Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation*[S]◆

Julien Valton[1,2], **Aurélie Dupuy**[1], **Fayza Daboussi, Séverine Thomas, Alan Maréchal, Rachel Macmaster, Kevin Melliand, Alexandre Juillerat, and Philippe Duchateau**[3]
From Cellectis S.A., 8 rue de la croix Jarry, 75013 Paris, France

Background: TALE-based technologies are poised to revolutionize the field of biotechnology; however, their sensitivity to cytosine methylation may drastically restrict their ranges of applications.
Results: TALE repeat N* proficiently accommodates 5-methylated cytosine.
Conclusion: Sensitivity of TALE to cytosine methylation can be overcome by using TALE repeat N*.
Significance: Utilization of TALE repeat N* enables broadening the scope of TALE-based technologies.

Within the past 2 years, transcription activator-like effector (TALE) DNA binding domains have emerged as the new generation of engineerable platform for production of custom DNA binding domains. However, their recently described sensitivity to cytosine methylation represents a major bottleneck for genome engineering applications. Using a combination of biochemical, structural, and cellular approaches, we were able to identify the molecular basis of such sensitivity and propose a simple, drug-free, and universal method to overcome it.

Transcription activator-like effectors (TALEs),[4] a group of bacterial plant pathogen proteins, have recently emerged as new engineerable scaffolds for production of tailored DNA binding domains with chosen specificities (1). Interest in these systems comes from the apparent simple cipher governing DNA recognition by their DNA binding domain (2, 3). The TALE DNA binding domain is composed of multiple TALE repeats that individually recognize one DNA base pair through specific amino acid di-residues (repeat variable di-residues or RVDs). The remarkably high specificity of TALE repeats and the apparent absence of context-dependent effects among repeats in an array allow modular assembly of TALE DNA binding domains able to recognize almost any DNA sequence of interest. Within the past 2 years, engineered TALE DNA binding domains have been fused to transcription activator (dTALEs) (4), repressor (5), or nuclease domains (TALENs) (6)

and used to specifically regulate or modify genes of interest (1). Although successfully used in different cellular contexts, engineered TALE DNA binding domains have recently been reported to be affected by the presence of 5-methylated cytosine (5mC) in their endogenous cognate target (7). Often considered as the fifth base, 5mC is found in about 70% of CpG dinucleotides in mammalian and plant somatic/pluripotent cells (8, 9) and has also been reported in 5-cytosine-phosphoadenine, 5-cytosine-phosphothymine, and 5-cytosine-phosphocytosine dinucleotides (10). Moreover, 5mC has been identified in CpG islands embedded in many promoters (11) and, to a higher extent, in proximal exons of several genes (12). These two critical regulatory regions are generally chosen by investigators to knock out genes of therapeutic and biotechnological interest or to modulate their expression using TALE-based technologies. The ubiquity of 5mC in different cell types and genomic kingdoms, its particular localization, and its negative impact on dTALE activity reported in Ref. 7 make this epigenetic modification a major drawback for all TALE-based technologies ranging from genome engineering, iPS production/differentiation, to therapeutic applications. Utilization of 5-aza-2-deoxycytidine (5-aza-dC) as a demethylating agent could be an option to overcome such a drawback as suggested earlier (7). However, its cytotoxicity and well known pleiotropic effect (13) make it unsuitable for precise therapeutic and biotechnological genome engineering applications. Therefore, alternative and harmless strategies are highly desired. One of the most relevant strategies would be to develop 5mC-insensitive TALE DNA binding domains. To date, the cipher governing TALE repeat/DNA base recognition has been thoroughly documented for the unmodified bases A, T, G, and C (2, 3). However, 5mC has not been reported to be specifically recognized by any of the naturally occurring TALE repeats described so far. Here we unravel the ability of TALE repeat N* to efficiently accommodate 5mC and describe a simple, drug-free, and universal approach to overcome TALE DNA binding domain sensitivity to cytosine methylation. Besides the important biotechnological prospects raised by our work, we provide the first example of TALEN-mediated processing of methylated *XPC* loci, paving the way for xeroderma pigmentosum-targeted gene therapy.

