Review

TALE: A tale of genome editing

Mingjie Zhang a,c,1, Feng Wang b,1, Shifei Li a, Yan Wang a, Yun Bai a,∗, Xueqing Xu a,∗

a Department of Medical Genetics, College of Basic Medical Science, Third Military Medical University, Chongqing 400038, People’s Republic of China
b Department of Clinical Laboratory, Institute of Surgery Research, Daping Hospital, Third Military Medical University, Chongqing 400042, People’s Republic of China
∗ Battalion 16 of Cadet Brigade, Third Military Medical University, Chongqing 400038, People’s Republic of China

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Transcription activator-like effectors
TALE transcription factors
TALE nucleases
Genome editing
Gene targeting

ABSTRACT

Transcription activator-like effectors (TALEs), first identified in Xanthomonas bacteria, are naturally occurring or artificially designed proteins that modulate gene transcription. These proteins recognize and bind DNA sequences based on a variable numbers of tandem repeats. Each repeat is comprised of a set of ~34 conserved amino acids; within this conserved domain, there are usually two amino acids that distinguish one TALE from another. Interestingly, TALEs have revealed a simple cipher for the one-to-one recognition of proteins for DNA bases. Synthetic TALEs have been used to successfully target genes in a variety of species, including humans. Depending on the type of functional domain that is fused to the TALE of interest, these proteins can have diverse biological effects. For example, after binding DNA, TALEs fused to transcriptional activation domains can function as robust transcription factors (TALE-TFs), while fused to restriction endonucleases (TALENs) can cut DNA. Targeted genome editing, in theory, is capable of modifying any endogenous gene sequence of interest; this can be performed in cells or organisms, and may be applied to clinical gene-based therapies in the future. With current technologies, highly accurate, specific, and reliable gene editing cannot be achieved. Thus, recognition and binding mechanisms governing TALE biology are currently hot research areas. In this review, we summarize the major advances in TALE technology over the past several years with a focus on the interaction between TALEs and DNA, TALE design and construction, potential applications for this technology, and unique characteristics that make TALEs superior to zinc finger endonucleases.

© 2013 Published by Elsevier Ltd.

Contents

1. Introduction ......................................................................................................................... 00
2. The discovery of TALEs ....................................................................................................... 00
3. Features of TALEs: recognition and specificity ................................................................. 00
4. Structure of TALE-DNA ...................................................................................................... 00
5. Magic of TALEs: regulation of gene transcription and genome editing ................................. 00
6. The making of TALEs: designing and engineering ............................................................ 00
7. Practical applications .......................................................................................................... 00
8. TALEN vs. ZFs .................................................................................................................... 00
9. Summary ............................................................................................................................. 00
   Conflict of interest ................................................................................................................ 00
   Acknowledgements ............................................................................................................... 00
   References ............................................................................................................................ 00

1. Introduction

Genome editing can be used to modify any gene sequence of interest in either cells or whole organisms. Currently, this technique is useful for functional genomic studies. In the future, it may
also be useful in the clinic for the treatment of heritable diseases. Ideally, genome editing technology should be specific for the desired target, efficient, and affordable. Target specificity is typically achieved by gene targeting, in which a specific genomic sequence is recognized, bound, and modified via homologous recombination. While this methodology is popular in the laboratory, its lack of efficiency and its high cost have limited its widespread use. Zinc finger nucleases (ZFNs) are engineered proteins comprised of two distinct domains: a zinc finger domain to recognize and bind to DNA and the endonuclease FokI, which cuts DNA at a specific sequence. ZFN dramatically increase the efficiency of genome editing. Interestingly, drugs based on this methodology are currently being tested in phase 2 clinical trials for the treatment of HIV. Despite its promise, ZFN technology is limited by design complexity and cost; moreover, several companies hold patents for this type of technology.

