Article

Efficient Genome Manipulation by Variants of Site-Specific Recombinases R and TD

Eugenia Voziyanova, Rachelle P. Anderson, Riddhi Shah, Feng Li and Yuri Voziyanov

School of Biosciences, *Louisiana Tech University, 1 Adams Boulevard, Ruston, LA 71272, USA*

Correspondence to Yuri Voziyanov: *School of Biological Sciences, Louisiana Tech University, 1 Adams Boulevard, Ruston, LA 71272, USA. voziyan@latech.edu*
http://dx.doi.org/10.1016/j.jmb.2015.11.002
Edited by T. K. Lu

Abstract

Genome engineering benefits from the availability of DNA modifying enzymes that have different target specificities and have optimized performance in different cell types. This variety of site-specific enzymes can be used to develop complex genome engineering applications at multiple loci. Although eight yeast site-specific tyrosine recombinases are known, only Flp is actively used in genome engineering. To expand the pool of the yeast site-specific tyrosine recombinases capable of mediating genome manipulations in mammalian cells, we engineered and analyzed variants of two tyrosine recombinases: R and TD. The activity of the evolved variants, unlike the activity of the native R and TD recombinases, is suitable for genome engineering in *Escherichia coli* and mammalian cells. Unexpectedly, we found that R recombinase benefits from the shortening of its C-terminus. We also found that the activity of wild-type R can be modulated by its non-consensus "head" sequence but this modulation became not apparent in the evolved R variants. The engineered recombinase variants were found to be active in all recombination reactions tested: excision, integration, and dual recombinase-mediated cassette exchange. The analysis of the latter reaction catalyzed by the R/TD recombinase pair shows that the condition supporting the most efficient replacement reaction favors efficient TD-mediated integration reaction while favoring efficient R-mediated integration and deletion reactions.

Introduction

Site-specific DNA recombination systems Flp/*FRT* (from a 2-μm plasmid of *Saccharomyces cerevisiae*) and Cre/*loxP* (from coliphage P1) are widely used to precisely and efficiently engineer genomes. These recombination systems are successfully applied to genetically modify bacterial, plant, invertebrate, and vertebrate cells. The exact genome modification outcome depends on the relative location and orientation of the targets that are recognized by the recombinases: it can be integration, excision, inversion, replacement, or translocation [1–10]. The use of a recombination system is usually limited to manipulating genome at one locus; thus, multiple recombination systems would allow the development of complex, temporally separated genome engineering applications. Recently, the ability of several Cre/*loxP*-like and Flp/*FRT*-like recombination systems to manipulate genomes was tested. These were bacterial systems Dre/*rox* from the P1-like phage D6; VCre/*VloxP* and SCre/*SloxP* from the plasmid p0908 of *Vibrio* sp. and the plasmid 1 of *Shewanella* sp. ANA-3, respectively; and Vika/*vox* from *Vibrio coralliilyticus* [11–14] and yeast systems B2/*B2RT* from the yeast *Zygosaccharomyces bailii* plasmid pSB2, B3/*B3RT* from the yeast *Zygosaccharomyces rouxii* plasmid pSB3, and KD/*KDRT* from the yeast *Kluyveromyces drosophilarum* plasmid pKD1 [15].

Here we report on the development and analysis of two Flp-like site-specific recombination systems: R from the yeast *Z. rouxii* plasmid pSR1 [16] and TD from the yeast *Torulaspora delbrueckii* plasmid pTD1 [17]. R and TD, like Flp, are the members of the yeast subfamily of site-specific recombinases of the tyrosine type that, to date, includes eight recombinases: Flp, R,

B2, B3, KD, KW, SM, and TD. These recombinases share sequence homology at the primary level and, most likely, a common reaction mechanism [18–22]. Wild-type R recombinase is used primarily to manipulate plant genomes [8] and was recently shown to function in fruit flies [15]. However, the low activity of R precludes its active use in genome engineering. In the present work, we tested several versions of R recombinase of different length but otherwise of the same wild-type primary sequence and found that R variants with shorter C-terminus have significantly higher activity than full-size R. We also found that the recombination activity of R is affected by its non-consensus "head" sequence. We demonstrate here that two engineered variants of R recombinase, R1-111 and R1-6, shortened to approximately the size of Flp and mutated at two positions each, mediate efficient recombination reactions in both *Escherichia coli* and mammalian cells.

The yeast *T. delbrueckii* plasmid pTD1 was identified by Blaisonneau *et al.* almost two decades ago [17] but the recombinase encoded by this plasmid, TD, is not yet actively utilized in genome engineering applications. One of the goals of the present study was to test the recombination activity of TD and, if the activity is low, to generate its variants with a higher activity. We show here that wild-type TD recombinase does not mediate efficient recombination even in *E. coli*. In contrast, the recombination activity of its variant TD1-40, shortened to approximately the size of Flp and mutated at two positions, is capable of supporting efficient recombination.

Using the evolved R and TD variants, we performed an extensive analysis of the dual recombinase-mediated cassette exchange (dual RMCE) reaction showing that the replacement reaction catalyzed by the R/TD recombinase pairs is the most efficient when TD variant is supplied at the concentration that supports efficient integration while R variants are supplied at the concentration that supports both efficient integration and deletion.

As the evolved variants of TD and R recombinases support efficient excision, integration, and dual RMCE reactions, they can help expand the repertoire of the genome engineering applications.

Results

Wild-type TD and R recombinases have low activity

Eight known yeast site-specific DNA recombinases of the tyrosine family are homologous proteins of different length, ranging from 413 amino acid residues (KW recombinase) to 568 amino acid residues (B3 recombinase) [17,19,20]. The primary sequences of the recombinases can be aligned almost up to the C-terminal end of Flp, with Tyr318 of Flp being the last invariant residue (Fig. 1a). Although Flp is not the shortest recombinase, it has the shortest C-terminus. On the other hand, the C-termini of the shortest recombinases—KW and SM—are four amino acid residues longer. The C-termini of KD, TD, R and B2, and B3 recombinases are 9, 35, 56, and 147 residues longer, respectively, than that of Flp. All recombinases can be aligned well at their N-termini. Although the exact position of the start codon of the R gene has not been experimentally determined, the gene is traditionally considered to start from the methionine codon that immediately precedes the one that aligns well with the start methionine codons of other yeast recombinases (Fig. 1a). Potential effects of the resulting 15-amino-acid "head" that is added to the consensus R sequence have been not yet investigated.

To assess the activity of the wild-type TD and R recombinases, we utilized our standard bacterial deletion reporter system [23]. In this system, the *lacZα* cassette, which is flanked by a set of respective compatible recombination targets, is excised if a recombinase is able to use these targets as substrates (Fig. 2a). If the recombinase is sufficiently active, all or almost all copies of the *lacZα* reporter in bacterial cells undergo deletion. Such cells are signaled as white when bacteria are plated in the presence of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β,D-galactopyranoside (Xgal) (Fig. 2c).

First, we analyzed the activity of the full-length, wild-type TD recombinase, TD(1-458), and a "matching" version of R recombinase, R(16-490), that has Met16 as its first amino acid residue (Fig. 1a). We found that these recombinases were not active enough to generate white colonies (Fig. 2b). In contrast, when we tested the "full-length" version of R recombinase, R(1-490), it was able to generate about 0.1% of the white colonies (Fig. 2b), suggesting that the 15-amino-acid "head" at the N-terminus of R can increase its deletion activity.

Deletion of C-terminal "tail" in R recombinase increases its activity

We next tested whether the removal of the C-terminal non-consensus "tails" of TD and R can increase the activity of the recombinases. The length of the "tail-less" versions of TD and R was chosen based on the length of the shortest recombinases: Flp and KW/SM (Fig. 1a). As the C-terminus of Flp is four amino acid residues shorter than that of KW/SM, we tested two versions of the "tail-less" variants of TD and R: with the "Flp-like" length and with the "KW/SM" length. The respective TD variants tested were TD(1-423) and TD(1-427). As R variants can either have or lack the 15-amino-acid N-terminal "head", the respective two groups of R variants tested were R(1-438) and R(16-438) and R(1-442) and R(16-442).

