ACS SyntheticBiology

pubs.acs.org/synthbio Research Article

# TALE.Sense: A Versatile DNA Sensor Platform for Live Mammalian Cells

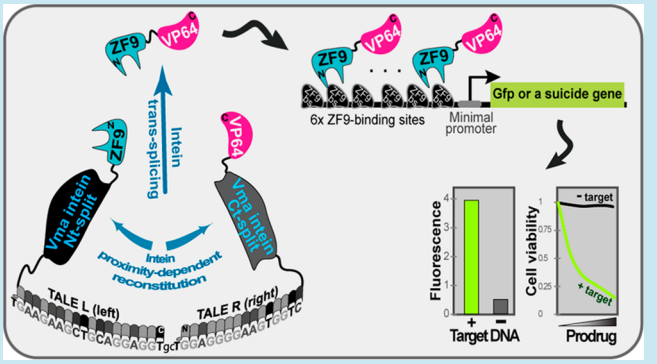
Aziz Taghbalout,* Nathaniel Jillette, and Albert W. Cheng*

Cite This: *ACS Synth. Biol.* 2022, 11, 116–124 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



**ABSTRACT:** Here we describe TALE.Sense, a versatile platform for sensing DNA sequences in live mammalian cells enabling programmable generation of a customable response that discerns cells containing specified sequence targets. The platform is based on the programmable DNA binding of transcription activator-like effector (TALE) coupled to conditional intein-reconstitution producing a trans-spliced ON-switch for a response circuit. TALE.Sense shows higher efficiency and dynamic range when compared to the reported zinc-finger based DNA-sensor in detecting same DNA sequences. Swapping transcriptional activation modules and introducing SunTag-based amplification loops to TALE.Sense circuits augment detection efficiency of the DNA sensor. The TALE.Sense platform shows versatility when applied to a range of target sites, indicating its suitability for applications to identify live cell variants with anticipated DNA sequences. TALE.Sense could be integrated with other cellular or synthetic circuits by using specified DNA sequences as control-switches, thus expanding the scope in connecting inducible modules for synthetic biology.




**KEYWORDS:** *DNA sensor, biosensor, transcription activator-like effector, TALE, DNA sequence-targeting, intein, trans-splicing*

The ability to identify and isolate live mammalian cell variants by detecting a specific DNA sequence in a subset of cell populations will be enormously advantageous to research investigations as well as therapy developments. This will require tools enabling effective sensing for DNA sequence targets and reporting by triggering a customized cellular response upon detection. The response could include an induced expression of a fluorescent protein as a detection signal, kill-switch inducing apoptotic cellular response, or targeted immune response in cells containing a specific DNA sequence target. The system could then be applied to a wide range of DNA sequences resulting from programmed genomic edits, aberrant chromosomal rearrangements, and viral infections or integrations in cancer cells.

The property of inteins to excise themselves and splice otherwise separated protein domains has been instrumental to developing tools for biological studies. Unlike spontaneous trans-splicing of highly active split inteins,[1,2] *Saccharomyces cerevisiae* Vma splits need to be brought to close proximity via a dimerization domain.[3,4] This proximity-triggered conditional protein trans-splicing of Vma intein, combined with the programmable feature of zinc-finger proteins (ZF) to bind DNA, was exploited to build a sequence-specific DNA-sensor that induces activation of a reporter gene, or reconstitution of luciferase activity upon sequence recognition.[5,6]

The ZF-based DNA sensor (ZF sensor) consists of a pair of ZF domains each N- or C-terminally linked to Vma N-terminal or C-terminal splits which are, respectively, attached to the ZF9 domain or VP64 activator components of the trans-spliced response activator (TSRA) of the system.[5] The binding of the ZF pair to proximal target sequences induces trans-splicing release of ZF9-VP64 TSRA that, in the presence of a reporter circuit, with ZF9-binding sites placed before a minimal promoter and GFP, activates GFP expression in cells containing the DNA sequence target.[5]

The ZF sensor system relies however on a cumbersome assembly of ZF pairs of DNA binding modules, hampering its general application to other sequence targets. Although there are platforms to help in the design and generation of ZF modules, validation of the ZF domains requires screening for specificity and affinity using *in vitro* and *in vivo* approaches.[7] We therefore sought to overcome ZF-assembly limitations and achieve an improved response of the DNA sensor to the detection of a DNA sequence in live cells.

Here we simplify engineering of DNA-sensors by replacing the ZF DNA-binding modules with the transcription activator-

