anti-R in DT40 cells combined with the enrichment of R residues within the new ORF region could serve as a way of controlling the rate of ribosome scanning or indeed its association with the TopBP1 RNA. Besides, it is known that tissue-specific expression of tRNA species is a general mechanism of regulating translation in vertebrates (Dittmar, Goodenbour et al. 2006).

What is more, there is evidence that codons other than AUG can be used for translation initiation. In fact, more than 21 alternative translation initiation codons (TICs) have been described in mammalian cells (Touriol, Bornes et al. 2003, Tikole and Sankararamakrishnan 2006). Therefore the newly identified ORF is not necessarily initiating at the ATG as it has been described so far. An interesting alternative is reading this ORF in a +1 frame (relatively to the one presented) and in this way a different ORF is created starting from a valine V (GUC) residue (**Figure 8.1 C**). This region indeed resembles an efficient consensus GTG(AG)CCGTCG (Kozak consensus is [GCC(**AG**)CC**AUGG**] (Kozak 1987)). This frame also creates two stop codons within the region corresponding to "new exon", which can again serve a functional role for the ribosome to stall/pause or fall off during scanning of the 5' end. Overall, it is possible that a combination of all the aforementioned mechanisms control the fate of the TopBP1 RNA. This is not something unheard of; translation of the different isoforms of the eIF4G (eukaryotic initiation factor 4G) results from a combination of multiple promoters, alternative splicing events and the use of an upstream non-canonical initiation codon (Coldwell, Sack et al. 2012).

But the most important finding strongly supported by the data herein, is that the newly identified ORF provides the function necessary to sustain viability of DT40 cells in the absence of the endogenous TopBP1 protein.

Taken together, I have created two TopBP1 knockout model systems useful for the study of both the human and the avian TopBP1 proteins. I have also identified the *bona fide* mRNA sequence of the *gallus gallus* TopBP1 required to sustain viability of DT40 cells, where all endogenous *TopBP1* alleles have been knocked out. This work opens up exciting possibilities concerning the functional importance of this upstream region of the TopBP1 mRNA in DT40 cells. Identifying the potential role of this region

in functionality of the TopBP1 protein itself, post-transcriptional control of the TopBP1 RNA, or indeed both would be necessary for a complete characterisation of the avian TopBP1.

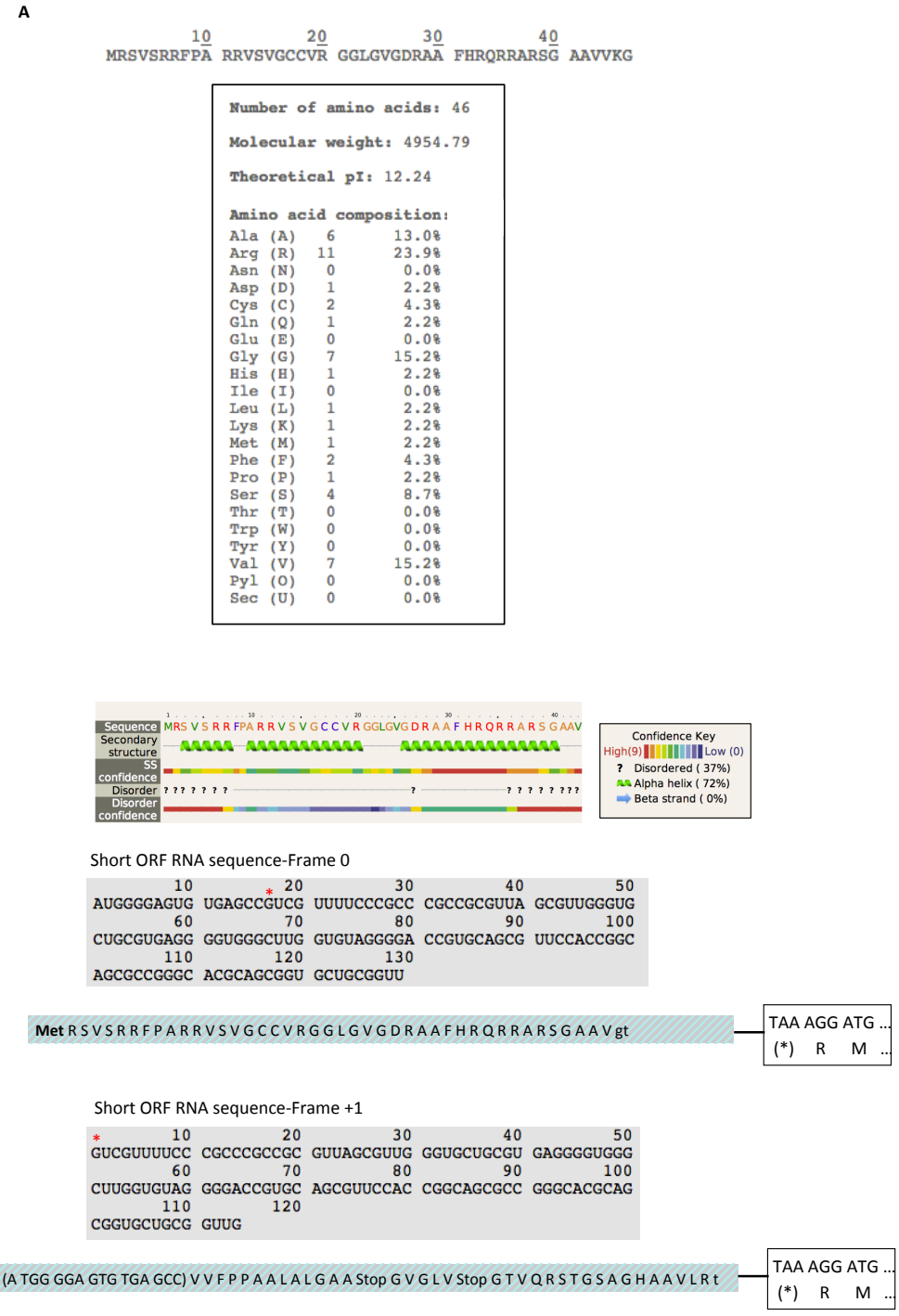


Figure 8.1: Characteristics of the novel 5'region of the TopBP1 RNA. A) Sequence and amino acid composition as calculated using the ExPasy ProtParam software. **B)** Secondary structure as predicted from the Phyre[2] software. **C)** Translation of the newly identified upstream RNA sequence of TopBP1 in two different frames. Frame 0 uses AUG as the translation initiation codon whereas Frame +1 uses a GUC downstream, which resembles more a Kozak consensus.

A novel knock in gene targeting system based on the gene dosage-dependent functions of TopBP1

To generate a knock in system for the study of TopBP1, an isogenic set of stable cell lines from the chicken B cell line DT40 was established by targeted deletion of the *TopBP1* alleles. Depending on the researchers needs, all three, two or one of the endogenous TOPBP1 alleles can be mutated in $TopBP1^{+/+/+}$, $TopBP1^{-/+/+}$ or $TopBP1^{-/-/+}$ cell lines, respectively, leading to the replacement of the total wild-type TopBP1 protein inside the cell by the mutant. Analysis of this system revealed that successive deletion of the avian *TopBP1* alleles led to a progressive reduction of the TopBP1 protein levels. This novel DT40 model system allowed the study of the kinetics of the events induced by progressive loss of function of TopBP1 in terms of checkpoint activation.

The very first intriguing observation that led to pursuing all work described in chapter 6 was the progressive reduction of TopBP1 protein levels with successive deletion of the *TopBP1* alleles. This suggested that there was no gene dosage compensation from the intact allele(s) for the loss of protein caused by *TopBP1* deletion.

Progressive gene dosage-dependent reduction of the TopBP1 levels leads to a progressively decreasing proliferation rate as well as decreased clonogenic potential of DT40. As explained in the introduction, $Dpb11^{TopBP1}$ –together with $Sld2^{RecQL4}$ and $Sld3^{Treslin/Ticrr}$– is one of the limiting factors for replication initiation. Overexpression of these factors in budding yeast is sufficient to convert late- into early-firing origins (Mantiero, Mackenzie et al. 2011). Conversely, low levels of these factors result to low levels of replication initiation (Zegerman 2015). Using a novel genetic system with varying TopBP1 levels, I provide *in vivo* evidence that the availability of TopBP1 determines the rate of replication in DT40 cells.

More importantly, gene dosage-dependent reduction of TopBP1 causes increasing sensitivity to killing by multiple DNA damaging agents. To investigate whether this sensitivity translates into genomic instability, $TopBP1^{+/+/+}$, $TopBP1^{-/+/+}$ and $TopBP1^{-/-/+}$ cells were examined for their respective checkpoint responses following replication stress or DNA damage. Analysis showed that both the replication and the DNA damage checkpoint pathways are *TopBP1* dosage- sensitive.

TopBP1 gene dosage reduction was associated with a partial defect in S345 Chk1 phosphorylation and presumably activation in response to replication stress as well as recovery from replication stress. There was also a partial defect in S139 γH2AX

phosphorylation at early stages of replicative stress. In contrast, prolonged replication stress resulted to increased pS139 γH2AX, indicative of persisting unrepaired damage. Additionally, reduction of the TopBP1 levels rendered some cells unable to recover from replication stress. These results show that TopBP1 is required for proper functioning of the ATR-Chk1 pathway in a gene dosage-dependent way. We speculate that the reduction in the protein levels of TopBP1 is associated with both a reduced activation of ATR (thus impairment of the auto-amplification loop) and a limited scaffolding function for assembly of the checkpoint apparatus, opening the gate to genetic instability. Therefore an intact ATR-Chk1 pathway is dependent on an optimal threshold of TopBP1 availability. Taken together these results suggest that activation of Chk1 in the context of prolonged replication stress is essential for suppression of DNA damage and this depends on *TopBP1* gene dosage. But why the majority of cells still manage to successfully arrest the cell cycle? This is probably due to compensation by the Chk2 kinase, which is known to enforce the S-M checkpoint in Chk1 knockout DT40 cells (Zachos, Rainey et al. 2003).

Gene dosage reduction of *TopBP1* resulted to decreased pS139 γH2AX following IR, indicative of defective ATM/ATR signaling. Interestingly, phosphorylation of H2AX is thought to be dispensable for activation of ATM substrates but necessary for recruitment of DNA repair factors (Fernandez-Capetillo, Chen et al. 2002). Furthermore, *TopBP1* gene dosage reduction caused abnormally increased levels of pS345 Chk1 following IR. We speculate that this is due to a repair defect resulting from limited TopBP1. In other words, reduced TopBP1 causes impaired ATM/ATR signaling and subsequently a defective assembly of damage repair machinery or defective repair *per se*. Besides, TopBP1 is known to be directly involved in repair via its association with the SLX4 repair scaffold as well as via its role in Rad51 loading (Gritenaite, Princz et al. 2014, Moudry, Watanabe et al. 2016). As a result, it is possible that when the available TopBP1 pool is limited, repair intermediates persist longer and lead to increased phosphorylation of the effector kinase.

A subset of *TopBP1*$^{-/-/+}$ cells was also unable to traverse from G1/S towards the G2 arrest following high doses of irradiation. This is possibly a result of unrepaired DNA damage preventing the S-M transition. Furthermore, the nocodazole-trap experiment showed that cells with reduced TopBP1 were unable both to properly hold the mitotic block in response to nocodazole treatment and to efficiently arrest at G2 when they have been previously irradiated.

Finally, depletion of TopBP1 using the miniAID-SMASh degron in RPE cells resulted in a complete inability to incorporate EdU, which is reminiscent of the essential function of TopBP1 during the initiation of DNA replication. Furthermore, TopBP1 depletion resulted to pS139 γH2AX accumulation, in otherwise untreated cells, pinpointing again to the role of TopBP1 in genome stability.

Taken together, *TopBP1* gene dosage is required for transmission of checkpoint signals through the ATR/ATM stress response kinases. The present study therefore complements the long-standing model that it is the threshold of TopBP1 that determines checkpoint signalling. The levels of TopBP1 are commonly disrupted in human cancers. These results might explain why genetic changes that alter TopBP1 levels are positively selected for during tumour evolution, since it fuels genomic instability thus tumour progression.

The establishment of this isogenic set of DT40 stable cell lines with varying copies of the *TopBP1* alleles defines a novel gene knock in platform for a structure-function analysis of Topbp1. Future work could also more precisely define the *TopBP1* dosage-dependent checkpoint defects at the molecular level.

Characterisation of the TopBP1 AAD using a knock-in point mutation strategy.

Central to the role of the TopBP1 protein within the DDR is its ability to act as a molecular scaffold for the assembly of the checkpoint apparatus as well as its ability to directly activate the stress response kinases ATM and ATR. Using a knock-in gene targeting strategy for the integration of *TopBP1* point mutants within the intact allele of *TopBP1*$^{-/-/+}$ cells, we have uncovered potential *in vivo* functions of the TopBP1 AAD both in the generation and transmission of the checkpoint signal as well as in the crosstalk between ATM and ATR activation routes. Our data corroborate previous studies, propose new possible explanations for the observed phenotypes and allow the study of TopBP1 mutants that have resulted to embryonic lethality in our mouse models, thus rendering the study of cellular phenotypes impossible (Zhou, Liu et al. 2013).

The S1132 residue within the AAD of TopBP1 has been described to have an essential role for the activation of ATR at DSBs but not stalled forks. In fact, ATM-catalysed phosphorylation of S1132 was shown to be necessary for activation of ATR in *Xenopus* egg extracts exposed to DNA damage (Yoo, Kumagai et al. 2007). This observation, however, was different to what Hashimoto et al (2006) has reported. To investigate this *in vivo*, I mutated S1132 at the chromosomal locus of *TopBP1*$^{-/-/+}$ cells, thus replacing the entire wild-type TopBP1 population with a TopBP1 S1132A mutant.

The results described in Chapter 7 propose a functional role of the TopBP1 S1132 during replication stress and DNA damage, albeit with different outcomes with respect to cell fate. S1132A mutant cells were less effective in phosphorylating S345 Chk1 and S139 γH2AX than the parental control cells. And this was observed both during continuous replicative stress and recovery as well as post-IR at all doses tested. It should be reminded, however, that the observed defects were more severe when cells were irradiated compared to HU-caused replication stress. Such observations help us speculate on the function of TopBP1 S1132. Although abolishment of TopBP1 S1132 leads to a slight impairment of checkpoint signaling following replication stress, this is not enough to impair successful cell cycle arrest. This is presumably because the abolishment of the ATM-Chk1 arm plays a comparatively minor role to the major ATR-Chk1 replication stress response mechanism. Therefore, the replication stress response mainly functions in a TopBP1 S1132-independent way to relay the checkpoint signal to the Chk1 effector kinase. And the presumably defective ATM-dependent and TopBP1

S1132-dependent phosphorylation of Chk1 is not sufficient to affect cell cycle arrest and cell fate.

In contrast, TopBP1 S1132 appears to play a more significant role during the DNA damage response, with mutant cells presenting low levels of nuclear fragmentation and defective cell cycle progression in addition to the signaling defects described earlier. We speculate that although the major IR response pathway ATM-Chk1(/Chk2) is functional, there is a TopBP1 S1132-dependent mechanism that is also important for successful G2 arrest and cell fate. One possibility is that ATR activation is required for the repair of a subset of IR-induced breaks (or indeed for cells traversing G1/S at the time of irradiation) and this activation depends on ATM-catalysed phosphorylation of TopBP1 S1132. Abolishment of this residue precludes its phosphorylation by ATM, impairs its interaction with ATR and thus overall prevents relaying of the signal from ATM to ATR. The other possibility which is more appealing but perhaps more speculative is that TopBP1 S1132 is required for activation of the ATM kinase itself.

