Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system

Madina Karimova¹, Josephine Abi-Ghanem², Nicolas Berger¹, Vineeth Surendranath³, Maria Teresa Pisabarro² and Frank Buchholz¹,*

¹Medical Systems Biology, Medical Faculty and University Hospital Carl Gustav Carus, University of Technology, Fetscherstrasse 74, 01307 Dresden, ²Structural Bioinformatics, BIOTEC TU Dresden, Tatzberg 47-51 01037 Dresden and ³Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany

ABSTRACT

Targeted genome engineering has become an important research area for diverse disciplines, with site-specific recombinases (SSRs) being among the most popular genome engineering tools. Their ability to trigger excision, integration, inversion and translocation has made SSRs an invaluable tool to manipulate DNA in vitro and in vivo. However, sophisticated strategies that combine different SSR systems are ever increasing. Hence, the demand for additional precise and efficient recombinases is dictated by the increasing complexity of the genetic studies. Here, we describe a novel site-specific recombination system designated Vika/vox. Vika originates from a degenerate bacteriophage of Vibrio coralliilyticus and shares low sequence similarity to other tyrosine recombinases, but functionally carries out a similar type of reaction. We demonstrate that Vika is highly specific in catalyzing vox recombination without recombining target sites from other SSR systems. We also compare the recombination activity of Vika/vox with other SSR systems, providing a guideline for deciding on the most suitable enzyme for a particular application and demonstrate that Vika expression does not cause cytotoxicity in mammalian cells. Our results show that Vika/vox is a novel powerful and safe instrument in the ‘genetic toolbox’ that can be used alone or in combination with other SSRs in heterologous hosts.

INTRODUCTION

Manipulation of DNA is arguably one of the central technologies used to obtain a better molecular understanding of biology and medicine. The discoveries of restriction enzymes (1) and DNA ligases (2) were fundamental achievements that launched the era of recombinant DNA (3,4). Understanding how these enzymes work provided the opportunity to manipulate genetic material obtained from different species, and with their help, it became possible to express heterologous genes to produce valuable proteins and to perform functional studies based on cut and paste principles of genetic material. Today, we have access to several hundred restriction enzymes that are instrumental for DNA manipulations in vitro (4).

Although restriction enzymes and DNA ligases have become invaluable to manipulate DNA in vitro, they have limited utility to precisely manipulate DNA in a living system. Restriction enzymes typically recognize target sequences of around four to eight nucleotides. Hence, any recognition site is likely present many times in a whole genome, compromising the ability to precisely rearrange DNA segments in vivo. The discovery of different classes of enzymes was necessary to turn genetic manipulations into genome engineering. Rare-cutting enzymes, such as homing endonucleases, represent such a tool because they recognize longer sequences that are typically not present in genomes (5). Other systems that can be used to specifically cleave sequences in vivo include engineered enzymes combining a DNA-binding domain with a nuclease domain, such as zinc finger nucleases (6) and transcription activator-like effector nucleases (7).

Another class of enzymes that have become indispensable for genome engineering are site-specific recombinases (SSRs) (8,9). Most SSRs require additional host factors for efficient catalysis, limiting their use for in vivo applications in heterologous hosts. However, some SSRs recombine their targets efficiently without the aid of accessory proteins. Hence, they can be used to rearrange DNA in living systems (10). Because of the simplicity and efficiency, SSRs now serve as ‘molecular scissors’ for robust

*To whom correspondence should be addressed. Tel: +49 351 46340288; Fax: +49 351 46340289; Email: Frank.Buchholz@tu-dresden.de

© The Author(s) 2012. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial reuse, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.
non-disruptive and reproducible genomic modifications. Their discovery and characterization have pushed the frontiers of molecular biology to allow conditional gene knockouts and sophisticated manipulations of whole genomes in vivo (1,9,11,12).

The tyrosine recombinases Cre and FLP are among the most popular and widely used SSRs for genome engineering. The Cre/loxP system originates from bacteriophage P1 (2,13) and FLP from 2μ plasmid of Saccharomyces cerevisiae (3,14). Both have demonstrated their utility in animal cells, carrying the 34-bp target sites loxP and FRT, respectively. Nevertheless, differences in enzymatic properties have been described for both enzymes (4,15), with Cre being more active in mammalian systems (4,16). For some experimental setups, a lower enzymatic activity with FLP has been enhanced through directed molecular evolution (5,17) and further been improved through mammalian applications high activities are demanded. The efficacy of Cre and FLP has been enhanced through mammalian systems (4,16). Nevertheless, differences in enzymatic properties have been described for both enzymes (4,15), with Cre being more active in mammalian systems (4,16). For some experimental setups, a lower enzymatic activity with FLP has been enhanced through directed molecular evolution (5,17) and further been improved through mammalian systems (4,16). Nevertheless, differences in enzymatic properties have been described for both enzymes (4,15), with Cre being more active in mammalian systems (4,16). For some experimental setups, a lower enzymatic activity with FLP has been enhanced through directed molecular evolution (5,17) and further been improved through mammalian systems (4,16). Nevertheless, differences in enzymatic properties have been described for both enzymes (4,15), with Cre being more active in mammalian systems (4,16).

