*Nucleic Acids Research, 2006, Vol. 34, No. 11* ***3345–3360***
*doi:10.1093/nar/gkl455*

# Homologous recombination is required for AAV-mediated gene targeting

**Ana Vasileva[1], R. Michael Linden[1] and Rolf Jessberger[1,2,*]**

[1]Department of Gene and Cell Medicine, Mount Sinai School of Medicine, New York, NY 10029, USA and
[2]Institute of Physiological Chemistry, Medical Faculty Carl Gustav Carus, Dresden University of Technology, 01307 Dresden, Germany

Received May 22, 2006; Revised May 22, 2006; Accepted June 13, 2006

## ABSTRACT

**High frequencies of gene targeting can be achieved by infection of mammalian cells with recombinant adeno-associated virus (rAAV) vectors [D. W. Russell and R. K. Hirata (1998) *Nature Genet.*, 18, 325–330; D. W. Russell and R. K. Hirata (2000) *J. Virol.*, 74, 4612–4620; R. Hirata *et al.* (2002) *Nat. Biotechnol.*, 20, 735–738], but the mechanism of targeting is unclear and random integration often occurs in parallel. We assessed the role of specific DNA repair and recombination pathways in rAAV gene targeting by measuring correction of a mutated enhanced green fluorescent protein (EGFP) gene in cells where homologous recombination (HR) or non-homologous end-joining (NHEJ) had been suppressed by RNAi. EGFP-negative cells were transduced with rAAV vectors carrying a different inactivating deletion in the EGFP, and in parallel with rAAV vectors carrying red fluorescent protein (RFP). Expression of RFP accounted for viral transduction efficiency and long-term random integration. Approximately 0.02% of the infected GFP-negative cells were stably converted to GFP positive cells. Silencing of the essential NHEJ component DNA-PK had no significant effect on the frequency of targeting at any time point examined. Silencing of the SNF2/SWI2 family members RAD54L or RAD54B, which are important for HR, reduced the rate of stable rAAV gene targeting ~5-fold. Further, partial silencing of the Rad51 paralogue XRCC3 completely abolished stable long-term EGFP expression. These results show that rAAV gene targeting requires the Rad51/Rad54 pathway of HR.**

## INTRODUCTION

The correction of mutations in genomic DNA *in situ* is one of the most attractive approaches for gene therapy of inherited single gene disorders. This process is inherently non-mutagenic and very likely to conserve appropriately regulated expression of the repaired gene product. Gene targeting through homologous recombination (HR) is the most accurate and versatile mechanism for such correction. However, HR occurs in human somatic cells with very low frequencies of about $10^{-7}$ (1). Efforts directed towards improving the frequency of gene repair have lead to the development of numerous oligonucleotide-based strategies, involving delivery of plain DNA by transfection or injection. Triplex-forming oligonucleotides (TFOs) can induce recombination between a polypurine/polypyrimidine rich DNA duplex and unlinked donor molecule by introducing damage to such duplex. The maximum frequency of targeted modification achieved with TFOs *in vivo* was $10^{-4}$ (2). Alternatively, the best frequency for gene editing with chimeric RNA-DNA oligonucleotides (RDOs) or chimeraplasts was 1% of all muscle fibers positive for the chimeraplast DNA mostly close to the injection site (3). Chimeraplast and single-stranded oligodeoxynucleotide-mediated gene repair (4) most likely involves a mismatch repair mechanism and is restricted to repair of point mutations. A major limitation of the ability of injected chimeraplast to promote gene conversion in muscle appears to be the restricted uptake of chimeraplasts into fibers due to inefficient delivery of the repair substrate *in vivo*. A short-fragment homologous replacement (SFHR) strategy allows alteration of as many as 4 bp, with a maximum reported frequency of about 0.4% (5).

Viral delivery systems based on retrovirus, adenovirus and adeno-associated virus (AAV) overcome some of these limitations by offering more efficient DNA delivery and by accommodating longer stretches of homology with the genomic locus to allow a broad spectrum of targeted modifications. The best targeting frequencies achieved by retrovirus and adenovirus vectors, however, were only comparable to those shown by DNA transfection (6,7). In contrast, recombinant adeno-associated virus (rAAV) vectors comprised of single-stranded DNA with unique inverted terminal repeats (ITRs), reach the nucleus in multiple copies and have been shown to target homologous sequences in cultured human cells with an efficiency of up to 1% [reviewed in (8,9)]. This property of rAAV combined with its broad host

*To whom correspondence should be addressed. E-mail: rolf.jessberger@tu-dresden.de

© 2006 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commerical use, distribution, and reproduction in any medium, provided the original work is properly cited.

range, the availability of multiple virus serotypes and recent advancements in viral re-targeting through capsid manipulation (10) make rAAV an attractive system for gene correction and, alternatively, for gene disruption aimed at generating knock-outs in cultured human primary cells or cell lines (11,12). Still, the rate of AAV-mediated gene targeting alone is not sufficiently high for gene therapy through gene correction *in vivo*, where the accompanying potentially mutagenic random integration events cannot be selected against. In cell culture, the rate of random integration was at 10%, and integration *in vivo* is known to occur predominantly in transcriptionally active genes (13,14).

Similar to conventional gene targeting the frequency of rAAV-mediated targeting was elevated significantly by the introduction of DNA double-strand breaks (DSBs) at the targeted site (15,16). This supports the hypothesis that components of DSB repair pathways, such as non-homologous end-joining (NHEJ) or HR are involved in rAAV-mediated gene targeting. Studies in SCID mice suggest that DNA-PK promotes circularization of linear double-stranded rAAV genomes (17). It has also been shown that proteins involved in both NHEJ and HR, such as KU86 and RAD52 bind to the AAV genome and affect the transduction efficiency, possibly via a modification of AAV DNA processing (18). However, it remains unclear which pathway and proteins mediate gene targeting by rAAV.

The unique structure of the rAAV vector DNA with an unusually short stretch of double-stranded (ds) DNA transitioning into the extended single-strand region between the ITRs, was found to be important for targeting. Addition of self-complementary double-stranded genomes to the native single-stranded rAAV DNA did not improve the targeting reaction, and dimeric vector molecules, which do not contain the characteristic ITRs and ss–dsDNA transition regions, failed to target efficiently (19). The largely single-stranded genomes represent the majority of vector forms in the infected cell for several days (20), a period of time that should be sufficient for targeting reactions to occur. Therefore, DNA repair factors or protein complexes involved in the cellular pathways of processing of ssDNA of specific DNA structures, such as ITRs, or of ss–ds transitional regions might be key to gene targeting by rAAV. Deciphering the mechanism of rAAV gene targeting may in the future allow manipulation of potentially competing pathways in order to both enhance targeting rates and lower random integration events. In this study we identified components of the cellular DNA repair/recombination machinery, which are essential for efficient rAAV-mediated gene targeting.

## MATERIALS AND METHODS

### Cloning of mutant enhanced green fluorescent protein (EGFP) target and substrate vectors

Standard molecular biology techniques were used to modify all plasmids (21). The pEGFP C1 expression plasmid (BD Biosciences Clontech, Palo Alto, CA) was used as a source of the EGFP gene and also as a plasmid backbone for the genomic target. Selection of the cells containing the target on hygromycin was made possible by replacing the neomycin phosphotransferase gene (*neo*) of pEGFP C1 with a hygromycin phosphotransferase gene (*hyg*) from pSecTag-Hygro (InVitrogen, Carlsbad, CA). A PCR fragment spanning bases 1–198 of the *EGFP* gene was generated using the following oligonucleotides A-NheI-forward: 5′-TGAACCGTCAGATCCGCTAG-3′ and A-SpeI-reverse 5′-GTACTAGTGGGTGGTCACGAGG-3′. Another PCR fragment spanning bases 230–740 was amplified using oligonucleotides C-SpeI-forward: 5′-GACTAGTTGAAGCAGCACGAC-3′ and B&C-XhoI-reverse: 5′-TGAGCTCGAGATCTGAGTCCG-3′. The EGFP gene resulting from joining these two PCR fragments via a SpeI endonuclease site contained a deletion of 32 bases and served as a mutant target. This EGFPΔ32 gene was used to replace the intact *EGFP* gene in pEGFP C1-Hyg linearized with NheI/XhoI to yield p EGFPΔ32 (Figure 1A).

Another mutant gene lacking the first 14 bases—Δ14EGFP (including the start codon) of the wild-type EGFP gene was generated by amplification from the pEGFP C1 template with oligonucleotides B-NotI-forward: 5′-CAAGCGGCCGCAGCTGTTCAC-3′ and B&C-XhoI-reverse shown above. This mutant was used to replace the GFP gene in NotI–XhoI digested proviral rAAV2 plasmid pTRUF11 (22). Both truncated EGFP mutants were confirmed by sequencing. The cytomegalovirus (CMV) enhancer and the CAAT and TATA signals of the chicken β-actin (CBA) promoter in pTRUF11 was replaced by a portion of the yeast *Pichia pastoris* AOX1 transcription termination sequence thus restricting additional homology between the viral template and the pEGFPΔ32 genomic target to 100 bp upstream of the EGFP start. The resulting plasmid pTR-Neo was used for the production of rAAV-Neo. Plasmid pTR-RFP was constructed by replacing the $P_{SV40}$-neo expression cassette on pTR-Neo with a red fluorescent protein (RFP) expressed under the mouse cytomegalovirus promoter (mCMV). This plasmid was used for preparation of rAAV-RFP. The RFP gene was amplified from pDsRed2.1 (BD Biosciences Clontech, Palo Alto, CA) and mCMV was derived from pSEAP-Hyg (gift from Dr Karen Anthony). Viruses containing substrate with extended homology to the target sequence were created by adding 1.009 kb fragment from the backbone of pEGFP C1 to the 3′ end of the GFP mutant gene on pTR-Neo or pTR-RFP. The resulting vectors pTR-NeoX and pTR-RFPX were used for production of rAAV-NeoX and rAAV-RFPX, respectively (Figure 1B and C). All pTRUF11-based plasmids were cloned and propagated in SURE2 cells (BD Biosciences Clontech, Palo Alto, CA). Plasmid DNA was purified using a kit (Qiagen Inc., Valencia, CA).

### Production, purification and titration of rAAV vector stocks

Recombinant AAV2 was produced in 293T cells using a dual-vector, calcium phosphate-mediated transfection approach and viral stocks were purified as described previously (22). Plasmid pDG carries the Rep and Cap genes necessary for rAAV2 packaging as well as the adenovirus helper functions (E2A, E4, VA). Iodixanol fractions containing the virus particles were placed directly in a 10 kDa cut-off Slide-A-Lyser cassette (Pierce, Rockford, IL) and dialyzed against DMEM at 4°C. The dialysis was complete when the refraction coefficient of the virus stock was equivalent to that of a pure DMEM medium. Virus stock aliquotes were

