Epigenetic reprogramming of endogenous genes for permanent modulation of gene expression
Huisman, Christian

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
General introduction and Scope of the thesis
General introduction

The human body is made up of more than 200 different cell types and each cell type has its own specific function. The genetic information, encoded in the DNA, instructs the cells how to perform its function. This genetic information contains roughly 20700 genes (1), which code for the functional units of the cell; the proteins. The specific protein expression pattern displayed by a cell determines which biological processes are taken place. Tight regulation of these processes, including growth, differentiation and proliferation ensure normal cell functioning. Usually only a subset of the genes is transcribed to proteins in order for the cell to perform its function.

The number of genes that are actively transcribed in a cell is constantly changing, and dynamically regulated by a layer on top of the DNA called the ‘epigenetic layer’. Epigenetics is the additional layer of heritable information on top of the base pair sequence of the genome (2). The structural units of the epigenetic layer are the nucleosomes (3), which are composed of several types of histone proteins. Histone proteins are highly tunable through their protruding tails which can be posttranslational modified on many positions. The nucleosomes organize and package the DNA into the nucleus of a cell into highly ordered structures allowing for flexible modulation of gene expression patterns.

The epigenetic state of a gene is an important factor that determines the expression level of a gene (4). The two main components of epigenetic regulation (Figure 1) are DNA methylation; the addition of methyl group on the 5th position of the cytosine ring within CpG dinucleotides (CpGs) (5), and posttranslational histone modifications (4). A variety of histone modifications have been identified, including histone methylation (e.g. trimethylation of lysine 9 of histone 3) and histone (de)acetylation (e.g. (de)acetylation of histone 3). Methylation of DNA and histones is usually associated with low levels of gene expression, while unmethylated DNA and histone acetylation is usually associated with high levels of gene expression (4). However, there are several exceptions, for example the very well-studied histone mark trimethylation of lysine 4 of histone H3 is usually associated with high levels of transcription (4). Especially epigenetic modifications found in the promoter region of a gene (region of DNA that initiates transcription) determine if a gene is transcribed or not. In 40% of the genes, promoter regions contain CpG islands (DNA sequences with a high frequency of CpGs) (6). The DNA methylation levels of these CpG islands often determine if the gene is ‘off’ or ‘on’, while the histone modifications may play a role in determining the precise levels of gene expression.

Epigenetics in cancer
It becomes increasingly clear that many diseases are associated with epigenetic alterations causing aberrant gene expression (7). A common example is cancer.
Every year, more than 10 million people die from this disease, and the number of new cases is increasing rapidly. Therefore, cancer is one of the leading causes of death. Cancer is characterized by uncontrolled cell growth, as genes that stimulate cell proliferation (oncogenes) are overexpressed or gained function, while genes inhibiting these processes (tumor suppressor genes (TSGs)) are downregulated or dysfunctional. Although cancer has long been considered as a set of diseases caused by genetic abnormalities, new insights after intensive research justify the concept that cancer is an epigenetic disease as well. Epigenetic deregulation can lead to genomic instability and influences the gene expression pattern displayed by cells. The epigenetic dysregulation may even be the key that drives carcinogenesis (8). Besides that, most cancers harbor inactivating mutations in genes that control the epigenome, which further contribute to epigenetic deregulation and cancer progression (8). The most observed epigenetic abnormality in malignant cells is promoter DNA hypermethylation of CpG islands, which in turn are accompanied by the presence of repressive histone marks and a loss of activating histone marks, which further inhibit gene expression (9). For example, in various types of cancer, *Maspin, EPB41L3, CDKN2A* and *RASSF1A* are among the genes which are down-regulated by epigenetic mechanisms (10, 11) and are subject of investigation in the current thesis.