Over the past several years, a second-generation artificial transcription activator-like effector nuclease (TALEN) has rapidly emerged as an alternative to ZFNs for genome editing. TALENs are similar to ZFNs in that they consist of a FokI nuclease domain fused to a customizable DNA-binding domain. The primary difference between the two is that the DNA-binding domain of TALEs is not a zinc finger domain. Instead, it is a multiple 33–35-amino-acid repeat domain that recognizes a single base pair. This technology implements a simple ‘protein–DNA code’ in which the modular DNA-binding TALE repeat domains are complementary to individual bases in a unique target-binding site. This feature makes the design and production of TALEs much easier and more efficient than the production of ZFNs. Over the course of a relatively short time period, TALENs have been successfully used to perform genome editing in plants, zebrafish, frogs, rats and pigs. The technology has also been used in human somatic and pluripotent stem cells. Importantly, this technology has proven successful in several cell lines and model organisms that have been described as extremely difficult or even impossible to genetically manipulate. In this review, we summarize the major advances in TALE technology over the years. We focus on the interaction between TALENs and DNA, the design and construction of these proteins, how they are superior to ZFNs, and their potential applications.

2. The discovery of TALEs

TALEs were first discovered in the bacteria Xanthomonas oryzae pv. oryzae (Xoo) and Xanthomonas oryzae pv. oryzicola (Xoc). These two pathogens are responsible for decreasing the world’s rice production by causing bacterial blight (BB) and bacterial leaf streak (BLS) in the rice species Oryza sativa (Bennetzen and Ma, 2003; Ronald and Leung, 2002). The effector proteins in these bacteria are encoded by a gene family called avrBs3/PthA (short for avirulence/pathogenicity A). AvrBs3 is the first member of the avrBs3/PthA family and encodes a 122kD protein with 17.5 repeats. Each repeat, except for the non-full-length repeat at the end (i.e. the 0.5 repeat), consists of 34 tandem amino acids (aa) (Bonas et al., 1989). AvrBs3-like genes with repeat domain diversity are found in 309 strains of Ralstonia solanacearum biovars (Heuer et al., 2007).

These natural effector proteins, termed Type III effectors, are secreted by Xanthomonas bacteria via their type III secretion system (T3SS) and delivered into host cells where they regulate the transcription of genes involved in disease pathogenesis (Szurek et al., 2002). The T3SS is encoded by the hrc (hypersensitive response and pathogenicity) gene cluster. Expression of genes from this cluster facilitates transfer of proteins from bacteria to eukaryotic cells (Cornelis and Van Gijsegem, 2000). Once these bacterial proteins have entered the nucleus of a host cell, they can bind to target DNA sequences and regulate gene expression. Owing to their similarity to eukaryotic transcription factors, these bacterial effector proteins were termed transcription activator-like (TAL) effectors (also referred to as TALEs). They generally contain carboxyl (C)-terminal nuclear localization signals (NLSs), an acidic transcriptional activation domain (AD) (Schornack et al., 2006; Zhu et al., 1998), and an N-terminal type III secretion and translocation signal (T3S) (Bogdanove et al., 2010) (Fig. 1A). The presence of NLSs (2 and 3) of the AvrBs3 protein contributes to the occurrence of hypersensitive response (HR) (Van den Ackerveken et al., 1996). In contrast, NLS1 is insufficient for full activity of the AvrBs3 protein (Szurek et al., 2001). The original AD region is required for the HR-inducing activity of AvrBs3; the foreign AD can compensate for the AD-deficient AvrBs3 protein to some extent (Szurek et al., 2001).

Thorough analyses of the 113 known TALEs from Xanthomonas, Boch and Bonas (Boch and Bonas, 2010) revealed that these molecules have nearly identical repeats of 30–42 amino acids, with 34 amino acids being most typical. The number of repeats for each TALE ranges from 1.5 to 33.5 (most commonly 17.5). The structure of these repeats is the basis of the simple ‘protein–DNA code’ discussed in detail below.

