In vitro transposon targeting using a catalytically inactive Cas9

Shivam Bhatt and Ronald Chalmers*

School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham, NG7 2UH, United Kingdom

* To whom correspondence should be addressed. Tel: +44 115 823 0087; Fax: +44 115 823 0087; Email: ronald.chalmers@nottingham.ac.uk

ABSTRACT

Transposases are attractive tools for the integration of therapeutic transgenes into the chromosome for gene therapy applications. Typically, transgenes can be flanked with inverted-terminal repeat sequences, which are recognised by the transposase and integrated at random sites. Minimising detrimental insertions of transgenes is a key goal in the development of gene delivery vectors for gene therapy. We fused the Hsmar1 transposase to a catalytically inactive Cas9. Our aim was to bias transposon insertions into the vicinity of the target site bound by a guide RNA-dCas9 ribonucleoprotein complex. Although we could not detect any targeted transposition events *in vivo,* we achieved a 15-fold enrichment of transposon insertions into a 600-bp target site in an *in vitro* plasmid-to-plasmid assay. Additionally, we show that among those integrations that were successfully targeted, the location is tightly constrained to a site immediately to one side of the guide RNA target site. We present an in vitro proof-of-concept study demonstrating that the transposase insertion profile can be biased using a catalytically inactive Cas9 variant as a programmable DNA-binding module. One factor that limits the utility of this approach is that the transposon continues to integrate randomly. Although the dCas9 domain can be targeted to chromosomal lacZ, as evidenced by transcriptional repression, we were unable to detect any targeted insertions in the vicinity of the target site. Any targeted insertions that did occur were masked be a much larger number of random insertions. It is therefore necessary to develop a method for the temporal control of the transposase to allow Cas9 time to locate its target site.

INTRODUCTION

Gene therapy aims to reestablish correct expression of an absent, mutated or misregulated gene through introduction of a foreign, therapeutic allele into somatic cells. Important steps for effective gene therapy are: i) passage of the therapeutic gene (transgene) across the cell membrane, ii) endosomal escape into the cytoplasm and iii) entry into the nucleoplasm (1,2). This makes viruses an attractive tool as they have adapted to invade host cells, escape the endosome and release their genetic cargo into the nucleus for expression. Retroviruses are particularly useful because of their ability to integrate DNA into the host chromosome, conferring long-term transgene expression (3). Indeed, viral vectors dominated the first-generation gene therapy systems. However, as the field matured, numerous drawbacks of viral vectors became apparent. Most have limited cargo capacity, a high potential for immunogenicity and a propensity to integrate into actively transcribed genes (4,5). This prompted research into non-viral alternatives. One development was the cationic liposome, whereby transgenes could be enveloped by a lipid bilayer that fuses with the plasma membrane and is endocytosed into the cell (6,7). In an effort to improve transgene delivery, different liposome compositions were tested. It was found that construction of the liposome with a combination of cationic amphiphiles and neutral lipids helped to destabilise the endosomal membrane aiding escape into the cytoplasm (8). The transgene remains in the cytoplasm until the nuclear envelope dissolves during S-phase. Only after cell division, when the nuclear envelope has reformed, is the transgene then enclosed with the chromosomes for expression. Thus, quiescent cells are difficult to engineer.

Without integration into the host chromosome, the transgene exists as an extrachromosomal copy with no guarantee of long-term persistence. This makes transposase enzymes a valuable tool as any DNA sequence can be chromosomally integrated provided it is flanked with the appropriate inverted terminal repeats (ITRs). Thus, transgenes can be flanked with ITRs and packaged within a liposome with a transposase expression plasmid. Once inside the nucleus, transposase is expressed and integrates the transgene into the chromosome in a semi-random manner.

Currently, no efficient non-homology-directed methods exist for precise genome engineering. In 2002, a gene therapy trial was conducted on 5 patients with X-linked severe combined immunodeficiency. A retroviral vector was used to integrate the gamma(c) transgene into CD34+ bone marrow cells ex-vivo. Immunodeficiency was corrected and no adverse effects were noted within four months. However, 30 months after treatment, it was discovered that for one patient, insertion of the transgene occurred within the LMO-2 proto-oncogene, leading to aberrant expression of the LMO-2 transcript, manifesting in acute lymphoblastic leukaemia (9,10). This study highlighted the need to integrate transgenes into genomic safe-harbours or ideally, precise replacement at the faulty gene locus.

In a transposase-based system, one way to target the integration reaction is by fusing it with a DNA binding domain (DBD) (11). The first proof-of-concept of a DBD-transposase was the fusion of the phage λ repressor – cI, to the IS30 transposase. The aim was to bias transposon insertions to a target plasmid harbouring the phage λ operator in an in vivo plasmid-to-plasmid assay. It was found that ~57% of all integrations occurred within 400-bp of the λ operator with the cI-transposase fusion and no integrations into the target plasmid with wildtype transposase (12). This study paved the way for second-generation targeting systems that use zinc-finger proteins (ZFPs) or transcription activator-like effector DNA-binding domains (TALEs), which can be programmed to target a user-defined sequence (Chandrasegaran et al 2016 and references therein). PiggyBac and Sleeping Beauty transposases were selected as attractive fusion partners due to their high in vivo activity and have been tested with a variety of DBDs. The success of each fusion varied and largely depended on the DBD used, the choice of target (plasmid vs. chromosome) and whether the target site was engineered or not.

Although ZFPs or TALEs are programmable, their development requires substantial protein engineering before a suitable candidate could be chosen (13). ZFPs are composed of independent zinc-finger arrays, which each recognise 3-bp. Once the target site is selected and the array is constructed, several iterative cycles of selection usually have to be performed to achieve high specificity (14). TALEs are 33-35 amino acid domains that use either residue 12 or 13 to recognise a single

DNA base. Consequently, a 33-35 amino acid DBD must be added for every base-pair within a target site. Thus, a DNA-binding protein with an easily interchangeable DNA-binding module would be highly desirable.

The Cas9 nuclease from Type II CRISPR-Cas systems is a promising candidate as a programmable DBD. Cas9 induces double stranded breaks at chromosomal loci where its interchangeable RNA guide is bound (15). For Type-II CRISPR-Cas systems, the RNA guide has two components; a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (trcrRNA). The crRNA contains a spacer sequence, which is the region of complementarity to target DNA. trcrRNA hybridises with non-spacer regions of crRNA and is processed by host factors to generate mature gRNA, which complexes with Cas9 (16). Artificial chimaeras of processed crRNA:trcrRNA hybrids can also be used, these are called single-guide RNAs (sgRNA). Stable binding of the *S. pyogenes* Cas9-gRNA ribonucleoprotein complex (RNC) only occurs when the target site is immediately followed by a 5'-NGG protospacer adjacent motif (PAM). Initially, the RNC randomly collides with DNA and quickly dissociates from DNA that does not contain a PAM. Once a PAM is detected, a one-dimensional scan occurs to find complimentary DNA (17). It was demonstrated that nuclease activity of Cas9 could be abolished by introducing point mutations whilst retaining targeting activity, essentially turning it into a programmable DBD (18). dCas9, a catalytically inactive Cas9 variant, was made by introducing D10A and H840A mutations that disable the nuclease activity of the RuvC-like and HNH domains, which would cleave the non-target and target DNA strands respectively (18). A dCas9-PiggyBac transposase fusion was made in an attempt to target transposon insertions, however no targeted integration events were detected (19).

In the present work, we demonstrate an in vitro proof-of-concept using dCas9 fused to the Hsmar1 transposase. The target site for integration was the early open reading frame of lacZ, which allowed for visualization of targeting through blue/white screening. We focus on the design of the dCas9-transposase fusions and the activities of each domain working independently and together. Although the dCas9 domain can be targeted to chromosomal lacZ, as evidenced by transcriptional

repression, we were unable to detect any targeted insertions in the vicinity of the target site. Any targeted insertions that did occur were masked be a much larger number of random insertions.