**Received:** May 11, 2021
**Published:** December 21, 2021




https://doi.org/10.1021/acssynbio.1c00212
*ACS Synth. Biol.* 2022, 11, 116–124

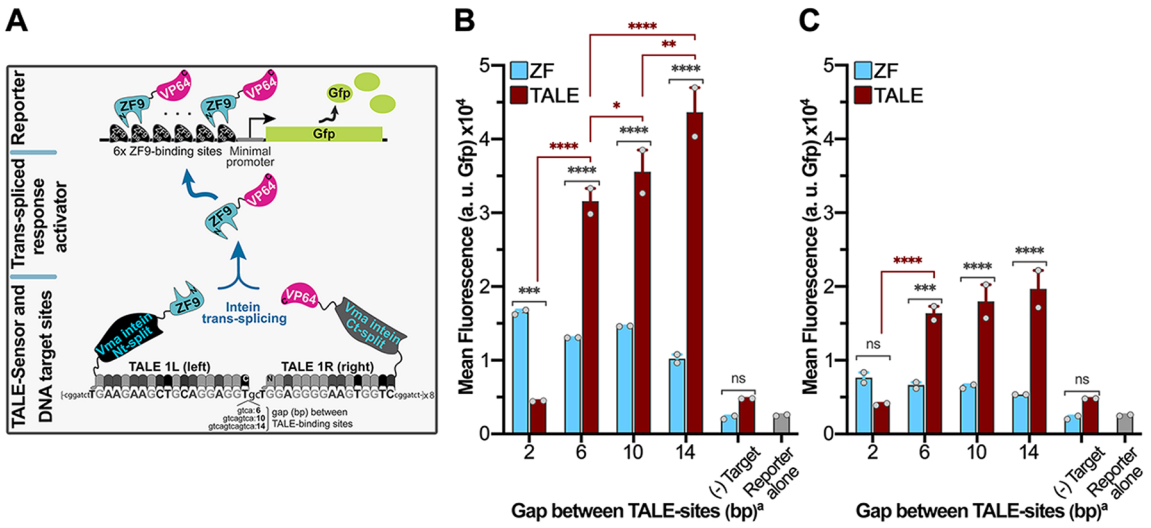



**Figure 1.** TALE.Sense system. (A) Schematic representation of TALE.Sense, a three-component system, showing TALE-Sensor, DNA target sequence1, and reporter circuit. The ZF9-VP64 trans-spliced activator (TSRA) product of proximity-mediated Vma trans-splicing which is conditioned by the binding of TALE 1L and TALE 1R arrays to their respective left and right sites that are separated by the indicated spacing. To meet TALE design requirement for a 5′ thymine preceding target sites, the TALE 1R site is 2 bp shorter than the ZF counterpart. The reporter circuit is based on six ZF9-binding sites as an operator for a minimal promoter driving TSRA-dependent GFP expression. Size, shape, amino (N) and carboxyl (C) termini of protein fusions and occupancy of the ZF9-binding sites are arbitrarily depicted. (B,C) Plots showing mean GFP-fluorescence intensity in arbitrary unit (au) from flow cytometry analysis of HEK293T cells transfected with ZF sensor or TALE-Sensor in the presence of plasmids encoding GFP-based reporter circuit or carrying target sequences when indicated. (−) Target represents empty vector that lacks the targeted DNA sequences. Replicative (with SV40-ori) (B), or nonreplicative (C) plasmid templates containing eight copies of target sequences for ZF sensor were used. TALE 1L and TALE 1R, DNA-binding modules of the TALE-Sensor, were programmed to bind to sequence targets reported for the ZF sensor. Means of two independent transfections (biological replicates) ± SEM are shown, and two-way ANOVA was performed (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p < 0.05$). [a]For ZF sensor, the actual gap size separating ZF-binding sites is 2 bp shorter than shown.

like effector (TALE) from *Xanthomonas sp.*[8,9] (Figure 1A). Advances in programming DNA-binding proteins using TALE arrays allow easier assembly of highly specific DNA-binding modules.[10−13] The TALE-based DNA sensor platform (TALE.Sense) shows higher efficiency and dynamic range when compared to the ZF sensor in detecting the same DNA target sequences. Swapping transcriptional activation domains, as well as introducing SunTag-based amplification loops to TALE.Sense circuits, further augments the detection efficiency of the DNA sensor. The TALE.Sense platform also shows an increased versatility when tested on a range of target sites and different cell types, thereby promising to be easily applied to the study of relevant problems requiring response circuits to distinguish cells containing predefined DNA sequences. TALE.Sense could be used as control switches that accept DNA sequences as input, expanding the scope of connectivity in synthetic gene circuits.

## RESULTS AND DISCUSSION

**TALE-Based DNA Sensor System—TALE.Sense.** To simplify the construction of DNA sensors inspired by the previous reported ZF-based sense-and-respond system,[5] we replaced the ZF domains responsible for target detection with TALE arrays (Figure 1A). We assembled a TALE pair, (TALE 1L, TALE 1R), to bind to the DNA sequence targets of the ZF sensor. The resulting TALE-Sensor and the reported ZF sensor were then compared for their efficiency to activate GFP expression when their three-plasmid components that contain a target DNA sequence, encode the two protein fusions of corresponding DNA sensor, or carry a GFP-based reporter circuit were introduced to HEK293T cells (Figure 1A). The DNA-binding sites for each arm of the DNA-sensors were separated by gaps of varying length on replicative (with SV40-ori) or nonreplicative plasmids as previously reported[5] (Figure 1A). Varying the spacing of DNA-binding sites allows for assessing proximity conditions for Vma splits protein fusions to achieve an efficient reconstitution of Vma activity. Flow cytometry analysis of the transfected cells showed substantial GFP fluorescence accumulation only when a DNA target was added, indicating detection of the DNA sequences by ZF sensor and TALE-Sensor (Figure 1B,C). However, TALE-Sensor showed significantly higher GFP expression with both replicative (Figure 1B) and nonreplicative (Figure 1C) plasmid templates containing DNA sequence targets, indicating that TALE modules of the sensor enabled improved efficiency to the system in detecting intended DNA sequences. Consistent with previous reports, the best signal observed with the ZF sensor was limited to DNA sequences with minimal spacing between ZF-binding sites. In contrast, TALE-Sensor produced a GFP signal that significantly increased as the spacing between TALE-binding sites within sequence targets increased, suggesting that the TALE-Sensor platform offers more flexibility and tunability in detecting DNA sequences in live cells. The higher GFP expression mediated by TALE.Sense was obtained with a range of target DNA concentrations and was not associated with differences in efficiency of transfection with components of the systems (Supplemental Figure 1A−C). Although other explanations may exist, the augmented efficiency observed with TALE-Sensor could be inherently associated with DNA-binding kinetics of TALE domains and/or a more favorable conformation of the TALE-Sensor, potentially leading to an increased TSRA yield by reconstituted


https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

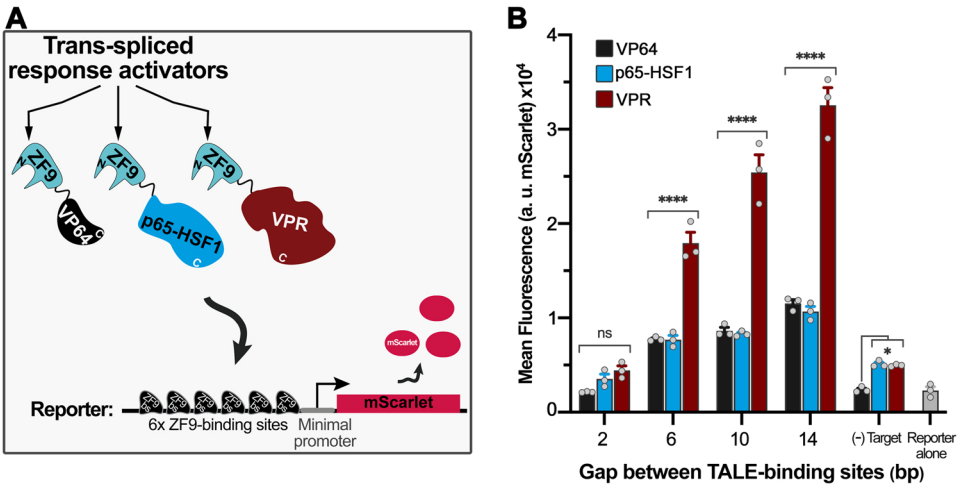



**Figure 2.** TALE.Sense with different transcription activation domains- (A) Schematic representation of TSRA generated upon DNA detection by TALE-Sensors that are based on VP64, p65-HSF1 or VPR activation domains. Size, shape, amino (N) and carboxyl (C) termini of protein fusions and reporter circuit are arbitrarily shown. (B) Comparison of VP64, p65-HSF1 and VPR-based TALE-Sensors. Mean mScarlet-fluorescence intensity in a. u. from flow cytometry analysis of HEK293T cells transfected with the indicated TALE-Sensors in the presence of mScarlet-based reporter circuit and replicative plasmids carrying target sequence1 with varying spacing (gap) between binding sites as indicated. Means of three biological replicates ± SEM are shown, and two-way ANOVA was performed (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$ and * $p < 0.05$).