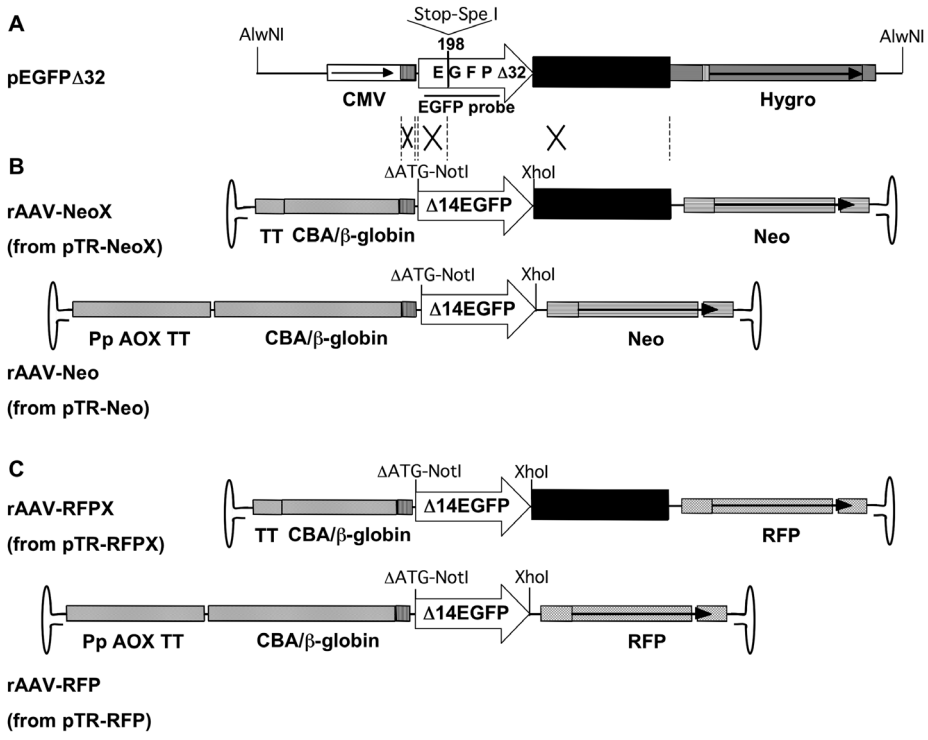


**Figure 1.** Schematic representation of the vectors used for gene targeting. (**A**) Plasmid pEGFPΔ32 containing the mutant target (thick white arrow) expressed under human CMV promoter was integrated in the genome of MO59K cells. Thirty-two base pair deletion at position 198 (5′-ctgacctacggcgtg-cagtgcttcagccgctaccccgaccacatgaa-3′) was replaced by an in-frame stop-codon supplied within a unique SpeI restriction site (5′-CACTAGTTGAAGCa-3′). Resistance to hygromycin was provided by a hygromycin phosphotransferase (Hygro) expressed from an SV40 promoter. (**B**) Schematic of rAAV used as repair substrates in gene targeting experiments. All viral vectors consisted of single-stranded DNA enclosed by native AAV subtype 2 inverted repeats. Both substrates (white arrows) lacked 14 bp from the 5′ end of the EGFP gene including the ATG codon (Δ14EGFP). A short and a long fragment from the *P.pastoris* alcohol oxidase termination sequence (Pp AOX TT and TT) replaced the CMV and the CAAT+TATA signals of the CBA/rabbit β-globin promoter (CBA/β-globin). The rAAV-NeoX contained additional homology to the genomic target (thick black line). The recombinant viruses also contained a neomycin phosphotransferase gene (Neo) expressed under a HSV-TK promoter. (**C**) Diagram of the vectors used to test cell infectivity. rAAV-RFPX and rAAV-RFP were identical to rAAV-NeoX and rAAV-Neo, respectively, except for the replacement of the neomycin phosphotransferase expression cassete with a RFP expression.

snap-frozen in liquid nitrogen for subsequent use. Physical titers of rAAV were determined by slot-blot hybridization and/or real-time PCR with GFP-specific primers (specified above) using the SYBR Green I QPCR kit (Qiagen Inc.). Samples were prepared as described previously (23). Throughout this study, MOIs were defined as the number of DNase-resistant rAAV genome particles per cell (gcp/cell).

## Cell culture and cell lines

Human glioma cells MO59K and MO59J (24), Phoenix-GP cells, 293 and 293T cells have been described previously. All cells were grown in a humidified chamber at 37°C and 5% $CO_2$. Cells were split and passaged by following standard procedures. 293, 293T and Phoenix-GP cells were grown in DMEM containing 4 g of glucose per liter (except for rAAV production where glucose was 1.5 g/l), 10% heat-inactivated calf serum (HyClone, Logan, UT), 2 mM L-glutamine-100 U of penicllin/ml and 100 μg of streptomycin/ml. MO59K cells were grown with the same supplements, except that the base medium was DMEM:F12 (Invitrogen). Cell line MO59KΔ32 was created by electroporation of $10^7$ cells with 20 μg of AlwNI linearized plasmid pEGFPΔ32. Drug-resistant cells were selected by culturing in 150 μg of hygromycin (Invitrogen)/ml from day 2 after electroporation until distinct colonies formed ∼14 days after electroporation. Minimum of 24 individual colonies were picked from each cell line, trypsinized and serially diluted for further purification to single cells in 48-well plates. Wells containing single colonies were transferred in duplicate to 24-well plates. After the expansion of these individual colonies replica plates were frozen for storage and the original plates used for DNA isolation for Southern analysis. In addition, early passage of the polyclonal MO59KΔ32 population was frozen in aliquots for subsequent rAAV infection and silencing experiments.

To enable infection with mouse retrovirus for RAD54L silencing clone MO59KΔ32 B was used to create a cell line expressing mouse retroviral ecotropic receptor. This cell line was generated by transduction with pEco (gift from Dr Saghi Gaffari) followed by G418 selection (0.3 mg of active compound per ml).

## DNA and RNA analysis

Genomic DNA and total RNA from hygromycin resistant clones were isolated using a DNA tissue kit or RNAeasy kit, respectively (Qiagen Inc.). Probes for Southern, northern

and slot-blots hybridizing with GFP, Neo or RFP were amplified from pEGFP C1, pTR-B and pTR-RFP, respectively. The probe hybridizing to XRCC3 was isolated as 1 kb EcoRI fragment from pGAGGS-XRCC3 (gift from Dr Maria Jasin). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe was amplified by RT–PCR from human total RNA. Northern and slot-blot hybridization signals were quantified by PhosphorImager analysis (Molecular Dynamics, Sunnyvale, CA).

### rAAV gene correction assays

EGFP targeted correction assays were performed in 100 mm culture dishes seeded with $5 \times 10^5$ live cells (determined by trypan-blue exclusion) per dish (70% confluence) at the time of viral infection. Unless specified otherwise, MO59KΔ32 cells were trypsinized, counted, seeded in triplicate and left to rest for an hour. The cells were transduced with rAAV carrying either minimal or extended homology to the target. The multiplicity of infection (MOI) was 1000, 10 000 or 50 000 genome containing gcp/cell. Twenty-four hours later the medium containing the virus was replaced. At 48 or 72 h after infection each plate was passaged and an aliquot (usually half of the cells by volume) analyzed by a fluorescence-activated cell scan (FACS, Becton Dickinson) with settings, FSC E-1, SSC 150, FL1 452 nm, FL2 410 nm and FL3 500 nm with AmpGain 6.23, 1.22, 1.15, 2.91 and 1, respectively; compensation FL1-1.8% FL2, FL2-41.7% FL1 and FL3-72.6%-FL2. Each EGFP positive cell was considered as a single gene correction event. Average gene correction efficiency was calculated from samples collected from triplicate transduction. Infected cells were then serially passaged on alternate days and analyzed by flow cytometry until the percentage of GFP positive cells stabilized. Gene repair assays comparing MO59KΔ32 silenced and untreated cells were similarly performed in triplicate in 60 mm plates seeded with $3 \times 10^5$ cells at the time of infection. In addition to the transduction with rAAV-Neo vectors, cells were transduced with rAAV-RFP or rAAV-RFPX that served as a control for the efficiency of infection. Cells transduced with rAAV-SEAP that expresses secreted alkaline phosphatase (AP) (gift from Dr Thomas Weber) were used instead of a non-transduced control.

### Discrete analysis of gene correction events

Cells from the MO59KΔ32 B clone ($3 \times 10^5$) were infected with rAAV-NeoX vector (MOI = 20 000 particles/cell) in a 60 mm plate. The virus was removed 24 h later. The number of EGFP positive cells was determined by FACS on day 3 when cells were expanded to $2 \times 100$ mm plates. EGFP positive cells were separated from the total infected population by fluorescence-activated cell sorting on a FACS Vantage SE with FACS Diva software (BD) on day 5. Single cells seeded onto 96-well plates were then expanded for cryopreservation and isolation of genomic DNA. The presence of full-length corrected EGFP was confirmed by PCR using primers spanning a sequence from the 3′ end of the CMV promoter (5′-TCAGATCCGCTAGCGCTACCGGTC-3′) and the EGFP gene (5′-GGTGCGCTCCTGGACGTAGCCTT-3′). The fidelity of repair was confirmed by sequencing of the PCR product. In addition, genomic DNA from eight independent GFP positive corrected clones was analyzed by Southern blot with EGFP-specific probe.

### Design of short-hairpin oligonucleotides and plasmids for RNA interference

A pool of oligonucleotides for transcriptional silencing was generated using the Whitehead Institute siRNA prediction software (http://jura.wi.mit.edu/siRNAext/). Analysis involving additional position and thermodynamic criteria narrowed the selection to 2–4 sequences for each target transcript. Unless stated otherwise, DNA oligonucleotides used to generate short-hairpin RNA (shRNA) were 64 bases long and shared common structure that differed only in the mRNA coding sequence: 5′-GATCccc(19N)ttcaagaga(19N)-ttttGGAAA-3′. To provide better export of the shRNA to the cytoplasm the loop of oligonucleotides for downregulation of Rad54L expression contained a native human sequence derived from miRNA-23 (5′-cttcctgtca-3′). The oligonucleotide targeting the DNA-PK mRNA starting from position 6684 had 20 nt of sense and 19 nt of antisense sequence. The siRNA starting at position 11 838 (or 11 740 on alternatively spliced) on the DNA-PKcs mRNA was described previously (25). XRCC3 mRNA was targeted for degradation with oligonucleotide sequences starting at positions 362, 367, 914 and 2059; Rad54B oligonucleotides started at positions 475, 1409, 1511 and 1512 of the mRNA and Rad54L oligonucleotides started at positions 1472 and 2281 of the mRNA.

A mouse stem-cell retrovirus (MSCV) vector system was used to downregulate the expression of target genes. Prior to cloning of the DNA oligonucleotides downstream of the human RNA pol III promoter H1, plasmid pSUPER (OligoEngine Inc., Seattle, WA) was modified by insertion of a 1 kb 'stuffer' sequence between the BamHI and HindIII restriction sites. The above oligonucleotides were cloned into pSuper-stuffer by replacing the 1 kb sequence with the 64 bp shRNA. shRNA expression cassette was then subcloned into a pMSCV-PAC vector (BD Biosciences) in the EcoRI/XbaI sites of the MCS. This enabled production of retroviruses carrying silencing shRNA and subsequent selection for silenced populations on puromycin (2 μg/ml). To prevent cleavage of hairpin structures *Escherichia coli* strain GT116 that lacks the sbcC and sbcD genes was used for plasmid propagation and maintenance. Prior to nucleofection (Amaxa™) into cells containing targeted substrate the silencing potential of individual shRNA sequences was evaluated in 293 cells. Linearized pMSCV vectors were introduced by $CaHPO_4$ transfection, positive cells selected on puromycin and nuclear extracts analysed by immunoblotting as described below.

### Retrovirus vector preparations

MSCViruses expressing shRNA hairpins targeting for degradation Rad54L RNA were made by transient transfection of Phoenix-GP cells with pMSCV-sh1472Rad54L and pMSCV-sh2281Rad54L, replacing the culture medium 16 h later, harvesting conditioned medium after 16 h of exposure to cells, and filtering through a 0.45 μm pore-size syringe filter. These preparations were then used to transduce MO59KΔ32+Eco cells in the presence of 4 μg of Polybrene (Sigma–Aldrich Corp., St. Louis, MO) per ml.

### Immunoblotting

Levels of the proteins of interest in cells relative to those in control cells were measured by densitometry of standard immunoblot analysis. Nuclear extracts were prepared from minimum $10^5$ cells as described previously (26) and protein concentration determined using Bradford Reagent (Sigma–Aldrich Corp.). Depending on the protein being analyzed 5 or 10 μg of total protein were loaded per lane for SDS–PAGE. Rad54L, Rad54B and actin were detected with polyclonal goat antibodies D-18, N-16 and I-19, respectively (Santa Cruz Biotechnology, CA). Anti-DNA-PKcs was a rabbit polyclonal (Ab-1, Calbiochem, San Diego, CA) and anti-RNA pol II was a mouse monoclonal (8WG16, Covance Research Products, Denver, PA). Polyclonal rabbit antibodies were used to detect SMC1α. Secondary detection was done with horseradish peroxidase (HRP)–labeled porcine anti-goat (Roche Biochemicals, Indianapolis, IN), goat anti-rabbit or goat anti-mouse IgG2a antibodies (Jackson ImmunoResearch Laboratories Inc., West Grove, PA). RNA pol II and SMC1α which has previously been shown to be unaffected by changes in DNA-PKcs (27) or β-actin, were used for the loading controls for the shRNA knock-down experiments. To accurately quantify the time-course of changes in DNA-PKcs, Rad54L or Rad54B after shRNA knock-down, enhanced chemiluminescence reagent was used for X-ray film detection and Enhanced Chemiluminescence Plus reagent was used for densitometric image analysis (Storm Scanner System, Amersham Biosciences). Quantification was performed by ImageQuant™ software.