---

* Cellectis S. A. is commercializing TALENs.
◆ This article was selected as a Paper of the Week.
[S] This article contains supplemental Tables I–V and Figs. 1–7.
[1] Both authors contributed equally to this work.
[2] To whom correspondence may be addressed. E-mail: julien.valton@cellectis.com.
[3] To whom correspondence may be addressed. E-mail: philippe.duchateau@cellectis.com.
[4] The abbreviations used are: TALE, transcription activator-like effector; dTALE, designer TALE; TALEN, TALE nucleases; RVD, repeat variable di-residue; CpG, 5-cytosine-phosphoguanine; 5mC, 5-methylated cytosine; 5-aza-dC, 5-aza-2-deoxycytidine; FAM, 6-carboxyfluorescein; TM, targeted modifications; iPS, induced pluripotent stem cells; SSA, single strand annealing; TM, targeted mutagenesis; EndoT7, T7 endonuclease I; *XPC*, xeroderma pigmentosum group C; XPCT, XPC TALEN; XPCT1L, XPCT1 left TALEN.

ASBMB

EXPERIMENTAL PROCEDURES

TALEN Constructs—XPCT1, XPCT2, and XPCT3 TALENs and their respective variants were purchased from Cellectis Bioresearch (Paris, France). Their amino acid sequences and their cognate DNA target are documented in supplemental Table III.

Cell Culture and Transfections—Human 293H cells (Life Technologies) and hamster CHO-KI cells (ATCC) were cultured at 37 °C with 5% CO_2 in complete medium DMEM or F12-K respectively, supplemented with 2 mM L-glutamine, 100 IU/ml penicillin, 100 μg/ml streptomycin, 0.25 μg/ml amphotericin B (Fungizone, Life Technologies) and 10% FBS.

Monitoring TALEN Extrachromosomal SSA Activity—CHO-KI cells were plated at 2500 cells/well in a 96-well plate. The next day, cells were cotransfected by increasing amounts of DNA encoding XPC TALEN (from 0 to 25 ng each) and a constant amount of XPC extrachromosomal unmethylated target (75 ng) using PolyFect transfection reagent (Qiagen) according to the manufacturer's protocol. TALENs single strand annealing (SSA) activities were determined according to the protocol described in Ref. 14.

Monitoring of TALEN-induced Targeted Mutagenesis (TM)—To evaluate the ability of different XPC TALENs to induce TM at their endogenous loci, 293H cells were first plated at a density of 1.2×10^6 cells/10-cm dish. The next day, cells were transfected with a total amount of 2, 5, or 10 μg of TALEN expressing vector or empty vector using Lipofectamine 2000 transfection reagent (Life Technologies) according to the manufacturer's protocol. Two or 3 days after transfection, genomic DNA was extracted, and the loci of interest were amplified with locus-specific primers (supplemental Table I) linked to adaptor sequences needed for the deep sequencing method. Amplicons were analyzed either by EndoT7 assay according to the protocol described in Ref. 15 or by deep sequencing using the 454 system (Roche Applied Science, an average of 5000 sequences/sample were analyzed).

5-Aza-2-deoxycytidine Treatment, Bisulfite Treatment, and DNA Sequencing—To investigate the influence of 5-aza-dC on the methylation status of *XPC* loci and on XPC TALENs activities, 293H cells were pretreated every day and for 2 days with 0.2 μM 5-aza-dC before transfection, and this treatment was maintained up to 48 h after transfection. To determine the level of DNA methylation, genomic DNA was extracted, treated by bisulfite according to the manufacturer's protocol (EZ DNA Methylation-Gold kit, Zymo Research), and then amplified by PCR using primers that were specific for the flanking regions of the bisulfite-treated *XPC* loci (supplemental Table I). PCR amplicons were then analyzed by regular sequencing or deep sequencing using the 454 system.

Toxicity Assay—The CHO-KI cell line was transfected in a 96-well plate as described above, with increasing amounts of TALEN expression vectors and a constant amount of GFP-encoding plasmid. GFP levels were monitored by flow cytometry (Guava EasyCyte, Guava Technologies) 1 and 6 days after transfection. Cell survival was calculated as a ratio (TALEN-transfected cells expressing GFP at day 6/control-transfected cells expressing GFP at day 6). Ratios were corrected for the transfection efficiency determined at day 1 and plotted as a function of final concentration of DNA transfected (in mM).

TALE DNA Binding Domain Overexpression and Purification—The coding sequences for XPCT1L_HD and N* recombinant TALE DNA binding domains were subcloned into the kanamycin-resistant pET-24 vector MCS located upstream a His_6 tag coding sequence. Recombinant proteins were overexpressed and purified according to the protocol described previously (16). Freshly purified recombinant XPCT1 left TALEN (XPCT1L) DNA binding domains (80% homogeneity, ~2 mg/ml) were used to perform equilibrium binding assays.

