Aus dem

Dr. Margarete Fischer-Bosch-Institut für Klinische Pharmakologie

Development of novel antineoplastic treatment approaches by targeted epigenetic silencing of the vascular endothelial growth factor A and its receptors

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Bashtrykov, Pavel

2021

Dekan:	Professor Dr. B. Pichler
1. Berichterstatter:	Professor Dr. U. Zanger
2. Berichterstatter:	Professor Dr. U. Lauer
Tag der Disputation:	21.06.2021

Dedicated to my beloved wife

TABLE OF CONTENTS

List of abbreviations

Name	Description
α3A3L	scFv-GCN4-DNMT3ACD-DNMT3LCD
AP-2	activation protein 2
ATAC-seq	assay for transposase-accessible chromatin using sequencing
BC	breast cancer
bp	base pair
Cas	CRISRP-associated
CGI	CpG island
CRC	colorectal cancer
CRISRP	clustered regularly interspaced short palindromic repeats
DBD	DNA binding domain
dCas	catalytically inactive Cas
dCas9S	dCas9-10xSunTag
DMEM	Dulbecco's Modified Eagle's Medium
DNase-seq	DNase I hypersensitive sites sequencing
DNMT	DNA methyltransferases
DNMT3ACD	the DNMT3A catalytic domain
DNMT3LCD	the DNMT3L C-terminus
dsRNA	double-stranded RNA
EC	endothelial cell
ECM	extracellular matrix
EMT	epithelial-mesenchymal transition
FACS	Fluorescence-activated cell sorting
FCS	fetal calf serum
FLT1	Fms-like tyrosine kinase 1
GCN4	general control protein 4
gDNA	genomic DNA
H3K27me3	histone H3 lysine 27 trimethylation
H3K4me3	histone H3 lysine 4 trimethylation

H3K9me2/3	histone H3 lysine 9 di- and trimethylation
H4K20m3	histone H4 lysine 20 trimethylation
HEK293	Human embryonic kidney 293
KDR	kinase domain region
LB	lysogeny broth
NGS	next-generation sequencing
PAM	proto-spacer adjacent motif
PBS	Dulbecco's Phosphate Buffered Saline
PTGS	post-transcriptional gene silencing
PTM	post-translational modification
RISC	RNA-induced silencing complex
RNAi	RNA interference
scFv	single-chain variable fragment
sgRNA	single guide RNA
siRNA	short interfering RNA
SP1	specificity protein 1
SP3	specificity protein 3
STAT-3	signal transducer and activator of transcription 3
TALE	transcription activator-like effector
TE	Tris-EDTA
TET	ten-eleven translocation
TPE	Tris-phosphate-EDTA
VEGF	vascular endothelial growth factor
VEGFA	vascular endothelial growth factor A
VEGFB	vascular endothelial growth factor B
VEGFC	vascular endothelial growth factor C
VEGFD	vascular endothelial growth factor D
VEGFR1	vascular endothelial growth factor receptor 1
VEGFR2	vascular endothelial growth factor receptor 2
ZFP	zinc finger protein

1 Introduction

1.1 Cancer

The term cancer covers multiple diseases, which are characterized by the appearance of abnormal cells harbouring "*essential alterations in cell physiology that collectively dictate malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis*", "*reprogramming of energy metabolism and evading immune destruction*" [1], [2]. Despite decades of intensive research aimed at the investigation of tumorigenesis and development of therapies, cancer occupies the second position of the leading causes of mortality with approximately 9.6 million deaths worldwide in 2018 [3].

Tumorigenesis is a multi-step process that requires accumulation of numerous mutations in cells [4]. There are two general mechanisms of how genetic abnormalities induce and promote tumorigenesis: (a) an increase in activity of growth-promoting genes, so-called oncogenes, and (b) a reduction of activity of tumor suppressor genes. Mutations can either change the expression of oncogenes and tumor suppressor genes or alter properties of the proteins they encode. Various mutations identified in cancer cells can be classified into two groups based on their impact on tumorigenesis: "driver" mutations, which bring growth advantages to cells and "passenger" mutations, which do not possess such properties [5]. Tumor sample whole-genome sequencing studies have identified at least 299 cancer-associated "driver genes", which are shared across tumor types tumorigenesis [6]. An average tumor contains two to eight mutations in "driver genes" and dozens of "passenger" mutations [7].

A modern anticancer treatment includes "classical" therapeutic approaches such as surgery, radiotherapy and chemotherapy, as well as many drugs of a new generation that have been developed to target specific molecules in cancer signaling. This so called targeted therapy includes tyrosine kinase inhibitors, immune checkpoint inhibitors and many others [8] (Figure 1). Concepts of

antiangiogenic therapy as the main focus of the current project will be discussed in more detail in the next paragraph.

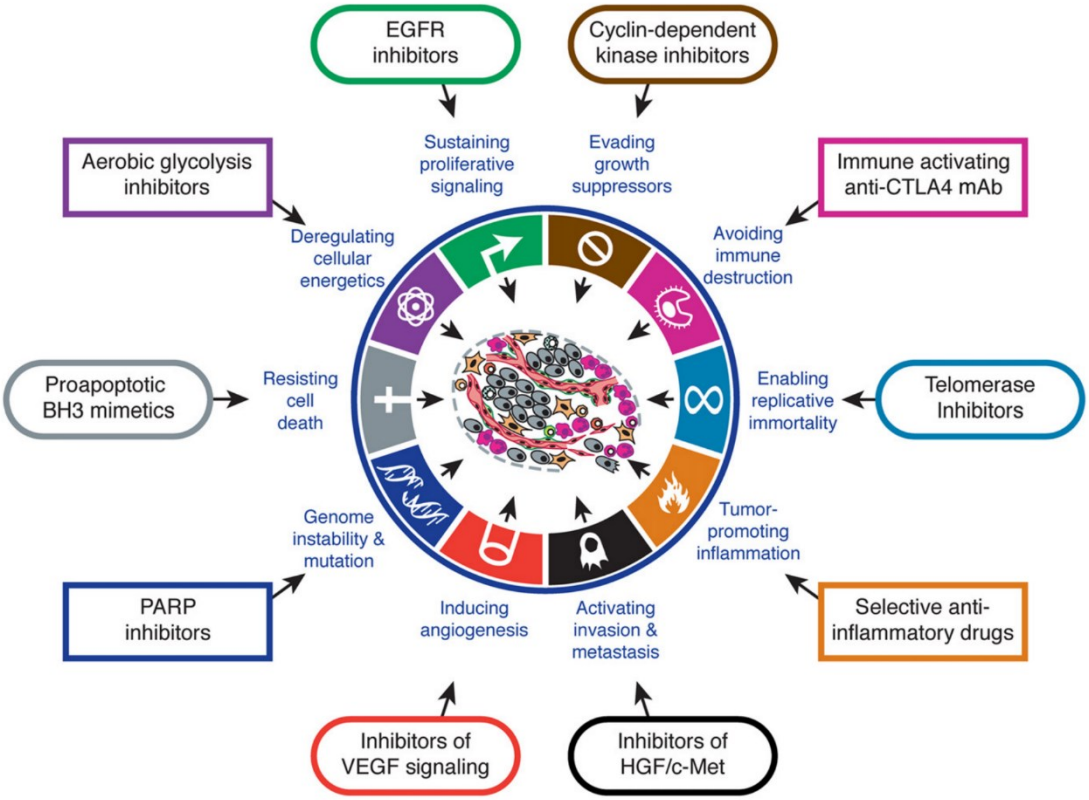

Figure 1. Examples of targeted anticancer therapy.
Exemplary groups of therapies are formed based on their targeting of specific cancer hallmarks. Figure taken from [2].

1.2 Angiogenesis, VEGFA and its receptors

Growth of tissues and organs during embryogenesis and afterward is supported by a network of blood vessels, which brings nutrients, oxygen, hormones and removes metabolites. The formation of blood vessels from pre-existing vessels, termed as angiogenesis, takes place mainly during body growth but in rare cases also in adults, such as healing of damaged tissue or growth of skeletal muscles [9]. Apart from normal physiological angiogenesis various pathophysiological processes include angiogenesis, for example age-related macular degeneration, inflammatory and autoimmune disorders and tumor growth [9], [10].

Angiogenesis is initiated in tissues, which under growth conditions experience hypoxia. To stimulate extension of neighboring vessels in their direction, they

secret pro-angiogenic factors. The most important regulator of angiogenesis is vascular endothelial growth factor A (VEGFA) discovered by Genentech scientist Napoleon Ferrara in 1989 [11]. It is secreted by various parenchymal cells and works in a paracrine manner: it diffuses into the extracellular matrix (ECM) and binds receptors on the cellular membrane of local endothelial cells (ECs). VEGFA stimulates migration and proliferation of endothelial cells, increases vascular permeability and mediates physiological and pathological angiogenesis [12]–[14]. VEGFA belongs to the family of VEGF growth factors, which also includes VEGFB [15], VEGFC [16], VEGFD [17] and placental growth factor [18]. Four main isoforms of VEGFA of 121, 165, 189, and 206 amino acids in length [19]–[21] and four additional variants, of 145, 162, 165b and 183 amino acids, which appear less frequently [20], [22]–[24], have been characterized. They all are products of one gene generated by alternative splicing.

Expression of VEGFA is regulated at the transcriptional level by a plethora of stimuli, including hormones [25], [26], cytokines [27], [28] and growth factors [29], [30]. The VEGFA promoter integrates all these different pathways via specific transcription factor binding sites, including hypoxia response elements [31], estrogen response elements [25] and binding sites for multiple transcription factors such as specificity protein 1/specificity protein 3 (SP1/SP3)[32], [33], activation protein 2 (AP-2)[33], [34] and signal transducer and activator of transcription 3 (STAT-3)[35].

After secretion, VEGFA interacts with the ECM and undergoes proteolytic cleavage by several proteases, including matrix metalloproteinases [36], urokinase [37] and plasmin [38]. The latter one has pleiotropic regulatory effects, such as activation and release of VEGFA from ECM storage or its degradation [39]. VEGFA acts as a homodimer via stimulation of the specific VEGF receptors (VEGFRs) located on the surface of target cells. VEGFRs are found in endothelial and various non-endothelial cells [40]. VEGFRs are receptor tyrosine kinases, binding of VEGFA leads to their dimerization, auto- and trans-phosphorylation, which initiates signal transduction [41]. VEGFR1 (the alternative name is Fms-like tyrosine kinase 1, FLT1) and VEGFR2 (also named as kinase domain region,

KDR) are the main mediators of the VEGFA signaling in angiogenesis [42], [43]. VEGFR2 possesses a stronger kinase activity and is the key mediator of VEGFA-dependent angiogenesis [44], whereas VEGFR1 has higher affinity to the VEGFA and in addition to its own signaling and is thought to modulate the activity of VEGFR2 [45], [46].

1.3 Role of the VEGFA/VEGFRs axis in cancer

Already two decades before, when VEGFA was discovered, experiments showed that solid tumors stop growth when reaching 2-3 mm in diameter but can continue growth after neovascularization. Tumor implants in mice stimulate ECs of neighbouring capillaries and venules, which is an extremely fast process that can be detected as soon as six hours after tumor cells transplantation and only three days are needed for new capillaries to penetrate the implant [47]. These early discoveries initiated studies on tumor angiogenesis and development of anti-angiogenic drugs for cancer treatment. Neovascularisation of tumors occurs via different mechanisms, including sprouting angiogenesis, intussusceptive angiogenesis, vasculogenesis, recruitment of endothelial progenitor cells, vascular mimicry and trans-differentiation of cancer stem cells, which are shown schematically on Figure 2.

Different factors are involved in tumor vascularisation but the leading role is taken by VEGFA. Significant VEGFA overexpression was detected in various cancers and associated with poor clinical prognosis [48], [49]. ErbB2, a transmembrane receptor tyrosine kinase, which is often overexpressed in human breast cancer [50] and stimulates expression of VEGFA via the hypoxia responsive element and the SP1 binding sites in the core promoter [51]. SP1 activates VEGFA expression in trastuzumab-resistant ovarian cancer cells [52].

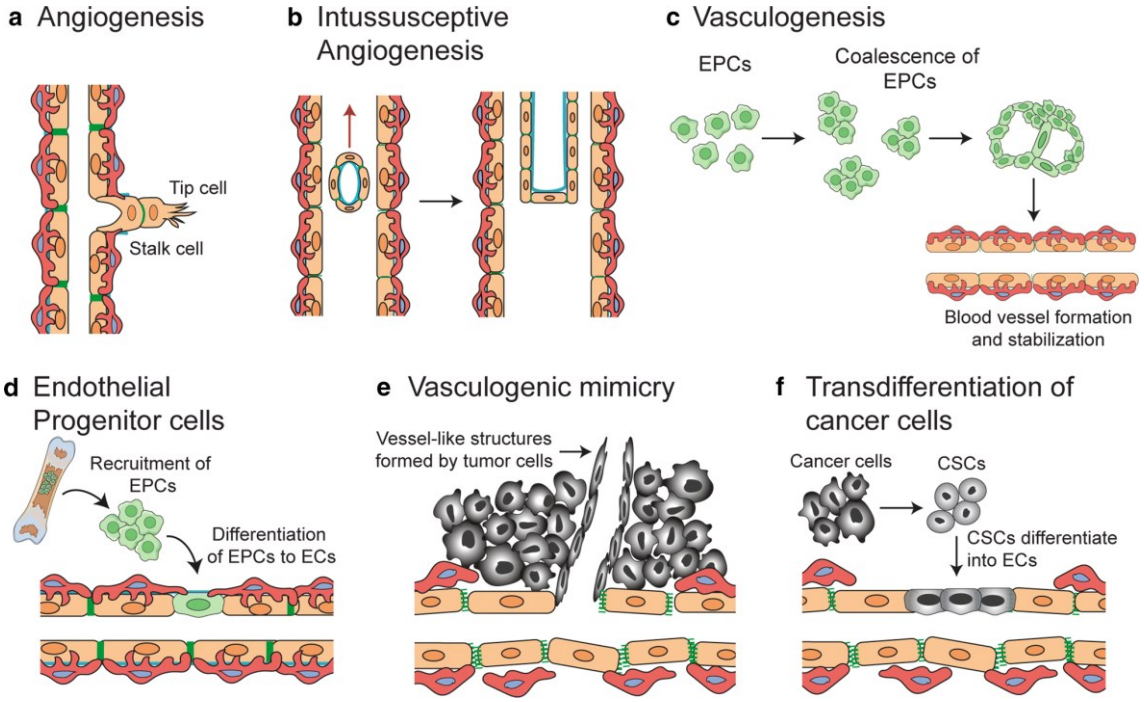

Figure 2. Different mechanisms of neovascularisation in tumors.
Vessel formation in tumors undergoes using one of the following pathways. **a** Outgrowth of new branches from existing vessels. **b** Bifurcation of an existing vessel in two. **c** Formation of vessels *de novo*: from endothelial progenitor cells (EPCs). **d** Extension of existing vessels upon recruitment of EPCs circulating in blood. **e** Vascular mimicry – formation of pseudovessels by tumor cells. **f** Differentiation of cancer stem cells (CSCs) to endothelial cells (ECs). Figure is taken from [53].

The initial paradigm assuming that the VEGFRs are expressed predominantly on endothelial cells of blood and lymphatic vessels and the VEGFA/VEGFRs axis regulates angiogenesis and vascular permeability, was expanded later. It was shown that ECs express VEGFA that acts in an autocrine manner and that is required for survival of ECs and the homeostasis of blood vessels [54]. Expression of VEFGA and its receptors was detected in various tumors including non-small cell lung carcinomas [55], urinary bladder cancer [56], breast cancer (BC) [57], colorectal cancer (CRC) [58] and many others. An autocrine mechanism of VEGFA on growth of meduloblastoma cells [59] and BC cell lines [57] was proposed via VEGFR1 and VEGFR2 signaling. It was shown that treatment of human pancreatic cell lines with VEGFA induces epithelial-mesenchymal transition (EMT) via VEGFR1 [56]. VEGFA is a poor prognosis factor in CRC and BC, and it stimulates migration and invasion of CRC cell lines via VEGFR1 [58], as well as invasion and proliferation of BC cell lines [57], [60].

VEGFR2 up-regulated in gastric cancer promotes tumorigenesis by stimulating cell proliferation and invasion [61].

1.4 Epigenetics and regulation of gene expression

The last decades of intense research have discovered a new layer of genome regulation, which is called epigenome. The term epigenome covers all aspects of regulation of gene expression leading to stable inheritable cellular phenotypes without underlying changes of the genome sequence. In other words, the field of epigenetics studies mechanisms how cells of multicellular organisms, such as humans, generate and maintain divers phenotypes having identical genetic material. This becomes possible because genomic DNA is stored within the cell nucleus as chromatin (Figure 3) [62]. Four types of histone proteins, H2A, H2B, H3 and H4, two copies of each, fold in an octamer, which binds a 145-147 base pair stretch of DNA to form a nucleosome, the minimal structural unit of chromatin [63]. The N-terminal tails of histone proteins pointing out of the packed histone core are freely accessible to chromatin-interacting proteins and undergo various post-translational modifications (PTMs) such as acetylation, methylation, phosphorylation, ubiquitination and many others [64].

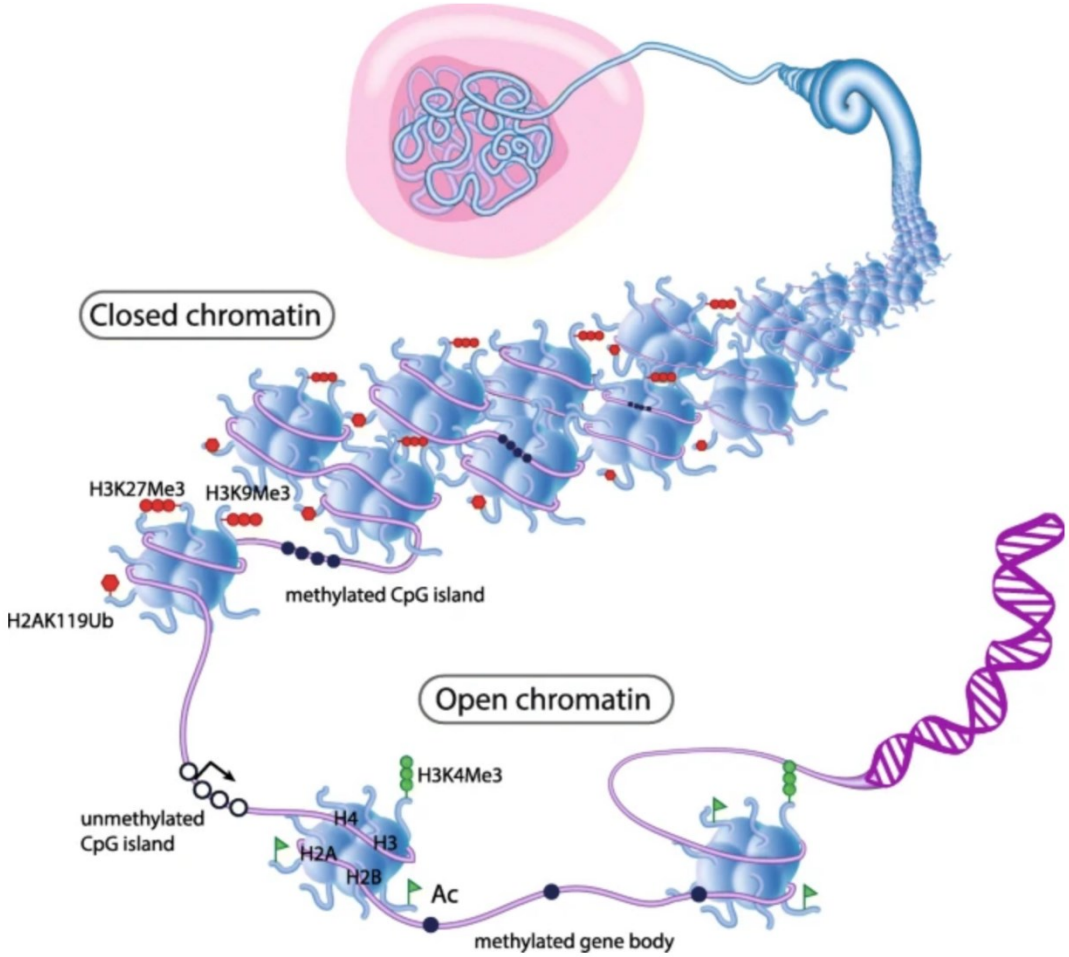

Figure 3. Chromatin structure modulated by epigenetic mechanisms regulates gene expression.
Two types of chromatin, closed, compact transcriptionally inactive heterochromatin and open, accessible for transcription factors euchromatin, are formed by chromatin marks, modifications of DNA and histone tails, regulating gene expression. Figure taken from [65] , which is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). Caption was modified.

The expression of genes is dynamically regulated by epigenetic signals also termed as chromatin marks or epigenome marks, which modulate accessibility of chromatin making it more open, transcriptionally active, or closed, where transcription is repressed or completely silenced [66]. The most studied chromatin marks are histone tail PTMs and methylation of DNA. Examples of active chromatin marks are histone H3 lysine 4 trimethylation (H3K4me3) and acetylation of lysines on histones H3 and H4 [67]. Repressive chromatin marks include methylation of histone H3 at lysine 9 (H3K9me2/3) and 27 (H3K27me3), methylation of lysine 20 at histone H4 (H4K20m3) and DNA methylation [68]. The chromatin marks are enzymatically deposited (by “writers”) and removed (by

"erasers") by chromatin-modifying enzymes and form a so called "histone code" [69], which is decoded by "readers" (Figure 4). All these complex mechanisms allow for a dynamic regulation of genome activity and access in response to internal and environmental signals [70].

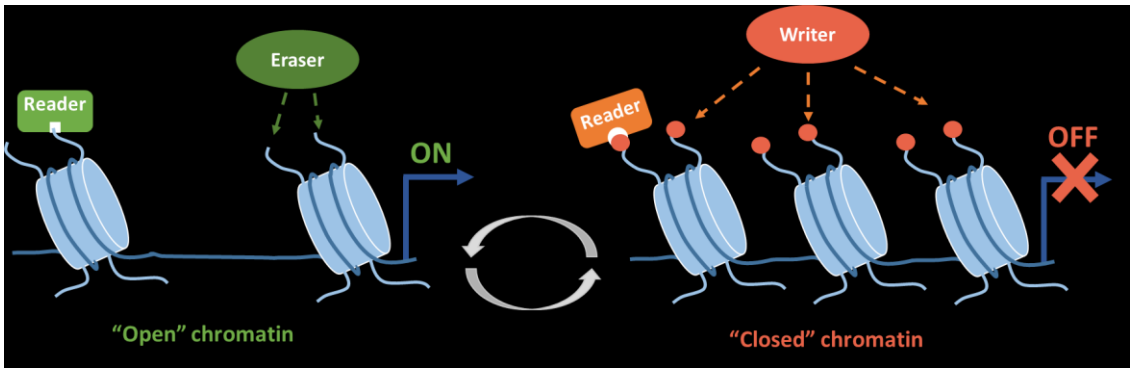

Figure 4. Writers, readers and erasers of chromatin marks.
Chromatin landscape is formed by chromatin-modifying enzymes: Write – sets a chromatin mark, Eraser – removes a chromatin mark. Information encoded in chromatin marks is interpreted by chromatin-interacting proteins - Readers.

1.4.1 DNA methylation

Methylation of human DNA takes place at cytosines, predominantly at CG dinucleotides also named as CpG sites [71]. About 1 % of cytosines are methylated, which corresponds to 60-80 % methylated CpG sites. Whole-genome DNA methylation profiling of different cell lines revealed that there is a bimodal distribution of DNA methylation in the human genome, meaning that regions tend to be either strongly methylated or unmethylated [72], [73]. Three DNA methyltransferases (DNMTs) were identified in human cells, DNMT1 [74], DNMT3A and DNMT3B [75]. DNA methylation is set in a cell-type specific manner during embryonic development by *de novo* enzymes DNMT3A and DNMT3B [76], [77]. Besides them, there is the DNMT3L protein, which has a similar structure, but does not possess catalytic activity [78], [79]. It forms complexes with DNMT3A and DNMT3B enzymes and enhances their catalytic activity [80], [81]. Crystallographic studies discovered that the DNMT3A catalytic domain (DNMT3ACD) and the C-terminal part of DNMT3L (DNMT3LCD) form a linear heterotetramer, where one DNA molecule is bound by the two DNMT3ACD molecules located centrally and interacting with each other and with two molecules on DNMT3LCD positioned at the sides of the tetramer thereby forming

a minimal enzymatic unit [82]. Moreover, DNMT3A alone forms homodimers, which can oligomerize in two directions, making protein filaments able to bind multiple DNA at the same time and along one DNA molecule, which facilitates methylation of target regions [82], [83]. DNA methylation patterns are copied during DNA replication by the maintenance enzyme DNMT1 [84]. It binds preferentially to hemimethylated CpG sites generated during DNA replication, and methylates in the unmethylated DNA strand, which leads to formation of fully methylated CpG sites [85], [86].

All methylated CpG sites of a cell constitute the cell-type specific methylome, also referred to as DNA methylation pattern, which plays a key role in genome usage [71], [87]. Multiple experimental evidences confirm that DNA methylation is a repressive epigenetic signal and methylation of CpG sites in gene promoters results in a stable gene silencing [87], [88]. Promoters with a very high or intermediate CpG density, termed CpG islands (CGIs) and weak CGIs, respectively, demonstrate the strongest response to methylation [72], [88]. Examples of DNA methylation-dependent transcription repression are gene silencing [89], [90], X-chromosome inactivation [91], suppression of transposable elements [92], [93] and genomic imprinting [94], [95]. Interestingly, the level of methylation in gene bodies positively correlates with their expression level, but the functionality of this phenomenon is not jet completely understood [96], [97]. There are two hypotheses on this, one proposes that gene-body methylation prevents an intragenic transcription initiation [98], while another speculates that it is a consequence of an open chromatin at the actively transcribed genes, which makes DNA more accessible to DNMTs [99].

1.5 Therapeutic targeting of the VEGFA/VEGFRs axis in cancer

As discussed earlier, VEGFA and its receptors contribute to tumorigenesis via angiogenesis-dependent and independent mechanisms, which made them promising targets for development of anticancer therapies. Inhibition of a protein function can be realized at three levels: at the protein itself, at the mRNA and at the gene level. At the protein or post-translational level, one may develop protein-specific inhibitors [100], [101], regulate protein degradation [102] or influence

activation of target proteins by PTMs [103]. At the mRNA or post-transcriptional level, one can inhibit translation from mRNA or decrease mRNA abundance inducing its degradation [104]. At the gene level, one would need to develop approaches for targeted and regulated silencing of the transcription of gene encoding target proteins [105]. Various therapeutic approaches targeting VEGFA, VEGFR1 and VEGFR2 at all three levels have been developed and some of them have been introduced into routine cancer treatment practice. The most interesting strategies, according to the author, will be described in details.

1.5.1 Inhibition of the VEGFA/VEGFRs axis at the post-translational level

The first example of VEGFA inhibiting drug was the discovery of anti-VEGFA monoclonal antibodies in 1993 by Genentech. These antibodies bind extracellular VEGFA and prevent its interaction with corresponding receptors, reducing tumor growth *in vivo* via inhibition of tumor angiogenesis [106]. The efficiency of the therapy was shown in mice after the subcutaneously injected tumor cells: three cell lines were tested and 70 to 95 % of growth inhibition was reported. In contrast, there were no inhibitory effects on tumor growth *in vitro*, which confirms an anti-angiogenesis-based mechanism not acting via the autocrine pathway. Bevacizumab (Avastin), a humanized version of the monoclonal anti-VEGFA antibodies, was generated in 1997 and confirmed findings observed with the original mouse anti-VEGFA antibodies (Figure 5) [107]. Since then, numerous clinical trials analysing efficiency and safety of Bevacizumab have been conducted (2624 entries at https://clinicaltrials.gov database, accessed on 11 August 2020) and it was approved for clinical use in a combination with antineoplastic agents as the first- and second-line treatment of multiple types of cancer [108]–[110]. Clinical benefits are achieved via improvement of overall and/or progression-free survival of patients.

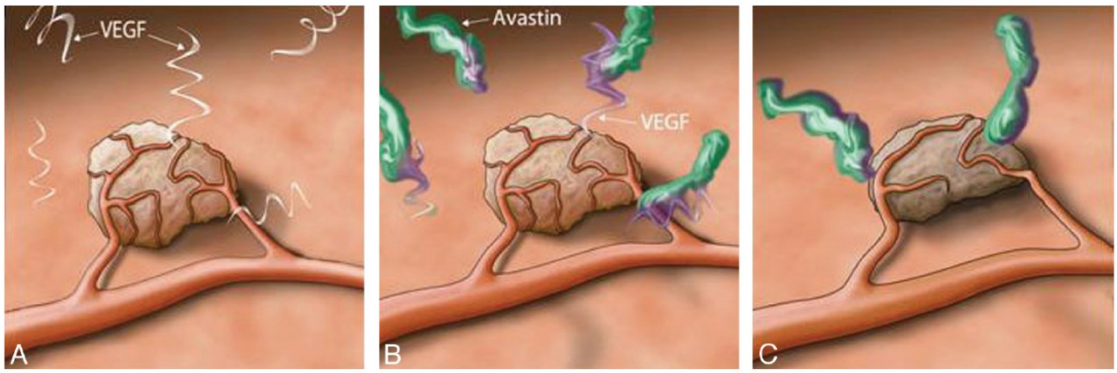

Figure 5. Mechanism of action of the antiangiogenic drug Bevacizumab.
Intravenously administrated Bevacizumab binds VEGFA, prevents its interaction with VEGFRs and inhibits tumor angiogenesis. Figure is taken from [111].