The growing list of applications is accompanied by sophistication in SSR-mediated genome engineering strategies (12,24). Combining several recombinases, and their ability to promote unidirectional recombination, has found widespread use in genome engineering. In particular, their ability to mediate locus-specific and stable genome insertions using for instance, recombinase-mediated cassette exchange, where two heterospecific target sites placed into the genome are used for precise targeted integration (8,20). In addition to tyrosine recombinases, genetic application of serine recombinases, and their ability to promote unidirectional recombination, has found widespread use in genome engineering. In particular, their ability to mediate locus-specific and stable genome insertions using for instance phiC31 (9,10,21,22) or Bxb1 (11,23) integrases has important implications in biotechnology and biomedicine.

The growing list of applications is accompanied by sophistication in SSR-mediated genome engineering strategies (12,24). Combining several recombinases, and their ability to carry out reversible rearrangements, makes it feasible to realize sequential gene switches on different alleles or multiple inducible knockouts (9). Few additional SSRs have recently been described (25–27), and first tests have documented their utility for applied uses (28). In this study, we describe a novel SSR system Vika/vox and show that it is a valuable genome engineering instrument in bacteria and in mammalian cells.

MATERIALS AND METHODS

Identification of Vika and its target site vox

The search for potential Cre-like SSRs was performed using the position-specific iterated Basic Local Alignment Search Tool (BLAST) algorithm [PSI-BLAST/National Center for Biotechnology Information (NCBI)] on the non-redundant sequences in the public DNA database of the NCBI (http://blast.ncbi.nlm.nih.gov/, date of search: 28 October 2010). Vika (ZP_05884863) was identified from Vibrio corallilyticus ATCC BAA-450 (GenBank: NZ_ACZN01000014.1). Initial recombinase alignments were done using ClustalW2 (29). The vox target site was identified using the SEARCH for LOX-like sequences (SeLOX) algorithm (30) using the following parameters: maximum 2 asymmetry positions, maximum 13 mismatches to the input loxP sequence allowed per half-site and spacer length of 8 ± 2 bp.

Structure-based modeling

A 3D model of Vika was generated by using the crystal structure of Cre [Protein Data Bank (PDB ID: 1KBU)] (2.2 Å) as template (31) and a sequence-to-structure alignment obtained by threading the Vika sequence against a fold library containing all structures available in the PDB [ProHit Professional V 2.2.2; ProCeryon Biosciences, (32)]. A seven-residue-long insertion in Vika that is not present in Cre was modeled by using the crystal structure of the lambda integrase (PDB ID: 1AE9) (1.9 Å) (33). The modeling was done with Modeler (34) embedded in Discovery Studio (Accelrys Software Inc., 2011). The model was refined with AMBER 10 (35) using the force field ff99SB (36). The model was neutralized with 13 Cl− molecules and hydrated with 1442 TIP3P water molecules (37,38) in a truncated octahedron. The water molecules and counterions were energy minimized and equilibrated at 100 K around the fixed complex for 100 ps, at constant volume and temperature in the NVT ensemble [moles (N), volume (V) and temperature (T) are conserved]. The entire system was then heated from 100 to 300 K in 10 ps by 5-K increments with harmonic positional restraints on the solute atoms (force constant of 5.0 kcal/mol/Å2). The simulation was continued in NPT ensemble [moles (N), pressure (P) and temperature (T) are conserved]. The positional restraints were gradually removed over 250 ps and followed by 1 ns of unrestrained simulations for further equilibration. Analysis of the simulation and the image rendering was done using Tachyon software in visual molecular dynamics (39).

Construction of expression plasmids

The Vika coding sequence was synthesized (GeneArt, Life Technologies Corporation) and cloned via BsrGI and XbaI New England Biolabs (NEB) into pEVO vectors as described previously (40). Two of the recombination sites, loxP, rox, VloxP or vox, were inserted into the pEVO vector, producing pEVO-loxP, pEVO-rox, pEVO-VloxP and pEVO-vox. A mutant variant of Vika in which Tyr343 was replaced by phenylalanine was constructed by site-directed mutagenesis. Primers are listed in Supplementary Table S1. For experiments in mammalian cells, the codon-optimized recombinase coding sequences (GeneArt, Life Technologies Corporation), including a nuclear localization signal (NLS), were cloned into the
expression plasmid pNPK (41). NLS-tagged recombinases were expressed from the phosphoglycerate kinase promoter. The pEVO-FLAG-Vika construct was constructed based on pEVO-Vika using primers listed in Supplementary Table S1. Western blot analysis was performed with a monoclonal mouse anti-FLAG antibody (1:2000, M2, Sigma-Aldrich) and goat anti-Mouse IgG (H+L)-HRP secondary antibody (1:5000, Bio-Rad).