### Isolation of rAAV DNA and preparation of total cell lysates for slot-blot analysis

The isolation of low molecular weight plasmid DNA from mammalian cells has been decscribed previously (28). Briefly, $3 \times 10^5$ transduced cells were lysed in the well [0.6% SDS, 10 mM EDTA and 10 mM Tris–Cl (pH7.4)], collected with a rubber policeman and digested with proteinase K (50 μg/ml) in the same buffer at 37°C overnight. The high molecular weight DNA was precipitated with NaCl (1 M final) on ice overnight and removed by subsequent centrifugation for 30 min at 4°C, 10 000*g*. The low molecular weight DNA in the supernatant was extracted with phenol, phenol–chloroform and chloroform. Viral DNA was then ethanol precipitated in presence of glycogen as a carrier (Roche Biochemicals, Indianapolis, IN) and dissolved in distilled water (20 μl). Hirt extracted viral DNA was used to transfect 293T cells by $CaHPO_4$ precipitation. Total cell lysates for slot-blot analysis were prepared from transduced cells at different times after infection. Briefly, cells were trypsinized, counted, washed once with 1× PBS and re-suspended in 400 μl lysis buffer (0.4 M NaOH and 10 mM EDTA). The cell lysate was stored at −70°C for subsequent analysis. For quantification of rAAV DNA lysates were thawn and 100 μl of each spotted onto a nylon membrane using a slot-blot assembly (Bio-Rad, Hercules, CA). Serial dilutions of appropriate plasmid DNA applied onto the same membrane served as a reference for virus copy number. Slots were washed twice with lysis buffer, the membrane rinsed in 2× SSC and processed for hybridization with an appropriate probe.

## RESULTS

### Generation of a gene targeting system for correction of a mutant EGFP

To define the mechanism underlying rAAV-mediated gene targeting we designed an appropriate model system for the detection of targeting events. rAAV targeting has been reported in systems that require selection for the targeting events, such as repair of the neomycin or HPRT gene (29). Others evaluated frequencies of targeting AP (19), lacZ (16) and GFP gene (30) under non-selective conditions. Antibiotic or chemical selection may lead to inaccurate representation of the number of targeting events. Therefore, we elected to measure rAAV targeting frequency by direct detection of EGFP fluorescence in living cells and in the absence of selection. Since deletions are corrected by rAAV targeting much more efficiently than insertions (19), we generated EGFPΔ32 by deleting a 32 bp sequence including the chromophore and simultaneously introduced an in-frame stop-codon within a SpeI recognition site (Figure 1A). As observed by microscopy (Supplementary Figure 1A) and flow cytometry (Supplementary Figure 1B) the MO59KΔ32 cells are GFP-negative, allowing the identification of GFP positive gene targeting products. Subsequently, polyclonal population or clones containing the integrated EGFPΔ32 expression cassette were selected based on their resistance to hygromycin. The presence of the targeting substrate was confirmed by Southern analysis of genomic DNA (data not shown). Total RNA from both a polyclonal population and individual clones, found to be positive for the insert, was subjected to northern analysis in order to verify expression of the mutant EGFPΔ32 mRNA. A significant proportion of the hygromycin resistant clones containing the construct lacked the message (Supplementary Figure 1C), which can be attributed to previously reported strong and efficient silencing of the hCMV promoter in human cells (31) or to a disruption of the EGFP gene following integration. Since gene repair events would be undetectable in such non-expressing cells in a polyclonal population, this could lead to a misrepresentation of the absolute targeting frequencies. Therefore, all of the following experiments were carried out in both an EGFPΔ32 polyclonal population and at least one clone with confirmed EGFPΔ32 mRNA expression. To exclude any possible locus-specific effects in the confirmed clone, recombination frequencies were also estimated in EGFPΔ32 polyclonal populations and their silenced derivatives.

### Mutant EGFP is corrected by gene targeting with homologous rAAV vectors

Gene targeting by rAAV vectors yielded variable success in different systems [reviewed in (9)]. The system we generated had to exhibit robust frequencies of gene targeting to enable identification of changes upon modulation of DNA repair pathways. Two rAAV vectors (rAAV-Neo, rAAV-NeoX) (Figure 1B), were evaluated for their ability to mediate repair of the integrated target gene (EGFPΔ32) in MO59KΔ32 cells. Prior to conducting targeting experiments wild-type MO59K cells were transduced with 10 000 MOI (as gcp/cell) of rAAV-Neo or rAAV-NeoX, and confirmed by flow cytometry as GFP-negative (data not shown).

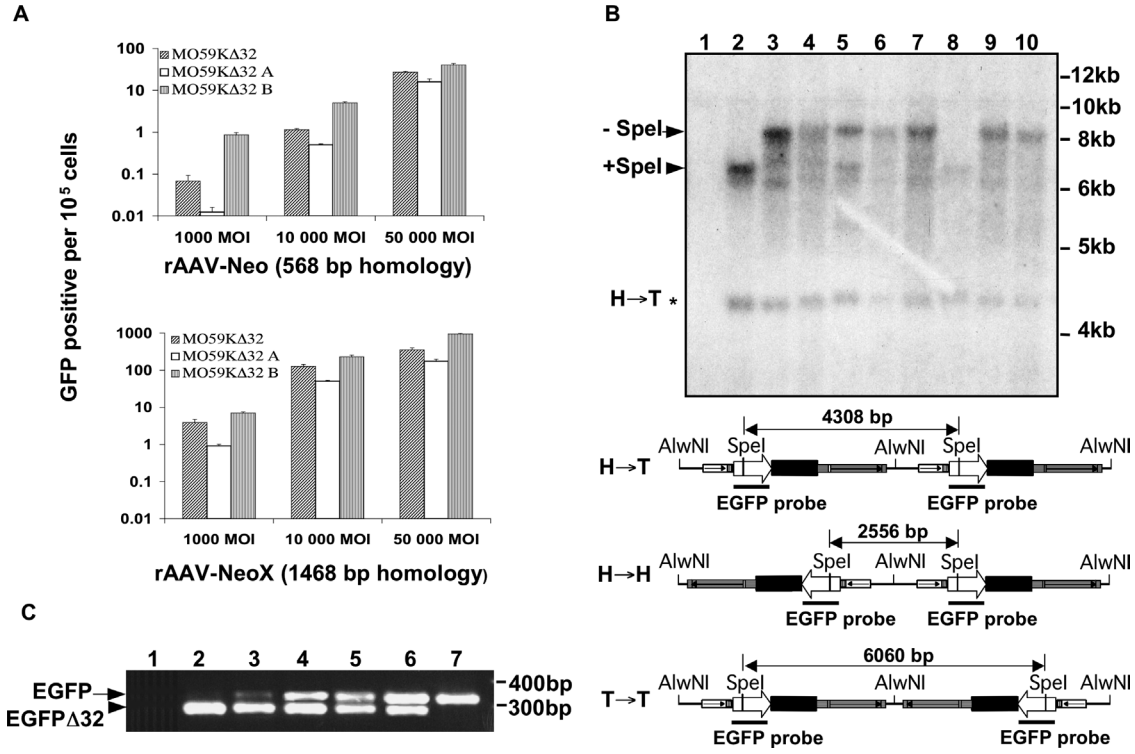


**Figure 2.** Evidence of accurate gene correction. (**A**) Optimization of rAAV infection. A total of $3 \times 10^5$ cells of MO59KΔ32 polyclonal population (diagonal stripes), clone MO59KΔ32 A (white columns) and clone MO59KΔ32 B (vertical stripes) were infected with rAAV-Neo (upper chart) or rAAV-NeoX (lower chart) at the indicated MOI. The virus contaning medium was removed 24 h later. EGFP positive cells were counted by flow cytometry on day 2. (**B**) Southern analysis of SpeI digested genomic DNA with a probe consisting of full-length EGFP. Wild-type MO59K cells (lane1) served as a negative control. DNA from uninfected MO59KΔ32 B clone was loaded in lane 2. Lanes 3 to 10 contain genomic DNA from individual EGFP positive clones derived from polyclonal rAAV-NeoX transduced population. DNA ladder sizes are shown to the right. A schematic bellow illustrates the variety of fragments that may arise from head-to-tail (H-T *), head-to-head (H-H) or tail-to-tail (T-T) dimers of the targeting substrate integrated in the same genomic locus of MO59KΔ32 B. The band arising from digestion of the EGFPΔ32 mutant with SpeI is marked by an arrowhead ('+SpeI'). The fragment containing repaired EGFP that lacks SpeI restriction site is marked with an arrow ('−Spe I'). (**C**) PCR of genomic DNA amplifying 320 bp of corrected EGFP and 288 bp of mutant EGFPΔ32 fragment. Lane 3 contains the products from rAAV-NeoX transduced polyclonal population. Lanes 4–6 represent product amplified from single clones derived from the same polyclonal population. DNA from wild-type MO59K cells (lane 1) and EGFP expressing cell line (lane 7) were used as a negative and positive control, respectively.

Previous report defined 150 bp of 5′ homology and 500 bp of 3′ homology as sufficient for efficient targeting (19). Taking this into account the vectors were constructed with 174 bp of homology to the target upstream of the mutation, and 568 to 1468 bp of downstream homology (Figure 1). It has been suggested that the efficiency of gene targeting is proportional to the virus load (29), and for most experimental systems the optimal MOI was defined empirically. We determined the optimal experimental conditions for gene targeting in the MO59KΔ32 cell line by transducing the polyclonal population and two individual clones expressing the EGFPΔ32 mRNA—MO59KΔ32 A and MO59KΔ32 B. The targeting rates were measured by counting the number of GFP positive cells 2 days after infection with either rAAV-Neo or rAAV-NeoX at an increasing MOI (Figure 2A). The infectivity of these two vectors was found to be the same as judged indirectly by transduction with rAAV-RFP or rAAV-RFPX (data not shown), which have an RFP expression cassette replacing the Neo gene expression (Figure 1C). In both polyclonal population and clones the increase of homology between the viral vector and the genomic target substantially enhanced the targeting efficiency. The difference between the short 3′-homology vector (rAAV-Neo) and the long 3′-homology vector (rAAV-NeoX) at an MOI of 1000 was 50- to 70-fold. In comparison, the increase in targeting rates was 46- to 100-fold at an MOI of 10 000 and 10- to 24-fold at an MOI of 50 000 when rAAV-NeoX was used instead of rAAV-Neo. At the same time, increasing the vector dose from 10 000 to 50 000 MOI improved the targeting frequency 8 to 32 times for rAAV-Neo and up to four times for rAAV-NeoX. This suggests that increased homology improves targeting efficiencies and that this effect is partially masked when higher virus load is applied to the cells. The lower targeting frequency detected in the polyclonal population compared to clone MO59KΔ32 B was expected as suggested by the northern data (Supplementary Figure 1C). The maximum EGFP repair frequency was 0.95% obtained with 50 000 MOI of rAAV-NeoX infection of the MO59KΔ32 B clone, which made it a good candidate for subsequent silencing experiments side by side with the polyclonal MO59KΔ32 population.

