New molecular biomarker discovery for diagnosis and prognosis in oral and oropharyngeal cancer
Melchers, Lieuwe Jurjen

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:
2014

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.

Download date: 25-12-2018
CHAPTER 4

EpCAM in carcinogenesis: the good, the bad or the ugly

B.T. van der Gun, L.J. Melchers, M.H. Ruiters, L.F. de Leij, P.M. McLaughlin, M.G. Rots

Chapter 4

Introduction

The epithelial cell adhesion molecule (EpCAM; CD326) is a membrane glycoprotein originally discovered on colon carcinomas\(^{184}\). EpCAM is expressed by the epithelium of healthy individuals, except by squamous epithelium, and some specific epithelial cell types, such as hepatocytes and keratinocytes\(^{185}\), but in most human carcinomas, EpCAM is overexpressed to varying degrees\(^{186}\). The diagnostic and prognostic characteristics of EpCAM have been demonstrated by many independent research groups\(^{186,187}\) and the EpCAM overexpression is exploited in several EpCAM directed antibody- or vaccine-based clinical trials for a wide variety of carcinomas\(^{188}\). Recently, EpCAM has been identified as an additional marker for cancer-initiating stem cells\(^{189}\), which makes it an even more interesting target for cancer therapy.

Several biological functions of EpCAM have been described: EpCAM is able to abrogate E-cadherin-mediated cell–cell adhesion by disrupting the link between a-catenin and F-actin thereby loosening cell–cell adhesion\(^{72}\). In addition, association of EpCAM with claudin-7 interferes with EpCAM-mediated homotypic cell–cell adhesion, promoting cell motility, proliferation, survival, carcinogenesis and metastasis formation\(^{73}\). Furthermore, it has been shown that upon intramembrane proteolysis of EpCAM, the intracellular domain functions as part of a transcriptional complex inducing c-myc and cyclin A and E expression\(^{190}\). These findings support a role for EpCAM as an oncogene. Indeed, EpCAM overexpression is associated with decreased overall survival of patients with different types of cancer\(^{191-194}\)(see EpCAM: the Bad).

In contrast to its promoting role regarding tumour formation, EpCAM is also described as a tumour suppressive protein. EpCAM was first proposed to function as a cell adhesion molecule since EpCAM is able to mediate homophilic adhesive interactions\(^{195}\), thereby preventing cell scattering. Due to these adhesive properties, EpCAM is likely to play a role in inhibition of invasion\(^{195,196}\). Indeed, loss of EpCAM contributed to increased migratory potential\(^{197}\) and EpCAM expression on metastases was lower compared with primary tumours\(^{198}\). Moreover, EpCAM overexpression in some carcinoma types is associated with improved patient survival\(^{199-203}\)(see EpCAM: the Good).

The dual role of EpCAM is also reflected by mechanistic studies investigating the role of EpCAM by enforced modulation of EpCAM expression. Murine colorectal carcinoma cells transfected with murine EpCAM complementary DNA (cDNA) increased cell–cell adhesion, attenuated tumour cell invasion in matrigel and decreased tumour incidence and metastasis when inoculated in the spleen of the mice\(^{196}\). These data suggest that EpCAM expression antagonizes tumour growth and metastasis. In contrast, induction of EpCAM expression into human embryonic kidney cells (HEK293) as well as into murine fibroblasts showed an enhanced metabolism and colony formation capacity compared with the empty vector–transfected cells\(^{204}\). Furthermore, in four different carcinoma types, forced downregulation of EpCAM expression utilizing antisense or small interfering RNA (siRNA) decreased cell proliferation, migration and invasiveness\(^{202-207}\). Whether EpCAM acts as a tumour suppressive gene or as an oncogene might depend on the microenvironment. Since epigenetic regulation is associated with aberrant EpCAM expression, recent advances in epigenetic interference\(^{208,209}\) might be a promising novel approach to either...

Abstract

The epithelial cell adhesion molecule (EpCAM) is a membrane glycoprotein that is highly expressed on most carcinomas and therefore of potential use as a diagnostic and prognostic marker for a variety of carcinomas. Interestingly, EpCAM is explored as target in antibody-based therapies. Recently, EpCAM has been identified as an additional marker of cancer-initiating cells. In this review, we describe the controversial biological role of EpCAM with the focus on carcinogenesis as an adhesion molecule, Ep-CAM mediates homophilic adhesion interactions, which in turn might prevent metastasis. On the other hand, EpCAM abrogates E-cadherin mediated cell–cell adhesion thereby promoting metastasis. Also, upon cleavage of EpCAM, the intracellular domain functions as a part of a transcriptional complex inducing c-myc and cyclin A and E. In line with these seemingly controversial roles, EpCAM overexpression has been associated with both decreased and increased survival of patients. Similarly, either induction or downregulation of EpCAM expression lowers the oncogenic potential depending on the cell type. As epigenetic dysregulation underlies aberrant EpCAM expression, we propose epigenetic editing as a novel approach to investigate the biological role of EpCAM, expanding the options for EpCAM as a therapeutic target in cancer.