Recombination reporter plasmids and recombination assays
Recombinase reporter plasmids were based on the plasmid pEVO (40) harboring the coding sequences for Cre, VCre, Dre and Vika cloned with BsrGI and XbaI (NEB) and the target sites loxP, VloxP, rox, and vox (for primer sequences see Supplementary Table S1). For comparison of the recombination activity on native target sites, three independent clones of Vika, Cre and VCre were tested on vox, loxP and VloxP, respectively. Recombinase expression was induced with L(+)-arabinose (Sigma-Aldrich) at 0, 1, 10 and 100 μg/ml.

LacZ-based reporter plasmids were constructed by inserting two of each recognition sites (loxP, VloxP and vox) into pSV-paX1 (42). Plasmid DNA from overnight cultures was isolated (Qiagen Miniprep kit) and electroporated into DH5α cultures was isolated (Qiagen Miniprep kit) and electroporated into DH5α (42). The neomycin cassette was amplified with oligo-neomycin plates. The mammalian recombinase expression constructs were based on pBabe-puro, containing an internal ribosomal entry site and GFP reporter gene (Addgene plasmid 1764). The plasmid was modified by introducing a unique XhoI restriction site upstream of the EcoRI site (pBabe-puroX). Protein coding sequences for Cre and CreR173K were amplified from LZRS-Cre and LZRS-Cre/R173K-ERT (43). ORFs of Vika and VikaY343F from pNPK-Vika and pNPK-VikaY343F, and all four were cloned via XhoI and EcoRI into pBabe-puroX. Virus was produced in Phoenix-GP cells (provided by G. Nolan, Stanford University, Stanford, CA, USA) as described previously (http://www.stanford.edu/group/nolan/protocols/pro_helper_dep.html). Cells (1 × 10⁶) were transfected with 18 μg of pBabe-puroX constructs containing various recombinases, 5 μg of p522 plasmid encoding ecotropic envelope and 10 μg of pR690, plasmid encoding gag-pol proteins, using Lipofectamine 2000 (Invitrogen). Viral supernatants were harvested 36 h after transfection. For infection, NIH3T3 cells (1 × 10⁵) were plated into six-well plates 24 h before retrovirus infection and then incubated with 1 ml retroviral supernatant in the presence of 4 μg/ml polybrene and 1 mM Hepes pH 7.25 for 4 h. At various time points, cells were trypsinized, and percentages of transduced cells (GFP-positive) were determined on a FACScan flow cytometer (Becton Dickinson) by using Cell Quest software. All retroviral infection experiments were repeated twice showing comparable results.

γ-H2AX assay
Induction of double-stranded DNA breaks was determined by the γ-H2AX assay. Seventy-two hours post-infection, transduced cells were fixed with 4% paraformaldehyde (Sigma-Aldrich) for 10 min at room temperature, and nuclei were stained with 1 μg/ml Hoechst 33342 (Invitrogen) for 30 min. For each of three biological replicates, nine images were acquired per condition on an IX81 microscope (Olympus), and cells were examined for EGFP expression. The distributions were compared using Student’s two-tailed t-test. Primers used for the construction of plasmids are listed in Supplementary Table S1.

Viral mammalian recombinase expression constructs were cloned via XhoI and EcoRI into pBabe-puroX. Protein coding sequences for Cre and CreR173K were amplified from LZRS-Cre and LZRS-Cre/R173K-ERT (43). ORFs of Vika and VikaY343F from pNPK-Vika and pNPK-VikaY343F, and all four were cloned via XhoI and EcoRI into pBabe-puroX. Virus was produced in Phoenix-GP cells (provided by G. Nolan, Stanford University, Stanford, CA, USA) as described previously (http://www.stanford.edu/group/nolan/protocols/pro_helper_dep.html). Cells (1 × 10⁶) were transfected with 18 μg of pBabe-puroX constructs containing various recombinases, 5 μg of p522 plasmid encoding ecotropic envelope and 10 μg of pR690, plasmid encoding gag-pol proteins, using Lipofectamine 2000 (Invitrogen). Viral supernatants were harvested 36 h after transfection. For infection, NIH3T3 cells (1 × 10⁵) were plated into six-well plates 24 h before retrovirus infection and then incubated with 1 ml retroviral supernatant in the presence of 4 μg/ml polybrene and 1 mM Hepes pH 7.25 for 4 h. At various time points, cells were trypsinized, and percentages of transduced cells (GFP-positive) were determined on a FACScan flow cytometer (Becton Dickinson) by using Cell Quest software. All retroviral infection experiments were repeated twice showing comparable results.

