Assembly of Designer TAL Effectors by Golden Gate Cloning

Ernst Weber, Ramona Gruetzner, Stefan Werner, Carola Engler, Sylvestre Marillonnet*

Icon Genetics GmbH, Halle/Saale, Germany

Abstract

Generation of customized DNA binding domains targeting unique sequences in complex genomes is crucial for many biotechnological applications. The recently described DNA binding domain of the transcription activator-like effectors (TALEs) from Xanthomonas consists of a series of repeats arranged in tandem, each repeat binding a nucleotide of the target sequence. We present here a strategy for engineering of TALE proteins with novel DNA binding specificities based on the 17.5 repeat-containing AvrBs3 TALE as a scaffold. For each of the 17 full repeats, four module types were generated, each with a distinct base preference. Using this set of 68 repeat modules, recognition domains for any 17 nucleotide DNA target sequence of choice can be constructed by assembling selected modules in a defined linear order. Assembly is performed in two successive one-pot cloning steps using the Golden Gate cloning method that allows seamless fusion of multiple DNA fragments. Applying this strategy, we assembled designer TALEs with new target specificities and tested their function in vivo.

Citation: Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011) Assembly of Designer TAL Effectors by Golden Gate Cloning. PLoS ONE 6(5): e19722. doi:10.1371/journal.pone.0019722

Editor: Mohammed Bendahmane, Ecole Normale Superieure, France

Received February 4, 2011; Accepted April 8, 2011; Published May 19, 2011

Copyright: © 2011 Weber et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by Icon Gentics GmbH. As all authors are employed by the funder, the funder played a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: All authors are employed by Icon Gentics GmbH. This does not alter the authors’ adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: marillonnet@icongenetics.de

Introduction

The development of synthetic nucleases that cleave unique genomic sequences in living cells provides powerful tools for genome engineering, allowing targeted gene knockout and gene replacement [1]. A key component of these artificial nucleases is the DNA binding domain which directs the nuclease to its target sequence. To date, the majority of customized DNA targeting domains used for genome engineering that have been made are based on engineered zinc-finger domains. However, the creation of new DNA binding specificities has proven to be technically challenging and time consuming. An alternative to zinc-finger domains may be the recently described DNA binding domain found in transcription activator-like effectors (TALEs) [2,3]. TALEs are virulence factors of plant pathogens from the genus Xanthomonas that are translocated via a type III secretion system inside the plant cell. The TALEs are then imported into the nucleus, where they bind to specific DNA sequences and transcriptionally activate gene expression [4,5]. DNA binding is mediated by a central repetitive region, formed by up to 33 tandem repeats of a 33 to 35 amino acid motif, each repeat corresponding to one DNA base pair of the target sequence. The amino acid sequences of the repeats are nearly identical, beside amino acid positions 12 and 13, the so-called repeat variable diresidues (RVD). Repeats with different RVDs show different DNA base pair preferences, and consecutive RVDs in a TALE correspond directly to the DNA sequence in the binding side, resulting in a simple one-repeat-to-one-base pair code [6,7]. Knowledge of this TALE recognition code has been used to predict the DNA binding specificity of native TALEs and to create designer TALEs (dTALEs) which transcriptionally activate user-defined promoter sequences [8,9]. Furthermore, several groups have combined dTALE DNA binding domains with the FokI derived DNA-cleavage domain, resulting in potent tools for genome engineering [10,11,12,13]. However, assembly of multiple repeats with highly identical sequences by standard cloning approaches is challenging and chemical synthesis of the entire repeat region expensive.

We present here an approach to assemble genes encoding TALE repeat domains based on the scaffold of AvrBs3, the first described and well characterized TALE family member [14]. For each of the 17 full repeats found in AvrBs3, four module types were generated, each with preference to one of the four DNA base pairs. With this set of 68 repeat modules, DNA recognition domains for any 17 nucleotide target sequence of choice can be assembled in two cloning steps. Both cloning steps use the Golden Gate cloning method that allows directional and seamless assembly of multiple DNA fragments [15,16]. As a proof of principle, we created three dTALE proteins designed to target the promoter of a reporter construct stably integrated in the Nicotiana benthamiana genome, and show that all three dTALEs are able to activate the reporter construct.