Taken together, we propose that activation of Chk1 in response to replication stress partly requires TopBP1 S1132 but failure to do so does not impact on cell fate. On the other hand, activation of Chk1 following DNA damage is to a greater extent dependent on TopBP1 S1132 and abolishment of this mechanism sensitises a subset of cells to apoptosis before they reach the G2 arrest. Future work is required to gain insight into the underling molecular mechanisms of the observed phenotypes.

Examination of the W1139 core residue of the TopBP1 AAD reveals an essential function during replication stress but not DNA damage, consistent with previous observations. First of all, the fact that W1139R cells have a longer cell cycle compared to their parental counterparts (15h versus 11.5h) suggests that this residue of the AAD is important, although not essential, during unperturbed growth. We speculate that this is related to the housekeeping function of ATR in the stabilisation of stalled forks even in undamaged cells. More importantly, our results provide *in vivo* evidence for an essential function of TopBP1 W1139 in checkpoint activation and cell fate following replication stress, which cannot be substituted for by other potential AAD domains in my assay and under the conditions tested. Abolishment of W1139 is sufficient to abolish phosphorylation of the effector kinase Chk1 on S345 in response to HU. Additionally, the altered mobility shift of the Chk1 protein itself in W1139R cells is presumably indicative of an inability to phosphorylate Chk1 on residues other than

S345 as well. Another possible explanation is that Chk1 becomes inactivated, as cells become apoptotic. In fact it has been documented that apoptosis is characterised by degradation of Claspin and subsequent inactivation of the Chk1 effector (Semple, Smits et al. 2007). This Chk1 phenotype correlates with the observation that prolonged exposure to replication stress –even at extremely low levels- renders W1139R cells unable to recover and induces nuclear fragmentation, a hallmark of apoptotic cell death. This apoptotic phenotype observed in the mutants is accompanied by an increased phosphorylation of pS345, presumably a result of severe damage remaining unrepaired.

Despite the severely impaired phosphorylation of S345 Chk1 (and S139 γH2AX) observed in W1139R cells that have been previously irradiated, cell fate is not affected. In fact, W1139R cells can successfully halt their cell cycle in response to IR, with very limited nuclear fragmentation observed. This suggests, that the ATM-Chk1(/Chk2) pathway functions –at least to a large extent- in a TopBP1 W1139-independent manner during the DNA damage checkpoint response.

Taken together, this data suggests that there are two mechanisms of checkpoint activation in this context. One mechanism operates during DNA damage and relays the signal to downstream effectors leading to successful cell cycle arrest. This is presumably an ATM-mediated pathway, which results to low levels of pS345 Chk1 and is not dependent on TopBP1 W1139, hence on activation of ATR. The other mechanism is the well-established mechanism of ATR activation by TopBP1, which is thought to provide a way of signal auto-amplification and full activation of the ATR pathway. This mechanism is W1139-dependent.

The fact that W1139R mutants completely fail to respond to replication stress and immediately become apoptotic suggests that during replication stress both of the aforementioned mechanisms do not operate. I speculate that this is due to the impaired activation of ATR by the mutant TopBP1 (W1139R) during HU leading to increased levels of ssDNA being generated. Consequently, coating of the extensive tracts of ssDNA with RPA leads to RPA exhaustion and replication catastrophy. In support of this hypothesis, the phenotypes observed after HU treatment of the W1139R mutants, in particular the accumulation of fragmented nuclei with sub-G1 content, resemble the phenotypes observed after loss of ATR kinase activity in human cell lines (Toledo *et al* 2013).

Another possibility is that the phenotypes may be due to TopBP1 W1139-dependent ATR activation being required for activation of ATM at broken forks. This

reveals a possible crosstalk route between the two kinases. So similarly to how TopBP1 S1132 is required for activation of ATR after DNA damage, it is possible that W1139 is required for activation of ATM during replication stress. In support of this idea, Chk1 is not essential in DT40 cells and in the absence of Chk1 compensatory mechanisms take over to elicit a checkpoint response (Zachos et al 2003). Here, we observe both an inability to phosphorylate Chk1 (which is dependent on TopBP1 W1139) and a complete absence of some compensatory mechanism that would act to at least prevent apoptosis.

Another interesting observation from the data presented herein is that no other protein factor seems to compensate for the mutated TopBP1 AAD domain following HU treatment. The recently characterised ATR activator ETAA1 would be an ideal candidate to rescue the apoptotic phenotype of W1139R cells through its ability to bind RPA and propagate ATR activation along stretches of ssDNA independently of the 9-1-1 complex (Bass et al 2016). But this is not the case in the system presented herein. This could be due to the AAD function of ETAA1 being in some way dependent on the AAD function of TopBP1, in DT40 at least. Alternatively, the W1139R mutation of TopBP1 might lock the protein on the lesion precluding the physical access to other AAD-containing proteins like ETAA1. In other words, the W1139R TopBP1 gets localised to the lesion as would the wild-type protein but no activation of ATR above the basal levels occurs.

Future work is necessary to gain insight into the mechanism of operation of the W1138 core AAD residue. For now, our data suggest that the ATR pathway determines cell fate in response to replication stress and this is dependent on TopBP1 W1138, in a way that is reminiscent of ATR deletion (Brown and Baltimore 2000, de Klein, Muijtjens et al. 2000). Therefore the TopBP1 W1139 residue performs a function that becomes essential during replication stress and possibly involves both an ability to activate the ATR kinase and the ability to signal to ATM.

Epilogue

The current work made use of novel genetically defined model systems to study the avian TopBP1 protein and its functions (**Figure 8.2**). Future use of these systems and the information obtained from this thesis will help gain insight into the characteristics of TopBP1 at the protein and RNA levels as well as its various roles of TopBP1 in the DNA Damage Response.

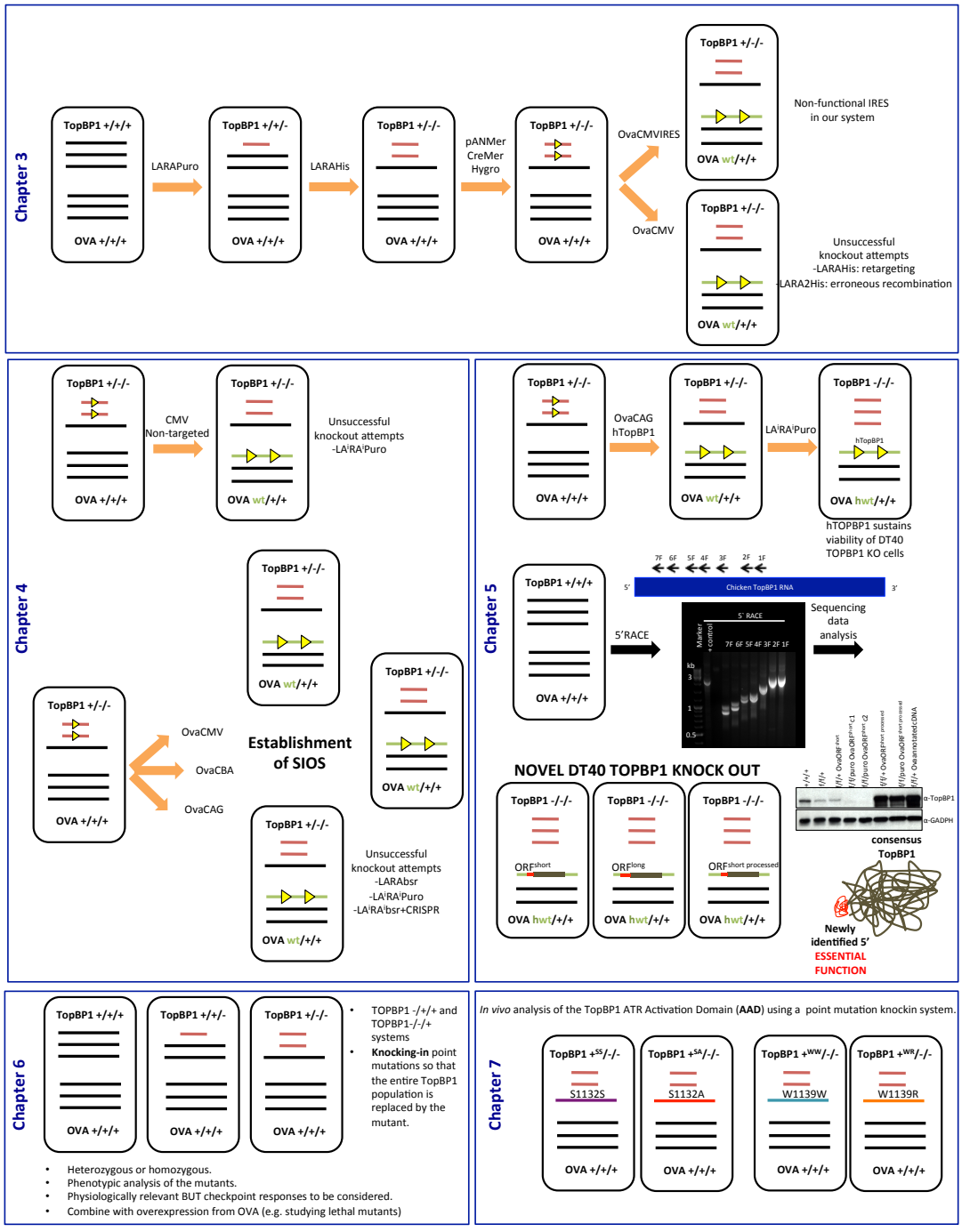


Figure 8.2: Summary of the work and approaches carried out in the current thesis.

There is still a lot to learn. The puzzle is far from being complete. The innate complexity of biological systems requires persistent research and faith towards a unifying understanding of living matter.

BIBLIOGRAPHY

Acevedo, J., et al. (2016). "Direct Binding to Replication Protein A (RPA)-coated Single-stranded DNA Allows Recruitment of the ATR Activator TopBP1 to Sites of DNA Damage." J Biol Chem **291**(25): 13124-13131.

Ahn, J. Y., et al. (2002). "Phosphorylation of threonine 68 promotes oligomerization and autophosphorylation of the Chk2 protein kinase via the forkhead-associated domain." J Biol Chem **277**(22): 19389-19395.

Alberts, B. (2002). Molecular biology of the cell. New York, Garland Science.

Alcasabas, A. A., et al. (2001). "Mrc1 transduces signals of DNA replication stress to activate Rad53." Nat Cell Biol **3**(11): 958-965.

Ali, A., et al. (2004). "Requirement of protein phosphatase 5 in DNA-damage-induced ATM activation." Genes Dev **18**(3): 249-254.

Araki, H. (2011). "Initiation of chromosomal DNA replication in eukaryotic cells; contribution of yeast genetics to the elucidation." Genes Genet Syst **86**(3): 141-149.

Araki, H., et al. (1995). "Dpb11, which interacts with DNA polymerase II(epsilon) in Saccharomyces cerevisiae, has a dual role in S-phase progression and at a cell cycle checkpoint." Proc Natl Acad Sci U S A **92**(25): 11791-11795.

Arakawa, H., et al. (2008). "Protein evolution by hypermutation and selection in the B cell line DT40." Nucleic Acids Res **36**(1): e1.

Arias, E. E. and J. C. Walter (2007). "Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells." Genes Dev **21**(5): 497-518.

Awasthi, P., et al. (2015). "ATM and ATR signaling at a glance." J Cell Sci **128**(23): 4255-4262.

Bakkenist, C. J. and M. B. Kastan (2003). "DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation." Nature **421**(6922): 499-506.

Balestrini, A., et al. (2010). "GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication." Nat Cell Biol **12**(5): 484-491.

Ball, H. L., et al. (2007). "Function of a conserved checkpoint recruitment domain in ATRIP proteins." Mol Cell Biol **27**(9): 3367-3377.

Balmain, A. (2001). "Cancer genetics: from Boveri and Mendel to microarrays." Nat Rev Cancer **1**(1): 77-82.

Bang, S. W., et al. (2013). "Oligomerization of TopBP1 is necessary for the localization of TopBP1 to mitotic centrosomes." Biochem Biophys Res Commun **436**(1): 31-34.

Bantele, S. C., et al. (2017). "Targeting of the Fun30 nucleosome remodeller by the Dpb11 scaffold facilitates cell cycle-regulated DNA end resection." Elife **6**.

Bao, S., et al. (2001). "ATR/ATM-mediated phosphorylation of human Rad17 is required for genotoxic stress responses." Nature **411**(6840): 969-974.

Bass, T. E., et al. (2016). "ETAA1 acts at stalled replication forks to maintain genome integrity." Nat Cell Biol **18**(11): 1185-1195.

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

Bentley, N. J., et al. (1996). "The Schizosaccharomyces pombe rad3 checkpoint gene." EMBO J **15**(23): 6641-6651.

Bermejo, R., et al. (2011). "The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores." Cell **146**(2): 233-246.

Bermudez, V. P., et al. (2003). "Loading of the human 9-1-1 checkpoint complex onto DNA by the checkpoint clamp loader hRad17-replication factor C complex in vitro." Proc Natl Acad Sci U S A **100**(4): 1633-1638.

Beucher, A., et al. (2009). "ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2." EMBO J **28**(21): 3413-3427.

Bhatti, S., et al. (2011). "ATM protein kinase: the linchpin of cellular defenses to stress." Cell Mol Life Sci **68**(18): 2977-3006.

Blackford, A. N., et al. (2015). "TopBP1 interacts with BLM to maintain genome stability but is dispensable for preventing BLM degradation." Mol Cell **57**(6): 1133-1141.

Blier, P. R., et al. (1993). "Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks." J Biol Chem **268**(10): 7594-7601.

Blow, J. J. (1993). "Preventing re-replication of DNA in a single cell cycle: evidence for a replication licensing factor." J Cell Biol **122**(5): 993-1002.

Blow, J. J. and R. A. Laskey (1986). "Initiation of DNA replication in nuclei and purified DNA by a cell-free extract of Xenopus eggs." Cell **47**(4): 577-587.

Bochkarev, A., et al. (1997). "Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA." Nature **385**(6612): 176-181.

Boisvert, F. M., et al. (2005). "Protein interfaces in signaling regulated by arginine methylation." Sci STKE **2005**(271): re2.

Boisvert, F. M., et al. (2005). "Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control." Genes Dev **19**(6): 671-676.
Boisvert, F. M., et al. (2005). "Methylation of MRE11 regulates its nuclear compartmentalization." Cell Cycle **4**(7): 981-989.

Bonilla, C. Y., et al. (2008). "Colocalization of sensors is sufficient to activate the DNA damage checkpoint in the absence of damage." Mol Cell **30**(3): 267-276.

Boos, D., et al. (2011). "Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans." Curr Biol **21**(13): 1152-1157.

Bosotti, R., et al. (2000). "FAT: a novel domain in PIK-related kinases." Trends Biochem Sci **25**(5): 225-227.