Similarly, antibodies inhibiting VEGFR1 and VEGFR2 were developed. Anti-VEGFR1 monoclonal antibodies suppressed growth of breast xenografts in mice increasing apoptosis of tumor cells [60]. Anti-VEGFR1 human monoclonal antibody, Icrucumab, inhibits stimulation by ligands and blocks downstream signaling of VEGFR1. A Phase I clinical study demonstrated potential for treatment of advanced solid malignancies, which do not respond to standard therapy and those for whom standard therapy was not available [112]. In total, 4 studies of Icrucumab have been conducted (https://clinicaltrials.gov, accessed on 28 Jul 2020). Specific anti-VEGFR2 antibodies, Ramucirumab, showed reduction of tumors in clinical trials in patients with advanced solid malignancies [113]–[115]. There are 117 studies for Ramucirumab of which 57 are active/recruiting patients (https://clinicaltrials.gov, accessed on 28 Jul 2020). It is approved as a single agent and in combination with other drugs for treatment of advanced or metastatic gastric cancer, some types of metastatic non-small cell lung cancer, metastatic colorectal cancer and hepatocellular carcinoma.

All these antibody-based therapeutics have a common approach – they aim for inhibition of extracellular target proteins but do not change their abundance and require regular intravenous injections.

1.5.2 RNA interference-based anti-VEGFA/VEGFRs therapeutics

RNA interference (RNAi) is a mechanism of gene silencing at the level of mRNA mediated by short complementary RNAs. Post-transcriptional gene silencing

(PTGS) effecting abundance of mRNAs was observed in plants [116], [117]. Experimentally PTSG can be induced by delivery to cells of plasmid vectors encoding antisense sequences of targeted mRNAs [118]. Finally, sequence-specific gene silencing upon delivery of short double-stranded RNAs, RNAi, was demonstrated for the first time in the nematode *Caenorhabditis elegans* [119]. Multiple studies in the following years showed that RNAi is a general natural mechanism used by a broad range of eukaryotes including *Homo sapiens* to control gene activity [120], [121]. Short interfering RNAs (siRNAs) are double-stranded RNAs of 21-23 nt with 2 nt overhangs at 3'-ends [122] (Figure 6). Endogenously, they are generated from long double-stranded RNAs (dsRNAs) by a nuclease Dicer [123], [124]. Exogenous or native siRNAs form a complex with proteins forming the RNA-induced silencing complex (RISC), where the sense strand is degraded and a mature complex guided by an antisense strand of the siRNA interacts with and cleaves a target mRNA by the Argonaute subunit of RISC [125]–[127].

RNAi was used in vitro to knock down VEGFA expression with siRNAs designed to target VEGFA reduced VEGFA expression in ovarian carcinoma and melanoma cells [128]. Anti-VEGF siRNAs were able to inhibit proliferation, migration and invasion of human hilar cholangiocarcinoma cell lines [129]. Systemic injection of siRNAs reduced fibrosarcoma cells growth and tumor vascularization in mice tumor model [130]. Direct intratumoral injection of anti-VEGF siRNAs inhibited VEGF secretion, tumor growth and tumor angiogenesis in a xenograft prostate cancer model [131]. Also, siRNA specific to VEGFR2 demonstrated down-regulation of VEGFR2, suppression of tumor growth and tumor angiogenesis after intravenous administration in a mouse tumor model [132]. Simultaneous targeting of VEGFA, VEGFR1/2 showed high efficiency of neovascularisation inhibition in ocular angiogenesis model [133].

At the moment there are no approved siRNA-based therapies for cancer treatment, but several drugs are in clinical trials testing their inhibitory effects on macular neovascularisation, such as AGN 211745, siRNA against VEGFR1, and

bevasiranib, anti-VEGFA siRNA (data are taken from https://clinicaltrials.gov, accessed on 12 August 2020).

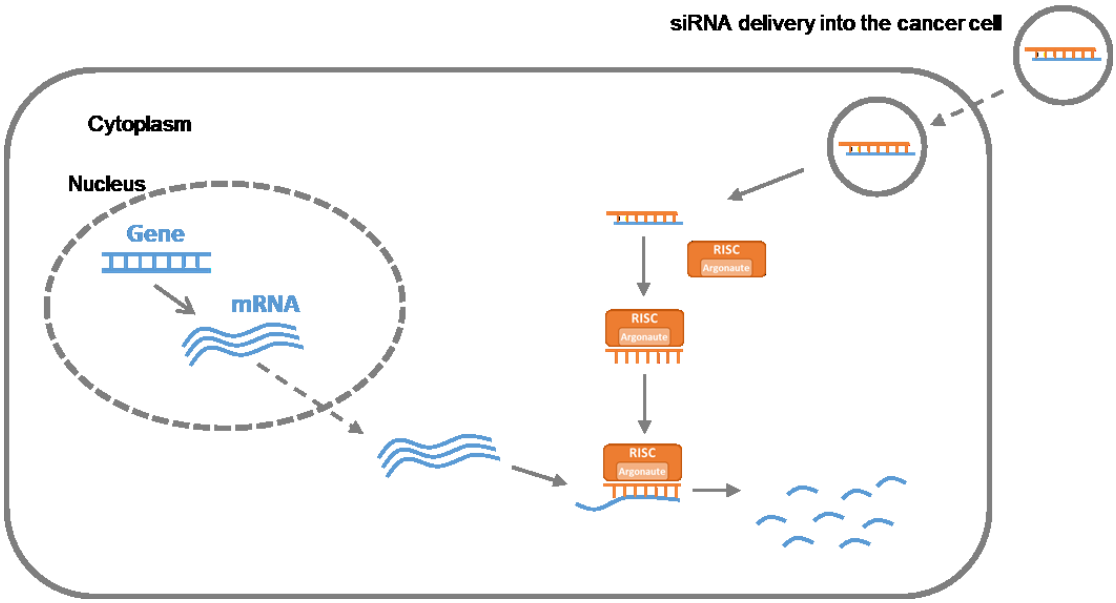

Figure 6. siRNA-mediated gene silencing.
Double-stranded siRNAs delivered into cells are processed by RISC, an antisense strand is used for degradation of a target mRNA catalyzed by Argonaute.

1.5.3 Repression of VEGFA and VEGFRs expression by epigenome editing

The next level, at which one may try to develop therapeutic intervention, is the regulation of expression of target genes. As discussed above, expression of genes is controlled via accessibility of chromatin to transcriptional factors, which is mediated by chromatin marks. The functions of many chromatin marks as well as of the enzymes, which deposit or remove these marks are well characterized, which led to the concept of epigenome/epigenetic editing. According to De Groote and colleagues epigenetic editing is defined as "*targeted rewriting of epigenetic marks to modulate expression of selected target genes*" [105]. However, the term epigenetic editing implies that introduced modifications of epigenome are principally heritable, which is not always the case, therefore epigenome editing will be used in this work.

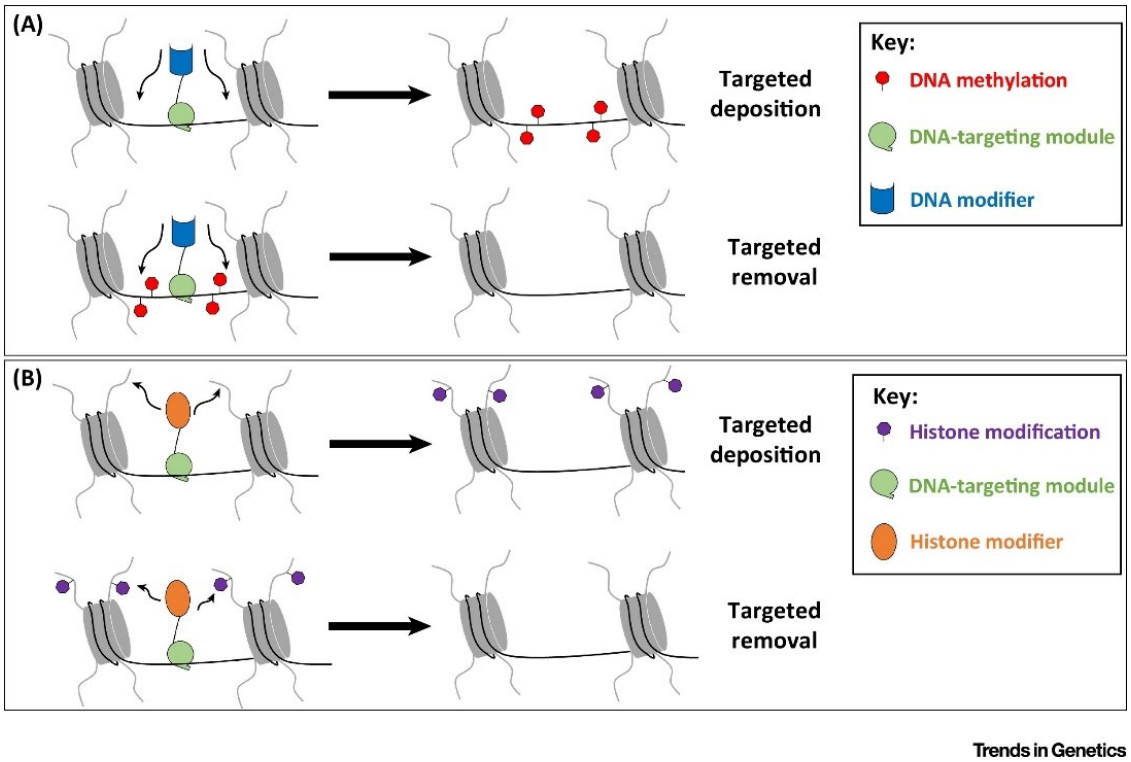

Figure 7. The concept of targeted epigenome editing.
EpiEditors are chimeric proteins, composed of DNA binding domain and a chromatin-modifying enzyme, designed to deposit or remove chromatin marks on DNA **(A)** or histone tails **(B)** at target genomic loci. Figure taken from [134].

Targeted epigenome editing can be achieved by designed chimeric proteins, also known as EpiEditors (Figure 7) that contain two functional units, a DNA binding domain (DBD) and a chromatin-modifying enzyme [134]. DBDs are constructed with DNA sequence-specificity allowing to bring the EpiEditor to the target genomic locus. The chromatin-modifying enzyme alters local histone tails or DNA by introducing or removing selected epigenome marks. The EpiEditor genes are delivered into target cells via plasmid DNA or viral vectors, where they are expressed and generate the desired chromatin state changing the transcription of the target genes.

DBDs of three types have been used until now, zinc finger proteins (ZFPs), transcription activator-like (TAL) effector (TALE) proteins and clustered regularly interspaced short palindromic repeats (CRISRP)-associated (Cas) proteins (Figure 8). ZFPs were the first utilized DBDs for targeting to specific genome regions, they consist of zinc-fingers, each of those binds 3 bps of DNA [135]. A lot of human ZFPs with different sequence specificity have been identified and

many new ones were generated. Multiple fingers can be fused together to produce a protein with a long recognition sequence to achieve a unique targeting in complex genomes as human [136]. TALE proteins were identified in bacteria of genus Xanthomonasin, they contain a central domain of tandem repeats, where one repeat recognizes one DNA base pair [137], [138]. The TALE's repeats can be assembled in arrays to generate DBDs with a customizable sequence specificity [138], [139]. Both ZFPs and repeats of TALEs utilize a modular structure but demand intensive protein engineering to generate a novel DBD with required sequence specificity.

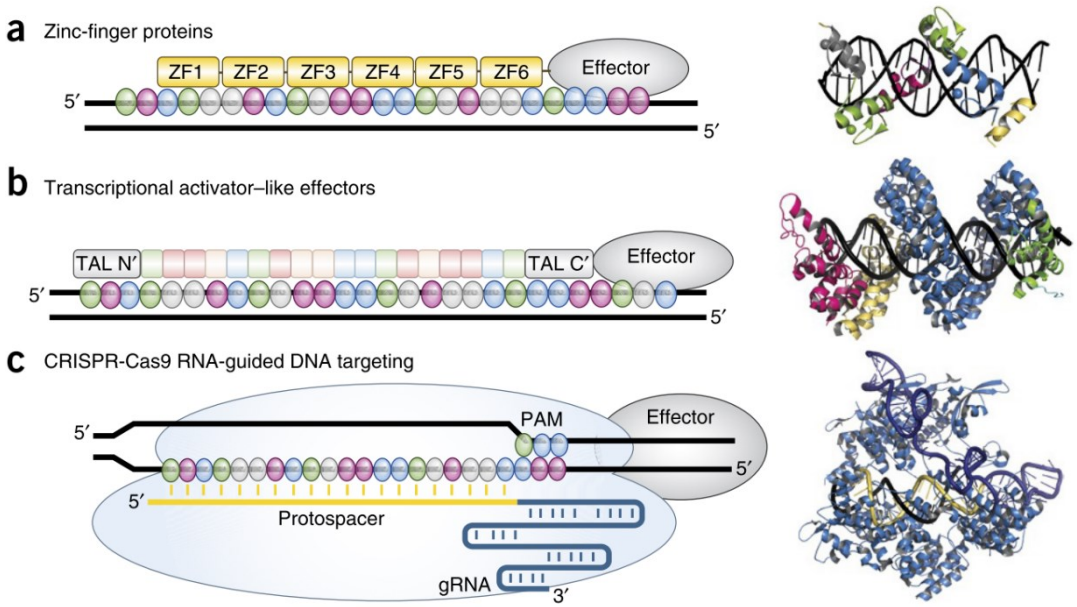

Figure 8. The most commonly used DBDs.
Schemes depicting functional elements of three types of DBDs and crystal structures of their complexes with DNA are shown. In the case of CRISPR/Cas9 a tripartite complex of Cas9 with DNA and a gRNA is shown in the crystal structure. a ZFPs. b TALEs. c CRISRP/Cas9. Figure taken from [140].

The last group of DBDs, Cas proteins, were discovered in bacteria as a component of their adaptive defense system CRISPR/Cas [141]. Cas proteins are nucleases which use short RNAs complementary to viral genomic DNA for targeting a DNA site. They form a DNA/RNA duplex and then cleave both DNA strands. Shortly after their discovery, CRISPR/Cas systems were adopted for the regulation of gene expression via targeting to a specific genomic locus [142]. In the current state of the art, Cas-based DBDs comprised of a catalytically inactive

Cas protein (dCas) and a single guide RNA (sgRNA) [143], [144]. The most widely used dCas protein is a dCas9 derived from *S. pyogenes* and to lesser extent another one cloned from *S. aureus*. Structurally sgRNAs contain a 20 nt guide sequence complementary to the target genomic DNA and a scaffold part required for the interaction with Cas proteins [143], [145]. The Cas/sgRNA complex requires the presence of proto-spacer adjacent (PAM) motifs for the interaction with DNA [146]. It is a short DNA motif (mostly 2-6 base pairs long) positioned directly at the 3'-end of the target DNA sequence, the sequence of which depends on the type of the CRISPR-Cas system. Of the three DBDs the dCas9/sgRNA system gained popularity very rapidly, since its retargeting is very straightforward and requires only to provide a new sgRNA with the desired sequence specificity [147]–[149]. Hence, it does not involve unpredictable protein engineering as for ZFPs and TALEs.

The second functional unit of EpiEditors is a chromatin modifying enzyme - or in a broader sense – an isolated catalytic domain or a protein/domain that can recruit a chromatin-modifying enzyme. Various proteins have been used for this purpose, demonstrating the general feasibility of this approach. Activation of gene expression was achieved via targeted acetylation of histone H3 lysine K27 at promoters or enhancers mediated by the catalytic core of the acetyltransferase p300 [150]. Genes silenced by promoter methylation were activated via DNA demethylation triggered by the ten-eleven translocation (TET) dioxygenase1 catalytic domain [151]–[153]. Targeted gene silencing has been demonstrated upon use of EpiEditors setting repressive chromatin marks as methylation of histone H3 lysine 9 [154]–[156], methylation of histone H3 lysine 27 [156], [157] and methylation of DNA.

The concept of targeted DNA methylation for gene silencing was proposed in 1997 [158] and the first artificial protein consisting of the ZFP and the M.*Sss*I methyltransferase from *Spiroplasma* species was produced as well. The fusion proteins methylated *in vitro* only substrates containing the ZFP binding site but not the substrate without this sequence. The first in vivo proved targeted activity was published in 2003 by Carvin and colleagues [159], who used a bacterial

methyltransferase M.*Cvi*PI fused to transcriptional activator PHO4 to target methylation to the PHO gene in yeast *Saccharomyces cerevisiae*. The same group showed later the *in vivo* activity of the ZFP fusions with M.*Sss*I or M.*Cvi*PI in yeast cells [160]. The first demonstration of targeted methylation in human cells was published in 2007 by two groups, in both cases designed EpiEditors used ZFPs as DBDs. One publication reported the catalytic domains of DNMT3A and DBNMT3B [161] and the second used mutants of the *Hpa*II and *Hha*I methyltransferases as effector domains [162].

The first prove of DNA methylation-based transcription repression after *de novo* methylation using ZFPs fused to DNMT3ACD and DNMT3BCD was done in HEK293 cells using a reporter plasmid [161]. Next, repression of transcription from the genomically integrated reporter via transient expression of ZFP-*Hpa*II fusion in NIH/3T3 cells was observed [163] and additional loss of H3K4me3 and gain of H3K9me2 was observed. The first silencing of endogenous genes Maspin and SOX2 in breast cancer cells, leading to a stable repression and reprogramming of cancer phenotype, was achieved using designed ZFPs-fused with to DNMT3ACD [164]. Afterward, several studies have reported targeted deposition of DNA methylation and silencing of corresponding target genes using EpiEditors based on TALE [165], [166] or CRISPR/dCas [153], [167]–[169] DBDs. The most recent modification of the dCas9-based targeting is its fusion with the so called SunTag [170]. The SunTag is a repeat of up to 24 peptides from the general control protein 4 (GCN4), which works as a binding platform for proteins fused to a single-chain variable fragment (scFv) antibody against the peptide. This strategy has been used for the epigenome editing and confirmed its efficiency due to recruitment of multiple chromatin-modifying enzymes [152], [171], [172]. Various DNA methyltransferases have been used as effector domains, the most frequent of those are the full-length DNMT3A, DNMT3ACD, and the DNMT3ACD-DNMT3LCD fusion [153], [166]–[169], [171]–[174].

One of the major problems of the approach is methylation of untargeted loci, also referred to as off-target activity. Untargeted DNA methylation was already documented in two early studies, which authors explained by interaction of

methyltransferases with DNA [159], [160]. To reduce off-target DNA methylation mutagenesis of the methyltransferases *Hpa*II and *Hha*I was done, it led to reduction of activity or reduction of affinity to DNA, which increased specificity [162]. Another approach tested to reduce off-target methylation is generation of a split methyltransferase. The coding sequence of the enzyme is divided into two parts, each of which is fused to two different ZFP, which bind neighboring DNA sequences. Because a functional methyltransferase assembles only upon binding of both fusion proteins to their target sequences, off-target binding is reduced [175]. This approach was tested in E. coli overexpressing fusion proteins and a plasmid with the target sites as a substrate for methylation. It is important to mention that despite simplified design of EpiEditors and improved editing specificity following the implementation of new DBDs such as TALEs and dCas9, several recent reports showed off-target methylation at many loci genome-wide originating from untargeted activity of DNA methyltransferases, indicating that still more specific EpiEditors are needed [171], [173], [174].

The first attempt to regulate expression of VEGFA was conducted in 2001, when the group of A. Wolffe designed several synthetic zinc finger proteins (ZFPs) targeting different sequences in DNaseI-hypersensitive regions within the VEGFA promoter [176]. They fused ZFPs with VP16 or p65 transcriptional activators, which upon recruitment to the promoter activated transcription of the VEGFA gene. Much later the same ZFP was used in another study aiming to silence VEGFA expression. DNMT3ACD or DNMT3ACD-DNMT3LCD fused to the ZFP were used for targeted methylation of the VEGFA in the ovarian cancer cell line SKOV3 [177]. The artificial fusion protein DNMT3ACD-DNMT3LCD was 10 times more active *in vitro* than the DNMT3ACD alone [177]. Both constructs demonstrated *de novo* methylation at the target region, addition of the DNMT3LCD increased deposition of DNA methylation by two-fold (up to 49%). The VEGFA expression was reduced by 36 and 56 % after transient transfection of the ZPF-DNMT3ACD and ZPF-DNMT3ACD-DNMT3LCD respectively. Later, the same ZPF-DNMT3ACD fusion protein was delivered into SKOV3 cells via adenoviral transduction and a time course showed that the methylation reached its maximum at day 5. Afterwards, however, it gradually decreased following the

loss of expression vector in the cells [178]. This lack of stability is not limited to the VEGFA gene but a more general phenomenon, thus indicating that further improvement of the technology is needed for stable gene repression.

1.6 Aims of the study

Anti-VEGFA and anti-VEGFRs therapy using various inhibitors demonstrated efficacy in multiple clinical trials and is now in routine use in clinical practice for the treatment of different malignancies. Targeting the VEGFA/VEGFRs axis has dual functions resulting in (a) reduction of tumor angiogenesis and slowing tumor growth and (b) blocking of autocrine and paracrine effects of VEGFA on cancer cells leading to inhibition of cell proliferation, invasiveness and EMT. Compared to inhibitors, which require regular repetitive administrations, epigenome editing to silence VEGFA, VEGFR1, VEGFR2 genes has the potential of simultaneous stable inactivation of these genes at the transcriptional level.

Previous *in vitro* studies successfully demonstrated that targeted DNA methylation of the VEGFA promoter using ZPF-based EpiEditors results in a reduction of VEGFA expression, but only moderate levels of methylation were achieved. In addition, multiple studies reported significant off-target effects of EpiEditing, which may potentially lead to silencing of undesired genes and prohibit prospective clinical application. Hence, the development of more specific EpiEditors is urgently needed. Lastly, a promising way to increase potency of the therapeutic approach is to target VEGFA and its receptors simultaneously, since this would inactivate three nodes in the pathway. Compared to the concomitant use of two or three anti-VEGFA/VEGFRs drugs in clinics, which would drastically increase treatment costs, gene silencing using epigenome editing has the capability to implement multiplex gene targeting for the cost of one gene.

Based on these considerations, the following aims have been set for the current study:

1) **Develop a strategy to achieve effective methylation of the VEGFA promoter**

The first step will be to change the genomic targeting module of EpiEditors from ZPF- to the CRISPR/Cas9 technology. This will allow to use and develop dCas9-based systems. For example, the dCas9-10xSunTag vector able to recruit up to ten effector domains will be applied and tested for its efficiency at the VEGFA locus. Additionally, the most active DNA methyltransferase, the artificial DNMT3ACD-DNMT3LCD, will be used as the effector domain.

2) **Establish multiplex methylation of the VEGFA, VEGFR1 and VEGFR2 promoters**

   Use of the CRISPR/Cas9 technology will also allow to set DNA methylation at multiple genes using dCas9 as the DBD and sgRNAs targeting individual genes. Towards this end, sgRNAs targeting the VEGFA-pathway genes will be designed and the efficiency tested by analyzing DNA methylation at the target loci. To increase the efficiency of multiplex editing sgRNAs further, targeting three genes will be cloned into one expression vector.

3) **Develop approaches for targeted DNA methylation with higher specificity than the published ZFP-based EpiEditors**

   The specificity of targeted epigenome editing is a consequence of the specificities of the DBD and the effector domain. The use of dCas9 instead of the ZFP as the DBD has the additional advantage that this will decrease the number of EpiEditor binding sites in the genome. Furthermore, two variants of the DNMT3ACD-DNMT3LCD effector domain, the wild type and its more specific mutant R887E will be used for DNA methylation. Their on-target and off-target activity will be compared.

4) **Detailed analysis of targeted DNA methylation patterns to extract guidelines for future experiments**

   To increase the efficiency of epigenome editing in future experiments, patterns of introduced DNA methylation will be analysed in detail. DNA methylation levels will be investigated on both DNA strands. Also, a relationship between methylation of individual CpG sites to their distance

from the sgRNA binding site and their flanking sequences will be studied. This should help to elucidate structural rules about epigenome editing.

2 Material and methods

2.1 Generation of vectors for the expression of single sgRNAs

2.1.1 Design of sgRNAs

A target sequence of the sgRNA binding at the VEFGA promoter overlapping with the ZFP binding sites was selected manually. Firstly, NGG PAM sequences of dCas9 from *S. pyogenes*, were found in the close proximity to the ZFP binding site on the antisense strand. Secondly, the first PAM sequence 3' from the ZFP binding site, was selected since the corresponding sgRNA sequence would overlap with the ZFP binding site. 19 nt 5' to the PAM sequence were used to generate sgRNA's target-specific sequence (Table 1).

sgRNAs targeting open chromatin regions within the VEGFA, VEGFR1 and VEGFR2 promoters were also selected manually. The desired target regions were chosen based on the DNase-seq and ATAC-seq data sets (GEO accession GSE108513). The PAM sequences were allocated on the flanks of the target regions keeping in mind a directionality of the dCas9/sgRNA binding to DNA. The PAM sequence for the 5' flank sgRNA was selected on the sense strand and the 3' sgRNA PAM was selected on the antisense strand. For all sgRNAs a 20 nt sequence 5' to the corresponding PAM sequence was used to generate an sgRNA's target-specific sequence (Table 1).

Table 1. sgRNA target-specific sequences and their genomic coordinates

	sgRNA target-specific sequence (5' to 3')	Genomic coordinates, based on hg19	Strand
VEGFA/ZFP	GGCGGTCACCCCCAAAAGC	chr6:43,738,374-43,738,392	antisense
VEGFA/sgRNA1	CAGAGTTTCCGGGGGCGGAT	chr6:43,737,695-43,737,714	sense
VEGFA/sgRNA2	GCCACGACCTCCGAGCTACC	chr6:43,737,968-43,737,987	antisense
VEGFR1/sgRNA1	CGCCCTGAGCGCCCGTCTCG	chr13:29,069,080-29,069,099	sense
FLT2/sgRNA2	GACCCCTTGACGTCACCAGA	chr13:29,069,334-29,069,353	antisense
VEGFR2/sgRNA1	CCAGCGCAGTCCAGTTGTGT	chr4:55,991,929-55,991,948	antisense
VEGFR2/sgRNA2	GGGCGTCTGCGGGTGCCGGT	chr4:55,991,624-55,991,643	sense

2.1.2 Cloning of sgRNA encoding vectors

To generate vectors for sgRNA expression, two different approaches were used. The VEGFA-targeting sgRNA at the ZFP binding site was cloned into the gRNA_Cloning_Vector kindly provided by George Church (obtained via Addgene, plasmid # 41824) following a published protocol [179]. First, the sgRNA target-specific sequence and its reverse complement were extended by adaptors (Table 2). Synthesis of obtained oligonucleotides was ordered at Integrated DNA Technologies. Oligonucleotides were annealed and extended in a thermocycler using the following conditions.

Reaction mixture		
	Concentration	Volume
Oligo 1	100 µM	1.0 µL
Oligo 2	100 µM	1.0 µL
dNTPs	10 mM	0.4 µL
Phusion HF buffer	5X	4.0 µL
Phusion pol.		1.0 µL
H_2O		12.6 µL

Reaction program	
95 ℃	Pause
95 ℃	1 min
Cooling to 20 ℃	0.03 ℃/sec
20 ℃	30 min
8 ℃	Pause

2 µL of reaction products were resolved using 10 % polyacrylamid gel electrophoresis in Tris-phosphate-EDTA (TPE) buffer to confirm successful synthesis. The rest was purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel GmbH) and eluted in 15 µL MilliQ water. DNA concentration was measured by NanoDrop (ThermoFisher Scientific).

gRNA_Cloning_Vector was linearized by AflII restriction enzyme for 2 h at 37 ℃ in the following reaction.

Reaction mixture	
Vector #41824	4 µg
10xCutSmart buffer	2 µL
AflII	2 µL
H_2O to final volume	20 µL

Products of digestion were resolved on 1 % agarose gel in TPE buffer for 1 h at 80 V. The band at 3.9 kb corresponding to linearized plasmid was cut from the gel, purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel GmbH) and eluted in 15 µL MilliQ water. DNA concentration was measured by NanoDrop (ThermoFisher Scientific).

Finally, the vector was generated using Gibson Assembly Master Mix (New England BioLabs Inc.) using following conditions.

Reaction mixture	
sgRNA coding dsDNA	25 ng
Linearized vector #41824	75 ng
H_2O	to 5 µL
2x Gibson Assembly Master mix	5 µL
Final volume	10 µL

Reaction program	
50 ℃	10 min
40 ℃	10 min
8 ℃	Pause

The sgRNAs targeting open chromatin regions within the promoters of VEGFA, VEGFR1 and VEGFR2 genes were cloned into the sgRNA-GGA2 cloning vector generated in the Department of Biochemistry at the Institute of Biochemistry and Technical Biochemistry at Stuttgart University (Catalogue #02_87). The target-specific sequences of sgRNAs and their revers complements were extended by adaptors required for cloning (Table 2) and these oligonucleotides were synthesized by Integrated DNA Technologies. The produced pairs of oligonucleotides were annealed in a thermocycler using following conditions.

Reaction mixture		
	Concentration	Volume
H_2O		16 µL
NEBuffer 2	10x	2 µL
Oligo 1	100 µM	1.0 µL
Oligo 2	100 µM	1.0 µL

Reaction program	
95 ℃	Pause
95 ℃	1 min
Cooling to 20 ℃	1 ℃ / min
20 ℃	1 min
8 ℃	Pause

sgRNA encoding vectors were cloned using Golden Gate Assembly protocol and following conditions.

