HHS Public Access

Author manuscript

Nat Biotechnol. Author manuscript; available in PMC 2020 May 11.

Published in final edited form as:

Nat Biotechnol. 2020 January ; 38(1): 50–55. doi:10.1038/s41587-019-0296-7.

Dose-dependent activation of gene expression is achieved using CRISPR and small molecules that recruit endogenous chromatin machinery.

Anna M. Chiarella[1], **Kyle V. Butler**[2], **Berkley E. Gryder**[3], **Dongbo Lu**[1], **Tiffany A. Wang**[1], **Xufen Yu**[2], **Silvia Pomella**[3,4], **Javed Khan**[3], **Jian Jin**[2,#], **Nathaniel A. Hathaway**[1,#]

[1]Division of Chemical Biology and Medicinal Chemistry, Center for Integrative Chemical Biology and Drug Discovery, UNC Eshelman School of Pharmacy, Chapel Hill, North Carolina 27599, United States.

[2]Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences and Oncological Sciences, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA 10029, United States.

[3]Oncogenomics Section, Genetics Branch, National Cancer Institute, National Institutes of Health (NIH), Bethesda, MD 20892, United States.

[4]Department of Pediatric Hematology and Oncology, Bambino Gesu' Children's Hospital, Rome, Italy.

Abstract

Gene expression can be activated or suppressed using CRISPR/Cas9 systems. However, tools that enable dose-dependent activation of gene expression without the use of exogenous transcription regulatory proteins are lacking. Here we describe chemical epigenetic modifiers (CEMs) designed to activate the expression of target genes by recruiting components of the endogenous chromatin activating machinery, eliminating the need for exogenous transcriptional activators. The system has two parts: catalytically inactive Cas9 (dCas9) in complex with FK506 binding protein (FKBP), and a CEM consisting of FK506 linked to a molecule that interacts with cellular epigenetic machinery. We show that CEMs upregulate gene expression at target endogenous loci up to 20-fold or more depending on the gene. We also demonstrate dose-dependent control of transcriptional activation, function across multiple diverse genes, reversibility of CEM activity, and specificity of our best in class CEM genome wide.

#Correspondence should be addressed to N.A.H (hathaway@unc.edu) or J.J. (jian.jin@mssm.edu).

Author contributions:

N.A.H., J.J., K.V.B., and A.M.C. conceived the project. N.A.H., J.J., K.V.B., A.M.C., D.L., B.E.G. and J.K. contributed to the experimental design. N.A.H., J.J., K.V.B., A.M.C., D.L., B.E.G., J.K. , T.A.W., X.Y., and S.P. contributed experimentally. K.V.B. synthesized the compounds. B.E.G. and J.K. carried out ChIP-seq and RNAseq experiments and analysis. N.A.H., J.J., K.V.B., A.M.C., D.L., B.E.G. wrote the manuscript. All authors edited the manuscript.

Competing interests:

N.A.H., J.J., K.V.B., and A.M.C. have filed a patent application for chemical epigenetic modifiers.

The eukaryotic genome is organized and packaged into chromatin with varying degrees of compaction, which contributes to the regulation of gene expression. A network of protein-protein and protein-DNA interactions regulates the proper levels of gene expression. Disruptions to this regulatory network drive many human diseases including cancer[1,2]. An important contributing factor that sculpts the chromatin landscape is post-translational histone tail modification. Lysine acetylation is one such modification that has both biophysical and indirect protein-recruitment effects. Protein families of writers (histone acetyl transferases, HATs), erasers (histone deacetylases, HDACs), and readers (bromodomains, chromodomains, etc.) intricately control gene expression[3,4].

Several groups have demonstrated the power of recruiting exogenous chromatin modifying machinery as a way to control expression levels in a gene-specific manner[5-11]. With major advances in the CRISPR-associated protein 9 (Cas9) and catalytically inactive Cas9 (dCas9) technology, the ability to precisely induce changes in expression has rapidly evolved. Pioneering work by Liszczak and colleagues has demonstrated the ability to recruit endogenous machinery to a reporter locus using a dCas9 system combined with conjugated inhibitors of chromatin regulatory proteins[12]. Other work in the Ansari group used programmable DNA-binding ligands coupled with bromodomain inhibitors to modulate transcription[13]. Inspired by these studies, we sought to develop a system capable of modulating gene expression of endogenous mammalian genes in a specific, dose-dependent manner using chemical entities.

We have previously demonstrated the ability of chemical epigenetic modifiers (CEMs) to modify chromatin and subsequently repress gene expression at engineered reporter loci[14]. In this study, we report CEM activating (CEMa) molecules that recruit endogenous gene activating machinery. Our CEMa family includes CEM87, CEM88, and CEM114 that each bind to different chromatin modifying enzymes from previously published bromodomain inhibitors of HATs or acetylated lysine reader proteins (Fig. 1a). CEM87 was created with iBet762, shown to bind BRD2, BRD3, and BRD4[15] (Fig. 1b, Supplementary Note). CEM88 was created with a 1,3-dimethyl benzimidazolone, previously shown to bind the BRPF1 bromodomain (Fig. 1b, Supplementary Note)[16]. Lastly, CEM114 was created with “compound 33”, previously shown to bind CBP (Fig. 1b, Supplementary Note)[17]. Here, we show our CEMa family is compatible with dCas9-FKBP-based systems, allowing us to direct CEMa activity to any gene.