γ-H2AX assay
Induction of double-stranded DNA breaks was determined by the γ-H2AX assay. Seventy-two hours post-infection, transduced cells were fixed with 4% paraformaldehyde and stained with a primary mouse phospho-H2AX antibody (clone JBW301, 1:1000 dilution, Upstate Biotechnology) and with a donkey anti-mouse TxA red conjugated secondary antibody (1:1000 dilution, Molecular Probes). DNA was stained with 4',6'-diamidino-2-phenylindole (1 μg/ml). Control wells were treated with the DNA-damaging agent camptothecin at 5 μM for 2 h before staining. Nine images per replica for each condition were acquired on an Olympus IX81 microscope and analyzed with the ScanR Analysis software (Olympus). Staining for each transduced cell line was repeated three times. P-values were calculated by Student’s t-test.
RESULTS

Identification and 3D model of a new tyrosine recombinase

Recent sequencing efforts offer a rich resource for identifying new enzymes based on sequence homology to known proteins. To unmask putative novel Cre-like recombinases, we performed a search in the non-redundant protein database at NCBI using the PSI-BLAST algorithm. Iterative alignment principle of the PSI-BLAST allows detecting distant sequence similarities and retrieving proteins of conserved structures. Several putative Cre-like recombinases were identified with this approach (Supplementary Table S2) with the protein ZP_05884863 (hereafter referred as Vika) emerging as a particularly promising candidate (Figure 1). Cre and many other Cre-like recombinases are bacteriophage enzymes involved in phage genome resolution (44). Vika originates from the gram-negative bacterium *V. coralliilyticus* ATCC BAA-450. Analysis of the genomic region surrounding the Vika sequence identified open reading frames with homologies to bacteriophage proteins, including a gp7-like capsid protein and Panton-Valentine leukocidin-like toxin along with the putative recombinase gene (Figure 1A). Hence, these sequences probably resemble a degenerate prophage present in the bacterial genome, and therefore, Vika likely originates from a bacteriophage.

To investigate in more detail the putative Cre-like fold of Vika anticipated by its sequence similarity to Cre (26 and 49% sequence identity and similarity, respectively), and to investigate its possible substrate recognition preferences, we built a 3D atomic model of Vika. We used threading to search for the best structural template for modeling Vika; a Cre crystallographic structure (see ‘Materials and Methods’ section). The obtained Vika 3D model presented a root-mean-square deviation of 2.4 ± 0.3 Å compared with the Cre template. The sequence-to-structure alignment obtained by threading with Vika and Cre was slightly different than the sequence-based obtained by PSI-BLAST (Figure 1C). In both alignments, sequence and structure based, important residues characteristic for catalysis are conserved in Vika (Figure 1B and C). To identify residues that contribute in DNA recognition, we mapped residues involved in DNA recognition in the Cre/loxP complex (PDB ID: 1CRX) (45) onto our Vika 3D model (Figure 1B and C). Interestingly, most residues involved in recognition of the minor groove were highly conserved between Vika and Cre, whereas a number of Vika residues involved in major groove recognition could only be predicted by using the 3D model, as these residues are not conserved in sequence but in structure (Figure 1C and Supplementary Table S3). Hence, the 3D model of Vika allowed us to determine structural and functional similarities between Cre and Vika, supporting that Vika may function in a Cre-like manner. Furthermore, our 3D model helped us to identify putative residues responsible for DNA recognition and to hypothesize on their possible role in target site recognition.

Vika does not recombine target sites of known recombinase systems

Recently, a Cre-like recombinase called VCre, and its recognition target site *VloxP*, has been identified on the *Vibrio* plasmid p0908 (26). Because both Vika and VCre originate from *Vibrio* strains, we compared both recombinases. An amino acid alignment between Vika and VCre showed that, like Cre, these recombinases share moderate sequence identity (37%, Supplementary Figure S1), possibly reflecting that they recognize divergent target sequences. To test whether Vika shows activity on target sites of other known recombinases, we cloned its coding sequence into *E. coli* recombination reporter plasmids harboring two target sites for Cre (*loxP*), VCre (*VloxP*) and Dre (*rox*) (25), a Cre-like recombinase from phage D6, with 23 and 37% amino acid identity to Vika and Cre, respectively (Figure 2A). In these constructs, Vika is expressed from the L-arabinose promoter, which allows inducible expression of the enzyme by addition of L(+)-arabinose to the growth medium (Supplementary Figure S2) (46). On recombination, an 1157-bp fragment is excised from the plasmid, signifying recombination activity. Isolation of the plasmid DNA after strong induction of Vika did not show a change in plasmid size of any of the used reporters (Figure 2B), demonstrating that Vika does not act on the target sites for established recombinase systems.

SeLOX computational target site prediction identifies the target site of Vika

Because of the structural resemblance of our Vika model to Cre-like recombinases (Figure 1), we predicted that Vika likely recognizes a similarly structured sequence. Cre and related tyrosine recombinases recognize a palindromic DNA sequence of two inverted 13-bp repeats that are typically separated by an 8-bp spacer (19). We anticipated that the genome sequence of *V. coralliilyticus* might contain the Vika target site. To examine the host genome to rationally predict putative targets, we used the SeLOX algorithm. SeLOX allows for a streamlined analysis of large DNA inputs for the presence of degenerate inverted repeats and is therefore ideally suited to identify putative target sites for recombinases (30). The analysis of the *V. coralliilyticus* ATCC BAA-450 genome predicted six palindromic sequences that matched the input criteria in a 10-kb area flanking the Vika coding sequence (see ‘Materials and Methods’ section, Supplementary Table S4). Only one of these sites was found in the vicinity of the recombinase and was called *vox* (Figure 3). The *vox* sequence is a 13-bp inverted repeat with one asymmetry separated by an 8-bp spacer and shows about 50% sequence homology to both *loxP* and *VloxP* and 33% sequence homology to *rox* (Figure 3B). Based on these analyses, we decided to test whether *vox* might be the target site of Vika.