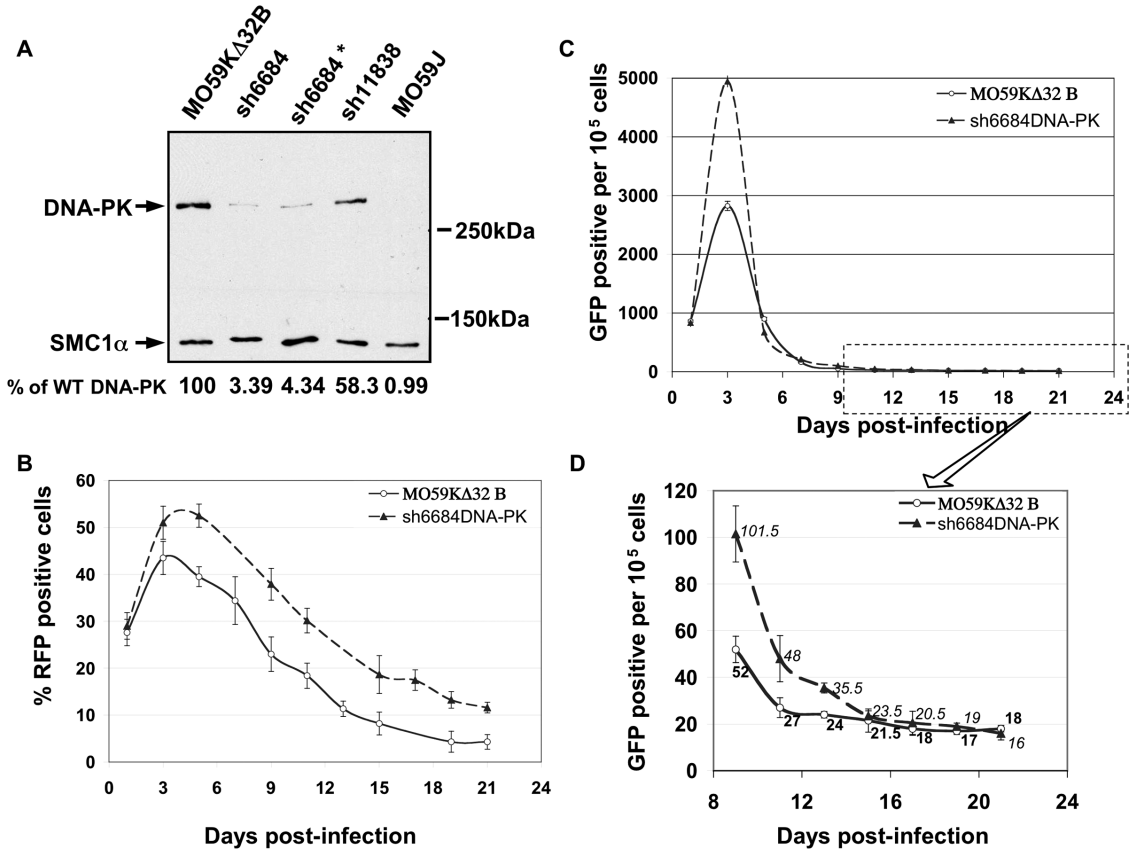


**Figure 3.** DNA-PKcs is not required for stable rAAV-mediated targeting. (**A**) Immunoblot analysis of specific RNAi-mediated silencing of DNA-PKcs in MO59KΔ32 cells. Lane sh6684 and sh6684* depict the downregulation of DNA-PKcs expression pre- and post-infection, respectively. Sh11838 represents inefficient downregulation of expression at the respective position of the mRNA. Nuclear extract from DNA-PKcs mutant MO59J cell line (derived from MO59K) was used as a negative control. The relative level of residual protein expression with respect to MO59KΔ32 cells is shown at the bottom of each lane. (**B**) Time-course analysis of rAAV infection. $3 \times 10^5$ cells were infected with rAAV-RFPX and the virus-containing medium removed 24 h later. Cells were passaged and the percentage of RFP positive cells determined by FACS in fraction of the cells. (**C** and **D**) Time-course analysis of rAAV gene targeting. $3 \times 10^5$ cells were infected with 50 000 MOI of rAAV-NeoX. The virus was removed 24 h later and half of the cells assayed by FACS for GFP expression. The rest of the cells were re-plated and subsequently used for analysis every other day.

## Repair of the genomic EGFPΔ32 target occurs with high fidelity

To analyze the outcome of the targeting events at a molecular level, GFP positive cells were separated from the total transduced MO59KΔ32 B population by single-cell sorting at day 5 post-infection. Genomic DNA from individual expanded colonies was analysed by Southern blot comparing the founder MO59KΔ32 B to targeted clones. SpeI-digested genomic DNA identified MO59KΔ32 B (lane 2) as a multicopy clone with respect to the targeting substrate (Figure 2B). Upon such restriction digest a single copy of EGFPΔ32 should produce two bands when hybridized to a full-length EGFP probe. In MO59KΔ32 B the 4.3 kb band corresponds to a SpeI-digested head-to-tail dimer integrated in the genome. The low intensity of the 6.2 kb band suggests it represents the 'head' part of the head-to-tail insert. This part shares only 190 bp of homology with the probe and is unlikely to hybridize strongly to the probe labeled by the random primer method. The more intense 6.8 kb upper band is likely to represent the 'tail' part of the digested head-to-tail dimer, which contains the remaining 500 bp of the EGFP gene thus producing much stronger signal (Figure 2B schematic). This band could also correspond in size with an additional copy of the target integrated independently elsewhere in the genome thus increasing the band intensity. However, multiple copies of the target cannot reconstitute functional GFP unless an appropriate rAAV vector introduces the correct sequence. This notion is confirmed by the SpeI digest of genomic DNA from EGFP positive targeted clones, where all clones show a substantial size-shift of a fragment resulting in a new band of ~8.2 kb, while some still retain the 6.8 kb band albeit with lower intensity. This shift indicates loss of the SpeI restriction site on a repaired EGFPΔ32 target. The appearance of an 8 kb instead of 8.2 kb band-shift in lane 10 probably represents a truncation that accompanied the repair event. The absence of the 8.2 kb fragment in one of the analyzed clones indicates lack of EGFPΔ32 repair (Figure 3B, lane 8). This clone was selected as GFP positive at day 5 after infection, but the fluorescence was lost after expansion. This suggests that a repair of the mutant Δ14EGFP gene present on rAAV episomes may occur instead of the genomic EGFPΔ32 target DNA. Subsequently, the repaired episome could be degraded or excluded during cell division. An additional band of about 5.3 kb in lane 5 represents an integration of virus DNA. Aided by the multicopy nature of the targets in the analyzed

clone, we demonstrate the complexity of events accompanying rAAV-mediated targeting. Notably, there was no change in the size of the head-to-tail dimer band suggesting that it is not a preferred target for the repair reaction and that most probably single-strand annealing is an unlikely mechanism of rAAV-mediated gene repair.

Next, the faithful repair of the EGFPΔ32 genomic target was confirmed by PCR analysis and sequencing of the amplified products. The primer design allowed for the amplification of both wild-type EGFP (320 bp fragment) and the mutant EGFPΔ32 (288 bp fragment) only from the integrated substrate and not from the viral DNA. As evident from agarose gel electrophoresis (Figure 2C) and sequencing (Supplementary Figure 2) of the PCR products, individual targeted EGFP positive clones contained the corrected fragment of the EGFP gene, i.e. lacking the 32 bp deletion. Together, these results demonstrate efficient and accurate repair of the EGFPΔ32 target with rAAV vectors.

### DNA-PK is not required for rAAV-mediated gene targeting

The unique hairpin-shaped ends of the AAV genome are essential for rAAV gene targeting (19). A number of reports show that DNA-dependent protein kinase (DNA-PKcs) mediates the removal of the free DNA ends generated by rAAV molecules by promoting their ligation into circular intermediates in muscle and liver cells (17,32). In addition, integration of AAV targeted to the AAVS1 site is increased after antibody-depletion of DNA-PKcs (33). Other components of the NHEJ pathway, such as KU86, have also been found to assist in the processing of the rAAV genome in cultured cells (18). Based on these observations we considered a role for DNA-PKcs in rAAV-mediated gene targeting. Since NHEJ and HR may be competing (34), we hypothesized that reduction of DNA-PKcs expression levels could result in shift of the DSB repair pathways from NHEJ to HR. In order to address any role of DNA-PKcs in rAAV targeting, expression of DNA-PKcs in MO59KΔ32 cells was downregulated by RNAi. For short-term silencing of gene expression both siRNA transfection and vector-mediated expression of shRNA have proven effective. The analysis of rAAV-mediated gene targeting was performed over an extended period of time, which required stable transcriptional knock-down of the gene. Therefore, populations with reduced expression of DNA-PKcs of MO59KΔ32 polyclonal population and the MO59KΔ32 B clone were generated by transfection with linearized retrovirus vector expressing shRNA from the RNA pol III promoter H1. Following selection on puromycin, nuclear extracts from these cells were assayed by immunoblot for the presence of DNA-PKcs, and for the constitutively expressed SMC1α (27) used as a loading control (Figure 3A). There was a dramatic reduction in the protein level of DNA-PKcs reduced to 3.4% of the non-silenced MO59KΔ32 B cells, using the shRNA starting at position 6684 of the mRNA. This silencing effect was stable for more than one month as confirmed in nuclear extracts of cells infected with rAAV-NeoX (Figure 3A). The shRNA at position 11 838 has been shown previously to be efficient in the form of transfected siRNA (25). However, this sequence was not as efficient in our system and yielded only 40% reduction in protein levels of DNA-PKcs.

The rates of gene targeting were compared in silenced and normal MO59KΔ32 (Supplementary Figure 3A) and MO59KΔ32 B cells (Figure 3). At day 0, $3 \times 10^5$ cells were transduced with 50 000 MOI of rAAV-NeoX. Alternatively, as a measure of infectivity, the same number of cells were transduced with rAAV-RFPX. Following infection, the number of GFP positive cells generated by gene targeting, was measured by flow cytometry in a fraction of the transduced cells. Gene correction frequencies were represented as the number of fluorescent cells per $10^5$ live cells counted. The infectivity of silenced MO59KΔ32 B cells was marginally higher, but not statistically significant at a transduction efficiency of ~50% (Figure 3B). The highest number of GFP positive cells was observed at day 3 after infection, i.e. 5% in cells with silenced DNA-PKcs and 3% in normal MO59KΔ32 B cells (Figure 3C). This number declined rapidly over time stabilizing around two weeks after infection to about 0.02% for both cell populations (Figure 3D). The large difference between the number of GFP positive cells at day 3 and day 14 probably reflects gene targeting that repairs the Δ14EGFP mutant present on the rAAV episomes.

### Preferential targeted repair of the GFP mutation carried by the virus vector

Data in Figure 3C show that gene targeting frequencies decline rapidly after infection. Theoretically, targeted repair of a mutant may occur with the same probability either on the genomic target or on the virus DNA. The kinetics of these two events can be monitored separately in our system (Figure 4A). Repair of the genomic target results in strong fluorescence from EGFP expression driven by the hCMV promoter and early enhancer (Figure 4A; dashed GFP gate). In contrast, weak EGFP expression of the repaired Δ14EGFP mutant from the incomplete CBA promoter present in the vector (see Materials and Methods) produced low fluorescence intensity (undetectable by microscopy) (Figure 4A; solid GFP gate). While up to day 7 the contribution of rAAV episomes to the total number of fluorescent cells was significant, targeting rates stabilized slowly thereafter and the frequency of genomic target repair remained stable. To evaluate the physical presence of EGFP expressing rAAV episomes we isolated low molecular weight DNA from a fraction of cells transduced with 10 000 MOI rAAV-NeoX or rAAV-RFPX at day 5. This DNA, a mixture of single-stranded and double-stranded rAAV genomes, was transfected in 293T cells and GFP fluorescence was assayed by flow cytometry 2 days later. Transfection with low molecular weight DNA prepared from a cell line, which stably expressed EGFP and was transduced with rAAV-SEAP served as a negative control for potential contamination with genomic DNA. Although ~40% of the cells infected with rAAV-RFPX expressed RFP, only $4.5 \times 10^{-3}$% of the transfected cells were RFP positive (Figure 4B). Hence it was not surprising that a gene targeting frequency of 0.93% translated in $4.33 \times 10^{-4}$ % of 293T GFP positive cells (Figure 4C). This assay only enabled visualization of episomal EGFP expression, but not direct quantitation of rAAV