To test for changes in gene expression, we transfected human embryonic kidney (HEK) 293T cells with a Green Fluorescent Protein (GFP) reporter gene downstream of a Tre3g promoter and performed flow cytometry on cells co-expressing the gRNA and dCas9 machinery[9]. The same targeting gRNA is used throughout our experiments with the Tre3g-GFP reporter and comprises a sequence that binds six interspaced sites within the Tre3g promoter. As a benchmark control for gene activation, we used a dCas9 (*Streptococcus pyogenes*, Sp) protein fused to p300, ten-eleven translocation (TET), or -VP64-p65-Rta (VPR), all previously shown to increase expression (Supplementary Fig. 1a)[11,18,19]. We then tested activating the reporter gene with a plasmid expressing dCas9–FKBPx1 or x2 and three of our predicted activating CEMs. After 48-hrs of exposure with 200 nM of indicated CEM normalized GFP, reporter expression significantly increased in all cases compared to

untreated cells (Fig. 1c). We next tested another commonly used dCas9 system adapted from *Staphylococcus aureus* (Sa). Cells expressing dCas9(Sa)-FKBPx1 or x2 were treated with the indicated CEM. Likewise, cells showed increased GFP expression demonstrating that the CEM technology is adaptable to multiple species of dCas9 (Fig. 1d). We continued further experiments with dCas9(Sp). To confirm that the CEMa system is activating GFP in a controlled, FKBP-dependent manner, we analyzed cells expressing dCas9 alone treated with 200 nM of CEMa for 48-hrs. As predicted, CEM treatment did not significantly change GFP expression (Fig. 1e). We also sought to validate that the activation was a result of the CEMa molecule as a whole, rather than any one component of the molecule. To test this, we expressed cells with dCas9-FKBPx2, and treated the cells with 200 nM of iBet762 (the inhibitor from which CEM87 was synthesized), 200 nM of FK506 (the FKBP-binding moiety), or 200 nM of CEM87. CEM87 treatment was the only condition that increased GFP expression (Fig. 1f).

Optimization of the dCas9-CEMa system was done by incorporating several dCas9-related systems that we adapted from work of others, specifically the “MS2-system” and the “dCas9-SunTag-system”[7,20]. The ms2-gRNAs have a modified stem-loop, capable of recruiting both a dCas9-fusion as well as a bacteriophage MS2 coat protein (MCP)-fusion[7,21]. By using a MCP-FKBP fusion, we were able to increase the number of recruited CEMs to the chromatin. The ms2[N55K] used was a previously published MCP mutant that provides strong binding of the coat protein to the stem loops, referred here on as just ms2[22]. To optimize our CEMa gene activation with the MS2-system, we transfected cells with dCas9-alone, targeting gRNA, and either ms2-FKBPx1 or ms2-FKBPx2. Of these approaches the ms2-FKBPx2 was the most effective (Supplementary Fig. 1b).

We next adapted the dCas9 “SunTag” system expressing an array of 10 or 24 yeast-specific gene control protein 4 (GCN4) peptides from the C-terminus of the dCas9[20,23]. We co-transfected with a single chain variable fragment (scFv), made to be GCN4-specific, fused to one copy of FKBP. Our results showed that the dCas9-SunTagx24 fusion was more effective (Supplementary Fig. 1c). Secondly, we tested dCas9-SunTagx24 with scFv-FKBPx1 or x2. Cells were again treated with 200 nM of indicated CEM. With most of the CEMs, the scFv-FKBPx2 provided a slightly higher advantage over the scFv-FKBPx1 (Supplementary Fig. 1d). Using a combinatorial approach of the SunTag and ms2 system, we created and tested fusions of the MCP to the SunTag peptide, in an attempt to recruit the additional scFv-FKBP fusions closer to the targeted chromatin (Supplementary Fig. 1e).

To directly compare these optimized dCas9 systems, we transfected cells with equal amounts of the gRNA, dCas9-fusion, scFv-fusion, ms2-fusion, or “filler” DNA. Cells were treated with 200nM of CEM87, CEM88, or CEM114 for 48-hrs and mean fluorescent values were normalized to untreated cells with the same transfection mixture (Supplementary Fig. 2a). The same transfection was conducted on cells using non-targeting (NT) gRNA. As expected, the GFP signal was consistently lower in the non-targeting gRNA control cells than in the cells expressing targeting gRNA (Supplementary Fig. 2b). Together, these data suggest that more strategically recruited FKBP fusions, rather than simply the number of FKBP, is important for improving the dCas9-CEMa system. With this in mind, we chose to

proceed with the dCas9-alone, bringing in the CEM binding site only through the ms2-FKBPx2.

With the optimized dCas9 system, we conducted a time course with the two most effective CEMa molecules, CEM87 and CEM114. We transfected cells with dCas9, ms2-FKBPx2, Tre3g-GFP, and either targeting or NT gRNA. Cells were treated with 200 nM of CEM87 at the indicated time prior to flow cytometry. The data was normalized to cells expressing NT gRNA. Compared to untreated cells, CEM87 and CEM114 both yielded high GFP expression after 24- and 48-hrs (Supplementary Fig. 2c, d). We next performed a dose curve with the same CEMa molecules measurement at 48-hrs. CEM87 and CEM114 treatment significantly increased GFP expression in direct response to CEM addition until ~200nM (Fig. 1g, h). Dose curves from CEM87 and CEM114 treatment demonstrate a hook effect near 400 nM – 800 nM treatments, where the CEMa treatment becomes less effective. This could be a result of active inhibition of the desired chromatin regulator machinery meant for recruitment. Alternatively, there could be toxicity at high concentrations (morphology indicating potential cytotoxicity was observed at high concentrations above 400nM, data not shown). Thus, the dCas9-CEMa platform can exhibit control of gene activity by varying compound dose between 6.25 nM and 200 nM in a dose-dependent manner. This could be useful for target validation studies, which need new tools to improve clinical success rates[24,25].

To confirm that CEMa-mediated activation of GFP is a result of CEMa binding directly to the ms2-FKBPx2 protein fusions through an FK506 binding site interaction, we tested if excess FK506 could outcompete CEMa binding, thereby diminishing activation. Cells were treated with 2000 nM of FK506, 200 nM of CEM114, or a combination of 2000 nM FK506 and 200 nM CEM114 for 48-hrs. Following flow cytometry, data was normalized to cells given the same treatment but transfected with NT gRNA. As expected, only the 200 nM CEM114 treatment significantly activated the GFP signal, suggesting that excess FK506 is able to outcompete CEM114 from the FKBP binding site (Supplementary Fig. 2e). We next tested if a combination of CEMa could yield a stronger level of gene activation. We treated cells for 48-hrs, performed flow cytometry and normalized the results to cells expressing NT gRNA. GFP expression was significantly increased in cells treated with 66.6 nM or 200 nM each of CEM87 + CEM88 + CEM114, but neither condition increased expression to the extent of CEM114 alone (Supplementary Fig. 2f). These data suggest that CEM114 yields the most effective activation in this particular assay.

