Published online 11 March 2016

Nucleic Acids Research, 2016, Vol. 44, No. 12 ***5615–5628***
doi: 10.1093/nar/gkw159

Repurposing the CRISPR-Cas9 system for targeted DNA methylation

Aleksandar Vojta[1,†], Paula Dobrinić[1,†], Vanja Tadić[1], Luka Bočkor[1], Petra Korać[1], Boris Julg[2], Marija Klasić[1] and Vlatka Zoldoš[1,*]

[1]Department of Biology, Division of Molecular Biology, University of Zagreb, Faculty of Science, Zagreb, HR-10000, Croatia and [2]Ragon Institute of MGHT, MIT and Harvard, 400 Technology Square, Cambridge, MA, USA

Received January 04, 2016; Revised February 29, 2016; Accepted March 01, 2016

ABSTRACT

Epigenetic studies relied so far on correlations between epigenetic marks and gene expression pattern. Technologies developed for epigenome editing now enable direct study of functional relevance of precise epigenetic modifications and gene regulation. The reversible nature of epigenetic modifications, including DNA methylation, has been already exploited in cancer therapy for remodeling the aberrant epigenetic landscape. However, this was achieved non-selectively using epigenetic inhibitors. Epigenetic editing at specific loci represents a novel approach that might selectively and heritably alter gene expression. Here, we developed a CRISPR-Cas9-based tool for specific DNA methylation consisting of deactivated Cas9 (dCas9) nuclease and catalytic domain of the DNA methyltransferase DNMT3A targeted by co–expression of a guide RNA to any 20 bp DNA sequence followed by the NGG trinucleotide. We demonstrated targeted CpG methylation in a ∼35 bp wide region by the fusion protein. We also showed that multiple guide RNAs could target the dCas9-DNMT3A construct to multiple adjacent sites, which enabled methylation of a larger part of the promoter. DNA methylation activity was specific for the targeted region and heritable across mitotic divisions. Finally, we demonstrated that directed DNA methylation of a wider promoter region of the target loci *IL6ST* and *BACH2* decreased their expression.

INTRODUCTION

Development of genome editing tools such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the CRISPR-Cas9 system (1–3) has opened new possibilities for analyses of gene function by targeting and altering almost any sequence in the genome of cultured cells or a whole organism (4,5). While the classical approach to functional analysis by gene knockout remains essential in unraveling different cellular processes, it lacks the ability to directly address the complexity of gene regulation. Therefore, some of the genome engineering tools have been altered and repurposed for targeted modification of histone tails and the DNA molecule, such as histone methylation (6) and demethylaton (7), histone acetylation (8); cytosine methylation (9–11) and hydroxymethylation (12,13), as well as for direct control of transcription either by VP64-mediated transactivation (14,15) or silencing via CRISPR interference (16).

Until recently, epigenetic studies relied only on correlations between certain epigenetic modifications and gene regulation (i.e. activation or silencing). Epigenome editing now enables direct study of functional relevance of certain epigenetic modification at a specific locus or a genomic region (7–12). The reversible nature of epigenetic modifications including DNA methylation (17,18) has been already exploited by the use of epigenetic inhibitors (such as inhibitors of DNA methyltransferases or histone deacetylases) in cancer therapy in order to remodel the aberrant epigenetic landscape (19,20). However, such approach can only be applied non-selectively at the global genome level and not at specific loci. The use of epigenome editing is therefore appealing in gene therapy, design of new smart drugs and stem cell differentiation, since it can target specific genes (21).

The basic layout of a tool for epigenome editing consists of two essential parts: a DNA-binding targeting domain and a functional domain. The targeting domain can be based on ZFNs, TALEs or the CRISPR-Cas9 system (22). While the first tools for epigenome editing relied on ZFNs for targeting (23), which remained a popular choice due to their compact size and good binding properties (9,10,12), TALE-based systems gained prominence in applications to epigenetic modifications (7,11,13) due to their easier assembly and targeting in comparison with ZFNs. However, TALE-based targeting suffers from sensitivity to CpG methylation that is difficult to overcome (24), which

*To whom correspondence should be addressed. Tel: +385 1 4606266; Fax: +385 1 4606 286; Email: vzoldos@biol.pmf.hr
†These authors contributed equally to the work as the first authors.

makes it inherently unsuitable for modification of mammalian gene promoter methylation. Finally, the choice of the targeting domain for both genome and epigenome editing seems to be shifting toward the CRISPR-Cas9 system (8,15), which is not surprising because of the ease of targeting by guide RNAs even to multiple sites (25) and insensitivity to CpG methylation.

The functional domain of such an epigenome editing tool carries out the desired epigenetic modification at the targeted site. Posttranslational modifications of histone tails and modification of DNA molecules by cytosine methylation are important epigenetic mechanisms involved in gene regulation. Both represent attractive targets for manipulation (6–8,26). However, DNA methylation is a more stable epigenetic mark with a potential for long-term effects on gene expression, which makes it an appealing modification for epigenome editing in functional studies and possibly even in gene therapy (9–13).

The catalytic domain of DNA methyltransferase 3A (DNMT3A) represents the functional domain of choice for targeted CpG methylation. DNMT3A is the most active *de novo* DNA methyltransferase isoform in humans and it preferentially methylates CpG sites *in vivo* (27). In addition, its catalytic domain alone displays enzymatic activity in transfected cells (28). Its interaction partner, the DNA methyltransferase 3-like (DNMT3L) protein is essential for stimulation of a high level of DNMT3A activity. While its external addition is indispensable for significant DNA methylation activity *in vitro* (29), DNMT3L seems to be efficiently recruited to an active complex by DNMT3A in a living cell (30).

Protein *N*–glycosylation is an important posttranslational modification affecting the structure and modulating the function of many eukaryotic proteins (31). Epigenetic regulation of this process is gaining prominence in the emerging big picture of cellular adaptation and integration of environmental information (32). Two genes, the *BACH2* and *IL6ST*, have been previously shown to be relevant for *N*–glycosylation of immunoglobulin G (IgG) and to be associated with autoimmune diseases such as inflammatory bowel disease (IBD) and lupus erythematosus (33). These genes also have CpG islands in their promoter region, which indicates that their expression may be controlled by CpG methylation. Changes in glycosylation observed in IBD (34) might be caused by deregulation of these two genes and we are currently studying association of promoter methylation of *BACH2* and *IL6ST* with IBD. However, the analysis of causality in these associations would require more sophisticated functional studies involving direct manipulation of the promoter methylation status of the investigated genes.

In this study, we constructed and characterized for the first time a CRISPR-Cas9-based tool for targeted CpG methylation by fusing the inactive Cas9 (dCas9) endonuclease (the targeting domain) with the DNMT3A catalytic domain (the functional domain) via a flexible Gly_4Ser linker. This tool combines the ease of targeting of the Cas9 protein using short complementary single guide RNA (sgRNA) with the already established DNMT3A activity in fusion constructs (9,10). We demonstrated high efficiency and specificity of this tool for methylation at targeted CpG sites within the promoter of two loci, the *BACH2* and *IL6ST*. Moreover, the activity of the dCas9-DNMT3A construct guided by multiple sgRNAs elevated the methylation level in a wider region within the promoters of the *IL6ST* and *BACH2* loci and lowered their expression level. The putative involvement of *BACH2* and *IL6ST* in the loss of immunosuppressive role of IgG in IBD suggests the CRISPR-Cas9 strategy as potentially applicable in locus-specific epigenetic therapeutics.

MATERIALS AND METHODS

Plasmid construction

Nickase activity of the plasmid pSpCas9n(BB)-2A-Puro (PX462) (Addgene plasmid # 48141, a gift from Feng Zhang) (2) encoding Cas9n D10A was abolished by introducing the H840A mutation into the Cas9 gene using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA). A short Gly_4Ser linker introducing the unique BamHI and FseI restriction sites was inserted at the C-terminus of the Cas9 gene.

Undesired BbsI restriction site within the catalytic domain of human DNMT3A from the plasmid pcDNA3/Myc-DNMT3A (Addgene plasmid # 35521, a gift from Arthur Riggs) (28) was removed by site-directed mutagenesis. The DNMT3A catalytic domain (amino acids P602-V912) was then amplified using the Herculase II Fusion DNA Polymerase (Agilent Technologies) with primers designed to enable insertion C–terminally from the Gly_4Ser linker, between the BamHI and FseI restriction sites. Catalytically inactive DNMT3A was generated by site-directed mutagenesis of the active site motif ENV to ANV (E756A) (28). EcoRI fragments encoding T2A-Puro and T2A-EGFP were cut from plasmids pSpCas9n(BB)-2A-Puro (PX462) (Addgene plasmid # 48141) and pSpCas9n(BB)-2A-GFP (PX461) (Addgene plasmid # 48140, a gift from Feng Zhang) (2), respectively and cloned at the C–terminus of the dCas9-DNMT3A fusion to allow for either puromycin selection or EGFP screening of transfected cells.

Guide sequences targeting the loci of interest were selected either manually or using the GT-Scan web-tool (http://gt-scan.braembl.org.au/gt-scan/) (35). Sequences of the non-targeting (NT) control guide RNAs, without a target in the human genome, were taken from the Human GeCKOv2 Library (36). All sgRNA genes (Supplementary Table S1) were synthesized by oligo annealing and cloned into expression vectors via BbsI sites essentially as described in (25) using BpiI/BbsI (Thermo Scientific, Waltham, Massachusetts, USA) and T4 DNA ligase (Takara Biotechnology, Dalian, China). Additionally, BACH2-sgRNA8 was cloned into H840A-mutated pSpCas9n(BB)-2A-Puro vector, expressing dCas9 without the DNMT3A catalytic domain, in order to analyze the sole effect of sgRNA-dCas9 binding on CpG methylation. The cloned guide RNAs were verified by sequencing using the U6–Fwd primer. Sequences of all oligonucleotides are listed in Supplementary Table S2.

Cell transfection

Human embryonic kidney cell line HEK293 was maintained in Dulbecco's Modified Eagle Medium (Lonza, Verviers, Belgium) supplemented with 10% heat inactivated fetal bovine serum (Biosera, Ringmer, UK), 4 mM L–glutamine (Lonza), 100 U/ml penicillin and 100 μg/ml streptomycin (Lonza). Cells were incubated at 37°C in a humidified 5% CO_2–containing atmosphere.