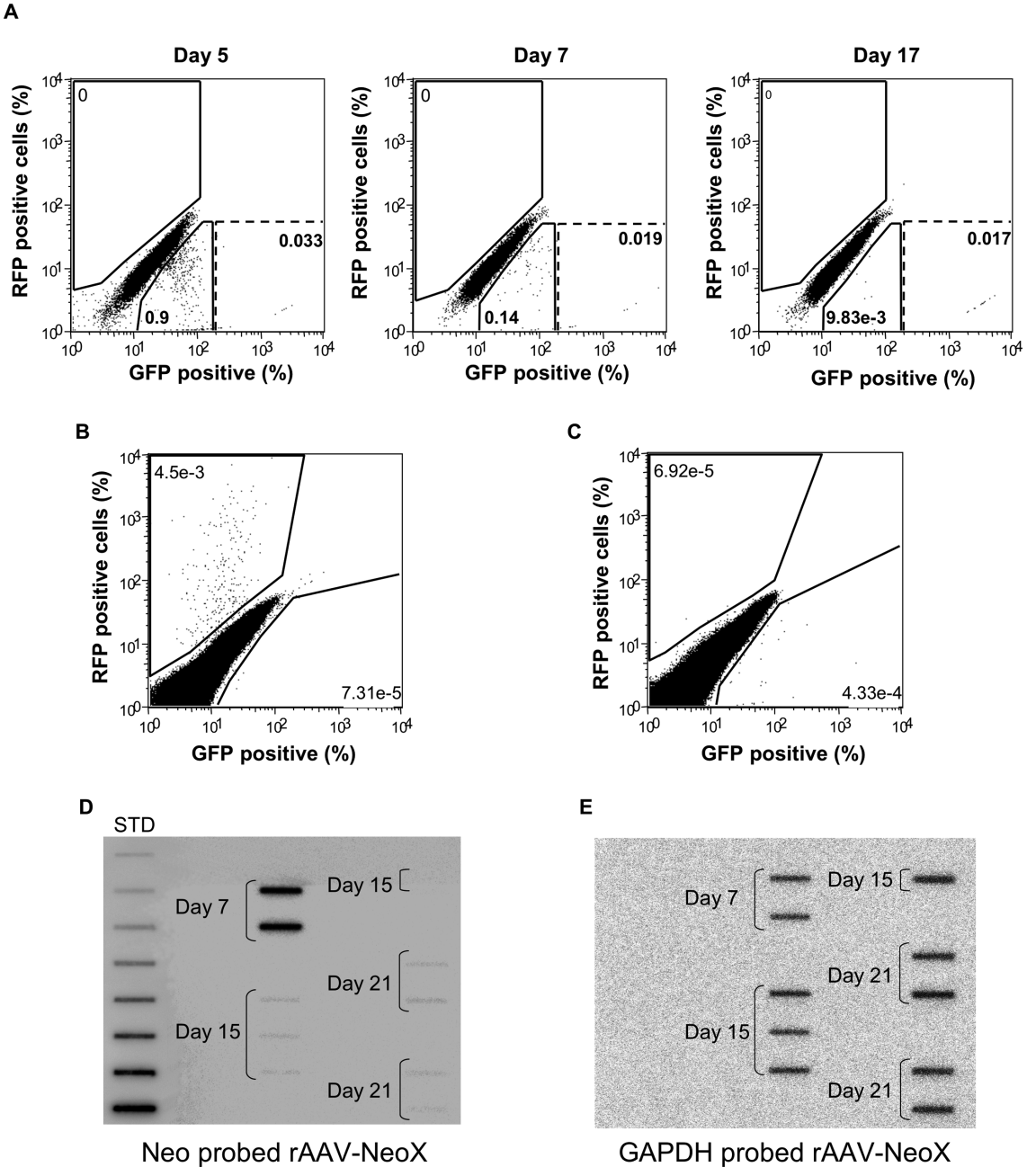


**Figure 4.** Rapid removal of rAAV episomes results in reduced targeting rates. (**A**) FACS analysis of gene targeting in MO59KΔ32 cells at the indicated days after transduction with rAAV-NeoX. GFP positive (%) gates define the number of cells with low and high fluorescence intensity of GFP. Targeted events occurring on the virus DNA are represented by cells exhibiting low intensity of fluorescence (full line) and repair of the genomic target is measured by cells exhibiting strong fluorescence (interrupted line). (**B**) FACS analysis of 293T cells transfected with low molecular weight DNA isolated from rAAV-RFPX infected cells. (**C**) FACS analysis of 293T cells transfected with low molecular weight DNA isolated from rAAV-NeoX infected cells. (**D**) and (**E**) Slot-blot analysis of of rAAV DNA isolated from total cell lysates at day 7, day 15 and day 21 post-infection. To quantify rAAV-NeoX DNA (**D**) a reference plasmid pTR-NeoX was used for standard curve (STD) and the blot was hybridized with Neo probe. (**E**) Slot-blot membrane re-probed with a GAPDH probe to quantify relative cell numbers.

genomes. To evaluate episome persistence total lysates from cells transduced with 50 000 MOI of rAAV-NeoX were prepared at one, two and three weeks post-infection. Aliquotes were spotted onto a membrane using a slot-blot adaptor and rAAV DNA was quantified by comparison with serial dilutions of a plasmid control by hybridization with a vector probe or a control probe (Figure 4D and E). The data from this assay are summarized in Table 1. As evident from the results, there were 215 viral DNA molecules per cell at day 7 and this number decreased drastically by day 15 to almost undetectable at day 21. However, the gene targeting efficiency remained constant after two weeks of culture and thus was sufficient to observe differences after silencing of recombination proteins.

**Table 1.** Stable gene targeting rates (starting at day 15) persist independently from the episomal rAAV-NeoX DNA

| Cell line | Day 7 Targeting events per $10^5$ cells | rAAV DNA/cell | Day 15 Targeting events per $10^5$ cells | rAAV DNA/cell | Day 21 Targeting events per $10^5$ cells | rAAV DNA/cell |
|---|---|---|---|---|---|---|
| MO59KΔ32 | 162.5 | 215.76 | 21.5 | 5.28 | 18 | 0.511 |
| sh6684DNA-PK | 210.0 | 238.90 | 23.5 | 8.48 | 16 | 4.016 |

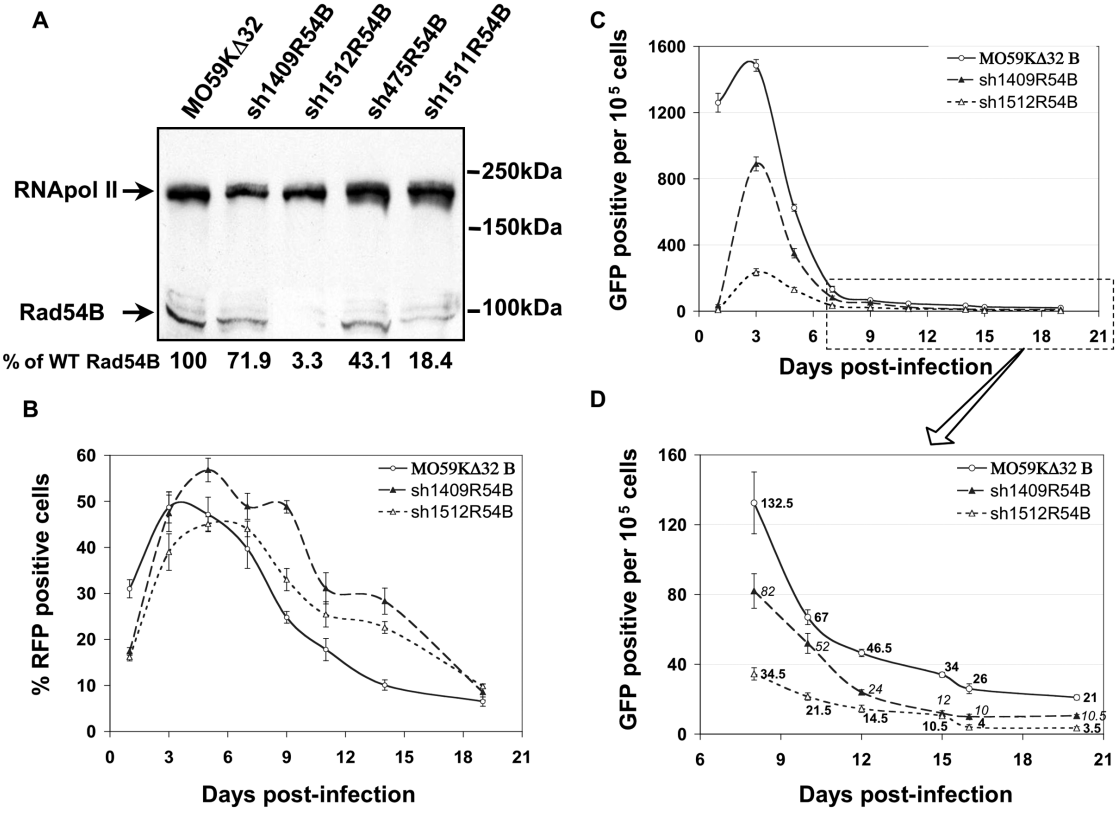


**Figure 5.** Gene targeting frequency of rAAV corellates with RAD54B expression levels. (**A**) Immunoblot analysis of RAD54B-specific silencing. MO59KΔ32 designates the founder EGFPΔ32 expressing cell line. sh1409R54B, sh1512R54B, sh475R54B and sh1511R54B indicate the start position of the shRNA oligo on the Rad54B mRNA. The relative level of residual protein expression with respect to the founder cells is shown at the bottom of each lane. (**B**) Monitoring of infection efficiency. Cells were transduced with rAAV-RFPX and the number of RFP positive cells was counted by flow cytometry in aliquots from passaged cells. (**C** and **D**) Time-course analysis of rAAV gene targeting. (D) shows a scaled-up version of the later time-points from (C).

## rAAV-mediated targeting rates correlate with RAD54B levels

Based on the above data indicating no requirement for DNA-PKcs in the targeting process, we elucidated possible factors from the HR pathway that are important for rAAV targeting. Most of the proteins in the RAD52 epistasis group were recognized by their ability to aid homology-directed repair of DNA DSBs induced by ionizing radiation [reviewed in (35)]. There are no reports on the role of these proteins in intracellular processing of the rAAV genome. Cell lines where one of the members of this group, Rad54B, has been eliminated, display greatly reduced frequency of gene targeting using linear plasmid DNA as compared to the wild-type cells (36). Therefore, we examined the potential role of RAD54B in the process of rAAV-mediated gene targeting. RNAi-mediated silencing of RAD54B has no effect on cell growth or cell survival. Four different shRNA sequences were tested for their efficiency in silencing the expression of RAD54B and the amount of remaining protein was assessed in nuclear extracts by immunoblot (Figure 5A). RNA pol II was used as a reference for protein loading.

The most potent oligonucleotide (starting position 1512 of the RAD54B mRNA) reduced expression levels to 3.3% of RAD54B expressed in the non-silenced MO59KΔ32 cells. The infectivity of silenced and non-silenced cell populations was assessed by transduction with rAAV-RFPX (Figure 5B). We compared the gene targeting frequency in non-silenced MO59KΔ32 B and derivatives of the same silenced with sh1512 (Figure 5C and D and Supplemetary Figure 3C and D). The targeting efficiency in cells with strongly silenced RAD54B was more than 6-fold lower than the frequency in non-silenced cells. To address whether the rate of gene targeting may depend on the level of protein expression we carried

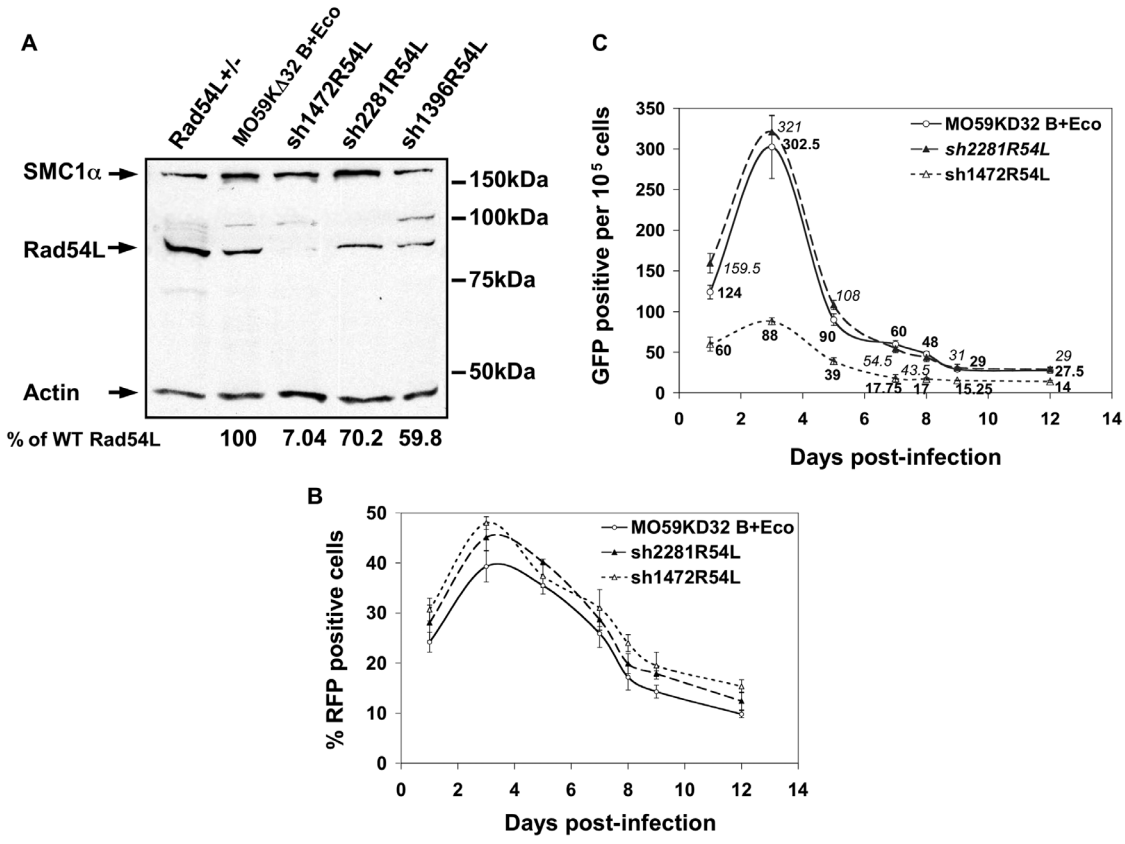


**Figure 6.** Strong inhibition of RAD54L expression leads to reduced rAAV targeting frequency. (**A**) Immunoblot analysis of Rad54L-specific silencing. MO59KΔ32 B+Eco designates the founder MO59KΔ32 B clone expressing Ecotropic receptor. sh1472R54L, sh2281R54L and sh1396R54L indicate the start position of the retrovirally expressed shRNA oligo on the RAD54L mRNA. SMC1α and β-actin were used as loading controls and nuclear extract from mouse ES Rad54± cells was used as a RAD54 size-reference. (**B**) Time-course analysis of rAAV infection. MO59KΔ32 B+Eco cells and RAD54L silenced polyclonal populations were transduced with rAAV-RFPX to test infectivity. The fraction of cells expressing RFP was determined every other day in an aliquot of passaged cells. (**C**) Gene targeting of MO59KΔ32 B+Eco and RAD54L silenced cells. $5 \times 10^5$ cells were transduced with 50 000 MOI of rAAV-NeoX and GFP positive cells detected by flow cytometry on alternate days starting 24 h after infection.

out additional gene targeting experiments with a polyclonal population with 71.9% residual RAD54B expression. Here, the targeting frequency with respect to non-silenced cells was reduced by 1.7- to 2-fold throughout the course of infection. The correlation between the protein levels of RAD54B and the frequency of gene targeting suggests the importance of HR for efficient targeting.