MATERIAL AND METHODS

All reagents were generally of the highest quality available and were from the following suppliers unless otherwise stated. Dry chemicals were from Sigma Aldrich, Missouri, USA. All enzymes were from New England Biolabs (NEB), Massachusetts, USA and plasmid purification kits were from Qiagen, Hilden, Germany. Hsmar1 transposase was purified and assayed as described in (20,21). Plasmid pdCas9 was Addgene plasmid # 46569 (18). Guide RNA (gRNA) candidates were initially 30-mer annealed oligos with BsaI overhangs cloned into the CRISPR locus of pdCas9. AvrII and AatII restriction sites were introduced at the N- or C-termini of the dCas9 coding sequence to clone in Hsmar1 transposase-linker or linker-Hsmar1 transposase coding sequences to dCas9 respectively. These were pSB1 (N-terminal fusion), pSB2 (C-terminal fusion), pSB3 (N-terminal fusion with gRNA) and pSB4 (C-terminal fusion with gRNA). The dCas9 coding sequence was deleted to introduce NdeI and SalI restriction sites to clone in the native Hsmar1 transposase coding sequence (pSB15). The dCas9 promoter was deleted to introduce NotI and XhoI restriction sites to generate the promoterless expression construct (pSB21) and subsequently the P2 and P1 promoters were cloned in to generate pSB31 and pSB32 respectively. The entire P1-dCas9-transposase expression cassette was amplified from pSB31 with SacII overhangs and cloned into a SacII-digested single-copy expression plasmid to generate pSB33. The ampicillin resistance gene and origin of replication from pBluescript SK+ was PCR amplified with primers that introduced PsiI and DraIII overhangs. The PCR product was digested with PsiI/DraIII and ligated into a similarly digested pBluescript SK+ to generate the small target plasmid pSB34. To generate the large target, the argE-argeH locus was amplified from the bacterial chromosome with PciI overhangs and ligated into a PciI-digested small target plasmid to generate pSB35. In vivo targeting activity was visualized by a blue/white screen or by a modified Miller assay in liquid culture (22).

Purification of dCas9-Hsmar1 transposase:

The dCas9-Hsmar1 transposase expression plasmid, pSB2, was transformed into E. coli NiCo21 cells. A single colony was picked to inoculate a 10 mL LB-Lennox starter culture supplemented with 30 µg/mL chloramphenicol and grown at 37 °C for 16 hours. After incubation, the starter culture was diluted 1:100 into fresh liquid broth with 30 µg/mL chloramphenicol and grown to OD600= 0.5 followed by incubation at 18 °C for 16 hours, 250 rpm. Bacterial cells were pelleted and resuspended in binding buffer (5 mM imidazole, 0.5 M sodium chloride, 2 mM DTT and 20 mM Tris at pH 8.0) followed by lysis using a French pressure cell press. dCas9-Hsmar1 transposase was purified from crude cell lysate by passing the lysate through a 1 mL HisTrap HP column (GE Life Sciences) on an AKTA FPLC (Amersham Pharmacia). The column was equilibrated with binding buffer and bound protein was washed with 40 column volumes of wash buffer (60 mM imidazole, 0.5 M sodium chloride, 2 mM DTT and 20 mM Tris at pH 8.0) before being eluted with elution buffer (1 M imidazole, 0.5 M sodium chloride, 2 mM DTT and 20 mM Tris at pH 8.0) over a 20 column-volume gradient. Fractions containing purified protein were pooled and analysed via SDS-PAGE.

Targeted integration assays:

In vitro targeting assays were modified hop assays based on the one described in (23). dCas9-Hsmar1 transposase was reconstituted with sgRNA as described by (Ref = Anders et al 2015) except the room temperature incubation time was increased from 10 minutes to 20 minutes. 4.5 nM of wild type transposase, dCas9-transposase or dCas9-transposase-sgRNA was mixed with 9 nM of target plasmid pSB34 or pSB35, 9 nM of transposon donor plasmid pRC704, 13.5 nM of decoy plasmid pRC1104, 1X Hsmar1 reaction buffer (5% glycerol, 100 mM sodium chloride, 2 mM DTT, 25 mM Tris-HCl at pH 8.0), 2.5 mM MgCl2 and brought to 15 µL with dH2O. The reaction was incubated at 37 °C for 20 minutes in the absence of transposon donor followed by addition of donor and incubation for 24 hours. After incubation, decoy plasmid was degraded by incubation with NheI restriction endonuclease and lambda exonuclease simultaneously in a 50 µL reaction for 16 hours. Following digestion, the reaction was deproteinated with 15 µL of proteinase

K stop solution (0.5 mg/mL proteinase K, 0.1% SDS, 50 mM EDTA in 1X Cutsmart buffer) at 60 °C for 2 hours and cleaned using the Qiagen PCR clean-up kit. Purified target plasmids were eluted from the clean-up column in 30 µL and one-tenth of the reaction was transformed into E. coli NEB5a for the small target (3 µL) and one-third transformed for the large target (10 µL). Transformants were plated on 1.5% LB-Lennox agar supplemented with 50 µg/mL kanamycin, 100 µg/mL ampicillin, 40 µg/mL X-gal and 0.1% lactose. The number of white colonies was divided by the total number of colonies and multiplied by 100 to calculate targeting efficiency. Target plasmid DNA was purified from white clones and digested with SalI to calculate the number of transposon insertions. Additionally, purified target plasmids were sequenced using the M13R primer to map the transposon integration site.
In vivo targeting assays were performed by transforming *E. coli* BL21 cells with the C-terminal fusion expression plasmid, pSB4, and made chemically competent. A non-replicating transposon donor plasmid, pRC704, was subsequently transformed into this strain and plated on 1.5% LB-Lennox agar supplemented with 30 µg/mL chloramphenicol, 50 µg/mL kanamycin, 40 µg/mL X-gal and 0.1% lactose. White, kanamycin resistant clones were selected for colony PCR scanning a ±1.1 kb window around the guide RNA target site.

RESULTS

Design of dCas9-Hsmar1 transposase fusions

In order to be targeted, dCas9 must form a ribonucleoprotein complex (RNC) with guide RNA (gRNA) that is complementary to the target site. However, not all gRNAs are equally efficient at targeting (24). Thus, in order to find an efficient guide, we screened ten gRNA candidates that target lacZ. Efficient targeting was visualised on a blue/white screen. Candidate gRNAs were constructed by annealing 30-mer oligo pairs and cloned into the CRISPR locus of plasmid pdCas9 (Figure 1A). Subsequent transcription and processing result in the formation of mature gRNA. *E. coli* BL21 cells were transformed separately with each dCas9-gRNA expression plasmid and spread on x-gal indictor plates. Efficient targeting produces white colonies due to target-bound dCas9 acting as a transcriptional roadblock (25). Of ten gRNA

candidates, two produced white colonies (Figure 1B). One gRNA targeted the lac operator and the other targeted the early open read frame (ORF) of lacZ. We performed later experiments with the gRNA targeting the ORF as any transposon integrations upstream of the operator, outside of the promoter, would not be detected. Subsequently, the Hsmar1 transposase coding sequence was fused to either the N- or C-terminus of dCas9 with- and without the gRNA sequence. Transposase and dCas9 were covalently linked using a 187 amino acid linker that had previously been used to generate a single-chain Hsmar1 dimer (26).

Targeting efficiency of the dCas9 domain in vivo

At an actively transcribed gene, target-bound dCas9 represses transcription by blocking transcript elongation by RNA polymerase (25). For lacZ, this can be measured by Miller assay, which is a colorimetric β-galactosidase assay in liquid culture. Thus, in order to determine targeting efficiency, we transformed the N-terminal fusion (transposase-dCas9) and C-terminal fusion (dCas9-transposase) expression plasmids (± gRNA) into E. coli BL21. The N- and C-terminal fusions repress β-galactosidase activity 4- and 8-fold respectively compared to untargeted dead Cas9 alone (Figure 2B). Thus, the C-terminal fusion was 2-fold more efficient at targeting than the N-terminal fusion. The efficiency of both fusion proteins in the absence of gRNA was comparable but slightly higher than the dCas9 control, which had comparable β-galactosidase activity to the BL21 positive control, as expected.