As was the case for the full-length, wild-type TD recombinase, the TD(1-423) and TD(1-427) variants

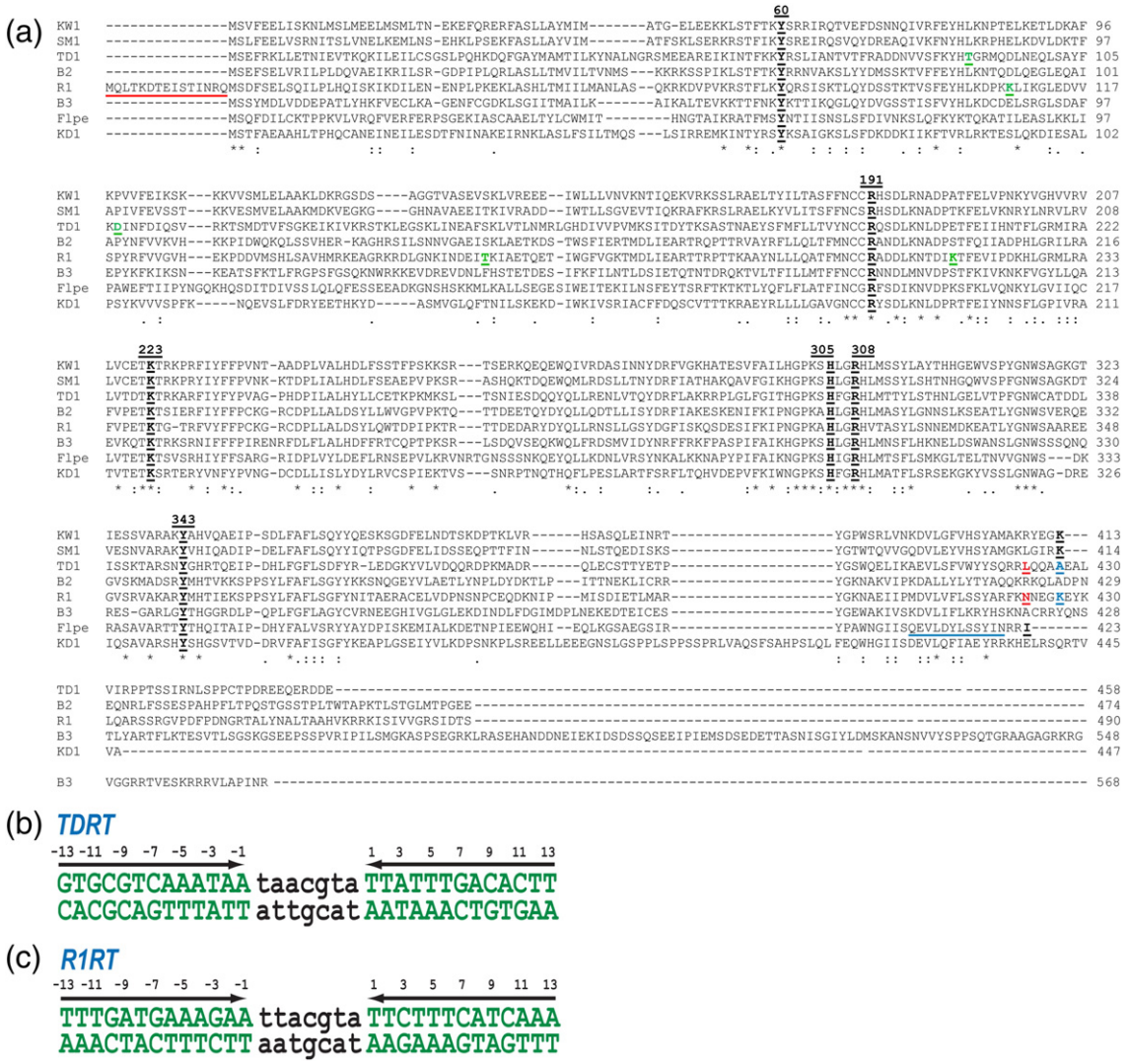

Fig. 1. (a) Alignment of the primary sequences of the yeast tyrosine recombinases Flp (Flpe sequence is shown), R, B2, B3, SM, KW, KD, and TD. Stars below the aligned sequences mark the positions of the invariant amino acid residues; colons and dots mark the positions of the amino acid residues that have similar properties. The catalytic invariant RHRY tetrad (positions 191, 305, 308, and 343 in Flp) and two other highly conserved residues (Tyr60 and Lys223 in Flp) are shown in boldface and underlined. The last amino acid residues in KW, SM, and Flp are shown in bold black and underlined. The residues in R and TD that correspond to the KW/SM-like end are shown in bold blue and underlined. The residues in R and TD that correspond to the Flp-like end are shown in bold red and underlined. The 15-amino-acid-residue "head" in R and the amino acid residues that correspond to helix Q in Flp are marked in red and blue, respectively. The amino acid residues that are modified in R1-111, R1-6, and TD1-40 are shown in bold green and underlined. (b and c) The recombination targets for TD and R that were used in this study: *TDRT* and *R1RT*, respectively.

were not sufficiently active to generate white colonies in the *lacZα* deletion assay (Fig. 2b). Similarly, two variants of R recombinase with the Flp-like length of their C-termini, R(1-438) and R(16-438), were not sufficiently active to generate white colonies. In contrast, the variants of R recombinase with the KW/SM-like length of their C-termini were more active than the "full-length" R: R(1-442) and R(16-442) generated ~5% and ~1% of white colonies, respectively (Fig. 2b).

Evolution of highly active TD and R variants

As wild-type TD and R were found to be poorly active, we wanted to evolve more efficient variants of these recombinases. In these evolution experiments, we used the better-performing R variants R(1-442) and R(16-442) and TD variants TD(1-423) and TD(1-427) as templates. The protein evolution strategy involved random mutagenesis of the recombinase genes, identification of the recombinase variants

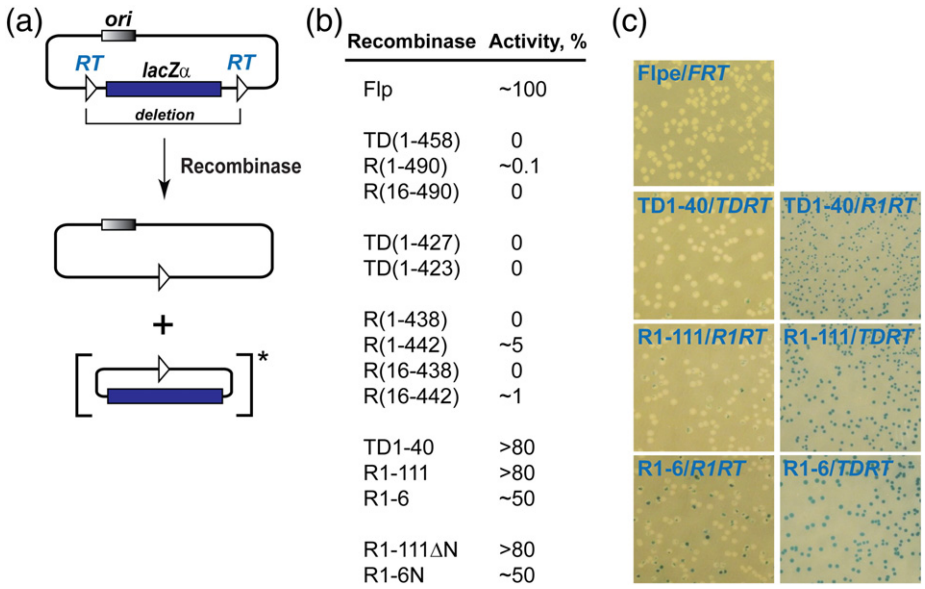

Fig. 2. The deletion activity of TD and R recombinases in *E. coli*. (a) Schematics of the deletion assay. Bacterial cells bear a reporter plasmid that contains the *lacZα* cassette flanked by the respective recombination targets (*RT*): either *TDRT* or *R1RT*. Upon transformation of these cells with a plasmid that expresses a recombinase or its variants (the expression and reporter plasmids have compatible origins of replication: p15A and ColE1, respectively), the *lacZα* cassette can be excised. The cells, in which the excision occurs, form white colonies when plated onto Xgal plates since the excised cassette is lost from the cells. (b) The deletion activity of the recombinases was analyzed by incubating the transformed bacterial cells with the inducer (0.1% of L-arabinose) for 2.5 h and then plating the cells onto the plates that contain Xgal. At 48 h after incubating the plates at 37 °C, the color of the colonies that formed was assessed and the ratio of the white to the blue colonies was calculated. (c) Examples of the bacterial colonies that formed in the experiments with Flpe, TD1-40, R1-111, and R1-6.

that are more active than the template recombinase variants, and shuffling of the resultant variant libraries followed by screening for the recombinase variants with the higher activity [24].