## RAD54L is necessary for efficient targeting

Disruption of Rad54 in mouse embryonic stem-cells and chicken DT40 cells impairs HR and results in sensitivity to ionizing radiation (37). Since our data show that RAD54B is involved in the targeting process we hypothesized a role for its homolog RAD54L, which is also a member of the SWI2/SNF2 subfamily of DNA-dependent ATPases. To test this hypothesis we used RNAi for silencing of RAD54L. To further improve the generation of silenced polyclonal populations we modified the MO59KΔ32 B clone by introducing an ecotropic receptor that enables retroviral entry. The new cell line MO59KΔ32 B+Eco was infected with retroviruses carrying shRNA designed to induce cleavage of the RAD54L mRNA. Retrovirus MSCV-GFP was used to monitor the efficiency of retrovirus infection that was for the most part ~50% (data not shown). RAD54L protein levels were determined by immunoblot of nuclear extracts from puromycin resistant polyclonal populations and compared to that of non-silenced cells. Both actin and SMC1α served as loading controls. Due to appearance of additional bands in human cells a nuclear extract from Rad54 heterozygous mouse ES cells was used to identify the RAD54L protein (83.4 kDa). Maximum reduction of RAD54L expression was achieved by retrovirus infection with MSCV-sh1472 (Figure 7A), which reduced RAD54L levels to 7% of that in non-silenced cells. As evident from data obtained after transduction with rAAV-RFPX, untreated and RAD54L silenced cells have the same infectivity (Figure 6B). We measured the frequency of gene targeting in MO59KΔ32 B+Eco and silenced population after transduction with 50 000 MOI of rAAV-NeoX (Figure 6C). Similar to RAD54B, strong silencing of RAD54L lead to impaired rAAV-mediated targeting, which translated to 3.5-fold reduction of targeting rates at the beginning of infection and stabilized at 2-fold by day 9 after infection. Weak silencing with sh2281 (70.2% of residual RAD54L expression) yielded no difference between the targeting rates in the founder cell line and its silenced counterpart. However, the lack of RAD54L had a pronounced effect on the fidelity of gene

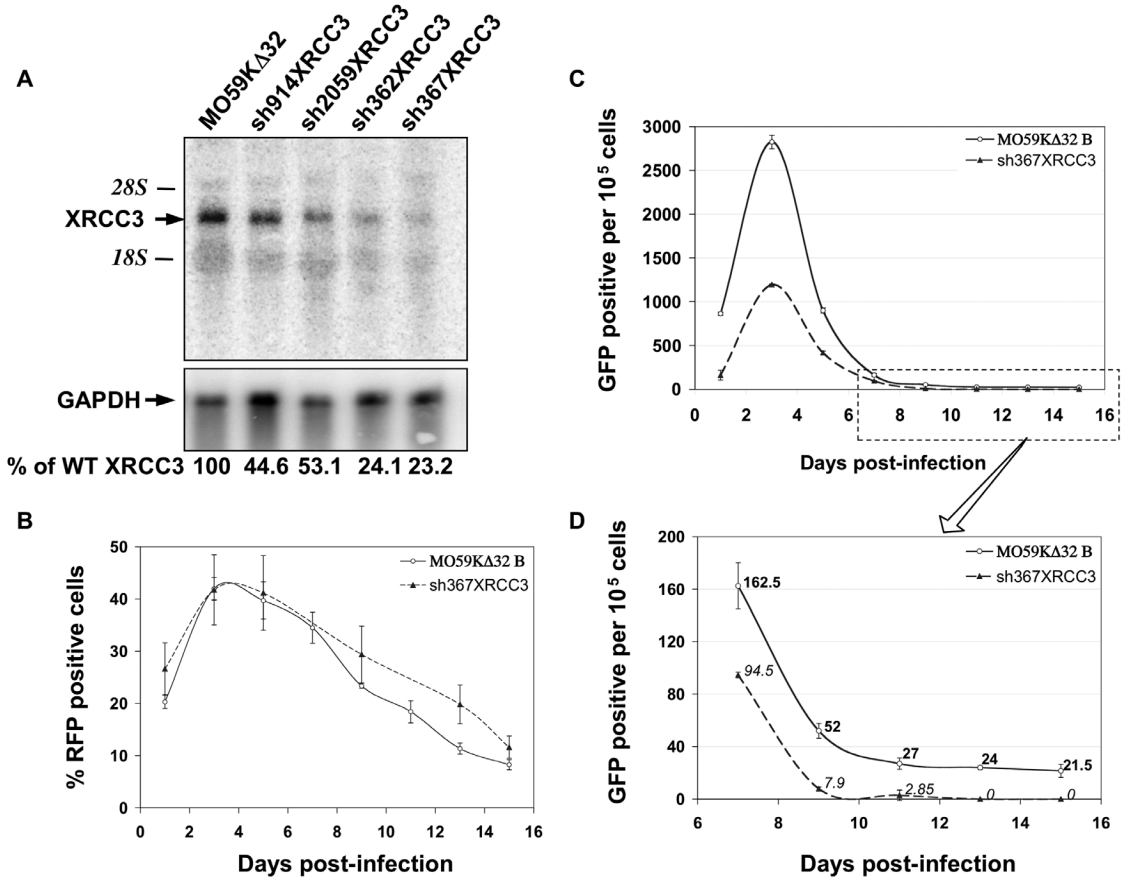


**Figure 7.** Suppression of XRCC3 abolishes long-term gene correction. (**A**) Northern analysis of XRCC3 silencing. MO59KΔ32 designates the founder cell line and sh914, sh2059, sh362 and sh367 indicate the type of shRNA used to downregulate XRCC3 expression in the respective puromycin resistant population the total RNA was isolated from. GAPDH was used as an internal control to normalize for loading. (**B**) Comparison of infectivity of silenced and non-silenced cells. MO59KΔ32 B and their sh367XRCC3 derivative were transduced with rAAV-RFPX. As a measure of infection efficiency the level of RFP expression in the total cell population was monitored on alternate days. (**C**) Gene targeting in non-silenced and XRCC3 silenced cells. MO59KΔ32 B and sh367XRCC3 cells were transduced with 50 000 MOI of rAAV-NeoX. The GFP positive cells were detected by flow cytometry every other day for 15 days after infection. (**D**) Enlarged version of (C) spanning day 7 to day 15 after infection, which enables viewing the number of stable targeting events in detail.

repair. Sequencing of the 'repaired' EGFP gene from the genomic DNA of sorted cells positive by FACS revealed high rate of base substitutions—three in the stretch of 36 bp (Supplementary Figure 2B). This observation indicates that in the absence of RAD54L and probably other HR proteins rAAV utilizes other DNA mechanisms of interaction with the cellular genome.

## XRCC3 deficiency blocks stable gene correction by rAAV

The above results indicate the dependence of rAAV-mediated targeting on components of the HR pathway of DNA DSB repair. For a factor further upstream in the Rad52 pathway, we silenced expression of the Rad51 paralog XRCC3, the absence of which did not affect the growth rate of irs1SF cells, a CHO-derived line (38).

MO59KΔ32 cells were transfected with shRNA expressing vectors targeting the XRCC3 mRNA at positions 362, 367, 914 and 2059. Since no reliable antibodies for XRCC3 detection by immunoblot were available to us, the level of downregulation of XRCC3 expression was determined by northern analysis of total RNA from puromycin resistant polyclonal populations (Figure 7A). The housekeeping gene GAPDH was used as an internal loading control. The most efficient silencing was achieved with the shRNA starting at position 367, which reduced XRCC3 mRNA level to 23.2% from that of untreated cells, and this silenced population was chosen for subsequent gene targeting experiments. Transduction of silenced and non-silenced cells with rAAV-RFPX showed they possess the same rate of infectivity (Figure 7B). Silencing of XRCC3 also did not affect the cell growth. A gene targeting experiment was carried out in parallel by transduction of silenced and non-silenced cells with 50 000 MOI of rAAV-NeoX (Figure 7C and D). Decrease of XRCC3 mRNA had a negative effect on both short- and long-term gene targeting, which was much more pronounced at later time-points. While the decrease in targeting efficiency was about 3-fold at days 3 and 5 after the infection (Figure 7C), it amounted to 10-fold at day 11 and there were no GFP positive cells after day 13 post-infection (Figure 7D). This rather dramatic effect of XRCC3 deficiency on rAAV targeting is consistent with the above results suggesting that the RAD51/RAD54 HR pathway controls this process.

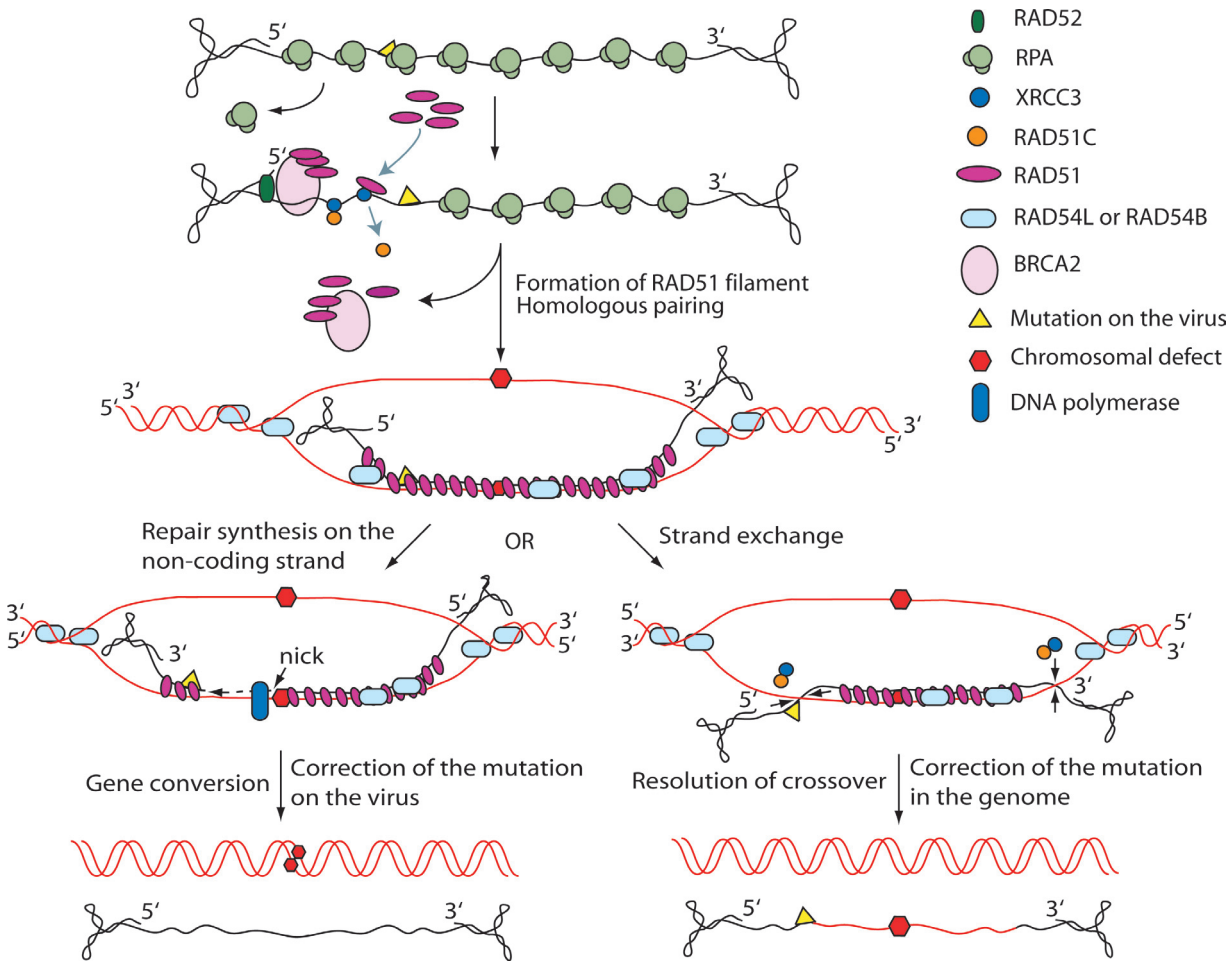


**Figure 8.** Possible mechanism of gene targeting by rAAV. After the viral capsid has been removed RPA may bind to the single-stranded AAV genome and support initiation of HR. The BRCA2-RAD51 complex, assisted by RAD51 paralogs, such as XRCC3, replaces RPA with RAD51. The nucleoprotein filament finds and pairs with the homologous chromosomal sequence via unwinding/supercoiling facilitated by RAD54. Non-crossover gene conversion by repair DNA synthesis can occur either on the virus DNA (as shown here) or on the chromosomal target. Alternatively, resolution of vector-chromosomal DNA intermediates (arrows mark the two possible options) results in introduction of targeted modification in the homologous chromosome by crossing over and exchange of sequences. Only two of four possible outcomes are shown. The images may not represent the actual sequence of events.