To determine how stable CEM87- and CEM114- activation is without constant CEMa treatment, we exchanged the CEM-media with untreated media or with excess FK506 (2000 nM). As a baseline, cells were treated with 200 nM CEM87 or 200 nM CEM114 for 48-hrs. Flow cytometry was performed on a subset of cells and the data was normalized to untreated cells. The remaining cells were then re-seeded with continued CEMa treatment, treatment of excess FK506 or untreated media. Initially, cells treated with CEM87 or CEM114 for 48-hrs significantly increased GFP expression. After two days of continued treatment, no treatment, or FK506 treatment, the same cells continued to express significantly activated levels of GFP expression. After four days, all conditions continued to have significant activation, except cells treated with CEM87 for two days and then treated with excess FK506 for four

days. After six days of excess FK506 in CEM87 and CEM114 cells, and after six days of no treatment in CEM87 cells, the level of activation was no longer significant (Supplementary Fig. 2g). These data suggest that the dCas9-CEMa system is reversible.

To target endogenous genes, we adapted our optimized dCas9-CEMa system to be delivered with lentiviral infection of the dCas9 machinery (dCas9 and ms2-FKBP) into HEK 293T cells. After the cells had been stably selected for integration of the dCas9/FKBP constructs, we transfected the gRNAs for the desired gene target. As a positive control, we transfected cells expressing dCas9 fused to activating domains with gRNAs targeted to the myogenic differentiation 1 (*MYOD1*) gene, extracted the RNA and performed qRT-PCR (Supplementary Fig. 3a). *MYOD1* was an ideal initial target for this question because it has been previously shown to be capable of modulation by transiently expressed dCas9-p300[11]. To determine the optimal time at which we observe activation, we transfected cells expressing dCas9 and ms2-FKBPx2 with gRNA and conducted RNA extraction after 24, 48, 72, 96, and 120-hrs post treatment. At 48-hrs of treatment, cells with *MYOD1*-targeting gRNA significantly increased expression with 200 nM treatment of CEM87 (22.7-fold, $p< 0.01$). CEM87 continued to yield the greatest level of *MYOD1* overexpression after 48-hrs and 72-hrs. After 96-hrs, CEM87 activation was no longer significant, likely due to high variability and/or plasmid loss after transfection (Fig. 2a, Supplementary Fig. 3b). To determine the concentration at which CEM87 is most effective in this assay, we performed a dose curve with CEM87 concentrations ranging from 50 nM to 400 nM. Cells expressing targeting gRNA showed dose-dependent activation of *MYOD1* after 48-hrs treatments of doses ranging from 50 nM to 400 nM. However, cells expressing NT gRNA showed lower non-specific activation of *MYOD1* after a 48-hr treatment at a dose of 400nM compared to untreated cells (Fig. 2b). Thus, CEM87 is most effective over a concentration range of 0-200nM. With increased activity coupled with more non-specific activation at 400nM, usage over 400nM lead to more variable results (data not shown).

To evaluate the versatility of the dCas9-CEMa system, we tested the dCas9-CEMa system with gRNAs targeting the Interleukin 1 Receptor Antagonist (*IL1RN*) locus, a lowly expressed gene. Compared to untreated cells, 200 nM of CEM87 increased *IL1RN* expression 92.5-fold ($p< 0.05$), while control cells expressing NT gRNA did not significantly change *IL1RN* expression (Supplementary Fig. 4a). We also tested activating the octamer-binding transcription factor 4 (*OCT4*) gene. After 48-hrs of treatment with 200 nM of CEM87, *OCT4* mRNA levels increased 4.9-fold ($p< 0.01$), whereas the levels did not significantly change in the NT gRNA control cells (Supplementary Fig. 4b). Next, we targeted the *MYC1* locus, an area of the genome more highly expressed compared to the other targets. CEM87 did not significantly increase expression in the *MYC1* targeting cells (Supplementary Fig. 4c). As a fourth test of the general utility of the system, we designed a series of gRNAs targeting the super enhancer (SE) network controlling *MYOD1*. We created one set of 4 gRNAs targeting the SE boundaries and a set of 6 gRNAs targeting the epicenters of this SE network. These gRNAs were induced in a rhabdomyosarcoma cancer cell line (RH4) that expresses *MYOD1* and possesses these SE networks (Supplementary Fig. 4d)[26]. After introducing dCas9 and ms2-FKBP into RH4 cells, we virally infected the gRNA sets overnight, replaced media and added CEM87 to recruit additional BRD4. We observed, after 24-hrs of CEM87, that *MYOD1* expression increased significantly, more so

when CEMa was guided to the SE epicenters (Supplementary Fig. 4d). Since BRD4 is already present at these epicenters, as confirmed in bulk ChIP-seq, we envision two possible alternatives: either this population average data masks the underlying heterogeneity of BRD4 binding to this SE, and CEM87 is activating *MYOD1* in cells that have incomplete BRD4 binding, or alternatively, even when a SE has endogenous BRD4 it may not have reached saturation. Taken together, these data suggest that dCas9-CEMa technology can robustly activate low-expressing genes, but to less of an extent highly expressed genes.