After the first round of random mutagenesis, we screened the resultant recombinase variant libraries and observed the characteristic white colonies in the R(1-442), R(16-442), and TD(1-427) libraries but not in the TD(1-423) libraries. These results suggested that TD variants with the Flp-like C-terminus, in contrast to the ones with the KW/SM-like C-terminus, cannot be easily generated. Therefore, we continued the protein evolution experiments only with the libraries of the recombinase variants that had the KW/SM-like C-termini: R(1-442), R(16-442), and TD(1-427).

After the second round of protein evolution, we chose three recombinase variants that exhibited higher activity than their respective templates (Fig. 2b): R1-111 and R1-6 [the derivatives of R(1-442) and R(16-442), respectively] and TD1-40 [the derivative of TD(1-427)]. R1-111 and R1-6 have two mutations each: (K108R and T161A) and (K108R and K217R), respectively (Fig. 1a). TD1-40 is also mutated at two positions: T91M and D107G (Fig. 1a). The primary sequences of TD1-40, R1-6, and R1-111 are deposited in GenBank under accession numbers GU075693, GU075694.1, and JX087436.1, respectively.

Among the evolved recombinase variants, TD1-40 has the highest deletion activity in *E. coli*, followed by R1-111 and R1-6 (Fig. 2b and c). Although the deletion activity of the variants is somewhat lower than the deletion activity of Flpe, it is high enough to recombine the reporters in at least a half of the cells (Fig. 2c).

To determine if the N-terminal "head" can modulate the activity of the evolved R variants, we engineered a derivative of R1-111 that lacks N-terminal "head" (dubbed R1-111ΔN) and a derivative of R1-6 with the N-terminal "head" added (dubbed R1-6N). Surprisingly, the activity of these derivatives was not notably different from that of the respective unmodified recombinase variants (Fig. 2b). These results indicated that the effect of the mutations that were acquired in R1-111 and R1-6 variants to support their high activity at 37 °C is much stronger than the modulatory effect of the N-terminal "head" that we observed in the wild type R environment (Fig. 2b).

The experiments to test the cross-reactivity of the TD and R variants with the non-cognate recombination targets, *R1RT*, *TDRT*, *FRT*, and *loxP*, yielded negative results indicating that the variants mainly recognize their own targets. Flp and Cre also did not recombine *R1RT* and *TDRT*. The results on the activity of TD1-40 on *R1RT* and R1-111 and R1-6 on *TDRT* are shown in Fig. 2c.

TD1-40, R1-111, and R1-6 are active in mammalian cells

We next tested the ability of TD1-40, R1-111, and R1-6 to integrate a reporter into genome of CHO Flp-In cells in the model setting (Fig. 3). A schematic of the integration assays is depicted in Fig. 3a. The integration platform reporters, which were delivered into the unique genomic *FRT* site of CHO Flp-In cells, express the DsRed gene under the control of the CMV promoter while the incoming integration reporters bear a promoter-less EGFP gene. In these reporters, the respective recombination targets (either *TDRT* or *R1RT*) are located upstream of the reporter genes. Upon integration of the incoming EGFP reporter into the platform DsRed reporter, the EGFP gene is placed under the control of the CMV promoter and its expression signals a successful integration event (Fig. 3a and b).

The integration experiments were performed by transfecting the respective incoming EGFP reporters and the vectors that express recombinase variants into the corresponding platform DsRed reporter cell lines. At 48 h post-transfection, the cells were expanded and the green cell colonies that formed were counted (Fig. 3b). Only these colonies and not individual green cells were considered as indicators of successful integration events. Ten randomly

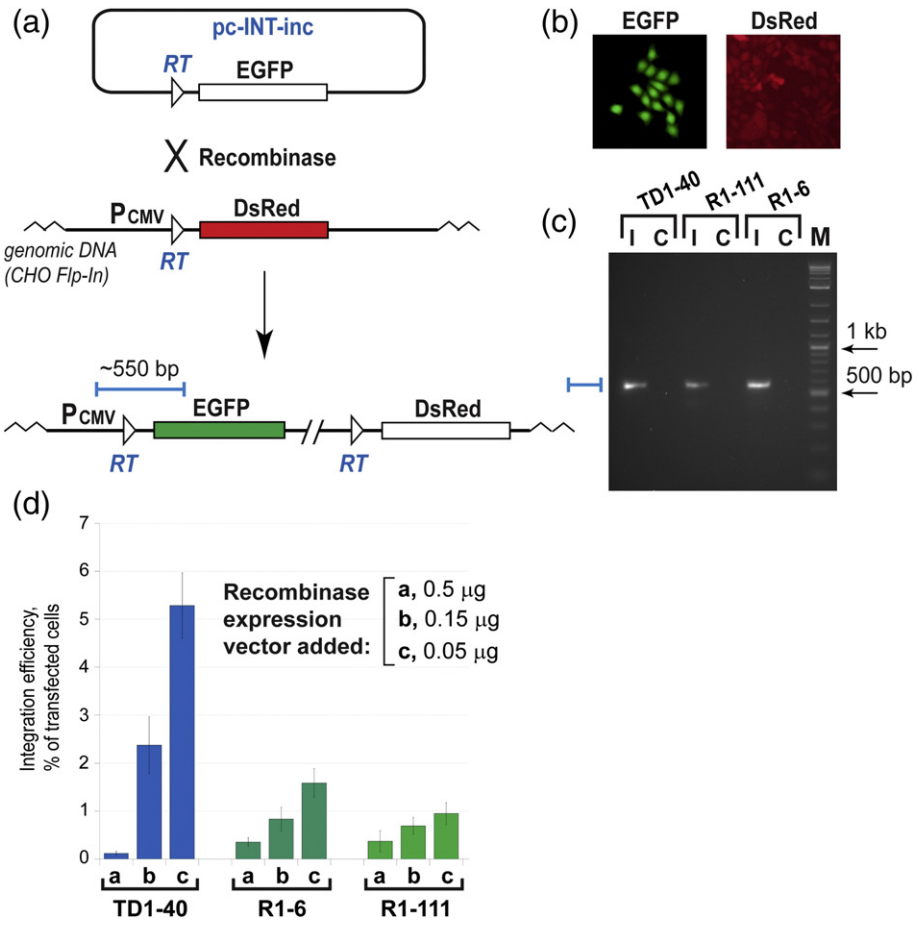

Fig. 3. TD-40, R1-111, and R1-6 are active in the integration reaction in mammalian cells. (a) Schematics of the assay. The integration platform reporter, which bears the DsRed gene under the control of the CMV promoter, is integrated into the genome of CHO Flp-In cells. Upon the recombinase-dependent integration of the incoming reporter that carries the promoter-less EGFP gene, the gene is positioned to be expressed from the CMV promoter. *RT*, either *TDRT* or *R1RT*. (b) An example of a green cell colony formed in the experiments with TD1-40. (c) The PCR analysis of the expanded green colonies formed in the experiments with TD1-40, R1-111, and R1-6. The blue bar in (a) and (c) represents the PCR product expected from a successful integration reaction. The sequencing of the PCR products confirmed their identity. Lane "I" shows the result of the PCR analysis of an expanded green colony. Lane "C" shows the control PCR analysis of the original platform reporter cells using the same primers as in lane "I". M, DNA ladder. (d) The analysis of the integration reaction catalyzed by TD1-40, R1-111, and R1-6 was performed in 24-well plates; the amount of the recombinase expression vector added per transfection is indicated. The efficiency of the integration reaction is represented by bars, which show the mean value of three experiments; the error bars indicate standard deviation.

chosen green colonies were expanded and subjected to the PCR analysis and sequencing confirming that the cells expressing EGFP are indeed the result of integration of the incoming reporter into the platform reporter (Fig. 3c).