Vma at target DNA sites. Western blot analysis showed a somewhat improved splicing activity with TALE.Sense as suggested by the ratios of a released product to precursor (Supplemental Figure 1D).

**TALE-Sensors with Different Transcription Activator Domains.** Recombinant transcription activator domains such as VP64, VPR, and p65-HSF1 were used to generate artificial transcription factors (ATF) that when introduced to cells gave varied amplitude in transcriptional activations depending on gene targets and delivery systems.[14−16] We therefore asked whether an improved output signal could be obtained upon DNA detection when the VP64 component of the TALE-Sensor is replaced by p65-HSF1 or VPR (Figure 2A). Transfection of HEK293T cells with each of the TALE-Sensor variants, producing trans-spliced response activators comprising p65-HSF1, VPR, or VP64, in the presence of DNA targets followed by flow cytometry analysis, showed that the VPR-based TALE-Sensor gave significantly higher mScarlet-fluorescence signal compared to VP64 and p65-HSF1 (Figure 2B). However, no significant changes were observed between p65-HSF1 and VP64-based TALE-Sensors. Transfections without target sequences gave background levels with the TALE-Sensor variants. The higher amplitude in response obtained with the VPR-based TALE-Sensor is consistent with the reported relative superiority of VPR-based ATF in mediating gene activations.[16] The same trend was obtained when VP64 and VPR-based TALE-Sensors were compared against a range of DNA target amounts in the presence of a GFP-based reporter circuit (Supplemental Figure S2A). Consistent with the higher fluorescence accumulation mediated by the VPR-based TALE-Sensor, Western blot analysis seems to suggest a higher splicing activity with VPR-based TALE.Sense as indicated by the improved ratio of a released product to precursor (Supplemental Figure S2B). Thus, swapping activation domains of the TALE-Sensor improved significantly the efficiency in producing a cellular response triggered by detection of the targeted DNA sequences.

**Coupling DNA Sequence Recognition to Cell Ablation.** Sensing of a specific DNA sequence could act as an ON-switch for targeted execution of cellular programs. Gene-directed enzyme prodrug therapy is an area of interest where metabolic conversion of a nontoxic prodrug to cytotoxic derivative leads to the killing of unwanted cells.[17] The NTR/CB1954 pair for example, is an enzyme/prodrug where the *E. coli* nitroreductase (NTR) mediates reduction of CB1954 to produce the metabolite that causes DNA interstrand cross-linking leading to apoptotic cellular response and cell-death.[18,19] To evaluate TALE.Sense in mediating DNA sequence-induced cellular ablation, we altered the reporter circuit to introduce a suicide gene encoding mScarlet-NTR protein fusion to enable dual detection of fluorescence signal and induced apoptotic response to prodrug (Figure 3A). CB1954-treatment of cells transfected with TALE-Sensor components, followed by MTT proliferation assay, showed significant cell death when the target DNA sequence was present (Figure 3B). Cells transfected with mScarlet-NTR response plasmid in the absence of TALE-Sensor, did not show any significant cell death in response to prodrug treatment, indicating that activation of NTR expression triggered by sequence detection of the TALE-Sensor is required for the observed cell death. The induced cell death was markedly higher with VPR as compared with VP64-based TALE-Sensor consistent with a VPR-driven higher gene activation[14,16] (Figure 3B). In the absence of target sequence and at high prodrug concentration, cells transfected with the VPR-based TALE-Sensor manifested low frequency of cell death attributable to a leaky effect amplified by the higher transactivation activity of VPR, as no such effect was observed with the VP64-based TALE-Sensor (Figure 3C). Flow cytometry and fluorescence microscopy analyses of transfected cells without CB1954-treatment showed the same trend obtained with the mScarlet response circuit, with VPR producing higher output signal than VP64 (Figure 3D,E). The TALE.Sense platform could potentially serve as a basis for developing tools enabling targeted cell ablation when