To compare RNA expression with single cell protein analysis, we targeted C-X-C chemokine receptor type 4 (*CXCR4*) and analyzed the level of expression in two ways. After 48-hrs of 200 nM CEM87 exposure, we conducted RNA extraction and qRT-PCR which showed increased expression of 12.2-fold compared to untreated cells ($p < 0.05$). The cells expressing the NT gRNA did not show a significant change in *CXCR4* expression (Fig. 2c). In addition, we used a fluorescent CXCR4 antibody and performed flow cytometry to test for changes in protein expression. Compared to untreated cells, the results showed a 5.6-fold ($p < 0.005$) increase in CXCR4, while cells expressing NT gRNA did not show any changes in CXCR4 protein levels (Fig. 2d). To complement our work in HEK 293T cell line models, we also performed a *CXCR4* targeting assay in human colorectal carcinoma HCT116 tissue. The *CXCR4*-targeting gRNAs that we used in the 293T cell line showed modest activation (data not shown). To improve the activation, we tested a different set of *CXCR4*-targeting gRNAs in a multiplexed gRNA backbone. We virally infected the HCT116 cells with dCas9 and ms2-FKBPx2. To use this cell line to benchmark our dCas9-CEMa with two widely used dCas9 based activators, -VPR and -p300, we also virally infected HCT116 cells with dCas9-VPR and dCas9-p300. Cells were transfected with the new *CXCR4*-targeting gRNA or NT gRNA. After 48-hrs treatment with CEM87 and CEM114, we found dCas9-CEMa also had a dose response of activation in this cell line peaking with 200nM CEM87 and increased *CXCR4* expression, which was in the range of what we saw with dCas9-p300 induced change, but less than that achieved with the VPR system (Fig. 2e). All cells expressing NT gRNA did not change *CXCR4* expression significantly. The ability to control expression levels was unique to dCas9-CEMa. For any given new cell application CEMa concentration should be titrated between 0-200nM CEM87.

To investigate the extent of potential off-target or indirect effects of CEM activation by recruitment of BRD4 to the *MYOD1* promoter, we assayed the whole transcriptome (RNA-seq) and genome-wide BRD4 binding profiles (ChIP-seq). BRD4 was recruited to the promoter of *MYOD1* only in the combined presence of both CEM87 and sgMYOD1 (Fig. 3a). BRD4 strongly positively regulates *MYOD1* in myogenic contexts[26], where thousands of downstream targets of MYOD are activated from BRD4-bound *cis*-regulatory elements, yet at 48-hrs *MYOD1* was one of the very few sites with a gain of BRD4 binding (Fig. 3b). Indeed, nearly all other BRD4 sites saw a reduction in binding, a known effect of small molecules engaging BET bromodomains[27]. Transcriptionally, *MYOD1* was selectively upregulated (Fig. 3c). Downstream myogenic targets of MYOD remained silent, potentially because of incomplete dosage and short duration. BRD4 inhibition also strongly upregulates *HEXIM1* regardless of biological context[28], which we observed in the RNA-seq from cells treated with CEM87 (Fig. 3c, d). To validate our ChIP sequencing results, we performed ChIP-qPCR with primers targeting the *MYOD1* locus. Using a primer set 1000 basepairs

(bp) upstream of the *MYOD1* transcriptional start site (TSS), BRD4 enrichment was increased 2.4-fold ($p < 0.05$) in *MYOD1*-targeting compared to NT cells with 48-hrs of 50 nM CEM87 treatment. When we used primers targeting 350 bp upstream of the TSS, BRD4 enrichment was increased 10.7-fold ($p < 0.05$) in targeting versus NT cells with the same treatment. Lastly, using a primer set 50 bp downstream of the TSS, BRD4 enrichment increased 5.5-fold (Supplementary Fig. 5a, $p < 0.05$).

In summary, we have designed, synthesized, and optimized a class of CEMa molecules capable of activating endogenous genes in a dose-dependent, gene-specific manner. By adapting the CEMa technology to dCas9 targeting constructs, we can use this system to theoretically target any gene in the genome by strategic gRNA design. We have demonstrated the ability to control the chromatin landscape and induce changes in the expression of endogenous mammalian, disease-related genes in a direct, biologically-relevant manner. This dCas9-CEMa technology paves the way in targeting disease relevant genes to ask specific, pointed questions about disease mechanisms of action. Because our gene activation platform is chemically regulated in a dose dependent manner, it will be useful in target validation work for visualization of trends between phenotype and gene dosage over a range of gene expression levels.

Methods

Chemical Synthesis:

See Supplementary Note. CEMa compounds were diluted in DMSO (Sigma D2650) and kept dry at −20 °C.

Statistical Analysis:

For all flow cytometry and RNA extractions followed by qRT-PCR, a student's t-test (two-tailed) was used to determine significance. In each experiment, three different cell culture replicates were used. Error bars represent the standard deviations.

Plasmid design:

To create the Sa-dCas9-compatible gRNAs, we used plasmid: Addgene # 64710. To create the Sp-dCas9 ms2-compatible gRNA plasmids, we used plasmid Addgene # 61427. For the multiplexed gRNAs, constructs were modified from Addgene Kit # 1000000055 in order to express gRNAs in a lenti-viral backbone, containing ms2-compatible stem loops, and lacking a Cas9 protein. All targeting gRNA sequences are references in the Supplementary Table 2.

All dCas9-, ms2-, and scFv- fusions were reconstructed to ensure proper comparison between effectors: Ef1α-dCas9(Sp)-HA ; Ef1α-dCas9(Sp)-FKBPx1 ; Ef1α-dCas9(Sp)-FKBPx2 ; Ef1α-dCas9(Sp)-SunTagx10 ; Ef1α-dCas9(Sp)-SunTagx24 ; Ef1α-dCas9(Sp)-p300 ; Ef1α-dCas9(Sp)-Tet ; Ef1α-dCas9(Sp)-VPR ; Ef1α-dCas9(Sa)-FKBPx1 ; Ef1α-dCas9(Sa)-FKBPx2 ; Ef1α-scFv-FKBPx1 ; Ef1α-scFv-FKBPx2 ; Ef1α-ms2[N55K]-FKBPx1 ; Ef1α-ms2[N55K]-FKBPx2 ; Ef1α-ms2[N55K]-SunTagx10 ; Ef1α-ms2[N55K]-

SunTagx24. Constructs were adapted from Hathaway et al., Braun et al., Tanenbaum et al., Thakore et al., Liu et al., Gao et al., and Hilton et al.[7,20,29-33].