The integration experiments were performed using different amounts of the expression vectors added at transfection (Fig. 3d). In these experiments, we noted an inverse dependence of the integration efficiency on the amount of the expression vectors: the lower the amount of the expression vector tested is, the higher the efficiency is (Fig. 3d). Under the best conditions, TD1-40 showed the highest integration efficiency: after expanding the transfected cells in 6-well plates, we observed, on average, 63 green colonies; this number corresponded to ~5% of the transfected cells. R1-6 was about three times less effective (20 colonies or ~1.5% of the transfected cells), followed by R1-111 (12 colonies or ~1% of the transfected cells).

In the experiments to test cross-reactivity of the recombinase variants, we observed a low number of pale green colonies when R variants were used to integrate the incoming *TDRT* reporter into the platform *TDRT* cells. The fraction of these colonies was about 1.0% of the green colonies that can be obtained when TD1-40 is used instead. We did not expand and analyze these green colonies because of their low number and considered them a background of the system. No green colonies were observed when TD1-40 was used to integrate the incoming *R1RT* reporter into the platform *R1RT* cells.

TD1-40 can be paired with R1-6 or R1-111 to catalyze dual RMCE

We next tested whether the TD variant can be paired with the R variants to perform dual RMCE [6,10,25,26]. In these experiments, we utilized the assay similar to that we used to analyze the replacement activity of the Flp/HK022 Int recombinase pair [26]. A schematic of this assay is shown in Fig. 4a. The neo^R-STOP cassette of the platform reporter and the EGFP-CMV cassette of the incoming reporter are flanked by *R1RT* and *TDRT*. Upon successful cassette exchange, the EGFP gene is transferred under the control of the EF1α promoter while the CMV promoter is positioned to drive the expression of the DsRed gene. Thus, the replacement event can be detected by examining the transfected cells for the expression of both EGFP and DsRed.

To assess the efficiency of the dual RMCE reaction catalyzed by the variants of R and TD, we transfected the respective platform reporter cells with three plasmids: the incoming replacement reporter and the vectors that express TD1-40 and either R1-6 or R1-111. As the integration assays, the replacement assays were performed using different amounts of the expression vectors added at transfection (Fig. 4c and d). At 48 h post-transfection, the cells were expanded and analyzed for the presence of three types of colonies: green, red, and both green and red (only colonies and not the individual cells were considered a result of a fruitful recombination event). These colonies were indicative of a successful recombination only at the *R1RT* site, only at the *TDRT* site, and at both sites (dual RMCE), respectively (Fig. 4a). To confirm that the colonies expressing both EGFP and DsRed reflect the dual RMCE outcome, we expanded 10 such colonies in the replacement experiments with the TD1-40/R1-6 pair and subjected them to the PCR analysis and sequencing. The analysis showed that these cells are indeed the product of the cassette exchange between the incoming and platform reporters (Fig. 4b).

We found that the efficiency of the replacement depends on the input of the expression vectors at transfection: the higher amount of the R variant expression vectors (0.5 μg) and the lower amount of the TD1-40 expression vector (0.05 μg and 0.015 μg) lead to the higher efficiency of the dual RMCE reaction (Fig. 4c and d). Under the optimal conditions, the TD1-40/R1-6 recombinase pair performed the cassette replacement more efficiently than did the TD1-40/R1-111 recombinase pair: the TD1-40/R1-6 recombinase pair yielded replacement in ~10% of the transfected cells (115 colonies, on average) while the TD1-40/R1-111 recombinase pair did so in only ~4% of such cells (42 colonies, on average) (Fig. 4c).

Role of TD and R variants in mediating dual RMCE

We noted that the amount of the TD1-40 expression vector that is optimal for replacement (0.05 μg) matches the amount of the vector that is optimal for integration in the assays described in the previous section (Figs. 3d and 4c). In contrast, the amount of the R expression vectors that is optimal for replacement (0.5 μg; Fig. 4c and d) is significantly higher than that optimal for integration: 0.05 μg (Fig. 3d). Since these data might reflect the mechanism of the dual RMCE reaction (replacement through two integrations *versus* through integration then deletion), we wanted to determine what amount of the expression vectors is optimal for the integration and deletion reactions in the dual RMCE reporter setting (Fig. 5).

To test the efficiency of the integration reaction, we transfected the replacement platform cells with the incoming replacement reporter and either the TD1-40 or R1-6 or R1-111 expression vectors and monitored the appearance of red or green colonies that would signify integration of the reporter into *TDRT* or *R1RT*, respectively (Fig. 4a). The results demonstrated that the lower replacement optimal

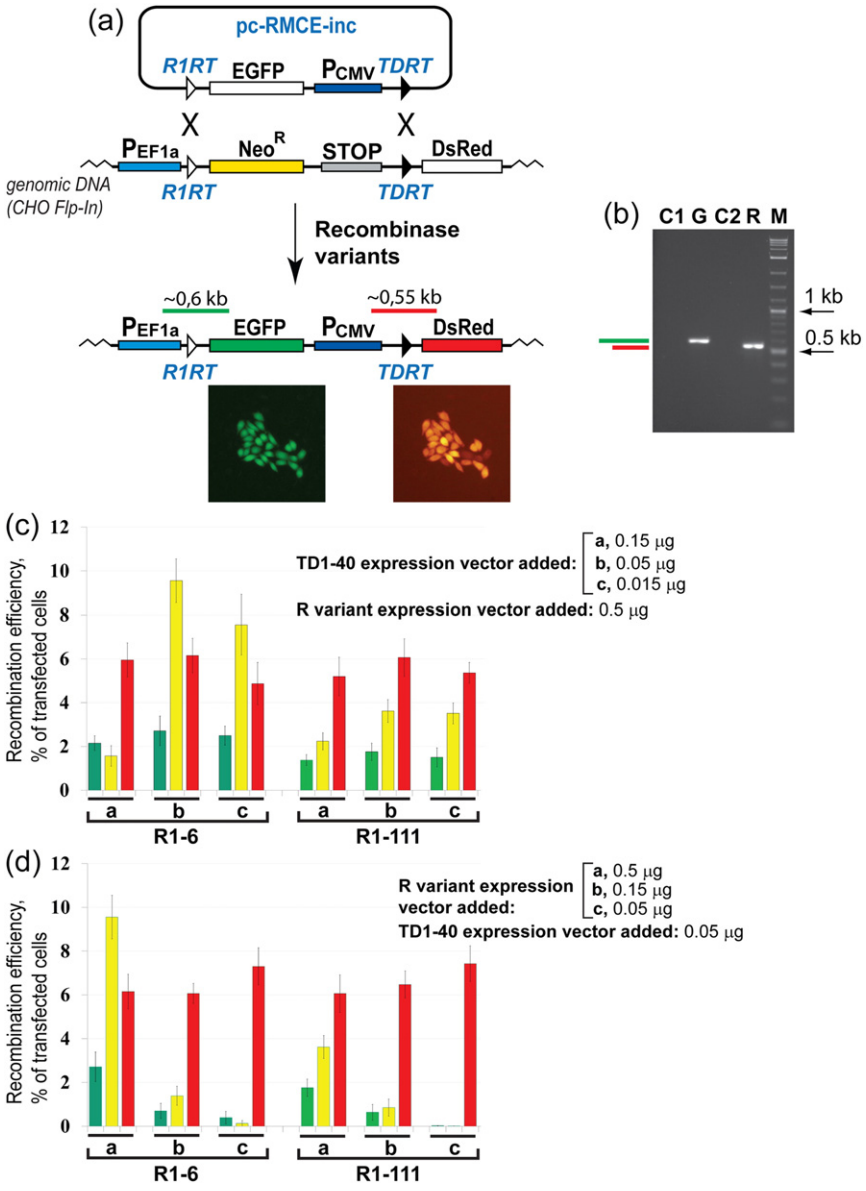

Fig. 4. Dual RMCE catalyzed by the R/TD recombinase variant pairs. (a) Schematics of the replacement assay. A swap between the neo^R-STOP cassette located in the pre-integrated platform reporter and the EGFP-CMV cassette located in the incoming reporter positions the EGFP and DsRed genes for the expression from the EF1α and CMV promoters, respectively. An example of a green/red colony formed in the R1-6/TD1-40 experiments is shown below the dual RMCE scheme. (b) The PCR analysis of a typical expanded green/red colony formed in the replacement experiments using the R1-6/TD1-40 pair. The green and red bars in (a) and (b) represent the PCR products that are indicative of a successful replacement reaction. Lanes "G" and "R" show the PCR analyses of the expanded green/red colony with the primers that anneal on the EF1α promoter and the EGFP gene, as well as on the CMV promoter and the DsRed gene, respectively. Lanes "C1" and "C2" show the control PCR analyses of the replacement platform reporter cells using the same set of primers as in lanes "G" and "R", respectively. M, DNA ladder. (c and d) The analysis of the replacement and integration reactions catalyzed by TD1-40, R1-111, and R1-6 and their respective pairs. The amounts of the recombinase expression vectors added at transfection are indicated. The efficiency of dual RMCE catalyzed by the pairs of the respective R variants and TD1-40 is represented by yellow bars. The efficiency of the integration reaction is represented by green bars (integration into *R1RT* site) or red bars (integration into *TDRT* site). The bars show the mean value of three experiments; the error bars indicate standard deviation.