Reaction mixture	
#02_87 plasmid	75 ng
Pre-annealed oligonucleotides from the previous step	2-fold molar excess over plasmid
T4 DNA ligase (New England BioLabs Inc.)	400 units
10x T4 DNA ligase buffer (New England BioLabs Inc.)	2 µL
BbsI-HF (New England BioLabs Inc.)	10 units

Reaction program		
37 ℃	Pause	
37 ℃	1 min	30 cycles
16 ℃	1 min	
37 ℃	10 min	
85 ℃	5 min	
8 ℃	Pause	

Table 2. Oligonucleotides used for cloning of sgRNAs

Name	Sequence (5’ to 3’)	Description
PB761	TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGGCGGTCACCCCCAAAAGC	VEGFA/ZFP sgRNA oligo 1
PB762	GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGCTTTTGGGGGTGACCGCCC	VEGFA/ZFP sgRNA oligo 2

PB817	accgCAGAGTTTCCGGGGGCGGAT	VEGFA/sgRNA1 oligo 1
PB818	aaacATCCGCCCCCGGAAACTCTG	VEGFA/sgRNA1 oligo 2
PB819	accGCCACGACCTCCGAGCTACC	VEGFA/sgRNA2 oligo 1
PB820	aaacGGTAGCTCGGAGGTCGTGG	VEGFA/sgRNA2 oligo 2
PB823	accgCGCCCTGAGCGCCCGTCTCG	VEGFR1/sgRNA1 oligo 1
PB824	aaacCGAGACGGGCGCTCAGGGCG	VEGFR1/sgRNA1 oligo 2
PB825	accGACCCCTTGACGTCACCAGA	VEGFR1/sgRNA2 oligo 1
PB826	aaacTCTGGTGACGTCAAGGGGT	VEGFR1/sgRNA2 oligo 2
PB827	accgCCAGCGCAGTCCAGTTGTGT	VEGFR2/sgRNA1 oligo 1
PB828	aaacACACAACTGGACTGCGCTGG	VEGFR2/sgRNA1 oligo 2
PB829	accGGGCGTCTGCGGGTGCCGGT	VEGFR2/sgRNA2 oligo 1
PB830	aaacACCGGCACCCGCAGACGCC	VEGFR2/sgRNA2 oligo 2

2.2 Generation of vectors for expression of multiple sgRNAs

2.2.1 Amplification of the sgRNA expression cassettes

To generate multiple sgRNA expressing vectors sgRNA expression cassettes were amplified from the single sgRNA encoding vectors. Each cassette was amplified with the unique combination of primer pairs containing overhangs with the BbsI restriction sites. The primers were synthesized by IDT (Integrated DNA Technologies), sequences are listed in Table 3. The following reaction conditions were used for PCR.

Reaction mixture		
	Concentration	Volume
H_2O		31 µL
Q5 buffer	5x	10 µL
dNTPs	10 mM	1 µL
Q5 pol.		0.5 µL
Primer 1	10 µM	2.5 µL
Primer 2	10 µM	2.5 µL
Plasmid DNA		2.5 ng
	Total volume	50 µL

Reaction program		
98 ℃	Pause	
98 ℃	2 min	
98 ℃	10 sec	35 cycles
57 ℃	15 sec	
72 ℃	20 sec	
72 ℃	2 min	
8 ℃	Pause	

5 µL of reaction products were resolved on 1 % agarose gel in TPE buffer for 1 h at 80 V to confirm that expected product was obtained (541 bp). The rest was digested with 1 µL DpnI enzyme in 1x CutSmart buffer (New England BioLabs Inc.) for 1 h at 37 ℃. Afterward, DNA was purified with NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel GmbH) and eluted in 40 µL MilliQ water. DNA concentration was measured by NanoDrop (ThermoFisher Scientific).

Table 3. Primers used for amplification of sgRNAs expression cassettes

Name	Sequence (5' to 3')	Description
Fragments for the multi-sgRNA1 vector assembly		
PB905	GGCTACgaagacTATGCCCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFR1
PB906	TTCTACgaagacCCCATAAATTTACGAGCTT TCTGG	reverse primer for gRNA1 VEGFR1
PB907	GGCTACgaagacTATATGCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFR2
PB908	TTCTACgaagacCCAGTTAATTTACGAGCTT TCTGG	reverse primer for gRNA1 VEGFR2
PB909	GGCTACgaagacTAAACTCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFA
PB916	TTCTACgaagacCCTCTGAATTTACGAGCT TTCTGG	reverse primer for gRNA1 VEGFA
Fragments for the multi-sgRNA2 vector assembly		
PB905	GGCTACgaagacTATGCCCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFR1
PB910	TTCTACgaagacCCGAATAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFR1
PB911	GGCTACgaagacTAATTCCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFR2
PB912	TTCTACgaagacCCCCTAAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFR2
PB913	GGCTACgaagacTATAGGCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFA
PB916	TTCTACgaagacCCTCTGAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFA
Fragments for the multi-sgRNA3 vector assembly		
PB905	GGCTACgaagacTATGCCCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFR1
PB906	TTCTACgaagacCCCATAAATTTACGAGCTT TCTGG	reverse primer for gRNA1 VEGFR1
PB907	GGCTACgaagacTATATGCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFR2
PB908	TTCTACgaagacCCAGTTAATTTACGAGCTT TCTGG	reverse primer for gRNA1 VEGFR2
PB909	GGCTACgaagacTAAACTCCAAACTCATCA ATGTATCT	forward primer for gRNA1 VEGFA
PB910	TTCTACgaagacCCGAATAATTTACGAGCT TTCTGG	reverse primer for gRNA1 VEGFA
PB911	GGCTACgaagacTAATTCCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFR1

PB912	TTCTACgaagacCCCCTAAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFR1
PB913	GGCTACgaagacTATAGGCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFR2
PB914	TTCTACgaagacCCCGTTAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFR2
PB915	GGCTACgaagacTAAACGCCAAACTCATCA ATGTATCT	forward primer for gRNA2 VEGFA
PB916	TTCTACgaagacCCTCTGAATTTACGAGCT TTCTGG	reverse primer for gRNA2 VEGFA

2.2.2 Assembly of multiple sgRNA expression cassettes in one vector

Amplified sgRNA expression cassettes were cloned into the pMulti-sgRNA-LacZ-DsRed vector [180] kindly gifted by Yujie Sun (obtained via Addgene, plasmid # 99914). In total 3 plasmids with 3, 3 and 6 sgRNA expression cassettes were produced by Golden Gate assembly using following conditions.

Reaction mixture for multi-sgRNA1	
10x T4 DNA ligase buffer (New England BioLabs Inc.)	2 µL
T4 DNA ligase, 400 units (New England BioLabs Inc.)	1 µL
BbsI-HF, 10 units (New England BioLabs Inc.)	1 µL
pMulti-sgRNA-LacZ-DsRed	20 ng
VEGFA sgRNA1 cassette	4 ng
VEGFR1 sgRNA1 cassette	4 ng
VEGFR2 sgRNA1 cassette	4 ng
H_2O	to 20 µL

Reaction mixture for multi-sgRNA2	
10x T4 DNA ligase buffer (New England BioLabs Inc.)	2 µL
T4 DNA ligase, 400 units (New England BioLabs Inc.)	1 µL
BbsI-HF, 10 units (New England BioLabs Inc.)	1 µL
pMulti-sgRNA-LacZ-DsRed	20 ng
VEGFA sgRNA2 cassette	4 ng
VEGFR1 sgRNA2 cassette	4 ng
VEGFR2 sgRNA2 cassette	4 ng
H_2O	to 20 µL

Reaction mixture for multi-sgRNA3	
10x T4 DNA ligase buffer (New England BioLabs Inc.)	2 µL
T4 DNA ligase, 400 units (New England BioLabs Inc.)	1 µL
BbsI-HF, 10 units (New England BioLabs Inc.)	1 µL
pMulti-sgRNA-LacZ-DsRed	20 ng
VEGFA sgRNA1 cassette	4 ng

VEGFR1 sgRNA1 cassette	4 ng
VEGFR2 sgRNA1 cassette	4 ng
VEGFA sgRNA2 cassette	4 ng
VEGFR1 sgRNA2 cassette	4 ng
VEGFR2 sgRNA2 cassette	4 ng
H_2O	to 20 µL

Reaction program		
37 ℃	Pause	
37 ℃	1 min	30 cycles
16 ℃	1 min	
37 ℃	10 min	
85 ℃	5 min	
8 ℃	Pause	

2.2.3 Transformation

Products of Gibson Assembly and Golden Gate Assembly reactions were diluted 3-fold with MilliQ water. 2 µL of diluted products were used for transformation of electrocompetent *E. coli* XL1-blue strain. 50 µL of cells were mixed with DNA and electroporated with 1.8 kV for 4 ms. Afterward, 1 mL of lysogeny broth (LB) medium was added and the cells were incubated for 1 h at 37 ℃ and 150 rpm. The entire suspension was plated on 1 % LB agar plate (w/v) containing 10 µg/mL tetracycline and 25 µg/mL kanamycin. Plates were let open to dry for 10 min, afterward they were incubated at 37 ℃ overnight.

2.2.4 Isolation of plasmid DNA

Obtained *E. coli* colonies were inoculated in 3 mL and 30 mL LB medium supplemented with 10 µg/mL tetracycline and 25 µg/mL kanamycin and grown at 37 ℃ overnight. Plasmid DNA was isolated from 3 mL overnight cultures using NucleoSpin Plasmid kit (Macherey-Nagel GmbH) and eluted with 30 µL Tris-EDTA (TE) pH 8.0 buffer. Obtained DNA was sequenced, correct clones were identified and expanded by midiprep plasmid isolation from 30 mL overnight cultures using QIAGEN Plasmid Plus Midi kit (QIAGEN GmbH). At last DNA was eluted with 100 µL TE buffer. DNA concentration was measured by NanoDrop (ThermoFisher Scientific).

2.2.5 Sanger sequencing of vectors

Obtained clones were sent for Sanger sequencing to confirm correct cloning. Sequencing was done at Microsynth AG. 1000 ng plasmid DNA was mixed in 1.5 mL Eppendorf tube with 2 µL 10 µM sequencing primers listed in Table 4. Sequencing results were analysed using the SnapGene program (GSL Biotech LLC).

Table 4. Sequencing primers

Name	Sequence (5' to 3')
For single sgRNA expression vectors	
PB330	GTGGTTTGTCCAAACTCATC
For multi-sgRNA1 and multi-sgRNA2 expression vectors	
PB330	GTGGTTTGTCCAAACTCATC
PB331	GTGGACTCTTGTTCCAAACTGG
For multi-sgRNA3 expression vector	
PB330	GTGGTTTGTCCAAACTCATC
PB331	GTGGACTCTTGTTCCAAACTGG
PB826	AAACTCTGGTGACGTCAAGGGGT
PB827	ACCGCCAGCGCAGTCCAGTTGTGT

2.3 Cell culture

2.3.1 Maintenance of cells

Human embryonic kidney 293 (HEK293) cells were cultured in T75 flasks in incubators at 37 °C and 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma-Aldrich, Inc.) supplemented with 10% fetal calf serum (FCS) (Sigma-Aldrich, Inc.), 20 ml/L L-glutamine (Sigma-Aldrich, Inc.) and 20 ml/L penicillin/streptomycin (Sigma-Aldrich, Inc.). Cells were subcultured when reaching 70-80% confluence. For this, growth medium was removed, cells were rinsed with Dulbecco's Phosphate Buffered Saline without calcium chloride and magnesium chloride (DPBS) (Sigma-Aldrich, Inc.), covered with 1 mL Trypsin-EDTA solution (Sigma-Aldrich, Inc.), incubated for 10 min in the CO2 incubator at 37 °C. Afterward, cells were resuspended in 9 mL and divided into 3 T75 flasks.

2.3.2 Cell counting in Neubauer haemocytometer

Cell resuspended after trypsinization (50 µL) were mixed with 0.4 % Trypan Blue Solution (50 µL) (ThermoFisher Scientific), mixed and 10 µL of the final

suspension were loaded onto a Neubauer haemocytometer. Viable unstained cells were counted in four 1 cm x 1 cm areas using EVOS microscope (ThermoFisher Scientific) under transmitted light settings, the average value was calculated and multiplied by 2x10,000 to get the number of cells in one milliliter.

2.3.3 Transient transfection of cells

Plasmids expressing dCas9-10xSunTag, scFv-GCN4-DNMT3ACD-DNMT3LCD (α3A3L) or α3A3L-R883E were published by us recently [171] and were available in the lab. HEK293 cell were harvested by trypsinization and the concentration of suspension was counted using Neubauer haemocytometer. 1.4 million HEK293 cells were seeded into 100 mm Petri dishes (ThermoFisher Scientific) in final 10 mL standard growth medium. 24 hours later, the medium was replaced by 8 mL growth medium. Plasmids were mixed for transfection in 840 µL serum free DMEM using the following recipe.

Plasmid name	Amount
dCas9-10xSunTag	6000 ng
wild type or R883E α3A3L	3000 ng
Multi-sgRNA1, 2 or 3	500 ng

Afterward, 27 µL FuGENE HD (Promega) transfection reagent was added and solution was mixed by pipetting multiple times. It was incubated for 20 minutes at room temperature for complex formation. Lastly, the obtained suspension was distributed drop-wise over one 100 mm Petri dish with adherent HEK293 cells. 24 hours later medium was removed and 10 mL fresh growth medium was added. Two individual transfections were conducted for each experimental condition, which were independently treated and analysed, leading to two biological replicates.

2.3.4 Fluorescence-activated cell sorting (FACS)

3 days after transfection cells were harvested by 1 mL trypsin, 1 mL of growth medium was added and cells were resuspended. Before sorting cells were filtered through 30 µL Pre-Separation Filters (Miltenyi Biotec) to get rid of cell aggregates. Cells were sorted on SH800S Cell Sorter (Sony Biotechnology) using 70-µm microfluidic sorting chips in Target mode. Since each of three co-transfected

plasmids has a unique reporter fluorescent protein, three lasers were used: 405 nm for tagBFP (dCas9-10xSunTag plasmid), 488 nm for superfoldedGFP (α3A3L plasmid), 561 nm for DgRed (multi-sgRNA1/2/3 plasmids). Untransfected HEK293 cells were used to set gaiting for negative and positive population. HEK293 cells transfected with single plasmids were used to calculate compensation parameters to exclude a fluorescence spillover between channels. Triple-positive cells were collected, centrifuged at 300 g for 5 min. Supernatant was removed and the cell pellets were processed for isolation of genomic DNA.

2.3.5 Genomic DNA isolation

Cell pellets were washed ones in PBS, pelleted by centrifugation for 5 min at 300 g. Genomic DNA was extracted using QIAamp DNA Blood Mini Kit (QIAGEN GmbH) following a manual. At the last step DNA was eluted by 100 µL Tris pH 8.0 buffer. DNA concentration was measured by NanoDrop (ThermoFisher Scientific). DNA was stored at -20 °C till the next step.

2.4 Locus-specific DNA methylation analysis

2.4.1 Bisulfite conversion

Genomic DNA (gDNA) was fragmented by enzymatic digestion with *Eco*RI restriction enzyme overnight at 37 °C in the following reaction mixture.

Reaction mixture		
	Concentration	Volume
CutSmart Buffer	10x	2 µL
*Eco*RI	20 units/µL	2 µL
gDNA		500 ng
H_2O		Up to 20 µL

Next day, fragmented DNA was bisulfite converted with EZ-DNA Methylation Lightning Kit (Zymo Research Corporation) following manufacture's protocol. Finally, DNA was eluted in 12 µL Tris pH 8.0 buffer. DNA was stored at -20 °C until use.

2.4.2 PCR1 with locus-specific primers

To generate libraries for sequencing a two-step PCR approach was used. At first target regions were amplified with from bisulfite-converted DNA primers (Table 5)

containing a sequence-specific part and additional overhangs with barcodes for labeling samples from different experiments, random pentamers N_5 to improve sequencing quality and adaptors for binding of the second PCR primers. For every set of primers an additional reaction was set as water control, where 1 µL of autoclaved water was added instead of DNA to control for possible sample cross-contamination.

Reaction mixture		
	Concentration	Volume
H_2O		14.4 µL
PCR buffer	10x	2 µL
dNTPs	10 mM	0.4 µL
HotStartTag pol.	5 U/µL	0.2 µL
Primer 1	10 µM	1 µL
Primer 2	10 µM	1 µL
DNA		1 µL
	Total volume	20 µL

Reaction program		
95 ℃	Pause	
95 ℃	15 min	
94 ℃	30 sec	35 cycles
X ℃, see below	30 sec	
72 ℃	1 min	
72 ℃	10 min	
8 ℃	Pause	

Annealing temperatures, X	
VEGFA sense	52 ℃
VEGFA antisense	50 ℃
VEGFR1 sense	50 ℃
VEGFR1 antisense	52 ℃
VEGFR2 sense	52 ℃
VEGFR2 antisense	50 ℃
ISG15	50 ℃

Table 5. PCR1 primer sequences

Name	**Sequence (5' to 3')**	**Description**
PB504	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTN NNNNAGAGCGTTTGTTATTTTTTATTTGAAT	VEGFA_ZFP_bis_fp
PB505	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNN NNAGCATAATCACTCACTTTACCCCTATC	VEGFA_ZFP_bis_rp
PB893	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNN NNAGCATTTTTAGGTTGTGAATTTTGGTG	VEGFA_UP_bis fp

PB894	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATCCTCCCCRCTACCAAC	VEGFA_UP_bis rp
PB895	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNTAGAGTTATTYGGTTGTTTTAAGTTT	VEGFA_DS_bis fp
PB896	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAATCRAACTTCCCCTTCAT	VEGFA_DS_bis rp
PB897	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNCTAGGTTTTAGTTAGGAGATAATTATTTT	VEGFR1_UP_bis fp
PB898	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCCCTTAACRTCACCAAA	VEGFR1_UP_bis rp
PB899	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNGCTGATTTTTTGAYGTTATTAGAAGG	VEGFR1_DS_bis fp
PB900	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCAAAAAACAACCACTTCC	VEGFR1_DS_bis rp
PB901	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNACGTAGGAGAGGATATTTAGGTTG	VEGFR2_UP_bis fp
PB902	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAACCCAACRCAATCCAA	VEGFR2_bis rp
PB903	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNTTCGGAAATGGGGAGATGTAAAT	VEGFR2_DS_bis fp
PB904	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATAAAAAAAAATATCCAAACTACC	VEGFR2_DS_bis rp
ISG15 fp	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNAGAGCTTAGGTGTTTTTAGGGTGTTGG	FP1 taken from [171]
ISG15 rp	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNAGCATCACAAACTCCTATACTAACAAAAATAAAT	RP1 taken from [171]

After PCR 5 µL of reaction products were resolved on 1 % agarose gel in TPE buffer of 40 min at 80 V.

2.4.3 PCR2 with indexing primers

Gels were examined and samples with efficient amplification resulting in generation of products with expected size and no secondary product and negative results in water control were selected. The reaction products were diluted 1:4 with autoclaved water and used as a template for the second PCR. This reaction introduces adaptors and indices for Illumina sequencing according to TruSeq protocol. Reactions were conducted with the custom set of primers (Table 6) following reaction conditions provided below. Again, a water control reaction was conducted with 1 µL of autoclaved water to control for possible sample and reagents cross-contamination.

Reaction mixture		
	Concentration	Volume
H_2O		12.8 µL
Q5 buffer	5x	4 µL
dNTPs	10 mM	0.4 µL
Q5 pol.		0.2 µL
Primer 1	10 µM	0.8 µL
Primer 2	10 µM	0.8 µL
PCR1 (1:4)		1 µL
	Total volume	20 µL

Reaction program		
98 ℃	Pause	
98 ℃	30 min	
98 ℃	10 sec	15 cycles
72 ℃	40 sec	
72 ℃	2 min	
8 ℃	Pause	

Table 6. PCR2 primer pairs sequences

Name	Sequence (5' to 3')	Description
PB611	AATGATACGGCGACCACCGAGATCTACACTCCGCGAA ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-710
PB650	AATGATACGGCGACCACCGAGATCTACACTCTCGCGC ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-711
PB652	AATGATACGGCGACCACCGAGATCTACACAGCGATAG ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-712
PB654	AATGATACGGCGACCACCGAGATCTACACTTCCTCCTA CACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-713
PB659	AATGATACGGCGACCACCGAGATCTACACTGCTTGCT ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-714
PB661	AATGATACGGCGACCACCGAGATCTACACGGTGATGA ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-715
PB667	AATGATACGGCGACCACCGAGATCTACACAACCTACG ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-716
PB673	AATGATACGGCGACCACCGAGATCTACACGGATCTGA ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-717
PB705	AATGATACGGCGACCACCGAGATCTACACTGATCACG ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-718
PB732	AATGATACGGCGACCACCGAGATCTACACAAGCGACT ACACTCTTTCCCTACACGACGCTCTTCCGATCT	i5-719
PB612	CAAGCAGAAGACGGCATACGAGATTTCGCGGAGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-710
PB651	CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGA CTGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-711
PB653	CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-712
PB655	CAAGCAGAAGACGGCATACGAGATAGGAGGAAGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-713
PB660	CAAGCAGAAGACGGCATACGAGATAGCAAGCAGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-714

PB662	CAAGCAGAAGACGGCATACGAGATTCATCACCGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-715
PB668	CAAGCAGAAGACGGCATACGAGATCGTAGGTTGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-716
PB674	CAAGCAGAAGACGGCATACGAGATTCAGATCCGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-717
PB706	CAAGCAGAAGACGGCATACGAGATCGTGATCAGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-718
PB733	CAAGCAGAAGACGGCATACGAGATAGTCGCTTGTGAC TGGAGTTCAGACGTGTGCTCTTCCGATCT	i7-719

After PCR 5 µL of reaction products were resolved on 1 % agarose gel in TPE buffer for 30 min at 110 V. Samples with negative water controls and prominent products at expected size were selected. Based on the band densities samples were pooled in one tube and purified using NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel GmbH) and eluted in 40 µL Tris pH 8.0 buffer. Concentrations of obtained libraries were measured by NanoDrop (ThermoFisher Scientific).

2.4.4 Next-generation sequencing

Libraries were sent to Admera Health Biopharma Services (USA) for NGS. Firstly, library quantification was done at the company using qPCR to determine molar concentration. Libraries were sequenced on Illumina MiSeq machine in 2x250 mode. Sets of paired reads of 250 bp were got from the company in fastqsanger format. Every pair of the fastqsanger files contains reads from one combination of indexes.

2.5 Bioinformatics analysis of NGS data

2.5.1 Analysis of DNA methylation at target loci

Obtained fastqsanger files were uploaded into the Galaxy server [181]. DNA methylation analysis have been conducted using the workflow described earlier [182]. Initially, the reads quality was analysed and Illumina adaptors and nucleotides with the score below 20 were trimmed by the Trim Galore! tool (developed by Felix Krueger, the Babraham Institute). Next, the paired reads were merged to produce one sequenced fragment using the Pear tool [183] under the standard settings and allowing minimum length of reads overlap of 20 bp. Further reads were filtered into separate pools based on barcodes used in the

first PCR and mapped onto reference sequences using the bwameth tool [184]. Generated bam file was processed further by MethylDackel (https://github.com/dpryan79/MethylDackel, developed by Devon Ryan) together with the reference sequences to extract methylation levels at each CpG sites. Data from the Galaxy servers were exported as tabular files and finally analysed and visualized in Microsoft Excel.

2.5.2 Extracting additional features of established DNA methylation patterns

Several bioinformatics analyses were conducted using Python scripts, which I wrote using freely available modules. First of all, all scripts were written and executed in Visual Studio Code (Microsoft Corporation). Python environment and package management was conducted with help of Anaconda (Anaconda Inc). Figure 21 and Figure 22 were generated using matplotlib v 3.3.1.

Two scripts were written by me to count occurrence of sequences in human genome.

Script #1. Count a sequence occurrence in human genome

```
"""
*********************************************************************
Reads FASTA file containing multiple sequences, counts an occurrence of
provided query in every sequence and sums up, saves statistics in csv
file.
v.200613
Pavel Bashtrykov
*********************************************************************

"""
from Bio import SeqIO
from Bio.Seq import reverse_complement
import csv
import time
#****************************************
#Input data:'
sequence = "GGGGGTGAC"
analyse_filename = "hg19.fasta"
output_filename = "report.csv"
#****************************************


def     counts_sequence_in_fasta(analyse_filename,     output_filename,
variants):
with open(output_filename, "a+", newline="") as csvfile:
filewriter = csv.writer(csvfile, delimiter=",")
for seq_record in SeqIO.parse(analyse_filename, "fasta"):
```

```
for search_seq in variants.keys():
occurrence = seq_record.seq.upper().count(search_seq)
variants[search_seq] += occurrence

all_observations = 0
for key, val in variants.items():
filewriter.writerow([key, val])
all_observations += val
filewriter.writerow(["Total", all_observations])


def generate_sequences(sequence):
"""Generates a reverse complement sequence.
Makes a list of queries: original sequence + reverse_complement.
"""
list_of_sequences = []
list_of_sequences.append(sequence)
rev_com = reverse_complement(sequence)
list_of_sequences.append(rev_com)
variants = {}
for var in list_of_sequences:
variants[var] = 0

return variants
#*********************************************************************
if __name__ == "__main__":
variants = generate_sequences(sequence)
counts_sequence_in_fasta(analyse_filename, output_filename, variants)
#*********************************************************************
```

Script #2. Count a sequence occurrence in human genome allowing a single nucleotide mismatch

```
"""
*********************************************************************
Reads FASTA file containing multiple sequences, counts an occurrence of
provided query in every sequence and sums up, saves statistics in a csv
file.
v.200613
Pavel Bashtrykov
*********************************************************************
"""

from Bio import SeqIO
from Bio.Seq import reverse_complement
import csv
import time


#****************************************
# Input data:'
sequence = "GGCGGTCACCCCCAAAAGC"
analyse_filename = "hg19.fasta"
output_filename = "report.csv"
#****************************************
```

```
def     counts_sequence_in_fasta(analyse_filename,     output_filename,
variants):
with open(output_filename, "a+", newline="") as csvfile:
filewriter = csv.writer(csvfile, delimiter=",")
for seq_record in SeqIO.parse(analyse_filename, "fasta"):
for search_seq in variants.keys():
occurrence = seq_record.seq.upper().count(search_seq)
variants[search_seq] += occurrence

all_observations = 0
for key, val in variants.items():
filewriter.writerow([key, val])
all_observations += val
filewriter.writerow(["Total", all_observations])


def generate_mut_sequences(sequence):
"""
Makes a list of sequences based on an original sequence by mutagenesis
of every single position.
"""
list_of_sequences = []
letters ="GATC"
length = len(sequence)
for n in range(length):
for i in letters:
new_sequence = sequence[0:n]+i+sequence[n+1:]
list_of_sequences.append(new_sequence)
rev_com = reverse_complement(new_sequence)
list_of_sequences.append(rev_com)
variants = {}
for var in list_of_sequences:
variants[var] = 0
return variants

#*********************************************************************

if __name__ == "__main__":
variants = generate_mut_sequences(sequence)
counts_sequence_in_fasta(analyse_filename, output_filename, variants)
#*********************************************************************
```

3 Results

3.1 Methylation of the VEGFA promoter using the dCas9-10xSunTag/ scFv-GCN4-DNMT3ACD-DNMT3LCD system

This work is a continuation of the research published earlier, where targeted methylation of the VEGFA promoter to silence the VEGFA expression had been achieved using fusion proteins ZFP-DNMT3ACD or ZFP-DNMT3ACD-DNMT3LCD (termed ZFP-3A3L from here) [177], [178]. The ZFP applied in these experiments recognizes the GGGGGTGAC sequence located 434 bp downstream of the VEGFA TSS [176]. The maximum methylation achieved in these experiments was 49 % using transient transfection of ZFP-3A3L [177] and 43 % using adenoviral delivery of ZFP-DNMT3ACD [178]. To achieve higher methylation efficiency, a recently developed dCas9-10XSunTag (termed dCas9S from here) system that can recruit up to ten scFv-GCN4-DNMT3ACD-DNMT3LCD effector domains (termed α3A3L from here) was used in this work. To compare the efficiency of targeted DNA methylation of the dCas9S/α3A3L system with the previously used ZFP-3A3L, which delivers only one 3A3L protein to a target site, a sgRNA binding to the same locus as the ZFP was selected (Figure 9a). The dCas9 used in this experiment was derived from *S. pyogenes* and requires an NGG trinucleotide sequence as PAM site. Multiple NGG motifs are present in the VEGFA promoter region including a few in the vicinity of the ZFP binding site. As shown in previous studies [169], the direct fusion protein dCas9-3A3L has a directionality of DNA methylation deposition, meaning that DNA methylation will appear mostly on one side relative to the dCas9/sgRNA complex binding site, namely the site of the PAM sequence. Keeping this in mind a sgRNA sequence was selected that covers the ZFP biding site and should lead to the deposition of DNA methylation mostly at the locus analysed in the previous studies [177], [178] (Figure 9b).