https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

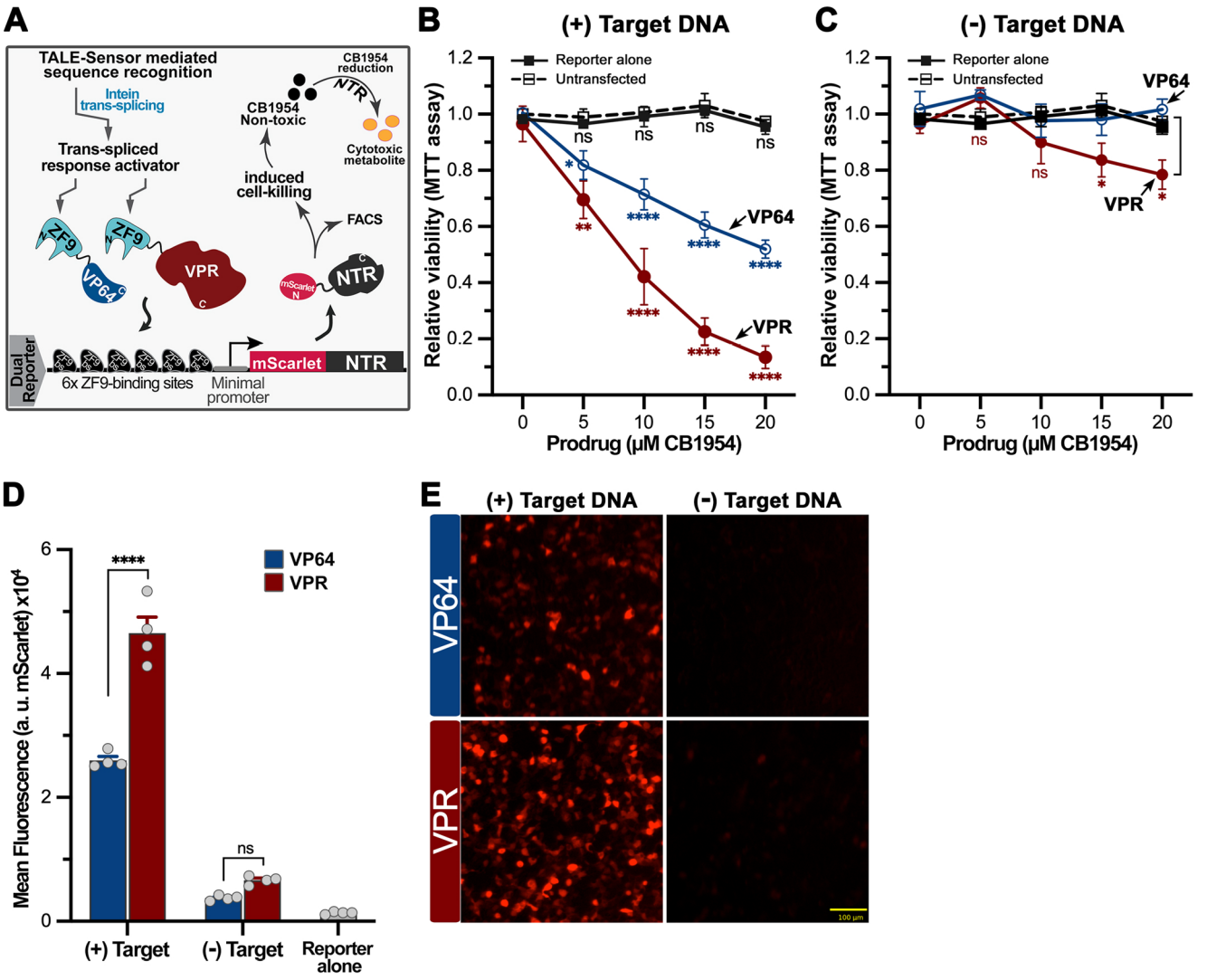



**Figure 3.** Induced cell ablation by DNA sequence recognition. (A) Drawing showing dual circuit for TALE.Sense enabling fluorescence and induced cell death mediated by expression of mScarlet-NTR fusion upon DNA recognition. TSRA-based on VP64, or VPR domains used in panels B−E are shown. NTR-mediated conversion of nontoxic CB1954 prodrug to cytotoxic metabolite is depicted. Shape, size, amino (N), and carboxyl (C) termini of protein fusions and dual reporter circuit are arbitrarily shown. (B−C) Comparison of TALE.Sense systems with TSRA-based on VP64 or VPR for induced cell-killing triggered by sequence-detection in the presence of CB1954 prodrug. Relative viability from MTT assays of HEK293T cells transfected with the indicated TALE-Sensors in the presence of dual reporter circuit, and replicative plasmid carrying the 14 bp-gap target sequence1 (B), or empty vector (C). Arrows indicate activation domains components of the TALE-Sensors. Absorbances obtained in MTT assays were normalized to untransfected cells. Means of three biological replicates ± SEM are shown, and two-way ANOVA was performed. (D) Mean mScarlet-fluorescence intensity au from flow cytometry analysis of HEK293T cells transfected with TALE-Sensors as in panel B in the absence of CB1954 treatment. Means of four biological replicates ± SEM are shown, and two-way ANOVA was performed. (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p < 0.05$). (E) Micrographs showing mScarlet fluorescence in HEK293T cells analyzed in panel D. Scale bar = 100 $\mu$m.

unwanted cells containing specific DNA sequences such as those produced in cancer cells by aberrant chromosomal rearrangements, viral infections or genomic integrations. Off-target and bystander effects are potential problems that could however be addressed by fine-tuning the response signal through rational design of the TALE modules of DNA sensors and the target sequences, as well as designing prodrugs that convert to less-diffusible products.

**TALE-Sensor with an Amplification Loop.** Here we sought to augment the TALE.Sense by introducing a SunTag-based[20] amplification loop in the reporter circuit where activation by a TSRA is coupled to P2A-mediated release of an amplifier module (Figure 4A). Binding of the amplifier module to the GCN4-array appended to TSRA could enable synergetic activation of mScarlet expression resulting in an amplified response upon DNA detection (Figure 4A). To establish which of the TSRA−amplifier module pairs produce substantial synergistic response, we compared TALE-Sensors when VP64, p65-HSF1, or VPR were linked to either the TSRA or amplifier module (Figure 4A). Transfection of components of TALE.Sense with or without amplification loop into HEK293T cells followed by flow cytometry analysis showed up to 4-fold signal amplification upon detection of target sequence within replicative or nonreplicative plasmid templates (Figure 4B−D). VP64-based TSRA when paired with VPR-amplifier produced the highest synergetic fold