Cell culture:

Low passage (p30-40) HEK 293T cells were cultured with high-glucose DMEM (Corning, 10-013-CV) supplemented with 10% FBS serum (Atlantic Biologicals, S11550), 10 mM HEPES (Corning, 25-060-Cl), NEAA (Gibco, 11140-050), Pen/Strep, and 55 μM 2-Mercaptoethanol. Cells were passaged every 2 – 5 days and maintained at 10-90% confluency in an incubator at 37 °C and 5% CO_2.

Cell Transfection:

The day after splitting cells, PEI (polyethlenimine, Polysciences #23966-1) transfection was done with 1:3 ratio of μg of DNA:μL of PEI, and 1:10 volume of μL of Optimem (Gibco, 31985070) to total media volume. Specific ratios of DNA are shown in Supplementary Table 3. For 6 well plates, cells were cultured and transfected with 2 mL of media. For 12 well plates, cells were cultured and transfected with 1 mL of media. For cells grown in a 15 cm plate, 20 mL of media was used. For every transfection experiment, media was changed 16-hrs later.

Flow cytometry:

Flow cytometry was performed with the Attune Nxt as previously described[34]. The data presented represents three separate transfections and significance was determined with the student's t-test. To compare CXCR4 protein levels, we used a protocol using a fluorescently conjugated primary antibody (APC-CXCR4, Biolegend catalogue #306510). After washing with 1X PBS, we incubated the cells with the antibody (8 μL/mL) in the dark at room temperature for 1-hr. An example of the gating strategy can be found at Supplementary Fig. 4f.

Lentiviral infection:

Lentivirus production for HEK 293T infection was done using LentiX 293T cells (Clontech). Low passage cells (8-20) were plated onto 15 cm cells such that they were 70% confluent 24-hrs later. Each plate was transfected with 18 μg of the plasmid of interest (dCas9-X), 13.5 μg of the Gag-Pol expressing plasmid (Addgene #12260), and 4.5 μg of the VSV-G envelope expressing plasmid (Addgene #12259). PEI transfection was done and 60-hrs after transfection, the virus was spun down at 20, 000 rpm for 2 ½-hrs at 4°C and then added to the HEK 293Ts in combination with 10 μg/mL Polybrene (Santa Cruz, sc-134220). The selection of lentiviral constructs was done with either hygromycin (200 μg/mL) or blasticidin (12 μg/mL).

RNA extraction:

For Fig. 2, Fig. 3, Supplementary Fig. 3 Supplementary Fig. 4, cells from the 12-well or 6-well plates were isolated. Cells were washed with 1X PBS, disassociated with 0.05% trypsin, quenched with media, centrifuged and washed with 1X PBS. RNA extraction was performed with a RNeasy Plus Mini Kit (Qiagen, 74134) and the relative enrichment of

mRNA was quantified with the RNA-to-C_T™ 1-step kit (Thermo Fisher Scientific, 4389986). The primers used are shown in Supplementary Fig. 4e.

RNA sequencing:

RNA extracted as above was quantified and profiled for quality using Agilent Bioanalyzer. Then, Poly-A enriched and Illumina barcoded libraries were made and sequenced on an Illumina NextSeq 500, to a depth of 30 million base pairs (150 bp paired end). Reads were mapped to UCSC reference for hg19 using STAR[35], then read abundance for each gene counted (TPM method) using RSEM. Data visualization was performed using custom R scripts (available here: https://github.com/GryderArt/VisualizeRNAseq/tree/master/RNAseq_Pipeline).

ChIP qPCR:

Chromatin isolation was performed on 293 HEK cells as described in Chiarella et al.[34]. Immunoprecipitation was performed with a BRD4 antibody (ab128874). qPCR was performed on the isolated DNA with SYBR green as previously described using the primers described in Supplementary Fig. 5b.

ChIP sequencing:

Chromatin bound to BRD4 was immunoprecipitated from fixed and sheared HEK cell nuclei, with or without CEM87 treatment, in cells bearing sgNT or sgMYOD1 (promoter). Enriched DNA from ChIP was barcoded and libraries amplified, followed by size selection (fragment ranges between 250 and 1000 bp) using AMPure XP beads (Beckman Coulter), and sequenced to a depth of 30 million base pairs (75 bp single end) on an Illumina NextSeq 500. Reads were mapped to hg19 using BWA, indexed with samtools, converted to TDF for visualization of browser tracks in IGV using igvtools. Peaks were identified using MACS2 with statistical thresholds of 1E-5 for CEM87 treated samples and 1E-7 for untreated samples, owing to the reduced signal to noise ratio under CEM87 conditions. Read coverage under all BRD4 peaks were calculated using bedtools (multicov mode, see details here: https://bedtools.readthedocs.io/en/latest/content/tools/multicov.html) and visualized in R (scripts available here: https://github.com/CBIIT/ChIP_seq).

Data and code availability:

Readers can view our code through the public link (https://github.com/GryderArt/ChIPseqPipe). There are no access restrictions. Genome-wide data generated herein is publicly available through the Gene Expression Omnibus (GEO) with the accession number GSE129407. All data presented in this manuscript are available from the corresponding authors upon reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank the members of the Hathaway, Jin, and Khan laboratories for many helpful discussions. We acknowledge E. Chory, S. Braun, G. Crabtree, R. Vale, F. Zhang, Z. Xie, and C. Gersbach for sharing plasmids used or adapted in this study. This work was supported in part by Grant R01GM118653 from the U.S. National Institutes of Health (to N.A.H.); and by Grants R01GM122749 and R01HD088626 from the U.S. National Institutes of Health (to J.J.). This work was further supported by a Tier 3 and a Student Grant from the UNC Eshelman Institute for Innovation (to N.A.H and A.M.C, respectively). T-32 GM007092 from the U.S. National Institutes of Health (to A.M.C) also supported early portions of this work. Additional support provided from an American Cancer Society postdoctoral fellowship PF-14-021-01-CDD (to K.V.B.). Flow cytometry data was obtained at the UNC Flow Cytometry Core Facility funded by P30 CA016086 Cancer Center Core Support Grant to the UNC Lineberger Comprehensive Cancer Center.