-53-
upregulate or downregulate EpCAM expression, depending on the tumour type. This review comprehensively summarizes the large body of evidence for EpCAM acting either as a tumour suppressor gene (the Good), an oncogene (the Bad) or both (the Ugly). We describe the increasing insights into (epi)genetic parameters involved in EpCAM regulation and discuss the carcinoma types that might benefit from future epigenetic approaches interfering with EpCAM gene expression, either inducing or repressing EpCAM expression.

**Biological role of EpCAM in carcinogenesis**

The highly overexpressed tumour-associated antigen on most carcinomas, currently referred to as EpCAM, has been ‘discovered’ multiple times\(^5\). With each discovery, EpCAM received the name of the respective monoclonal antibody or cDNA clone, leading to many synonyms. The name EpCAM reflects its function as a homophilic intercellular adhesion molecule as demonstrated by Litvinov et al.\(^6\). It has been suggested that the adhesive properties of EpCAM might prevent metastasis because intercellular adhesion should be reduced to gain the ability to migrate. Metastasis involves loss of cell-cell adhesion and cell polarity, an increase in cell motility and invasion. Development of such a mesenchymal phenotype is known as epithelial mesenchymal transition (EMT)\(^7\) and seems to require downregulation of EpCAM followed by re-expression at the site of the future metastasis. Indeed, EpCAM downregulation has been associated with EMT\(^7\) and in mice with colon carcinoma, small metastases were EpCAM negative but large metastases in the same mouse exhibited intense uniform membranous overexpression, frequently also associated with cytoplasmic staining\(^8,9\). In addition, EpCAM was found to be hyperglycosylated in carcinoma tissue as compared with healthy autologous epithelia\(^10\). The heavily glycosylated extracellular domain (EpEX) of EpCAM is a transmembrane protein consisting of an extracellular domain, a single transmembrane domain and a short 26-amino acid intracellular domain (EpICD)\(^11\). The extracellular domain comprises an epidermal growth factor-like domain, a thyroglobulin repeat domain followed by a cysteine-poor domain. The epidermal growth factor-like and thyroglobulin domains form a globular structure and are required for the homophilic cell-cell adhesion of EpCAM\(^12\). It has been suggested that the adhesive properties of EpCAM might prevent metastasis because intercellular adhesion should be reduced to gain the ability to migrate. Metastasis involves loss of cell-cell adhesion and cell polarity, an increase in cell motility and invasion.

EpCAM: the Good?

The name EpCAM reflects its function as a homophilic intercellular adhesion molecule as demonstrated by Litvinov et al.\(^6\). From this, one might expect EpCAM to prevent metastases and as such acting as a good guy. However, only for two tumour types, high EpCAM expression has been consistently associated with improved clinical outcome. In metastases of renal clear cell carcinomas, EpCAM was less frequently expressed than in primaries\(^13,14\) and high EpCAM expression in the primary was associated with improved patient survival\(^15,16\). Similarly, in thyroid carcinoma, EpCAM expression was lower in less differentiated tumours and high expression correlated with improved survival\(^17,18\). Downregulation of EpCAM in these tumours might reflect general tumour dedifferentiation, rather than a functional downregulation for EMT. However, the recent study by Ralhan et al.\(^19\) indicates that loss of membranous EpCAM in anaplastic (undifferentiated) thyroid carcinoma might be due to cleavage of EpCAM. More details of these studies are presented in table 4.1.

EpCAM: the Bad?

Based on the cell signalling role, EpCAM might very well play a promoting role in carcinogenesis as described for many tumour types (table 4.2). High EpCAM expression has been associated with decreased overall survival in carcinomas of the bladder\(^5\), gall bladder\(^5\) as well as of the...
In breast carcinoma, high EpCAM expression was observed in less differentiated tumours and associated with larger tumours, nodal metastasis and worse survival. Moreover, for this type of carcinoma, EpCAM was upregulated in metastases compared with node-negative disease and poor differentiation.

In squamous cell carcinoma of the lung, high EpCAM expression was associated with a trend towards a shorter survival in patients with high EpCAM expression. In patients with adenocarcinoma of the lung, no clear correlations of EpCAM expression with cancer progression, metastasis or survival have been found. Although, EpCAM has been shown to be an accurate diagnostic marker for reverse transcription-polymerase chain reaction-based identification of lymph node micrometastasis and the presence of EpCAM-positive tumour cells in lymph nodes correlated with reduced survival rates. The lack of a clear effect of EpCAM on survival in lung carcinoma might reflect different effects of EpCAM in various treatment subgroups. Subgroup analysis, as performed for breast carcinoma, might provide more insights.

For many other tumour types, high EpCAM expression has been associated with poor prognosis, but contradictory reports underline a dual role of EpCAM. In these tumour types for which both protecting and promoting roles for EpCAM have been described, we consider EpCAM to be an ugly player.

EpCAM: the Ugly?

Table 4.1. Protective (-) role of EpCAM in carcinomas.