In Vitro Binding Assay—3′-Carboxyfluorescein (FAM)-labeled DNA single strand oligonucleotide, corresponding to XPC1 left target (XPC1L), was synthesized and HPLC-purified by Eurogentec (supplemental Table II). To prepare DNA duplex corresponding to XPC1L double strand DNA, XPC1L_Forward labeled with FAM on its 3′ end was mixed with 1 eq of XPC1L_Reverse in 100 mM Tris-HCl, 50 mM EDTA, 150 mM NaCl, p H8. The mixture was heated to 95 °C for 2 min and then cooled down to 25 °C over 1 h. XPC1L duplex and FAM final concentrations were assessed by spectrophotometry using their respective extinction coefficients at 260 and 495 nm. As expected, a ratio [XPC1L duplex]/[FAM] ~1 was obtained. This procedure was used to prepare all other DNA targets used in our experiments. Oligonucleotide sequences are documented supplemental Table II. To investigate the binding of TALE DNA binding domain to XPCT1L duplex, 5 nM XPC1L duplex was incubated with increasing concentrations of XPCT1L TALE DNA binding domain (from 0 to 250 nM) in binding buffer (10 mM Tris-HCl, 175 mM NaCl, pH 8) at 25 °C. After a 5-h incubation, necessary to reach equilibrium, the fluorescence anisotropy of the mixture was recorded with a PHERAstar Plus (BMG Labtech) operating in fluorescence polarization end point mode with excitation and emission wavelengths set to 495 and 520 nm, respectively.

To avoid any experimental bias due to the presence of protein aggregates in our purified protein preparations and to compare binding isotherms obtained from one protein with those obtained from another, the fraction of active recombinant proteins (*i.e.* folded proteins that are competent for DNA binding) was determined using a stoichiometry binding assay (17). Using three different concentrations of XPC1L unmethylated DNA duplex (100, 200, and 300 nM), we were able to determine that the mean percentages of active recombinant XPCT1L-HD and N* proteins were 9 and 13%, respectively. Variations of fluorescence anisotropy were represented as a function of active XPCT1L-HD or N* concentrations. All binding isotherms displayed positive cooperativity ($[XPCT1L]_{0.9}/[XPCT1L]_{0.1}$ ~10) consistent with Ref. 18. Apparent dissociation constants (K_d) were determined with the Simfit/Sffit programs using a cooperative ligand binding saturation function (19).