References:

1. Dawson MA The cancer epigenome: Concepts, challenges, and therapeutic opportunities. Science. 355, 1147–1152 (2017). [PubMed: 28302822]
2. MacDonald IA & Hathaway NA Epigenetic roots of immunologic disease and new methods for examining chromatin regulatory pathways. Immunol. Cell Biol 93, 261–70 (2015). [PubMed: 25533290]
3. Zeng L & Zhou MM Bromodomain: an acetyl-lysine binding domain. FEBS Lett. 513, 124–8 (2002). [PubMed: 11911891]
4. de Ruijter AJM, van Gennip AH, Caron HN, Kemp S & van Kuilenburg ABP Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem. J 370, 737–49 (2003). [PubMed: 12429021]
5. Gao Y et al. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. Nat. Methods 13, 1043–1049 (2016). [PubMed: 27776111]
6. Chen T et al. Chemically Controlled Epigenome Editing through an Inducible dCas9 System. J. Am. Chem. Soc 139, 11337–11340 (2017). [PubMed: 28787145]
7. Braun SMG et al. Rapid and reversible epigenome editing by endogenous chromatin regulators. Nat. Commun 8, (2017).
8. Gao D & Liang FS Chemical Inducible dCas9-Guided Editing of H3K27 Acetylation in Mammalian Cells. in 429–445 (2018). doi:10.1007/978-1-4939-7774-1_24
9. Ma D, Peng S & Xie Z Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells. Nat. Commun 7, 13056 (2016). [PubMed: 27694915]
10. Shrimp JH et al. Chemical Control of a CRISPR-Cas9 Acetyltransferase. ACS Chem. Biol 13, 455–460 (2018). [PubMed: 29309117]
11. Hilton IB et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol 33, 510–517 (2015). [PubMed: 25849900]
12. Liszczak GP et al. Genomic targeting of epigenetic probes using a chemically tailored Cas9 system. PNAS 114, 681–686 (2017). [PubMed: 28069948]
13. Erwin GS et al. Synthetic transcription elongation factors license transcription across repressive chromatin. Science. 6370, 1617–1622 (2017).
14. Butler KV, Chiarella AM, Jin J & Hathaway NA Targeted Gene Repression Using Novel Bifunctional Molecules to Harness Endogenous Histone Deacetylation Activity. ACS Synth. Biol 7, 38–45 (2018). [PubMed: 29073761]
15. Chung C et al. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. J. Med. Chem 54, 3827–3838 (2011). [PubMed: 21568322]
16. Demont EH et al. 1,3-Dimethyl Benzimidazolones Are Potent, Selective Inhibitors of the BRPF1 Bromodomain. ACS Med. Chem. Lett 5, 1190–1195 (2014). [PubMed: 25408830]
17. Hay DA et al. Discovery and Optimization of Small-Molecule Ligands for the CBP/ p300 Bromodomain. J. Am. Chem. Soc 26, 9308–9319 (2014).
18. Chavez A et al. Highly efficient Cas9-mediated transcriptional programming. Nat. Methods 12, 326–328 (2015). [PubMed: 25730490]

19. Morita S et al. Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat. Biotechnol 34, 1060–1065 (2016). [PubMed: 27571369]
20. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS & Vale RD A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell 159, 635–646 (2014). [PubMed: 25307933]
21. Konermann S et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature 517, 583–588 (2015). [PubMed: 25494202]
22. Lim F et al. Altering the RNA binding specificity of a translational repressor. J Biol Chem. 12, 9006–9010 (1994).
23. Chavez A et al. Comparison of Cas9 activators in multiple species. Nat. Methods 13, 563–567 (2016). [PubMed: 27214048]
24. Begley CG & Ellis LM Raise standards for preclinical cancer research. Nature 483, 531–533 (2012). [PubMed: 22460880]
25. Hay M, Thomas DW, Craighead JL, Economides C & Rosenthal J Clinical development success rates for investigational drugs. Nat. Biotechnol 32, 40–51 (2014). [PubMed: 24406927]
26. Gryder BE et al. PAX3–FOXO1 Establishes Myogenic Super Enhancers and Confers BET Bromodomain Vulnerability. Cancer Discov. 7, 884–899 (2017). [PubMed: 28446439]
27. Lovén J et al. Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers. Cell 153, 320–334 (2013). [PubMed: 23582323]
28. Lin X et al. *HEXIM1* as a Robust Pharmacodynamic Marker for Monitoring Target Engagement of BET Family Bromodomain Inhibitors in Tumors and Surrogate Tissues. Mol. Cancer Ther 16, 388–396 (2017). [PubMed: 27903752]
29. Hathaway N. a. et al. Dynamics and memory of heterochromatin in living cells. Cell 149, 1447–1460 (2012). [PubMed: 22704655]
30. Thakore PI et al. RNA-guided transcriptional silencing in vivo with S. aureus CRISPR-Cas9 repressors. Nat. Commun 9, 1674 (2018). [PubMed: 29700298]
31. Liu S et al. Editing DNA methylation in the mammalian genome. Cell 167, 233–247 (2016). [PubMed: 27662091]
32. Gao Y et al. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. Nat. Methods 13, 1043–1049 (2016). [PubMed: 27776111]
33. Hilton IB et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol 33, 510–517 (2015). [PubMed: 25849900]
34. Chiarella AM et al. Cavitation Enhancement Increases the Efficiency and Consistency of Chromatin Fragmentation from Fixed Cells for Downstream Quantitative Applications. Biochemistry 57, 2756–2761 (2018). [PubMed: 29658277]
35. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]