3. Features of TALEs: recognition and specificity

TALEs are proteins with customizable DNA binding domains (Romé et al., 2007; Voytas and Joung, 2009) and are distinct from other DNA binding factors, including zinc finger proteins (ZFP) (Gommans et al., 2005), helix-turn-helix (HTH) proteins (Aravind et al., 2005), homeodomain (HD) factors (Mannervik, 1999), leucine zippers (Elhiti and Stasolla, 2009), and basic helix-loop-helix (bHLH) proteins (Jones, 2004). The extremely diverse TALE is the ‘protein-DNA code’ they follow, which allows specific recognition between one TALE and one nucleotide. TALEs differ from one another in the number and order of their repetitive sequences. Each repeat consists of approximately 34 amino acids (termed a monomer), which are highly conserved except for two hypervariable residues at positions 12 and 13 (Boch and Bonas, 2009) (Fig. 1A). These two amino acids (termed repeat variable diresidues, RVDs) determine the nucleotide-binding specificity of each repeat (Boch and Bonas, 2009). Some amino acid repeats do not recognize a specific nucleotide sequence; instead, some are more promiscuous. For example, NN and N* can recognize either G/A or C/T, respectively (Moscou and Bogdanove, 2009). Furthermore, a single nucleotide may be recognized by multiple repeats, albeit with different efficiencies. For example, NN is more efficient when binding with G instead of N (Christian et al., 2012). Amino acid repeats differing in the non-RVD regions have only minimal effect on the defined RVD specificity (Morbitzer et al., 2010) (summarized in Fig. 1B). Streubel et al. (2012) recently described new parameters to better characterize the specificity and efficiency of RVDs. Unlike earlier reports, the authors of this paper classified RVDs into weak, intermediate, and strong to describe their binding efficiencies; they suggested that these classifications should help researchers design more efficient TALEs (Streubel et al., 2012).

Since some repeats are nonspecific, how can one be sure that the TALE in question will actually target the sequence of interest in the genome? Previous work has shown that a minimum of 16.5 bases, with at least 6.5 continuous repeats from the N-terminus, is needed for specificity to activate expression (Boch et al., 2009). While TALEs with few repeats may also activate gene expression, those with repeats less than the threshold number display severely attenuated activity. Morbitzer et al. (2010) described that many mismatching repeats at the C-terminus limit activity and may also impair TALE integrity and stability. Such problems can generally be fine-tuned by increasing the number of repeat units.
The extent to which the position of mismatching repeats affects TALE activity is currently unknown. Some studies have shown that the position of some mismatches attenuate DNA-protein binding, while mismatches at other positions seem to have little to no effect (Romer et al., 2009b). Moreover, the N- and C-terminal capping regions are likely necessary for TALE activity when truncations are made near the loop regions (Zhang et al., 2011). The activity of TALEs may also be influenced by epigenetic mechanisms; for example, certain genomic regions may be inaccessible to TALE binding because of a condensed chromatin conformation. Although different TALEs display differential dependence on their N- and C-termini for activity, for most TALEs, the N-terminus of the repeat seems to be more important to the TALE-DNA interaction (Mussolino et al., 2011). At least four parameters have been described to influence the repeat-DNA recognition: (1) the number of minimal repeats needed to activate gene expression; (2) a strong negative effect of mismatching repeat-DNA combinations; (3) different repeat types with different specificity; and (4) mismatched repeat-DNA combinations with position- and context-dependence (Scholze and Boch, 2010).

The first nucleotide (a thymine, corresponding to repeat 0 at the TALE N-terminus) was shown to contribute to hax3 activation in N. benthamiana (Boch et al., 2009). Importantly, most natural TALE binding sites within the plant genome begin with a thymine residue (Moscou and Bogdanove, 2009). Thus, whether the 0th position is strictly limited to thymine, or whether this is a plant-specific phenomenon is currently unknown. For example, Briggs et al. (2012) found that some TALENs with different bases at the 0th position display similar activity, suggesting that variability at this position may be well tolerated. As another example, TALENs preceded by A, C, or G can all recognize and cleave the human beta-globin (HBB) gene responsible for sickle cell disease with equal efficacy (Sun et al., 2012).

