RESEARCH ARTICLE

The interaction profile of homologous recombination repair proteins RAD51C, RAD51D and XRCC2 as determined by proteomic analysis

Changanamkandath Rajesh[1], Aaron M. Gruver[2], Venkatesha Basrur[3] and Douglas L. Pittman[1]

[1] Department of Pharmaceutical and Biomedical Sciences, South Carolina College of Pharmacy, University of South Carolina, Columbia, SC, USA
[2] Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH, USA
[3] Department of Pathology, University of Michigan, Ann Arbor, MI, USA

The RAD51 family of proteins is involved in homologous recombination (HR) DNA repair and maintaining chromosome integrity. To identify candidates that interact with HR proteins, the mouse RAD51C, RAD51D and XRCC2 proteins were purified using bacterial expression systems and each of them used to co-precipitate interacting partners from mouse embryonic fibroblast cellular extracts. Mass spectroscopic analysis was performed on protein bands obtained after 1-D SDS-PAGE of co-precipitation eluates from cell extracts of mitomycin C treated and untreated mouse embryonic fibroblasts. Profiling of the interacting proteins showed a clear bias toward nucleic acid binding and modification proteins. Interactions of four candidate proteins (SFPQ, NONO, MSH2 and mini chromosome maintenance protein 2) were confirmed by Western blot analysis of co-precipitation eluates and were also verified to form *ex vivo* complexes with RAD51D. Additional interacting proteins were associated with cell division, embryo development, protein and carbohydrate metabolism, cellular trafficking, protein synthesis, modification or folding, and cell structure or motility functions. Results from this study are an important step toward identifying interacting partners of the RAD51 paralogs and understanding the functional diversity of proteins that assist or regulate HR repair mechanisms.

1 Introduction

Homologous recombination (HR) is an error free DNA repair pathway that is involved in the repair of DNA damage generated by double strand breaks (DSB), collapsed repli-

Correspondence: Dr. Douglas L. Pittman, Department of Pharmaceutical and Biomedical Sciences, South Carolina College of Pharmacy, University of South Carolina Campus, 715 Sumter Street, Columbia, SC 29208, USA
E-mail: pittman@cop.sc.edu
Fax: +1-803-777-8356

Abbreviations: DSB, double strand break; **EMSA,** electrophoretic mobility shift assay; **HR,** homologus recombination; **IBs,** inclusion bodies; **ICL,** inter-strand cross link; **MCM,** mini chromosome maintenance; **MEF,** mouse embryonic fibroblast; **MMC,** mitomycin C; **MMR,** mismatch repair; **PER,** protein extraction reagent

cation forks and unresolved strand break intermediates. Through this mechanism, damaged genetic information is restored and genomic instability prevented [1]. During HR, the site of the break is processed into 3'-overhanging ends, and a complex of proteins polymerize along the ssDNA forming a nucleoprotein filament. This complex directs the strand invasion onto a homologous stretch of DNA that serves as a replication template [2]. Key proteins involved in the HR repair reaction are the RAD51 family, which include RAD51, RAD51B, RAD51C, RAD51D, XRCC2 and XRCC3. RAD51 proteins function as genomic caretakers, and over-expression or disruption of their function leads to genomic instability [3, 4].

Targeted gene disruption of the Rad51 paralogs confers embryo lethality likely due to chromosome rearrangements [5–8]. Data from BRCA2 and RAD51 paralog deficient DT40 cells suggest the RAD51 proteins function at an early step

during HR repair with their loss stimulating the alternative error-prone non-homologous end joining pathway [9]. The paralogs also appear to have a role in later steps during HR, contributing to the resolution of Holliday junctions and branch migration [10, 11]. The reduced RAD51 foci formation and severely impaired HR in *Rad51* paralog mutants also suggest an important role for the paralogs in assisting RAD51 during HR repair [12]. Our previous studies with *Rad51d*-deficient mouse embryonic fibroblast (MEF) cell lines ($Rad51d^{-/-} Trp53^{-/-}$) demonstrated that RAD51D is essential for repairing the DNA inter-strand crosslinks (ICL) generated upon treatment with mitomycin C (MMC), and that RAD51D is critical for telomere maintenance [13]. It is becoming increasingly clear that repairing multiple types of DNA damage also requires HR repair proteins [14, 15]. In spite of these evidences in favor of an indispensable role for RAD51 paralogs in DNA repair, there is little data that can assign a specific function at each precise stage of HR. This is mostly due to the lack of knowledge of its various interacting partners that may control and assist the DNA repair functions of the RAD51 family.

Two distinct RAD51 paralog complexes have been identified. One complex consists of RAD51B-RAD51C-RAD51D-XRCC2 (BCDX2) and the second of RAD51C-XRCC3 [16]. The BCDX2 complex binds to ssDNA, performs strand annealing, and also binds to Holliday junctions formed during damage resolution [17]. The RAD51C-XRCC3 complex is suspected to be part of the Holliday junction resolvase complex with additional components [18, 19]. RAD51D has been demonstrated to interact with the N-terminal domain of Bloom helicase, and the RAD51D-XRCC2 complex stimulates Bloom helicase to disrupt Holliday junction structures [20]. Although these two RAD51 paralog protein complexes play roles during HR, it is clear that they do not independently carry out the intricate repair process and require the presence of additional proteins [19]. Therefore, to clearly decipher the mechanism of action of this pathway, it is important to identify and perform functional analyses of other HR components that work in concert to repair these DNA lesions.

Large-scale systematic identification of RAD51 interacting candidate proteins has not been performed. Here, we report an approach to identify RAD51 paralog interacting factors by a high-throughput proteomics screen using purified mouse RAD51C, RAD51D and XRCC2 proteins. This screen was extended to identify the differences in the interaction profiles following treatment of cells with a DNA crosslinking agent, MMC, either due to the increase in the expression levels of proteins or due to the modifications occurring in the proteins following DNA damage. A total of 75–140 interacting proteins were identified with each protein co-precipitation, with more than two-thirds having a role in nucleic acid modification or metabolism. Further dissection of the known functional properties reveal that these nucleic acid associated proteins are involved in DNA repair, ssDNA stabilization or processing, DNA unwinding,

transcription and translation. Four of these interactions, SFPQ, NONO, mini chromosome maintenance (MCM)2 and MSH2 were further verified by Western blot analysis of co-precipitation samples. The presence of the interactions *ex vivo* was confirmed by immunoprecipitation of these proteins in complex with RAD51D in cell extracts. This study will therefore serve to help decipher the coordination of RAD51 paralogs with other proteins within HR as well as identify new roles for these proteins in additional DNA repair pathways or cellular functions.

2 Materials and methods

2.1 Cloning and expression of MmRAD51D, MmRAD51C and MmXRCC2

Full-length *MmRad51d*, *MmRad51c* and *MmXrcc2* cDNA were PCR amplified from pUC19HARAD51D, pcDNA3.1-HARAD51C Hygro(+) and pcDNA 3.1HAXRCC2 Hygro (+), respectively, and cloned in-frame into *Nco I-Sal I* digested pET28b and pET32b vectors (Novagen/EMD biosciences, CA, USA) to create plasmids pET28RAD51D(His), pET28RAD51C(His), pET28XRCC2(His), pET32Trx-RAD51D(His), pET32Trx-RAD51C(His) and pET32TrxXRC-C2(His), respectively. Sequences were confirmed by restriction enzyme analysis and DNA sequencing.

Escherichia coli T7 express cells (New England Biolabs, MA, USA) were transformed with each plasmid and expressed at an A600 nm of 0.6–0.8 at 37°C, 32°C, 25°C and 15°C with 200 μM, 400 μM, 600 μM, 800 μM and 1 mM IPTG for 30 min, 60 min, 90 min and 2 h. Overnight Express induction system 1 (OnEX auto-induction system, Novagen) was also used for expression at the different temperatures. The cell pellet was used for protein extraction using the B-protein extraction reagent (PER) (Pierce Biotech, IL, USA) according to the manufacturer's protocol. The soluble fraction and the pellet, along with the whole cell lysate from above, were analyzed by 10% SDS-PAGE.

2.2 Purification of MmRAD51D, MmRAD51C and MmXRCC2 proteins

pET32Trx-RAD51D(His), pET32Trx-RAD51C(His) and pET28XRCC2(His) were transformed into *E. coli* T7 express cells (New England Biolabs). Proteins were expressed in 250 mL each of OnEX auto-induction system containing ampicillin (100 μg/mL) for 16 h at 32°C, shaking at 200 rpm, and the cells were pelleted at 5000 × *g* for 10 min at 4°C. The pellets were suspended in 10 mL of B-PER-II extraction reagent (Pierce Biotech) and inclusion bodies (IBs) prepared following the manufacturer's protocol. The IBs were suspended in 50 mM CAPS buffer pH 11.0 containing 0.3% *N*-lauroyl sarcosine detergent by gentle pipetting to a final concentration of 10 mg/mL w/v. The solution was incubated

at RT for 15 min and debris removed by centrifugation at $10\,000 \times g$ for 10 min at 4°C.. Clear supernatant was then dialyzed for 18 h in 20 mM Tris-HCl pH 8.0 with two changes of buffer.

The refolded protein was slowly loaded onto a 1 mL Ni-NTA column (Qiagen, CA, USA) at ~1 mL/min. The bound protein was washed with ten-column volumes of 50 mM Tris pH 7.0 containing 150 mM NaCl and 10 mM imidazole and 12-column volumes of 50 mM Tris containing 300 mM NaCl and 25 mM imidazole. The protein was then eluted using 50 mM Tris pH 7.0 containing 300 mM NaCl, 200 mM imidazole and 10% glycerol. Protein samples Trx-MmRAD51D-(His)$_6$ (MmRAD51D, 55 kDa), Trx-MmRAD51C-(His)$_6$ (MmRAD51C, 59 kDa) and MmXRCC2-(His)$_6$ (MmXRCC2, 32 kDa) were analyzed by SDS-PAGE followed by concentration estimation using the BCA protein assay (Pierce Biotech), and protein was stored as aliquots at −20°C.