<table>
<thead>
<tr>
<th>Carcinoma</th>
<th>Carcinogenesis</th>
<th>Progression</th>
<th>Metastasis</th>
<th>Survival</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renal Cell ca.</td>
<td>-p</td>
<td>-p</td>
<td>O&quot;</td>
<td>-p</td>
<td>Seligson, 2004,207</td>
</tr>
<tr>
<td>Thyroid ca.</td>
<td>-p</td>
<td>-p</td>
<td>O&quot;</td>
<td>-p</td>
<td>Klatte, 2009,211</td>
</tr>
</tbody>
</table>

-: a protective role or longer survival associated with EpCAM expression; O: no (significant) role found; p: in patient material; r: in mice/rats; c: in cell lines; ca.: carcinoma.

In the prostate, EpCAM expression was significantly increased from normal via prostatic intraepithelial neoplasia to adenocarcinoma, but expression in adenocarcinoma was not associated with differentiation grade or clinical outcome. Interestingly, hormone refractory carcinomas were found to express EpCAM in a significantly higher frequency than untreated carcinomas, but this finding was not confirmed in another study.

In squamous cell carcinoma (SCC) of the lung, high EpCAM expression was associated with nodal metastasis, high-stage disease and poor differentiation. A more recent study found only a trend towards a shorter survival in patients with high EpCAM expression. In patients with adenocarcinoma of the lung, no clear correlations of EpCAM expression with cancer progression, metastasis or survival have been found. Although, EpCAM has been shown to be an accurate diagnostic marker for reverse transcription-polymerase chain reaction-based identification of lymph node micrometastasis and the presence of EpCAM-positive tumour cells in lymph nodes correlated with reduced survival rates. The lack of a clear effect of EpCAM on survival in lung carcinoma might reflect different effects of EpCAM in various treatment subgroups. Subgroup analysis, as performed for breast carcinoma, might provide more insights.

For many other tumour types, high EpCAM expression has been associated with poor prognosis, but contradictory reports underline a dual role of EpCAM. In these tumour types for which both protecting and promoting roles for EpCAM have been described, we consider EpCAM to be an ugly player.

EpCAM: the Ugly?

For several tumour types, the reported role of EpCAM seems contradictory (table 4.3). In gastric cancer, Du et al. reported that EpCAM expression was associated with nodal metastasis and higher EpCAM expression showed an increase in the proliferating cell nuclear antigen. Survival was not analysed here, but in a different study EpCAM-positive disseminated tumour cells in pathological tumourfree lymph nodes were an independent prognostic factor for reduced survival. Yet, other studies found no significant relations with expression. However, a protective role for EpCAM was reported by Songun et al.: patients with high EpCAM expression had a better survival.
expression had a significantly better 10 years survival and loss of EpCAM identified aggressive tumours in early stage disease. Also for colorectal cancer (CRC), contradictory results have been reported. A reduced EpCAM expression at the invasive margin of CRC specimens correlated significantly with higher extent of tumour budding, tumour grade and risk of local recurrence. Interestingly, this finding was associated with nuclear localization of b-catenin, consistent with the signal transducer function of EpCAM. However, a significant positive correlation of EpCAM expression with survival has only been found in a subgroup of moderately differentiated colon cancers. The interaction of EpCAM with the cell–matrix adhesion molecule CD44v6 and the tight junction molecule claudin-7 in association with the tetraspanin CO-029 in tetraspanin-enriched membrane microdomains was initially found in CRC. Co-expression and complex formation of EpCAM and its partners (but not EpCAM alone) in liver metastases was accompanied by a significantly decreased disease free survival. Independently, CD44 and EpCAM were identified as markers of a subpopulation with greatly enhanced tumourigenicity in a human CRC—mouse xenograft model. In addition, EpCAM proved to be a good marker for reverse transcription-polymerase chain reaction–polymerase chain reaction to detect micrometastases in lymph nodes.

In specimens of head and neck squamous cell carcinoma (HNSCC), EpCAM messenger RNA (mRNA) expression increased from hyperplasia via dysplasia to tumour, which might suggest a role for EpCAM in carcinogenesis. EpCAM was also identified as a very good reverse transcription–polymerase chain reaction marker to detect micrometastases in lymph nodes and disseminated HNSCC cells. EpCAM is expressed de novo in HNSCC, but most studies do not find any relation with clinicopathologic variables, including differentiation and survival. However, in a study looking specifically at tongue SCC, EpCAM expression was associated with larger tumour size, nodal metastasis and tumour dedifferentiation. Interestingly, recently in a Taiwanese series of oral SCC, EpCAM expression was reported to decrease from normal via dysplasia to carcinoma, and lower EpCAM labeling index was associated with, among others, larger tumour size and presence of nodal metastasis. These adverse findings might be due to the large number of areca quid chewers in Taiwan. Areca quid has been shown to increase tumor necrosis factor-a production and therefore might downregulate EpCAM. In general, a lack of consistent association of EpCAM expression in HNSCC might be attributable to the heterogeneity of tumours included in these studies. In a group of esophageal cancer (mainly SCC) patients, high EpCAM expression indicated a significantly higher survival rate. In contrast, Stockeckl et al. identified high EpCAM as an independent prognostic factor for decreased survival. Others found no correlations with grade, stage or disease-specific survival. Furthermore, the presence of EpCAM-positive cells in pathological tumour-free lymph nodes was an independent indicator for a poor prognosis.