4. Structure of TALE-DNA

Structural studies of the TALE-DNA interaction are necessary to facilitate a more complete understanding of the mechanism of DNA recognition by TALE repeats. The RVD is contained within a short loop structure that separates the TALE into two helices, which, themselves, are right-handed, superhelical structures (Boch and Bonas, 2010). The 13th residue is responsible for recognizing the specific nucleotide, whereas the 12th residue stabilizes the proper conformation of the RVD loop, which contacts the major groove of the DNA duplex. Similar findings have also been reported based on TALE-DNA binding simulations (Bradley, 2012). Hydrogen (H) bonding between the atoms of the 13th residue and base pair satisfactorily accounts for binding specificity of HD, NG, and NS for C, T, and A, respectively. Similar structural characteristics have also been described for the TALE PthXo1 from Xanthomonas oryzae; for this molecule, both H bonding and van der Waals forces play a key role in associations between HD, NG (or HG), NN, and NI with C, T, and A, respectively (Mak et al., 2012).

Deletion within a repeat at position 13 may result in a failure to recognize the primary base. As an example, when the G and D were deleted from the NG and ND repeat, respectively, specificity was altered to C/T. Change is specified to be influenced by all of the amino acids involved (Boch and Bonas, 2010). Interestingly, Yan, Shi, Zhu and colleagues (Deng et al., 2012) used electrophotoretic mobility shift assay (EMSA) to show that 5-methylcytosine (mC) can be recognized by the TALE code NG. This is the first description of TALEs binding to modified DNA. TALEs have also been shown to bind to DNA-RNA hybrids, protecting the structure from RNase H-mediated degradation (Yin et al., 2012). Research continues to identify new applications for TALE technology.

5. Magic of TALEs: regulation of gene transcription and genome editing

An important topic to address is the mechanism by which TALEs regulate transcription in the host. TALEs translocate to the nucleus in the host cell via T3SS. They bind to UPT (up regulated by TALEs) boxes in target genes and regulate their expression (Morbitzer et al., 2010). This type of gene regulation does not appear to be species-specific. For example, the Xoo AvrXa27 protein can bind to and regulate both the rice Xa27 promoter and the pepper Bs3 gene.
intact, TALEs appear to retain some transcriptional activation po-
eration (Fig. 2). Importantly, even with the activation domain
recruit multiprotein transcription complexes to regulate gene
expression (De Souza et al., 2012). These proteins are engineered to form heterodimers that target both sides of a specific sequence. While the specific locus to
induce DNA double-strand breaks (DSBs) via type II restriction
endonuclease activity (Christian et al., 2010). Following the
creation of these breaks, two major pathways are employed to repair
the damage (Huertas; 2010) — either nonhomologous end-joining
(NHEJ) (Lieber, 2008, 2010) or homologous directed recombination
(HDR) (Hockemeyer et al., 2011; Miller et al., 2011). Fig. 2). NHEJ is more common in mammalian cells and generally results in
gene disruption. Li et al. (2011a) have constructed the TALEN hybrid
proteins FN-AvrXa7, AvrXa7-FN, and PthXo1-FN. These hybrids
consist of the TALE from Xoo (either AvrXa7, 26 repeats or PthXo1,
24 repeats) and the FokI DNA-cleavage domain (FN). The authors
demanded that these TALENs recognize specific effector binding
elements (EBE) and exhibit nuclease function in a yeast cell based
assay. Similar approaches using TALENs in the more complex hu-
man genome have also been described (Miller et al., 2011). In
addition to FokI, other nucleases have also been used. One such
element is I-TevI, a nuclease that has been shown to participate in
gene editing (Kleinstiver et al., 2012). The length of spacer be-
tween chimeric nucleases binding sites greatly influences the ac-
tivity and efficiency of cleavage, and it is partially determined by
the distance between the binding and cleavage domains (termed
the linker region) (Bibikova et al., 2001) (Fig. 3). Similar findings
have been reported for AvrB3 and PthXo1 (Christian et al., 2010)
and for Hax3 (Mahfouz et al., 2011). The spacer length depends on
the affinity of the RVD for the genetic locus of interest, the specific

Taken together, TALEs recognize target DNA sequences and re-
cruit multiprotein transcription complexes to regulate gene
expression (Fig. 2). Importantly, even with the activation domain
intact, TALEs appear to retain some transcriptional activation po-
tential (Geissler et al., 2011). Such findings suggest that the AD is
not the only domain that influences TALE transcriptional activity
and other unknown mechanisms probably also play a role.

