Robust, synergistic regulation of human gene expression using TALE activators

Morgan I. Maeder¹-⁴, Samantha J. Linder¹–3,6, Deepak Reyon¹–3,5,6, James F. Angstman¹–3, Yanfang Fu¹–3,5, Jeffry D. Sander¹–3,5 & J. Keith Joung¹–5

Artificial activators designed using transcription activator–like effector (TALE) technology have broad utility, but previous studies suggest that these monomeric proteins often exhibit low activities. Here we demonstrate that TALE activators can robustly function individually or in synergistic combinations to increase expression of endogenous human genes over wide dynamic ranges. These findings will encourage applications of TALE activators for research and therapy, and guide design of monomeric TALE-based fusion proteins.

Rapid advances in Xanthomonas sp.–derived TALE technology have enabled any researcher to construct customizable DNA-binding domains with broad potential uses for targeted alteration of gene sequence or expression. Highly conserved 33–35–amino-acid TALE repeat domains each bind one nucleotide of DNA with specificity dictated by two hypervariable residues. This one-to-one code allows one to design proteins with desired DNA-binding specificities by simply joining TALE repeats into an array. Recently, considerable effort has been focused on dimeric TALE nucleases (TALENs), artificial proteins composed of customized TALE repeats fused to a nuclease domain, which enable targeted modification of endogenous genes in a variety of organisms and cell types. Engineered TALEs have also been fused to heterologous transcriptional activation domains to construct artificial monomeric TALE activators (Supplementary Fig. 1). In terms of the capability to increase target gene expression, however, monomeric TALE activators reported thus far have often exhibited modest activities at best compared to the high efficiencies reported for dimeric TALENs. For example, 22 of 26 published TALE activators (for which quantitative information is available) did not induce target endogenous gene expression by fivefold or more (Supplementary Table 1). The use of multiple different architectures makes it difficult to ascertain what parameters may or may not influence the activities of the TALE activators tested in these previous studies (Supplementary Table 1).

Here we leveraged our recently developed fast ligation-based automatable solid-phase high-throughput (FLASH) assembly method to systematically test the activities of TALE activators. In initial experiments, we constructed many TALE activators composed of variable numbers of TALE repeats and tested the capabilities of these activators to stimulate expression of the endogenous human VEGFA gene. We targeted nine regions in a DNase I hypersensitive site located ~500 base pairs downstream of the VEGFA transcription start site (Fig. 1a) because...
previously published work has shown that targeting sequences in DNase I hypersensitive sites is an important parameter for elevating expression of a target gene with artificial zinc finger–based transcriptional activators\(^1\). We constructed sets of six variable-length TALE activators (composed of 14.5, 16.5, 18.5, 20.5, 22.5 or 24.5 TALE repeats) for each of the nine target regions. All 54 proteins had a common architecture similar to the one previously described\(^2\) but one that harbors a VP64 (instead of a VP16) activation domain (Online Methods and Supplementary Fig. 1). We found that 53 of these 54 TALE activators induced significant increases in VEGF-A protein expression ranging from 5.3- to 114-fold (mean of 44.3-fold activation; Fig. 1b) and that these activities did not depend on which strand of DNA was bound. We do not know why we observed variability in the extent of activation but possible explanations include differences in protein expression or in the DNA-binding activities of the arrays. Regardless of mechanism, our results suggest that TALE activators can function efficiently when they are targeted to a DNase I hypersensitive site.

To test the robustness of TALE activators, we built proteins targeted to six additional sites in the human VEGFA promoter, five sites in the human NTF3 promoter and five sites in the promoter of the microRNA promoter MIR302b,c,d–MIR367 (miR-302/367). All of these TALE activators comprised 16.5 or 17.5 TALE repeats and were targeted to sites within DNase I hypersensitive sites (Supplementary Fig. 2 and Online Methods). Notably, all six TALE activators targeted to the VEGFA promoter and four of five TALE activators targeted to the miR-302/367 promoter induced significant increases in target gene expression in human HEK293 and primary BJ fibroblasts, respectively (Fig. 2a,b). Because NTF3 mRNA was expressed at an essentially undetectable level in the HEK293 cells used for our experiments, we could not quantify fold-activation values for proteins targeted to this gene, but all five TALE activators significantly increased expression of NTF3 relative to a GAPDH control (Fig. 2c). Overall, 15 of 16 TALE activators (~94%) induced significant increases in expression of their endogenous gene targets (Fig. 2a–c).