FoldX Modeling of TALE Repeat HD and N in Complex with 5mC*—TALE repeat N* and HD tridimensional structures were obtained from PthXO TAL effector crystal structure (Protein Data Bank (PDB) ID 3UGM) (20). TALE repeat N* 7 and TALE repeat HD 8 in complex with their respective cytosines were used as templates to model methylation in position 5 of each
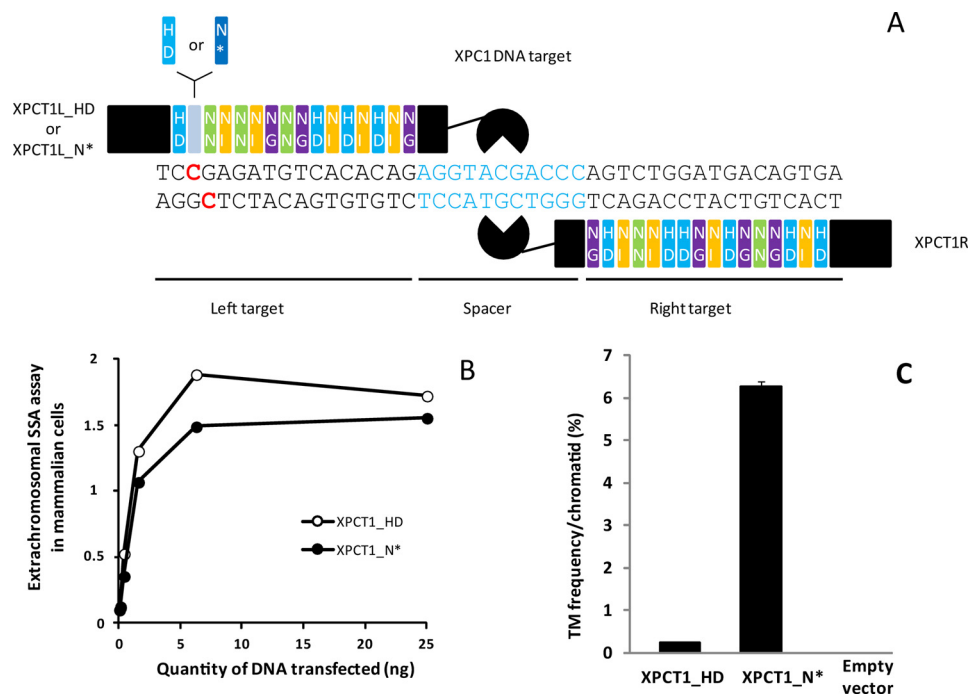
FIGURE 1. **TALE repeat N* enables overcoming XPCT1 TALEN sensitivity to 5mC in 293H cells.** *A*, schematic representation of the XPCT1 TALEN model used to investigate the influence of TALE repeat N* on TALE DNA binding domain sensitivity to 5mC. Sequence of the XPC1 DNA target is indicated; the 5mC located at position +2 of the left target is colored in *red*, and the 11-bp spacer between the right and left targets is colored in *blue*. TALE repeat array sequences bearing HD (*blue*), NI (*yellow*), NN (*green*), and NG (*purple*) and constituting the TALE DNA binding domains of XPCT1 left (XPCT1L-HD or N*) and right (XPCT1R) are indicated. N-terminal, C-terminal, and FokI domains are colored in *black*. *B*, intrinsic nuclease activity of XPCT1-HD or N* TALENs investigated by SSA assay using an extrachromosomal unmethylated XPC1 target (14). *C*, TM of endogenous methylated XPC1 target induced by 5 µg of XPCT1-HD- or N* TALEN-encoding plasmids in 293H cells, determined by deep sequencing (*n* = 2). *Error bars* indicate S.D.

cytosine, using the FoldX software. It should be noted that the FoldX software modeling algorithm did not allow for conformational change of TALE repeat backbone.

U2OS Methylome Survey—The U2OS methylome documented by Ohm *et al.* (21) under the sample and series names GSM739944 and GSE29872, respectively, was surveyed using a BLAST search to look for the TALEN DNA targets documented by Reyon *et al.* (15) and determine their methylation status. The β value associated with the retrieved sequences (hits) indicates their methylation status. A score of 1 is expected for full methylation, 0 for absence of methylation, and $0 \leq \beta \leq 1$ for all signals between (21).

RESULTS AND DISCUSSION

Although there was no prior evidence that the TALE repeat N* could bind 5mC, our interest in exploring this possibility was driven by its dual ability to bind cytosine and thymine with the same efficiency (2). Due to its missing glycine residue in position 13, the RVD loop of TALE repeat N* does not extend as deep into the DNA major groove as the other TALE repeats (20). Consistent with its apparent dual specificity, this aspect of its structure suggested that N* could avoid steric clashes with the additional methyl moiety present at position 5 of thymine and, by extension, potentially accommodate the methyl present in 5mC. This property was unlikely to be shared by TALE repeat HD, which displayed hydrophobic interactions and hydrogen bonding with the aromatic ring and the amine moieties of cytosine (20, 22). Therefore, as suggested by previous observations (7), TALE DNA binding domains bearing the TALE repeat HD

are likely to be sensitive to cytosine methylation, and we envisioned that substituting HD repeats by N* repeats could overcome methylation sensitivity. To examine this hypothesis, we engineered two TALEN variants (XPCT1-HD and XPCT1-N*) containing either HD or N* at position +2 of their left TALE DNA binding domain (Fig. 1*A*) and compared their ability to introduce targeted DNA modifications into the methylated *XPC* locus, involved in the development of the human genetic disease xeroderma pigmentosum (23).