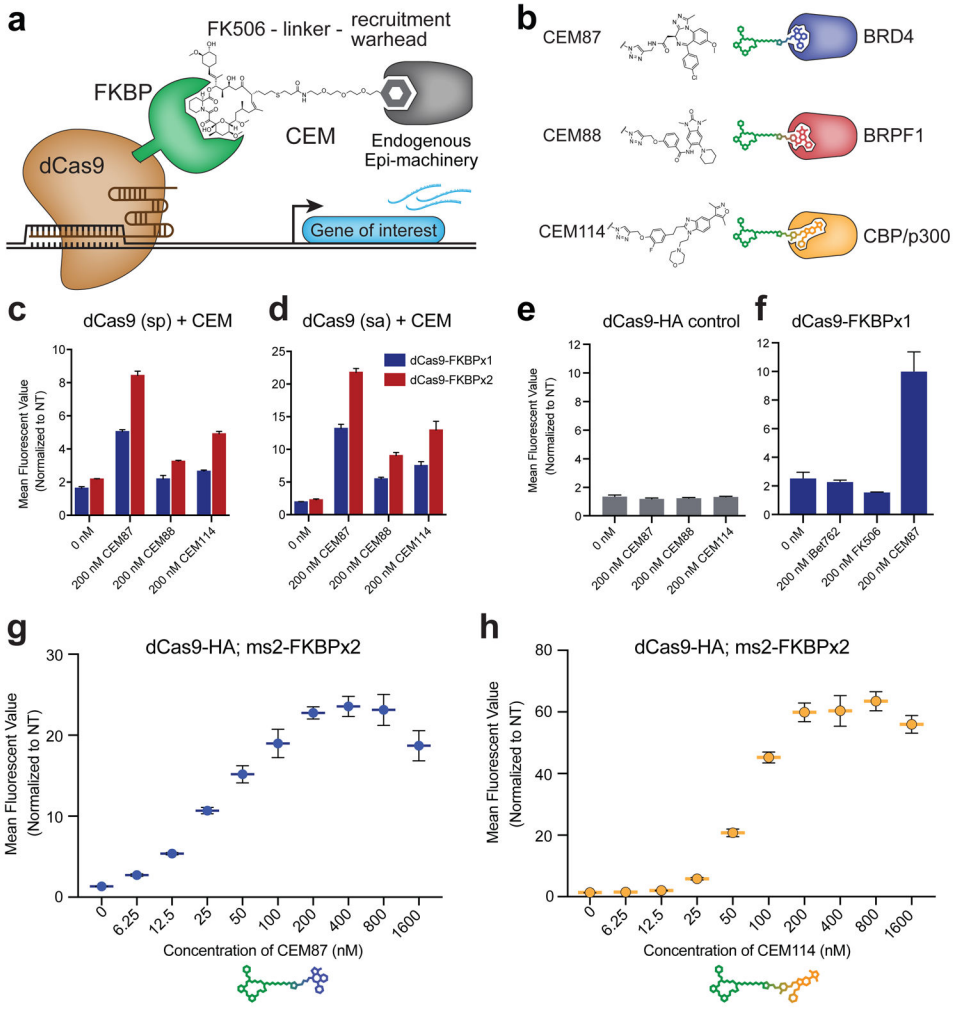

Fig. 1. Using Chemical Epigenetic Modifiers (CEMs) to increase gene expression.

a, A dCas9-FKBP fusion protein is used to target the CEMs to our gene of interest for activation. **b,** CEM87, CEM88, and CEM114 are predicted to bind and recruit BRD4, BRPF1, and CBP/p300, respectively. **c,** dCas9 (*S. pyogenes*) fused to one (blue bars) or two (red bars) copies of FKBP was recruited to a GFP reporter and flow cytometry was performed. CEM87, CEM88, and CEM114 significantly increased expression (as quantified by mean fluorescent value) compared to untreated cells. **d,** dCas9 (*S. aureus*) fused to one (blue bars) or two (red bars) copies of FKBP was recruited to a GFP reporter and flow cytometry was performed. CEM87, CEM88, and CEM114 significantly increased expression compared to untreated cells. **e,** Using a dCas9 fused to an HA tag, and no FKBP, CEM87, CEM88 and CEM114 did not significantly change expression compared to untreated cells. **f,** Treatment with the individual recruitment components of CEM87 (iBet762 and FK506) did not significantly change gene expression, activation was only observed with the bifunctional CEM87. **g,** A dose curve of CEM87 was performed with CEM concentration ranging from 6.25 nM to 1600 nM. Mean fluorescent value was normalized to cells expressing non-targeting gRNA. Compared to untreated cells, expression increased after 6.25 nM CEM87 treatment for 48 hours. **h,** A dose curve of CEM87 was performed

with CEM concentration ranging from 6.25 nM to 1600 nM. Mean fluorescent value was normalized to cells expressing non-targeting gRNA. Compared to untreated cells, expression increased after 6.25 nM CEM114 treatment for 48 hours. **c-h,** Significance was determined with three different samples, using a two-tailed student's t-test, fold change and significance reported in Supplementary Table 1. Error bars represent the standard deviation.

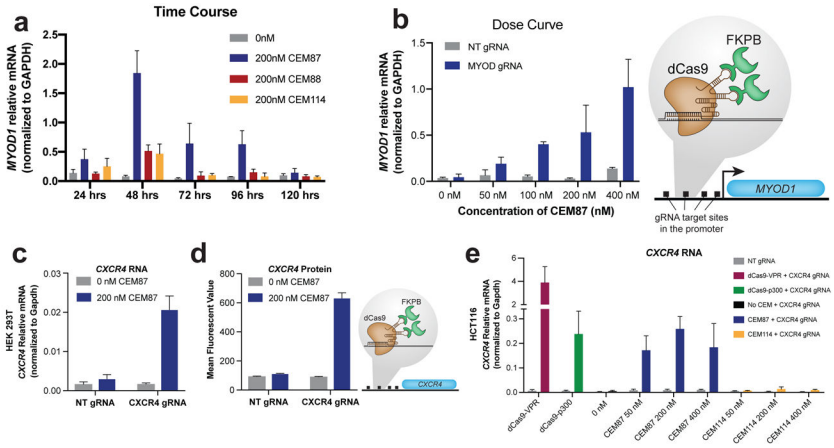

Fig 2. Evaluating dCas9-CEMa over time, dose range, and Benchmarking dCas9-CEMa system with current dCas9 activating technologies.