Transpositional activity of the Hsmar1 transposase domain in vivo

After establishing that dCas9 targets the lacZ ORF efficiently, we next examined the transpositional activity of the transposase domain in vivo (Figure 2A). When guide RNA is present, we observed a decrease in transpositional activity of both the N- and C-terminal fusions. Both were 2-fold less active than their counterparts lacking gRNA. Compared to native Hsmar1 transposase, the activity of the C-terminal fusion with gRNA was comparable, whereas the activity of the N-terminal fusion with gRNA was 2-fold higher. In the absence of guide RNA, the activity of the N-terminal fusion was 9-fold higher than that of the C-terminal fusion. No transpositional activity with dCas9 alone was detected.

Attempt at chromosomal integration

The C-terminal fusion was more efficient at targeting than the N-terminal fusion and had similar transpositional activity to native Hsmar1 transposase. Thus, this fusion was selected for in vivo targeted transposition in E. coli BL21. Cells were transformed with the C-terminal fusion expression plasmid, pSB4, and made chemically competent. A non-replicating transposon donor plasmid, pRC704, was subsequently transformed into this strain. In this assay, colonies only grow if a transposition event has occurred, moving the kanR transposon from the donor plasmid to the chromosome. White colonies would be produced if the transposon integrated into lacZ. Most colonies that grew were blue, however several white colonies also formed. After extended incubation, all white colonies eventually turned blue. Initial white colonies were subject to colony PCR scanning a ±1.1 kb window around the guide RNA target site. An amplicon of 2.2 kb is produced in the absence of any transposon integration and we did not detect any amplicon at 4.5 kb that would be indicative of targeted integration (Supplementary Figure S1).

Impact of protein expression level on targeting efficiency and transpositional activity

Although no targeted chromosomal integrations could be detected, kanR colonies still grew. This suggested that off-target integrations were occurring. We hypothesized that a high transposase concentration was responsible for the off-target integrations. Thus, a series of weakened dCas9-transposase expression plasmids were constructed with variable promoters and plasmid copy number. These were P2, P1, promoterless and P1-single copy in decreasing expression level respectively. A positive correlation was found between expression level and targeting. As the expression level decreased, the targeting efficiency also decreased. Targeting efficiency decreased by 1.4-fold from the native promoter to P2, 4.7-fold from native to P1 and 7.7-fold from native to P1-single copy. For transpositional activity, a negative correlation was observed. As the expression level decreased, transpositional activity increased. Activity was comparable between the native promoter and P2, a 1.9-fold increase from native to P1 and a 9.1-fold increase from native to P1-single copy.

Target DNA binding assays of guide RNA-dCas9

In order to demonstrate targeting in vitro, we purified dCas9-Hsmar1 transposase and first examined target DNA binding via electrophoretic mobility shift assay (EMSA). The guide RNA from the in vivo analysis was commercially synthesised as sgRNA and was reconstituted with purified protein to form a RNC. RNC was mixed with oligoduplex DNA that was complementary to the sgRNA in a binding reaction. The products of the reaction were separated on a non-denaturing polyacrylamide gel and visualised on a Fujifilm FLA-3000 phosphorimager (Figure 3D-G). A DNA-RNC complex is first observed as a doublet band at 50 nM RNC, where approximately 50% of the free DNA substrate has been consumed. As the concentration of RNC rises further to 100- and 250 nM, all free DNA becomes bound (Figure 3E). In the absence of sgRNA, some free DNA becomes bound but no stable complex is detected (Figure 3D). Unlabelled, non-specific DNA was introduced into the binding reaction to observe any change in specific DNA binding activity (Figure 3F). No significant change in binding characteristics was detected. Lastly, we examined the impact of transposon DNA in the binding reaction (Figure 3G). We observed a significant decrease in substrate binding where the DNA-RNC complex could only be detected at the highest enzyme concentration of 250 nM.

Transposon-end binding assays of Hsmar1 transposase

Binding of transposase to transposon-end was also examined by EMSA. In this binding reaction, oligoduplex DNA was composed of a single inverted-terminal repeat, target-site duplication and flanking DNA. dCas9-Hsmar1 transposase was compared to an MBP-Hsmar1 transposase control. With MBP-Hsmar1 transposase (Figure 3A), we detected substrate binding at the lowest transposase concentration of 10 nM. Initially, this is a doublet band with a prominent lower band. The lower band is SEC1, where a single monomer of transposase is bound to a single transposon end. The higher band is SEC2, involving a dimeric transposase bound to the transposon end. As the concentration of transposase rises, we observed a gradual transition from SEC1 to SEC2. With dCas9-Hsmar1 transposase (Figure 3C), we detected binding of the substrate at 10 nM to produce a triplet band. As the concentration of dCas9-Hsmar1 rises, the intensities of all three bands increased. In

both gels, the amount of substrate that gets bound at identical enzyme concentrations is consistent, with more than 50% substrate bound at 40 nM enzyme and almost all bound at 80 nM and higher. We assigned the middle and lower bands of the dCas9-Hsmar1 transposase gel to be SEC2 and SEC1 respectively. We do not, however, see the gradual transition of band intensity from SEC1 to SEC2 as we see with MBP-Hsmar1 transposase. Above SEC2 there is a third band, which could represent either a structural isoform of SEC2 or potentially the paired-ends complex (PEC).

In vitro transposition assay of the Hsmar1 transposase domain

To ensure transposase catalytic activity is retained in vitro, the dCas9-Hsmar1 transposase fusion was subject to the in vitro transposition assay (Figure 4). The substrate for the transposase domain is 6.5 nM of supercoiled transposon donor plasmid – pRC650, and the reaction products were electrophoretically separated on an agarose gel. In this assay, consumption of the supercoiled substrate and production of plasmid backbone are the two main determinants of transpositional activity. MBP-Hsmar1 transposase was used as a control (Figure 4, left). As the concentration of MBP-Hsmar1 transposase rises, the amount of supercoiled substrate depletes until an optimal concentration of enzyme is found, that being 40 nM. Beyond 40 nM, we observed over-production inhibition, which is typical for mariner transposases. At the optimal 40 nM, almost all of the supercoiled substrate has been converted into backbone, that is, the leftover linear plasmid produced once the transposon has been excised. We also saw the 4.7 kb linearised substrate plasmid, which occurs upon production of a single-end break. The two closest bands above 10 kb are the nicked-circular substrate plasmid and a supercoiled substrate plasmid as indicated on the gel. At 6.7 kB, We observed the 2X linear product, which arises from an excised transposon integration into an unreacted substrate plasmid. Below the plasmid backbone, between 1 and 2 kb are the autointegration products wherein the excised transposon has integrated into itself. The number of supercoiled nodes trapped between each transposon end gives rise to the variation in autointegration products. When the dCas9-Hsmar1 transposase fusion ± sgRNA is examined (Figure 4, second and third from left resp.), we observed a decrease in

catalytic activity at any enzyme concentration used when compared to MBP-Hsmar1 transposase. We saw lower consumption of supercoiled substrate and backbone production. When the substrate for the RNC is introduced (Figure 4, right), we observed an increase in catalytic activity up to 20 nM enzyme when compared to dCas9-Hsmar1 and the RNC. However, above 20 nM we see a reduction in the amount of backbone produced, not as a result of OPI, but due to an increased DNA nicking activity. Taken as a whole, we demonstrate that the transposase domain of the fusion protein retains catalytic activity albeit with reduced activity.