## DISCUSSION

Understanding the mechanisms underlying the process of rAAV-mediated gene targeting shown previously to reach unusually high frequencies (29,39) will likely be beneficial for gene therapy applications and will contribute to our knowledge of virus-host interactions and DNA dynamics. We utilized RNA interference to examine the role of cellular proteins involved in DNA repair and recombination pathways for rAAV-mediated gene targeting. Compared with the non-silenced parent populations, all polyclonal silenced populations proliferated with the same rate and contained similar numbers of RFP positive cells indicative of infection. Therefore, the silencing procedure does not seem to affect parameters critical for a valid comparison of targeting rates, such as the kinetics of rAAV transduction, surface receptor expression, or growth properties of the cells.

Several types of mutations have been efficiently corrected using rAAV vectors [reviewed in (9)]. In our system, rAAV targeting vectors achieve high fidelity correction of a 32 bp deletion at a stable, long-term targeting frequency of 0.02% consistent with previous observations made under non-selective conditions. Also consistent with earlier reports, we observed significant improvement in the frequency of repair with the increase of the vector load. Besides efficient delivery of vector DNA, the length of homology between vector and target proved important. A 2-fold extension of the 3′ homology between the chromosomal substrate and the recombinant viral vector resulted in an increase in the repair frequency of up to 73-fold at low, 101-fold at intermediate and 23-fold at high vector dose. The length of homology that can be included in the rAAV system is limited by the packaging capacity of the virus, which is about 4.7 kb, and for maximal targeting efficiency at low MOI the longest possible homology should be used. Still, the rAAV system requires less extended homology than conventional targeting systems at mostly higher targeting frequencies.

We have tested the effect of silencing several genes that are key to NHEJ and HR, including DNA-PK, RAD54L, RAD54B or XRCC3. Cell lines deficient in either Rad54L

or Rad54B or carrying non-functional mutants of DNA-PK and XRCC3 are viable (24,36,38,40), rendering RNA interference feasible.

Absence of DNA-PK reduces circularization of linear dsrAAV genomes by about 50% (17) and one member of the DNA–PK complex, KU86, apparently binds to rAAV DNA, and acts as an inhibitor of rAAV transduction (18). In this report, we demonstrate that rAAV gene targeting does not depend on NHEJ since transcriptional silencing of DNA-PKcs does not affect targeting frequencies. It is likely that the higher number of GFP positive cells in the DNA-PKcs silenced population observed at day 3 (Figure 3C) is due to reduced circularization and abundance of linear rAAV monomer genomes, which have been shown to be effective in targeting (19).

However, we identified components of the RAD51/RAD54 pathway of HR as central to rAAV-mediated gene targeting. We found that reduced protein levels of the two Rad54-homologous genes in human cells, RAD54L and RAD54B result in decreased stable rAAV-mediated targeting rates by 2- and 6-fold, respectively. RAD54L and RAD54B are members of the RAD52 epistasis group and belong to the SNF2/SWI2 family of proteins that dissociate and remodel protein complexes on DNA (40,41). Our data are consistent with up to 10-fold reduced targeting frequencies in colon cancer cell lines with inactivated RAD54B (36). RAD54L facilitates strand-exchange by RAD51, which binds to RAD54L and stimulates its ATPase activity (42). Since the yeast homologue of RAD54B, TID1/RDH54 acts in the same recombinational repair pathway as RAD54L via roles partially overlapping with those of RAD54 the residual levels of targeting we observed in the absence of either protein were to be expected. Lack of reports on the viability and proliferative capacity of cells deficient in both RAD54L and RAD54B, renders the outcome of a potential double silencing experiment to test this functional redundancy in HR very uncertain.

RAD51 plays a central role in HR by replacing RPA from single-stranded DNA and forming a nucleoprotein filament that participates in homologous pairing (35). Disruption of the RAD51 gene leads to cell death (43), whereas RAD51 paralogs including XRCC3 are not essential for cell viability. The XRCC3-deficient X-ray sensitive hamster cell line, irs1SF shows increased chromosome mis-segregation (44) and has a 25-fold decrease in the frequency of error-free homology-directed repair of DNA DSBs (38). Both can be restored to wild-type levels through XRCC3 expression (44). However, in contrast to RAD51, XRCC3 is known to act in the early as well as late stages of HR (45,46). Following reduction of XRCC3 mRNA levels we observed nearly complete elimination of rAAV-mediated gene targeting, which may reflect the multiple roles XRCC3 plays in HR. Although the reduction in XRCC3 mRNA was about 80% compared to mRNA levels in non-silenced cells, the remaining XRCC3 protein may be below a critical concentration. A complex between XRCC3 and another RAD51 paralog, RAD51C associates specifically with single-stranded DNA (47) and modulates replication fork progression by slowing it down after DNA damage (48). The complex also plays a role probably downstream of RAD54 in the HR pathway, i.e. as a resolvase of Holliday junctions (49). Thus, resolution of recombination intermediates and/or triggering of DNA damage-related signaling present two of the possible roles for the XRCC3/RAD51C complex in rAAV targeting. A recent report reveals that XRCC3 forms nuclear foci independently of RAD51 very early during homology-directed DSB repair, associates directly with DNA, and promotes formation of the RAD51 nucleoprotein filament (45). One may hypothesize, that the absence of XRCC3 affects potential binding of RAD51 to the single-stranded AAV vector genome and subsequently its homologous pairing with the chromosomal locus during the targeting reaction. Based on this hypothesis we propose a molecular model for the mechanism of rAAV-mediated targeting in human cells (Figure 8). Removal of capsid proteins exposes the single-stranded AAV genome, allowing its degradation by nucleases. There are several processes that could prevent this degradation: annealing of positive and negative viral genomes, second-strand synthesis, or coating of the single-stranded region by high-affinity single-stranded DNA-binding proteins, such as RPA. RPA, a heterotrimer with multiple roles in DNA replication and repair, may ensure proper initiation of HR by preventing secondary structure formation from single-stranded DNA, thus supporting even coating by RAD51. It is unknown, however, whether RPA affects the structure of the rAAV ITRs. The RAD51C/XRCC3 heterodimer binds single-stranded AAV DNA at a terminal repeat site that resembles a DNA break (50), and through direct interactions between XRCC3 and RAD51 mediates the formation of the RAD51 nucleoprotein filament. RAD51 is likely derived from a mobile pool or from a BRCA2 complex, which also has a high binding affinity for ds- to single-stranded DNA transitions (51,52). Following ATP binding and hydrolysis by XRCC3 the RAD51C/XRCC3 dimer may be destabilized (53). RAD54L and probably RAD54B interact with the RAD51 filament, unwind the ds genomic target DNA, and may remove nucleosomes from the homologous locus in order to prepare for pairing with the rAAV DNA. After pairing with the homologous target the resulting intermediate can be resolved either by crossing over and Holliday junction resolution or by the predominant pathway for DSB repair in mammalian cells—non-crossover gene conversion (54,55). The rAAV DNA serves as a template for repair of the chromosomal target, however, gene conversion can occur on the virus with the genomic sequence as a template.

Our system shares significant similarity with a yeast model for gene targeting where donation of information to an unbroken chromosome is facilitated by a linearized plasmid (56). The authors suggest this process occurs via the formation of a heteroduplex intermediate and the efficiency of donation depends on the spatial distribution of heteroduplex DNA. The frequency of donation was proportional to the increased distance from the break to the mutation and always resulted from a single gene conversion event, which is very likely to be the case with rAAV targeting.

The conclusion that rAAV targeting depends on HR is in concert with recent evidence that targeting occurs with higher frequency in actively replicating cells, presumably the S phase of the cell cycle (30,57). NHEJ and HR function in a complementary, overlapping manner (58) and the cell cycle phase determines the prevalence of one or the other repair process. A predominant role for NHEJ early in the cell cycle in mammalian cells is supported for example by the

high IR sensitivity of NHEJ-defective cells in the $G_1$ phase (59), while HR seems to occur mostly during S and $G_2$/M phases (58).

A notable finding in our study is the significant decrease of GFP positive cells within 5 days after infection. This phenomenon has not yet been reported. Some of the rAAV targeting systems measured the number of events after an extended period of antibiotic selection (29). Other systems are based on staining of the cells at a particular time point after infection (16,19). In yet another system, which uses GFP as a readout in living cells, the results were given for only one time point, i.e. as positive foci at day 7 (30). The higher number of GFP positive cells at early time-points could also imply asymmetric recognition of the rAAV ITRs by a component of the cellular recombination machinery. That is, gene conversion on the viral vector instead of on the genomic target may have occurred. This may happen, for example, if a nick occurs at a hot-spot on the viral DNA within the double-stranded (60,61) or single-stranded region with subsequent invasion and copying of the chromosomal DNA sequence to repair the vector mutation. The unusual structure of the ITRs and the ss–ds transitional regions may be likely targets for nucleases.

Our results together with previous data provide the groundwork for improvement of the rAAV targeting frequency by shifting the balance from the NHEJ to the HR pathway. Since DNA-PK does not affect gene targeting, specific inhibition of DNA-PK with chemically synthesized small molecules (62), which has been shown to enhance gene conversion (63), combined with overexpression of proteins promoting HR (64) or synchronization of cells in S phase (for targeting of stem-cells) would enhance further targeting efficiencies increasing the potential for therapeutic applications.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