To first control whether substitution of HD to N* affected the intrinsic nuclease activity of XPCT, we performed an SSA assay in Chinese hamster ovary (CHO) cells (14), using an unmethylated extrachromosomal XPC1 target (Fig. 1). Our results showed that both TALENs displayed similar intrinsic activity (Fig. 1*B*). We then assessed the ability of these TALENs to disrupt the endogenous methylated XPC1 target in 293H cells by TM. TALEN-induced TM, consisting of small insertion or deletion of nucleotide generated via imprecise nonhomologous end joining, was determined by an EndoT7 assay and/or by deep sequencing as described previously (15, 24). We observed that although both TALENs share similar intrinsic nuclease activities, XPCT1-N* induced at least 17 times more TM events than XPCT1-HD (Fig. 1*C* and supplemental Fig. 1*A*). The difference of TM frequency induced by the two TALENs was due neither to variation of transfection efficiency from one TALEN to another nor to a difference of protein expression (supplemental Fig. 1*B*), thus suggesting a possible involvement of 5mC in the observed difference on TM efficiency. To evaluate

whether this difference of activity originated from the presence of 5mC in the endogenous XPC1 target, we tested the ability of 5-aza-dC, a well known demethylating agent, to rescue XPCT1-HD activity. Interestingly, treatment of 293H cells by 200 nM 5-aza-dC resulted in about 20% of XPC1 demethylation (supplemental Fig. 2A) and induced more than 6-fold increase of XPCT1-HD dependent TM frequency (supplemental Fig. 2B, left panel) without influencing its protein expression (supplemental Fig. 2C). In contrast, XPCT1-N* activity remained unchanged (supplemental Fig. 2B, right panel), indicating that TALE repeat N* binding capacity was not influenced by the presence of the methyl moiety of 5mC. Together, our results showed that XPCT1-HD was sensitive to cytosine methylation and that such sensitivity can be efficiently overcome by substituting the TALE repeat HD with N* at position +2 of its DNA binding domain.

To investigate the molecular mechanism underlying sensitivity of XPCT1-HD to 5mC, the DNA binding affinity of the XPCT1L was investigated *in vitro* in the presence of unmethylated or methylated DNA targets. TALE DNA binding domains of XPCT1L-HD and XPCT1L-N* fused to an N-terminal His tag were overexpressed in *Escherichia coli*, purified by affinity column chromatography, and used to perform equilibrium binding assays with fluorescent XPCT1L DNA targets (supplemental Fig. 3). Binding isotherms, obtained from the variation of DNA fluorescence anisotropy upon increasing protein concentration, showed that full methylation of XPC1L DNA target impaired the DNA binding capacity of XPCT1L-HD (supplemental Fig. 3A, compare *open circles* and *open triangles*; supplemental Table IV). This impairment was essentially due to the presence of 5mC at position +2 of the sense strand (supplemental Fig. 3A, *closed triangles*; supplemental Table IV), in full agreement with the general DNA binding mode of the TALE repeat, which is known to interact exclusively with the sense strand (2, 3, 20, 22). In contrast, the DNA binding capacity of XPCT1L-N* remained unaffected by the presence of 5mC on both DNA strands (supplemental Fig. 6B). We thus conclude that the additional methyl moiety present in 5mC induces a binding penalty that prevents proficient binding by the TALE repeat HD. Such a binding penalty is alleviated in the case of TALE repeat N*, probably because of its peculiar RVD loop structure. The availability of three-dimensional coordinates of TALE repeats HD and N* in complex with cytosine (20) allowed us to rationalize this assumption at a molecular level. Indeed, when the TALE repeat HD was modeled in the presence of 5mC, an obvious steric clash could be observed between the $C\beta$ of aspartate 13 (D) and the methyl moiety of 5mC (distance between $D\_C\beta$ and $C\_C5 < 3$ Å, supplemental Fig. 4B), thus implying that RVD HD loop must operate conformational changes to accommodate 5mC. TALE repeat N*, however, remained free of any steric hindrance with regard to 5mC (supplemental Fig. 4C), explaining its ability to efficiently accommodate the 5mC in position +2 of the endogenous XPC1 target. Interestingly, our *in vitro* results also showed that XPCT1-N* and XPCT1-HD harbored the same affinity for unmethylated XPC1L target (supplemental Fig. 3 and supplemental Table IV). This indicates that the binding penalty caused by the loss of interactions between aspartate 13 and

cytosine (supplemental Fig. 4, B and C, top) is not sufficient enough to affect the overall TALE DNA binding affinity.

In light of these structural and biochemical considerations, we hypothesized that other naturally occurring TALE repeats, either lacking or harboring small side chain residues at position 13, could also bind 5mC. To confirm this, we assessed the ability of TALE repeats H* and NG to substitute HD within XPCT1 TALE DNA binding domain and rescue its activity toward its endogenous methylated locus in 293H cells (supplemental Fig. 5). Our results showed that both TALE repeats H* and NG could rescue XPCT1 activity, with a clear advantage for H*, which was almost as efficient as N*. We thus conclude that although small amino acids in position 13 can accommodate 5mC, absence of such amino acids, the hallmark of the TALE repeat *, leads to more proficient 5mC recognition.