Cells were seeded into a 24–well plate and transfected the next day at 70–90% confluence using Lipofectamine 3000 (Life Technologies, Carlsbad, California, USA) according to the manufacturer's protocol. Transfections were done with 100 ng of plasmids co–expressing dCas9-DNMT3A and a chimeric sgRNA, in biological triplicates. Cells transfected with dCas9-DNMT3A-EGFP plasmid were fixed 24 h after transfection and the percentage of cells expressing GFP was assessed by fluorescent microscopy, to ensure consistent level of transfection efficiency in all experiments. One day after transfection, cells were passaged to 6–well plates and selected for 48 h with 1.5 μg/ml puromycin. Eight days after transfection (or at other defined time points for the time-course experiment), cells were harvested and incubated at 37°C overnight in lysis buffer (50 mM Tris pH 8.5, 1 mM EDTA, 0.5% Tween 20) with proteinase K (400 μg/ml) yielding crude cell lysates.

Analysis of DNA methylation

DNA from crude cell lysates was bisulfite converted and purified using EZ DNA Methylation-Gold Kit (Zymo Research Europe, Freiburg, Germany) according to the manufacturer's protocol. Genomic fragments encompassing sgRNA target sites were amplified from bisulfite converted DNA using the PyroMark PCR kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Methylation of repetitive LINE-1 elements was analyzed as described previously (37). The cycling was performed as follows: initial denaturation for 15 min at 95°C; 50 cycles of 30 s at 95°C, 30 s at 48°C (BACH2–A and BACH2–B fragments), 52°C (IL6ST–A fragment) or 64°C (LINE-1) and 30 s at 72°C; final extension for 10 min at 72°C. In order to quantify methylation levels at individual consecutive CpG sites, PCR amplicons were sequenced using the PyroMark Q24 Advanced pyrosequencing system (Qiagen). Sequences of PCR primers and pyrosequencing primers are listed in Supplementary Table S2.

Dose-response experiment

Titration of dCas9-DNMT3A plasmid in HEK293 cells was performed in 24-well plates using 500, 250, 100 or 25 ng of plasmid DNA. Cells were transfected with the targeting BACH2-sgRNA3 constructs (wild-type dCas9-DNMT3A or ANV-mutant) or with two different non-targeting sgRNA constructs. After selection, CpG methylation in the BACH2-B gene fragment was assayed as described above.

Targeted promoter methylation and expression analysis

The promoter regions of the *IL6ST* and *BACH2* genes were targeted using either individual or pooled sgRNAs co-expressed from dCas9-DNMT3A-PuroR plasmids. Eight different sgRNAs were used to target the *BACH2* promoter, and four sgRNAs were used to target the *IL6ST* promoter (Supplementary Table S1). Additionally, the *BACH2* promoter was targeted with the sgRNA8 co-expressed with dCas9 protein, in order to investigate if the sole binding of sgRNA-dCas9, in the absence of the DNMT3A domain, affects CpG methylation level.

Cells were transfected in 24-well plates using in total 100 ng of plasmid DNA per transfection, selected and harvested as described above. Experimental groups were cells transfected with either active or inactive constructs co–expressing targeting sgRNAs, while the control groups were mock-transfected cells (without plasmid DNA) and cells co-expressing the active dCas9-DNMT3A and the non-targeting sgRNA. Transfection was done in biological triplicates. Total RNA was isolated at the same time as DNA using RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. One microgram of RNA was used as a template for reverse transcription using the PrimeScript RTase (Takara) and random hexamer primers (Invitrogen, Paisley, UK). Quantitative real-time PCR (RT–qPCR) was performed on cDNA from cells transfected with pooled sgRNAs co–expressing either the active or the inactive construct and both controls (mock transfection and active non–targeting sgRNA). For RT–qPCR, we used hydrolysis probes in the 7500 Fast Real-Time PCR system (Applied Biosystems, Carlsbad, California, USA) in technical triplicates with TaqMan Gene Expression Master Mix and the following TaqMan Gene Expression Assays: Hs00174360_m1 (*IL6ST*), Hs00222364_m1 (*BACH2*) and Hs02800695_m1 (*HPRT1*) (Applied Biosystems). The expression of the target genes *BACH2* and *IL6ST* was normalized to the reference gene *HPRT1* and analyzed using the comparative C_q method (38). Expression was shown as fold change (FC) relative to mock-transfected cells.

Time course experiment

HEK293 cells were transfected in triplicates in 6-well plates with 500 ng of dCas9-DNMT3A-PuroR plasmid encoding BACH2-sgRNA8 or IL6ST-sgRNA3, or their DNMT3A-mutated counterparts. One day after transfection, triplicates were pooled and re-plated in 6-well plates. Cells were selected as described above. Aliquots of the transfected cells were collected daily until the 9th day and after that less frequently until the 42nd day. Those samples were used for CpG methylation analysis by pyrosequencing as well as for detection of plasmid DNA by qPCR. Methylation of the BACH2–A and IL6ST–A fragments was assayed as described above.

The experiment was repeated with two sets of samples taken on days 1, 3, 6, 7, 10 and 15 after transfection when cells were fixed on microscopy slides. The samples were used for direct detection of plasmid DNA in cells by fluorescent *in situ* hybridization (FISH) and for detection of Cas9 expression by immunofluorescence (IF).

Relative quantification of plasmid DNA by qPCR

qRT-PCR was performed on crude cell lysates from cells transfected in the time course experiment. We ana-

lyzed samples transfected with plasmids encoding BACH2-sgRNA8 (active or inactive construct), collected on days 1, 2, 3, 4, 7, 9 and 16 after transfection, and mock transfected cells.

The amount of plasmid encoding dCas9-DNMT3A was determined relative to the genomic *RAG1* sequence, amplified using a pair of previously published primers (39). Sequences of qPCR primers are listed in Supplementary Table S2. Reactions were run in technical duplicates using Power SYBR Green PCR Master Mix (Applied Biosystems) and 50 nM primers for Cas9 or *RAG1*. To assess primer efficiency we prepared standard curves of known amounts of the linearized plasmid and purified genomic DNA isolated from HEK293 cells. Cycling conditions were 10 min at 95°C followed by 40 cycles at 95°C for 15 s and 62°C for 1 min, followed by melt curve analysis of amplified targets. PCR was performed in 7500 Fast Real-Time PCR System (Applied Biosystems) using standard method. The amount of plasmid DNA relative to genomic DNA was calculated essentially as described in (40).

EGFP fluorescence analysis

HEK293 cells were transfected with 500 ng of the plasmid co–expressing Cas9-DNMT3A-EGFP fusion with different guide RNAs targeting the *BACH2* (sgRNA8) or *IL6ST* (sgRNA3) promoter regions. Constructs either co–expressing non–targeting guide RNA or having inactive DNMT3A domain were used as controls.

Images were acquired under the same conditions using Zeiss Axiovert 40CFL microscope with AxioCam MRm in the AxioVision software (Carl Zeiss Microscopy, LLC, USA) at specified time points (1, 4, 6, 7 and 10 days after transfection). EGFP fluorescence, number of EGFP positive cells per field and total number of cells per field were quantified using the ImageJ software (41).

Immunofluorescence

At each time point, expression of dCas9-DNMT3A protein was analyzed in cells transfected with BACH2-sgRNA8, BACH2-sgRNA8 ANV and mock-transfected cells. Fixed cells were washed with PBS, incubated with primary anti-tag antibody (MA1–91878, FLAG™ Epitope Tag (DYKDDDDK) Monoclonal Antibody, 1:5000, Thermo Fisher Scientific, Waltham, MA, USA) for 90 min and afterward with secondary antibody labeled with FITC (Goat polyclonal Secondary Antibody to Mouse IgG – H&L (FITC), 1:200, Abcam, Cambridge, UK) for 60 min. Cells were mounted using anti-fade Fluorescence Mounting Medium (Dako, Glostrup, Denmark) containing 8 μg/ml DAPI. Results were analyzed using Olympus BX51 microscope. Fluorescence intensities were measured using ImageJ software (41) and percentage of cells expressing Cas9 protein per sample was calculated.

Fluorescent *in situ* hybridization

The Cy3-labeled probe for the plasmid encoding dCas9-DNMT3A was prepared using Nick Translation Kit (Roche, Indianapolis, Indiana, USA) and subsequently purified with the ChargeSwitch PCR Clean-Up Kit (Invitrogen, Paisley, UK) according to the manufacturer's protocols. For detection of genomic DNA, the centromeric chromosome 7 enumeration probe was used (Vysis CEP7 SpectrumGreen; Abbott Molecular, Des Plaines, IL, USA). At each time point, mock-transfected cells and cells transfected with BACH2-sgRNA8 and BACH2-sgRNA8 ANV were analyzed. Chromosome slides containing the mixture of probes specific for the plasmid and the chromosome 7 were denatured together at 85°C for 5 min and hybridized at 37°C overnight. The next day, slides were washed three times in preheated 0.1x saline sodium citrate (SSC) buffer at 65°C and mounted with anti-fade Fluorescence Mounting Medium (Dako) containing 8 μg/ml DAPI. The images were acquired using Olympus BX51 microscope. Number of signals per cell representing plasmid DNA encoding dCas9-DNMT3A was determined on each slide and used as approximate value for evaluation of plasmid DNA stability.

Western blot analysis of total protein extracts

HEK293 cells were co–transfected with 100 ng of plasmid expressing BACH2-sgRNA8 constructs with puromycin selection marker (wild-type dCas9-DNMT3A or ANV-mutant) and 20 ng of plasmid expressing EGFP (pcDNA3-EGFP, Addgene plasmid # 13031, a gift from Doug Golenbock). The cells were selected with puromycin as described above and collected 1, 3, 7, 11 and 15 days following transfection and frozen in liquid nitrogen. Cell lysates were homogenized in RIPA buffer containing 150 mM NaCl, 1% NP–40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris–HCl pH 8 and 2x cOmplete™ protease inhibitor cocktail (Roche, Basel, Switzerland). Subsequently, 10 μg of protein extract was run on 10% SDS-PAGE.