amount of the TD1-40 expression vector (0.05 μg) and the higher replacement optimal amount of the R1-6 and R1-111 expression vectors (0.5 μg) are also optimal for the integration reaction in the replacement reporter vector setting (Figs. 4c and d and 5a). The data on the R-mediated integration in

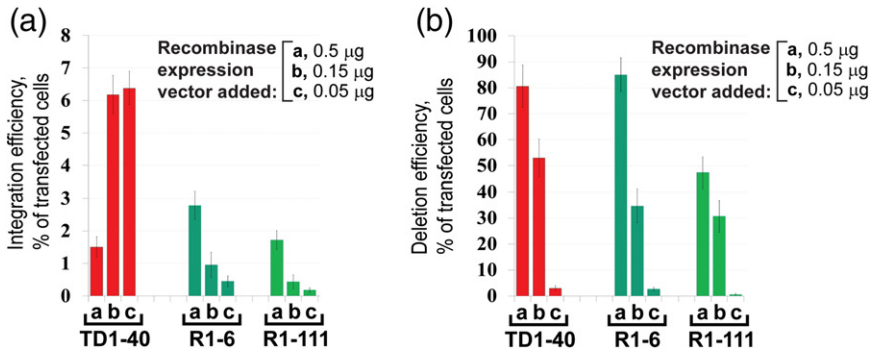

Fig. 5. The analysis of the integration and deletion reactions catalyzed by TD1-40, R1-111, and R1-6 in the dual RMCE reporter setting. (a) Integration reactions were performed by transfecting the replacement platform cells with the replacement incoming reporter and the vector that expresses either TD1-40 or R1-111 or R1-6. The amounts of the recombinase expression vectors added at transfection are indicated. The efficiency of the integration reaction is represented by red bars (TD1-40-mediated integration) or by green bars (R1-6- and R1-111-mediated integration). The bars show the mean value of three experiments; the error bars indicate standard deviation. (b) Deletion reactions were performed by transfecting the expanded green or red colonies obtained in the respective integration experiments. These cells were transfected with the vector that expresses either TD1-40 (green cells) or R1-111 (red cells) or R1-6 (red cells). The amounts of the recombinase expression vectors added at transfection are indicated. The efficiency of the deletion reaction is represented by red bars (TD1-40-mediated deletion) or by green bars (R1-6- and R1-111-mediated deletion). The bars show the mean value of three experiments; the error bars indicate standard deviation.

different vector settings (Figs. 3d and 5a) point to their apparent functional differences that have to be taken into account as they can affect the efficiency of the recombination reaction under analysis; these differences need to be further investigated.

The integration experiments described above allowed us to test cross-reactivity of the recombinase variants in the dual RMCE reaction. In the experiments with R1-6 or R1-111, we did not observe green/red or just red colonies. In the experiments with TD1-40, a low fraction of green/red and just green colonies was noted: ~1.0% of the number of the respective colonies that can be obtained when a respective recombinase pair is used to perform the replacement reaction. We did not expand and characterize these colonies due to their low numbers and considered them a background of the system.

In the experiments to assess the deletion activity of the recombinase variants in the dual RMCE reporter setting, we used the expanded green and red colonies obtained after integrating the incoming reporter into the platform reporter cells using just R1-6 and or just TD1-40, respectively (Fig. 4a). To test the deletion activity, we transfected the green cells with different amounts of the vector that expresses TD1-40 and we transfected the red cells with different amounts of the vector that expresses either R1-6 or R1-111. The results show that all recombinase variants exhibit the highest deletion activity when the expression vector is supplied at the highest amount tested: both TD1-40 and R1-6 were able to perform deletion in about 80–85% of the transfected cells (on average, we observed 1015 red and 1074 green colonies, respectively). R1-111 was about twice as less effective: deletion was seen in about 50% of the transfected cells (on average, 629 green colonies were observed; Fig. 5b).

The integration and deletion experiments pointed to the inverse recombination optima for TD1-40: the replacement and integration reactions were the most efficient at the lowest supply of the TD1-40 expression vector while the efficiency of the deletion reaction was the best at the highest supply of the vector (Figs. 4c and 5). In contrast, all recombination reactions mediated by R variants were the most efficient at the highest supply of the respective expression vectors. We will discuss how the results of the integration and deletion experiments can fit into the alternative models of the cassette replacement in the next section.

Discussion

Our work on engineering active variants of site-specific tyrosine recombinases R and TD uncovered several unexpected features of R recombinase and helped to analyze the mechanism of the dual RMCE reaction.

We found that wild-type R and TD recombinases are not sufficiently active even in *E. coli*, and thus, they do not appear to be very useful in genome engineering applications, particularly in mammalian cells. Before experimenting with generating more active variants of the recombinases, we wanted to test whether the N-terminal “head” of the “full-length”

R and the C-terminal "tails" of R and TD (Fig. 1) affect the efficiency of the recombinases since, ultimately, we wanted to obtain the shortest recombinase variants possible.

The comparison of the activity of the "full-length" R and its "head-less" version in bacterial cells showed that the "head" sequence of R significantly enhances the deletion activity of the recombinase (Fig. 2b). We also found that R positively responds to the shortening of its C-terminus to the longer KW/SM-like end but not to the shorter Flp-like end (Figs. 1a and 2b): the respective R variants—R(1-442) and R(16-442)—were about 50-fold and 10-fold more active than the "full-length" R in the deletion assay (Fig. 2b). In contrast, wild-type TD recombinase did not respond positively to the shortening of its C-terminus to either end. However, the subsequent experiments with the libraries of the randomly mutated TD variants shortened either to the Flp-like end or to the KW/SM-like end showed that the variants of TD that are sufficiently active to be detected by our bacterial deletion screening system [23] can be generated only when TD is shortened to the KW/SM-like end. These results suggest that TD shares the R's preference for the longer KW/SM-like end over the shorter Flp-like end.

The reason for this preference is not apparent. In Flp, the very last amino acid residues form helix Q [27], which contains two invariant residues and four homologous residues (Fig. 1). Helix Q, like "pig in a blanket", is surrounded by helixes P, O, and I and the base of the beta sheet formed by beta strands 1, 2, and 3 (Fig. 6d). Helix Q directly interacts with the recombinase structures that participate in the monomer–monomer interactions and catalysis (Fig. 6a and d). The importance of helix Q is signified by the observation that the deletion of the last six amino acids in Flp abolishes its activity [28]. We hypothesize that the four residues located between the shorter Flp-like end and the longer KW/SM-like end in R and TD recombinases interact with the structures that, in Flp, correspond to the loop between beta strand 3 and helix I and/or with helix O (Fig. 6d). These potential interactions in R and TD might be preserved in the variants of the recombinases that are shortened to the KW/SM-like end but not to the Flp-like end. Alternatively, the shortening to the Flp-like end might affect the folding and/or stability of the recombinase variants.

Both R and TD recombinases responded positively to the random mutagenesis yielding variants that are active in bacterial and mammalian cells (Figs. 2–4). R variants R1-111 and R1-6 have one common mutation: at position 108. In Flp, the amino acid residue at this position, Ile88 (Figs. 1a and 6a and b), is located in helix C that participates in monomer–monomer interactions by directly interacting with helix D from a neighboring monomer [27]. The other mutations in R1-111 and R1-6 are at positions 161 and 217, respectively (Fig. 1a). In Flp, the amino acid residue at the position that corresponds to the position 161 in R is Leu144 (Figs. 1a and 6a and c). This residue is part of helix E and is located inside a pocket formed by Ile295, Pro302, Ile306, His309 of one monomer, and two residues from a neighboring monomer: Thr342 and His345 [27]. This array of the amino acid residues interacts with the catalytic residues His305, His308, and Tyr343 either directly or through other residues (Fig. 6c). It is worth noting that one of the mutations in Flpe—a variant of Flp that was evolved to be more active at 37 °C [29]—is located in the same group of residues: at position 294.