Botuyan, M. V., et al. (2006). "Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair." Cell **127**(7): 1361-1373.

Boveri T. Über Mehrpolige Mitosen als Mittel zur Analyse des Zellkerns. Verhandlungen der Physikalische-medizinischen Gesellschaft zu Würzburg. 1902;35:67–90.

Broderick, R., et al. (2015). "TOPBP1 recruits TOP2A to ultra-fine anaphase bridges to aid in their resolution." Nat Commun **6**: 6572.

Brown, E. J. and D. Baltimore (2000). "ATR disruption leads to chromosomal fragmentation and early embryonic lethality." Genes Dev **14**(4): 397-402.

Brown, E. J. and D. Baltimore (2003). "Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance." Genes Dev **17**(5): 615-628.

Brooks, A. R., et al. (2004). "Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle." J Gene Med **6**(4): 395-404.

Bryant, J. A. and S. J. Aves (2011). "Initiation of DNA replication: functional and evolutionary aspects." Ann Bot **107**(7): 1119-1126.

Brooks, A. R., et al. (2004). "Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle." J Gene Med **6**(4): 395-404.

Bueno, A., et al. (1991). "A fission yeast B-type cyclin functioning early in the cell cycle." Cell **66**(1): 149-159.

Buerstedde, J. M. and S. Takeda (1991). "Increased ratio of targeted to random integration after transfection of chicken B cell lines." Cell **67**(1): 179-188.

Bugaut, A. and S. Balasubramanian (2012). "5'-UTR RNA G-quadruplexes: translation regulation and targeting." Nucleic Acids Res **40**(11): 4727-4741.

Burma, S., et al. (2001). "ATM phosphorylates histone H2AX in response to DNA double-strand breaks." J Biol Chem **276**(45): 42462-42467.

Calonge, T. M. and M. J. O'Connell (2008). "Turning off the G2 DNA damage checkpoint." DNA Repair (Amst) **7**(2): 136-140.

Canman, C. E., et al. (1998). "Activation of the ATM kinase by ionizing radiation and phosphorylation of p53." Science **281**(5383): 1677-1679.

Carr, A. M. (1996). "Checkpoints take the next step." Science **271**(5247): 314-315.

Carr, A. M. (2000). "Cell cycle. Piecing together the p53 puzzle." Science **287**(5459): 1765-1766.

Carr, A. M. (2002). "DNA structure dependent checkpoints as regulators of DNA repair." DNA Repair (Amst) **1**(12): 983-994.

Caspari, T. and A. M. Carr (1999). "DNA structure checkpoint pathways in Schizosaccharomyces pombe." Biochimie **81**(1-2): 173-181.

Caspari, T., et al. (2000). "Characterization of Schizosaccharomyces pombe Hus1: a PCNA-related protein that associates with Rad1 and Rad9." Mol Cell Biol **20**(4): 1254-1262.

Casper, A. M., et al. (2002). "ATR regulates fragile site stability." Cell **111**(6): 779-789.

Cescutti, R., et al. (2010). "TopBP1 functions with 53BP1 in the G1 DNA damage checkpoint." EMBO J **29**(21): 3723-3732.

Chandana, T. and Y. P. Venkatesh (2016). "Occurrence, Functions and Biological Significance of Arginine-Rich Proteins." Curr Protein Pept Sci **17**(5): 507-516.

Chehab, N. H., et al. (2000). "Chk2/hCds1 functions as a DNA damage checkpoint in G(1) by stabilizing p53." Genes Dev **14**(3): 278-288.

Choi, J. H., et al. (2010). "Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling." Proc Natl Acad Sci U S A **107**(31): 13660-13665.

Choi, J. H., et al. (2009). "Cooperative activation of the ATR checkpoint kinase by TopBP1 and damaged DNA." Nucleic Acids Res **37**(5): 1501-1509.

Choi, S. H. and H. Y. Yoo (2016). "Mdc1 modulates the interaction between TopBP1 and the MRN complex during DNA damage checkpoint responses." Biochem Biophys Res Commun **479**(1): 5-11.

Chowdhury, A., et al. (2010). "The DNA unwinding element binding protein DUE-B interacts with Cdc45 in preinitiation complex formation." Mol Cell Biol **30**(6): 1495-1507.

Chowdhury, P., et al. (2014). "Targeting TopBP1 at a convergent point of multiple oncogenic pathways for cancer therapy." Nat Commun **5**: 5476.

Chuang, L. C., et al. (2009). "Phosphorylation of Mcm2 by Cdc7 promotes pre-replication complex assembly during cell-cycle re-entry." Mol Cell **35**(2): 206-216.

Chung, H. K., et al. (2015). "Tunable and reversible drug control of protein production via a self-excising degron." Nat Chem Biol **11**(9): 713-720.

Ciccia, A. and S. J. Elledge (2010). "The DNA damage response: making it safe to play with knives." Mol Cell **40**(2): 179-204.

Cimprich, K. A. and D. Cortez (2008). "ATR: an essential regulator of genome integrity." Nat Rev Mol Cell Biol **9**(8): 616-627.

Coldwell, M. J., et al. (2012). "Multiple isoforms of the translation initiation factor eIF4GII are generated via use of alternative promoters, splice sites and a non-canonical initiation codon." Biochem J **448**(1): 1-11.

Coleman, T. R., et al. (1996). "The Xenopus Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts." Cell **87**(1): 53-63.

Cortez, D., et al. (2004). "Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases." Proc Natl Acad Sci U S A **101**(27): 10078-10083.

Cortez, D., et al. (2001). "ATR and ATRIP: partners in checkpoint signaling." Science **294**(5547): 1713-1716.

Cortez, D., et al. (1999). "Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks." Science **286**(5442): 1162-1166.

Cotta-Ramusino, C., et al. (2011). "A DNA damage response screen identifies RHINO, a 9-1-1 and TopBP1 interacting protein required for ATR signaling." Science **332**(6035): 1313-1317.

Dalton, S. (2015). "Linking the Cell Cycle to Cell Fate Decisions." Trends Cell Biol **25**(10): 592-600.

Dart, D. A., et al. (2004). "Recruitment of the cell cycle checkpoint kinase ATR to chromatin during S-phase." J Biol Chem **279**(16): 16433-16440.

de Klein, A., et al. (2000). "Targeted disruption of the cell-cycle checkpoint gene ATR leads to early embryonic lethality in mice." Curr Biol **10**(8): 479-482.

Deegan, T. D. and J. F. Diffley (2016). "MCM: one ring to rule them all." Curr Opin Struct Biol **37**: 145-151.

Deegan, T. D., et al. (2016). "Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation." EMBO J **35**(9): 961-973.

Delacroix, S., et al. (2007). "The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1." Genes Dev **21**(12): 1472-1477.

Derheimer, F. A. and M. B. Kastan (2010). "Multiple roles of ATM in monitoring and maintaining DNA integrity." FEBS Lett **584**(17): 3675-3681.

Dibitetto, D., et al. (2016). "Slx4 and Rtt107 control checkpoint signalling and DNA resection at double-strand breaks." Nucleic Acids Res **44**(2): 669-682.

Dimitrova, D. S. and D. M. Gilbert (2000). "Temporally coordinated assembly and disassembly of replication factories in the absence of DNA synthesis." Nat Cell Biol **2**(10): 686-694.

Dittmar, K. A., et al. (2006). "Tissue-specific differences in human transfer RNA expression." PLoS Genet **2**(12): e221.

Doi, A., et al. (2006). "A combined pathway to simulate CDK-dependent phosphorylation and ARF-dependent stabilization for p53 transcriptional activity." Genome Inform **17**(1): 112-123.

Doil, C., et al. (2009). "RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins." Cell **136**(3): 435-446.

Donnelly, M. L., et al. (2001). "Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'." J Gen Virol **82**(Pt 5): 1013-1025.

Duursma, A. M., et al. (2013). "A role for the MRN complex in ATR activation via TOPBP1 recruitment." Mol Cell **50**(1): 116-122.

Edwards, R. J., et al. (1999). "A Rad3-Rad26 complex responds to DNA damage independently of other checkpoint proteins." Nat Cell Biol **1**(7): 393-398.

Ellison, V. and B. Stillman (2003). "Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA." PLoS Biol **1**(2): E33.

Enoch, T. and P. Nurse (1990). "Mutation of fission yeast cell cycle control genes abolishes dependence of mitosis on DNA replication." Cell **60**(4): 665-673.

Evans, T., et al. (1983). "Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division." Cell **33**(2): 389-396.

Evrin, C., et al. (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**(48): 20240-20245.

Eykelenboom, J. K., et al. (2013). "ATR activates the S-M checkpoint during unperturbed growth to ensure sufficient replication prior to mitotic onset." Cell Rep **5**(4): 1095-1107.

Falck, J., et al. (2005). "Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage." Nature **434**(7033): 605-611.

Fanning, E., et al. (2006). "A dynamic model for replication protein A (RPA) function in DNA processing pathways." Nucleic Acids Res **34**(15): 4126-4137.

Fenech, M., et al. (1991). "Cloning and characterization of the rad4 gene of Schizosaccharomyces pombe; a gene showing short regions of sequence similarity to the human XRCC1 gene." Nucleic Acids Res **19**(24): 6737-6741.

Feng, S., et al. (2016). "Ewing Tumor-associated Antigen 1 Interacts with Replication Protein A to Promote Restart of Stalled Replication Forks." J Biol Chem **291**(42): 21956-21962.

Fernandez-Capetillo, O., et al. (2002). "DNA damage-induced G2-M checkpoint activation by histone H2AX and 53BP1." Nat Cell Biol **4**(12): 993-997.

Field, S. J., et al. (1996). "E2F-1 functions in mice to promote apoptosis and suppress proliferation." Cell **85**(4): 549-561.

Flatt, P. M., et al. (2000). "p53 regulation of G(2) checkpoint is retinoblastoma protein dependent." Mol Cell Biol **20**(12): 4210-4223.

Flynn, R. L. and L. Zou (2011). "ATR: a master conductor of cellular responses to DNA replication stress." Trends Biochem Sci **36**(3): 133-140.

Foltman, M., et al. (2013). "Eukaryotic replisome components cooperate to process histones during chromosome replication." Cell Rep **3**(3): 892-904.

Forsburg, S. L. and P. Nurse (1991). "Identification of a G1-type cyclin puc1+ in the fission yeast Schizosaccharomyces pombe." Nature **351**(6323): 245-248.

Foster, F. M., et al. (2003). "The phosphoinositide (PI) 3-kinase family." J Cell Sci **116**(Pt 15): 3037-3040.

Friedrich, G. and P. Soriano (1991). "Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice." Genes Dev **5**(9): 1513-1523.

Fukagawa, T., et al. (2001). "CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells." EMBO J **20**(16): 4603-4617.

Fukuura, M., et al. (2011). "CDK promotes interactions of Sld3 and Drc1 with Cut5 for initiation of DNA replication in fission yeast." Mol Biol Cell **22**(14): 2620-2633.

Furuya, K., et al. (2004). "Chk1 activation requires Rad9 S/TQ-site phosphorylation to promote association with C-terminal BRCT domains of Rad4TOPBP1." Genes Dev **18**(10): 1154-1164.

Gallina, I., et al. (2016). "TopBP1-mediated DNA processing during mitosis." Cell Cycle **15**(2): 176-183.

Gambus, A., et al. (2011). "MCM2-7 form double hexamers at licensed origins in Xenopus egg extract." J Biol Chem **286**(13): 11855-11864.

Gambus, A., et al. (2009). "A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome." EMBO J **28**(19): 2992-3004.

Garcia, V., et al. (2005). "Identification and functional analysis of TopBP1 and its homologs." DNA Repair (Amst) **4**(11): 1227-1239.

Germann, S. M., et al. (2011). "Dpb11/TopBP1 plays distinct roles in DNA replication, checkpoint response and homologous recombination." DNA Repair (Amst) **10**(2): 210-224.

Germann, S. M., et al. (2014). "TopBP1/Dpb11 binds DNA anaphase bridges to prevent genome instability." J Cell Biol **204**(1): 45-59.

Geva-Zatorsky, N., et al. (2006). "Oscillations and variability in the p53 system." Mol Syst Biol **2**: 2006 0033.

Ghosal, G. and J. Chen (2013). "DNA damage tolerance: a double-edged sword guarding the genome." Transl Cancer Res **2**(3): 107-129.

Godin, K. S. and G. Varani (2007). "How arginine-rich domains coordinate mRNA maturation events." RNA Biol **4**(2): 69-75.

Gong, Z., et al. (2010). "BACH1/FANCJ acts with TopBP1 and participates early in DNA replication checkpoint control." Mol Cell **37**(3): 438-446.

Gorgoulis, V. G., et al. (2005). "Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions." Nature **434**(7035): 907-913.

Gossen, M., et al. (1995). "Transcriptional activation by tetracyclines in mammalian cells." Science **268**(5218): 1766-1769.

Greenberg, R. A., et al. (2006). "Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes." Genes Dev **20**(1): 34-46.

Greenwell, P. W., et al. (1995). "TEL1, a gene involved in controlling telomere length in S. cerevisiae, is homologous to the human ataxia telangiectasia gene." Cell **82**(5): 823-829.

Greer, D. A., et al. (2003). "hRad9 rapidly binds DNA containing double-strand breaks and is required for damage-dependent topoisomerase II beta binding protein 1 focus formation." Cancer Res **63**(16): 4829-4835.

Gritenaite, D., et al. (2014). "A cell cycle-regulated Slx4-Dpb11 complex promotes the resolution of DNA repair intermediates linked to stalled replication." Genes Dev **28**(14): 1604-1619.

Guleria, A. and S. Chandna (2016). "ATM kinase: Much more than a DNA damage responsive protein." DNA Repair (Amst) **39**: 1-20.

Guo, Z., et al. (2000). "Requirement for Atr in phosphorylation of Chk1 and cell cycle regulation in response to DNA replication blocks and UV-damaged DNA in Xenopus egg extracts." Genes Dev **14**(21): 2745-2756.

Haahr, P., et al. (2016). "Activation of the ATR kinase by the RPA-binding protein ETAA1." Nat Cell Biol **18**(11): 1196-1207.

Haber, J. E. (2000). "Recombination: a frank view of exchanges and vice versa." Curr Opin Cell Biol **12**(3): 286-292.

Hagstrom, K. A. and B. J. Meyer (2003). "Condensin and cohesin: more than chromosome compactor and glue." Nat Rev Genet **4**(7): 520-534.

Hartlerode, A. J. and R. Scully (2009). "Mechanisms of double-strand break repair in somatic mammalian cells." Biochem J **423**(2): 157-168.

Hartwell, L. H. and T. A. Weinert (1989). "Checkpoints: controls that ensure the order of cell cycle events." Science **246**(4930): 629-634.

Hauf, S., et al. (2001). "Cohesin cleavage by separase required for anaphase and cytokinesis in human cells." Science **293**(5533): 1320-1323.

Hayles, J., et al. (1994). "Temporal order of S phase and mitosis in fission yeast is determined by the state of the p34cdc2-mitotic B cyclin complex." Cell **78**(5): 813-822.

Herold, S., et al. (2002). "Negative regulation of the mammalian UV response by Myc through association with Miz-1." Mol Cell **10**(3): 509-521.