Epithelial ovarian cancer, EpCAM is highly overexpressed compared with normal ovarian surface epithelium and no differences in EpCAM expression were observed among different histological subtypes and grades in two independent studies. In one of these studies, with all-

<table>
<thead>
<tr>
<th>Carcinoma</th>
<th>Carcinogenesis</th>
<th>Progression</th>
<th>Metastasis</th>
<th>Survival</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head &amp; Neck:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral SCC</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Laimer, 2008§</td>
</tr>
<tr>
<td>- disseminated cells</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Went, 2008§</td>
</tr>
<tr>
<td></td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Stockeckl, 2006§</td>
</tr>
<tr>
<td></td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Kimura, 2007§</td>
</tr>
<tr>
<td>Gastric ca.</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Hosch, 2000§</td>
</tr>
<tr>
<td>- disseminated cells</td>
<td>+p</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>Du, 2009§</td>
</tr>
<tr>
<td>Colorectal ca.</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Devoci, 2007§</td>
</tr>
<tr>
<td>- disseminated cells</td>
<td>+p</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>Songun, 2005§</td>
</tr>
<tr>
<td>Gynecological:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovarian ca.</td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Heinzelmann, 2004§</td>
</tr>
<tr>
<td></td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Kim, 2003§</td>
</tr>
<tr>
<td></td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Spizzo, 2006§</td>
</tr>
<tr>
<td></td>
<td>+p</td>
<td>0p</td>
<td>0p</td>
<td>0p</td>
<td>Bellone, 2009§</td>
</tr>
</tbody>
</table>

Table 4.3: Protective (-) or promoting (+) role of EpCAM in carcinomas.

+, a promoting role in carcinogenesis (e.g. higher expression in tumour compared to normal), tumour progression (higher in larger tumours), metastasis (higher in metastasized tumours) or shorter survival associated with EpCAM expression; - a protecting role or longer survival associated with EpCAM expression; O: no (significant) role found; p: in patient material; m: in mice/rats; c: in cell lines; ca: carcinoma; *: effect on survival only in subgroup of moderately differentiated tumours; $: >$ as complex.

most half of the tumours being of the borderline type (low malignant potential), International Federation of Gynecology and Obstetrics stage III/IV showed lower EpCAM expression than stage I/I. However, in the other study, International Federation of Gynecology and Obstetrics stage III/IV showed significant higher EpCAM expression than stage I/I disease suggesting
that a higher expression of EpCAM correlates with tumour progression, but no correlation with relapse-free survival or disease-specific survival was found204. A more recent study without borderline tumours did find differences among histological subtypes and a significantly higher EpCAM expression in poorly differentiated tumours and overexpression correlated with decreased overall survival205. In a study that also looked into the expression in metastatic and recurrent disease, these tumours were found to express significantly higher levels of EpCAM compared with primaries206.

Studies regarding the expression of EpCAM suffer from the use of different antibodies, scoring methods and heterogeneous groups of tumours analyzed, which might lead to differing, possibly even contrasting results. Even when consistent results have been found, interpretation is not straightforward, as for example loss of EpCAM might be due to active downregulation during EMT or an effect of general tumour dedifferentiation. Nevertheless, it is quite clear that EpCAM plays a role in carcinogenesis, tumour progression and metastasis in various carcinoma types, providing opportunities for diagnosis and therapeutic interventions.

**Modulation of EpCAM expression to address the biological role of EpCAM**

The function of EpCAM as an adhesion molecule was discovered by induction of EpCAM in non-EpCAM-expressing cells196. Transfection of EpCAM murine cDNA in fibroblast and mammary carcinoma cell lines resulted in aggregates of cells caused by increased intercellular adhesion. Moreover, the EpCAM-positive transfectants segregated from the EpCAM-negative parental cells and EpCAM expression inhibited invasive growth in cell colonies. Additional evidence supporting the protective role of EpCAM in carcinogenesis has been obtained by either induction of EpCAM expression in colon or reduction of EpCAM in lung adenocarcinoma cell lines. Murine colorectal carcinoma cells transfected with cDNA encoding the murine EpCAM showed significant lower growth rates, colony formation and invasion through matrigel in vitro compared with the vector-only-transfected cells196. Also cells transfected with cDNA encoding human EpCAM showed reduced invasion through matrigel196. In syngeneic immunodeficient and immunocompetent mice, the EpCAM-transfected murine colorectal cells showed a reduction in metastatic potential compared with the control-transfected cells. In a lung adenocarcinoma cell line, reduction of EpCAM expression using short hairpin RNA showed an elevated cell invasion196.