Proteins were transferred onto a nitrocellulose membrane for Western blot analysis. The membrane was cut into pieces that were handled separately. To detect Cas9, the membrane was blocked overnight in 5% milk-PBS (5% nonfat dry milk in 1x phosphate buffer saline: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4), then incubated overnight with anti-Cas9 antibody 1:2000 (ab204448; Abcam) in 5% milk-PBS containing 0.2% Tween 20, followed by incubation with anti-rabbit HRP-conjugated secondary antibody 1:2000 (Dako). Membranes used to detect tubulin (using the E7 antibody, 1:1000, Developmental Studies Hybridoma Bank, Iowa City, IA, USA; deposited by prof. Michael Klymkowsky, University of Colorado, Boulder, CO, USA) and EGFP (using the sc–9996 antibody, 1:1000, Santa Cruz Biotechnology, Dallas, TX, USA) were blocked for one hour at room temperature in 5% BSA–PBS, then incubated overnight with corresponding antibodies in 5% BSA–PBS (5% bovine serum albumin in 1x PBS) containing 0.2% Tween 20, followed by incubation with anti-mouse HRP-conjugated secondary antibody 1:3000 (Dako). The antibody-tagged protein bands were visualized by enhanced chemiluminescence (ECL) system (Pierce Biotechnology, Waltham, MA, USA) according to manufacturer's protocol. ECL signals captured on film were quantified using the ImageJ software (41).

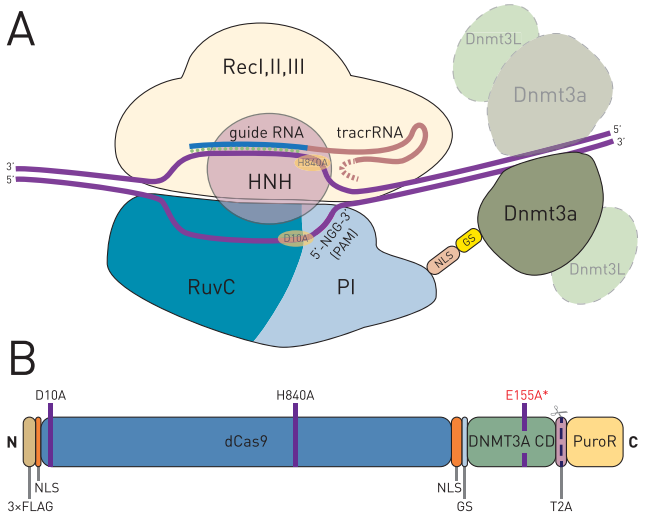

Figure 1. (**A**) Schematic representation of the dCas9-DNMT3A fusion protein in complex with sgRNA and its target DNA sequence. The sgRNA is bound in a cleft between the recognition lobe (RecI, II and III domains) and the nuclease lobe (HNH, RuvC and PI domains) of Cas9 protein. The C–terminus of Cas9 is located on the PAM–interacting (PI) domain and faces the side where the bound genomic DNA protrudes with its 3′ end relative to the sgRNA sequence. The sgRNA is a synthetic fusion between bacterial crRNA and tracrRNA, with guide sequence and tracrRNA part shown in different colors. The catalytic domain of DNMT3A recruits its partner for dimerization and DNMT3L proteins *in vivo* (dashed lightened symbols). NLS, nuclear localization signal; GS, Gly_4Ser peptide linker. (**B**) Domain structure of the dCas9–DNMT3A fusion protein. The nuclease-inactivating mutations D10A and H840A of *Streptococcus pyogenes* Cas9 are indicated. Deactivated Cas9 was fused to the catalytic domain of the human de novo DNA methyltransferase 3A (DNMT3A CD) using a short Gly_4Ser peptide (GS). The dCas9–DNMT3A is expressed as a bicistronic mRNA, along with puromycin resistance gene (PuroR, shown) or EGFP gene, thus enabling selection of transfected cells. The PuroR (or EGFP) moiety is separated during translation by action of the T2A self-cleaving peptide. The inactive fusion methyltransferase (dCas9-DNMT3A-ANV) for use as a negative control contains an additional substitution ($E155A^*$) in the active site of DNMT3A. 3x FLAG, epitope tag; NLS, nuclear localization signal.

RESULTS

Construction of the dCas9–DNMT3A fusion protein

The dCas9–DNMT3A fusion protein was constructed by addition of DNMT3A catalytic domain to the C–terminus of the inactive Cas9 (dCas9) via a short Gly_4Ser linker. The predicted structure of the fusion product based on the Cas9 crystal structure (42) with the DNMT3A catalytic domain protruding at the PAM–interacting (PI) domain side of Cas9 is depicted in Figure 1A. Four types of the construct were made: a construct with either active or inactive DNMT3A catalytic domain, each fused to either puromycin resistance or EGFP gene, as depicted in Figure 1B.

The final dCas9–DNMT3A expression plasmids contained sgRNA expression cassettes compatible with the oligo cloning protocol using BpiI/BbsI (25). The plasmids are deposited at Addgene (plasmid numbers #71667, #71684, #71666, #71685). The amino acid sequences of constructed fusion proteins are shown in Supplementary Figure S1.

The dCas9-DNMT3A construct induced targeted CpG methylation at the *BACH2* locus

We modified the CRISPR-Cas9 tool previously developed for genome editing (2) to be used for targeted DNA methylation of specific loci in mammalian cells. We fused the DNMT3A catalytic domain to the C-terminus of dCas9 via a short flexible linker to enable methylation of DNA sequences adjacent to the sgRNA binding site. In order to determine the optimal amount of dCas9-DNMT3A-encoding plasmid for HEK293 transfection, we performed a dose-response experiment. The plasmid amount eliciting a saturation level of on-target activity and acceptable non-specific activity arising from DNA methyltransferase overexpression (100 ng for transfection in a 24–well plate or 500 ng for a 6–well plate) was used in all subsequent experiments (Supplementary Figure S2).

Two genomic fragments within the *BACH2* promoter, unmethylated in HEK293 cell line, were targeted with several different sgRNAs (Figure 2), which bind in different orientations and at distinct positions relative to the 13 or 14 CpG sites covered either by the BACH2-A or the BACH2-B pyrosequencing assay, respectively. All five sgRNAs targeting the intronic BACH2-B fragment induced a strong increase of CpG methylation relative to the mock-transfected cells (Figure 3A, Supplementary Figure S3). Among them, sgRNA3 showed the highest activity, increasing the methylation level of certain CpG sites upstream and downstream of its binding site up to 34%. Two out of the three sgRNAs targeting the upstream BACH2-A fragment induced similar or even higher increase of the methylation level relative to the mock-transfected cells (Supplementary Figure S3). Methylation level of CpG sites located downstream from the PAM sequence was increased up to 33% (sgRNA7) or 59% (sgRNA8). Methylation level of CpGs within the sgRNA binding sites remained unchanged. Non-targeting dCas9-DNMT3A and targeted inactive constructs (dCas9-DNMT3A-ANV), used as negative controls, did not induce any significant change in methylation level, confirming that both sgRNA-mediated binding and intact methyltransferase activity of dCas9-DNMT3A are required for specific DNA methylation of targeted loci. We observed only a slight increase in CpG methylation level with the catalytically inactive BACH2-sgRNA8 construct (up to ∼10% with the inactive versus ∼55% with the active construct; Figure 3A left panel, Supplementary Figure S3B) mirroring the activity pattern of the active construct, there was no increase in methylation when we targeted the dCas9-only construct (i.e. without the DNMT3A domain) using the BACH2-sgRNA8, thus ruling out any source of CpG methylation activity other than the DNMT3A catalytic domain (Supplementary Figure S3C).

The dCas9–DNMT3A construct specifically methylates a ∼35 bp region centered at 27 bp downstream from the PAM sequence

Various individual sgRNAs targeted the dCas9–DNMT3A construct to different CpG sites within the promoter regions of the *BACH2* and *IL6ST* genes. Changes in CpG methylation level in cells transfected with the dCas9-DNMT3A/sgRNA-expressing plasmids were assessed by