Targeted transposition in vitro

In order to assay for targeted integration, where the two domains of the fusion protein would work in tandem, we used a modified hop assay (23). In this assay, transposase catalyses the movement of a kanR transposon from a donor plasmid to a target plasmid harbouring the sgRNA target site (Figure 5A). This site is the coding sequence for the β-galactosidase alpha-peptide in a modified pBluescript plasmid. The reaction was digested to remove decoy plasmid, deproteinated, cleaned and subsequently transformed into E. coli NEB5α for alpha complementation. The number of white, ampR/kanR colonies was divided by the total number of colonies and multiplied by 100 to determine targeting efficiency (Figure 5C). We detected 9.3-fold and 8.2-fold higher targeting efficiencies using the RNC compared to wild type transposase and dCas9-Hsmar1 transposase respectively. We modified the assay by doubling the size of the target plasmid. Using the large target plasmid, we detected a 14.8-fold higher targeting efficiency using the RNC compared to wild type transposase. A targeting efficiency of 0 % was calculated for dCas9-Hsmar1 transposase as no white colonies were formed. Thus, the fold-increase in targeting efficiency of the RNC compared to dCas9-Hsmar1 transposase was incalculable. Collectively, 60 white colonies were sampled across each enzyme set and target plasmids were purified. Plasmids were digested with a single cutter restriction enzyme and the size difference between the parental target and integrated target was used to calculate the number of integrations each plasmid had received. Out of 60 target plasmids, 49 had received a single transposon insertion and 11 had received two (81.7% vs. 18.3% respectively).

Transposon insertion sites from the small target plasmid were mapped by DNA sequencing. Ten insertion sites were mapped for all enzymes. For the RNC, all 10 integrations were within 22 bp to one side of the sgRNA binding site. Seven transposon integrations occurred at position 190, the first TA dinucleotide 5' of the sgRNA binding site. The other three occurred at position 187, the second closest TA dinucleotide. For dCas9-Hsmar1 transposase, one integration occurred at position 29, three occurred at position 187, two at position 190, three at position 215 and the last one at position 255. For wild type transposase, I recovered two transposon integrations at position 29, one at position 39, three at position 83, one at position 89, two at position 187 and the last at position 190.

DISCUSSION

Almost all transposons are integrated in a fairly random manner by their cognate transposases. Thus, a necessary step for the evolution of transposon technology in gene therapy revolves around reducing, if not eliminating, random integration. The consequences of random integration can be poor or unstable transgene expression, insertional mutagenesis and oncogenesis. Attempts at targeting transposon integrations using DBD-transposase fusions have yielded encouraging results in *in vivo* plasmid-to-plasmid assays (27-30). Recently, Luo et al. attempted to demonstrate chromosomal targeting of a dCas9-PiggyBac transposase fusion but could not detect any targeted integrations. Here, we report our attempt at targeted chromosomal integrations as well an in vitro plasmid to plasmid based assay using dCas9 fused to the mariner Hsmar1 transposase.

We first designed 10 gRNA candidates and found that two were efficient at binding the target site (Figure 1). Both gRNAs targeted the non-template strand, which is consistent with several groups demonstrating that targeting the non-template strand when using dCas9 yields improved transcriptional repression (31,32).

dCas9-transposase is composed of two DNA-binding modules, one being formed when dCas9 complexes with gRNA and the other being the two helix-turn-helix DNA-binding domains located on the N-terminus of the transposase moiety. Thus, we

suspected that a conflicted search might occur where the RNC is searching for target DNA whilst the transposase is searching for the transposon-end. The consistency of the 2-fold reduction implies that the same activity-reducing mechanism is occurring at both ends, with the only variable being the presence of gRNA.

During examination of transpositional activity, we noted that the activity of the N-terminal fusion (transposase-dCas9) was ~4-fold higher than native Hsmar1 transposase (Figure 2A). This could be due to physical constraints placed upon the C-terminal catalytic domain of Hsmar1. With Hsmar1 transposase (and other mariner transposases), a phenomenon called over-production inhibition occurs when there are too many dimers of transposase present (33). These dimers saturate both ends of the transposon and do not permit efficient synapsis. The underlying mechanism behind OPI is called assembly-site occlusion (ASO). ASO can also be augmented by the allostery between transposase domains whereby upon SEC2 formation, information is passed from the bound monomer to the other in the form of a conformational change, which lowers the affinity of the unoccupied monomer for the unbound ITR (34). The conduit for this exchange of information is through the conserved WVPHEL-YSPDL motifs, which forms the clamp-loop structure connecting the two transposase monomers that extends off of the C-terminal catalytic domain (34). It is possible that the presence of the large dCas9 on the Hsmar1 C-terminus relieves OPI by sterically hindering any conformational change required to relay information exchange across the clamp-loop dimer interface.

Whereas several attempts by other groups have utilised an artificially engineered target site (28-30,35), we decided to leave the endogenous site intact to mimic the environment a targeted transgene integration system would most likely be in. Although we could target the endogenous lacZ site, no targeted transposition events *in vivo* could be detected (Supplementary Figure S1). We observed only the wildtype amplicon when colony PCR was performed on white, kanR colonies. This raised several questions regarding protein expression level, target copy number and OPI. Constitutive expression of dCas9-Hsmar1 transposase would lead to large numbers of enzyme present in the cytoplasm dependant on the strength of the *SpCas9*

promoter and plasmid copy number. Theoretically, only one of these molecules can occupy the single-copy lacZ target site. This molecule would have to either a) bind the lacZ target site then carry out transposition whilst locked into the target-bound configuration, placing physical constraints on the transposase or b) bind the transposon end, perform excision and integrate the transposon before the target site is occupied by another molecule. Coupled to this is the impact of overproduction inhibition, causing competition between the one bound molecule and many more unbound molecules for the second transposon end. It therefore seems likely that any transposition events that do occur would be off-target if the chromosomal target site were occupied.

The transpositional activity of Hsmar1 follows a bell-shaped curve that supports OPI. Over-saturation of transposon ends occurs when transposase is expressed from a strong promoter. Weakening the promoter and reducing plasmid copy number reduce the concentration of transposase within the cell until the optimal concentration is reached providing peak activity. This is followed by production of sub-optimal levels of transposase leading to a decline in transpositional activity. For targeted transposition *in vivo*, it seemed appropriate to use the optimal, if not sub-optimal, transposase expression level to minimise the impact of OPI and limit off-target integrations. However, when the experiment was repeated to measure targeting efficiency, as the protein expression level decreased, the targeting efficiency also decreased (Figure 2C vs. 2D). This suggests that a high concentration of dCas9 is required to maintain transcriptional repression. It became clear that *in vivo*, a balancing issue is raised. A high concentration of dCas9-Hsmar1 is required for effective targeting but this leads to OPI and off-target integrations (and vice versa).

To detect targeted transposition events *in vitro,* we performed an unbiased, in vitro plasmid-to-plasmid targeted integration assay. It is unbiased because the target plasmid contains two copies of the plasmid origin of replication (ori) and antibiotic resistance. This ensures that transposition events can be detected throughout the target plasmid except in rare circumstances where two transposition events have

occurred into both ampR and ori. Additionally, we repeated the assay using a target plasmid that was twice as large, which dilutes the target site. The reaction was supplemented with a 3-fold excess of a third, decoy plasmid with respect to the transposase. This provided the transposase with a choice for integration. Two-fold less transposase was used with respect to transposon donor and target plasmid to ensure that transposase remains at sub-saturating levels to both the target-site and transposon ends. Additionally, the transposase was pre-incubated with target plasmid thereby allowing all molecules of the RNC to have docked onto the target site. When compared to wild type transposase, the RNC is 9.3-fold more effective at integrating the transposon into the 600-bp target with the small target. When using the larger target, the targeting efficiency is preserved using the RNC (~55 %) because it is directed to the target site. Comparatively, the targeting efficiency of wild type and dCas9-transposase decreased. This is because wild type and dCas9-transposase are integrating into the target at random. By diluting the target site by adding more DNA (and therefore more TA dinucleotides for integration), the chance of random integration into the target decreases proportionally. There are 33 TA dinucleotides within the target site out of 252 within the entire small target. Therefore, random integration into the target should occur at a frequency of 13 %. However, this frequency is reduced to 6 % for wild type transposase due to the presence of decoy plasmid in the reaction. Alternatively, the frequency may be lower if the target site is a ‘cold-spot’ for integration. With the large target, the frequency of random integration into the target should be 7.7 %, and in the presence of decoy was calculated to be 3.8 %. Thus, if the decoy is responsible for reducing the frequency of random integration, it does so by approximately 50% when in a 3-fold excess to the transposase. In terms of window specificity, 100 % of integrations were detected within 22-bp of the sgRNA binding site using the RNC.