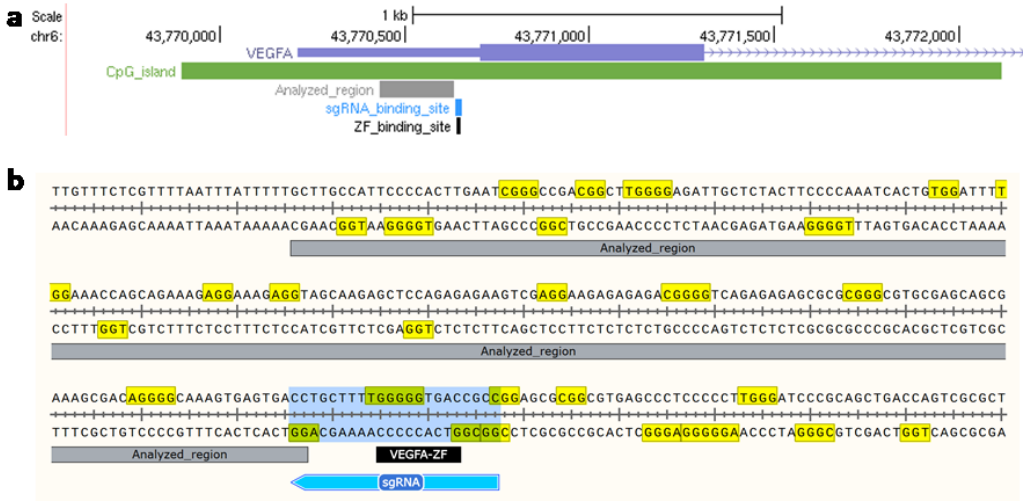

Figure 9. Design of the sgRNA targeting the VEGFA promoter.
a UCSC Genome Browser view showing the VEGFA promoter region with the CpG island (green bar), binding sites of the ZFP (black bar) and the sgRNA (blue bar), and the region analysed by bisulfite NGS (grey bar). **b** Partial sequence of the VEGFA promoter. *S. pyogenes* NGG PAM sequences are highlighted in yellow, the region analysed by bisulfite sequencing is shown as a grey bar, the binding site of the ZFP is shown as a black bar and the manually selected sgRNA binding site is highlighted with blue color.

Oligonucleotides required for the sgRNA cloning were synthesized, annealed and extended to produce a 100 bp double-stranded DNA fragment. The latter one was used in a Gibson Assembly reaction with the sgRNA cloning plasmid resulting in a vector for expression of the VEGFA sgRNA. The sequence of the obtained plasmid was confirmed by Sanger sequencing (Figure 10).

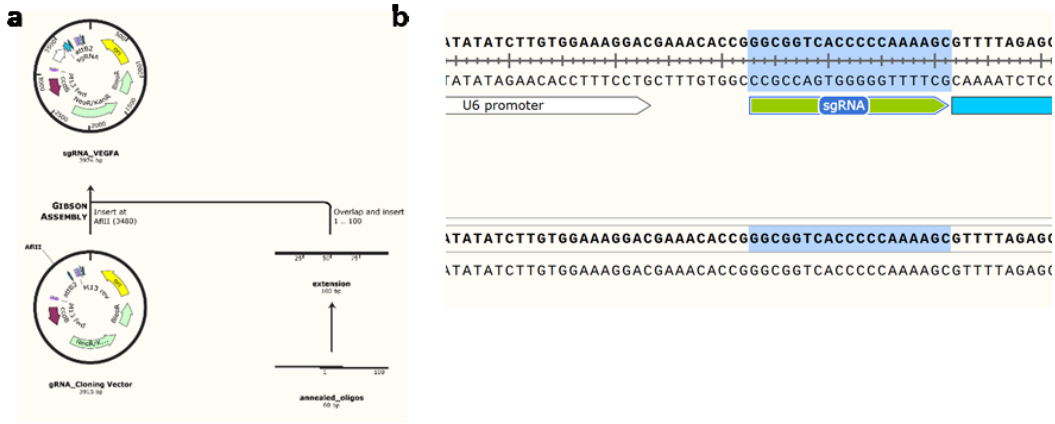

Figure 10. Cloning of the sgRNA targeting the VEGFA promoter.
a Scheme of the cloning strategy used to produce a VEGFA sgRNA expression vector. **b** Sanger sequencing results confirming correct sequence of the vector with the VEGFA sgRNA.

To determine the DNA methylation status of the native VEGFA promoter, gDNA of untreated HEK293 cells was analysed, revealing less than 1 % methylated CpG sites (Figure 11a). To study the efficiency of targeted DNA methylation, HEK293 cells were transfected with the cocktail of plasmids encoding dCas9S, α3A3L and sgRNA. Treatment of cells resulted in methylation of the analysed region to various extent (Figure 11a). The methylation level varied between CpG sites with a minimum value of 38 % and a maximum of 88 %. The average DNA methylation calculated based on the 12 CpG sites present in the analysed region was 69 %. Two independent transfections and downstream DNA methylation analysis were conducted and showed reproducible results with the mean DNA methylation of the targeted region equal to 69 % (Figure 11b). Thus, the dCas9S/α3A3L EpiEditor recruiting multiple effector domains was able to methylate the targeted genomic region and as expected the level of DNA methylation was higher than in earlier experiments used the ZFP-based targeting.

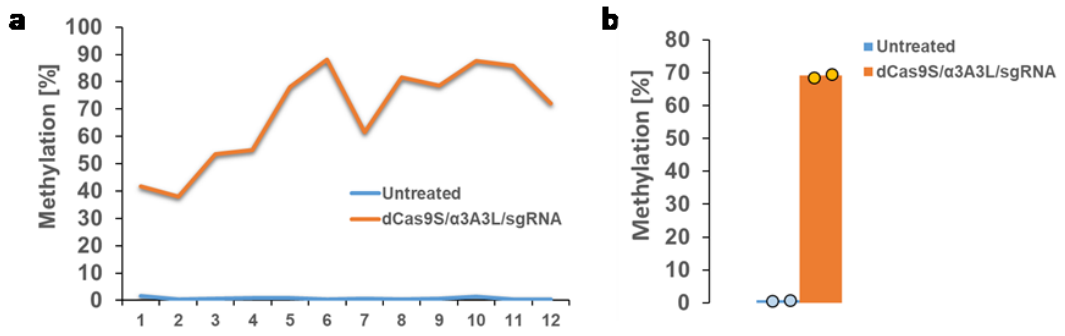

Figure 11. Targeted methylation of the VEGFA promoter.

a Line chart showing the methylation levels of individual CpG sites in the VEGFA region analysed by bisulfite NGS in untreated HEK293 cells (blue line) and 3 days after transfection with the dCas9S/α3A3L/sgRNA (orange line) (data show one biological experiment). **b** DNA methylation levels of the analysed region in untreated and treated HEK293 cells (average methylation and individual values are shown, n=2 biological replicates).

3.2 Targeting multiple genes of the VEGFA pathway

To increase the antineoplastic effect, three main directions could be used, namely, targeting multiple pathways, targeting multiple nodes in one pathway or a combination of both. In order to investigate the potential for improvement within

a given pathway, the second approach was implemented here. It is known that VEGFA is a secreted protein, which works as paracrine and autocrine growth factor stimulating cell migration, proliferation and angiogenesis, while effects of VEGFA are mediated via the VEGFR1 and VEGFR2 receptors. Thus, depletion not only of VEGFA but also of both main receptors of its signaling pathway should in theory enhance the efficiency of the whole strategy. To silence all three genes VEGFA, VEGFR1 and VEGFR2 simultaneously, an appropriate DNA methylation strategy was developed.

3.2.1 Design of sgRNAs targeting the VEGFA, VEGFR1 and VEGFR2 promoters

DNA methylation represses genes via two mechanisms, generation of more condensed chromatin and prevention of transcription factors binding to their recognition sequences. Many transcription factors have CpG sites in their recognition sequences and some have reduced affinity to the sequences if the CpG sites are methylated [185]. Thus methylation of transcription factor binding sites in gene promoters should be expected to reduce transcription rate. Usually multiple transcription factors influence expression of one gene in a cell line dependent way. It is a difficult task to identify the main TFs involved in the regulation of a particular gene in a particular cell line. To simplify the task one may look at regions where most of the TFs bind. Often, these are the regions of open chromatin, where DNA is more accessible for DNA interacting proteins. These regions can be identified mainly by two techniques, namely DNase I hypersensitive sites sequencing (DNase-seq) and assay for transposase-accessible chromatin using sequencing (ATAC-seq). Data produced by using these techniques for HEK293 cells were identified in the GEO database (GEO accession GSE108513), downloaded and visualized using the UCSC Genome Browser (Figure 12). As one can see at the VEGFA promoter region (Figure 12a), there is an open chromatin area (orange dashed frame) about 250 bp upstream of the region analysed by bisulfite NGS (old target region) that was discovered by both assays. Additionally, it had been published that the TF Sp1 promotes angiogenesis and migration of SKOV3-T ovarian cancer cells [52] via binding sites that are also located in this open chromatin region [51]. Based on these data

this area was selected as new target region for the dCas9-based EpiEditors. Retargeting was attempted by selecting two new sgRNAs in the vicinity of the new target region (Figure 12a). Following this strategy, open chromatin regions at the VEGFR1 and VEGFR2 promoters were also analysed. Targeting regions were selected based on the overlap of the DNase-seq and ATAC-seq data sets. Two sgRNAs were designed per gene using the NGG sequences available in the vicinity of the target region ends (Figure 12b,c).

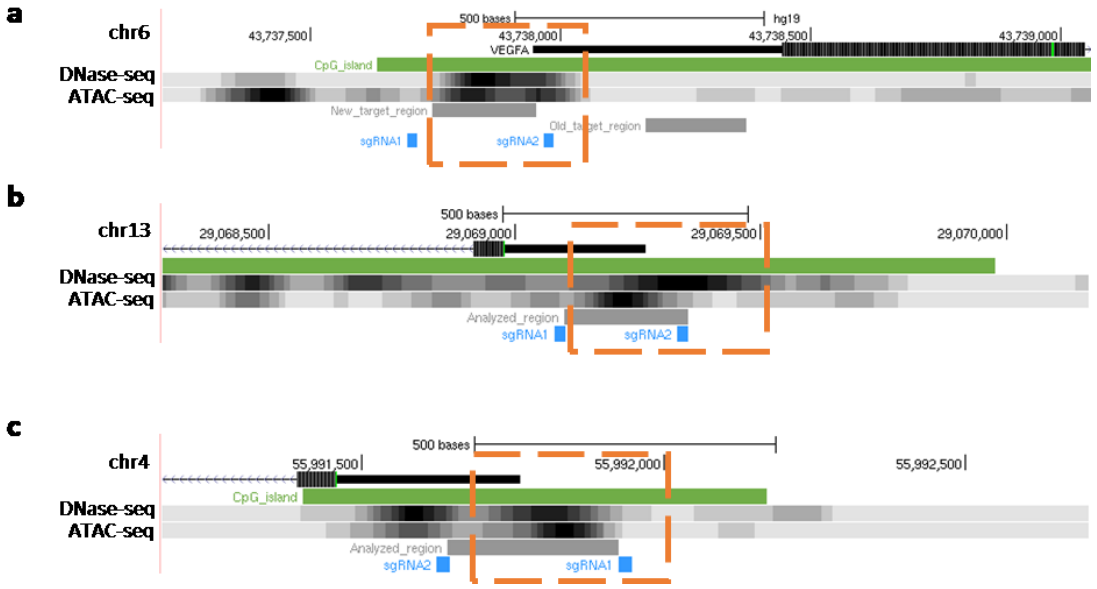

Figure 12. Selection of the target regions for DNA methylation of the VEGFA, VEGFR1 and VEGFR2 promoters.

UCSC Genome Browser views showing the VEGFA (**a**), VEGFR1 (**b**) and VEGFR2 (**c**) promoter regions with the CpG islands (green bar), two data tracks depicting DNase-seq and ATAC-seq of the HEK293 cells. Open chromatin regions are shown by orange dashed frames and sgRNA binding sites are shown as blue bars. Target regions for DNA methylation and subsequent bisulfite sequencing are shown as grey bars and labeled as “Analysed region”. In **a** the former target region used for ZFP-targeted methylation (Old target region) and the one designed for the current study (New target region) are indicated.

At first, six single sgRNA vectors, two for each gene, were cloned (Figure 13). To achieve multiple gene methylation, co-transfection of cells with a mix of the three sgRNA expressing vectors would be one practical possibility. However, in order to ensure that cells get all three sgRNAs one could conduct an additional cloning step to assemble all three sgRNA expression cassettes in one vector. This approach was utilized in this study and the multi-sgRNA1 vector was cloned for expression of three sgRNA1 targeting the genes VEGFA, VEGFR1 and VEGFR2

(Figure 14a,b). Using the same approach, the three sgRNA2 targeting the same genes were cloned into multi-sgRNA2 vector (Figure 14c).

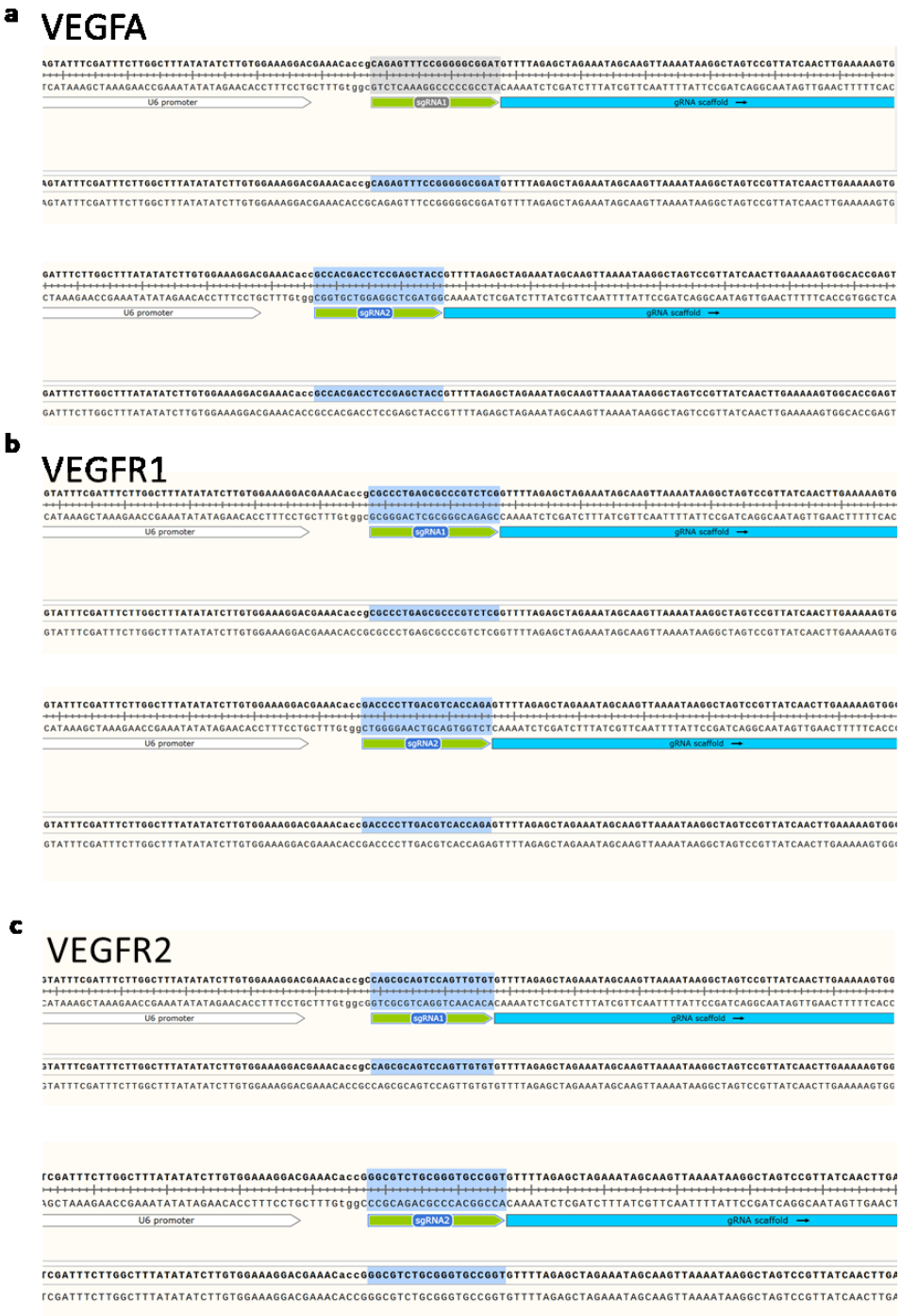

Figure 13. sgRNAs targeting the VEGFA, VEGFR1 and VEGFR2 promoters cloned into the single sgRNA expression vectors.
Sanger sequencing results of vectors expressing sgRNAs targeting the VEGFA (**a**), VEGFR1 (**b**) and VEGFR2 (**c**) promoters. Screen shots showing SnapGene views with the scheme of the sgRNA expression cassettes (colored arrows) underneath the vector sequence (bold font). Below an exemplary sequencing result is shown in normal font.

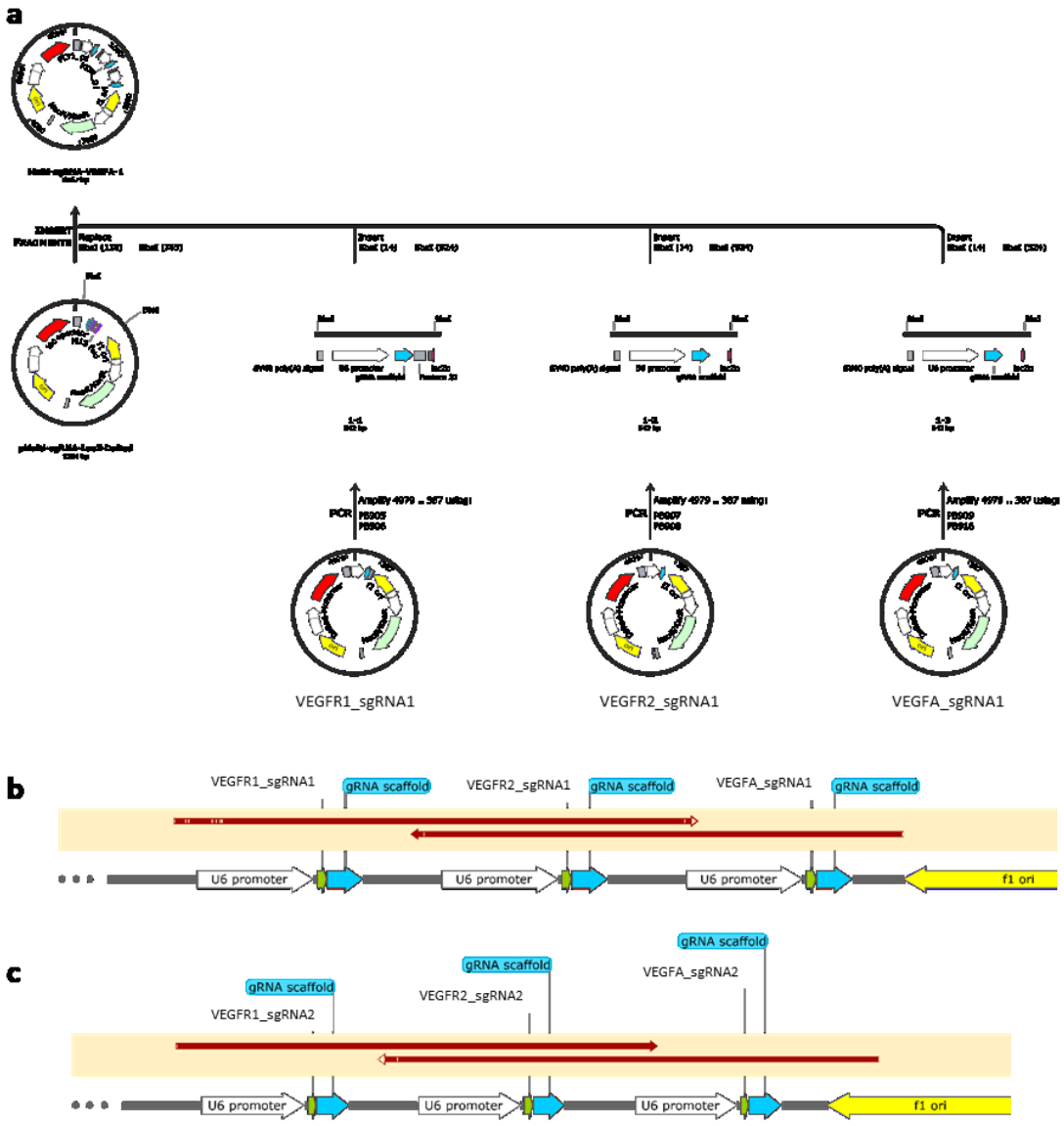

Figure 14. Cloning of the multiple sgRNAs expression vectors.

a Cloning strategy used to generate the multiple sgRNAs expression vectors. This example illustrates the generation of the multi-sgRNA1 vector containing the VEGFR1_sgRNA1, VEGFR2_sgRNA1 and VEGFA_sgRNA1. Sanger sequencing results of the multi-sgRNA1 (**b**) and multi-sgRNA2 (**c**) vectors are shown using SnapGene. Alignments of two sequencing reads per plasmid are shown as red arrows.

3.2.2 Simultaneous methylation of three genes of the VEGFA pathway

First, basal levels of DNA methylation at the three target regions in the HEK293 cell line were analysed by bisulfite sequencing. The data showed that the level of methylation at each of the three loci was less than 1.2 % (Figure 15, blue data sets). Targeted DNA methylation of multiple genes was conducted using the dCas9S, α3A3L and multi-sgRNA1 vectors and DNA methylation was analysed

three days after treatment. As shown in Figure 15 (orange data sets), all three designed sgRNA1 were able to recruit the dCas9S/α3A3L complex to their respective target genes and increase DNA methylation in all tested regions. Every region was characterized by a unique DNA methylation pattern as observed previously, i.e. individual CpG sites had different methylation levels. In the VEGFA promoter the maximum methylation reached was 55 % at the CpG site #6 (Figure 15a), whereas CpG sites #13-18 showed hardly any methylation introduced by the EpiEditor. Based on the two biological replicates, the mean level of DNA methylation of the whole region delivered by the EpiEditor was about 19 %. DNA methylation of the VEGFR1 promoter was more efficient (Figure 15b), with methylation levels of individual CpG sites ranging from 22 to 92 %. Calculations of the average methylation levels of the whole analysed region revealed around 60 %, which was much higher compared to the levels achieved for the VEGFA promoter. Analysis of the VEGFR2 region (Figure 15c) revealed a maximum methylation level of 86 % at the CpG site #19 and 41 % for the overall region. Thus, using a vector encoding three sgRNAs targeting different genes allowed multiplexed methylation of corresponding targets. DNA methylation conducted using the multi-sgRNA2 vector, encoding the second set of sgRNAs, and dCas9S, α3A3L vectors were also performed. Bisulfite sequencing revealed that all three sgRNAs were able to recruit the EpiEditor to the target regions (Figure 15, grey data sets). Methylation profiles and average levels of the VEGFA and VEGFR1 promoters were similar to the sgRNAs set #1 and reached about 19 % and 52 %, respectively (Figure 15a,b; grey data sets). Use of the sgRNA2 lead to a different methylation profile at the VEGFR2 promoter compared to the sgRNA1 and a higher methylation level of 61 % (Figure 15c; grey data set). Thus, both designed multi-sgRNA expressing vectors were efficient and the goal of setting DNA methylation at multiple targets was achieved.

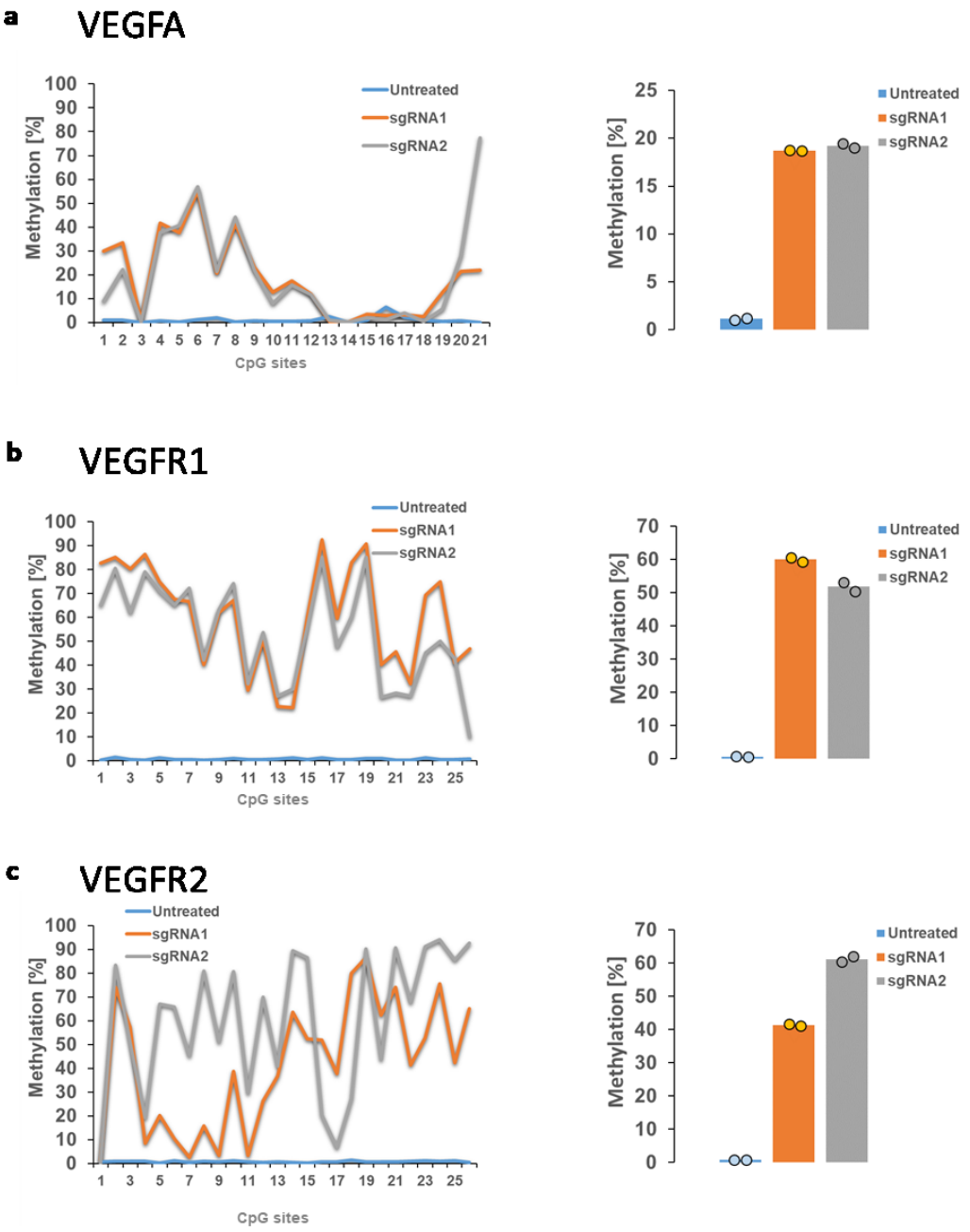

Figure 15. Methylation of multiple genes of the VEGFA pathway.

Results of simultaneous methylation of the VEGFA (**a**), VEGFR1 (**b**) and VEGFR2 (**c**) promoters. Line charts show methylation of individual CpG sites in corresponding regions of one biological sample analysed by bisulfite NGS. Bar diagrams depict mean DNA methylation level of the whole region calculated from two biological samples, individual values are shown as well. Methylation of untreated HEK293 cells (blue data set), and cells 3 days after transfection with the dCas9S/α3A3L and multi-sgRNA1 (orange data set) or multi-sgRNA2 (grey data set) are shown.

3.3 Development of approaches for targeted DNA methylation with higher specificity than the published ZFP-based EpiEditors

One further aim of the study was to increase the specificity of targeted methylation compared to the ZFP-based strategy used in the previous study [178]. To achieve this goal, the ZFP was replaced by the dCas9/sgRNA, which binds to a 20 bp target region and should thus reduce off-target effects originating from the DBD. Additionally, we and another group had shown that in the dCas9 system off-target DNA methylation comes from unspecific binding of the α3A3L to DNA and not from off-target binding of the dCas9/sgRNA complex [171], [174]. In that study several mutants of DNMT3ACD within the α3A3L protein had been designed aiming to reduce DNA interaction and their activity and specificity had been characterized [171]. The variant of the α3A3L with R887E mutation demonstrated higher specificity and slightly reduced activity compared to the wild type protein. This variant was used in the current study for the targeted methylation of the VEGFA pathway genes.

3.3.1 Comparison of the occurrence of the dCas9 and ZFP binding sites in the human genome

The specificity of targeted DNA methylation depends on many properties, one of which is the length of a recognition sequence of the DNA-binding domain. The longer the recognition sequence, the less frequently it occurs in the human genome, leading to a lower number of off-target sites. To compare the specificity of the ZFP and the dCas9/sgRNA targeting the VEGFA promoter, a program was written using Python programming language to count the occurrence of the recognition sequences in the human genome. Firstly, the occurrence of the ZFP target sequence and the VEGFA sgRNA binding sequence (from Figure 9) in the human genome assembly hg19 was calculated. As one can see, the short ZFP recognition sequence (9 nt) was found 12,980 times and the nineteen nucleotides long sgRNA binding sequence was found only once (Table 7).

Table 7. Target sequences of the ZFP and sgRNA targeting VEGFA and their occurrence in human genome.

Sequence	Perfect match, counts	One mismatch allowed, counts
GGGGGTGAC	12,890	404,477
GGCGGTCACCCCCAAAAGC	1	2

Assuming that binding specificity is not absolute and proteins can bind to sequences containing a single nucleotide mismatch, this would give more binding sites. A modified version of the program was written and the occurrences of sequences with these parameters was analysed. As one can see in Table 7, more than 400,000 degenerate sites were found for the ZFP protein but only one additional binding site appeared for the sgRNA. An additional requirement for the dCas9/sgRNA complex binding to the target site is the presence of a PAM sequence 3' to the sgRNA binding site. The existence of the NGG PAM sequence of *S. pyogenes* dCas9 used in the study was checked. None of the four NGG sequences was found at the binding sites with one nucleotide mismatch. This means, that there is a unique binding site in the human genome for sgRNA targeting the VEGFA promoter, if only one mismatch is allowed, confirming that the dCas9 is a more specific DNA-binding module than this particular ZFP protein.