Vika specifically recombines *vox* sites

The 34-bp *vox* sequence was cloned into the recombination reporter vectors harboring coding sequences for
Figure 1. Sequence and structural model of Vika recombinase. (A) Localization of Vika (red) and its target site vox site (red triangle) in the genome of *V. coralliilyticus* ATCC BAA-450. Putative prophage genes flanking the Vika coding sequence are depicted by gray arrows. (B) Structural model of Vika and comparison with the X-ray structure of Cre (PDB ID: 1KBU). Residues known to interact with DNA in the Cre/LoxP system and their putative counterparts in the Vika/vox system are highlighted (catalytic residues in red, residues with major or minor groove contacts in yellow and residues contacting the phosphate backbone in green). (C) Sequence alignment of Vika and Cre. Residues are highlighted as mentioned earlier. A secondary structure prediction alignment using Jpred (jpred) and a structure-based alignment using threading (1KBU) are depicted. α-Helixes are in yellow and red β-sheets in cyan.
Cre, Dre, VCre and Vika recombinases (Figure 2A). The reporter plasmids were transferred to E. coli, and recombinase expression was induced with L(+)-arabinose at different concentrations. None of the Cre, VCre and Dre recombinases exhibited recombination activity on vox, indicating that they do not catalyze recombination of vox sites (Figure 4A). In contrast, Vika efficiently and specifically recombined vox sites on expression of the recombinase (Figure 4A). Hence, the vox sequence is the substrate for the Vika recombinase. To confirm high specificity of the novel recombinase system, we investigated Vika, Cre and VCre in two additional assays. First, we used a LacZ reporter assay to visualize recombination activity (Figure 4C). The advantage of the LacZ-based assay lies in the possibility to detect weak/promiscuous recombination activity in order to precisely assess the cross-recombination of different recombinases and their target sites. After extended expression of the recombinases, the plasmids were harvested, and E. coli cells were retransformed and plated on X-gal plates. Each colony was evaluated by color, providing a highly sensitive assay. Non-recombined reporter plasmids express LacZ and thus form blue colonies when plated on X-gal plates, whereas cells harboring recombined plasmids appear white on the same plates. When Vika was expressed at high levels in cells withloxP or VloxP reporters, all colonies were blue, demonstrating that these sequences cannot be recombined by Vika. In contrast, almost all colonies appeared white, signifying efficient

![Figure 2](image_url)

Figure 2. Evaluation of Vika recombination activity on heterologous target sites. (A) Scheme of the recombination assay. Site-specific recombination excises a 1.2-kb DNA fragment, flanked by the recombinase target sites (triangles) from the plasmid. Protein coding genes are depicted by arrows and the origin of replication (ori) is shown as a box. CmR, chloramphenicol resistance gene; AraC, gene coding arabinose promoter regulator protein and recombinase, recombinase coding gene. (B) Lack of recombination activity of Vika on theloxP,rox and VloxP target sites. Site-specific recombination of native targets by CreloxP, Drerox and VCre (VloxP) served as positive control. The − and + indicate presence or absence of L(+)-arabinose (100 µg/ml) in the growth medium. Non-recombined and recombined plasmids are denoted as two triangles and one triangle, respectively. M, marker; 2-log DNA ladder, NEB.

![Figure 3](image_url)

Figure 3. Identification and comparison of therox target site. (A) Workflow of the SeLOX algorithm and important input parameters to predictlox-like sites in the genome of V. coralliilyticus strain NZ_ACZN01000014.1. The SeLOX results are shown in Supplementary Table S4. (B) Nucleotide alignment of vox with recombination sites for CreloxP, Drerox and VCre (VloxP). Nucleotides that break perfect palindromic symmetry of thelox sites are colored in red. Asterisks depict conservation among all sites.
Figure 4. Evaluation of Vika recombination activity and specificity in E. coli. (A) Recombination specificity of indicated recombinases on vox. The − and + indicates absence or presence of L(+)arabinose (100 μg/ml) in the growth medium. Non-recombined and recombined plasmids are denoted as two triangles and one triangle, respectively. M, marker; 2-log DNA ladder, NEB. (B) Matrix of the recombination specificities of indicated recombinase systems. (C) Schematic view of LacZ-based reporter assay. The non-recombined reporter plasmid constitutively expresses β-galactosidase, inferring blue color to cells on X-gal plates. On recombination, the promoter driving LacZ expression is removed, resulting in white colonies on X-gal plates. Triangles depict lox sites. Prom = promoter. (D) Recombination activity of indicated recombinases on the loxP, voxP and vox LacZ-based reporters. Note that white colonies, signifying recombination, only appear when a recombinase is expressed together with its corresponding reporter. Positive controls for the non-recombined (pSVpaX), and recombined form of the reporter (pSVpaXΔ), is shown as Mock. (E) Scheme of the integration assay. The host plasmids (pH) carry the coding sequence for the used recombinases, the chloramphenicol resistance (continued)
recombination, when Vika was used in conjunction with the vox reporter (Figure 4D). Corresponding results were observed for Cre and VCre on their native targets loxP and VloxP, respectively (Figure 4D). Therefore, these results confirm that Vika, Cre and VCre recognize their targets in a highly specific manner and do not cross-react with non-native sites.