For comparison, the Sleeping Beauty transposase showed a 10% targeting efficiency when redirected to its artificially engineered chromosomal target site (35). Gal4-Mos1 transposase and Gal4-PB transposase fusions could be biased towards a UAS-containing target plasmid with 12.7- and 11.6-fold efficiency respectively compared to unfused transposase (27). A ZFP-PB transposase fusion could bias PB insertions

within a 500-bp window of the ZFP-binding site at a frequency of 74% in an in vivo plasmid-to-plasmid assay. However, it was also detected that native PB integrated into the same window with a frequency of 50% (28). In another in vivo plasmid-to-plasmid assay, a Gal4-PB transposase fusion was able to bias integrations around a UAS-containing target plasmid with ~4-fold enrichment. 87% of integrations occurred within 800-bp of the UAS compared to 57% with native PB transposase (29). After engineering the HEK-293 cell genome to contain 1-2 Gal4-UAS target sequences, Gal4-PB transposase integrated transposon DNA within a 1.8-kb window of the target 32% of the time compared to 8% for native PB transposase (29). It was also demonstrated that a TALE-PB fusion could direct PB insertions into the first intron of the human CCR5 gene, which were detected at a frequency of ~0.010-0.014% of total stably transfected cells (36). The zinc-finger protein, zif268, was fused to the C-terminus of the ISY100 transposase and it was shown that 50% of all integrations occurred within 20-bp of the target site. However, the target site on the target plasmid was engineered to contain 9 tandemly repeated TANN transposase integration sites adjacent to the ZFP binding site (30). In a more recent study, a TALE-PB transposase and ZFP-PB transposase fusion could target and integrate transposon DNA into the human HPRT gene at a frequency of 0.97% and 0.42% respectively (37). Taken as a whole, our results fit well with the evidence from the literature given that our target site remains unmodified. The two caveats however, were that the plasmid-to-plasmid assay was performed in vitro, and the relatively small sample size taken when mapping transposon integrations. We present here a proof-of-concept that dCas9 can indeed be fused to a transposase to bias its integration pattern. Future work would include being able to establish chromosomal integrations as well as improving the targeting efficiency to avoid off-target integrations.

ACKNOWLEDGEMENT

FUNDING

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) research training support grant number RS8654 to SB. The funders had no role in study design, data collection and analysis, decision to publish

or preparation of the manuscript.

CONFLICT OF INTEREST

None declared

REFERENCES

1. Katare, D.P. and Aeri, V. (2010) Progress in gene therapy: A Review. *IJTPR*, **1**, e41.
2. Nayerossadat, N., Maedeh, T. and Ali, P.A. (2012) Viral and nonviral delivery systems for gene delivery. *Adv Biomed Res*, **1**, 27.
3. Bouard, D., Alazard-Dany, D. and Cosset, F.L. (2009) Viral vectors: from virology to transgene expression. *Br J Pharmacol*, **157**, 153-165.
4. Thomas, C.E., Ehrhardt, A. and Kay, M.A. (2003) Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet*, **4**, 346-358.
5. Vannucci, L., Lai, M., Chiuppesi, F., Ceccherini-Nelli, L. and Pistello, M. (2013) Viral vectors: a look back and ahead on gene transfer technology. *New Microbiol*, **36**, 1-22.
6. Zou, Y., Zong, G., Ling, Y.-H. and Perez-Soler, R. (2000) Development of cationic liposome formulations for intratracheal gene therapy of early lung cancer. *Cancer gene therapy*, **7**, 683.
7. Semple, S.C., Akinc, A., Chen, J., Sandhu, A.P., Mui, B.L., Cho, C.K., Sah, D.W., Stebbing, D., Crosley, E.J. and Yaworski, E. (2010) Rational design of cationic lipids for siRNA delivery. *Nature biotechnology*, **28**, 172.
8. Yusa, K., Zhou, L., Li, M.A., Bradley, A. and Craig, N.L. (2011) A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci U S A*, **108**, 1531-1536.
9. Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.P., Thrasher, A.J., Wulffraat, N., Sorensen, R., Dupuis-Girod, S. *et al.* (2002) Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med*, **346**, 1185-1193.
10. Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M. *et al.* (2003) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med*, **348**, 255-256.
11. Kaminski, J.M., Huber, M.R., Summers, J.B. and Ward, M.B. (2002) Design of a nonviral vector for site-selective, efficient integration into the human genome. *Faseb j*, **16**, 1242-1247.
12. Szabo, M., Muller, F., Kiss, J., Balduf, C., Strahle, U. and Olasz, F. (2003) Transposition and targeting of the prokaryotic mobile element IS30 in zebrafish. *FEBS Lett*, **550**, 46-50.
13. Beumer, K.J., Trautman, J.K., Mukherjee, K. and Carroll, D. (2013) Donor DNA Utilization During Gene Targeting with Zinc-Finger Nucleases. *G3 (Bethesda)*, **3**, 657-664.
14. Li, X., Burnight, E.R., Cooney, A.L., Malani, N., Brady, T., Sander, J.D., Staber, J., Wheelan, S.J., Joung, J.K., McCray, P.B. *et al.* (2013) *piggyBac* transposase tools for genome engineering. *Proceedings of the National Academy of Sciences*, **110**, E2279-E2287.
15. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, **8**, 2281-2308.
16. Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. and Charpentier, E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, **471**, 602-607.
17. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C. and Doudna, J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, **507**, 62-67.
18. Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F. and Marraffini, L.A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*, **41**, 7429-7437.

19. Luo, W., Galvan, D.L., Woodard, L.E., Dorset, D., Levy, S. and Wilson, M.H. (2017) Comparative analysis of chimeric ZFP-, TALE- and Cas9-piggyBac transposases for integration into a single locus in human cells. *Nucleic Acids Res*, **45**, 8411-8422.
20. Claeys Bouuaert, C., Liu, D. and Chalmers, R. (2011) A simple topological filter in a eukaryotic transposon as a mechanism to suppress genome instability. *Mol Cell Biol*, **31**, 317-327.
21. Claeys Bouuaert, C. and Chalmers, R.M. (2010) Gene therapy vectors: the prospects and potentials of the cut-and-paste transposons. *Genetica*, **138**, 473-484.
22. Miller, J. (1972) Experiments in Molecular Genetics. *Cold Spring Harbor Laboratory*.
23. Munoz-Lopez, M., Siddique, A., Bischerour, J., Lorite, P., Chalmers, R. and Palomeque, T. (2008) Transposition of Mboumar-9: identification of a new naturally active mariner-family transposon. *J Mol Biol*, **382**, 567-572.
24. Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J. and Takumi, T. (2016) Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Sci Rep*, **6**, 19675.
25. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S. and Qi, L.S. (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*, **8**, 2180-2196.
26. Claeys Bouuaert, C. and Chalmers, R. (2017) A single active site in the mariner transposase cleaves DNA strands of opposite polarity. *Nucleic Acids Res*, **45**, 11467-11478.
27. Maragathavally, K.J., Kaminski, J.M. and Coates, C.J. (2006) Chimeric Mos1 and piggyBac transposases result in site-directed integration. *FASEB J*, **20**, 1880-1882.
28. Kettlun, C., Galvan, D.L., George, A.L., Jr., Kaja, A. and Wilson, M.H. (2011) Manipulating piggyBac transposon chromosomal integration site selection in human cells. *Mol Ther*, **19**, 1636-1644.
29. Owens, J.B., Urschitz, J., Stoytchev, I., Dang, N.C., Stoytcheva, Z., Belcaid, M., Maragathavally, K.J., Coates, C.J., Segal, D.J. and Moisyadi, S. (2012) Chimeric piggyBac transposases for genomic targeting in human cells. *Nucleic Acids Res*, **40**, 6978-6991.
30. Feng, X., Bednarz, A.L. and Colloms, S.D. (2010) Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic Acids Res*, **38**, 1204-1216.
31. Lei, Y., Zhang, X., Su, J., Jeong, M., Gundry, M.C., Huang, Y.-H., Zhou, Y., Li, W. and Goodell, M.A. (2017) Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nature communications*, **8**, 16026.
32. Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Holper, S., Kruger, M. and Stainier, D.Y. (2015) Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, **524**, 230-233.
33. Claeys Bouuaert, C., Lipkow, K., Andrews, S.S., Liu, D. and Chalmers, R. (2013) The autoregulation of a eukaryotic DNA transposon. *Elife*, **2**, e00668.
34. Liu, D. and Chalmers, R. (2014) Hyperactive mariner transposons are created by mutations that disrupt allosterism and increase the rate of transposon end synapsis. *Nucleic Acids Res*, **42**, 2637-2645.
35. Ivics, Z., Katzer, A., Stuwe, E.E., Fiedler, D., Knespel, S. and Izsvak, Z. (2007) Targeted Sleeping Beauty transposition in human cells. *Mol Ther*, **15**, 1137-1144.
36. Owens, J.B., Mauro, D., Stoytchev, I., Bhakta, M.S., Kim, M.S., Segal, D.J. and Moisyadi, S. (2013) Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. *Nucleic Acids Res*, **41**, 9197-9207.
37. Luo, J., Torres-Kolbus, J., Liu, J. and Deiters, A. (2017) Genetic Encoding of Photocaged Tyrosines with Improved Light-Activation Properties for the Optical Control of Protease Function. **18**, 1442-1447.