https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

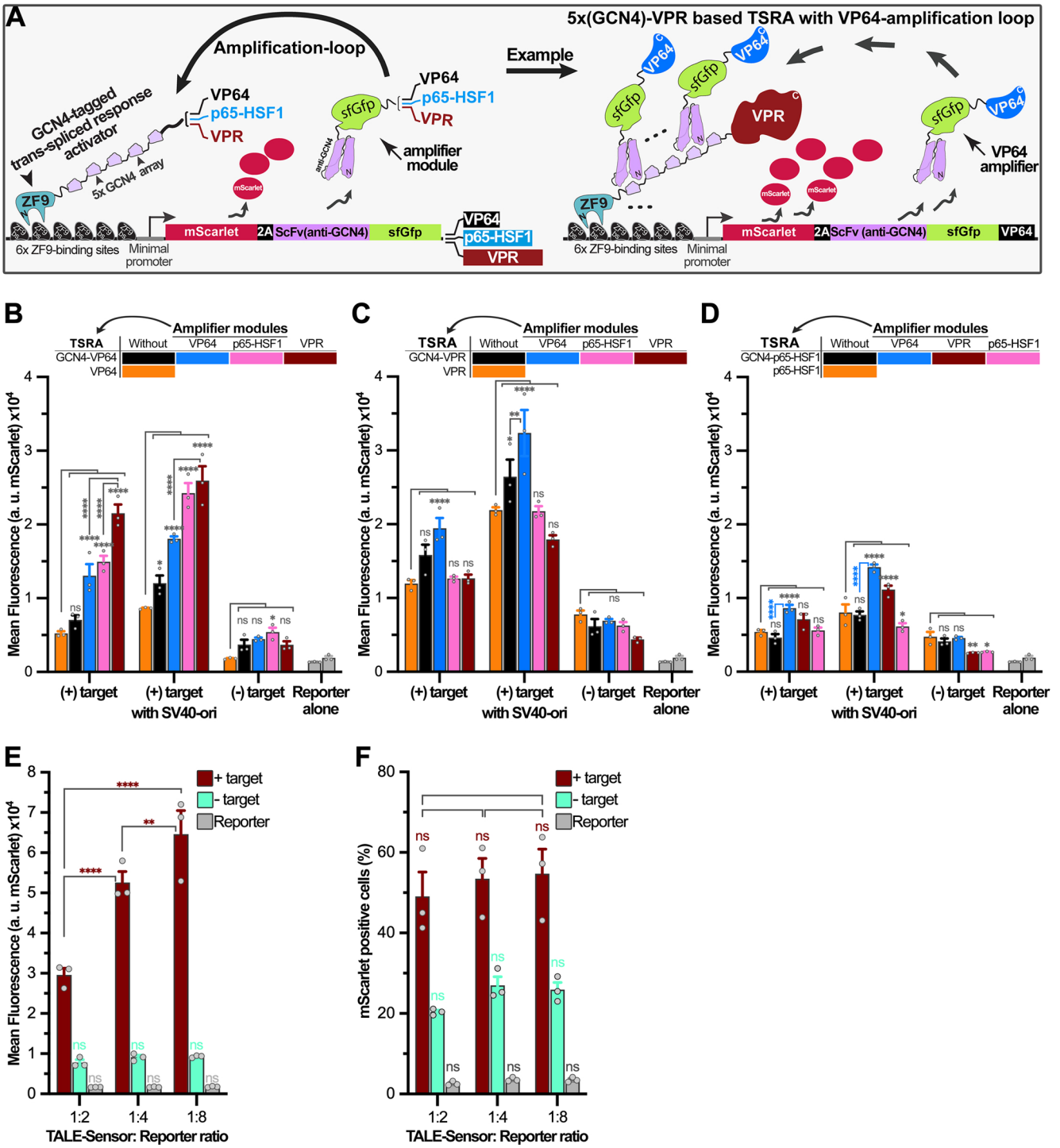



**Figure 4.** TALE.Sense with amplification circuit. (A) Drawing of TALE.Sense with amplification loop showing TSRA generated upon DNA detection by TALE-Sensors containing GCN4-tagged VP64, p65-HSF1, or VPR domains, and amplifier modules based on anti-GCN4 ScFv of the SunTag system linked to VP64, p65-HSF1, or VPR. TALE-Sensor with VPR-based TSRA and VP64-amplifier is illustrated as an example. Size, shape, amino (N) and carboxyl (C) termini of protein fusions, ZF9-binding sites and GCN4 occupancy, and reporter circuit are arbitrarily depicted. (B–D) Evaluation of amplification circuits of TALE.Sense when TSRA based on VP64 (B), VPR (C), or p65-HSF1 (D) and amplifier modules were linked to VP64, p65-HSF1, or VPR as depicted above the column plots. Plots showing mean mScarlet-fluorescence intensity in au from flow cytometry analysis of HEK293T cells transfected with the indicated TALE-Sensors and reporters, in the presence of replicative (with SV40-ori) or nonreplicative plasmids with the 10 bp-gap target sequence1, or empty vector. (E,F) Flow cytometry analysis of TALE.Sense transfected HEK293T cells by using varied plasmid ratios of VPR-based TALE-Sensor and reporter circuit containing VP64-amplifier, in the presence of a fixed amount of replicative plasmid carrying 14 bp-gap target sequence1, or empty vector. Plots show mean mScarlet-fluorescence intensity in au (E) and frequency of mScarlet positive cells (F). Means of three biological replicates ± SEM are shown, and two-way ANOVA was performed (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p < 0.05$).



https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116–124

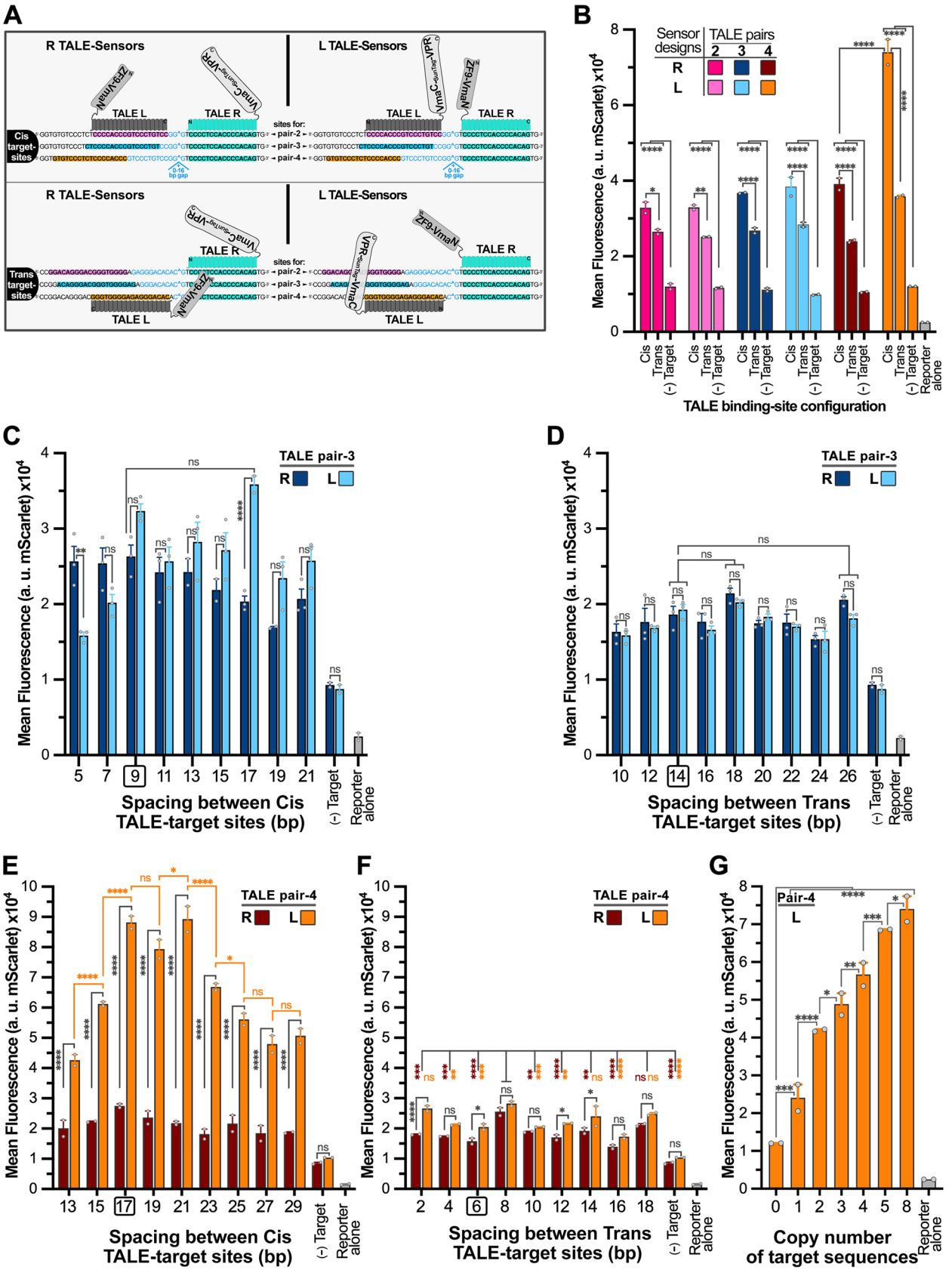



**Figure 5.** TALE-Sensors and target DNA configurations. (A) Illustration of TALE-Sensors design and their binding to DNA sequence targets. Binding-sites for TALE L (left) and TALE R (right) are shaded within the bipartite target sequence. Template sequence2 is shown as cis and trans configurations with their intercalating sequences and gap inserts (0−16 bp). L and R TALE-Sensor designs, when $Vma^{C}$-VPR component is linked to TALE L or TALE R arrays, respectively, and TALE-pair modules for each TALE-Sensor are illustrated. Size, conformation, amino (N) and carboxyl (C) termini of protein fusions are arbitrarily shown. (B) L and R TALE-Sensors targeting DNA binding-sites in cis or trans configurations.