2.3 ATPase assay

A total of 1 µM MmRAD51D and 500 µM ATP substrate were incubated at 37°C for 1 h in 200 µL of 50 mM Tris-HCl buffer containing 2.5 mM MgCl$_2$, and the free inorganic P$_i$ released was measured by a colorimetric ATPase assay (Innova Biosciences, Cambridge, UK). The P$_i$ concentration was determined using the P$_i$ standard curve. Fifty micromolar nucleotides of single stranded 63 mer oligonucleotide and double stranded annealed oligonucleotides [21] were included as co-factors. The same concentration of trypsinized MmRAD51D protein was used as negative control and purified RecA (New England Biolabs) as positive control. As an additional negative control, protein was purified exactly as described for MmRAD51D from *E. coli* T7 express cells transformed with pET32b vector and the same volume as MmRAD51D used to test for ATPase activity. ATPase inhibitors, including a non-hydrolyzable ATP analog, β-, γ-Methyleneadenosine 5′-triphosphate disodium salt (AMP-PCP, Sigma-Aldrich, MO, USA), and two natural ATPase inhibitor antibiotics, geldanamycin (Invivogen, CA, USA) and radicicol (EMD Biosciences), were included in the reaction mixture to determine inhibition of activity [22].

2.4 DNA binding assay

DNA binding electrophoretic mobility shift assay (EMSA) was performed by mixing 10 µM single stranded 63 mer oligo with increasing concentrations of purified MmRAD51D, MmRAD51C and MmXRCC2 as previously described [21]. The reactions were resolved on a 5% polyacrylamide non-denaturing TBE gel and stained for nucleic acids using SyBr Green Gold stain (Invitrogen, CA, USA). The presence of protein was verified by staining with Sypro Ruby following the manufacturer's protocol (EMSA kit, Invitrogen). Identical concentrations of BSA and trypsinized MmRAD51D were used as negative controls for the DNA binding assays.

2.5 Co-precipitation and Western analysis of proteins from MEF extracts

Wild-type MEF, Rad51d$^{−/−}$ Trp53$^{−/−}$ MEF (MEF168) and Rad51d$^{−/−}$ Trp53$^{−/−}$ HA-MmRAD51d MEF (MEF172AG) cell lines were previously described [23]. Whole cell extract was prepared from each 80–90% confluent plate of cells using 500 µL of Mammalian-(M-PER, Thermo Scientific, IL, USA) containing Protease inhibitor cocktail (Complete Mini, Roche Life Sciences, IN, USA) following the manufacturer's protocol. Five-hundred micrograms of whole cell extract proteins from wild-type MEF cells was incubated with 50 µL of Ni-NTA agarose beads (Qiagen) and incubated at 4°C for 1 h with slow shaking. The beads were pelleted and the pre-cleared supernatant incubated for 16 h at 4°C with 10 µg of purified MmRAD51D, MmRAD51C and MmXRCC2, respectively. The same concentration of whole cell extract without bait protein was used as a control. The bound complexes were co-precipitated using 50 µL of Ni-NTA agarose beads after incubation for 30 min followed by three washes of 20 mM Tris-buffered saline containing 0.05% Tween-20 pH 7.6. The bound proteins were then eluted by boiling in Laemmelli buffer and resolved on a 10% SDS-PAGE gel. The bands detected after staining with Imperial protein stain (Thermo Scientific) were excised for analysis by MS. For MMC-treated samples, cells were treated with 10 µg/mL of MMC for 12 h at 37°C prior to protein extraction. The same samples were transferred onto nitro-cellulose membrane (GE Healthcare, NJ, USA) and subjected to immunoblotting with 1:5000 dilution rabbit polyclonal anti-SFPQ, 1:2000 dilution of goat monoclonal anti-NONO, 1:1000 mouse polyclonal anti-MCM2, 1:500 dilution of goat polyclonal anti-RAD51B (AbCam, MA, USA), 1:100 dilution of rabbit polyclonal anti-MLH1 and 1:2000 mouse polyclonal anti-MSH2 antibody (Santa Cruz Biotechnology, CA, USA), respectively. This was followed by incubation with 1:5000 dilution of the respective HRP-conjugated secondary antibody (Santa Cruz Biotechnology), and the blot finally developed using west pico chemiluminescent substrate (Thermo Scientific).

Confirmatory immunoprecipitations for *ex vivo* complexes were performed using 500 µg of whole cell extracts prepared from MEF cell lines, MEF168 (Rad51d$^{−/−}$ Trp53$^{−/−}$) and MEF172AG (Rad51d$^{−/−}$ Trp53$^{−/−}$ HA-MmRAD51d) [23] incubated with 100 µL of anti-HA sepharose beads (3F10, Roche Biotech) for 16 h at 4°C. The precipitated eluates were used for immunoblotting as described above.

2.6 Identification of proteins by mass spectroscopy (MALDI-TOF and LTQ-MS)

For MALDI-TOF, bands from each lane of MmRAD51D, MmRAD51C and MmXRCC2 eluates in the Coomasie stained gel were excised. Proteins were then trypsinized with 300 ng trypsin (Promega, WI, USA) at 37°C overnight, peptides extracted and dried. Samples were cleaned up

using Zip-Tip C18 (Millipore, MA, USA) and eluted in 50% ACN/0.1% TFA and 0.5 µL was spotted onto a Pre-Spotted Anchor Chip target plate (Bruker Daltonics, MA, USA). A Bruker Daltonics Ultra-flex II TOF-TOF mass spectrometer outfitted with a solid-state smartbeam laser was used for analysis. The samples were run in the delayed extraction, reflectron mode at an accelerating voltage of 25 kV and 10.0 Hz. Spectra were collected from 600–6000 m/z and were summed to 500 laser shots. The data were analyzed using FlexAnalysis 2.4 (Bruker Daltonics), and peptide mass fingerprint searches were run in MASCOT, using a 0.1 Da error tolerant search in the Swiss-Prot database. MS/MS was performed on several intense peaks in the spectrum and analyzed by an MS/MS MASCOT search with an error tolerance of 0.5 Da for fragment ions.

For the LTQ analysis, gel plugs were washed, de-stained using 25 mM ammonium bicarbonate in 50% ACN for 15 min, dehydrated with 100% ACN for 15 min and dried. Each gel plug was trypsinized (Sigma) at 37°C overnight, peptides extracted with one wash of 25 mM ammonium bicarbonate for 20 min and three washes of 5% formic acid, 50% ACN. Each supernatant was collected, pooled and dried to ~2 µL, then reconstituted with 10 µL of 2% ACN and 0.2% formic acid prior to LC/MS analysis. Samples were analyzed *via* LC-ESI-MS/MS on a linear ion trap mass spectrometer (LTQ, Thermo Finnigan, TX, USA) coupled to an LC Packings nano LC system. A 75 µ C-18 reversed phase LC column (Micro-Tech Scientific, CA, USA) was utilized with a 30 min gradient from 2% ACN, 0.2% formic acid to 70% ACN, 0.2% formic acid. Data Dependant Analysis was utilized on the LTQ to perform MS/MS on all ions above an ion count of 1000. Dynamic Exclusion is set to exclude ions from MS/MS selection for 3 min after being selected two times in a 30 s window. The MS/MS data were searched against the NCBI mouse database utilizing Thermo Finnigan Bioworks 3.3 software.

3 Results

3.1 Expression and purification of mouse RAD51D, RAD51C and XRCC2 proteins

MmRAD51D, MmRAD51C and MmXRCC2 proteins were purified from a bacterial expression system and used to precipitate interacting partners from MEF whole cell protein extracts. The entire coding region of each gene was cloned into pET vectors, pET28 and pET32, and expression confirmed. The protein expression was performed at temperatures as low as 15°C using IPTG concentrations from 0.2 to 1 mM for induction times varying from 30 min to overnight in order to improve the amount of soluble protein (See Section 2). However, the amount of soluble protein was less than 5% of the total amount of induced protein under all conditions tested (Fig. 1A). This is consistent with previous reports describing expression and purification of these paralogs [20, 21]. The optimal expres-

sion of MmRAD51D and MmRAD51C in the pET32 vector and MmXRCC2 in the pET28 vector was obtained by overnight induction at 32°C (Fig. 1A). Owing to the low concentration of soluble protein, each protein was purified from IBs by *in vitro* refolding. Solubilization was performed under mild denaturing conditions of alkaline pH in the presence of 0.3% detergent, *N*-lauroyl sarcosine. The solubilized protein was slowly allowed to refold by overnight dialysis against Tris buffer pH 8.0 and purified on a Ni-NTA agarose column. This purification protocol yielded higher protein levels than previously reported from soluble fractions: ~23.5 mg of MmRAD51D from a 500 mL culture, 12 mg of MmRAD51C from a 250 mL culture, and 6 mg of MmXRCC2 from a 250 mL culture (Fig. 1B).