Evidence supporting the promoting role of EpCAM in carcinogenesis has also been reported: stable transfection of EpCAM cDNA in HEK293 cells and murine fibroblast cells resulted in an increased metabolic activity and formation of larger and more colonies compared with the empty vector-transfected cells. Moreover, EpCAM expressing HEK293 induced the expression of c-myc and cyclins A and E196. Also silencing of EpCAM expression by siRNA in breast cancer cell lines showed inhibition of proliferation, migration and invasion of the treated cells196, which reflects the correlation between high EpCAM expression and poor prognosis in breast cancer patients. In agreement with its promoting role in patients with hepatocellular carcinoma, the EpCAM gene promoter has been shown to be overexpressed in primary tumours. Furthermore, these treated cells with lower EpCAM expression showed a reduced tumour growth in nude mice196, and a tail vein metastatic assay showed that intravenous inoculation of EpCAM siRNA-treated gastric carcinoma cells led to significantly less visible tumours in the liver compared with non-treated cells196. Inhibition of EpCAM expression by antisense mRNA in an HNCC cell line showed changes in morphology and reduced proliferation and metabolism196, indicating a promoting role for EpCAM in HNCC.

The above forced modulation studies suggest that EpCAM, whether it is a good or a bad guy for a given tumour, can serve as a therapeutic target for upregulation (good guy) or downregulation (bad guy).

**Regulation of EpCAM expression**

To better understand why EpCAM is overexpressed in carcinomas, more insights in the regulation of the epcam gene itself are required; therefore the (epi)genetic events involved in EpCAM regulation will now be described.

**Genetics**

The EpCAM protein is encoded by the TACSTD1 gene originally reported as the GA733-2 gene262. In the present manuscript, however, we prefer to use epcam as gene name because tumour-associated calcium signal transducer protein-1 precursor does not properly reflect the function of the encoded protein196. The epcam gene has a minimal estimated size of ~14 kb and is located on chromosome 2p21256. The epcam gene consists of a total of nine exons196, the mRNA is ~7.5 kb (NCBI: AHO03574); all reported open reading frames of EpCAM are identical and encode a protein of 314 amino acids196. No splicing variants were found, although a large number of carcinoma cell lines were screened262. To our knowledge, mutations in the epcam gene have only been identified in patients suffering from Lynch syndrome or congenital tufting enteropathy. In Lynch syndrome, different heterozygous germ line deletions disrupt the 3′-end of the epcam gene and lead to inactivation of the adjacent MSH2 gene through methylation induction of its promoter in tissues expressing EpCAM266. In congenital tufting enteropathy, four different point mutations have been described resulting in decreased or no expression of EpCAM on protein level266.

The EpCAM promoter region that controls the expression of the gene has been cloned and characterized266. The sequence upstream of the transcription start site (TSS) has been defined266 (NCBI: AF148093). A 3.4 kb fragment of this EpCAM 5′-regulatory sequence is capable of directing heterologous gene expression and the promoter activity is restricted to EpCAM-expressing cells266. A complementary study confirmed that the transcriptional activity of a 1.1 kb EpCAM fragment starting ~700 bp upstream of the TSS directly correlated with the amount of EpCAM expression196. In silico analysis of the EpCAM promoter revealed several homologies to known
transcriptional regulatory sequences and putative transcription-binding sites\(^{264}\). Although no TATA or CAAT boxes were found, the position of the consensus initiator element (Inr) matches with the putative TSS based on 5′-untranslated region sequencing studies\(^{265,266}\). By deletion analysis, it was established that 177 bp of the 5′-flanking sequence are sufficient to drive reporter gene expression, whereas the region 687–341 bp upstream of the TSS appeared to be responsible for epithelial-specific expression\(^{266}\).

**Transcription factors**

Several putative transcription-binding sites within the EpCAM promoter have been reported\(^{255,256,265,266}\) (Figure 4.1). For ovarian cancer, we have confirmed binding of several transcription factors by chromatin immunoprecipitation\(^{265}\). Unfortunately up till now, little biological data supporting an actual role for these transcription factors in epcam gene expression have been described. Indirect evidence has been reported for ESE-1 (epithelial-specific Ets-1): upregulation of ESE-1 in metastatic lymph nodes from lung, breast and pancreas cancers correlated well with the expression of EpCAM\(^{266}\). An indication that Sp1 plays an active role in EpCAM regulation was demonstrated by reporter gene analysis: after transfection with an EpCAM promoter fragment (-250 to +90, relatively to TSS) containing putative binding sites for Sp1 (Figure 4.1), an elevated promoter activity was observed in the presence of Sp1 compared with the activity in the absence of Sp1\(^{254}\).

Recently, it has been shown that β-catenin activation induced EpCAM transcription via binding of TCF/Lef at -489 upstream of the EpCAM TSS\(^{255}\). Interestingly, TCF/Lef and β-catenin are also involved in nuclear signalling by EpCAM itself\(^{255}\): proteolytic cleavage of EpCAM releases EpiCD, which forms a complex with β-catenin and TCF/Lef that contacts DNA at the Lef consensus sites. Therefore the authors suggested that EpiCD may impose a positive-feedback loop on EpCAM expression at the level of gene transcription\(^{255}\).