The ability of TALE repeat N* to overcome TALE DNA binding domain sensitivity to 5mC was then challenged using two other engineered TALENs, XPCT2 and XPCT3, specifically designed to process different methylated endogenous XPC targets (XPC2 and XPC3). These targets contained one and two 5mCs, respectively, located at different positions (Fig. 2A), making it possible to evaluate the influence of the number and position of N* repeats in a TALE DNA binding domain. TALEN activities of XPCT2-N* and XPCT3-N* were determined in 293H cells according to the protocol described above and then compared with their HD counterparts (Fig. 2B). Our EndoT7 assays and deep sequencing results showed that N* variants were always the most active, indicating that TALE repeat N* is able to successfully bind 5mC in different contexts. Interestingly, the basal activities of TALEN-HD variants and their enhancement observed after HD/N* substitution were different from one TALEN to another (from 0.7 to 3% of TM frequency and from 2- to 17-fold enhancement respectively, Fig. 2B). These results suggest that the binding penalty induced by 5mC and the activity enhancement promoted by HD/N* substitution depend on the number and the position of 5mC within the TALE DNA binding site. However, we cannot rule out the influence of TALE DNA target sequence on the overall sensitivity of TALENs to 5mC, and further experiments are needed to address this point.

Finally, we verified that HD/N* substitution within TALE DNA binding domains did not increase TALEN-induced toxicity in CHO cells using the protocol described by Grizot *et al*. (14). For all TALENs tested, we found that the presence of single or multiple TALE repeats N* did not influence TALEN-induced toxicity as seen by similar cell survival patterns obtained between HD and N* variants (supplemental Fig. 6). In full agreement with their lack of toxicity, TALEN-N* variants displayed similar TM frequencies in 293H cells, 3 or 7 days after transfection (less than 2-fold difference of TM frequencies was observed between day 3 and day 7, data not shown). Consistent with this absence of toxicity, naturally occurring TAL effectors were reported to bear up to 23% of TALE repeat N* (2, 3) within their DNA binding domain, whereas retaining high specificity. Therefore, taken together, our results showed that the TALE repeat N* could be used as a universal 5mC binding module without affecting toxicity of engineered TALE DNA binding domains.
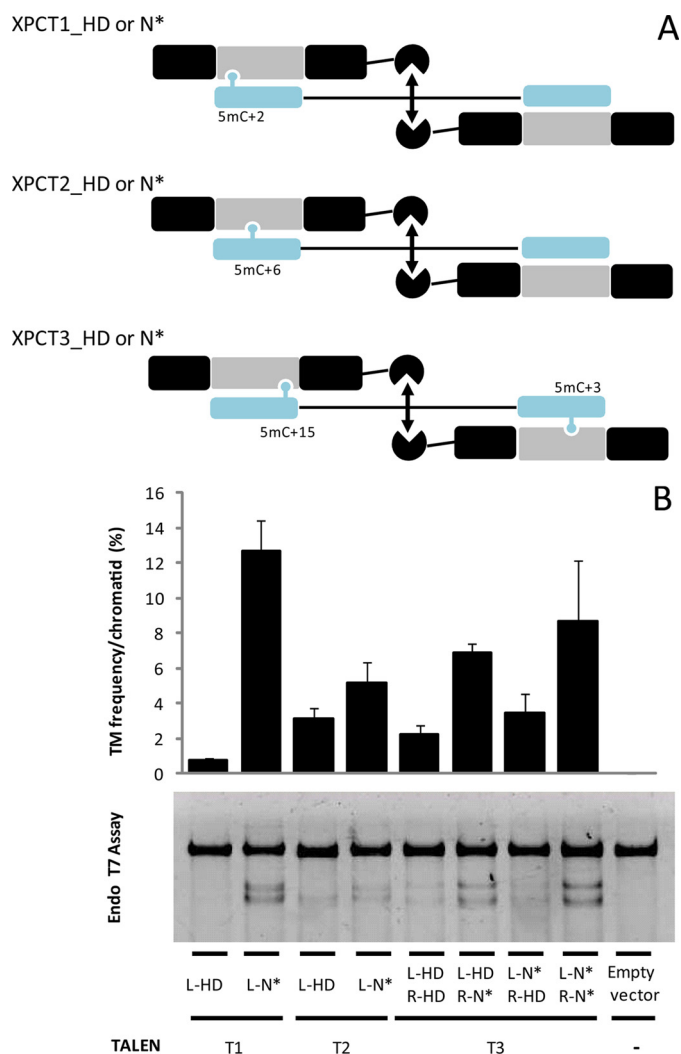
FIGURE 2. **TALE repeat N*, a universal 5mC binding module.** *A*, schematic representation of the XPCT1, XPCT2, and XPCT3 TALENs used to challenge the ability of TALE repeat N* to overcome TALE DNA binding domain sensitivity to 5mC. XPC1, XPC2, and XPC3 DNA targets are colored in *light blue*, and the positions of 5mCs are indicated by *dots*. TALE DNA binding domains are colored in *gray*, and the N-terminal, C-terminal, and FokI domains are colored in *black*. *B*, TM of endogenous methylated XPC1, XPC2, and XPC3 targets induced by 10 μg of DNA plasmids encoding their respective TALENs, determined by deep sequencing ($n = 2$, top) and EndoT7 assay (*bottom*). TALENs containing different combinations of TALE repeats HD or N* on their right (*R*) and left (*L*) DNA binding domains were assayed. As an example for the sake of clarity, XPCT3 bearing TALE repeat HD and N* on its right and left DNA binding domains, respectively, is indicated as XPCT3 *R*-HD, *L*–N*. *Error bars* indicate S.D.