3.3.2 Multiplex methylation of the VEGFA pathway genes using a more specific mutated EpiEditor

The experiment was conducted with the more specific variant α3A3L-R887E using the same conditions as for the wild type EpiEditor described above to allow direct comparison of these two effector domains. Both sgRNAs were able to recruit dCas9S/α3A3L-R887E to the VEGFA region (Figure 16a). The DNA methylation profile per CpG site looked quite similar to the one obtained with the wild type variant, but overall a lower level of methylation was reached. A maximum DNA methylation of 48 % was detected at the CpG site #21 and some methylation delivered at CpG sites # 1-8 (Figure 16a, left panel). No editing was observed at the CpG sites # 10-19. The mean DNA methylation of the targeted region was about 7 and 6 % for sgRNA1 and sgRNA2, respectively, which is

about 3 times lower than for the wild type construct (Figure 16a, right panel). Similarly, the dCas9S/α3A3L-R887E variant deposited methylation at the VEGFR1 and VEGFR2 promoters targeted by either sgRNA1 or sgRNA2. The maximum methylation of the VEGFR1 region was achieved at the CpG site #1, which was 78 % (Figure 16b, left panel) and thus nearly as high as observed for the wild type α3A3L (83 %, Figure 15b, left panel). But overall methylation of the VEGFR1 promoter dropped significantly to about 30 and 14 % for sgRNA1 and sgRNA2, respectively, i.e. 2.0 and 3.8 times lower than the values obtained with the wild type Ab-3a3l construct (Figure 16b, right panel). The highest levels of methylation of the VEGFR2 promoters were detected at the 3' end of the analysed region, where for 3 CpG sites (#21, 24 and 26) more than 75 % methylation was achieved if α3A3L-R887E was used in combination with the sgRNA2 (Figure 16c, left panel). By comparison, the same CpG sites were methylated by the wild type α3A3L to about 90 % (Figure 15c, left panel). Average methylation of the VEGFR2 region also showed a more than 2-fold reduction compared to the wild type Ab-3a3l and was 15 and 27 % for the sgRNA1 and sgRNA2, respectively (Figure 16c, right panel). Thus, the more specific mutant α3A3L-R887E was able to simultaneously set methylation at multiple target genes, but with lower efficiency.

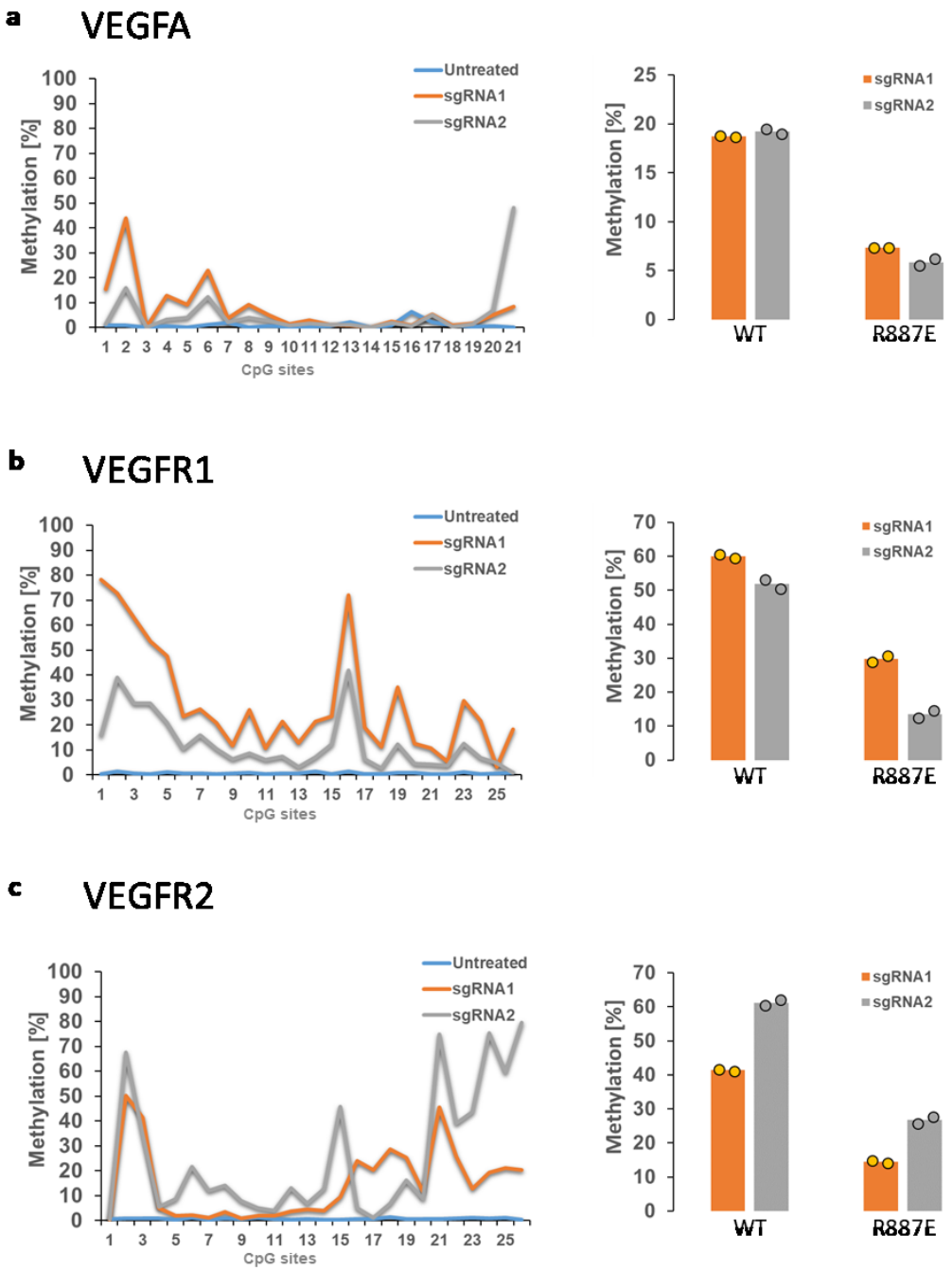

Figure 16. Simultaneous methylation of the VEGFA pathway genes with the enhanced specificity EpiEditor variant.
Methylation of the VEGFA (**a**), VEGFR1 (**b**) and VEGFR2 (**c**) promoters. Line charts show methylation of individual CpG sites in corresponding regions taken from one biological replicated analysed by bisulfite NGS. Methylation of untreated HEK293 cells (blue data set, taken from Figure 15 for comparison), and cells 3 days after transfection with the dCas9S/α3A3L and multi-sgRNA1 (orange data set) or multi-sgRNA2 (grey data set) are shown. Bar diagrams depict mean DNA methylation levels of the whole region calculated from two biological replicates using the enhanced specificity EpiEditor (R887E) and the wild type variant (WT, data taken from Figure 15 for comparison); values of individual replicates are shown as dots.

3.3.3 Comparison of off-target editing activity of the wild type and R887E EpiEditors

As mentioned above, the R887E variant had been shown to have reduced off-target DNA methylation in comparison to the wild type α3A3L [171]. To validate this in the current study, DNA methylation analysis of one off-target genomic region was conducted. For this purpose the promoter of the ISG15 gene was selected, since it had been shown to have a permissive chromatin and was efficiently methylated by EpiEditors [171]. gDNA from the experiments described in chapters 3.2.2 and 3.3.2 was analysed and data showed that the R887E variant had approximately 5-fold lower off-target methylation compared to the wild type α3A3L (Figure 17).

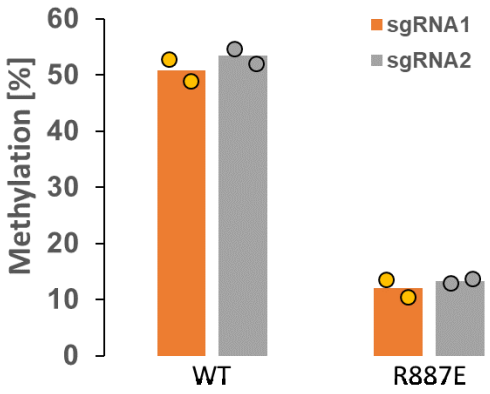

Figure 17. Off-target DNA methylation of the wild type and R887E variants of the EpiEditor. Off-target DNA methylation introduced by the wild type (WT) and R887E variants of the EpiEditor was analysed at the ISG15 promotor region. The bar diagram shows average DNA methylation level of the sequenced region based on two biological replicates, which are shown as dots. Data from experiments conducted using the multi-sgRNA1 (sgRNA1) and multi-sgRNA2 (sgRNA2) vectors are shown.

3.3.4 Multiplex methylation of three genes using double sgRNA targeting

The experiment described in paragraph 3.3.2 showed that the α3A3L-R887E variant can be used for multiplex methylation of multiple genes, but levels of achieved methylation were lower compared to the original wild type construct. To increase DNA methylation of targets with the R887E variant, an additional approach was tested. Since two sets of sgRNAs were designed for every promoter, use of both of them to recruit two dCas9S to one target region should lead to a higher number of α3A3L-R887E molecules bound to the locus. In effect,

this oligomerisation of α3A3L-R887E along the DNA may result in higher methylation. Since sgRNA binding sites were selected in a way that PAM sites face each other inversely, this setting may lead to preferential recruitment of α3A3L-R887E between dCas9S molecules, which could also facilitate methylation of the target region. Firstly, all six sgRNAs prepared earlier were cloned into one multiple sgRNA expression vector using the same protocol (Figure 18a). The multi-sgRNA3 plasmid was sequenced to confirm correct assembly of the vector (Figure 18b).

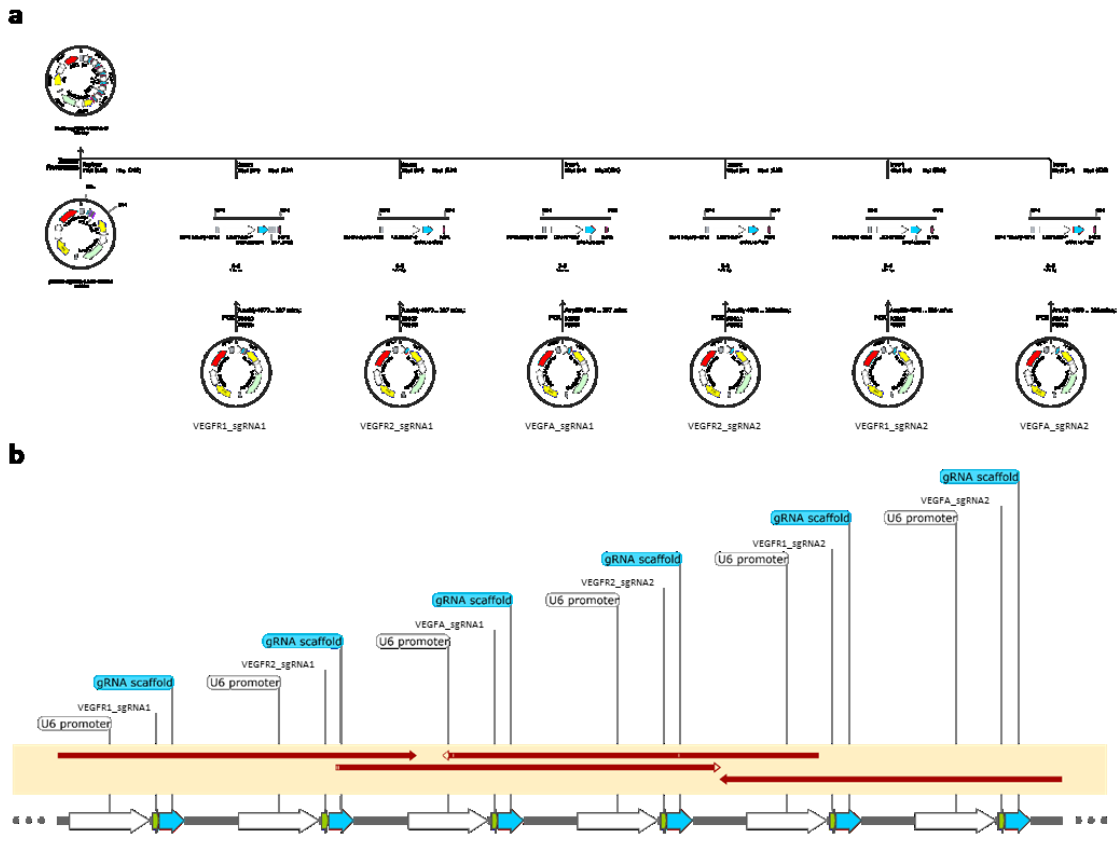

Figure 18. Cloning of six sgRNAs into one multiple sgRNA expression vector.
a Cloning scheme to generate six sgRNA expression vector using Golden Gate Assembly protocol. Produced using SnapGene. **b** Alignment of Sanger sequencing results performed to confirm a sequence of the obtained multi-sgRNA3 plasmid. Image generated using SnapGene.

DNA methylation of the VEGFA, VEGFR1 and VEGFR2 promoters was conducted using dCas9S, α3A3L-R887E and multi-sgRNA3 vectors. Appearances of *de novo* DNA methylation was analysed following the standard procedure. Results are shown in Figure 19 (green data set). As one can see, use of the multi-sgRNA3 was successful and lead to the deposition of DNA

methylation at all three target regions. Interestingly, the levels of DNA methylation at single CpG cites were at least as high as obtained with the single sgRNA protocol (data sets of multi-sgRNA1 and multi-sgRNA2 shown in orange and grey are taken from Figure 16 to simplify a direct comparison with the multi-sgRNA3). Collectively, the average methylations of the different target regions were higher than those achieved by use of single sgRNAs although differences were in some cases modest (Figure 19, right panels).

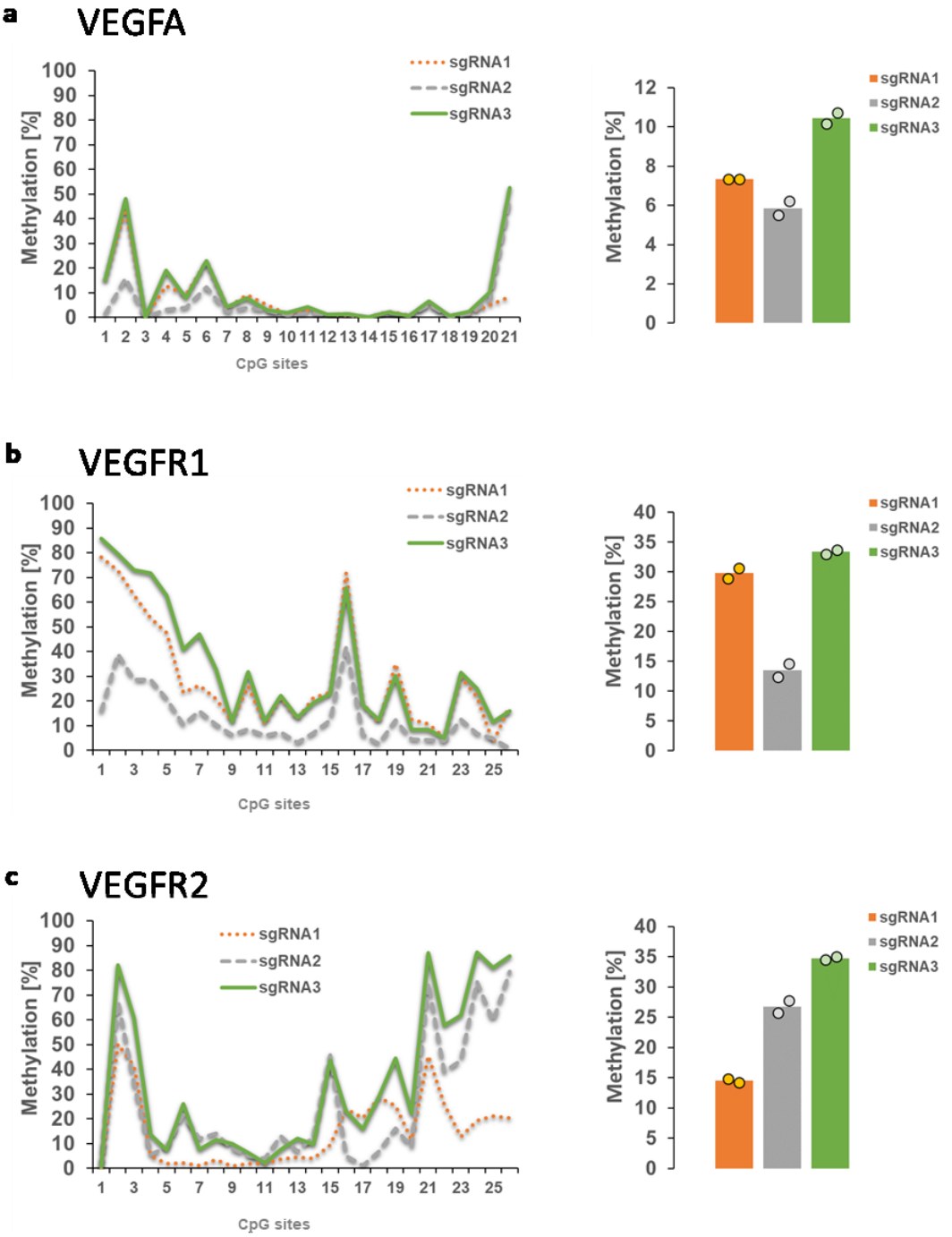

Figure 19. Methylation of the VEGFA pathway genes using the enhanced specificity EpiEditor and two sgRNAs per target.
Results of simultaneous methylation of the VEGFA (**a**), VEGFR1 (**b**) and VEGFR2 (**c**) promoters. Methylation data obtained with the dCas9S/α3A3L-R887E variant and the multi-sgRNA3 vector expressing two sgRNAs per gene are presented by the green data set (sgRNA3). Methylation achieved with the multi-sgRNA1 (sgRNA1, orange data set) and multi-sgRNA2 (sgRNA2, grey data set) and dCas9S/α3A3L-R887E was taken from Figure 16 and shown to simplify a direct comparison with the multi-sgRNA3 vector. Line charts show methylation of individual CpG sites in corresponding regions (represent one biological replicate). Bar diagrams depict mean DNA methylation level of the whole regions calculated from two biological replicates, the individual data points are shown as dots.

It is known that the efficiency of epigenome editing depends on the expression levels of its components, in this case dCas9S, α3A3L and sgRNAs. The obtained results could be explained by lower expression levels of these components compared to the previous single sgRNA per gene experiment. Since a co-transfection of cells followed by the FACS enrichment of triple-positive population was performed in all experiments, the expression levels of fluorescent proteins was compared. Due to design of the vectors tagBFP and sfGFP expression levels directly correlate with the expression of dCas9S, α3A3L proteins respectively. Expression of DsRed, which is present of the multi-sgRNA plasmids does not directly link to expression of sgRNAs but reflects the amount of the corresponding plasmid in cells and may thus be used as an indirect indicator of sgRNA expression. As one can see in both experimental replicates, the multi-sgRNA3 cells showed the same expression level of tagBFP and even higher expression of sfGFP, compared to multi-sgRNA1 or multi-sgRNA2 cells (Figure 20). Expression of DsRed in multi-sgRNA3 cells has a similar level to the one of the replicates of multi-sgRNA1 cells, showing the lowest expression level. Thus all these data demonstrated that the benefit of using two sgRNAs per target cannot be explained by low expression of EpiEditors.

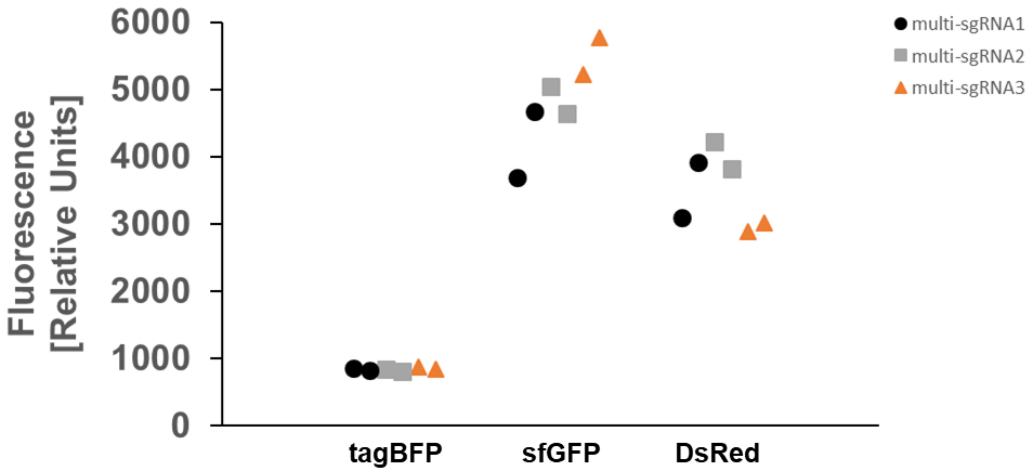

Figure 20. Comparison of EpiEditors expression levels.
Expression of three fluorescent proteins present on three plasmids encoding components of EpiEditor are plotted as relative fluorescent units. Expression of tagBFP and sfGFP is a direct measure of the dCas9S and α3A3L expression levels. Expression of DsRed correlates with the amount of sgRNA coding plasmid in cells. Every marker represents one of two biological experiment conducted with multi-sgRNA1, multi-sgRNA2 and multi-sgRNA3.

3.4 Analysis of patterns of targeted DNA methylation

3.4.1 Comparison of DNA methylation of both DNA strands

CpG sites have three states of methylation, unmethylated, hemimethylated, and fully methylated and all three forms exist in the human methylome. Methylation states are recognized by readers, for example, proteins containing a methyl-CpG-binding domain (MBD) for interaction with methylated CpG sites [186]. It is known that there are MBDs with preferential binding to fully methylated CpG sites [187]. Since some MBD-containing proteins are involved in maintenance of DNA methylation and silenced chromatin states, the status of methylation introduced by epigenome editing may affect the stability of DNA methylation and efficacy of target gene silencing. In the experiments described earlier, only one strand of the target regions was analysed. To get a complete understanding of the methylation state of CpG sites the complementary DNA strand has to be analysed too. Therefore, a second set of primers binding to the complementary strand of the target regions was designed and bisulfite converted DNA from experiments presented earlier was used for the library generation. Sequencing of samples α3A3L and α3A3L-R887E in combination with multi-sgRNA1 or multi-sgRNA2 was conducted. Methylation of the sense and antisense strands of the VEGFA and VEGFR1 promoters is shown in Figure 19 as bar diagram with overlaid datasets. Unfortunately, all attempts to amplify the antisense strand of the VEGFR2 were not successful and this data set had to be excluded from the analyses.

In case of the VEGFA region slightly different regions were analysed due to limitations in primer design for PCR on bisulfite-converted DNA. Therefore, the CpG site #1 was only analysed in the sense strand, and the CpG sites #22-24 were only analysed in the antisense strand. The amplicons designed for the VEGFR1 region covered identical CpG sites on both DNA strands (Figure 21). Overall, DNA methylation was obtained on both DNA strands no matter what Ab-3a3l variant or sgRNA was used. Correlation analysis of methylation levels on both DNA strands aiming to look at the prospective relationships between them

showed that methylation levels are very similar at the VEGFR1 region with Pearson's r values of 0.96-0.97. This suggests that most CpG sites are in a fully methylated state. By contrast, methylation levels of opposite strand CpG sites in the VEGFA region differed significantly and lower Pearson's r values 0.67-0.85 were obtained, which probably means that many CpG sites were hemimethylated and *de novo* methylation was set in the antisense strand.

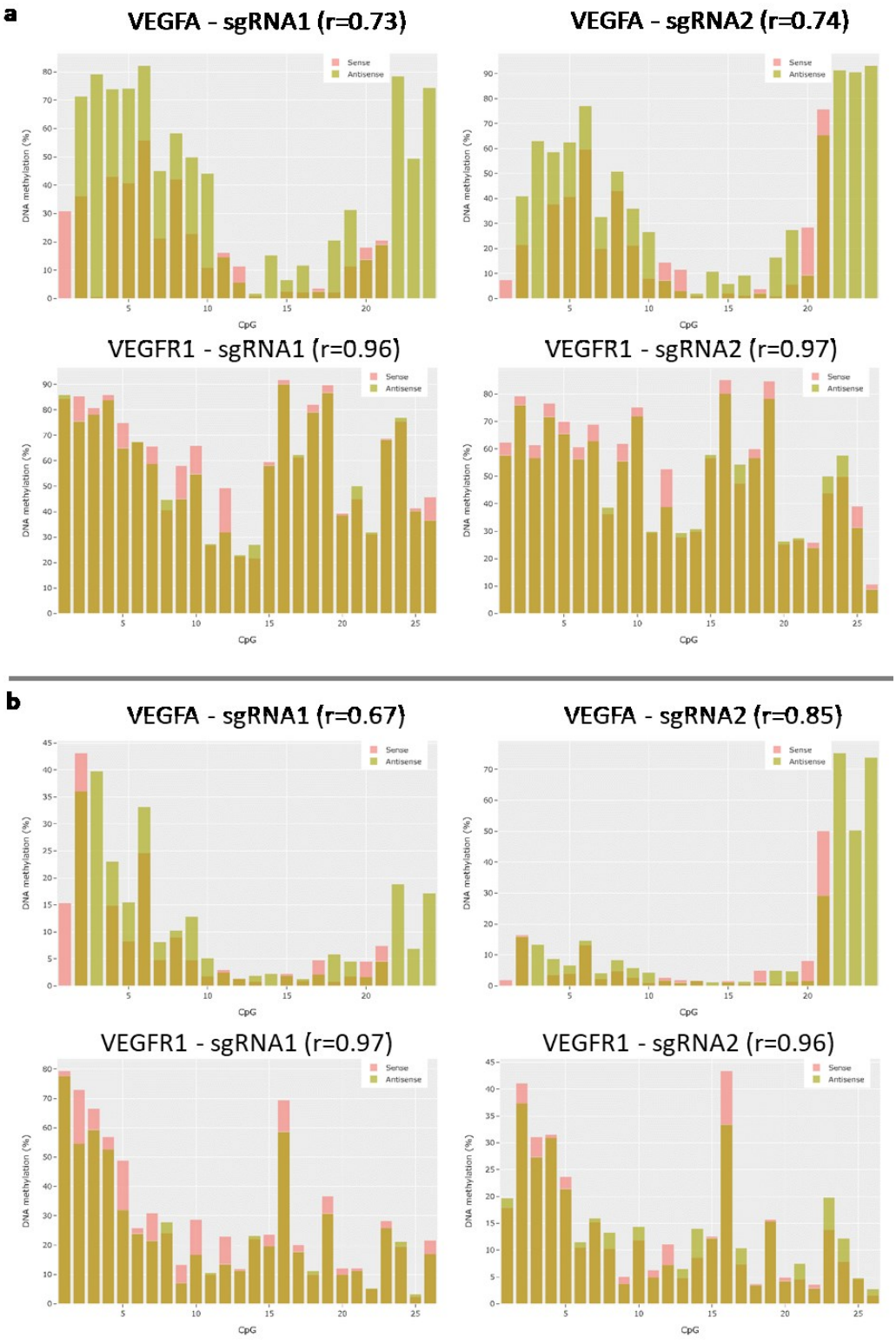

Figure 21. Comparison of DNA methylation introduced on the sense and antisense strands. Bar diagrams show methylation of individual CpG sites on the sense and antisense strands of the VEGFA and VEGFR1 promoters after targeted methylation using dCas9S/α3A3L (**a**) or dCas9S/α3A3L-R887E (**b**) variants and multi-sgRNA1 and multi-sgRNA2 vectors. The values are means of two biological replicates. Pearson correlation coefficients r of DNA methylation levels of CpG site on two strands are shown.

3.4.2 Dependence of DNA methylation efficiency on the distance from the PAM site

The dCas9/sgRNA complex binds to its target locus and recruits the α3A3L DNA methyltransferase, which will methylate CpG sites in close proximity. One can expect that there is an optimum distance from the sgRNA target site, where maximum methylation activity is possible and that levels of DNA methylation will decrease with distance. This knowledge if confirmed would provide rules for the design of DNA binding sites to achieve maximum efficiency of DNA methylation at smaller defined target loci, for example TF binding sites. To analyse this hypothesis, methylation of the individual CpG sites and their distance from the PAM sequence were derived from the available sequencing data. DNA methylation levels of two replicates were used to calculate average values, which were plotted against the distance from the PAM sequence (Figure 22). Data analysed after treatment of cells with the α3A3L-R887E variant guided by the multi-sgRNA1 or multi-sgRNA2 to the VEGFA and VEGFR1 promoters were combined in one plot and they are shown for the sense (upper left panel) and the antisense strand (upper right panel). As one can see, there is a very weak anti-correlation between the DNA methylation level and the distance from the PAM site. The highest *de novo* methylation was observed around 20 to 50 bp from the PAM site and it declined when moving away from the PAM site. Experiments conducted using two sgRNAs per target were excluded from the study since there are two complexes bound to the locus and they both introduce DNA methylation. Analogously, the same analysis was conducted for the wild type α3A3L variant, which showed no correlation between the efficiency of *de novo* DNA methylation and distance from the PAM sequence (Figure 22, low panels).