3.2 Functional analysis of purified proteins

To verify proper folding of the purified proteins, the ATPase activity of the *E. coli* purified MmRAD51D protein was measured (Fig. 2A). RAD51D has two ATPase motifs required for resistance to DNA ICL agents [24, 25]. Human
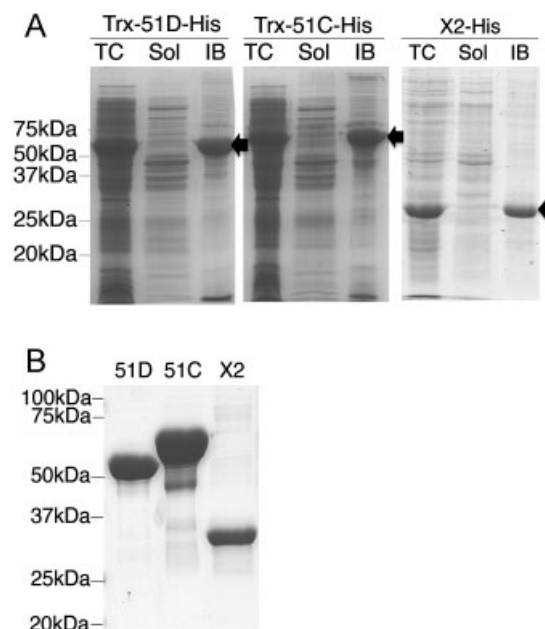
Figure 1. Expression and purification of MmRAD51D, MmRAD51C and MmXRCC2. (A) Coomassie Blue stained 10% SDS-PAGE gels of extracts from *E. coli* expressing TrxMmRAD51D(His) (Trx-51D-His), TrxMmRAD51C(His) (Trx-51C-His) and MmXRCC2(His) (X2-His) proteins, respectively. TC – total cell extract, Sol – soluble fraction, IB – insoluble inclusion bodies. The respective induced bands of epitope tagged MmRAD51D (55 kDa), MmRAD51C (~59 KDa) and MmXRCC2 (~32 kDa) are indicated by the arrows and molecular weight markers (kDa) are shown on the left. (B) Coomassie Blue stained 10% SDS gel of the pooled fractions of Nickel-NTA agarose purified MmRAD51D (51D), MmRAD51C (51C) and MmXRCC2 (X2) proteins. The sizes of molecular weight standards are indicated on the left.

RAD51D was also demonstrated previously to have moderate ATPase activity that increased in the presence of a ssDNA co-factor [10]. In the present study, similar ATPase activity was observed for purified mouse RAD51D (7 µM/h) with a 1.5-fold increase upon addition of a single stranded oligonucleotide (10.5 µM/h) (Fig. 2A). The ATPase activity of MmRAD51D in the presence of ssDNA was similar to that of the RecA protein without addition of a DNA co-factor. To rule-out the possibility of ATPase activity arising due to co-purified contaminant proteins, E. coli transformed with vector alone was processed as described above, and only substantially low levels of ATPase activity was detected (Fig. 2A, Vector). To determine the specificity of MmRAD51D ATPase activity, inhibitors that prevent ATP binding and hydrolysis were used [22]. The ATPase activity was inhibited in the presence of the non-hydrolysable ATP analog, AMP-PCP, and the natural inhibitors of ATPase activity, geldanamycin and its chemical analog radicicol (Fig. 2B). Therefore, MmRAD51D purified using the in vitro refolding technique retains its ssDNA binding stimulated ATPase activity, indicating proper folding.

MmRAD51D, MmRad51C and MmXRCC2 were also verified for DNA binding activity. Human RAD51D [11] and RAD51C [21] bind to ssDNA, 3′-tailed DNA and dsDNA with different affinities, and each was found to bind most strongly to ssDNA. However, human XRCC2 may have substantially less DNA binding activity [26]. In this study, ssDNA binding activity of MmRAD51D, MmRAD51C and MmXRCC2 was determined by incubating increasing concentration of each protein with a 50 bp single stranded oligonucleotide. As shown in Fig. 2C, the three purified proteins bound to the oligonucleotide and formed an aggregate-DNA complex. Purified human RAD51C was previously reported to bind DNA as a multimeric protein-DNA complex and an aggregate complex [20]. Similar aggregate-DNA complexes predominantly formed with all three paralogs, while lower molecular weight complexes formed at higher concentrations of MmRAD51C and MmRAD51D (>1 µM). DNA binding activity of MmXRCC2 was detectable but appears to be less than the other paralogs. Therefore, DNA binding was observed for all three paralogs at the lowest protein concentration tested (500 nM) with a proportional increase for MmRAD51C and MmRAD51D with increasing protein concentration. There were no protein-DNA complexes observed in the negative controls using the same protein concentrations of BSA and inactivated MmRAD51D. Post-staining with Sypro-Ruby verified the presence of proteins in each lane (result not shown). These results suggest that the purified proteins retain their proper folding and DNA binding activity.
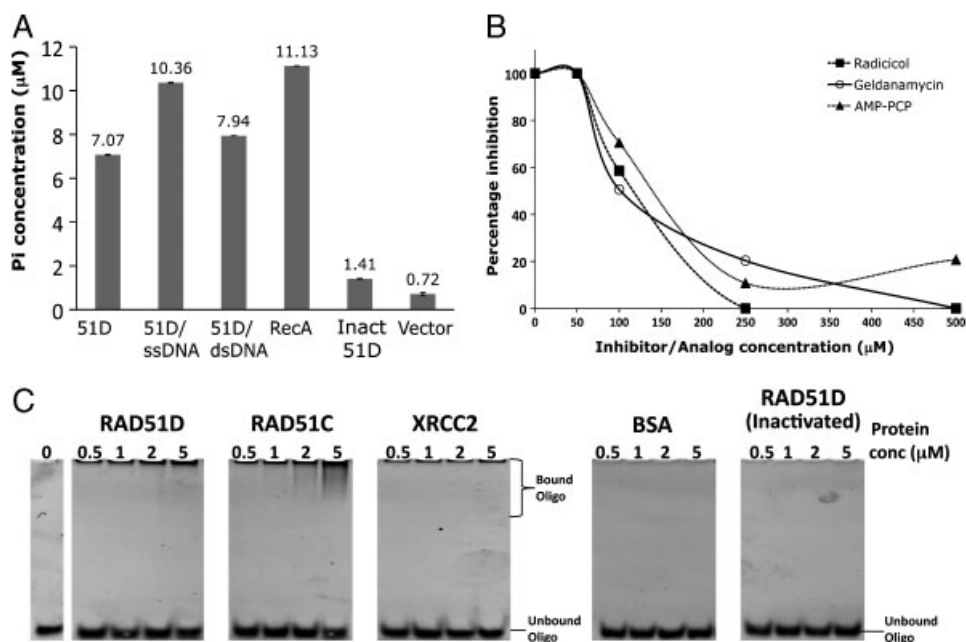

Figure 2. Functional analysis of purified proteins. (A) ATPase activity of purified MmRAD51D. The ATP hydrolysis was measured by the amount of inorganic phosphate (P_i concentration, µM) released upon incubation with 1 µM of MmRAD51D in 1 h. There was a 1.5-fold increase in ATP hydrolysis when single stranded oligonucleotide (51D/ssDNA) was used as a co-factor. Inactivated MmRAD51D Protein (Inact 51D) and identically purified vector transformed cell extract (Vector) were used as negative controls. (B) Inhibition of MmRAD51D ATPase activity. MmRAD51D ATP hydrolysis activity was inhibited upon addition of a non-hydrolyzable ATP analog, AMP-PCP, and ATPase natural product inhibitors, geldanamycin and radicicol. (C) DNA binding activity of the purified proteins. Electrophoretic mobility shift assay analysis was performed by incubation of 10 µM of single stranded oligonucleotide with the indicated amounts of purified MmRAD51D (RAD51D), MmRAD51C (RAD51C) and MmXRCC2 (XRCC2), respectively. The bound protein-DNA complex (bound oligo) was seen for all the three proteins tested, and no bound complexes were observed with negative controls BSA and inactivated MmRAD51D.

3.3 Mass spectroscopic identification of interacting proteins

Each of the (His)$_6$-tagged proteins, MmRAD51D, MmRAD51C and MmXRCC2, were mixed separately with 500 μg of pre-cleared MEF whole cell extract. The protein complexes bound to the His-tagged RAD51 paralogs were co-precipitated from the MEF extract by addition of Ni-NTA agarose beads to create MmRAD51D, MmRAD51C and MmXRCC2 "affinity matrices" (Fig. 3A). After washing off unbound proteins, the complexes were eluted from the beads, resolved by 1-D SDS-PAGE and visualized by silver staining and Coomassie Blue staining (Fig. 3B).

We have previously described MEFs that do not express RAD51D, and hence are deficient in HR as well as highly sensitive (up to 17.6-fold) to the DNA ICL agent MMC [23]. To potentially identify differences in the interaction profiles that could arise upon treatment of cells with MMC, MEFs were treated with 10 μg/mL of MMC for 12 h. Whole cell extracts were prepared from these cells and used for co-precipitation as described above (Fig. 3B, MMC-treated). Using this approach, we expected to identify variations in the binding pattern of proteins with the RAD51 paralogs upon DNA damage that could lead to identification of new candidates.

The protein bands from each co-precipitation lane of Coomassie stained gels were excised, trypsinized and the peptides subjected to MS analysis on Ultra-flex II TOF-TOF MS. The bands corresponding to those seen clearly in the control lane (Fig. 3B, control) were avoided during excision and not used for MS analysis to minimize the number of non-specific protein interactions detected. Protein identity was matched from the monoisotopic peaks obtained, with a $p < 0.05$, which suggests identity or extensive homology of a protein. If identity was not conclusive by the MS, an MS/MS tandem spectroscopy analysis was performed on intense peaks. The sequences obtained were then BLASTed to the NCBI, Swiss-Prot and TrEMBL protein databases. Twenty to twenty-five proteins were identified from each co-precipitation (MALDI-TOF proteins in Supporting Information table). Several low intensity peaks were detected, but the sequence information could not be verified by MS/MS to conclusively identify each protein (results not shown). For example, five peaks obtained for one band analyzed for MmRAD51D co-precipitation matched that of XRCC2 (part of BCDX2 complex involving RAD51D and XRCC2) [16]. Further analysis of the *m/z* values of these peaks using the MS Fit algorithm (PROTEIN PROSPECTOR at www.prospector.ucsf.edu) yielded XRCC2 as one of the most probable hits. This indicates that although XRCC2 might be present, the conclusive sequence identification required a more sensitive MS instrument.