We next explored whether replacing the VP64 domain in our TALE activators with an NF-kB p65 domain would lead to consistently higher or lower target gene expression. For the 15 target sites for which we obtained active TALE activators, the mean fold activation induced was lower with the NF-kB p65–containing TALE activators than with their matched VP64-containing counterparts (here called p65 and VP64 TALE activators, respectively;
sequence architecture of the TALE DNA-binding domains used (Supplementary Discussion). Our findings also expand the types of genes and the range of DNA sequences that can be targeted by TALE activators. To our knowledge, our results provide the first demonstration that it is possible to activate a non-coding gene, thereby broadening the range of potential targets for TALE activators. In addition, analysis of our data suggests that there are no substantial limitations in the range of sequences we can target (Supplementary Discussion and Supplementary Table 2). Taken together, our results provide strong experimental support for the use of TALE activators to control expression of essentially any gene.

We showed that TALE activators can be used to regulate target genes across a wide dynamic range of expression, an important capability that will enable a broader range of applications for this technology. Our studies suggest multiple potential approaches that might be used to fine-tune the level of gene expression induced by TALE activators (Supplementary Discussion). The finding that TALE activators can synergistically activate transcription broadens the range of gene expression changes that can be achieved with this platform and raises the exciting possibility that target genes might be made responsive to multiple inputs, as has recently been shown with engineered zinc-finger transcription factors. The greater targeting range of engineered TALE activators relative to that of engineered zinc-finger activators provides a substantial advantage for enabling synthetic biology applications in which artificial circuits are designed to interface with endogenous genes.

Here we demonstrated that TALE activators should function efficiently to regulate essentially any protein-coding or noncoding gene in human cells. This capability provides a useful complement to regulation strategies involving cDNA overexpression or RNA interference for studying gene function and to previously described synthetic biology strategies for regulating endogenous gene expression. An important area for future investigation will be the potential off-target effects of TALE activators in human cells (Supplementary Discussion). Our successes with TALE activators in human cells should encourage the use of these proteins in other cell types and organisms. More importantly, our findings should inspire the generation of other monomeric TALE-based fusion proteins that might be used to rationally alter expression and/or the epigenetic status of genes. Thus, our findings should stimulate efforts to expand the repertoire of engineered TALE-based tools available for research, synthetic biology and therapeutic applications.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by a US National Institutes of Health (NIH) Director’s Pioneer Award DP1 OD006862 (J.K.J.), NIH P50 HG005550 and R01 NS073124 (J.K.J.), the Jim and Ann Orr Massachusetts General Hospital Research Scholar Award (J.K.J.), a US National Science Foundation Graduate Research Fellowship (M.L.M.), and NIH T32 CA009216 (J.D.S.). We thank J. Foley for technical assistance with construction of TALE-activator plasmids, R. Mylvaganam for technical assistance with flow cytometry, and the Massachusetts General Hospital Nucleic Acid Quantitation Core (supported by NIH P30 NS45776) for assistance with performing real-time reverse-transcription–PCR assays.