The transcription factors nuclear factor-κB (NF-κB) and p53 have been described as transcriptional repressors of the epcam gene: treatment of EpCAM-positive SCCs with tumor necrosis factor-α and interferon alpha resulted in a reduced endogenous EpCAM expression\(^{265,266}\). Inhibition of the activation of NF-κB by cotransfection of a plasmid coding for the dominant-negative NF-κB-inhibiting inhibitor-1-kappaB and a luciferase reporter plasmid under control of the EpCAM promoter, supported a direct role for NF-κB as a repressor of the EpCAM promoter. A second repressor of EpCAM promoter activity is the tumour suppressor gene p53\(^{255,266}\). Induction of wild-type p53 (WT p53) was associated with a dose-dependent decrease in EpCAM expression, whereas ablation of p53 expression was associated with an increase in EpCAM expression. Ten putative binding sites for p53 in the epcam gene were identified and by chromatin immunoprecipitation, the binding of WT p53 to a site located within intron 4 was confirmed\(^{266}\). Interestingly, simultaneous silencing of p53 and EpCAM expression via stable transduction of short hairpin RNA prevented the increase of EpCAM expression caused by ablation of p53 expression and decreased the invasiveness of the breast cancer cells\(^{266}\).

**Epigenetics**

Accessibility of transcription factors to the specific binding sites within the epcam gene depends on the chromatin structure, which is affected by DNA methylation and histone modifications\(^{266}\). Modifications of DNA and histones thus have profound impact on gene expression. Here, we will focus on DNA methylation and histone modifications, also because these epigenetic events are potentially reversible by drug treatments.

DNA methylation. Already in 1994, it was described that DNA methylation prevents amplification of the epcam gene\(^{266}\). Loss of DNA methylation in the epcam gene, caused by inactivation of the p53 gene, resulted in epcam gene amplification\(^{266}\). In view of these findings, the observation that downregulation of p53 caused upregulation of EpCAM expression is noteworthy\(^{266}\).

In humans, DNA methylation occurs mainly on cytosines within cytosine-guanine dinucleotides (CpGs). CpGs are relatively rare in the genome but tend to cluster in islands which are usually located in the 5′-regulatory region of many genes. Methylation of CpG islands in promoters leads to transcriptional silencing of genes. Several studies have reported that EpCAM expression is associated with DNA methylation\(^{264,270,271}\). In cell lines of different origin, high EpCAM expression was associated with hypomethylation and no EpCAM expression was associated with hypermethylation\(^{264,270,271}\). Interestingly, the CpG within the putative binding site for Sp1 (-231) was methylated in EpCAM-negative cell lines and not methylated in EpCAM-positive cell lines, whereas around the putative binding site for activator protein 1 (-125), the CpGs were unmethylated in all cell lines analysed\(^{264}\). Modulation by epigenetic drugs confirmed the correlation between EpCAM expression and the DNA methylation status of the epcam gene. Treatment of EpCAM-negative cell lines with a DNA demethylating agent (5-aza-2'-deoxycytidine) induced EpCAM expression de novo, both on mRNA and protein level and caused upregulation of EpCAM expression in EpCAM-positive cell lines\(^{264,270,271}\). However, in the EpCAM-negative leukaemia K562 (hypermethylated) and the liver HepG2 (CpGs were 50% methylated) cell lines, no EpCAM re-expression was observed after 5-aza treatment, although most methylated CpGs were converted to unmethylated CpGs\(^{264}\). In addition, upon 5-aza treatment of the EpCAM-negative lung carcinoma cell line GLC-1, of which part of the epcam gene (-830 to +282) is intermediated methylated, no de novo induction of EpCAM expression was detected\(^{264}\). Alternatively, we demonstrated that endogenous
Methylation status of the EpCAM promoter. In contrast, in breast cancer tissue, no correlation was observed between DNA methylation and EpCAM expression, reflecting the observed methylation status. Also in lung adenocarcinoma tissue -265 to -100, EpCAM expression was significantly associated with the methylation status of the EpCAM promoter. In contrast, in breast cancer tissue, no correlation was found between EpCAM protein expression and EpCAM promoter methylation for six CpGs measured. However, in the same study, they found the promoter of EpCAM-negative breast cancer cell line to be methylated to a higher degree as compared with an EpCAM-positive cell line. The discrepancy found between breast cancer cell lines versus tissue might be due to the number of CpGs analysed: MethyLight technology analysing six CpGs and bisulfite sequencing analysing 64 CpGs, respectively (table 4.4). Since in lung adenocarcinoma as well as in oral squamous cell carcinoma tissue, high EpCAM expression indeed correlated with a low DNA methylation level, the association between DNA methylation of the epcam gene and EpCAM expression in patient samples appears to depend on the tissue type (table 4.4). In normal colon tissues, 50% of the tested CpGs were methylated, whereas in colon cancer tissues, most CpGs were unmethylated. The expression level of EpCAM was 1000-fold higher in colon cancers than in normal colon tissue, reflecting the observed methylation status. Also in lung adenocarcinoma tissue and in oral squamous cell carcinoma, EpCAM expression was significantly associated with the methylation status of the EpCAM promoter. In contrast, in breast cancer tissue, no correlation was found between EpCAM protein expression and EpCAM promoter methylation for six CpGs measured. However, in the same study, they found the promoter of EpCAM-negative breast cancer cell line to be methylated to a higher degree as compared with an EpCAM-positive cell line. The discrepancy found between breast cancer cell lines versus tissue might be due to the number of CpGs analysed: MethyLight technology analysing six CpGs and bisulfite sequencing analysing 64 CpGs, respectively (table 4.4). Since in lung adenocarcinoma as well as in oral squamous cell carcinoma tissue, high EpCAM expression indeed correlated with a low DNA methylation level,

Table 4.4. Determination of molecular epigenetic marks for indicated regions of the EpCAM gene.