Therefore, the co-precipitation bands were analyzed via more sensitive LC-ESI-MS/MS on a linear ion trap mass spectrometer (LTQ). Protein identifications chosen were those that meet the minimum criteria of a Protein Prob-ability of 1.0 E-3 or greater and have an Xcorr *versus* charge state > 1.5, 2.0, or 2.5 for +1, +2, and +3 ions, respectively, with at least two unique peptides matching the protein. Another selection criterion employed was a good match for at least four consecutive *y* or *b* ion series from the MS/MS spectra (Supporting Information table). Using this, interacting partners of the RAD51B-RAD51C-RAD51D-XRCC2 (BCDX2) complex were identified, which further validated the approach (Tables 1–3). Eighty-five proteins were identified with MmRAD51D, 140 proteins with MmRAD51C, and 131 proteins with MmXRCC2 co-precipitations, respectively (Supporting Information Table). Protein overlaps of the same protein were occasionally detected in multiple excised bands of different molecular sizes for each co-precipitation, (Supporting Information data) most likely due to modified forms of the protein or carry over from excision of the closely located bands on the acrylamide gel. For example, Elongation factor Tu, (110 kDa) and NEDD4 (103 kDa) proteins were detected in two different band sizes screened for MmRAD51D precipitations, 100–125 kDa and 125–140 kDa for (Supporting Information table). The MS analysis of blanks between samples was always clean and did not give any peptide hits.

3.4 Functional profiling of the interacting partners

The proteins from the MmRAD51D, MmRAD51C and MmXRCC2 pulldowns, respectively, were assigned broad classification based upon known or predicted functions from the Swiss-Prot and TrEMBL database entry, functional annotations using Gene Ontology at MGI database, and NCBI data entries. Proteins with unknown functions were omitted from this classification, and proteins with multiple functions were assigned to both groups. As predicted by the known functions of these three proteins in DNA repair, profiling of interacting proteins showed a clear bias toward DNA/RNA binding and modification proteins, constituting more than one-third (33–47%) of precipitated proteins in all the co-precipitations (Fig. 3C). Other proteins identified could be functionally classified to be involved in protein synthesis/folding and modification (14–20%), cell division and development (3–14%), cellular trafficking (3–9%), amino acid (1–4%) and carbohydrate (9–11%) metabolism, cell signaling (1–3%) and cytoskeletal proteins (10–22%) for each co-precipitation. A summary of the selected proteins is listed in Table 1 (MmRAD51D), Table 2 (MmRAD51C) and Table 3 (MmXRCC2). Further classification of the DNA/RNA binding/modification proteins shows that a majority of proteins in this group comprise ssDNA/RNA processing proteins (>25%) and helicases (~10%), both important groups of proteins that may conceptually assist the RAD51 paralogs in their HR DNA repair functions (Fig. 3D). Additional functional groups include DNA damage repair, RNA stabilization, nucleotide biosynthesis, nuclease, and transcription and translation proteins (Table 1–3).
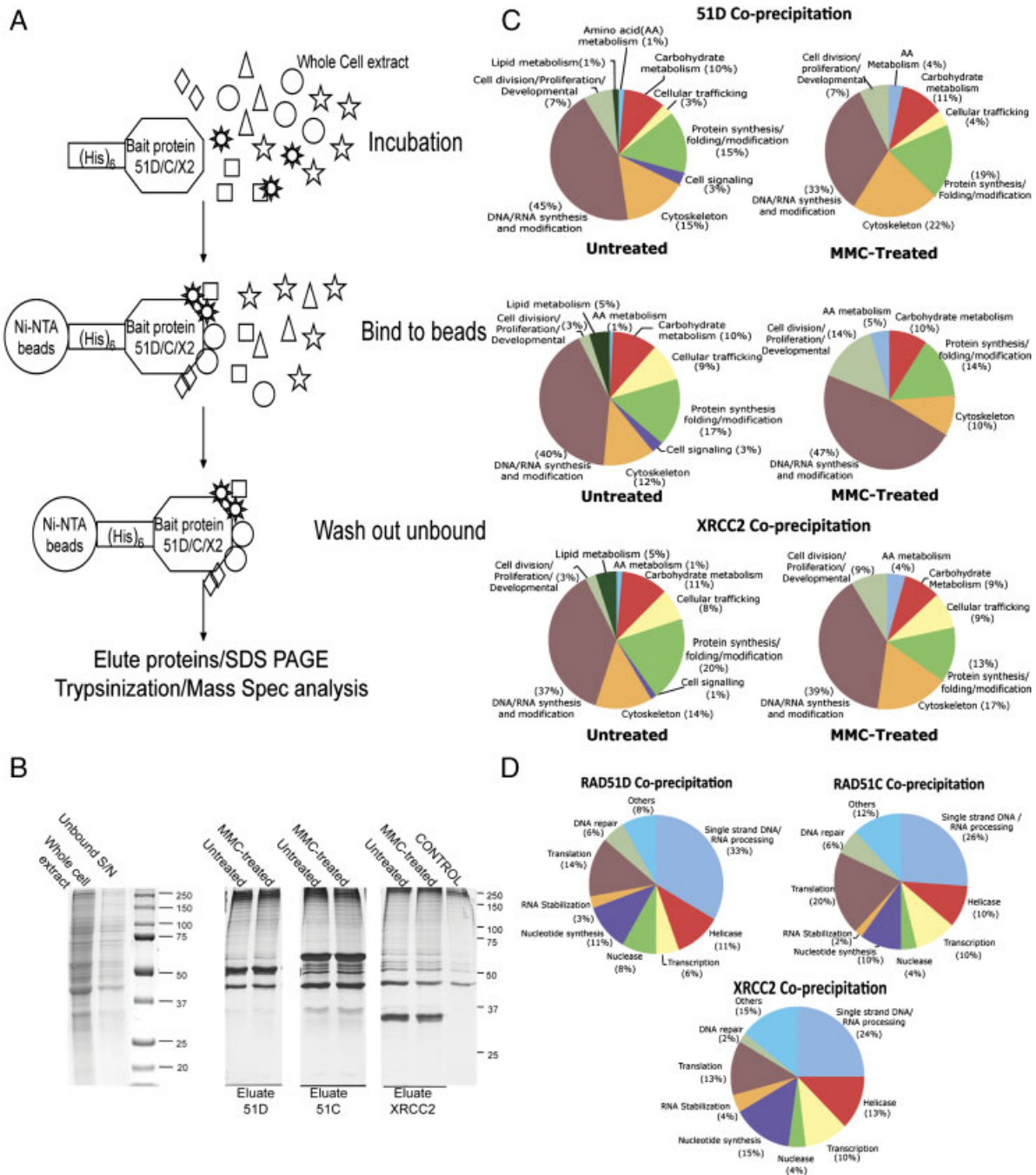
Figure 3. Co-precipitation of MmRAD51D, MmRAD51C and MmXRCC2 binding partners. (A) Schematic representation of co-precipitation procedure used to identify interacting partners of MmRAD51D (51D), MmRAD51C (51C) and MmXRCC2 (X2) as described in the methods section. $(His)_6$-tagged bait protein 51D or 51C or X2 used separately, binds different proteins in the whole cell MEF extract (represented by different symbols). Following overnight incubation, the unbound proteins are washed away and the bound complexes eluted from the matrix and resolved on 1-D SDS PAGE as indicated. (B) Coomassie stained gel of the whole cell extract and unbound supernatant (Unbound S/N) are shown on the left. A Silver stained 10% SDS-PAGE gel (right) showing the eluate fractions of untreated MEFs extracts (Untreated) and MEFs treated with MMC (MMC-treated) for 51D (Eluate 51D), 51C (Eluate 51C) and XRCC2 (Eluate XRCC2). CONTROL-Eluate fractions of co-precipitation without any bait protein. Bands visualized on each eluate lane were excised, trypsinized, extracted and analyzed by mass spectroscopy. The molecular size is indicated to the right of the gels. (C) Functional profiles of all interacting partners identified by the proteomic screen. Distribution in broad functional clusters of all the RAD51 paralog binding proteins identified was based upon known or predicted functions from the Swiss-Prot and TrEMBL database entry and functional annotations using Gene Ontology at MGI database and NCBI data entries. Numbers represent the percentage of the total in each specific functional group. (D) Functional distribution of nucleic acid binding proteins. Further classification of a more specific role for each protein in the nucleic acid binding and metabolism protein category of section C.