Hirano, T., et al. (1986). "Isolation and characterization of Schizosaccharomyces pombe cutmutants that block nuclear division but not cytokinesis." EMBO J **5**(11): 2973-2979.

Hirao, A., et al. (2002). "Chk2 is a tumor suppressor that regulates apoptosis in both an ataxia telangiectasia mutated (ATM)-dependent and an ATM-independent manner." Mol Cell Biol **22**(18): 6521-6532.

Hochegger, H., et al. (2006). "Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells." EMBO J **25**(6): 1305-1314.

Hochegger, H., et al. (2008). "Cyclin-dependent kinases and cell-cycle transitions: does one fit all?" Nat Rev Mol Cell Biol **9**(11): 910-916.

Hodgson, B., et al. (2007). "Mrc1 and Tof1 regulate DNA replication forks in different ways during normal S phase." Mol Biol Cell **18**(10): 3894-3902.

Hoeijmakers, J. H. (2001). "Genome maintenance mechanisms for preventing cancer." Nature **411**(6835): 366-374.

Hong, J. H., et al. (1994). "G2/M-phase arrest and release in ataxia telangiectasia and normal cells after exposure to ionizing radiation." Radiat Res **140**(1): 17-23.

Hopkins, K. M., et al. (2004). "Deletion of mouse rad9 causes abnormal cellular responses to DNA damage, genomic instability, and embryonic lethality." Mol Cell Biol **24**(16): 7235-7248.

Houldsworth, J. and M. F. Lavin (1980). "Effect of ionizing radiation on DNA synthesis in ataxia telangiectasia cells." Nucleic Acids Res **8**(16): 3709-3720.

Hunter, T. (1995). "When is a lipid kinase not a lipid kinase? When it is a protein kinase." Cell **83**(1): 1-4.

Huppert, J. L., et al. (2008). "G-quadruplexes: the beginning and end of UTRs." Nucleic Acids Res **36**(19): 6260-6268.

Huyen, Y., et al. (2004). "Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks." Nature **432**(7015): 406-411.

Iiizumi, S., et al. (2006). "Simple one-week method to construct gene-targeting vectors: application to production of human knockout cell lines." Biotechniques **41**(3): 311-316.

Ikura, T., et al. (2007). "DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics." Mol Cell Biol **27**(20): 7028-7040.

Irion, S., et al. (2007). "Identification and targeting of the ROSA26 locus in human embryonic stem cells." Nat Biotechnol **25**(12): 1477-1482.

Jackman, M., et al. (2003). "Active cyclin B1-Cdk1 first appears on centrosomes in prophase." Nat Cell Biol **5**(2): 143-148.

Jackson, S. P. (2002). "Sensing and repairing DNA double-strand breaks." Carcinogenesis **23**(5): 687-696.

Jackson, S. P. and J. Bartek (2009). "The DNA-damage response in human biology and disease." Nature **461**(7267): 1071-1078.

Jeggo, P. A., et al. (1998). "Splitting the ATM: distinct repair and checkpoint defects in ataxia-telangiectasia." Trends Genet **14**(8): 312-316.

Jeon, Y., et al. (2011). "TopBP1 deficiency causes an early embryonic lethality and induces cellular senescence in primary cells." J Biol Chem **286**(7): 5414-5422.

Jeon, Y., et al. (2007). "Human TopBP1 participates in cyclin E/CDK2 activation and preinitiation complex assembly during G1/S transition." J Biol Chem **282**(20): 14882-14890.

Jiang, K., et al. (2003). "Regulation of Chk1 includes chromatin association and 14-3-3 binding following phosphorylation on Ser-345." J Biol Chem **278**(27): 25207-25217.

Jiang, W., et al. (1999). "Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication." EMBO J **18**(20): 5703-5713.

Jinek, M., et al. (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Science **337**(6096): 816-821.

Johnson, R. E., et al. (2015). "A Major Role of DNA Polymerase delta in Replication of Both the Leading and Lagging DNA Strands." Mol Cell **59**(2): 163-175.

Johnston, K., et al. (2010). "Vertebrate kinetochore protein architecture: protein copy number." J Cell Biol **189**(6): 937-943.

Jones, R. M. and E. Petermann (2012). "Replication fork dynamics and the DNA damage response." Biochem J **443**(1): 13-26.

Kaidi, A. and S. P. Jackson (2013). "KAT5 tyrosine phosphorylation couples chromatin sensing to ATM signalling." Nature **498**(7452): 70-74.

Kamimura, Y., et al. (2001). "Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in Saccharomyces cerevisiae." EMBO J **20**(8): 2097-2107.

Kastan, M. B. and J. Bartek (2004). "Cell-cycle checkpoints and cancer." Nature **432**(7015): 316-323.

Kastan, M. B., et al. (1991). "Participation of p53 protein in the cellular response to DNA damage." Cancer Res **51**(23 Pt 1): 6304-6311.

Kastan, M. B., et al. (1992). "A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia." Cell **71**(4): 587-597.

Khanna, K. K., et al. (1995). "Nature of G1/S cell cycle checkpoint defect in ataxia-telangiectasia." Oncogene **11**(4): 609-618.

Khanna, K. K., et al. (2001). "ATM, a central controller of cellular responses to DNA damage." Cell Death Differ **8**(11): 1052-1065.

Kim, C., et al. (1994). "Interactions of human replication protein A with oligonucleotides." Biochemistry **33**(47): 14197-14206.

Kim, J. H., et al. (2011). "High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice." PLoS One **6**(4): e18556.

Kim, J., et al. (2014). "TopBP1 deficiency impairs V(D)J recombination during lymphocyte development." EMBO J **33**(3): 217-228.

Kobayashi, M., et al. (2013). "NBS1 directly activates ATR independently of MRE11 and TOPBP1." Genes Cells **18**(3): 238-246.

Kolas, N. K., et al. (2007). "Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase." Science **318**(5856): 1637-1640.

Kozak, M. (1987). "An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs." Nucleic Acids Res **15**(20): 8125-8148.

Kozlov, S. V., et al. (2011). "Autophosphorylation and ATM activation: additional sites add to the complexity." J Biol Chem **286**(11): 9107-9119.

Kozlov, S. V., et al. (2006). "Involvement of novel autophosphorylation sites in ATM activation." EMBO J **25**(15): 3504-3514.

Kramer, A., et al. (2004). "Centrosome-associated Chk1 prevents premature activation of cyclin-B-Cdk1 kinase." Nat Cell Biol **6**(9): 884-891.

Kumagai, A. and W. G. Dunphy (2000). "Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts." Mol Cell **6**(4): 839-849.

Kumagai, A. and W. G. Dunphy (2003). "Repeated phosphopeptide motifs in Claspin mediate the regulated binding of Chk1." Nat Cell Biol **5**(2): 161-165.

Kumagai, A., et al. (2006). "TopBP1 activates the ATR-ATRIP complex." Cell **124**(5): 943-955.

Kumagai, A., et al. (2011). "Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication." J Cell Biol **193**(6): 995-1007.

Kumar, S. and P. M. Burgers (2013). "Lagging strand maturation factor Dna2 is a component of the replication checkpoint initiation machinery." Genes Dev **27**(3): 313-321.

Kunkel, T. A. and P. M. Burgers (2014). "Delivering nonidentical twins." Nat Struct Mol Biol **21**(8): 649-651.

Labib, K. (2010). "How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells?" Genes Dev **24**(12): 1208-1219.

Lambert, S. and A. M. Carr (2005). "Checkpoint responses to replication fork barriers." Biochimie **87**(7): 591-602.

Lambert, S. and A. M. Carr (2013). "Impediments to replication fork movement: stabilisation, reactivation and genome instability." Chromosoma **122**(1-2): 33-45.

Lander, E. S. (2016). "The Heroes of CRISPR." Cell **164**(1-2): 18-28.

Lavin, M. F. (2008). "Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer." Nat Rev Mol Cell Biol **9**(10): 759-769.

Lavin, M. F. and K. K. Khanna (1999). "ATM: the protein encoded by the gene mutated in the radiosensitive syndrome ataxia-telangiectasia." Int J Radiat Biol **75**(10): 1201-1214.

Lee, J. and W. G. Dunphy (2013). "The Mre11-Rad50-Nbs1 (MRN) complex has a specific role in the activation of Chk1 in response to stalled replication forks." Mol Biol Cell **24**(9): 1343-1353.

Lee, J., et al. (2001). "Positive regulation of Wee1 by Chk1 and 14-3-3 proteins." Mol Biol Cell **12**(3): 551-563.

Lee, J. H., et al. (2010). "53BP1 promotes ATM activity through direct interactions with the MRN complex." EMBO J **29**(3): 574-585.

Lee, Y., et al. (2012). "Neurogenesis requires TopBP1 to prevent catastrophic replicative DNA damage in early progenitors." Nat Neurosci **15**(6): 819-826.

Lee, Y. C., et al. (2016). "RPA-Binding Protein ETAA1 Is an ATR Activator Involved in DNA Replication Stress Response." Curr Biol **26**(24): 3257-3268.

Lees, E. (1995). "Cyclin dependent kinase regulation." Curr Opin Cell Biol **7**(6): 773-780.

Lehmann, A. R. (1993). "Duplicated region of sequence similarity to the human XRCC1 DNA repair gene in the Schizosaccharomyces pombe rad4/cut5 gene." Nucleic Acids Res **21**(22): 5274.

Leman, A. R. and E. Noguchi (2013). "The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication." Genes (Basel) **4**(1): 1-32.

Lempiainen, H. and T. D. Halazonetis (2009). "Emerging common themes in regulation of PIKKs and PI3Ks." EMBO J **28**(20): 3067-3073.

Leung, C. C., et al. (2011). "Molecular basis of BACH1/FANCJ recognition by TopBP1 in DNA replication checkpoint control." J Biol Chem **286**(6): 4292-4301.

Lin, S. J., et al. (2012). "The Rad4(TopBP1) ATR-activation domain functions in G1/S phase in a chromatin-dependent manner." PLoS Genet **8**(6): e1002801.

Lin, W. C., et al. (2001). "Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation." Genes Dev **15**(14): 1833-1844.

Lindsey-Boltz, L. A., et al. (2014). "Coupling of human DNA excision repair and the DNA damage checkpoint in a defined in vitro system." J Biol Chem **289**(8): 5074-5082.

Lindsey-Boltz, L. A. and A. Sancar (2011). "Tethering DNA damage checkpoint mediator proteins topoisomerase IIbeta-binding protein 1 (TopBP1) and Claspin to DNA activates ataxia-telangiectasia mutated and RAD3-related (ATR) phosphorylation of checkpoint kinase 1 (Chk1)." J Biol Chem **286**(22): 19229-19236.

Liu, K., et al. (2009). "Regulation of p53 by TopBP1: a potential mechanism for p53 inactivation in cancer." Mol Cell Biol **29**(10): 2673-2693.

Liu, K., et al. (2013). "Akt switches TopBP1 function from checkpoint activation to transcriptional regulation through phosphoserine binding-mediated oligomerization." Mol Cell Biol **33**(23): 4685-4700.

Liu, K., et al. (2003). "Regulation of E2F1 by BRCT domain-containing protein TopBP1." Mol Cell Biol **23**(9): 3287-3304.

Liu, K., et al. (2004). "TopBP1 recruits Brg1/Brm to repress E2F1-induced apoptosis, a novel pRb-independent and E2F1-specific control for cell survival." Genes Dev **18**(6): 673-686.

Liu, K., et al. (2006). "Regulation of TopBP1 oligomerization by Akt/PKB for cell survival." EMBO J **25**(20): 4795-4807.

Liu, Q., et al. (2000). "Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint." Genes Dev **14**(12): 1448-1459.

Liu, S., et al. (2013). "Nek1 kinase associates with ATR-ATRIP and primes ATR for efficient DNA damage signaling." Proc Natl Acad Sci U S A **110**(6): 2175-2180.

Liu, S., et al. (2011). "ATR autophosphorylation as a molecular switch for checkpoint activation." Mol Cell **43**(2): 192-202.

Lonn, U. and S. Lonn (1988). "Extensive regions of single-stranded DNA in aphidicolin-treated melanoma cells." Biochemistry **27**(2): 566-570.

Lopez-Girona, A., et al. (2001). "Serine-345 is required for Rad3-dependent phosphorylation and function of checkpoint kinase Chk1 in fission yeast." Proc Natl Acad Sci U S A **98**(20): 11289-11294.

Lou, Z., et al. (2006). "MDC1 maintains genomic stability by participating in the amplification of ATM-dependent DNA damage signals." Mol Cell **21**(2): 187-200.

Lovejoy, C. A. and D. Cortez (2009). "Common mechanisms of PIKK regulation." DNA Repair (Amst) **8**(9): 1004-1008.

Lukas, C., et al. (2001). "DNA damage-activated kinase Chk2 is independent of proliferation or differentiation yet correlates with tissue biology." Cancer Res **61**(13): 4990-4993.

Lukas, C., et al. (2003). "Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage." Nat Cell Biol **5**(3): 255-260.

MacDougall, C. A., et al. (2007). "The structural determinants of checkpoint activation." Genes Dev **21**(8): 898-903.

Mahaney, B. L., et al. (2009). "Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining." Biochem J **417**(3): 639-650.

Mailand, N., et al. (2007). "RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins." Cell **131**(5): 887-900.

Mailand, N., et al. (2000). "Rapid destruction of human Cdc25A in response to DNA damage." Science **288**(5470): 1425-1429.

Mailand, N., et al. (2002). "Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability." EMBO J **21**(21): 5911-5920.
Maiorano, D., et al. (2000). "XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis." Nature **404**(6778): 622-625.

Majka, J., et al. (2006). "Replication protein A directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions." J Biol Chem **281**(38): 27855-27861.

Makiniemi, M., et al. (2001). "BRCT domain-containing protein TopBP1 functions in DNA replication and damage response." J Biol Chem **276**(32): 30399-30406.

Malumbres, M. (2014). "Cyclin-dependent kinases." Genome Biol **15**(6): 122.

Malumbres, M. and M. Barbacid (2005). "Mammalian cyclin-dependent kinases." Trends Biochem Sci **30**(11): 630-641.

Manke, I. A., et al. (2003). "BRCT repeats as phosphopeptide-binding modules involved in protein targeting." Science **302**(5645): 636-639.

Mantiero, D., et al. (2011). "Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast." EMBO J **30**(23): 4805-4814.

Marechal, A., et al. (2014). "PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry." Mol Cell **53**(2): 235-246.

Marechal, A. and L. Zou (2013). "DNA damage sensing by the ATM and ATR kinases." Cold Spring Harb Perspect Biol **5**(9).

Marechal, A. and L. Zou (2015). "RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response." Cell Res **25**(1): 9-23.

Marteijn, J. A., et al. (2014). "Understanding nucleotide excision repair and its roles in cancer and ageing." Nat Rev Mol Cell Biol **15**(7): 465-481.

Masai, H., et al. (2006). "Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin." J Biol Chem **281**(51): 39249-39261.

Masumoto, H., et al. (2002). "S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast." Nature **415**(6872): 651-655.

Matsuoka, S., et al. (2007). "ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage." Science **316**(5828): 1160-1166.

Meilinger, D., et al. (2009). "Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells." EMBO Rep **10**(11): 1259-1264.

McGlynn, P. and R. G. Lloyd (2002). "Replicating past lesions in DNA." Mol Cell **10**(4): 700-701.