<table>
<thead>
<tr>
<th>Study</th>
<th>Material</th>
<th>Region</th>
<th>Technique</th>
<th>CpGs analysed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spizzo, 2007^273</td>
<td>breast cancer cell line</td>
<td>-356 to +361</td>
<td>Bisulfite sequencing</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>paraffin embedded breast cancer</td>
<td>-135 to -37 no</td>
<td>MethyLight</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>tissue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tai, 2007^274</td>
<td>lung aden. bladder, colon, germ cell</td>
<td>-265 to -100</td>
<td>Methylation Specific PCR</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>ovary carcinoma cell lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>lung adenocarcinoma tissue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>aChRbO EpCAM pos. H3K9me+</td>
<td>-265 to -100 yes</td>
<td>Methylation Specific PCR</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>EpCAM neg</td>
<td>-682 to -540</td>
<td>Chromatin ImmunoPrecipitation</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-356 to -140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yu, 2008^274</td>
<td>colon, prostate, breast, liver, haematological tumor cell lines</td>
<td>-321 to +790</td>
<td>Bisulfite sequencing</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>colon cancer tissue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van der Gun, 2008^275</td>
<td>lung, ovarian, colon carcinoma and human embryonic kidney, glioblastoma cell lines</td>
<td>-321 to +790 yes</td>
<td>Bisulfite sequencing</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-830 to +283</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shih, 2008^274</td>
<td>oral squamous cell carcinoma tissue</td>
<td>-265 to -100</td>
<td>Methylation Specific PCR</td>
<td>6</td>
</tr>
</tbody>
</table>

Positions are relatively to the transcription start site. Between brackets the number of CpGs analyzed by the indicated technique. The remark 'yes' or 'no' indicates the correlation between DNA methylation and EpCAM expression examined in patient tissue.

EpCAM expression can be actively downregulated in a persistent manner via induced DNA methylation^275. After delivery of the DNA methyltransferase M.SssI into EpCAM-positive ovarian carcinoma cells, methylation of the epcam gene resulted in reduced EpCAM expression, which maintained through successive cell divisions as the reduced EpCAM expression persisted for at least 17 days^276. The association between DNA methylation of the epcam gene and EpCAM expression in patient tissue appears to depend on the tissue type (table 4.4). In normal colon tissues, 50% of the tested CpGs were methylated, whereas in colon cancer tissues, most CpGs were unmethylated. The expression level of EpCAM was 1000-fold higher in colon cancers than in normal colon tissue, reflecting the observed methylation status. Also in lung adenocarcinoma tissue and in oral squamous cell carcinoma, EpCAM expression was significantly associated with the methylation status of the EpCAM promoter. In contrast, in breast cancer tissue, no correlation was found between EpCAM protein expression and EpCAM promoter methylation for six CpGs measured. However, in the same study, they found the promoter of EpCAM-negative breast cancer cell line to be methylated to a higher degree as compared with an EpCAM-positive cell line. The discrepancy found between breast cancer cell lines versus tissue might be due to the number of CpGs analysed: MethyLight technology analysing six CpGs and bisulfite sequencing analysing 64 CpGs, respectively (table 4.4). Since in lung adenocarcinoma as well as in oral squamous cell carcinoma tissue, high EpCAM expression indeed correlated with a low DNA methylation level,

Modulation of EpCAM by epigenetic editing

Although the exact biological role of EpCAM is not clear, yet the growth inhibitory effect of EpCAM overexpression or silencing is established for a list of different tumour types. For such types, modulation of EpCAM expression provides a promising approach to interfere with the oncogenic potential of these tumour cells. Since the aberrant expression of EpCAM on carcinomas seems to be associated with epigenetic mutations without underlying genetic defects, modulation of EpCAM expression by epigenetic interference opens up new possibilities to modify expression levels. Unlike genetic mutations, epigenetic mutations are reversible. The reversible nature of epigenetic mutations is currently exploited by genome-wide epigenetic drugs, for example to re-express tumour suppressor genes.

In epigenetic editing, molecular epigenetic marks (DNA methylation, post-translational histone modifications) are overwritten by targeting an epigenetic effector domain to the gene of interest using a sequence specific DNA-binding motif. Three classes of DNA-binding motifs are easily available to direct attached epigenetic effector domains to a specific sequence^277. These mo-
tifs are either based on synthetic polyamides, on designed recombinant zinc finger moieties or on oligonucleotides, which can form triple helices with the target double-strand DNA.