Next, we tested properties of the different enzymes for recombinase-mediated integration. For these experiments, we used a two-plasmid assay that reports integrative recombination [Figure 4E (41)]. Here, the integration property of the recombinase is tested by its ability to use two target sites located on different plasmids to merge them into one plasmid through the site-specific recombination reaction. The assay is designed with vectors based on two types of origin of replication: on p15A for the host pH plasmid (chloramphenicol resistant), and on R6K for the donor pD-plasmid (kanamycin resistant) (Figure 4E). Replication of the R6K-plasmid is pi-protein dependent, and therefore, this plasmid cannot replicate in a pir-negative bacterial strains. Hence, kanamycin resistance conveyed by the donor plasmid depends on its integration into a host plasmid through a lox site to form the co-integrant and consequent replication of the kanamycin resistance gene in pir-negative cells. When plated on medium with chloramphenicol and kanamycin, only the combinations of the recombinase and its native sites produced colonies. Vika promoted co-integration of two plasmids only if vox sites were present on both vectors, whereas abolishing integration if one of the sites was heterologous (Figure 4F). Equivalent results were obtained for Cre, Dre and VCre, demonstrating that donor and recipient plasmids have to carry the corresponding target sites to achieve co-integration. These data suggest that Vika, like Cre, Dre and VCre, is a recombinase that can use the native target site (vox) for intermolecular integration. Moreover, Vika only recognizes vox target sites, whereas Cre, Dre and VCre show no activity on vox. In conclusion, we find that Vika/vox is a novel SSR system with no cross-reactivity with other known recombinase systems.

**Comparative analysis of different recombinase systems in E. coli**

In the experiments for excision and integration, the enzymes demonstrated high target site specificity. However, we also noticed that the tested recombinases displayed different enzymatic activities. To obtain a more quantitative comparison of the activities of Cre, VCre and Vika in vivo, we used various induction levels for the recombinase expression to dissect the peculiarities of the recombination catalysis. For this purpose, we used the established recombinase reporter assay (Figure 2A) with Cre, VCre and Vika recombinases cloned into the vectors containing loxP, VloxP and vox, respectively. The plasmids were then propagated in E. coli at different L(+) arabinose concentrations from 0 to 100 µg/ml of L(+) arabinose. All recombinases fully recombined their corresponding target sites at 100 µg/ml of L(+)-arabinose, demonstrating that they efficiently recombine their targets when expressed at high levels (Figure 5). In contrast, differences in recombination efficiencies were evident at lower expression levels. Almost full recombination of the targets was observed for Cre and Vika already at 10 µg/ml, whereas for VCre, only ~50% of its substrate was recombined at this concentration (Figure 5). Furthermore, no recombination was observed for VCre at 0 and 1 ug/ml, whereas Cre and Vika recombined their targets to ~50% at these concentrations (Figure 5). The apparent recombination of Cre and Vika even without the addition of L(+)-arabinose indicates the high activity of these enzymes because the ‘leakiness’ of the arabinose promoter is relatively low (46). Consistent with the excision assay, the integration assay revealed a similar activity difference. Although hundreds of colonies appeared for Vika and Cre in this assay, only few colonies were counted when VCre was co-transfected with its donor vector pD-VloxP (Figure 4F and G). Overall, these experiments identify a qualitative difference in in vivo recombination activity of the studied recombinases, with Vika and Cre being more active than VCre.

**Figure 5.** Evaluation of the recombination activity of indicated recombinase systems in E. coli. One and two triangles mark the size of the recombinated and non-recombinated plasmids, respectively. M = marker, 2-log DNA ladder, NEB, no ind = no addition of L(+) arabinose to the growth medium, 1, 10 and 100 = growth of the cells in the presence of 1 µg/ml, 10 µg/ml and 100 µg/ml L(+) arabinose, respectively.
Vika efficiently and specifically recombines vox in mammalian cells