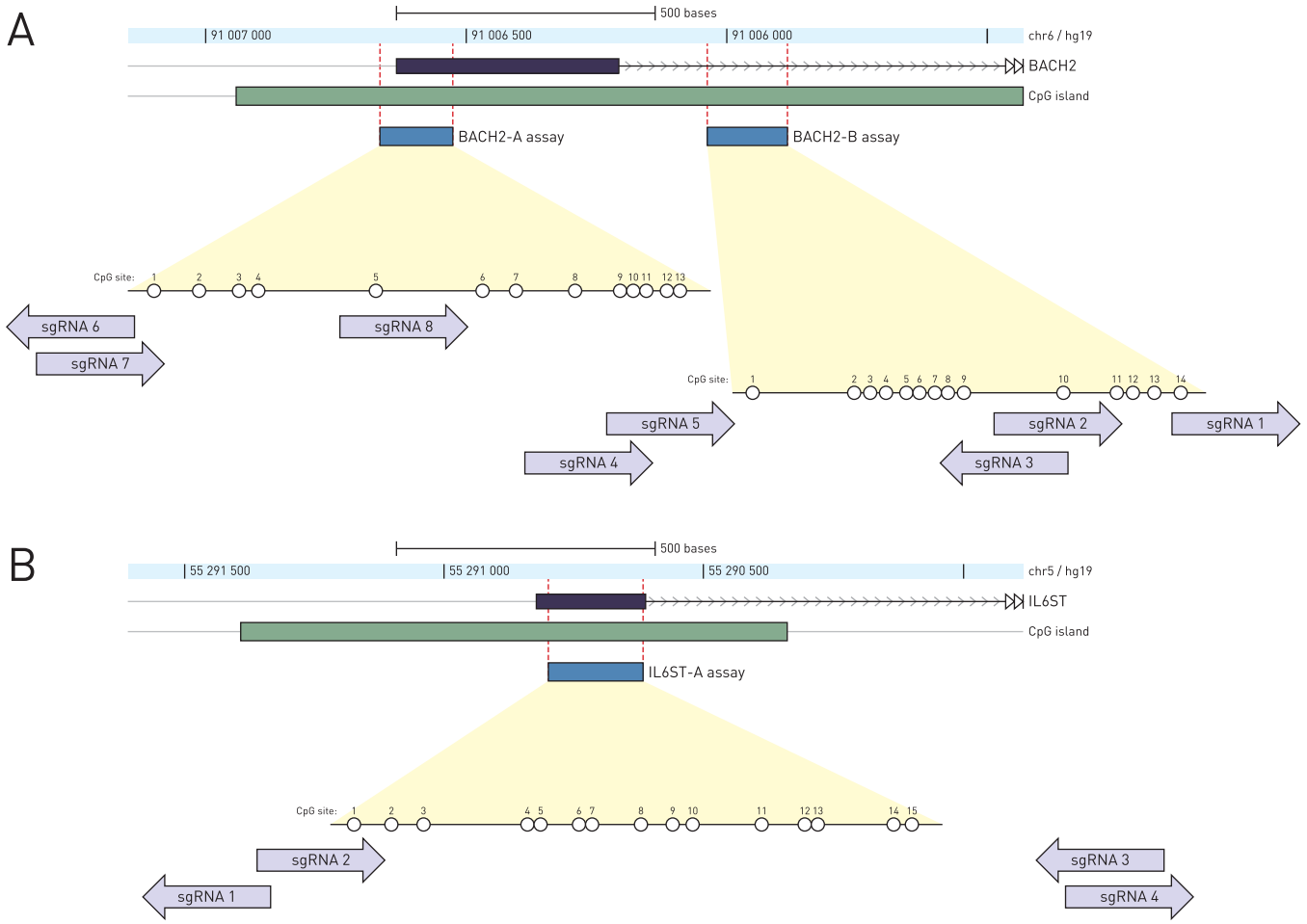

Figure 2. Targeting the dCas9-DNMT3A tool to the *BACH2* and *IL6ST* regulatory regions. The human (**A**) *BACH2* and (**B**) *IL6ST* loci are shown with positions of the pyrosequencing assays (BACH2-A assay, BACH2-B assay and IL6ST-A assay) indicated by blue rectangles. Magnified insets show individual CpG sites analyzed by pyrosequencing, with arrows (aligned to the magnified regions) indicating 20 bp binding sites of sgRNAs used to guide the dCas9-DNMT3A construct. Arrows point toward the PAM sequence.

pyrosequencing of the targeted regions. Targeting the construct to different positions and in different orientations relative to the pyrosequencing assays (and the CpG sites covered by those assays, Figure 2) enabled us to characterize the summary DNA methyltransferase activity profile of the dCas9–DNMT3A construct (Figure 3). We used results from multiple experiments with different combinations of binding sites and the 'windows' where methylation levels were assessed, thus covering regions about 100 bp upstream and downstream from the binding site in an overlapping and redundant manner. The genomic regions were initially unmethylated and no significant increase in methylation level was detected with either of the two negative controls, represented by non–targeting sgRNA combined with the active dCas9–DNMT3A or by sequence-specific sgRNAs with inactive dCas9–DNMT3A–ANV constructs (Figure 3A, Supplementary Figure S3). The experiments using non–targeting sgRNA showed a non–significant increase in methylation level of 1.71 ± 3.28% compared to mock-transfected cells. The summary of the CpG methylation level increase for all sgRNAs aligned relative to their binding sites and orientations (Figure 3B) showed peak activity of about 35% on average, reaching up to 60% for some individual targeted CpG sites. We consistently found a 25–35 bp wide peak of methylation activity centered at 27 bp downstream from the protospacer adjacent motif (PAM) sequence. A much lower peak of methylation activity could be detected approximately at the same distance upstream from the sgRNA binding site. Inactive (dCas9–DNMT3A–ANV) constructs did not induce any statistically significant differences in CpG methylation levels compared to mock–transfected cells (Student's *t*–test, $\alpha = 0.95$ corrected for multiple testing), although a slight increase paralleling the peak of activity for the active constructs can be observed on the graph (Figure 3B, dotted curve).

Relative positions and orientations of individual sgRNAs in reference to the pyrosequencing assays within the *BACH2* and *IL6ST* genes are depicted in Figure 2. The profile did not critically depend on any single experiment or a particular sgRNA, which we validated by omitting each single sgRNA in turn and constructing the activity profiles, which were similar to each other and showed only a minimal deviation from the summary profile (Supplementary Figure S4).

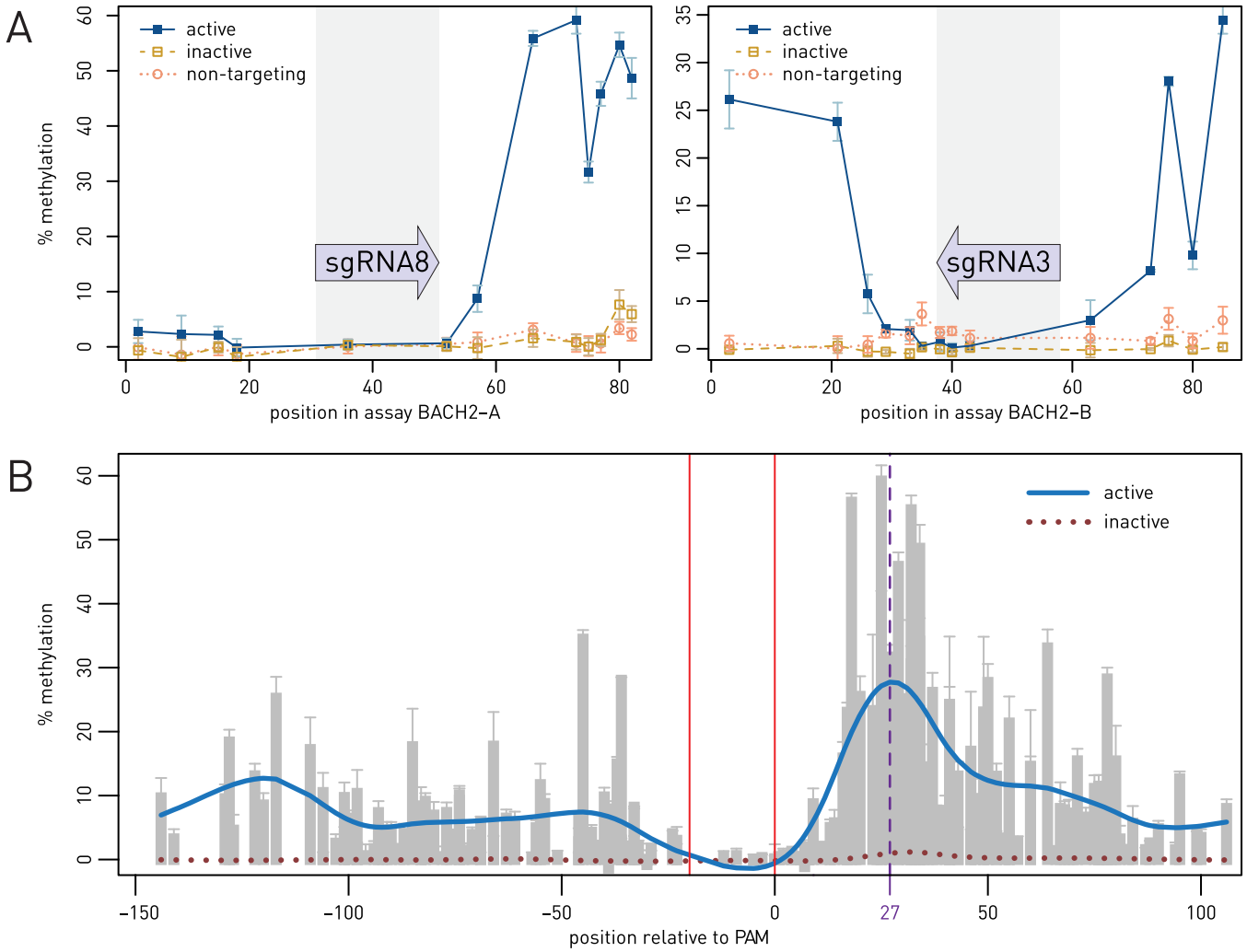

Figure 3. (**A**) Activity of the dCas9-DNMT3A tool guided by sgRNA8 or sgRNA3 was quantified using the BACH2-A and BACH2-B assays, respectively. The graphs show increase in CpG methylation level relative to the mock–transfected cells. Increase in the methylation level for active constructs could be observed some distance downstream from the PAM sequence (e.g. sgRNA8) or at both sides of the binding site (e.g. sgRNA3) but not at the binding site itself (shaded region indicated with arrows labeled with sgRNA names). Inactive (dCas9–DNMT3A–ANV) or non-targeting (sgRNA with no binding site in the human genome) did not show any significant increase in methylation. (**B**) Summary profile of the dCas9–DNMT3A activity is shown as absolute methylation fraction increase (compared to mock–transfected cells) relative to the distance of a CpG site from the PAM sequence. The summary activity profile is based on the activity of all sgRNAs targeting the *BACH2* promoter and the *IL6ST* promoter. Vertical solid red lines represent the binding region complementary to the sgRNA. Results of all experiments were integrated by orienting the sgRNAs in the same direction and aligning the PAM sequence to position zero. Different relative positions of binding sites and pyrosequencing assays enabled construction of an activity profile covering a wide region when all available experimental data were used. The peak of CpG methylation activity extends over about 25–30 nucleotide pairs centered at the 27th nucleotide (vertical dashed purple line) downstream from the targeted PAM sequence. Another much smaller peak was consistently observed at the approximately same distance upstream from the sgRNA binding site. Gray bars represent the CpG methylation level increase observed within a single experiment (error bars show standard deviation). Each gray bar summarizes data for one CpG site position and also serves as a visual guide showing the density of coverage with experimental points. The blue curve shows LOESS smoothing of the data from multiple experiments. The brown dotted curve shows smoothed data for CpG methylation level increase by the inactive construct.