McIntosh, D. and J. J. Blow (2012). "Dormant origins, the licensing checkpoint, and the response to replicative stresses." Cold Spring Harb Perspect Biol **4**(10).

Meilinger, D., et al. (2009). "Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells." EMBO Rep **10**(11): 1259-1264.

Melixetian, M., et al. (2009). "NEK11 regulates CDC25A degradation and the IR-induced G2/M checkpoint." Nat Cell Biol **11**(10): 1247-1253.

Melo, J. A., et al. (2001). "Two checkpoint complexes are independently recruited to sites of DNA damage in vivo." Genes Dev **15**(21): 2809-2821.

Meselson, M. and F. W. Stahl (1958). "The Replication of DNA in Escherichia Coli." Proc Natl Acad Sci U S A **44**(7): 671-682.

Miyabe, I., et al. (2015). "Polymerase delta replicates both strands after homologous recombination-dependent fork restart." Nat Struct Mol Biol **22**(11): 932-938.

Mondesert, O., et al. (1996). "Cig2, a B-type cyclin, promotes the onset of S in Schizosaccharomyces pombe." Mol Cell Biol **16**(4): 1527-1533.

Mordes, D. A., et al. (2008). "TopBP1 activates ATR through ATRIP and a PIKK regulatory domain." Genes Dev **22**(11): 1478-1489.

Mordes, D. A., et al. (2008). "Dpb11 activates the Mec1-Ddc2 complex." Proc Natl Acad Sci U S A **105**(48): 18730-18734.

Morishima, K., et al. (2007). "TopBP1 associates with NBS1 and is involved in homologous recombination repair." Biochem Biophys Res Commun **362**(4): 872-879.

Moudry, P., et al. (2016). "TOPBP1 regulates RAD51 phosphorylation and chromatin loading and determines PARP inhibitor sensitivity." J Cell Biol **212**(3): 281-288.

Murray, J. M., et al. (1991). "Cloning and characterisation of the rad9 DNA repair gene from Schizosaccharomyces pombe." Nucleic Acids Res **19**(13): 3525-3531.

Myers, J. S. and D. Cortez (2006). "Rapid activation of ATR by ionizing radiation requires ATM and Mre11." J Biol Chem **281**(14): 9346-9350.

Nakada, D., et al. (2005). "Role of the C terminus of Mec1 checkpoint kinase in its localization to sites of DNA damage." Mol Biol Cell **16**(11): 5227-5235.

Nam, E. A. and D. Cortez (2011). "ATR signalling: more than meeting at the fork." Biochem J **436**(3): 527-536.

Navadgi-Patil, V. M. and P. M. Burgers (2008). "Yeast DNA replication protein Dpb11 activates the Mec1/ATR checkpoint kinase." J Biol Chem **283**(51): 35853-35859.

Navadgi-Patil, V. M. and P. M. Burgers (2009). "The unstructured C-terminal tail of the 9-1-1 clamp subunit Ddc1 activates Mec1/ATR via two distinct mechanisms." Mol Cell **36**(5): 743-753.

Nghiem, P., et al. (2001). "ATR inhibition selectively sensitizes G1 checkpoint-deficient cells to lethal premature chromatin condensation." Proc Natl Acad Sci U S A **98**(16): 9092-9097.

Nick McElhinny, S. A., et al. (2000). "Ku recruits the XRCC4-ligase IV complex to DNA ends." Mol Cell Biol **20**(9): 2996-3003.

Nishimura, K., et al. (2009). "An auxin-based degron system for the rapid depletion of proteins in nonplant cells." Nat Methods **6**(12): 917-922.

Nishimura, K. and M. T. Kanemaki (2014). "Rapid Depletion of Budding Yeast Proteins via the Fusion of an Auxin-Inducible Degron (AID)." Curr Protoc Cell Biol **64**: 20 29 21-16.

Nishitani, H., et al. (2000). "The Cdt1 protein is required to license DNA for replication in fission yeast." Nature **404**(6778): 625-628.

Nurse, P. (1997). "Regulation of the eukaryotic cell cycle." Eur J Cancer **33**(7): 1002-1004.

Nurse, P., et al. (1998). "Understanding the cell cycle." Nat Med **4**(10): 1103-1106.

O'Connell, M. J., et al. (2000). "The G2-phase DNA-damage checkpoint." Trends Cell Biol **10**(7): 296-303.

O'Donnell, M. and H. Li (2016). "The Eukaryotic Replisome Goes Under the Microscope." Curr Biol **26**(6): R247-256.

Oda, K., et al. (2000). "p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53." Cell **102**(6): 849-862.

Ohashi, E., et al. (2014). "Interaction between Rad9-Hus1-Rad1 and TopBP1 activates ATR-ATRIP and promotes TopBP1 recruitment to sites of UV-damage." DNA Repair (Amst) **21**: 1-11.

Ohouo, P. Y., et al. (2010). "DNA damage signaling recruits the Rtt107-Slx4 scaffolds via Dpb11 to mediate replication stress response." Mol Cell **39**(2): 300-306.

Ohouo, P. Y., et al. (2013). "DNA-repair scaffolds dampen checkpoint signalling by counteracting the adaptor Rad9." Nature **493**(7430): 120-124.

Painter, R. B. and B. R. Young (1980). "Radiosensitivity in ataxia-telangiectasia: a new explanation." Proc Natl Acad Sci U S A **77**(12): 7315-7317.

Pardee, A. B. (1974). "A restriction point for control of normal animal cell proliferation." Proc Natl Acad Sci U S A **71**(4): 1286-1290.

Parrilla-Castellar, E. R., et al. (2004). "Dial 9-1-1 for DNA damage: the Rad9-Hus1-Rad1 (9-1-1) clamp complex." DNA Repair (Amst) **3**(8-9): 1009-1014.

Paull, T. T. (2015). "Mechanisms of ATM Activation." Annu Rev Biochem **84**: 711-738.

Paull, T. T. and J. H. Lee (2005). "The Mre11/Rad50/Nbs1 complex and its role as a DNA double-strand break sensor for ATM." Cell Cycle **4**(6): 737-740.

Paulsen, R. D. and K. A. Cimprich (2007). "The ATR pathway: fine-tuning the fork." DNA Repair (Amst) **6**(7): 953-966.

Paweletz, N. (2001). "Walther Flemming: pioneer of mitosis research." Nat Rev Mol Cell Biol **2**(1): 72-75.

Pedersen, R. T., et al. (2015). "TopBP1 is required at mitosis to reduce transmission of DNA damage to G1 daughter cells." J Cell Biol **210**(4): 565-582.

Peng, C. Y., et al. (1997). "Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216." Science **277**(5331): 1501-1505.

Peng, M., et al. (2006). "BACH1 is a DNA repair protein supporting BRCA1 damage response." Oncogene **25**(15): 2245-2253.

Perry, J. and N. Kleckner (2003). "The ATRs, ATMs, and TORs are giant HEAT repeat proteins." Cell **112**(2): 151-155.

Petermann, E., et al. (2008). "Claspin promotes normal replication fork rates in human cells." Mol Biol Cell **19**(6): 2373-2378.

Peters, J. M. (2002). "The anaphase-promoting complex: proteolysis in mitosis and beyond." Mol Cell **9**(5): 931-943.

Pfander, B. and J. F. Diffley (2011). "Dpb11 coordinates Mec1 kinase activation with cell cycle-regulated Rad9 recruitment." EMBO J **30**(24): 4897-4907.

Plans, V., et al. (2006). "The RING finger protein RNF8 recruits UBC13 for lysine 63-based self polyubiquitylation." J Cell Biochem **97**(3): 572-582.

Polo, S. E. and S. P. Jackson (2011). "Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications." Genes Dev **25**(5): 409-433.

Powell, S. K., et al. (2015). "Viral expression cassette elements to enhance transgene target specificity and expression in gene therapy." Discov Med **19**(102): 49-57.

Prakash, R., et al. (2015). "Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins." Cold Spring Harb Perspect Biol **7**(4): a016600.

Puddu, F., et al. (2008). "Phosphorylation of the budding yeast 9-1-1 complex is required for Dpb11 function in the full activation of the UV-induced DNA damage checkpoint." Mol Cell Biol **28**(15): 4782-4793.

Qin, J. Y., et al. (2010). "Systematic comparison of constitutive promoters and the doxycycline-inducible promoter." PLoS One **5**(5): e10611.

Qu, M., et al. (2013). "Phosphorylation-dependent assembly and coordination of the DNA damage checkpoint apparatus by Rad4(TopBP1)." Mol Cell **51**(6): 723-736.

Ramirez-Lugo, J. S., et al. (2011). "CtIP interacts with TopBP1 and Nbs1 in the response to double-stranded DNA breaks (DSBs) in Xenopus egg extracts." Cell Cycle **10**(3): 469-480.

Ran, F. A., et al. (2013). "Genome engineering using the CRISPR-Cas9 system." Nat Protoc **8**(11): 2281-2308.

Rappas, M., et al. (2011). "Structure and function of the Rad9-binding region of the DNA-damage checkpoint adaptor TopBP1." Nucleic Acids Res **39**(1): 313-324.

Recolin, B., et al. (2014). "Molecular mechanisms of DNA replication checkpoint activation." Genes (Basel) **5**(1): 147-175.

Reini, K., et al. (2004). "TopBP1 localises to centrosomes in mitosis and to chromosome cores in meiosis." Chromosoma **112**(7): 323-330.

Rhind, N. and P. Russell (2000). "Chk1 and Cds1: linchpins of the DNA damage and replication checkpoint pathways." J Cell Sci **113 (Pt 22)**: 3889-3896.

Rialland, M., et al. (2002). "Essential role of human CDT1 in DNA replication and chromatin licensing." J Cell Sci **115**(Pt 7): 1435-1440.

Rozenzhak, S., et al. (2010). "Rad3 decorates critical chromosomal domains with gammaH2A to protect genome integrity during S-Phase in fission yeast." PLoS Genet **6**(7): e1001032.

Russell, P. and P. Nurse (1986). "cdc25+ functions as an inducer in the mitotic control of fission yeast." Cell **45**(1): 145-153.

Russell, P. and P. Nurse (1987). "Negative regulation of mitosis by wee1+, a gene encoding a protein kinase homolog." Cell **49**(4): 559-567.

Ryan, M. D., et al. (1991). "Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence." J Gen Virol **72 (Pt 11)**: 2727-2732.

Saka, Y., et al. (1997). "Damage and replication checkpoint control in fission yeast is ensured by interactions of Crb2, a protein with BRCT motif, with Cut5 and Chk1." Genes Dev **11**(24): 3387-3400.

Saka, Y. and M. Yanagida (1993). "Fission yeast cut5+, required for S phase onset and M phase restraint, is identical to the radiation-damage repair gene rad4+." Cell **74**(2): 383-393.

Sakabe, K. and R. Okazaki (1966). "A unique property of the replicating region of chromosomal DNA." Biochim Biophys Acta **129**(3): 651-654.

Sanchez, Y., et al. (1997). "Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25." Science **277**(5331): 1497-1501.

Savic, V., et al. (2009). "Formation of dynamic gamma-H2AX domains along broken DNA strands is distinctly regulated by ATM and MDC1 and dependent upon H2AX densities in chromatin." Mol Cell **34**(3): 298-310.

Savitsky, K., et al. (1995). "A single ataxia telangiectasia gene with a product similar to PI-3 kinase." Science **268**(5218): 1749-1753.

Schupbach, M. (1971). "The isolation and genetic classification of UV-sensitive mutants of Schizosaccharomyces pombe." Mutat Res **11**(4): 361-371.

Semple, J. I., et al. (2007). "Cleavage and degradation of Claspin during apoptosis by caspases and the proteasome." Cell Death Differ **14**(8): 1433-1442.

Sheng, Z. Z., et al. (2011). "Functional Evolution of BRCT Domains from Binding DNA to Protein." Evol Bioinform Online **7**: 87-97.

Sheu, Y. J. and B. Stillman (2006). "Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression." Mol Cell **24**(1): 101-113.

Shiloh, Y. and Y. Ziv (2013). "The ATM protein kinase: regulating the cellular response to genotoxic stress, and more." Nat Rev Mol Cell Biol **14**(4): 197-210.

Shin, M. H., et al. (2012). "ATM-dependent phosphorylation of the checkpoint clamp regulates repair pathways and maintains genomic stability." Cell Cycle **11**(9): 1796-1803.

Shiotani, B., et al. (2013). "Two distinct modes of ATR activation orchestrated by Rad17 and Nbs1." Cell Rep **3**(5): 1651-1662.

Shiotani, B. and L. Zou (2009). "Single-stranded DNA orchestrates an ATM-to-ATR switch at DNA breaks." Mol Cell **33**(5): 547-558.

Simon, A. C., et al. (2016). "Structure of human Cdc45 and implications for CMG helicase function." Nat Commun **7**: 11638.

Smith, G. C., et al. (1999). "Purification and DNA binding properties of the ataxia-telangiectasia gene product ATM." Proc Natl Acad Sci U S A **96**(20): 11134-11139.

Smith, J. and D. W. Burt (1998). "Parameters of the chicken genome (Gallus gallus)." Anim Genet **29**(4): 290-294.

Smith, J., et al. (2010). "The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer." Adv Cancer Res **108**: 73-112.

Smits, V. A. (2006). "Spreading the signal: dissociation of Chk1 from chromatin." Cell Cycle **5**(10): 1039-1043.

Smits, V. A., et al. (2006). "Rapid PIKK-dependent release of Chk1 from chromatin promotes the DNA-damage checkpoint response." Curr Biol **16**(2): 150-159.

So, S., et al. (2009). "Autophosphorylation at serine 1981 stabilizes ATM at DNA damage sites." J Cell Biol **187**(7): 977-990.

Sobhian, B., et al. (2007). "RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites." Science **316**(5828): 1198-1202.

Sonoda, E., et al. (2006). "Differential usage of non-homologous end-joining and homologous recombination in double strand break repair." DNA Repair (Amst) **5**(9-10): 1021-1029.

Sonoda, E., et al. (1998). "Rad51-deficient vertebrate cells accumulate chromosomal breaks prior to cell death." EMBO J **17**(2): 598-608.

Sorensen, C. S., et al. (2004). "ATR, Claspin and the Rad9-Rad1-Hus1 complex regulate Chk1 and Cdc25A in the absence of DNA damage." Cell Cycle **3**(7): 941-945.

Speck, C., et al. (2005). "ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA." Nat Struct Mol Biol **12**(11): 965-971.

St Onge, R. P., et al. (2003). "A role for the phosphorylation of hRad9 in checkpoint signaling." J Biol Chem **278**(29): 26620-26628.

Stoecklin, G. and S. Diederichs (2014). "tRNAs: new tricks from old dogs." EMBO J **33**(18): 1981-1983.

Stucki, M., et al. (2005). "MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks." Cell **123**(7): 1213-1226.

Sui, G., et al. (2004). "Yin Yang 1 is a negative regulator of p53." Cell **117**(7): 859-872.

Sun, J., et al. (2016). "The eukaryotic CMG helicase pumpjack and integration into the replisome." Nucleus **7**(2): 146-154.

Sun, Y., et al. (2007). "DNA damage-induced acetylation of lysine 3016 of ATM activates ATM kinase activity." Mol Cell Biol **27**(24): 8502-8509.