https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

**Figure** 5. continued

TALE-Sensors are based on the TALE-pairs as depicted in panel A. Mean mScarlet-fluorescence intensity in au from flow cytometry analysis of HEK293T cells transfected with the indicated TALE-Sensors that are based on VPR-TSRA and VP64-amplifier, in the presence of replicative plasmids containing eight copies of cis or trans-configured target sequence2 with 4 bp gap insert, or empty vector that lacked the target sequence. (C−F) Spacing determinant of TALE-binding sites. Mean mScarlet-fluorescence intensity in au from flow cytometry analysis of HEK293T cells transfected with the indicated TALE-Sensors that are based on VPR-TSRA and VP64-amplifier, in the presence of replicative plasmids containing five copies of cis (C, E) or trans (D, F) target sequences with varying size of the gap insert, or empty vector. Empty squares below the $x$ axis of panels C−E represent the spacing match with the DNA targets in panel (A) for the corresponding TALE-Sensors. Means ± SEM are shown, and two-way ANOVA was performed (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p < 0.05$). (G) Target sequence copy number determinant. Flow cytometry analysis of HEK293T cells transfected with L TALE-Sensor, based on TALE pair-4 and VPR-TSRA/VP64-amplifier, in the presence of plasmids containing 0−8 copies of the target sequence with cis configuration and 17 pb spacing between the binding-sites. Means of mScarlet-fluorescence intensity ± SEM are shown, and two-way ANOVA was performed (ns: $p \geq 0.05$, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p < 0.05$).

change, whereas the reciprocal TSRA−amplifier pair gave the highest signal intensity as TSRA without amplifier module showed superiority with VPR over VP64 activation domains (Figure 4B,C). p65-HSF1- and VPR-based amplifiers when paired with TSRA containing VPR or p65-HSF1 produced weak or no amplification probably due to a steric hindrance associated with the relatively larger size of the associated activation domains (Figure 4C,D). Evidence indicating that the system could be applied to other cell types was obtained when K562 and PC3 cells were transfected with TALE.Sense components (Supplemental Figure 3). Taken together the data indicate that an augmented efficiency in producing a cellular response triggered by DNA sequence detection can be obtained by equipping TALE.Sense platform with an amplification loop enabling synergistic transcriptional activation by combined recruitment of VP64 and VPR domains to the promoter of the reporter gene.

To optimize response of the VPR-based TALE-Sensor with VP64 amplification loop, we transfected cells with varying TALE-Sensor to reporter plasmid ratios in the presence of a fixed amount of target-sequence plasmid. Flow cytometry analysis showed that the system performed better at higher ratios as significant increase in mScarlet signal was observed when compared to that of lower plasmid ratios (Figure 4E). In contrast, the background signal obtained with a vector that lacked the target sequence did not show any significant changes (Figure 4E). The increase in output signal in response to TALE-Sensor/reporter ratios did not result from differences in transfection efficiency as no significant changes in the frequencies of mScarlet positive cells were obtained (Figure 4F). Thus, the presence in excess of reporter circuit improved the response of the TALE.Sense system to DNA sequence input.

**TALE-Sensors and Target DNA Configurations.** To determine the requirements for designing TALE-Sensors and selecting their targeted DNA sequences, we assembled and validated four TALE arrays targeting three bipartite target sites, of which the left and right sides contain binding-sites for TALE L and TALE R of each TALE-pair, respectively (Figure 5A, Supplemental Figure 4). TALE-Sensors can be made as L or R configuration depending on whether TALE L or TALE R module is linked to the VmaC-VPR arm of the TALE-Sensor, respectively (Figure 5A). Also, target sites with cis or trans configurations can be encountered as the binding-sites since a TALE-pair can occupy the same or separate strands of the DNA target, respectively (Figure 5A). We therefore compared three TALE-Sensors under L or R configurations against cis or trans DNA target sites, in the presence of a reporter circuit with VP64-based amplifier. We found for the target DNA configuration, cis showed significantly stronger response than trans as indicated by flow cytometry analysis of cells transfected with each of the six TALE-Sensors (Figure 5A,B). For L versus R design of the TALE-Sensors, no significant differences were observed between TALE-Sensors based on TALE pair-2 and pair-3. However, TALE pair-4 showed a much stronger response with L compared to R configuration of the TALE-Sensor in flow cytometry analysis, although the R sensor gave comparable results to other similarly configured DNA sensors from TALE pair-2 and pair-3 (Figure 5A,B). Thus, the intensity of response signal from an R or L-configured TALE-Sensor could considerably vary depending on intrinsic properties of the TALE arrays, suggesting that optimal outcomes would require comparing both configurations of the TALE-Sensors.

Because DNA sensor system is based on proximity mediated Vma trans-splicing, we asked whether the obtained trend in response efficiencies with TALE-Sensors and target sites configurations came from the inherent difference in spacing between the Vma splits within the DNA-bound TALE-Sensors. We thus compared TALE-Sensors based on TALE pair-3 and pair-4 on target sequences with varying spacing separating the TALE-binding sites (Figure 5A). This showed that the trend observed in Figure 5B for TALE-Sensors and target sites configurations was maintained for most of the target DNA sequences independently of the spacing span of the TALE-binding sites (Figure 5C−F). However, spacing variation was associated with some increases in the intensity of mScarlet signal with substantial changes observed with L TALE-Sensor from TALE pair-4 and cis target configuration (Figure 5E). Thus, the optimal spacing of DNA-binding sites can significantly vary between TALE-pairs as their corresponding Vma fusions may adopt different conformations/topologies resulting in different requirements for optimal proximity to efficiently reconstitute intein activity.