Fig. 2. Designer TALE transcription factors and nucleases activate endogenous DNA transcription and editing. Footnotes: TALEs can be used to generate TALE transcription factors (TALE-TFs) that regulate the transcription of target genes. The TALE DNA-binding domain is fused to the synthetic VP64 transcriptional activator and assembles transcriptional multiprotein complexes with RNA polymerase II. TALE nucleases (TALENs)—induced DNA double-strand breaks (DSBs) typically result in one of two major DNA damage repair
pathways: nonhomologous end joining (NHEJ) or homologous directed recombination (HDR) repair. NHEJ directly ligates the ends of the broken DNA strands and introduces an
error-prone deletion or insertion at the repair site by DNA repair proteins and ligase. Gene editing through HDR from the donor template can result in gene additions and re-
placements (or corrections). TSS, transcription start site; NHEJ, nonhomologous end joining; HDR homologous directed recombination.

Please cite this article in press as: Zhang, M., et al., TALE: A tale of genome editing, Progress in Biophysics and Molecular Biology (2013), http:// dx.doi.org/10.1016/j.pbiomolbio.2013.11.006
features of the sequence, and the length of the TALE-DNA binding region. Thus, the linker length has a positive relationship with the spacer length (Mahfouz and Li, 2011). Along these lines, a long linker length requires a long spacer length to maintain robust activity (Sun et al., 2012). Importantly, a spacer length between 10 and 30bp has been identified as optimal for TALEN function (Bogdanove and Voytas, 2011; Sun et al., 2012).

6. The making of TALEs: designing and engineering

The commercial production of TALEs is becoming more and more popular. However, because of their repetitive nature, it is difficult to construct TALEs with definite sequence specificity. To try to overcome this technical difficulty, Zhang et al. (2011) developed a hierarchical ligation-based strategy. The principle is to first form three tetramers and then to reassemble them into a linear tandem complex based on the unique sticky-end surrounding each monomer junction. Weber et al. (2011) generated TALEs using the Golden Gate cloning technology, which requires type IIS enzymes and occurs in two successive steps. First, there is a BsaI-based pre-assembly reaction of selected repeat modules, and next, there is a BpiI-based final assembly reaction of the preassembled repeat blocks. This modular cloning approach has been used in other studies as well (Li et al., 2011b; Morbitzer et al., 2011). Li et al. (2012) described a rapid and highly efficient approach to construct the backbones of both TALE transcription factors (TALE-TFs) and TALENs in a one-step sequential restriction—ligation reaction using type II restriction enzymes. Sanjana et al. (2012) have further modified TALE-TFs to regulate transcription of human genes and TALENs to mediate gene knockout or insertion; these techniques are constantly improving (Cong et al., 2012). These research findings are also available online (http://taleffectors.com/). This online access allows rapid communication among researchers working with TALEs to greatly improve the technology. Garg et al. (2012) developed an algorithm to computationally design TALEs with cognate binding sites that are at least a given number of mismatches away from putative human promoter sequences. They proved that synthetic TALEs can effectively activate transcription from specific target sites without influencing expression of the closest off-target genes. Cermak et al. (2011) have expanded the utility and flexibility of TALEs by describing a methodology that does not rely on PCR. Another method uses fast ligation-based automatable solid-phase high-throughput (referred to as FLASH) technology to accomplish large-scale assembly of TALENs (Reyon et al., 2012). This method builds TALENs by sequential ligation of pre-assembled 4-mers on magnetic beads using a liquid-handling instrument. Briggs et al. (2012) created an approach called iterative capped assembly (ICA) used to construct custom TALENs fairly successfully and rapidly through sequential ligation of repeat monomers; this method differs from FLASH in that it is based on a
solid support that increases the frequency of final products up to 21 monomers. It also uses hairpin capping oligonucleotides to block the growth of incomplete chains. Recently, Schmid-Burgk et al. (2013) created a ligation-independent cloning (LIC) technique. They generated a 5-mer TALE repeat fragment library for the synthesis of TALEs. This method has the benefits of speed, simplicity, high fidelity, high throughput, and automation. Despite technological advances, off-target effects continue to be a major problem (Mussolino et al., 2011; Tesson et al., 2011). Lack of sequence recognition, context-dependence, and the formation of homodimers can all reduce TALE specificity and induce cytotoxicity due to cleavage at off-target sites. There is no doubt that decreased off-target effects will lead to more widespread applications of TALE technology.