Targeted CpG methylation of the *IL6ST* and *BACH2* promoters using the dCas9-DNMT3A tool reduces gene expression

The locus *IL6ST*, located on a chromosome different from the one containing the *BACH2* locus, was targeted with dCas9-DNMT3A construct in HEK293 cells where the *IL6ST* gene promoter is unmethylated. When cells were transfected with four different sgRNAs independently, the methylation level at the targeted CpG sites in the *IL6ST* promoter region was elevated up to 35% as confirmed by pyrosequencing of this region using the assay IL6ST–A (Figure 4A). Co–transfection of pooled sgRNAs 1–4 resulted in markedly larger increase in methylation level compared to the individual sgRNAs reaching up to 55% at some CpG sites. Inactive constructs (used either individually or pooled) had no detectable impact on CpG methylation (Figure 4A). Pooled sgRNAs 1–4 targeting the *IL6ST* promoter region paired with the active dCas9–DNMT3A construct elicited statistically significant decrease in *IL6ST* mRNA level compared to mock-transfected cells, as de-

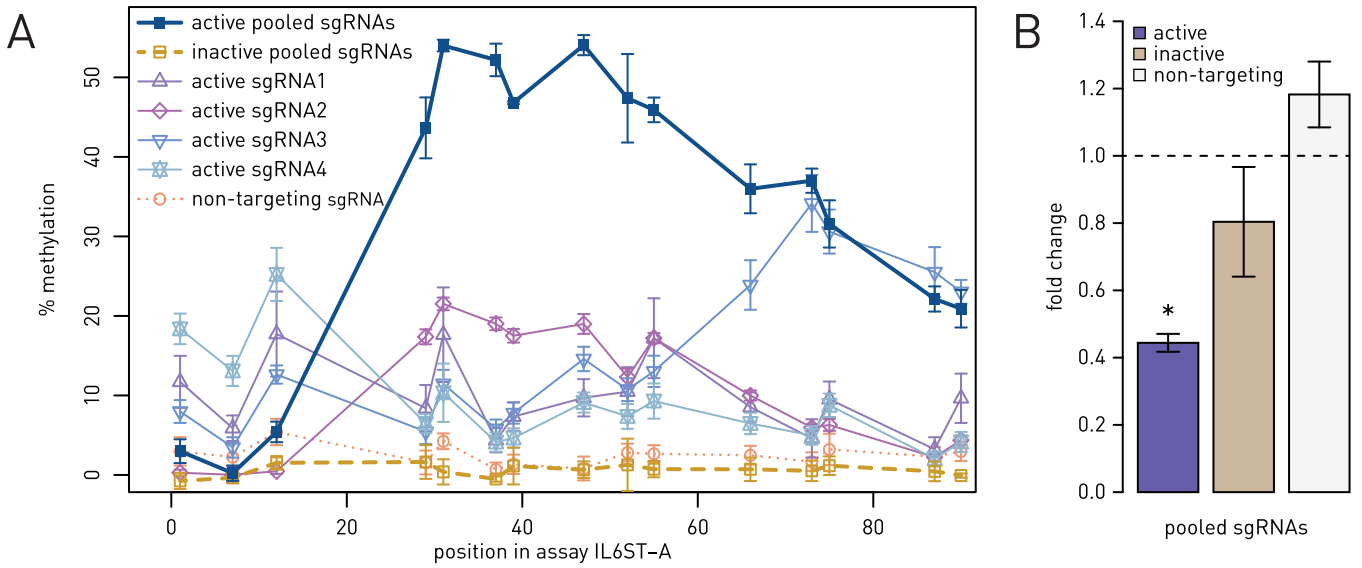

Figure 4. Targeted CpG methylation and transcriptional silencing of the *IL6ST* gene by the dCas9–DNMT3A tool. (**A**) Increase in CpG methylation level relative to the mock–transfected cells in the *IL6ST* promoter region targeted by either individual sgRNAs (1–4) or by pooled sgRNAs 1–4. (**B**) Expression level of the *IL6ST* gene as measured using RT–qPCR revealed a statistically significant decrease ($P < 0.05$) in the transcript level following transfection with pooled sgRNAs 1–4. Fold change is relative to mock-transfected cells. Error bars represent standard deviation. Non-targeting sgRNA served as negative control.

tected by qPCR (Student's t–test, $P = 0.028$, Figure 4B). No statistically significant decrease in mRNA level was detected with pooled sgRNAs 1–4 paired with the inactive construct ($P = 0.209$). Similarly, active constructs with non–targeting sgRNA did not decrease *IL6ST* gene expression ($P = 0.230$).

Similar results were obtained when we targeted the *BACH2* promoter with pooled sgRNAs 1–8. As outlined in Figure 2A, sgRNAs 6–8 bind in the region covered by the pyrosequencing assay BACH2–A, while sgRNAs 1–5 bind in the region covered by the BACH2–B pyrosequencing assay. Pooled sgRNAs showed increased total CpG methylation activity (up to 64% at some CpG sites) in both assayed regions when compared with individual sgRNAs (Figure 5A and B). Catalytically inactive and non-targeting constructs did not show any increase in methylation levels. When we tested expression level of the *BACH2* gene following transfection with pooled sgRNAs, we found a statistically significant decrease in expression level (compared to mock-transfected cells) with the active construct ($P = 0.018$) and some decrease in expression level (still statistically significant, $P = 0.032$) with the inactive construct (Figure 5C). No statistically significant change in expression level was observed with the non-targeting control ($P = 0.217$, Figure 5C).

Methylation level of LINE-1 elements, which was used for estimation of the global genome methylation, showed no change when we transfected cells with either individual or pooled sgRNAs (Figure 6), indicating absence of unspecific genome-wide CpG methylation activity.

Methylation level at the targeted loci peaks at 25–55% on days 6 or 7 in transfected HEK293 cells

Changes in methylation level were followed during 42 days after transfection of HEK293 cells with the dCas9-DNMT3A construct. The profile of the methylation activity is summarized in Figure 7A. The initial increase in methylation level was rapid, reaching its maximum at days 7 (for the *BACH2* gene, Figure 7A, left panel) or 6 (for the *IL6ST* gene, Figure 7A, right panel), then slowly decreasing in the following 10 to 15 days. In the days 20 to 42 after the transfection with the dCas9-DNMT3A construct, a low basically unchanging level of increased methylation could be detected.

Number of HEK293 cells containing the dCas9-DNMT3A construct decreased within 10 days after transfection

Relative amount of plasmid DNA determined by qPCR. Number of copies of the Cas9-encoding plasmid (active or inactive dCas9–DNMT3A construct) relative to the number of genomes of the host HEK293 cells in the 16 days following transfection is shown in Figure 7B. The number of plasmid molecules per cell rapidly decreased until day 7 and by day 9 dropped essentially to zero. There was no statistically significant difference between the active and the inactive constructs.

Fluorescence of the dCas9-DNMT3A-EGFP construct. Active and inactive constructs expressing dCas9-DNMT3A-EGFP fusion protein were targeted to different genomic sites (the *BACH2* and *IL6ST* loci) with the active construct co–expressing non–targeting guide RNA used as an additional negative control. EGFP fluorescence in individual cells and the number of cells showing fluorescence were measured during 14 days. We observed a trend