Cervical cancer is one of cancers which is extensively studied in this thesis, and is one of the top killers of women by cancer worldwide. In the last decade, genome-wide mapping of DNA methylation in normal- and cervical cancer
genomes revealed that a large number of genes in cervical cancer are associated with hypermethylated gene promoters (12), (11). Besides their functional implications of gene inactivation in tumor development, these methylation patterns can serve as diagnostic biomarkers for cervical cancer. Current screenings programs rely on the pap-smear assays which have poor sensitivity. Our group and others have reported several gene promoters that are specifically hypermethylated in cervical cancer. One of the hypermethylation markers we identified was *C13ORF18* (10), but we and others also found that *CCNA1*, *EPB41L3* and *TFPI2* methylation levels were strongly associated with cancer (11, 13-17). As promoter hypermethylation is often seen for TSGs, such potent diagnostic marker genes might exhibit tumor suppressive activities upon re-expression. Therefore, one of the goals in this thesis was to re-express these silenced marker genes and study their tumor suppressive activity in cervical cancer cell lines. Key TSGs for cervical cancer growth may subsequently be used as therapeutic target to decrease tumor growth in this malignancy.

It has already been shown that re-introduction of single genes into cancer cells where the gene is inactivated showed potent anti-tumor effects, as demonstrated for *CDKN2A* (encoding p16Ink4A) (18) or *TP53* (19). Contrary to genes that are inactivated through genetic mutations, epigenetic deregulation is a reversible process and can be targeted by epigenetic drugs (20). Current clinical drugs, such as the DNA methylation inhibitor Decitabin (5-aza-deoxycytosine) and the histone deacetylase inhibitor Vorinostat (suberoylanilide hydroxamic acid, SAHA), induce re-expression of a variety of hypermethylated silenced genes (21). So far, four epigenetic drugs have been FDA approved for subtypes of leukemias and lymphomas, and ongoing clinical trials in solid tumors indicate some beneficial effects for these patients as well (22). However, the disadvantages of these drugs include their genome-wide effects (20). In that respect it is important to note that epigenetic drugs have been shown to drive tumor promotion and metastasis by enhanced expression of pro-metastatic genes (23). Therefore, gene-targeted re-activation of TSGs is of great importance. A novel way to achieve gene-targeted re-activations is by the use of artificial transcription factors (ATFs) (24).

**Artificial transcription factors (ATFs)**
An interesting development in the last years for the gene-targeted re-expression of silenced genes or the downregulation of overexpressed genes is the use of ATFs (24). ATFs consist of a transcriptional regulatory domain coupled to an engineered DNA binding domain (Figure 2), such as a Zinc Finger Protein (ZFP), a transcription activator-like effector (TALE) protein or a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) associated 9 (Cas9) protein. Such DNA binding domains can be designed to target virtually any gene in the human genome (25, 26). While the TALE and CRISPR/CAS9 system are relative new targeting platforms, ZFP-mediated targeting has already been explored for more than a decade, and
this system is best characterized for its ability to target and modulate gene expression. ZFPs are natural occurring transcription factors containing several DNA binding domains ("fingers"). Each finger is around 30 amino acids and recognizes three base pairs. Individual fingers can be stitched together to increase their specificity. Typically, ZFPs are designed to target an unique DNA stretch of 18 base pairs within the genomic DNA. An amino acid- DNA binding code has been generated allowing ZFPs to target the desired triplet (27). The most common transcriptional regulatory domains are the gene-activator VP64 (a tetramer of the herpes simplex virus gene activator VP16) and the gene-repressor SKD.

**Zinc Finger Protein**: designed and engineered to target 18 bp in the chosen gene promoter (this thesis: EPCAM, ICAM-1, C13orf18, CCNA1, TFPI2, Maspin, EPB41L3, Her2 and PLOD2).

**Figure 2. Artificial transcription factors (ATFs).** ATFs are fusion proteins consisting of a DNA binding domain (DBD), e.g. an engineered Zinc Finger Protein (ZFP), fused to a transcriptional effector domain. The DBD is targeted to the gene of interest. The effector domain enforces the modulation of gene expression (VP64, SKD). The approach can be redesigned to directly modulate the epigenetic landscape (Epigenetic Editing) depending on which effector domain (e.g. Tet2, G9a, SUV39H1 and M.Sss1) is used.