8. Practical applications

Custom-designed TALE-TFs have been successfully generated to target the A. thaliana genes EGL3 and KNAT1 (Motorbitzi et al., 2010). Homo sapiens genes KLF4, Sox2 (Zhang et al., 2011), NF13 (Miller et al., 2011), PUMA, IFNα1, and IFNβ1 (Geissler et al., 2011). TALENs have been successfully used in nematodes (Wood et al., 2011), zebrabean (Sander et al., 2011), rats (Tesson et al., 2011), silkworms (Ma et al., 2012), mammals (Carlson et al., 2012), crickets (Watanabe et al., 2012), fruit flies (Li et al., 2012), humans (Miller et al., 2011), and in pluripotent stem cells (Geissler et al., 2011; Hockemeyer et al., 2011). Despite these successes, there is no evidence directly linking TALE-mediated gene regulation to TALE-mediated gene cleavage (Mussolino et al., 2011).

Reyon et al. (2012) used the FLASH system to produce 48 TALENs. They tested the activities of the proteins in a reporter screen based on eGFP (enhanced green fluorescent protein) expression. In that study, they also examined the effect of TALENs on 96 endogenous cancer-related genes. They found a high success rate for TALEN activity in both assays.

TALEs can have different functions depending on the domain they are fused to. Types of domains that have been used in the past include activation domains, repression domains, nucleases, chromatin modifying domains, and fluorescent proteins. TALEs have also been used to construct transgenic animals, to generate isogenic pluripotent stem cells (PSCs), and for use in vivo gene therapy (Hsu and Zhang, 2012). TALENs have been used for therapeutic applications, including the treatment of sickle cell disease (Sun et al., 2012), chronic viral infections (Schiffer et al., 2012), and neurodegenerative genetic disorders (Hsu and Zhang, 2012). TALENs have the potential to be applied to many more diseases in the future.

Many factors must be considered before TALEs can be widely used either in basic research or in clinical applications. Some of these factors include the specificity and efficiency of RVDs, off-target effects, structural stability of repeat modules, TALE-DNA binding activity, and linker and spacer lengths. In addition, context dependence and cytotoxicity of TALEs require detailed examination. Importantly, viral infections can cause genetic mutations in the host (Schiffer et al., 2012). Thus, for therapeutic implications, TALEN delivery should be studied in more detail. Additionally, since TALEs are only found in plant pathogens, immune responses in clinical treatment are quite possible. These are all factors that must be considered before TALEs can be used clinically.

8. TALEN vs. ZFs

Technology based on both TALENs and ZFNs has attracted tremendous attention in recent years. ZFNs precisely target nucleic acid sequences of interest and have been widely used for genome editing in a variety of cells; they have also been used therapeutically (Urnov et al., 2010). Despite the success of ZFNs, a lot of research suggests that TALENs are superior to ZFNs (Christian et al., 2010; DeFrancesco, 2011; Motorbitzi et al., 2010; Mussolino et al., 2011; Ramirez et al., 2008) (Table 1). However, before such claims can be substantiated, more work needs to be done. In contrast to the study of ZFNs, which has been around for many years, the study of TALEs really only began in 2007. Thus, much more research is necessary.