For the applied use in genome engineering, the recombinase needs to work in heterologous organisms. To investigate whether Vika can be used in organisms other than the natural host and *E. coli*, we tested the recombinase in human cells. HeLa cells were co-transfected with GFP reporter plasmids and recombinase expression plasmids (Figure 6A). Twenty-four hours post-transfection, the cells were monitored for GFP expression. Many GFP-positive cells were observed, when the Vika expression plasmid was co-transfected with the vox reporter (Figure 6B), signifying that Vika efficiently functions in mammalian cells. In contrast, no green cells were observed with the reporters for *loxP* and *VloxP* (Figure 6B), demonstrating Vika’s specificity also in mammalian cells. Corresponding results were seen with Cre and VCre, where these enzymes showed activity only on their target sites without recombining the target sites of the other recombinases (Figure 6B). Hence, these enzymes can be used in combination without cross-activities. To compare the activities of the different recombinases also in mammalian cells, we quantified the number of GFP-positive cells from the co-transfection experiments. Similar to the activity differences measured in *E. coli*, we also observed a difference in activity in HeLa cells. In this host, Cre showed the highest activity followed by Vika that exhibited about half of the activity of Cre. As in *E. coli*, the number of GFP-positive cells for VCre was markedly lower (~10%), indicating that this recombinase also possesses less activity in mammalian cells (Figure 6C).

Vika expression does not cause cytopathic effects in human and mouse cells

The successful use of recombinases in heterologous hosts rests on their restricted activity on the introduced target sites. Important concerns are cryptic (pseudo) recombination sites in the genome, which could potentially have cytopathic consequences. Active pseudo-*loxP* sites have indeed been described in human and mouse (47). As a consequence, expression of Cre in human and mouse cells can cause cytotoxicity (43,48–51).

To investigate potential site effects of Vika, we searched for pseudo-vox sites in the mouse and human genome using the SeLOX algorithm (30). In both species, lower overall numbers of pseudo-vox sites were uncovered than pseudo-loxP sites (Supplementary Figure S3). We then tested pseudo-vox sites that most closely resembled vox sites (Supplementary Figure S4) experimentally. In contrast to Cre, which showed prominent activity on pseudo-loxP sites, Vika did not display measurable activity on pseudo-vox sites.
Figure 7. Evaluation of genotoxicity and cytotoxicity of Vika. (A) Recombination specificity of Vika and Cre on the respective cryptic human and mouse chromosomal target sites. The - and + indicates absence or presence of L(+)-arabinose (100 μg/ml) in the growth medium. Non-recombined and recombined plasmids are denoted as two triangles and one triangle, respectively. M, marker; 2-log DNA ladder, NEB. (B) Evaluation of DNA damage induction by γ-H2AX assay 72h after infection with indicated recombinase-expressing virus, n = 3. (C) Quantification of the γ-H2AX-positive cells either infected with virus for respective recombinase expression or treated with campthothecin for 2h. Statistically significant increase of γ-H2AX signals is indicated by asterisks. Error bars indicate standard deviation of the mean value (** indicate P = 0.01). (D) Proliferation effects on expression of indicated recombinases in mouse NIH3T3 cells. Cells were infected with bicistronic retroviruses expressing respective recombinase linked to GFP. Every 72h, cells were analyzed by flow cytometry. Error bars indicate standard deviation of the mean value, n = 2.
activity on pseudo-vox sites in these assays (Figure 7A). To examine the influence of prolonged expression of Vika in primary cells, we generated mouse embryonic stem (ES) cells stably expressing Vika. No effect on cell growth and morphology was observed, indicating that Vika is active when selected for stable expression in mouse ES cells and is well tolerated (Supplementary Figure S5). Next, we investigated the potential impact of recombinase expression on DNA damage. NIH3T3 cells were infected with GFP-bicistronic retroviral particles encoding Cre, Vika or controls. Three days post-infection, γ-H2AX signals were investigated in fixed cells. Cre expression caused a marked increase in γ-H2AX signals signifying induction of DNA damage. In contrast, Vika had no influence on the amount of γ-H2AX counted (Figure 7B and C), indicating that Vika expression does not lead to increased DNA damage in these cells. Finally, we monitored the percentage of GFP-positive cells over time in the populations as an indicator for cytotoxicity. As published previously (43), Cre expression led to a rapid decline of GFP-positive cells in the population (Figure 7D and Supplementary Figure S6), whereas no decline in GFP-positive cells was observed in Vika-expressing cells. These experiments indicate that high levels of Cre are cytotoxic, whereas Vika expression is well tolerated in the same setting.

We conclude that Vika is an active and specific SSR that should be useful for genome engineering purposes in mammalian cells.

**DISCUSSION**

In this study, we describe the discovery of Vika/vox, a new tyrosine family site-specific recombination system. Vika was identified based on sequence alignments and protein fold predictions from sequences and structures of known SSRs. Based on known structures of SSRs, we generated a structural model, which predicted that Vika has Cre-like structural properties. We then implemented a rational approach for the search of Vika’s target site and identified its native recombination site, vox. We demonstrate that Vika/vox does not cross-recombine with either the Cre, VCre or Dre recombinase systems in bacteria or in human cultured cells. Therefore, these SSR systems can be combined for advanced genome engineering approaches.