(Kruppel associated box repressor domain). ATFs technology has proven itself as very useful tool for the modulation of gene expression in many disease models (24) and ZFP-targeting is currently the most advanced DNA targeting approach. This technology has the potency to activate gene expression, but depending on the transcriptional regulatory domain, can also down-regulate gene-expression (28). Compared to the current drugs which target only 2.5% of the proteins, the ability to target any gene using ATFs, including previously undruggable targets (29, 30), is a huge advantage. Traditional cDNA expression approaches have the disadvantage that only one splice variant is expressed, but ATFs enable expression from the endogenous gene locus allowing for all splice variants to be expressed in natural ratios. The importance of expressing alternative splice variants have been demonstrated for the VEGF gene (31) and was the basis for the first clinical trial based on ATFs. Furthermore, ATFs achieve efficient gene silencing by targeting the
DNA molecule directly, thereby overcoming the difficulties of RNA interference which needs to target continuously the newly produced mRNA molecules. The above mentioned advantages make ATFs suitable to study gene function or to modulate gene expression for therapeutic purposes. Interestingly, the ZFPs can also be attached to nucleases for site-specific DNA editing (changing of the genetic sequence at a specific site). Genome editing with engineered nucleases was declared ‘Method of the Year 2011’ by Nature Methods (32) and currently several clinical phase II ZF-nuclease trials are ongoing.

Epigenetic editing

The transcriptional effects of VP64 and SKD are traditionally considered transient in differentiated cells, as no epigenetic marks are actively rewritten, and genes likely return to the ‘normal’ state after the ATF is removed (33, 34). For ‘permanent’ modulation of gene expression, the epigenetic marks need to be rewritten, which may be achieved by a technique called Epigenetic Editing (35). Epigenetic editing is the gene-specific targeting of epigenetic enzymes to overwrite epigenetic signatures and to achieve long-term a) silencing of actively transcribed genes or b) re-expression of epigenetically silenced genes. Targeted epigenetic enzymes have the potency to change local DNA methylation or histone modifications at the site, which may result in transcriptional modulation. Epigenetic editing offers attractive prospects as the introduced epigenetic changes may achieve stable reprogramming, even after clearance of the drug (hit and run approach). So far, this exciting novel strategy has already resulted in the efficient transcriptional modulation of various genes targeting either histone modifications or DNA methylation. VEGF was the first gene modified by epigenetic editing; the catalytic domains of histone methyl transferases G9a and SUV39-H1 were targeted to the VEGF promoter, resulting in efficient downregulation by the induced repressive histone marks (36). The subsequent targeting of methyltransferase (DNA methylating enzyme) to the promoter regions of Sox2, Maspin and VEGF demonstrated that downregulation can also be achieved by targeted DNA methylation (37, 38). In this thesis, we further explored such targeting of histone modifying enzymes and DNA methylating enzymes for their repressive effects on gene expression.

To achieve long term re-expression of epigenetically silenced genes by means of epigenetic editing, epigenetically activating enzymes or enzymes that remove repressive marks, such as DNA methylation, can be targeted to silenced promoters. Recently, the ten eleven translocation (Tet) enzymes have been identified as actively DNA demethylators (39). Therefore, Tet proteins can be fused to ZFPs to achieve site-specific demethylation potentially resulting in gene-activation. For a long time, gene activation by targeted DNA demethylation seemed unfeasible in mammalian cells as no enzymes were identified with the capacity to actively demethylate DNA, although various biological processes
hinted to the existence of such mechanism, including the rapid demethylation of the paternal genome in the zygote hours after fertilisation (40). With the identification of the Tet-enzymes and other enzymes implicated in the DNA demethylating pathway, the feasibility of targeted DNA demethylation to activate gene expression is under exploration and receives widespread attention (41). Recently, the first studies appeared which achieved demethylation and upregulation of genes by targeting DNA demethylases to hypermethylated promoters (42, 43). Also in this thesis, we investigated the ability of Tet-enzymes (Tet2) to demethylate and activate hypermethylated genes.