Results

dTALE assembly strategy

The dTALE assembly strategy described here uses the Golden Gate cloning method, which is based on the ability of type IIIS enzymes to cleave outside of their recognition site. When type IIIS recognition sites are placed to the far 5’ and 3’ end of any DNA
Assembly of Designed TALE Effectors

To test functionality of the assembled dTALEs, we used transgenic N. benthamiana plants containing a stably integrated GFP reporter construct (Fig. 2A). This construct consists of a tobacco mosaic virus-based viral vector under control of the alcA promoter from Aspergillus nidulans [18,19]. Since the alcR transcriptional activator that is required for activation of the alcA promoter is not present in the transgenic plants, the alcA promoter can be considered here as a minimal promoter. Three sequences were chosen from the promoter, all starting with a thymidine as defined by the specificity of repeat 0. The target sequence chosen for dTALE-1 is overlapping with the alcR binding site in the alcA promoter (bp -143 to -127), while the target sequences for dTALE-2 and dTALE-3 consist of bp -61 to -45 and bp -69 to -53 respectively (target site positions numbered relative to the viral vector transcription start, Fig. 2A). dTALE-4 was constructed as a negative control and targets a randomly selected sequence not found in the promoter region.

For construction of the 4 dTALE constructs, 12 parallel BsaI-based Golden Gate cloning reactions were set up with selected modules and the respective preassembly vectors pL1-TA1 to 3. For each reaction, plasmid DNA from two colonies was purified and sequenced, and all plasmids were found to contain the correct sequence. Preassembled repeat blocks were assembled to the final constructs dTALE-1 to 4 using a second BpiI-based Golden Gate cloning reaction (Fig. 2B). Eleven out of 12 colonies analyzed contained a correct construct. After sequence verification, the constructs dTALE-1 to 4 were transformed into A. tumefaciens constructs dTALE-1 to 4 were transformed into A. tumefaciens and inoculated into leaves of transgenic N. benthamiana plants containing the GFP reporter construct. All three dTALEs with DNA binding domains designed to target sequences in the alcA promoter induced GFP expression in infiltrated leaf areas, with expression from dTALE-1 being the weakest. In contrast, dTALE-4 did not induce any GFP expression from the reporter construct (Fig. 2C).

Discussion

We have shown here that constructs for dTALE proteins containing a 19 base DNA binding domain (consisting of 17 engineered full repeats, repeat 0 and the half repeat 17.5) can be easily assembled by two successive one-pot Golden Gate cloning reactions. We have prepared a set of 68 repeat modules that allows construction of DNA binding domains for any 17 base user-defined target sequence. The native half repeat 17.5 of AvrBs3, which contains a RVD specific for thymidine, was included in the C-terminal fragment of the final assembly vector. It would however be possible to also make half repeat modules with different RVD types to improve the binding of dTALE proteins for target sequences that do not have a T at this position. Such repeats could be assembled together with repeats 13 to 17 in a new preassembly vector replacing pL1-TA3. A new compatible final assembly vector lacking the half repeat should also be made. In case 17 repeats are not sufficient to provide specific binding, dTALE proteins with additional repeats could easily be constructed. In order to expand the TALE modular cloning system to more than 17 repeats, new unique fusion sites have to be defined for each additional repeat, and one or more new preassembly vectors specific for the added fusion sites have to be constructed. A further option to increase dTALE specificity is the replacement of the NN RVD, which has an equal preference to A and G, by the highly G-specific NK RVD [9,12].

The Golden Gate cloning method provides a perfect fit for dTALE protein engineering because it allows directional and
seamless assembly of multiple DNA fragments. In addition, this cloning method is sequence-independent and allows assembly of repeats with identical or highly homologous sequences, since only the 4 base pair fusion sites at the end of the repeats have to be unique. Selection of fusion sites with unique sequence at the ends of successive repeats can be easily accomplished by either changing the codon usage of the ends of the repeats, or by shifting the fusion sites a few nucleotides at the ends of the various repeats. Since a complementary shift can be selected at the beginning of each following repeat (as shown in the result section/supporting information), seamless assembly of direct repeats can then be easily achieved.

Other alternative methods for seamless assembly of multiple DNA fragments include SLIC [20], SOEing [21] and ssDNA oligonucleotide assembly [22]. These methods are however limited by the homology present among the repeats since they either involve PCR steps [21] or require annealing of single-stranded DNA fragments [20,22], both of which run the risk of deleting some of the repeats by recombination during amplification and/or cloning. Codon optimization may nevertheless be used to minimize the risk of loss of repeats during cloning. A recently published protocol combines the use of type IIS enzymes and PCR amplification of codon-optimized repeats, and was shown to allow assembly of dTALEs containing 12.5 repeats [23]. This protocol is however...
more laborious, as it requires two rounds of PCR amplification and several purification steps and, since PCR is involved, some of the final constructs may be expected to contain mutations derived either from polymerase amplification or from the primers.

In conclusion, the cloning system described here provides a simple and economical way of assembling constructs encoding dTALE proteins for genome engineering and other biotechnological applications.