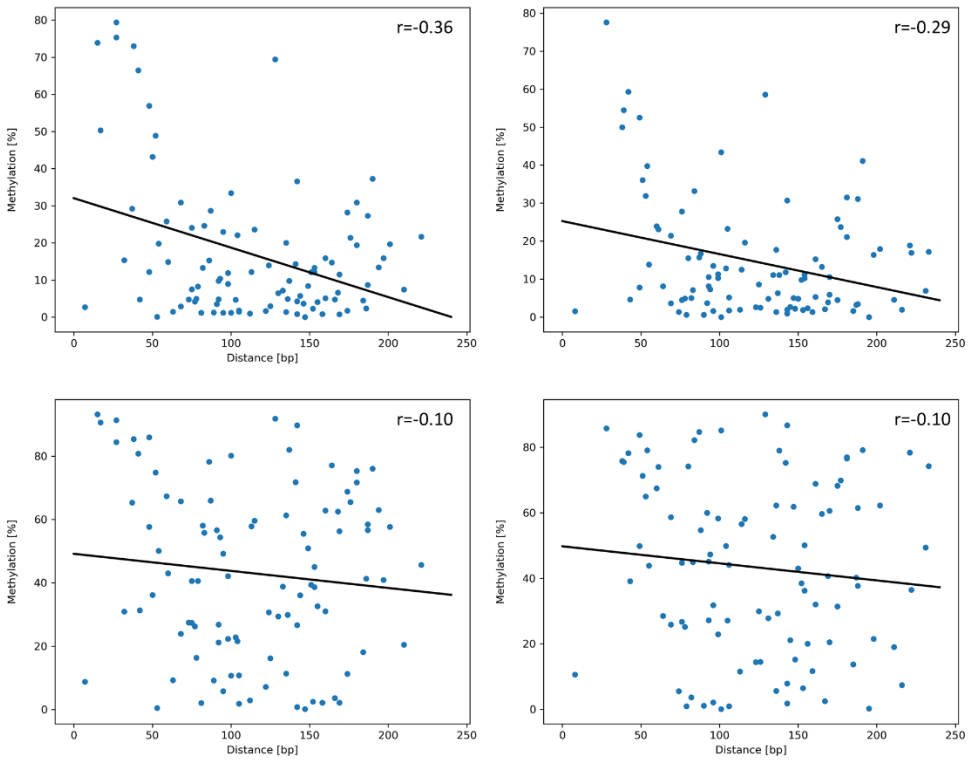

Figure 22. Dependence of DNA methylation efficiency on the distance from the PAM site. Scatter plots show methylation of individual CpG sites and their distance from the PAM sequence. A linear regression line and Pearson's r correlation coefficient are shown. The two upper plots show data for the α3A3L-R887E variant and DNA methylation of the sense strand (left panel) and antisense strand (right panel). The two low plots show data for the α3A3L wild type and DNA methylation of the sense and antisense strands (left and right panels correspondingly).

3.4.3 Comparison of the DNA methylation efficiency with the flanking sequence preferences of DNMT3A and DNMT1 DNA methyltransferases

Targeted methylation of three promoters conducted in the current study revealed the great diversity of methylation levels of individual CpG sites within the analysed loci. A possible explanation of this effect could be the flanking sequence preference of DNA methyltransferases. It has been shown that the activity of DNA methyltransferases DNMT3A, DNMT3B, and DNMT1 depends on the flanking sequence of CpG sites [188], [189]. This means that these enzymes contact nucleotides outside of CG dinucleotides and the nature of the flanking nucleotides makes a particular CpG site a preferred or disfavored substrate. The maximum effect on methylation activity was shown for flanking nucleotides at the positions from -3 to -1 and +1 to +3 relative to a CG site. In the current work, the catalytic domain of the DNMT3A methyltransferase was used to set DNA methylation, thus

it makes sense to compare levels of methylation of individual CpG sites with DNMT3A flanking preferences. DNA methylation patterns are maintained in cells by DNMT1, which has its own preferences of flanking sequences, so some *de novo* methylated CpG sites might be maintained better than others. Therefore, a comparison of obtained DNA methylation with the DNMT1 flanking preference had to be conducted too.

Flanking sequences of analysed CpG sites were extracted from target sequences that lead to generation of octamer sequences NNNCGNNN, where N are flanking nucleotides. Methylation levels for all CpG sites and their flanks were analysed for the sense and antisense strands and the wild type α3A3L and the α3A3L-R887E variants. Flanking sequence preference scores of the DNMT3A and DNMT1 were taken from published data [188], [189].

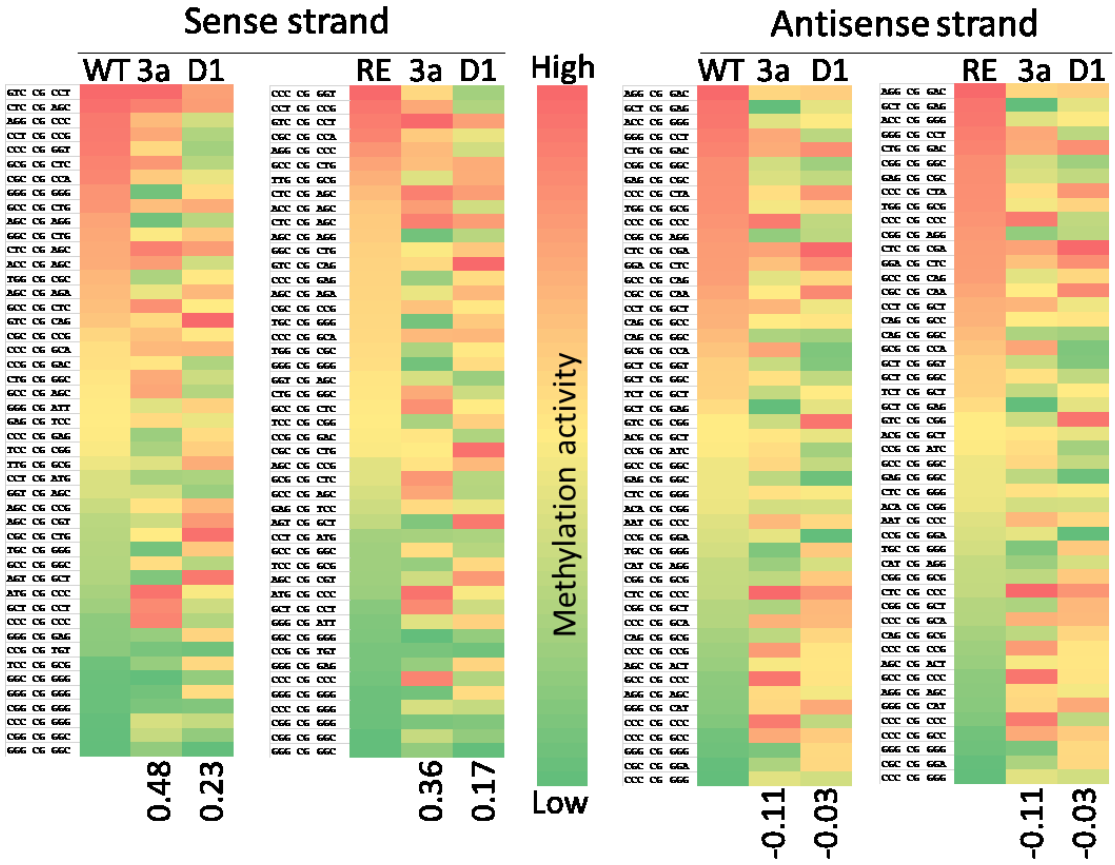

Figure 23. Comparison of DNA methylation levels of individual CpG sites with the flanking sequence preferences of DNMT3A and DNMT1.
Heatmaps show flanking sequences of individual CpG sites and their methylation preferences by the wild type (WT) or α3A3L-R887E (RE) variants. CpGs are sorted by methylation preferences

and compared with the methylation preference scores of DNMT3A (3a) and DNMT1 (D1). CpG sites of the sense and antisense strands were analysed separately. Pearson's correlation coefficients for pairs of data sets (WT or RE with 3a or D1) are shown.

Comparison of DNA methylation preferences versus published DNMT3a data revealed a moderate correlation between methylation introduced by the wild type α3A3L (WT) at the sense strand and flanking preference of DNMT3A (3a) with Pearson's coefficient of 0.48 (Figure 23). A somewhat weaker correlation was observed for the activity of the α3A3L-R887E (RE) variant on the sense strand and the flanking preference of DNMT3A (Pearson's coefficient 0.36). In contrast, none of the methylation patterns showed correlation with the flanking preference of DNMT1 (D1). Interestingly, the extent of methylation of CpG sites on the antisense strand was not correlated to either flanking preference of DNMT3A or DNMT1 (Figure 23).

4 Discussion

This project aimed to enhance editing efficiency and specificity at the VEGFA promoter by application of CRISPR/Cas9 technology instead of the previously used ZPF-technique. This allowed to i) increase efficiency of on-target DNA methylation, ii) to increase prospective antineoplastic efficacy by multiplex editing of the three genes VEGFA and its receptors VEGFR1 and VEGFR2, iii) to decrease off-target effects, and finally, iv) to conduct in-depth analysis of introduced DNA methylation patterns to be able to achieve more efficient epigenome editing in the future.

4.1 Methylation efficiency

The first aim was to increase the efficiency of DNA methylation at the VEGFA promoter. To approach this aim, the effector domain with the highest known DNA methyltransferase activity and the recruitment system able to bind several effector domains published recently [171] were used. The CRISPR/dCas9 construct with ten SunTag repeats, which are theoretically able to recruit up to ten DNMT3ACD-DNMT3LCD effector domains was targeted to the VEGFA promoter. As shown by targeted bisulfite sequencing, DNA methylation of up to 88 % was achieved at several CpG sites, resulting in an average methylation of the analysed region of 69 %. The two previous studies have reported that 43 and 49 % of methylation of the VEGFA promoter was introduced using single DNMT3ACD and DNMT3ACD-DNMT3LCD effector domains respectively [177], [178]. Although one cannot directly compare these results, since different cell lines HEK293 and SKOV3 cells were used, the data of this study showed that higher methylation levels of the VEGFA locus than the published ones are possible and the 10xSunTag was able to enhance on-target epigenome editing. Hence, the goal to increase on-target methylation efficiency of the VEGFA promoter was achieved by the combination of two strategies: the use of the highly active methyltransferase effector domain and the 10xSunTag system for activity amplification.

There are only a few studies that reported the implementation of the 10xSunTag system for epigenome editing to amplify on-target activity of DNA methylation. It had been used in combination with the DNMT3ACD-DNMT3LCD [171], the DNMT3ACD [173], and DNMT3A enzymes [172]. In all three cases, the efficiency of editing with and without the 10xSunTag was directly compared and only the latter study showed a significant increase of on-target methylation associated with the 10xSunTag system. Levels of DNA methylation reached in this study are quite high compared to results published by other groups, which typically reported a range from 5 to 50 % [153], [167], [173], [177], [178]. This can be explained by the higher efficiency of the applied system and by differences in experimental conditions such as the selected cell line, target gene, EpiEditor expression levels. While the latter is not possible to assess without additional experiments, the activity of various methyltransferases had been compared by several research groups. The DNMT3ACD-DNMT3LCD protein selected for this study outperforms the DNMT3ACD and the full-length DNMT3A [169][177][172] and its combination with the 10xSunTag had given a very promising results [171].

Several groups showed that the introduced DNA methylation is transient and is lost a few days after epigenome editing [177], [178]. One of the possible explanations of the transiency is that moderate levels of introduced DNA methylation are not maintained in cells. Genome-wide analysis of native methylomes shows generally bimodal distributions [72], [73]: many regions have low or high methylation level and very few an intermediate one, suggesting that high methylation levels may be more stable compared to intermediate levels (P. Bashtrykov, unpublished work). Thus the higher methylation efficiency achieved in the current study may help to overcome the problem of transiency. This hypothesis requires experimental validation and will be tested in the future.

4.2 Multiplex targeting

Targeted silencing of multiple “druggable” genes, for example, oncogenes, is a very attractive and successful approach to achieve a synergistic therapeutic effect. However, multiplex targeting of cancer-associated genes has a very strong economic burden, since providing drugs for several targets will increase final

treatment costs. Moreover, the approach also bears therapeutic and health risks due to the possibility of drug-drug interactions (DDI). In the case of targeted epigenome editing, the generation of additional sgRNAs required for targeting multiple genes and cloning them into a combined therapeutic vector could have minimal extra costs and DDIs are avoided. The unique nature of the CRISPR/Cas9 system, which is guided by sgRNAs, makes it an ideal tool to realize multiplex gene editing, since several sgRNAs can be delivered into cells to recruit the dCas9-based EpiEditor to diverse targets. By contrast, ZFP- and TALE-based DBDs would require generation of an array of proteins targeting individual loci. In this project, the second aim was to set DNA methylation at three genes, VEGFA and its receptors, which was successfully achieved by the use of vectors expressing multiple sgRNAs binding the targets and co-transfected with the EpiEditor. Two sgRNAs were designed and tested for each gene and resulting in average methylation levels of 19 %, 41 % and 60 % for VEGFA, VEGFR2 and VEGFR1, respectively. Interestingly, the level of VEGFA methylation was roughly 3-times lower compared to the downstream region targeted in the first experiment, which may be explained by different chromatin environments at these loci.

The multiplex targeting approach had been demonstrated in several previous studies. For example, epigenome editing at two promoters was achieved with several sgRNAs and the dCas9-M.*Sss*I mutant, but significant levels of methylation were seen only at maximum two CpG sites close to the sgRNA binding sites [190]. Another study published multiplex methylation of two genes by using a mix of two single sgRNA coding vectors [169]. Multiplex targeting of three loci, two CTCF binding sites and the UNC5C promoter, had been shown by use of dCas9-10xSunTag/DNMT3ACD [173], where methylation of three-four CGs at the CTCF sequences reached around 42 % and the promoter got an average 12.3 % methylation. The main difference of the current study from the mentioned ones is that several sgRNAs were cloned into one vector and not delivered as a mix of single sgRNA expression vectors. In the latter case, there is only a fraction of cells who gets all sgRNAs as a result of transfections. “The one plasmid-multiple sgRNAs” approach ensures that all cells that get the sgRNA

expressing vector get all sgRNAs, which should enhance multiplex epigenome editing efficiency.

Targeting multiple genes designed in this study aiming for amplification of therapeutic effects, had been also utilized to inhibit inflammatory signaling networks in dorsal root ganglion neurons which represent an in vitro model of the degenerative intervertebral disc induced back pain. In that study, three genes, interleukin-6, tumor necrosis factor receptor 1 and Interleukin 1 receptor type I, where silenced using dCas9-KRAB epigenome editing [191]. The neurons were transduced with three lentiviral vectors each of which encoded the EpiEditor plus one sgRNA targeting one gene. Comparing single versus multiplex targeting, the authors showed that only simultaneous silencing of three genes led to significant reduction of the redundant signaling pathways. Interestingly, a single gene targeting resulted in 90 % reduction of mRNA expression, whereas triple gene targeting caused only 75 % downregulation.

In conclusion, the current study realized for the first time multiplex epigenome editing of the three genes of the VEGFA pathway and applied the efficient method for the delivery of several sgRNAs into cells.

4.3 Specificity of epigenome editing

Specificity of therapeutics is one of the key parameters evaluated during drug development. In case of targeted DNA methylation aiming for gene silencing one has to carefully control off-target DNA methylation since aberrant epigenome editing may lead to modulation of expression of untargeted genes. Several studies including recent ones reported genome-wide off-target DNA methylation after use of EpiEditors [171], [173], [174], emphasizing the importance of controlling it and the need to develop more specific tools. As mentioned earlier, off-target DNA methylation originates from the interaction of DBDs and DNA methyltransferases with off-target loci. In this study two improvements relative to the preceding VEGFA DNA methylation reports [177], [178] were implemented addressing both issues.

Firstly, the ZFP used in that studies was replaced by dCas9, which increased the length of the target sequence recognized by the DBD from nine to twenty base pairs. Bioinformatics analysis conducted here showed that the selected sgRNA (the one targeting the same sequence as the ZFP) has a unique binding site in the human genome, whereas the ZFP recognition site has nearly thirteen thousand perfect matches. Motif search with one mismatch with the target sequence identified more than 400,000 sites for the ZFP and still only one unique sgRNA binding site. Of cause, this additional analysis is based on the assumption that the ZFP and the sgRNA tolerate all single nucleotide mismatches equally at all positions, which is not correct, but demonstrates that in general a long recognition sequence of 18-20 base pairs is sufficient to generate a DBD targeting a unique genomic locus.

Secondly, the off-target activity of the effector DNA methyltransferase domain can be minimized by decreasing its affinity to DNA. This approach had been implemented recently, where mutagenesis of arginines and lysines at the DNMT3ACD surface involved in DNA interaction was conducted [171]. Several single amino acid mutants were generated in that study and their on- and off-target activity was tested and compared to the wild type scFv-GCN4-DNMT3ACD-DNMT3LCD. The R887E variant showed a drastic reduction of off-target activity, namely 88 % less methylation than the wild type at a single off-target locus and a 7.8-fold decrease on genome-wide scale. This advanced EpiEditor was compared to the wild type in the current study for multiplex methylation of the VEGFA/VEGFRs axis genes with the multi-sgRNA expression plasmids. *De novo* DNA methylation was introduced at all target loci, although at lower levels compared to the wild type construct. For specificity analysis, off-target methylation of the wild type and R887E variants was compared at an untargeted locus (ISG15) that was chosen because it had been proved to be easily methylatable in our previous work and thus can serve as a sensitive reference [171]. Indeed, a 5-fold reduced methylation at the untargeted locus indicated increased specificity of epigenome editing, although extended analyses, such as genome-wide DNA methylation and whole transcriptome studies, may be warranted before therapeutic implementation.

Similar approaches used to decrease off-target editing of EpiEditors based on bacterial methyltransferases had demonstrated that mutations mitigating DNA binding affinity of the enzyme increase specificity but also lead to loss of on-target activity [192][190]. An alternative strategy to increase specificity proposed the development of split proteins, which had been realized in two variants based on the split M.*Sss*I methyltransferase and the ZFP [193] or dCas9 [194] as DBDs. In the first variant, two non-functional parts of M.*Sss*I methyltransferase were fused to two ZFPs targeted to the same genomic locus in a very close proximity; reconstitution of the functional M.*Sss*I resulted in deposition of DNA methylation between the ZFPs [193]. In the second variant, one part of the split M.*Sss*I was fused to dCas9 and a second freely diffused without DBD. Reconstitution of the functional methyltransferase at the dCas9-target locus led to a very specific methylation deposition, which was observed at 12 and 22-23 base pairs distance from the PAM site [194]. Both variants had improved specificity but had a very low on-target editing activity. The approach may be useful only if single CpG site methylation is needed, for example to interrogate specific TFs binding sites [185], [195] or insulators sensitive to DNA methylation as CTCF [196]. Several studies had reported attempts to increase on-target efficiency by using up to ten sgRNAs per target region. Most of them showed no or poor improvement [168], [169], [173], [190], [194] and only one documented a benefit of this strategy at only on of tested loci [167].

Thus the current study showed that the specificity of targeted DNA methylation can be improved but it is accompanied by reduced on-target editing, which appears to be a general phenomenon observed by many groups and in different systems. Further studies will be necessary to solve this problem.

4.4 Analysing patterns of targeted DNA methylation

To achieve high levels of DNA methylation at specific locations, one has to know patterns of *de novo* DNA methylation set by EpiEditors. For example, dCas9-based EpiEditors leave footprints of about 20 bps, these are the regions of the DNA occupied by dCas9/sgRNA complex and thereby protected from methylation [169], [190]. Direct fusions of dCas9 with DNMT3A showed high levels of editing

in the area close to the sgRNA binding site and it decreased further away from it [167], [168]. Similar information was not available for the dCas9-SunTag'ed EpiEditors in the only three studies published their use [171]–[173]. Therefore, the last aim of the current study was to analyse patterns of DNA methylation established by dCas9-10xSunTag/DNMT3ACD-DNMT3LCD EpiEditors used here by three parameters: methylation levels of both DNA strands, the dependency of methylation level from the distance from the PAM site and flanking sequence preferences of DNMTs. Such information may improve the design of future experiments and enhance the efficiency and stability of DNA methylation.

The vast majority (if not all) of studies published results of targeted DNA methylation analysed only one DNA strand, the obtained information does not allow to say whether CpG sites have a hemi- or fully methylated state. This knowledge is very important, since it may explain the transiency of introduced DNA methylation, because hemimethylated states, if not converted to fully methylated by endogenous DNMT1 enzyme, may be less efficiently maintained in cells than fully methylated ones. Moreover, methylation of DNA by the DNMT3ACD/DNMT3LCD heterotetramer generates two hemimethylated CpG sites spaced by 8-10 bps [197]. To understand patterns of DNA methylation introduced by the EpiEditors used in this study, both DNA strands were analysed. The obtained data showed a very strong correlation between methylation levels of two DNA strands of individual CpG sites. The technology utilized in this experiment, where two strands are amplified and sequenced independently, does not allow to conclude that CpG sites are in a fully methylated state, for such purpose a hairpin-ligation-based bisulfite sequencing should be conducted allowing simultaneous analysis of methylation at both strands of one DNA molecule [198], [199]. Nevertheless, one can speculate that it is very likely especially in the case of the VEGFR1 promoter for which Pearson's correlation coefficient r was around 0.96-0.97 that most CpG sites were fully methylated. In contrast, a significant portion of CpG sites in the VEGFA promoter was in the hemimethylated state. Thus, the methylation state of CpG sites differs between analysed regions and the hemimethylated CpG sites generated by the EpiEditor

during the first catalysis could be methylated further to the fully methylated state by the EpiEditor or endogenous DNMTs, most likely DNMT1.

The length of the region which can be effectively methylated is an important parameter to be considered during experiment design, therefore levels of DNA methylation of individual CpG sites obtained in this study were analysed taking into account their distance from the PAM site. The 10xSunTag utilized here is a 288 amino acids long tail, which theoretically should increase this length compared to the direct dCas9-DNMT3A fusions. Interestingly, the peak of *de novo* methylation was identified at a distance of 20-50 bp from the PAM sites on both DNA strands, as it had been shown for the dCas9 without the 10xSunTag [168], [169]. Levels of methylation decreased further away from this peak in case of the R887E DNMT3ACD-DNMT3LCD variant but stayed nearly the same for the wild type variant of the effector domain. The wild type version demonstrated a different pattern, which can be explained by several observations. Firstly, it had been shown that the wild type introduces intensive off-target methylation [171], which may be the reason of high methylation at longer distances from the PAM sites. Secondly, homotetramers of DNMT3A and heterotetramers of DNMT3A/DNMT3L can oligomerize along DNA, which boosts methylation spreading [197], [200], [201]. This phenomenon had been observed by Stepper and colleagues, who reported spreading of methylation up to 1200 bp from the PAM site using dCas9-DNMT3ACD-DNMT3LCD EpiEditor and a single sgRNA [169]. They showed that this effect was completely lost by a single mutation R832E in the DNMT3A catalytic domain, which disables oligomerization. The R887E mutation used in this study is not located on the multimerization surface [197] but its reduced affinity to DNA could impair cooperative multimerization. Thus, the current study showed that the 10xSunTag does not add extra flexibility and does not increase the length of a genomic region (compared to the dCas9-DNMT3A direct fusion EpiEditors), which can be efficiently methylated by the advanced effector domain.

Lastly, the flanking sequences of CpG sites of the target region may influence DNA methylation efficiency, since mammalian DNMTs have preferences to

certain flanks [188], [189], [202]. Flanking sequences of CpG sites in the targeted regions were analysed and compared with the published flanking sequence preferences of DNMT3A (its catalytic domain is used in the effector domain) and DNMT1 enzymes (the main enzyme maintaining DNA methylation in human cells) [188], [189]. Interestingly, there were some correlations found but only on the sense strand. CpG sites with flanks preferred by DNMT3A were slightly more methylated than the disfavored ones by the wild type DNMT3ACD-DNMT3LCD effector domain. Keeping in mind that there was no correlation found between the distance from the PAM site and methylation efficiency demonstrated by this EpiEditor, one may speculate that many distant CGs were methylated in an off-target mode and the level of methylation was determined mainly by the flanking sequence preference of DNMT3A. The R887E mutant of DNMT3ACD-DNMT3LCD showed even weaker correlation with the flanking preference of DNMT3A, either because the mutation may change the flanking preference or because the distance from the PAM site has more influence, since it showed more specific deposition of DNA methylation. Interestingly, there were no correlations found between the methylation profiles and the flanking preferences of the DNMT1 enzyme. One possible explanation of this is that the expression of the EpiEditor is very high at this early stage after transfection and its activity dominates the footprint of DNA methylation deposition, and only at later time points one may see a stronger impact of maintenance activity of DNMT1. Therefore, there is a weak correlation of DNA methylation patterns obtained during targeted epigenome editing and flanking sequence preferences of the methyltransferase, at least at the analysed time point.

4.5 Closing remarks

One of the key characteristics of epigenome-based gene regulation of cells is heritability, which means that transcriptional states, active or repressed genes, can be stably maintained by cells and even inherited after cell divisions. This specific property makes targeted epigenome editing an attractive approach for the development of therapeutics for various disorders including cancer. Unfortunately, the stable editing has not been realized yet, except for isolated

examples. The main goal of this study was to develop efficient and specific multiplex epigenome editing for transcriptional repression of the VEGFA axis genes, which are very important targets in anticancer therapy. The study led to the following accomplishments. Firstly, using a recently published combination of dCas9-10xSunTag with DNMT3ACD-DNMT3LCD methylation of the VEGFA promoter was achieved, exceeding all previously published results for this gene by 40 %. Secondly, multiplex methylation of promoters of VEGFA and its receptors VEGFR1 and VEGFR2 was established for the first time. The vectors encoding several sgRNAs targeting all genes were generated and used to assure the maximum efficiency of simultaneous editing, that had not been used for this purpose before. Thirdly, utilization of dCas9 and the R887E DNMT3ACD-DNMT3LCD variant with reduced off-target methylation increased the specificity of the VEGFA axis genes editing compared to previously published results. Fourthly, the in-depth analysis of the DNA methylation patterns generated by the EpiEditors revealed that a) CpG sites within target regions obtained hemi- and fully methylated states, and this varied depending on the genomic locus, b) the methylation peaked at 20-50 bp from the PAM site, c) methylation levels of individual CpG sites moderately correlated with the flanking sequence preference of the effector domain used for methylation. These characteristics will be very helpful in the design of future experiments.

5 Summary

Decades of efforts of clinicians and scientists to fight against cancer resulted in the development of multiple targeted therapies, which improved clinical outcomes for many types of tumors. Nevertheless, cancer is still the second leading cause of death worldwide, thus more efficient therapeutic approaches are urgently needed. Inhibitors of the VEGFA/VEGFRs axis have shown efficiency against various solid malignancies via reduction of neoangiogenesis of tumors and inhibition of autocrine stimulation of proliferation, migration and invasion of tumor cells by VEGFA. Monoclonal inhibitory antibodies against VEGFA or its receptors efficiently block signaling along the VEGFA/VEGFRs axis, but are very expensive and require repetitive drug injections since they function at the posttranslational level.

Targeted epigenome editing is a new emerging technology allowing to control the expression of selected genes at the transcriptional level. Regulation of gene expression is achieved via rewriting of chromatin marks at their cis-regulatory elements. For example, setting of DNA methylation at promoters may lead to stable silencing of corresponding genes. Epigenome editing can be achieved by EpiEditors, artificial chimeric proteins composed of a DNA-binding domain and a DNA methyltransferase, designed to set DNA methylation at target genomic loci. Previous studies demonstrated that methylation of the VEGFA promoter to approximately 50 % resulted in 70 % decrease of VEGFA gene expression in the ovarian cancer cell line SKOV3. This result was promising but only a moderate level of methylation was achieved. Additionally, off-target activity, epigenome editing at non-target genomic loci, has been reported in several studies, and has to be eliminated.

This project aimed to develop the technology further to

i) improve on-target editing efficiency at the VEGFA promoter to gain higher methylation level;
ii) establish multiplex editing to methylate promoters of VEGFA and its receptors VEGFR1, VEGFR2 for simultaneous silencing of all three genes;
iii) decrease off-target editing and

iv) analyse established DNA methylation patterns.

These goals were approached as follows:

i) The DNA genomic targeting technique of EpiEditors was changed from ZPF- to the CRISPR/Cas9 technology. This allowed to employ the recently published EpiEditor composed of the dCas9-10xSunTag protein and the anti-SunTag antibody fused with the highly active DNMT3ACD-DNMT3LCD chimeric methyltransferase. The SunTag allows signal amplification by recruiting up to 10 effector domains, which lead to 40 % higher DNA methylation of the VEGFA promoter compared to the EpiEditors using a single effector domain published previously.

ii) Multiplex methylation of the VEGFA, VEGFR1 and VEGFR2 promotors was established for the first time. Targeting of the dCas9-based EpiEditor to multiple loci was realized using vectors expressing several sgRNAs targeting these genes, which theoretically increases editing efficiency compared to co-transfection of vectors expressing single sgRNA used in previous reports.

iii) Implementation of dCas9/sgRNA instead of ZFP used previously for targeting the VEGFA promoter significantly improved editing specificity by reducing of off-target effects originating from the DBD. In addition, use of more specific mutant version R887E of EpiEditor showed 5-fold lower off-target activity at a single representative locus.

iv) An in-depth analysis of the introduced DNA methylation patterns revealed that the degree of methylation of individual CpG sites depends on several parameters, such as their distance from the sgRNA binding site and the flanking sequence preference of the effector domain. Furthermore, CpG sites can be in hemi- and fully methylated state and the predominance of one or the other state depends on the target locus.

The current study led to the development of multiplex epigenome editing of VEGFA and its receptors with the efficiency and specificity superior to previous reports and revealed insights into established DNA patterns which will enhance future design to achieve stable genes silencing and desired antineoplastic therapeutic effects.