Table 1. Selected proteins that co-precipitated with MmRAD51D

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
DNA Repair							
RAD51D	14	90.32	1.17E-13	RAD51D	14	116.31	2.21E-12
RAD51C	9	90.27	3.42E-09	NONO	11	88.21	1.07E-10
NONO	4	40.16	3.47E-08	SFPQ	10	84.23	6.21E-08
Ribosomal protein S3	3	20.28	7.48E-10	Ribosomal protein S3	8	70.22	1.23E-09
SFPQ (MALDI-TOF: score-41; seq. coverage-14%)							
FANCA (MALDI-TOF: score-17; seq. coverage-6%)							
ssDNA/RNA processing							
G protein/p205	5	50.23	5.48E-09	NONO	11	88.21	1.07E-10
Splicing Factor 3b	5	50.23	2.47E-09	SFPQ	10	84.23	6.21E-08
NONO	4	40.16	3.47E-08	G protein/p205	6	60.22	1.54E-07
PolyA binding protein	4	40.26	1.73E-10	Splicing factor 3b	5	50.21	6.05E-11
hnRNPK	3	30.29	7.19E-11	hnRNPK	2	20.16	9.13E-09
hnRNPL	2	20.29	1.83E-10	RNA 3-terminal phosphate cyclase-like protein (MALDI-			
ELAV	2	20.19	1.38E-04	TOF: score-32; seq. coverage-21%)			
SFPQ (MALDI-TOF: score-41; seq. coverage-14%)							
Helicase							
DEAD box polypeptide 17	2	16.22	2.16E-09	DHX15	6	60.22	5.26E-08
DEAD/H box polypeptide 3	2	20.2	3.40E-08	DDX54 (MALDI-TOF:score-39; seq. coverage-8%)			
DHX15 (MALDI-TOF: score-18;seq. coverage-13%)				DDX58(MALDI-TOF:score-39; seq. coverage-11%)			
Nucleotide synthesis							
Inosine 5′-phosphate dehydrogenase	3	30.35	1.05E-09				
GMP synthetase	2	20.23	1.51E-05				
Transcription							
IFIT-1	7	70.23	3.48E-12	Elongation factor Tu	7	60.24	4.50E-11
IQ motif containing GTPase activating prot. 1	6	60.18	6.97E-07	Transducer-like enhancer protein	2	20.17	3.95E-11
Elongation factor Tu	2	20.20	6.13E-07	Suppressor of tumorigenicity protein 18 (MALDI-TOF: seq			
TIF1-β/KRIP1	2	18.14	1.36E-05	coverage-6%)			
Cell division/proliferation/developmental							
NEDD4	4	36.16	8.75E-05	MCM3	2	20.14	1.53E-04
Ribosomal protein S6	2	20.28	7.48E-10	MCM5	2	20.15	5.70E-05
ELAV	2	20.19	1.38E-04	Stress-70 protein/PBP74	2	20.19	1.42E-10
Stress-70 protein/PBP74	2	20.16	1.16E-06	Tripartite motif protein 16	2	20.12	1.34E-04
MCM2 (MALDI-TOF: score-27; seq. coverage-11%)				Gtf2ird1 (MALDI-TOF: seq. coverage-1%)			
Cellular trafficking							
HSP 60	6	60.28	1.43E-10	ADP/ATP translocase 2	4	40.18	6.08E-07
Lectin, mannose binding	5	50.27	4.66E-13	Lectin, mannose binding	2	20.2	8.88E-15
Cell signaling							
IQ motif containing GTPase activating prot. 1	6	60.18	6.97E-07	G protein/p205	6	60.22	1.54E-07
G protein/p205	5	50.23	5.48E-09				
Protein synthesis/folding/modification							
ERp72	15	80.30	7.58E-12	HSP70 protein5/BiP	12	114.28	1.89E-14
HSP70 protein5/BiP	13	110.2	4.83E-10	Serpin H1 precursor/HSP47	9	80.21	2.41E-08
Serpin H1 precursor/HSP47	11	70.27	1.55E-13	DnaJ/HSP40	4	34.18	4.83E-07
Elongation factor-α1	6	50.18	4.44E-06	Elongation factor-α1	3	30.15	3.26E-04
40S ribosomal prot. S3a	4	40.21	9.73E-07	ERp44	3	20.12	2.85E-04
TCP-1 complex proteins	5	40.21	2.42E-09	TCP-1 complex protein	3	30.23	2.36E-10
ERO1-La	4	40.25	6.12E-09	ERp72	3	30.25	6.69E-07

Table 1. Continued

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
Ribophorin I/RPN-I	2	20.19	1.59E-07	AMPK-α1	2	20.14	9.60E-06
NEDD4	4	36.16	8.75E-05				
Amino acid/carbohydrate metabolism							
δ-Pyrroline-5-carboxylate synthetase	12	100.3	1.11E-14	δ-pyrroline-5-carboxylate synthetase	11	90.29	9.71E-10
GAPDH	11	70.38	9.99E-15				
Lactate dehydrogenase	10	98.23	3.98E-10	Pyruvate kinase M	6	50.22	3.46E-08
Transketolase	10	30.29	2.33E-14	Lactate dehydrogenase	4	40.26	1.11E-15
					2	20.11	1.60E-05
Cytoskeleton							
Myosin	17	160.3	9.99E-15	Vimentin	21	166.26	3.16E-09
Fibronectin	10	100.3	1.36E-13	Myosin	17	140.27	7.63E-12
Filamin	9	90.24	1.76E-07	Palladin	4	30.24	5.56E-10
Plectin 1	9	90.2	2.07E-09	α-Tubulin	4	30.28	7.68E-10
α-Tubulin	8	90.24	1.12E-10	Gelsolin	3	20.29	1.27E-07
Vimentin	6	60.25	2.65E-11				

(Complete list in Supporting Information table RAD51D).

3.5 Confirming binding partners by Western analysis and *ex vivo* immunoprecipitation

In addition to identification of known binding partners from the BCDX2 complex, a set of several new interacting partners of the RAD51 paralogs was identified. Antibodies against four novel interacting partners were selected to validate their presence in the co-precipitation samples: SFPQ/PSF (splicing factor proline and glutamate-rich/PTB associated splicing factor), NONO (Non POU domain containing octamer binding protein/p54^{nrb}), DNA replication licensing factor MCM2 and DNA mismatch repair (MMR) protein MSH2. Each was confirmed to be present in all three co-precipitations with MmRAD51D, MmRAD51C and MmXRCC2, respectively (Fig. 4A). The MSH2 concentration, based on the band intensity, was low in MmRAD51D co-precipitation eluate lane when compared with that of MmRAD51C and MmXRCC2 suggesting that the interaction of MSH2 with the RAD51D may not be direct. In the control lanes, no enrichment of the target proteins is observed. However, each of the four proteins was enriched in the co-precipitation lanes of all the three paralogs, suggesting the validity and specificity of these interactions. RAD51B, another RAD51 paralog that is part of the BCDX2 complex and directly interacts with RAD51C, was used as a positive control. The MMR protein MLH1 that was not identified in the screen was used as a negative control. RAD51B was present in the MmRAD51C co-precipitation eluate consistent with their direct interaction. RAD51B was not detected in the MmRAD51D and MmXRCC2 eluates with which it does not interact directly [16]. However, RAD51B was identified in the mass spectroscopy analysis from the MmXRCC2 co-precipitation eluates. The MLH1

protein was not detected in any of the co-precipitation eluates, supporting the specificity of the interactions identified using this approach.

To further determine the specificity of the newly identified binding partners, *ex vivo* interaction of these same candidates were investigated in MEF cell lines using *Rad51d*-deficient MEFs stably expressing an epitope tagged RAD51D. The expression of full length RAD51D in these cells was previously determined to complement the cell-sensitivity phenotype of *Rad51d*-deficient cells to DNA damaging agents including MMC [23]. Immunoprecipitations using anti-HA matrix was performed using whole cell extracts from the stably transfected HA-RAD51D expressing and *Rad51d*-deficient cell lines. The four proteins SFPQ, NONO, MCM2 and MSH2 were each co-precipitated in eluates from the complemented cell lines only expressing HA-RAD51D and not from cells deficient for RAD51D (Fig. 4B). The proteins were co-precipitated both in the MMC-treated and untreated cell extracts, and there were no substantial differences between either eluates. Although the band intensity of MSH2 protein may be increased in eluates from extracts treated with MMC (Fig. 4B, MSH2) supporting the MS identification of MSH2 only in MMC-treated co-precipitation eluates (Table 2). The MCM2 and MSH2 protein bands in the co-precipitation eluate also appear to be less enriched, suggesting that their interaction with RAD51D might be indirect or mediated by the other paralogs. Further investigation is being performed to decipher the direct interaction patterns of each of these proteins and to study the variation of expression and interaction upon treatment with DNA damaging agents. The results presented in this section further validate the specificity of the candidates identified by the proteomic screen.