Methods

Molecular biology reagents

Restriction enzymes used in this study were purchased from New England Biolabs (Ipswich, MA) and Fermentas (Burlington, Canada). T4 DNA ligase was purchased from Promega (Fitchburg, WI). Plasmid DNA preparations were made by using the NucleoSpin Plasmid Quick Pure kit (Macherey-Nagel, Düren, Germany) following the manufacturer protocol. Plasmid DNA concentration was measured using a Nano Drop® Spectrophotometer ND-2000 (Peglab, Erlangen, Germany). DNA sequences for the AvrBs3 N- and C-termine were codon-optimized using the Nicotiana tabacum codon usage (GENEius software from MWG Eurofins, Ebersberg, Germany) and were synthesized by this company. Both synthesized fragments do not contain any BpiI or BsaI restriction sites. Sequences of the codon-optimized avrBs3 gene and of the 68 repeat modules, as well as primer sequences necessary for construction of the destination plasmids are listed in Supporting Information S1.

Vector construction

The repeat modules were made by annealing two partially overlapping primers and filling the single-stranded extensions using KOD polymerase (Merck, Darmstadt, Germany). The double-stranded products were digested with XhoI and cloned in the SalI site of a pUC19-derived vector conferring spectinomycin resistance and lacking BpiI and BsaI restriction sites. For construction of the preassembly vectors pL1-TA1-3, a lacZ reporter gene and of the 68 repeat modules, as well as primer sequences required for synthesis of the TALE repeat construction. (C) Primer sequences necessary for construction of the destination plasmids are listed in Supporting Information S1.

Supporting Information

Supporting Information S1

Sequence of the codon-optimized avrBs3 gene and of the primers required for synthesis of the TALE repeats and of the preassembly vectors. (A) Sequence of the codon-optimized avrBs3 gene. The sequences selected as fusion sites for assembly of dTALEs are shown in bold and underlined. (B) Primer sequences required for TALE repeat construction. (C) Primer sequences for construction of preassembly vectors pL1-TA1-3.

Author Contributions

Conceived and designed the experiments: EW SM. Performed the experiments: EW RG SW CE. Wrote the paper: EW SM.

References

1. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11: 636–646.
2. Kay S, Hahn S, Marcus E, Haase G, Bonas U (2007) A bacterial effector acts as a phytohormone receptor and induces a cell death response in Arabidopsis. Science 318: 648–651.
3. Romer P, Hahn S, Jordan T, Strauss T, Boch J, et al. (2007) Plant pathogen recognition mediated by promoter activation of the pepper B3 resistance gene. Science 318: 645–648.
4. Boch J, Hahn S, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Plant Physiol 48: 419–436.
5. Bogdanove AJ, Schornack S, Lahaye T (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326: 1509–1512.
6. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326: 1501.
7. Romer P, Recht S, Strauss T, Elsasser J, Schornack S, et al. (2009) Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, Xanthomonas oryzae pv. oryzae. New Phytol 180: 1048–1057.
8. Moehlman R, Romer P, Heck J, Lahaye T (2010) Optimization of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci U S A 107: 21617–21622.
9. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, et al. (2010) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res.
10. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, et al. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186: 757–761.
11. Miller JC, Tian S, Qiao Z, Barlow KA, Wang J, et al. (2010) A TAL nuclease architecture for efficient genome editing. Nat Biotechnol.
12. Mahfouz MM, Li L, Shaminmazazian M, Wibowo A, Fang X, et al. (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nucleases with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A 108: 2623–2628.
13. Bonas U, Stall RE, Staskawicz B (1989) Genetic and structural characterization of the avirulence gene avrBs3 from Xanthomonas campestris pv. vesicatoria. Mol Gen Genet 218: 127–136.
14. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3: e3647.
15. Engler C, Kretzmer R, Kandzia R, Marillonnet S (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type II restriction enzymes. PLoS One 4: e5553.
16. Lebedenko EN, Birikh KR, Putalov OV, Berlin Yu A (1991) Method of artificial DNA splicing by directed ligation (SDL). Nucleic Acids Res 19: 6757–6761.
17. Felenbok P (1991) The ethanol utilization regulon of Aspergillus nidulans: the alcA-alcR system as a tool for the expression of recombinant proteins. J Bacteriol 171: 11–17.
18. Marillonnet S, Thoringer G, Kandzia R, Klimyk V, Gheva Y (2005) Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. Nat Biotechnol 23: 718–723.
19. Li MZ, Elledge SJ (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 4: 251–256.
20. Horton RM, Cai ZL, Ho SN, Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. Biotechniques 6: 328–385.
21. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods 6: 343–345.
22. Zhang F, Cong L, Ladozo S, Kousiri S, Church GM, et al. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol.
23. Weber E, Engler C, Gruetzmacher R, Werner S, Marillonnet S (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. PLOS One.