FIGURES LEGENDS

Figure 1. Plasmid pdCas9 and chromosomal *lacZ* target

A) Plasmid pdCas9 was a gift from Addgene plasmid # 46569. This plasmid expresses a catalytically inactive Cas9 variant harbouring the D10A and H840A mutations, which abolish nuclease activity. This plasmid codes for two separate transcripts that form the trans-activating CRISPR RNA and CRISPR RNA, which hybridise and are processed to form mature guide RNA. The CRISPR RNA coding sequence is found in the minimal CRISPR locus of pdCas9, composed of a promoter, leader and a repeat-spacer-repeat configuration. The spacer is the user-defined sequence that confers targeting to any predetermined target site. B) For these experiments, the target site for dCas9 binding was the *E. coli lacZ* gene. Ten candidate spacers were designed to target the early open reading frame of *lacZ.* A blue/white screen was used to determined whether each spacer conferred efficient targeting to *lacZ,* which manifests as a white colony on x-gal indicator plates. Of these 10 candidates, only spacers 4 and 7 conferred efficient targeting.

Figure 2. Transpositional and targeting activity of dCas9-transposase fusions.

A) We found that the N-terminal fusion (transposase-dCas9) had 4-fold higher transpositional activity than wildtype whereas the C-terminal fusion (dCas9-transposase) had similar activity to wildtype. In both fusions, the presence of guide RNA reduces the activity by 2-fold. B) β-galactosidase activity was measured using a colorimetric assay as described by (Miller 1972). The targeting efficiency is given as a function of β-galactosidase activity where a higher targeting efficiency results in lower Miller units (the unit of β-galactosidase activity). C) lacZ transcriptional repression by dCas9 as a function of promoter strength from weakest to strongest was P1 to P5. D) Conversely, when transpositional activity was examined, we found that activity increased as the expression level decreased.

Figure 3. Substrate binding assays of dCas9-transposase.

A) Electrophoretic mobility shift assay of purified Hsmar1 transposase with an

inverted terminal repeat-containing DNA substrate. As the transposase concentration rises, the amount of free substrate DNA decreases proportionately. Binding is initially detected at 10 nM transposase to form a doublet band, which we ascribe SEC1 and SEC2. B) Schematic illustration of protein-DNA complexes formed during the transposition reaction and decomposition of the paired-end complex to form SEC1 during electrophoresis. C) Electrophoretic mobility shift assay of dCas9-transposase with an inverted terminal repeat-containing DNA substrate. D-G) Electrophoretic mobility shift assays of dCas9-transposase using the guide RNA target site as DNA substrate. We found that binding only occurs in the presence of guide RNA (D), and that binding occurs to form two protein-DNA complexes (E). Binding proceeds to occur even in the presence of non-specific DNA (F), and that binding is hindered by the presence of competitor DNA containing an inverted terminal repeat (G).

Figure 4. *In vitro* transposition assays of dCas9-transposase
Purified transposase was mixed with 6.5 nM of supercoiled plasmid substrate containing a transposon. As the transposition reaction progresses, supercoiled substrate is converted to linear backbone, the end product of the reaction. In this titration experiment, 40 nM of wildtype transposase is sufficient for conversion of almost all substrate to product after 4 hours. However, dCas9-transposase displays reduced transpositional activity, being able to convert a maximum of 50% of substrate to product at the highest concentration of 60 nM after 4 hours. A similar phenotype is displayed by the dCas9-sgRNA-transposase ribonucleoprotein complex. When the target of Cas9 is supplemented into the reaction, activity of the transposase domain increases with a concomitant increase in nicking activity.

Figure 5. Targeted transposition *in vitro*.
A) Schematic representation of the *in vitro* hop assay. The assay includes a transposon donor plasmid, target plasmid and a decoy plasmid (not shown). Purified transposase catalyses the movement of the kan^R transposon into the target or decoy plasmid. The reaction was composed of a 1:2:2:3 ratio of transposase:target:donor:decoy respectively. Targeting was assessed after deproteination and transformation into *E. coli* strain suitable for *lacZ* α-

complementation. B) Colonies from blue/white screen and kanamycin selection for targeted transposition. C) The number of white colonies formed was divided by the total number of colonies and multiplied by 100 to provide a measure for targeting efficiency. With the small target, we detected a 9.3- and 8.2-fold increase in targeting efficiency of the RNC compared to wild type transposase and dCas9-transposase respectively. Error-bars were calculated using standard error of the mean where n = 6. Ordinary one-way ANOVA multiple comparison: Small target: WT vs. dCT = not significant (P-value = 0.9595), WT vs. RNC = significant (P-value = <0.0001), dCT vs. RNC = significant (P-value = <0.0001). Large target: WT vs. dCT = not significant (P-value = 0.1352), WT vs. RNC = significant (P-value = <0.0001), dCT vs. RNC = significant (P-value = <0.0001). D) Target plasmids were harvested from 10 white, kan^R/amp^R colonies from all plates. These plasmids were sequenced to map transposon insertion sites. Of the ten RNC-catalysed insertions, all occurred within 22-bp to one side of the sgRNA target site. E) Expanded sequence view of the *lacZ* target with TA dinucleotides highlighted and integration sites marked with an arrow.