AUTHOR CONTRIBUTIONS

M.L.M., S.J.L., D.R., J.F.A., Y.F., J.D.S. and J.K.J. designed the experiments. M.L.M. and J.K.J. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at http://www.nature.com/doifinder/10.1038/nmeth.2366. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Joung, J.K. & Sander, J.D. Nat. Rev. Mol. Cell Biol. 14, 49–55 (2012).
2. Mussolino, C. & Cathomen, T.Curr. Opin. Biotechnol. 23, 644–650 (2012).
3. Miller, J.C. et al. Nat. Biotechnol. 29, 143–148 (2011).
4. Zhang, F. et al. Nat. Biotechnol. 29, 149–153 (2011).
5. Gelisler, R. et al. PloS ONE 6, e19509 (2011).
6. Gang, A., Lohmueller, J.J., Silver, P.A. & Armel, T.Z. Nucleic Acids Res. 40, 7584–7595 (2012).
7. Tremblay, J.P., Chapdelaine, P., Coulombe, Z. & Rousseau, J. Hum. Gene Ther. 23, 883–890 (2012).
8. Wang, Z. et al. Angew. Chem. Int. Edn. Engl. 51, 8505–8508 (2012).
9. Cong, L., Zhou, R., Kuo, Y.C., Cunniff, M. & Zhang, F. Nat. Commun. 3, 968 (2012).
10. Bultmann, S. et al. Nucleic Acids Res. 40, 5368–5377 (2012).
11. Cermak, T. et al. Nucleic Acids Res. 39, e82 (2011).
12. Reyon, D. et al. Nat. Biotechnol. 30, 460–465 (2012).
13. Liu, P.O. et al. J. Biol. Chem. 276, 11323–11334 (2001).
14. Carey, M., Lin, Y.S., Green, M.R. & Ptashne, M. Nat. Biotechnol. 30, 361–364 (1990).
15. Khalil, A.S. et al. Cell 150, 647–658 (2012).
16. Deans, T.L., Cantor, C.R. & Collins, J.J. Cell 130, 363–372 (2007).
17. Tigges, M., Marquez-Lago, T.T., Stelling, J. & Fussenegger, M. Nature 457, 309–312 (2009).
18. Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Science 333, 1307–1311 (2011).
19. Culler, S.J., Hoff, K.G. & Smolke, C.D. Science 330, 1251–1255 (2010).
20. Auslander, S., Auslander, D., Muller, M., Wieland, F. & Fussenegger, M. Nature 487, 123–127 (2012).
**ONLINE METHODS**

**Selection of TALE-activator binding sites.** For the human VEGFA gene, target sites were chosen that fall within DNase I hypersensitive sites previously described for HEK293 cells. For the NTF3 and miR-302/367 cluster genes, target sites were chosen within DNase I hypersensitive regions identified from University of Washington Encyclopedia of DNA Elements (ENCODE) data using the University of California Santa Cruz (UCSC) Genome Browser (http://genome.ucsc.edu/); we targeted these regions because they have been identified as DNase I hypersensitive sites in multiple different cell types, and we therefore reasoned that these areas had a high probability of being in open chromatin.

**Construction of TALE activators.** DNA fragments encoding TALE repeat arrays were generated using the FLASH method. These fragments were cloned using overhangs generated by digestion with BsmBI restriction enzyme into expression vectors containing the EEF1A1 (EF1a) promoter and sequence encoding the N-terminal and C-terminal TALE-derived domains from the previously described TALE activator NT-L+95 (ref. 3). As previously defined, Δ152 refers to a truncated N-terminal TALE-derived domain lacking the first 152 amino acids, and +95 refers to the C-terminal TALE-derived domain containing only the first 95 amino acids following the C-terminal TALE repeat. NF-κB p65 and VP64 activation domains were fused directly to the C-terminal end of the +95 domain, and all fusion proteins contained a nuclear localization signal. Plasmids for constructing VP64 and p65 TALE activator expression vectors are available from Addgene (http://www.addgene.org/talengineering/taletfs/).