To determine the detection sensitivity of the TALE.Sense system for copy number of target sequences, we assessed L TALE-Sensor from pair-4 against plasmid templates containing 0−8 copies of the target sequence. Flow cytometry analysis showed the detection range of the system with fluorescence accumulation increasing with a higher copy number of the target sequence. This also showed that the system discerned cells containing a template with a single copy of the target sequence (Figure 5G). The data show the versatility of the TALE.Sense system in detecting DNA sequence targets and could serve as a reference point to designing TALE-Sensors to target other DNA sequences. The use of TALE arrays as DNA


https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

binding modules for DNA-sensors could have potential limitations as TALE-binding requires a 5′ thymine preceding the target sequences.[8,21,22] However, evolved TALE surrogates whose N-terminal domains acquired mutations enable unconstrained TALE-based targeting of any DNA sequence.[22]

## CONCLUSIONS

Here we repurposed TALE arrays to construct sequence-specific DNA sensors for live cells. We coupled the TALE-mediated programmable DNA binding activity to conditional intein trans-splicing to enable execution of customized cellular responses facilitating identification of live-cell sequence variants. TALE.Sense is not only relatively more practical to implement, but it also enables DNA detection with higher efficiency than previously reported ZF-based DNA sensor,[5] as indicated by the augmented range in the obtained response signals. TALE.Sense in contrast to the ZF-sensor enables increased flexibility and tunability in targeting DNA sequences with varying spacing between DNA binding sites, providing hence a handle to tune the strength of a response signal upon sequence detection in relevant applications. Introduction of an amplification loop to TALE.Sense circuit enables synergistic gene activation that further enhanced the cellular response upon detection of DNA sequences. The TALE.Sense could serve as a toolbox for detection of live-cell sequence variants in research investigations and therapy developments, as well as application of DNA sequences as control switches in synthetic biology.

## METHODS

**Cell Culture and Transfection.** HEK293T cells obtained from ATCC were cultivated in Dulbecco's modified Eagle's medium (Sigma) with 10% fetal bovine serum (FBS)(Lonza), 4% Glutamax (Gibco), 1% sodium pyruvate (Gibco), and penicillin-streptomycin (Gibco) in a humidified incubator set to 37 °C and 5% $CO_2$. K562 cells were obtained from ATCC and cultivated in Iscove's Dulbecco's modified medium (ThermoFisher) supplemented with L-glutamine (ATCC), 10% FBS, and 0.1 mM MEM nonessential amino acids (ThermoFisher). PC3 cells were a gift from Dr. Roel Verhaak laboratory and cultivated in F-12K (ATCC) supplemented with 10% FBS. Cells were seeded into 96-well 1 day before being transfected with three plasmids components of the DNA-sensor system including TALE-Sensor or ZF sensor (75 ng), reporter (150 ng), and target sequence templates (75 ng) in the presence of Lipofectamine 3000 transfection reagents (HEK293T and PC3) or Lipofectamine 2000 (K562) according to manufacturer's instructions (ThermoFisher). An empty vector was used as a template in experiments without DNA target sequence. The amounts of plasmid DNA in transfection mixes were doubled when cells were seeded into 24-well plates. The plasmid DNA amounts and ratios in transfection mixes were determined in preliminary experiments using the ZF sensor components and were used throughout this study, except when otherwise indicated.

**Plasmid Constructions.** A plasmid list with links to their Addgene entries are provided in supplemental Table 1. The plasmid sequences are available upon request. Description and protein sequences for relevant plasmids are given in supplemental Table 2. The assembly of TALE arrays was performed by using Golden Gate assembly plasmid kit (gift from Daniel Voytas and Adam Bogdanove, Addgene kit #100000024).[10] DNA sequence targets for TALE-arrays are listed in supplemental Table 3. The plasmids pCAG-dCas9-5xPlat2AflD and pCAG-scFvGCN4sfGFPTET1CD used as templates for PCR amplification of sequence for components of the SunTag system were gifts from Izuho Hatada (Addgene no. 82560 and 82561, respectively).[20] The plasmids pVITRO1-SS-269, pGL4.26-SS-192, pBW121-SS-309, pBW121-SS-310, pBW121-SS-287, pBW121-SS-311 and pBW121-SS-315 (Addgene no. 68771, 68759, 68777, 68778, 68779, 68780, and 68786, respectively) were gifts from James Collins.[5]

**Cell Proliferation Assays.** Cell were transfected in 24-well plates followed by medium changes after 24-h using CB1954-supplemented media as indicated. The cells were incubated for 96 h then subjected to MTT assay according to manufacturer's instructions (Thermo Fisher Scientific). Absorbance at 540 nm was recorded using a SpectraMax M5 plate reader (Molecular devices).

**Flow Cytometry, and Fluorescence Microscopy.** For flow cytometry, cells were collected 48 h after transfection. Cells were treated with trypsin-EDTA, suspended in the culture medium, and then analyzed on FACSymphony A5 flow cytometer (BD Bioscience) using a high-throughput plate sampler and FACSDiVA software (BD Bioscience). Twenty thousand events were collected in each run and untransfected cells were used to set the gates. For fluorescence microscopy, images were taken 24 h after transfection with the iRiS Digital Cell Imaging System (Logos Biosystems).

**Western Blot.** Protein extracts from transfected cells (40 $\mu$g) were separated by electrophoresis on 4−20% SDS-polyacrylamide gels (BioRad) and then transferred to nitrocellulose membranes at 100 V for 45 min using Bjerrum Schafer-Nielsen buffer with SDS. Blocked membrane in 5% Blotting-Grade Blocker (BioRad) in TBS-T (50 mM Tris pH 7.6, 200 mM NaCl, 0.1% Tween 20) were incubated overnight at 4 °C with the indicated antibodies, and then protein bands were detected using horseradish peroxidase-conjugated secondary antibodies (Sigma) and Clarity Western ECL Substrate according to manufacturer's instructions (BioRad). Monoclonal anti-HA (dilution 1:1000, Sigma cat. no. H3663) and monoclonal anti-$\beta$-actin (dilution 1:5000, Sigma cat. no. MAB 1501) antibodies were used according to manufacturers' instructions. Blots were imaged using a G:Box (Syngene).

**Statistical Analyses.** Information on replication, statistical tests, and presentation are given in the respective figure legends. GraphPad Prism 9 was used to perform the indicated tests. Differences in all comparisons were considered significant if the obtained $P$ values were less than 0.05.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.1c00212.

TALE.Sense performance under different amounts of target DNA sequence, Western blot analysis, evaluation of VP46 and VPR-based TALE.Sense when different amounts of target DNA sequence and GFP-reporter circuit were used, TALE.Sense in other cell types, and functional validation of TALE-arrays; plasmid list and sequence information (PDF)



https://doi.org/10.1021/acssynbio.1c00212
ACS Synth. Biol. 2022, 11, 116−124

**Special Issue Paper**
Invited contribution from the 12th International Workshop on Bio-Design Automation.

## AUTHOR INFORMATION

### Corresponding Authors

**Albert W. Cheng** − *The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut 06032, United States; The Jackson Laboratory Cancer Center, Bar Harbor, Maine 04609, United States; Department of Genetics and Genome Sciences, University of Connecticut Health, Farmington, Connecticut 06030, United States; Institute for Systems Genomics, UConn Health Science Center, Farmington, Connecticut 06030, United States;* Present Address: School of Biological and Health Systems Engineering, Arizona State University, Tempe, Arizona 85281, United States; orcid.org/0000-0003-2653-5743; Email: albert@cheng.bio

**Aziz Taghbalout** − *The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut 06032, United States*; Email: aziz.taghbalout@jax.org

### Author

**Nathaniel Jillette** − *The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut 06032, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.1c00212

### Author Contributions

A.T. and A.C. conceived and designed the study. A.T. and N.J. performed the experiments. A.T. and A.C. analyzed data and wrote the manuscript.