Scope of the thesis
This thesis describes how ZFP-based ATF-technology can be applied to unravel a biological function of silenced genes currently arising from biomarker searches in cervical cancer and other malignancies. Furthermore, this thesis aims to provide more insights into causal effects of epigenetic marks on gene expression patterns (DNA methylation, histone modifications) and crosstalk in-between. Moreover, we explored the potencies of epigenetic editing to achieve stable expression modulation and we contribute to unravelling the underlying epigenetic rules for such epigenetic reprogramming. Current approaches for stable re-expression modulation require the potentially hazardous integration of viral DNA into the host genome, which limit their clinical use. The presented hit-and-run strategy circumvents these integrations, and assists to realize the curable genome concept, allowing a major role in modulation of expression of any gene at will, including currently undruggable (non)protein-coding genes.

In total, we utilized 38 ZFPS (34 self-engineered ZFPs) targeted to eight different genes (Table I), fused them to transcriptional expression modulators (VP64, SKD) or epigenetic enzymes (Tet2, G9A, SUV39H1 and M.Sss1) and studied the effect on various cell biological features, including target gene expression modulation, cell growth, DNA methylation and histone modifications.

Targeting the EpCAM and ICAM-1 promoter to decrease tumor growth in ovarian cancer
The first gene we modulated by ATF-technology was our model gene epithelial cell adhesion molecule (EpCAM), which is overexpressed in most carcinomas. Dependent on the tumor type, its overexpression is either associated with improved or worse patient survival. For ovarian cancer however, the role of EpCAM remains unclear. We used previously engineered ATFs for both the up- and downregulation of EpCAM. The bidirectional ATF-based approach is uniquely suited to study cell-type-specific biological effects of EpCAM expression. Using this approach, the function of EpCAM in ovarian cancer was investigated (Chapter 2).
**Table I.** The gene-targeted ATFs used in this thesis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>ATFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPCAM</td>
<td>Calcium-independent cell adhesion molecule overexpressed on cancer cells</td>
<td>2</td>
</tr>
<tr>
<td>ICAM-1</td>
<td>Cell adhesion molecule involved in immune response (44)</td>
<td>1*</td>
</tr>
<tr>
<td>C13ORF18</td>
<td>Putative TSG with unknown function, possibly involved in autophagy regulation</td>
<td>5</td>
</tr>
<tr>
<td>CCNA1</td>
<td>Cyclin involved in control of (germ line meiotic) cell cycle</td>
<td>6</td>
</tr>
<tr>
<td>TFPI2</td>
<td>Linked to regulation of plasmin-mediated matrix remodeling</td>
<td>11</td>
</tr>
<tr>
<td>Maspin</td>
<td>Inhibits cell motility and invasion, inhibitor of HDAC-1</td>
<td>2*</td>
</tr>
<tr>
<td>EPB41L3</td>
<td>Cell adhesion, cell motility and cell growth</td>
<td>2</td>
</tr>
<tr>
<td>HER2</td>
<td>Enhances activation of signaling pathways promoting cell proliferation (45)</td>
<td>1*</td>
</tr>
<tr>
<td>PLOD2</td>
<td>Catalyses the hydroxylation of lysyl residues in collagen-like peptides</td>
<td>8</td>
</tr>
</tbody>
</table>

*from literature

In the next chapter (Chapter 3), we investigated whether ICAM-1 can function as TSG in ovarian cancer in the absence of immune cells. ICAM-1 is a protein which is commonly expressed on endothelial cells and cells of the immune system. In ovarian cancer, ICAM-1 is downregulated compared with healthy ovarian cells and expression is correlated with decreased tumorigenicity (46). Whereas ICAM-1 expression on tumor cells is of importance for attracting immune cells, ICAM-1 might also induce tumorigenicity and chemoresistance. In ovarian cancer, such a role of ICAM-1 is unclear. We investigated whether ICAM-1 has a tumor suppressive role in this malignancy by re-expression of ICAM-1 using previously published ICAM-1-targeting ATFs (44).