Zusammenfassung

Jahrzehntelange Bemühungen von Klinikern und Wissenschaftlern zur Bekämpfung von Krebs führten zur Entwicklung mehrerer zielgerichteter Therapien, die die klinischen Ergebnisse für viele Arten von Tumoren verbesserten. Dennoch ist Krebs nach wie vor die zweithäufigste Todesursache weltweit, weshalb effizientere Therapieansätze dringend erforderlich sind. Inhibitoren der VEGFA / VEGFR-Achse haben Wirksamkeit gegen verschiedene solide maligne Erkrankungen durch Reduktion der Neoangiogenese von Tumoren und Hemmung der autokrinen Stimulation der Proliferation, Migration und Invasion von Tumorzellen durch VEGFA gezeigt. Monoklonale inhibitorische Antikörper gegen VEGFA oder seine Rezeptoren blockieren effizient die Signalübertragung entlang der VEGFA / VEGFR-Achse, sind jedoch sehr teuer und erfordern wiederholte Arzneimittelinjektionen, da sie auf posttranslationaler Ebene funktionieren.

Die gezielte Manipulation von Epigenomen ist eine neue Technologie, mit der die Expression ausgewählter Gene auf Transkriptionsebene gesteuert werden kann. Die Regulation der Genexpression wird durch Umschreiben der Chromatinmarkierungen an ihren cis-regulatorischen Elementen erreicht. Beispielsweise kann das Setzen der DNA-Methylierung an Promotoren zu einer stabilen Stummschaltung entsprechender Gene führen. Die Epigenom-Editierung kann durch EpiEditors erreicht werden, künstliche chimäre Proteine, die aus einer DNA-Bindungsdomäne und einer DNA-Methyltransferase bestehen und die DNA-Methylierung an genomischen Zielorten festlegen sollen. Frühere Studien zeigten, dass die Reduktion der Methylierung des VEGFA-Promotors auf ungefähr 50% zu einer 70% igen Abnahme der VEGFA-Genexpression in der Eierstockkrebs-Zelllinie SKOV3 führte. Dieses Ergebnis ist vielversprechend, es wurde jedoch nur ein mäßiger Methylierungsgrad erreicht. Darüber hinaus wurde in mehreren Studien über Off-Target-Aktivität, also Epigenom-Editing an nicht beabsichtigten Genomloci, berichtet, die es zu verhindern gilt.

Dieses Projekt zielt darauf ab, die Technologie weiterzuentwickeln durch

i) Verbesserung der zielgerichteten Bearbeitungseffizienz am VEGFA-Promotor, um einen höheren Methylierungsgrad zu erreichen;
ii) Etwicklung einer Multiplex-Technik zur Methylierung der Promotoren von VEGFA und seinen Rezeptoren VEGFR1 und VEGFR2 zur gleichzeitigen Stummschaltung aller drei Gene;
iii) Verringerung der unspezifischen Methylierung und
iv) Analyse induzierter DNA-Methylierungsmuster.

Diese Ziele wurden wie folgt angegangen:
i) Die DNA-Targeting Technik für EpiEditoren wurde von ZPF- auf CRISPR/Cas9-basierte Technologie umgestellt. Dies ermöglichte den Einsatz eines kürzlich veröffentlichte EpiEditors, der aus dem dCas9-10xSunTag-Protein und der an den Anti-SunTag-Antikörper fusionierten hochaktiven DNMT3ACD-DNMT3LCD-chimären Methyltransferase besteht. Der SunTag ermöglicht die Signalverstärkung durch Rekrutierung von bis zu 10 Effektordomänen, was zu einer 40% höheren DNA-Methylierung des VEGFA-Promotors im Vergleich zu den EpiEditors unter Verwendung einer zuvor veröffentlichten einzelnen Effektordomäne führt.
ii) Zum ersten Mal wurde eine Multiplex-Methylierung der VEGFA-, VEGFR1- und VEGFR2-Promotoren etabliert. Das Targeting des dCas9-basierten EpiEditors auf mehrere Loci wurde unter Verwendung von Vektoren realisiert, die mehrere sgRNAs exprimieren, die auf diese Gene abzielen, was theoretisch die Editiereffizienz im Vergleich zur bisher verwendeten Co-Transfektion von Vektoren für einzelne sgRNAs erhöht.
iii) Die Implementierung von dCas9 / sgRNA anstelle von ZFP, das zuvor für das Targeting des VEGFA-Promotors verwendet wurde, verbesserte die Editierspezifität signifikant, indem die von der DBD ausgehenden Off-Target-Effekte reduziert wurden. Zusätzlich zeigte die Verwendung einer spezifischeren mutierten Version R887E von EpiEditor eine 5-fach geringere Off-Target-Aktivität an einem einzelnen repräsentativen Ort.
iv) Eine eingehende Analyse der eingeführten DNA-Methylierungsmuster ergab, dass der Methylierungsgrad einzelner CpG-Stellen von mehreren Parametern

abhängt, wie z. B. ihrem Abstand von der sgRNA-Bindungsstelle und der Präferenz der flankierenden Sequenz der Effektordomäne. Darüber hinaus können sich CpG-Stellen in einem hemi- und vollständig methylierten Zustand befinden, und das Vorherrschen des einen oder anderen Zustands hängt vom Zielort ab.

Die aktuelle Studie führte zur Entwicklung einer Multiplex-Epigenom-Bearbeitung von VEGFA und seinen Rezeptoren mit einer Effizienz und Spezifität, die früheren Berichten überlegen ist. Auch enthüllte sie Einblicke in etablierte DNA-Muster, die helfen können, das zukünftige Design zu verbessern, um eine stabile Gen-Stummschaltung und gewünschte antineoplastische therapeutische Wirkungen zu erzielen.

6 List of figures

7 List of tables

8 References

[1] D. Hanahan and R. A. Weinberg, “The Hallmarks of Cancer,” *Cell*, vol. 100, no. 1, pp. 57–70, Jan. 2000, doi: 10.1016/S0092-8674(00)81683-9.

[2] D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: The next generation,” *Cell*, vol. 144, no. 5, pp. 646–674, 2011, doi: 10.1016/j.cell.2011.02.013.

[3] W. S. Bernard and P. W. Christopher, *World cancer report 2020*. 2020.

[4] K. R. Loeb and L. A. Loeb, “Significance of multiple mutations in cancer,” *Carcinogenesis*, vol. 21, no. 3, pp. 379–385, Mar. 2000, doi: 10.1093/carcin/21.3.379.

[5] M. R. Stratton, P. J. Campbell, and P. A. Futreal, “The cancer genome,” *Nature*, vol. 458, no. 7239, pp. 719–724, Apr. 2009, doi: 10.1038/nature07943.

[6] M. H. Bailey *et al.*, “Comprehensive Characterization of Cancer Driver Genes and Mutations,” *Cell*, vol. 173, no. 2, pp. 371-385.e18, 2018, doi: 10.1016/j.cell.2018.02.060.

[7] B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, and K. W. Kinzler, “Cancer genome landscapes,” *Science (80-.).*, vol. 340, no. 6127, pp. 1546–1558, 2013, doi: 10.1126/science.1235122.

[8] L. Falzone, S. Salomone, and M. Libra, “Evolution of cancer pharmacological treatments at the turn of the third millennium,” *Front. Pharmacol.*, vol. 9, no. NOV, 2018, doi: 10.3389/fphar.2018.01300.

[9] P. Carmeliet, “Angiogenesis in life, disease and medicine,” *Nature*, vol. 438, no. 7070, pp. 932–936, 2005, doi: 10.1038/nature04478.

[10] A. S. Chung and N. Ferrara, “Developmental and Pathological Angiogenesis,” 2011, doi: 10.1146/annurev-cellbio-092910-154002.

[11] N. Ferrara and W. J. Henzel, “Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells,” *Biochem. Biophys. Res. Commun.*, vol. 161, no. 2, pp. 851–858, Jun. 1989, doi: 10.1016/0006-291X(89)92678-8.

[12] N. Ferrara *et al.*, “Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene,” *Nature*, vol. 380, no. 6573, pp. 439–442, Apr. 1996, doi: 10.1038/380439a0.

[13] P. Carmeliet *et al.*, “Abnormal blood vessel development and lethality in embryos lacking a single vascular endothelial growth factor allele,” *Nature*, vol. 380, no. 1, pp. 435–439, 1996.

[14] N. Ferrara, “Vascular endothelial growth factor: Basic science and clinical progress,” *Endocr. Rev.*, vol. 25, no. 4, pp. 581–611, 2004, doi: 10.1210/er.2003-0027.

[15] B. Olofsson *et al.*, “Vascular endothelial growth factor B, a novel growth factor for endothelial cells,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 6,

pp. 2576–2581, 1996, doi: 10.1073/pnas.93.6.2576.

[16] V. Joukov *et al.*, “A novel vascular endothelial growth factor, VEGF-C, is a ligand for the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases,” *EMBO J.*, vol. 15, no. 2, pp. 290–298, 1996, doi: 10.1002/j.1460-2075.1996.tb00359.x.

[17] M. G. Achen *et al.*, “Vascular endothelial growth factor D (VEGF-D) is a ligand for the tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4),” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 2, pp. 548–553, 1998, doi: 10.1073/pnas.95.2.548.

[18] D. Maglione, V. Guerriero, G. Viglietto, P. Delli-Bovi, and M. G. Persico, “Isolation of a human placenta cDNA coding for a protein related to the vascular permeability factor,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 20, pp. 9267–9271, 1991, doi: 10.1073/pnas.88.20.9267.

[19] E. Tischer *et al.*, “The human gene for vascular endothelial growth factor: Multiple protein forms are encoded through alternative exon splicing,” *J. Biol. Chem.*, vol. 266, no. 18, pp. 11947–11954, 1991.

[20] K. A. Houck, N. Ferrara, J. Winer, G. Cachianes, B. Li, and D. W. Leung, “The vascular endothelial growth factor family: Identification of a fourth molecular species and characterization of alternative splicing of rna,” *Mol. Endocrinol.*, vol. 5, no. 12, pp. 1806–1814, 1991, doi: 10.1210/mend-5-12-1806.

[21] P. Vempati, A. S. Popel, and F. Mac, “Cytokine & Growth Factor Reviews Extracellular regulation of VEGF : Isoforms , proteolysis , and vascular patterning,” *Cytokine Growth Factor Rev.*, vol. 25, no. 1, pp. 1–19, 2014, doi: 10.1016/j.cytogfr.2013.11.002.

[22] L. Jingjing, Y. Xue, N. Agarwal, and R. S. Roque, “Human Muller cells express VEGF183, a novel spliced variant of vascular endothelial growth factor,” *Investig. Ophthalmol. Vis. Sci.*, vol. 40, no. 3, pp. 752–759, 1999.

[23] T. Lange, N. Guttmann-Raviv, L. Baruch, M. Machluf, and G. Neufeld, “VEGF162, a new heparin-binding vascular endothelial growth factor splice form that is expressed in transformed human cells,” *J. Biol. Chem.*, vol. 278, no. 19, pp. 17164–17169, 2003, doi: 10.1074/jbc.M212224200.

[24] D. O. Bates *et al.*, “VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, is down-regulated in renal cell carcinoma,” *Cancer Res.*, vol. 62, no. 14, pp. 4123–4131, 2002.

[25] H. Buteau-Lozano, M. Ancelin, B. Lardeux, J. Milanini, and M. Perrot-Applanat, “Transcriptional regulation of vascular endothelial growth factor by estradiol and tamoxifen in breast cancer cells: A complex interplay between estrogen receptors α and β,” *Cancer Res.*, vol. 62, no. 17, pp. 4977–4984, 2002.

[26] J. Wu, J. Richer, K. B. Horwitz, S. M. Hyder, and G. D’Andrilli, “Progestin-dependent induction of vascular endothelial growth factor in human breast cancer cells: Preferential regulation by progesterone receptor B,” *Women’s*

Oncol. Rev., vol. 4, no. 2, pp. 97–98, 2004, doi: 10.1080/14733400410001728500.

[27] M. Ryuto *et al.*, “Induction of vascular endothelial growth factor by tumor necrosis factor α in human glioma cells: Possible roles of SP-1,” *J. Biol. Chem.*, vol. 271, no. 45, pp. 28220–28228, 1996, doi: 10.1074/jbc.271.45.28220.

[28] J. Gille, R. A. Swerlick, and S. W. Caughman, “Transforming growth factor-α-induced transcriptional activation of the Vascular permeability factor (VPF/VEGF) gene requires AP-2-dependent DNA binding and transactivation,” *EMBO J.*, vol. 16, no. 4, pp. 750–759, 1997, doi: 10.1093/emboj/16.4.750.

[29] G. Finkenzeller, A. Sparacio, A. Technau, D. Marmé, and G. Siemeister, “Sp1 recognition sites in the proximal promoter of the human vascular endothelial growth factor gene are essential for platelet-derived growth factor-induced gene expression,” *Oncogene*, vol. 15, no. 6, pp. 669–676, 1997, doi: 10.1038/sj.onc.1201219.

[30] K. Reisinger, R. Kaufmann, and J. Gille, “Increased Sp1 phosphorylation as a mechanism of hepatocyte growth factor (HGF/SF)-induced vascular endothelial growth factor (VEGF/VPF) transcription,” *J. Cell Sci.*, vol. 116, no. 2, pp. 225–238, 2003, doi: 10.1242/jcs.00237.

[31] J. A. Forsythe *et al.*, “Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1.,” *Mol. Cell. Biol.*, vol. 16, no. 9, pp. 4604–4613, 1996, doi: 10.1128/mcb.16.9.4604.

[32] T. Tanaka *et al.*, “Induction of VEGF gene transcription by IL-1β is mediated through stress-activated MAP kinases and Sp1 sites in cardiac myocytes,” *J. Mol. Cell. Cardiol.*, vol. 32, no. 11, pp. 1955–1967, 2000, doi: 10.1006/jmcc.2000.1228.

[33] J. Milanini, F. Viñals, J. Pouysségur, and G. Pagès, “p42/p44 MAP kinase module plays a key role in the transcriptional regulation of the vascular endothelial growth factor gene in fibroblasts,” *J. Biol. Chem.*, vol. 273, no. 29, pp. 18165–18172, 1998, doi: 10.1074/jbc.273.29.18165.

[34] L. Yen *et al.*, “Differential regulation of tumor angiogenesis by distinct ErbB homo- and heterodimers,” *Mol. Biol. Cell*, vol. 13, no. 11, pp. 4029–4044, 2002, doi: 10.1091/mbc.E02-02-0084.

[35] G. Niu *et al.*, “Constitutive Stat3 activity up-regulates VEGF expression and tumor angiogenesis,” *Oncogene*, vol. 21, no. 13, pp. 2000–2008, 2002, doi: 10.1038/sj.onc.1205260.

[36] S. Lee, S. M. Jilani, G. V. Nikolova, D. Carpizo, and M. L. Iruela-Arispe, “Processing of VEGF-A by matrix metalloproteinases regulates bioavailability and vascular patterning in tumors,” *J. Cell Biol.*, vol. 169, no. 4, pp. 681–691, May 2005, doi: 10.1083/jcb.200409115.

[37] J. Plouët *et al.*, “Extracellular Cleavage of the Vascular Endothelial Growth Factor 189-Amino Acid Form by Urokinase Is Required for Its Mitogenic

Effect," *J. Biol. Chem.*, vol. 272, no. 20, pp. 13390–13396, May 1997, doi: 10.1074/jbc.272.20.13390.

[38] B. A. Keyt *et al.*, "The Carboxyl-terminal Domain(111165) of Vascular Endothelial Growth Factor Is Critical for Its Mitogenic Potency," *J. Biol. Chem.*, vol. 271, no. 13, pp. 7788–7795, Mar. 1996, doi: 10.1074/jbc.271.13.7788.

[39] P. Vempati, A. S. Popel, and F. Mac Gabhann, "Extracellular regulation of VEGF: Isoforms, proteolysis, and vascular patterning," *Cytokine Growth Factor Rev.*, vol. 25, no. 1, pp. 1–19, 2014, doi: 10.1016/j.cytogfr.2013.11.002.

[40] S. Koch, S. Tugues, X. Li, L. Gualandi, and L. Claesson-Welsh, "Signal transduction by vascular endothelial growth factor receptors," *Biochem. J.*, vol. 437, no. 2, pp. 169–183, 2011, doi: 10.1042/BJ20110301.

[41] M. Simons, E. Gordon, and L. Claesson-Welsh, "Mechanisms and regulation of endothelial VEGF receptor signalling," *Nature Reviews Molecular Cell Biology*, vol. 17, no. 10. Nature Publishing Group, pp. 611–625, 01-Oct-2016, doi: 10.1038/nrm.2016.87.

[42] G. H. Fong, J. Rossant, M. Gertsenstein, and M. L. Breitman, "Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium," *Nature*, vol. 376, no. 6535, pp. 66–70, 1995, doi: 10.1038/376066a0.

[43] F. Shalaby *et al.*, "Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice," *Nature*, vol. 376, no. 6535, pp. 62–66, Jul. 1995, doi: 10.1038/376062a0.

[44] H. Gerhardt *et al.*, "VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia," *J. Cell Biol.*, vol. 161, no. 6, pp. 1163–1177, 2003, doi: 10.1083/jcb.200302047.

[45] F. Mac Gabhann and A. S. Popel, "Model of competitive binding of vascular endothelial growth factor and placental growth factor to VEGF receptors on endothelial cells," *Am. J. Physiol. - Hear. Circ. Physiol.*, vol. 286, no. 1 55-1, 2004, doi: 10.1152/ajpheart.00254.2003.

[46] M. Autiero *et al.*, "Role of PlGF in the intra- and intermolecular cross talk between the VEGF receptors Flt1 and Flk1," *Nat. Med.*, vol. 9, no. 7, pp. 936–943, 2003, doi: 10.1038/nm884.

[47] J. Folkman, "Tumor Angiogenesis: Therapeutic Implications," *N. Engl. J. Med.*, vol. 285, no. 21, pp. 1182–1186, Nov. 1971, doi: 10.1056/NEJM197111182852108.

[48] J. C. Lee, N. H. Chow, S. T. Wang, and S. M. Huang, "Prognostic value of vascular endothelial growth factor expression in colorectal cancer patients," *Eur. J. Cancer*, vol. 36, no. 6, pp. 748–753, Apr. 2000, doi: 10.1016/S0959-8049(00)00003-4.

[49] A. Kaya *et al.*, "The prognostic significance of vascular endothelial growth factor levels in sera of non-small cell lung cancer patients," *Respir. Med.*,

vol. 98, no. 7, pp. 632–636, Jul. 2004, doi: 10.1016/j.rmed.2003.12.017.

[50] F. Révillion, J. Bonneterre, and J. P. Peyrat, “ERBB2 oncogene in human breast cancer and its clinical significance,” *European Journal of Cancer*, vol. 34, no. 6. pp. 791–808, 1998, doi: 10.1016/S0959-8049(97)10157-5.

[51] R. M. B. Loureiro, A. S. R. Maharaj, D. Dankort, W. J. Muller, and P. A. D’Amore, “ErbB2 overexpression in mammary cells upregulates VEGF through the core promoter,” *Biochem. Biophys. Res. Commun.*, vol. 326, no. 2, pp. 455–465, Jan. 2005, doi: 10.1016/j.bbrc.2004.11.053.

[52] F. Su *et al.*, “SP1 promotes tumor angiogenesis and invasion by activating VEGF expression in an acquired trastuzumab-resistant ovarian cancer model,” *Oncol. Rep.*, vol. 38, no. 5, pp. 2677–2684, Nov. 2017, doi: 10.3892/or.2017.5998.

[53] R. Lugano, M. Ramachandran, and A. Dimberg, “Tumor angiogenesis: causes, consequences, challenges and opportunities,” *Cell. Mol. Life Sci.*, vol. 77, no. 9, pp. 1745–1770, May 2020, doi: 10.1007/s00018-019-03351-7.

[54] S. Lee *et al.*, “Autocrine VEGF Signaling Is Required for Vascular Homeostasis,” *Cell*, vol. 130, no. 4, pp. 691–703, Aug. 2007, doi: 10.1016/j.cell.2007.06.054.

[55] M. Decaussin *et al.*, “Expression of vascular endothelial growth factor (VEGF) and its two receptors (VEGF-R1-Flt1 and VEGF-R2-Flk1/KDR) in non-small cell lung carcinomas (NSCLCs): correlation with angiogenesis and survival,” *J. Pathol.*, vol. 188, no. 4, pp. 369–377, Aug. 1999, doi: 10.1002/(SICI)1096-9896(199908)188:4<369::AID-PATH381>3.0.CO;2-X.

[56] K. Sato *et al.*, “Expression of Vascular Endothelial Growth Factor Gene and Its Receptor (flt-1) Gene in Urinary Bladder Cancer.,” *Tohoku J. Exp. Med.*, vol. 185, no. 3, pp. 173–184, 1998, doi: 10.1620/tjem.185.173.

[57] D. J. Price, T. Miralem, S. Jiang, R. Steinberg, and H. Avraham, “Role of vascular endothelial growth factor in the stimulation of cellular invasion and signaling of breast cancer cells.,” *Cell Growth Differ.*, vol. 12, no. 3, pp. 129–35, Mar. 2001.

[58] F. Fan *et al.*, “Expression and function of vascular endothelial growth factor receptor-1 on human colorectal cancer cells,” *Oncogene*, vol. 24, no. 16, pp. 2647–2653, Apr. 2005, doi: 10.1038/sj.onc.1208246.

[59] M. L. Slongo *et al.*, “Functional VEGF and VEGF receptors are expressed in human medulloblastomas,” *Neuro. Oncol.*, vol. 9, no. 4, pp. 384–392, Oct. 2007, doi: 10.1215/15228517-2007-032.

[60] Y. Wu *et al.*, “The vascular endothelial growth factor receptor (VEGFR-1) supports growth and survival of human breast carcinoma,” *Int. J. Cancer*, vol. 119, no. 7, pp. 1519–1529, Oct. 2006, doi: 10.1002/ijc.21865.

[61] L. Lian *et al.*, “VEGFR2 promotes tumorigenesis and metastasis in a pro-angiogenic-independent way in gastric cancer,” *BMC Cancer*, vol. 19, no.

1, p. 183, Dec. 2019, doi: 10.1186/s12885-019-5322-0.

[62] J. J. Parmar, M. Woringer, and C. Zimmer, “How the Genome Folds: The Biophysics of Four-Dimensional Chromatin Organization,” *Annu. Rev. Biophys.*, vol. 48, pp. 231–253, 2019, doi: 10.1146/annurev-biophys-052118-115638.

[63] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, “Crystal structure of the nucleosome core particle at 2.8 Å resolution,” *Nature*, vol. 389, no. 6648, pp. 251–260, 1997, doi: 10.1038/38444.

[64] A. J. Bannister and T. Kouzarides, “Regulation of chromatin by histone modifications,” *Cell Res.*, vol. 21, no. 3, pp. 381–395, 2011, doi: 10.1038/cr.2011.22.

[65] H. Holliday, L. A. Baker, S. R. Junankar, S. J. Clark, and A. Swarbrick, “Epigenomics of mammary gland development,” *Breast Cancer Res.*, vol. 20, no. 1, pp. 1–11, 2018, doi: 10.1186/s13058-018-1031-x.

[66] E. R. Gibney and C. M. Nolan, “Epigenetics and gene expression,” *Heredity (Edinb).*, vol. 105, no. 1, pp. 4–13, Jul. 2010, doi: 10.1038/hdy.2010.54.

[67] L. A. Gates, C. E. Foulds, and B. W. O’Malley, “Histone Marks in the ‘Driver’s Seat’: Functional Roles in Steering the Transcription Cycle,” *Trends Biochem. Sci.*, vol. 42, no. 12, pp. 977–989, 2017, doi: 10.1016/j.tibs.2017.10.004.

[68] K. Hyun, J. Jeon, K. Park, and J. Kim, “Writing, erasing and reading histone lysine methylations,” *Exp. Mol. Med.*, vol. 49, no. 4, 2017, doi: 10.1038/emm.2017.11.

[69] B. D. Strahl and C. D. Allis, “The language of covalent histone modifications,” *Nature*, vol. 403, no. 6765, pp. 41–45, 2000, doi: 10.1038/47412.

[70] N. Ahuja, H. Easwaran, and S. B. Baylin, “Harnessing the potential of epigenetic therapy to target solid tumors,” *J. Clin. Invest.*, vol. 124, no. 1, pp. 56–63, Jan. 2014, doi: 10.1172/JCI69736.

[71] M. Ehrlich *et al.*, “Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells,” *Nucleic Acids Res.*, vol. 10, no. 8, pp. 2709–2721, 1982, doi: 10.1093/nar/10.8.2709.

[72] M. Weber *et al.*, “Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome,” *Nat. Genet.*, vol. 39, no. 4, pp. 457–466, Apr. 2007, doi: 10.1038/ng1990.

[73] A. Meissner *et al.*, “Genome-scale DNA methylation maps of pluripotent and differentiated cells,” *Nature*, vol. 454, no. 7205, pp. 766–770, 2008, doi: 10.1038/nature07107.

[74] T. Bestor, A. Laudano, R. Mattaliano, and V. Ingram, “Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells,” *J. Mol. Biol.*, vol. 203, no. 4, pp. 971–983, Oct. 1988, doi: 10.1016/0022-2836(88)90122-2.

[75] M. Okano, S. Xie, and E. Li, “Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases,” *Nat. Genet.*, vol. 19, no. 3, pp. 219–220, Jul. 1998, doi: 10.1038/890.

[76] M. Okano, D. W. Bell, D. A. Haber, and E. Li, “DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development,” *Cell*, vol. 99, no. 3, pp. 247–257, Oct. 1999, doi: 10.1016/S0092-8674(00)81656-6.

[77] M. Kaneda *et al.*, “Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting,” *Nature*, vol. 429, no. 6994, pp. 900–903, 2004, doi: 10.1038/nature02633.

[78] U. Aapola *et al.*, “Isolation and Initial Characterization of a Novel Zinc Finger Gene, DNMT3L, on 21q22.3, Related to the Cytosine-5-Methyltransferase 3 Gene Family,” *Genomics*, vol. 65, no. 3, pp. 293–298, May 2000, doi: 10.1006/geno.2000.6168.

[79] D. Bourc’his, G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor, “Dnmt3L and the establishment of maternal genomic imprints,” *Science (80-.).*, vol. 294, no. 5551, pp. 2536–2539, Dec. 2001, doi: 10.1126/science.1065848.

[80] F. Chédin, M. R. Lieber, and C. L. Hsieh, “The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 26, pp. 16916–16921, 2002, doi: 10.1073/pnas.262443999.

[81] H. Gowher, K. Liebert, A. Hermann, G. Xu, and A. Jeltsch, “Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L,” *J. Biol. Chem.*, vol. 280, no. 14, pp. 13341–13348, 2005, doi: 10.1074/jbc.M413412200.

[82] D. Jia, R. Z. Jurkowska, X. Zhang, A. Jeltsch, and X. Cheng, “Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation,” *Nature*, vol. 449, no. 7159, pp. 248–251, Sep. 2007, doi: 10.1038/nature06146.

[83] R. Z. Jurkowska *et al.*, “Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: Heterochromatic localization and role of Dnmt3L,” *J. Biol. Chem.*, vol. 286, no. 27, pp. 24200–24207, 2011, doi: 10.1074/jbc.M111.254987.

[84] H. Leonhardt, A. W. Page, H.-U. Weier, and T. H. Bestor, “A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei,” *Cell*, vol. 71, no. 5, pp. 865–873, Nov. 1992, doi: 10.1016/0092-8674(92)90561-P.

[85] S. Pradhan, A. Bacolla, R. D. Wells, and R. J. Roberts, “Recombinant Human DNA (Cytosine-5) Methyltransferase,” *J. Biol. Chem.*, vol. 274, no. 46, pp. 33002–33010, Nov. 1999, doi: 10.1074/jbc.274.46.33002.

[86] M. Fatemi, A. Hermann, S. Pradhan, and A. Jeltsch, “The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an

allosteric activation of the enzyme after binding to methylated DNA," *J. Mol. Biol.*, vol. 309, no. 5, pp. 1189–1199, Jun. 2001, doi: 10.1006/jmbi.2001.4709.

[87] J. A. Law and S. E. Jacobsen, "Establishing, maintaining and modifying DNA methylation patterns in plants and animals," *Nature Reviews Genetics*, vol. 11, no. 3. Nature Publishing Group, pp. 204–220, 09-Mar-2010, doi: 10.1038/nrg2719.

[88] A. P. Bird, "CpG-rich islands and the function of DNA methylation," *Nature*, vol. 321, no. 6067, pp. 209–213, May 1986, doi: 10.1038/321209a0.

[89] H. Cedar, "DNA methylation and gene activity," *Cell*, vol. 53, no. 1. Cell Press, pp. 3–4, 08-Apr-1988, doi: 10.1016/0092-8674(88)90479-5.

[90] K. Opdecamp, M. Rivière, M. Molné, J. Szpirer, and C. Szpirer, "Methylation of an α-foetoprotein gene intragenic site modulates gene activity," *Nucleic Acids Res.*, vol. 20, no. 2, pp. 171–178, 1992, doi: 10.1093/nar/20.2.171.

[91] B. Panning and R. Jaenisch, "RNA and the Epigenetic Minireview Regulation of X Chromosome Inactivation (Figure 1). In invertebrates, X-linked numerator and au- tosomally encoded denominator genes encode proteins which differ in relative abundance in XX versus XY or XO cells. A 2-fold," *Cell*, vol. 93, no. Xic, pp. 305–308, 1998, doi: 10.1016/S0092-8674(00)81155-1.

[92] L. En, C. Beard, and R. Jaenisch, "Transcription of IAP endogenous retroviruses is constrained by cytosine methylation," *Nature*, vol. 366, no. october, pp. 362–365, 1993.

[93] T. H. Bestor and D. Bourc'his, "Transposon silencing and imprint establishment in mammalian germ cells," *Cold Spring Harb. Symp. Quant. Biol.*, vol. 69, pp. 381–387, 2004, doi: 10.1101/sqb.2004.69.381.