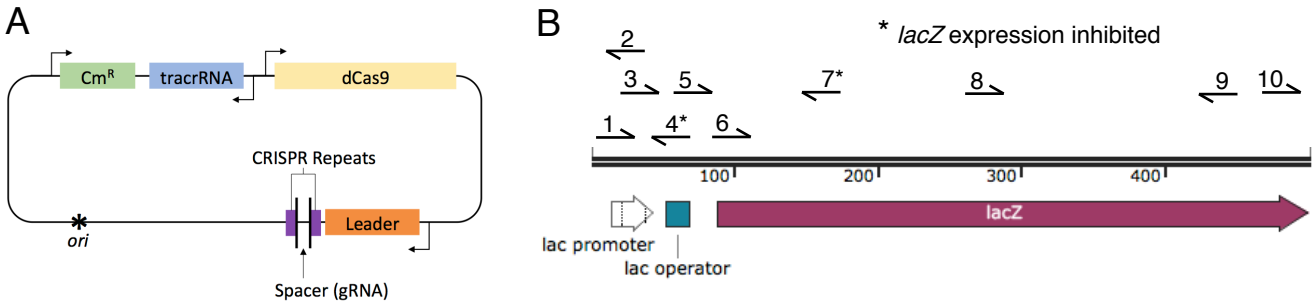
A
Cm^R
tracrRNA
dCas9
CRISPR Repeats
Leader
ori
Spacer (gRNA)
B
* lacZ expression inhibited
2
3
5
7*
8
9
10
1
4*
6
100
200
300
400
lacZ
lac promoter
lac operator

Figure 1

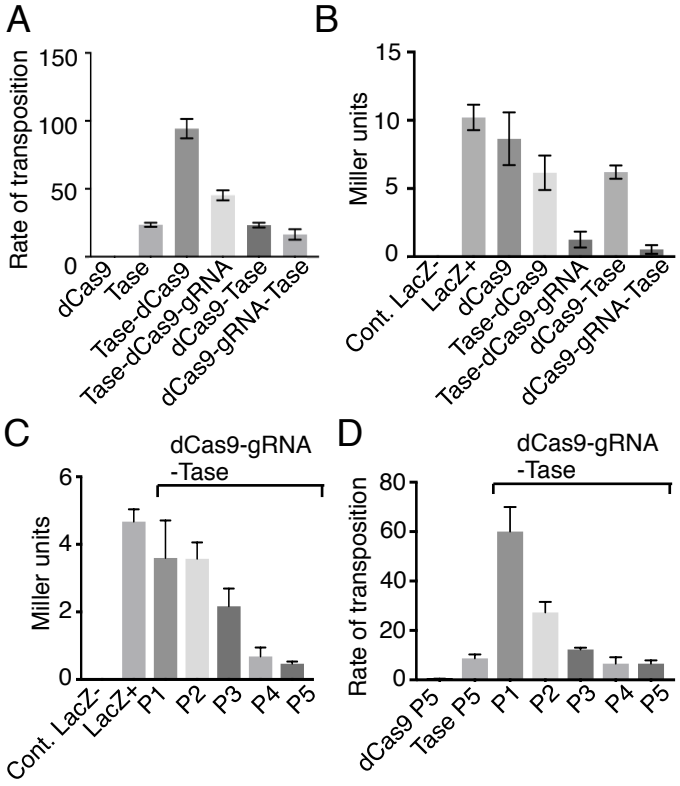
A
Rate of transposition
150
100
50
0
dCas9
Tase
Tase-dCas9
Tase-dCas9-gRNA
dCas9-Tase
dCas9-gRNA-Tase
B
Miller units
15
10
5
0
Cont. LacZ-
LacZ+
dCas9
Tase-dCas9
Tase-dCas9-gRNA
dCas9-Tase
dCas9-gRNA-Tase
C
dCas9-gRNA
-Tase
Miller units
6
4
2
0
Cont. LacZ-
LacZ+
P1
P2
P3
P4
P5
D
dCas9-gRNA
-Tase
Rate of transposition
80
60
40
20
0
dCas9 P5
Tase P5
P1
P2
P3
P4
P5

Figure 2

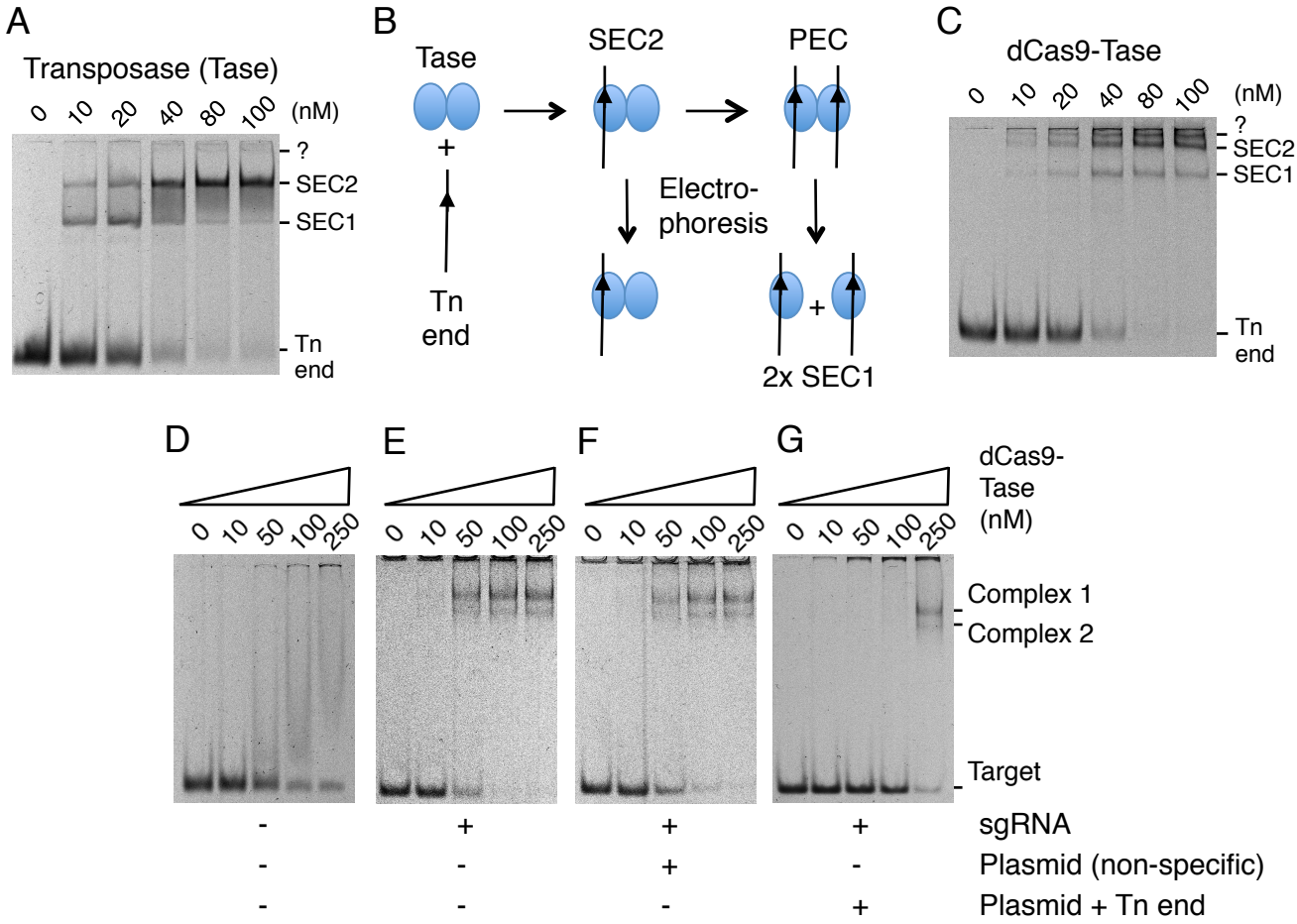
A
Transposase (Tase)
0 10 20 40 80 100 (nM)
?
SEC2
SEC1
Tn end
B
Tase
+
Tn end
SEC2
PEC
Electro-phoresis
2x SEC1
C
dCas9-Tase
0 10 20 40 80 100 (nM)
?
SEC2
SEC1
Tn end
D
E
F
G
0 10 50 100 250
dCas9-Tase (nM)
Complex 1
Complex 2
Target
sgRNA
Plasmid (non-specific)
Plasmid + Tn end