Syljuasen, R. G., et al. (2005). "Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage." Mol Cell Biol **25**(9): 3553-3562.

Szuts, D., et al. (2006). "Role for RAD18 in homologous recombination in DT40 cells." Mol Cell Biol **26**(21): 8032-8041.

Takai, H., et al. (2002). "Chk2-deficient mice exhibit radioresistance and defective p53-mediated transcription." EMBO J **21**(19): 5195-5205.

Takai, H., et al. (2000). "Aberrant cell cycle checkpoint function and early embryonic death in Chk1(-/-) mice." Genes Dev **14**(12): 1439-1447.

Takao, N., et al. (1999). "Disruption of ATM in p53-null cells causes multiple functional abnormalities in cellular response to ionizing radiation." Oncogene **18**(50): 7002-7009.

Tanaka, K. and P. Russell (2001). "Mrc1 channels the DNA replication arrest signal to checkpoint kinase Cds1." Nat Cell Biol **3**(11): 966-972.

Tanaka, S. and H. Araki (2010). "Regulation of the initiation step of DNA replication by cyclin-dependent kinases." Chromosoma **119**(6): 565-574.

Taricani, L. and T. S. Wang (2006). "Rad4TopBP1, a scaffold protein, plays separate roles in DNA damage and replication checkpoints and DNA replication." Mol Biol Cell **17**(8): 3456-3468.

Tercero, J. A., et al. (2000). "DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p." EMBO J **19**(9): 2082-2093.

Teschendorf, C., et al. (2002). "Comparison of the EF-1 alpha and the CMV promoter for engineering stable tumor cell lines using recombinant adeno-associated virus." Anticancer Res **22**(6A): 3325-3330.

Tikole, S. and R. Sankararamakrishnan (2006). "A survey of mRNA sequences with a non-AUG start codon in RefSeq database." J Biomol Struct Dyn **24**(1): 33-42.

Toledo, L. I., et al. (2013). "ATR prohibits replication catastrophe by preventing global exhaustion of RPA." Cell **155**(5): 1088-1103.

Touriol, C., et al. (2003). "Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons." Biol Cell **95**(3-4): 169-178.

Toyoshima-Morimoto, F., et al. (2002). "Plk1 promotes nuclear translocation of human Cdc25C during prophase." EMBO Rep **3**(4): 341-348.

Trenz, K., et al. (2008). "Plx1 is required for chromosomal DNA replication under stressful conditions." EMBO J **27**(6): 876-885.

Tsai, A. G. and M. R. Lieber (2010). "Mechanisms of chromosomal rearrangement in the human genome." BMC Genomics **11 Suppl 1**: S1.

Tsaponina, O. and A. Chabes (2013). "Pre-activation of the genome integrity checkpoint increases DNA damage tolerance." Nucleic Acids Res **41**(22): 10371-10378.

Ueda, S., et al. (2012). "Two serine phosphorylation sites in the C-terminus of Rad9 are critical for 9-1-1 binding to TopBP1 and activation of the DNA damage checkpoint response in HeLa cells." Genes Cells **17**(10): 807-816.

Unsal-Kacmaz, K., et al. (2007). "The human Tim/Tipin complex coordinates an Intra-S checkpoint response to UV that slows replication fork displacement." Mol Cell Biol **27**(8): 3131-3142.

van Deursen, F., et al. (2012). "Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation." EMBO J **31**(9): 2195-2206.

Velez-Cruz, R., et al. (2016). "RB localizes to DNA double-strand breaks and promotes DNA end resection and homologous recombination through the recruitment of BRG1." Genes Dev **30**(22): 2500-2512.

Villa, F., et al. (2016). "Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase." Mol Cell **63**(3): 385-396.

Wahl, G. M. and A. M. Carr (2001). "The evolution of diverse biological responses to DNA damage: insights from yeast and p53." Nat Cell Biol **3**(12): E277-286.

Wang, B. and S. J. Elledge (2007). "Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage." Proc Natl Acad Sci U S A **104**(52): 20759-20763.

Wang, J., et al. (2013). "TopBP1 controls BLM protein level to maintain genome stability." Mol Cell **52**(5): 667-678.

Wang, J., et al. (2011). "MDC1 collaborates with TopBP1 in DNA replication checkpoint control." J Cell Biol **193**(2): 267-273.

Wang, X., et al. (2006). "Rad17 phosphorylation is required for claspin recruitment and Chk1 activation in response to replication stress." Mol Cell **23**(3): 331-341.

Wang, X. Q., et al. (2006). "ATR dependent activation of Chk2." J Cell Physiol **208**(3): 613-619.

Ward, I. M. and J. Chen (2001). "Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress." J Biol Chem **276**(51): 47759-47762.

Wardlaw, C. P., et al. (2014). "TopBP1: A BRCT-scaffold protein functioning in multiple cellular pathways." DNA Repair (Amst) **22**: 165-174.

Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-738.

Watts, F. Z. and N. C. Brissett (2010). "Linking up and interacting with BRCT domains." DNA Repair (Amst) **9**(2): 103-108.

Weinert, T. A. and L. H. Hartwell (1988). "The RAD9 gene controls the cell cycle response to DNA damage in Saccharomyces cerevisiae." Science **241**(4863): 317-322.

Wohlschlegel, J. A., et al. (2000). "Inhibition of eukaryotic DNA replication by geminin binding to Cdt1." Science **290**(5500): 2309-2312.

Wollmann, Y., et al. (2007). "The DNA topoisomerase IIbeta binding protein 1 (TopBP1) interacts with poly (ADP-ribose) polymerase (PARP-1)." J Cell Biochem **102**(1): 171-182.

Wright, R. H., et al. (2006). "TopBP1 contains a transcriptional activation domain suppressed by two adjacent BRCT domains." Biochem J **400**(3): 573-582.

Wu, L., et al. (2014). "Mechanism of chromosomal DNA replication initiation and replication fork stabilization in eukaryotes." Sci China Life Sci **57**(5): 482-487.

Wu, L., et al. (2008). "MDC1 regulates intra-S-phase checkpoint by targeting NBS1 to DNA double-strand breaks." Proc Natl Acad Sci U S A **105**(32): 11200-11205.

Xia, Z., et al. (2001). "Negative cell cycle regulation and DNA damage-inducible phosphorylation of the BRCT protein 53BP1." J Biol Chem **276**(4): 2708-2718.

Xu, B., et al. (2001). "Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation." Mol Cell Biol **21**(10): 3445-3450.

Xu, X., et al. (2008). "The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling." Mol Cell Biol **28**(24): 7345-7353.

Yabuuchi, H., et al. (2006). "Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins." EMBO J **25**(19): 4663-4674.

Yamane, K., et al. (1997). "A DNA-topoisomerase-II-binding protein with eight repeating regions similar to DNA-repair enzymes and to a cell-cycle regulator." Eur J Biochem **250**(3): 794-799.

Yamane, K. and T. Tsuruo (1999). "Conserved BRCT regions of TopBP1 and of the tumor suppressor BRCA1 bind strand breaks and termini of DNA." Oncogene **18**(37): 5194-5203.

Yamane, K., et al. (2002). "A DNA damage-regulated BRCT-containing protein, TopBP1, is required for cell survival." Mol Cell Biol **22**(2): 555-566.

Yamazoe, M., et al. (2004). "Reverse genetic studies of the DNA damage response in the chicken B lymphocyte line DT40." DNA Repair (Amst) **3**(8-9): 1175-1185.

Yan, S., et al. (2006). "Direct requirement for Xmus101 in ATR-mediated phosphorylation of Claspin bound Chk1 during checkpoint signaling." J Cell Biol **173**(2): 181-186.

Yan, S. and J. Willis (2013). "WD40-repeat protein WDR18 collaborates with TopBP1 to facilitate DNA damage checkpoint signaling." Biochem Biophys Res Commun **431**(3): 466-471.

Yao, N. and M. O'Donnell (2016). "Bacterial and Eukaryotic Replisome Machines." JSM Biochem Mol Biol **3**(1).

Yarden, R. I., et al. (2002). "BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage." Nat Genet **30**(3): 285-289.

Yoo, H. Y., et al. (2007). "Ataxia-telangiectasia mutated (ATM)-dependent activation of ATR occurs through phosphorylation of TopBP1 by ATM." J Biol Chem **282**(24): 17501-17506.

Yoo, H. Y., et al. (2009). "The Mre11-Rad50-Nbs1 complex mediates activation of TopBP1 by ATM." Mol Biol Cell **20**(9): 2351-2360.

Yoo, H. Y., et al. (2004). "Mcm2 is a direct substrate of ATM and ATR during DNA damage and DNA replication checkpoint responses." J Biol Chem **279**(51): 53353-53364.

Yuan, Z., et al. (2016). "Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation." Nat Struct Mol Biol **23**(3): 217-224.

Zachos, G., et al. (2003). "Chk1-deficient tumour cells are viable but exhibit multiple checkpoint and survival defects." EMBO J **22**(3): 713-723.

Zegerman, P. and J. F. Diffley (2007). "Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast." Nature **445**(7125): 281-285.

Zegerman, P. and J. F. Diffley (2010). "Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation." Nature **467**(7314): 474-478.

Zegerman, P. (2015). "Evolutionary conservation of the CDK targets in eukaryotic DNA replication initiation." Chromosoma **124**(3): 309-321.

Zhang, Y., et al. (1996). "Inducible site-directed recombination in mouse embryonic stem cells." Nucleic Acids Res **24**(4): 543-548.

Zhang, X. P., et al. (2009). "Cell fate decision mediated by p53 pulses." Proc Natl Acad Sci U S A **106**(30): 12245-12250.

Zhou, Z. W., et al. (2013). "An essential function for the ATR-activation-domain (AAD) of TopBP1 in mouse development and cellular senescence." PLoS Genet **9**(8): e1003702.

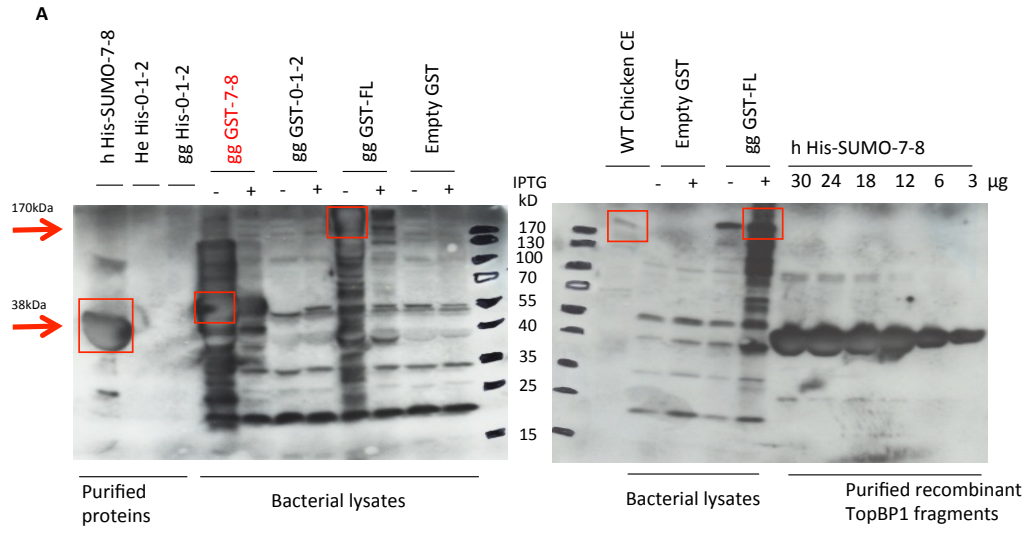
Zimmermann, K., et al. (2002). "Targeted disruption of the GAS41 gene encoding a putative transcription factor indicates that GAS41 is essential for cell viability." J Biol Chem **277**(21): 18626-18631.

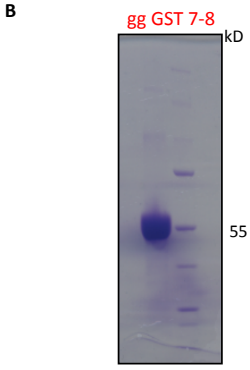
Zou, L. and S. J. Elledge (2003). "Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes." Science **300**(5625): 1542-1548.

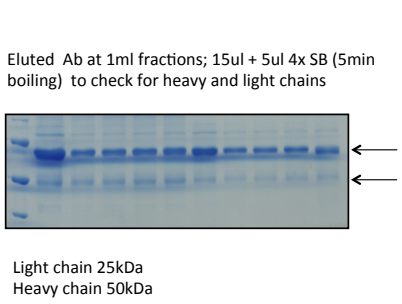
APPENDIX 1

Generation of a polyclonal anti-TopBP1 antibody

To study the function of the avian TopBP1, an antibody against a C-terminal peptide of the *gallus gallus* protein was raised in rabbits by Eurogentec. This corresponded to 55kDa. To first test the specificity of the antibody, I performed some Western blot (as shown in **A**) analysis using recombinant fragments (or the corresponding bacterial lysates) of either terminus of the protein as well as a full length clone to quickly test whether the serum could specifically recognize the peptide used to immunize the rabbits. Once specificity was confirmed, this peptide was recombinantly expressed in bacteria (as shown in **B**) and used to purify the TopBP1 antibody from the rabbit serum as explained in the Material and Methods section. The gel in C shows the heavy and light chains stained with Coomassie blue following elution of the antibody from the column. All work in this thesis makes use of this antibody.