Table 2. Selected proteins that co-precipitated with MmRAD51C

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
DNA Repair							
RAD51C	18	130.31	1.22E-14	RAD51C	20	118.27	7.79E-12
XRCC2	5	50.81	3.49E-06	Ribosomal protein S3	9	80.22	2.43E-10
Ribosomal protein S3	5	50.18	1.74E-11	SFPQ	6	54.23	4.20E-09
Pigpen protein	2	20.21	3.09E-06	MSH2 (MALDI-TOF: score:20; seq coverage: 10%)			
				MSH6 (MALDI-TOF: score:19; seq coverage: 4%)			
ssDNA/RNA processing							
G protein/p205	9	70.23	6.90E-10	G protein/p205	6	60.23	1.55E-08
IGF2 mRNA binding protein 2	9	50.19	3.19E-08	SFPQ	6	54.23	4.20E-09
Ribosomal protein S6	5	40.28	1.49E-09	Splicing factor 3a	3	28.21	1.64E-06
FUSE binding protein 3	4	40.18	1.26E-06				
hnRNPA3	4	40.26	2.00E-14				
EMAP-II/p43	4	40.21	6.07E-08				
PolyA binding protein	4	40.25	2.38E-09				
Poly(rC) binding protein	3	30.24	6.52E-11				
Y-box protein 1	3	30.39	4.44E-15				
Helicase							
DDX5	3	20.14	5.99E-04	DHX15	6	60.21	7.38E-06
DDX1	3	20.19	1.58E-05	DEAD/H box polypeptide 3 (MALDI-TOF MS/MS:			
DHX15	2	20.18	1.24E-05	No. of assigned peaks- 36; score-38.72)			
Nucleotide synthesis							
GMP synthetase	4	40.18	1.24E-07	ATP synthase	8	70.26	1.32E-10
CTP synthetase	3	30.21	4.73E-10	Nucleoside diphosp. Kinase	3	28.18	5.50E-07
Transcription							
STAT 1	6	58.26	1.52E-12	Elongation factor Tu	7	70.24	5.92E-13
Myb-binding prot. 1a (P160)	6	60.22	8.94E-10				
TIF1-β/KRIP1	4	40.17	2.74E-11				
Elongation factor Tu	5	50.2	1.83E-06				
Nucleolin	2	20.22	8.29E-08				
Cell division/proliferation/developmental							
Vcp97	9	90.28	2.95E-10	Stress-70 protein/PBP74	7	70.12	1.43E-10
NEDD4	6	58.2	4.22E-09	Src substrate cortactin	6	50.16	1.13E-08
GTPase activating protein 1	6	60.27	3.46E-12	NEDD4	3	30.2	2.37E-07
Ribosomal protein S6	5	40.28	1.49E-09	MCM6	2	18.15	2.54E-06
FUSE binding protein 3	4	40.18	1.26E-06	Vcp97	2	20.13	4.21E-04
Desmoyokin	4	30.19	4.38E-07	SET binding protein (MALDI-TOF MS/MS:			
Nestin	3	30.26	3.90E-08	No. of assigned peaks-16; score-9.21)			
ELAV	3	20.14	3.94E-07				
Debrin A	2	20.17	3.18E-08				
Nucleolin	2	20.22	8.29E-08				
Cellular trafficking							
GPI anchored protein p137	4	30.15	3.29E-07	Solute carrier family 23	3	30.17	2.79E-07
Solute carrier family 23	3	30.17	5.71E-07	Vesicle docking protein	3	30.21	2.31E-09
SEC23A	3	30.2	3.43E-08				
Karyopherin-β1	3	30.18	8.09E-07				
ERGIC 1	2	20.18	5.59E-07				
Cell signaling							
G protein/p205	9	70.23	6.90E-10	G protein/p205	6	60.23	1.55E-08
Guanine nucleotide binding protein G(i)	3	30.22	2.94E-07	IQ motif containing GTPase activating protein 1	2	20.16	2.94E-04

Table 2. Continued

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
Protein synthesis/folding/modification							
ERP72	22	210.29	8.90E-12	HSP70 protein5/BiP	22	160.29	2.54E-11
Serpin H1 precursor/HSP47	18	146.24	1.04E-10	Elongation factor 1-α	14	48.18	3.36E-06
HSP 90-β	17	122.43	3.47E-11	HSP 90-β	11	84.26	1.18E-11
HSP70 protein5/BiP	14	140.28	7.02E-13	Serpin H1 precursor/HSP47	10	70.22	3.28E-12
Calnexin	10	110.23	2.57E-11	ERP72	8	60.26	8.73E-10
NEDD4	6	58.2	4.22E-09	Elongation factor-α1	2	20.18	4.03E-06
Elongation factor-γ1	5	50.21	1.74E-08				
HSP 70.2	4	30.23	9.38E-09				
Ubiquitin-activating enz. E1	4	38.19	2.96E-04				
DnaJ/HSP40	2	20.21	1.27E-09				
Amino acid/carbohydrate metabolism							
GAPDH	22	196.26	5.33E-14	GAPDH	15	102.22	2.20E-08
Aldolase 1	20	118.26	2.66E-14	Pyrroline-5-carboxylate synthetase	10	88.24	1.50E-09
Lactate dehydrogenase	14	128.23	1.68E-11	Glutamate dehydrogenase	8	70.27	1.39E-07
Pyrroline-5-carboxylate synthetase	9	78.28	8.44E-14				
Cytoskeleton							
Plectin 1	50	460.34	8.88E-15	Tubulin	17	158.36	3.79E-13
Myosin	23	200.3	4.52E-15	Vimentin	14	100.28	2.85E-10
Filamin A	20	196.28	1.98E-09	β-actin	10	72.23	9.97E-11
β-actin	10	88.28	5.73E-10	Myosin	7	70.25	5.67E-09
Vimentin	9	58.26	5.67E-10	Gelsolin	2	20.18	1.09E-07

(Complete list in Supporting Information table RAD51C).

4 Discussion

RAD51 paralogs act as genomic caretakers by playing a central role during HR repair. They are involved in DNA damage recognition, strand invasion into an undamaged DNA template, branch migration and resolution of Holliday junctions [9, 10, 20]. In spite of their indispensible role during DNA repair, few binding partners of the RAD51 paralogs have been identified. Yeast two-hybrid analysis, co-immuno-precipitation, mutation analysis and domain mapping of the RAD51 protein family have revealed protein-protein interactions among these gene products, but these are insufficient to explain the mechanism of action of these proteins in HR repair [16, 27–31]. In this study, we have used three different RAD51-like proteins, RAD51D, RAD51C and XRCC2, each known to function as a complex and participate in HR repair, for co-precipitation followed by mass spectroscopic analysis of their interacting partners. By this method, new proteins that may assist or regulate the RAD51 paralogs in their function were identified and will likely help to decipher new roles of these important proteins.

The interacting proteins were identified for each paralog from whole cell protein extracts of untreated MEFs and MEFs treated with the DNA crosslinking agent MMC. By using the treated cell extract, we expected to identify additional interactions that arise due to modifications in response to DNA

damage. Eighty-five interacting proteins were identified from RAD51D, 140 proteins from RAD51C and 131 proteins from XRCC2 co-precipitations, respectively. RAD51B, RAD51C and XRCC2, which constitute part of the BCDX2 complex [16], were identified during the screen, which supported the validity of this approach for identifying binding partners. As expected, several of the proteins identified were common to each of the three co-precipitation eluates, since RAD51C, RAD51D and XRCC2 form a complex and are expected to pull down proteins interacting with each other. Therefore, each of the proteins identified will not necessarily directly interact with the RAD51 paralog used as bait and could be part of a complex involving these proteins or through accessory intermediate proteins.

The functional pattern of interacting proteins identified shows a clear bias toward those involved in nucleic acid modification, with several of them having the potential to affect DNA repair and related functions of RAD51 paralogs. Further analyses of the functional profile of the nucleic acid modification proteins suggest that more than one-fourth of these proteins are associated with ssDNA binding and processing function, which could have important roles in the stabilization of ssDNA or in finding and pairing with homologous DNA strands during HR. Another major class of the nucleic acid modification proteins identified were helicases, which could play important roles in relaxing the DNA configuration for HR repair to occur [32, 33]. There

Table 3. Selected proteins that co-precipitated with MmXRCC2

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
DNA Repair							
XRCC2	21	86.28	4.35E-08	XRCC2	17	70.3	4.64E-07
SFPQ	3	30.3	6.03E-13	Ribosomal protein S3	9	60.22	1.26E-07
				SFPQ	8	50.32	4.70E-13
				RAD51B(MALDI-TOF: score-19.8; seq. cov.- 17 %)			
				RAD54(MALDI-TOF: score-23.7; seq. cov.-5.6 %)			
				RAD54B(MALDI-TOF: score-35.5;seq. Cov.-11.1%)			
ssDNA/RNA processing							
Splicing factor 3b	8	78.23	2.09E-11	Splicing factor 3b	12	110.22	1.42E-09
EMAP-II/p43	5	50.22	6.12E-10	G protein/p205	4	40.23	1.58E-07
IGF2 mRNA binding protein 2	4	40.17	8.17E-06	Prohibitin	2	20.18	2.45E-06
PolyA binding protein	3	30.25	3.22E-11	PolyA binding protein	2	18.25	1.57E-09
mFxr1p	3	30.18	5.46E-06				
hnRNPA3	2	18.21	2.11E-07				
Y-box protein 1	2	20.32	4.24E-13				
Nucleolin	2	20.22	8.27E-09				
Splicing factor Prp8	2	20.16	1.77E-06				
Helicase							
DEAD/H box polypeptide 3	3	30.18	3.17E-06	DHX15	5	50.24	3.50E-08
MOV-10 helicase	2	20.2	2.77E-10	DEAD/H box polypeptide 3	2	20.18	1.23E-06
DHX15	2	20.18	1.24E-05	DDX58 (MALDI-TOF); score-23.8; seq. Cov.-12 %			
Nuclear DNA helicase II	2	20.2	2.37E-05	DNA topoisomerase (MALDI-TOF: score-39.8; seq. cov.-20 %)			
Nucleolar RNA helicase II	2	20.16	7.14E-05				
Nucleotide synthesis							
Prolyl tRNA synthase	7	68.26	2.85E-07	Prohibitin	2	20.18	2.45E-06
GMP synthetase	3	28.18	1.49E-04				
Transcription							
STAT 1	6	50.21	3.76E-07	Myb-binding prot. 1a (P160)	2	20.22	1.88E-06
Elongation factor Tu	5	50.2	1.83E-06				
Myb-binding prot. 1a (P160)	5	50.19	2.40E-08				
TIF1-β/KRIP1	3	30.23	2.10E-10				
Cell division/proliferation/developmental							
NEDD4	6	60.22	2.98E-09	Src substrate cortactin	5	48.23	3.03E-06
Stress-70 protein/PBP74	5	50.23	1.03E-08	MCM2	3	30.16	6.92E-05
GTPase activating protein 1	5	50.23	3.05E-08	Stress-70 protein/PBP74	2	20.19	2.29E-05
Nestin	3	28.24	2.20E-06	Prohibitin	2	20.18	2.45E-06
MOV-10 helicase	2	20.2	2.77E-10	Arsenate resistance protein 2	2	20.23	1.42E-07
Nucleolin	2	20.22	8.27E-09	Prohibitin	2	20.18	2.45E-06
Cellular trafficking							
SEC23A	4	40.3	5.28E-12	Lectin, mannose binding	5	50.26	5.61E-13
Vescicular docking protein	4	40.2	5.47E-09	Solute carrier family 25	4	40.15	9.54E-08
Tpr	2	20.15	2.27E-05	SEC23A	3	30.23	1.88E-11
Cell signaling							
Pfetin/KCTD12	2	20.18	4.50E-07	G protein/p205	4	40.23	1.58E-07
A kinase (gravin)	2	20.15	1.17E-05	β-ARK1	2	20.13	1.14E-05
Protein synthesis/folding/modification							
ERp72	26	148.28	1.77E-12	HSP70 protein5/BiP	16	106.26	1.33E-10