The negative impact of cytosine methylation on TALEN activity could also be found in the recently published work of Reyon *et al.* (15). This extensive work reported that 12 out of 96 different TALENs tested in U2OS cells were inactive, without clearly identifying the reason of this failure. The presence of multiple CpG dinucleotides in almost all DNA sequences targeted in this study prompted us to look for possible correlation between TALEN activity and CpG methylation of their respective target. A BLAST survey of U2OS cells methylome (21) enabled us to identify 15 different TALEN targets and determine their methylation status (supplemental Table V). Among them, two were found to be fully methylated. Interestingly, these two methylated sequences, named HOXD11 and TLX3,

fell into the group of unprocessed targets reported by Reyon *et al.* (15), which was consistent with the inhibitory effect of cytosine methylation on TALEN activity (supplemental Fig. 7). The 13 other unmethylated targets were reported to be processed with widely variable outcomes (15). The failure to process the unmethylated TGFBR2 target represents an extreme case indicating that other factors, including chromatin compaction, target sequence, or TALEN expression as the most obvious ones, could also significantly influence TALEN activity.

The identification of TALE repeat N* as a universal 5mC binding module unmasks exciting prospects for the fields of genome engineering, gene regulation, and gene therapy. Among the most promising ones is the ability to reprogram somatic cells into iPS cells by dTALE-mediated activations of pluripotency genes, known to be silenced by hypermethylation (c-Myc and Oct4) (7). A similar approach could also be used to prevent development of a wide variety of malignant cell lines by promoting activation of hypermethylated genes involved in tumor suppression (25, 26). Finally, the successful TALEN-mediated processing of methylated *XPC* loci paved the way for *XPC* gene correction in a patient cell line, a step further in xeroderma pigmentosum-targeted gene therapy.

In summary, our work demonstrates the ability of TALE repeat N* to efficiently accommodate 5mC. Based on this finding, we present a simple, efficient, and universal method to overcome TALE DNA binding domain sensitivity to cytosine methylation. Such a method presents three major advantages. First, it allows one to bypass the need for chemical demethylation of endogenous targets, which is unsuitable for cell engineering and therapeutic applications. Second, it is readily applicable to all TALE-derived proteins, and in particular, to engineered transcription activators, thus potentially enabling site-specific activation of methylated promoters responsible for gene silencing. Third, it is transposable to the broad range of cellular systems including ES, iPS mammalian cells, and plant cells that have already been shown to be engineerable with TALE-based technologies.