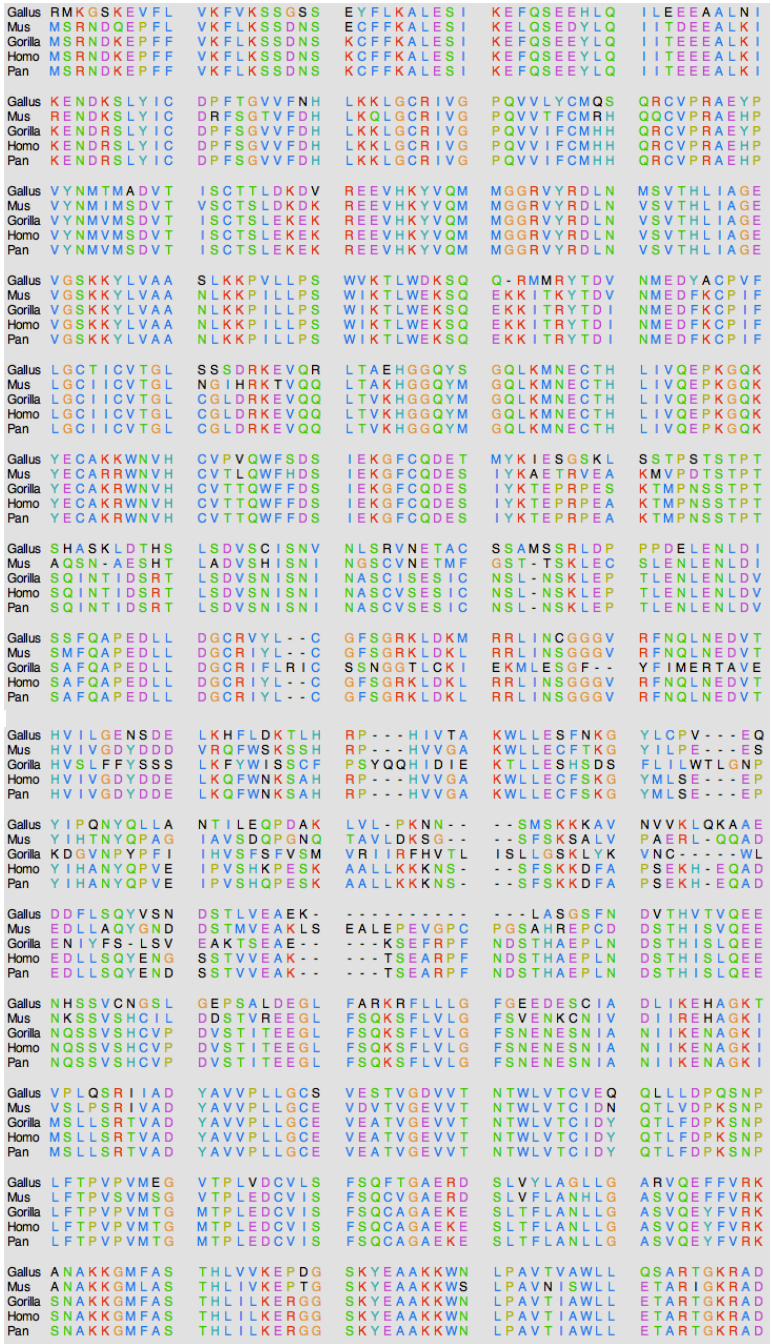


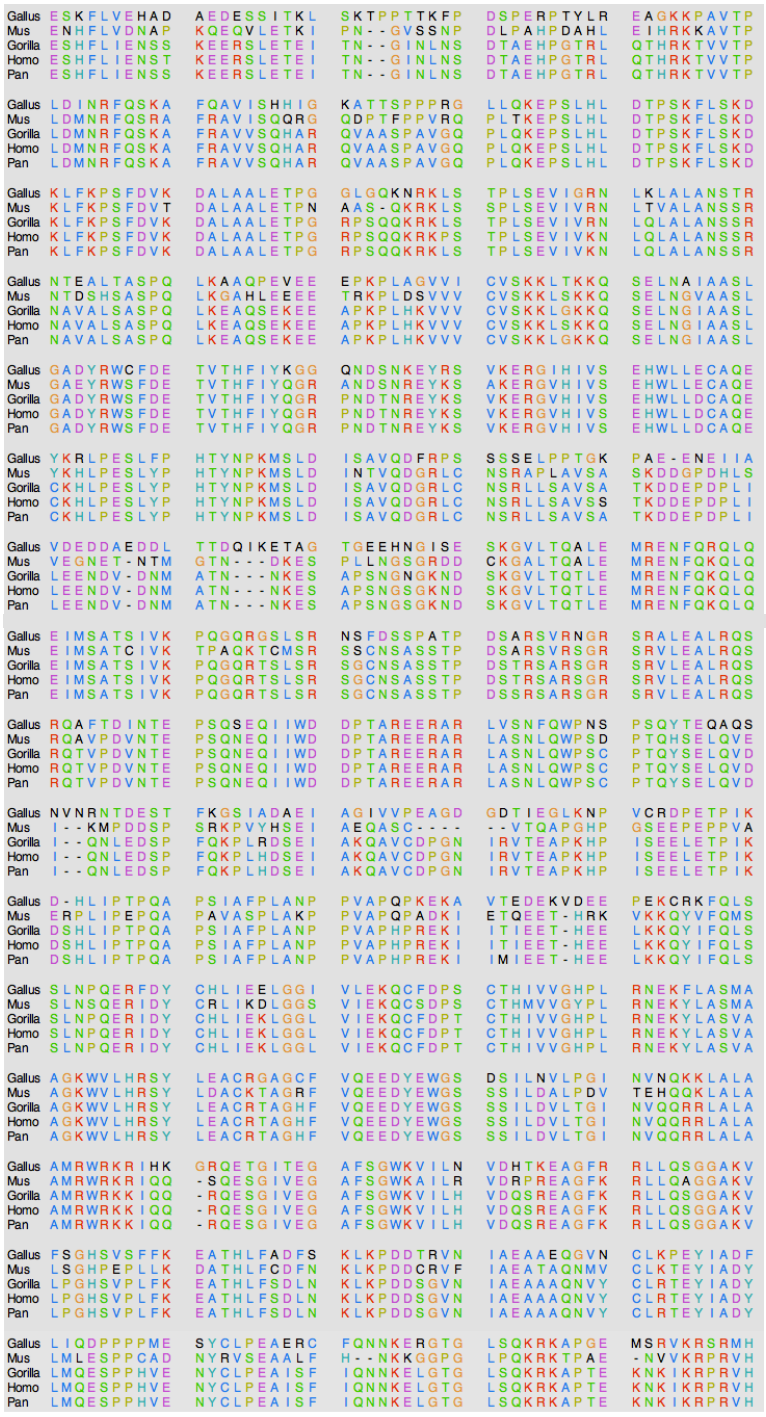




APPENDIX 2

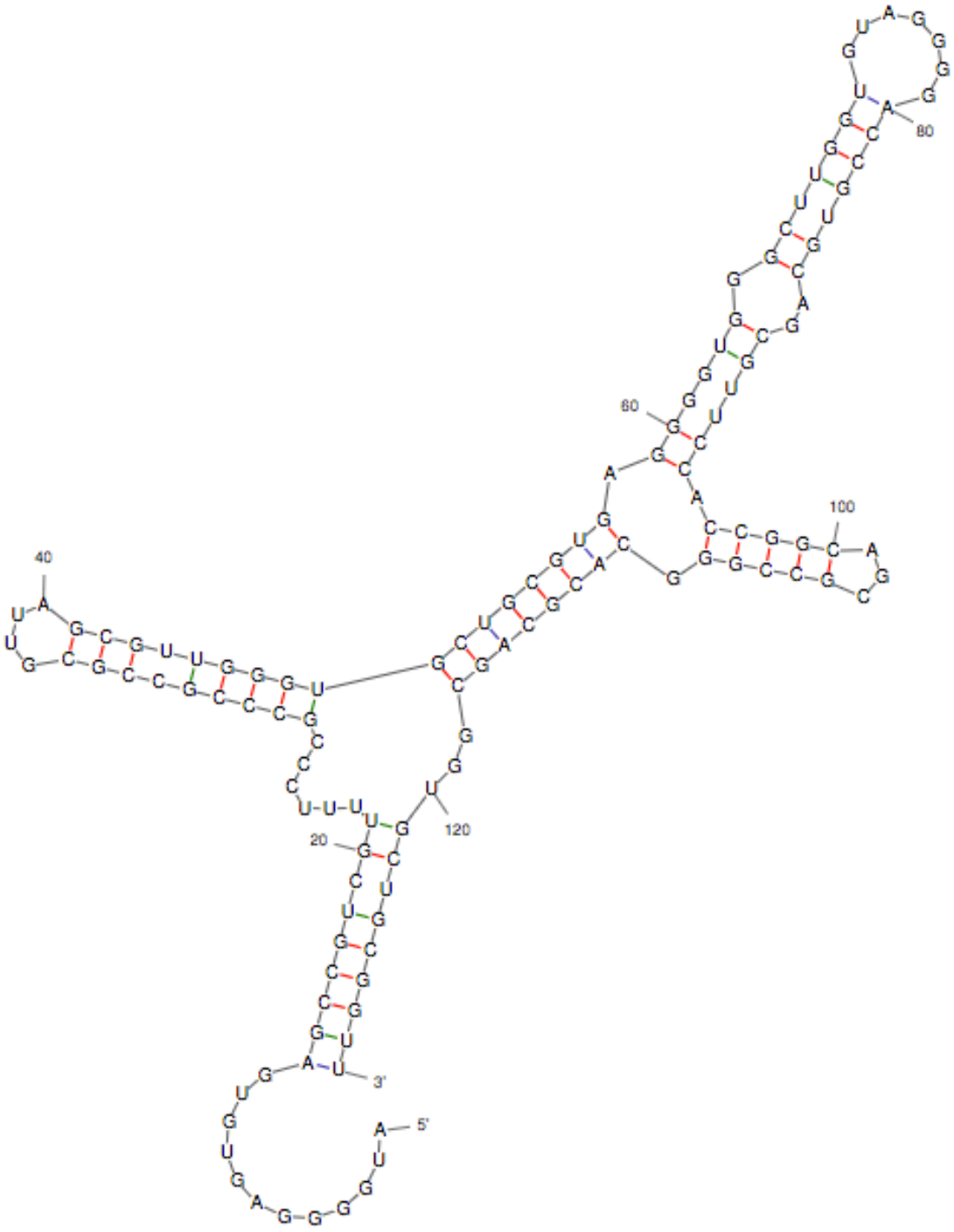
TopBP1 protein sequence alignment in different organisms as generated by the ClustalW sequence alignment software