[94] A. C. Ferguson-Smith, H. Sasaki, B. M. Cattanach, and M. A. Surani, "Parental-origin-specific epigenetic modification of the mouse H19 gene," *Nature*, vol. 362, no. 6422, pp. 751–755, Apr. 1993, doi: 10.1038/362751a0.

[95] E. Li, C. Beard, and R. Jaenisch, "Role for DNA methylation in genomic imprinting," *Nature*, vol. 366, no. 6453, pp. 362–365, Dec. 1993, doi: 10.1038/366362a0.

[96] A. Hellman and A. Chess, "Gene body-specific methylation on the active X chromosome," *Science (80-.).*, vol. 315, no. 5815, pp. 1141–1143, 2007, doi: 10.1126/science.1136352.

[97] D. Aran, G. Toperoff, M. Rosenberg, and A. Hellman, "Replication timing-related and gene body-specific methylation of active human genes," *Hum. Mol. Genet.*, vol. 20, no. 4, pp. 670–680, 2011, doi: 10.1093/hmg/ddq513.

[98] F. Neri *et al.*, "Intragenic DNA methylation prevents spurious transcription initiation," *Nature*, vol. 543, no. 7643, pp. 72–77, Mar. 2017, doi: 10.1038/nature21373.

[99] D. Jjingo, A. B. Conley, S. V. Yi, V. V. Lunyak, and I. King Jordan, “On the presence and role of human gene-body DNA methylation,” *Oncotarget*, vol. 3, no. 4, pp. 462–474, 2012, doi: 10.18632/oncotarget.497.

[100] T. Vu and F. X. Claret, “Trastuzumab: Updated mechanisms of action and resistance in breast cancer,” *Front. Oncol.*, vol. 2 JUN, no. June, pp. 1–6, 2012, doi: 10.3389/fonc.2012.00062.

[101] Q. Jiao, L. Bi, Y. Ren, S. Song, Q. Wang, and Y. shan Wang, “Advances in studies of tyrosine kinase inhibitors and their acquired resistance,” *Mol. Cancer*, vol. 17, no. 1, pp. 1–12, 2018, doi: 10.1186/s12943-018-0801-5.

[102] C. M. Crews, “Inducing Protein Degradation as a Therapeutic Strategy,” *J. Med. Chem.*, vol. 61, no. 2, pp. 403–404, Jan. 2018, doi: 10.1021/acs.jmedchem.7b01333.

[103] B. Donati, E. Lorenzini, and A. Ciarrocchi, “BRD4 and Cancer: Going beyond transcriptional regulation,” *Mol. Cancer*, vol. 17, no. 1, pp. 1–13, 2018, doi: 10.1186/s12943-018-0915-9.

[104] R. L. Setten, J. J. Rossi, and S. ping Han, “The current state and future directions of RNAi-based therapeutics,” *Nat. Rev. Drug Discov.*, vol. 18, no. 6, pp. 421–446, 2019, doi: 10.1038/s41573-019-0017-4.

[105] M. L. De Groote, P. J. Verschure, and M. G. Rots, “Epigenetic Editing: Targeted rewriting of epigenetic marks to modulate expression of selected target genes,” *Nucleic Acids Res.*, vol. 40, no. 21, pp. 10596–10613, 2012, doi: 10.1093/nar/gks863.

[106] K. J. Kim *et al.*, “Inhibition of vascular endothelial growth factor-induced angiogenesis suppresses tumour growth in vivo,” *Nature*, vol. 362, no. 6423, pp. 841–844, Apr. 1993, doi: 10.1038/362841a0.

[107] L. G. Presta *et al.*, “Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders,” *Cancer Res.*, vol. 57, no. 20, pp. 4593–4599, 1997.

[108] J. C. Yang *et al.*, “A Randomized Trial of Bevacizumab, an Anti–Vascular Endothelial Growth Factor Antibody, for Metastatic Renal Cancer,” *N. Engl. J. Med.*, vol. 349, no. 5, pp. 427–434, Jul. 2003, doi: 10.1056/NEJMoa021491.

[109] F. Kabbinavar *et al.*, “Phase i I, randomized trial comparing bevacizumab plus fluorouracil (FU)/Leucovorin (LV) with FU/LV alone in patients with metastatic colorectal cancer,” *J. Clin. Oncol.*, vol. 21, no. 1, pp. 60–65, 2003, doi: 10.1200/JCO.2003.10.066.

[110] G. M. Keating, “Bevacizumab: A review of its use in advanced cancer,” *Drugs*, vol. 74, no. 16, pp. 1891–1925, 2014, doi: 10.1007/s40265-014-0302-9.

[111] S. K. Mukherji, “Bevacizumab (Avastin),” *Am. J. Neuroradiol.*, vol. 31, no. 2, pp. 235–236, Feb. 2010, doi: 10.3174/ajnr.A1987.

[112] P. M. Lorusso *et al.*, “Icrucumab, a fully human monoclonal antibody against the vascular endothelial growth factor receptor-1, in the treatment

of patients with advanced solid malignancies: A Phase 1 study," in *Investigational New Drugs*, 2014, vol. 32, no. 2, pp. 303–311, doi: 10.1007/s10637-013-9998-8.

[113] A. Landgraf Oholendt, PharmD, BCOP and J. L. Zadlo, PharmD, BCOP, "Ramucirumab: A New Therapy for Advanced Gastric Cancer," *J. Adv. Pract. Oncol.*, vol. 6, no. 1, Feb. 2015, doi: 10.6004/jadpro.2015.6.1.8.

[114] R. Wadhwa, T. Taketa, K. Sudo, M. Blum-Murphy, and J. A. Ajani, "Ramucirumab: A novel antiangiogenic agent," *Futur. Oncol.*, vol. 9, no. 6, pp. 789–795, Jun. 2013, doi: 10.2217/fon.13.68.

[115] E. B. Garon *et al.*, "Ramucirumab plus docetaxel versus placebo plus docetaxel for second-line treatment of stage IV non-small-cell lung cancer after disease progression on platinum-based therapy (REVEL): A multicentre, double-blind, randomised phase 3 trial," *Lancet*, vol. 384, no. 9944, pp. 665–673, 2014, doi: 10.1016/S0140-6736(14)60845-X.

[116] H. Vaucheret, C. Béclin, and M. Fagard, "Post-transcriptional gene silencing in plants," *J. Cell Sci.*, vol. 114, no. 17, pp. 3083–3091, 2001.

[117] C. Napoli, C. Lemieux, and R. Jorgensen, "Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans," *Plant Cell*, vol. 2, no. 4, pp. 279–289, 1990, doi: 10.2307/3869076.

[118] A. Fire, D. Albertson, S. W. Harrison, and D. G. Moerman, "Production of antisense RNA leads to effective and specific inhibition of gene expression in C. elegans muscle," *Development*, vol. 113, no. 2, pp. 503 LP – 514, Oct. 1991.

[119] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans," *Nature*, vol. 391, no. 6669, pp. 806–811, Feb. 1998, doi: 10.1038/35888.

[120] Y.-L. Chiu and T. M. Rana, "RNAi in Human Cells," *Mol. Cell*, vol. 10, no. 3, pp. 549–561, Sep. 2002, doi: 10.1016/S1097-2765(02)00652-4.

[121] M. T. McManus and P. A. Sharp, "Gene silencing in mammals by small interfering RNAs," *Nat. Rev. Genet.*, vol. 3, no. 10, pp. 737–747, Oct. 2002, doi: 10.1038/nrg908.

[122] S. M. Elbashir, "RNA interference is mediated by 21- and 22-nucleotide RNAs," *Genes Dev.*, vol. 15, no. 2, pp. 188–200, Jan. 2001, doi: 10.1101/gad.862301.

[123] R. F. Ketting, S. E. J. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, and R. H. A. Plasterk, "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans," *Genes Dev.*, vol. 15, no. 20, pp. 2654–2659, 2001, doi: 10.1101/gad.927801.

[124] E. Bernstein, A. A. Caudy, S. M. Hammond, and G. J. Hannon, "Role for a bidentate ribonuclease in the initiation step of RNA interference," *Nature*, vol. 409, no. 6818, pp. 363–366, Jan. 2001, doi: 10.1038/35053110.

[125] S. M. Hammond, E. Bernstein, D. Beach, and G. J. Hannon, “An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells,” *Nature*, vol. 404, no. 6775, pp. 293–296, Mar. 2000, doi: 10.1038/35005107.

[126] A. Nykänen, B. Haley, and P. D. Zamore, “ATP Requirements and Small Interfering RNA Structure in the RNA Interference Pathway,” *Cell*, vol. 107, no. 3, pp. 309–321, Nov. 2001, doi: 10.1016/S0092-8674(01)00547-5.

[127] J. J. Song, S. K. Smith, G. J. Hannon, and L. Joshua-Tor, “Crystal structure of argonaute and its implications for RISC slicer activity,” *Science (80-.).*, vol. 305, no. 5689, pp. 1434–1437, 2004, doi: 10.1126/science.1102514.

[128] J. Q. Yin, J. Gao, R. Shao, W.-N. Tian, J. Wang, and Y. Wan, “siRNA agents inhibit oncogene expression and attenuate human tumor cell growth,” *J. Exp. Ther. Oncol.*, vol. 3, no. 4, pp. 194–204, Jul. 2003, doi: 10.1046/j.1359-4117.2003.01092.x.

[129] K. E. Xiao, Z. Ouyang, and H. H. Tang, “Inhibiting the proliferation and metastasis of hilar cholangiocarcinoma cells by blocking the expression of vascular endothelial growth factor with small interfering rna,” *Oncol. Lett.*, vol. 16, no. 2, pp. 1841–1848, 2018, doi: 10.3892/ol.2018.8840.

[130] S. Filleur *et al.*, “SiRNA-mediated inhibition of vascular endothelial growth factor severely limits tumor resistance to antiangiogenic thrombospondin-1 and slows tumor vascularization and growth,” *Cancer Res.*, vol. 63, no. 14, pp. 3919–3922, 2003.

[131] Y. Takei, K. Kadomatsu, Y. Yuzawa, S. Matsuo, and T. Muramatsu, “A small interfering RNA targeting vascular endothelial growth factor as cancer therapeutics,” *Cancer Res.*, vol. 64, no. 10, pp. 3365–3370, 2004, doi: 10.1158/0008-5472.CAN-03-2682.

[132] R. M. Schiffelers *et al.*, “Cancer siRNA therapy by tumor selective delivery with ligand-targeted sterically stabilized nanoparticle.,” *Nucleic Acids Res.*, vol. 32, no. 19, pp. 1–10, 2004, doi: 10.1093/nar/gnh140.

[133] B. Kim *et al.*, “Inhibition of Ocular Angiogenesis by siRNA Targeting Vascular Endothelial Growth Factor Pathway Genes,” *Am. J. Pathol.*, vol. 165, no. 6, pp. 2177–2185, Dec. 2004, doi: 10.1016/S0002-9440(10)63267-1.

[134] G. Kungulovski and A. Jeltsch, “Epigenome Editing: State of the Art, Concepts, and Perspectives,” *Trends Genet.*, vol. 32, no. 2, pp. 101–113, Feb. 2016, doi: 10.1016/j.tig.2015.12.001.

[135] S. A. Wolfe, L. Nekludova, and C. O. Pabo, “DNA Recognition by Cys 2 His 2 Zinc Finger Proteins,” *Annu. Rev. Biophys. Biomol. Struct.*, vol. 29, no. 1, pp. 183–212, Jun. 2000, doi: 10.1146/annurev.biophys.29.1.183.

[136] Q. Liu, D. J. Segal, J. B. Ghiara, and C. F. Barbas, “Design of polydactyl zinc-finger proteins for unique addressing within complex genomes,” *Proc. Natl. Acad. Sci.*, vol. 94, no. 11, pp. 5525–5530, May 1997, doi: 10.1073/pnas.94.11.5525.

[137] K. Herbers, J. Conrads-Strauch, and U. Bonas, “Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein,” *Nature*, vol. 356, no. 6365, pp. 172–174, Mar. 1992, doi: 10.1038/356172a0.

[138] J. Boch *et al.*, “Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors,” *Science (80-.).*, vol. 326, no. 5959, pp. 1509–1512, Dec. 2009, doi: 10.1126/science.1178811.

[139] Y. Kim *et al.*, “A library of TAL effector nucleases spanning the human genome,” *Nat. Biotechnol.*, vol. 31, no. 3, pp. 251–258, Mar. 2013, doi: 10.1038/nbt.2517.

[140] P. I. Thakore, J. B. Black, I. B. Hilton, and C. A. Gersbach, “Editing the epigenome: technologies for programmable transcription and epigenetic modulation,” *Nature Methods*, vol. 13, no. 2. Nature Publishing Group, pp. 127–137, 01-Feb-2016, doi: 10.1038/nmeth.3733.

[141] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, “A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity,” *Science (80-.).*, vol. 337, no. 6096, pp. 816–821, Aug. 2012, doi: 10.1126/science.1225829.

[142] L. S. Qi *et al.*, “Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression,” *Cell*, vol. 152, no. 5, pp. 1173–1183, Feb. 2013, doi: 10.1016/j.cell.2013.02.022.

[143] L. Cong *et al.*, “Multiplex Genome Engineering Using CRISPR/Cas Systems,” *Science (80-.).*, vol. 339, no. 6121, pp. 819–823, Feb. 2013, doi: 10.1126/science.1231143.

[144] A. E. Briner *et al.*, “Guide RNA functional modules direct Cas9 activity and orthogonality,” *Mol. Cell*, vol. 56, no. 2, pp. 333–339, 2014, doi: 10.1016/j.molcel.2014.09.019.

[145] P. D. Hsu *et al.*, “DNA targeting specificity of RNA-guided Cas9 nucleases,” *Nat. Biotechnol.*, vol. 31, no. 9, pp. 827–832, Sep. 2013, doi: 10.1038/nbt.2647.

[146] F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, and C. Almendros, “Short motif sequences determine the targets of the prokaryotic CRISPR defence system,” *Microbiology*, vol. 155, no. 3, pp. 733–740, Mar. 2009, doi: 10.1099/mic.0.023960-0.

[147] L. A. Gilbert *et al.*, “Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation,” *Cell*, vol. 159, no. 3, pp. 647–661, Oct. 2014, doi: 10.1016/j.cell.2014.09.029.

[148] P. Mali *et al.*, “CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering,” *Nat. Biotechnol.*, vol. 31, no. 9, pp. 833–838, Sep. 2013, doi: 10.1038/nbt.2675.

[149] S. Konermann *et al.*, “Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex,” *Nature*, vol. 517, no. 7536, pp. 583–588, Jan. 2015, doi: 10.1038/nature14136.

[150] I. B. Hilton *et al.*, “Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers,” *Nat. Biotechnol.*, vol. 33, no. 5, pp. 510–517, 2015, doi: 10.1038/nbt.3199.

[151] S. R. Choudhury, Y. Cui, K. Lubecka, B. Stefanska, and J. Irudayaraj, “CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter,” *Oncotarget*, vol. 7, no. 29, pp. 46545–46556, Jul. 2016, doi: 10.18632/oncotarget.10234.

[152] S. Morita *et al.*, “Targeted DNA demethylation in vivo using dCas9–peptide repeat and scFv–TET1 catalytic domain fusions,” *Nat. Biotechnol.*, vol. 34, no. 10, pp. 1060–1065, Oct. 2016, doi: 10.1038/nbt.3658.

[153] X. S. Liu *et al.*, “Editing DNA Methylation in the Mammalian Genome,” *Cell*, vol. 167, no. 1, pp. 233-247.e17, Sep. 2016, doi: 10.1016/j.cell.2016.08.056.

[154] M. Gasperini *et al.*, “A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens,” *Cell*, vol. 176, no. 1–2, pp. 377-390.e19, Jan. 2019, doi: 10.1016/j.cell.2018.11.029.

[155] A. Amabile *et al.*, “Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing,” *Cell*, vol. 167, no. 1, pp. 219-232.e14, 2016, doi: 10.1016/j.cell.2016.09.006.

[156] H. O’Geen *et al.*, “Ezh2-dCas9 and KRAB-dCas9 enable engineering of epigenetic memory in a context-dependent manner,” *Epigenetics Chromatin*, vol. 12, no. 1, p. 26, Dec. 2019, doi: 10.1186/s13072-019-0275-8.

[157] H. S. Fukushima, H. Takeda, and R. Nakamura, “Targeted in vivo epigenome editing of H3K27me3,” *Epigenetics and Chromatin*, vol. 12, no. 1, pp. 1–12, 2019, doi: 10.1186/s13072-019-0263-z.

[158] G. L. Xu and T. H. Bestor, “Cytosine methylation targetted to pre-determined sequences,” *Nature Genetics*, vol. 17, no. 4. p. 378, 1997, doi: 10.1038/ng1297-376.

[159] C. D. Carvin, A. Dhasarathy, L. B. Friesenhahn, W. J. Jessen, and M. P. Kladde, “Targeted cytosine methylation for in vivo detection of protein-DNA interactions,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 13, pp. 7743–7748, 2003, doi: 10.1073/pnas.1332672100.

[160] C. D. Carvin, R. D. Parr, and M. P. Kladde, “Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins,” *Nucleic Acids Res.*, vol. 31, no. 22, pp. 6493–6501, 2003, doi: 10.1093/nar/gkg853.

[161] F. Li *et al.*, “Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes,” *Nucleic Acids Res.*, vol. 35, no. 1, pp. 100–112, Jan. 2007, doi: 10.1093/nar/gkl1035.

[162] A. E. Smith and K. G. Ford, “Specific targeting of cytosine methylation to DNA sequences in vivo,” *Nucleic Acids Res.*, vol. 35, no. 3, pp. 740–754, 2007, doi: 10.1093/nar/gkl1053.

[163] A. E. Smith, P. J. Hurd, A. J. Bannister, T. Kouzarides, and K. G. Ford, “Heritable gene repression through the action of a directed DNA methyltransferase at a chromosomal locus,” *J. Biol. Chem.*, vol. 283, no. 15, pp. 9878–9885, 2008, doi: 10.1074/jbc.M710393200.

[164] A. G. Rivenbark *et al.*, “Epigenetic reprogramming of cancer cells via targeted DNA methylation,” *Epigenetics*, vol. 7, no. 4, pp. 350–360, 2012, doi: 10.4161/epi.19507.

[165] D. L. Bernstein, J. E. Le Lay, E. G. Ruano, and K. H. Kaestner, “TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts,” *J. Clin. Invest.*, vol. 125, no. 5, pp. 1998–2006, May 2015, doi: 10.1172/JCI77321.

[166] C.-L. Lo, S. R. Choudhury, J. Irudayaraj, and F. C. Zhou, “Epigenetic Editing of Ascl1 Gene in Neural Stem Cells by Optogenetics,” *Sci. Rep.*, vol. 7, no. 1, p. 42047, Feb. 2017, doi: 10.1038/srep42047.

[167] A. Vojta *et al.*, “Repurposing the CRISPR-Cas9 system for targeted DNA methylation,” *Nucleic Acids Res.*, vol. 44, no. 12, pp. 5615–5628, Jul. 2016, doi: 10.1093/nar/gkw159.

[168] J. I. McDonald *et al.*, “Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation,” *Biol. Open*, vol. 5, no. 6, pp. 866–874, Jun. 2016, doi: 10.1242/bio.019067.

[169] P. Stepper *et al.*, “Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase,” *Nucleic Acids Res.*, vol. 45, no. 4, pp. 1703–1713, 2017, doi: 10.1093/nar/gkw1112.

[170] M. E. Tanenbaum, L. A. Gilbert, L. S. Qi, J. S. Weissman, and R. D. Vale, “A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging,” *Cell*, vol. 159, no. 3, pp. 635–646, Oct. 2014, doi: 10.1016/j.cell.2014.09.039.

[171] D. Hofacker, J. Broche, L. Laistner, S. Adam, P. Bashtrykov, and A. Jeltsch, “Engineering of effector domains for targeted DNA methylation with reduced off-target effects,” *Int. J. Mol. Sci.*, vol. 21, no. 2, 2020, doi: 10.3390/ijms21020502.

[172] Y.-H. Huang *et al.*, “DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A,” *Genome Biol.*, vol. 18, no. 1, p. 176, Dec. 2017, doi: 10.1186/s13059-017-1306-z.

[173] C. Pflueger *et al.*, “A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs,” *Genome Res.*, vol. 28, no. 8, pp. 1193–1206, Aug. 2018, doi: 10.1101/gr.233049.117.

[174] C. Galonska *et al.*, “Genome-wide tracking of dCas9-methyltransferase footprints,” *Nat. Commun.*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-017-02708-5.

[175] W. Nomura and C. F. Barbas, “In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase,” *J. Am. Chem. Soc.*, vol. 129,

no. 28, pp. 8676–8677, 2007, doi: 10.1021/ja0705588.

[176] P. Q. Liu *et al.*, “Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions: Activation of vascular endothelial growth factor A,” *J. Biol. Chem.*, vol. 276, no. 14, pp. 11323–11334, Apr. 2001, doi: 10.1074/jbc.M011172200.

[177] A. N. Siddique *et al.*, “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.*, vol. 425, no. 3, pp. 479–491, Feb. 2013, doi: 10.1016/j.jmb.2012.11.038.

[178] G. Kungulovski, S. Nunna, M. Thomas, U. M. Zanger, R. Reinhardt, and A. Jeltsch, “Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained,” *Epigenetics Chromatin*, vol. 8, no. 1, p. 12, Dec. 2015, doi: 10.1186/s13072-015-0002-z.

[179] P. Mali *et al.*, “RNA-guided human genome engineering via Cas9,” *Science (80-.).*, vol. 339, no. 6121, pp. 823–826, Feb. 2013, doi: 10.1126/science.1232033.

[180] S. Shao, L. Chang, Y. Sun, Y. Hou, X. Fan, and Y. Sun, “Multiplexed sgRNA Expression Allows Versatile Single Nonrepetitive DNA Labeling and Endogenous Gene Regulation,” *ACS Synth. Biol.*, vol. 7, no. 1, pp. 176–186, 2018, doi: 10.1021/acssynbio.7b00268.

[181] E. Afgan *et al.*, “The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update,” *Nucleic Acids Res.*, vol. 46, no. W1, pp. W537–W544, 2018, doi: 10.1093/nar/gky379.

[182] P. Bashtrykov and A. Jeltsch, “DNA methylation analysis by bisulfite conversion coupled to double multiplexed amplicon-based next-generation sequencing (NGS),” in *Methods in Molecular Biology*, vol. 1767, Humana Press Inc., 2018, pp. 367–382.

[183] J. Zhang, K. Kobert, T. Flouri, and A. Stamatakis, “PEAR: a fast and accurate Illumina Paired-End reAd mergeR,” *Bioinformatics*, vol. 30, no. 5, pp. 614–620, Mar. 2014, doi: 10.1093/bioinformatics/btt593.

[184] B. S. Pedersen, K. Eyring, S. De, I. V. Yang, and D. A. Schwartz, “Fast and accurate alignment of long bisulfite-seq reads,” vol. 00, no. 00, pp. 1–2, Jan. 2014.

[185] Y. Yin *et al.*, “Impact of cytosine methylation on DNA binding specificities of human transcription factors,” *Science (80-.).*, vol. 356, no. 6337, 2017, doi: 10.1126/science.aaj2239.

[186] I. Ohki *et al.*, “Solution Structure of the Methyl-CpG Binding Domain of Human MBD1 in Complex with Methylated DNA,” *Cell*, vol. 105, no. 4, pp. 487–497, May 2001, doi: 10.1016/S0092-8674(01)00324-5.

[187] A. Bellacosa *et al.*, “MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1,” *Proc. Natl. Acad. Sci.*, vol. 96, no. 7, pp. 3969–3974, Mar. 1999, doi: 10.1073/pnas.96.7.3969.

[188] L. Gao *et al.*, “Comprehensive structure-function characterization of DNMT3B and DNMT3A reveals distinctive de novo DNA methylation mechanisms,” *Nat. Commun.*, vol. 11, no. 1, p. 3355, Dec. 2020, doi: 10.1038/s41467-020-17109-4.

[189] S. Adam *et al.*, “DNA sequence-dependent activity and base flipping mechanisms of DNMT1 regulate genome-wide DNA methylation,” *Nat. Commun.*, vol. 11, no. 1, p. 3723, Dec. 2020, doi: 10.1038/s41467-020-17531-8.

[190] Y. Lei *et al.*, “Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein,” *Nat. Commun.*, vol. 8, no. 1, pp. 1–10, Jul. 2017, doi: 10.1038/ncomms16026.

[191] J. D. Stover, N. Farhang, B. Lawrence, and R. D. Bowles, “Multiplex Epigenome Editing of Dorsal Root Ganglion Neuron Receptors Abolishes Redundant Interleukin 6, Tumor Necrosis Factor Alpha, and Interleukin 1β Signaling by the Degenerative Intervertebral Disc,” *Hum. Gene Ther.*, vol. 30, no. 9, pp. 1147–1160, Sep. 2019, doi: 10.1089/hum.2019.032.

[192] G. E. Meister, S. Chandrasegaran, and M. Ostermeier, “An engineered split M.HhaI-zinc finger fusion lacks the intended methyltransferase specificity,” *Biochem. Biophys. Res. Commun.*, vol. 377, no. 1, pp. 226–230, Dec. 2008, doi: 10.1016/j.bbrc.2008.09.099.

[193] B. Chaikind and M. Ostermeier, “Directed Evolution of Improved Zinc Finger Methyltransferases,” *PLoS One*, vol. 9, no. 5, p. e96931, May 2014, doi: 10.1371/journal.pone.0096931.

[194] T. Xiong *et al.*, “Targeted DNA methylation in human cells using engineered dCas9-methyltransferases,” *Sci. Rep.*, vol. 7, no. 1, p. 6732, Dec. 2017, doi: 10.1038/s41598-017-06757-0.

[195] Z. Zuo, B. Roy, Y. K. Chang, D. Granas, and G. D. Stormo, “Measuring quantitative effects of methylation on transcription factor–DNA binding affinity,” *Sci. Adv.*, vol. 3, no. 11, p. eaao1799, Nov. 2017, doi: 10.1126/sciadv.aao1799.

[196] H. Wang *et al.*, “Widespread plasticity in CTCF occupancy linked to DNA methylation,” *Genome Res.*, vol. 22, no. 9, pp. 1680–1688, Sep. 2012, doi: 10.1101/gr.136101.111.

[197] A. Rajavelu, R. Z. Jurkowska, J. Fritz, and A. Jeltsch, “Function and disruption of DNA Methyltransferase 3a cooperative DNA binding and nucleoprotein filament formation,” *Nucleic Acids Res.*, vol. 40, no. 2, pp. 569–580, Jan. 2012, doi: 10.1093/nar/gkr753.

[198] C. D. Laird *et al.*, “Hairpin-bisulfite PCR: Assessing epigenetic methylation patterns on complementary strands of individual DNA molecules,” *Proc. Natl. Acad. Sci.*, vol. 101, no. 1, pp. 204–209, Jan. 2004, doi: 10.1073/pnas.2536758100.

[199] P. Giehr and J. Walter, “Hairpin bisulfite sequencing: Synchronous methylation analysis on complementary DNA strands of individual

chromosomes," in *Methods in Molecular Biology*, vol. 1708, 2018, pp. 573–586.

[200] M. Emperle, A. Rajavelu, R. Reinhardt, R. Z. Jurkowska, and A. Jeltsch, "Cooperative DNA Binding and Protein/DNA Fiber Formation Increases the Activity of the Dnmt3a DNA Methyltransferase," *J. Biol. Chem.*, vol. 289, no. 43, pp. 29602–29613, Oct. 2014, doi: 10.1074/jbc.M114.572032.

[201] R. Z. Jurkowska *et al.*, "Formation of nucleoprotein filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L," *Nucleic Acids Res.*, vol. 36, no. 21, pp. 6656–6663, Dec. 2008, doi: 10.1093/nar/gkn747.

[202] M. Dukatz, S. Adam, M. Biswal, J. Song, P. Bashtrykov, and A. Jeltsch, "Complex DNA sequence readout mechanisms of the DNMT3B DNA methyltransferase," *Nucleic Acids Res.*, no. 14, pp. 1–15, Oct. 2020, doi: 10.1093/nar/gkaa938.

9 Erklärung zum Eigenanteil der Dissertationsschrift

Die vorliegende Dissertation wurde am Dr. Margarete Fischer-Bosch Institut für Klinische Pharmakologie (Stuttgart) und in der Abteilung Biochemie des Instituts für Biochemie und Technische Biochemie an der Universität Stuttgart unter der Betreuung von Herrn Prof. Dr. Ulrich M. Zanger durchgeführt.

Alle Experimente und Datenanalysen wurden von mir durchgeführt.

Ich versichere, das Manuskript selbstständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

_____

Dr. rer. nat. Pavel Bashtrykov

10 Acknowledgements

First of all, I would like to thank my doctoral supervisor Prof. Dr. Ulrich M. Zanger for the agreement to mentor me during the studies, for teaching me scientific writing and data presentation, and for the comprehensive support.

I am thankful to Prof. Dr. Albert Jeltsch for giving me an opportunity to conduct these studies, for reading my manuscript, and for giving me lots of helpful critical comments.

Furthermore, I want to thank employees of the Dr. Margarete Fischer-Bosch Institute for Clinical Pharmacology, especially Dr. Maria Thomas, Dr. Kathrin Klein, for their great support of my experiments in the S2 lab.

I am very grateful to my colleagues, with whom we have been developing Epigenome editing tools, especially to Julian Broche, Daniel Hofacker, and Laura Laistner. It's a great pleasure to work with a team of young, smart, and creative scientists.

And last but not least, my deepest respect and words of gratitude to my wife Inna Bashtrykova for all these years of support, wisdom, love, patience, thanks to which I was able to go this way.