REFERENCES

1. Bogdanove, A. J., and Voytas, D. F. (2011) TAL effectors: customizable proteins for DNA targeting. *Science* **333,** 1843–1846
2. Moscou, M. J., and Bogdanove, A. J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* **326,** 1501
3. Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326,** 1509–1512
4. Morbitzer, R., Römer, P., Boch, J., and Lahaye, T. (2010) Regulation of selected genome loci using *de novo*-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 21617–21622
5. Cong, L., Zhou, R., Kuo, Y. C., Cunniff, M., and Zhang, F. (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3,** 968
6. Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., and Voytas, D. F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186,** 757–761
7. Bultmann, S., Morbitzer, R., Schmidt, C. S., Thanisch, K., Spada, F., El-

saesser, J., Lahaye, T., and Leonhardt, H. (2012) Targeted transcriptional activation of silent *oct4* pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.* **40,** 5368–5377

8. Jaenisch, R., and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33,** (suppl) 245–254

9. Vanyushin, B. F., and Ashapkin, V. V. (2011) DNA methylation in higher plants: past, present, and future. *Biochim. Biophys. Acta* **1809,** 360–368

10. Ziller, M. J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C. B., Bernstein, B. E., Lengauer, T., Gnirke, A., and Meissner, A. (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* **7,** e1002389

11. Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., Johnson, B. E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V. M., Rowitch, D. H., Xing, X., Fiore, C., Schillebeeckx, M., Jones, S. J., Haussler, D., Marra, M. A., Hirst, M., Wang, T., and Costello, J. F. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466,** 253–257

12. Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A. J., Socci, N. D., and Scandura, J. M. (2011) DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* **6,** e14524

13. Palii, S. S., Van Emburgh, B. O., Sankpal, U. T., Brown, K. D., and Robertson, K. D. (2008) DNA methylation inhibitor 5-aza-2′-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. *Mol. Cell Biol.* **28,** 752–771

14. Grizot, S., Smith, J., Daboussi, F., Prieto, J., Redondo, P., Merino, N., Villate, M., Thomas, S., Lemaire, L., Montoya, G., Blanco, F. J., Pâques, F., and Duchateau, P. (2009) Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res.* **37,** 5405–5419

15. Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D., and Joung, J. K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* **30,** 460–465

16. Valton, J., Daboussi, F., Leduc, S., Molina, R., Redondo, P., Macmaster, R., Montoya, G., and Duchateau, P. (2012) 5′-Cytosine-phosphoguanine (CpG) methylation impacts the activity of natural and engineered meganucleases. *J. Biol. Chem.* **287,** 30139–30150

17. Arosio, D., Costantini, S., Kong, Y., and Vindigni, A. (2004) Fluorescence anisotropy studies on the Ku-DNA interaction: anion and cation effects. *J. Biol. Chem.* **279,** 42826–42835

18. Murakami, M. T., Sforça, M. L., Neves, J. L., Paiva, J. H., Domingues, M. N., Pereira, A. L., Zeri, A. C., and Benedetti, C. E. (2010) The repeat domain of the type III effector protein PthA shows a TPR-like structure and undergoes conformational changes upon DNA interaction. *Proteins* **78,** 3386–3395

19. Bardsley, W. G., and McGinlay, P. B. (1987) The use of non-linear regression analysis and the F test for model discrimination with dose-response curves and ligand binding data. *J. Theor. Biol.* **126,** 183–201

20. Mak, A. N., Bradley, P., Cernadas, R. A., Bogdanove, A. J., and Stoddard, B. L. (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335,** 716–719

21. Ohm, J. E., Mali, P., Van Neste, L., Berman, D. M., Liang, L., Pandiyan, K., Briggs, K. J., Zhang, W., Argani, P., Simons, B., Yu, W., Matsui, W., Van Criekinge, W., Rassool, F. V., Zambidis, E., Schuebel, K. E., Cope, L., Yen, J., Mohammad, H. P., Cheng, L., and Baylin, S. B. (2010) Cancer-related epigenome changes associated with reprogramming to induced pluripotent stem cells. *Cancer Res.* **70,** 7662–7673

22. Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J. K., Shi, Y., and Yan, N. (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335,** 720–723

23. Khan, S. G., Muniz-Medina, V., Shahlavi, T., Baker, C. C., Inui, H., Ueda, T., Emmert, S., Schneider, T. D., and Kraemer, K. H. (2002) The human *XPC* DNA repair gene: arrangement, splice site information content and influence of a single nucleotide polymorphism in a splice acceptor site on alternative splicing and function. *Nucleic Acids Res.* **30,** 3624–3631

24. Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39,** 9283–9293

25. Berdasco, M., and Esteller, M. (2010) Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell* **19,** 698–711

26. Jones, P. A., and Baylin, S. B. (2007) The epigenomics of cancer. *Cell* **128,** 683–692