The second amino acid residue that is mutated in R1-6—at position 217—corresponds to lysine at position 201 in Flp (Fig. 1). Lys201 is located in beta strand 1 that forms beta sheet with beta strands 2 and 3 (Fig. 6a and d). This amino acid is part of the group of the conserved residues called Box I that contains catalytic Arg191 [18,19]. Lys201 and the nearby residues interact with the Flp regions that are distant in the primary structure: with beta strand 5 and with the loop between helixes P′ and Q. Because of its location, a mutation at position 201 can affect communications between different structures within the recombinase and/or influence its folding.

The two mutations in TD1-40—at positions 91 and 107—correspond to the residues 83 and 99 in Flp, in which they are located between beta strand 2n and helix C and between helix C and beta strand 3n, respectively (Fig. 6a and b). Beta strands 1n, 2n, and 3n form beta sheet that interacts with helix C, which, in turn, directly interacts with helix D of the neighboring monomer [27] suggesting that the amino acids at positions 91 and 107 in TD1-40 influence monomer–monomer interactions.

Taken together, the analysis of the location of the mutations in R1-6, R1-111, and TD1-40 indicates that R and TD recombinases respond to the selective pressure of being more active at 37 °C by modifying the amino acid residues that are located primarily at the monomer–monomer interfaces and in the vicinity of the catalytic residues.

Modification of the monomer–monomer interfaces in the evolved recombinase variants can, in principle, relax their target specificity as compared to the wild-type enzymes [30]. For R1-6, R1-111, and TD1-40 that are not extensively mutated, this potential effect is unlikely a concern since these variants do not cross-react with their respective non-cognate targets although *TDRT* and *R1RT* are quite similar (Fig. 1b and c). Also, the extensively utilized Flpe that was evolved using an approach that is similar to that used in the present work [29] has mutations at the monomer–monomer interfaces and apparently does not exhibit abnormal specificity compared to Flp.

It was unexpected to find that the N-terminal "head" of the "full-length" R can improve the efficiency of a

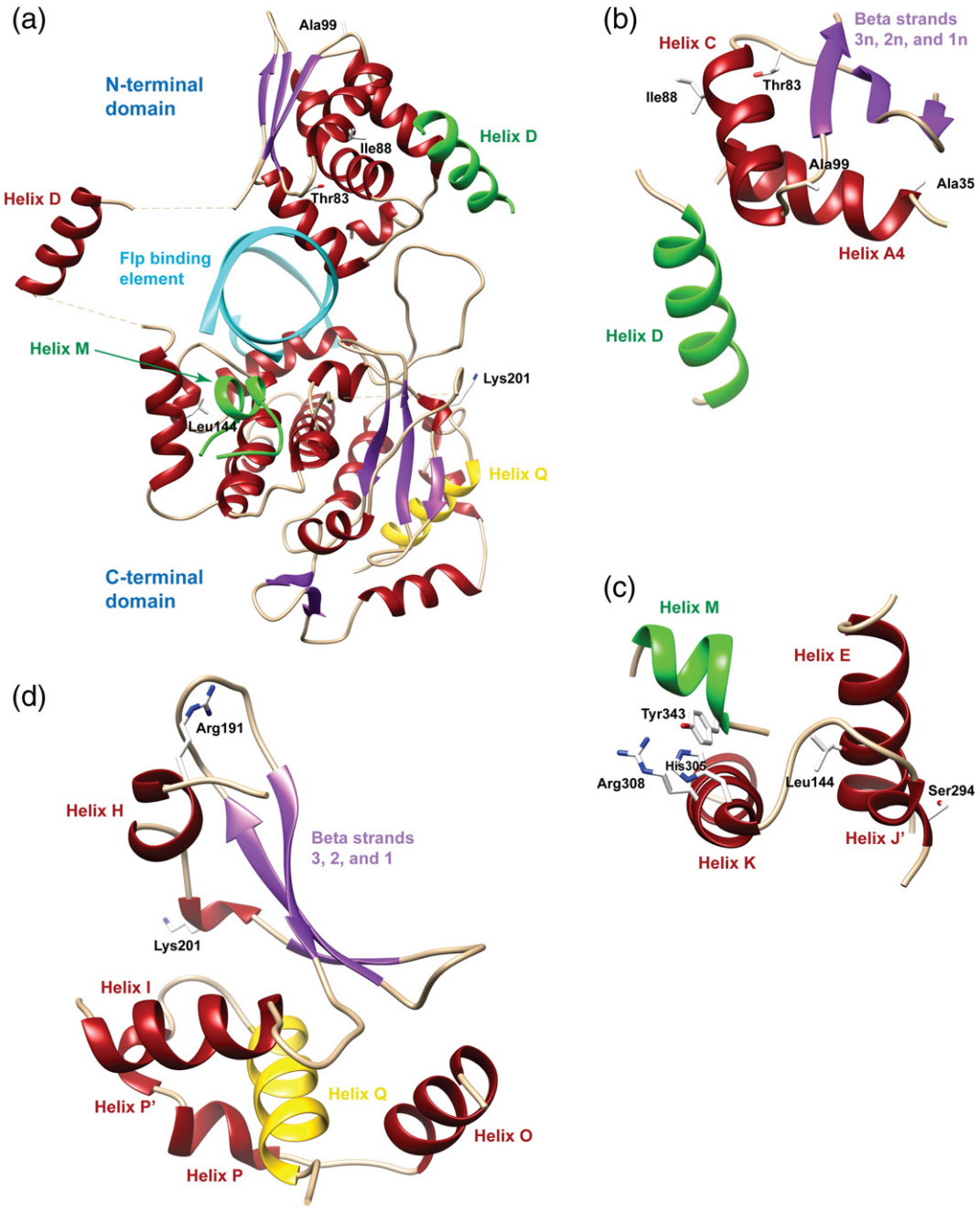

Fig. 6. Relative location of Flp amino acids that correspond to the amino acids that are mutated in R1-6, R1-111, and TD1-40. The coordinates of Flp/DNA co-crystal (PDB code 1FLO [27]) were used to visualize Thr83 (corresponds to Thr91 in TD), Ile88 (corresponds to Lys108 in R), Ala99 (corresponds to Aps107 in TD), Leu144 (corresponds to Thr161 in R), and Lys201 (corresponds to Lys217 in R). (a) Flp monomer A bound to Flp binding element and interacting with helix D from Flp monomer C and helix M from Flp monomer D (both helixes are show in green). Helix Q of Flp monomer A is shown in yellow. (b) Close-up view of Thr83 (Thr91 in TD), Ile88 (Lys108 in R), and Ala99 (Aps107 in TD). (c) Close-up view of Leu144 (Thr161 in R). Catalytic residues His305, Arg308, and Tyr343 are shown along with Ser294 that is mutated in Flpe [29]. (d) Close-up view of helix Q and Lys201 (Lys217 in R). Catalytic residue Arg191 is also shown. The color coding in (b), (c), and (d) is the same as in (a).

recombination reaction (Fig. 2b). It was even more puzzling to see that the effect was prominent in the wild-type R environment but was not noticeable in the evolved R variants. We currently do not have a clear explanation for this phenomenon. It is possible that the "head" sequences some weak tetramer stabilizing

activity but this activity is shadowed by a much stronger tetramer stabilizing activity of the mutations in R1-6 and R1-111 variants. The *in vitro* studies are apparently needed to understand the role of the "head" sequence.

An important result of the present work is the demonstration that the evolved variants of R and TD recombinases can catalyze an efficient dual RMCE (Fig. 4). Since these variants do not cross-react with *FRT* and *loxP* and since Flp and Cre do not react with *R1RT* or *TDRT*, the evolved variants of R and TD can be therefore paired with these dual RMCE-proved recombinases [6,10,25,26] to mediate multi-faceted replacement of genomic fragments.