To optimally use the different SSR systems, it is essential to characterize their applied properties. The determination of the in vivo recombinase reactivity and maximum reachable recombination rate may be crucial for planning an experiment. Sophisticated genetic schemes that include simultaneous modifications (e.g. double knockouts) will require recombinases of comparable activity, whereas for sequential DNA rearrangements (e.g. recombinase-mediated cassette exchange), it may be more important to design the experiment in a manner that intermediate recombination events could be selected for. We compared recombination efficiencies of the new recombinase Vika with Cre and another recently reported Cre-like recombinase VCre. Although VCre was described to display efficient recombination rates (26), our results indicate that VCre recombines its target VloxP at considerably lower rate when directly compared with Cre/loxP and Vika/vox. In our work, we used assays that detect various recombination stages based on induction levels. We observed a consistent trend for differences between the tested recombinases. Cre and Vika were consistently more active than VCre. The comparison of Vika with Cre and VCre should be useful for considerations when planning experiments with several independent recombinases. For instance, in development of a double conditional strategy, described recombinases can be combined in efficiency correlated ways, e.g. Cre + FLPs and Vika + VCre, where Cre and Vika are dedicated to conditional knockouts, whereas FLPs and VCre are reserved for excision of selection marker cassettes. Alternatively, molecular evolution strategies could be used to enhance the activity of recombinases, as has been shown for the FLP recombinase (17), if a multitude of highly efficient recombinases is required.

To integrate Vika recombinase for the applied use in higher organisms, we have addressed the question of cytotoxicity. Vika, as other SSRs, acts on DNA, and therefore, undesired binding to non-target sequences may cause undesired events. With the full genome sequences available for many organisms, it is possible to analyze beforehand whether there are potential binding sites for DNA-binding proteins. We have analyzed and tested pseudo-vox sites for human and mouse in this study and did not observe recombination on these sites. Similar analyses can be carried out for other sequenced species to evaluate the suitability of Vika in a particular organism. We have also compared high-level expression of Vika and Cre in mouse cells via retroviral gene delivery and observed that cells expressing Vika recombinase, unlike Cre, did not exhibit growth inhibition. In addition, in contrast to Cre, Vika did not cause DNA damage when expressed from the retroviral constructs. Therefore, Vika might become an additional tool or even a valuable substitute for experimental settings where Cre is not applicable.

To expand the utility of SSRs, enzymes with novel properties have been developed using directed molecular evolution. Several studies have reported alteration of substrate specificity for both Cre and FLP (52–56). Molecular evolution approaches provide new perspectives for potentially therapeutic applications, as these approaches offer design of customized recombinases. Indeed, eradication of a Human immunodeficiency virus provirus from infected human cells has already been demonstrated with the evolved Cre-like recombinase Tre, which recombines a defined sequence within the Human immunodeficiency virus-1 long terminal repeat (54). However, the development of such recombinases remains tedious and time consuming. The discovery of novel SSR systems should be helpful in two ways to accelerate the development of therapeutically relevant enzymes. First, the detailed investigation of DNA binding should lead to a better understanding of how these enzymes specifically recognize their targets. With this knowledge, rational design of enzymes with novel specificities might become possible. Second, different naturally occurring recombinases could be useful to accelerate the directed evolution process. It has been demonstrated that shuffling of related genes can
speed up the evolution process (57). Hence, new recombinases with desired properties might become available in shorter time through family shuffling.

Up to date, several enzymatic systems have been successfully developed for genome engineering purposes (58, 59). However, the rapid growth and scale of genetic studies impose the need for a broad spectrum of genetic instruments. Nature is a rich source for discovering DNA-modifying enzymes with desired properties. Large-scale sequencing projects, in particular of environmental samples, provide a rich source for such discoveries (60).

With the identification of the Vika/vox system and the proposed recombinase systems in Supplementary Table S2, we demonstrate the utility of these resources if appropriate bioinformatics tools are available. We predict that many more useful genome engineering tools, including additional SSRs with unique specificities, will be discovered using similar approaches in the future.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–4 and Supplementary Figures 1–6.

ACKNOWLEDGEMENTS

The authors thank Francis Stewart and Jun Fu, BioTec, Dresden, Germany, for providing the R6K-neo vector and Anton Berns and Jos Jonkers, Netherlands Cancer Institute, Amsterdam, The Netherlands, for sharing retrovirus plasmids (LZRS-Cre, LZRS-Cre173 and LZRS-GFP). They also thank Ralf Gey for technical virus plasmids (LZRS-Cre, LZRS-Cre173 and LZRS-GFP). They also thank Ralf Gey for technical support and the ZIH at the Technische Universität Dresden for computational resources and assistance and the members of the Buchholz and Pisabarro groups for helpful discussions.

FUNDING

J Klaus Tschira Stiftung gGmbH (to A.-G.); European Union’s Seventh Framework Programme for Research (EU FP7) SyBoSS [242129]; Deutsche Forschungsgemeinschaft (German Research Foundation) [SPP1356 BU1400/3-1, SFB655]. Funding for open access charge: Medical Systems Biology, University Hospital and Medical Faculty, Carl Gustav Carus, University of Technology, Dresden, Germany.

Conflict of interest statement. None declared.

REFERENCES