Trimeric and hexameric zinc finger proteins have been designed to target the EpCAM promoter and when fused to a transcriptional repressor or an activation domain, these artificial transcription factors have been shown to modulate the EpCAM promoter activity. Recently, we also designed an EpCAM-specific triple helix-forming oligonucleotide, which when coupled to a mutator methyltransferase is able to target methylation predominantly to a specific DNA sequence in the EpCAM promoter without significant background methylation. Alternatively, histone modifiers like the histone methyltransferase SUV39H and G9a have been successfully used as epigenetic effectors domains to silence genes: a minimal catalytic domain of the histone methyltransferase linked to a zinc finger targeting the vascular endothelial growth factor gene showed enrichment of H3K9 methylation associated with the vascular endothelial growth factor promoter, resulting in transcriptional repression of the vascular endothelial growth factor gene. Similarly, epigenetically silenced genes can be re-expressed: the hypermethylated tumour suppressor gene maspin was reactivated by an engineered zinc finger protein targeting the maspin promoter gene is a promising new tool in unravelling the role of EpCAM, opening up novel approaches in exploiting EpCAM as an anti-carcinoma therapeutic.

EpCAM in perspective

Whether EpCAM is a good or bad guy in cancer appears to be dependent on the cancer type. The ‘ugly’ role of EpCAM is reflected by studies describing both a protective and a promoting role within the very same cancer type. Tumours are very heterogeneous, consisting of phenotypically diverse cells, which is also reflected in the EpCAM expression. It might be that the role of EpCAM is not tumour type dependent but that the cell environment determines which function of EpCAM will dominate, resulting in the balance to either shift to a protective or promoting role in cancer. Eventually, because of this delicate balance, unravelling what triggers the cancer-promoting role will be a challenge.

The identification of EpCAM as a signal transducer has been a step forward towards elucidating a direct promoting role of EpCAM in carcinogenesis. As suggested by Munz et al., studies examining clinicopathologic correlations with surface EpCAM expression should be expanded with EpICD staining to better elucidate the signal-transducing role of EpCAM in tumour tissues. It would be interesting to know for example whether the reduced EpCAM expression observed at the invasive margin of rectal tumours is a result of cleavage of EpEX, followed by cleavage and translocation of EpICD to the nucleus promoting a more tumorigenic phenotype. Such translocation is indeed suggested in rectal tumours by increased cytoplasmic EpCAM staining and nuclear translocation of b-catenin, a partner of the EpICD transcription complex. Nuclear localization of EpICD might have been unambiguously proven by using an EpICD-specific antibody. Interestingly, very recently, a study on thyroid carcinomas reported that loss of membrane EpEX was accompanied by increased cytoplasmic and nuclear EpICD and nuclear b-catenin localization. Moreover, both loss of membranous EpEX as well as the presence of nuclear EpICD correlated with reduced patient survival. To our knowledge, this is the first study on a series of patient material showing the nuclear presence of EpICD. This finding suggests that the loss of membranous EpCAM expression as also observed in another thyroid carcinoma study might be due to cleavage more than functional downregulation of EpCAM as proposed in EMT. On the other hand, proteolysis of EpCAM could also explain the association of high EpCAM expression with poor survival. It has been suggested that EpICD in its transcription complex might induce EpCAM transcription via binding of Lef to the EpCAM promoter. Proteolysis of EpCAM induces shedding of EpEX followed by increased nuclear EpICD, which in turn might increase EpCAM expression. However, it is not clear why re-expression of membranous EpCAM appears not to take place in thyroid cancer, which directs to a tumour type dependent role for EpCAM.

In addition to taking EpICD into account, analysis for concomitant presence of claudin 7, CO-029, CD44v6 and EpCAM expression might give more information regarding patient survival, since not the solitary expression, but the presence of all four molecules in a complex formation has been shown to facilitate metastasis. Future clinicopathologic studies assessing co-expression of EpCAM in combination with its partners and nuclear EpICD staining may solve the ugly role of EpCAM in cancer.

In this review, we classified the tumour types for which EpCAM acts as potential diagnostic marker with some prognostic significance. EpCAM functions as a target in antibody-based clinical trials and in 2009, the European Medicines Agency approved the use of trifunctional bispecific antibody catumaxomab, which binds to EpCAM and enhances the immunological response against EpCAM-positive cells in malignant ascites. The emerging function of EpCAM in cell proliferation, migration and possibly cancer initiation broadens the interest to use EpCAM not only as an immunotarget but also as a target for epigenetic silencing. Cancer stem cells expressing EpCAM are more tumorigenic than EpCAM-negative stem cells and because cancer stem cells are radiation and drug resistant, targeting EpCAM might be a promising approach to stop tumour initiation and progression. Since epigenetic dysregulation seems to underly aberrant EpCAM expression, epigenetic editing provides unique tools to elucidate the promoting or protective role of EpCAM by upregulate or downregulate EpCAM expression.

Acknowledgements: We apologize to our colleagues that due to size constraints, we could not profoundly describe all their work on the role of EpCAM in cancer.

Funding: This study has been partially funded by EU FP7 EuroTransBio: ProTuMa ETB09008.