a, A time course was performed after transfection of *MYOD1*-targeting gRNAs in cells that were virally induced with dCas9-HA and ms2-FKBPx2. Cells were treated with 200nM of CEM87, CEM88, or CEM114, mRNA was isolated and qRT-PCR was done every 24 hours (following initial media change post transfection) for 5 days. Relative mRNA levels were compared to untreated cells. CEM87 activated endogenous *MYOD1* the most effectively at 48 hours. **b,** A dose curve of CEM87-mediated activation was performed with concentrations ranging from 50 nM to 400 nM. CEM87 treatment began and mRNA was isolated after 48 hours of treatment. **c,** HEK 293T cells were transfected with *CXCR4*-targeting gRNA or non-targeting gRNA and treated with 200 nM of CEM87 for 48 hours. RNA was isolated and qRT-PCR was performed. Significant levels of *CXCR4* activation were only observed in cells expressing the targeting gRNA. **d,** HEK 293T cells were transfected with *CXCR4*-targeting gRNA or non-targeting gRNA and treated with 200 nM of CEM87 for 48 hours. Flow cytometry was performed with a fluorescently conjugated CXCR4 antibody. **e,** HCT116 cells were infected with dCas9-VPR, dCas9-p300 and dCas9-HA, ms2-FKBPx2. Cells were transfected with *CXCR4*-targeting gRNA or non-targeting gRNA. Infected dCas9-CEMa cells were treated with 0 nM, 50 nM, 200 nM and 400 nM of CEM87 and CEM114 for 48 hours. RNA was isolated and qRT-PCR was performed. Significant levels of *CXCR4* activation were observed in dCas9-VPR, dCas9-p300 cells and all doses of CEM87 treated cells. Significant levels but lower fold changes of *CXCR4* activation were observed in cells treated with 200nM and 400nM of CEM114. No significant *CXCR4* activation were observed in cells transfected with *CXCR4*-targeting gRNA and treated with 0 nM of CEM. **a-e,** Significance was determined with three different samples, using a two-tailed student's t-test, fold change and significance reported in Supplementary Table 1. Error bars represent the standard deviation.

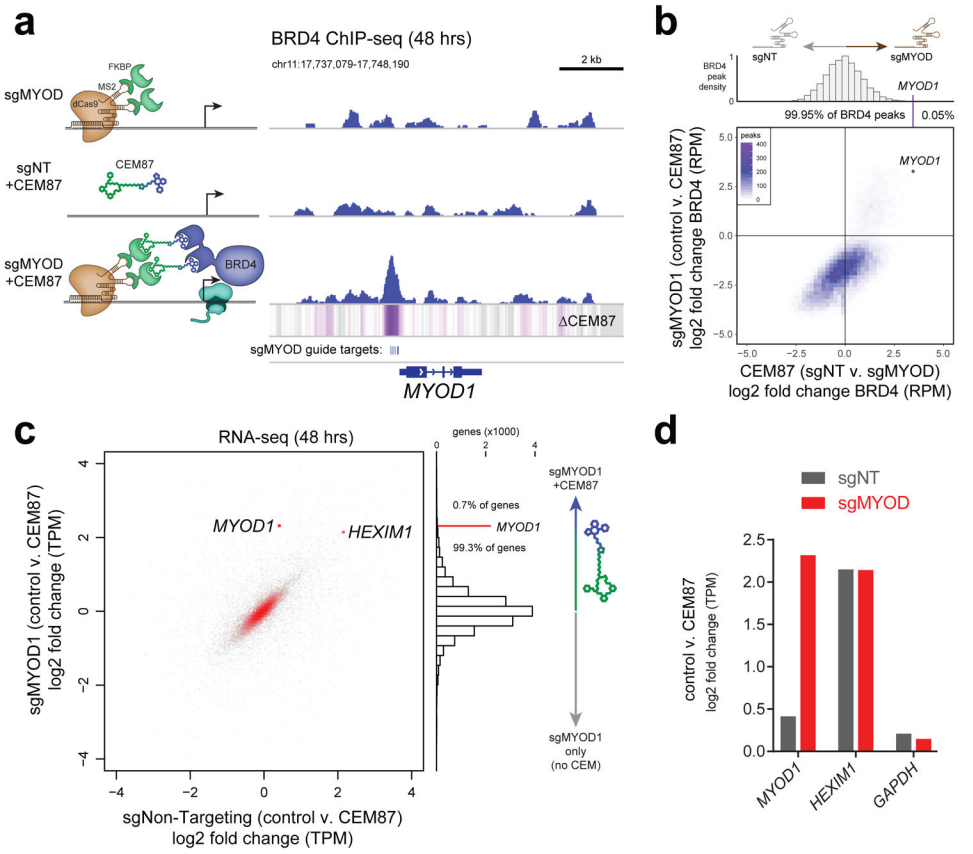

Fig 3. Whole-genome analysis of the dCas9-CEMa system.

a, Genome-wide binding of BRD4 assayed by ChIP-seq in HEK cells containing dCas9-HA and ms2-FKBP, treated with either 4x sgRNAs targeting the *MYOD1* promoter (sgMYOD), the BRD4 recruitment molecule CEM87 with non-targeting sgRNA (sgNT), or both sgMYOD and CEM87, for 48 hours. Delta CEM87 genomic heatmap was generated by subtracting the reads per million mapped reads (RPM) of BRD4 ChIP-seq in sgMYOD with CEM87 minus sgMYOD without CEM87. **b,** Quantification of change in ChIP-seq signal (log2 fold change of RPM) for each called BRD4 peaks, compared for changes induced by CEM87 (y-axis) and changes induced by sgMYOD (x-axis). Histogram of BRD4 peak density (normalized to 1) for sgMYOD specific changes in the presence of CEM87 (compared to sgNT) is shown above. BRD4 peak at *MYOD1* promoter is highlighted. **c,** RNA-seq quantified by Transcripts Per Million (TPM) in HEK cells treated as in (a), with CEM87 induced changes with sgMYOD (y-axis) compared with CEM87 induced transcriptional changes with sgNT (x-axis). Histogram of CEM87 induced gene expression changes in cell with sgMYOD is shown on the right. *MYOD1* and *HEXIM1* are highlighted, and bar charts of their log2 fold changes in TPM are shown in **d.**