Table 3. Continued

Untreated				MMC-treated			
Protein	Pep	Score	Prob.	Protein	Pep	Score	Prob.
Serpin H1 precursor/HSP47	19	130.25	3.73E-12	δ-Pyrroline-5-carboxylate synthetase	14	100.29	6.99E-14
δ-Pyrroline-5-carboxylate synthetase	15	86.3	6.88E-15	ERp72	7	48.26	1.04E-09
TCP-1 complex proteins	11	84.25	3.38E-08	TCP-1 complex proteins	6	60.32	3.79E-13
HSP 60	10	70.32	4.44E-15	HSP 90-β	2	18.24	9.21E-04
HSP 90-β	10	66.24	1.44E-10				
HSP70 protein5/BiP	7	70.21	2.48E-08				
Elongation factor 1-α	6	40.17	1.16E-06				
NEDD4	6	60.22	2.98E-09				
Protein Tyr kinase 7	5	50.20	5.40E-10				
Ubiquitin-activating enz. E1	4	38.21	1.07E-06				
Amino acid/carbohydrate metabolism							
Aldolase 1	20	100.28	2.22E-15	GAPDH	9	58.22	6.55E-08
Pyruvate kinase M	9	50.19	1.40E-06	Pyruvate kinase M	4	38.25	1.00E-30
GAPDH	4	40.21	1.94E-08	Isocitrate dehydrogenase	2	20.20	2.11E-05
Transketolase	3	28.16	1.26E-06				
Cytoskeleton							
Myosin	37	338.31	1.00E-30	Vimentin	32	128.29	4.63E-12
Actinin-α	23	158.28	2.10E-12	Myosin	16	120.28	4.86E-11
β-Actin	16	62.25	1.93E-10	β-Tubulin	15	100.22	2.90E-09
Vimentin	11	106.27	2.95E-10	Filamin A	11	110.32	1.24E-13
Cytosk. assoc. prot. P63	6	60.2	1.44E-07	Gelsolin	2	20.24	1.19E-07

(Complete list in Supporting Information table XRCC2).

were other proteins involved in DNA repair, transcription, translation, and nucleotide biosynthesis. Several may now be considered as networking partners of the HR repair machinery. Such *in vitro* interactions now need to be confirmed *ex vivo* and further characterized by functional studies to determine significance.

Four of the interacting proteins identified by the MS screen of the eluates, SFPQ, NONO, MCM2 and MSH2 were confirmed by immunoblot analysis. The NONO protein, also known as p54[nrb] in humans, was identified previously as a marker in a subset of human breast cancer and thought to play a role in transcriptional regulation [34, 35]. SFPQ (PSF) is a paralog of NONO and was initially discovered to be associated with pre-mRNA splicing [36]. However, later evidences demonstrated that each of these proteins have multiple functions in RNA processing, transcriptional regulation, DNA unwinding and homologous pairing in the formation of D-Loop structure, one of the first structural intermediates of HR repair [37, 38]. As a stable heteromeric complex, SFPQ-NONO has been identified as a candidate for DNA repair by efficient rejoining of DSB [39]. The homologous pairing along with the strand rejoining function of these proteins strongly suggests a role in HR mediated DNA repair. Interestingly, a previous proteomic screen identified predicted splicing factor proteins as ATM and ATR phosphorylated substrates in response to DNA damage [40]. These findings are also consistent with RNA posttranscriptional modification proteins being involved in DNA damage response and DSB repair [41].

The discovery of the MutS-α complex, MSH2 and MSH6, involved in recognition and processing of DNA mismatches and insertion-deletion by the DNA MMR machinery, suggests an integration of HR and MMR pathways. This is the first report of mammalian RAD51 proteins being associated with MMR proteins, indicating that a reductionist paradigm of isolated DNA repair pathways is oversimplified [15]. Rather, these data support that dynamic interconnections exist among the repair pathways. Another proteomic study to dissect out the interactome of human MutL homologues demonstrated that NONO, discussed above, interacts with another MMR protein, PMS2 [42], further supporting that crosstalk occurs among DNA repair pathways.

An additional group of proteins identified by this study was the DNA replication-licensing factor, MCM proteins. MCM2, MCM3 and MCM5 were identified with RAD51D; MCM6 was identified with RAD51C; and MCM2 was identified with XRCC2 co-precipitations, respectively. These findings are consistent and build on the known connections between DNA damage response and DNA replication [40], suggesting a functional interaction to either assist in restarting broken replication forks and/or initiate replication following DNA damage. Interestingly, each of these interactions except MCM2-RAD51D was identified from the MMC-treated MEF extracts alone. The identification of these proteins (as well as the MutL-α complex) upon MMC treatment suggest either an overexpression of these proteins upon MMC treatment or secondary modifications upon DNA damage, enhancing their binding affinity of these to
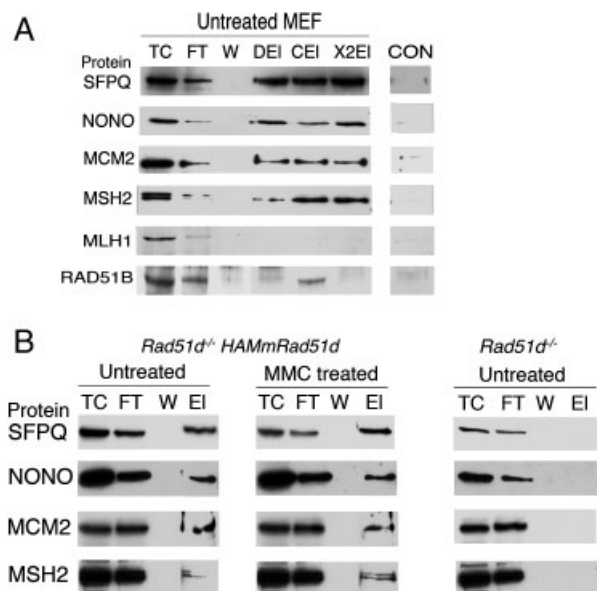
Figure 4. Validation of RAD51 paralog interacting candidates. (A) Four candidates identified by mass spectroscopy, SFPQ, NONO, MCM2 and MSH2, were verified by Western analysis using the respective antibody in each co-precipitation eluate. Anti-MLH1 was used as a negative control to verify the specificity of the above interactions. RAD51B was used as a positive control and detected specifically in the RAD51C eluate. Also note that RAD51B was identified by mass spectroscopy analysis in the XRCC2 eluate. TC-whole cell extract; FT-co-precipitation supernatant; W- Co-precipitation washes; EI-eluates from the respective co-precipitation of RAD51D (DEI), RAD51C (CEI) and XRCC2 (X2EI). CON: Eluate of precipitation without any bait protein. Representative lanes of only one co-precipitation FT and W sample are shown since these samples did not differ between co-precipitation experiments. (B) *Ex vivo* interaction validation using Anti-HA antibody immunoprecipitation eluates of extracts of *Rad51d*[−/−] HA-MmRAD51d and *Rad51d*[−/−] MEFs. Untreated-whole cell extracts from untreated MEFs. MMC-treated-whole cell extracts from cells treated with MMC. TC-whole cell extract; FT-immunoprecipitation supernatant; W-immunoprecipitation washes; EI-eluates of immunoprecipitation.

the RAD51 paralogs. The six related MCM proteins (MCM2–MCM7) form a complex that is an essential replicative helicase required during DNA synthesis both at the initiation and elongation stages [43]. This complex is important for maintaining genomic stability at the replication forks, and our results indicate a role of these MCM proteins in HR mediated DNA repair or suggest a role for the RAD51 paralogs in DNA repair during replication.

The findings presented here are the first report of RAD51 paralogs interacting with proteins involved in MMR (MSH2) and proteins with a suggested role in non-homologous end joining (SFPQ-NONO) mediated repair. This is consistent with the hypothesis that RAD51-like proteins function as multiple complexes at different stages of HR mediated DNA repair and is also suggestive of their role in multiple repair pathways [15]. These results also suggest important roles for RAD51 paralogs

in additional cellular activity that will require further investigations. The DNA damage response was previously linked to a large number of protein networks and modules pointing to a broader landscape of proteins involved in repairing the damage [40]. Additional candidates and protein functional classes identified by this screen that function in RNA processing, protein modifications and cellular trafficking are now being implicated for their direct or regulatory role in DNA repair pathways [44–46]. The findings reported here also highlight the fact that it is difficult to delineate the complexity of cellular functions, as several proteins with multiple roles must act in cohesion. Future work will include validating these candidate protein interactions and deciphering the complexities of networking proteins involved in DNA damage repair and the maintenance of genomic stability.

5   References

[1] Kowalczykowski, S. C., Structural biology: snapshots of DNA repair. *Nature* 2008, *453*, 463–466.

[2] Baumann, P., Benson, F. E., West, S. C., Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions *in vitro*. *Cell* 1996, *87*, 757–766.

[3] Richardson, C., Stark, J. M., Ommundsen, M., Jasin, M., Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. *Oncogene* 2004, *23*, 546–553.

[4] Bertrand, P., Lambert, S., Joubert, C., Lopez, B. S., Over-expression of mammalian Rad51 does not stimulate tumorigenesis while a dominant-negative Rad51 affects centrosome fragmentation, ploidy and stimulates tumorigenesis, in p53-defective CHO cells. *Oncogene* 2003, *22*, 7587–7592.