**Re-expression of TSGs in cervical cancer**

To investigate whether we could achieve re-expression of genes which were aberrantly silenced by promoter hypermethylation in cervical cancer, we performed three different studies. In the first study, we aimed to re-express C13ORF18 in cervical cancer cell lines by five engineered ATFs and studied the effect of re-expression on cancer-related cell biological features. Currently, the function of C13ORF18 is unknown, but this gene is specifically methylated in cervical cancer compared to normal cervix (10). This supposed silenced state of C13ORF18 in cervical cancer hints towards a tumor suppressive role for this gene.
in this malignancy, which was elaborately explored in this chapter. Moreover, in this chapter we describe the effects of targeting ATFs to the C13ORF18 promoter on the local DNA methylation levels and histone modifications (Chapter 4).

In the second study (Chapter 5), we aimed to re-activate the hypermethylated candidate tumor suppressor genes C13ORF18, CCNA1, TFPI2 and Maspin by Tet-induced demethylation. To target Tet2 to hypermethylated CpGs, we used engineered zinc fingers proteins (ZFPs), which were all validated for gene regulatory capacity by fusions to VP64. We also compared the hypothesized tumor suppressive effects of these genes by the engineered gene-targeting constructs as well as by cDNA overexpression. We used primary spherical growing cell lines to study the functional effects, which are a better representative compared to established cell lines, and may gap the bridge with real tumors. Furthermore, we addressed the binding of ZFPs to their intended sequence and throughout the genome using ChIP-Seq and elaborate on the ‘questionable’ role of Maspin as TSG (47) in (cervical) cancer.

In the third study, we aimed to fully reprogram the biomarker EPB41L3 in order to improve the sustainability of gene re-expression using the synergistic effects of ATFs with epigenetic drugs (5-aza-dC or TSA). Until now, no methods exist to permanently re-activate single silenced genes, and this chapter describes a novel way to improve the sustainability of targeted re-expression. We also addressed the potential synergy of ATFs with epigenetic drugs. Interestingly, EPB41L3 was recently identified as the best marker for detecting CIN2/3 (13, 17), but its function in cervical cancer has not been explored yet (Chapter 6).

**Downregulation of gene-expression by epigenetic editing**

In order to down-regulate gene expression, we targeted Human Epidermal Growth Factor Receptor 2 (HER2) linked to repressive epigenetic editing domains. HER2 is an important member of the epidermal growth factor receptor family and overexpressed in several cancer types. Downregulation of this oncogene dramatically decreases cancer cell metabolism, growth and proliferation. In this chapter, we targeted the HER2 promoter using a previously described ZFP specific for the HER2 promoter (45) and fused it to histone methyl transferases G9a and SUV39H1. Subsequently, we investigated if we could overwrite the histone modifications in the HER2 promoter and whether this results in HER2 expression modulation and reduced tumorigenicity (Chapter 7).

To further exploit the potentials of ATFs and Epigenetic Editing for the downregulation of gene expression, we set out to target procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2). Excessive collagen deposition is the hallmark of fibrosis (48), and downregulation of PLOD2 is believed to inhibit this process. For the silencing of target genes, both targeted methylation of DNA and histones has been explored. However, insights into sustainability of gene expression modulation by such methods are largely lacking. In this chapter, we
designed a system to compare the effects of a strong DNA methylator (M.SssI) and a strong gene repressor (SKD) on sustainability of gene repression of PLOD2. We engineered stable inducible human skin fibroblasts to express ZFP-SKD or ZFP-M.SssI fusions targeting the PLOD2 promoter. After short term expression of the PLOD2-targeting constructs, we investigated the effects on gene expression and local epigenetic features (DNA methylation and histone marks). Finally, we aimed to downregulate PLOD2 expression in a profibrotic environment (mimicked by addition of TFGb) to demonstrate the therapeutic potential of our tools (Chapter 8). In Chapter 9 the results from chapter 2-8 are summarized and discussed.
References