Figure 3

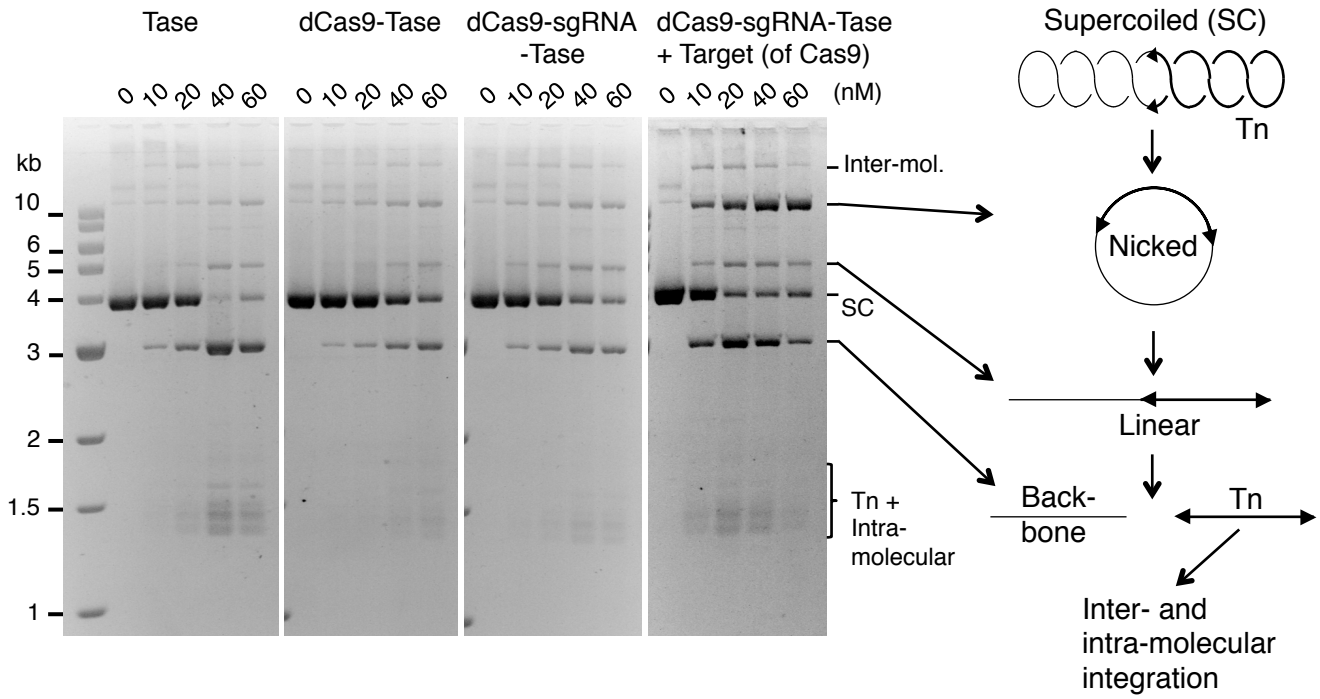
Tase
dCas9-Tase
dCas9-sgRNA -Tase
dCas9-sgRNA-Tase + Target (of Cas9)
0 10 20 40 60
(nM)
kb
10
6
5
4
3
2
1.5
1
Inter-mol.
SC
Tn + Intra-molecular
Supercoiled (SC)
Tn
Nicked
Linear
Back-bone
Tn
Inter- and intra-molecular integration

Figure 4

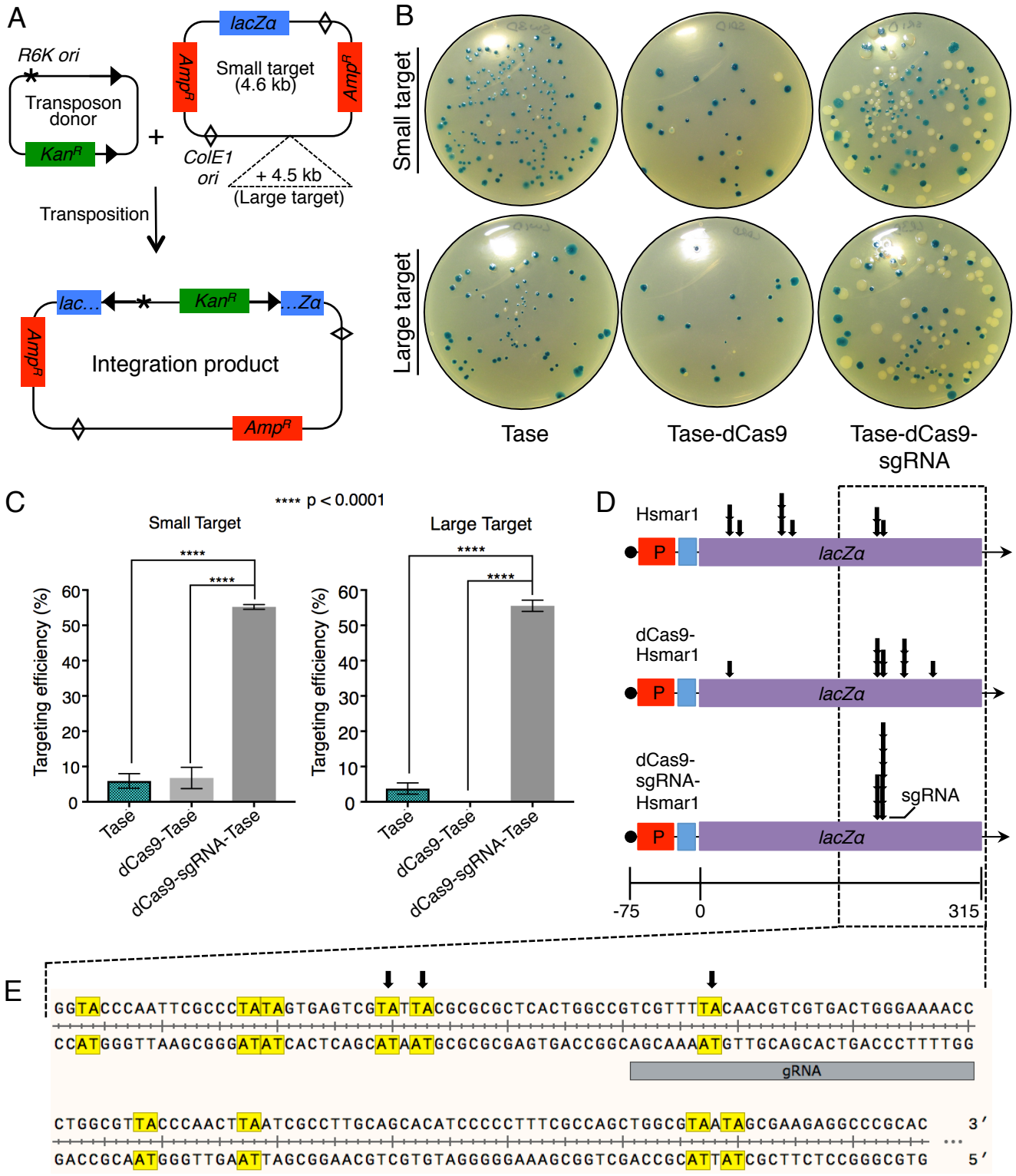
A
R6K ori
Transposon donor
Kan^R
lacZα
Amp^R
Small target (4.6 kb)
Amp^R
ColE1 ori
+ 4.5 kb
(Large target)
Transposition
lac...
Kan^R
...Zα
Amp^R
Integration product
Amp^R
B
Small target
Large target
Tase
Tase-dCas9
Tase-dCas9-sgRNA
C
**** p < 0.0001
Small Target
Large Target
Targeting efficiency (%)
Tase
dCas9-Tase
dCas9-sgRNA-Tase
D
Hsmar1
dCas9-Hsmar1
dCas9-sgRNA-Hsmar1
P
lacZα
sgRNA
-75
0
315
E
GGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACC
CCATGGGTTAAGCGGGATATCACTCAGCATAATGCGCGCGAGTGACCGGCAGCAAAATGTTGCAGCACTGACCCTTTTGG
gRNA
CTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC 3′
GACCGCAATGGGTTGAATTAGCGGAACGTCGTGTAGGGGGAAAGCGGTCGACCGCATTATCGCTTCTCCGGGCGTG 5′

Figure 5