### Notes

The authors declare the following competing financial interest(s): A.W.C., A.T., and N.J. filed PCT patent application (WO/2019/090287) for the invention.

## ACKNOWLEDGMENTS

This work has been supported by internal funds provided by the Jackson Laboratory as well as grants from the National Cancer Institute (P30CA034196) and the National Human Genome Research Institute (R01-HG009900).

## REFERENCES

(1) Shah, N. H.; Muir, T. W. Inteins: Nature's Gift to Protein Chemists. *Chem. Sci.* **2014**, *5* (1), 446−461.
(2) Iwai, H.; Zuger, S.; Jin, J.; Tam, P. H. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS Lett.* **2006**, *580* (7), 1853−8.
(3) Mootz, H. D.; Muir, T. W. Protein splicing triggered by a small molecule. *J. Am. Chem. Soc.* **2002**, *124* (31), 9044−5.
(4) Brenzel, S.; Kurpiers, T.; Mootz, H. D. Engineering artificially split inteins for applications in protein chemistry: biochemical characterization of the split Ssp DnaB intein and comparison to the split Sce VMA intein. *Biochemistry* **2006**, *45* (6), 1571−8.
(5) Slomovic, S.; Collins, J. J. DNA sense-and-respond protein modules for mammalian cells. *Nat. Methods* **2015**, *12* (11), 1085−90.
(6) Huang, X.; Narayanaswamy, R.; Fenn, K.; Szpakowski, S.; Sasaki, C.; Costa, J.; Blancafort, P.; Lizardi, P. M. Sequence-specific biosensors report drug-induced changes in epigenetic silencing in living cells. *DNA Cell Biol.* **2012**, *31* (Suppl1), S2−10.
(7) Hossain, M. A.; Barrow, J. J.; Shen, Y.; Haq, M. I.; Bungert, J. Artificial zinc finger DNA binding domains: versatile tools for genome engineering and modulation of gene expression. *J. Cell. Biochem.* **2015**, *116* (11), 2435−44.
(8) Boch, J.; Scholze, H.; Schornack, S.; Landgraf, A.; Hahn, S.; Kay, S.; Lahaye, T.; Nickstadt, A.; Bonas, U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **2009**, *326* (5959), 1509−12.
(9) Moscou, M. J.; Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **2009**, *326* (5959), 1501.
(10) Cermak, T.; Doyle, E. L.; Christian, M.; Wang, L.; Zhang, Y.; Schmidt, C.; Baller, J. A.; Somia, N. V.; Bogdanove, A. J.; Voytas, D. F. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **2011**, *39* (12), e82.
(11) Zhang, S.; Chen, H.; Wang, J. Generate TALE/TALEN as Easily and Rapidly as Generating CRISPR. *Mol. Ther.–Methods Clin. Dev.* **2019**, *13*, 310−320.
(12) Ma, A. C.; McNulty, M. S.; Poshusta, T. L.; Campbell, J. M.; Martinez-Galvez, G.; Argue, D. P.; Lee, H. B.; Urban, M. D.; Bullard, C. E.; Blackburn, P. R.; Man, T. K.; Clark, K. J.; Ekker, S. C. FusX: A Rapid One-Step Transcription Activator-Like Effector Assembly System for Genome Science. *Hum. Gene Ther.* **2016**, *27* (6), 451−63.
(13) Ding, Q.; Lee, Y. K.; Schaefer, E. A.; Peters, D. T.; Veres, A.; Kim, K.; Kuperwasser, N.; Motola, D. L.; Meissner, T. B.; Hendriks, W. T.; Trevisan, M.; Gupta, R. M.; Moisan, A.; Banks, E.; Friesen, M.; Schinzel, R. T.; Xia, F.; Tang, A.; Xia, Y.; Figueroa, E.; Wann, A.; Ahfeldt, T.; Daheron, L.; Zhang, F.; Rubin, L. L.; Peng, L. F.; Chung, R. T.; Musunuru, K.; Cowan, C. A. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* **2013**, *12* (2), 238−51.
(14) Chavez, A.; Scheiman, J.; Vora, S.; Pruitt, B. W; Tuttle, M.; P R Iyer, E.; Lin, S.; Kiani, S.; Guzman, C. D; Wiegand, D. J; Ter-Ovanesyan, D.; Braff, J. L; Davidsohn, N.; Housden, B. E; Perrimon, N.; Weiss, R.; Aach, J.; Collins, J. J; Church, G. M Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **2015**, *12* (4), 326−8.
(15) Casas-Mollano, J. A.; Zinselmeier, M. H.; Erickson, S. E.; Smanski, M. J. CRISPR-Cas Activators for Engineering Gene Expression in Higher Eukaryotes. *CRISPR J.* **2020**, *3* (5), 350−364.
(16) Chavez, A.; Tuttle, M.; Pruitt, B. W.; Ewen-Campen, B.; Chari, R.; Ter-Ovanesyan, D.; Haque, S. J.; Cecchi, R. J.; Kowal, E. J. K.; Buchthal, J.; Housden, B. E.; Perrimon, N.; Collins, J. J.; Church, G. Comparison of Cas9 activators in multiple species. *Nat. Methods* **2016**, *13* (7), 563−567.
(17) Zhang, J.; Kale, V.; Chen, M. Gene-directed enzyme prodrug therapy. *AAPS J.* **2015**, *17* (1), 102−10.
(18) Knox, R. J.; Friedlos, F.; Boland, M. P. The bioactivation of CB 1954 and its use as a prodrug in antibody-directed enzyme prodrug therapy (ADEPT). *Cancer Metastasis Rev.* **1993**, *12* (2), 195−212.
(19) Grohmann, M.; Paulmann, N.; Fleischhauer, S.; Vowinckel, J.; Priller, J.; Walther, D. J. A mammalianized synthetic nitroreductase gene for high-level expression. *BMC Cancer* **2009**, *9*, 301.
(20) Morita, S.; Noguchi, H.; Horii, T.; Nakabayashi, K.; Kimura, M.; Okamura, K.; Sakai, A.; Nakashima, H.; Hata, K.; Nakashima, K.; Hatada, I. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nat. Biotechnol.* **2016**, *34*, 1060.
(21) Bogdanove, A. J.; Voytas, D. F. TAL effectors: customizable proteins for DNA targeting. *Science* **2011**, *333* (6051), 1843−6.
(22) Lamb, B. M.; Mercer, A. C.; Barbas, C. F., 3rd Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. *Nucleic Acids Res.* **2013**, *41* (21), 9779−85.