**Cell culture and transfection.** Human Flp-In T-Rex HEK293 cells and primary human BJ fibroblasts were maintained in Advanced DMEM supplemented with 10% FBS, 1% penicillin-streptomycin and 1% Glutamax (Life Technologies). Cells were transfected using either Lipofectamine LTX (Life Technologies) or Nucleofector technology (Lonza) according to manufacturer’s instructions. Briefly, for experiments targeting VEGA and NTF3, 160,000 Flp-In T-Rex HEK293 cells were seeded in 24-well plates and transfected the following day with 300 ng of plasmid encoding the TALE activator (except in the reduced-concentration VEGF-A experiments (Fig. 2d) in which 50 ng of plasmid encoding TALE activator and 250 ng of a control plasmid expressing only the activation domain were transfected), 30 ng of pmaxGFP plasmid (Lonza), 0.5 μl Plus Reagent and 1.65 μl Lipofectamine LTX. For the synergy experiments with VEGA and NTF3, we transfected cells with 300 ng of TALE activator–encoding plasmids (50 ng of each of six TALE activator–encoding plasmids for VEGA and 60 ng of each of five TALE activator–encoding plasmids for NTF3). For experiments targeting miR-302/367 cluster expression, 5 × 10⁵ BJ fibroblasts were transfected using Nucleofector technology with 10 μg of plasmid encoding TALE activator and 500 ng of pmaxGFP plasmid using the NHDF kit (Lonza) and program U-023 on the Nucleofector 2b device. For the synergy experiment with miR-302/367, we transfected cells with 10 μg of TALE activator–encoding plasmids (2 μg of each of five TALE activator–encoding plasmids).

**Enzyme-linked immunosorbent assays.** Flp-In T-REx HEK293 cells were transfected with plasmids encoding TALE activators targeted to the human VEGFA gene. All transfections were performed in triplicate. Cell medium was collected 40 h after transfection, and secreted VEGF-A protein levels in the medium were assayed using a Human VEGF-A ELISA kit (R&D Systems). All samples were measured according to the manufacturer’s instructions. Fold-activation values were calculated by dividing mean VEGF-A amounts in medium collected from cells transfected with plasmids expressing TALE activators by mean VEGF-A amounts in medium from cells transfected with plasmid expressing only the VP64 or NF-κB p65 activation domain.

**Quantitative reverse transcription–PCR assays.** To measure NTF3 mRNA levels, cells were collected 2 d after transfection, and total RNA was isolated using the TRIzol Plus RNA purification system (Ambion). RNA was reverse-transcribed using SuperScript III First-Strand Synthesis SuperMix and oligo(dT) primer (Life Technologies). Quantitative PCR was then performed using the following Taqman primer/probe sets, as previously described except with the modification that the GAPDH probe was labeled with a 5′ hexachlorofluorescein (HEX) fluorophore to allow for multiplexing. NTF3 forward primer: 5′-GATAAACACTGGAATCTCATGGCA-3′; NTF3 reverse primer: 5′-GCCACCCCCAGAATTTGTTGT-3′; NTF3 Taqman probe: 5′/56-FAM/CACACCTACZN/ ZEN/ GTCCGAGCAGTACTCAGAG/31AbkFQ/-3′; GAPDH forward primer: 5′-CATGTACTGTCATGGGTGTGA-3′; GAPDH reverse primer: 5′-CATGGACTGTGGTCATGAGT-3′; GAPDH Taqman probe: 5′/56FAM/CACACCTACZN/ ZEN/ACAACCTCATTCAG/C31AbkFQ/-3′. All TALE activator–encoding plasmids and control plasmids were introduced into cells by transfection using Nucleofector technology in triplicate and quantitative reverse transcription–PCR (RT-PCR) was performed in triplicate on each sample.

To measure miR-302a transcript levels, cells were collected 3 d after transfection and GFP-positive cells were isolated by flow cytometry. Total microRNA was isolated using the mirVana miRNA Isolation kit (Ambion). Reverse transcription and quantitative PCR were performed according to manufacturer’s instructions using Applied Biosystems Taqman microRNA assays (#000529 for has-miR-302a and #001006 for RNU48 control). Fold activation of miR-302a RNA transcripts was calculated by comparing transcript amounts in BJ fibroblasts transfected with plasmids encoding TALE activators to transcript amounts in BJ fibroblasts transfected with control plasmids expressing only the VP64 or NF-κB p65 activation domains and using the comparative cycle threshold (ΔΔCt) method. All TALE activators and controls were introduced into cells by transfection using Nucleofector technology, in triplicate, and quantitative RT-PCR for miR302a transcript and small RNA control SNORD48 (RNU48) were performed in triplicate on each sample.

21. Rosenbloom, K.R. et al. Nucleic Acids Res. 40, D912–D917 (2012).