Both the integration and the replacement activity of the R and TD variants are comparable to that of Flp and HK022 [25,26]. In fact, under optimal conditions, the integration efficiency of TD-40 and R1-6 is higher than that of any other tyrosine recombinase we experimented with: ~5–6% and ~1.5–3% of the transfected cells, respectively (Figs. 3d and 5a). In contrast, the integration efficiency of HK022 Int is about 0.8% of transfected cells [26]. Flp, as a control, routinely gives us integration in ~0.1% of the transfected cells (as we did not optimize the parameters of this reaction, it is possible that Flp can mediate integration more efficiently).

Although the efficiency of the cassette exchange by the TD1-40/R1-6 pair is lower than that of the other recombinase pairs we tested, ~10% (Fig. 4c) *versus* ~12% by the Flp/HK022 Int pair [26] *versus* 35–45% by the Flp/Cre pair [25], it is high enough to identify the positive clones even without a selective pressure.

The cumulative data on the integration, deletion, and replacement activity of TD1-40 in the dual RMCE reporter setting (Figs. 4 and 5) indicate that the ability of TD1-40 to efficiently mediate dual RMCE positively correlates with its ability to support efficient integration but not deletion. These data suggest that, during dual RMCE, this recombinase mainly performs integration of the incoming reporter into the platform reporter. The functional behavior of R variants is different: R1-6 and R1-111 efficiently mediate dual RMCE under the condition that supports both integration and deletion reactions. The data on R, therefore, cannot help to discriminate between two alternative models of the cassette replacement: two simultaneous integrations (crossovers) at the flanks of the replacement cassettes *versus* integration of the incoming reporter and then deletion of the respective segments of the reporters. We cannot rule out the possibility that, for the dual RMCE vector setting we used in this study, both models are correct and which dual RMCE pathway is responsible for the replacement depends on whether two recombinases are able to act simultaneously at the flanks of the reporter cassette or one at a time.

The ability of TD1-40 to efficiently deliver vectors into a unique, pre-integrated *TDRT* site was successfully utilized to construct TD-In cell lines [25,26]. TD1-40 and R variants described in the present study, although tested only in *E. coli* and mammalian cells, should be also useful to engineer genomes in plants and fruit flies.

Materials and Methods

Bacterial experiments

The *E. coli* strain NEB 10-beta [*araD139* Δ(*ara-leu*)*7697 fhuA lacX74 galK* (*φ80* Δ(*lacZ*)*M15*) *mcrA galU recA1 endA1 nupG rpsL* (*StrR*) Δ(*mrr-hsdRMS-mcrBC*)] from New England Biolabs was used in all bacterial experiments. The deletion assays in *E. coli* were performed essentially as described in Voziyanov *et al.* [23].

DNA mutagenesis and shuffling

PCR mutagenesis of the wild-type recombinase genes and DNA shuffling of the variant recombinase genes were performed as described in Bolusani *et al.* [24]. In brief, mutagenic PCR was performed using Taq polymerase in the Mg-free buffer supplemented with 0.25 mM $MnCl_2$, 0.5 mM $MgCl_2$, 200 μM dNTP, 100 nM of each primer, and ~100 ng of DNA template. After amplification, the PCR product was digested with SacI and SphI (TD recombinase variant gene libraries) or with SacI and HindIII (R recombinase variant gene libraries) and was ligated to pBAD33 [31] digested with the respective enzymes.

To perform DNA shuffling, we amplified the genes of interest using Taq polymerase under non-mutagenic conditions. The PCR products were mixed in an equimolar ratio and fragmented with DNase I for 5 min on ice in the buffer containing 50 mM Tris–HCl (pH 7.4) and 10 mM $MnCl_2$. The resultant DNA fragments were reassembled in the two-step PCR reaction: first, using Pfu-Ultra polymerase without the addition of specific primers and then using Taq polymerase with specific primers that anneal outside the coding region of the genes. The amplified recombinase gene libraries were cloned as described above.

Cell lines

Derivatives of Chinese hamster ovary cells, CHO Flp-In cells (Invitrogen), were used as model mammalian cells. CHO Flp-In cells were propagated in F-12K medium supplemented with zeocin (100 mg/l). After the integration and the dual RMCE platform reporter vectors were delivered into genome of CHO Flp-In, the resultant cell lines were propagated in F-12K medium supplemented with hygromycin (550 mg/l).

Vectors for recombination experiments in mammalian cells

Vectors that were used to express recombinase genes and to report recombination events in mammalian cells were

based on the vectors of the Flp-In system: pOG44 and pcDNA5/FRT, respectively. To express TD and R variants in mammalian cells, we cloned their genes into pOG44 in place of the Flp-F70L variant gene as in Anderson *et al.* [25].

pcDNA5/FRT was used to construct two types of reporter vectors: to report integration and dual RMCE reactions. To construct the platform integration reporters, pc-INT-plt, we PCR amplified the DsRed-neo^R cassette from pIRES2-DsRed-Express (Clontech) and cloned it into pcDNA5/FRT between NheI and XhoI; either *TDRT* or *R1RT* was placed upstream of the DsRed-neo^R cassette during its amplification. The incoming integration reporters were constructed by cloning the PCR amplified EGFP-neo^R cassette from pIRES2-EGFP (Clontech) into pcDNA5/FRT between NheI and EcoRI located in the Hygro^R gene; the respective recombination targets were placed upstream of the EGFP-neo^R cassette during its amplification. In the resultant reporters, the CMV promoter was deleted by digesting the plasmids with NheI and MluI, blunting the ends with Klenow and self-ligating to obtain the incoming integration reporters pc-INT-inc.

The platform and incoming reporters for the dual RMCE experiments, pc-RMCE-plt and pc-RMCE-inc, respectively, were constructed by modifying the corresponding reporters described in Voziyanova *et al.* [26]. The modification included flanking the neo^R-STOP and EGFP-CMV cassettes with *R1RT* and *TDRT* and introducing the *FRT* sequence upstream of the hygro^R gene in the platform reporter to match the pcDNA5/FRT configuration.

Construction of platform reporter cell lines

The platform reporter cell lines for the integration and replacement experiments were constructed as follows. CHO Flp-In cells were co-transfected, in 6-well plates using Polyfect (Qiagen), with the respective platform reporters (0.2 μg) and pOG44 (2 μg), which expresses Flp variant Flp(F70L). At 48 h post-transfection, 1/6 of the cells were transferred into a 100-mm plate containing medium supplemented with hygromycin (550 mg/l). About 10 days later, several hygromycin-resistant colonies were transferred into 96-well plate and their sensitivity to zeocin was tested. Zeocin-sensitive colonies were expanded and used in the subsequent experiments.

Integration experiments

The experiments to test the integration activity of TD and R variants in CHO cells were performed in 24-well plates. The incoming reporters (0.5 μg) were co-transfected using DNA-In transfection reagent (Molecular Transfers) into the respective integration platform reporter cells with the vectors that express the recombinase variants to be analyzed; the amount of the vectors added at transfection is indicated in Fig. 3. The variants of the incoming integration reporters that express EGFP from the CMV promoter were used to estimate the rate of transfection, which was, on average, 50–60%. At 48 h post-transfection, 1/16 of the cells were transferred into 6-well plates, the cells were allowed to become confluent, and the number of the green colonies was counted. Several integration-positive colonies from these plates were expanded and analyzed.

Dual RMCE experiments

The replacement experiments were performed similar to the integration experiments. The incoming reporter (0.5 μg) and the vectors that express TD1-40 and R recombinase variants were co-transfected into the replacement platform reporter cells using DNA-In reagent; the amounts of the vectors added at transfection are indicated in Fig. 4. The efficiency of transfection was estimated as in the integration experiments. Two days after transfection, 1/16 of the cells were transferred into 6-well plates, expanded until confluent, and the green/red colonies, just green colonies and just red colonies that formed, were counted. Several green/red colonies from these plates were expanded and analyzed.

The integration and the deletion experiments in the dual RMCE reporter setting were performed similar to the replacement experiments.