## ACKNOWLEDGEMENTS

The authors would like to thank Drs Nathalie Dutheil and Peter Ward for critical reading of the manuscript and valuable discussions. Thanks to Dr Nathalie Clement for providing plasmids and protocols for rAAV production. This work was supported by a grant from PrimeBiotech Inc., Paris, France. Funding to pay the Open Access publication charges for this article was provided by internal funds of the Institute of Physiological Chemistry, Dresden University of Technology.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Thomas,K.R., Deng,C. and Capecchi,M.R. (1992) High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol. Cell. Biol.*, **12**, 2919–2923.
2. Vasquez,K.M., Narayanan,L. and Glazer,P.M. (2000) Specific mutations induced by triplex-forming oligonucleotides in mice. *Science*, **290**, 530–533.
3. Rando,T.A., Disatnik,M.H. and Zhou,L.Z. (2000) Rescue of dystrophin expression in mdx mouse muscle by RNA/DNA oligonucleotides. *Proc. Natl Acad. Sci. USA*, **97**, 5363–5368.
4. Bertoni,C., Morris,G.E. and Rando,T.A. (2005) Strand bias in oligonucleotide-mediated dystrophin gene editing. *Hum. Mol. Genet.*, **14**, 221–233.
5. Tsuchiya,H., Harashima,H. and Kamiya,H. (2005) Increased SFHR gene correction efficiency with sense single-stranded DNA. *J. Gene Med.*, **7**, 486–493.
6. Ellis,J. and Bernstein,A. (1989) Gene targeting with retroviral vectors: recombination by gene conversion into regions of nonhomology. *Mol. Cell. Biol.*, **9**, 1621–1627.
7. Mitani,K. and Kubo,S. (2002) Adenovirus as an integrating vector. *Curr. Gene Ther.*, **2**, 135–144.
8. Hendrie,P.C. and Russell,D.W. (2005) Gene targeting with viral vectors. *Mol. Ther.*, **12**, 9–17.
9. Vasileva,A. and Jessberger,R. (2005) Precise hit: adeno-associated virus in gene targeting. *Nature Rev. Microbiol.*, **3**, 837–847.
10. White,S.J., Nicklin,S.A., Buning,H., Brosnan,M.J., Leike,K., Papadakis,E.D., Hallek,M. and Baker,A.H. (2004) Targeted gene delivery to vascular tissue *in vivo* by tropism-modified adeno-associated virus vectors. *Circulation*, **109**, 513–519.
11. Kohli,M., Rago,C., Lengauer,C., Kinzler,K.W. and Vogelstein,B. (2004) Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res.*, **32**, e3.
12. Samuels,Y., Diaz,L.A.,Jr, Schmidt-Kittler,O., Cummins,J.M., Delong,L., Cheong,I., Rago,C., Huso,D.L., Lengauer,C., Kinzler,K.W. *et al.* (2005) Mutant PIK3CA promotes cell growth and invasion of human cancer cells. *Cancer Cell*, **7**, 561–573.
13. Russell,D.W. (2003) AAV loves an active genome. *Nature Genet.*, **34**, 241–242.
14. Nakai,H., Montini,E., Fuess,S., Storm,T.A., Grompe,M. and Kay,M.A. (2003) AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nature Genet.*, **34**, 297–302.
15. Porteus,M.H., Cathomen,T., Weitzman,M.D. and Baltimore,D. (2003) Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol. Cell. Biol.*, **23**, 3558–3565.
16. Miller,D.G., Petek,L.M. and Russell,D.W. (2003) Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol. Cell. Biol.*, **23**, 3550–3557.
17. Duan,D., Yue,Y. and Engelhardt,J.F. (2003) Consequences of DNA-dependent protein kinase catalytic subunit deficiency on recombinant adeno-associated virus genome circularization and heterodimerization in muscle tissue. *J. Virol.*, **77**, 4751–4759.
18. Zentilin,L., Marcello,A. and Giacca,M. (2001) Involvement of cellular double-stranded DNA break binding proteins in processing of the recombinant adeno-associated virus genome. *J. Virol.*, **75**, 12279–12287.
19. Hirata,R.K. and Russell,D.W. (2000) Design and packaging of adeno-associated virus gene targeting vectors. *J. Virol.*, **74**, 4612–4620.
20. Ferrari,F.K., Samulski,T., Shenk,T. and Samulski,R.J. (1996) Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.*, **70**, 3227–3234.
21. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Plainview, NY.
22. Zolotukhin,S., Byrne,B.J., Mason,E., Zolotukhin,I., Potter,M., Chesnut,K., Summerford,C., Samulski,R.J. and Muzyczka,N. (1999) Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Ther.*, **6**, 973–985.
23. Rohr,U.P., Wulf,M.A., Stahn,S., Steidl,U., Haas,R. and Kronenwett,R. (2002) Fast and reliable titration of recombinant adeno-associated virus type-2 using quantitative real-time PCR. *J. Virol Meth.*, **106**, 81–88.
24. Lees-Miller,S.P., Godbout,R., Chan,D.W., Weinfeld,M., Day,R.S. IIIrd., Barron,G.M. and Allalunis-Turner,J. (1995) Absence of p350 subunit of DNA-activated protein kinase from a radiosensitive human cell line. *Science*, **267**, 1183–1185.
25. Peng,Y., Zhang,Q., Nagasawa,H., Okayasu,R., Liber,H.L. and Bedford,J.S. (2002) Silencing expression of the catalytic subunit of DNA-dependent protein kinase by small interfering RNA sensitizes human cells for radiation-induced chromosome damage, cell killing, and mutation. *Cancer Res.*, **62**, 6400–6404.
26. Andrews,N.C. and Faller,D.V. (1991) A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.*, **19**, 2499.

27. Stursberg,S., Riwar,B. and Jessberger,R. (1999) Cloning and characterization of mammalian SMC1 and SMC3 genes and proteins, components of the DNA recombination complexes RC-1. *Gene*, **228**, 1–12.
28. Hirt,B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.*, **26**, 365–369.
29. Russell,D.W. and Hirata,R.K. (1998) Human gene targeting by viral vectors. *Nature Genet.*, **18**, 325–330.
30. Liu,X., Yan,Z., Luo,M., Zak,R., Li,Z., Driskell,R.R., Huang,Y., Tran,N. and Engelhardt,J.F. (2004) Targeted correction of single-base-pair mutations with adeno-associated virus vectors under nonselective conditions. *J. Virol.*, **78**, 4165–4175.
31. Teschendorf,C., Warrington,K.H. Jr, Siemann,D.W. and Muzyczka,N. (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**, 3325–3330.
32. Nakai,H., Fuess,S., Storm,T.A., Meuse,L.A. and Kay,M.A. (2003) Free DNA ends are essential for concatemerization of synthetic double-stranded adeno-associated virus vector genomes transfected into mouse hepatocytes *in vivo*. *Mol. Ther.*, **7**, 112–121.
33. Song,S., Lu,Y., Choi,Y.K., Han,Y., Tang,Q., Zhao,G., Berns,K.I. and Flotte,T.R. DNA-dependent PK inhibits adeno-associated virus DNA integration. *Proc. Natl Acad. Sci. USA*, **101**, 2112–2116.
34. Allen,C., Halbrook,J. and Nickoloff,J.A. (2003) Interactive competition between homologous recombination and non-homologous end joining. *Mol. Cancer Res.*, **1**, 913–920.
35. Symington,L.S. (2002) Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.*, **66**, 630–670.
36. Miyagawa,K., Tsuruga,T., Kinomura,A., Usui,K., Katsura,M., Tashiro,S., Mishima,H. and Tanaka,K. (2002) A role for RAD54B in homologous recombination in human cells. *EMBO J.*, **21**, 175–180.
37. Essers,J., Hendriks,R.W., Swagemakers,S.M., Troelstra,C., de Wit,J., Bootsma,D., Hoeijmakers,J.H. and Kanaar,R. (1997) Disruption of mouse RAD54 reduces ionizing radiation resistance and homologous recombination. *Cell*, **89**, 195–204.
38. Fuller,L.F. and Painter,R.B. (1988) A Chinese hamster ovary cell line hypersensitive to ionizing radiation and deficient in repair replication. *Mutat. Res.*, **193**, 109–121.
39. Hirata,R., Chamberlain,J., Dong,R. and Russell,D.W. (2002) Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat. Biotechnol.*, **20**, 735–738.
40. Kanaar,R., Troelstra,C., Swagemakers,S.M., Essers,J., Smit,B., Franssen,J.H., Pastink,A., Bezzubova,O.Y., Buerstedde,J.M., Clever,B. *et al.* (1996) Human and mouse homologs of the *Saccharomyces cerevisiae* RAD54 DNA repair gene: evidence for functional conservation. *Curr. Biol.*, **6**, 828–838.
41. Hiramoto,T., Nakanishi,T., Sumiyoshi,T., Fukuda,T., Matsuura,S., Tauchi,H., Komatsu,K., Shibasaki,Y., Inui,H., Watatani,M. *et al.* (1999) Mutations of a novel human RAD54 homologue, RAD54B, in primary cancer. *Oncogene*, **18**, 3422–3426.
42. Sigurdsson,S., Van Komen,S., Petukhova,G. and Sung,P. (2002) Homologous DNA pairing by human recombination factors Rad51 and Rad54. *J. Biol. Chem.*, **277**, 42790–42794.
43. Lim,D.S. and Hasty,P. (1996) A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Mol. Cell. Biol.*, **16**, 7133–7143.
44. Griffin,C.S., Simpson,P.J., Wilson,C.R. and Thacker,J. (2000) Mammalian recombination-repair genes XRCC2 and XRCC3 promote correct chromosome segregation. *Nature Cell. Biol.*, **2**, 757–761.
45. Forget,A.L., Bennett,B.T. and Knight,K.L. (2004) Xrcc3 is recruited to DNA double strand breaks early and independent of Rad51. *J. Cell. Biochem.*, **93**, 429–436.
46. Brenneman,M.A., Wagener,B.M., Miller,C.A., Allen,C. and Nickoloff,J.A. (2002) XRCC3 controls the fidelity of homologous recombination: roles for XRCC3 in late stages of recombination. *Mol. Cell.*, **10**, 387–395.
47. Masson,J.Y., Tarsounas,M.C., Stasiak,A.Z., Stasiak,A., Shah,R., McIlwraith,M.J., Benson,F.E. and West,S.C. (2001) Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Genes Dev.*, **15**, 3296–3307.
48. Henry-Mowatt,J., Jackson,D., Masson,J.Y., Johnson,P.A., Clements,P.M., Benson,F.E., Thompson,L.H., Takeda,S., West,S.C. and Caldecott,K.W. (2003) XRCC3 and Rad51 modulate replication fork progression on damaged vertebrate chromosomes. *Mol. Cell*, **11**, 1109–1117.
49. Liu,Y., Masson,J.Y., Shah,R., O'Regan,P. and West,S.C. (2004) RAD51C is required for Holliday junction processing in mammalian cells. *Science*, **303**, 243–246.
50. Masson,J.Y., Stasiak,A.Z., Stasiak,A., Benson,F.E. and West,S.C. (2001) Complex formation by the human RAD51C and XRCC3 recombination repair proteins. *Proc. Natl Acad. Sci. USA*, **98**, 8440–8446.
51. Yang,H., Jeffrey,P.D., Miller,J., Kinnucan,E., Sun,Y., Thoma,N.H., Zheng,N., Chen,P.L., Lee,W.H. and Pavletich,N.P. (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science*, **297**, 1837–1848.
52. Yang,H., Li,Q., Fan,J., Holloman,W.K. and Pavletich,N.P. (2005) The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature*, **433**, 653–657.
53. Yamada,N.A., Hinz,J.M., Kopf,V.L., Segalle,K.D. and Thompson,L.H. (2004) XRCC3 ATPase activity is required for normal XRCC3-Rad51C complex dynamics and homologous recombination. *J. Biol. Chem.*, **279**, 23250–23254.
54. Johnson,R.D. and Jasin,M. (2000) Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO J.*, **19**, 3398–3407.
55. Johnson,R.D. and Jasin,M. (2001) Double-strand-break-induced homologous recombination in mammalian cells. *Biochem. Soc. Trans.*, **29**, 196–201.
56. Roitgrund,C., Steinlauf,R. and Kupiec,M. (1993) Donation of information to the unbroken chromosome in double-strand break repair. *Curr. Genet.*, **23**, 414–422.
57. Trobridge,G., Hirata,R.K. and Russell,D.W. (2005) Gene targeting by adeno-associated virus vectors is cell-cycle dependent. *Hum. Gene Ther.*, **16**, 522–526.
58. Rothkamm,K., Kruger,I., Thompson,L.H. and Lobrich,M. (2003) Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.*, **23**, 5706–5715.
59. Lee,S.E., Mitchell,R.A., Cheng,A. and Hendrickson,E.A. (1997) Evidence for DNA-PK-dependent and -independent DNA double-strand break repair pathways in mammalian cells as a function of the cell cycle. *Mol. Cell. Biol.*, **17**, 1425–1433.
60. Miller,D.G., Petek,L.M. and Russell,D.W. (2004) Adeno-associated virus vectors integrate at chromosome breakage sites. *Nature Genet.*, **36**, 767–773.
61. Nakai,H., Wu,X., Fuess,S., Storm,T.A., Munroe,D., Montini,E., Burgess,S.M., Grompe,M. and Kay,M.A. (2005) Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. *J. Virol.*, **79**, 3606–3614.
62. Hollick,J.J., Golding,B.T., Hardcastle,I.R., Martin,N., Richardson,C., Rigoreau,L.J., Smith,G.C. and Griffin,R.J. (2003) 2,6-disubstituted pyran-4-one and thiopyran-4-one inhibitors of DNA-Dependent protein kinase (DNA-PK). *Bioorg. Med. Chem. Lett.*, **13**, 3083–3086.
63. Delacote,F., Han,M., Stamato,T.D., Jasin,M. and Lopez,B.S. (2002) An xrcc4 defect or Wortmannin stimulates homologous recombination specifically induced by double-strand breaks in mammalian cells. *Nucleic Acids Res.*, **30**, 3454–3463.
64. Yanez,R.J. and Porter,A.C. (1999) Gene targeting is enhanced in human cells overexpressing hRAD51. *Gene Ther.*, **6**, 1282–1290.