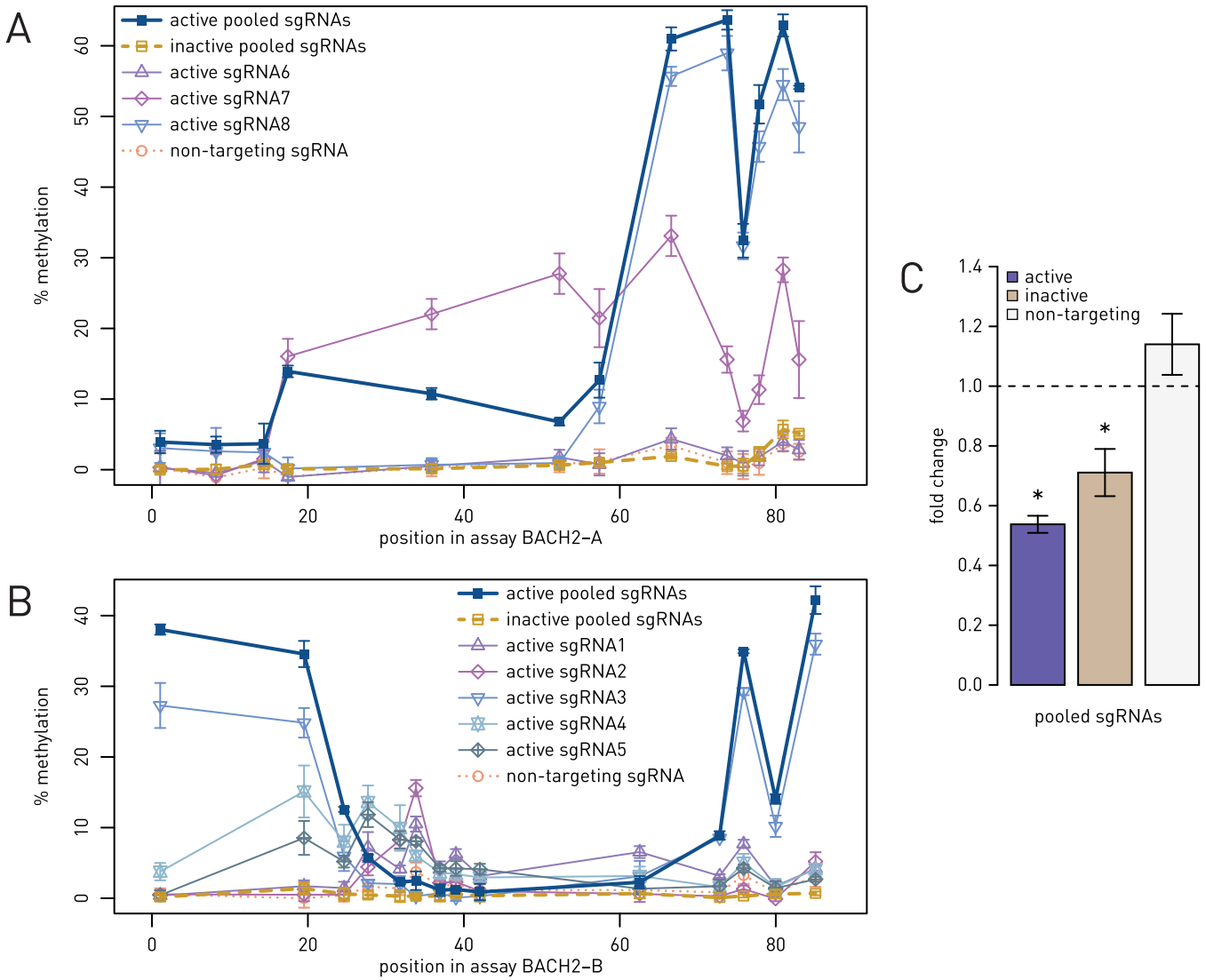

Figure 5. Targeted CpG methylation and transcriptional silencing of the *BACH2* gene using the dCas9–DNMT3A tool. (**A**) CpG methylation level increased relative to mock-transfected cells in the *BACH2* promoter region when targeted by either individual sgRNAs (6–8) or pooled sgRNAs 1–8. The region is covered by the pyrosequencing assay BACH2–A. Note that sgRNAs 6–8 bind close to the region covered by the BACH2–A assay, while sgRNAs 1–5 bind further downstream (see also Figure 2 for reference). (**B**) CpG methylation level increased relative to mock-transfected cells in the *BACH2* promoter region (covered by the pyrosequencing assay BACH2–B) when targeted by either individual sgRNAs (1–5) or pooled sgRNAs 1–8. Note that sgRNAs 1–5 bind close to the region covered by the BACH2–B assay, while sgRNAs 6–8 bind further upstream (see also Figure 2 for reference). (**C**) Expression level of the *BACH2* gene as measured using RT–qPCR revealed a statistically significant ($P < 0.05$) decrease in the transcript level following transfection with dCas9–DNMT3A and pooled sgRNAs 1–8. A lower but statistically significant decrease in expression was observed with matching pooled sgRNAs 1–8 used with inactive dCas9–DNMT3A constructs, which is consistent with CRISPR interference. Fold change is relative to mock-transfected cells. Error bars represent standard deviation. Non-targeting sgRNA served as negative control.

of increased expression for all constructs from day 1 to day 6. After day 6 expression of EGFP per cell did not change (Figure 7C). Importantly, both active and inactive isoforms were expressed equally, suggesting that EGFP turnover is not affected by the inactivating mutation in DNMT3A. We detected a sharp decrease in the proportion of EGFP positive cells, which decreased to almost zero by day 10 (Supplementary Figure S5). Still, several individual cells were observed showing EGFP fluorescence signal even at day 23 (data not shown).

Immunofluorescence. Cells transfected with either active or inactive BACH2-sgRNA8 plasmids collected on days 1, 3 and 6 showed green immunoflourescent signal indicating Cas9 expression (Supplementary Figure S6). About 1.5% of the cells collected on day 1 showed detectable fluorescent signal. No fluorescence was detected in the mock-transfected cells. We found 7% or 4% of transfected cells showing immunofluorescent signal on days 3 and 6, respectively, while the negative control (mock transfection) remained without detectable fluorescence. There was no statistically significant difference in fluorescence intensity either between cells transfected with plasmids carrying ac-

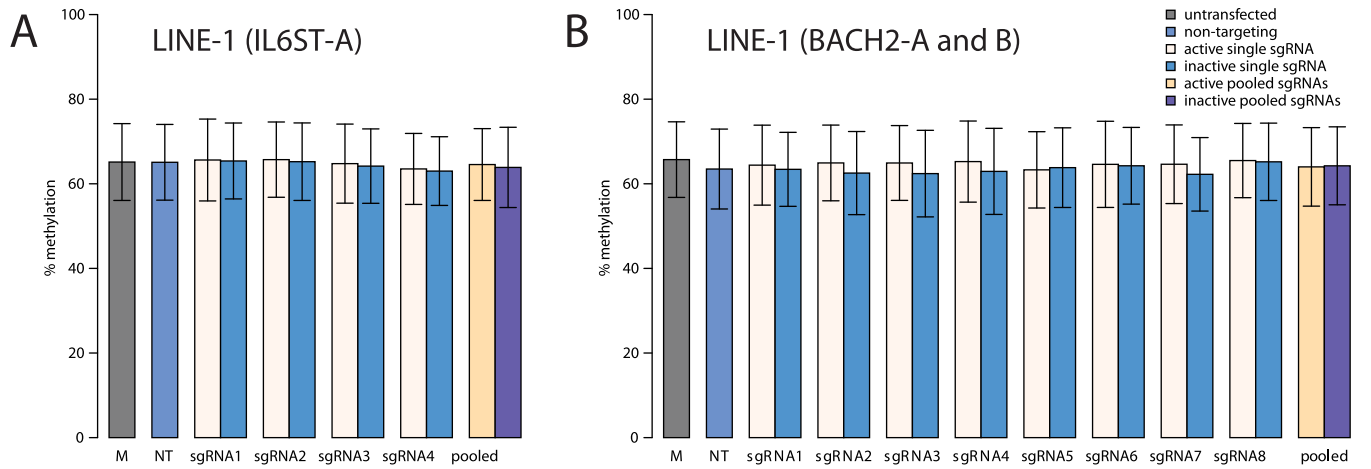

Figure 6. Methylation level of LINE-1 elements reflecting genome-wide methylation. There was no difference in LINE-1 methylation between the mock-transfected cells (M) and those transfected with non–targeting (NT) sgRNA or any of the sgRNAs targeting (**A**) the *IL6ST* promoter region or (**B**) the *BACH2* promoter region. Same results were obtained for either individual or pooled sgRNAs. For targeted constructs, the lighter shaded columns represent the active construct, while the darker columns correspond to controls without methyltransferase activity. Columns represent the mean methylation level across 6 CpG sites. Error bars show standard deviation.

tive and inactive DNMT3A catalytic domain sequences or between cells collected at different time points. No fluorescence could be observed on days 7, 10 and 15 in either group.

Fluorescent in situ hybridization. Each cell on a microscope slide prepared at each time point showed two or more green fluorescent signals located in the nucleus, representing the centromeric region of the chromosome 7. Less than 10% of all cells collected on days 1, 3 and 6, transfected with either active or inactive BACH2-sgRNA8 plasmid showed between 4 and 14 additional red signals specific for the Cas9 DNA located in both the cytosol and the nucleus (Supplementary Figure S7). We could also observe cell divisions with plasmid molecules being distributed between daughter cells (Supplementary Figure S7). There was no statistically significant difference in the number of plasmid signals between cells transfected with active or inactive BACH2-sgRNA8 construct collected on days 1 and 3. On day 6, only several individual transfected cells showed between 4 and 9 plasmid signals. Cells collected on days 7, 10 and 15 transfected either with active or inactive BACH2-sgRNA8 plasmids, as well as mock transfected cells collected at all time-points, showed no detectable plasmid signals.

Western blot analysis. We detected Cas9 protein in cells transfected with the inactive construct on days 1, 3 and 7 following transfection and no expressed protein could be found on days 11 and 15 following transfection (Supplementary Figure S8). Mock-transfected cells used as negative control did not show any Cas9 signal.

DISCUSSION

We have constructed a CRISPR-Cas9-based tool to specifically increase CpG methylation of a targeted DNA sequence in the genome of mammalian cells. The CRISPR-Cas9 technology enables easy and multiplexed targeting of genomic DNA sequences thus facilitating epigenetic silencing by specific methylation of gene promoters or other regulatory elements.

The CRISPR-Cas9 system can target any 20 bp genomic DNA sequence which is followed by the PAM sequence 5′-NGG-3′, which occurs on average every 8–12 bp in the human genome (2). While the Cas9 nuclease cleaves 3 bp upstream from the PAM sequence within the region complementary to sgRNA, we have found the highest CpG methylation activity of the dCas9–DNMT3A construct centered at 27 bp downstream from the PAM sequence, which is in line with the predicted structure of the fusion protein, where the C–terminally fused DNMT3A catalytic domain protrudes from the PAM–interacting (PI) domain of Cas9 facing the side downstream from the PAM (Figure 1A). Activity was virtually absent at the sequence bound to the dCas9 part of the construct, which was expected because the bound DNA should be inaccessible to the DNMT3A catalytic domain. Similarly to a recently developed TALEN-based DNMT3A construct (11), the dCas9–DNMT3A fusion showed some level of activity in a wide area around the binding site. Interestingly, there was also a detectable increase in CpG methylation activity at about 25 bp upstream from the sgRNA binding sequence. This secondary activity could be explained by a limited interaction of the DNA molecule upstream from the dCas9 binding site with the DNMT3A catalytic domain, which was oriented pointing downstream from the PAM sequence (Figure 1) concentrating its activity on that side. Additionally, co-operative binding with other cell-endogenous DNMTs cannot be ruled out.