A natural RNA-guided DNA nuclease system has been recently identified (Cho et al., 2013; van der Oost, 2013). The clustered regularly spaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) protein system was discovered in bacteria and archaea for adapted immunity. The system consists of the typical nuclease Cas9, two RNAs called CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), and host factor ribonuclease (RNase) III. The crRNA is complementary to foreign DNA sequences and directs Cas9 to cleave these sites. Some groups have demonstrated that the Cas9 system provides greater specificity and efficiency, affordable engineering, reduced off-target effects, and decreased cytotoxicity when applied to mammalian genomes (Cong et al., 2013; Mali et al., 2013). In addition, RNA-guided endonucleases (RGENs) can induce multisite cleavage in a single cell, which may be difficult to achieve by TALENs or ZFNs. However, recent reports also suggest significant off-target mutagenesis induced by RGENs (Fu et al., 2013). Whether RGENs will prove to be more promising than TALENs and ZFNs remains to be determined.

9. Summary

Nuclease-mediated gene editing can generate cells or organisms with novel functions through activation, replacement, removal, or addition of genomic elements (McMahon et al., 2012). In addition, gene editing can be used to simulate or interfere with off-target mutagenesis induced by RGENs (Fu et al., 2013). Whether RGENs will prove to be more promising than TALENs and ZFNs remains to be determined.

Table 1 Comparison of TALENs and ZFNs.

<table>
<thead>
<tr>
<th>Items</th>
<th>TALENs</th>
<th>ZFNs (commonly C2H2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central domain</td>
<td>Tandem repeats of ~34-aa</td>
<td>Zinc fingers of ~30-aa</td>
</tr>
<tr>
<td>Determining region</td>
<td>RVD of 12th and 13th</td>
<td>Alpha-1,3,6 amino acid</td>
</tr>
<tr>
<td>Secondary structure</td>
<td>Two helices connected by a short RVD-containing loop</td>
<td>Beta folding-beta folding-alpha helix (β-β-x)</td>
</tr>
<tr>
<td>DNA binding pattern</td>
<td>One repeat for one nucleotide</td>
<td>One zinc finger for one trinucleotide</td>
</tr>
<tr>
<td>Validity</td>
<td>Very little apparent context dependence</td>
<td>Context dependence</td>
</tr>
<tr>
<td>Specificity</td>
<td>Relatively high</td>
<td>Relatively low (off-target effect)</td>
</tr>
<tr>
<td>Degeneracy</td>
<td>Exising same repeat for distinct nucleotides</td>
<td>Having the affinity for similar but not identical sequences their intended targets</td>
</tr>
<tr>
<td>Toxicity</td>
<td>No or little (if any) detectable cytotoxicity</td>
<td>Yes</td>
</tr>
<tr>
<td>Optimized dimers</td>
<td>10-30 bp spacers with corresponding linkers</td>
<td>18–24 bp DNA linkers with 4–7 bp spacers</td>
</tr>
<tr>
<td>Applicable sites</td>
<td>Arbitrary theoretically</td>
<td>Existing modules can’t target every possible sequence</td>
</tr>
<tr>
<td>Predictability</td>
<td>More flexibility and repeats</td>
<td>Less flexibility and fingers</td>
</tr>
<tr>
<td>Intellectual property</td>
<td>Qualifed</td>
<td>Sangamo and Sigma</td>
</tr>
<tr>
<td>Cost</td>
<td>Relatively low-cost and time-saving</td>
<td>Relatively costly and time-consuming</td>
</tr>
</tbody>
</table>

TALENS: transcription activator-like effectors; ZFNs: zinc finger nucleases, C2H2/2-cysteine and 2-histidine, RVD: repeat variable diresidues.
differentiation of pluripotent stem cells and may be used as ther-
apies for genetic diseases. TALEs have proved to be a driving force in
the genome editing field. Because of DNA recognition specificity,
TALEs can be used to promote gene transcription or to knockout a
specific gene of interest. TALEs may be useful for therapeutic
application in the future. However, before this can happen, the
mechanism of action governing TALEs must be further elucidated.
Specifically, we much achieve a decrease in off-target effects, and
future research efforts should be aimed at this purpose.

Conflict of interest

The authors do not declare any conflict of interest relevant to
this manuscript.