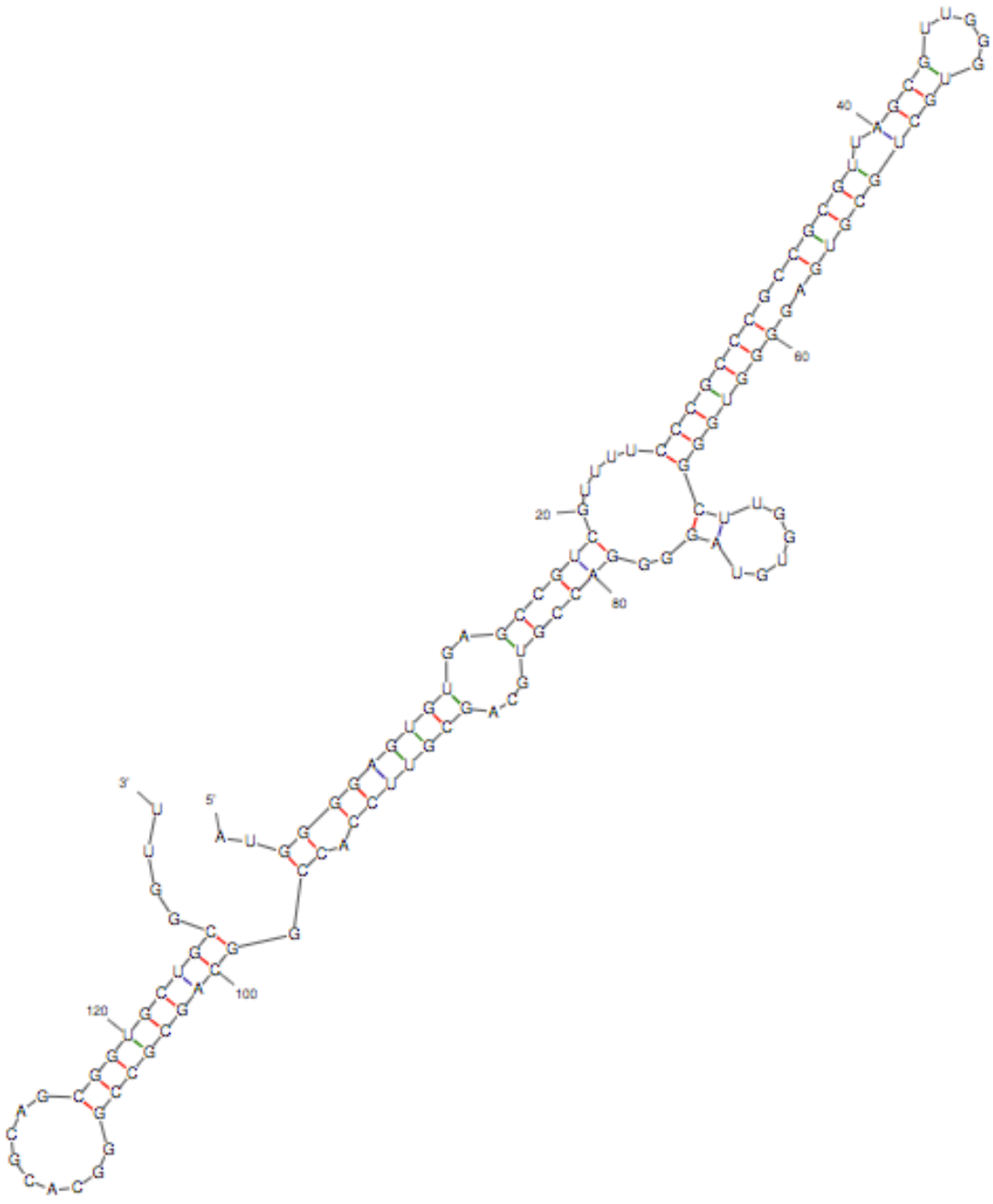


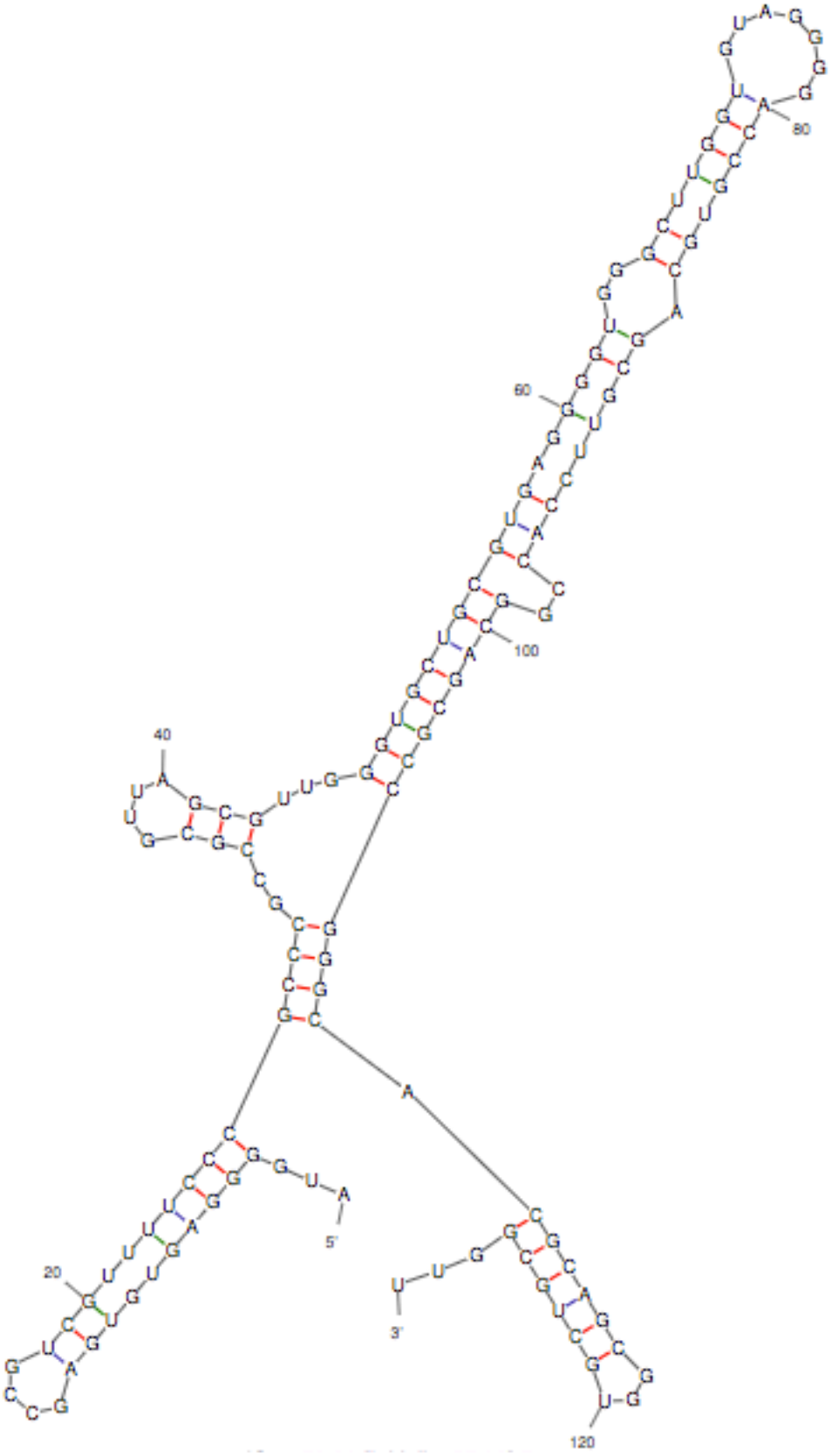

Gallus ESKFLVEHAD AEDESSITKL SKTPPTTKFP DSPERPTYLR EAGKKPAVTP
Mus ENHFLVDNAP KQEQVLETKI PN--GVSSNP DLPAHPDAHL EIHRKKAVTP
Gorilla ESHFLIENSS KEERSLETEI TN--GINLNS DTAEHPGTRL QTHRKTVVTP
Homo ESHFLIENST KEERSLETEI TN--GINLNS DTAEHPGTRL QTHRKTVVTP
Pan ESHFLIENSS KEERSLETEI TN--GINLNS DTAEHPGTRL QTHRKTVVTP
Gallus LDINRFQSKA FQAVISHHIG KATTSPPPRG LLQKEPSLHL DTPSKFLSKD
Mus LDMNRFQSRA FRAVISQQRG QDPTFPPVRQ PLTKEPSLHL DTPSKFLSKD
Gorilla LDMNRFQSKA FRAVVSQHAR QVAASPAVGQ PLQKEPSLHL DTPSKFLSKD
Homo LDMNRFQSKA FRAVVSQHAR QVAASPAVGQ PLQKEPSLHL DTPSKFLSKD
Pan LDMNRFQSKA FRAVVSQHAR QVAASPAVGQ PLQKEPSLHL DTPSKFLSKD
Gallus KLFKPSFDVK DALAALETPG GLGQKNRKLS TPLSEVIGRN LKLALANSTR
Mus KLFKPSFDVT DALAALETPN AAS-QKRKLS SPLSEVIVRN LTVALANSSR
Gorilla KLFKPSFDVK DALAALETPG RPSQQKRKLS TPLSEVIVRN LQLALANSSR
Homo KLFKPSFDVK DALAALETPG RPSQQKRKPS TPLSEVIVKN LQLALANSSR
Pan KLFKPSFDVK DALAALETPG RPSQQKRKLS TPLSEVIVKN LQLALANSSR
Gallus NTEALTASPQ LKAAQPEVEE EPKPLAGVVI CVSKKLTKKQ SELNAIAASL
Mus NTDSHSASPQ LKGAHLEEEE TRKPLDSVVV CVSKKLSKKQ SELNGVAASL
Gorilla NAVALSASPQ LKEAQSEKEE APKPLHKVVV CVSKKLGKKQ SELNGIAASL
Homo NAVALSASPQ LKEAQSEKEE APKPLHKVVV CVSKKLSKKQ SELNGIAASL
Pan NAVALSASPQ LKEAQSEKEE APKPLHKVVV CVSKKLGKKQ SELNGIAASL
Gallus GADYRWCFDE TVTHFIYKGG QNDSNKEYRS VKERGIHIVS EHWLLECAQE
Mus GAEYRWSFDE TVTHFIYQGR ANDSNREYKS AKERGVHIVS EHWLLECAQE
Gorilla GADYRWSFDE TVTHFIYQGR PNDTNREYKS VKERGVHIVS EHWLLDCAQE
Homo GADYRWSFDE TVTHFIYQGR PNDTNREYKS VKERGVHIVS EHWLLDCAQE
Pan GADYRWSFDE TVTHFIYQGR PNDTNREYKS VKERGVHIVS EHWLLDCAQE
Gallus YKRLPESLFP HTYNPKMSLD ISAVQDFRPS SSSELPPTGK PAE-ENEIIA
Mus YKHLPESLYP HTYNPKMSLD INTVQDGRLC NSRAPLAVSA SKDDGPDHLS
Gorilla CKHLPESLYP HTYNPKMSLD ISAVQDGRLC NSRLLSAVSA TKDDEPDPLI
Homo CKHLPESLYP HTYNPKMSLD ISAVQDGRLC NSRLLSAVSS TKDDEPDPLI
Pan CKHLPESLYP HTYNPKMSLD ISAVQDGRLC NSRLLSAVSA TKDDEPDPLI
Gallus VDEDDAEDDL TTDQIKETAG TGEEHNGISE SKGVLTQALE MRENFQRQLQ
Mus VEGNET-NTM GTN---DKES PLLNGSGRDD CKGALTQALE MRENFQKQLQ
Gorilla LEENDV-DNM ATN---NKES APSNGNGKND SKGVLTQTLE MRENFQKQLQ
Homo LEENDV-DNM ATN---NKES APSNGSGKND SKGVLTQTLE MRENFQKQLQ
Pan LEENDV-DNM ATN---NKES APSNGSGKND SKGVLTQTLE MRENFQKQLQ
Gallus EIMSATSIVK PQGQRGSLSR NSFDSSPATP DSARSVRNGR SRALEALRQS
Mus EIMSATCIVK TPAQKTCMSR SSCNSASSTP DSARSVRSGR SRVLEALRQS
Gorilla EIMSATSIVK PQGQRTSLSR SGCNSASSTP DSTRSARSGR SRVLEALRQS
Homo EIMSATSIVK PQGQRTSLSR SGCNSASSTP DSTRSARSGR SRVLEALRQS
Pan EIMSATSIVK PQGQRTSLSR SGCNSASSTP DSSRSARSGR SRVLEALRQS
Gallus RQAFTDINTE PSQSEQIIWD DPTAREERAR LVSNFQWPNS PSQYTEQAQS
Mus RQAVPDVNTE PSQNEQIIWD DPTAREERAR LASNLQWPSD PTQHSELQVE
Gorilla RQTVPDVNTE PSQNEQIIWD DPTAREERAR LASNLQWPSC PTQYSELQVD
Homo RQTVPDVNTE PSQNEQIIWD DPTAREERAR LASNLQWPSC PTQYSELQVD
Pan RQTVPDVNTE PSQNEQIIWD DPTAREERAR LASNLQWPSC PTQYSELQVD
Gallus NVNRNTDEST FKGSIADAEI AGIVVPEAGD GDTIEGLKNP VCRDPETPIK
Mus I--KMPDDSP SRKPVYHSEI AEQASC---- --VTQAPGHP GSEEPEPPVA
Gorilla I--QNLEDSP FQKPLRDSEI AKQAVCDPGN IRVTEAPKHP ISEELETPIK
Homo I--QNLEDSP FQKPLHDSEI AKQAVCDPGN IRVTEAPKHP ISEELETPIK
Pan I--QNLEDSP FQKPLHDSEI AKQAVCDPGN IRVTEAPKHP ISEELETPIK
Gallus D-HLIPTPQA PSIAFPLANP PVAPQPKEKA VTEDEKVDEE PEKCRKFQLS
Mus ERPLIPEPQA PAVASPLAKP PVAPQPADKI ETQEET-HRK VKKQYVFQMS
Gorilla DSHLIPTPQA PSIAFPLANP PVAPHPREKI ITIEET-HEE LKKQYIFQLS
Homo DSHLIPTPQA PSIAFPLANP PVAPHPREKI ITIEET-HEE LKKQYIFQLS
Pan DSHLIPTPQA PSIAFPLANP PVAPHPREKI IMIEET-HEE LKKQYIFQLS
Gallus SLNPQERFDY CHLIEELGGI VLEKQCFDPS CTHIVVGHPL RNEKFLASMA
Mus SLNSQERIDY CRLIKDLGGS VIEKQCSDPS CTHMVVGYPL RNEKYLASMA
Gorilla SLNPQERIDY CHLIEKLGGL VIEKQCFDPT CTHIVVGHPL RNEKYLASVA
Homo SLNPQERIDY CHLIEKLGGL VIEKQCFDPT CTHIVVGHPL RNEKYLASVA
Pan SLNPQERIDY CHLIEKLGGL VIEKQCFDPT CTHIVVGHPL RNEKYLASVA
Gallus AGKWVLHRSY LEACRGAGCF VQEEDYEWGS DSILNVLPGI NVNQKKLALA
Mus AGKWVLHRSY LDACKTAGRF VQEEDYEWGS SSILDALPDV TEHQQKLALA
Gorilla AGKWVLHRSY LEACRTAGHF VQEEDYEWGS SSILDVLTGI NVQQRRLALA
Homo AGKWVLHRSY LEACRTAGHF VQEEDYEWGS SSILDVLTGI NVQQRRLALA
Pan AGKWVLHRSY LEACRTAGHF VQEEDYEWGS SSILDVLTGI NVQQRRLALA
Gallus AMRWRKRIHK GRQETGITEG AFSGWKVILN VDHTKEAGFR RLLQSGGAKV
Mus AMRWRKRIQQ -SQESGIVEG AFSGWKAILR VDRPREAGFK RLLQAGGAKV
Gorilla AMRWRKKIQQ -RQESGIVEG AFSGWKVILH VDQSREAGFK RLLQSGGAKV
Homo AMRWRKKIQQ -RQESGIVEG AFSGWKVILH VDQSREAGFK RLLQSGGAKV
Pan AMRWRKKIQQ -RQESGIVEG AFSGWKVILH VDQSREAGFK RLLQSGGAKV
Gallus FSGHSVSFFK EATHLFADFS KLKPDDTRVN IAEAAEQGVN CLKPEYIADF
Mus LSGHPEPLLK DATHLFCDFN KLKPDDCRVF IAEATAQNMV CLKTEYIADY
Gorilla LPGHSVPLFK EATHLFSDLN KLKPDDSGVN IAEAAAQNVY CLRTEYIADY
Homo LPGHSVPLFK EATHLFSDLN KLKPDDSGVN IAEAAAQNVY CLRTEYIADY
Pan LPGHSVPLFK EATHLFSDLN KLKPDDSGVN IAEAAAQNVY CLRTEYIADY
Gallus LIQDPPPPME SYCLPEAERC FQNNKERGTG LSQKRKAPGE MSRVKRSRMH
Mus LMLESPPCAD NYRVSEAALF H--NKKGGPG LPQKRKTPAE -NVVKRPRVH
Gorilla LMQESPPHVE NYCLPEAISF IQNNKELGTG LSQKRKAPTE KNKIKRPRVH
Homo LMQESPPHVE NYCLPEAISF IQNNKELGTG LSQKRKAPTE KNKIKRPRVH
Pan LMQESPPHVE NYCLPEAISF IQNNKELGTG LSQKRKAPTE KNKIKRPRVH

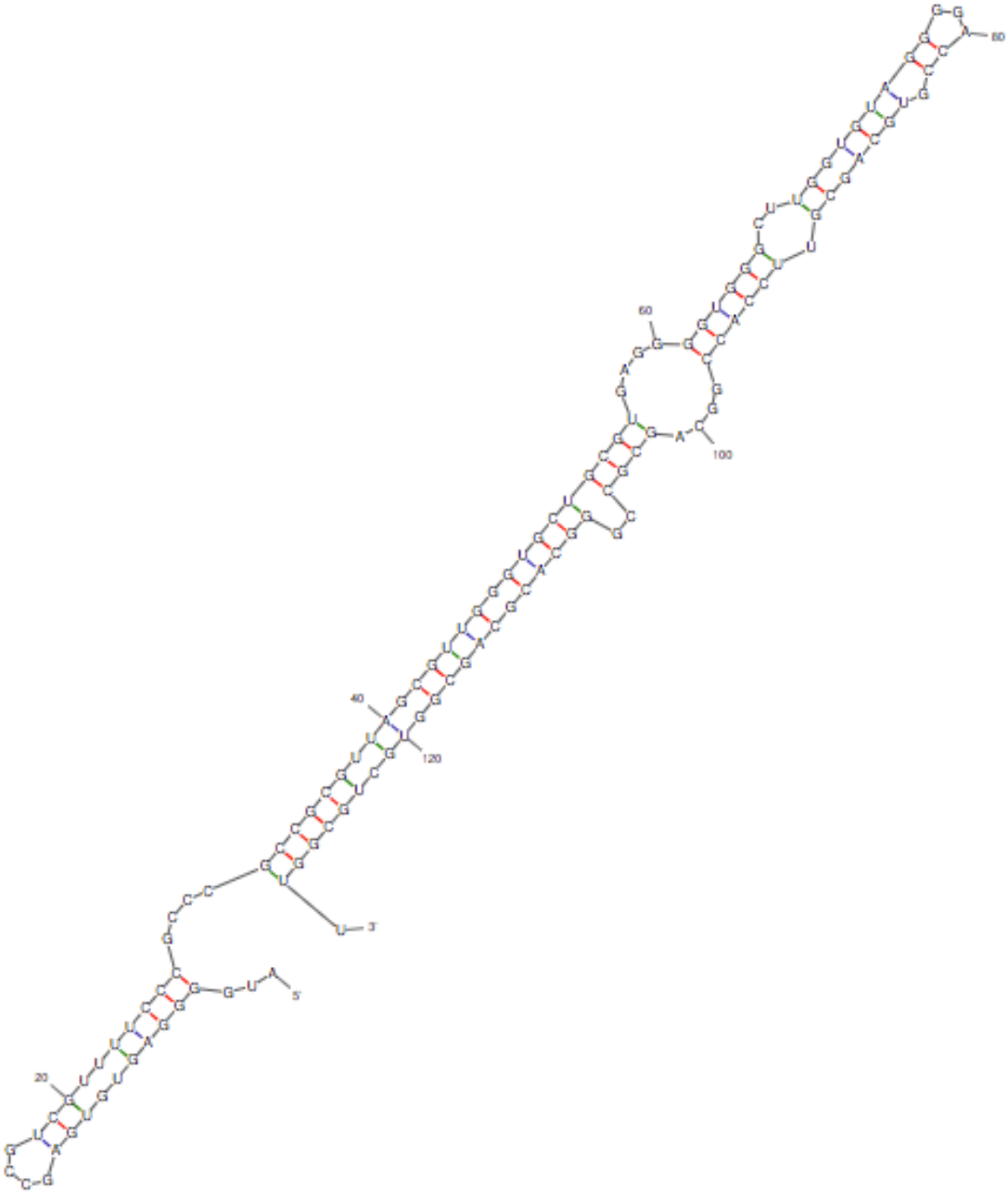
APPENDIX 3

Predicted secondary structures of the new 5' end of the gallus gallus TopBP1 RNA




3'
5'
20
40
60
80
100
120


5'
3'
20
40
60
80
100
120


5'
3'
20
40
60
80
100
120