Non–targeting sgRNA did not induce an increase in CpG methylation level in the sequences analyzed by pyrosequencing. Those sequences were methylated only when properly targeted by specific complementary sgRNAs, thus demonstrating dependence of dCas9–DNMT3A activity on specific targeting. The inability of the inactive (DNMT3A–ANV) constructs to increase methylation level at targeted sequences clearly demonstrated that the increase in methy-

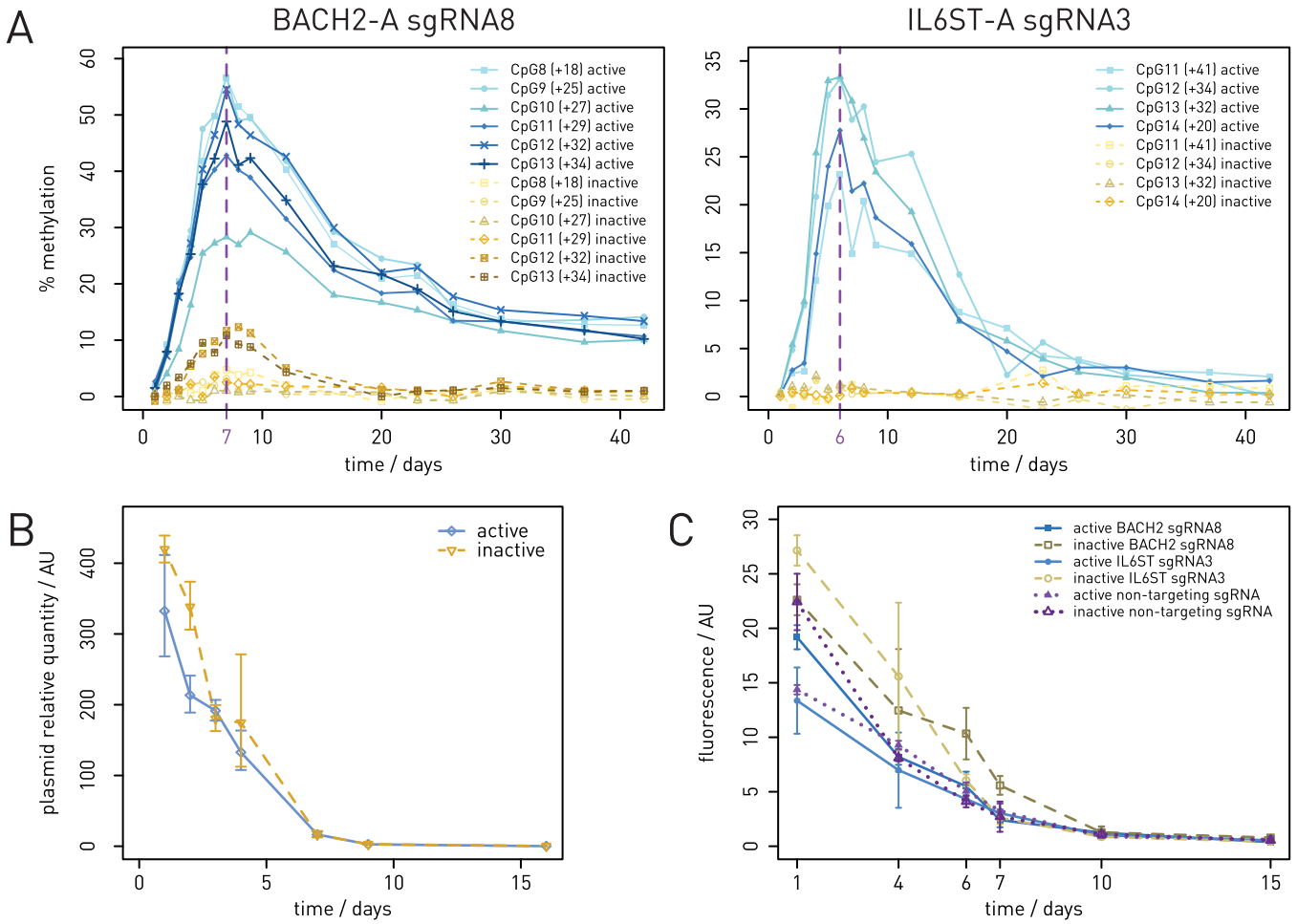

Figure 7. (**A**) Time-course evaluation of targeted CpG methylation in cells transfected with active or inactive dCas9-DNMT3A construct co-expressed with individual targeting sgRNAs. CpG methylation level increase was measured at different time points after transfection over a time period of 40 days. We analyzed CpG sites falling within the peak of methylation activity 18–42 bp downstream from the PAM sequence (positions of CpG sites relative to the PAM sequence are given in brackets). The time point when DNA methylation reaches its maximum is marked with a purple dashed vertical line. Left panel: Methylation of the BACH2-A fragment in cells co-expressing BACH2-sgRNA8. Right panel: Methylation of the IL6ST-A fragment in cells co-expressing IL6ST-sgRNA3. (**B**) Relative amount of plasmid DNA per cell, as determined by qPCR, decreases sharply in the 10 days following transfection. (**C**) Even without puromycin selection (using the construct co–expressing EGFP instead of puromycin resistance gene), the detected expression of the Cas9-DNMT3A construct (measured via co-expression of EGFP) decreases within 10 days after transfection. Arbitrary units (AU) represent fluorescence intensity per field of view normalized to the total number of cells counted under visible light.

lation level was indeed conferred by the enzymatic activity of the DNMT3A catalytic domain. The slight on–target methylation activity of the catalytically inactive construct (Figure 3B, dotted curve) could be explained by dimerization of the inactive DNMT3A catalytic domain from the construct with an active cell–endogenous interaction partner recruited from the nucleoplasm along with two DNMT3L subunits. Building of such tetrameric complex is essential for the observed high level of DNA methylation by the active construct (29). Dimerization of the inactive DNMT3A catalytic domain and the active protein in the nucleoplasm is a plausible explanation for the observed residual DNA methylation activity (43). We ruled out any ncRNA-like or other source of methylation activity distinct from the DNMT3A catalytic domain by conducting additional experiments with a construct lacking the DNMT3A domain (Supplementary Figure S3C).

To show that overexpression of a DNA (cytosine-5-)-methyltransferase alone does not lead to unspecific global genome methylation, we tested LINE–1 methylation levels in cells transfected with active and non-targeting constructs (expressing the active catalytic DNMT3A domain) and found no difference when compared to the inactive construct or the mock-transfected cells (Figure 6). The LINE–1 methylation is an established method for assessing global genome methylation (44), although newer and more accurate methods for this purpose, such as reduced representation bisulfite sequencing (RRBS) or methylation arrays, would be more appropriate to rule out any unspecific activity.

In order to prove the concept of epigenetic silencing by targeted CpG methylation using the dCas9–DNMT3A construct, we analyzed transcription levels of the *IL6ST* and *BACH2* genes in the cells where we used the dCas9-DNMT3A tool for targeted methylation within their CpG

islands. Since we had no prior information about the relevant CpG sites regulating the *IL6ST* and *BACH2* gene transcription, we used all available sgRNAs simultaneously (pooled sgRNAs) in an attempt to increase the methylation level of a larger part of the CpG island located within the *IL6ST* or *BACH2* gene promoters. The approach based on broad targeting using multiple sgRNAs seems to exhibit a synergistic effect on CpG methylation, since it resulted in a much higher overall increase of the methylation level in the whole targeted region although the total amount of transfected plasmid DNA was equal to that used for each experiment with a dCas9-DNMTA plasmid expressing a single sgRNA. Interestingly, the methylation level increase in the *IL6ST* targeted region elicited with pooled sgRNAs (compared to individual sgRNA 1, 3 and 4) was lower in the region where sgRNA2 binding was expected, indicating possible steric interference and hinting at stable binding (Figures 2 and 4A, see positions 0 to 10 in Figure 4A). The more than 2-fold decrease in the *IL6ST* gene transcript level following promoter methylation clearly demonstrated that we could induce gene silencing by increasing DNA methylation using the dCas9-DNMT3A construct guided by multiple sgRNAs, with the magnitude of the silencing effect similar to that induced using other DNMT3A–based tools published previously (9,10). Similar results were obtained for the *BACH2* locus. Inactive constructs targeted to the *IL6ST* promoter region with multiple sgRNAs did lead to a slightly lower average transcript level, although the change was not statistically significant. The effect with inactive constructs targeting the *BACH2* promoter region was of similar magnitude as in the case of *IL6ST*, but statistically significant. Even when lower transcript levels were confidently detected after using the inactive construct, they could be explained by CRISPR interference (16), where inactive Cas9 competes with transcription factors for binding to the promoter region. The relatively higher number of sgRNA targets in the *BACH2* promoter compared to the *IL6ST* promoter (8 versus 4, respectively) might account for the higher statistical significance of gene silencing with inactive construct in the case of *BACH2*. Still, the level of silencing with the inactive construct was consistently lower than for the active construct (fold change ∼0.5 for active and ∼0.75 for inactive construct).

A Cas9-based construct is easily targeted to almost any position in the genome by selecting the appropriate sgRNA complementary to the desired binding site. This is much easier than targeting similar TALE- or ZF-based constructs where a fusion protein needs to be specifically designed for each targeted sequence. However, the most important advantage of targeting by sgRNAs is the possibility of efficient targeting to multiple sites simultaneously. The ability to target many adjacent CpG sites seems to have a synergistic effect on methylation increase of a larger genomic region (Figure 4A) and may be essential for efficient silencing of gene expression, as we have demonstrated with both the *IL6ST* and *BACH2* loci. In addition, multiple targeting could be used to target several distinct gene promoters simultaneously.

CpG methylation achieved using the active dCas9-DNMT3A construct was not stable in cultured cells, as shown in the time course experiment. We found a 'window' of high methylation activity between days 5 and 15 after transfection when functional studies can be carried out. Cultured cells lost the construct after about 10 days following transfection with the dCas9-DNMT3A plasmid. Changes in CpG methylation level followed (with a delay of several days) the level of dCas9-DNMT3A plasmid as determined by qPCR (Figure 7B) or by FISH (Supplementary Figures S6 and S7), which was in turn closely mirrored by the fusion protein expression as determined by EGFP fluorescence (Figure 7C) or immunofluorescence (Supplementary Figure S6). Immunofluorescence results were confirmed by Western blot analysis which detected Cas9 expression on days 1 to 7 following transfection, but not on day 11 and later (Supplementary Figure S8). This also provided an additional control of expression of the inactive construct, confirming that the lack of activity was not due to the lack of expression in case of the inactive construct. The number of cells showing co–expression of Cas9 and EGFP from the construct as determined either by EGFP fluorescence (without puromycin selection) or by immunofluorescence detecting Cas9 protein (after puromycin selection) seemed to decrease, but the signal intensity of individual positive cells remained at a constant level. Yet, it seemed (as shown by FISH analysis) that the total number of plasmid molecules per transfected cell decreases with each cell division. This could be explained by strong expression of the construct and sgRNA that remained at a high level as long as there was a single plasmid molecule in the cell.

As demonstrated by experiments including constructs with EGFP fluorescence, transfection efficiency is limited, which is probably due to the construct's size of almost 10 kilobases. This highlights the importance of selection, either by antibiotic or fluorescence-based sorting. The relatively high methylation activity obtained in our experiments demonstrates effectiveness of puromycin selection.