Acknowledgements

This work was supported by National Natural Science Foundation
of China grant 3107123 and Chongqing Science & Technology
Commission grant CSTC2010BB5026 (Xueqing Xu)

References

of the helix-turn-helix domain: transcription regulation and beyond. FEMS
Microbiol. Rev. 29, 231–262.
Bennetzen, J.L., Ma, J., 2003. The genetic colinearity of rice and other cereals on the
Boch, J., Bonas, U., 2010. Xanthomonas AvrBs3 family-type III effectors: discovery
Bonas, U., Stall, R.E., Staskawicz, B., 1989. Genetic and structural characterization of
21, 471–478.
Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA
339, 1593–1596.
Biotechnol. (Advance online publication).
Transcriptional activators of human genes with programmable DNA-specificity. PLoS
One 6, e16509.
transcription factors: the therapeutic relevance of switching endogenous gene
Hall, D.B., Struhl, K., 2002. The VP16 activation domain interacts with multiple
transcriptional components as determined by protein-protein Cross-linking
of avrBs3-like genes in Ralstonia solanacearum strains and association with host preferences in the pathogen. Appl. Environ. Microbiol. 73, 4379–4384.
Hsu, P.D., Zhang, F., 2012. Dissecting neural function using targeted genome engi-
5, 229.
A 109, 8061–8066.
transcriptional regulators and nucleases for genome modulation. Plant Mol.
TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI
transcriptional regulators and nucleases for genome modulation. Plant Mol.
Highly efﬁcient and speciﬁc genome editing in silkworm using custom TALENs.
Crops 2, 59–103.
Mahfouz, M.M., Li, L., Piatek, M., Fang, X., Mansour, H., Bangarusamy, D.K.,
engineering of TAL effector/nucleaselike effectors (TALE) hybrid nucleases with
Mak, A.S., Bradley, P., Cerda, R., Bogdanove, A.J., Stoddart, B.L., 2012. The crystal
structure of TAL effector PthXo1 bound to its DNA target. Science 335, 716–719.
Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E.,
339, 823–826.
Sander, J.D., Cade, L., Khayter, C., Reyon, D., Peterson, R.T., Joung, J.K., Yeh, J.-R.J.,
Ronald, P., Leung, H., 2002. The rice genome. The most precious things are not jade
Romer, P., Strauss, T., Hahn, S., Scholze, H., Morbitzer, R., Grau, J., Bonas, U.,
Ramirez, C.L., Foley, J.E., Wright, D.A., Müller-Lerch, F., Rahman, S.H., Cornu, T.I.,
dx.doi.org/10.1016/j.pbiomolbio.2013.11.006
Please cite this article in press as: Zhang, M., et al., TALE: A tale of genome editing, Progress in Biophysics and Molecular Biology (2013), http://
geted DNA mutagenesis for the cure of chronic viral infections. J. Virol. 86,
Protoc. 7, 171
A transcription activator-like effector toolbox for genome engineering. Nat.
TALENs. Nat. Biotechnol. 29, 697
and pearls. Science 296, 58
1712.
171
Lahaye, T., 2009b. Recognition of AvrBs3-like proteins is mediated by speci
cification. J. Virol. 86, 4430.
A novel TALE nuclease scaffold enables high genome editing activity in combi-
nation with low toxicity. Nucleic Acids Res. 39, 9283
A transcription activator-like effector toolbox for genome engineering. Nat.
Schiffer, J.T., Aubert, M., Weber, N.D., Mintzer, E., Stone, D., Jerome, K.R., 2012. Target-
mediated recognition of nuclear-targeted AvrBs3-like bacterial effector pro-
mination of a hemimetabolous insect using zinc-
pathway. Mol. Cell. 36, 255
Eukaryotic features of the Xanthomonas type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. Plant J. 26, 523–534.
1492.
1316.
11, 534
696.
192.
11, 824
533.
16765.
266.
192.
432.
44.
1491.
1255.
81.
1492.
81.
153.
153.
512.
11, 832.
75.
596.
1492.
1316.
1255.
8936.