Other methods

Multiple sequence alignment was performed using ClulstalW2 (European Bioinformatics Institute, European Molecular Biology Laboratory). Plasmid DNA was isolated using E.Z.N.A. Plasmid Mini Kit (Omega Bio-Tek). Amplification of the DNA fragments used for cloning was performed using Pfu-Ultra polymerase (Agilent). PCR analysis of the mammalian genomic DNA was performed using Taq polymerase (New England Biolabs). Genomic DNA from CHO cells was isolated using GeneJET Genomic DNA Purification Kit (Thermo). General genetic engineering experiments were performed as described in Molecular Cloning Manual [32]. Flp three-dimensional structure was analyzed using Swiss-PdbViewer [33] and visualized to prepare Fig. 5 using UCSF Chimera [34].

Acknowledgements

This work was supported in part by the National Institutes of Health grant R01GM085848 to Y.V. and by Louisiana Tech University.

Author Contributions: Y.V. designed the research; E.V, R.P.A., R.S., F.L., and Y.V. performed the research; Y.V. analyzed the data and wrote the manuscript.

Conflict of Interest: The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2015.11.002.

Received 17 January 2015;
Received in revised form 17 October 2015;
Accepted 2 November 2015
Available online 7 November 2015

Keywords:
site-specific recombination;
tyrosine recombinases;
genome engineering;
RMCE;
protein engineering

Abbreviations used:
dual RMCE, dual recombinase-mediated cassette exchange; Xgal, 5-bromo-4-chloro-3-indolyl-β,D-galactopyranoside.

References

[1] K.G. Golic, S. Lindquist, The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome, Cell 59 (1989) 499–509.
[2] N.J. Kilby, M.R. Snaith, J.A. Murray, Site-specific recombinases: Tools for genome engineering, Trends Genet. 9 (1993) 413–421.
[3] A.A. Mills, A. Bradley, From mouse to man: Generating megabase chromosome rearrangements, Trends Genet. 17 (2001) 331–339.
[4] F. Buchholz, Y. Refaeli, A. Trumpp, J.M. Bishop, Inducible chromosomal translocation of AML1 and ETO genes through Cre/loxP-mediated recombination in the mouse, EMBO Rep. 1 (2000) 133–139.
[5] D. Werdien, G. Peiler, G.U. Ryffel, FLP and Cre recombinase function in *Xenopus* embryos, Nucleic Acids Res. 29 (2001) e53.
[6] M. Lauth, F. Spreafico, K. Dethleffsen, M. Meyer, Stable and efficient cassette exchange under non-selectable conditions by combined use of two site-specific recombinases, Nucleic Acids Res. 30 (2002) e115.
[7] S. Glaser, K. Anastassiadis, A.F. Stewart, Current issues in mouse genome engineering, Nat. Genet. 37 (2005) 1187–1193.
[8] K. Nanto, K. Yamada-Watanabe, H. Ebinuma, Agrobacterium-mediated RMCE approach for gene replacement, Plant Biotechnol. J. 3 (2005) 203–214.
[9] D. Wirth, L. Gama-Norton, P. Riemer, U. Sandhu, R. Schucht, H. Hauser, Road to precision: Recombinase-based targeting technologies for genome engineering, Curr. Opin. Biotechnol. 18 (2007) 411–419.
[10] M. Osterwalder, A. Galli, B. Rosen, W.C. Skarnes, R. Zeller, J. Lopez-Rios, Dual RMCE for efficient re-engineering of mouse mutant alleles, Nat. Methods 7 (2010) 893–895.
[11] B. Sauer, J. McDermott, DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages, Nucleic Acids Res. 32 (2004) 6086–6095.
[12] K. Anastassiadis, J. Fu, C. Patsch, S. Hu, S. Weidlich, K. Duerschke, et al., Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice, Dis. Model. Mech. 2 (2009) 508–515.
[13] E. Suzuki, M. Nakayama, VCre/VloxP and SCre/SloxP: New site-specific recombination systems for genome engineering, Nucleic Acids Res. 39 (2011) e49.
[14] M. Karimova, J. Abi-Ghanem, N. Berger, V. Surendranath, M.T. Pisabarro, F. Buchholz, Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system, Nucleic Acids Res. 41 (2012) e37.
[15] A. Nern, B.D. Pfeiffer, K. Svoboda, G.M. Rubin, Multiple new site-specific recombinases for use in manipulating animal genomes, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 14198–14203.
[16] H. Araki, N. Nakanishi, B.R. Evans, H. Matsuzaki, M. Jayaram, Y. Oshima, Site-specific recombinase, R, encoded by yeast plasmid pSR1, J. Mol. Biol. 225 (1992) 25–37.
[17] J. Blaisonneau, F. Sor, G. Cheret, D. Yarrow, H. Fukuhara, A circular plasmid from the yeast *Torulaspora delbrueckii*, Plasmid 38 (1997) 202–209.
[18] I. Utatsu, S. Sakamoto, T. Imura, A. Toh-e, Yeast plasmids resembling 2 micron DNA: Regional similarities and diversities at the molecular level, J. Bacteriol. 169 (1987) 5537–5545.
[19] J.W. Chen, B.R. Evans, S.H. Yang, H. Araki, Y. Oshima, M. Jayaram, Functional analysis of box I mutations in yeast site-specific recombinases Flp and R: Pairwise complementation with recombinase variants lacking the active-site tyrosine, Mol. Cell. Biol. 12 (1992) 3757–3765.
[20] J. Lee, M.C. Serre, S.H. Yang, I. Whang, H. Araki, Y. Oshima, et al., Functional analysis of Box II mutations in yeast site-specific recombinases Flp and R. Significance of amino acid conservation within the Int family and the yeast sub-family, J. Mol. Biol. 228 (1992) 1091–1103.
[21] X.J. Chen, Y.S. Cong, M. Wesolowski-Louvel, Y.Y. Li, H. Fukuhara, Characterization of a circular plasmid from the yeast *Kluyveromyces waltii*, J. Gen. Microbiol. 138 (1992) 337–345.
[22] S.H. Yang, M. Jayaram, Generality of the shared active site among yeast family site-specific recombinases. The R site-specific recombinase follows the Flp paradigm [corrected], J. Biol. Chem. 269 (1994) 12789–12796.
[23] Y. Voziyanov, A.F. Stewart, M. Jayaram, A dual reporter screening system identifies the amino acid at position 82 in Flp site-specific recombinase as a determinant for target specificity, Nucleic Acids Res. 30 (2002) 1656–1663.
[24] S. Bolusani, C.H. Ma, A. Paek, J.H. Konieczka, M. Jayaram, Y. Voziyanov, Evolution of variants of yeast site-specific recombinase Flp that utilize native genomic sequences as recombination target sites, Nucleic Acids Res. 34 (2006) 5259–5269.
[25] R.P. Anderson, E. Voziyanova, Y. Voziyanov, Flp and Cre expressed from Flp-2A-Cre and Flp-IRES-Cre transcription units mediate the highest level of dual recombinase-mediated cassette exchange, Nucleic Acids Res. 40 (2012) e62.
[26] E. Voziyanova, N. Malchin, R.P. Anderson, E. Yagil, M. Kolot, Y. Voziyanov, Efficient Flp-Int HK022 dual RMCE in mammalian cells, Nucleic Acids Res. 41 (2013) e125.
[27] Y. Chen, U. Narendra, L.E. Iype, M.M. Cox, P.A. Rice, Crystal structure of a Flp recombinase-Holliday junction complex: Assembly of an active oligomer by helix swapping, Mol. Cell 6 (2000) 885–897.
[28] A.A. Amin, P.D. Sadowski, Synthesis of an enzymatically active FLP recombinase *in vitro*: Search for a DNA-binding domain, Mol. Cell. Biol. 9 (1989) 1987–1995.
[29] F. Buchholz, P.O. Angrand, A.F. Stewart, Improved properties of FLP recombinase evolved by cycling mutagenesis, Nat. Biotechnol. 16 (1998) 657–662.
[30] J.H. Konieczka, A. Paek, M. Jayaram, Y. Voziyanov, Recombination of hybrid target sites by binary combinations of Flp variants: Mutations that foster interprotomer collaboration and enlarge substrate tolerance, J. Mol. Biol. 339 (2004) 365–378.
[31] M.E. Kovach, P.H. Elzer, D.S. Hill, G.T. Robertson, M.A. Farris, R.M. Roop II, et al., Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes, Gene 166 (1995) 175–176.

[32] J. Sambrook, D.W. Russell, Molecular Cloning: A Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001.
[33] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[34] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, et al., UCSF Chimera—A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.