In this work we have created and characterized a novel CRISPR-Cas9-based epigenome editing tool, the dCas9-DNMT3A, which enabled targeted and specific CpG methylation at the promoter of two loci, the *BACH2* and the *IL6ST*. The dCas9-DNMT3A construct was easily targeted to single as well as to multiple sites by co-expressed sgRNA molecules, thus changing methylation level at a wider region within the CpG island of both the *BACH2* and *IL6ST* promoters. We have demonstrated the ability of the dCas9-DNMT3A construct to silence gene expression – therefore, this tool complements other recently developed DNMT3A-based tools (9–11) with potential for more efficient epigenome editing. Since the *BACH2* and *IL6ST* loci were previously associated with IgG glycosylation and inflammatory as well as autoimmune diseases in a GWAS study (33), the dCas9-DNMT3A construct, possibly combined with other CRISPR-Cas9 tools (such as CRISPR-Cas9-based acetyltransferase and others), could be useful in our prospective studies of epigenetic regulation of *N*–glycosylation and its implications in development of these diseases.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

European Commission projects IBD-BIOM [contract #305479]; RegPot Integra-Life [contract #315997]; Croatian national science foundation project EpiGlycoIgG [Epigenetic regulation of IgG glycosylation; contract #3361]. Funding for open access charge: European Commission projects IBD-BIOM [contract #305479]; RegPot Integra-Life [contract #315997]; Croatian national science foundation project EpiGlycoIgG [Epigenetic regulation of IgG glycosylation; contract #3361].

Conflict of interest statement. None declared.

REFERENCES

1. Doyle,E.L., Stoddard,B.L., Voytas,D.F. and Bogdanove,A.J. (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. *Trends Cell Biol.*, **23**, 390–398.
2. Ran,F.A., Hsu,P.D., Wright,J., Agarwala,V., Scott,D.A. and Zhang,F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.*, **8**, 2281–2308.
3. Urnov,F.D., Rebar,E.J., Holmes,M.C., Zhang,H.S. and Gregory,P.D. (2010) Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.*, **11**, 636–646.
4. Carlson,D.F., Tan,W., Lillico,S.G., Stverakova,D., Proudfoot,C., Christian,M., Voytas,D.F., Long,C.R., Whitelaw,C.B. and Fahrenkrug,S.C. (2012) Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 17382–17387.
5. Dahlem,T.J., Hoshijima,K., Jurynec,M.J., Gunther,D., Starker,C.G., Locke,A.S., Weis,A.M., Voytas,D.F. and Grunwald,D.J. (2012) Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.*, **8**, e1002861.
6. Falahi,F., Huisman,C., Kazemier,H.G., van der Vlies,P., Kok,K., Hospers,G.A. and Rots,M.G. (2013) Towards sustained silencing of HER2/neu in cancer by epigenetic editing. *Mol. Cancer Res.*, **11**, 1029–1039.
7. Mendenhall,E.M., Williamson,K.E., Reyon,D., Zou,J.Y., Ram,O., Joung,J.K. and Bernstein,B.E. (2013) Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.*, **31**, 1133–1136.
8. Hilton,I.B., D'Ippolito,A.M., Vockley,C.M., Thakore,P.I., Crawford,G.E., Reddy,T.E. and Gersbach,C.A. (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.*, **33**, 510–517.
9. Nunna,S., Reinhardt,R., Ragozin,S. and Jeltsch,A. (2014) Targeted methylation of the epithelial cell adhesion molecule (EpCAM) promoter to silence its expression in ovarian cancer cells. *PloS One*, **9**, e87703.
10. Rivenbark,A.G., Stolzenburg,S., Beltran,A.S., Yuan,X., Rots,M.G., Strahl,B.D. and Blancafort,P. (2012) Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics*, **7**, 350–360.
11. Bernstein,D.L., Le Lay,J.E., Ruano,E.G. and Kaestner,K.H. (2015) TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts. *J. Clin. Invest.*, **125**, 1998–2006.
12. Chen,H., Kazemier,H.G., de Groote,M.L., Ruiters,M.H., Xu,G.L. and Rots,M.G. (2014) Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. *Nucleic Acids Res.*, **42**, 1563–1574.
13. Maeder,M.L., Angstman,J.F., Richardson,M.E., Linder,S.J., Cascio,V.M., Tsai,S.Q., Ho,Q.H., Sander,J.D., Reyon,D., Bernstein,B.E. *et al.* (2013) Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat. Biotechnol.*, **31**, 1137–1142.
14. Konermann,S., Brigham,M.D., Trevino,A.E., Hsu,P.D., Heidenreich,M., Cong,L., Platt,R.J., Scott,D.A., Church,G.M. and Zhang,F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, **500**, 472–476.
15. Perez-Pinera,P., Kocak,D.D., Vockley,C.M., Adler,A.F., Kabadi,A.M., Polstein,L.R., Thakore,P.I., Glass,K.A., Ousterout,D.G., Leong,K.W. *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods*, **10**, 973–976.
16. Qi,L.S., Larson,M.H., Gilbert,L.A., Doudna,J.A., Weissman,J.S., Arkin,A.P. and Lim,W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.
17. Barres,R., Yan,J., Egan,B., Treebak,J.T., Rasmussen,M., Fritz,T., Caidahl,K., Krook,A., O'Gorman,D.J. and Zierath,J.R. (2012) Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metab.*, **15**, 405–411.
18. Tsaprouni,L.G., Yang,T.P., Bell,J., Dick,K.J., Kanoni,S., Nisbet,J., Vinuela,A., Grundberg,E., Nelson,C.P., Meduri,E. *et al.* (2014) Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics*, **9**, 1382–1396.
19. Heerboth,S., Lapinska,K., Snyder,N., Leary,M., Rollinson,S. and Sarkar,S. (2014) Use of epigenetic drugs in disease: an overview. *Genet. Epigenet.*, **6**, 9–19.
20. Rius,M. and Lyko,F. (2012) Epigenetic cancer therapy: rationales, targets and drugs. *Oncogene*, **31**, 4257–4265.
21. 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. *et al.* (2013) A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*, **12**, 238–251.
22. Strauss,A. and Lahaye,T. (2013) Zinc fingers, TAL effectors, or Cas9-based DNA binding proteins: what's best for targeting desired genome loci? *Mol. Plant*, **6**, 1384–1387.
23. Li,F., Papworth,M., Minczuk,M., Rohde,C., Zhang,Y., Ragozin,S. and Jeltsch,A. (2007) Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res.*, **35**, 100–112.
24. Valton,J., Dupuy,A., Daboussi,F., Thomas,S., Marechal,A., Macmaster,R., Melliand,K., Juillerat,A. and Duchateau,P. (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem.*, **287**, 38427–38432.
25. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
26. Kungulovski,G., Nunna,S., Thomas,M., Zanger,U.M., Reinhardt,R. and Jeltsch,A. (2015) Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin*, **8**, 12.
27. Oka,M., Rodic,N., Graddy,J., Chang,L.J. and Terada,N. (2006) CpG sites preferentially methylated by Dnmt3a in vivo. *J. Biol. Chem.*, **281**, 9901–9908.
28. Chen,Z.X., Mann,J.R., Hsieh,C.L., Riggs,A.D. and Chedin,F. (2005) Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *J. Cell. Biochem.*, **95**, 902–917.
29. Jia,D., Jurkowska,R.Z., Zhang,X., Jeltsch,A. and Cheng,X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature*, **449**, 248–251.
30. Siddique,A.N., Nunna,S., Rajavelu,A., Zhang,Y., Jurkowska,R.Z., Reinhardt,R., Rots,M.G., Ragozin,S., Jurkowski,T.P. and Jeltsch,A. (2013) Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J. Mol. Biol.*, **425**, 479–491.
31. Lauc,G., Pezer,M., Rudan,I. and Campbell,H. (2015) Mechanisms of disease: The human N-glycome. *Biochim. Biophys. Acta*, doi:10.1016/j.bbagen.2015.10.016.
32. Lauc,G., Vojta,A. and Zoldos,V. (2014) Epigenetic regulation of glycosylation is the quantum mechanics of biology. *Biochim. Biophys. Acta*, **1840**, 65–70.
33. Lauc,G., Huffman,J.E., Pucic,M., Zgaga,L., Adamczyk,B., Muzinic,A., Novokmet,M., Polasek,O., Gornik,O., Kristic,J. *et al.* (2013) Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.*, **9**, e1003225.
34. Trbojevic Akmacic,I., Ventham,N.T., Theodoratou,E., Vuckovic,F., Kennedy,N.A., Kristic,J., Nimmo,E.R., Kalla,R., Drummond,H., Stambuk,J. *et al.* (2015) Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin G glycome. *Inflamm. Bowel Dis.*, **21**, 1237–1247.

35. O'Brien,A. and Bailey,T.L. (2014) GT-Scan: identifying unique genomic targets. *Bioinformatics*, **30**, 2673–2675.
36. Sanjana,N.E., Shalem,O. and Zhang,F. (2014) Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods*, **11**, 783–784.
37. Zoldos,V., Horvat,T., Novokmet,M., Cuenin,C., Muzinic,A., Pucic,M., Huffman,J.E., Gornik,O., Polasek,O., Campbell,H. *et al.* (2012) Epigenetic silencing of HNF1A associates with changes in the composition of the human plasma N-glycome. *Epigenetics*, **7**, 164–172.
38. Schmittgen,T.D. and Livak,K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.*, **3**, 1101–1108.
39. van Dongen,J.J., Langerak,A.W., Bruggemann,M., Evans,P.A., Hummel,M., Lavender,F.L., Delabesse,E., Davi,F., Schuuring,E., Garcia-Sanz,R. *et al.* (2003) Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98–3936. *Leukemia*, **17**, 2257–2317.
40. Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, e45.
41. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
42. Nishimasu,H., Ran,F.A., Hsu,P.D., Konermann,S., Shehata,S.I., Dohmae,N., Ishitani,R., Zhang,F. and Nureki,O. (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, **156**, 935–949.
43. Van Emburgh,B.O. and Robertson,K.D. (2011) Modulation of Dnmt3b function in vitro by interactions with Dnmt3L, Dnmt3a and Dnmt3b splice variants. *Nucleic Acids Res.*, **39**, 4984–5002.
44. Yang,A.S., Estecio,M.R., Doshi,K., Kondo,Y., Tajara,E.H. and Issa,J.P. (2004) A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res.*, **32**, e38.