[5] Pittman, D. L., Schimenti, J. C., Midgestation lethality in mice deficient for the RecA-related gene, Rad51d/Rad51l3. *Genesis* 2000, *26*, 167–173.

[6] Shu, Z., Smith, S., Wang, L., Rice, M. C., Kmiec, E. B., Disruption of muREC2/RAD51L1 in mice results in early embryonic lethality which can Be partially rescued in a p53(−/−) background. *Mol. Cell. Biol.* 1999, *19*, 8686–8693.

[7] Deans, B., Griffin, C. S., Maconochie, M., Thacker, J., Xrcc2 is required for genetic stability, embryonic neurogenesis and viability in mice. *EMBO J.* 2000, *19*, 6675–6685.

[8] Kuznetsov, S. G., Haines, D. C., Martin, B. K., Sharan, S. K., Loss of Rad51c leads to embryonic lethality and modulation

of Trp53-dependent tumorigenesis in mice. *Cancer Res.* 2009, *69*, 863–872.

[9] Hatanaka, A., Yamazoe, M., Sale, J. E., Takata, M. *et al.*, Similar effects of Brca2 truncation and Rad51 paralog deficiency on immunoglobulin V gene diversification in DT40 cells support an early role for Rad51 paralogs in homologous recombination. *Mol. Cell. Biol.* 2005, *25*, 1124–1134.

[10] Yokoyama, H., Sarai, N., Kagawa, W., Enomoto, R. *et al.*, Preferential binding to branched DNA strands and strand-annealing activity of the human Rad51B, Rad51C, Rad51D and Xrcc2 protein complex. *Nucleic Acids Res.* 2004, *32*, 2556–2565.

[11] Braybrooke, J. P., Spink, K. G., Thacker, J., Hickson, I. D., The RAD51 family member, RAD51L3, is a DNA-stimulated ATPase that forms a complex with XRCC2. *J. Biol. Chem.* 2000, *275*, 29100–29106.

[12] Takata, M., Sasaki, M. S., Tachiiri, S., Fukushima, T. *et al.*, Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol. Cell. Biol.* 2001, *21*, 2858–2866.

[13] Tarsounas, M., Munoz, P., Claas, A., Smiraldo, P. G. *et al.*, Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* 2004, *117*, 337–347.

[14] Helleday, T., Pathways for mitotic homologous recombination in mammalian cells. *Mutat. Res.* 2003, *532*, 103–115.

[15] Wyatt, M. D., Pittman, D. L., Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chem. Res. Toxicol.* 2006, *19*, 1580–1594.

[16] Masson, J. Y., Tarsounas, M. C., Stasiak, A. Z., Stasiak, A. *et al.*, Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.* 2001, *15*, 3296–3307.

[17] Masson, J. Y., Stasiak, A. Z., Stasiak, A., Benson, F. E., West, S. C., Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 8440–8446.

[18] Liu, Y., Tarsounas, M., O'regan, P., West, S. C., Role of RAD51C and XRCC3 in genetic recombination and DNA repair. *J. Biol. Chem.* 2007, *282*, 1973–1979.

[19] Ip, S. C., Rass, U., Blanco, M. G., Flynn, H. R. *et al.*, Identification of holliday junction resolvases from humans and yeast. *Nature* 2008, *456*, 357–361.

[20] Braybrooke, J. P., Li, J. L., Wu, L., Caple, F. *et al.*, Functional interaction between the bloom's syndrome helicase and the RAD51 paralog, RAD51L3 (RAD51D). *J. Biol. Chem.* 2003, *278*, 48357–48366.

[21] Lio, Y. C., Mazin, A. V., Kowalczykowski, S. C., Chen, D. J., Complex formation by the human Rad51B and Rad51C DNA repair proteins and their activities *in vitro. J. Biol. Chem.* 2003, *278*, 2469–2478.

[22] Cheung, K. M., Matthews, T. P., James, K., Rowlands, M. G. *et al.*, The identification, synthesis, protein crystal structure and *in vitro* biochemical evaluation of a new 3,4-diarylpyrazole class of Hsp90 inhibitors. *Bioorg. Med. Chem. Lett.* 2005, *15*, 3338–3343.

[23] Smiraldo, P. G., Gruver, A. M., Osborn, J. C., Pittman, D. L., Extensive chromosomal instability in Rad51d-deficient mouse cells. *Cancer Res.* 2005, *65*, 2089–2096.

[24] Gruver, A. M., Miller, K. A., Rajesh, C., Smiraldo, P. G. *et al.*, The ATPase motif in RAD51D is required for resistance to DNA interstrand crosslinking agents and interaction with RAD51C. *Mutagenesis* 2005, *20*, 433–440.

[25] Wiese, C., Hinz, J. M., Tebbs, R. S., Nham, P. B. *et al.*, Disparate requirements for the Walker A and B ATPase motifs of human RAD51D in homologous recombination. *Nucleic Acids Res.* 2006, *34*, 2833–2843.

[26] Shim, K. S., Schmutte, C., Tombline, G., Heinen, C. D., Fishel, R., hXRCC2 enhances ADP/ATP processing and strand exchange by hRAD51. *J. Biol. Chem.* 2004, *279*, 30385–30394.

[27] Schild, D., Lio, Y. C., Collins, D. W., Tsomondo, T., Chen, D. J., Evidence for simultaneous protein interactions between human Rad51 paralogs. *J. Biol. Chem.* 2000, *275*, 16443–16449.

[28] Miller, K. A., Sawicka, D., Barsky, D., Albala, J. S., Domain mapping of the Rad51 paralog protein complexes. *Nucleic Acids Res.* 2004, *32*, 169–178.

[29] Richardson, C., RAD51, genomic stability, and tumorigenesis. *Cancer Lett.* 2005, *218*, 127–139.

[30] Henson, S. E., Tsai, S. C., Malone, C. S., Soghomonian, S. V. *et al.*, Pir51, a Rad51-interacting protein with high expression in aggressive lymphoma, controls mitomycin C sensitivity and prevents chromosomal breaks. *Mutat. Res.* 2006, *601*, 113–124.

[31] Gruver, A. M., Yard, B. D., McInnes, C., Rajesh, C., Pittman, D. L., Functional characterization and identification of mouse Rad51d splice variants. *BMC Mol. Biol.* 2009, *10*, 27.

[32] Ouyang, K. J., Woo, L. L., Ellis, N. A., Homologous recombination and maintenance of genome integrity: cancer and aging through the prism of human RecQ helicases. *Mech. Ageing Dev.* 2008, *129*, 425–440.

[33] Mimitou, E. P., Symington, L. S., Nucleases and helicases take center stage in homologous recombination. *Trends Biochem. Sci.* 2009, *34*, 264–272.

[34] Pavao, M., Huang, Y. H., Hafer, L. J., Moreland, R. B., Traish, A. M., Immunodetection of nmt55/p54nrb isoforms in human breast cancer. 2001, *1*, 15.

[35] Traish, A. M., Huang, Y. H., Ashba, J., Pronovost, M. *et al.*, Loss of expression of a 55 kDa nuclear protein (nmt55) in estrogen receptor-negative human breast cancer. *Diagn. Mol. Pathol.* 1997, *6*, 209–221.

[36] Patton, J. G., Porro, E. B., Galceran, J., Tempst, P., Nadal-Ginard, B., Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev.* 1993, *7*, 393–406.

[37] Shav-Tal, Y., Zipori, D., PSF and p54(nrb)/NonO-multifunctional nuclear proteins. *FEBS Lett.* 2002, *531*, 109–114.

[38] Akhmedov, A. T., Lopez, B. S., Human 100-kDa homologous DNA-pairing protein is the splicing factor PSF and promotes DNA strand invasion. *Nucleic Acids Res.* 2000, *28*, 3022–3030.

[39] Bladen, C. L., Udayakumar, D., Takeda, Y., Dynan, W. S., Identification of the polypyrimidine tract binding protein-associated splicing factor.p54(nrb) complex as a candidate DNA double-strand break rejoining factor. *J. Biol. Chem.* 2005, *280*, 5205–5210.

[40] Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R. r. *et al.*, ATM and ATR substrate analysis reveals

extensive protein networks responsive to DNA damage. *Science* 2007, *316*, 1160–1166.

[41] Li, L., Monckton, E. A., Godbout, R., A role for DEAD box 1 at DNA double-strand breaks. *Mol. Cell. Biol.* 2008, *28*, 6413–6425.

[42] Cannavo, E., Gerrits, B., Marra, G., Schlapbach, R., Jiricny, J., Characterization of the interactome of the human MutL homologues MLH1, PMS1, and PMS2. *J. Biol. Chem.* 2007, *282*, 2976–2986.

[43] Forsburg, S. L., Eukaryotic MCM proteins: beyond replication initiation. *Microbiol. Mol. Biol. Rev.* 2004, *68*, 109–131.

[44] de Souza-Pinto, N. C., Mason, P. A., Hashiguchi, K., Weissman, L. *et al.*, Novel DNA mismatch-repair activity involving YB-1 in human mitochondria. *DNA Repair (Amst)* 2009, *8*, 704–719.

[45] Knudsen, N. O., Andersen, S. D., Lutzen, A., Nielsen, F. C., Rasmussen, L. J., Nuclear translocation contributes to regulation of DNA excision repair activities. *DNA Repair (Amst)* 2009, *8*, 682–689.

[46] Huang, J., Huen, M. S., Kim, H., Leung, C. C. *et al.*, RAD18 transmits DNA damage signalling to elicit homologous recombination repair. *Nat. Cell Biol.* 2009, *11*, 592–603.